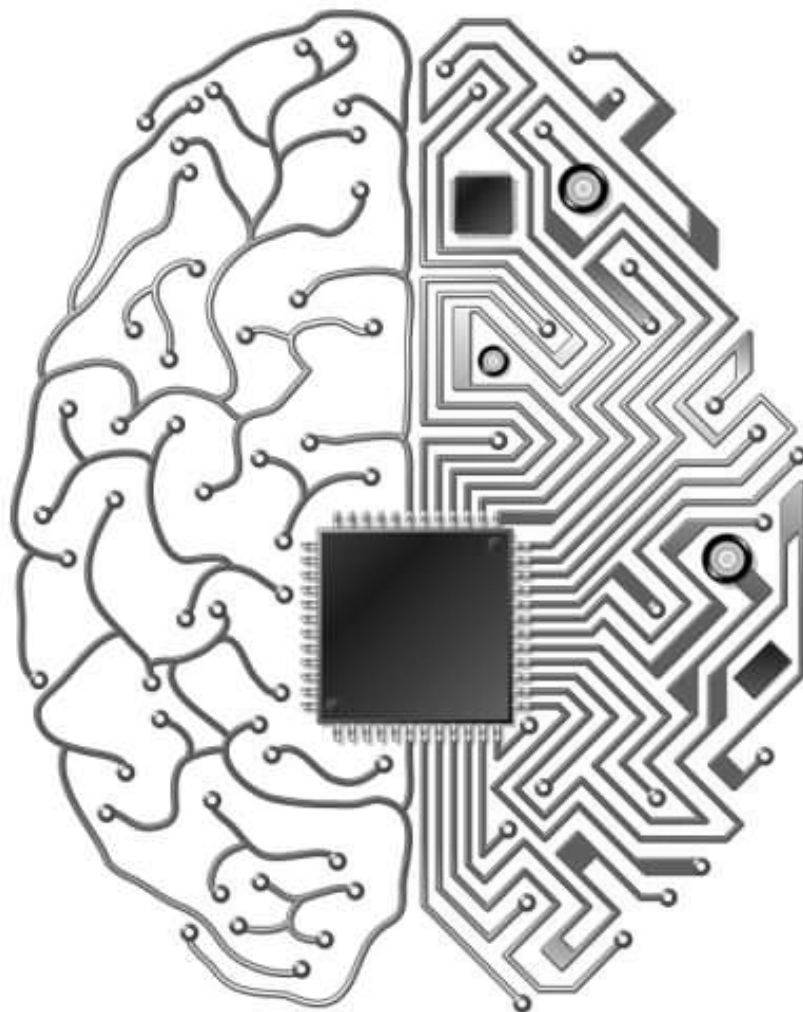


Ю.С. Синєкоп, А.М. Продеус, Є.Я. Швець,
Є.М. Кісельов, М.М. Баран

ЕКСПЕРТНІ СИСТЕМИ В МЕДИЦИНІ



Міністерство освіти і науки України
Національний технічний університет України
«Київський політехнічний інститут»
Запорізька державна інженерна академія

ЕКСПЕРТНІ СИСТЕМИ В МЕДИЦИНІ

Навчальний посібник
для студентів вищих навчальних закладів

Рекомендовано

Міністерством освіти і науки України

№1/11-2642 від 17.02.2014 р.

Експертні системи в медицині: Навчальний посібник / Продеус А. М., Синєкоп Ю. С., Швець Є. Я., Кісельов Є. М., , Баран М. М. – Запоріжжя: Видавництво ЗДІА, 2014. – 332 с.: іл.

ISBN

У навчальному посібнику розглянуті питання технології експертних систем медичного призначення. При цьому враховано, що найбільшого поширення набуло трактування ЕС як різновидів систем штучного інтелекту. Така точка зору є синтетичною, оскільки узагальнює досягнення декількох технологій - програмування, системотехніки, інженерії знань, обчислювальної техніки, - що істотно використовують апарат математичної статистики, формальної логіки, теорії множин і ряду інших математичних дисциплін. Розглянуті питання практичної реалізації експертних систем, приведені численні приклади практичної реалізації медичних експертних і інформаційно-довідкових систем. У додатках дані описи комп'ютерних практикумів (лабораторних робіт), реалізованих в середовищах Excel і Matlab, виконання яких дозволить краще засвоїти елементи теорії і практики створення експертних систем медичного призначення.

Навчальний посібник призначений для студентів, аспірантів, фахівців в області біомедичної електроніки, а також для медичних працівників, що цікавляться автоматизацією ухвалення рішень.

Рекомендовано Міністерством освіти і науки України у якості навчального посібника для студентів вищих навчальних закладів України. Гриф надано 17.02.2014, лист №1/11-2642.

Рецензенти:

А.В. Чалий, доктор фізико – математичних наук, професор, завідувач кафедри медичної і біологічної фізики Національного медичного університету ім. А.А. Богомольця, член – кореспондент Академії педагогічних наук України, заслужений діяч науки і техніки України

В.І. Бойко, доктор технічних наук, професор, завідувач кафедри електроніки Дніпродзержинського державного технічного університету, заслужений діяч науки і техніки України

В.О. Гусев, доктор технічних наук, професор, завідувач кафедри електронної техніки Севастопольського національного технічного університету

Зміст

Передмова	8
Частина 1 Технологія експертних систем медичного призначення.....	10
Розділ 1 Медичні інформаційні системи (МІС)	10
1.1. Визначення, цілі, задачі МІС	10
1.2. Класифікація МІС	11
1.3. Принципи розробки МІС	13
1.4. Проблеми розробки і впровадження МІС	15
1.5. Інформаційна база МІС	16
1.6. Математичне забезпечення МІС	17
1.7. Організація обробки біомедичної інформації.....	18
1.8. Автоматизація медичної діагностики	20
1.8.1. Задача медичної діагностики.....	22
1.8.2. Види медичної діагностики	24
1.8.3. Діагностичні алгоритми	26
1.8.4. Достовірність діагностики з використанням МІС	28
1.9. Експертні системи: концепції.....	29
1.10. Експертні системи медичного призначення	31
Короткі висновки.....	35
Література до Розділу 1	36
Контрольні запитання до Розділу 1	37
Розділ 2 Моделювання «клінічного мислення» лікаря.....	38
2.1. Природний інтелект: психологічний аспект.....	38
2.2. Природний інтелект: анатомо-фізіологічний аспект	44
2.3. Моделювання мислення: короткий історичний екскурс	45
2.4. Моделювання інтелекту лікаря – кібернетичний підхід	46
2.4.1. Автоматизація формування висновку лікаря	46
2.4.2. Інформативність симптомів.....	48
2.4.3. «Медична пам'ять» - модель пам'яті лікаря	49
2.4.4. Моделювання логічних міркувань лікаря	50
2.5. Системи штучного інтелекту в медицині	55
2.5.1. Сучасні дослідження в галузі медичних систем ШІ.....	56
2.5.2. KliniC - методологічна модель прийняття медичних рішень	57
2.5.3. Патологічні стани як елементи декларативних знань	59
2.5.4. Стверджуючі знання як зв'язки між станами.....	60
Короткі висновки.....	62
Література до Розділу 2	63
Контрольні запитання до Розділу 2	64
Розділ 3 Представлення знань в ЕС.....	65
3.1. Знання та дані.....	65
3.2.Класифікація знань.....	69
3.3. Основні формалізми для представлення знань	71

3.3.1. Правило продукції	71
3.3.2. Байєсова стратегія як форма реалізації правил продукції.....	73
3.3.3. Семантичні мережі	77
3.3.4. Фрейми.....	80
3.4. Проблеми використання Байєсових стратегії (на прикладі іридодіагностичних ЕС).....	82
3.4.1. Метод іридодіагностики	82
3.4.2. Статистична залежність ознак.....	85
3.4.3. Неповнота апіорних відомостей	87
3.4.4. "Дефекти" даних	90
3.4.5. Чинники, що "заважають"	90
3.5. Правила продукції з елементами Байєсівської стратегії	92
3.6. Системи нечіткого виводу.....	95
3.6.1. Основні поняття систем нечіткого виводу.....	95
3.6.2. Основні етапи нечіткого висновку	96
3.6.3. Основні алгоритми нечіткого виводу.....	102
3.7. Моделювання систем нечіткого виведення в середовищі MATLAB	103
3.7.1. Інструментарій для моделювання систем нечіткого виведення в інтерактивному режимі	103
3.7.2. Приклад моделювання в інтерактивному режимі	103
Короткі висновки.....	111
Література до Розділу 3	112
Контрольні запитання до Розділу 3	113
Розділ 4 Вимірювання і діагностика в медичних інформаційних системах ..	114
4.1. Структура і зміст теорії вимірювання.....	114
4.1.1. Теорія вимірювання як "сплав" приватних теорій	114
4.1.2. Рівні теорії вимірювання.....	114
4.1.3. Основні напрями фундаментальної теорії вимірювання.....	115
4.1.4. Основні напрями прикладної теорії вимірювання	116
4.2. Фізичні і математичні моделі, що використовуються при вимірюванні. 117	
4.3. Вимірювання елементарні і комплексні	117
4.4. Метрологія, природничі і гуманітарні науки	118
4.5. Вимірювання в медицині і психології.....	119
4.5.1. Про джерела похибок при апаратурних вимірюваннях в медицині ..	120
4.5.2. Вимірювальні шкали в психології і медицині	122
4.6. Шкали вимірювальні	130
4.6.1. Поняття шкали в фізичному напрямі теорії вимірювання	131
4.6.2. Поняття шкали в математичному напрямі теорії вимірювання	132
4.6.3. Види шкал.....	133
4.6.4. Шкали в психології.....	134
4.6.4.1. Первинні і вторинні оцінки.....	135
4.6.4.2. Форми представлення шкал.....	137
4.6.4.3. Зв'язок між шкалами	139

4.6.5. Рівень ("сила") шкали.....	141
4.7 . Медична діагностика як вимірювальна задача.....	141
4.7.1 . Класифікація як розбиття тезауруса на підмножини.....	143
4.7.2 . Організація класів в задачі діагностики.....	144
4.7.3 . Позначення класів	145
4.7.4 . Ознаки і параметри.....	146
4.7.5. Класи і ознаки - відносність понять	147
Короткі висновки.....	148
Література до Розділу 4	150
Контрольні запитання до Розділу 4	151
Частина 2 Практична реалізація експертних систем	152
Розділ 5 Засоби побудови ЕС	152
5.1. Особливості експертної системи як програмного продукту.....	152
5.2. Класифікація ЕС	154
5.3. Класифікація інструментальних засобів побудови ЕС.....	158
5.3.1. Мови програмування для додатків в області ЕС.....	159
5.3.2. Мови інженерії знань	171
5.3.3. Допоміжні засоби	175
5.3.4. Засоби підтримки.....	176
Короткі висновки.....	177
Література до Розділу 5	178
Контрольні запитання до Розділу 5	178
Розділ 6 Розробка ЕС	180
6.1. Об'єктивні передумови для розробки ЕС.....	180
6.2. Умови доцільності розробки експертної системи.....	181
6.3. Рівні розробки ЕС	182
6.4. Етапи розробки ЕС	183
6.5. Вибір інструментального засобу побудови ЕС	185
6.6. Робота з експертами (вилучення знань).....	186
6.7. Класифікація розробників ЕС	187
6.8. Труднощі розробки ЕС	188
6.9. Приклад розробки ЕС.	193
6.10. Експертні системи та ринок	197
6.10.1. Плюси і мінуси розробки та придбання ЕС.....	197
6.10.2. Як придбати готову ЕС?	198
Короткі висновки.....	199
Література до Розділу 6	200
Контрольні запитання до Розділу 6	201
Частина 3 Приклади практичної реалізації медичних інформаційних систем. 202	
Розділ 7 Лікувально-діагностичні центри та експертні системи.....	202
7.1 Мікропроцесорні системи домашнього і віддаленого моніторингу	202
7.2. Лікувально-діагностичні системи з комплексуванням методів традиційної і нетрадиційної медицини.....	207

7.2.1. Технологія лікувально-діагностичного та реабілітаційного центру "ЛИДИР"	208
7.2.1.1. Технічне рішення на комп'ютерній платформі Macintosh.....	208
7.2.1.2. Технічне рішення на базі IBM-сумісних комп'ютерів	211
7.3. Технологія багатофункціонального комп'ютерного лікувально-діагностичного комплексу "АМЕКС"	214
Короткі висновки.....	215
Література до Розділу 7	216
Контрольні запитання до Розділу 7.....	217
Розділ 8 Експертні та інформаційно довідкові системи.....	218
8.1. Іридодіагностичні комп'ютерні комплекси та програми	218
8.1.1. Експертна система для іридодіагностики ЭСИД	222
8.1.2. Експертна система для іридодіагностики ЭСИД-2.....	227
8.1.3. Програмний комплекс для іридодіагностики: ЭСИД-3, електронні атласи і графічний редактор GIDRA	229
8.1.4. Труднощі і перспективи комп'ютерної іридодіагностики.....	235
8.2. Інформаційно-довідкові системи для гомеопатії	237
8.2.1. Комп'ютерна програма «REPER» для підбору гомеопатичних препаратів	239
8.2.2. Методика призначення гомеопатичних препаратів за допомогою програми REPER.....	239
8.2.3. Облік сили впливу гомеопатичних препаратів	242
8.2.4. Інтерфейс - коротка характеристика.....	243
Короткі висновки.....	243
Література до Розділу 8	244
Контрольні запитання до Розділу 8.....	245
Розділ 9 Оболонки експертних та інформаційно-довідкових систем.....	246
9.1. Оболонка експертних систем "Універсал"	246
9.1.1. Можливості та режими функціонування	246
9.1.2. Приклад рішення задачі прогнозування з використанням ОЕС "Універсал"	252
9.1.3. Валеоаналіз уроку з використанням ОЕС «Універсал»	256
9.1.4. Використання ОЕС «Універсал» для самоорганізації здорового способу життя	258
9.1.5. Рання діагностика судинних захворювань мозку з використанням ОЕС «Універсал»	259
9.2. Оболонки інформаційно-довідкових систем МЕД і МEG	261
9.2.1. Основні характеристики оболонок	263
9.2.2. Бази даних та знань	265
9.2.3. Особливості інтерфейсу.....	265
9.2.4. Принципи формування висновку.....	267
9.2.5. Досвід експлуатації та перспективи використання.....	267
Короткі висновки.....	268

Література до Розділу 9	268
Контрольні запитання до Розділу 9	269
Розділ 10 Вибрані питання теорії експертних систем	270
10.1. Вилучення знань у експертів.....	270
10.1.1. Експертне оцінювання як процес вимірювання	270
10.1.2. Методи вимірювання ступеня впливу об'єктів.....	271
10.1.2.1. Метод ранжування	272
10.1.2.2 . Метод парних порівнянь	273
10.1.2.3 . Метод безпосередньої оцінки.....	274
10.1.3 . Формування групи експертів.....	275
10.2. Обробка експертних оцінок.....	278
10.2.1. Завдання обробки	278
10.2.2. Обробка парних порівнянь	281
10.2.3. Визначення узагальнених ранжировок	286
10.3 . Експертні системи з невизначеними знаннями	288
10.3.1 . Проблеми , властиві ЕС з невизначеними знаннями	288
10.3.2 . Суб'єктивні ймовірності.....	289
10.3.3 . Байєсівське оцінювання	290
10.4 . Ймовірнісний логічний висновок	291
10.4.1 . Найпростіший логічний висновок	291
10.4.2. Діагностичний алгоритм із застосуванням формули Байєса	293
10.4.3. Поширення ймовірностей в ЕС	296
10.4.4. Послідовне поширення ймовірностей	298
Короткі висновки.....	299
Література до Розділу 10	301
Контрольні запитання до Розділу 10	301
Додатки.....	303
Лабораторна робота 1 Оцінювання інформативності діагностичних ознак ..	303
Лабораторна робота 2_Ранговая кореляція діагностичних ознак	309
Лабораторна робота 3_Рекурентна форма формули Байєса	316
Лабораторна робота 4 Адаптивне вимірювання розподілу апіорних ймовірностей.....	319
Лабораторна робота 5 Адаптивне вимірювання розподілу умовних ймовірностей.....	323
Лабораторна робота 6 Методи нечіткої логіки в задачах підтримки прийняття рішення.....	327
Лабораторна робота 7 Моделювання систем нечіткої логіки в середовищі Matlab	331

Передмова

Поява перших «справжніх» медичних експертних систем (ЕС) відходить до початку 70-х років. Тим не менш, інтерес до експертних систем як до різновиду систем штучного інтелекту не слабшає. За минулі роки накопичений значний досвід розробки експертних систем (медичні системи займають серед них одну з ключових позицій), узагальнений в численних статтях, монографіях і довідниках. Однак і сьогодні побудова ефективних експертних систем різного призначення, і в тому числі, медичних систем, залишається завданням актуальним і все ще далеким від свого завершення. Пройдено лише перший відрізок довгого шляху, труднощі подолання якого обумовлені, з одного боку, надзвичайною складністю об'єкту, що моделюється з використанням експертних систем - мозку людини, а з іншого боку - високою складністю завдань, вирішення яких покладається на медичні експертні системи.

Даний навчальний посібник складається з десяти розділів, згрупованих у три частини, і Додатку, що містить інструкції та теоретичні відомості для виконання семи комп'ютерних практикумів (лабораторних робіт).

У першій частині, що складається з чотирьох розділів, розглянуті питання технології експертних систем медичного призначення. При цьому враховано, що найбільшого поширення набуло трактування ЕС як різновиду систем штучного інтелекту. Така точка зору є синтетичною, оскільки підсумовує досягнення декількох технологій - програмування, системотехніки, інженерії знань, обчислювальної техніки, - істотно використовують апарат математичної статистики, формальної логіки, теорії множин і ряду інших математичних дисциплін. Разом з тим, така точка зору представляється історично закономірною, дозволяючи углядіти успадкування медичними експертними системами основних досягнень кібернетики взагалі і медичної кібернетики зокрема, включаючи, в першу чергу, досягнення теорії розпізнавання образів, теорії шкал і алгоритмічних вимірювань.

Друга частина, що складається з двох розділів, присвячена питанням практичної реалізації експертних систем. При цьому зіставлені різні інструментальні засоби розробки експертних систем, і в першу чергу - мови програмування для додатків у вигляді ЕС, мови інженерії знань, зокрема, оболонки ЕС. Крім того, розглянуто ряд важливих питань, пов'язаних як з розробкою, так і з придбанням ЕС: об'єктивні передумови та умови доцільності розробки ЕС; рівні і етапи розробки ЕС; вибір інструментального засобу побудови ЕС; робота з експертами; плюси і мінуси розробки та придбання ЕС.

У третій, заключній, частині наведені приклади практичної реалізації медичних експертних та інформаційно-довідкових систем у вигляді систем прикійкового моніторингу, комплексів з поєднанням методів традиційної і нетрадиційної медицини, іридодіагностичних систем, гомеопатичних інформаційно-довідкових систем та ін. Розглянуто ряд прикладів швидкого створення порів-

няно простих, проте вельми корисних експертних систем за допомогою оболонки «Універсал». Показані корисність і зручність експлуатації оболонок інформаційно-довідкових систем «MED» і «MEG». Крім того, більш докладно (ніж у першій частині) розглянуті деякі важливі і складні питання створення експертних систем, такі як витяг знань у експертів і обробка експертних оцінок, а також детально, з використанням безлічі прикладів, розглянуті основи створення так званих «байєсівських» ЕС.

Даний навчальний посібник значною мірою являє собою результат переробки та розширення матеріалів однойменного навчального посібника А.Н. Продєуса і Е.Н. Захрбової (вид. К: «ВЕК+», 1998). Необхідність такої переробки досить очевидна. По-перше, за минулий період рівень інформаційних технологій істотно підвищився. По-друге, істотно зріс і рівень розуміння теоретиками експертних систем того важливого місця, яке займає, серед інших механізмів логічного висновку, байєсівська стратегія. Так, здавалося вельми наївним і некоректним припущення про статистичну незалежність діагностичних ознак, в кінцевому рахунку, дозволило створювати системи, чиї рішення мало відрізнялися від рішень значно складніших оптимальних або близьких до них систем. Крім того, в навчальний посібник включені матеріали, присвячені системам нечіткого виводу, а також матеріали, в яких детально висвітлені питання формалізації вилучення знань у експертів і комплексування експертних оцінок. Навчальний посібник закінчується Додатком, в якому дано опис семи комп'ютерних практикумів (лабораторних робіт), виконання яких дозволить краще засвоїти елементи теорії і практики створення експертних систем медичного призначення.

Навчальний посібник призначений для студентів, аспірантів, фахівців у галузі біомедичної електроніки, а також для медичних працівників, які цікавляться автоматизацією прийняття рішень.

Автори дякують за підтримку завідувачу кафедри фізичної та біомедичної електроніки доктору технічних наук, професору В.І.Тимофєєву, завідувачу кафедри акустики і акустоелектроніки доктору технічних наук, професору В.С.Дідковському.

Крім того, автори дякують технічним редакторам і перекладачам, багатьом співробітникам видавництва, чия праця допомогла надати книзі її остаточний вигляд.

Оскільки повністю уникнути неточностей, а можливо, і деяких помилок, на-вряд чи вдалося, автори будуть вдячні всім, хто надішле свої зауваження за адресами: aprodeus@gmail.com, mihail.baran@gmail.com.

Частина 1 Технологія експертних систем медичного призначення

Розділ 1 Медичні інформаційні системи (МІС)

Експериментальні дослідження процесів розумової діяльності людини показали, що 85% часу йде на створення умов для власне розумової роботи, що зводяться до знаходження необхідної інформації, друкування, побудови графіків та іншої "канцелярської" роботи, яку, в принципі, можуть виконувати різні технічні пристрої. До роботи медичного дослідника або практикуючого лікаря цей висновок цілком можна застосувати.

Інструментом, що удосконалює діяльність медичних працівників у сфері переробки інформації, управління і планування виступає медична інформаційна система (МІС).

1.1. Визначення, цілі, задачі МІС

Відомі, щонайменше, два підходи до визначення МІС.

У широкому сенсі під медичною інформаційною системою розуміють форму організації діяльності в медицині та охороні здоров'я, яка об'єднує медиків, математиків, інженерів, техніків з комплексом технічних засобів і забезпечує збір, зберігання, переробку і видачу медичної інформації різного профілю в процесі вирішення певних завдань медицини та охорони здоров'я.

У вузькому сенсі медичну інформаційну систему називають комплексом технічних засобів і математичного забезпечення, призначеного для збору, аналізу медико - біологічної інформації та видачі результатів в зручному для користувача вигляді.

Надалі, якщо це не буде спеціально обумовлено, ми будемо дотримуватися другого визначення.

Створення медичної інформаційної системи переслідує декілька цілей:

- підвищення якості діяльності медичних працівників та установ охорони здоров'я шляхом організації досконалої (відповідає рівню використовуваних технічних засобів) обробки медичної інформації, в тому числі шляхом вдосконалення процесів управління і планування;
- полегшення праці медичних працівників, ліквідація трудомістких і мало-ефективних процесів ручної обробки та аналізу медичних даних;
- забезпечення ефективного обміну інформацією з іншими інформаційними системами.

Найбільш загальні завдання МІС в клінічних установах:

- об'єктивізація трактування результатів досліджень (за деякими даними, помилково позитивні або помилково негативні, клінічно недостатні або неправильно витлумачені результати рентгенологічного, електрокардіологічного та лабораторних досліджень приводять в 31% випадків до помилкового діагнозу);
- автоматизація обробки інформації на етапі попередньої роботи медичного персоналу з визначення діагнозу і вироблення тактики лікування (лікар приймає остаточне рішення з питань діагностики та лікування хворого);
- автоматизація лабораторних досліджень: біохімічних, електрофізіологічних, рентгенорадіологічних та інших;
- створення баз (банків) даних: накопичення відомостей про кожного хворого для подальшого аналізу матеріалу, організація обробки цієї інформації відповідним математичним забезпеченням (в тому числі системами управління базами даних - СУБД);
- створення баз знань: накопичення знань експертів в галузі медицини та охорони здоров'я, необхідних для розробки експертних систем діагностики, лікування і реабілітації, профоглядів, експертизи, планування і управління;
- впорядкування потоку інформації всередині установи (завдання організаційного управління, завдання кадрові, матеріально-технічного постачання, статистичні звіти, оцінка діяльності відділень лікарень за деякими поглибленими показниками та інше).

Практично всі перераховані задачі можуть бути ефективно вирішені із застосуванням експертних систем - одного з прогресивних напрямків сучасних комп'ютерних технологій, що суттєво базується на ідеях створення штучного інтелекту (ШІ).

1.2. Класифікація МІС

Класифікацію МІС проводять, використовуючи різні класифікаційні ознаки.

Залежно від ступеня автоматизації процесів збору та обробки інформації МІС діляться на автоматизовані і автоматичні. В автоматизованих системах частина операцій щодо збору та обробки інформації виконується людиною. Автоматичні системи припускають повне виключення людини з процесів збору та обробки інформації.

Залежно від типу інформаційної бази МІС діляться на системи, які оперують з даними і зі знаннями. Системи другого типу називають експертними системами - по-перше, їх функціонування істотно спирається на знання, отримані від експертів, а по-друге, результати їх функціонування у відомому сенсі близькі результатами функціонування експертів.

Залежно від виду вирішуваних задач МІС можна розділити на три групи:

- інформаційно-довідкові - системи автоматизованого пошуку, вимірювальні системи;
- інформаційно-логічні - діагностичні системи, які прогнозують, системи, що стежать (системи моніторингу);

- керуючі - автоматизовані системи управління.

Інформаційно-довідкова система, крім пошуку інформації, здатна зробити певні перетворення інформації і сформувати довідковий документ.

Інформаційно-логічна система призначена для перетворення інформації, що зберігається в її пам'яті таким чином, щоб можна було отримати нову інформацію, відсутню раніше в інформаційному масиві.

У системах управління реалізується принципово нова функція - вироблення керуючих впливів.

Найбільш широке поширення у медичних закладах отримали інформаційно-пошукові системи (ІПС), які, залежно від характеру інформації, що зберігається, діляться на системи документальні та фактографічні.

Фактографічні ІПС містять інформаційні масиви фактичних даних. Аналогами таких систем виступають "паперові" довідники, каталоги, технічні паспорти. У комп'ютерних ІПС фактичні дані зазвичай зберігаються в базах даних (БД) і являють собою таблиці, в колонках яких перераховуються різні характеристики об'єктів, а в рядках наведено описи (значення характеристик) цих об'єктів. З використанням такого роду таблиць може бути сформована, так звана, "медична пам'ять" - таблиця, де в стовпцях зазначаються найменування захворювань, у рядках - найменування симптомів, а на місці перетину - значення частоти народження даного симптому при даному захворюванні (прикладом такої "медичної пам'яті" може служити табл. 1.1 для диференціальної діагностики геморагічного та ішемічного інсульту, де оцінка частоти народження має вербальну форму; далі будуть наведені приклади, коли ця оцінка проведена в числовій формі).

Таблиця 1.1. Диференціальна діагностика геморагічного і ішемічного інсульту (за Мисюк Н.С.)

ОЗНАКА	ІНСУЛЬТ геморагічний	ІНСУЛЬТ ішемічний
Передвісники	Малохарактерні	Характерні
Раптовий розвиток	Характерний	Менш характерний
Повільний розвиток	Нехарактерний	Характерний
Втрата свідомості	Характерна	Менш характерна
Блідість обличчя	Нехарактерна	Характерна
Гіперемія	Характерна	Нехарактерна
Артеріальна гіпотонія	Нехарактерна	Характерна
Артеріальна гіпертензія	Характерна	Нехарактерна
Гіпертермія	Характерна	Нехарактерна
Менінгізм	Характерний	Нехарактерний
Кров у лікворі	Характерна	Нехарактерна
Крові у лікворі немає	Нехарактерно	Характерно

Документальні ІПС оперують з інформацією у вигляді документів - прикладами таких систем можуть бути бібліографічна картотека, картотека з історіями хвороб, інші картотеки. Виконуючи пошук, документальна ІПС видає або номери документів, або їх заголовки, або адреси знаходження шуканих документів. При цьому оцінку інформації, що міститься в знайдених документах, дає людина.

Діагностичні системи до певної міри нагадують ІПС, їх можна розглядати як "складні" ІПС, що здійснюють вирішення складної (у плані виконуваної математичної обробки інформації) діагностичної задачі - пошук найменування передбачуваного захворювання по заданих найменуваннями спостережуваних симптомів. Діагностичні системи зазвичай будуються таким чином, що комп'ютер висуває ряд правдоподібних варіантів діагнозу, завдання лікаря - обрати правильний діагноз. Обчислювальна діагностика не передбачає виключення лікаря зі сфери діагностичної діяльності, її головне завдання - створення нових, більш точних і досконалих методів прийняття рішення на основі математичних методів і обчислювальної техніки.

Керуючі системи реалізують збір інформації про об'єкт управління, переробку інформації, передачу даних до органу управління, формування керуючого рішення, видачу керуючого впливу на об'єкт управління. Для збору і переробки інформації до певної міри можуть бути використані ІПС, інформаційно-вимірювальні та діагностичні системи - в цьому сенсі вони є складовою частиною керуючих систем.

Керуючі системи реалізують збір інформації про об'єкт управління, переробку інформації, передачу даних до органу управління, формування керуючого рішення, видачу керуючого впливу на об'єкт управління. Для збору і переробки інформації до певної міри можуть бути використані ІПС, інформаційно-вимірювальні та діагностичні системи - в цьому сенсі вони є складовою частиною керуючих систем.

1.3. Принципи розробки МІС

Розробка МІС - комплексна робота фахівців в області медицини і охорони здоров'я, системотехніки та обчислювальної техніки. При цьому різні групи фахівців виконують різні функції.

До функцій фахівців в області медицини і охорони здоров'я входять:

- формулювання мети (цілей) та основних завдань системи;
- розробка номенклатур методів лікування, діагнозів, ускладнень, ознак (показників), методів дослідження;
- розробка інформаційних моделей (модель хвороби, установи, мережі лікувально-профілактичних установ);
- визначення потоків інформації системи;
- створення стандартизованих документів.

Фахівці з системотехніки та обчислювальної техніки визначають склад технічної частини системи: підсистеми збору та підготовки даних, канали передачі ін-

формації, конфігурація комп'ютера, конфігурація локальної комп'ютерної мережі, підсистеми видачі результатів користувачам. Крім того, системотехніки, математики і програмісти визначають математичні методи, необхідні для вирішення завдань системи, вибирають способи підвищення надійності обробки інформації, розробляють алгоритми вирішення цих завдань, створюють програмне забезпечення (ПЗ), проводять аналіз функціонування та доопрацювання системи.

Розробка будь-якої медичної інформаційної системи повинна спиратися на деякі принципи побудови систем взагалі (вивченням і формулюванням цих принципів займається системотехніка, що іменується також "теорією великих систем", "теорією складних систем").

Принцип формулювання мети системи полягає в тому, що повинна бути чітко сформульована мета побудови та використання конкретної системи, представлена вся послідовність завдань і умов, виконання яких приведе до досягнення мети.

Принцип нових завдань полягає в тому, щоб більшість нових задач, що вирішуються системою, були б принципово нездійсненні або неефективні при їх вирішенні без МІС. Дотримуючись цього принципу, розробка МІС повинна привести до підвищення ефективності вирішення поставленого завдання.

Принцип комплексного (системного) підходу передбачає ретельне узгодження вирішуваних завдань і набору технічних засобів. Складові МІС повинні задовольняти певному набору вимог (точність апаратури, способи представлення даних, типи обчислювальних пристроїв тощо), що забезпечують найбільш раціональне рішення завдань.

Принцип типізації проектних рішень полягає у використанні позитивного досвіду, накопиченого при розробці аналогів, що пройшли успішну перевірку практикою.

Принцип безперервного розвитку системи означає, що в проекті МІС повинні бути передбачені можливості модернізації системи з метою підвищення її ефективності.

Принцип автоматизації документообігу передбачає застосування уніфікованих документів - стандартизованих історій хвороби, висновків. Дотримання цього принципу різко підвищує ефективність системи, полегшує виконання неавтоматизованих ручних операцій, що залишилися, по збору даних.

Принцип єдиної і гнучкої інформаційної бази припускає ефективність оперування даними, що зберігаються в інформаційних масивах, - цей принцип зазвичай реалізують з використанням баз даних і математичного забезпечення, що дозволяє ефективно оперувати цими базами (СКБД).

Принцип комплектності завдань і робочих програм реалізується шляхом створення пакетів прикладних програм (ППП), які об'єднують завдання на основі єдиних методів рішення й взаємозалежної інформації.

Принцип економії ресурсів при введенні, зберіганні та виведенні інформації означає, що ці операції повинні бути організовані таким чином, щоб мінімізувати час і зусилля обслуговуючого персоналу на їх здійснення шляхом викорис-

тання ефективних технічних засобів (сканери, принтери, програми автоматичного розпізнавання текстових і графічних документів і т.д.).

Принцип узгодження пропускних здібностей різних частин системи - наприклад, тимчасові характеристики підсистем обробки даних повинні бути узгоджені зі швидкістю введення-виведення інформації через канали зв'язку.

1.4. Проблеми розробки і впровадження МІС

При розробці МІС першорядне значення має використання знань, які має сучасна медицина. Практично це означає постійні консультації з провідними фахівцями (клініцистами, організаторами охорони здоров'я, гігієністами і т.д.), проведення експертних оцінок всіх рішень, що стосуються питань медицини або організації охорони здоров'я на кожному етапі створення системи.

Розробляючи конкретну діагностичну або прогноуючу систему, необхідно послідовно вирішити ряд питань:

- вибір мети і визначення основного призначення системи;
- вибір структурної схеми системи;
- складання переліку нозологічних форм та збір статистично достовірної інформації про симптоматику станів;
- побудова вирішального правила для вирішення завдань оцінки медичної інформації і видачі висновків з діагностики та прогнозуванню, тобто розробка алгоритмічної основи системи.

Основні проблеми, що виникають при розробці МІС, можна згрупувати наступним чином.

1. Розробка форм документів, зручних для фіксування медичної інформації, пошуку та обробки інформації. Питання стандартизації форм медичних документів. Класифікація даних. Розробка інформаційної медичної мови (використання медичної термінології).

2. Вибір раціональних методів організації медичних даних, що забезпечують ефективний пошук, зберігання, оновлення, вилучення інформації з пам'яті машини.

3. Розробка комплексу програмно-технічних засобів, що забезпечують передачу даних в середині системи, обмін інформацією з зовнішніми системами, аналіз інформації (графічної інформації, результатів лабораторних даних і т.д.).

4. Впровадження та експлуатація МІС. Методичні вказівки по впровадженню системи, порядок заповнення стандартизованих медичних документів, розробка інструкцій з експлуатації МІС.

Проблема впровадження та експлуатації МІС - одна з найбільш складних і гострих, оскільки тільки на практиці можна перевірити дієвість ідей, реалізованих при розробці МІС. Тому завжди слід мати на увазі, що впровадження МІС пов'язане не тільки з отриманням позитивного ефекту, а й з рядом неминучих витрат. Ціна впровадження комп'ютерних технологій в медичну діагностику визначається наступними факторами:

1) додаткові витрати на придбання, встановлення та супровід необхідного програмно-технічного забезпечення;

2) збільшення чисельності персоналу за рахунок необхідності утримання інженерно-технічних працівників;

3) необхідність навчання лікарів, як мінімум, елементам комп'ютерної грамотності і, як максимум, правилам ефективної експлуатації МІС;

4) недосконалість програмно-технічного забезпечення (гіпердіагностика, помилки функціонування програмного забезпечення, збої апаратури ...);

5) психологічний бар'єр з боку лікарів по відношенню до технічних нововведень, особливо із застосуванням комп'ютерних технологій;

б) дуже швидкий моральний і фізичний знос апаратно-програмних засобів і, як наслідок, необхідність витрат на їх модернізацію.

Слід пам'ятати, що впровадження будь-якого різновиду МІС завжди зачіпає сферу діяльності керівництва медичного закладу, лікарів, середнього медичного персоналу. Тому в експлуатацію повинні передаватися абсолютно закінчені і неодноразово перевірені рішення (документи, програми, методики збору, обробки і передачі інформації).

Вкрай небажано "допрацьовувати" при впровадженні, переривати функціонування МІС з причини недосконалості алгоритмів і програм, ненадійною зв'язку, відсутності підготовленого персоналу.

Позитивний ефект від функціонування МІС повинен бути зрозумілий і відчутний для всіх співробітників, що безпосередньо беруть участь в діяльності МІС.

1.5. Інформаційна база МІС

Сукупність даних і знань, використовуваних системою при вирішенні завдань, становить інформаційну базу системи. Складовою частиною інформаційної системи є також способи подання інформації, її зберігання та перетворення.

Розробка інформаційної бази включає наступні розділи:

- визначення сукупності показників, які використовуються в даній МІС;
- класифікація інформації, створення (вибір) класифікаторів і словників системи;
- кодування інформації;
- створення стандартизованих і єдиних для системи форм документів;
- виділення інформаційних зв'язків між розділами системи;
- оцінка об'ємно-часових характеристик інформації;
- організація інформаційних масивів;
- визначення способів контролю інформації.

Інформаційне забезпечення МІС складають: історії хвороби, виписки з історій хвороби, епікризи, стандартизовані карти обстеження, діагностичні та інформативні оцінки показників і станів ("медична пам'ять"), критерії ефективності обстеження та лікування, каталог медичних понять і термінів.

МІС характеризуються наявністю, як правило, великих обсягів даних і знань. Обробка даних і знань зводиться до трьох основних етапів. На першому етапі елементи інформації розміщуються в певних структурах - базах даних і базах знань (БЗ). На другому етапі БД і БЗ піддаються впорядкуванню: змінюється їх структура, порядок розміщення інформації, характер взаємозв'язків між елементами інформації. На третьому етапі здійснюють експлуатацію БД: пошук потрібної інформації, прийняття рішень, редагування баз даних і знань.

1.6. Математичне забезпечення МІС

Математичне забезпечення МІС складають комплекси спеціальних алгоритмів і програм - в цьому сенсі можна говорити також про алгоритмічне забезпечення (АЗ) і програмне забезпечення (ПЗ).

До складу АЗ МІС зазвичай входять такі алгоритми:

- спеціальне введення інформації (тут враховуються особливості задачі, що розв'язується системою, характер інформації, що вводиться);
- спеціальний вивід інформації (наприклад, передбачена подача сигналів тривоги, екстрених повідомлень і т.д.);
- біометрія, статистичний аналіз інформації;
- спеціальний аналіз інформації;
- діагностика;
- прогнозування станів хворого;
- безперервне стеження за станом хворого;
- управління станом хворого;
- управління спеціальною медичною апаратурою;
- обробка документальної інформації;
- управління установою або групою установ;
- рішення адміністративно-оперативних задач.

Програмне забезпечення (ПЗ) МІС являє собою комплекс спеціальних програм і підпрограм, баз даних і баз знань, що реалізують рішення поставлених перед МІС завдань відповідно до заданого комплексу алгоритмів. Склад і основні функції ПЗ МІС широкого профілю показані на рис.1.1.

Відповідно до термінології сучасних комп'ютерних технологій, що реалізуються на базі IBM-сумісних комп'ютерів, ПЗ МІС належить до групи "Прикладні програми". Крім того, слід пам'ятати про існування двох інших важливих груп ПЗ: операційна система (ОС) і службові програми. Прикладами ОС IBM-сумісних комп'ютерів є MS-DOS, Windows, а прикладами службових програм можуть служити такі утиліти DOS як FORMAT.COM, CHKDSK.COM та ін., програми-оболонки Norton Commander, Total Commander, Windows 3.1, Windows 3.11, Far.

Починаючи з Windows 95, операційна система і службові програми об'єднані всередині однієї загальної програми-оболонки. Роль службових програм при цьому виконують, наприклад, засоби швидкого пошуку та перегляду докумен-

тів різних форматів, засоби установки і відключення різних програмних і апаратних компонентів, засоби обміну поштовими і факсимільними повідомленнями, засоби встановлення зв'язку між різними комп'ютерами.



Рис. 1.1

1.7. Організація обробки біомедичної інформації

У різних системах обробки біомедичної інформації використовується два підходи до організації медичних даних: фрагментованість і комплексування. При фрагментованості спільне завдання обробки даних розділяється на окремі частини з метою більш ефективного її вирішення.

При комплексуванні, навпаки, вирішення приватних завдань об'єднуються в більші розділи - таке комплексування доцільно, якщо різні розділи тісно взаємопов'язані і дані, що виводяться, мають спільний додаток.

У практиці розробки МІС знаходять застосування обидва підходи, оскільки дані різних досліджень тісно пов'язані один з одним, а результати обробки використовуються, як правило, для досягнення будь-якої спільної мети: встановлення діагнозу, вибору методу знеболення, методики лікування.

На рис.1.2 представлена схема ієрархічної організації клінічної інформації: масиви первинних даних (відліки ЕКГ, результати аналізів крові, сечі тощо) є

окремими файлами нижнього рівня; після обробки цих відомостей утворюються файли симптомів; потім об'єднуються результати обробки більших розділів (тем) - утворюються файли синдромів.

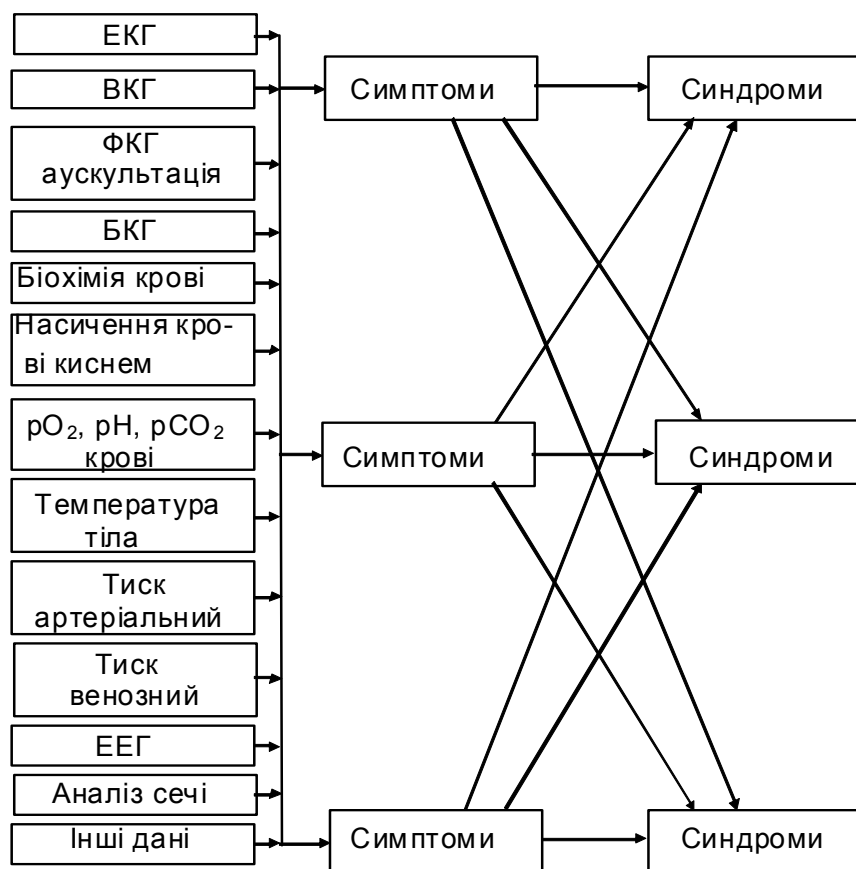


Рис. 1.2

Така організація інформації суттєво використовується в МІС, які вирішують питання комплексування інформації, отриманої з різних інформаційних каналів. Як приклад можна навести широко поширені МІС моніторингу, призначені для роботи в режимі діалогу з медичним персоналом, що обслуговує палати інтенсивної терапії, хірургічні операційні та блоки реанімації. Такі МІС являють собою приліжкові монітори, які працюють автономно або в комплексі з центральним постом медсестри, які мають у своєму складі дисплей і спеціальний реєстратор для роздрукування інформації та реалізують такі функції:

- контроль за станом хворих в декількох десятках палат;
- можливість внесення отриманої інформації в історію хвороби;
- контроль за станом ЕКГ, кров'яного тиску, дихання, температури, електроенцефалограми і т.д.

Різновидом систем такого роду є мікропроцесорна система домашнього та віддаленого моніторингу, описана в розділі 7.1, ключовою ланкою якої є "електронний комбайн" - спеціалізована ЕОМ, що займає проміжне місце (за функціональними можливостями) між добре відомими телефонними апаратами з автоматичним визначенням номера (АВН) і персональними комп'ютерами.

Ще одним прикладом фрагментування і комплексування інформації є технологія Київського лікувально-діагностичного та реабілітаційного центру "ЛИДИР" (див. розділ 7.2.1), що базується на об'єднанні в рамках єдиної локальної комп'ютерної мережі чотирьох автоматизованих робочих місць (АРМ'ів):

- АРМ іридолога (загальна оцінка стану органів і систем організму людини, виявлення спадкової схильності до захворювань, доклінічна діагностика);
- АРМ фахівця УЗД (приватний висновок і рекомендації фахівця-УЗД);
- АРМ фахівця-кардіолога (приватний висновок і рекомендації кардіолога);
- АРМ терапевта (інтегрально-аналітична оцінка стану здоров'я; паспорт здоров'я; рекомендації з дообстеження і консультації; рекомендації з лікування).

Збільшена схема організації обробки медичної інформації в часі досить тривіальна і являє собою послідовність трьох фаз: отримання вихідної інформації, обробку інформації згідно певної програми і видачу результатів. Ця схема справедлива для будь-якого рівня вищенаведеної ієрархічної структури.

Якщо під МІС розуміють форму організації діяльності в медицині та охороні здоров'я, що об'єднує медиків, математиків, інженерів, техніків з комплексом технічних засобів (визначення МІС в широкому сенсі), тоді можна розрізнити МІС наступних рівнів: державного, регіональних (обласних), районних, міських і селищних, клінічних, медичних центрів різного рівня і профілю. При цьому істотний інтерес представляє проблема взаємодії фахівців різного профілю в рамках МІС. Як приклад такої взаємодії, на рис.1.3 представлена функціональна схема клінічної МІС.



Рис. 1.3

1.8. Автоматизація медичної діагностики

Вже на перших етапах освоєння і розвитку обчислювальної техніки висловлювалися припущення про можливість їх застосування в медичній діагностиці. У 1962 році на конференції Американської кардіологічної асоціації відбулася

перша публічна демонстрація системи автоматичної діагностики електрокардіограм із застосуванням ЕОМ. А наприкінці 60-х років такі системи вже практично використовувалися різними комерційними організаціями США.

У СРСР початок машинної діагностики також датується 1962р.: В Інституті хірургії ім. А.В.Вишневського ЕОМ використана для виявлення вроджених вад серця (пізніше там же досліджувалася можливість машинного діагностування механічної жовтяниці, стенозу соска Фатера, раку жовчного міхура), а в м. Мінську на Першій Білоруської наукової конференції з кібернетики зроблено доповідь про застосування ЕОМ для діагностики пухлин головного мозку.

Чим викликаний інтерес лікарів до діагностики із застосуванням ЕОМ? На думку ініціаторів та наукових організаторів першої в Радянському Союзі лабораторії медичної кібернетики А.А. Вишневського та М.Л. Биховського, застосування ЕОМ спрямоване на усунення наступних недоліків діагностичного процесу, здійсненого лікарем:

- діагноз рідкісного захворювання може виявитися невстановленим внаслідок незнайомства клініциста з даною нозологічною формою;
- лікар не завжди може згадати те, що зберігається в його пам'яті;
- важко ефективно переробити значний обсяг інформації, яка збирається про хворого;
- лікар досить швидко втомлюється, внаслідок чого якість його діагностики погіршується.

У цьому випадку цікавий досвід, проведений американським кардіологом Ц.Каересом: 50 кардіограм, попередньо проаналізованих експертами, були спрямовані 6 кардіологам і 4 терапевтам - виявилось, що лише в 50% випадків їх висновок співпадав з думкою експертів.

А як справи з достовірністю комп'ютерної діагностики? За даними літератури, надійність комп'ютерної діагностики досягає 70-90%, а часом буває і вище. Наприклад, перевіряючи ефективність експертної системи MYCIN-TEIREISAS, п'ятеро фахівців, що не входять в один науковий колектив, провели 15 сеансів випробувань системи, під час яких перевірялися правильність лікувальних рекомендацій, доречність заданих системою питань, наявність незаданих, але важливих питань. У 72% випадків експерти висловили згоду з системою, в більшості інших випадків вони самі не змогли прийти до єдиної думки.

Переваги машинної діагностики особливо наочні при розпізнаванні рідкісних захворювань. Так, стеноз соска Фатера майже ніколи не встановлюється лікарями до операції. Застосування ЕОМ дозволяє в 6 з 11 випадків розпізнати це захворювання.

Крім того, попередній машинний діагноз часто дозволяє уникнути хірургічних ("кривавих") методів дослідження, часом зі смертельним результатом. Наприклад, при діагнозі вроджених вад серця вдається уникнути зондування та ангіографії, досягаючи 86% правильного діагнозу, що всього лише на 5-7% нижче, ніж із застосуванням цих методів.

Разом з тим, вирішення проблеми взаємодії фахівців-медиків з ЕОМ неодмінно має включати облік професійних особливостей медичного персоналу, існування відомого психологічного бар'єру, породженого рядом причин. Це, перш за все, мала компетенція (або відсутність такої) в питаннях можливостей ЕОМ і відсутність практичного досвіду використання обчислювальної техніки. У той же час у ряду фахівців-медиків виникає хибне уявлення про те, що обчислювальна техніка "все може", а використання її в медицині і охороні здоров'я не представляє особливих труднощів. До обставин, що утрудняє впровадження діагностичних систем, можна віднести також небажання доручати машині вироблення рішень в силу виняткової відповідальності лікаря за результати діагностики та лікування. Це - наслідок нерозуміння того, що лікарський висновок є і залишиться завжди остаточним. Нарешті, в ряді випадків істотну роль можуть грати міркування престижу і авторитету, коли фахівці-медики бояться проявити низьку професійну кваліфікацію в порівнянні з ЕОМ.

1.8.1. Задача медичної діагностики

Задача медичної діагностики відноситься до класу так званих "зворотніх" задач, оскільки практично всяке захворювання проявляється у вигляді тих чи інших ознак. Природно припустити можливість прийняття рішення про наявність захворювання по наявності, ступеня вираженості і поєднанню спостережуваних ознак.

Термінологічно медичні діагностичні ознаки обозначають по-різному: знаки, ознаки, симптоми, синдроми.

Розрізняють симптоми суб'єктивні, засновані на описі хворими своїх відчуттів (відчуття болю, тяжкості і т.д.), і об'єктивні, отримані при обстеженні хворого (рентгенологічна ознака "ніші" при виразці шлунка).

Крім того, розрізняють симптоми клінічні та доклінічні. Клінічні - ознаки наявності функціональних або органічних змін. Доклінічні - ознаки ймовірного розвитку захворювання внаслідок генетичної слабкості органу / системи або несприятливих умов життя і праці.

Можна розрізняти ознаки за ступенем "складності". До "елементарних" ознак відносять симптоми, до "комплексних" - синдроми. Синдроми, як поєднання симптомів, зумовлені єдиним патогенезом, найчастіше розглядають як самостійне захворювання (наприклад, синдром Мен'єра) або як стадію (форму) яких-небудь захворювань (наприклад, нефротичний синдром, уремія при хронічному нефриті).

Нарешті, найчастіше ознаки групують як основні та додаткові. Так, наприклад, для методу іридодіагностики основними є такі ознаки як колір, щільність структури райдужної оболонки, зміщення зіниці, наявність і місце розташування пігментних і токсичних плям та ін, а додатковими - результати анамнезу, фізіогномічного аналізу, аналізу результатів співбесіди з пацієнтом, спрямованого на облік суб'єктивних ознак.

Термінологічно захворювання також позначають по-різному: захворювання, хвороба, нозологічна форма. Хворобою називають порушення нормальної життєдіяльності організму, обумовлене функціональними або (і) морфологічними змінами. Виникнення хвороби обумовлено генетичними дефектами організму, способом життя і умовами життя. Нозологічною формою іменують певне захворювання, що виділяється на основі встановленої причини (етіологія), особливостей розвитку (патогенез), типових зовнішніх проявів і характерного ураження органів і тканин. Очевидно, слід розрізняти найменування захворювання та його опис (нозологічну форму).

За захворювання також класифікують, об'єднуючи їх у групи і підгрупи по ієрархічній схемі (рис.1.4).

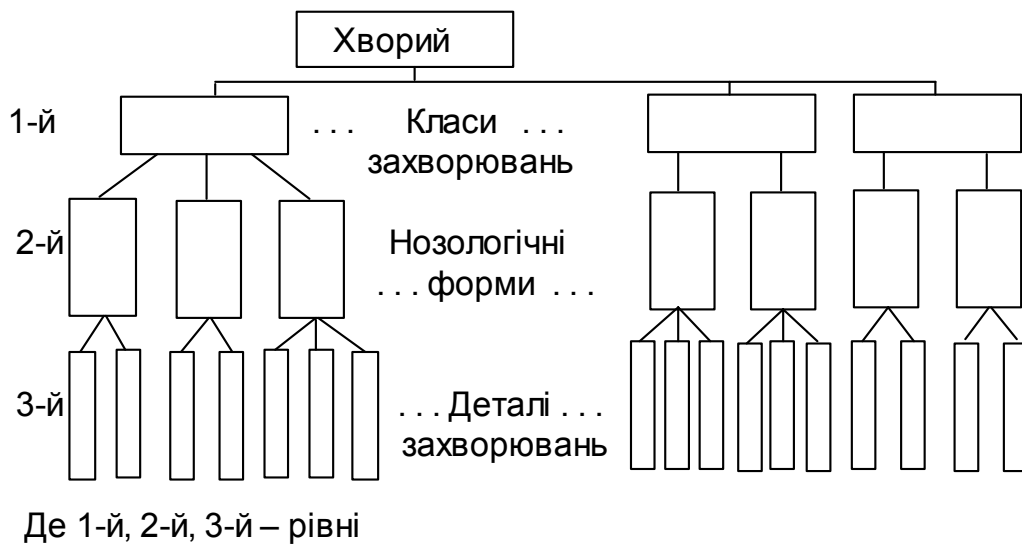


Рис. 1.4

Такій схемі організації підпорядкований, наприклад, міжнародний класифікатор хвороб МКБ-10, фрагмент якого має наступний вигляд:

- Інфекційні та паразитарні хвороби (кл.І)
- Новоутворення (кл.ІІ)
- Хвороби крові та кровотворних органів (кл.ІІІ)
- Хвороби ендокринної системи, порушення обміну речовин і імунітету (кл.ІV)

- Кл.І
- Кишкові інфекції (А00-А09)
- Туберкульоз (А15-А19)
- Бактеріальні зоонози (А20-А28)

- (А00-А09)
- Холера (А00)

Черевний тиф і паратифи (A01)
Інші сальмонельозні токсикоінфекції (A02)
Шигельоз (A03)
.....

A01:

Черевний тиф (A01.0)
Паратиф (A01.1)
Паратиф (A01.2)
Паратиф (A01.3)
Паратиф неуточнений (A01.4)

A03:

Шигельоз викликаний *Shigella dysenteriae* (A03.0)
Шигельоз викликаний *Shigella flexneri* (A03.1)
Шигельоз викликаний *Shigella boydii* (A03.2)
Шигельоз викликаний *Shigella sonnei* (A03.3)
Інші шигельози (A03.8)
Неуточнений шигельоз (A03.9)

Таким чином, завдання діагностики можна сформулювати як виділення з можливих станів хворого стану (або комбінацію декількох, розглянуту як окремий стан), в якому знаходиться даний хворий. У поняття "стан" можна вкладати ту чи іншу нозологічну форму даного класу захворювань, коли є необхідність діагностики всередині класу захворювань. Але під терміном "стан" можна розуміти і весь клас захворювань. У такому випадку це буде попередня діагностика. Наприклад, коли потрібно з'ясувати, чи відноситься даний хворий до групи хворих пухлинами головного мозку, гострими порушеннями мозкового кровообігу або спадковими захворюваннями нервової системи, цей процес кваліфікується як зсув рівня діагностики "вгору". Подібним чином можна зміщувати рівні діагностики і "вглиб" - у бік більшої деталізації стану хворого. Наприклад, у якості різних станів можна взяти різні ступені прояву або "відтінки" своєрідності даного захворювання. За таким принципом проводиться детальна (диференціальна) діагностика всередині даного захворювання по різних ступеням тяжкості та ін.

1.8.2. Види медичної діагностики

Розглянемо наступні види діагностики: загальну, диференціальну, приватну (системну) і прогностичну діагностику.

Загальна та диференціальна діагностика подібні в тому сенсі, що в обох випадках за заданими ознаками необхідно визначити вид захворювання (рис. 1.5).

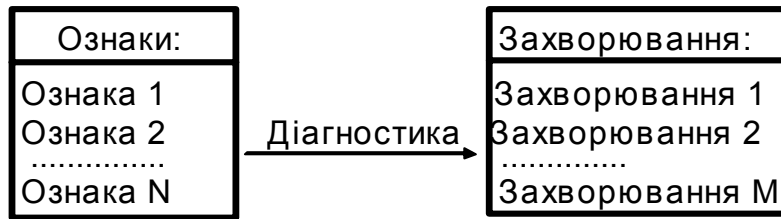


Рис. 1.5

Різниця між ними в тому, що загальна діагностика "обслуговує" верхні рівні ієрархічної класифікаційної структури захворювань, показаної на рис.1.4 - вибір класів захворювань і вибір нозологічної форми. У цьому сенсі диференціальна діагностика - задача "низького" рівня, присвячена з'ясуванню деталей виявлених захворювань.

В результаті формальна, тобто чисто зовнішня, відмінність між цими задачами проявляється в тому, що перелік діагностованих захворювань при диференціальній діагностиці, як правило, істотно менший ніж при загальній діагностиці.

Мета приватної (системної) діагностики - застосування деякої стратегії для перевірки єдиної гіпотези на тлі інших правдоподібних гіпотез. Таку діагностику можна представити у вигляді двох етапів: на першому етапі для заданого захворювання формують (обирають) відповідний перелік ознак, а на другому етапі для обраного переліку ознак вирішують задачу власне діагностики, яка подібна до задач загальної або диференціальної діагностики (рис.1.6).

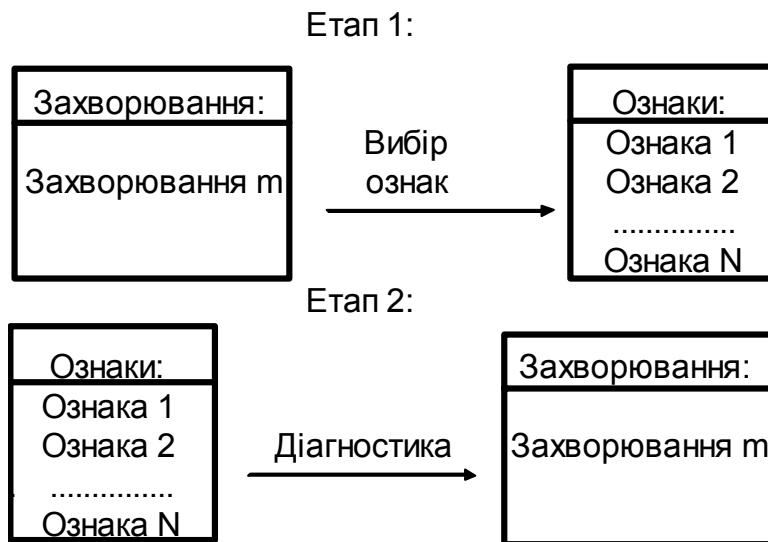


Рис. 1.6

Прогностична діагностика націлена на визначення схильності до захворювання при відсутності клінічних ознак (у таких випадках говорять про "донозологічну", "доклінічну" або "ранню" діагностику), - при цьому використовують спеціальні системи діагностичних ознак, - або на передбачення розвитку захворювання за наявними клінічними ознаками. Формально цей вид діагностики подібний з задачами загальної та диференціальної діагностики - принципова різниця полягає в

тому, що при прогнозуванні факти спостереження ознак і можливого прояву захворювання рознесені в часі, тоді як при інших видах діагностики ці факти віднесені до одного і того ж моменту часу.

Приклади опису змістовної частини всіх перерахованих видів діагностики знаходимо в різній медичній літературі. Наприклад, у монографії Е.С.Вельховера "Клінічна іридологія" змістом розділу "іридодіагностична семіологія" є опис комплексів іридогенетичних синдромів, топічних і тополабільних ознак, яка в стислій формі може бути представлено таблицею відповідності "ознака → захворювання" і використано для загальної та диференціальної діагностики. У розділі "Системна іридодіагностика" іридоознаки систематизовані по захворюваннях деяких найважливіших систем і органів: нервової, серцево-судинної та гепатобіліарної систем, шлунково-кишкового тракту, нирок, при цьому для опису ознак використана схема "захворювання → ознака", зручна при приватній діагностиці. Нарешті, в розділах "Профілактична іридодіагностика" і "Зміни райдужки в динаміці", де наведені оцінки можливості раннього, доклінічного розпізнавання захворювань при іридодіагностичних оглядах, а також можливості контролю ефективності лікувальних заходів, фактично використано модель прогностичної діагностики.

1.8.3. Діагностичні алгоритми

Існують різні алгоритми (вирішальні правила) медичної діагностики. Найпростішим алгоритмом, який часто не вимагає застосування ЕОМ, є так званий "табличний алгоритм", що базується на використанні спеціальних діагностичних таблиць, які визначають відповідність між ознаками що спостерігаються і відповідними їм захворюваннями.

В якості прикладу в табл. 1.2 наведено фрагмент діагностичної таблиці для спільної іридодіагностики.

Зауважимо дві особливості наведеної діагностичної таблиці. По-перше, вона не націлена на постановку остаточного діагнозу: "ознаки епілепсії", але не "епілепсія" і т.д. По-друге, вона дозволяє вирішувати завдання не тільки загальної діагностики, але і прогнозування: "схильність до ...". Зазначені особливості обумовлені, з одного боку, бажанням уникнути гіпердіагностики, а з іншого боку - специфічними властивостями іридодіагностики як методу (ознаками служать результати не прямих, а опосередкованих спостереженнях стану органів і систем людського організму).

У табл. 1.1 наведено ще один приклад - діагностична таблиця для диференціальної діагностики геморагічного та ішемічного інсульту. Особливість табл. 1.1 - зв'язок між захворюваннями і відповідними їм ознаками описана у вигляді "нечітких" формулювань: не "є" або "немає", а "характерно", "малохарактерно" і т.д. Це дозволяє, з одного боку, також уникнути гіпердіагностики. А з іншого боку, така форма таблиці може служити першим кроком на шляху до реалізації імовірнісного підходу до вирішення задачі діагностики - дійсно, "характерно"

означає "зустрічається дуже часто", "малохарактерно" означає "зустрічається рідко" і т.д.

Таблиця 1.2. Загальна іридодіагностика (фрагмент)

ОЗНАКА	ЗАХВОРЮВАННЯ
ЗІНИЦЯ, ЩО "СКАЧЕ"	Ознаки вегетосудинної дистонії
	Ознаки депресивного стану
	Ознаки епілепсії
	Онконасторожений. в області головн. мозку
	Схильність до розсіяного склерозу
	Схильність до менінгіту
	Виснаження НС
ЗІНИЦЯ НЕРУХОМА	Можливо параліч окорухового нерва
	Тенденція до сухотки спинного мозку
	Можлива втрата функцій зору
АНИЗОКОРІЯ З ТЕНДЕНЦІЄЮ ДО МІОЗУ	Можливий ранній розвиток цук. діабету
	Схильність к акромегалії
	Схильність до хвороби Іценко-Кушинга
	Схильність до вузликowego періартриту
АНИЗОКОРІЯ З ТЕНДЕНЦІЄЮ ДО МІДРІАЗУ	Схильність до глистяної інвазії
	Схильність до гіповітамінозу РР
	Можливо отруєння свинцем
	Схильність до ТБС
	Можливо плеврит
	Схильність до неврастенії
	Схильність до пошкодження аорти
	Тенденція до сухотки спинного мозку

Маючи в своєму розпорядженні такого роду діагностичну таблицю, лікар по вихідному переліку спостережуваних ознак формує результуючий перелік - список можливих захворювань. Зрозуміло, на цьому діагностика не завершується - зазвичай буває необхідно також:

- виділити найбільш важливі гіпотези про наявність захворювання;
- переконатися в достовірності висунутих гіпотез (за додатковими даними - результатами додаткових аналізів, обстежень та консультацій).

Відомі й інші, потенційно більш ефективні, види медичних діагностичних алгоритмів:

- імовірнісний метод, що складається в обчисленні ймовірностей захворювань за формулою Байєса;
- метод послідовного статистичного аналізу Вальда;

- метод пошуку клінічного прецеденту (іменований іноді "статистичним методом");
- метод фазового інтервалу, заснований на поділі багатовимірного розподілу ознак на зони, відповідні здоровому і хворому організму;
- метод логічного базису - відрізняється від методу фазового інтервалу використанням апарату математичної логіки для метризації простору ознак.

Ці алгоритми більш трудомісткі і зазвичай реалізуються з використанням ЕОМ. Крім того, із застосуванням ЕОМ реалізуються діагностичні алгоритми на базі моделей ансамблів штучних нейронів - нейронних мереж.

1.8.4. Достовірність діагностики з використанням МІС

Наведемо цікаву цитату з книги М.С.Абрамова "Гридодіагностика з комп'ютером": "Автоматичні системи можуть забезпечувати 100% точність діагностики. Це можливо при виконанні важливої умови: не допускати помилок в складаються програмах, в управлінні роботою ЕОМ та інтерпретації отриманих результатів".

Небезпека подібного роду заяв полягає в тому, що у неспеціалістів в області теорії розпізнавання образів може створитися оманливе враження, ніби діагностика з 100%-вою достовірністю - задача легко досяжна, варто лише "навалитися" на її рішення всією сучасною науково-технічною міццю.

Тому зазначимо, що принципова особливість будь, навіть з використанням потужних технічних засобів, діагностики: досягти 100%-вої достовірності діагнозу практично неможливо в силу принципово складних причин як об'єктивного, так і суб'єктивного характеру.

До причин об'єктивного характеру відносяться індивідуальні особливості організму, нетиповий перебіг хвороби, помилки вимірювання ознак, перекриття областей простору ознак.

Причинами суб'єктивного характеру є помилки лікаря, зумовлені неправильною організацією обстежень, невірним тлумаченням результатів обстежень внаслідок втоми, психофізіологічних особливостей або некомпетентності лікаря. Це джерело помилок можна для стислості називати "людським фактором".

Про "немедичні" причини об'єктивного характеру можна коротко сказати наступне.

При проведенні обстежень з використанням технічних засобів неминуче виникають помилки оцінювання ознак, обумовлені наявністю:

- перешкод зовнішніх (наведення, вібрації тощо);
- перешкод внутрішніх (шуму, нестабільність параметрів і т.д.);
- апаратних похибок вимірювань (систематичні і випадкові складові похибки, властиві обраному способу технічної реалізації даного методу вимірювань);
- методичних похибок вимірювань (систематична і випадкова складові).

Перекриття областей простору ознак, що характеризують різні захворювання, призводить до необхідності додаткової діагностики - фактично це означає не-

обхідність збільшення кількості діагностичних ознак. Однак навіть збільшуючи кількість діагностичних ознак, не завжди вдається домогтися поліпшення якості діагностики. Більш того, збільшення кількості діагностичних ознак ускладнює процедуру прийняття рішення - в теорії розпізнавання образів цей феномен називають "прокляттям розмірності" (мається на увазі надмірно висока розмірність простору ознак, утрудняє прийняття рішення навіть з використанням потужних обчислювальних засобів).

1.9. Експертні системи: концепції

Історію розвитку напрямку "експертні системи" коротко можна представити таким чином.

В кінці 70-х років фахівцями в галузі інформатики, іменованої "штучний інтелект" (ІІ), була сформульована концепція: ефективність комп'ютерної програми при вирішенні задач залежить від знань, якими вона володіє, а не тільки від формалізмів і схем виводу, які вона використовує. Реалізація цієї концепції призвела до розвитку спеціалізованих програмних систем, кожна з яких є "експертом" в деякій вузькій предметній області - ці програми отримали назву експертних систем (ЕС).

Накопичення та організація знань - одна з найважливіших характеристик ЕС. Наслідки такого підходу виходять за межі побудови програми, призначеної для вирішення деякого класу задач. Причина в тому, що знання - основа експертних систем - є явними і доступними, що і відрізняє ці системи від більшості традиційних програм.

Базова структура ЕС, показана на рис.1.7, демонструє наявність у інтерфейсу такої комп'ютерної програми двох основних функцій: здобувати знання в експерта і вести діалог з користувачем, тобто давати поради, пояснення, ставити запитання користувачеві.

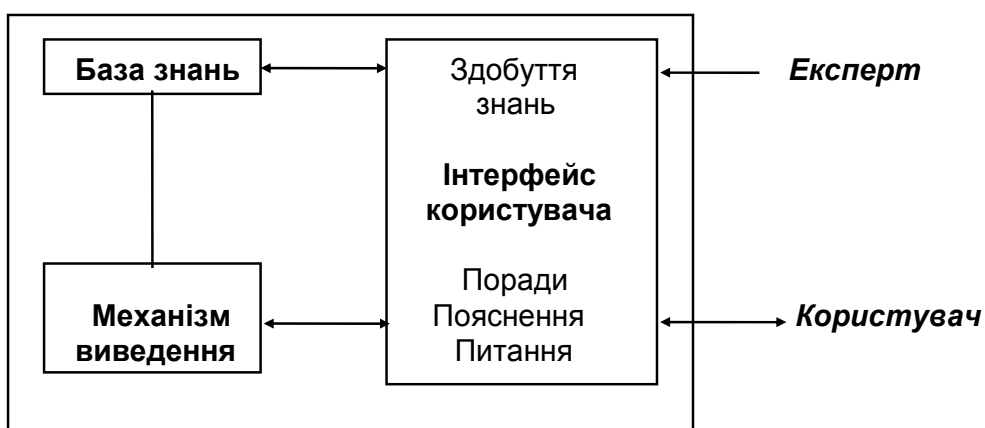


Рис. 1.7

Найбільш корисною характеристикою експертної системи є те, що вона застосовує для вирішення проблеми високоякісний досвід. Цей досвід може представляти рівень мислення найбільш кваліфікованих експертів у даній області, що ве-

де до творчих, точних і ефективних рішень. За визначенням Пола І. Джонсона, "Експерт - це людина, яка завдяки навчанню і досвіду може робити те, що ми всі, решта людей, робити не вміємо; експерти працюють не просто професійно, але до того ж впевнено і ефективно. Експерти ... вміють швидко переворухити масу несуттєвої інформації, щоб дістатися до головного, і добре вміють розпізнавати в проблемах, з якими стикаються, приклади тих типових проблем, з якими вони вже знайомі".

Іншою важливою характеристикою експертної системи є її гнучкість – під цим розуміють високу модернізаційну здатність бази знань експертної системи. Система може збільшуватися поступово відповідно до потреб замовника або користувача. Це важливо і в економічному плані, оскільки означає, що у створенні системи можна спочатку вкласти скромні кошти, а потім нарощувати її можливості по мірі необхідності.

Технологія побудови експертних систем отримала назву "інженерії знань". Особливість цієї технології полягає в тому, що якийсь фахівець, іменованій "інженером по знаннях" і що є, по суті, системотехніком, здійснює посередницьку діяльність між експертами у цій проблемній області, з одного боку, і програмістами (чи програмними засобами), які вкладають знання експертів в створювану програму (створюють базу знань), з іншого.

Очікувалося, що створений в результаті такої колективної діяльності програмний продукт-інструмент для оперування знаннями, що зберігаються в базі знань (БЗ), дозволить:

- підтримувати фахівців при вирішенні складних задач, що вимагають значних інтелектуальних зусиль - задач інтерпретації, діагностики, прогнозування, планування, проектування та управління;
- зберігати і накопичувати досвід дуже кваліфікованих фахівців;
- навчати і тренувати фахівців.

Великі надії на експертні системи різного призначення покладалися як потенційними "користувачами", так і фахівцями в галузі комп'ютерних технологій. На жаль, втрату почуття міри в оцінці потенційних можливостей експертних систем страждають не тільки недосвідчені користувачі, а й фахівці. У цьому зв'язку в передмові до книги Д. Уотермена "Керівництво експертними системами" В.Л. Стефанюк пише: "У нас створено чимало експертних систем різного призначення ... Великий інтерес викликають відповідні семінари та симпозиуми. Однак ця популярність породила і деякі негативні компоненти ... Багато звертаються до "модних" експертних систем, не маючи на те серйозних підстав, не проаналізувавши, чи можливе створення для їх задач адекватних експертних систем ". Далі для демонстрації складності побудови експертних систем В.Л. Стефанюк наводить приклад однієї з найостанніших американських експертних систем - системи Syntel, призначеної для використання у світі фінансів і бізнесу. У її розробці брали участь найбільш авторитетні вчені (Р. Дуда, П. Харт, Р. Ребо, Дж. Рейтер, Т. Ріш), досвід кожного з яких достатній для того, щоб очо-

лити самостійний аналогічний проект - проте вони все ж вважали за доцільне об'єднати свої зусилля.

Нарешті, в 1995р. у статті Тома Давенпорта з промовистою назвою "Епітафія експертним системам" підводяться своєрідні підсумки більш ніж двадцятирічного циклу розвитку експертних систем:

- припинено розробку ЕС Xcon фірми Digital Equipment (ЕС для вибору конфігурації комп'ютерів);
- відставлена стратегія фірми DuPont (побудова великого числа дрібних додатків);
- припинили існування начмнаючі фірми Symbolics і Pal-ladian.

Аналіз, проведений автором статті, свідчить, що головні проблеми виникають при впровадженні експертних систем:

- багатьом ЕС потрібна спеціальна платформа (апаратне забезпечення), перенесення їх на універсальні платформи призводить до зниження продуктивності;
- громіздкі ЕС важко керуючі - якщо вони містять десятки тисяч правил і для їх управління необхідно кілька операторів, тоді їх не тільки складно адаптувати до конкретних організацій або ситуацій, але і їх структуру ще й важко освоїти;
- повільність деяких ЕС - з цієї причини персонал іноді відмовлявся використовувати ЕС;
- необхідність суттєвих організаційних змін при впровадженні ЕС - ця вимога виявилася неприйнятним для менеджерів більшості компаній.

Висновки Т. Давенпорта:

- технологія не прийметься, якщо вимагає нової платформи, серйозних змін у кадрах і у виробничому процесі;
- комп'ютери гарні для накопичення знань, але не для заміни спеціалістів;
- невгамовний і перебільшений захват рідко виправдовуються згодом.

Незважаючи на те, що Давенпорт безпосередньо не торкався медичної проблематики, спільність його висновків дозволяє вважати їх справедливими і для експертних систем медичного призначення.

1.10. Експертні системи медичного призначення

Медичні ЕС розробляються для вирішення різного роду проблем, але основні типи їх діяльності можна згрупувати в категорії, наведені в табл. 1.3

Експертні системи, що виконують інтерпретацію, як правило, використовують для опису ситуації інформацію від сенсорів спостерігаючих систем (наприклад, значення пульсу, кров'яного тиску тощо), щоб встановити діагноз або тяжкість захворювання. Нижче наведено типові правила для медичної ЕС системи SPE, яка інтерпретує свідчення скануючого денситометра для різних випадків запальних процесів у хворих.

ЯКЩО: профіль щільності є "асиметричний гамма-пік"
і площа під кривою відповідає віковій нормі,

ТО: концентрація гаммаглобуліну в межах норми.

Таблиця 1.3. Типові категорії способів застосування медичних ЕС

Категорія	Проблема, що вирішується
Інтерпретація	Опис ситуації по інформації, яка надходить від сенсора
Діагностика	Виявлення причин захворювання
Прогноз	Визначення ймовірних наслідків даних ситуацій
Лікування, реабілітація	Виконання послідовності запропонованих дій, спрямованих на приведення до норми
Навчання	Діагностика і корекція поведінки учня
Планування	Визначення послідовності дій
Управління	Управління станом об'єкта

Діагностичні системи служать для визначення причин захворювання за симптомами, які спостерігаються у пацієнтів, часто є консультантами, які допомагають за результатами діагностики зробити прогноз, підібрати курс лікування або реабілітації. Вони можуть взаємодіяти з користувачем, щоб надати допомогу при діагностиці, а потім запропонувати порядок дій з проведення лікувальних заходів. У медичній галузі було розроблено більше діагностичних систем, ніж у будь-якій іншій окремо взятій предметній області.

Приклад правила діагностичної системи MYCIN для діагностики бактеріальної інфекції у госпіталізованого хворого наведено нижче.

- ЯКЩО:** 1) Забарвлення грам-позитивних бактерій
 2) Морфологія бактерій характерна для коків
 3) Форма колоній - ланцюжки,

ТО: Є підстави вважати (з ймовірністю 0.7), що вид бактерій - стрептокок.

Експертні системи, що здійснюють прогноз, визначають ймовірні наслідки заданих ситуацій. Відзначається, що фахівці ШІ поки що розробили порівняно мало прогнозуючих систем, тому що такі системи найчастіше базуються на використанні імітаційного моделювання - між тим задача реалізації імітаційного моделювання досить непроста.

Проте, в тих випадках, коли таке моделювання необов'язкове, розробка експертних систем цілком реальна. Так, наприклад, пропонуються діагностичні карти для прогнозування мозкових інсультів і гострих інфарктів міокарда. У цих картах містяться переліки факторів ризику, ступінь їх вираженості і відповідні оцінки вкладів кожного фактора в загальну суму балів, на підставі якої робиться висновок про ймовірність захворювання в найближчі кілька років. Використовуючи такі карти, зовсім неважко розробити відповідну ЕС. Одним з правил такої ЕС може бути наступне.

ЯКЩО: Спостерігається почастищення або поява нападів стискаючої болі (стенокардія),

ТО: Ризик виникнення інсульту або гострого інфаркту міокарда становить 2 бали (при пороговому значенні 9 балів загальної суми балів)

Експертні системи для лікування та реабілітації часто містять діагностичні компоненти, як, втім, і діагностичні системи містять компоненти для видачі рекомендацій з лікування. Це цілком зрозуміло, оскільки процеси діагностики та лікування взаємопов'язані. Приклад правила експертної системи ONCOCYN, що допомагає лікувати онкологічних хворих, що піддаються хіміотерапії, наведено нижче. Хворий з лімфою піддається хіміотерапії з використанням адрианоміцину.

ЯКЩО: 1) Хворий отримав хіміотерапію та
2) Показники крові дають підстави зменшити дозу,

ТО: Висновок, що поточна зменшена доза є попередньої дозою, зменшеною на мінімум з величини зменшення дози, обумовленої зниженням числа білих клітин крові, і величини її зменшення, зумовленої зниженням числа кров'яних пластинок.

Навчальні системи створюють модель того, що той хто навчається знає як він ці знання застосує до вирішення проблеми. Системи діагностують і вказують учню його помилки, аналізуючи модель і будуючи плани виправлень зазначених помилок. Приклад правила навчальної експертної системи GUIDON, яка навчає студентів-медиків правилам вибору антибактеріальної терапії інфекційних хворих, наведено нижче. Правило містить міркування про медичні (предметні) правила, які міг вивчити студент, і модифікує ступінь впевненості в тому, що він дійсно їх вивчив.

ЯКЩО: 1) Ви вважаєте, що студент розглянув конкретне правило предметної області, і
2) Це правило видає значення, яке співпадає зі значенням, запропонованим студентом, і
3) Ви припускаєте, що студент розглянув якесь інше правило, яке приводить до того ж значення,

ТО: Збільшити сумарну впевненість у тому, що студент розглянув це правило, на 0.40.

Плануючі ЕС визначають повну послідовність дій, перш ніж почнеться їх виконання. Прикладами можуть служити планування курсу лікування або ходу операції. Вони найчастіше повинні мати здатність до повернення, тобто відки-

дати деяку послідовність міркувань чи частина плану через порушення обмежень завдання і повертати управління назад до більш ранньої точки або ситуації, з якої аналіз повинен початися заново. Повернення може дорого коштувати, особливо при вирішенні задач медичної проблематики, де ціною помилки може бути життя пацієнта, тому в деяких плануючих системах задачі планування розбивається на підпроблеми і робиться спроба впорядкувати їх так, щоб уникнути перепланування.

Адаптивні ЕС керують станом об'єкта - наприклад, станом хворих при інтенсивній терапії. Тому вони повинні включати компоненти, що інтерпретують (спостерігають) системи, щоб відстежувати стан об'єкта протягом часу, але вони можуть мати потребу також і в інших компонентах для виконання будь-яких або всіх з розглянутих вище типів задач: діагностики, прогнозування, навчання, планування.

Два приклади правил з управляючої експертної системи VM наведені нижче.

Система управляє доглядом за оперованими хворими при інтенсивній терапії. Правило 1 (застосовне до хворих, які знаходяться на режимах штучного дихання VOLUME, CMV ASSIST і T-PIECE).

- ЯКЩО:** 1) Серцевий ритм є прийнятним, і
2) Частота пульсу не змінилася більш ніж на 20 ударів / хв. протягом 15 хвилин, і
3) Середній артеріальний тиск є прийнятним, і
4) Середній артеріальний тиск не змінилося більш ніж на 15 мм рт.ст. протягом 15 хвилин, і
5) Систолічний кров'яний тиск є прийнятним,

ТО: Гемодинаміка є стійкою.

Правило 2 (застосовується до всіх хворим, які знаходяться на режимі контрольованої примусової вентиляції - CMV).

- ЯКЩО:** 1) Хворий переведений з VOLUME на CMV, або
2) Хворий переведений з ASSIST на CMV,

ТО: Очікується наступне:

- 1) Прийнятне середній тиск 75 - 110;
- 2) Прийнятний серцевий ритм 60 - 110;
- 3) Зміст CO₂ в видиху 28 - 42;

Правило 1 ілюструє, як управляюча система стежить за станом хворого, застосовуючи терміни, значення яких змінюються в залежності від клінічного контексту (наприклад, "прийнятно"). Правило 2 показує, як система використо-

вує свої очікування в деякому контексті, щоб точно визначити значення термінів на кшталт "прийнятний середній артеріальний тиск".

Короткі висновки

Медична інформаційна система у вузькому значенні - це називають комплекс технічних засобів і математичного забезпечення, призначений для збору, аналізу медико-біологічної інформації та видачі результатів в зручному для користувача вигляді.

Залежно від виду вирішуваних завдань МІС можна розділити на три групи: інформаційно-довідкові, інформаційно-логічні, керуючі.

Розробка будь-якої медичної інформаційної системи повинна спиратися на деякі принципи побудови систем, вивченням і формулюванням яких займається системотехніка.

При розробці МІС першорядне значення має використання знань, які має сучасна медицина.

Розробляючи конкретну діагностичну або прогнозуючу систему, необхідно послідовно вирішити ряд питань:

- 1) вибір мети і визначення основного призначення системи;
- 2) вибір структурної схеми системи;
- 3) складання переліку нозологічних форм та збір статистично достовірної інформації по симптоматиці станів;
- 4) побудова вирішального правила для вирішення завдань оцінки медичної інформації і видачі висновків з діагностики та прогнозування, тобто розробка алгоритмічної основи системи.

МІС характеризуються наявністю, як правило, великих обсягів даних і знань. Обробка даних і знань зводиться до трьох основних етапів. На першому етапі елементи інформації розміщуються в певних структурах - базах даних і базах знань (БЗ). На другому етапі БД і БЗ піддаються впорядкуванню: змінюється їх структура, порядок розміщення інформації, характер взаємозв'язків між елементами інформації. На третьому етапі здійснюють експлуатацію БД: пошук потрібної інформації, прийняття рішень, редагування баз даних і знань.

Програмне забезпечення МІС являє собою комплекс спеціальних програм і підпрограм, баз даних і баз знань, що реалізують рішення поставлених перед МІС завдань відповідно до заданих комплексом алгоритмів.

У різних системах обробки біомедичної інформації використовується два підходи до організації медичних даних: фрагментованість і комплексування.

Існують різні алгоритми медичної діагностики. Найпростішим алгоритмом, що часто не вимагає застосування ЕОМ, є так званий "табличний алгоритм", що базується на використанні спеціальних діагностичних таблиць. Інші, потенційно більш ефективні, види медичних діагностичних алгоритмів:

- 1) імовірнісний метод, що полягає в обчисленні ймовірностей захворювань за формулою Байєса;

- 2) метод послідовного статистичного аналізу Вальда;
- 3) метод пошуку клінічного прецеденту;
- 4) метод фазового інтервалу;
- 5) метод логічного базису;
- 6) алгоритми на базі нейронних мереж.

Принципова особливість будь-якої, навіть з використанням потужних технічних засобів, діагностики: досягти 100 %-ної достовірності діагнозу практично неможливо в силу принципово складноусувних причин як об'єктивного, так і суб'єктивного характеру. Причини об'єктивного характеру - індивідуальні особливості організму, нетиповий перебіг хвороби, помилки вимірювання ознак, перекриття областей простору ознак. Причини суб'єктивного характеру - це помилки лікаря, зумовлені низкою факторів.

Експертні системи - це не тільки технологія, а й філософія використання медичних знань. Найбільш корисною характеристикою експертної системи є те, що вона застосовує для вирішення проблеми високоякісний досвід медичних фахівців. Іншою важливою характеристикою експертної системи є її гнучкість, під якою мають на увазі високу модернізаційну здатність бази знань експертної системи.

Медичні ЕС розробляються для вирішення різного роду проблем, таких як інтерпретація, діагностика, прогноз, лікування, реабілітація, планування, управління.

Література до Розділу 1

1. Мельников В.Г. Медицинская кибернетика. – К., "Вища школа", 1978, 238 с.
2. Мисюк Н.С., Гурленя А.М., Лозовик В.В. Диагностические алгоритмы. – Минск, "Высшейшая школа", 1970, 119с.
3. Абрамов М.С. Иридодиагностика с компьютером. Ташкент, изд-во им.Ибн Сины, 1991, 194с.
4. Поспелов Г.С. Искусственный интеллект – основа новой информационной технологии. М., "Наука", 1988. – 269стр.
5. Давенпорт Т. "Эпитафия экспертным системам". – М., "Computer Week", #27, 1995.
6. Уотермен Д. Руководство по экспертным системам. – М., "Мир", 1989, 388 с.
7. Мисюк Н.С. ЭВМ в диагностике нервных заболеваний. – Минск, "Беларусь", 1978, 158с.
8. Мисюк Н.С. Системы прогнозирования мозговых инсультов. Методическое пособие. – Минск, Мединститут, 1986, 40 с.
9. Вельховер Е.С. Клиническая иридология. – М., "Орбита", 1992, 431с.
10. Лорьер Ж.-Л. Системы искусственного интеллекта. – М., Мир, 1991, 566с.

Контрольні запитання до Розділу 1

1. Що мають на увазі під МІС в широкому і вузькому сенсі?
2. Які види МІС існують?
3. Які завдання вирішує розробник у процесі розробки медичної діагностичної системи?
4. Назвіть етапи обробки медичних даних і знань при створенні МІС.
5. Які структура і розв'язувані завдання у програмного забезпечення МІС?
6. Як між собою пов'язані два підходи до організації медичних даних і знань - фрагментованість і комплексування?
7. Назвіть відомі вам алгоритми медичної діагностики і дайте їм коротку характеристику.
8. Назвіть причини, за якими практично неможливо досягти 100%-ої достовірності медичної діагностики.
9. Дайте визначення експертної системи і вкажіть, у чому її принципова відмінність від звичайної комп'ютерної програми.
10. Вкажіть проблеми, для вирішення яких використовуються ЕС медичного призначення.

Розділ 2

Моделювання «клінічного мислення» лікаря

До появи ЕОМ медична діагностика удосконалювалася переважно шляхом ретельної розробки методів збору анамнестичних даних, виявлення суб'єктивних і об'єктивних ознак хвороби (симптомів). Особливо велика увага приділялася опису нових ознак різних захворювань.

Однак внаслідок накопичення клінічних даних створилися і певні труднощі, особливо для молодих лікарів, яким потрібно було не тільки запам'ятати, але й зуміти правильно оцінити їх у кожному конкретному випадку. І якщо запам'ятати клінічні дані можна за порівняно короткий час, то навчитися «клінічному мисленню» набагато важче. *Молодому лікарю зазвичай бракує не стільки вміння зібрати анамнестичні відомості і об'єктивні дані про хворого, скільки досвіду в осмисленні (переробки) й оцінки цих даних для постановки точного діагнозу.*

У цьому зв'язку величезний неминучий інтерес викликає ідея математичного моделювання функцій інтелекту лікаря в процесі переробки і оцінки медичної інформації з використанням ЕОМ. І хоча до теперішнього часу накопичений великий досвід створення автоматизованих комп'ютерних систем, включаючи і розробку низки великих експертних систем медичного призначення, задач побудови «клінічно думаючих» діагностичних систем все ще далека від свого завершення. На думку деяких фахівців, що зіткнулися на цьому шляху з великою кількістю технічних та методологічних проблем, «в комп'ютер потрібно закласти, принаймні, інститутську програму підготовки лікаря, щоб він міг давати кваліфікований висновок».

2.1. Природний інтелект: психологічний аспект

Вивчаючи проблему моделювання інтелекту, доцільно спочатку розглянути сам об'єкт моделювання - природний інтелект.

Прояви інтелекту різноманітні, але їм властиво те загальне, що дозволяє відрізнити їх від інших особливостей поведінки людини. Цим спільним є активізація тих психічних функцій, які забезпечують пізнання навколишнього світу.

Аналізуючи прийняті в ужитку вирази типу «високий інтелект», «низький інтелект» і т.п., неважко прийти до висновку про можливість і необхідність вимірювання інтелекту. Точно так само можна зробити висновок, що інтелект - поняття збірне, що характеризує такі психічні функції як пам'ять, здатність до асоціацій, абстракції, узагальнень і т.д.

Погляд на інтелект як на об'єкт вимірювань знайшов своє відображення в численних тестах для оцінки різних властивостей інтелекту - це тести логічного мислення, смислової і асоціативної пам'яті, арифметичної, просторової візуалізації і т.д. Цей ряд тестів («очищених» від вимірюють м'язової силу, гостроти слуху та зору, часу реакції і ін), що утворилися емпірично, досить чітко відме-

жований від інших методик вимірювання індивідуально-психологічних особливостей - тестів особистості, спрямованих на вивчення емоцій, інтересів, а також особливостей поведінки в певних соціальних ситуаціях.

Розглянемо коротко еволюцію підходів до визначення і вимірювання інтелекту.

У 1905 році А. Біне і Т. Сімоном була розроблена так звана «шкала розумового розвитку Біне-Сімона» - тест, призначений для виявлення розумово неповноцінних дітей. У первинному виді цей тест був сукупністю 30 задач, розташованих у міру зростання труднощів. Обстежувані з різними видами недоумства не могли вирішити завдання за межами певного рівня складності - на цій основі і вироблялася диференціація.

У 1908-1911г.г. тест Біне-Сімона був модернізований: крім диференціації дітей на нормальних і недоумкуватих виділялися різні рівні інтелектуального розвитку нормальних дітей, з угрупованням завдань за віковими рівнями, а потім шкала була продовжена до рівня дорослого. Поняття розумового рівня в результаті різних перекладів і переробок тесту Біне-Сімона було замінено поняттям *розумовий вік*, під яким мався на увазі найвищий віковий рівень тих завдань, з якими він справляється.

Згодом у тесті Станфорд-Біне (перший варіант тесту розроблений Л.М.Терменом в 1916р.) поняття розумового віку було уточнено: «розумовий вік» дорівнює «базовому віку» («базовий вік» - максимальний віковий рівень, абсолютно всі завдання якого виконуються) з додатком до нього додаткових місяців за правильно вирішені завдання в рівнях, що знаходяться вище.

Абсолютною мірою інтелекту виступає різниця між розумовим віком і віком хронологічним. Однак ця різниця для різних вікових груп має неоднакове значення, тому що розвиток інтелекту йде нерівномірно - один рік випередження або відставання в інтелектуальному розвитку для чотирирічної дитини має набагато більше значення, ніж для дванадцятирічного. У зв'язку з цим В.Штерном в 1912 р. було запропоновано визначати не абсолютну міру інтелекту - різниця, а відносну - коефіцієнт інтелекту IQ:

$$IQ = \frac{\text{Розумовий вік}}{\text{хронологічний вік}} \cdot 100.$$

Однак і цього виявилось недостатньо для усунення впливу віку - тому згодом відносний IQ був замінений стандартним IQ, які представляють собою шкалу оцінок з середнім значенням 100 і стандартним відхиленням 16 (іноді вказують значення 15). У даному випадку під стандартизацією розуміють перетворення нормальної (або штучно нормалізованої) шкали оцінок у нову шкалу, засновану вже не на кількісних емпіричних значеннях досліджуваного показника, а на його відносному місці в розподілі результатів у вибірці піддослідних.

Технічно стандартизацію здійснюють у три етапи. На першому етапі встановлюють нормальність розподілу первинних оцінок - якщо це не так, тоді вда-

ються до штучної нормалізації шляхом нелінійного перетворення вихідної шкали. На другому етапі шкалу первинних оцінок перетворюють в z-шкалу шляхом центрування і нормування:

$$z_i = \frac{x_i - \bar{x}}{\bar{\sigma}}$$

де x_i - значення первинної оцінки; \bar{x} - оцінка середнього значення; $\bar{\sigma}$ - оцінка середньоквадратичного відхилення. На третьому етапі z-шкалу трансформують у зручну для практичного використання шкалу шляхом вибору інших значень математичного сподівання M_S і середньоквадратичного відхилення σ_S :

$$z_{iS} = \sigma_S z_i + M_S$$

У результаті таких перетворень вдається формалізувати широко поширені в медицині, психології, та й у побуті, поняття «норми» і «відхилення від норми». Ідейною основою для такої формалізації служить передумова, що «нормальні» люди становлять більшість. Класифікація IQ-показників по Векслеру служить прикладом диференціації пацієнтів з використанням, подібним чином, певного поняття «норма» (табл. 2.1).

Таблиця 2.1. Класифікація IQ-показників за Векслером

IQ - показник	Рівень інтелектуально-го розвитку	Процент виявлення (по вибірці 1,7 тис. осіб від 16 до 64 років)
130 і вище	Досить високий інтелект	2,2
120–129	Високий інтелект	6,7
110–119	«Хороша» норма	16,1
90–109	Середній рівень	50,0
80–89	Знижена норма	16,1
70–79	Пограничний рівень	6,7
69 і нижче	Розумовий дефект	2,2

Ця ж ідея лежить в основі однієї з різновидів п'ятибальною оціночної шкали, яка формується наступним чином. За результатами тестування великої кількості пацієнтів будують гістограму, яку апроксимують дзвіноподібною кривою Гауса. Відраховуючи від медіанної точки, що ділить всю сукупність значень навпіл, розставляють опорні точки А, В, С і D: точки В і С по обидві сторони від медіанної точки ділять навпіл площу кожної половини, а точки А і D ділять навпіл площу крайніх чвертей.

У результаті отримують медіанну формулу

$$MF: A < B < C < D,$$

по якій будують оціночну шкалу:

- 0,0 - 1,5 балів - 1/8 всіх значень (менше А);
- 1,5 - 2,5 балів - 1/8 значень (лежать між А і В);
- 2,5 - 3,5 балів - 1/2 значень (лежать між В і С);
- 3,5 - 4,5 балів - 1/8 значень (лежать між С і D);
- 4,5 - 5,0 балів - 1/8 значень (перевищують D).

Практично медіанну формулу використовують наступним способом. Якщо, той хто тестується, набрав 2,5-3,5 балів, його результат відносять до "норми". В іншому випадку фіксують відхилення від "норми", тим даліше, чим віддаленіший результат тестування від кордонів зазначеного інтервалу.

Розглянемо тепер, які властивості інтелектуальної діяльності людини оцінюються при психологічному тестуванні. Група тестів Векслера (вперше запропонована в 1939 р., остання редакція опублікована в 1981 р.) складається з 11 субтестів, що становлять вербальну (словесну, мовну) і невербальну (дії) шкали. У табл. 2.2 дано короткий опис властивостей інтелекту, які оцінюються за Векслером.

Відповідно до методики Векслера, окремо оцінюються вербальний (перші п'ять субтестів з табл. 2.2), невербальний (решта субтестів) і загальний IQ.

У тесті ММРІ (Мінесотський багатоаспектний особистісний тест) комплексний тест "Інтелект" включає 4 основних і 6 похідних показників. Основні показники: ТБ (творчий потенціал - рівень асоціативного мислення), ЕР (ерудиційний потенціал - здатність запам'ятовувати факти), РУ (схильність до рутинної діяльності) і АК (рівень інтелектуальної активності - прагнення до новизни).

Похідні показники: $ІК=ТБ+ЕР$, $ІП=ІК \times АК$, $ТП=ТБ \times АК$, де ІК - інтелектуальний показник, ІП - інтелектуальна продуктивність, ТП - творча продуктивність. Що стосується ще однієї трійки похідних показників у вигляді НП (науковий потенціал - здатність до аналітичного мислення), АР (артистичний потенціал - рівень інтуїції) та РП (релігійний потенціал - рівень некритичності мислення), їх обчислюють таким чином: суму їх прирівнюють до ІК, тобто $НП + АР + РП = ІК$, а величину кожного компонента визначають виходячи з їх співвідношення. Про тип мислення судять по відносній домінанті однієї з показників: НП, АР або РП. Співвідношення $НП > АР > РП$ в поєднанні з високим ТБ, а ще краще - з високим ТП, свідчать про наявність важливих для дослідника якостей.

Ідея представити інтелекту у вигляді одного, генерального, компонента була вперше висунута Ч. Спірмену у 1904р. На основі статистичного аналізу тестів Ч.Спірмен висунув теорію інтелекту, згідно з якою існують фактор генеральний (фактор «g») і фактори специфічні ($S_1, S_2, S_3, \dots, S_n$), властиві кожному з використовуваних тестів. Схематично це може бути показано наступним чином:

Фактор					
Тест	g	S ₁	S ₂	S ₃	S ₄
1	x	X	—	—	—
2	x	—	x	—	—
3	x	—	—	x	—
4	x	—	—	—	x

Таблиця 2.2. Властивості інтелекту, які оцінюються за Векслером

Субтест	Властивості, що вивчаються	Фактор, що впливає на результат
Поінформованість	Зберігання в пам'яті матеріал тривалий час. Асоціації і організація досвіду	Культурне середовище, інтереси
Розуміння	Абстрактне мислення. Організація знань. Формування понять.	Можливість залучення до культури. Реакція на реальні ситуації. Об'єм уваги
Арифметичний	Збереження в пам'яті арифметичних операцій	Можливість оволодіти основними арифметичними операціями
Встановлення подібності	Аналіз зв'язків, взаємозалежностей. Формування вербальних понять	Можливість залучення до культури.
Мовний	Мовний розвиток. Формування понять	Можливість залучення до культури.
Числові ряди	Негайне відтворення. Слухові образи (частково - зорові образи)	Об'єм уваги
Послідовні картини	Зорове сприйняття відносин (візуальна інтуїція)	Можливість залучення до культури.
Відсутні деталі	Зорове сприйняття, аналіз. Зорові образи	Досвід сприйняття навколишнього середовища
Складання фігур	Зорове сприйняття, синтез. Візуально-моторна інтеграція	Ступінь і точність рухової активності
Кубики Коса	Сприйняття форми. Зорове сприйняття, аналіз. Візуально-моторна інтеграція	Ступінь рухової активності. Рівень кольорового зору
Кодування	Негайне відтворення. Візуально-моторна інтеграція. Зорові образи	Ступінь рухової активності.

У концепції Ч. Спірмена позитивні кореляції пояснюються тільки наявністю фактора «g», причому чим сильніша насиченість тестів цим фактором, тим вище кореляція між ними. Фактори специфічні грають ту ж роль, що і помилки вимірювання. Виходячи з цього, теорію Ч. Спірмена слід вважати *монофакторною*.

Найбільш вузька інтерпретація фактора «g» полягає в тому, що цей фактор притаманний усім вимірюванням інтелекту. Більш широке трактування фактора «g» - як «розумової енергії». На підставі аналізу тестів, максимально «навантажених» чинником «g», Ч.Спірмен прийшов до висновку про те, що *цей фактор в основному пов'язаний з осягненням зв'язків і відносин між предметами і яви-*

щами дійсності, а також можливістю відтворення цих відносин по певній закономірності.

Втім, вже в ранніх дослідженнях Ч. Спірменом було встановлено, що кореляції між тестами не можуть бути пояснені тільки наявністю фактора «g». Було припущено, що крім факторів генерального і специфічних існують проміжні, названі *чинниками угруповання*. Розвиток Л.Терстоуном в 1931р. цієї ідеї привело до виявлення та опису цих факторів, стало основою багатфакторних теорій інтелекту. Більше того, Л.Терстоун фактично запропонував концепцію, протилежну концепції фактора «g», відповідно до якої загальна основа інтелекту заперечується, а сам інтелект представляється у вигляді сукупності окремих, не пов'язаних один з одним здібностей - цю теорію інтелекту називають *багатофакторною*.

З виділених Л.Терстоуном «первинних чинників» найбільш чітко визначаються і підтверджуються іншими дослідниками наступні:

Позначення фактору	Сутність фактору
V	Словесне розуміння
W	Швидкість мови
N	Легкість (швидкість і точність) оперування числовим матеріалом
S	Просторова орієнтація
M	Асоціативна пам'ять
P	Швидкість сприйняття
I	Індуктивне мислення

Нарешті, серед концепцій природи інтелекту значне місце займає концепція *генетичної обумовленості інтелекту*. Так, Г.Айзенк стверджує (1979 р.), що приблизно 80% варіацій IQ слід віднести за рахунок генетичних відмінностей між людьми. Разом з тим, ряд фахівців висловлює протилежну думку - визначальне значення у формуванні інтелекту мають соціальні фактори, такі як якість освіти, приналежність до того чи іншого соціального прошарку, культури взагалі.

На підкріплення останньої позиції висуваються такі аргументи. По-перше, вже в тестах Біне-Сімона неважко бачити, що при їх розробці автори виходили з *соціальних вимог*, що висуваються до учнів шкіл того часу. По-друге, ігнорування умов культури, яка активізує, наприклад, розвиток не вербально-логічного, а просторово-образного типу переробки інформації, може призвести до постановки помилкового діагнозу розумової відсталості.

Відзначаючи справедливність наведених аргументів, проте зауважимо, що у вітчизняній психології проблема взаємодії природного і придбаного в інтелекті довгий час була однією з найбільш болючих, оскільки до певної міри корелювала з концепцією расової нерівності. Мабуть, прагнення до ідеологічної лояльності та, разом з тим, небажання применшувати дійсно значну роль спадковості

призвело в результаті до своєрідного компромісу - *вченню про задатки як передумови розвитку*. Відповідно до цього вчення, генетично детерміновані та вроджені анатомо-фізіологічні особливості мозку і нервової системи (задатки) є лише умовами формування інтелекту, що безпосередньо його не визначаючи.

Таким чином, *інтелект - це відносно самотійна, динамічна структура пізнавальних властивостей особистості, що виникає на основі спадково закріплених (і вроджених) анатомо-фізіологічних особливостей мозку і нервової системи (задатків), що формується у взаємозв'язку з ними. Він проявляється в діяльності, обумовленої культурно-історичними умовами і переважно забезпечує адекватну взаємодію з навколишньою дійсністю, її спрямоване перетворення.*

2.2. Природний інтелект: анатомо-фізіологічний аспект

Дослідження міжпівкульного розподілу функцій головного мозку свідчать, що здібності людини до того чи іншого виду діяльності значною мірою визначаються структурою його мозку. Ліва півкуля спеціалізується на аналітичній діяльності, її роль досить істотна при виконанні мовних функцій. При виконанні синхронного перекладу з однієї мови на іншу задіяні обидві півкулі, проте їх функції різні: ліва виконує власне переклад, а права здійснює контроль і розподіл уваги. Взагалі ж права півкуля "спеціалізується" на вирішенні просторових завдань, забезпечуючи швидку обробку багатовимірних даних. Тому "образне мислення" - прерогатива правої півкулі.

За допомогою комп'ютерної томографії встановлено, що між півкулями існує не тільки функціональна, але і чисто об'ємна асиметрія: обсяг лівої півкулі дещо більший обсягу правої, причому це відноситься і до передніх, і до задніх відділів мозку. Об'ємна асиметрія мозку розвивається у людини поступово, у міру його зростання, причому ліва півкуля розвивається раніше правої. Ступінь асиметрії дуже індивідуальна і визначається як генетичними факторами, так і умовами розвитку, тренуванням у вирішенні тих чи інших завдань.

У цьому зв'язку ряд дослідників вважає, що сучасні стереотипи навчання розвивають в основному лівопівкульні можливості особистості, знижуючи тим самим її творчі здібності (можливо, цим пояснюється інстинктивний інтерес ряду людей до застосування різних психотехнічних і навіть психофармакологічних засобів, що дозволяють активізувати правопівкульні функції образного мислення).

До проблем подібного роду відноситься і питання, переучувати чи ні вроджених лівшів, оскільки функціональне розходження правої і лівої рук обумовлено саме функціональним розходженням півкуль головного мозку.

Отримавши в даний час велику популярність метод нейролінгвістичного програмування, - метод психологічної корекції стану і поведінки людини, - суттєво використовує зазначені функціональні відмінності півкуль головного мозку. Доктор медичних наук, проф. А.Биховський наступним чином відгукнувся про один з різновидів методу нейролінгвістичного програмування - методі

СОЕВУС (словесно-образне емоційно-вольове управління станом людини), який розроблений доктором психологічних наук Г.Н.Ситінім: "Американський вчений Нобелівський лауреат Сперрі відкрив асиметрію мозку. Ситін відкрив, що причина багатьох соматичних захворювань у нерівномірному навантаженні мозку - ліва, розсудлива, півкуля вантажиться більше ... " Цінність методу СОЕВУС полягає насамперед у можливості здійснити психокорекцію соматичних структур самому хворому: метод реалізується шляхом використання словесних наклепів, прототипом яких є старовинні змови (наговори мають багато схожого з молитвами і за формою, і за змістом - відмічено, до речі, що сильним каталізатором для методу є саме наявність релігійної віри).

2.3. Моделювання мислення: короткий історичний екскурс

Вибір моделі мислення істотно залежить від рівня технічних досягнень.

Грецький лікар Гален (II століття н.е.) одним з перших анатомував мозок людини і тварин. Головним технічним досягненням його часу були водопровід і каналізаційна система, засновані на принципах механіки рідин. Тому навряд чи можна вважати випадковим переконавання Галена, що в мозку важливу роль відіграє не сама речовина, а заповнені рідиною порожнини (сьогодні ці порожнини відомі як система мозкових шлуночків, а рідина, що міститься в них, називається цереброспінальною рідиною). Це представлення, поряд з представленнями про чотири основні рідини, - крові, флегми (слизу), чорної жовчі і жовтої жовчі, - як ті, що визначають стан організму людини, протрималося майже 1,5 тисячі років.

У XVII столітті в зв'язку з промисловою революцією, з настанням "століття машин", прийшло переконавання, що все можна пояснити з позицій механіки. Німецький астроном Йоганн Кеплер висловив думку, що око діє по суті як звичайний оптичний інструмент. Англійська анатом Томас Вілліс описав механізм внутрішнього вуха - було визнано, що слух заснований на перетворенні звуку. Дія нервової системи пояснювалося як дія якоїсь пневматичної системи: нерви представлялися у вигляді порожніх трубок, заповнених газом. Пізніше газ в порожнистих нервах замінили на "життєву рідину" (оскільки експериментальні операції під водою показали, що розсічення тіла не супроводжується виділенням газових бульбашок), яка втікала в м'язи, змішувалася з їх рідинами і викликала різкі скорочення.

Англійський фізик Ісаак Ньютон не погодився з теорією "життєвих рідин" - на його думку, передачу впливу здійснювала віброуюче "ефірне середовище".

За допомогою примітивних приладів XVIII і XIX століть була доведена електрична природа взаємодії нервів і м'язів: італійський вчений Луїджі Гальвані виявив ефект впливу електричного потенціалу на м'язи, а німецький біолог Еміль Дюбуа-Реймон вперше виміряв електричні потенціали нервів і м'язів і першим з учених спробував пояснити всі функції мозку на основі законів хімії та фізики.

Американський вчений Норберт Вінер в 40-х роках XX століття дійшов висновку, що "... багато нинішніх автоматів ... містять органи чуття, виконавчі ор-

гани і якийсь еквівалент нервової системи, що об'єднує передачу інформації від перших до других ... Не дивно, що автомати і фізіологічні системи можна охопити однією теорією ...".

Н.Вінер не тільки відзначив, що "... сучасна надшвидка обчислювальна машина в принципі є прекрасною центральною нервовою системою для пристроїв автоматичного керування ...", але й одним з перших сформулював вимоги до такого роду обчислювальної машини:

- центральні підсумовуючі і розмножувальні пристрої (зараз ми говоримо "центральный процесор") повинні бути цифровими, а не аналоговими - вимога точності обчислень);
- ці пристрої повинні складатися з електронних ламп (зараз це надвеликі інтегральні мікросхеми), а не з зубчастих передач або електромеханічних реле - вимога швидкодії;
- повинна використовуватися двійкова, а не десяткова система числення - вимога економічності;
- всі арифметичні і логічні операції повинні виконуватися автоматично - вимога відмови від участі людини в процесі обчислень;
- машина повинна мати пристрій для запасаєння даних - вимога наявності машинної пам'яті.

Пік захоплення ідеями кібернетики припав на 40-50-ті роки. Пізніше були ще дві хвилі в області моделювання мислення: завдання автоматичного розпізнавання образів (одне з перших технічних пристроїв для її вирішення, - перцептрон, - було запропоновано американським ученим Ф.Розенблаттом в 1957 р.) і, нарешті, завдання створення штучного інтелекту.

2.4. Моделювання інтелекту лікаря – кібернетичний підхід

Моделювання інтелекту лікаря - задача не нова, її рішенням займалися колективи лікарів і кібернетиків з початку 60-х років. У цьому зв'язку закономірно запитання - чи призвела поява напрямків експертних систем до революційних змін у поглядах фахівців, які професійно займаються розробкою та практичною експлуатацією таких моделей?

У пошуках відповіді на це питання розглянемо ситуацію з моделюванням інтелекту лікаря «до ери» експертних систем, тобто в 60-70-ті роки.

У поняття «інтелект лікаря» при цьому будемо вкладати наступний сенс — це інтелект, сформований в процесі навчання і функціонування лікаря. Схематично узагальнену структуру технології навчання та функціонування лікаря можна зобразити за допомогою табл. 2.3.

2.4.1. Автоматизація формування висновку лікаря

Зусилля багатьох розробників моделей інтелекту лікаря спрямовані на автоматизацію формування висновку лікаря. Це зрозуміло, оскільки висновок лікаря у відомому сенсі є кінцевим продуктом, "вершиною" його діяльності. При

цьому передбачається, що така автоматизація повинна забезпечувати підвищення достовірності висновку за рахунок зниження його залежності від таких чинників, що заважають як різниця рівнів кваліфікації лікарів і мінливості їх психофізіологічного стану.

Таблиця 2.3. Узагальнена структура технології навчання і функціонування лікаря

Етапи технології	Зміст технології
Придбання загальних і спеціальних теоретичних знань.	Класифікація і симптоматика захворювань: вивчення нозологічних форм на базі анатомічних, фізіологічних, біологічних, мікробіологічних і ін. відомостей. Освоєння методів діагностики і лікування.
Практична діяльність під управлінням досвідченого спеціаліста.	Поглиблення і закріплення знань про нозологічні форми, методів діагностики і лікування.
Накопичення власного досвіду практичної роботи.	Поглиблення і модернізація знань про нозологічні форми. Удосконалення методів і алгоритмів діагностики і лікування.

Про підвищення якості форми висновку за рахунок використання потужних комп'ютерних засобів редагування і роздруківки графічних і текстових документів, а також про підвищення швидкості видачі висновку в 60-70-ті роки не йшлося - рівень комп'ютерних технологій був такий, що час введення-виведення інформації в ЕОМ вимірювалося годинами, а можливість інтерактивної взаємодії користувачів з ЕОМ була дуже обмеженою.

Проте, застосування ЕОМ вже в той час істотно просунуло задачу формалізації медичної документації взагалі і висновку лікаря зокрема.

Кожен вид медичної діагностики пред'являє до висновку лікаря свої вимоги. Разом з тим, у цих вимог багато спільного, що дозволяє виділити наступні ключові розділи практично будь-якого висновку:

- дані про пацієнта (анамнез, у тому числі сімейний; основні антропометричні та соціальні характеристики);
- результати огляду (перелік спостережуваних симптомів і синдромів);
- власне висновок (оцінка стану органу, системи, організму в цілому; діагноз; прогноз);
- рекомендації (напрямок на додаткове обстеження, призначення курсу профілактичних, лікувальних, реабілітаційних заходів).

Аналізуючи дану структуру, неважко побачити, що процедура формування висновку лікаря, - а саме її необхідно автоматизувати, - може бути збільшено представлена наступними основними етапами:

- оцінювання інформативних симптомів;

- обробка результатів оцінювання симптомів відповідно до діагностичним алгоритмом;
- прийняття рішення (постановка діагнозу, вибір методів лікування).

2.4.2. Інформативність симптомів

Оцінка інформативності симптомів є одним з найважливіших задач, що передують власне діагностиці і розв'язуваних колективами медичних експертів, математиків та інженерів.

Оцінку інформативності симптомів роблять у два етапи:

- на основі наявних знань, досвіду і літературних даних експерти виділяють перелік симптомів, характерних для даного захворювання або групи захворювань;
- кількісно оцінюють діагностичну значимість симптомів, що належать до виділеного переліку.

При побудові густин розподілів симптомів, необхідних для кількісної оцінки діагностичної значимості симптомів, враховують такі вимоги. Повинні бути ретельно відібрані однорідні групи хворих. Вибір шкали можливих значень оцінюваних симптомів повинен бути проведений провідними фахівцями - від якості даного етапу у великій мірі залежить ефективність всієї процедури діагностики. При цьому враховують, що симптоми можуть бути кількісними та якісними, простими і складними, залежними і незалежними. Остання властивість симптомів, - їх статистична залежність, - вимагає особливої уваги від фахівців, оскільки застосування потужних ймовірнісних методів для діагностики вимагає статистичної незалежності діагностичних ознак. Між іншим, симптоми, як правило, статистично залежні, бо є проявом одного і того ж захворювання. У цьому випадку можна вчинити по-різному: працювати з ними як зі статистично незалежними, приймаючи до уваги лише одномірний розподіл ймовірностей симптомів, або будувати багатовимірні щільності розподілу ймовірностей. Другий шлях математично коректний, але трудомісткий. Перший - некоректний, але зате більш простий для подальших обчислень. Тому частіше вибирають саме його в надії, що при вирішенні практичних діагностичних задач він забезпечить прийнятні результати - часто такий підхід цілком себе виправдовує.

В якості запобіжної діагностичної значущості ознаки використовують різні характеристики. Зокрема, *інформаційна вага симптому* обчислюють за формулою:

$$I_{ij} = \log_2 \frac{P(X_i / Y_j)}{P(X_i)} = \log_2 \frac{P(Y_j / X_i)}{P(Y_j)},$$

де $P(X_i / Y_j)$ - ймовірність (її оцінкою є частота, що зустрічається) наявності ознаки X_i за наявності захворювання Y_j ; $P(X_i)$ - ймовірність спостереження ознаки X_i взагалі; $P(Y_j / X_i)$ - ймовірність захворювання Y_j при спостереженні ознаки X_i ; $P(Y_j)$ - ймовірність захворювання Y_j .

Згідно з методикою Бродмена, діагностична цінність симптому обчислюють за формулою:

$$B_{ij} = \frac{P(X_i/Y_j) - P(X_i)}{2\sqrt{P(X_i)}} \pm 1.$$

Крім того, інформативність симптому оцінюють з використанням дивергенції Кульбака

$$K_{ijt} = [P(X_i/Y_j) - P(X_i/Y_t)] \log_2 \frac{P(X_i/Y_j)}{P(X_i/Y_t)},$$

або кількості інформації по Шеннону

$$S_{ij} = -P(X_i/Y_j) \log_2 P(X_i/Y_j).$$

2.4.3. «Медична пам'ять» - модель пам'яті лікаря

Наступним важливим поняттям, введеним в ужиток розробниками перших діагностичних комп'ютерних програм, є «медична пам'ять» — одна або сукупність декількох «діагностичних матриць», що зводять воедино переліки діагностованих захворювань і відповідні їм ознаки, з вказівкою в явній формі зв'язків між елементами цих переліків (табл. 2.4).

Таблиця 2.4

Найменування ознак	Значення ознак	Y ₁	Y ₂	...	Y _k
X ₁	x ₁₁	P(X ₁₁ /Y ₁)
	x ₁₂	P(X ₁₂ /Y ₁)
	x ₁₃	P(X ₁₃ /Y ₁)

	Всього	1,0	1,0	1,0	1,0
X ₂	x ₂₁	P(x ₂₁ /Y ₁)
	x ₂₂	P(x ₂₂ /Y ₁)

	Всього	1,0	1,0	1,0	1,0
X _r	x _{r1}	P(x _{r1} /Y ₁)

	x _{rm}	P(x _{rm} /Y ₁)
	Всього	1,0	1,0	1,0	1,0

Зауважимо, однак, що ймовірна форма опису симптомів не єдина — можуть застосовуватися і вербальні (словесні), і кодовані числами характеристики симптомів, що зустрічаються. Так, в універсальній діагностичній програмі «Ундіано» для ЕОМ «Мінськ-22» використовувалися характеристики, наведені в табл. 2.5. Очевидно, таке уявлення відрізняється від наведеного вище у вигляді значень умовних ймовірностей $P(X_i/Y_j)$ лише за формою - за змістом ці уявлення ідентичні.

Таблиця 2.5

ХАРАКТЕРИСТИКА СИМПТОМА	КОД
Патогномонічний для діагнозу	4
Дуже характерний	3
Характерний	2
Нехарактерний	1
Індиферентний	0

Прикладом «медичної пам'яті» з використанням вербальної форми опису частоти ознак, що зустрічаються, є табл. 1.1, призначена для диференціальної діагностики геморагічного та ішемічного інсульту. Цю ж таблицю після числового кодування можна представити у вигляді табл. 2.6, де 1 означає наявність симптому, 0 - його відсутність.

Таблиця 2.6

Симптоми	Геморагічний інсульт	Ішемічний інсульт
Передвісники	0	1
Раптовий розвиток	1	0
Повільний розвиток	0	1
Втрата свідомості	1	0
Блідість обличчя	0	1
Гіперемія	1	0
Артеріальна гіпотонія	0	1
Артеріальна гіпертензія	1	0
Гіпертермія	1	0
Менінгізм	1	0
Кров у лікворі	1	0
Крові у лікворі немає	0	1

Очевидна «заглибленість» табл. 2.6 - можна було, використовуючи кодування, аналогічне наведеному в таблиці 2.5, отримати більш «зважені» числові характеристики ознак, що зустрічаються. Дійсно, в табл. 2.7 відображено лише невеликий фрагмент такого роду таблиці.

2.4.4. Моделювання логічних міркувань лікаря

Перейдемо тепер до розгляду питання, яким чином здійснюється моделювання логічних міркувань лікаря в процесі діагностики. Говорячи про логічні міркування, можна було б очікувати, що під цим мається на увазі лише процес оперування з «медичною пам'яттю» (вихідними даними для такого оперування являються симптоми, які спостерігаються). У цьому сенсі «медична пам'ять»

первинна, а логічні міркування - вторинні. Однак попередньо зауважимо, що питання про «первинність» і «вторинності» на перевірку виявляється досить умовним - «стиль» мислення може мати значний вплив на структуру і характер «медичної пам'яті». Сенс цього зауваження стане більш зрозумілим далі, при розгляді конкретних алгоритмів прийняття рішення.

Найбільшого поширення в медицині та психології, аж до останнього часу, отримали так звані табличні алгоритми прийняття рішення - алгоритми, що базуються на обчисленнях з використанням таблиць виду 2.6-2.7. Підсумовуючи для ознак числа, що спостерігалися і наведені у стовпцях цих таблиць, отримують кілька сум. Діагноз ставлять, орієнтуючись на найбільшу суму.

Перевагою табличних алгоритмів є їх виняткова простота, що дозволяє обходитися без ЕОМ. Недолік - чутливість до ситуацій, коли деякі ознаки що з тих чи інших причин невимірні - може виявитися, що ці ознаки якраз є найбільш інформативними.

Таблиця 2.7

№ п/п	Симптоми	Гемораг., субархної дальн.	Гемораг. в мозок	Гемораг. в шлуночок	Ішеміч., тромбоз	Ішеміч., емболія
1	Сильний головний біль	5	3	0	1	2
2	Психомоторне збудження	3	1	0	1	2
3	Гіперемія обличчя	3	3	2	0	1
4	Блідість обличчя	0	0	0	3	1
5	Свідомість збережено	2	0	-10	3	2
6	Свідомість втрачено швидко	0	3	4	1	2
7	Кома	1	2	3	1	0
8	Артеріальна гіпертонія	2	3	3	1	0
9	Артеріальна гіпотонія	0	0	0	4	1
10	Пульс напружений	3	3	3	1	0
11	Серцева діяльність ослаблена	0	0	0	4	1
12	Дихання розстроєне	0	3	4	0	0
...
28	Інфаркт міокарда	0	0	0	10	0
29	Ревмокардит	0	0	0	10	0
30	Вік до 50 років	3	2	1	0	5
31	Вік старше 50 років	2	3	2	5	0

Значною мірою вільні від цього недоліку алгоритми, які використовують ймовірнісний підхід до прийняття рішення. Обчислення при цьому здійснюються за формулою Байєса або з використанням методу послідовного статистичного аналізу (методу Вальда). Порівняльна складність обчислень передбачає використання ЕОМ. Розглянемо в загальних рисах ці методи.

При обчисленнях за формулою Байєса в якості міри вірогідності висновку про наявність тієї чи іншої патології використовується умовна ймовірність $P(Y_j / X_i)$:

$$P(Y_j / X_i) = \frac{P(Y_j)P(X_i / Y_j)}{P(X_i)}$$

Для множини ознак $X = \{X_1, \dots, X_I\}$ формула Байєса набуває вигляду:

$$P(Y_j / X_1, \dots, X_I) = \frac{P(Y_j)P(X_1, \dots, X_I / Y_j)}{P(X_1, \dots, X_I)} \quad (2.2)$$

або, для статистично незалежних ознак,

$$P(Y_j / X_1, \dots, X_I) = P(Y_j) \prod_{i=1}^I \frac{P(X_i / Y_j)}{P(X_i)} \quad (2.3)$$

Вирішальне правило при використанні формули Байєса полягає у пошуку максимуму функції $P(Y_j / X_1, \dots, X_I)$.

У додатках зручніше рекурентний варіант формули Байєса:

$$P(Y_j / X_1, \dots, X_I) = P(Y_j / X_1, \dots, X_{I-1}) \frac{P(X_I / Y_j)}{P(X_I)}, \quad (2.4)$$

що дозволяє робити обчислення по мірі оцінювання нових симптомів, не чекаючи моменту, коли будуть оцінені всі I симптоми. Цей принцип дозволяє припинити облік нових симптомів, якщо лікар вважатиме оцінку ймовірності аналізованої гіпотези досить високою. Слід лише пам'ятати, що формула (2.4) являє собою варіант формули (2.3), тобто математично коректна лише за умови статистичної незалежності ознак.

В окремому випадку діагностики одного з двох можливих захворювань і в припущенні $P(Y_1) = P(Y_2)$ справедливе співвідношення:

$$\frac{P(Y_1 / X_1, \dots, X_I)}{P(Y_2 / X_1, \dots, X_I)} = \frac{P(X_1, \dots, X_I / Y_1)}{P(X_1, \dots, X_I / Y_2)},$$

яке для статистично незалежних ознак можна переписати у вигляді:

$$\frac{P(Y_1 / X_1, \dots, X_I)}{P(Y_2 / X_1, \dots, X_I)} = \prod_{i=1}^I \frac{P(X_i / Y_1)}{P(X_i / Y_2)}$$

або, після логарифмування,

$$u_I = \ln \frac{P(Y_1 / X_1, \dots, X_I)}{P(Y_2 / X_1, \dots, X_I)} = \sum_{i=1}^I \ln \frac{P(X_i / Y_1)}{P(X_i / Y_2)} = \sum_{i=1}^I \ln z_i. \quad (2.5)$$

В рекурентній формі:

$$u_I = u_{I-1} + \ln z_I . \quad (2.6)$$

Вирішальне правило в цьому випадку має вигляд:

$$\begin{aligned} u_I \geq 0 &\rightarrow X \in Y_1 , \\ u_I < 0 &\rightarrow X \in Y_2 . \end{aligned} \quad (2.7)$$

У поєднанні з виразом (2.6), це вирішальне правило можна трактувати таким чином: якщо після врахування чергового симптому знак величини u_I не змінився, є підстави для припинення процесу діагностики. Хоча «для страховки» можна розглянути ще кілька симптомів, переконавшись, що величина u_I дійсно продовжує віддалятися від граничного значення 0.

Недоліком такого роду міркувань є відсутність міри віддаленості величини u_I від значення 0, що можна трактувати як можливість свавілля з боку особи, що приймає рішення - що одному лікарю здасться достатнім, то іншому - ні.

Суть методу послідовного аналізу Вальда полягає в тому, що така міра вводиться, в результаті чого замість (2.7) використовується вирішальне правило вигляду:

$$\begin{aligned} u_I \geq a &\rightarrow X \in Y_1 , \\ u_I \leq b &\rightarrow X \in Y_2 , \\ b < u_I < a , \end{aligned} \quad (2.8)$$

де величини a і b , що визначають межі «коридору» в околиці 0, обчислюються виходячи із заданих значень ймовірностей помилок ε_1 і ε_2 - помилок прийняття рішення про захворювання Y_1 за наявності захворювання Y_2 , і навпаки.

Крім того, Вальд суворо математично довів доцільність впорядкування діагностичних ознак в порядку зменшення їх інформативності. Втім, теза про необхідність враховувати спочатку більш інформативні ознаки, а потім - менш інформативні, досить зрозуміла і на інтуїтивному рівні (з позицій «здорового глузду»).

У практичній діагностиці іноді використовують метод В.С.Генеса, сутність якого полягає в тому, що перелік «елементарних» симптомів зводять в один комплексний симптом (синдром) - у цьому випадку обчислення надзвичайно спрощуються, оскільки мова тепер йде про одномірні щільності розподілу ймовірності $P(Y_j / X)$ і $P(X / Y_j)$. Зауважимо лише, що цей метод можна застосовувати для діагностики невеликої групи захворювань в певному лікувальному закладі в межах чітко вираженого відрізка часу.

Покажемо тепер, що «стиль» мислення при прийнятті рішення дійсно позначається і на формі «медичної пам'яті». Відома, наприклад, діагностична таблиця Сано, що припускає проведення обчислень за формулою Байєса, з тією лише

відмінністю, що завдяки логарифмуванню операція множення замінюється складанням (табл. 2.8).

Що стосується методу послідовного статистичного аналізу можна використовувати «медичну пам'ять», представлена таблицею 2.9.

Таблиця 2.8

Найменування ознак	Значення ознак	Y_1	Y_2	...	Y_k
X_1	x_{11}	$\ln P(x_{11} / Y_1)$
	x_{12}	$\ln P(x_{12} / Y_1)$
	x_{13}	$\ln P(x_{13} / Y_1)$

X_2	x_{21}	$\ln P(x_{21} / Y_1)$
	x_{22}	$\ln P(x_{22} / Y_1)$

X_r	x_{r1}	$\ln P(x_{r1} / Y_1)$

	x_{rm}	$\ln P(x_{rm} / Y_1)$
		$\sum_i \ln P(X_i / Y_1)$			

Таблиця 2.9

Найменування ознак	Значення ознак	$\ln z_i$
X_1	x_{11}	$\ln \frac{P(x_{11} / Y_1)}{P(x_{11} / Y_2)}$
	x_{12}	$\ln \frac{P(x_{12} / Y_1)}{P(x_{12} / Y_2)}$
	x_{13}	...

X_2	x_{21}	$\ln \frac{P(x_{21} / Y_1)}{P(x_{21} / Y_2)}$
	x_{22}	$\ln \frac{P(x_{22} / Y_1)}{P(x_{22} / Y_2)}$

...

При реалізації методу лінійних дискримінантних функцій зручна медична пам'ять, показана в табл. 2.10.

Взагалі, метод лінійних дискримінантних функцій призначений для діагностики двох захворювань, сутність його зводиться до пошуку рівняння площини в багатовимірному просторі ознак. Площину вибирають таким чином, щоб якість вирішення завдання діагностики було задовільним у заздалегідь заданому сенсі.

Таблиця 2.10

Найменування ознак	Дискримінантні коефіцієнти	Значення ознак	Діагностичні індекси
X_1	a_1	x_{11} x_{12} x_{13} ...	a_1x_{11} a_1x_{12} a_1x_{13} ...
X_2	a_2	x_{21} x_{22} ...	a_2x_{21} a_2x_{22} ...
			$z = \sum_i a_i X_i$

Один із способів розрахунку дискримінантних коефіцієнтів a_i описується співвідношенням:

$$a_i = \frac{\bar{X}_i(Y_1) - \bar{X}_i(Y_2)}{\sigma_i^2},$$

де $\bar{X}_i(Y_j)$ – умовне середнє значення симптому X_i при захворюванні Y_j ; σ_i^2 – сумарна дисперсія симптому X_i при обох захворюваннях.

2.5. Системи штучного інтелекту в медицині

Становлення "індустрії інтелектуальних систем" наприкінці 70-х, початку 80-х років обумовлено:

- успіхами у сфері комп'ютерних технологій, що призвели до масового впровадження обчислювальних машин та роботів при вирішенні завдань проектування, конструювання та управління;
- високим рівнем конкуренції в розвитку комп'ютерних технологій.

Про успіхи у сфері комп'ютерних технологій відомо добре - уособленням їх може служити масове тиражування малих за габаритами, недорогих і разом з тим досить потужних настільних і переносних комп'ютерів.

Швидкий розвиток ринку персональних комп'ютерів призвів до високого рівня конкуренції не тільки між окремими виробниками, а й між національними програмами розвитку. Ініціатором була японська програма ЕОМ п'ятого покоління, за нею з'явилися стратегічна комп'ютерна програма США (1983г.), національні програми Англії, Франції, ФРН, міждержавний західноєвропейський проект ESPRIT. Країнами соцтабору також була прийнята "Комплексна програма науково-технічного прогресу країн-членів РЕВ до 2000 року", однією з пріоритетних позицій якої було створення супер-ЕОМ нового покоління з швидкістю більше 10 млрд. операцій в секунду з використанням принципів штучного інтелекту.

Дослідження в галузі штучного інтелекту (ШІ) можна згрупувати за такими чотирма напрямками:

- моделювання на ЕОМ окремих творчих процесів: ігри, автоматичне доведення теорем, автоматичний синтез програм, аналіз і синтез музичних творів тощо;
- зовнішня інтелектуалізація ЕОМ (на базі ЕОМ існуючої архітектури): дослідження в області створення діалогового інтерфейсу;
- внутрішня інтелектуалізація ЕОМ: створення ЕОМ нової архітектури з використанням принципів штучного інтелекту;
- цілеспрямована поведінка роботів.

Автоматизація медичної діагностики є елементом другого напрямку - найбільш інтенсивно розвивається і найбільш важливого для прикладних цілей.

Головною причиною очікуваного різкого підвищення ефективності роботи користувачів називалося застосування ними інтелектуального інтерфейсу, що дає можливість вирішувати складні завдання, не виходячи за межі мови своєї предметної області. Іншими словами - ефективно використовувати комп'ютерні засоби без залучення професійних програмістів.

Досягнення таких результатів планувалося шляхом створення наступних типів програмно-апаратного інструментарію:

- інформаційно-пошукові системи;
- інтелектуальні пакети прикладних програм і розрахунково-логічні системи;
- експертні системи;
- інтегровані експертні системи (об'єднання традиційних експертних систем з розрахунково-логічними).

2.5.1. Сучасні дослідження в галузі медичних систем ШІ

Поява перших робіт, присвячених застосуванню методів ШІ, датується 1972-1976 роками, можна відзначити, що незважаючи на значущість минулого відрізка часу, системи підтримки прийняття рішення, засновані на знаннях, застосовуються в медичній практиці в дуже обмежених масштабах. Навіть часто згадуються в літературі медичні експертні системи MYCIN і INTERNIST насправді застосовуються обмежено, для експериментально-дослідних цілей - у звичайній медичній практиці вони не знайшли широкого застосування. А доведені до комерційного рівня системи ILIAD і QMR придатні переважно для цілей навчання. Системи підтримки прийняття рішення, успішно впроваджених у медичну практику, вкрай мало.

Говорячи про причини цього феномену, вказується в першу чергу на причини методологічні, з посиланням на наступне висловлення Хукленброха: «Медичні експертні системи ... тільки тоді будуть задовільними в теоретичному і практичному аспектах, коли будуть базуватися на методологічно коректному та адекватному реконструюванні лікарських знань і мислення». Таким чином, сучасні дослідники не задоволені станом справ у галузі моделювання інтелекту лікаря. Що ж пропонується нам тепер, через такий значний відрізок часу? У якійсь мірі

отримати уявлення про це можна, розглянувши питання, що піднімаються в проєкті KliniC, хоча, зрозуміло, при цьому слід мати на увазі, що мова йде про роботи лише одного з сучасних європейських дослідницьких центрів.

2.5.2. KliniC - методологічна модель прийняття медичних рішень

KliniC - це методологічна модель прийняття медичних рішень, що включає всі фази процесу медичного рішення на основі різних інформаційних структур і видів знань, що використовуються в медичній практиці. Вона базується на роботах Хукленброха, присвячених методологічним і теоретичним основам медицини, претендуючи на те, щоб стати базовою (еталонною) моделлю, що забезпечує розвиток медичних баз знань і систем, заснованих на знаннях. Поки що, однак, KliniC є скоріше інформаційною, вербальною моделлю. Перетворення її в формалізовану модель здійснюється шляхом розробки комп'ютерних програм, ефективність яких перевіряється експериментально.

У моделі KliniC розрізняють наступні чотири інформаційні структурні компоненти медичного прийняття рішення, об'єднані у дві групи:

Епістемологічні знання	Деонтологічні знання
Факти	Результати вимірювання
Діагностичні гіпотези	Результати прийняття рішень

В результаті досліджень і розробок, що передували моделі KliniC, були створені:

- банк медичних термінів;
- модель інформаційної структури даних про пацієнта.

Прототипом банку медичних термінів послужив Німецький словник медичних термінів (хвороб, знаків, симптомів, процедур, якісних і кількісних характеристик і т.д.). В якості програмного інструментарію для створення комп'ютерної програми редагування і пошуку медичних термінів обраний Smalltalk / V. У банк включені 8 тисяч термінів. Структура банку досить гнучка для взаємодії з вже існуючими термінологічними системами, номенклатурами і словниками.

Модель інформаційної структури даних про пацієнта являє собою фреймоподібну структуру «прецедент-дані-листок», створену для представлення інформації про прецеденти в рамках архітектури «об'єкт - атрибут - значення». В результаті історію хвороби можна представити сукупністю «прецедент-дані-листки», в яких використані терміни з банку медичних термінів.

У рамках моделі KliniC в якості її підсистеми розроблено модель «профілю хвороби» - надалі передбачається на її основі розробити алгоритми формування висновку лікаря, що містять дані про пацієнтів з профілями хвороби і використовуються медична термінологія з наявного банку медичних термінів.

Замість понять «симптом» і «діагноз» запропоновано використовувати концепцію «патологічного стану». Причин для такого нововведення декілька.

По-перше, терміном «хвороба» позначають занадто розмиту концепцію, захоплюючи такі терміни як діагноз, синдром, розлад. І хоча загальноприйнято під синдромом мати на увазі чітко визначену групу симптомів, безвідносно до

їх етіології або патогенезісу, ці вимоги часто не дотримуються - прикладами можуть служити синдром Дауна і синдром імунодефіциту (СНІД).

По-друге, в деяких випадках важко сказати, з чим ми маємо справу - з ознакою або хворобою. Наприклад, гіпертензія може бути і тим, і іншим. Без додаткової інформації однозначна відповідь неможлива. Аналогічно ренальну недостатність можна розглядати і як хворобу, яку потрібно лікувати, і як симптом інших захворювань. Хвороби, подібно симптомам і синдромам, можуть бути елементами медичної історії (анамнезу).

Нарешті, класична схема медичного рішення (від медичної історії, знаків і симптомів - до діагнозу, а потім до відповідної терапії), запропонована в більшості медичної літератури та медичних експертних системах, не відповідає, на думку Шульца, повсякденній медичній практиці. Особливо при первинному наданні допомоги і при екстреному лікуванні, коли медичне втручання значно частіше проводиться на підставі знаків і симптомів, ніж за результатами діагностичних рішень. Крім того, у переважній більшості медичних спеціальностей лікарі приділяють більше часу організаційним проблемам, ніж діагностичним рішенням.

Справедливості заради відзначимо, що міркування Шульца не відрізняються великою новизною. В якості одного з доказів можна розглянути таблицю для диференціальної експрес-діагностики гострих порушень мозкового кровообігу, розроблену Н.С. Мисюк і І.М. Довнар в 1968 році, де поруч з найменуваннями анамнестичних даних і симптомів містяться найменування захворювань - всі вони в даному випадку грають роль вихідних даних задачі діагностики. Що ж до прийняття рішення в екстраординарних ситуаціях, як доказ можна навести структурну схему логіко-імовірнісного алгоритму (ЛВА), використану Н.С. Мисюк на початку 70-х років в діагностичній програмі «Ундіано» для ЕОМ «Мінськ-22» (рис.2.1).



Рис. 2.1

Відповідно до цієї схеми, особливо виділені випадки, коли у вихідних даних про хворого є ознаки, властиві тільки конкретному захворюванню. Ці ознаки класифікуються як патогномонічні, їм присвоюється найбільшу вагу (табл. 2.5), при їх виявленні ставиться діагноз не по імовірній логіці, як для «звичайних» ознак, а по детермінованій логіці.

Можна, зрозуміло, заперечити, що наведений приклад свідчить всього лише про можливість прискорити прийняття діагностичного рішення в екстрених си-

туаціях - однак етап діагностики при цьому все одно присутній, передуючи етапу лікування. Тим часом як Шульц чи не заперечує класичну схему медичного рішення (симптоми - діагноз - терапія), наводячи як аргумент застосування в екстрених ситуаціях альтернативну схему «симптоми - терапія».

Неважко, однак, зауважити, що «виключення» з правил, що наводиться, насправді таким не є - етап діагностики, як проміжний між етапами оцінювання симптоматики і терапії, в даному випадку просто «прихований» від спостереження за цілком зрозумілих причин - в екстрених ситуаціях фактор швидкості застосування лікувальних заходів грає головну роль. Однак неявна присутність етапу діагностики негайно виявляється, варто лише задуматися про обґрунтування лікувальних заходів, що вживаються.

Таким чином, коментуючи «розгром» Шульцем класичної схеми медичного рішення, слід, з одного боку, відзначити перебільшеність звинувачень в «неповноцінності», пред'явлених класичною схемою. З іншого боку, не можна не погодитися і з тим, що припустимо говорити про існування не однієї, а двох схем медичного рішення. До речі, гарною ілюстрацією життєздатності другої, скороченої, схеми є метод гомеопатичного лікування, коли лікувальні препарати підбираються переважно за результатами оцінювання симптомів (див. розділ 8.2), а не на підставі попередньої діагностики.

2.5.3. Патологічні стани як елементи декларативних знань

На думку німецьких фахівців, ізольована модель профілю хвороби незручна. Для її застосування потрібно залучати додаткові моделі сутностей синдромів, симптомів, знаків і т.д., з чіткими відмінностями між цими категоріями. Виходом з цієї ситуації є введення більш загального терміна для позначення всіх видів медичних відхилень і пов'язаних з ними станів, а саме: *патологічні стани*.

Знову відзначимо, що і ця ідея не нова - Н.С.Місюк також використовував термін «стан», вкладаючи в нього подібний сенс: це не тільки та чи інша нозологічна форма, коли необхідно провести діагностику усередині класу захворювань. Під терміном «стан» можна розуміти і весь клас захворювань - в такому випадку це буде попередня діагностика. Або навпаки, розглядаючи в якості станів різні ступені прояву або «відтінки» своєрідності даного захворювання, приходимо до диференціальної діагностики.

Кілька прикладів, що ілюструють діапазон можливих патологічних станів:

- 1) гострий інфаркт міокарда (класичний діагноз);
- 2) тупий біль в нижньому правому квадранті живота протягом більше однієї години (комплекс симптомів);
- 3) підвищений рівень холестерину (результат лабораторних аналізів);
- 4) апендектомія більш ніж 10 років тому (медична історія);
- 5) алергічна готовність (знак схильності до хвороби);
- 6) нетримання стільця (діагноз медсестри);

7) наявність 6 з 11 АРА-показників (оцінка показників за класифікацією АРА - американської асоціації ревматологів);

8) протеїнурія (типовий симптом).

Ці приклади свідчать, що категорія «патологічний стан» виробляє елементи, корисні для опису конкретних клінічних випадків і абстрактних медичних знань. Вона відповідає концепції «епістемологічних знань» моделі KliniC.

Відповідно до цієї моделі, будь-який патологічний стан можна охарактеризувати спеціальним фактором, інтерпретується як апостеріорна ймовірність по відношенню до даних, або як апріорна ймовірність по відношенню до узагальнених знань про хвороби.

Будь-який патологічний стан може також стимулювати показання, протипоказання, рішення і дії, що відповідає «деонтологічному» рівню моделі KliniC. Створення узагальненої моделі цих «медичних процедур» з'явиться вмістом наступного кроку в реалізації KliniC.

Патологічні знання являють знання, іменовані *декларативними*. Частина їх можна розглядати як *термінологічні* знання, лексично представлені ідентифікаційними медичними термінами (приклад 8, «протеїнурія»). Навпаки, синтаксичні конструкції, подібні 2), 4), і 7), є прикладами декларативних, але не термінологічними знаннями.

Триплети об'єкт-атрибут-значення є елементарними структурними утвореннями. Кожен триплет належить класу «дітей», що породжувані об'єктом. Дозволяється об'єднання атрибутів в групи. Значення можуть бути кількісними та якісними. Кожне значення може заперечуватися. Обмеження на зв'язку об'єкт-атрибут, атрибут-атрибут і атрибут-значення встановлюються на підставі існуючого банку термінів (тобто атрибут «локалізація» не може мати значення «15 г / л»). Зовнішні відносини представляються аналогічно (наприклад: AIDS - ICD9 - 042). На рис.2.2 наведено приклад такої структури.

Можна показати, що дана елементарна структура достатня для даних, що описують прості випадки. Для більш загальної моделі патологічних станів, тим не менш, потрібна додаткова інформація про логічні зв'язки між об'єктами типу І, АБО, ПОСЛІДОВНІСТЬ, N з M (див. вище приклад 7). Така структура дозволяє описувати патологічні стани як концептуальні графи, що допускають множинні успадкування. Наприклад, патологічний стан, показаний на рис.2.2 (назвемо його PS_001), може бути доповнено інформацією про те, що біль неперервна - при цьому створюється об'єкт PS_002, який визначається триплетом «PS_001 - модель часу - безперервна».

2.5.4. Стверджуючі знання як зв'язки між станами

Знання, що описують явище, називаються ствердними. У семантичній мережі декларативні знання відповідають вузлам, ствердні знання - семантичним зв'язкам. Приклад використання моделі стверджувальних знань наведено на рис. 2.3.

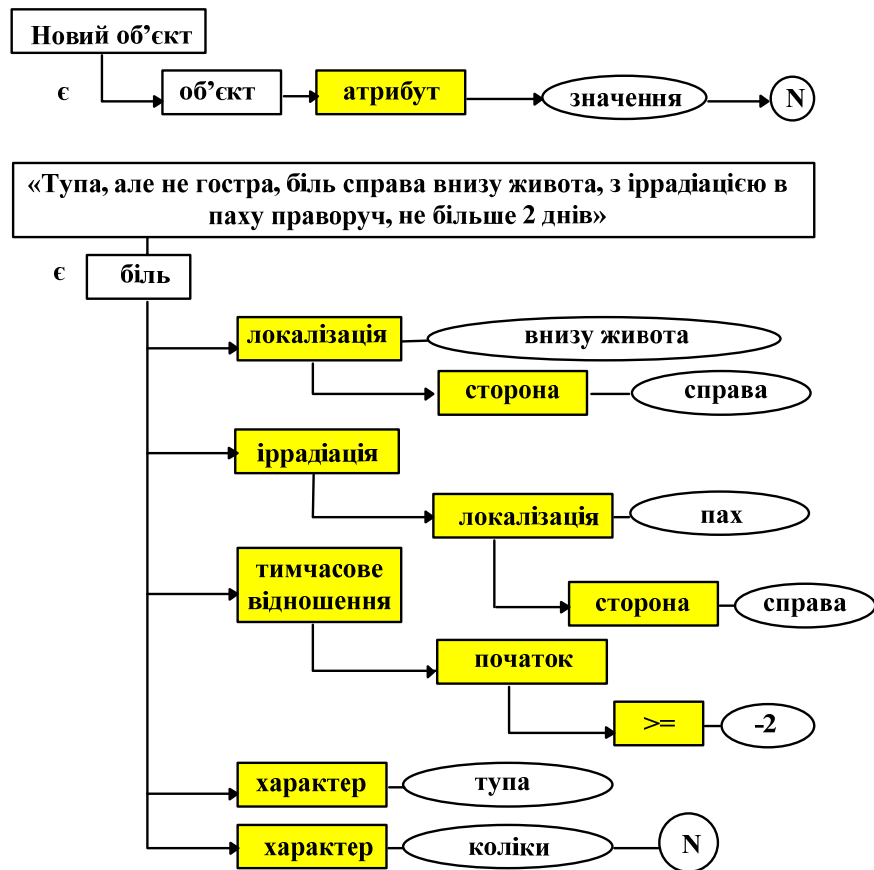


Рис. 2.2

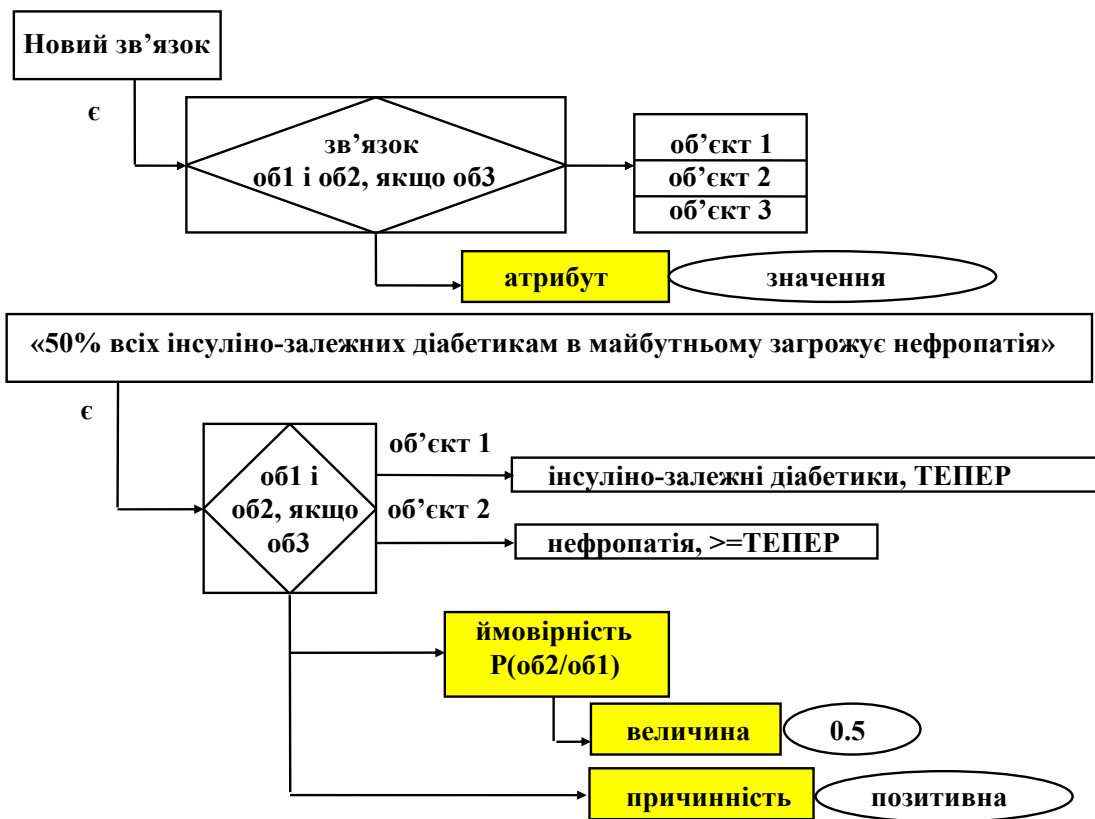


Рис. 2.3

Таке представлення базується на архітектурі імовірнісних мереж, хоча необхідно відмітити, що в медичній літературі дуже мала частка зв'язків між патологічними станами описується в кількісній формі. Найбільш поширеним видом семантичного зв'язку є ставлення «Властивість Чого» (приклад: протеїнурія - «Властивість Чого» - нефрит), зазвичай характеризується вербально (такими словами як «характерно», «часто», «типово» і т.д.), переклад якого в кількісну форму для подальшого застосування Байєсових стратегії являє собою складну задачу.

Таки чином, реалізація моделі KliniC зводиться до зв'язування патологічних станів за допомогою стверджувальних знань. При цьому між двома об'єктами встановлюються кількісні або якісні зв'язки, з можливою умовою наявності третього об'єкта. У тому числі з використанням такого важливого атрибуту як причинність.

Короткі висновки

Експертні системи являють собою комп'ютерну модель мислення лікаря.

У поняття «інтелект лікаря» доцільно вкладати наступний сенс - це інтелект, сформований в процесі навчання і функціонування лікаря.

Зусилля багатьох розробників моделей інтелекту лікаря спрямовані на автоматизацію формування висновку лікаря, оскільки висновок лікаря у відомому сенсі є кінцевим продуктом, "вершиною" його діяльності. При цьому передбачається, що така автоматизація повинна забезпечувати підвищення достовірності укладання за рахунок зниження його залежності від таких чинників, що заважають як різниця рівнів кваліфікації лікарів і мінливість їх психофізіологічного стану.

Процедура формування висновку лікаря, яка зазвичай є об'єктом автоматизації, може бути представлена наступними основними етапами:

- 1) оцінювання інформативних симптомів;
- 2) обробка результатів оцінювання симптомів відповідно до діагностичного алгоритму;
- 3) прийняття рішення (постановка діагнозу, вибір методів лікування).

Оцінка інформативності симптомів є одним з найважливіших завдань, що передують власне діагностиці і розв'язуваних колективами медичних експертів, математиків та інженерів.

Оцінку інформативності симптомів роблять у два етапи:

- 1) на основі наявних знань, досвіду і літературних даних експерти виділяють перелік симптомів, характерних для даного захворювання або групи захворювань;
- 2) кількісно оцінюють діагностичну значимість симптомів, що належать виділеному переліку.

Важливим поняттям, введеним в побут розробниками перших діагностичних комп'ютерних програм, є «медична пам'ять» - одна або сукупність декількох «діагностичних матриць», що зводять воедино переліки діагностованих захво-

рювань і відповідних їм ознак, з вказівкою в явній формі зв'язків між елементами цих переліків.

При цьому можуть використовуватися різні форми опису таких зв'язків:

- 1) якісна вербальна ;
- 2) кількісна частотно - імовірнісна чи бальна.

Найбільшого поширення набули так звані табличні алгоритми прийняття рішення, перевагою яких є виняткова простота, що дозволяє обходитися без ЕОМ. Недолік - чутливість до ситуацій, коли деякі ознаки що виміряні з тих чи інших причин, а в результаті може виявитися, що ці ознаки якраз є найбільш інформативними.

Значною мірою вільні від цього недоліку алгоритми, які використовують імовірнісний підхід до прийняття рішення. Обчислення при цьому здійснюються за формулою Байєса або з використанням методу послідовного статистичного аналізу (методу Вальда). Порівняльна складність таких обчислень передбачає використання ЕОМ.

Незважаючи на значущість відрізка часу, що пройшов після появи перших медичних діагностичних систем (1972 – 76 рр.), системи підтримки прийняття рішення, засновані на знаннях, і донині застосовуються в медичній практиці в дуже обмежених масштабах.

Причина цього феномена, на думку багатьох фахівців, полягає в тому, що медичні експертні системи тільки тоді будуть задовільними в теоретичному і практичному аспектах, коли будуть базуватися на методологічно коректному та адекватному реконструюванні лікарських знань і мислення.

Література до Розділу 2

1. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психологической диагностике. – Киев, Наукова думка, 1989. – 196с.
2. Кураев Г., Пожарская Е., Глумов А. Межполушарное распределение функций. "Известия вузов. Северокавказский регион. Естественные науки", #2, 1996г.
3. Блум Ф., Лейзерсон А., Хофстедтер Л. "Мозг, разум и поведение". М., Мир, 1988. – 248с.
4. Винер Н. Кибернетика. М., Сов.радио, 1968. – 323 с.
5. Поспелов Г.С. Искусственный интеллект – основа новой информационной технологии. – М., "Наука", 1988. – 269с.
6. Мельников В.Г. Медицинская кибернетика. – К., Вища школа, 1978. – 238с.
7. Мисюк Н.С., Гурленя А.М., Лозовик В.В. Диагностические алгоритмы. – Минск, Вышэйшая школа, 1970. – 107с.
8. Мисюк Н.С. ЭВМ в диагностике нервных заболеваний. – Минск, Беларусь, 1978. – 158с.

9. Мисюк Н.С. Системы прогнозирования мозговых инсультов. (Методическое пособие). - Минский госуд. мед. институт, 1983. – 59с.

10. Петенко О.В., Гречишникова Н.И. Практическое пособие по иридодиагностике. – Душанбе: Изд-во "ИРФОН", 1991.

11. Продеус А.Н., Сядро Т.А. Проблемы реализации Байесовской стратегии при автоматизации иридодиагностики. Сб. «Электроника и связь», №2, часть II. – К., НТУУ (КПИ), 1997г., с.312–316.

12. Schulz S. Knowledge Representation of Pathological States (1995). In: Proc. 19th Annual Conference of the Gesellschaft fur Klassifikation. University of Basel, 447–456.

Контрольні запитання до Розділу 2

1. Вкажіть етапи придбання знань медиками у вузах. В якій мірі експертні системи повинні відображати ці знання?

2. Чому основні зусилля розробників моделей інтелекту лікаря спрямовані на автоматизацію формування висновку лікаря?

3. З яких етапів складається процедура формування висновку лікаря? Наведіть відповідний приклад зі своєї практики спілкування з лікарями.

4. Чому важливо мати інформацію про діагностичну цінність ознак? Яким чином можна отримати таку інформацію?

5. Чому таблиці зв'язків можливих патологій і супутніх їх симптомів іменують «медичною пам'яттю»?

6. Які види діагностичних алгоритмів вам відомі? Які діагностичні алгоритми можна використовувати без ЕОМ, а які - з застосуванням ЕОМ?

7. Чому системи підтримки прийняття рішення, засновані на знаннях, і донині застосовуються в медичній практиці в дуже обмежених масштабах?

Розділ 3 Представлення знань в ЕС

3.1. Знання та дані

Питання про співвідношення понять «знання» і «дані» завжди піднімається при розгляді сутності експертних систем. При цьому спостерігається відсутність однотайності в інтерпретації цих понять. Цікаво тому зіставити кілька точок зору з тим, щоб по можливості ближче підійти до істини, яка «посередині».

Г.С.Поспелов. З розробкою і використанням обчислювальної техніки традиційно пов'язані такі поняття як програми і дані.

На ранніх етапах розвитку обчислювальної техніки програміст розробляв програму і сам вводив в неї необхідні дані.

Потім відбулася велика зміна - дані були відокремлені від програм шляхом використання баз даних (БД) різної структури (табличні, ієрархічні, мережеві) і систем управління базами даних (СКБД).

Змінювалося і представлення даних у зовнішній пам'яті ЕОМ. Фундаментальним поняттям тут стало поняття файлу - інформаційного масиву, що має ім'я і містить всі необхідні структуровані записи даних про різні об'єкти, з якими веде роботу система. Еволюція уявлення даних у зовнішній пам'яті пройшла наступні етапи:

1) способи формування записів даних у файлах, ведення файлів і організація доступу до них повністю визначалися в конкретних програмах користувачів;

2) управління файлами і організацію доступу до них стали здійснювати операційні системи ЕОМ;

3) створення БД і СУБД, що забезпечують ефективну роботу з великими обсягами різних даних.

Відділення даних від програм не тільки підвищило ефективність обробки даних, але і породило нові ситуації. В принципі, стало можливим викликати програму, виходячи з даних, а не навпаки, як раніше.

З появою систем ШІ з'явилося нове поняття - «база знань». Виникла необхідність якось співставити звичні поняття «дані і база даних» з поняттями «знання і база знань». Безсумнівно, що дані і структура бази даних певною мірою відображають знання про предметну область і її структуру. Тим не менш, можна виділити наступні чотири ознаки, що відрізняють знання від даних:

- внутрішня інтерпретованість;
- структурованість;
- зв'язність;
- активність.

У якійсь мірі деякі з перерахованих ознак властиві базам даних. Наприклад, інтерпретованість явно спостерігається в реляційній (табличній) БД, де імена стовпців є атрибутами відносин, імена яких вказані в рядках. А структурованість (використання відносин «частина-ціле», «клас-під ⊂ клас», «рід-вид» і т.п.) зустрічається в ієрархічних, мережевих БД, і при необхідності може бути реалізована в табличних БД.

Для третьої ознаки - зв'язність - практично не можна знайти аналогів в БД, які згадувалися, оскільки наші знання пов'язані не тільки в сенсі структури, але також відображають закономірності щодо фактів, процесів, явищ і причинно-наслідкові відносини між ними.

Що ж до четвертого ознаки - активність - то по ньому знання принципово відмінні від даних. Дані пасивно зберігаються в пам'яті ЕОМ. Знання активні в тому сенсі, що виявлення протиріч у них стає причиною їх подолання та появи нових знань. Стимулом, що спонукає до активності, є також неповнота знань.

Незважаючи на наявність відмінних ознак, чітку грань між даними і знаннями провести можна не завжди.

В.Томов. В основному згоден з точкою зору Г.В.Поспелова, що впливає з його посилання на роботу Г.В.Поспелова.

У даних немає структури. Вони можуть бути структуровані в рамках БД. Але навіть після такого структурування вони не мають горизонтальних (ситуаційних, причинно-наслідкових) і вертикальних (спадкування, рід-вид, загальноприватне) зв'язків.

Відносини між поняттями «дані» і «знання» - одна з головних проблем в комп'ютерній інформатики.

Знання - це загальна, відносно постійна і незмінна частина інформації (для класів об'єктів, законів, відносин і т.д.). Дані - це більш динамічна частина інформації. Приклад: факт про те, що крива K є колом з $R=1$ см - це дані, а факт, що довжину кола можна обчислити за формулою $C=2\pi R$ - це знання.

Між поняттями «дані» і «знання» немає суворого розмежування. У певних випадках знання можуть стати даними. Приклад: в задачі загального опису різних фігур той факт, що для кола справедливе співвідношення $C=2\pi R$ - це вже дані, поряд з іншими даними у вигляді формул довжини інших геометричних фігур: еліпса, квадрата, ромба...

З даними оперують, а знання застосовують для прийняття рішень. У цьому сенсі знання є даними в системах обробки знань.

Д.І.Шишков. Не в усьому згоден з В.Томовим і Г.С.Поспеловим. Тема відносини дані-знання - одна з не самих істотних тем штучного інтелекту. Крім того, ця тема розглядається в літературі по ШІ не зовсім коректно.

Загальне неблагополуччя в трактуванні відносини дані-знання викликано не зовсім зрозумілою термінологією, відсутністю математичного визначення поняття «дані».

Пропонується визначити дані в такий спосіб.

Дане - це слово над кінцевим алфавітом, яке розглядається разом з операціями, що здійснюються над ним (і над іншими даними його типу й іншого типу, з результатом його типу або іншого типу даних), незалежно від сенсу в реальному світі (в природі чи суспільстві), яке він може мати (йому можна приписати).

Точне математичне визначення абстрактного і конкретного типів даних можна дати, використовуючи математичну структуру (по Бурбак)

$$MS = \{Obj, R, Ax\},$$

де Obj – множина об'єктів; R – множина відносин (в т.ч. операцій); Ax – кінцева множина аксіом, які зв'язують об'єкти і відносини між ними.

Абстрактний тип даних:

$$ADT = \{D_{DT}, R, (Op,) AX\}, Op \subseteq R,$$

де D_{DT} – множина даних цього типу з невизначеною машинною реалізацією; Op – множина операцій; ADT – математична структура по Бурбаки і клас еквівалентності.

Конкретний (візуальний) тип даних:

$$AD = \{A, Og, D_{DT}, F, R, (Op,) AX\},$$

де A – кінцевий алфавіт; Og – множина операцій, що виникли; F – формат (кінцева множина правил, які форматують дані із D_{DT} , тобто. групують літери даного і надають групам спеціальний модельний сенс).

Дане певного типу завжди має два сенсу: модельний (формальний, математичний, комп'ютерний, синтаксичний) і реальний (у природі і суспільстві).

Що ж стосується поняття «знання», Шишков обмежується повторенням «стандартного» його визначення:

$$\text{знання} = \text{інформація} + \text{мета},$$

мотивуючи це тим, що ні інформація, ні мета не є обумовленими (математичними) поняттями. Вони повинні бути прийняті аксіоматично (незважаючи на те, що вимірами інформації займаються давно і з корисним ефектом).

Відштовхуючись від запропонованих їм визначення поняття «дане», Шишков стверджує наступне.

Знання та дані – поняття, які не можуть порівнюватись.

Дані - формальне (в комп'ютерах - математичне) поняття, слово над якимось алфавітом. Дані - це носії інформації, але самі по собі вони не є інформацією. Навпаки, знання - це інформація, яку можуть отримати люди, які знають сенс відповідних даних. Наприклад, для людей, які не знають китайської мови, китайський текст - всього лише дані у вигляді сукупності ієрогліфів; для тих, хто знає китайську мову, цей же текст - вже інформація.

Все, що в пам'яті комп'ютера - це дані. Один з найбільших міфів комп'ютерної інформатики полягає в тому, що дані обробляються, а машинні команди інтерпретуються - насправді властивість інтерпретації належить тільки людині.

Комп'ютер, в кінцевому рахунку, обробляє машинні команди як дані - в пам'яті машини дані і машинні команди не можна розмежувати за допомогою формального правила, що випливає з третього принципу Джона фон Неймана «І дані, і програма повинні знаходитися в пам'яті одночасно».

Знанням як таким (інформація плюс мета) немає місця в комп'ютерному світі. Тільки у людини є знання (і деякі дані - їх носії).

Тому замість некоректних у вказаному вище сенсі відносин «дані-знання» слід розглядати інше відношення: «знання (факт) на більш низькому рівні - знання (факт) на більш високому рівні».

Ж.-Л.Лор'єр. Його позиція у відомому сенсі подібна з позицією Д.І.Шишкова. Замість відносин «дані - знання» він розглядає відношення «процедурний підхід - декларативний підхід».

Процедурний підхід властивий класичному програмуванню, запис алгоритму мовою програмування при цьому:

- 1) складається виключно з інструкцій;
- 2) являє собою строго певну послідовність інструкцій.

Перша проблема програмування на процедурній мові полягає в тому, що програмуючи, ми вирішуємо справжню головоломку - ми складаємо алгоритм з розрізнених дрібних шматочків, тоді як насправді наше знання модульно і кожен модуль складається з чотирьох-шести таких шматочків.

Друга, і більш істотна, проблема полягає в тому, що зазвичай ми хочемо отримати від програми більше, ніж вона це може зробити після закінчення її розробки. Пояснити це можна на прикладі програмування закону Ома: «ЕРС U , опір R і струм I , що протікає по електричного кола при кімнатній температурі, пов'язані між собою співвідношенням $U = R I$ ». Складаючи обчислювальну програму, необхідно врахувати, що не тільки U є функцією R і I , але і R - функція U і I , і I - функція U і R . Навіть якщо з трьох значень U , R і I невідомими є два, то все одно можна використовувати деякі співвідношення, наприклад для двох паралельно включених опорів справедливо $R1 / R2 = I2 / I1$. Очевидно, в принципі врахувати все це при процедурному підході до програмування можна, якщо заздалегідь передбачити всі мислимі бажання кінцевого користувача. На практиці ж так не чинять, тому очевидно, до якої міри властивості процедурного підходу стискають кінцевого користувача - варто йому поставити нове питання, як знову необхідно змінювати написану перш програму.

При декларативному ж підході керуюча частина програми відокремлена від частини, що зберігає інформацію про правила. Тому правила можуть вводитися в міру необхідності, в безладді. Головна властивість декларативного підходу тому - відсутність впорядкованості (у зазначеному сенсі). Воно дозволяє взаємодіяти людині з програмою найбільш простим чином, на рівні знань, не вдаючись до послуг професійних програмістів - володіють такою здатністю програми іменуються експертними системами.

3.2. Класифікація знань.

Класифікуючи знання, Ж.-Л.Лор'єр виділяє вісім основних типів знань.

1. Базові елементи, об'єкти реального світу. Пов'язані з безпосереднім сприйняттям, не вимагають обговорення і додаються до бази фактів у тому вигляді, в якому вони отримані.

2. Твердження й визначення. Засновані на базових елементах і заздалегідь розглядаються як достовірні.

3. Концепції. Представляють собою перегрупування або узагальнення базових об'єктів. Для побудови кожної концепції використовуються свої прийоми. Наприклад, вони можуть будуватися на основі прикладів, контрприкладів, приватних випадків, більш загальних або аналогічних концепцій.

4. Відносини. Висловлюють як елементарні властивості базових елементів, так і відносини між концепціями. Крім того, до властивостей відносин відносяться їх більшу або меншу правдоподібність, більша або менша зв'язок з даною ситуацією.

5. Теорема і правила перезапису. Є окремим випадком продукційних правил з цілком певними властивостями. Теорема не уявляє ніякої користі без експертних правил їх використання.

6. Алгоритми рішення. Є знанням особливого типу, оскільки обумовлена ними послідовність дій виявляється оформленою в строго необхідному порядку на відміну від інших типів знання, де елементи інформації можуть з'являтися і розташовуватися без зв'язку один з одним. Використання чистих алгоритмів обмежено дуже окремими випадками, більша частина яких має справу з обробкою числової інформації.

7. Стратегії та евристика. Цей тип являє собою вроджені або придбані правила поведінки, засновані на використанні інформації в порядку, зворотному тому, в якому вона була отримана. Прикладом є міркування типу: «Я знаю, що ця дія призводить до такого-результату, тому, якщо я хочу отримати саме цей результат, я можу розглянути це дія».

8. Метазнання. Знання про знаннях, відводить місце даного знання в ряду іншої інформації. Забезпечує організацію кожного типу знань і вказівок, коли і як вони можуть бути використані.

При цьому Ж.-Л.Лор'єр одночасно стосується питання про організацію цих знань в експертних системах, відзначаючи, що подання знань в експертних системах близько до моделям, використовуваним в базі даних. Таким шляхом побудована реляційна (узагальнена) модель БД в системах OPS, PROLOG, SNARK і TANGO.

Так, говорячи про відносини, наголошується, що пара понять «властивість - значення» добре відома в БД типу семантичних мереж. А такі БД як фрейми і скрипти є не чим іншим, як найбільш простими бінарними відносинами.

Що ж стосується відмінності експертних систем від СУБД, то зазначається, що головною відмінною рисою ЕС є явне присутність теорем - в класичних СУБД теорема або відсутні, або програмуються. І хоча модифікація або дода-

вання нових теорем в ЕС є дуже трудомісткою процедурою, однак вона необхідна для хорошого структурованого управління базою даних і оптимізації отримання відповідей.

Неважко бачити, що чіткого визначення, що ж таке «база знань», Ж.-Л.Лор'єр так і не дає. Замість цього фактично пропонується аналізувати існуючі експертні системи. Ситуація ця досить типова: точно визначити, що таке база знань, вельми складно - поняття це розмите і широке, трактується по-різному.

Між іншим, схожість словосполучень "база знань" і "база даних" вже неминуче породжує питання про зв'язок цих понять по суті. Один з підходів до визначення БЗ запропонував С.С.Лавров.

Розпочинаючи опис підходу С.С.Лаврова, нагадаємо, що експертні системи іноді іменують системами, що базуються на знаннях (СБЗ), підкреслюючи тим самим відмітна властивість ЕС - наявність у її складі бази знань (БЗ). Інші складові частини ЕС - механізм отримання рішень (МР) і інтерфейс (ІФ). Базова структура ЕС, показана на рис 1.7, таким чином, може бути описана співвідношенням

$$ЕС = БЗ + МР + ІФ.$$

Згідно С.С. Лаврову, БЗ є інтегрована система, що складається з трьох компонентів:

- бази даних;
- моделі предметної області (МПО);
- пакети прикладних програм (ППП).

Отже, при такому підході БД розглядається як складова частина БЗ:

$$БЗ = БД + МПО + ППП.$$

Кожна складова частина БЗ при цьому призначена для зберігання одного з трьох видів знань - предметного (БД), понятійного (МПО), процедурного (ППП).

Крім того, С.С.Лавров розрізняє алгоритмічні та неалгоритмічні знання.

Алгоритмічні (або процедурні) знання - це алгоритми (програми, процедури), що обчислюють функції, які виконують перетворення, вирішальні точно певні конкретні завдання. Базою алгоритмічних знань можна (БАЗ) вважати будь-яку (бібліотеку) програм. Кожна система програмування і операційна система включають в себе базу алгоритмічних знань. Основою кожного ППП або проблемно-орієнтованої системи (ПОС) є база алгоритмічних знань у конкретній прикладній області. Відомі приклади ЕС (MACSYMA, DENDRAL), БЗ яких складається головним чином з алгоритмічних знань - такі ЕС іноді вважають нетиповими або взагалі не ЕС.

Для ЕС характерна наявність бази неалгоритмічних знань (БНЗ), що містить два види неалгоритмічних знань:

- концептуальне (понятійне), що складається з понять (уявних об'єктів, що мають ім'я, визначення, структуру) і зв'язків між поняттями;

- фактуальне (предметне), утворене "звичайними" БД.

Таким чином, запропоновану вище «формулу» БЗ можна переписати у вигляді:

$$БЗ = БНЗ + БАЗ,$$

де

$$БНЗ = БД + МПО, БАЗ = ППП.$$

Слід зауважити, що знання, втілені в поняттях, не зводяться до моделей предметних областей - є ще математичні знання, також складаються з математичних понять, зв'язків між ними та тверджень про них.

Крім того, в багатьох ЕС розрізняють такі види знань як «факти» і «правила». Факти - це елементарні «одиниці знання» (прості твердження про характеристики об'єктів). Правила служать для вираження зв'язків, залежностей між фактами та їх комбінаціями.

Лавровим пропонується наступна класифікація видів знань:

- 1) поняття (математичні та нематематичні);
- 2) факти;
- 3) правила, залежності, закони, зв'язку;
- 4) алгоритми, процедури.

Зазначимо кращу, на наш погляд, структурованість класифікаційної схеми Ж.-Л.Лор'єра.

3.3. Основні формалізми для представлення знань

В ЕС використовуються такі основні формалізми для подання знань:

- продукційні правила;
- семантичні мережі;
- фрейми.

Продукційні правила зручні для формалізації рішення діагностичні завдань і легко реалізуються засобами реляційних СУБД.

Формалізми типу семантичних мереж і фреймів базуються на використанні БД як реляційного, так і нереляційного типу. І ті, й інші використовують загальне поняття мережі - структури у вигляді вузлів і зв'язують їх дуг (іменованої також графом). Такого роду формалізми дуже зручні для цілей аналізу і синтезу мовної і візуальної інформації, для планування з використанням поняття "мережевий графік".

3.3.1. Правило продукції

Правило продукції "якщо А, то В" ("якщо є подія А, тоді є подія В") володіє незаперечними перевагами перед іншими формалізмами, оскільки легко сприймається користувачами, зокрема, лікарями при вирішенні задачі діагностики. Схема міркувань при цьому, в першому наближенні, має вигляд: "якщо спостерігається сукупність ознак А, тоді з великою часткою ймовірності має місце група захворювань В". Оскільки в самому простому вигляді правило продукції

близьке за змістом логічної операції імплікації, для правила продукції часто застосовують позначення

$$A \rightarrow B$$

або, трактуючи A як сукупність деяких заданих умов $P_i, i = 1, 2, \dots, n$,

$$P_1 \wedge P_2 \wedge \dots \wedge P_n \rightarrow B,$$

де \wedge – символ кон'юнкції.

Кілька прикладів правил продукції, застосованих в різних ЕС медичного призначення, наведено в п.1.1.10. Зокрема, такий:

ЯКЩО: профіль щільності є "асиметричний гамма-пік"
і площа під кривою відповідає віковій нормі,

ТО: концентрація гаммаглобуліну в межах норми.

Неважко бачити, що подія A в цьому прикладі є кон'юнкція двох «елементарних» подій, тому цей же приклад можна переписати у вигляді:

ЯКЩО: профіль щільності є "асиметричний гамма-пік"
І

ЯКЩО: площа під кривою відповідає віковій нормі,

ТО: концентрація гаммаглобуліну в межах норми.

Система продукцій зручна для вираження знань, які можуть приймати форму переходів між станами (ситуація \rightarrow дію, посилка \rightarrow висновок, причина \rightarrow слідство).

Розрізняють продукційні системи, керовані даними (передумов правил) і продукційні системи, керовані цілями (діями правил).

База знань продукційної ЕС складається з безлічі правил продукції (бази правил) і кінцевого набору фактів (бази фактів).

Продукційні системи мають переваги і недоліки. Почнемо з недоліків. Виділимо з них три, з якими пов'язані певні обмеження, які у основі використовуваного формалізму. Вони відносяться до концепції, формулюванні та використанні правил.

1. Труднощі складання продукційного правила, відповідного елементу знання. Потрібно, щоб розглянута область вже була достатньо вивчена і встановлені хороші примітиви. І разом з тим, рівень деталізації не повинен бути надмірно докладним, інакше треба було б мати по одному правилу на кожен ситуацію, що істотно ускладнило б можливість розібратися людині в великій кількості інформації.

2. Труднощі запису правила. Єдиний формат запису ЯКЩО-ТО призводить до громіздким виразами в лівій частині і повторення тих же посилки у схожих ситуаціях; з його допомогою важко висловити складні правила.

3. Труднощі використання, пов'язані не з самими правилами, а з єдиною сис-

темою їх зв'язку. При цьому база фактів грає роль короткочасної пам'яті. У продукційній системі заборонений взаємний прямий виклик одного правила з іншого, і вона незручна для виконання алгоритмів у звичайному сенсі слова.

До переваг відносяться наступні.

1. Модульність. Системи влаштовані так, що кожна одиниця інформації може бути видалена, змінена або додана незалежно від всіх інших. У цьому сенсі знання вводяться невпорядковано. Практика показує, що це - природний спосіб поповнення знань для експертів.

2. Модифікованість. Якщо додається або модифікується будь-яке правило, то все, що було вже зроблено раніше, залишається в силі і до нового правилом не відноситься. Таким чином, кожна зміна має властивість адитивності і локальності. На відміну від цього, в «звичайних» комп'ютерних програмах, написаних на процедурному мовою, будь-яка модифікація є важким, що загрожує серйозними помилками.

3. Доступність читання і для людини, і для машини. Оскільки наше власне знання є модульним, продукційні системи нам здаються більш легкими для читання (у порівнянні з системами, заснованими на семантично мережах і фреймах) - це дуже важливо для таких кінцевих користувачів як лікарі, біологи, психологи. Для машини база знань у формі правил більш легко піддається перевірці, ніж у формі процедур. Суперечності, неузгодженість, надмірність легко визначаються за допомогою синтаксичної обробки.

4. Здатність до самопояснення. Ця властивість пов'язана одночасно і з правилами, і з їх структурами зовнішнього управління. Система легко простежує ланцюжок правил, яку вона використовувала для отримання висновку, так само як і що використовувалися метаправила.

5. Ефективність. Практика довела гнучкість і компетентність таких систем як MYCIN, PROSPECTOR. Відомий приклад випробувань системи MYCIN, коли п'ятеро незалежних фахівців у 72% випадків погодилися з висновками цієї системи, а в інших випадках самі не змогли прийти до єдиної думки. Також відомий приклад, коли система PROSPECTOR передбачила наявність покладів молібдену, а наступне розвідувальне буріння підтвердило передбачення.

3.3.2. Байєсова стратегія як форма реалізації правил продукції

Введене в рамках медичної кібернетики поняття «медичної пам'яті» (табл. 2.4) і поширені в медичній практиці діагностичні таблиці (прикладами є табл. 1.1, 2.6-2.10) можна трактувати як бази фактів, що встановлюють залежність між симптомами (ознаками) і захворюваннями.

Структура таблиць типу 1.1 і 2.6 зручна для реалізації загальної та диференціальної діагностики, яким властиві міркування виду "ознака → захворювання", реалізують "пряме" напрям логічних міркувань. При приватній діагностиці, коли потрібно перевірити наявність передбачуваного захворювання, реалізується "зворотне" напрям міркувань "захворювання →ознака" для перевірки наявності ознак, що підтверджують гіпотезу про наявність захворювання. Для цієї мети

також можна використовувати таблиці вищенаведених типів, хоча, для скорочення часу пошуку інформації, їх можна трансформувати. Наприклад, табл. 1.1 можна перетворити в табл. 3.1.

Разом з тим, таблиці виду 3.1 «незручні» для ЕОМ, що оперують з числами. Вони незручні і для самих лікарів, оскільки не ясно, наскільки істотно відрізняються, наприклад, градації «нехарактерно» і «малохарактерні». А кодування вербальних характеристик цілими числами, як це зроблено, наприклад, у табл. 2.5-2.7, хоча і зручно з точки зору ручних обчислень, однак лише в межах певних ситуацій, коли для діагностики використовується невелика кількість симптомів. Крім того, «пропадають» напрацьовані в процесі наукових досліджень численні статистичні дані, що зберігаються у звітах, дисертаціях, науковій літературі.

Таблиця 3.1

Захворювання	Ознаки	Характерність ознак
Інсульт геморагічний	Передвісники	Малохарактерні
	Раптовий розвиток	Характерний
	Повільний розвиток	Нехарактерний
	Втрата свідомості	Характерна
	Блідість обличчя	Нехарактерна
	Гіперемія	Характерна
	Артеріальна гіпотонія	Нехарактерна
	Артеріальна гіпертензія	Характерна
	Гіпертермія	Характерна
	Менінгізм	Характерний
	Кров у лікворі	Характерна
	Крові у лікворі немає	Нехарактерно

Найбільш природним виходом з цієї ситуації видається використання імовірнісної форми правила продукції. Тим більше, що при розробці алгоритмів медичної діагностики знайшли достатньо широке застосування:

- імовірнісний метод, що складається в обчисленні ймовірностей захворювань за формулою Байєса;
- метод послідовного статистичного аналізу Вальда;
- метод пошуку клінічного прецеденту (іменований іноді "статистичним методом").

З точки зору форми представлення кінцевого результату, принципової відмінності між перерахованими методами немає - при практичній реалізації всі вони зводяться до оцінювання умовної ймовірності $P(Y_j/X_i)$, де Y_j - найменування j -того захворювання, X_i - i -тої ознаки.

Витоки такого підходу можна помітити, по-перше, в типах, які використовуються в медицині, психіатрії та психології шкал ознак і нозологічних форм -

це шкали найменувань, для яких підрахунок частоти (ймовірності) чи не єдиний тип допустимих математичних операцій. Друга причина - існування добре розробленого математичного апарату трансформації шкал найменувань, іменованою формулою Байєса і яка забезпечує оптимальність (у сенсі максимізації достовірності) прийнятого рішення:

$$P(Y_j / X_i) = \frac{P(Y_j)P(X_i / Y_j)}{P(X_i)}, \quad (3.1)$$

де $P(Y_j)$ и $P(X_i)$ – апіорні ймовірності захворювань Y_j і ознак X_i ; $P(X_i/Y_j)$ – умовна ймовірність появи ознаки - симптому X_i для захворювання Y_j ; $P(Y_j/X_i)$ – апостеріорна ймовірність захворювання Y_j при спостереженні ознаки X_i .

Співвідношення (3.1) є одним з правил бази правил (прикладами інших правил є співвідношення (2.2)-(2.8)) експертної системи, що розробляється з використанням елементів Байєсових стратегій.

Зазначимо властивість симетрії формули Байєса: з (3.1) легко вивести співвідношення

$$P(X_i / Y_j) = \frac{P(X_i)P(Y_j / X_i)}{P(Y_j)}. \quad (3.2)$$

Зіставляючи (3.1) і (3.2), приходимо до висновку, що формула Байєса дозволяє практично за однією і тією ж схемою обчислень здійснювати висновки як в прямому ("ознака → захворювання"), так і в зворотному ("захворювання → ознака") напрямках. Причиною тому - принципова особливість формули Байєса, яка встановлює зв'язок не між "причиною" і "слідством", а між двома довільними подіями. Співвідношення (3.2) можна, в принципі, використовувати для ранжування ознак окремих захворювань (при приватній діагностиці, наприклад). Крім того, зазначена властивість симетрії дозволяє трохи відійти від звичного відстеження причин і наслідків, згадавши про таке явище, як синхронізм явищ у природі (наприклад, одна і та ж спалах на Сонці може породити як загострення існуючих захворювань, так і послужити "спусковим гачком" для виникнення нових захворювань в ослабленому організмі; тому не слід поспішати оголошувати виявлені залежності як причинно-наслідкові - залежності такого роду можуть бути і відсутніми).

Для множини статистично незалежних ознак $X = \{X_1, \dots, X_l\}$ формулу Байєса в рекурентному її варіанті

$$P(Y_j / X_1, \dots, X_l) = P(Y_j / X_1, \dots, X_{l-1}) \frac{P(X_l / Y_j)}{P(X_l)} \quad (3.3)$$

можна розглядати як своєрідний ланцюжок висновків наступного виду.

Нехай лікар спостерігає деяку ознаку X_l . Проводячи обчислення по формулі (3.3), отримуємо апостеріорний розподіл ймовірностей $P(Y_j/X_l)$, $j=1, \dots, J$, орієнтуючись на максимум якого можемо зробити висновок про наявність деякого

захворювання Y_j . Символічно правило продукції при цьому можна записати наступним чином: $X_1 \rightarrow Y_j$.

Однак значення апостеріорної ймовірності $P(Y_j/X_1)$ при цьому може виявитися недостатньо високим для впевненого висновку про наявність захворювання. Тоді необхідно по результатам спостереження наступної ознаки X_2 повторити обчислення по формулі (3.3), спостерігаючи за змінами максимуму апостеріорного розподілу ймовірностей $P(Y_j/X_1, X_2)$. В загальному випадку цей максимум може зміститися – тоді можемо зробити висновок про наявність іншого захворювання Y_r . Схема висновків при цьому набуває вигляду, як показано на рис.3.1.

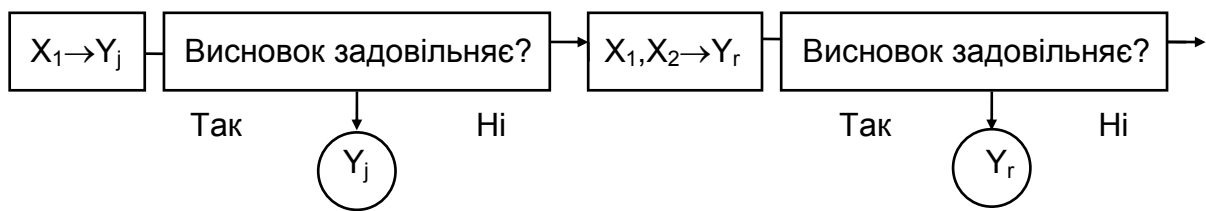


Рис. 3.1

Така схема виведення справедлива для випадку, коли рівень пошуку незмінний - згідно зі схемою рис. 1.4, це може бути рівень класів захворювань при попередній діагностиці, або рівень нозологічних форм при загальній діагностиці, або розташовуються ще нижче рівні диференціальної діагностики.

При переході до все більш детальної діагностики відбувається зміна рівнів, як показано на рис.3.2.

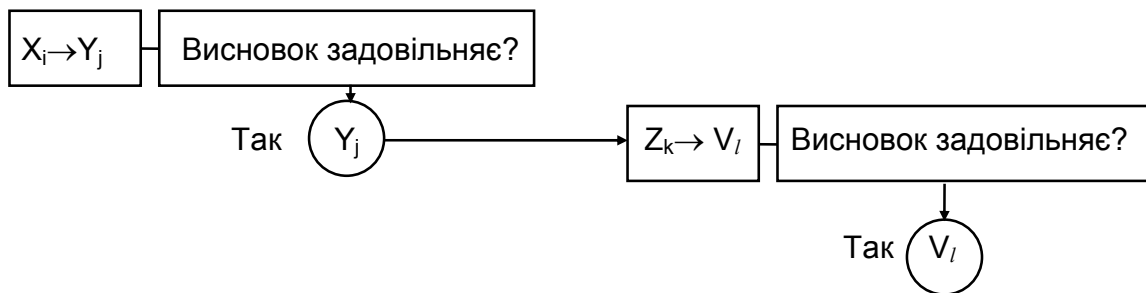


Рис. 3.2

Зміна рівня діагностики, природно, призводить до зміни множин симптомів і відповідних їм захворювань. Ланцюжка міркувань, подібні наведеним вище, є *прямими ланцюжками*, що забезпечують просування від симптомів до захворювань.

Разом з тим, при приватній діагностиці, коли перевіряється припущення про наявність деякого захворювання, виникає необхідність у використанні *зворотних ланцюжків міркувань*. При цьому обчислення проводяться з використанням співвідношень, подібних (3.2).

По суті, як показано на рис.1.6 і рис.3.3, приватну діагностику можна пред-

ставити у вигляді двох етапів:

1) використовуючи зворотний хід міркувань, по заданому припущенням про захворювання Y_j обчислюють розподіл ймовірностей $P(X_i/Y_j)$ - тим самим визначають не тільки перелік симптомів $X=\{X_1, \dots, X_I\}$, супутніх даному захворюванню, а й отримують інформацію про найбільш діагностично цінних симптомах (їм відповідають найбільші значення розподілу $P(X_i/Y_j)$);

2) виробляючи прямий хід міркувань і починаючи з перевірки найбільш діагностично цінних симптомів, обчислюють значення ймовірності $P(Y_j/X_1, X_2, \dots)$ діагностується захворювання Y_j .

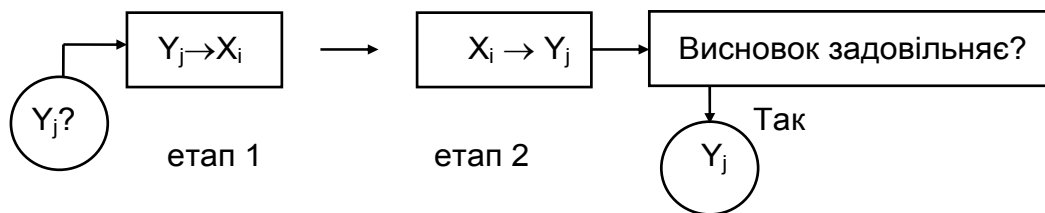


Рис. 3.3

3.3.3. Семантичні мережі

Семантичні мережі - різновид баз знань, спочатку при змінених в психології для моделювання людської пам'яті. Тепер це один із стандартних методів подання знань у системах ШІ і ЕС.

Семантичні мережі є різновид графів - об'єктів у вигляді "вузлів", пов'язаних "дугами". Вузли використовують, наприклад, для моделювання об'єктів, концепцій, подій, робіт, операцій ... Дуги відповідають відносинам між вузлами - типи відносин залежать від виду знань.

Звичайні дуги, використовувані для представлення ієрархії, включають дуги типу ϵ і має частину (рис.3.4).

Пропозиції «Національний технічний університет України (КПІ) є технічним вузом» і «Кожен технічний вуз входить в систему вищої освіти» представлені на рис.3.4 у вигляді елементів мережі з використанням важливого типу дуг: ϵ . Оскільки ми знаємо, що транзитивність є властивість дуг ϵ , ми можемо вивести з мережі третього твердження: «Національний технічний університет України (КПІ) входить в систему вищої освіти». Таким чином, ставлення ϵ , як і ставлення *має частину*, встановлює властивість *ієрархії спадкоємства* в мережі. Це означає, що елементи нижчого рівня в мережі можуть успадковувати властивості елементів більш високого рівня в мережі. Це економить пам'ять, оскільки інформацію про пов'язані вузлах не потрібно повторювати для кожного вузла мережі.

Наприклад, у семантичній мережі, що становить систему вищої освіти, такі його частини, як факультет, кафедра, бібліотека, включені один раз на високому рівні ієрархії. В результаті, використовуючи знання мережі, неважко встановити факти на кшталт «Медичний університет має бібліотеку». Семантичні ме-

режі є одним з найбільш вдалих методів подання знань з добре встановленої таксономією і дозволяють істотно спростити пошук необхідної інформації.

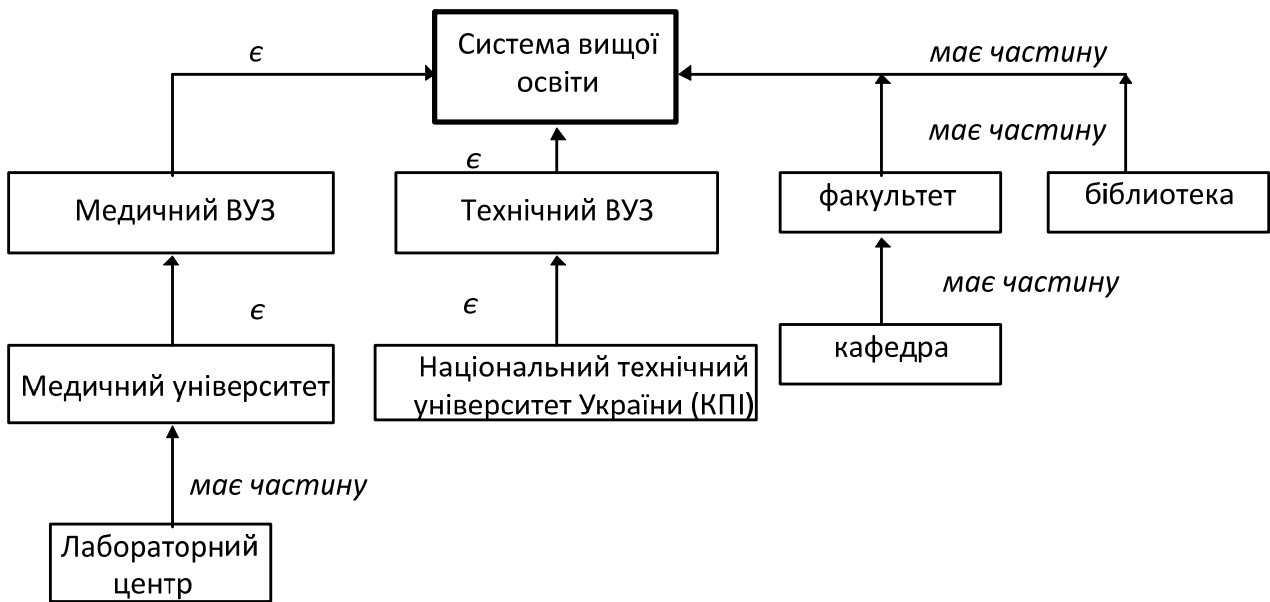


Рис. 3.4

Семантичні мережі також успішно використовуються в роботах по природному мови для представлення складних граматичних конструкцій - з цією метою використовують дуги типу *агент*, *об'єкт*, *реципієнт*.

Приклад подання пропозиції «Викладач допустив студента X до іспиту» за допомогою семантичної мережі показаний на рис.3.5.

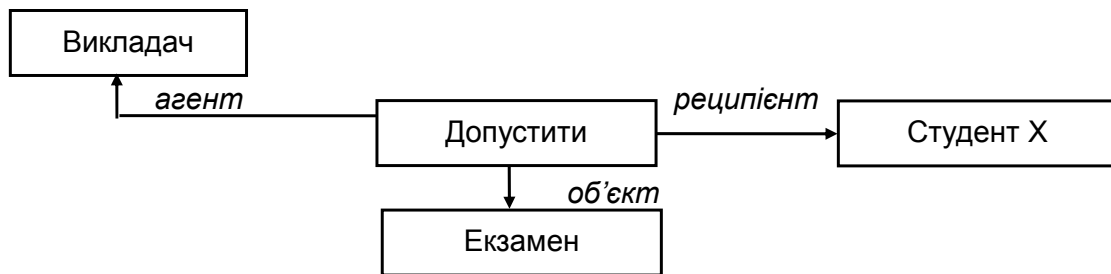


Рис. 3.5

Зазначимо, що тут дуги визначають відносини між присудком (ДОПУСКАТИ) і поняттями (такими як СТУДЕНТ X і ІСПИТ), пов'язаними з цим присудком. Цей же метод можна використовувати для представлення більш складної пропозиції «Викладач повідомив в деканат, що він не допустив студента X до іспиту», показаного на рис.3.6.

Подання через семантичну мережу корисно, оскільки воно забезпечує стандартний метод аналізу сенсу пропозиції. Крім того, воно вказує схожість в сенсі пропозицій, тісно пов'язаних, але що володіють різною структурою. Хоча про-

позиції на рис.3.5 та 3.6 виглядають дуже відрізняються один від одного, але семантичні мережі, що представляють зміст цих пропозицій, схожі. Так, семантична мережа на рис.3.5 цілком міститься в мережі рис.3.6.

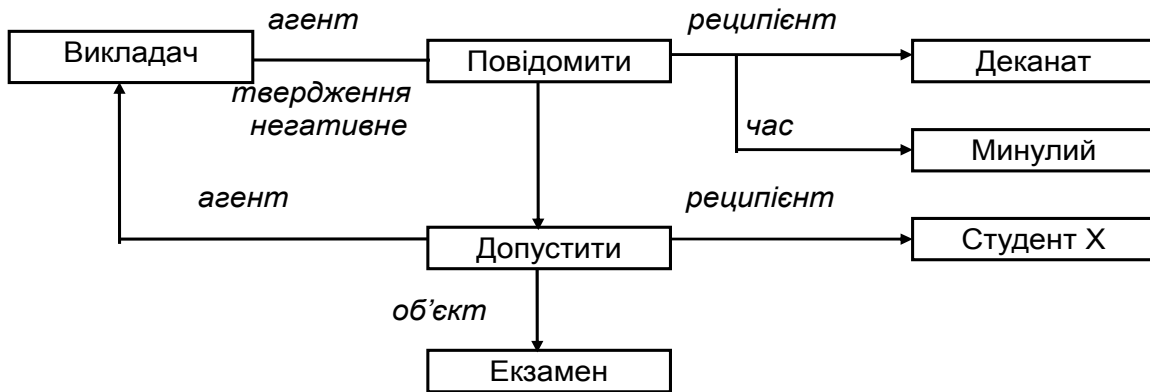


Рис. 3.6

В ЕС для планування та управління широко використовуються мережі типу «вершини - роботи». Можливі типи відносин у цьому випадку: «фініш-старт», «старт-старт», «фініш-фініш». Приклад гіпотетичного календарного плану у формі логічної мережі такого роду, з відносинами типу «фініш-старт» між роботами і з можливими затримками початку робіт (показаними у відповідних дуг), представлений на рис.3.7, де прийнята наступна система позначень:

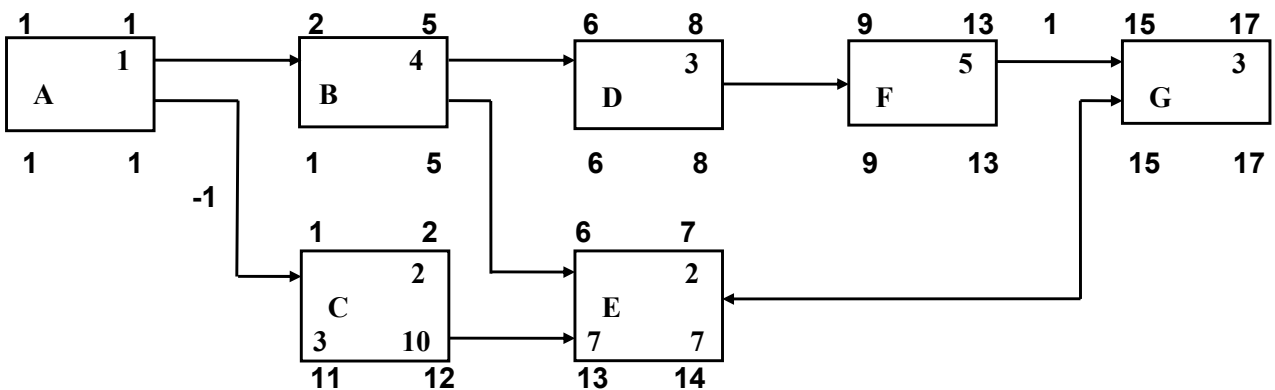
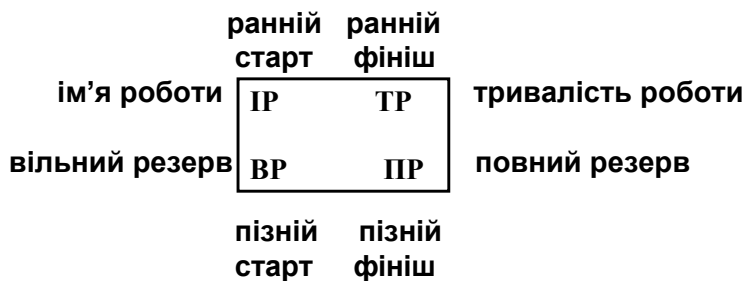


Рис. 3.7

Календарні плани такого типу використовуються проєкт-менеджерами для управління часом, людьми і ресурсами (матеріальними та грошовими) при реалізації великих проєктів в галузі медицини, освіти, промисловості та ін.

У силу значної складності такого управління створюються спеціальні СУБЗ, що представляють собою, по суті, ЕС.

Наприклад, світовим лідером з розробки таких ЕС є американська фірма Primavera, яка створила за останнє десятиліття ряд потужних ЕС: P3, SureTrak та ін. Ціна таких систем від \$ 700 до \$ 4000, на їх розробку витрачалося від 10 до 15 людино-років, витрати на розробку однієї системи склали \$ 11-16 млн.

Як приклад високої ефективності застосування ЕС такого роду можна привести використання мережевої версії програми P3 (на 8 проєкт-менеджерів) для оперативної ліквідації наслідків терористичного акту в торговому центрі Манхеттена (США, 1995р.). Роботи з відновлення торгового центру були завершені протягом 2-х тижнів, при цьому витрачено \$ 22 млн. - настільки значні витрати обумовлені не тільки великими масштабами руйнувань, але і жорсткими тимчасовими обмеженнями.

Переваги семантичних мереж:

- наочність представлення відносин;
- цілісність опису ситуації у всій її складності;
- швидкість проектування бази знань.

Недоліки:

- занадто довільна структура мереж;
- різнотипність вузлів і відносин, що призводить до великої різноманітності процедур обробки інформації;
- складність і різноманіття типів СУБЗ.

3.3.4. Фрейми

Фрейми по своїй організації багато в чому схожі на семантичні мережі особливого типу:

1) з ієрархічною організацією вузлів і, як наслідок, з успадкуванням вузлами нижніх рівнів властивостей вузлів верхніх рівнів - цим досягається економія пам'яті ЕОМ;

2) фрейм є сукупність слотів (атрибутів), тобто має структуру виду $f[⟨v_1, g_1⟩, ⟨v_2, g_2⟩, \dots, ⟨v_n, g_n⟩]$, де f - ім'я фрейму; пара $⟨v_i, g_i⟩$ - i -тий слот; v_i - ім'я слота, g_i - його значення;

3) фрейм має властивість вкладеності: значенням слота може бути система імен слота більш глибокого рівня, посилання на інші слоти того ж кадру і т.д.;

4) з кожним слотом можна пов'язати будь-яку кількість процедур, з яких частіше за інших зустрічаються:

- "якщо додано" - виконується, коли нова інформація поміщається в слот;
- "якщо видалено" - виконується, коли нова інформація видалена із слота;
- "якщо треба" - виконується, коли запитується інформація з слота, а він порожній.

Для описових наук фрейми - один з небагатьох способів формалізації, створення понятійного апарату.

Розглянемо наступний приклад. Припустимо, якийсь керівник запитує у ЕС з фреймовий уявленнями баз знань: "Мені потрібен звіт про просування проекту з модернізації іридодіагностичної експертної системи ЭСИД-2".

Інтерфейс аналізує запит і вносить фразу "Проект іридодіагностичної системи ЭСИД-2" в слот "тема" порожнього вузла гілки "звіт про просування". Далі все відбувається автоматично (див. рис.3.8):

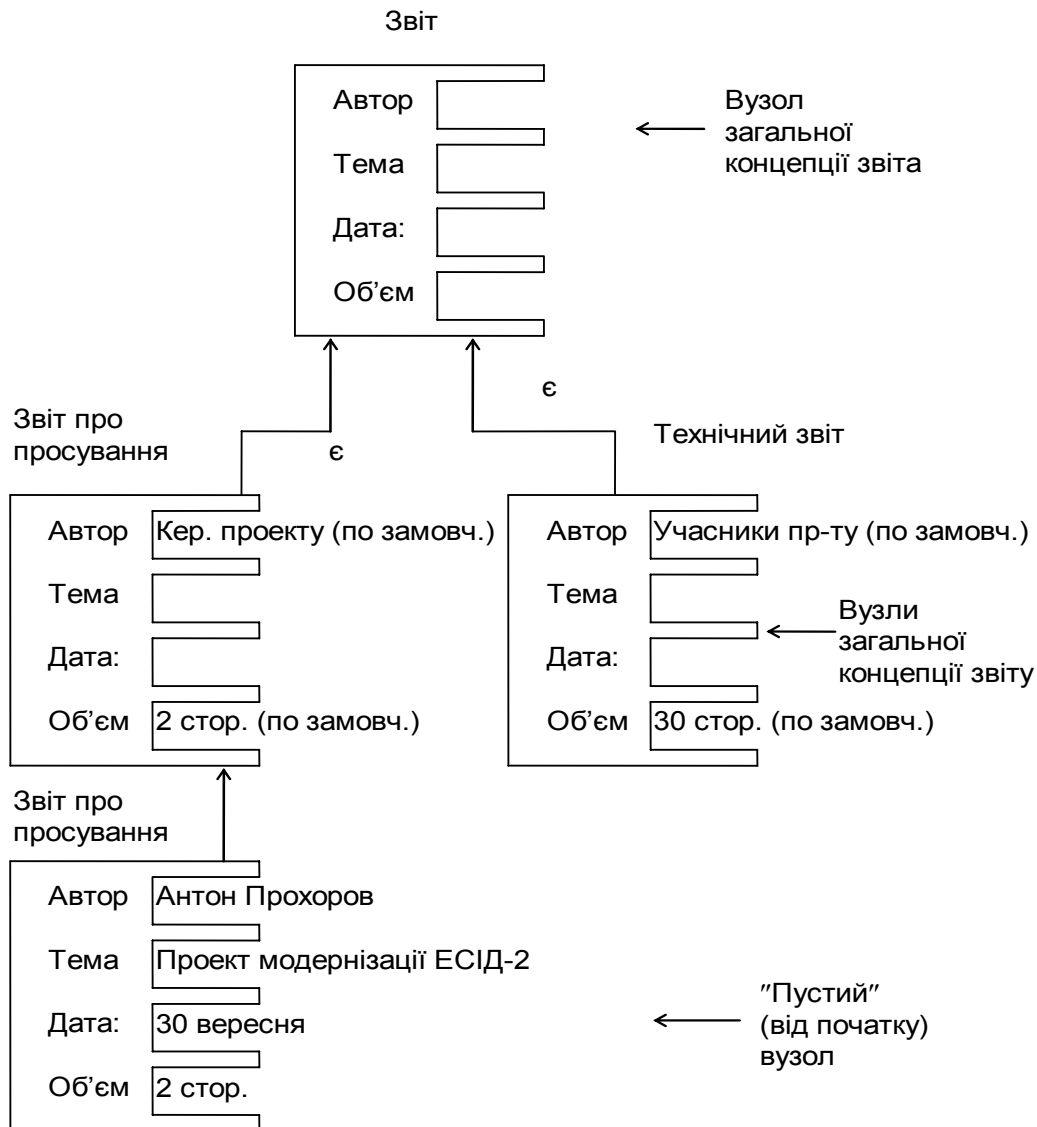


Рис. 3.8

- процедура "якщо додано" слоту "тема" здійснює пошук у БД імені керівника проекту - припустимо, його звать Антон Прохоров;

- процедура "якщо додано" слоту "автор" починає складати повідомлення Антону Прохорову, але виявляє відсутність значень "дата подання" і "обсяг", після чого активізує процедуру "якщо треба", яка: а) вибере дату, орієнтуючись на по-

точну дату і на календар типових дат подання звітів, вписавши її в слот "дата", б) запозичує значення обсягу з вузла загальної концепції звіту про просування;

- складається повідомлення "Антон Прохоров, будь ласка, закінчите звіт про просування робіт за проектом модернізації ЭСИД-2 до 30 вересня", яке автоматично відправляється Антону Прохорову.

3.4. Проблеми використання Байєсових стратегії (на прикладі іридодіагностичних ЕС)

Закономірне питання: якщо Байєсова стратегія оптимальна, чому замість неї в медичній діагностиці, як правило, воліють використовувати завідомо гірші методи обробки типу табличних алгоритмів?

Для визначеності розглянемо ситуацію в іридодіагностиці: у літературі міститься, з одного боку, величезна статистика зустрічальності іридологических ознак при різних захворюваннях, що створює серйозні передумови для автоматизації іридодіагностики з використанням Байєсових стратегії. З іншого боку, в ряді випадків теоретичні дослідження завершуються описом кількох комп'ютерних програм, що реалізують при діагностуванні тільки метод підсумовування балів, принципово поступається Байєсових стратегії в плані надійності діагностування.

Парадоксальність зазначеної ситуації особливо очевидна, якщо врахувати надзвичайно високу потужність сучасних персональних комп'ютерних засобів, що дозволяють зберігати і швидко обробляти величезні обсяги даних і для яких реалізація обчислень відповідно до Байєсових стратегією представляється завданням цілком вирішуваною.

Разом з тим, ретельний аналіз ситуації дозволяє виявити причини використання в іридодіагностиці простих малоефективних табличних алгоритмів замість більш складних, але і більш ефективних, алгоритмів, заснованих на Байєсових стратегії.

Перелік цих причин:

- статистична залежність між іридопрізнаками;
- необхідність знання апіорних ймовірностей $P(Y_j)$ захворювань Y_j (принаймні, такі знання дуже бажані);
- "дефекти" даних, обумовлені неоднорідністю і неповнотою даних;
- існування зовнішніх і внутрішніх "заважають" факторів.

Зробимо детальний аналіз цих причин.

3.4.1. Метод іридодіагностики

Метод іридодіагностики, заснований на сигнальній функції екстерорецепторів райдужки ока, будучи одним з найбільш інформативних і достовірних методів раннього виявлення генетичних і патологічних зрушень в організмі, займає помітне місце в сучасному медичному світі. Відсутність будь-яких протипоказань (крім епілепсії, як відносного протипоказання у зв'язку з провокуючим напад дією світла), повна безпека і нетравматичність в поєднанні з високою дос-

товірністю і об'єктивністю одержуваних результатів обстеження забезпечили широке впровадження цього методу в систему охорони здоров'я.

Ґрунтуючись на концепції П.Анохіна про систему кільцевих рефлекторних зв'язків, теорії функціональних систем організму, нових досягнень у галузі вивчення функціонування вегетативної нервової системи, ретикулярної формації головного мозку, біоенергетиці організму і величезного, більш ніж двовікового експериментального матеріалу вчених Європи, Америки (В.Йенсен, Ф.Віда і Л.Дек, К.Гюнтер, Ф.Робертс, А.Кастро, П.Дімков, Р.Бурдіоль, Й.Анжерер, Е.Вельховер, Ф.Ромашов та ін), в даний час обґрунтовано механізм відображення на райдужці ока генетичних особливостей людини, механізм появи і формування іридоознак на райдужці, що відображають різні патологічні стани в організмі.

Меланоцити райдужки, представляючи анатомічно окреслені виходи на периферію вісцеральних рефлекторно-кільцевих систем, виконують найважливішу місію рецепторів-блокаторів, які через системи трійчастого, симпатичного і парасимпатичного нервів здійснюють зв'язок навколишнього світлової середовища з трігемноретикулярним комплексом, найважливішими симпатичними гангліями і через них - з різними внутрішніми органами. У гострій стадії захворювання больові імпульси з осередку ураження, проходячи через трігемноретикулярні центри, впливають на свої меланоцити на райдужці і змушують їх посилено функціонувати, що призводить до появи ділянок просвітлення в певних зонах на райдужці ока, топичні пов'язаних з ураженими органами. При переході захворювання в хронічну стадію змінюються умови функціонування всієї кільцевої рефлекторної системи. Імпульси, що надходять з ураженого органу до своїх меланоцитам, викликають їх перезбудження і спазм дрібних судин, в результаті чого настає гіпоксія і трофічні розлади в меланоцитах і тканинах райдужки. Пошкоджені меланоцити лізуються макрофагами, перетворюючись у функціонально недієльному гіперпігментовані плями, завдяки яким частково перекривається доступ світловим імпульсам, тобто протікає виборча захисна реакція організму, спрямована на оберігання ураженого органу від світлової дії.

При огляді пацієнта лікар іридолог, оцінюючи структурний стан райдужки ока, оцінюючи адаптаційно-трофічні зміни, які у ній в часі і просторі, маючи можливість одномоментного огляду в одному полі зору проєкційних зон (один з різновидів схем проєкційних зон тіла людини на райдужній оболонці ока, що належить Дімкову, наведена на рис.3.9) всього організму, діагностує з досить високою точністю спадкові особливості пацієнта, функціональну та органну слабкість певних органів і систем, що дозволяє в кінцевому рахунку зробити висновки про резервних можливостях організму, скласти прогноз, тобто побудувати вектор майбутнього стану здоров'я організму обстежуваного. На підставі поєднань різних іридоознак на райдужці іридолог діагностує місце розташування патологічного процесу і, в деякій мірі, його характер.

Різноманіття схем проєкційних зон, іноді суперечать один одному, індивідуальні особливості будови нервової системи людини, що призводять до зсуненості проєктуємих на райдужку знаків, наявність великої кількості взаємопов'яза-

них комбінацій іридологічних знаків - різнокольорових пігментних плям, кільць, лакун, змін у формі автономного кільця, зіниці, зіничної облямівки і т.д. - Всі ці фактори роблять діагностику за іридологічними ознаками досить складною для практичного лікаря, вимагаючи від нього високої кваліфікації і постійної, максимально можливої психофізіологічної мобілізації.

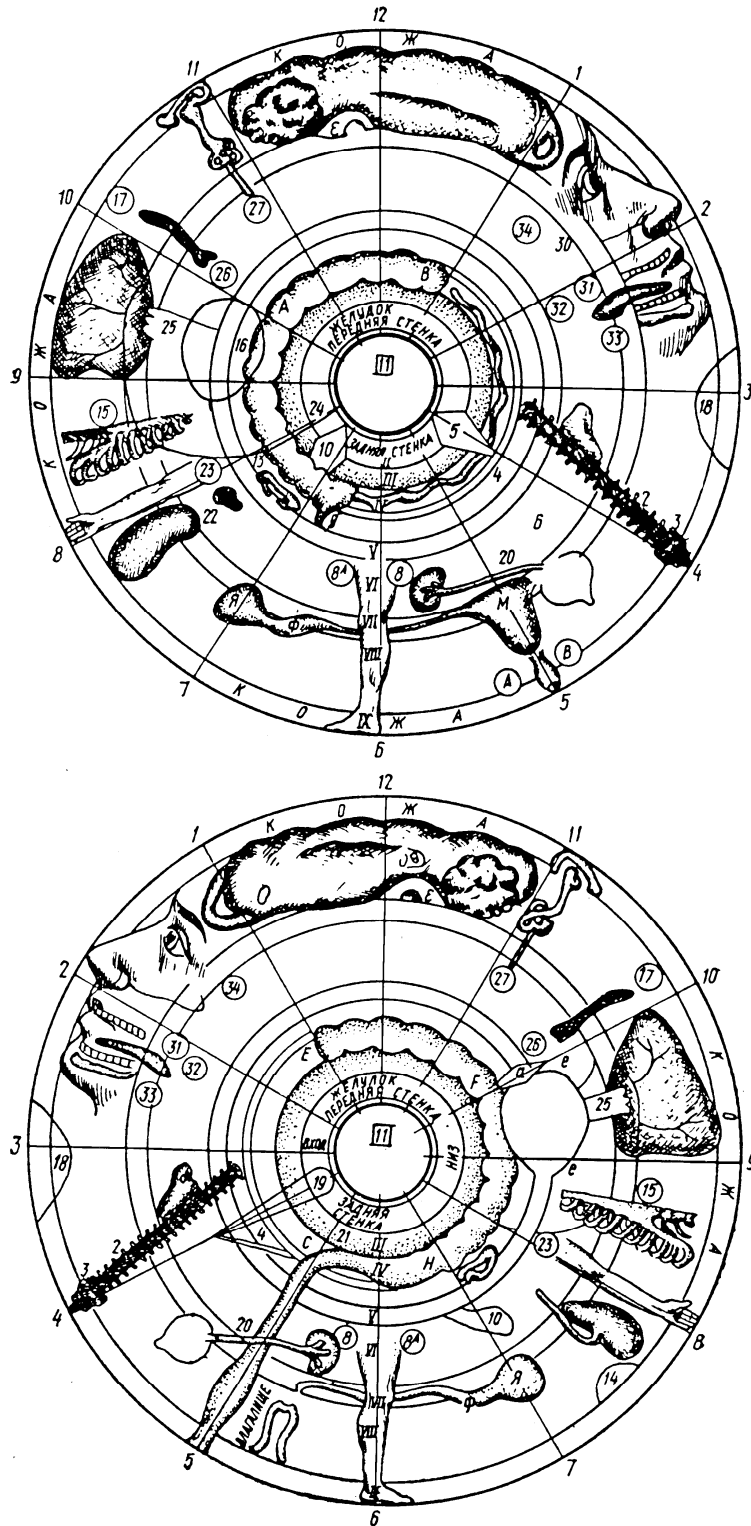


Рис. 3.9

Разом з тим, як зазначалось вище, в літературі наводиться величезна статистика характерності іридологічних ознак при різних захворюваннях, що створює серйозні передумови для розробки іридодіагностичних експертних систем, що базуються на використанні Байєсової стратегії.

3.4.2. Статистична залежність ознак

Прості та зручні для обчислень співвідношення (2.3), (2.4), (3.2) справедливі в припущенні статистичної незалежності використовуваних ознак. У разі статистично залежних ознак необхідно використовувати більш складне вираз (1.2), складність якого полягає в необхідності знання багатовимірних густин розподілу ймовірностей $P(X_1, X_2, \dots, X_l)$ і $P(X_1, X_2, \dots, X_l/Y_j)$.

Між іншим, в літературі знаходимо, головним чином, оцінки одновимірних розподілів ймовірностей $P(X_i)$ і $P(X_i/Y_j)$, які придатні лише для обчислень за формулами (2.3), (2.4), (3.2), тобто у припущенні статистичної незалежності ознак. І разом з тим, в літературі зазначається наявність статистичної залежності між ознаками - як наслідок, формули (2.3), (2.4), (3.2) не можуть бути використані.

На підтвердження наведемо кілька прикладів.

По-перше, розглянемо фрагмент таблиці "Частота народження різних типів райдужки у людей з різним кольором очей" (табл.3.2), аналіз якого дозволяє зробити висновок про існування залежності між кольором і типом за Е.С.Вельхверу (між щільністю райдужної оболонки і формою її волокон): частота народження радіально-гомогенного (найбільш "благополучного" в плані рівня здоров'я, обумовленого генетичними факторами) типу райдужки тим вище, чим темніше райдужка.

Таблиця 3.2

Колір райдужки	Радіальний тип	Радіально-хвильовий тип	Радіально-гомогенний тип	Радіально-лакунарний тип	Лакунарний тип
Блакитний	5,6	78,0	1,7	8,7	6,0
Синій	6,1	66,9	—	17,0	10,0
Сірий	4,9	81,6	—	8,2	5,3
Світло-коричневий	1,2	12,0	83,8	2,1	0,9
Темно-коричневий	—	—	88,0	7,8	4,2

Залежність між кольором і щільністю райдужної оболонки зустрічаємо також в описі ознаки "щільність райдужки" по В. Jensen, а в описі типів райдужки по J. Deck виявляємо існування залежності не тільки між кольором, щільністю і формою волокон райдужки, а й іншими іридоознаками: лімфатичних розарієм ("гідрогеноїдний" тип), кільцями освітлення навколо автономного кільця ("чисто лімфатичний" і "гідрогеноїдний" типи), кільцями білувато-сірих платівок по

периферії райдужки (тип "сечокислий діатез"), зашлакованістю автономного кільця (тип "феррум-хроматоз") і т.д.

Крім того, встановлено наявність залежності частоти виникнення нервових кілець від кольору райдужки, від товщини і зашлакованості автономного кільця, від форми зрачкової облямівки.

При зіставленні форми зрачкової облямівки (ЗК) з формою АК і зрачкового поясу (ЗП) виявлено, що чистому, з ясним малюнком і кордонами АК і ЗП відповідає густа і широка ЗК. І навпаки, зашлакованим, гіпертрофованим АК і ЗП відповідає менш густа і більш вузька ЗК.

При аналізі залежності між кольором райдужної оболонки і формою ЗК виявлено, що в молодому віці у світлооких людей ширина ЗК в 1,5-2 рази більше, ніж у людей з темними очима (після 35-40 років ця різниця нівелюється).

З приводу інформативності пігментних плям відзначено, що найчастіше одиначне пляма не містить в собі інформації про наявність захворювання, але якщо воно розташоване поблизу або в товщі іншого іридологічного знаку (в глибині лакуни, між розщепленими трабекулами і т.д.), то набуває діагностичний сенс. Існує залежність між лакунами, пігментними плямами і кольором райдужної оболонки: у бідних пігментом світлих райдужках переважають морфологічні зміни, тоді як в багатих пігментом темних райдужках - колірні.

Резюмуючи наявні дані про статистичної залежності іридоознак, відзначимо:

1) між іридоознаками існує статистична залежність, причому можна виділити принаймні два основних механізми виникнення такої залежності: "фізіологічний" і "математичний"; в першому випадку залежність зумовлена або проявом одного і того ж захворювання у вигляді декількох ознак, або проявом ознак кількох залежних захворювань, у другому випадку це залежність між "сумою" і "доданками", тобто між комплексною іридоознакою, утвореною сукупністю елементарних іридоознак, і які входять до його складу елементарних іридоознак;

2) до теперішнього часу в найбільшою мірою вивчена залежність між різними іридоознаками і ознакою "колір райдужки", що, мабуть, пояснюється й не так інформативністю ознаки "колір райдужки", скільки простотою і легкістю його оцінювання.

Фізична природа статистичної залежності ознак може бути різною. Відзначимо лише два найбільш відомих і вивчених типу такої залежності:

- каузальність (причинно-наслідкова залежність);
- синхронізм.

У першому випадку поява ознаки X тягне за собою, з деякою вірогідністю, появу іншої ознаки Y . У другому випадку мається на увазі наявність третього, прихованого від спостереження (латентного) або просто ігнорованої події Z , каузально пов'язаної з ознаками X і Y , які в результаті такого зв'язку виявляються статистично залежними.

Перейдемо тепер до математичної сторони статистичної залежності.

Оцінити характер і ступінь статистичної залежності ознак X і Y можна, використовуючи поняття регресії і коефіцієнтів регресії. Так, регресією Y на X називають

вається умовне математичне очікування (МО) випадкової величини (ВВ) Y при фіксованому значенні $X=x$:

$$E\{Y(x)\}=E\{Y/X=x\}.$$

Лінією регресії Y на X називається це МО, що розглядається як функція змінної x . Аналогічно визначається регресія X на Y .

Лінії регресії Y на X і X на Y , взагалі кажучи, не збігаються. Регресія називається лінійною, якщо лінія регресії пряма. Для незалежних ВВ лінії регресії перетворюються на прямі, паралельні координатним осям.

Якщо в розглянутій вище табл. 3.2 позначити колір райджі символом X , а тип райджі - символом Y , тоді легко бачити, що в табл. 3.2 наведені значення умовної щільності $P(Y/X)$. Враховуючи істотну нерівномірність цієї функції Y (при фіксованих значеннях X), наближено оцінити її середнє значення (математичне очікування) не представляє праці - йому відповідає максимум щільності $P(Y/X)$ як функції Y . У результаті приходимо до наведеного вище висновку: частота народження радіально-гомогенного (найбільш "благополучного" в плані рівня здоров'я, обумовленого генетичними факторами) типу райджі тим вище, чим темніше райджка. Тим самим наявність статистичної залежності між кольором райджної оболонки і типом її по Е.С.Вельховеру можна вважати встановленим.

Характер статистичної залежності між ознаками може бути як лінійним, так і нелінійним. У разі лінійної залежності зручно користуватися поняттям "коефіцієнт кореляції"

$$r_{XY}=E\{(X-E\{X\})(Y-E\{Y\})\}/\sigma_X\sigma_Y,$$

де σ_X і σ_Y – середньквдратичне відхилення ВВ X і Y :

$$\sigma_X=[E\{X-E\{X\}\}^2]^{1/2}; \sigma_Y=[E\{Y-E\{Y\}\}^2]^{1/2}.$$

Це зручно, оскільки $|r_{XY}|<1$, причому рівністю $r_{XY}=0$ має сенс для некорельованих (і незалежних – у випадку нормального розподілу X і Y) ВВ, а $|r_{XY}|=1$ – для лінійно залежних детермінованих ВВ.

3.4.3. Неповнота апіорних відомостей

Другою істотною перешкодою на шляху реалізації обчислень за формулою Байєса є необхідність знання апіорних ймовірностей $P(Y_j)$ захворювань Y_j . Принаймні, такі знання дуже бажані: якщо їх немає, можна прийняти всі гіпотези рівноімовірними, тобто $P(Y_j) = 1/J$, де J - кількість альтернативних захворювань. Однак це призведе або до недостатньо високої достовірності висновків (при фіксованій кількості спостережуваних іридоознак), або потребуватиме збільшення обсягу спостережень (при фіксованій достатньо високій достовірності висновків). Зауважимо, що в деяких літературних джерелах недостатня обґрунтованість апіорного розподілу взагалі оголошується основною перешкодою на шляху використання формули Байєса в додатках. На наш погляд, це явна пере-

оцінка ролі апіорної інформації, однак сумнівів не викликає і той факт, що цінність такої інформації дуже висока.

Як приклад наведемо спробу "перекладення" на мову формули Байеса статистичної інформації про деякі ознаки ниркової патології. "... Характерним для ниркової патології симптомом був лімфатичний розарій, який виявляли у обстежених хворих у 57% випадків ... При захворюваннях легенів, шлунково-кишкового тракту, серцево-судинної та нервової систем лімфатичний розарій виявляли набагато рідше, ніж при хворобах нирок, в 9-22% випадків ... Вказана обставина дозволяє лікарю міркувати таким чином: при всякому виявленні лімфатичного розарію на райдужці можна припускати, але жодною мірою не можна стверджувати, що у даного хворого є зміни з боку нирок".

У даному конкретному випадку явно не враховується частота народження ниркової патології взагалі, безвідносно до будь-якої сукупності діагностичних ознак, отже апіорні ймовірності гіпотез Y_1 ="є захворювання нирок" і Y_2 = "немає захворювань нирок" можна прийняти однаковими: $P(Y_1)=P(Y_2)=0,5$.

Умовні ймовірності $P(X_i/Y_j)$ дані і є рівними:

$$P(X_1/Y_1) = 0,57; P(X_1/Y_2) = 0,155,$$

де X_1 означає "є лімфатичний розарій", а значення величини $P(X_1/Y_2)=0,155$ отримано як середнє арифметичне значення 0,09 і 0,22.

Враховуючи (3.1),

$$P(Y_1 / X_1) = \frac{P(Y_1)P(X_1 / Y_1)}{P(X_1)} \quad (3.4)$$

Оскільки

$$P(X_i) = \sum_j P(Y_j)P(X_i / Y_j), \quad (3.5)$$

з (3.4) слід:

$$P(Y_1 / X_1) = \frac{P(Y_1)P(X_1 / Y_1)}{\sum_j P(Y_j)P(X_1 / Y_j)} = \frac{P(X_1 / Y_1)}{\sum_j P(X_1 / Y_j)} = \frac{0,57}{0,57 + 0,155} = 0,8.$$

Аналогічно

$$P(Y_2 / X_1) = \frac{P(X_1 / Y_2)}{\sum_j P(X_1 / Y_j)} = \frac{0,155}{0,57 + 0,155} = 0,2.$$

Таким чином, нерівність $P(Y_1/X_1) > P(Y_2/X_1)$ відповідає виразу "можна припускати наявність захворювання нирок", а той факт, що $P(Y_1/X_1) < 1$, відповідає фразі "але жодною мірою не можна стверджувати". Слова "жодною мірою" - свідчення надзвичайної обережності автора: адже все-таки у 8 випадках з 10

твердження виявиться справедливим. Разом з тим, цю обережність можна зрозуміти - йдеться про достовірність діагностування захворювання, яке потім доведеться лікувати.

Врахуємо тепер апіорну інформацію у вигляді розподілу $P(Y_j)$ і оцінимо її вплив на вагомість одержуваних висновків. За результатами профілактичного огляду школярів у віці 12-17 років відомо, що патологія нирок притаманна 68% обстежених. Приймаючи у виразі (3.4) $P(Y_1) = 0,68$ і $P(Y_2) = 0,32$, отримаємо

$$P(Y_1 / X_1) = \frac{0,68 \cdot 0,57}{0,68 \cdot 0,57 + 0,32 \cdot 0,155} = 0,9.$$

$$P(Y_2 / X_1) = 1 - P(Y_1 / X_1) = 0,1.$$

Таким чином, завдяки обліку апіорної інформації ймовірність виведення на користь гіпотези $Y_1 =$ "є захворювання нирок" зросла на 0,1. Це зовсім не мало, якщо врахувати, що на таку ж величину зменшилася ймовірність $P(Y_2/X_1)$ - в результаті відносна надійність виведення, яка характеризується відношенням $P(Y_1/X_1) / P(Y_2/X_1)$, збільшилася з 4 до 9, тобто більш ніж в 2 рази.

Очевидно, вигравш від використання апіорної інформації тим істотніше, чим більше обсяг такої інформації, тобто ніж нерівномірніше розподілення $P(Y_j)$. Наприклад, за результатами іншого профілактичного огляду, де середній вік обстежуваних становив 46,5 років, за допомогою аналогічних обчислень для $P(Y_1) = 0,82$ і $P(Y_2) = 0,18$, отримуємо

$$P(Y_1/X_1) = 0,94, P(Y_2/X_1) = 0,06.$$

Відношення $P(Y_1/X_1) / P(Y_2/X_1)$ тепер рівне близько 16, тобто збільшилося у порівнянні з початковим у 4 рази.

З урахуванням апіорної інформації про те як часто зустрічається ниркова патологія, наведений вище висновок щодо спостереження лімфатичного розарію можна було б перефразувати так: "оскільки частота народження ниркової патології серед населення з середнім віком 14-46 років досить висока і становить 70-80%, при всякому виявленні лімфатичного розарію на райдужці можна припускати зі значним ступенем впевненості (приблизно 92% проти 8%), що у хворого, що належить до тієї ж вікової категорії, є зміни з боку нирок, хоча категорично цього стверджувати не можна".

Резюмуючи питання використання апіорної інформації у вигляді різних захворювань, що зустрічаються, відзначимо:

1) облік апіорної ймовірності суттєво впливає на достовірність висновків, тому завжди слід прагнути до отримання та обліку апіорних відомостей про захворювання, що зустрічаються;

2) апіорна інформація, взагалі кажучи, не має відношення до того чи іншого виду діагностування - це інформація загальнономедичного характеру, що зберігається в певних документах МОЗ та відображає, можливо, залежність $P(Y_j)$ від

багатьох факторів і умов (соціальних, територіальних, кліматичних, екологічних, санітарно-епідеміологічних і т.п.);

3) при діагностиці лікар практично завжди використовує апріорну інформацію, проте зазвичай він це робить на інтуїтивному рівні.

3.4.4. "Дефекти" даних

Під "дефектами" даних будемо розуміти неоднорідність і неповноту даних.

Одним з проявів неоднорідності даних є опис ірідознак, що зустрічаються як в кількісній (числовій), так і в якісній (вербальній) формі. Причинами такої неоднорідності можуть бути або недостатнє дослідження ознак, або звичка практикуючих лікарів до вербального їх опису. Наприклад, ознака "тип райдужки" по J.Deck описаний у вербальній формі, тоді як інші ознаки описані переважно в числовій формі, у вигляді одновимірних розподілів умовних ймовірностей $P(Y_i/X_j)$. В результаті найважливіша ірідогенетична ознака "тип райдужки" випадає з переліку ознак із заданими значеннями $P(Y_i/X_j)$ і, отже, не може бути використана в обчисленнях за формулою Байєса. Ця ознака не єдина: якісно описані також щільність райдужки по B.Jensen, рельєф райдужки по G.Jausas, децентрація зіниць по R.Bourdiol. Іноді якісний опис поєднується з кількісним, що, в принципі, дозволяє встановити, нехай і наближено, однозначна відповідність між якісною і кількісною шкалами. Дійсно, градаціями шкал ознаки "тип райдужки" по J.Deck є, в основному, такі: "характерна схильність", "не виключена ймовірність", "є схильність" і т.п. В описі нейрогенного типу спостерігаємо іншу ситуацію, коли якісної градації ставиться у відповідність кількісна: "... нейрогенний тип ... зустрічається при бронхіальній астмі і виразковій хворобі досить часто - у 75% випадків". Завдяки цьому, знаходимо, що градації "достатньо часто" шкали найменувань відповідає градація "75%" числової шкали. Труднощі подібної "оцифровки" шкали значень ознак не тільки в рідкості подібних "знахідок" в авторському тексті, але і в близькості одержуваних результатів.

Неповнота даних може бути пояснена або недостатньою дослідженістю ірідознак, або недбалістю авторів публікацій.

Наприклад, J.Deck "забув" про існування ірідогенетичних типів для людей з високим рівнем здоров'я. З іншого боку, аналогічна "забудькуватість" властива й іншим авторам, які відзначають, що симптом "адаптаційні дуги, деформовані в проекційній зоні нирок" зустрічається у 2/3 хворих з нирковими патологіями - при цьому аналогічна оцінка для здорових людей не наводиться.

3.4.5. Чинники, що "заважають"

Аналізуючи чинники заважають", умовно розділимо їх на зовнішні та внутрішні (по відношенню до методу ірідодіагностики). До зовнішніх віднесемо неоднорідність складу пацієнтів (відмінності за статтю, віком, освітою, місцем проживання, соціальним станом тощо), неоднорідність складу і стану лікарів (рівень кваліфікації, психофізіологічний стан лікаря в момент обстеження), неоднорідність умов обстеження (кліматичних, екологічних, санітарно-

епідеміологічних і інше), включаючи умови праці лікаря (освітленість, вентиляція, площа приміщень, тривалість періодів праці та відпочинку тощо).

Незважаючи на достатню очевидність більшості згаданих чинників, наведемо, тим не менш, деякі приклади. Сенільне кільце (симптом атеросклерозу) зустрічається у чоловіків в 1,5 рази частіше, ніж у жінок, і зазвичай стає вираженим після 20 років. До 40 років воно спостерігається у 10% населення, а до 65 років - у 45%. Лімфатичний розарій у французів зустрічається в 2 рази частіше, ніж у москвичів. Утворенню лімфатичного розарію "сприяють" професії, пов'язані зі значними фізичними навантаженнями: ковалів, вантажників, спортсменів.

Достовірність і швидкість формування висновку іридолога зростають із збільшенням рівня кваліфікації спеціаліста і, разом з тим, залежать від його особистих психофізіологічних характеристик - обсягу і виду домінуючої пам'яті, концентрації уваги, типу мислення, рівнів інтелекту, екстраверсії, нейротизму і інше. Значну роль при цьому грають економічні, соціальні, санітарно-гігієнічні та інші характеристики навколишнього середовища іридолога - характер міжособистісних відносин у колективі, зручність розташування робочого місця, ступінь його оснащення оргтехнікою тощо.

До внутрішніх чинників, що "заважають" віднесемо проблеми іридології як науки. Першою серед проблем такого типу назвемо неможливість диференціювання одними лише засобами візуального аналізу моменту появи захворювання - у більшості випадків тільки з урахуванням клінічних даних і, отже, тільки шляхом діалогу з пацієнтом, іридолог може встановити, ознаки якого захворювання спостерігаються: минулого, теперішнього або майбутнього. Конструктивність такого діалогу, очевидно, може істотно залежати практично від усіх перерахованих вище зовнішніх факторів, що "заважають".

По-друге, до цих пір не створена універсальна іридологічна схема проєкційних зон, яка задовольнила б усіх іридологів і, ймовірно, така схема ніколи не буде створена з причини принципово тяжко вирішуваних труднощів:

1) формування нервових шляхів у кожної людини індивідуально, у зв'язку з чим кожен орган має не точну, а ймовірнісну проєкцію (особливо яскравим прикладом такого роду є різне формування іридоознак у правшів, лівшів та осіб з амбідекстрією - при цьому топографія іридоознак у осіб з амбідекстрією може формуватися настільки складним чином, що фактична топографія іридоознак не відповідає ні одній схемі);

2) можливість виникнення ознак, пов'язаного не з патологією органів, а з патологією наведених симпатичних і парасимпатичних нервових шляхів, причому досить надійно подиференціювати ці дві причини буває складно і не завжди можливо;

3) можливість виникнення іридоознак, зв'язаних не з патологією органу, а за типом відбитої рефлекторної іррадіації (наприклад, виникнення лакун у зоні проєкції лівої верхньої кінцівки при стенокардії).

Тому при коментуванні самої ідеї комп'ютерної підтримки іридодіагностики в літературі зазначається, що механічне накладення зображення на іридосхему

не може дати високу якість розшифровки. Для правильної інтерпретації іридоознач потрібно клінічне мислення лікаря, щоб подиференціювати помилкові ознаки від істинних, виявити механізм їх утворення, тобто в комп'ютер потрібно закласти принаймні інститутську програму підготовки лікаря, щоб він міг давати достатньо кваліфікований висновок.

Саме наявністю внутрішніх проблем іридології можна пояснити ту обставину, що в практичній діяльності, при оцінці генетичних і характерологічних якостей людини, іридолог вдаються до використання додаткових симптомокомплексів допоміжного апарату очей, окремих параіридальних симптомів: видів і форми брів, підочноямкових мішків, очних щілин і повік, форми судин бульбарної кон'юнктиви і т.п. На допомогу залучаються також фізіономічні, медико-астрологічні, хіромантичні, анатомічні, психологічні та інші комплекси ознак, в результаті застосування яких метод іридодіагностики втрачає "чистоту", але зате забезпечує прийнятну якість діагностування.

3.5. Правила продукції з елементами Байєсівської стратегії

Як вже вказувалося, застосування Байєсівської стратегії прийняття рішення в її класичній формі усилено з ряду причин: статистична залежність ознак; неповнота апріорних даних; неповнота і неоднорідність іридологічної інформації; наявність чинників, що заважають. Проблему статистичної залежності ознак при цьому можна вирішувати двома шляхами:

1) не помічати її і робити обчислення так, як якщо б ознаки були незалежні, в надії, що з практичної точки зору результати будуть задовільними;

2) враховуючи залежність ознак, робити обчислення з відступом від класичної формули Байєса, зберігаючи, однак, елементи Байєсівської стратегії.

Прикладом реалізації другого варіанту є механізм виведення, прийнятий у іридодіагностичній експертній системі ЭСИД: за наявності i -ї ознаки відповідно до формули (3.1) здійснюють обчислення ймовірності j -тій гіпотезі про наявність захворювання (у припущенні рівномірності апріорних розподілів гіпотез для i -ї ознаки):

$$P(Y_j / X_i) = \begin{cases} \frac{P(X_i / Y_j)}{\sum_j P(X_i / Y_j)}, & J_i > 1, \\ P(X_i / Y_j), & J_i = 1. \end{cases} \quad (3.6)$$

Ситуація $J_i = 1$ (тобто i -тій ознаці відповідає лише одна гіпотеза) введена примусово, вона відображає ситуацію "пробілів" даних, коли оцінка умовних ймовірностей $P(X_i / Y_j)$ відома лише для єдиної гіпотези.

При наявності декількох ознак усереднюють оцінки апостеріорних ймовірностей, отриманих у відповідності з виразом (3.6):

$$P(Y_j / X_1, \dots, X_I) = \frac{1}{I} \sum_{i=1}^I P(X_i / Y_j) \quad (3.7)$$

Враховуючи наявні відступи від Байєсівської стратегії, доцільно трактувати співвідношення (3.7) не як імовірнісну оцінку достовірності гіпотез, а як результат "зваженого голосування" ознак. Співвідношення (3.6) при цьому дозволяє визначити приватні рейтинги гіпотез про наявність захворювань, а співвідношення (3.7) формує загальний, усереднений, рейтинг.

По суті, ідея такого відступу не нова, її основні риси можна углядіти в деяких табличних діагностичних алгоритмах, розроблених ще в 70-і роки.

Дійсно, зіставимо наступні два фрагменти 1.1ф і 2.6ф табл. 1.1 і 2.6, відповідно (про фрагментарну характеристику свідчить індекс «ф»).

Таблиця 1.1ф

Симптоми	Геморагічний інсульт	Ішемічний інсульт
Передвісники	Малохарактерні	Характерні
Раптовий розвиток	Характерний	Менш характерний
Повільний розвиток	Нехарактерний	Характерний
Втрата свідомості	Характерна	Менш характерна

Таблиця 2.6ф

Симптоми	Геморагічний інсульт	Ішемічний інсульт
Передвісники	0	1
Раптовий розвиток	1	0
Повільний розвиток	0	1
Втрата свідомості	1	0

Неважко бачити, що якісні характеристики табл. 1.1 в табл. 2.6 замінені кількісними з великим ступенем закруглення. Тим часом, можна було б, у принципі, експериментально оцінити відповідні частоти виникнення різних ознак для даних захворювань, звівши результати в табл. виду 3.3.

Таблиця 3.3

Вербальна оцінка	Числова оцінка
Характерно	0,8 – 1
Менш характерний	0,5 – 0,8
Малохарактерний	0,2 – 0,5
Нехарактерна	0 – 0,2

Застосовуючи дані табл. 3.3, можливо було б замість табл. 2.6ф отримати більш точну табл. 3.4.

Таблиця 3.4

Симптоми	Геморагічний інсульт	Ішемічний інсульт
Передвісники	0,35	0,9
Раптовий розвиток	0,9	0,65
Повільний розвиток	0,1	0,9
Втрата свідомості	0,9	0,65

У клітинах табл. 3.4 проставлені середні значення частот зустрічальності симптомів з діапазонів, зазначених у табл. 3.3.

Якщо тепер, вважаючи для простоти захворювання рівноімовірними, за формулою Байеса (3.1) оцінити значення $P(Y_j/X_i)$ і розмістити їх у відповідних клітинах табл. 3.4, отримаємо нову табл. 3.5, що дозволяє робити обчислення, аналогічні (3.6) і (3.7).

Таблиця 3.5

Симптоми	Геморагічний інсульт	Ішемічний інсульт
Передвісники	0,28	0,72
Раптовий розвиток	0,58	0,42
Повільний розвиток	0,1	0,9
Втрата свідомості	0,58	0,42

Дійсно, нехай спостерігається дві ознаки: «передвісники» і «повільний розвиток». Складаючи «в стовпчик» числа з відповідних клітин і ділячи результати на два (кількість спостережуваних ознак), отримаємо рейтинги геморагічного та ішемічного інсульту, рівні 0,19 і 0,81, відповідно. В результаті маємо, що ішемічний інсульт є кращим діагнозом.

Нескладно побачити, що табл. 3.5 фактично є добре нам знайомою діагностичною таблицею, з тією лише відмінністю, що в її клітинах стоять не цілі числа (що зручно для ручних обчислень), а дробові. Очевидно, це розходження легко усувається шляхом множення всіх дрібних чисел на один і той же коефіцієнт, наприклад на 10, з подальшим округленням результатів. У результаті таких дій отримаємо «справжню» діагностичну табл. 3.6, придатну для ручних обчислень.

Таблиця 3.6

Симптоми	Геморагічний інсульт	Ішемічний інсульт
Передвісники	3	7
Раптовий розвиток	6	4
Повільний розвиток	1	9
Втрата свідомості	6	4

Наведені вище результати, з одного боку, демонструють наявність глибокої внутрішньої зв'язку між добре відомими табличними діагностичними алгоритмами і Байєсівської стратегією. З іншого боку, доведена в клінічній практиці досить висока ефективність табличних діагностичних алгоритмів може служити аргументом на користь правила продукції з елементами Байєсівської стратегії, описуваного співвідношеннями (3.6) - (3.7).

3.6. Системи нечіткого виводу

Системи нечіткого виведення - це такі системи, в яких умови і укладення окремих правил формулюються у формі нечітких висловлювань щодо значень тих чи інших лінгвістичних змінних.

3.6.1. Основні поняття систем нечіткого виводу

Розглянемо такі основні поняття систем нечіткого виведення як *нечітке лінгвістичне висловлювання і нечітке правило продукції*.

Нечітким лінгвістичним висловлюванням називають висловлювання наступних видів.

1. Вислів « $\beta \in \alpha$ », де β - найменування лінгвістичної змінної, α - її значення, якому відповідає окремий лінгвістичний терм лінгвістичної змінної β (приклад: «тиск крові високий»).

2. Вислів « $\beta \in \nabla\alpha$ », де ∇ - модифікатор, відповідний таким словам, як: «дуже», «більш-менш», «багато більше» та ін. (Приклад: «рівень холестерину дуже високий»).

3. Складові висловлення, утворені з висловлювань видів 1 і 2 та нечітких логічних операцій у формі зв'язувань: «І», «АБО», «ЯКЩО-ТО», «НЕ» (приклад: «тиск крові високий» І «рівень холестерину дуже високий»).

Далі нечіткі лінгвістичні висловлювання будемо називати просто нечіткими висловлюваннями.

Розглянемо кілька прикладів нечітких висловлювань. На рис. 3.10, а наведено вид функції приналежності $\mu_A(x)$ для висловлення «*вік пацієнта середній*». Нечіткий вислів «*вік пацієнта дуже середній*» означає перетворення функції приналежності $\mu_A(x)$ в функцію $\mu_C(x)$ (рис. 3.10, б) відповідно до деякої формули, наприклад

$$\mu_C(x) = \mu_A(x)^2 .$$

А нечітке висловлювання виду «*вік пацієнта більш чи менш середній*» може означати перетворення функції приналежності $\mu_A(x)$ в функцію $\mu_B(x)$ (рис. 3.10,в) у відповідно до формули

$$\mu_B(x) = \mu_A(x)^{0,5} .$$

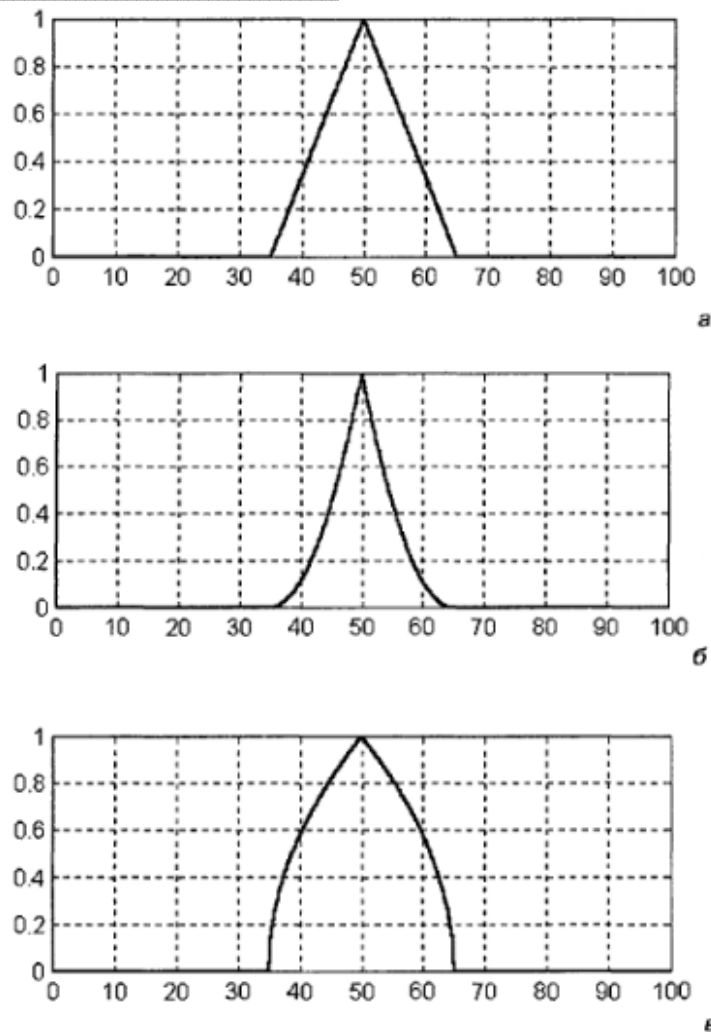


Рис. 3.10

Нечітке правило продукції (або просто нечітка продукція) дозволяє записувати такі логічні висновки як «якщо вік пацієнта середній, тоді ризик інсульту середній» в формі:

ПРАВИЛО <#>: ЯКЩО « $\beta_1 \in \alpha$ » ТО « $\beta_2 \in \alpha$ »

або в еквівалентній формі:

ПРАВИЛО_#: ЯКЩО «Умова_#» ТО «Висновок_#»,
 RULE_#: IF «Condition_#» THEN «Conclusion_#».

3.6.2. Основні етапи нечіткого висновку

Системи нечіткого висновку призначені для перетворення значень вхідних змінних у вихідні змінні на основі використання нечітких правил продукції.

Основними етапами нечіткого висновку є (рис. 3.11):



Рис. 3.11

- формування бази правил;
- фазифікації вхідних змінних;
- агрегування підумови в нечітких правилах продукції;
- активізація або композиція підзаклучень;
- акумулювання висновків в нечітких правилах продукції;
- дефазифікації вихідних змінних.

Оскільки зазвичай правил багато, їх об'єднують в безліч, іменоване базою правил. Найбільш часто база правил представляється у формі структурованого тексту:

ПРАВИЛО_1: ЯКЩО «Умова_1» ТО «Висновок_1»,
 ПРАВИЛО_2: ЯКЩО «Умова_2» ТО «Висновок_2»,
 ...
 ПРАВИЛО_N: ЯКЩО «Умова_N» ТО «Висновок_N»,

Фазифікацією називається процес або процедура знаходження значень функцій належності нечітких множин (термів) на основі звичайних («чітких») вихідних даних. Метою фазифікації є встановлення відповідності між конкретним (зазвичай - чисельним) значенням вхідної змінної і значенням функції приналежності відповідного їй терма вхідної лінгвістичної змінної.

Наприклад, якщо вхідна лінгвістична змінна «температура води» застосовуємо значення «низька», «середня» і «висока», тоді за допомогою набору функцій приналежності $y = \mu_i(x)$, $i = 1, 2, 3$, показаного на рис. 3.12, для значення температури $t = 55^\circ\text{C}$ результатом фазифікації являється трійка чисел $0; 0,67; 0$.

Агрегування являє собою процедуру визначення ступеня істинності умов по кожному з правил системи нечіткого виводу.

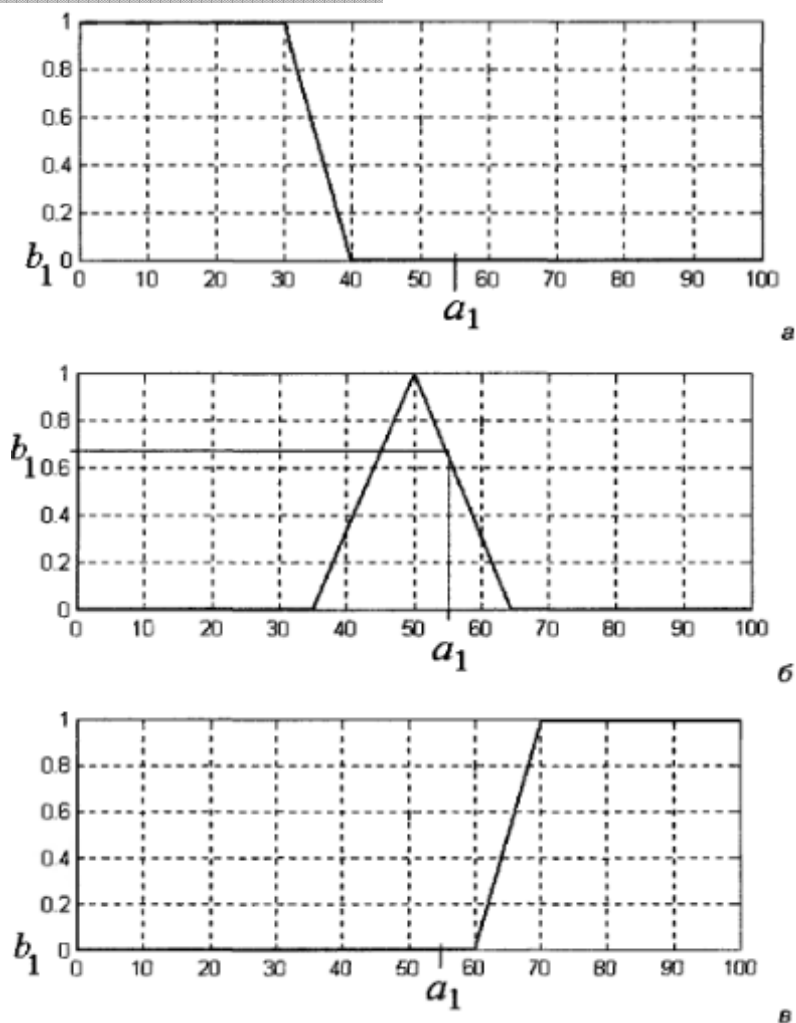


Рис. 3.12

Для ілюстрації сутності даного етапу розглянемо приклад агрегування двох нечітких висловлювань: «температура води середня» І «кава гаряча» і «температура води середня» ЧИ «кава гаряча». При цьому припустимо, що температури води дорівнює $t_1 = 55^\circ\text{C}$, а температура кави дорівнює $t_2 = 70^\circ\text{C}$.

Агрегування першого нечіткого висловлювання з використанням операції нечіткої кон'юнкції

$$T(A \wedge B) = \min\{T(A), T(B)\}$$

дає в результаті число $b_1'' = 0,67$, яке є мінімальним значенням із пари значень 0,67 і 0,8 (рис. 3.13, а).

Агрегування другого нечіткого висловлювання з використанням операції нечіткої диз'юнкції

$$T(A \vee B) = \max\{T(A), T(B)\}$$

дає в результаті число $b_1'' = 0,8$, яке є максимальним значенням із пари значень 0,67 і 0,8 (рис. 3.13, б).

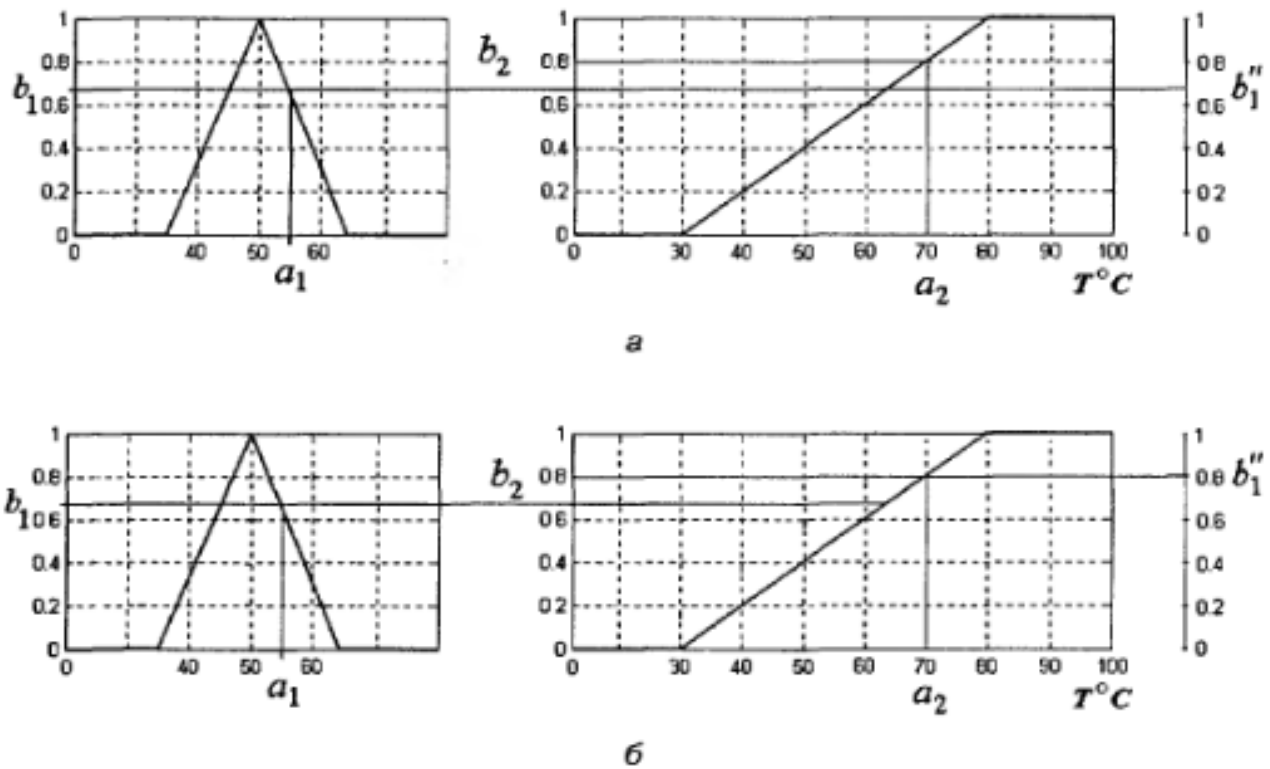


Рис. 3.13

Активізація являє собою процедуру або процес знаходження ступеня істинності кожного з підзаключення правил нечіткої продукції. Знаходяться всі значення ступенів істинності підзаключення для кожного з правил, що входять до бази правил. Після знаходження множини $C = \{c_1, c_2, \dots, c_q\}$ визначаються функції приналежності кожного з підзаключення для розглянутих вихідних лінгвістичних змінних. Для цього використовується один з методів активізації:

$$\begin{aligned} \mu'(y) &= \min\{c_i, \mu(y)\} - \text{min-активізація;} \\ \mu'(y) &= c_i \cdot \mu(y) - \text{prod- активізація;} \\ \mu'(y) &= 0,5 \cdot (c_i + \mu(y)) - \text{average- активізація,} \end{aligned}$$

де $\mu(x)$ - функція приналежності.

Проілюструємо етап активізації на прикладі активізації висновку в наступному правилі нечіткої продукції:

ЯКЩО «температура води середня» ТО «кава гаряча».

Якщо температура води дорівнює 55 градусів ($a_1 = 55$), тоді агрегування умови цього правила дає $b_1'' = 0,67$. Це значення буде використано для отримання результату активізації. Результат, отриманий методом min-активізації, зображений більш темним кольором на рис. 3.14, а, як і результат, отриманий методом prod-активізації (рис. 3.14, б).

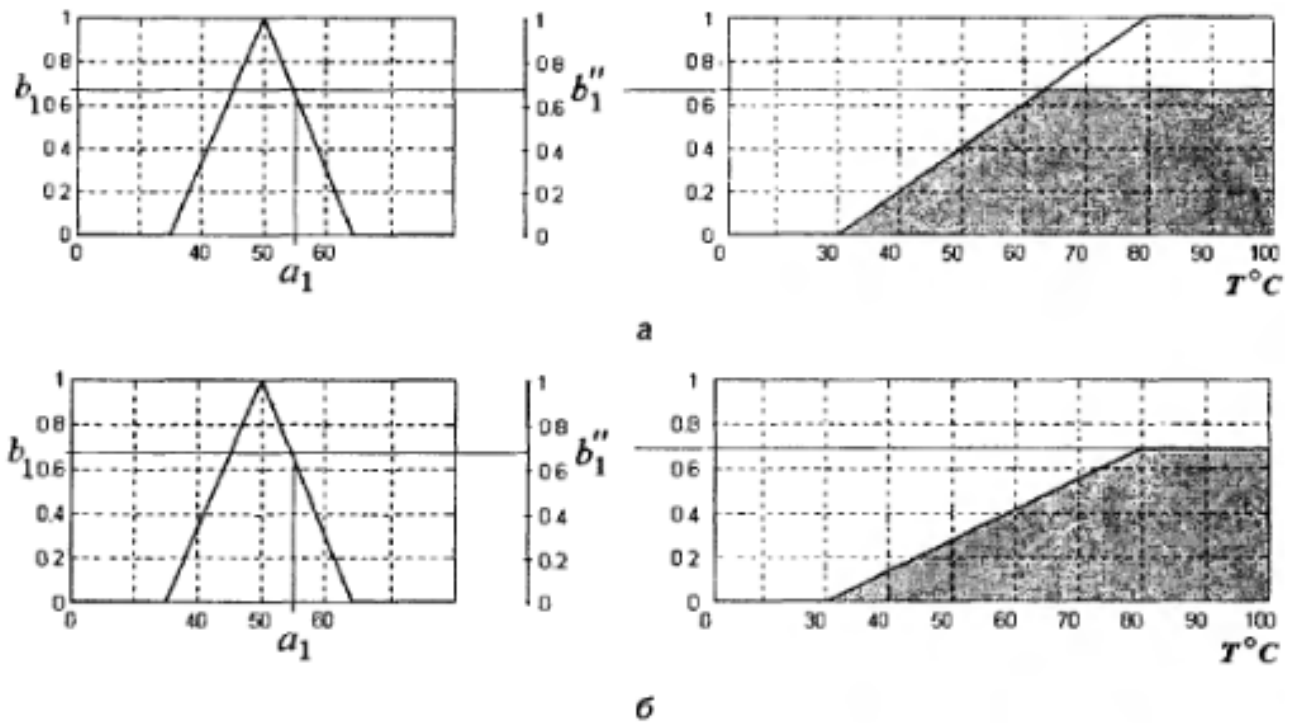


Рис. 3.14

Акумуляція (акумулявання) являє собою процедуру або процес об'єднання всіх ступенів істинності висновків (подзаклученій) для знаходження функції приналежності для кожної з вихідних лінгвістичних змінних. Приклад акумуляції висновків для трьох нечітких множин C_{11}, C_{12}, C_{13} , отриманих в результаті виконання процедури активізації для вихідної лінгвістичної змінної «температура води», представлений на рис. 3.15. У цьому випадку функція належності вихідної змінної відповідає множині C'_1 , при цьому акумуляція множин C_{11}, C_{12}, C_{13} , проведена методом max-об'єднання.

Дефазифікація являє собою процедуру або процес знаходження звичайного «чіткого» значення для кожної з вихідних лінгвістичних змінних. Дефазифікації виробляють з використанням деяких аналітичних виразів, іменованих методами дефазифікації. До таких методів належать: метод центру ваги, метод центру площі, метод лівого (або правого) модального значення. Зрозуміло, застосування різних методів призводить і до різних, взагалі кажучи, результатів.

Центр ваги, наприклад, розраховується за формулою:

$$y = \frac{\int_{Min}^{Max} x \cdot \mu(x) dx}{\int_{Min}^{Max} \mu(x) dx} ,$$

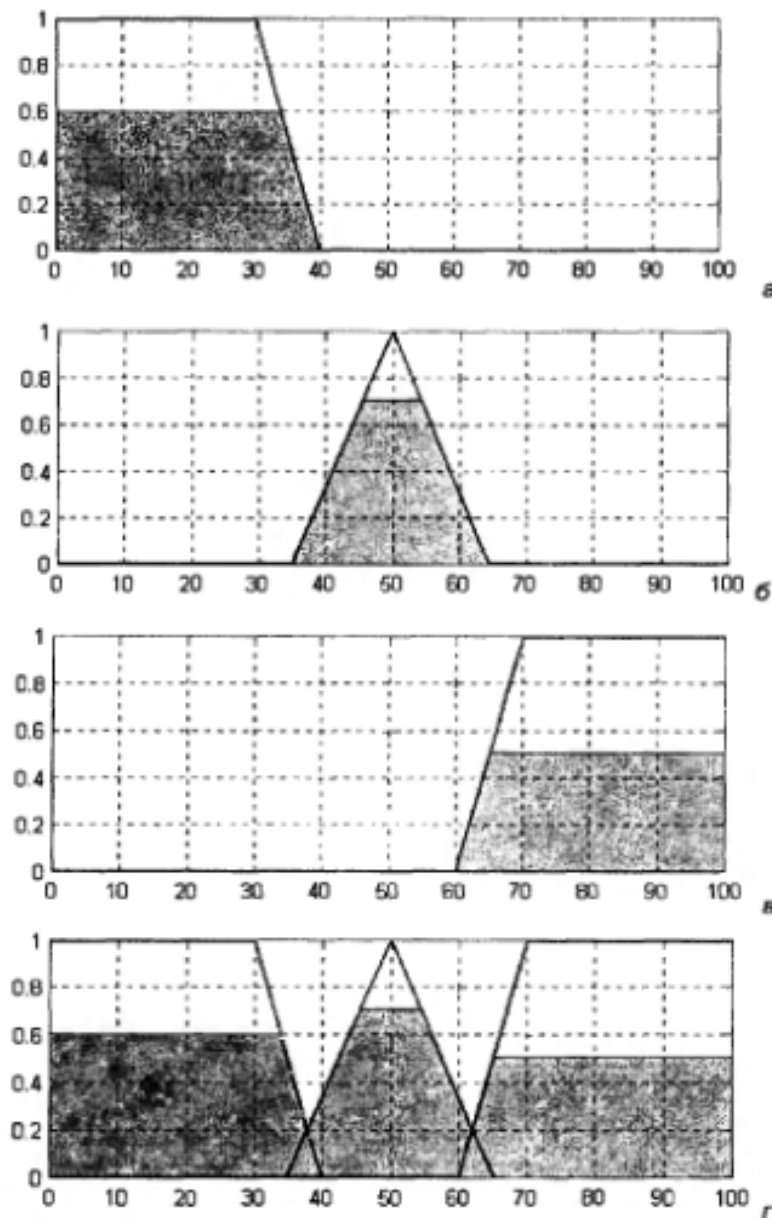


Рис. 3.15

а центр площі рівний $Y = u$, де u визначається із рівності:

$$\int_{Min}^u \mu(x) dx = \int_u^{Max} \mu(x) dx$$

На рис. 3.16,а представлений результат дефазифікації методом центра маси, а на рис. 3.16,б – методом центра площі.

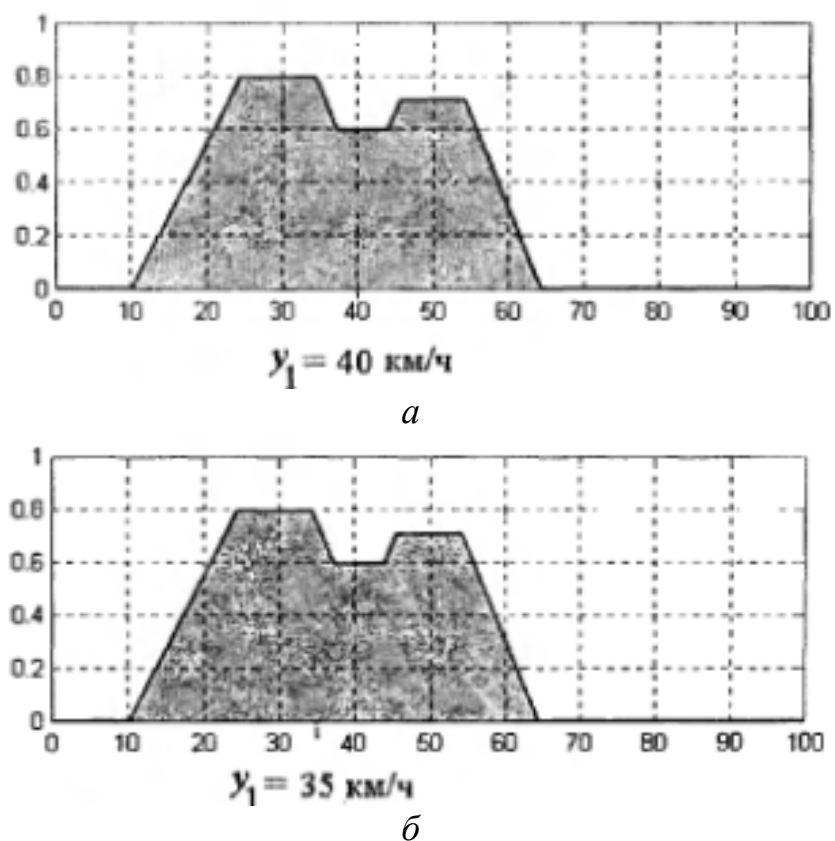


Рис. 3.16

3.6.3. Основні алгоритми нечіткого виводу

Розглянуті вище етапи нечіткого виведення можуть бути реалізовані різними способами. Вибір поєднання конкретних способів реалізації різних етапів означає формування конкретного алгоритму нечіткого виводу. Існує якась «канонізована» кількість алгоритмів нечіткого виводу, до яких відносяться: алгоритм Мамдані (Mamdani), алгоритм Цукамото (Tsukamoto), алгоритм Ларсена (Larsen), алгоритм Сугено (Sugeno), спрощений алгоритм нечіткого виводу.

При вирішенні практичних завдань можливе поєднання різних алгоритмів.

У підрозділі 3.7.2 наведено приклад моделювання системи нечіткого виведення в середовищі Matlab, в якому використаний алгоритм Мамдані. У цьому зв'язку доцільно ознайомитися з наступними особливостями етапів алгоритму Мамдані.

1. Особливості формування бази правил і фазифікації вхідних змінних збігаються з розглянутими вище при описі даних етапів.

2. При агрегування підумови використовуються парні нечіткі логічні операції. Правила, ступінь істинності яких відмінна від нуля, вважаються активними і використовуються для розрахунків.

3. Активізація підвисновків відбувається по формулі $\mu'(y) = \min\{c_i, \mu(y)\}$, при цьому для скорочення часу виведення враховуються тільки активні правила.

4. Акумуляція висновків відбувається по формулі

$\mu_D(x) = \max\{\mu_A(x), \mu_B(x)\}, (\forall x \in X)$, для об'єднання нечітких множин.

5. Дефазифікація вихідних змінних відбувається методом центра мас або методом центра площі.

3.7. Моделювання систем нечіткого виведення в середовищі MATLAB

Для моделювання систем нечіткого виведення в середовищі Matlab призначений спеціальний пакет розширення Fuzzy Logic Toolbox. В рамках цього пакету користувач може працювати в двох режимах:

- в інтерактивному режимі, за допомогою графічних засобів редагування і візуалізації;

- в командному режимі, в командному вікні системи Matlab.

Враховуючи, що інтерактивний режим краще підходить для отримання першого уявлення про прикладний аспект методології побудови систем нечіткого виводу, обмежимося розглядом саме цього режиму.

3.7.1. Інструментарій для моделювання систем нечіткого виведення в інтерактивному режимі

Для розробки і застосування систем нечіткого виведення в інтерактивному режимі можуть бути використані наступні засоби графічного інтерфейсу:

- редактор систем нечіткого виведення FIS (FIS Editor) або скорочено - редактор FIS, що активізується командою **fuzzy**;
- редактор функцій належності (Membership Function Editor), що активізується командою **medit**;
- редактор правил (Rule Editor), що активізується командою **ruleedit**;
- засіб перегляду правил вивід (Rule Viewer), викликаного командою **ruleview**;
- засіб перегляду поверхневого виведення (Surface Viewer), викликаного командою **surfview**.

Крім того, в пакет Fuzzy Logic Toolbox входять спеціальні програми:

- редактор адаптивних нейронних систем нечіткого виведення (Adaptive Neuro-Fuzzy Interface System Editor) - скорочено редактор ANFIS, що активізується командою **anfisedit**;
- програма нечіткої кластеризації методом нечітких-середніх (Fuzzy C-Means Clustering), активізується командою **findcluster**.

З особливостями даного інструментарію зручніше знайомитися, розглядаючи конкретний приклад моделювання системи нечіткого виведення в інтерактивному режимі.

3.7.2. Приклад моделювання в інтерактивному режимі

Перш ніж самостійно приступати до створення системи нечіткого виведення в інтерактивному режимі, доцільно попередньо ознайомитися з демонстраційними прикладами системи Matlab. Один з таких прикладів - це досить простий,

а тому і зручний для ознайомлення з можливостями програмного інструментарію, приклад «Чайові в ресторані».

Кілька слів про даній задачі. Згідно з традиціями США, в ресторані, після закінчення обслуговування відвідувача, прийнято залишати офіціантові чайові. Сума чайових не є фіксованою і залежить від ряду факторів, в першу чергу таких як якість їжі та якість обслуговування. Тому база правил може бути представлена наступним чином:

1. Якщо обслуговування погане або вечеря підгоріла, то чайові - малі.
2. Якщо обслуговування хороше, то чайові - середні.
3. Якщо обслуговування відмінне або вечеря чудова, то чайові - щедрі.

Наведена база правил не вільна від критики, однак вона досить проста і тому зручна для подальшого розгляду (зауважимо, до речі, що і ми в побуті керуємося найчастіше не дуже досконалими правилами, приймаючи ті чи інші рішення).

В якості вхідних змінних в даному випадку мається 2 нечіткі лінгвістичні змінні: «якість обслуговування» і «якість їжі», а в якості вихідних змінних - єдину нечітку лінгвістичну змінну «величина чайових».

Значення вхідної змінної «якість обслуговування»: {«погане», «середнє», «відмінне»}. Значення вхідної змінної «якість їжі»: {«підгоріла», «чудова»}. Значення вихідної змінної «величина чайових»: {«мала», «середня», «щедра»}. При цьому кожне із значень вхідних змінних будемо оцінювати за 10-бальною порядковою шкалою, при якій числу 0 відповідає найгірша оцінка, а числу 10 - найкраща. Що стосується значень вихідної змінної, будемо припускати, що малі чайові складають близько 5% від вартості замовлених страв, середні чайові - близько 15%, а щедрі - не більше 30%.

З урахуванням зроблених зауважень, наведена вище база правил може бути переписана у формі:

ПРАВИЛО_1: ЯКЩО «якість обслуговування погана» АБО «їжа підгоріла» ТО «величина чайових мала»

ПРАВИЛО_2: ЯКЩО «якість обслуговування хороша» ТО «величина чайових середня»

ПРАВИЛО_3: ЯКЩО «якість обслуговування відмінна» АБО «їжа чудова» ТО «величина чайових щедра»

Ознайомимося тепер з графічним інтерфейсом програмного інструментарію. Задамо в командному вікні середовища Matlab команду **fuzzy**. У результаті на моніторі з'явиться вікно FIS Editor з ім'ям системи нечіткого висновку Untitled (рис. рис. 3.17).

Як слід з тексту «FIS Type: mamdani», за замовчуванням користувачеві пропонується працювати з алгоритмом Мамдані.

Не будемо створювати систему нечіткого виведення «з нуля», а обмежимося тим, що активізуємо демонстраційний приклад «Чайові в ресторані». Для цього слід у меню вікна редактора FIS вибрати пункт **File>Import>From File**, після чого провести пошук папки **toolbox\fuzzy\fuzdemos** (зазвичай знаходиться на диску **C:** всередині папки **Program Files\MATLAB**) і файлу з іменем **tipper.fis**.

В результаті на моніторі появиться вікна FIS Editor з іменем системи нечіткого виведення tipper (рис. 3.18).

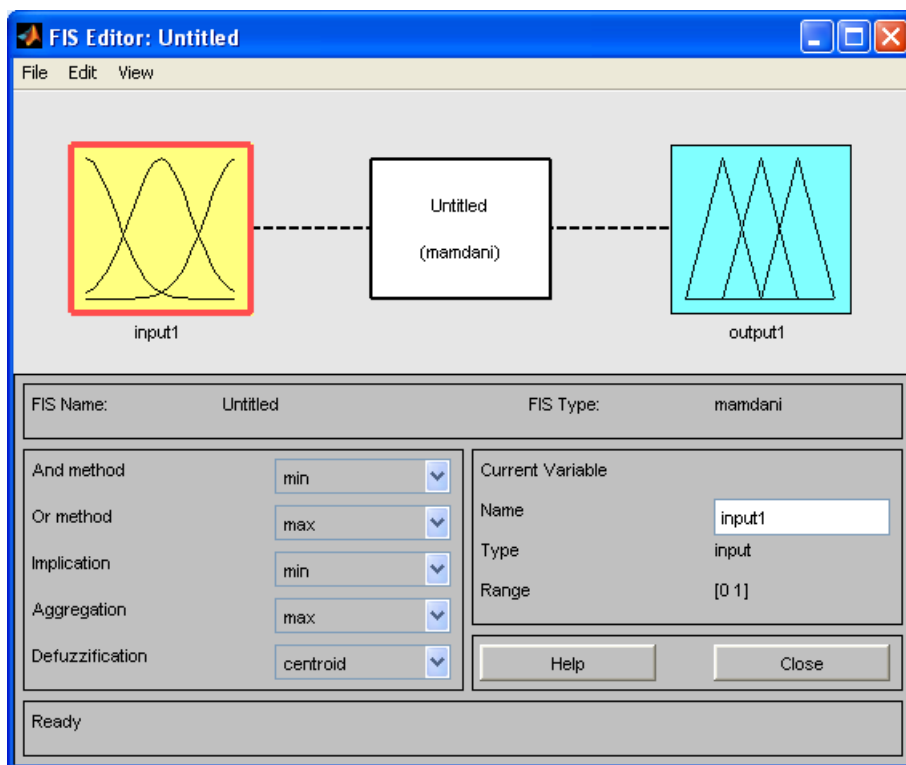


Рис. 3.17

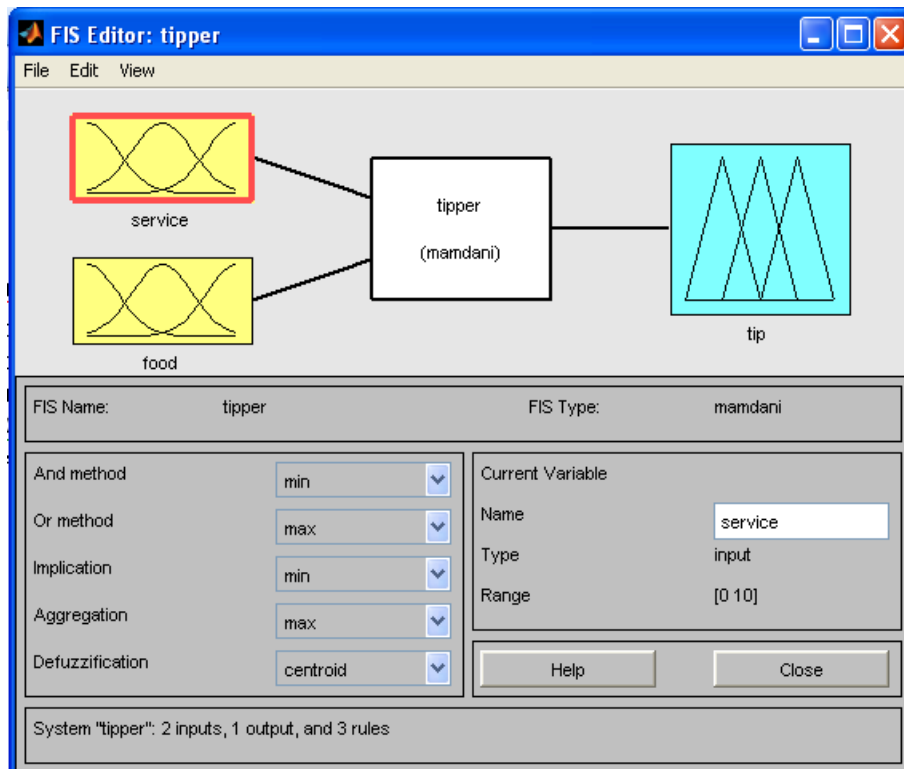


Рис. 3.18

У верхній частині вікна редактора FIS зображена схема системи нечіткого виведення на ім'я *tipper*. Власне система зображена у вигляді білого квадрата, усередині якого зазначені ім'я системи і алгоритм її реалізації. Деталі етапів цього алгоритму можна бачити (і редагувати) у лівій нижній частині вікна редактора FIS, а в нижній частині вікна редактора розташована статусна область з коротким описом системи нечіткого виведення (система «*tipper*»: 2 входи, 1 вихід, 3 правила).

Вхідні лінгвістичні змінні системи зображені ліворуч від білого квадрата, у вигляді двох жовтих квадратів, під якими підписані імена змінних. Клацнувши (одноразове клацання!) мишкою по верхньому жовтому квадрату з ім'ям *service* («обслуговування»), в правій нижній частині вікна редактора FIS (рис. 3.18) спостерігаємо (і можемо редагувати) характеристики даної змінної (тип - вхідна, область можливих числових значень - 0 ... 10).

Вихідна лінгвістична змінна системи, природно, зображена праворуч від білого квадрата, у вигляді квадрата блакитного кольору. Клацнувши (також одноразове клацання!) мишою на цьому квадраті (ім'я змінної *tip* («чайові»)), в правій нижній частині вікна редактора FIS можемо спостерігати (і редагувати) характеристики даної змінної.

Таким чином, приходимо до висновку, що у вікні редактора FIS можна виконувати наступні дії:

- створювати «з нуля» або редагувати раніше створену систему нечіткого виведення;
- редагувати параметри алгоритму виведення;
- створювати і видаляти змінні, задавати і редагувати імена змінних.

Крім того, використовуючи меню редактора FIS, можна викликати редактор функцій приналежності (**Edit>Membership Functions**), редактор правил (**Edit>Rules**), а також вікно перегляду правил (**View>Rules**) і перегляду поверхні (**View>Surface**).

Вікно редактора функцій належності **Membership Function Editor** можна активізувати також подвійним (!) клацанням по квадрату, відповідному вхідний чи вихідний змінної. Незалежно від способу активації, вікно редактора функцій належності має вигляд, показаний на рис. 3.19.

Про призначення та можливості даного вікна легко здогадатися не тільки за його назвою, а й за його структурі і елементам управління. Вибравши (у лівій верхній частині вікна) ту чи іншу змінну, в правій верхній частині вікна спостерігаємо набір функцій належності, що відповідають значенням вибраної змінної. Ім'я, вид і параметри функцій належності задаються і редагуються елементами управління, розташованими в правій нижній частині вікна. Задавати і редагувати діапазон можливих значень змінних можна за допомогою елементів управління, розташованих в лівій нижній частині вікна.

Реалізувати більш «серйозні» операції, такі як додавання або видалення функцій належності, або додавання користувача функції приналежності, можна тільки за допомогою меню (позиція *Edit*). Третім, і останнім, графічним редак-

тором є редактор правил **Rule Editor** (рис. 3.20), який можна активізувати подвійним клацанням по центральному білому квадрату схеми редактора FIS.

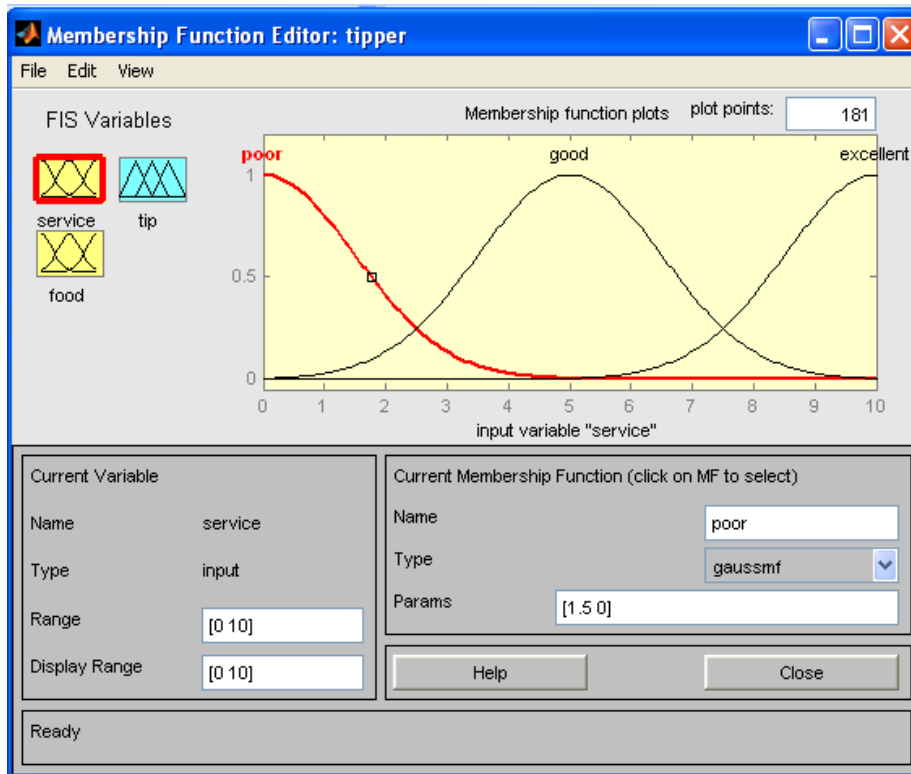


Рис. 3.19

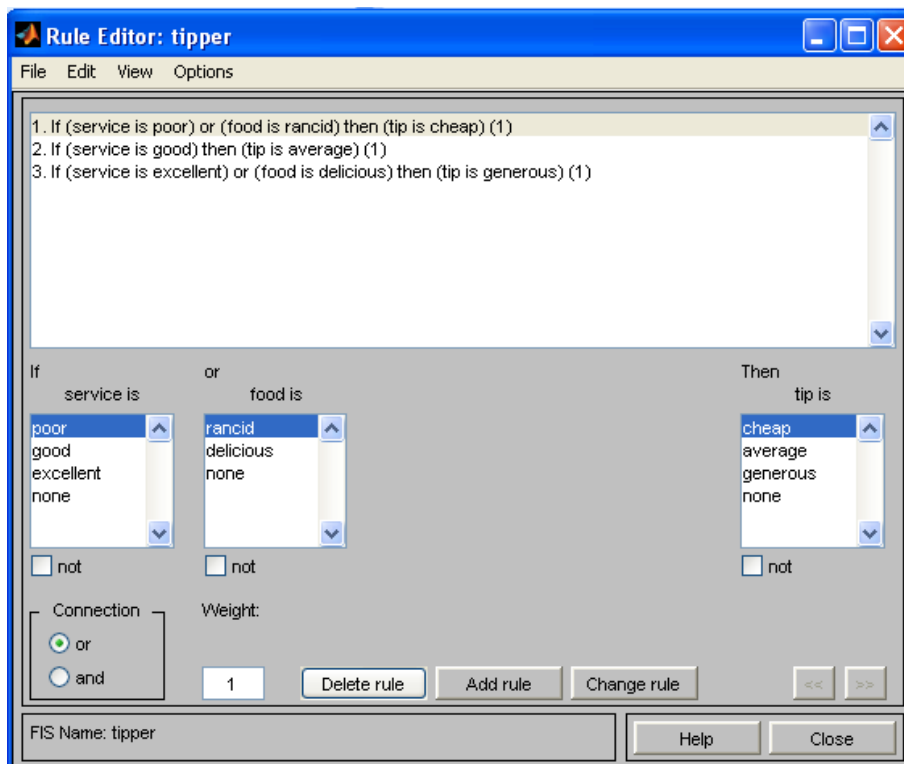


Рис. 3.20

У верхній частині вікна редактора правил знаходиться список правил, а в нижній частині вікна розташовані елементи управління, за допомогою яких цей список створюється і редагується. При такому підході до створення і редагування правил вдається позбавити користувача від набору тексту вручну, за допомогою клавіатури. У результаті, до певної міри гарантується коректність функціонування створюваної системи нечіткого виведення. Простота і наочність даного інтерфейсу сприяють швидкому його освоєнню, навіть на інтуїтивному рівні.

Результати роботи створюваної або редагованої системи нечіткого висновку можна спостерігати за допомогою вікна **Rule Viewer** (Рис. 3.21). При першому виклику даного вікна перегляду, за замовчуванням задаються середні значення вхідних змінних, рівні [5 5]. Цим значенням відповідають положення вертикальних червоних візирів на графіках функцій належності вхідних змінних.

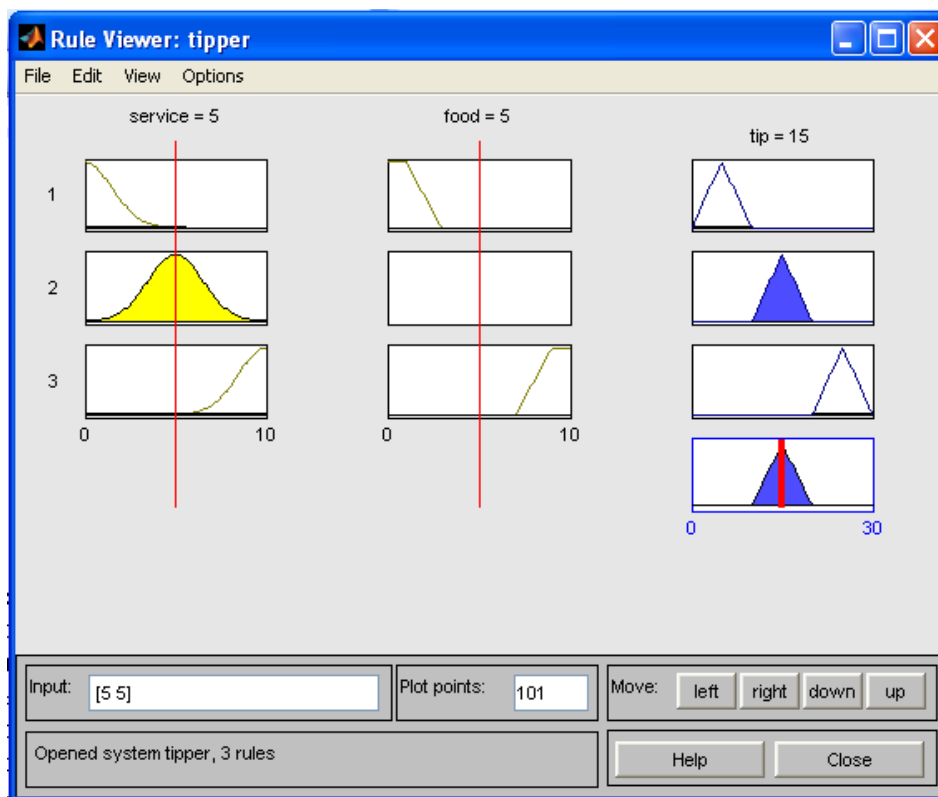


Рис. 3.21

Під трьома графіками функцій належності вихідної змінної представлений ще один, четвертий, графік, на якому зображені результати акумуляції (трикутник, залитий синім кольором) і дефазифікації (жирна вертикальна червона лінія). Рис. 3.21 відповідає ситуації, коли якість сервісу і якість їжі оцінені в 5 балів - в результаті система радить заплатити чайові у розмірі 15%.

Змінюючи значення вхідних змінних, виражені в балах, можна простежити, як ці зміни позначаються на значенні вихідної змінної, тобто на виведенні системи. Змінити значення вхідних змінних можна двома способами. Наприклад, «зачепившись» курсором мишки за вертикальний червоний візир, можна зру-

шити візир в потрібну сторону. Інший спосіб - за допомогою рядка введення Input - дозволяє більш точно встановити значення вхідних змінних.

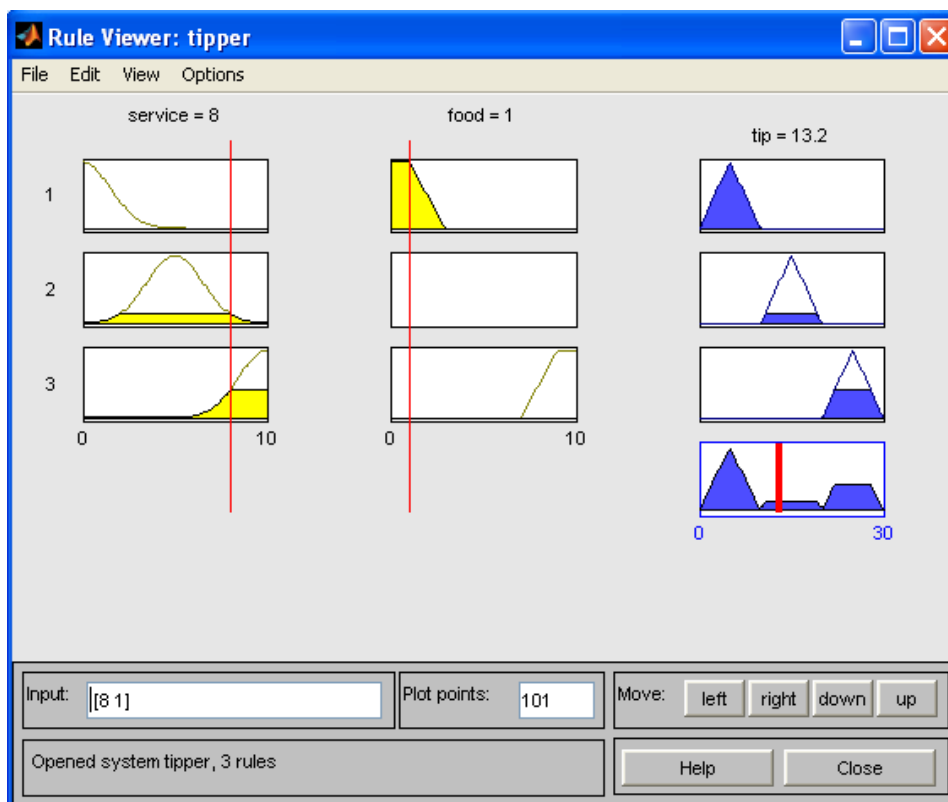


Рис. 3.22

На рис. 3.22 представлений результат таких змін для вектора значень [8 1] вхідних змінних - в даному випадку система рекомендує заплатити чайові у розмірі 13,2%.

«Граючи» зі значеннями вхідних змінних, можна оцінити якість розробленої системи і прийняти рішення про необхідність її модернізації.

Корисним також є вікно перегляду **Surface Viewer** так званої «поверхні нечіткого виведення» (рис. 3.23), також дозволяє (як і інформація у вікні **Rule Viewer**) судити про адекватність рішень, прийнятих системою нечіткого виведення. Дійсно, одного побіжного погляду достатньо, щоб дійти висновку, що дана система рекомендує варіювати розмір чайових в діапазоні 5 ... 25%. Зрозуміло, ситуація буде набагато складніше для більшого, ніж два, кількості вхідних змінних, однак в будь-якому випадку така візуалізація «станів» системи дуже корисна.

Може також виявитися корисним переглянути залежність вихідної змінної тільки від однієї вхідної змінної. У цьому випадку в рядку **Y (input)** встановлюємо значення **none**, а в рядку **X (input)** вибираємо ім'я потрібної вхідної змінної. Якщо задати **X (input) = service**, тоді результат прийме вид, як показано на рис. 3.24. Це перетин поверхні нечіткого виведення, відповідне значенню **food = 5 балів** (середнє значення з діапазону можливих значень змінної **food**). Для

отримання інших перерізів потрібно відповідним чином змінити вміст рядка введення **Ref. input**.

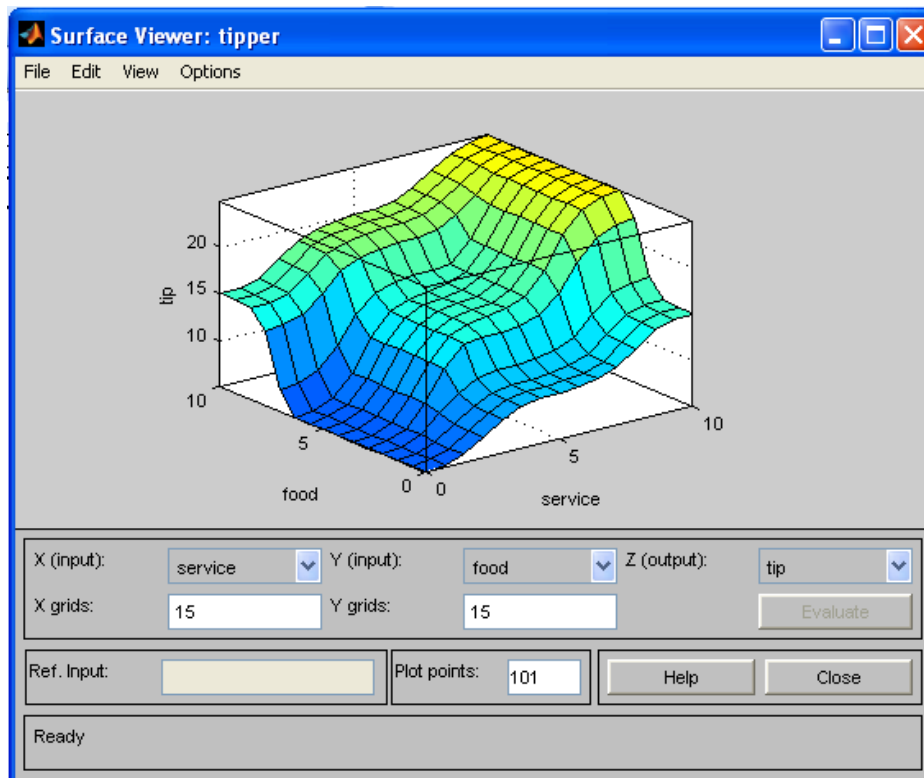


Рис. 3.23

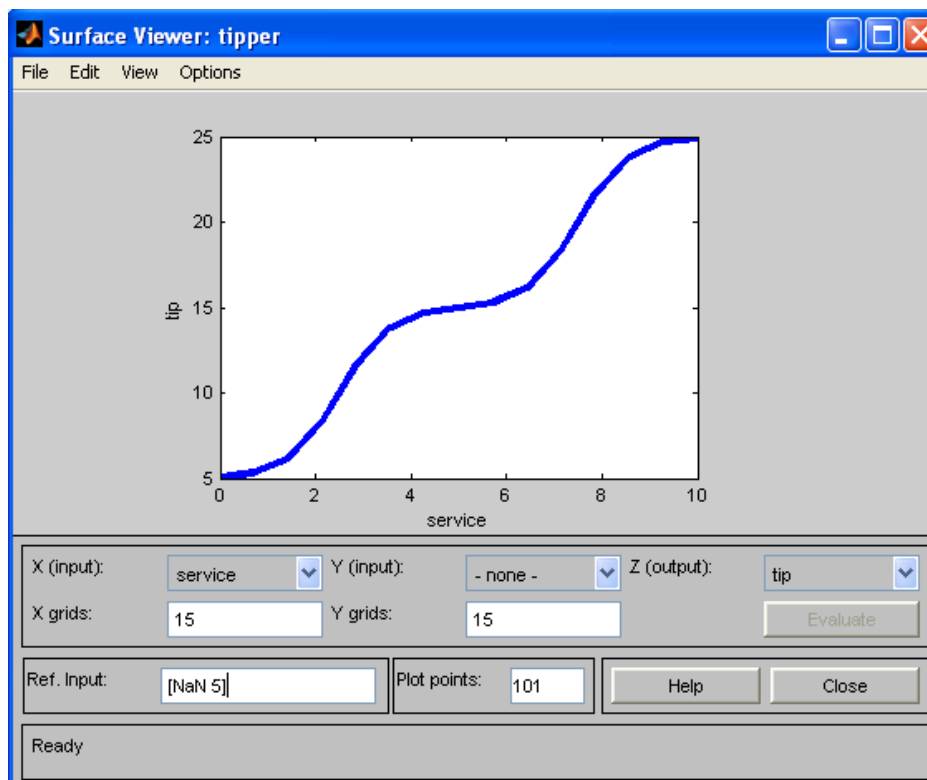


Рис. 3.24

На завершення ще раз зазначимо корисність розглянутого прикладу, уважний аналіз якого дозволяє згодом приступити до створення власних систем нечіткого виведення. У Додатку 7 пропонується комп'ютерний практикум по створенню власної найпростішої системи нечіткого висновку для медичної діагностики.

Короткі висновки

В ЕС використовуються такі основні формалізми для подання знань:

- 1) продукційні правила;
- 2) семантичні мережі;
- 3) фрейми.

Правило продукції "якщо А, то В" ("якщо є подія А, тоді є подія В") володіє незаперечними перевагами перед іншими формалізмами, оскільки легко сприймається користувачами, зокрема, лікарями при вирішенні задачі діагностики.

База знань продукційної ЕС складається з безлічі правил продукції (бази правил) і кінцевого набору фактів (бази фактів).

Продукційні системи мають переваги і недоліки.

До недоліків відносяться:

- 1) труднощі складання продукційного правила, відповідного елементу знання (потрібно, щоб розглянута область вже була достатньо вивчена, і разом з тим, рівень деталізації не повинен бути надмірно докладним);
- 2) труднощі запису правила з використанням єдиного формату запису ЯКЩО - ТО;
- 3) труднощі, пов'язана із заборонаю взаємного прямого виклику одного правила з іншого.

До переваг відносяться:

- 1) модульність (можливість неупорядкованого поповнення або видалення знань);
- 2) модифікованість (додавання нових знань не зачіпає наявних знань);
- 3) інтерпретованість (зрозумілість людині і машині);
- 4) здатність до пояснення зроблених висновків (на підставі яких отримано даний висновок);
- 5) ефективність (висока достовірність висновків).

Використання правила продукції в ймовірнісній формі, у вигляді формули Байєса, вельми зручно і перспективно, оскільки дозволяє практично за однією і тією ж схемою обчислень здійснювати висновки як в прямому ("ознака > захворювання"), так і в зворотному ("захворювання > ознака") напрямках. Причиною тому - принципова особливість формули Байєса, яка встановлює зв'язок не між "причиною" і "наслідком", а між двома довільними подіями.

Перелік причин, що ускладнюють практичне використання Байєсівської стратегії в медичних системах:

- 1) наявність статистичної залежності між ознаками;
- 2) необхідність знання апріорних ймовірностей $P(Y_j)$ захворювань Y_j (принаймні, такі знання дуже бажані);

- 3) "дефекти" даних, обумовлені неоднорідністю і неповнотою даних;
- 4) існування зовнішніх і внутрішніх "заважаючих" факторів.

Системи нечіткого виводу - це такі системи, в яких умови і укладення окремих правил формулюються у формі нечітких висловлювань щодо значень тих чи інших лінгвістичних змінних.

Основними етапами нечіткого висновку є:

- 1) формування бази правил;
- 2) фазифікація вхідних змінних;
- 3) агрегування підумови в нечітких правилах продукції;
- 4) активізація або композиція підвисновків;
- 5) акумулювання висновків в нечітких правилах продукції;
- 6) дефазифікація вихідних змінних.

Література до Розділу 3

1. Поспелов Г.С. Искусственный интеллект – основа новой информационной технологии. М., "Наука", 1988. – 269с.
2. Шишков Д.И. Данные и знания. – Искусственный интеллект, №1, 1996, с.121-127.
3. Ж.Л.Лорьер. Системы искусственного интеллекта. – М., Мир, 1991. – 566с.
4. Системы управления базами данных и знаний. Под ред. А.Н. Наумова. – М: "Финансы и статистика", 1991. – 348с.
5. Уотермен Д. Руководство по экспертным системам. М., "Мир", 1989г. – 388с.
6. Крисевич В.С. и др. Экспертные системы для персональных компьютеров. Методы, средства, реализации. Справочное пособие. – Минск, Вышэйшая школа, 1990,
7. Продеус А.Н., Сядро Т.А. Проблемы реализации Байесовской стратегии при автоматизации иридодиагностики. Сб. «Электроника и связь», №2, часть II. – К., НТУУ (КПИ), 1997г., с.312-316.
8. Вельховер Е.С. Клиническая иридология. - М., "Орбита", 1992. – 431с.
9. Математический энциклопедический словарь. М: "Советская энциклопедия", 1988. – 833 с.
10. Петенко О.В., Гречишников Н.И. Практическое пособие по иридодиагностике. – Душанбе: Изд-во "ИРФОН", 1991.
11. Мельников В.Г. Медицинская кибернетика. – К., Вища школа, 1978. – 238с.
12. Мисюк Н.С., Гурленя А.М., Лозовик В.В. Диагностические алгоритмы. – Минск, Вышэйшая школа, 1970. – 107с.
13. Мисюк Н.С. ЭВМ в диагностике нервных заболеваний. – Минск, Беларусь, 1978. – 158с.
14. Мисюк Н.С. Системы прогнозирования мозговых инсультов. (Методическое пособие). Минский госуд.мед.институт, 1983. – 59с.
15. Сядро Т.А., Продеус А.Н., Стукалин В.А. Компьютерная поддержка иридолога в практической деятельности, при обучении и в научных исследованиях. Тезисы докладов на II Украинской научно-практической конференции по народной и нетрадиционной медицине. Днепропетровск, 1996г., с.58-59.

16. Продеус А.Н. Компьютерная программа "ESID-2" для иридодиагностических исследований. – Народна та нетрадиційна медицина України, Вип.1(2), Київ, 1997, с.9-11.

17. Продеус А.М., Сядро Т.А., Верба А.І., Зафіріді О.Г., Мартиненко В.О. Свідоцтво про державну реєстрацію прав автора на твір, ПА#262. Комп'ютерна програма "ESID-2". Дата реєстрації 14 жовтня 1996р.

18. Кривенко В.В., Лисовенко Г.С., Потебня Г.П., Сядро Т.А. Иридодиагностика. Справочник. Киев, "Украинская Советская Энциклопедия" им. М.П. Бажана, 1991. – 133 с.

19. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – Санкт-Петербург: БХВ-Петербург, 2003. – 719 с.

Контрольні запитання до Розділу 3

1. Які основні формалізми використовуються в ЕС для представлення знань?
2. Що собою представляє так зване «правило продукції»? Які інші схеми логічного висновку вам відомі?
3. Вкажіть недоліки продукційних систем.
4. Перерахуйте переваги продукційних систем.
5. У чому полягає зручність і перспективність використання правила продукції в ймовірнісній формі, у вигляді формули Байєса?
6. Вкажіть причини, що ускладнюють практичне використання Байєсівської стратегії в медичних системах.
7. Що являють собою системи нечіткого виводу?
8. Що таке лінгвістична змінна? Що таке функція приналежності?
9. Перерахуйте основні етапи нечіткого виводу і вкажіть їх зміст.

Розділ 4

Вимірювання і діагностика в медичних інформаційних системах

Вимірюванням називається "... пізнавальний процес, в результаті якого можна отримати опис досліджуваного об'єкта в кількісних термінах, тобто - модель об'єкта або пов'язані з нею окремі характеристичні властивості об'єкта".

Роль вимірювань у діяльності людини величезна. Д.І. Менделєєв: "Наука починається з тих пір, як починають вимірювати. Точна наука немислима без міри". Сучасний польський метролог Я. Піотровський: "Вимірність явищ, досліджуваних наукою, характеризує міру її точності та рівня розвитку".

Теорія і практика створення систем штучного інтелекту - напрямок відносно молодий. Його поява обумовлена швидким розвитком в останні десятиліття електронних засобів автоматизації праці людини - комп'ютерної техніки та відповідного програмного забезпечення. Такі різновиди систем штучного інтелекту як системи розпізнавання образів, прийняття рішення, діагностики, що отримали останнім часом особливо широке поширення завдяки високій їх ефективності при виконанні ряду рутинних операцій - читанні та первинному перекладі текстів, швидкому пошуку і впорядкуванню інформації, прийнятті рішень в стандартних умовах і т.п. - Всі ці завдання можуть розглядатися як різновид вимірювань.

4.1. Структура і зміст теорії вимірювання

4.1.1. Теорія вимірювання як "сплав" приватних теорій

Кожна точна наука в якості найбільш істотної властивості вимірювання виділяє ту його сторону, яка є для неї найбільш важливою, і вивчення цього боку вимірювання призводить до появи тієї чи іншої приватної теорії вимірювань (табл. 4.1).

Загальна теорія вимірювань, таким чином, є якимось "сплавом" приватних теорій.

4.1.2. Рівні теорії вимірювання

Подібно до того як психічні функції людини прийнято відносити до "високого" рівня, а фізичні - до "низького", в теорії вимірювань розрізняють два рівня вивчення вимірів: "фундаментальний" і "прикладний".

Схема взаємовідносин між цими рівнями і точними науками представлена на рис.4.1.

Таблиця 4.1

Наука	Виділена властивість вимірювання	Приватна теорія вимірювання
Метрологія	Похибка вимірювань	Теорія похибок
Соціологія, психологія, системотехніка, кібернетика	Проблема вибору шкали для вимірюваної величини	Теорія шкал
Квантова фізика	Взаємодія мікрооб'єкта з вимірювальним приладом	Теорія квантовомеханічних вимірювань
Інформатика	Похибка вимірювань, розглянута як "перешкода" у "вимірювальному каналі"	Інформаційна теорія вимірювальних пристроїв
Цифрова обчислювальна техніка	Спосіб отримання числового результату вимірювань	Алгоритмічна теорія вимірювань

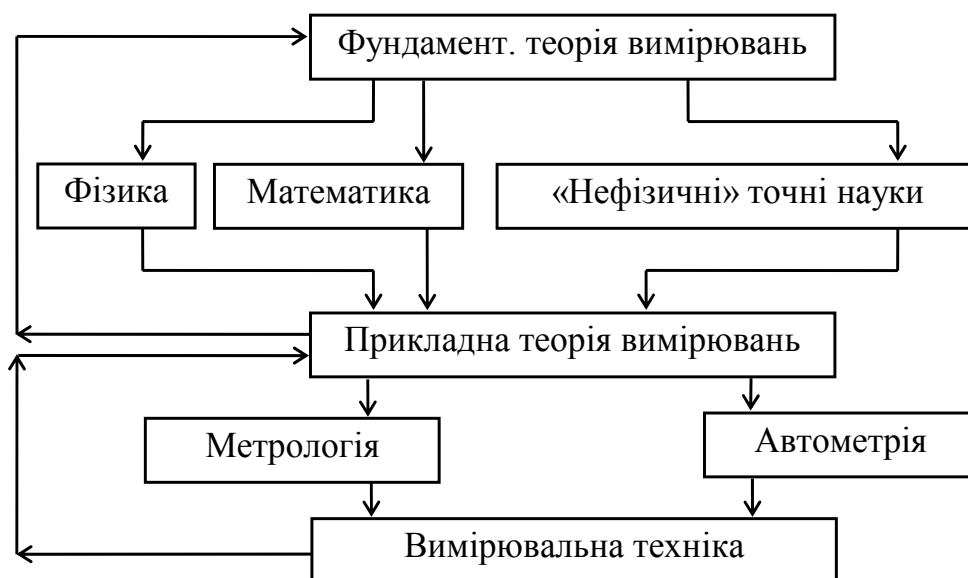


Рис. 4.1

Відповідно до цієї схеми, теорія вимірювань є "первинною". Разом з тим, про відносність «первинності» свідчить наявність зворотних зв'язків між рівнями, що значною мірою має зняти ідеологічну напруженість між представниками різних наук, що розглядають дану схему.

4.1.3. Основні напрями фундаментальної теорії вимірювання

Прийнято виділяти два основних напрямки розвитку фундаментальної теорії вимірювань: фізичне і математичне. Спрощена схема такого поділу наведена на рис.4.2.

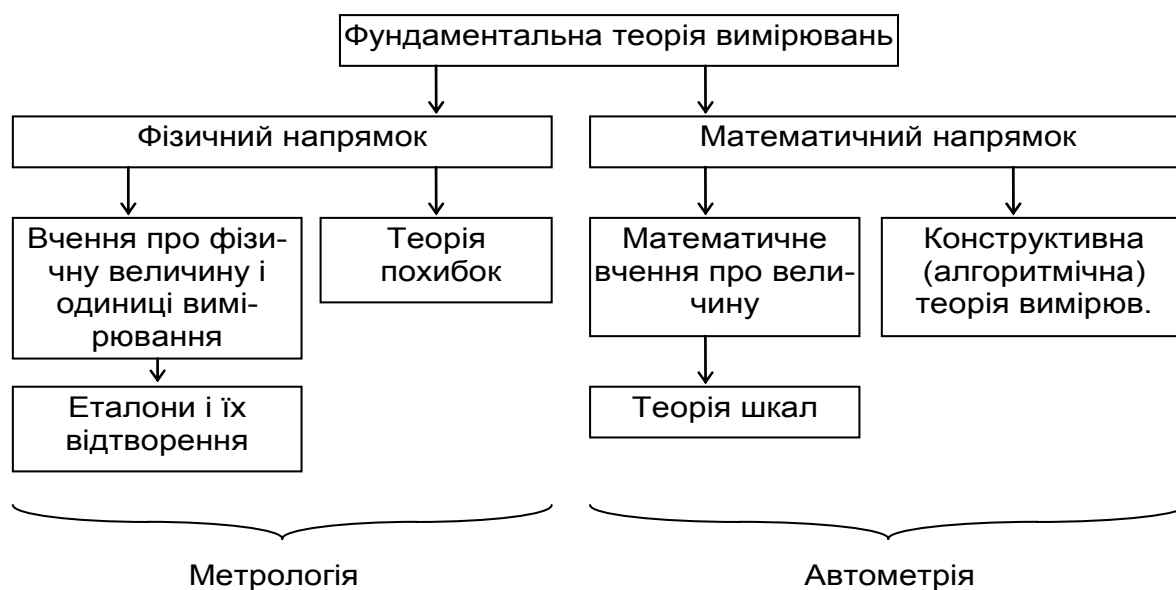


Рис. 4.2

Причину такого поділу можна побачити в існуванні двох "світів": предметного фізичного і абстрактного математичного. Протиставляючи ці світи, можна говорити про суперечливий характер вимірювання, яке є процедурою і фізичною, і математичною: починаючись як фізичний процес, воно завершується як математичний. Тут, однак, знову на передній план виступає проблема "первинності». Тому переважно, мабуть, говорити не стільки про суперечливість, скільки про взаємне доповнення цих двох напрямків.

Так, математичне вчення про величину є розширювальним по відношенню до вчення про фізичну величину, допускаючи принципову вимірність "нефізичних" величин, які не можна порівняти з еталонами в силу відсутності таких (еталоном іменують фізичну модель - річ цілком матеріальну) - дійсно, немає еталонів для таких об'єктів як здоров'я, інтелект, економіка і т.п. Цією обставиною можна пояснити важливість місця, займаного теорією шкал в математичному напрямку фундаментальної теорії вимірювань - воно аналогічне місцю, займаного проблемою еталонів у фізичному напрямку.

Нарешті, конструктивну (алгоритмічну) теорію вимірювань можна розглядати як розширення теорії похибок, спрямоване на зрівноважування важливого показника ефективності у вигляді похибки вимірювань іншим не менш важливим показником - часом вимірювань. Сенс такого врівноваження простий - нікому не потрібні високоточні результати вимірювань тоді, коли в них вже відпала необхідність.

4.1.4. Основні напрями прикладної теорії вимірювання

Двома основними напрямками фундаментальної теорії вимірювань, - фізичній та математичного, - відповідають два основних напрямки прикладної теорії вимірювань: метрологія та автоматрія.

Метрологія - наука про вимірювання, методи і засоби забезпечення їх єдності та способи досягнення необхідної точності.

Автометрія - наукова дисципліна, що займається розробкою і дослідженням методів аналізу та синтезу вимірювальних інформаційних систем в цілому, а також їх найважливіших функціональних вузлів.

Обидва напрямки взаємно доповнюють один одного і відображають два найважливіших процеси, що визначають прогрес вимірювальної техніки: "фізикалізм" (використання досягнень фізики для вдосконалення засобів вимірювань) і "кібернетизації" (застосування методів теорії інформації, теорії статистичних і спектральних вимірювань, методів і алгоритмів вимірювань, методів автоматичної корекції похибок вимірювальних пристроїв).

4.2. Фізичні і математичні моделі, що використовуються при вимірюванні

Розрізняють фізичні та математичні моделі - поняття, істотно використовувани при описі процедури вимірювань.

Математична модель - це сукупність абстрактних математичних об'єктів (число, вектор, тензор, множина, група тощо) і математичних відносин (функція, еквівалентність, порядок, істинність і т.д.), що замінює об'єкт реального світу. Безліч математичних моделей утворює тезаурус M - якийсь словник мови, наявного в розпорядженні теоретиків-дослідників.

Фізична модель - матеріалізована математична модель досліджуваного об'єкта. Сукупність фізичних моделей, що іменуються "еталонами", утворює множину E . Фізичні пристрої, які здійснюють утворення математичної моделі, називаються "мірою" (прикладі заходів: стандартний елемент для відтворення одиниці напруги; кварцовий генератор для відтворення одиниць часу або частоти; вимірювальний генератор синусоїдального сигналу з регульованими амплітудою, частотою і фазою; зразкові речовини; зразкові вироби; шкала кольорів і т.д.).

Підкреслимо взаємно однозначну відповідність елементів множин M і E .

При медичних вимірах використовують обидва типи моделей. Так, при вимірах фізичних величин (вага, довжина, частота, електричний потенціал і т.д.) використовують фізичні моделі (еталони). При вимірі ж таких характеристик як стан здоров'я, стан імунної, серцево-судинної чи іншої системи, рівень інтелекту, глибина пам'яті і т.п. використання еталонів у вигляді фізичних моделей неможливо - їх не існує. У даному випадку можна говорити лише про математичні моделі у вигляді описів, сукупності характеристик - часто їх також іменують еталонами, що зручно, хоча і не зовсім коректно з метрологічної точки зору. У цьому сенсі нозологічна форма є різновидом математичної моделі хвороби.

4.3. Вимірювання елементарні і комплексні

Розрізняють вимірювання елементарні і комплексні. Елементарне вимір являє собою процедуру порівняння значення оцінюваної характеристики (струм, напруга, сила, потужність, освітленість, вологість і т.д.) з набором еталонних значень, в результаті чого вибирається єдине еталонне значення, мінімально ві-

дрізняється від значення вимірюваної характеристики. Структура процедури елементарних вимірювань показана на рис.4.3.

Комплексне вимірювання головною своєю метою переслідує вибір виду математичної моделі, одночасно вирішуючи і завдання конкретизації значень її параметрів (завдання елементарних вимірювань). Схема процедури комплексних вимірювань, представлена на рис.4.4, таким чином, включає в себе як окремий випадок і схему процесу елементарних вимірювань.

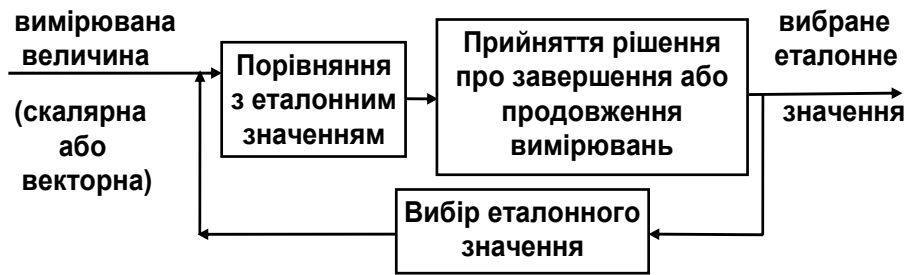


Рис. 4.3



Рис. 4.4

По суті, елементарні вимірювання відрізняються від комплексних лише рівнем складності використовуваних математичних моделей. Так, в задачі медичної діагностики до елементарних можна віднести вимірювання температури, ваги, частоти серцевих скорочень та інших характеристик фізичного і психічного стану людини, що представляють собою скалярні і векторні величини. Нарешті, до елементарних вимірювань, за визначенням, належать і завдання медичної, психологічної діагностики - в якості вимірюваної величини тут фігурує стан органу, системи, організму, психіки людини, що представляється безліччю (вектором) значень симптомів, а результатом вимірювань є нозологічна форма (математична модель хвороби).

Прикладом складних математичних моделей можуть служити зміни патологічних станів у часі, генезис, етіологія, прогнозування таких змін.

4.4. Метрологія, природничі і гуманітарні науки

На думку деяких фахівців у галузі теорії вимірювань, природничі науки мають більше "свободи": вони можуть ставити завдання або підвищення вимірюваності (збільшення кількості вимірюваних в метричних шкалах величин та збі-

льшення точності їх вимірювань), або завдання формулювання образів, що спрощують складні явища, але зручних для практичної діяльності людини.

Проблема вимірності явищ, якими займаються гуманітарні науки, більш складна. Принциповою особливістю вимірювань в гуманітарних науках є та обставина, що людина - складова частина засоби вимірювання, іноді є і єдиною. Відмінна особливість людини від технічних засобів вимірювань - нестабільність якостей. Рівень фізичного і психічного здоров'я, рівень кваліфікації, адаптаційні можливості, концентрація уваги, обсяг пам'яті, рівень інтелекту, особистісні та міжособистісні характеристики, і т.д. - всі ці властивості змінюються в часі для конкретної особистості, а в даний фіксований момент часу різні для різних людей.

Інша проблема - людина як вимірювальна система оперує з "нефізичними" моделями, з моделями математичними. Чи припустимо взагалі в цьому випадку говорити про виміри? Адже фізичні моделі (еталони), стан яких відносно легко піддається контролю, в цьому випадку відсутній.

Однієї цієї обставини (відсутність еталонів у вигляді фізичних моделей) виявилось достатньо, щоб "відлучити" гуманітарні науки від метрології. І проте, вимірювання рівня здоров'я, інтелекту, споживчих якостей товарів, стану економіки регіону або країни, художньої цінності творів літератури і мистецтва і т.д. виробляються і свідчать швидше про відоме обмеження самої метрології:

1) метрологія - наука більшою мірою фізична, ніж математична, бо "воліє" мати справу не з математичними моделями, а з фізичними;

2) ставлячи на чільне місце забезпечення високої точності та єдності вимірювань, метрологія віддає перевагу метричним шкалами, тоді як аналіз властивостей складних біологічних об'єктів з необхідністю призводить до оперування якісних, неметричних шкал.

4.5. Вимірювання в медицині і психології

При проведенні вимірювань медичного характеру можна умовно виділити дві ситуації, у відповідності з тим, використовуються при вимірюваннях фізичні моделі або не використовуються.

Перша ситуація означає застосування для діагностики деяких фізичних приладів, відповідним чином каліброваних і таким чином забезпечують прийнятну похибку вимірювань: до таких приладів належать, зокрема, кардіографи, УЗД-апарати, електроенцефалографи та ін

У другому випадку, в силу відсутності фізичних моделей, замість похибки вимірювань бажано оперувати іншими мірами близькості, що базуються на статистичному трактуванні подій. До цієї категорії належить ситуація оцінювання тих чи інших фізіологічних параметрів з використанням органів чуття людини: тактильна пульсодіагностика, психодіагностика, прослуховування тонів серця і шумів легенів, візуальне оцінювання зовнішніх характеристик людини (колір обличчя, форма і колір підочноямкових мішків, іридодіагностичні ознаки і т.п.) і ряд інших.

Приклади типових проблем вимірювального характеру, що виникають у кожній з цих ситуацій, наведені нижче. Перший приклад стосується вимірів артеріального тиску, другий - до вирішення психологічного завдання оцінки та підбору персоналу.

4.5.1. Про джерела похибок при апаратурних вимірюваннях в медицині

У медичних вимірах, вироблених із застосуванням фізичних вимірювальних приладів, беруть участь чотири об'єкти:

- вимірювальний прилад;
- пацієнт;
- лікар;
- умови, за яких проводяться вимірювання.

Кожен з цих об'єктів є джерелом похибок вимірювань. Проаналізуємо, на прикладі вимірювання артеріального тиску, внесок кожного з джерел в результуючу похибку вимірювань - важливість даної проблеми очевидна, оскільки від точності вимірювань фізіологічних характеристик залежить, в кінцевому рахунку, достовірність медичного діагнозу.

У клінічній практиці найпоширенішим є непрямий метод визначення АТ за допомогою надувної манжети, манометра і фонендоскопа. Оскільки при даного виду вимірювання використовується фізичний прилад, забезпечений метричною шкалою вимірів (манометр, одиниця вимірювань - мм.рт.ст.), дана вимірювальна задача відноситься до класу метрологічних, а основною проблемою її вирішення є похибка вимірювань (див. табл. 2.1) .

Суть методу непрямого визначення АТ полягає в тому, що накладену на плече манжету накачують повітрям до зникнення тонів Короткова, після чого повітря поступово випускають. Поява тонів Короткова (1 фаза) відповідає систолі (САТ), а зникнення (V фаза) - діастолічному (ДАТ) тиску.

Джерела похибок при такому вимірі АТ:

1) помилки *методу* вимірювання:

- найбільш точні значення артеріального тиску (АТ) досягаються при вимірюванні внутрішньоартеріального, або прямого АТ, за допомогою жорсткого катетера;

2) *апаратурні* похибки вимірювань:

- точність вимірювань АТ залежить від технічного стану приладів: наприклад, калібрування aneroidних манометрів рекомендують проводити не менше 2-х разів на рік;
- більше третини лікарів користуються aneroidними приладами, 30% яких через неточність налаштування занижують результати на 10 і більше мм рт.ст.;
- низькочастотний компонент тонів Короткова краще вловлюється безмембранним наконечником фонендоскопа, однак при користуванні ним можуть виникати похибки через необхідність щільно притискати його до руки обстежуваного (результати вимірювань занижуються); мембранний

наконечник гірше сприймає низькочастотні коливання, але не дає вищевказаних спотворень;

3) похибки через недотримання *технології* вимірювання:

- вимірювання АТ самим хворим недостатньо достовірні, оскільки здійснюються зазвичай вдома у вечірній час, коли АТ дещо нижче денного;
- одяг не повинен здавлювати руку, а сама рука повинна перебувати на рівні серця;
- надто повільна, як і надто швидка, декомпресія спотворює обидві цифри АТ;
- при неможливості скористатися фонендоскопом САД визначають методом пальпації - в цьому випадку результат вимірювань буде занижений в середньому на 7 мм рт.ст.;
- результати вимірювань залежать від індивідуального підбору розміру манжети під діаметр руки обстежуваного;

4) "*фізіологічне*" джерело похибок - особливості організму людини:

- АТ різний на правій і лівій руках; тому, наприклад, якщо відмінність більше 20 мм рт.ст. для САТ і 10 мм рт.ст. для ДАТ, рекомендують вимірювати АТ на обох руках; в іншому випадку вимірювання проводять на тій руці, де АТ вище;
- варіабельність АТ вище при повторних оглядах, ніж при повторних вимірах під час одного огляду - тому рекомендують підвищувати точність вимірювань шляхом збільшення числа оглядів, а не числа вимірів під час одного огляду;
- тихі тони Короткова, імовірно, призводять до заниження САТ і завищення ДАТ;
- у ряду обстежуваних осіб (діти, вагітні жінки, особи похилого віку, при анемії та ряді інших станів) тони Короткова при прослуховуванні повністю не зникають; в цих умовах для визначення ДАТ орієнтуються не на V, а на IV фазу, коли тони Короткова стають приглушеними, або, якщо тони Короткова взагалі слабкі, використовують прийоми їх посилення (наприклад, стиснути 10 разів руку в кулак, та ін);
- наявність аритмії, насамперед мерехтіння передсердь, істотно впливає на результати вимірювань АТ: пропущений аускультативний інтервал може значно, на 10-50 мм рт.ст., занизити САД (ДАТ в таких випадках завищується, але тільки в рідкісних випадках значно); в цих випадках рекомендують якомога повільніше випускати повітря з манжети;• у хворих, які приймають антигіпертензивні засоби, АТ слід вимірювати перед прийомом чергової дози препарату; у цих пацієнтів, крім того, слід вимірювати АТ у вертикальному положенні (через можливу ортостатичну гіпотонію), у вертикальному положенні слід вимірювати АТ і в осіб старше 65 років і хворих на цукровий діабет;

5) "*психологічне*" джерело похибок:

- існує феномен "лікарняної" гіпертензії, що виражається в стабільному перевищенні АТ під час перебування в стаціонарі, в порівнянні зі значеннями АТ поза стаціонару; цей феномен настільки міцний, що навіть коли хворий сам вимірює собі АТ в присутності лікаря, воно виявляється вище, ніж зазвичай; "лікарська" гіпертензія зустрічається у 10-40% осіб і призводить до завищення САД на 11-28 мм рт.ст., а ДАТ - на 3-15 мм рт.ст.;
 - придбання звички до вимірювань АТ зазвичай супроводжується зниженням АТ - найбільш яскраво це проявляється у недовірливих людей;
- б) правильність *підготовки* обстежуваного до вимірювань:
- перед вимірюванням артеріального тиску пацієнт повинен протягом принаймні 30 хвилин утриматися від куріння, прийому кофеїновмісних напоїв, адреностимуляторів; протягом 1 години перед вимірюванням АТ не рекомендуються фізичні вправи; крім того, значення АТ завищують також прийом алкоголю, переповнений сечовий міхур, бесіда під час вимірювань, тривожний стан обстежуваного, наявність у обстежуваного болювих відчуттів;
- 7) вплив *лікаря* на результати вимірів:
- точність вимірювань істотно залежить від особистості лікаря: розкид часто досягає 10/8 мм рт.ст.;
 - багато практикуючих лікарів не дотримуються критерії діагностики артеріальної гіпертензії, діагностуючи гіпертензію за результатами одного зареєстрованого підвищення АТ, хоча існує правило, згідно з яким необхідні як мінімум 3 вимірювання з перервами між ними в 1 тиждень і більше; крім того, більше третини німецьких і британських лікарів визначають ДАТ не по стандартним вимогам, орієнтуючись на IV, а не на V фазу тонів Короткова;
 - упередженість лікаря може призводити до завищення обох цифр АТ;
 - поганий слух лікаря іноді обумовлює заниження САТ і завищення ДАТ;
- 8) похибки через вплив умов, при яких проводяться вимірювання:
- холодна погода (порівняно з теплою) призводить до завищених значень САД в середньому на 6, а ДАТ - на 3-10 мм рт.ст.

4.5.2. Вимірювальні шкали в психології і медицині

Особливість вимірювань в психології полягає у відсутності фізичних вимірювальних приладів, тобто в процесі вимірювань беруть участь три об'єкти:

- респондент (пацієнт);
- психолог;
- умови, за яких проводяться вимірювання.

«Фізичності» стимульного матеріалу у вигляді бланків з питаннями, картинок, предметів тощо, а також використання комп'ютерів та інших програмно-апаратних засобів для цілей психодіагностики, не є еквівалентом фізичних вимірювальних приладів через відсутність фізичних моделей вимірюваних психологічних характеристик.

Розглянемо, на прикладі професіографії, як здійснюються вимірювання в психології в відсутність фізичних моделей вимірюваних величин - важливість даної проблеми очевидна, оскільки від вибору математичних моделей психологічних характеристик залежить зрештою достовірність психологічного діагнозу.

Професіограма все ширше використовується сучасними психологами, медиками та управлінцями при вирішенні такого різновиду діагностичних завдань як завдання оцінки та підбору персоналу. Причину цього можна углядіти в природному бажанні підвищити ефективність вирішення зазначених завдань за рахунок формалізації описів професій і спеціальностей. Дійсно, хороше структурування інформації шляхом продуманої системи вимірювальних шкал є першим необхідним кроком на шляху автоматизації процесу обробки інформації, що забезпечує швидкість і висока якість прийнятого рішення.

З іншого боку, ідея автоматизації професіографічної досліджень приваблива і для фахівців у галузі сучасних комп'ютерних технологій, оскільки служить гарним стимулом для подальшого розвитку теорії та програм систем штучного інтелекту.

Таким чином, актуальність проблеми формалізації вирішення завдань оцінки та підбору персоналу не викликає сумнівів. Разом з тим, пошук шляхів формалізації свідчить про наявність низки суттєвих труднощів, обумовлених, зокрема, фрагментарністю і незавершеністю результатів професіографічних досліджень і розробок. Зокрема, відсутні систематизовані і верифіковані описи професій та спеціальностей ("бібліотеки еталонних професіограм"), не стандартизовані вимоги до методик формування професіограм, не втрималась відповідна термінологія. Суттєвою перешкодою є також нечіткість позиції багатьох психологів, фактично не розрізняють професіограму - еталон і професіограму, що представляє результати обстежень конкретної людини. Причиною такої ситуації є, мабуть, нерозуміння того, що професіографія - лише одна з форм реалізації класичної схеми вимірів (рис.4.3), істотним елементом якої є процедура порівняння результату поточних вимірювань з набором еталонів.

Методики формування професіограми. Можна виділити, принаймні, три підходи до формування професіограми:

- емпіричний;
- аналітичний;
- дорадчий.

До методик першого типу, що базуються на постановці експерименту (схема якого наведена на рис.4.5) відносяться, наприклад, методики О. Ліпмана і К.К. Платонова. Реалізація цих методик полягає у використанні опитувального листа, наданого групі з 3-5 експертів, фахівців даної професії, і містить перелік професійно важливих якостей (ПВЯ). У опитувальному листі Ліпмана перераховані 151 ПВЯ, в опитувальному листі Платонова - 93. Рівень важливості кожного з ПВЯ оцінюється експертами з використанням якісних та кількісних

шкал. У методиці Ліпмана для цієї мети використовуються три шкали, з трьома градаціями кожна:

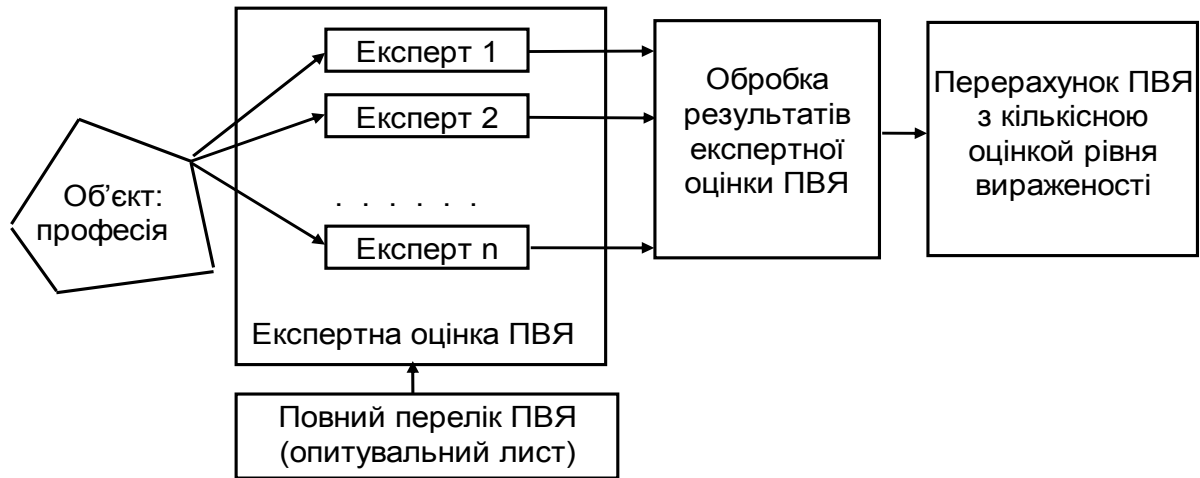


Рис. 4.5

- чи є дана якість необхідною, бажаною або байдужою (оцінки кодуються числами 2, 1 або 0, відповідно);
- чи необхідна ця якість при роботі в даній професії завжди, іноді або взагалі не потрібна (коди 2, 1 або 0, відповідно);
- чи можливо розвиток або вдосконалення цієї якості у високому ступені, в слабкому або взагалі неможливо (коди 2, 1 або 0, відповідно).

У методиці Платонова важливість ПВЯ оцінюють за єдиною десятибальною (від 1 до 10) шкалою.

Експерти виробляють оцінювання самостійно, без взаємних консультацій. По його закінченні опитувальні листи збираються і на підставі експертних оцінок розраховується усереднений бал по кожній якості. Цей усереднений бал повідомляється всім експертам з тим, щоб вони відкоригували, при бажанні, свої попередні оцінки. Якщо після повторного оцінювання все ще спостерігається розбіжність оцінок, усереднений бал вибирають шляхом колегіального обговорення.

На цьому процедуру формування професіограми, що представляє собою перелік ПВЯ, завершують. Особливістю методики Платонова при цьому є "стиснення" переліку з 93 ПВЯ в перелік з 33 ПВЯ, при якому об'єднують подібні у деякому сенсі якості (наприклад, чотири якості "рухливість", "спритність", "пластичність", "розважливість рухів" замінюють однією якістю "моторика"). Середній бал результуючої якості при цьому обчислюють як середнє арифметичне середніх балів кожного з компонентів.

В результаті застосування методик Ліпмана і Платонова користувач отримує опис професії, не отримуючи, проте, його аналітичного обґрунтування: знання, на підставі яких виставлені оцінки рівнів важливості ПВЯ, залишилися у експертів.

Інакше виглядає методика Е.М.Іванової, що представляє аналітичний підхід і

базується на логічному аналізі операційно-технологічної структури професії. Сутність цієї методики полягає у поетапному переході від опису професії до опису вимог до людини (рис.4.6). Для реалізації її досить, в принципі, одного висококваліфікованого експерта, який тісно взаємодіє з професійним психологом. Результатом їх спільної діяльності є "аналітична професіограма", що має вигляд таблиці, в останній колонці якої наведено перелік ПВЯ, а в попередніх її колонках міститься обґрунтування цього переліку.

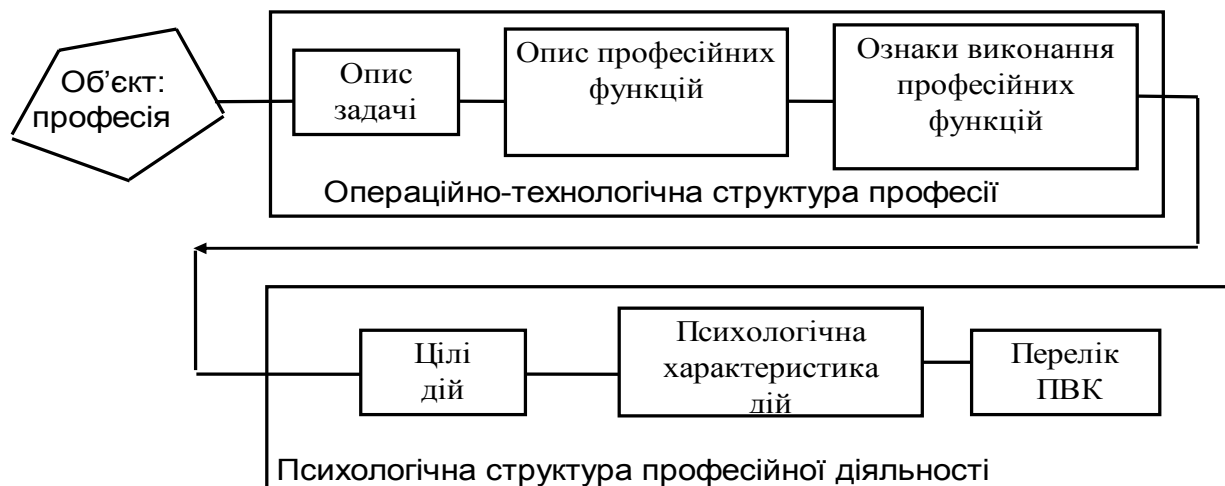


Рис. 4.6

Аналітичному підходу, таким чином, властиві такі особливості:

- знання експерта доступні для спостерігачів, оскільки зафіксовані на папері у вигляді опису операційно-технологічної структури професії;
- перелік ПВЯ отримують шляхом "вирощування потрібного", а не шляхом "видалення зайвого", як при емпіричному підході.

Дорадчий підхід полягає у проведенні колегіального обговорення експертами та психологами можливого змісту переліку ПВЯ. Особливістю такого підходу є якісний характер і неповнота одержуваного переліку ПВЯ, з неоднаковою глибиною аргументації складових його якостей.

Алгоритми розв'язання задач оцінки та підбору персоналу. Процедури вирішення завдань оцінки та підбору персоналу можуть бути представлені структурними схемами на рис.4.7-4.8.

У кожній з цих задач використовуються два переліки ПВЯ: будемо називати їх еталонним і поточним. Еталонний перелік - це перелік ПВЯ, сформований заздалегідь і службовець зразком. Поточний перелік ПВЯ - це оцінка вираженості ПВЯ у конкретної людини, що піддається випробуванню з метою видачі рекомендацій щодо розвитку ПВЯ (при оцінці персоналу) або для прийняття рішення про його професійну придатність (при відборі персоналу).

Неважко побачити, що поточний перелік ПВЯ являє собою різновид такого широко вживаного в психології поняття як «профіль особистості», а еталонний перелік за аналогією може бути названий "профілем професії". Оперування

двома зазначеними видами переліків ПВЯ необхідно для порівняння їх між собою, в результаті чого може бути сформовано висновок психолога.

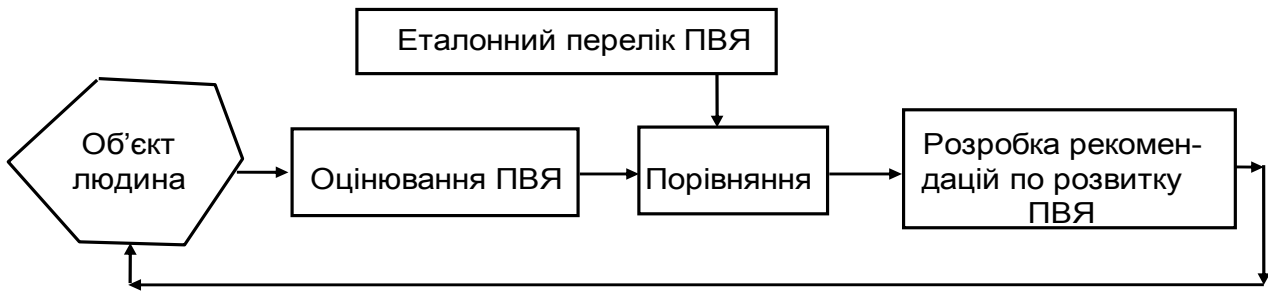


Рис. 4.7

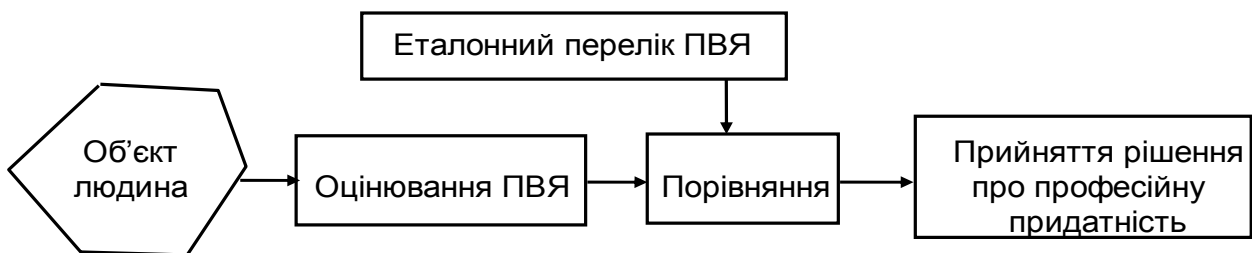


Рис. 4.8

Зіставляючи схеми на рис. 4.7 і 4.8, неважко помітити, що відмінність завдань оцінювання і підбору персоналу складається, головним чином, у відмінності алгоритмів формування висновку психолога. Вихідними даними в обох випадках є результат покоординатного порівняння еталонного і поточного переліків ПВЯ:

$$D[n] = A[n] - E[n], n = 1, 2, \dots, N.$$

Узагальнений алгоритм формування рекомендацій (в задачі оцінки персоналу) формально може бути описаний такими процедурами:

1) формують кодовану форму R рекомендацій шляхом упорядкування (по контексту) елементів множини $\{D[n]\}$:

$$R = U\{D[n]\},$$

де $U\{-\}$ – оператор упорядкування;

2) перетворюють кодовану форму рекомендацій R у вербальну форму $V = U\{V[n]\}$, ґрунтуючись на заздалегідь заданій множині однозначних відповідностей типу:

$$D[n] \leftrightarrow V[n].$$

Узагальнений алгоритм прийняття рішення про професійну придатність (в задачі підбору персоналу), у свою чергу, може бути описаний таким чином:

1) використовуючи попередньо обрану міру близькості в N -вимірному просторі, обчислюють відстань $\rho(A, E)$ між еталонним $E[n]$ і поточним $A[n]$ перелі-

ками ПВЯ; наприклад, у просторах L_1 і L_2 маємо відповідно (для дійсних функцій $E[n]$ і $A[n]$):

$$\rho_1(A,E) = |a_1 - e_1| + |a_2 - e_2| + \dots + |a_N - e_N|,$$

$$\rho_2(A,E) = [|a_1 - e_1|^2 + |a_2 - e_2|^2 + \dots + |a_N - e_N|^2]^{1/2},$$

де a_n і e_n , $n=1, 2, \dots, N$ – значення $A[n]$ і $E[n]$, відповідно;

- 2) приймаємо рішення, співставляючи величину ρ з порогом Π :
 якщо $\rho \leq \Pi$, приймаємо рішення про профпридатність;
 якщо $\rho > \Pi$, приймаємо рішення про профнепридатність;

Уніфікація і спрощення методик формування професіограми. При реалізації наведених вище узагальнених алгоритмів неминуче виникає проблема оптимізації, що стосується вибору бажаних методик формування еталонного і поточного переліків ПВЯ, вибору кількості шкал для оцінювання рівня важливості ПВЯ, кількості градацій кожної з шкал. Одночасно виникає природне запитання про можливості уніфікації та спрощення, в раціональних межах, існуючих методик.

Найбільш "природно" до вирішення згаданих питань вдається підійти, зіставляючи прийняті в різних методиках підходи до кількісного оцінювання рівня важливості ПВЯ. У методиці Ліпмана для цього використовують складову, багатокomпонентну оцінку, яку математично можна описати із залученням поняття гіперкомплексних чисел:

$$v = a + ib + jc, \quad a=0, 1, 2; \quad b=0, 1, 2; \quad c=0, 1, 2 \quad (4.1)$$

де a – оцінка ступені необхідності професіональної якості; b – оцінка того, як часто необхідна ця якість; c – оцінка можливості розвитку або удосконалення якості; i і j – уявні одиниці.

В методиці Платонова для такого оцінювання використовують єдину дійсну величину з 10 можливими значеннями:

$$v = a_1, \quad a_1 = 1, 2, \dots, 10,$$

а в методиці Іванової – єдину дійсну величину з двома можливими значеннями:

$$v = a_2, \quad a_2 = 0, 1,$$

(може створитися оманливе враження, ніби в методиці Іванової рівень важливості ПВЯ взагалі не оцінюють - непотрібні професійні якості в такому переліку ПВЯ взагалі відсутні, а ранжування наявних ПВЯ за важливістю не проводиться; проте це не так, оскільки вербальній оцінці "важливо" може бути поставлено у відповідність число "1", а оцінці "не важливо" - число "0").

Таким чином, у наявності різноманітність підходів до вибору кількості і типів шкал, що використовуються при оцінюванні рівня важливості ПВЯ, а також до вибору кількості градацій цих шкал.

У силу комплексності величин, оцінюваних за методикою Ліпмана, ця методика в її "дійсному" вигляді придатна лише для розробки еталонних переліків ПВЯ. Використовувати її при вирішенні задач оцінки і підбору персоналу не просто: по-перше, при формуванні поточного переліку ПВЯ втрачають сенс величини b ("як часто необхідно якість") і c ("можливість розвитку якості"), по-друге, формально виникає завдання порівняння комплексних функцій - завдання, розв'язане не повністю (можна, наприклад, попарно порівнювати дійсні і уявні компоненти комплексних функцій $A[n]$ і $E[n]$, оскільки відносини "більше" і "менше" не визначені на множині гіперкомплексних чисел.

У якійсь мірі уникнути зазначених труднощів можна, якщо при формуванні поточної професіограми $A[n]$ взагалі не оцінювати величини b і c , а при формуванні еталонної професіограми $E[n]$ трактувати їх як якісь вагові коефіцієнти, що підсилюють або ослаблюють ефект, який вимірюється за використанням величини a ("скільки необхідно дана якість для даної професії"). Тоді замість гіперкомплексної величини (4.1) буде використана дійсна величина

$$v=abc, a=0, 1, 2, \quad (4.2)$$

з раціональним вибором діапазонів можливих значень величин b і c (наприклад, мінімальне значення величини c має бути відмінним від нуля з тим, щоб "не вбивати" важлива якість лише на тій підставі, що воно не піддається розвитку або вдосконаленню).

При цьому, однак, виникає проблема, обумовлена неоднаковістю масштабування координат функцій $A[n]$ і $E[n]$, оскільки множення значень функції $E[n]$, на коефіцієнти b і c є ні що інше як масштабування.

Обійти цю проблему можна, взагалі відмовившись від оцінювання b і c . Тим більше що існують вагомні аргументи на користь такої відмови. Наприклад, неважко бачити тісну кореляцію значень a і b : як правило, необхідна якість необхідна часто, бажане - нечасто, байдуже - навряд чи колись буде потрібно. Тому питання "як часто потрібно дану якість?" практично без шкоди може бути вилучене з опитувальника Ліпмана. Крім того, питання про можливість розвитку або вдосконалення якості також не завжди потрібне - воно зайве в задачі "жорсткого" або "швидкого" підбору персоналу, коли достатньо знати, чи підходить дана людина для даної роботи в даний момент часу.

Зауважимо, однак, що питання про можливість розвитку або вдосконалення ПВЯ дуже важливе в задачі оцінки персоналу, коли людині потрібні рекомендації щодо коригування рівня цих якостей. У цьому випадку розумним представляється наступний вихід: значення величин c враховувати на етапі реалізації алгоритму формування рекомендацій, а не на етапі підготовки вихідних даних для цього алгоритму. До речі, такий підхід може бути узагальнений і поширений на завдання підбору персоналу - змістовна сторона завдання при цьому залишається незмінною, коригується лише розрахункова частина алгоритму прийняття рішення.

Інша важлива проблема - оптимізація кількості градацій використовуваних шкал. Вирішувати її, мабуть, доцільно окремо для кожного конкретного класу задач шляхом аналітичних або емпіричних досліджень, складність і громіздкість яких досить очевидна. Тому, щоб не відволікатися у бік вирішення оптимізаційних завдань, відзначимо лише, керуючись досить загальними міркуваннями експертного характеру, що для практичних цілей природно виходити з вимоги мінімальної кількості градацій, при якому забезпечується прийнятна якість виконання завдання. У цьому сенсі, можливо, методика Платонова "страждає" помітною надмірністю (10 градацій). З іншого боку, навіть при вирішенні практичних завдань використання 2-х градацій (в методиці Іванової) може надто загрубіть результати оцінювання. Таким чином, прийняті в методиці Ліпмана три градації - можливо, близьке до оптимального, за критерієм якість/трудомісткість, значення. В окремих випадках можна рекомендувати оцінювання у кілька етапів: спочатку по дрібно градованій шкалі, потім отримані результати заокруглюють, приводячи до більш грубої шкали.

Таким чином, методики Ліпмана і Платонова можуть бути модернізовані з метою їх уніфікації та спрощення, з подальшим використанням для формування як еталонного, так і поточного переліків ПВЯ. Методика Іванової, яка, в силу аналітичності прийнятого в неї підходу, придатна лише для формування еталонних переліків ПВЯ. Цю обставину необхідно враховувати при спробах використання методики Іванової для вирішення завдань оцінки та підбору персоналу.

Приклад рішення задачі оцінки персоналу. У 1996-97 роках у Фінансовому ліцеї Шевченківського району м. Києва вирішувалося завдання видачі психологічних рекомендацій випускникам.

Специфіка такого завдання полягала насамперед у тому, що стосовно до випускників ліцею уявлялося розумним говорити про вибір напрямку професійної діяльності, що охоплює безліч професій. Формально сформувавши еталонний перелік ПВЯ для настільки широкого напрямку як фінансово-економічна діяльність можна було б шляхом логічного підсумовування переліків ПВЯ кожного зі складових цього напрямку професій:

$$\text{ПВЯ}_{\text{фэ}} = \text{ПВЯ}_1 \vee \text{ПВЯ}_2 \vee \dots \vee \text{ПВЯ}_M,$$

де $\text{ПВЯ}_{\text{фэ}}$ – перелік ПВЯ фінансово-економічного профілю; ПВЯ_m , $m=1,2,\dots,M$ – переліки ПВЯ для кожної з професій цього профілю; \vee – символ логічного підсумовування. Однак на практиці здійснити таку процедуру виявилось неможливим через відсутність систематизованих "бібліотек професіограм". Тому був обраний інший шлях розробки еталонного переліку $\text{ПВЯ}_{\text{фэ}}$ - дорадчий. При цьому істотно використовувалися дані літературних джерел: рекомендації інституту діагностики менеджменту в Гамбурзі, переліки якостей менеджера з Брокхофу, за результатами опитування 1000 менеджерів Європи вищої та середньої ланки.

Враховуючи головну мету формованих рекомендацій, що складається в морально-психологічній підготовці 15-17-річних молодих людей до зміни стереотипів мислення і поведінки (обумовленої переходом від отримання загальної середньої освіти до отримання освіти більш високого рівня або переходом до трудової діяльності), а також враховуючи структуру державного класифікатора професій України, було вирішено структурувати видані рекомендації шляхом виділення трьох розділів:

- майбутньому працівникові фінансово-економічного профілю (14 ПВЯ);
- майбутньому управлінцю (11 ПВЯ);
- для роботи з людьми (4 ПВЯ).

При формуванні еталонного і поточного переліків ПВЯ рівень важливості та вираженості кожного з 29 якостей оцінювався за шкалами з трьох градаций "знижений-помірно-підвищено" шляхом стиснення результатів комп'ютерної, бланкової та візуальної психодіагностики, зафіксованих з використанням 42 первинних шкал. Принциповою особливістю оцінювання поточного переліку ПВЯ було істотне використання пакета комп'ютерних програм для психодіагностики, що дозволило замінити 5-7 експертів, які добре знають кожного учня, одним комп'ютером - переваги такої заміни досить очевидні. З 42 первинних якостей таким чином оцінені 28 якостей (67%), 3 якості відображали результати бланкового тестування (за період останнього року навчання), 2 - містили інформацію про успішність, 9 - оцінювалися з використанням методів і засобів візуальної психодіагностики.

Загальну структуру процедури формування психологічних рекомендацій можна показати схемою на рис.4.9.

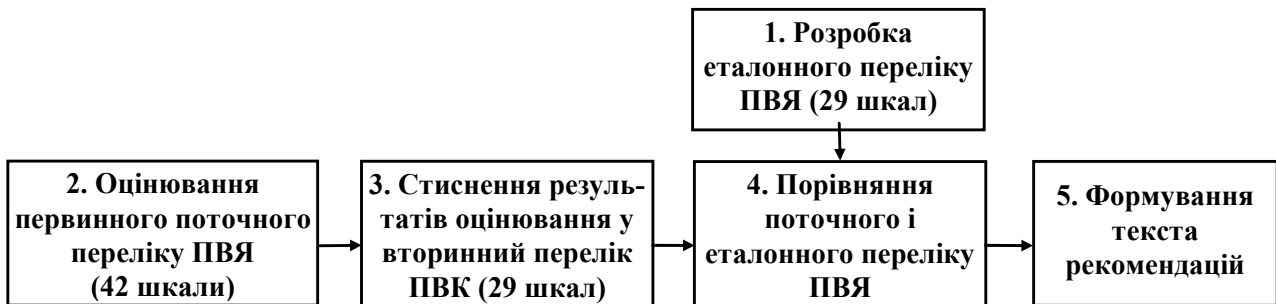


Рис. 4.9

Таким чином, ефективність вирішення завдань оцінки та підбору персоналу може бути підвищена за рахунок підвищення рівня формалізації процедур вирішення зазначених завдань, істотно базується на використанні класичної схеми вимірів (рис.4.3).

4.6. Шкали вимірювальні

У побуті під шкалою частіше мають на увазі об'єкт фізичний - частина засобу вимірювань.

Тим часом, при вимірах характеристик як фізичних, так і нефізичних об'єктів (у медицині та психології нефізичними, в сенсі відсутності їх фізичних еталонів, є, наприклад, такі об'єкти як здоров'я, інтелект, характер тощо) велике значення має розгляд шкали як математичного об'єкта.

Зокрема, при психологічному тестуванні, як вже зазначалося в розділі 4.5, об'єкт досліджень настільки складний, що говорити про можливість фізичного моделювання психічних явищ, нехай навіть з використанням сучасної обчислювальної техніки, можна лише з надзвичайною обережністю. Проте, вимірювання подібного роду настільки важливі і необхідні, що відкидати їх ніяк неможливо. До певної міри вихід з цієї складної ситуації вдається намітити шляхом розробки теорії шкал - перші роботи в цьому напрямку були написані відносно недавно і стимулювалися якраз необхідністю теоретичного обґрунтування психологічних вимірювань.

4.6.1. Поняття шкали в фізичному напрямі теорії вимірювання

У ГОСТ 16263-70 визначені два види шкал - шкала фізичної величини (шкала величини) і шкала засобів вимірювання (шкала). Шкала величини - це "послідовність значень, присвоєних у відповідності з правилами, прийнятими за угодою, послідовності однойменних фізичних величин різного розміру". Шкала засобів вимірювання - це "частина відлікового пристрою засоби вимірювань".

Коментуючи ці визначення, відзначимо, що шкала величин - поняття швидше математичне, незважаючи на те, що мова йде про фізичну величину, тоді як шкала засобів вимірювання - річ цілком матеріальна, "фізична".

Неоднозначність визначення поняття «шкала» проглядається і в підходах різних фахівців-метрологів.

У літературі поняття "шкала" згадується в двох контекстах. По-перше, вказується, що шкалою може бути забезпечена міра (нагадаємо, що міра - фізичний пристрій, що відтворює фізичні моделі, іменовані еталонами). По-друге, вказується, що шкалою може бути забезпечений компаратор - фізичний пристрій, що реалізує порівняння значення вимірюваної фізичної величини з набором еталонів. При цьому вказується на існування двох видів шкал:

1) шкали, проградуїровані так, що положення деякого покажчика порівнюється з положенням найближчого ділення шкали (прикладом є стрілочні ваги в продуктових магазинах, що мають шкалу і показник-стрілку);

2) шкали з довільними розподілами, одне з яких відповідає нульовий різниці між значенням вимірюваної величини і одним з еталонних значень (прикладом є рівноплечові ваги з врівноваженими гирями і так-звані "мости" для вимірювання електричних величин - опорів, ємностей, індуктивностей і т. п.).

У математичній літературі шкалою також іменується взаємно-однозначне відображення f множини станів $A=(a_1, \dots, a_N, \dots)$ характеристики Q об'єкта O в множини значень величин $B=(b_1, \dots, b_N, \dots)$ величини Q_1 образи об'єкта O_1 :

$$b = f(a).$$

Таким чином, у наявності двозначність ставлення до поняття "шкала", обумовлена подвійністю самої природи людини, - фізичної та психічної, - його здатністю оперувати як з матеріальними об'єктами, так і з абстрактними.

4.6.2. Поняття шкали в математичному напрямі теорії вимірювання

Точка зору математичного напрямку теорії вимірювань на процедуру вимірів до певної міри відображена в наступному формулюванні Кемпбелла: "вимір являється процесом присвоєння чисел речам (подіям) згідно деякій системі правил".

Зовні це визначення істотно відрізняється від визначення метрологів, схематично відображеного на рис.4.3:

1) відсутнє в явному вигляді уявлення про похибки вимірювань, обумовлене відмінністю вимірюваної характеристики і одним з безлічі її еталонних значень;

2) не вказана ітераційність (циклічність) характеру процедури вимірювань, яка завершується за умови, що похибка вимірювань стає менше заданої величини;

3) не відображені роль і місце класифікації моделей при вимірах, нехай навіть мова йде не про фізичні, а про математичні моделі.

У цьому зв'язку закономірне запитання: невже вимірювання в природничих та гуманітарних науках (а математичний напрям покликаний на допомогу саме представниками гуманітарних наук) відрізняються настільки сильно по суті? Чи ця відмінність чисто формальна, обумовлена відомою "самодостатністю" математики, представники якої не побажали приводити свої визначення у відповідність з визначеннями метрологів?

Відповідь на поставлені питання спробуємо знайти, звернувшись до визначень шкали, де вгадується прагнення узагальнити поняття вимірювань таким чином, щоб вона влаштувала представників як природних, так і гуманітарних наук.

У математиці вимірювальною шкалою називають "гомоморфності відображення властивостей математичним простором". У роботах математиків П.Суппеса і Дж.Зінеса (що викликали початок сучасної математичної психології та розвинули теорію шкал, в основі якої теоретико - множинний апарат відносин), шкалою називають сукупність правил, які дозволяють виконати зіставлення W (встановити гомоморфізм) емпіричної системи E в числову систему N ($N = \{M, P_i\}$, де M - безліч чисел, а P_i - відносини на числах), тобто шкала представляється трійкою $\langle E, N, W \rangle$.

Оскільки "сукупність правил" є необхідна передумова створення алгоритмів, істотно використовується поняття "алгоритмічне вимір" як деякий спосіб (алгоритм) одержання (знаходження) символу (числа, функції, алгебраїчного елемента), моделює властивість досліджуваного об'єкта відповідно до конкретної вимірювальної шкали. При цьому невідомі властивості об'єкта порівнюються з еталонними властивостями до тих пір, поки не знайдеться символ, що моделює досліджувану властивість.

При використанні таких, більш "компромісних", визначень понять шкали не-

важко бачити принципову схожість поглядів на вимірювання прихильників фізичного і математичного напрямків. Наявні відмінності хоча й істотні (замість фізичних моделей використовуються математичні), проте неминучі в силу специфічної природи об'єктів вимірювань (не можна "помацати", наприклад, такі об'єкти як праця, популярність, егоїзм і т.п.), з якими доводиться мати справу при економічних, соціальних, психологічних і т.п. "нефізичних" вимірах.

4.6.3. Види шкал

Розрізняють, головним чином, чотири види шкал:

- 1) найменувань (номінальна, номінативна);
- 2) порядку (порядкова, ординальна);
- 3) інтервалів (інтервальна);
- 4) відносин (масштабна).

Іноді до них додають п'яту шкалу: абсолютну (натуральну).

Інтенсивно розвинута в останні роки теорія шкал вимірювань базується на наступній аксіоматичній основі.

Тотожність:

1. Або $A = B$, або $A \neq B$.
2. Якщо $A = B$, то $B = A$.
3. Якщо $A = B$ і $B = C$, то $A = C$.

Ранговий порядок:

4. Якщо $A > B$, то $B < A$ (асиметричність).
5. Якщо $A > B$ і $B > C$, то $A > C$ (транзитивність).

Адитивність:

6. Якщо $A = P$ і $B > 0$, то $A + B > P$.
7. $A + B = B + A$
8. Якщо $A = P$ і $B = Q$, то $A + B = P + Q$.
9. $(A + B) + C = A + (B + C)$.

Шкала найменувань базується на аксіомах 1, 2, 3. Це якісна, тобто нечислова шкала, що представляє собою безліч якихось найменувань. Допускає деякі статистичні операції: підрахувати число індивідів кожного найменування, визначити частоту (відносна кількість) індивідів кожного найменування, знайти найменування з найбільшою кількістю (мода) індивідів. Допустимі будь-які перетворення шкали.

Шкала порядку базується на аксіомах 1-5 і також не є числовою (приклад: упорядкування по порядку зростання номерів, в алфавітному порядку тощо). Розрізняють три види шкал порядку:

- 1) шкала простого порядку;
 - 2) шкала слабкого порядку:
- 4'. Або $A \geq B$, або $A \leq B$ (антисиметричність).

5'. Якщо $A \geq B$ і $B \geq C$, то $A \geq C$.

3) шкала часткового порядку (деякі елементи множин непорівнянні).

Допустимі математичні операції: частоти, моди, медіани, коефіцієнти рангової кореляції. Інтервали між елементами шкал порядку різні. Допустимі монотонні перетворення шкали (логарифмування, зведення в степінь і ін.).

Прикладами застосування порядкових шкал в медицині є: шкала стадій гіпертонічної хвороби по М'ясникову, шкала ступенів серцевої недостатності по Стражеска-Василенко-Лангу, шкала вираженості коронарної недостатності по Фогельсону.

Шкала інтервалів є числовою рівномірною: інтервали між елементами шкали рівні. Більше того, ці інтервали самі можуть бути впорядковані (двічі впорядковані шкали). Шкала допускає перетворення виду $y = ax + b$, де $a > 0$, b - дійсне число, що означає відносність нульової точки шкали. Як наслідок - вимірювані за цією шкалою величини не можна ділити, множити, додавати і віднімати (наприклад, питання "у скільки разів гарячіше?" Некоректний, оскільки вимірювання, наприклад, за шкалами Цельсія і Фаренгейта будуть призводити до різних відповідей). Допустимі математичні операції ті ж, що і для шкали порядку, а також: математичне сподівання, середньоквадратичне відхилення, коефіцієнт асиметрії, змішані моменти. Неприпустимо обчислення коефіцієнту варіації через залежність результату від вибору нульової точки шкали.

Шкала відносин базується на аксіомах 1-9, являється числовою, рівномірною, нуль шкали "природний". Допустимі всі арифметичні і статистичні операції. Допустимо перетворення шкали виду $y = ax$, де $a > 0$.

Абсолютна шкала є числовою, не піддається ніяким перетворенням, використовується в метрології для вимірювання квантових постійних і відносних величин (коефіцієнти підсилення, тертя, добротність, ймовірність, к.к.д., інформація тощо), її використання не вимагає наявності еталонів (фізичних моделей).

Шкали найменувань і порядку відносяться до емпіричних (неметричних) шкал і характеризуються одиницями вимірювань, що не входять в систему одиниць СІ. Абсолютна шкали інтервалів є метричною - одиниці вимірювань за цією шкалою входять до складу системи одиниць СІ. (Зауважимо, що система СІ має відношення тільки до вимірювань фізичних величин).

4.6.4. Шкали в психології

Розгляд системи психологічних шкал дуже корисно і для задачі медичної діагностики:

- шкальне оцінювання базується на статистичному підході до визначення поняття «норма»;
- в задачі медичної діагностики поняття «норма» і «відхилення від норми» також базуються на статистичному підході;
- в психологічній діагностиці напрямок шкального оцінювання розвинено найбільш сильно з тієї простої причини, що це єдиний спосіб кількісного опису об'єктів дослідження, що має нефізичний характер (відсутні фізичні моделі сміливості, тривожності, інтелекту і т.п.);

- у медичній діагностиці об'єкт досліджень, як правило, цілком «фізичний», тому є можливість використання фізичних вимірювальних засобів, відтворюючих фізичні моделі вимірюваних фізичних величин (електричний потенціал, електричний опір, вага, довжина і т.п.), проте це справедливо лише для етапу збору вихідних даних - на етапах ж аналізу даних і прийняття рішення використовуються такі математичні моделі як «стан», «синдром» і т.п.

4.6.4.1. Первинні і вторинні оцінки

Розрізняють дві великі групи шкал - шкали для *первинних оцінок* і шкали для *вторинних оцінок*.

Первинні оцінки іменують також «сирими» - зазвичай це відомості про кількість правильно вирішених завдань, зокрема спроб при їх вирішенні, частоти звернень до окремих «тем», рідше - про час виконання завдання. У певній групі методик (функціональні проби, методики з якісним аналізом і інтерпретацією результатів) первинні оцінки є остаточним результатом, на підставі якого здійснюється інтерпретація даних і формування висновку.

Однак у більшості психологічних тестів (психометричні методики, особистісні опитувальники) шкали первинні трансформуються у вторинні шкали. Вторинні оцінки при цьому іменуються шкальними (на відміну від «сирих», первинних).

Шкальні оцінки в психології мають чітко визначений кількісний зміст і спрямовані на досягнення двох основних цілей:

- визначення місця індивідуального результату в стандартному розподілі групових даних;
- забезпечення порівняння кількісних результатів різних тестів, можливості їх спільної інтерпретації, відомості до єдиної системи.

Шкали при цьому містять дані про внутрішньогрупові норми виконання даної методики у вибірці стандартизації. Так, індивідуальне виконання завдань (первинні оцінки піддослідних) порівнюється з їх виконанням в порівнянні нормативної групі (наприклад, досягнутий результат порівнюється з показниками осіб в заданих вікових межах).

Процентилі. Процентиль - це відсоткова частка індивідів з вибірки стандартизації, результат якої нижче даного первинного показника. Шкала процентилей - рангова, при числі рангів 100 і відліку від 1-го рангу, відповідного самому низького результату; 50-й процентиль відповідає медіані розподілу результатів.

Процентилі не слід змішувати зі звичайними відсотковими показниками, котрі представляють собою частку правильних рішень із загальної кількості завдань тесту в індивідуальному результаті.

Основний недолік процентильних шкал полягає в нерівномірності їх градуювання (рис.4.10). Так, при нормальному розподілі, окремі значення тісно групуються в центрі шкали, розсіюючись в міру віддалення до країв. Тому рівним частотам випадків поблизу центру відповідають більш короткі інтервали.

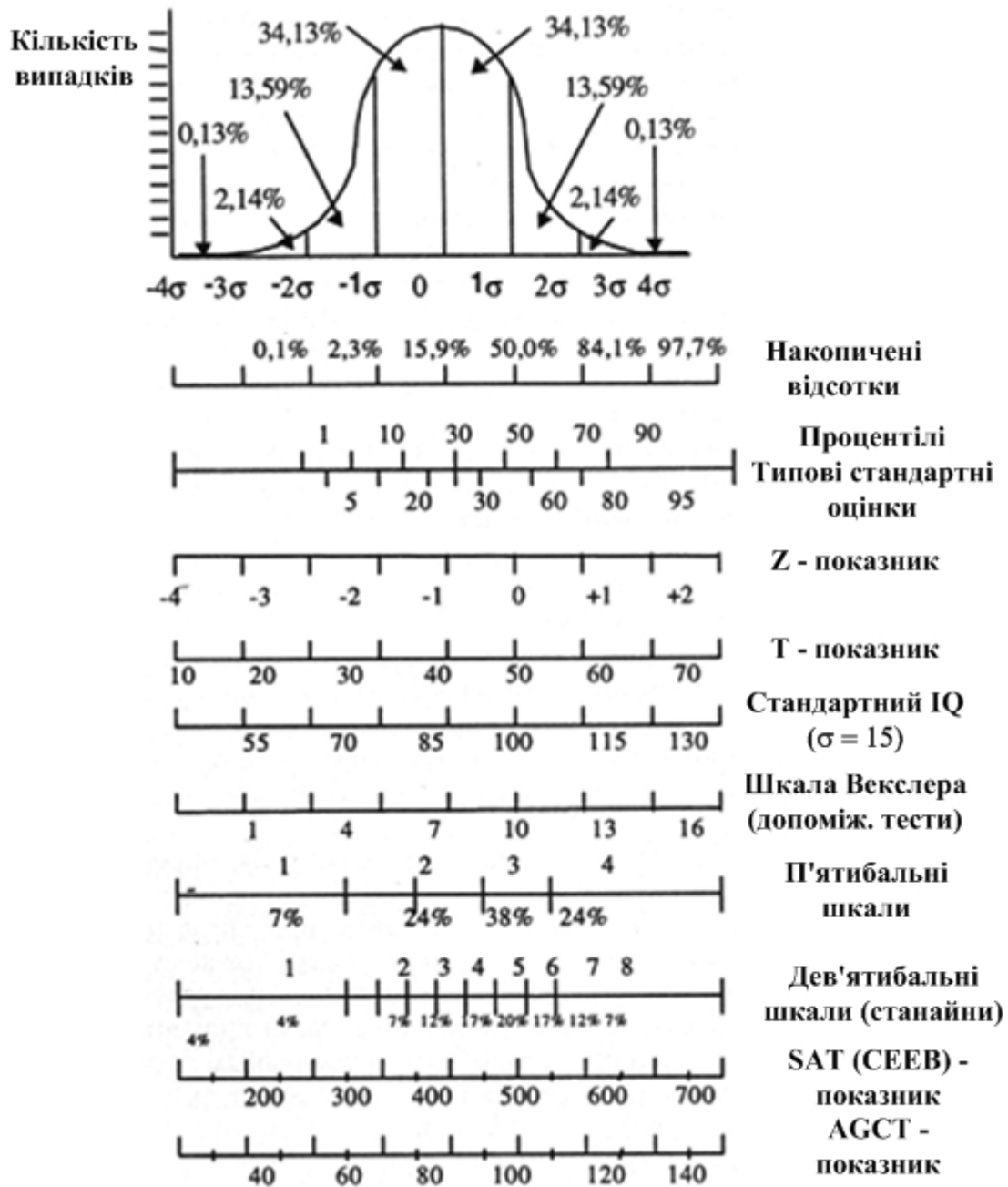


Рис. 4.10

Разом з тим, процентильні оцінки мають і низку переваг - вони легко доступні розумінню користувачів, універсальні по відношенню до різних типів методик і легко розраховуються.

Процентильні оцінки не відносяться до типових шкальним показникам. Більш широке поширення одержали стандартні показники.

Стандартні шкали. Виходять шляхом лінійного або нелінійного перетворення результатів первинних вимірювань.

Технічно стандартизацію здійснюють у три етапи. На першому етапі встановлюють нормальність розподілу первинних оцінок - якщо це не так, тоді вдаються до штучної нормалізації шляхом нелінійного перетворення вихідної шкали.

На другому етапі шкалу первинних оцінок перетворюють в z-шкалу шляхом центрування і нормування:

$$z_i = \frac{x_i - \bar{x}}{\bar{\sigma}},$$

де x_i – значення початкової оцінки; \bar{x} – оцінка середнього значення; $\bar{\sigma}$ – оцінка середньоквадратичного відхилення. На третьому етапі z – шкалу трансформують в зручну для практичного користування шкалу шляхом вибору інших значень математичного очікування M_S і середньоквадратичного відхилення σ_S :

$$z_{iS} = \sigma_S z_i + M_S.$$

Перевагою z-шкали є рівномірність її градування - одиниця вимірювання в шкалі z дорівнює 1σ стандартного (одиничного) нормального розподілу. Не завжди зручно те, що одиниці виміру занадто великі: в межах інтервалу $[-3, +3]$ уміщається переважна більшість (97,8%) результатів вимірювань. На усунення цього недоліку і спрямований третій етап.

Прикладами змінних від z-шкали є шкала IQ-показника ($M_S=100$, $\sigma_S \approx 16$), шкала Векслера ($M_S=10$, $\sigma_S \approx 3$), SAT (CEEВ)- шкала ($M_S=500$, $\sigma_S \approx 100$), T-шкала ($M_S=50$, $\sigma_S \approx 10$).

Станайни. Це своєрідний «гібрид» шкали процентилей і стандартної шкали (англ. standard nine - стандартна дев'ятка, $M_S=5$, $\sigma_S \approx 2$), який отримує все більше поширення і поєднує в собі переваги стандартних шкальних показників і простоту процентилей. Техніка формування такої шкали надзвичайно проста: первинні оцінки ранжуються за зростанням результатів, з наступним розбиванням їх на 9 груп, як показано в табл. 2.2.

Таблиця 4.2

Відсоток дослідження в виборці стандартизації	4	7	12	17	20	17	12	7	4
Станайн	1	2	3	4	5	6	7	8	9

4.6.4.2. Форми представлення шкал

Прийнято розрізняти такі форми подання шкал:

- вербальна;
- графічна;
- числова;
- таблична.

Вербальна шкала. Являє собою набір суджень про наявність або ступінь вираженості оцінюваної ознаки. Розрізняють два типи вербальних шкал:

- згорнуті;

- розгорнуті.

Згорнута шкала - це, по суті, лише найменування вербальної шкали. Приклади згорнутих шкал: «лабільність-ригідність», «екстраверсія - інтроверсія» і т.п.

Розгорнута вербальна шкала включає визначення ступеня вираженості ознак. Наприклад, характеризуючи результат операції, можна використовувати шкалу з градаціями: відмінно-задовільно-незадовільно.

Крім того, розрізняють вербальні шкали:

- монополярні;
- біполярні.

Монополярні (уніполярні) - характеризують ступінь вираженості властивості. Біполярні - характеризують наближення до одного з протилежних полюсів властивостей.

Прикладом уніполярної шкали є шкала оцінювання ступеня уваги: дуже стійке-стійке-лабільне-розсіяне. Прикладом біполярної шкали є шкала виду: здоровий - і не здоровий, і не хворий - хворий.

Переваги вербальних шкал - легкість сприйняття випробуваним.

Недолік - неоднозначність трактування випробуваним, що ускладнює застосування вербальних шкал при створенні точного психометричного інструменту.

Графічна шкала. Це форма фіксації значення ознаки у вигляді таких графічних об'єктів як лінія, деяка фігура. Графічна шкала рідко використовується в чистому вигляді - зазвичай вона комплексуються з вербальної і (або) числової шкал.

Розрізняють графічні шкали:

- горизонтальна;
- вертикальна.

Прикладом горизонтальної шкали є група завдань методики дослідження самооцінки:

Здоровий	3 2 1 0 1 2 3	Хворий
Свіжий		Втомлений
Витривалий		Виснажений
Бадьорий		В'ялий
Напружений		Розслаблений

і т.д.

У цьому прикладі горизонтальна графічна шкала застосовується в поєднанні з вербальною та числовою.

Приклад вертикальної шкали наведено на рис.4.11. Завдяки безперервності (по вертикалі) шкали з'являється можливість отримувати більш дробові, диференційовані значення оцінок.

Числова шкала. Являє собою обмежену безліч послідовних чисел. Ця форма вагів найбільш зручна для подальшої статистичної обробки результатів дослідження. Оскільки суб'єктивне сприйняття такої форми шкали для досліджуваних досить ускладнено, в чистому вигляді числова шкала майже не застосову-

ється, а об'єднується з вербальними і графічними шкалами. У будь-якому випадку числова шкала вимагає пояснень у вербальній формі числових градацій або позначення полюсів шкали. Прикладом змішаної вербально-числової шкали є наведена вище група завдань методики дослідження самооцінки.

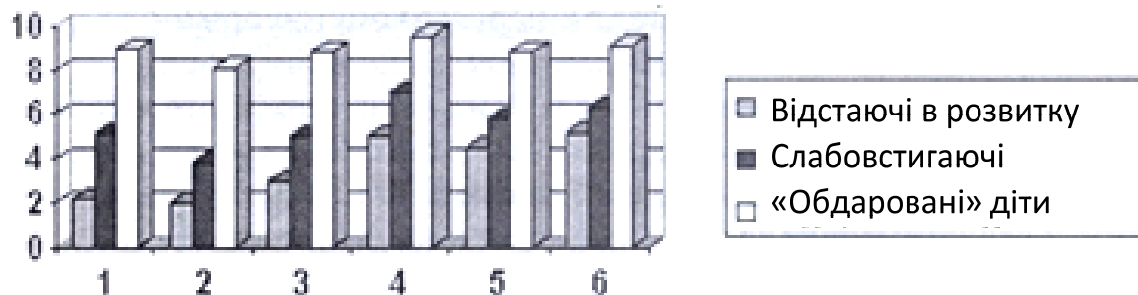


Рис. 4.11

Таблична шкала. Являє собою обмежену безліч послідовних чисел або найменувань. Зручна для подальшої ручної обробки результатів дослідження.

Зазвичай об'єднується з вербальними і числовими шкалами. Прикладом табличної шкали є 12-бальна шкала Бофорта для визначення сили вітру (див. табл. 4.3).

4.6.4.3. Зв'язок між шкалами

Як впливає з наведених вище прикладів, часто, представляючи ту чи іншу шкалу, пропонують її в сукупності з іншими шкалами - при цьому втрачається "чистота форми", зате поглиблюється і конкретизується зміст (так у розмові вдаються до оборотів "те-є", "іншими словами", "точніше" і т.п.). У цьому випадку перевагою є говорити не стільки про зв'язок, скільки про принципову сумісність різних форм.

Про зв'язок між шкалами доцільно говорити, маючи на увазі, по-перше, шкали первинних і вторинних оцінок, а по-друге, шкали різних видів: найменувань, порядку, інтервальну і відносин.

Прикладом існування такого зв'язку може бути рис.4.10, де самий верхній графік нормальної щільності розподілу ймовірностей представляє результати первинного оцінювання і, взагалі кажучи, може відноситися до шкал будь-якого виду (хоча шкала найменувань, елементи якої допускають довільну перестановку, стоїть дещо окремо - із-за такої перестановки форма розподілу ймовірностей буде змінюватися).

Всі інші графіки є прикладами шкал для вторинних оцінок. Серед них до шкал порядку відносяться шкали центилей, станайнів, п'ятибальних відсоткових оцінок. До інтервальних шкал відносяться z-шкала, T-шкала, IQ-шкала, шкала Векслера, SAT (СЕЕВ)-шкала і AGCT-шкала.

Домовимося називати такий тип зв'язку між шкалами «вертикальним». Це зручно, оскільки можна говорити про «горизонтальний» зв'язках між шкалами, маючи на увазі шкали, які використовуються для опису вихідних даних якоїсь

задачі, і шкали вихідних даних (результатів) цього завдання. Дійсно, якщо завдання зобразити у вигляді якогось прямокутника («ящика») із входом і виходом, як це роблять в теорії систем, тоді це завдання можна розглядати як процедуру трансформації шкал вихідних даних у шкали вихідних даних (рис.4.12).

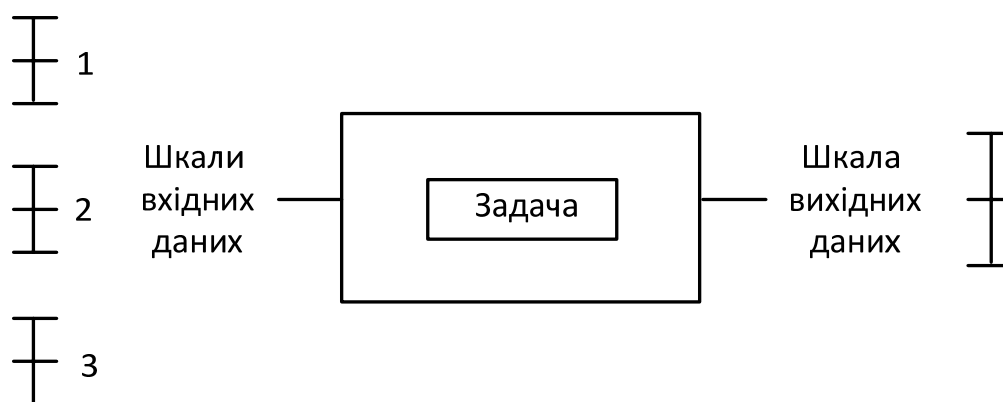


Рис. 4.12

Під завданням, зокрема, може матися на увазі завдання психологічної діагностики - але це може бути і завдання медичної діагностики і, взагалі кажучи, практично будь-яка задача перетворення інформації.

При розгляді питання про вимірювальні шкали в літературі часто наводять як приклад емпіричної шкали 12-бальну шкалу Бофорта для визначення сили вітру, фрагмент якої (опущені бали 4-12) наведено в табл. 4.3.

Таблиця 4.3

Бал за шкалою Бофорта	Назва вітру	Зовнішні ознаки у відкритому морі	Зовнішні ознаки на суші	Швидкість вітру, м/с
0	Штиль	Море гладке, як дзеркало	Дим піднімається вертикально	0-0,2
1	Тихий вітер	Дрібна рябь	Дим піднімається не зовсім вертикально	0,3-1,5
2	Легкий вітер	Дуже маленькі і короткі хвилі, які не заломлюються	Відчуття подиху на обличчі	1,6-3,3
3	Слабкий вітер	Маленькі і короткі хвилі, які заломлюються гребенями і дзеркальною піною	Ворушаться листки на деревах	3,4-5,4

Використовуємо цю таблицю для демонстрації зазначених вище зв'язків між шкалами. Так, перша колонка являє шкалу порядку. Друга, третя і четверта колонки, представлені окремо від першої - шкали найменувань. П'ята колонка, розглянута окремо, також є шкалою найменувань, отриманої з шкали відносин шляхом виділення її ділянок. Об'єднання цих п'яти колонок в рамках однієї таблиці, проте, дозволяє говорити в кінцевому рахунку про шкалу порядку в табличній формі. Користь такого подання шкали Бофорта очевидна: формулю-

вання 1-4-ї колонок визначають градації сили вітру з достатньою для практики мореплавання точністю, при цьому числова форма градування шкали (перша колонка) найбільш лаконічна, а вербальна форма (інші колонки) - найбільш змістовна. Наявність п'ятої колонки свідчить про принципову можливість вимірювань сили вітру зі значно більшою точністю - в даному випадку з точністю до десятих часток метра / сек.

Цю ж таблицю можна розглядати і як діагностичну (зв'язок між шкалами «по горизонталі»), що дозволяє вирішувати задачу діагностики. Дійсно, в колонках з третьої по п'яту представлені шкали діагностичних ознак, тоді як в першій і другій колонках - шкали результатів діагностики.

На закінчення розгляду цього прикладу відзначимо, що п'ята колонка в табл. 4.3 відіграє роль «п'ятої колони», допускаючи двояке трактування - і як шкали ознак, і як шкали результатів діагностики. Дійсно, якщо ми маємо в своєму розпорядженні прилад для вимірювань швидкості вітру, тоді це ознака, на підставі вимірів якого ми приймаємо рішення про силу вітру. Якщо таким приладом ми не володіємо, тоді ознаками можуть служити лише результати візуальних спостережень - в цьому випадку ми також приймаємо рішення про силу вітру, одночасно роблячи висновки про приблизну швидкість вітру. Зазначена властивість п'ятої колонки дозволяє нам припустити, що поняття ознака і клас (гіпотеза), застосовувані в задачі діагностики - це поняття не абсолютні, а відносні: те, що в одній ситуації є ознакою, в іншій ситуації може бути класом, і навпаки.

4.6.5. Рівень ("сила") шкали

Чим вище рівень шкали, тим більше статистичних і математичних дій можна виконати над отриманими при вимірах числами.

Зі згаданих у п.4.6.3, шкала відношення є найбільш "сильною", а шкала найменувань - найбільш "слабкою". У поняття "сили" шкали при цьому вкладають наступний сенс: властивості, виражені в шкалі більш високого рангу, містять всю інформацію про ці ж властивості, виражених в шкалах більш низького рангу - але не навпаки.

Для отримання інформації про властивості, вимірюваних в сильних шкалах, потрібні, взагалі кажучи, більш складні і дорогі вимірювальні прилади та процедури. Перевагами слабких шкал, таким чином, є їх відносна простота, а також дешевизна відповідних вимірювальних засобів. У слабких шкал є ще одна істотна перевага: вони мають більш високу заводо захищеність (ця властивість слабких шкал суттєво використовується при розробці сучасних систем обробки інформації).

4.7 . Медична діагностика як вимірювальна задача

Як зазначалося в п.4.3 , розрізняють вимірювання елементарні і комплексні .

При комплексному вимірі , головною метою якого є вибір виду математичної моделі , також вирішується і завдання конкретизації значень її параметрів (завдання елементарних вимірювань) . На рис.4.13-б представлена більш докладна, в порівнянні з рис.4.4 , схема процедури комплексних вимірювань.

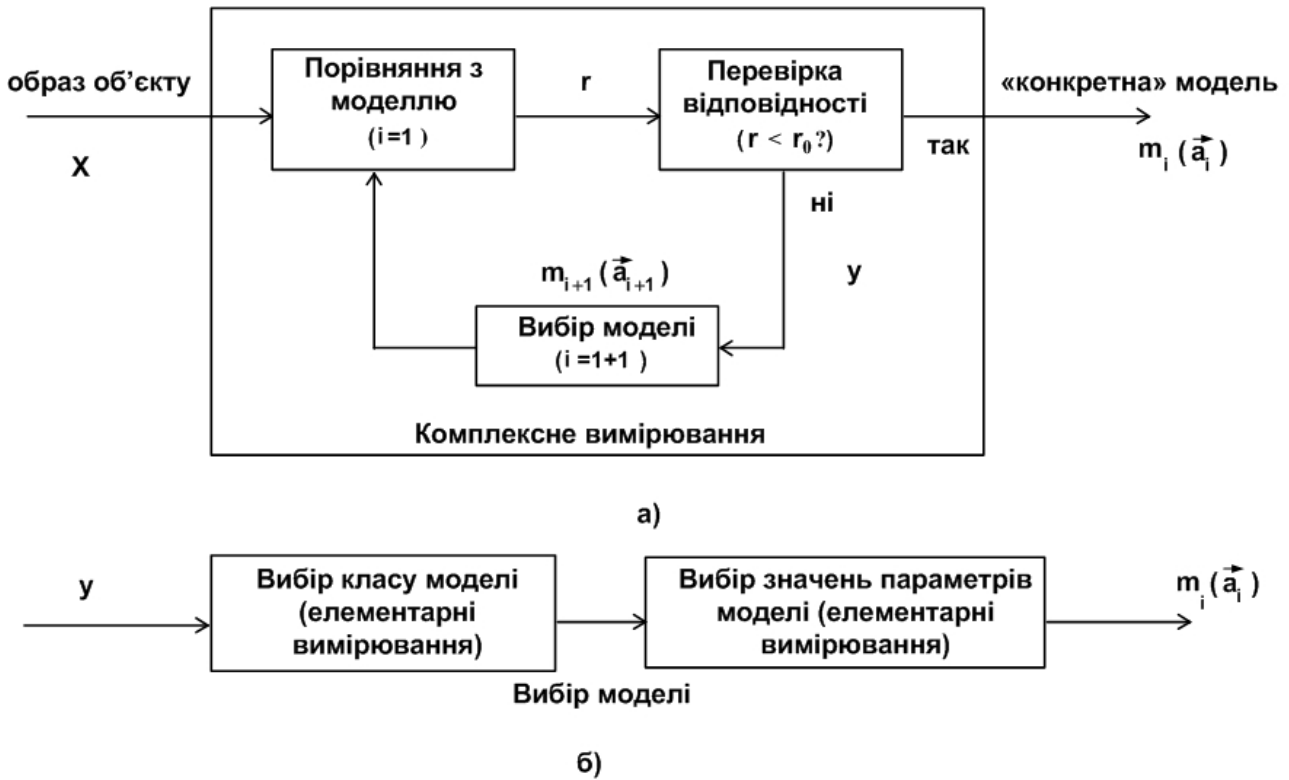


Рис. 4.13

Вибір моделі складається з двох етапів: вибір (виділення, пошук) класу моделі M_i і вибір конкретної моделі $m_i(a_i)$ в цьому класі (вибір значень параметрів a_i), як показано на рис.4.13 (вибір ініціюється керуючим впливом y). Слід зауважити, що вектор a_i не обов'язково тотожно дорівнює вектору X - це може бути інший, можливо, перетинаючий з X , набір параметрів.

Кожен із зазначених двох етапів відноситься до елементарних вимірів - і в цьому сенсі елементарні вимірювання є складовими частинами комплексного вимірювання. Відсутність хоча б одного з цих етапів зводить комплексне вимірювання до елементарного. Так, відсутність етапу вибору класу математичної моделі (наприклад, цей етап може бути опущений при апріорі відомому класі моделі) призводить до елементарних вимірів у вигляді задачі оцінювання векторного або скалярного параметра. У медицині це вимірювання температури, артеріального тиску, зняття кривих ЕКГ, отримання зображень УЗД і т.п.

З іншого боку, за відсутності етапу вибору значень параметрів моделі отримуємо різновид елементарних вимірювань, іменовану "прийняттям рішення" - до цього різновиду відносять задачі розпізнавання образів, виявлення об'єкта (сигналу на тлі шумів), контролю систем, медичної та технічної діагностики. У математичній статистиці ці завдання відомі як завдання перевірки гіпотез.

Стосовно до показаної на рис.1.4 схемою завдання медичної діагностики вибір класу моделі M_i можна трактувати по-різному. При фіксованому рівні ієрархічної схеми процедури діагностики: на етапі попередньої діагностики це ви-

бір класу захворювання, на етапах загальної та диференціальної діагностики це вибір нозологічної форми. При розгляді всієї схеми в цілому це поетапний вибір нозологічної форми.

4.7.1 . Класифікація як розбиття тезауруса на підмножини

У термін " класифікація" різні дослідники вкладають неоднаковий зміст.

Фахівці в області теорії розпізнавання образів зазвичай ототожнюють його з терміном " розпізнавання " , іноді супроводжуючи цю дію деякими несуттєвими застереженнями типу : при класифікації досить визначити тільки номер класу i , до якого належить аналізований об'єкт , тоді як при розпізнаванні необхідно вказати найменування M_i відповідного класу (наприклад , " літера А " при розпізнаванні тексту).

Метрологи в термін " класифікація" вкладають дещо інший сенс : класифікація (або кінцеве розбиття) - це кінцеве сімейство $\{ M_1 , \dots , M_I \}$ непорожніх попарно непересічних підмножин (іменованих " класами ") тезауруса M . Як випливає з наведеного визначення , класифікацією іменують кінцевий результат розбиття тезауруса M на підмножини $\{ M_1 , \dots , M_I \}$. При цьому , мабуть , терміном " класифікація" природно визначити і сам процес такого розбиття .

Після того як кінцеве розбиття зроблено , може початися новий процес : вибір з сімейства $\{ M_1 , \dots , M_I \}$ єдиної підмножини M_i , елементом якої є модель m , що володіє i - тою характеристичною властивістю ($i = 1 , \dots , I$) . Такий процес називають виділенням (або пошуком) класу. Схематично виділення класу відбувається як показано на рис. 4.14 .

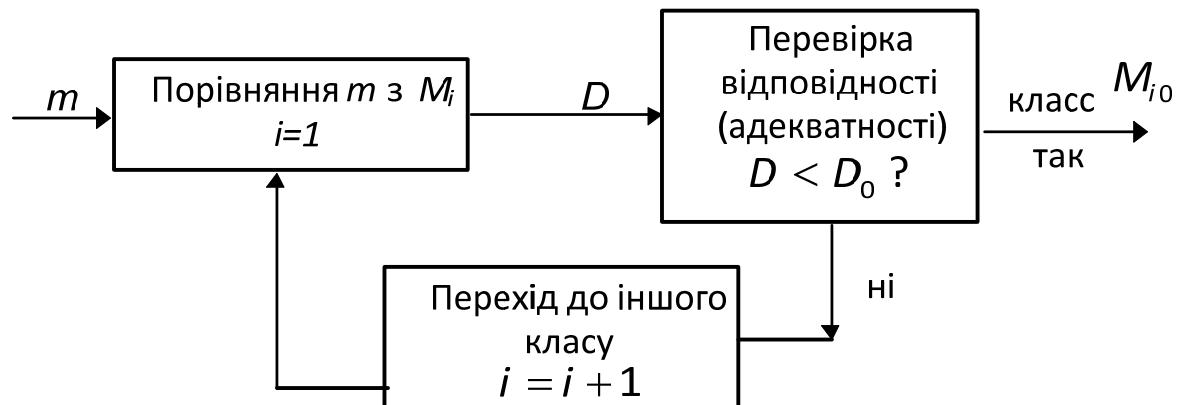


Рис. 4.14

Аналітично цей процес виглядає як вирішення завдання мінімізації функціонала D , що характеризує ступінь відповідності (адекватності) моделі m множини M_i :

$$M_{i_0} = \operatorname{argmin} D(m, M_i) . \quad (2.3)$$

Зовнішня схожість схем 4.13 і 4.14, мабуть, є однією з причин ототожнення завдань класифікації та розпізнавання образів. В обох випадках завдання вирішується шляхом мінімізації функціоналу, що характеризує ступінь відмінності між якимись "вихідними даними" і класом моделей. Однак при виділенні класу роль вихідних даних виконує одна з моделей m , тоді як при розпізнаванні образів вихідними даними є образ X якогось об'єкта.

Таким чином, з метрологічної точки зору, терміни "класифікація" і "розпізнавання" образів аж ніяк не тотожні.

Термін "класифікація" характеризує стан впорядкованості моделей, що входять в тезаурус M (в результаті "хорошої" впорядкованості пошук потрібної моделі відбувається швидко) або процес створення такої впорядкованості

Надалі, вживаючи термін «класифікація», ми будемо дотримуватися саме такого його тлумачення.

4.7.2 . Організація класів в задачі діагностики

Завдання діагностики, з точки зору теорії розпізнавання образів, полягає в знаходженні деякого класу (у медичній діагностиці це знаходження імені або коду захворювання або групи захворювань) серед безлічі інших класів, об'єкти яких відрізняються від об'єктів шуканого класу значеннями ознак (симптомів).

Очевидно, що швидкість пошуку потрібного класу залежить від того, як організовані класи.

Розглянемо в цьому зв'язку два основних способи організації - організацію "списком" і "деревом".

Формально організація списком може бути представлена як "просте" перерахування класів: $\{ M_i \}$, $i = 1, \dots, I$. При цьому класи нумеруються або їм присвоюються символічні імена (номерами звичайно є цілі позитивні числа в деякому діапазоні значень).

Організація деревом встановлює відносини панування-підпорядкування між класами. Добре відомим прикладом організації деревом є матеріал у книгах, організований так, що класи-сторінки підпорядковуються класам - підрозділам, ті підпорядковуються класам-розділам, які в свою чергу підпорядковані класам-главам.

Для нумерації класів, організованих деревом, використовують не одне число, а код у вигляді сукупності чисел (k, l, \dots, p) , позиції яких відповідають рівням ієрархії: наприклад, код 2.1.3 може бути номером третього підрозділу першого розділу другої глави книги.

Процедура пошуку потрібного класу при організації деревом проводиться поетапно: спочатку по компоненту k (верхній рівень ієрархічної драбини), потім по наступного компоненту (нижчий рівень) і т.д., завершуючись компонентом p (найнижчий рівень), як показано на рис. 4.15.

Цим вона відрізняється від одноетапної процедури пошуку при організації списком, що зводиться до перебору всіх наявних класів.

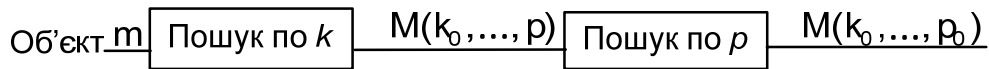


Рис. 4.15

Легко перевірити, що організація деревом дозволяє істотно заощадити час пошуку потрібного класу, причому ступінь економії швидко зростає із збільшенням кількості класів I - в цьому сенсі організація деревом має переваги над організацією списком. Ступінь виграшу в швидкості тим вище, чим з більшою кількістю класів потрібно мати справу: для "бінарного" дерева пошуку, кожному вузлу якого підпорядковані два вузли нижнього рівня (за винятком самого нижнього рівня), виграш пропорційний величині

$$B = I / \log_2 I$$

Стосовно до завдання медичної діагностики прикладами ієрархічної організації класів є структура міжнародного класифікатора хвороб (МКБ-9, МКБ-10), а також наведена на рис.1.4 узагальнена структура процедури діагностики.

4.7.3 . Позначення класів

Кожен клас з сімейства $\{ M_i \}$, $i = 1, \dots, I$, має номер i та найменування M_i , де i - число або код, M_i - символ, слово, фраза або їх комбінація. Який спосіб використовувати для позначення класів з метою їх формального розрізнення - питання зручності. Людині, як правило, зручніше працювати з найменуваннями, комп'ютеру - з номерами.

Простота організації класів списком веде і до простоти позначень. Для наведеного вище прикладу формальною відмінною ознакою класів - сторінок є їх номери (наприклад, "135") або найменування у вигляді слова і числа ("сторінка 135").

Складніше йде справа з позначеннями класів при організації деревом: нумерації класів в цьому випадку зазвичай надають вид позиційного коду. Форма найменування класів при цьому може бути різною для різних рівнів ієрархічної схеми. Так, найменування частин книги зазвичай складаються з поєднання слова "частина", номера частини римськими цифрами, потім ставиться крапка і приводиться назва частини, в загальних рисах розкриває її зміст (приклад: "Частина II. Додаток загальної теорії до окремих випадків"). Аналогічно іменуються глави книги - з тією лише відмінністю, що нумерація проводиться арабськими цифрами. Той же принцип використовується для найменування розділів і підрозділів, однак і тут можуть бути свої відмінності (у науково-технічній літературі ці номери, як правило, подаються у вигляді числового позиційного коду, що вказує номери глави, розділу і підрозділу).

Інший приклад - адресація файлів на магнітному диску: адреса $d:\text{top}\backslash\text{next}\backslash\text{text1}$ означає, що файл по імені text1 знаходиться в підрозділі next

розділу top логічного диска d.

Ще один корисний приклад - доменна структура імен в комп'ютерних мережах, зокрема, в Internet. Так, ім'я

alex@phbme.polytech.kiev.ua

підпорядковане наступній ієрархічній організації:

- країна (ua - код України)
- місто (kiev - Київ)
- організація (polytech - Національний технічний університет України (КПІ))
- структурний підрозділ (phbme - кафедра фізичної та біомедичної електроніки)
- конкретний користувач (alex - ім'я користувача).

4.7.4 . Ознаки і параметри

Ознака - поняття в задачі діагностики (розпізнавання образів) настільки ж фундаментальне, як і поняття "клас" .

У метрології ознакою іменують якісну характеристичну властивість, що дозволяє розрізнити класи $\{ M_i \}$, $i = 1 , \dots , I$, між собою. До якісних при цьому відносять характеристичні властивості, описані нечисловою шкалою лише одного типу - шкалою найменувань (з вербальними градаціям типу "так" , "ні" , "червоний" , "синій" , "дитина" , "дорослий" і т.п .).

Поряд з поняттям "ознака" в метрології широко використовують поняття "параметр" - кількісна характеристична властивість у вигляді вектора, описуваного сукупністю числових метричних шкал, що дозволяє розрізнити між собою моделі, що входять в клас M_i . Тому коли говорять про " конкретної моделі " , мають на увазі , що для деякого фіксованого значення $i = i_0$ необхідно здійснити пошук ще й значення вектора параметрів $a_i = a_{i_0}$ - тоді опис моделі буде повним в тому сенсі, що буде містити значення як якісних, так і кількісних характеристичних властивостей .

Коментуючи наведений вище поділ характеристичних ознак на якісні та кількісні, відзначимо його відому умовність. По-перше, прийняте в математиці визначення параметра як "величина , значення якої служить для розрізнення елементів деякої множини між собою" досить широко, дозволяючи приймати параметру як числові значення (таким чином вираз "числовий параметр" не є тавтологією), так і нечислові.

По-друге, в теорії розпізнавання образів настільки ж широко трактують поняття "ознака", маючи на увазі під ним якусь числову або нечислову величину, використовувану для розпізнавання елементів деякої множини".

Виходячи зі сказаного, в принципі, можна було б і не робити особливої різниці між поняттями "ознака" і "параметр". Однак іноді подібний поділ зручно. Наприклад, при розробці ОЕС "Універсал" (див. Розділ 9) виявилось зручним поняттям "ознака" і "параметр" додати наступний зміст:

- ознака - елемент опису об'єктів, що дозволяє розрізнити їх між собою;

• параметр - варіант вимірювань ознаки, вибраний з урахуванням деяких обмежувальних умов.

Наприклад, у медичній діагностиці при описі деякого захворювання такими обмежувальними умовами є стать, вік, професія та ін. У цьому випадку однією оцінюваною ознакою ("головний біль", наприклад) будуть відповідати різні параметри, тобто варіанти "головний біль у чоловіка" і "головний біль у жінки". Ще одним прикладом обмежувальних умов є використання для вимірювань ознаки того чи іншого методу, способу, приладу (пряме чи непряме вимірювання артеріального тиску, мембранний або безмембранний наконечник фонендоскопа і т.п.).

Як і класи, ознаки можна організувати або списком, або у вигляді ієрархічної структури. Перший спосіб зручний для невеликої кількості ознак (порядку 10). Якщо ознак багато (на практиці-більше двох десятків), доцільна їх ієрархічна організація, що дозволяє значно швидше здійснити пошук потрібної ознаки.

4.7.5. Класи і ознаки - відносність понять

Поняття "ознака" і "клас" різнополюсні і співвідносяться приблизно так, як співвідносяться вихідні дані і результат рішення деякої задачі.

Однак дане твердження справедливе лише у відносній ступені: п'ята колонка табл. 4.3 є прикладом ситуації, коли одна і та ж величина може трактуватися і як ознака, і як перелік класів.

Подібна ситуація досить типова для медичної діагностики: в табл. 2.7, що представляє фрагмент діагностичної таблиці для диференціальної діагностики гострих порушень мозкового кровообігу, в якості ознак фігурують артеріальна гіпертонія, інфаркт міокарда, ревмокардит - однак очевидно, що в задачі діагностики патологій серцево-судинної системи це вже діагнози (класи захворювань).

Таким чином, питання про те, з чим ми маємо справу, з ознакою або з класом - це питання про вид розглянутої діагностичної задачі (рис.4.16).

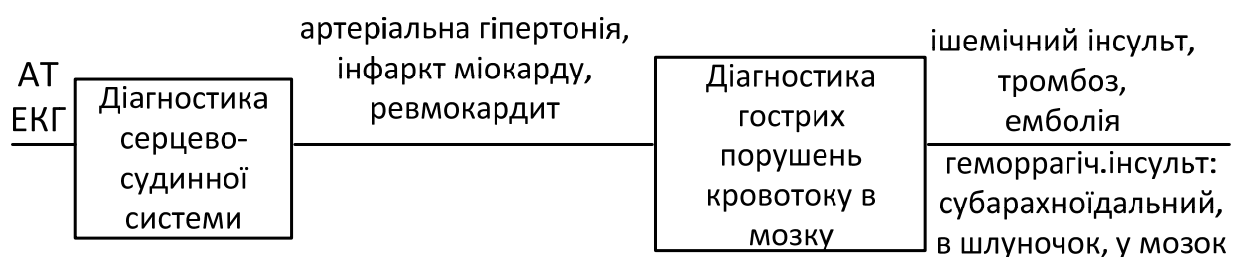


Рис. 4.16

Нарешті, при ієрархічній організації інформації одна і та ж величина формально може трактуватися і як ознака, і як клас.

На рис.4.17 наведено відповідний приклад: характеристика "колір, відтінок обличчя" може розглядатися як ознака по відношенню до характеристики "ар-

теріальний тиск", і разом з тим - як список класів по відношенню до величини "довжина хвилі".

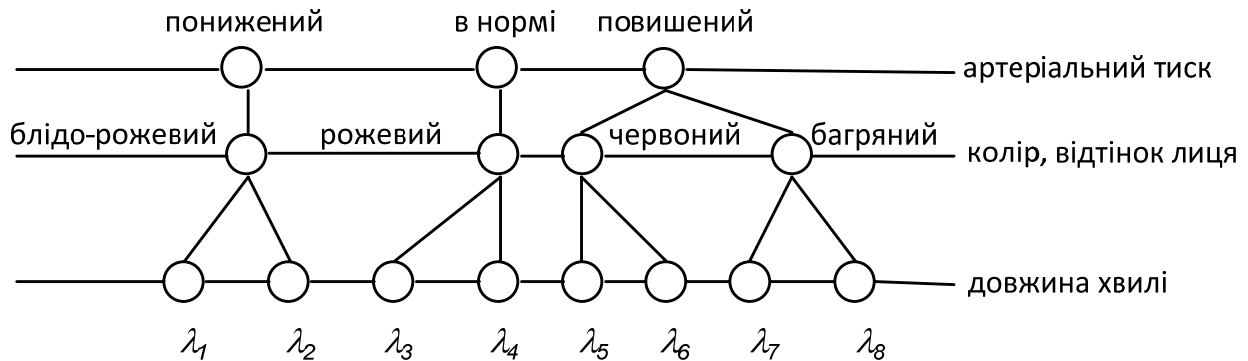


Рис. 4.17

Шкала класів "артеріальний тиск" може в свою чергу стати шкалою значень ознаки, якщо її підпорядкувати, наприклад, шкалою станів зі значеннями «артеріальна гіпертонія», «інфаркт міокарда» і т.п. (див. рис.4.16).

Таким чином, потенційна можливість двоякого трактування величин, що утворюють діагностичний ланцюжок (ланцюжок правил продукції «ЯКЩО-ТО»), впливає з тієї обставини, що формально визначення класів і ознак збігаються: в обох випадках мова йде про величини, що використовуються для розрізнення елементів деяких множин. По суті ж джерелом двозначності є та обставина, що одні й ті ж дані є результатом розв'язання однієї задачі (більш низького рівня) і вихідними даними іншого завдання (більш високого рівня).

Короткі висновки

У теорії вимірювань розрізняють два рівня вивчення вимірів: "фундаментальний" і "прикладний".

У фундаментальній теорії вимірювань прийнято виділяти два основних напрямки розвитку: фізичний і математичний.

Двом основним напрямкам фундаментальної теорії вимірювань, - фізичному та математичному, - відповідають два основних напрямки прикладної теорії вимірювань: метрологія та автометрія.

Розрізняють фізичні та математичні моделі, що використовуються при вимірюваннях.

Математична модель - це сукупність абстрактних математичних об'єктів і математичних відношень, що замінює об'єкт реального світу. Фізична модель - це матеріалізована математична модель досліджуваного об'єкту.

Розрізняють вимірювання елементарні і комплексні. Елементарне вимірювання являє собою процедуру порівняння значення оцінюваної характеристики з набором еталонних значень, в результаті чого вибирається єдине еталонне значення, яке мінімально відрізняється від значення вимірюваної характеристи-

ки. Комплексне вимірювання головною своєю метою переслідує вибір виду математичної моделі, попутно вирішуючи і завдання конкретизації значень її параметрів.

У медичних вимірюваннях, вироблених із застосуванням фізичних вимірювальних приладів (при інструментальній діагностиці), беруть участь чотири об'єкти:

- 1) вимірювальний прилад;
- 2) пацієнт;
- 3) лікар;
- 4) умови, при яких проводяться вимірювання.

Кожен з цих об'єктів є джерелом похибок вимірювань. У ряді випадків (при суб'єктивній діагностиці) роль вимірювального приладу виконує лікар, використовуючи для оцінки стану пацієнта свої органи чуття.

При вимірах характеристик як фізичних, так і нефізичних об'єктів (у медицині та психології нефізичними, в сенсі відсутності їх фізичних еталонів, є, наприклад, такі об'єкти як здоров'я, інтелект, характер тощо) велике значення має поняття шкали.

Розрізняють, головним чином, чотири види шкал:

- 1) найменувань (номінальна, номінативна);
- 2) порядку (порядкова, ординальна);
- 3) інтервалів (інтервальна);
- 4) відношень (масштабна).

Іноді до них додають п'яту шкалу: абсолютну (натуральну).

Шкали найменувань і порядку відносяться до емпіричних (неметричних) шкал і характеризуються одиницями вимірювань, що не входять в систему одиниць СІ. Шкали інтервалів, відношень і абсолютна є метричними - одиниці вимірювань за цими шкалами входять до складу системи одиниць СІ.

Прийнято розрізняти такі форми подання шкал:

- 1) вербальна;
- 2) графічна;
- 3) числова;
- 4) таблична.

Чим вище рівень шкали, тим більше статистичних і математичних дій можна виконати над отриманими при вимірюваннях числами. Так, шкала відношень є самою "сильною", а шкала найменувань - самою "слабкою". Для отримання інформації про властивості, вимірюваних в сильних шкалах, потрібні більш складні і дорогі вимірювальні прилади та процедури. Достоїнствами слабких шкал є їх відносна простота, а також дешевизна відповідних вимірювальних засобів. У слабких шкал є ще одна істотна перевага: вони мають більш високу перешкодо-захищеність.

Завдання діагностики, з точки зору теорії розпізнавання образів, полягає в знаходженні деякого класу (імені або коду захворювання або групи захворювань) серед безлічі інших класів, об'єкти яких відрізняються від об'єктів шуканого класу значеннями ознак (симптомів). Швидкість пошуку потрібного класу залежить

від того, як організовані класи - "списком" і "деревом". Організація деревом краща, оскільки дозволяє істотно заощадити час пошуку потрібного класу, причому ступінь економії швидко зростає із збільшенням кількості класів.

Поняття "ознака" і "клас" співвідносяться приблизно так, як співвідносяться вихідні дані і результат рішення деякої задачі. Разом з тим, при ієрархічній організації інформації одна і та ж величина формально може трактуватися і як ознака, і як клас.

Література до Розділу 4

1. Розенберг В.Я. Введение в теорию точности измерительных систем. – М., "Сов.радио", 1975. - 302с.
2. Стахов А.П. Введение в алгоритмическую теорию измерения. М., "Сов.радио", 1977. - 286с.
3. Пиотровский Я. Теория измерений для инженеров. М., "Мир", 1989, 333с.
4. Предприниматель: экономико-психологический профиль. – Психологический журнал, т.13, №3, 1992. - с.42-53.
5. Щекин Г.В. Практическая психология менеджмента. – К., «Україна», 1994. - с.152-169.
6. Иванова Е.М. Аналитическая профессиограмма как средство обеспечения профессиональной диагностики кадров. – Вестник Московского университета, сер.14, Психология, №3, 1989. - с.13-20.
7. Державний класифікатор України. Класифікатор професій. ДК 003-95. – Держстандарт України, К., 1995. - 150 с.
8. АРМ психолога. Руководство пользователя, – К., 1996. - 23 с.
9. Человек и бизнес. Путь совершенства. – М., БАРС, 1995. - 228 с.
10. Щекин Г.В. Визуальная психодиагностика и ее методы. – Всесоюзный заочный университет управления персоналом, чч.І,ІІ, К., 1992. - 365 с.
11. Математический энциклопедический словарь. – М., «Советская энциклопедия», 1988. - 845 с.
12. И.Березняков. Как измерять артериальное давление. "Все о медицине", #12, март'97.
13. Суппес П.,Зинес Д. Основы теории измерений//Психологические измерения. М., "Мир", 1967.
14. ГОСТ 16263-70 "Государственная система обеспечения единства измерений. Метрология. Основные термины и определения. М., 1970
15. Cambell N.R. Measurement and Calculation. London, Longmans, Green & Co., 1928
16. Яворский Я. Математические модели измерительных процедур и систем. Варшава, 1977.
17. Арутюнов П.А. Теория и применение алгоритмических измерений. М., Энергоатомиздат, 1990. - 256с.

18. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психологической диагностике. – К., Наукова думка, 1989. - 197с.

Контрольні запитання до Розділу 4

1. Що являють собою фізичні та математичні моделі, використовувани при вимірах?
2. Які чотири об'єкти беруть участь в медичних вимірах, вироблених із застосуванням фізичних вимірювальних приладів, тобто при інструментальній діагностиці?
3. Що таке суб'єктивна діагностика і яким чином проводяться вимірювання при суб'єктивній діагностиці?
4. Що таке вимірювальна шкала і навіщо вона потрібна?
5. Які види вимірювальних шкал вам відомі?
6. Які математичні операції допустимі на шкалі відношень?
7. Які математичні операції допустимі на шкалі порядку?
8. Які математичні операції допустимі на шкалі найменувань?
9. Наведіть приклади використання при медичних вимірах різних видів вимірювальних шкал.
10. Що таке «сила шкали»? Які недоліки і переваги «слабких» шкал?
11. Наведіть приклади, коли один і той же об'єкт є класом і ознакою.

Частина 2 Практична реалізація експертних систем

Розділ 5 Засоби побудови ЕС

Закономірність питання про узгодження властивостей засобів розробки з властивостями предмета розробки майже очевидна, оскільки обумовлена високим ступенем його спільності, справедливістю для широкого кола проблем. У цьому зв'язку розглянемо спочатку, що собою являє ЕС як програмний продукт, і ознайомимося з існуючими різновидами ЕС (класифікація ЕС).

5.1. Особливості експертної системи як програмного продукту

Природно припускати, що вибір засобів розробки ЕС повинен відбуватися з урахуванням особливостей ЕС як програмного продукту. Тому розглянемо спочатку, в чому ж полягають ці особливості.

У літературі наводиться кілька визначень цих особливостей. По-перше, найголовніша відмінність ЕС полягає в тому, що ЕС маніпулюють знаннями, тоді як звичайні програми маніпулюють даними (див. пп.3.1, 3.2). Інші відмінності можуть бути представлені у вигляді табл. 5.1.

Таблиця 5.1

Обробка даних	Інженерія знань
Представлення і використання даних	Представлення і використання знань
Алгоритми	Евристика
Повторний прогін	Процес логічного виводу
Ефективна обробка великих баз даних	Ефективна обробка великих баз знань

Характеризуючи уявлення фахівців у галузі ШІ про властивості, якими повинна володіти ЕС, в літературі наводиться схема, показана на рис.5.1.

Прокоментуємо рис.5.1. Під компетентністю мається на увазі, що ЕС повинна досягати в конкретній предметній області того ж рівня професіоналізму, що й експерти-люди, повинна бути вмілою в тому сенсі, що повинна отримувати правильне рішення швидко, уникаючи непотрібних дій, а також знати предмет не тільки «в глибину», але і «в ширину» (володіти робастністю). Остання властивість найбільш важко піддається реалізації, але саме їм успішно користуються експерти-люди.

Під здатністю до символічних міркувань розуміється здатність вести діалог з користувачем на природній мові завдяки використанню спеціальних форм

представлення знань, а також, при необхідності, переформулювати погано сформульовану задачу. І тут остання властивість належить більше до побажань, ніж до реальної ситуації - більшість існуючих в даний час ЕС не володіють цією властивістю.



Рис.5.1

Під глибиною мається на увазі здатність вирішувати не «іграшкові», гранично спрощені, завдання, а реальні складні завдання, що вимагають для свого вирішення великого і складно влаштованого простору пошуку, з безліччю можливих проміжних і остаточних рішень.

Нарешті, під самосвідомістю розуміється наявність метазнань - система повинна знати, що вона знає - це необхідно для самоперевірки отриманих рішень і для їх аргументації за допомогою якогось механізму пояснень.

Неважко бачити, що наведені вище вимоги до ЕС - це до певної міри вимоги «із запасом», «на виріст» - ці вимоги у всій повноті слід пред'являти перспективним ЕС і, якоюсь мірою, сучасним, так званим «великим» ЕС. Що стосується «малих» і «середніх» ЕС, вимоги до них можна дещо знизити:

- тип розв'язуваної задачі - прийняття рішення в умовах можливої неповноти інформації;
- наявність режиму аргументації прийнятого рішення;
- можливість поповнення та модернізації бази знань;
- можливість зміни стратегії прийняття рішення

Коментуючи ці вимоги, зазначимо, що принципова можливість прийняття

рішення в умовах неповноти інформації забезпечується наявністю неалгоритмічних знань в БЗ (див. п. 3.2).

Аргументація прийнятого рішення необхідна, щоб:

- виключити недовіру до отриманого комп'ютером вирішенню;
- виключити надмірну довіру до отриманого комп'ютером рішенням.

Поповнення і модернізація БЗ - реалізація властивості "активність" БЗ, спрямованого на прогресивний саморозвиток системи, у вузькому сенсі розуміється як усунення внутрішніх суперечностей і неповноти БЗ.

Зміна стратегії прийняття рішення можливе за досить розвиненому МР, що містить різні механізми виведення, і може переслідувати наступні цілі:

- проведення наукових досліджень з метою вибору найбільш ефективного механізму виведення;
- вибір найбільш ефективного механізму прийняття рішення при практичному використанні ЕС.

5.2. Класифікація ЕС

Прийнято виділяти такі три групи критеріїв класифікації ЕС:

- за характеристиками предметної області;
- за організацією спілкування з користувачем;
- за реалізацією внутрішньої структури.

Крім того, класифікація можлива за реалізацією зовнішньої структури (програмної архітектури), за типом використовуваних комп'ютерів (часто розробляють ЕС на спеціальній платформі або, як ще кажуть, для робочих станцій, маючи на увазі особливості апаратної архітектури), з використовуваних операційних систем і, нарешті, за вартістю.

Класифікація за характеристиками предметної області.

Відповідна класифікаційна схема наведена на рис.5.2.

Основні типи розв'язуваних завдань були розглянуті в п.1.10 (табл. 1.3). До них відносяться ЕС, вирішальні завдання: інтерпретації даних; діагностики; прогнозування; проектування; планування; моніторингу; налагодження; навчання; управління.

ЕС, що виконують інтерпретацію, як правило, використовують інформацію від датчиків для опису ситуації - прикладом може служити система медичного моніторингу за станом хворого. Специфіка умов роботи ЕС такого типу - можливий низький рівень надійності даних (зашумленість, нестабільність через поганий контакт з тілом пацієнта і т.п.). У силу безперервності потоку даних і значних часових проміжків спостереження за пацієнтом в таких системах необхідне застосування спеціальних методів стиснення, обробки і реєстрації даних і результатів їх інтерпретації.

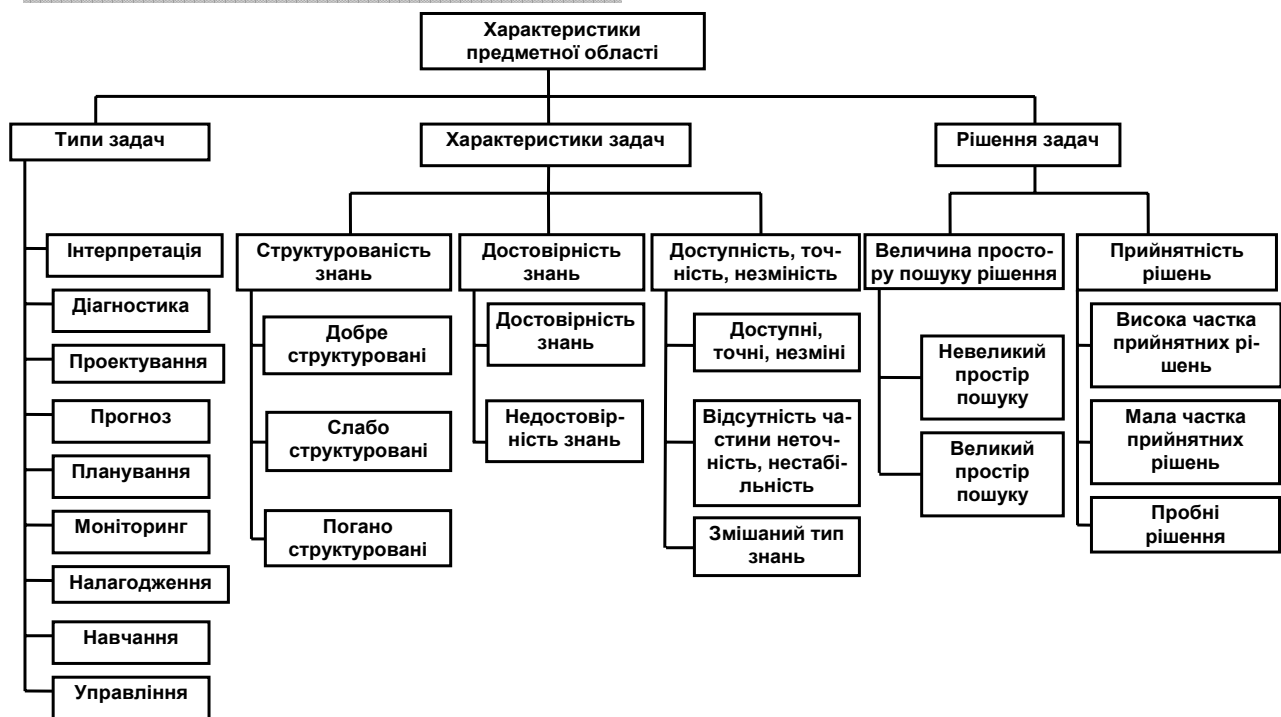


Рис. 5.2

Застосування діагностичних ЕС цілком природно в медицині - у цій області розроблено більше діагностичних систем, ніж у будь-якій іншій.

Природним виглядає також застосування прогностичних медичних ЕС: завжди важливо знати, наскільки "правильно" поводить ся організм після проведення лікувальних заходів або яке найбільш ймовірне його стан в майбутньому, в найближчому (дні, тижні) або віддаленому (місяці, роки).

Проектування - завдання більш віддалене від потреб практикуючих лікарів, вона може вирішуватися, наприклад, при розробці нового медичного обладнання.

Завдання планування є типовою для будь-якої системи управління, в тому числі - і для медичних адміністративно-управлінських структур. Вельми важливим є, наприклад, правильний розрахунок лікарських препаратів, необхідних аптекам міста, району, області. Планування підлягають і конкретні лікувальні заходи, та науково-дослідні роботи.

Моніторингові системи забезпечують спостереження поведінки об'єктів - зазвичай це системи приліжкового моніторингу.

Системи налагодження, призначені для видачі рекомендацій щодо усунення поганого функціонування, зазвичай містять елементи, властиві системам прогнозування, планування, управління - в цьому сенсі їх можна розглядати як комбіновані системи.

Навчальні системи застосовуються для передачі «законсервованих» знань малокваліфікованим користувачам - при цьому фактично вирішується завдання, обернене завданню отримання знань від висококваліфікованих фахівців - експертів.

Управляючі ЕС є найбільш складними серед перерахованих, оскільки, щоб ефективно управляти, система повинна вирішувати безліч перерахованих вище завдань - завдання інтерпретації, прогнозування, діагностики, планування і контролю.

Коментуючи класифікацію за характеристиками вирішуваних завдань, зауважимо, що недостатню достовірність знань і даних можна враховувати по-різному. Іноді для цієї мети вводять систему чисел (ваг, коефіцієнтів впевненості, ступеня вираженості і т.п.), які можуть бути пов'язані з імовірнісними характеристиками, використовуваними в механізмі логічного висновку (зокрема, в системі «ЭСИД» для цієї мети використовується числовий коефіцієнт $0 \leq r \leq 5$ - див п.8.1.1). Як добре теоретично обґрунтований підхід до міркувань з використанням недостовірних даних можна розглядати нечітку логіку Заде.

Що ж до класифікації за характеристиками одержуваних рішень, відзначимо особливо, що при невеликих просторах пошуку рішення (при роботі з БД - це невелика кількість БД з невеликою кількістю записів) можливе використання простого перебору в пошуках потрібних рішень. Однак для великих просторів пошуку доводиться вдаватися до ієрархічної організації просторів.

Класифікація за організації спілкування з користувачем.

Класифікаційна схема наведена на рис.5.3.



Рис. 5.3

Торкаючись видів спілкування, відзначимо, по-перше, що серед ЕС переважають системи консультаційного типу - зокрема, медичні діагностичні системи

відносяться саме до цього класу. Системи для придбання знань виділені умовно в тому сенсі, що рідко будуються ЕС для проведення досліджень механізмів придбання знань у експертів - значно частіше розробляють такі ЕС, де придбання знань є лише однією з безлічі необхідних функцій системи.

Значна різноманітність, як впливає з схеми на рис.5.3, спостерігається серед мов спілкування. Найчастіше в системах реалізується не одна, а кілька перерахованих можливостей.

Важливим елементом ЕС є підсистема пояснень. Установка ступеня допомоги передбачає надання користувачеві можливості вибору ступеня подробности одержуваних пояснень. Ретроспективні пояснення іменують також трасуванням - сенс таких пояснень у простежуванні історії виведення рішення. Вбудовані пояснення зазвичай реалізуються у формі контекстної допомоги - текстів, «прив'язаних» до кожної ситуації, і спричинених при необхідності. Гіпотетичні пояснення являють собою відповіді на питання «що, якщо ...» і служать для аналізу можливих наслідків у припущенні, що деякий факт мав би місце або деякий правило змінилося б.

Класифікація за реалізації внутрішньої структури. Схема такої класифікації представлена на рис.5.4.

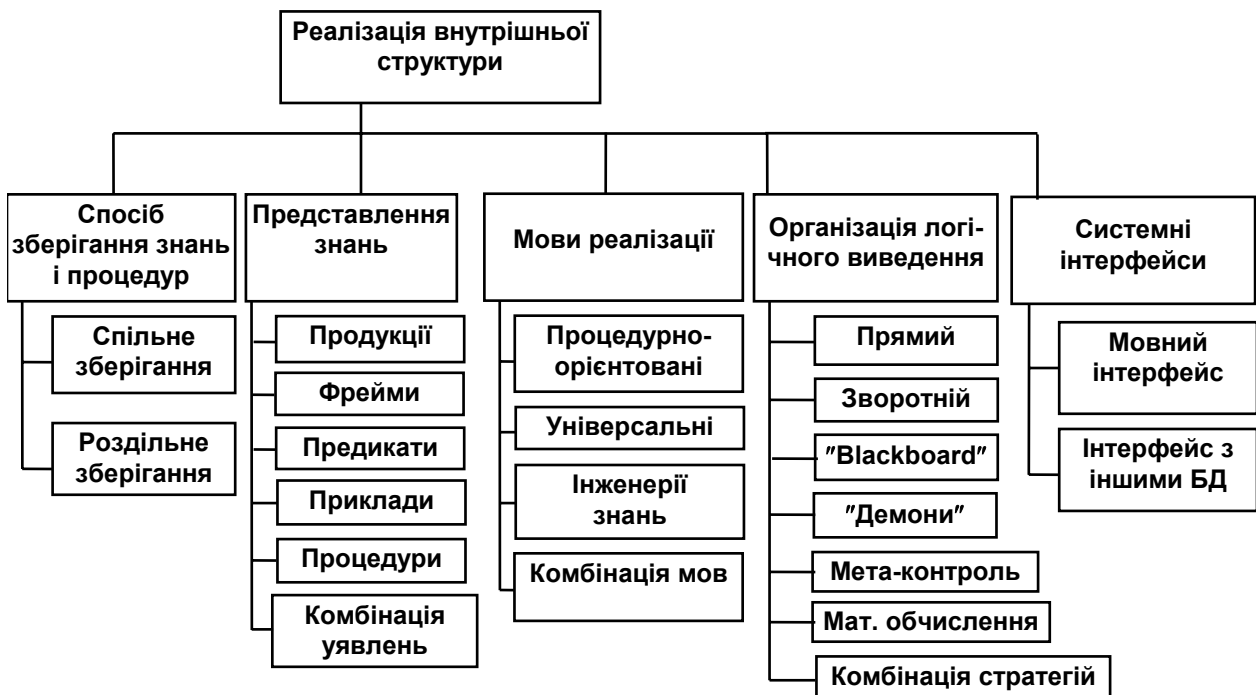


Рис. 5.4

З приводу форм організації логічного висновку відзначимо, що blackboard-механізм (механізм «дошки оголошень») являє собою спосіб представлення та управління знаннями, заснований на використанні незалежних груп правил, званих джерелами знань, які «спілкуються» через центральну БД, яку називають дошкою оголошень. «Демонами» іменують процедури, які "включаються" в ро-

боту при дотриманні певних умов. Прикладом мета-контролю є установка пріоритетів правил. Математичними обчисленнями іменують, наприклад, реалізацію процедур логічного висновку з використанням Байєсівської стратегії або її елементів, підсумовування балів відповідно до табличними алгоритмами та ін.

Програмна архітектура. За даним критерієм будемо розрізняти такі різновиди експертних систем:

- 1) спеціалізовані ЕС;
- 2) оболонки ЕС (порожні ЕС);
- 3) розподілені ЕС (мережеві ЕС);
- 4) гібридні ЕС (поєднання моделей прийняття рішення);
- 5) узагальнені ЕС (розподілені ЕС з сильним обчислювальним компонентом).

Різниця між спеціалізованими ЕС і оболонкам ЕС демонструється за допомогою табл. 5.2.

Таблиця 5.2

Спеціалізовані ЕС	Оболонки ЕС
Рішення задач однієї проблемної області	Рішення задач деякого безлічі проблемних областей
Глибина опрацювання методів вирішення завдань	Менша глибина опрацювання методів вирішення завдань
Складність перенастроювання на вирішення завдань іншої проблемної області	Легкість перенастроювання на вирішення завдань іншої проблемної області

Розподілені (мережеві) ЕС призначені для застосування в ситуаціях, коли для вирішення завдання необхідно одночасна участь кількох ОПР - прикладом такої ситуації в медицині може служити комплексне діагностування, при якому кожен фахівець-діагност має власний АРМ, результати приватних діагностичних обстежень за допомогою комп'ютерної мережі об'єднуються і аналізуються для прийняття остаточного рішення (див. п. 7.2.1).

Гібридні ЕС відрізняються від традиційних високим рівнем розвиненості БАЗ, а узагальнені - розвиненістю розрахунково-обчислювальної частини.

5.3. Класифікація інструментальних засобів побудови ЕС

Засоби побудови експертних систем - це системи програмування, які спрощують проектування експертних систем. Діапазон можливих засобів простягається від мов високого рівня до засобів підтримки низького рівня. У літературі інструментальні засоби побудови експертних систем поділяють на чотири категорії:

- мови програмування;
- мови інженерії знань;
- допоміжні засоби;
- засоби підтримки.

Охарактеризуємо коротко перераховані засоби.

Мови програмування, що застосовуються для розробки експертних систем, - це, як правило, або проблемно-орієнтовані мови, такі як Фортран або Паскаль, або мови обробки текстів, такі як Лісп або Пролог. Втім, в літературі можна зустріти твердження, що справа не стільки в мові програмування, - ним може бути чи не будь-яку мову, - скільки в тому, наскільки розв'язувана задача придатна для вирішення її засобами експертних систем.

Мови інженерії знань - мови більш високого рівня, що представляють собою якусь середу у вигляді однієї або декількох програм для розробки ЕС. Фактично це якісь СУБЗ, функції яких визначені специфікою розв'язуваних завдань.

Допоміжні засоби - це програми для підтримки процесу придбання знань у експерта-людини і програми, що допомагають розробити проекти ЕС.

Засоби підтримки - це засоби, що полегшують програмування і мають вигляд налагоджувальних засобів, редакторів баз знань, засобів швидкої розробки механізмів введення / виводу і пояснення.

Відома також дещо інша класифікаційна схема інструментальних засобів побудови ЕС:

- мови програмування;
- інструментальні системи.

При цьому до інструментальним систем відносять:

- «порожні» ЕС;
- оболонки (shells) ЕС;
- інструментальні ЕС.

Дана класифікаційна схема істотно перекривається з наведеної вище: мови програмування фігурують в обох схемах, а поняття мов інженерії знань еквівалентно поняттю інструментальних систем.

5.3.1. Мови програмування для додатків в області ЕС

Як відмічається багатьма розробниками, вибір мови програмування для розробки ЕС - це не тільки питання про найкраще співвідношення форми і змісту, а й питання про те, якою мовою програмування вже володіють програмісти, чії послуги доступні розробникам ЕС.

Проблемно-орієнтовані мови - мови, розроблені для певного класу задач. Наприклад, Фортран і Бейсік досить зручні для вирішення обчислювальних завдань. Системи управління базами даних (СКБД) типу Clipper, Paradox, Oracle, Delphi і ін зручні при роботі з даними, організованими у вигляді таблиць (реляційні СУБД), - зазвичай це інформація довідкового характеру.

Прикладами застосування мови програмування Clipper є експертна система для іридіодіагностики ЭСИД (див. п.8.1.1), оболонка ЕС "Універсал" (п.9.1), оболонки інформаційно-довідкових систем МЕД, МЕГ (п.9.2), інформаційно-довідкова система для гомеопатії REPER (п.8.2). Такий вибір мови програмування обумовлений, по-перше, можливістю розгляду функції пошуку інформа-

ції як однієї з найважливіших функцій простих консультаційних експертних систем продукційного типу і тісно пов'язаних з ними інформаційно-довідкових систем. Правило продукції "якщо А, тоді В" при цьому може бути витлумачено так: необхідно знайти інформацію про подію В на підставі інформації про подію А. Суть справи не змінюється навіть у тому випадку, якщо інформація про зв'язок між А і В має імовірнісну форму: наприклад, відома умовна ймовірність $P(A|B)$ настання події А при настанні події В - такого типу зв'язок використовується в системах ЭСИД і "Универсал".

Значна кількість експертних систем створюється з використанням мови Сі - цей високоєфективний засіб застосовано, зокрема, при розробці системи ЭСИД-2 (див. п.8.1.2). Ряд фірм-розробників з метою підвищення ефективності роботи своїх ЕС, що добре зарекомендували себе на ринку, все частіше переводять комерційні ЕС на більш ефективні (з точки зору збільшення кола використовуваних комп'ютерів, швидкодії і економії пам'яті) мови програмування - і найчастіше такою мовою виявляється Сі .

Мови обробки текстів розроблені для прикладних областей штучного інтелекту - до таких мов відносяться Лісп і Пролог.

Лісп. Дамо загальне уявлення про можливості та особливості мови Лісп.

У завданнях ШІ можливість маніпулювання з символами настільки важлива, що вже в 1956р. Ньюелл, Шоу і Саймон розробили мову програмування ІПЛ для обробки списків, який і з'явився попередником мови Лісп, розробленого Дж. Маккарті у 1960р.

Серед мов програмування Лісп займає особливе місце. Найбільш популярним він став у середовищі фахівців з ШІ в США, де в якості стандартного отримав поширення його діалект COMMONLISP. У нашій країні з Лісп знайомий лише вузьке коло фахівців, що частково пояснюється відсутністю хороших підручників.

Вирази мови Лісп називаються також *S-виразами*. Такий S-вираз являє собою або *атом*, або *список*.

Атом представляється яким небудь алфавітно-цифровим ім'ям. Список являє собою деяку внутрішню частину, укладену в дужки.

Синтаксис мови Лісп у формі Бекса-Наура. Визначається наступним чином:

```

<S- вираз>           :=<атом>
                        :=<список>
<список>             :=< внутрішня частина >
<внутрішня частина> :=<порожньо>
                        :=<S- вираз>
                        :=<S- вираз> <внутрішня частина>
<порожньо>           :=
<атом>               := ланцюжок алфавітно-цифрових
                        символів без пробілів або
                        спеціальних символів ( ) , ; ' .
    
```

Таким чином, всякий S-вираз являє собою набір атомів і списків, наприклад:
(ИНФАРКТ_МИОКАРДА ИНСУЛЬТ) (ГАЙМОРИТ) (РИНИТ)) В.-
С._ДИСТОНІЯ)

Як і у всіх мовах програмування, деякі попередньо певні атоми є функціями, аргументи яких представляють собою наступні за ними S-вирази. У свою чергу аргумент сам може бути функцією, яку потрібно обчислити. При цьому треба мати можливість визначити, що собою представляє даний елемент - значення виразу L або ж символічне ім'я L якогось S-виразу.

У першому випадку перед вираженням ставиться апостроф або ж пишуть у розгорнутій формі:

(QUOTE L).

Для завдання виразу використовується функція SETQ:

(SETQ <атом> '<S-вираз >) → <S-вираз >,

де значення S-виразу не обчислюються.

Дві фундаментальні функції CAR та CDR дозволяють здійснювати обробку списків, якщо останні не порожні. Функція CAR та CDR визначаються таким чином:

(CAR '<не порожній список>) → перший підвираз підсписку

(CDR '<не порожній список>) → список, складений з того доповнення внутрішньої частини, яке залишилося б після застосування до вихідного списку функції CAR

У Лісп є функція CONS, дія якої протилежна дії функцій CAR та CDR:

(CONS <S-вираз ><список> → список

Функція EVAL (S) - її реалізація заснована на використанні внутрішнього представлення в Лісп, в якому дужки не застосовуються - у мові Лісп виконання якої-небудь команди (яка сама є виразом) означає просто обчислення значення відповідного вхідного вираження.

Необхідно відзначити, що в мові Лісп будь-яка програма сама по собі є S-виразом. Зокрема, в Лісп дані мають ту ж структуру, що і програми, і також є виразами. Результат виконання програми в Лісп також є програмою. Відповідно цьому S-вираз в Лісп є програмою. Ця властивість мови Лісп дозволяє створювати програми, які можуть перебудовувати самі себе.

Предикати. Приймають значення T або NIL:

(ATOM <S-вираз>) → T, якщо і тільки якщо значення S-вирази є атом або ім'я функції, інакше NIL;

(NULL < S-вираз >) → Т, якщо і тільки якщо значення S-вирази є NIL, інакше NIL;

(EQUAL < S-вираз 1> <S- вираз 2>) → Т, якщо і тільки якщо значення обох виразів рівні;

(AND < S-вираз >) зазначено деяке число аргументів, які обчислені зліва направо; якщо при цьому зустрінуто NIL, повертається значення NIL, інакше значення останнього виразу;

(OR < S-вираз >) на відміну від попереднього, якщо при перегляді зустрінуто значення, що відрізняється від NIL, воно повертається, інакше NIL;

(NOT < S-вираз >) → Т, якщо і тільки якщо значення S-вирази є NIL.

Функції. Вибір між кількома можливостями здійснюється за допомогою функції COND:

(COND (<t1> <r1>)
(<t2> <r2>)
(<tn> <rn>))

Символи t вказують на вирази, пов'язані з умовами, що перевіряється, а символи r - на вираження, пов'язані з результатами.

Функції побудови списків. Дві функції APPEND і LIST так само, як і функція CONS, формують списки.

Функція APPEND:

(APPEND <список 1> <список 2> ... <список n>) – список, утворений зі списків 1, 2, ..., n, з яких видалені дужки першого рівня.

Функція LIST:

(LIST <S-вираз 1> < S-вираз 2> ... < S-вираз n>) →
→ (S-вираз 1 S-вираз 2 ... S-вираз n)

У даному випадку, на відміну від функції APPEND, дужки залишаються.

Функція LENGTH:

(LENGTH <S-вираз>) → число «братів» першого рівня в S-виразі

Функція REVERSE:

(REVERSE < S-вираз >) → будує S-вирази з «братів» першого рівня вихідного рівня в зворотному порядку.

Функція SUBST (функція підстановки):

(SUBST < S-вираз 1> < S-вираз 2> < S-вираз 3>) → результат заміщення виразом 1 всіх входжень виразу 2 до виразу 3. При цьому вираз 2 повинен бути атомом.

Функція MEMBER:

(MEMBER < атом><список>) → NIL, якщо і тільки якщо вихідний атом не з'являється в якості «брата» першого рівня у вихідному списку.

Функція RPLACA:

(RPLACA <список> <S-вираз>) → заміщає CAR списку на заданий вираз.

Функція RPLACA – симетрична по відношенню до функції RPLACA, впливаючи на CDR свого першого аргументу.

Ітерація MAPCAR:

(MAPCAR ‘<функція> <S-вираз>) – вираз, утворений застосуванням першого аргументу (який повинен бути функцією, відомій системі) до всіх елементів заданого S-виразу

Ітерація APPLY:

(APPLY ‘<функція> <S-вираз>) – тут початкова функція застосовується до зазначеного рівняння.

Лябда - функція LAMBDA:

(LAMBDA (X) < S- вираз 1> <S- вираз 2>) → значення < S- вираз 1>, в якому всі входження X заміщені значеннями <S- вираз 2>, де X – змінна, S-вираз 1 – тіло лябда-функції, S-вираз 2 – аргумент.

Функція PUTPROP:

(PUTPROP < атом 1> <список> < атом 2>) – «Пов'язує» властивість, яка виражена ім'ям атома 2, з об'єктом, представленим атомом 1, і з конкретними значеннями властивості, представленими у вихідному списку.

Функція GET:

(GET < атом 1> < атом 2>) → значення властивості, що задається атомом 2 в точці, яка визначається атомом 1, якщо воно існує, інакше NIL.

Ця функція дозволяє дізнатися, чи володіє об'єкт даними властивостями.

Математичні (числові) функції. Це арифметичні, алгебраїчні, тригонометричні функції: додавання, віднімання, зведення в ступінь і т. п.

Приклад запису висловлювання «Гострий біль в області серця» мовою Лісп наведено нижче.

Природна мова

Лісп

Гострий біль в області серця (МІСЦЕЗНАХОДЖЕННЯ (ГОСТРИЙ БІЛЬ)
(ОБЛАСТЬ СЕРЦЯ))

Елемент МІСЦЕЗНАХОДЖЕННЯ діє як відношення, яке вказує, що його перший аргумент розміщується в його другому аргументі. Використовуючи цю ж схему, можна записати «якщо гострий біль в області серця, зверніться до кардіолога» у Лісп - подібній формі, як показано нижче:

Природна мова

Лісп

Якщо гострий біль в області серця, (ЯКЩО (МІСЦЕЗНАХОДЖЕННЯ
ГОСТРОГО БОЛЮ СЕРЦЕ)
зверніться до кардіолога ТО (ЗВЕРНІТЬСЯ ДО КАРДІОЛОГА))

Неважко бачити, що перша частина прикладу - це дані, тоді як друга - це правило.

Очевидно, уявлення інформації мовою Лісп не можна вважати «правильним» або «найкращим» - це просто один з багатьох способів програмування. Інженер знань може привести Лісп - програми в ідеальну відповідність зі своїми потребами, якщо тільки він в змозі точно уявити, в чому ці потреби складаються і який з численних способів подання інформації на Лісп найкраще цим потребам відповідає.

Надаючи максимальну гнучкість розробнику ЕС, мова Лісп ніяк не "підказує" йому, як представляти знання або як побудувати механізм доступу до бази знань. Цього недоліку позбавлені мови інженерії знань, які, у свою чергу, позбавлені тієї гнучкості, якою володіють мови програмування.

Пролог. Дамо загальне уявлення про можливості та особливості мови Пролог.

Пролог з'явився майже чотири десятиліття тому. Його теоретичною основою були ідеї логічного програмування, запропоновані Р.Ковальським і П.Хейсом. Перший інтерпретатор Прологу створений в Марселі в 1973р. А.Колмерое. Наступна версія - Единбурзька реалізація Прологу на DEC-10, виконана Д.Уорреном, висунула логічне програмування на одне з провідних місць і зробила його корисним інструментом для різних додатків. Ще більшу популярність в самих різних областях інформатики Пролог набув після того, як в 1981 р. японські вчені вибрали його в якості базової мови для проекту систем п'ятого покоління.

В даний час одним з основних додатків Прологу є побудова ЕС та інструментальних засобів створення ЕС. Крім того, Пролог використовується для вирі-

шення інших завдань у різних галузях ШІ, таких, як розуміння природної мови, машинне проектування та ін.

Обчислення предикатів як математична основа мови. В основі Прологу, що є мовою програмування логічного типу, лежить поняття відносини, що сходять до предикатним логікам. Слово «предикат» походить від лат. *proedictum* - присудок. Обчислення предикатів - розділ математичної логіки, в якому досліджуються операції з логічними висловлюваннями. У логіці предикатів під предикатом розуміється деяка властивість, логічна функція від будь-якого числа аргументів, приймаюча лише два значення - «істина» або «брехня».

Логіка предикатів є в деякому роді формалізацією людського мислення, зокрема формалізацією доказів в математиці, що робить засновані на ній мови програмування логічного типу найбільш придатними для роботи зі знаннями.

Основними елементами, що утворюють обчислення предикатів, є:

- 1) константні терми c_1, c_2, \dots ;
- 2) змінні терми x_1, x_2, \dots ;
- 3) функціональні букви f_1, f_2, \dots ;
- 4) предикатні букви p_1, p_2, \dots ;
- 5) логічні символи $\rightarrow, \&, \vee, \neg, \forall, \exists$;
- 6) спеціальна висловлювання Ω .

Окремі висловлювання, звані елементарними, що складаються з предиката і пов'язаних з ним термів, можуть об'єднуватися в складні висловлювання за допомогою логічних зв'язок. До їх складу входять логічні зв'язки: «І» (and, $\&$), «АБО» (or, \vee), «НЕ» (not, \neg) і імплікація (\rightarrow). Особливо важливе значення має імплікація, оскільки саме ця зв'язка застосовується для формування правил і читається «ЯКЩО ..., ТО ...».

Для того, щоб у численні предикатів можна було маніпулювати змінними термами, введена додаткова структура - квантор.

Квантори служать для вказівки заходи, в якій значення змінних повинні бути істинними для того, щоб у цілому висловлення стало істинним. Розрізняють квантор спільності, що позначається \forall , і квантор існування - \exists . Для квантора спільності $\forall (X)$ всі значення змінної в дужках, пов'язані з деякою предметною областю, повинні бути щирі; для квантора існування $\exists (X)$ потрібно, щоб тільки деякі з таких значень були істинні. При вживанні квантора спільності перед висловлюванням можна подумки додавати вираз «для всіх X », а при використанні квантора існування - вираз «існує X ».

Число термів для кожного предиката фіксоване і називають його арністю. Терми визначаються таким чином:

- 1) константний терм є терм;
- 2) змінний терм є терм;
- 3) для функціональної, а також для предикатної букви визначена арність як ціле невід'ємне число. Якщо арність функціональної букви $f \in n$, а t_1, t_2, \dots, t_n - терми, то $f(t_1, t_2, \dots, t_n)$ також терм.

Приписуючи терми предикатним буквам, отримуємо такі формули:

- 1) Ω є формула;
- 2) якщо арність предиката $p \in n$, то $p(t_1, t_2, \dots, t_n)$ є формула;
- 3) для двох формул p_1 і p_2 $p_1 \rightarrow p_2$ - формула;
- 4) якщо формула p містить вільну змінну X , то $X * p(X)$ - формула. При цьому X в цій формулі називається пов'язаною змінною. Не пов'язані змінні є вільними змінними.

Формулу, що не містить вільних змінних, називають замкнутою формулою або висловлюванням. Замкнута формула може бути істинною або помилковою. Символ Ω є помилковою замкнутою формулою; він виражає поняття «протиріччя». Формула $p(t_1, t_2, \dots, t_n)$ говорить про те, що між термами t_1, t_2, \dots, t_n виконується «ставлення».

$A \rightarrow B$ означає, що якщо істинно A , то істинно B . $A \rightarrow \Omega$ означає, що A хибно. Це записується так: $\neg A$.

Серед формул виділяються тотожно істинні формули або аксіоми. Приклад аксіом:

$$\neg\neg A \text{ еквівалентно } A$$

$$A \rightarrow (B \rightarrow C) \text{ еквівалентно } B \rightarrow (A \rightarrow C)$$

та ін. Остання аксіома, наприклад, говорить про незалежність порядку проходження посилок.

Квантори спільності і квантори існування можуть перемежовуватися в межах одного і того ж виразу. У цьому випадку порядок, відповідно до якого вводяться квантифіковані змінні, може впливати на зміст твердження. Наприклад, вираз

$$\forall (X) \exists (Y) (\text{школяр}(X) \rightarrow \text{улюблена_книга}(Y, X))$$

може бути інтерпретовано так: «У кожного школяра є улюблена книга».

Якщо ж змінити порядок проходження кванторів, наприклад:

$$\exists (Y) \forall (X) (\text{школяр}(X) \rightarrow \text{улюблена_книга}(Y, X)),$$

то зміниться і твердження: «Є така книга, яку люблять всі школярі».

Використовуючи такий механізм подання, можна виразити в стандартному вигляді багато складних пропозицій, які властиві розмовній мові. При цьому можна уникнути двозначності таких виразів. Засоби числення предикатів дозволяють, не змінюючи сенсу, перетворити їх у форму, зручну для обробки за допомогою ЕОМ. Таке перетворення стає можливим завдяки тому, що одні логічні зв'язки можуть бути виражені через інші.

Прикладом правила виведення є правило *modus ponens*: якщо тотожно істинна формула A і при цьому тотожно істинна формула $A \rightarrow B$, то формула B тотожно істинна.

Суворий опис маніпуляцій з формулами і правилами виведення дається в літературі з логіки. Наведемо основні з них:

$$A \ \& \ B \text{ еквівалентно } \neg(\neg A \vee \neg B),$$

$$A \rightarrow B \text{ еквівалентно } \neg A \vee B.$$

Останнє перетворення має велике значення при автоматизації операцій обчислення предикатів, хоча воно і важко для розуміння.

Нехай є кілька справжніх висловлювань, що задаються триарним предикатом $p(a_i, b_i, c_i)$, де a_i, b_i, c_i - константні терми.

$$\begin{aligned} & p(a_1, b_1, c_1) \\ & p(a_2, b_2, c_2) \\ & \quad \cdot \\ & \quad \cdot \\ & \quad \cdot \\ & p(a_n, b_n, c_n) \end{aligned}$$

Цей запис можна представити у вигляді таблиці. У ній зведені разом елементи відношення, що задаються предикатом p і справедливого для a, b, c та інших трійок термів.

$p:$

	1	2	3
a1			
a2			
	·	·	·
	·	·	·
	·	·	·
a _n			
b _n			
c _n			

Повна таблиця, що виражає істинні висловлювання, називається базою даних. Оскільки в ній основним поняттям є поняття відносини, її називають реляційною базою даних.

Теорія реляційних баз даних є частиною числення предикатів. Сукупність операцій, що породжують нові таблиці з сукупності існуючих таблиць, які задають відношення, визначають реляційну алгебру. Реляційні операції діляться на дві великі групи: пошук (запит) і оновлення.

Вирази Хорна і алгоритм уніфікації. Якщо використовувати в мовах програмування числення предикатів (1-го порядку) в безпосередньому вигляді, то це призведе до надзвичайно великих обсягів обчислень, що непридатне на практиці. Тому в мові Пролог обмежуються спеціальними виразами, званими виразами Хорна, і тільки до них застосовують описаний нижче метод резолюції.

Вираз Хорна має вигляд:

$$\neg P_1 \vee \neg P_2 \vee \dots \vee \neg P_n \vee P_m,$$

де P_1, P_2, \dots, P_m – літерали.

Атоми називаються позитивними літералами, а їх заперечення - негативними. У цьому виразі є єдина логічна формула без заперечення, і всі формули пов'язані

ні знаком логічної суми «АБО». Це можна описати в іншому вигляді:

$$P_1 \& P_2 \& \dots \& P_n \rightarrow P_m.$$

Вираз читається так: «З P_1 і P_2 і так далі до P_n впливає P_m ». Тут «I» відповідає англійському слову «AND». Мовою Пролог в синтаксисі системи Турбо-Пролог це записується в іншому вигляді:

$$P_m: \neg P_1, P_2, \dots, P_n.$$

Для будь-якої системи логічного програмування характерно та обставина, що для виконання програм застосовується так звана система автоматичного пошуку виводу. Механізм пошуку виведення, використовуваний в Пролозі, бере свій початок від методу резолюцій. Формулами методу резолюцій є диз'юнктив, що мають вигляд:

$$A_1 \vee A_2 \vee \dots \vee A_n,$$

Порядок літералів в диз'юнктив неістотний. Використовувані в Пролозі формули

$$A \text{ і } A_1 \& \dots \& A_n \rightarrow B,$$

еквівалентні хорновським диз'юнктив

$$A \text{ і } \neg A_1 \vee \dots \vee \neg A_n \vee B.$$

Принцип резолюції має на увазі використання регулярної процедури послідовного порівняння із зразком.

Метод порівняння із зразком отримав назву алгоритму уніфікації. За наявності декількох змінних або логічних формул, що містять функції, ці змінні або функції виділяються, частина їх піддається процедурі уніфікації та проводиться їх підстановка в вихідні формули. Якщо при цьому серед них виявляються тотожно істинні формули виду «А або НЕ-А» ($A \vee \neg A$), то вони видаляються. Метод резолюції полягає в тому, щоб, повторюючи цю операцію, показати неможливість безлічі вихідних формул.

Пояснимо на прикладах принцип резолюції та уніфікацію. Запишемо вираз: «Х - лікар» у вигляді $\text{doctor}(X)$. Тут doctor - предикатна буква, X - змінний терм. Крім того, запишемо вираз «Х - хірург» у вигляді $\text{surgeon}(X)$. Таким чином, вираз «хірург є лікарем» можна записати у вигляді:

$$\text{surgeon}(X) \rightarrow \text{doctor}(X).$$

Спробуємо тепер, задавши вираз «Амосов - хірург», тобто $\text{surgeon}(\text{Амосов})$, де Амосов - константний терм, довести методом резолюції вираз $\text{doctor}(\text{Амосов})$.

Насамперед додамо до вихідної множини формул заперечення того твердження, яке ми хочемо довести, а саме, $\neg \text{doctor}(\text{Амосов})$, як показано нижче:

$$\neg \text{surgeon}(X) \vee \text{doctor}(X) \quad (5.1)$$

$$\text{surgeon}(\text{Амосов}) \quad (5.2)$$

$$\neg \text{doctor}(\text{Амосов}). \quad (5.3)$$

Формула $\neg \text{surgeon}(X) \vee \text{doctor}(X)$ логічно еквівалентна формулі $\text{surgeon}(X) \rightarrow \text{doctor}(X)$. Тому необхідно здійснити підстановку X , що уніфікує: Амосов щодо формули (5.1) або (5.2). В результаті віддаляється тотожно істинна формула

$$\neg \text{surgeon}(\text{Амосов}) \vee \text{surgeon}(\text{Амосов})$$

і залишається лише вираз $\text{doctor}(\text{Амосов})$:

$$\text{ф.}(5.1) \vee \text{ф.}(5.2) \neg \text{surgeon}(X) \vee \text{doctor}(X) \vee \text{surgeon}(\text{Амосов})$$

↓ уніфікація X і Амосов

$$\text{ф.}(5.1) \vee \text{ф.}(5.2) \neg \text{surgeon}(\text{Амосов}) \vee \text{doctor}(\text{Амосов}) \vee \text{surgeon}(\text{Амосов})$$

↓

$$\text{ф.}(5.1) \vee \text{ф.}(5.2) \text{ doctor}(\text{Амосов})$$

Після цього, приєднавши до вираження $\text{doctor}(\text{Амосов})$ через зв'язку «АБО» формулу (5.3), знову отримуємо тотожний справжній вираз

$$\text{doctor}(\text{Амосов}) \vee \neg \text{doctor}(\text{Амосов}),$$

яке також видаляється, даючи при цьому порожню множину формул:

$$\text{ф.}(5.1) \vee \text{ф.}(5.2) \vee \text{ф.}(5.3) \text{ doctor}(\text{Амосов}) \neg \text{doctor}(\text{Амосов})$$

↓ видалення істинної і помилкової формул

$$\text{ф.}(5.1) \vee \text{ф.}(5.2) \vee \text{ф.}(5.3) \text{ порожній диз'юнкт}$$

Це показує нездійсненність формул (5.1) - (5.3). Звідси методом обґрунтування висновку приходимо до формули $\text{doctor}(\text{Амосов})$. Даний приклад мовою Пролог виразиться в наступному вигляді:

$$\text{doctor}(X): \neg \text{surgeon}(X)$$

$$\text{surgeon}(\text{Амосов})$$

$$? \neg \text{doctor}(\text{Амосов}).$$

Вираз $\text{surgeon}(\text{Амосов})$ говорить про те, що предикат $\text{surgeon}(\text{Амосов})$, безумовно, виконується (є фактом). Вираз

$$? \neg \text{doctor}(\text{Амосов})$$

називається цільовим твердженням і є скороченням виразу

Ω : \neg doctor (Амосов).

Даний запис означає, що якщо виконується предикат doctor (Амосов), то не виконується нічого, отже, не виконується і саме вираження doctor (Амосов). Це говорить про те, що в мові Пролог, так само як і в методі резолюції, взявши за відправну точку заперечення формули doctor (Амосов), можливо методом обґрунтування висновку довести твердження doctor (Амосов).

Насамперед у лівій частині проводиться пошук предиката, що збігається з цільовим твердженням doctor (Амосов), і виробляється уніфікація змінної X і константного терму Амосов. Потім шукається вираз, що містить у лівій частині предикат, що співпадає з новою метою (в даному випадку doctor (Амосов), що утворюється в правій частині цього виразу внаслідок її уніфікації. Потім аналогічно виробляються подальші уніфікації. Програма мовою Пролог являє собою деяку множину виразів Хорна, і виконується вона шляхом міркувань від цільового затвердження в зворотному напрямку. Якщо логічних формул, що підлягають уніфікації, багато, то проводиться уніфікація для однієї з них. Якщо уніфікація задовільна, то програма виконується далі. Якщо ні, відбувається повернення (backtracking) і виконуються інші можливі уніфікації.

Таким чином, у мові Пролог міркування ведеться у зворотному напрямку методом проб і помилок.

Порівняння мов Лісп і Пролог. Пролог часто називають мовою штучного інтелекту. Таке формулювання, проте, некоректне.

Лісп - це найбільш відома мова ШІ. Якщо визначити аналогічним чином сферу застосування мови Пролог, то неминуче почнеться зіставлення цих двох мов з метою виявити, який з них краще. З точки зору практики для кожної мови існує своя методологія програмування. Єдино прийнятний спосіб порівняння мов, що дає підстави заявити, що одна мова «краще» іншого, полягає в обліку «людських» чинників, тобто у порівнянні читабельності програм, легкості у користуванні мовами і т.д. Американські фахівці з ШІ віддають перевагу Лісп швидше за звички до нього, ніж через недоліки Прологу як такого. Цим і пояснюється здивування багатьох американських фахівців з ШІ, коли японськими фахівцями в якості основної методології розробки програмних засобів для проекту ЕОМ п'ятого покоління було обрано логічне програмування.

Більш важливим моментом є те, що сфера застосування Прологу не обмежується штучним інтелектом. Ця мова успішно використовувався як для вирішення багатьох традиційних прикладних завдань (програмна інженерія, інтерфейси баз даних, системи допомоги у прийнятті рішень), так і для вирішення завдань ШІ (експертні системи, системи обробки природної мови).

5.3.2. Мови інженерії знань

Мови інженерії знань - це зазвичай майстерні інструментальні засоби розробки експертних систем, що складаються з мови побудови експертних систем, зануреного в обширне підтримуюче оточення.

Розглянемо два підходи до класифікації мов інженерії знань. Відповідно до першого з них, розрізняють мови інженерії знань:

- скелетні;
- універсальні.

Скелетна мова інженерії знань є просто "роздягнутою" (або "порожньою") ЕС, тобто ЕС без спеціальних предметних знань. Скелетні системи зазвичай застосовні до досить вузького класу проблем, що стримує можливості розробника ЕС.

До «порожніх» ЕС відносять, наприклад, систему KAS - скелетна мова для діагностики та розпізнавання, отриманий з часто описуваної в літературі ЕС PROSPECTOR. Аналогічно, на базі системи MYCIN для діагностики та лікування бактеріальних інфекцій згодом розробили скелетну систему EMYCIN (EMPTY MYCIN), а система CASNET - консультаційна система по глаукомі - стала скелетною системою EXPERT.

Універсальна мова інженерії знань відрізняється від скелетної здатністю осягнути більш широке коло завдань - досягається це за рахунок відомого їх ускладнення в частині організації пошуку інформації, різноманітності механізмів виводу.

Відповідно до другого підходу, розрізняють:

- «порожні» експертні системи (мета-експертні системи);
- оболонки (shells) експертних систем;
- інструментальні ЕС.

Враховуючи, що в обох класифікаційних схемах на першому місці фігурують «порожні» експертні системи як найпростіша різновид мов інженерії знань, приходимо до висновку, що об'єднання прийняті у другій схемі категорій «оболонки ЕС» і «інструментальні ЕС» еквівалентні категорії «універсальні», прийнятої в першій класифікаційній схемі.

Інструментальні ЕС - системи, орієнтовані на створення «порожніх» ЕС, які після наповнення знаннями перетворюються на прикладні ЕС. Прикладами такого підходу є проект американської фірми Teknowledge Inc. і розробка АН СРСР «ПЕС» - проте про ефективність цих систем судити важко зважаючи на недолік відповідної інформації.

Значно більше відомостей про оболонки експертних систем (ОЕС) - спеціальному програмному оточенні розробки систем, орієнтованих на знання. До складу такого програмного оточення зазвичай відносяться:

- інтерпретатори і (або) компілятори для обробки знань, представлених у вигляді логічних виразів, продукційних правил, фреймових моделей;
- механізми логічного висновку, що обробляють вищеперелічені моделі знань, і засоби контролю за коректністю рішення;
- засоби управління базою даних;

- засоби для отримання знань і модифікації баз знань;
- засоби для побудови інтерфейсу з користувачем в проблемно-орієнтованій ЕС.

З часом розробники ОЕС стали переходити від універсальності розробок до все більшого ступеня спеціалізації, оскільки вузькоспеціалізовані системи практичніше при розробці ЕС. Крім того, для цих систем підтримки інженерії знань характерні використання досить простих засобів представлення знань продукційного і (або) фреймового типу, застосування вбудованої в систему стратегії виведення рішень, а також наявність екранного редактора правил, компілятора або інтерпретатора правил, розвинених засобів трасування процесу вирішення завдань, що спираються на багатовіконну графіку.

Приклади зарубіжних ОЕС. У літературі можна знайти ряд прикладів зарубіжних ОЕС. Розглянемо деякі з них (в дужках після найменування ОЕС вказана фірма-розробник).

1-st CLASS (Program in Motion). Це система для побудови дерев рішень, що складаються з правил. Система будує правила з прикладів, що подаються у формі електронних таблиць. Завдання може бути розбите на модулі, кожен з яких складається з набору прикладів. Система логічного висновку пов'язує такі модулі, використовуючи пряму або зворотну стратегію. Крім того, є можливість безпосередньої побудови правил у графічній формі. При виведенні може використовуватися один з алгоритмів: система порівнює запити до прикладів, які є в базі даних, або використовує вже побудовані дерева правил, або працює в режимі оптимізації, коли при консультації системою перебудовується дерево правил для завдання питань в більш природному порядку. Швидкість системи досить велика, оскільки в її складі є компілятор правил. У системі є інтерфейси з іншими мовами і базами даних. Використовується операційна система MS DOS. Мова розробки - Паскаль.

EXPERT EDGE (Human Edge Software Corp.). Система для створення ЕС класифікаційного типу. Використовує базу правил, зворотний ланцюжок виводу, імовірнісний підхід, байєсівську формулу підрахунку впевненості. У системі є багаті можливості для реалізації інтерфейсу користувача (строкове введення, меню, редактор БЗ). Є природно-мовний інтерфейс і дуже потужні можливості налагодження. Деякі версії системи для професійних додатків включають можливості для математичних обчислень. Мова реалізації - Сі, операційна система - MS DOS, має мовні інтерфейси та інтерфейси з базами даних.

KES (Software Architecture and Engineering). Система, що підтримує продукційні правила, які породжують правила і байєсівський тип правил для областей, в яких знання можуть бути представлені ймовірно. KES застосовується переважно для вирішення задач класифікації і може бути включена в інші системи. Засіб написано мовою Сі, працює в операційних системах MS DOS або PC DOS, є версії для машин серії VAX.

MICROEXPERT (Digital Equipment Corp.). Система, що дозволяє будувати

діагностичні експертні системи. База знань являє собою мережу продукційних правил, оперує з байесівською формулою підрахунку впевненості, чіткою логікою і числовими даними, підтримує прямий та зворотній виводи. Інтерфейс побудований на основі строкового введення, є можливість трасування, пояснення і протоколювання ходу консультації на принтер. Система створена на мові UCSD-PASCAL, працює в операційній системі MS DOS типу IBM PC.

NEXPERT OBJECT (Neuron Data Corp.). Система, що базується на використанні правил. Підтримує різні типи правил (до 2000) і комбінацію прямого і зворотного висновків. Система може автоматично будувати мережі правил у графічному зображенні, що дозволяє спостерігати, яким чином правила пов'язані один з одним. Система має можливості представлення обох видів фреймів (з успадкуванням і без) і механізм зіставлення з образом. У цій системі велику увагу приділено графічному представленню як баз даних, так і пояснюючих можливостей, що робить можливим організацію природничих та дружніх інтерфейсів розробника і кінцевого користувача. Засіб написано мовою Сі, використовується на IBM-сумісних ПК з операційною системою MS DOS або VMS, має інтерфейси з Сі і Паскалем, а також з dBASE III.

Приклади вітчизняних ОЕС

СПЭИС (ВНИИ АН СРСР). Система призначена в основному для вирішення завдань діагностичного типу. Є середовищем проектування ЕС з відкритою архітектурою. Має три рівня доступу користувача: 1) «порожня» ЕС з наявним мовою подання знань (ЯПЗ) і механізмом виведення; 2) об'єктно-орієнтована мова програмування, що підтримує фрейми з приєднаними процедурами і продукціям; 3) мова сімейства Лісп - muLISP. Перший рівень проектування ставить мінімальні вимоги до кваліфікації розробника як програміста, третій - максимальні. Представлення знань - трирівневе: 1) мета-знання - верхній рівень БЗ, що містить знання про те, як вирішувати завдання; 2) подання знань про стани, які розпізнаються - фрейми, звані Н-концепціями; 3) подання знань про діагностичні ознаки - фрейми, іменовані Р-концепціями. ЯПЗ - це Р-концепції та Н-концепції, що представляють собою інформаційні структури у вигляді фреймів. Інформація при цьому зберігається в слотах - інформаційних підструктурах. Механізм логічного висновку побудований з використанням правила продукції. Крім того, в системі є спеціальні предикативні функції, за допомогою яких можна аналізувати процес пошуку рішень. Інтерфейс системи побудований на основі багатовіконною системи меню. Система СПЕІС написана мовою Лісп.

ОЕС ВИНТИ АН СРСР. Розроблена під керівництвом В.К.Фінна. Реалізує деякі принципи, покладені в основу радянсько-угорського проекту побудови логіко-інформаційних обчислювальних систем (ЛІВС), альтернативних відомому японському проекту розвитку логічних засобів ЕОМ п'ятого покоління: 1) створення логіко-математичного подання предметної області (на противагу традиційному підходу до створення ЕС, при якому програмно фіксується особистісний досвід експертів), 2) об'єднання позаособистісних і особистісних знань, 3) відкритість системи, що забезпечує швидку модернізацію БЗ; 4) включення ло-

гічного апарату в механізмі висновків; 5) обґрунтування результатів з використанням імітаційних моделей предметної області. У системі реалізований спеціальний метод міркувань - ДСМ- метод (названий так на честь Джон Стюарта Мілля, дослідника індуктивних міркувань), що представляє собою не продукційний підхід, а вирішувач типу «правдоподібний висновок + достовірний висновок» для виявлення причинно-наслідкових залежностей на безлічі нечислових даних в умовах неповноти інформації). Це формальний апарат багатозначної логіки предикатів, що використовує поняття подібності об'єктів, яке визначається за допомогою різних предикатів подібності.

ОЕС ФИАКР. Засіб створення діагностичних ЕС. Знання в БЗ організовані за допомогою «дескрипторів» властивостей - трійок вигляду «підсистема-атрибути-значення». Ознаки-атрибути бувають 2-х видів-декларативні (описуються безліччю можливих значень) і процедуральні (ім'я ознаки ототожнюється з ім'ям деякої обчислювальної програми, в якості вихідних даних якої використовується значення ознаки). Для організації діалогу використовується поняття «прапора» - записи БЗ, аналогічної дескриптору властивостей. Прапори розрізняються за призначенням: 1) прапори існування підсистеми або атрибута - вказують, чи потрібно взагалі активізувати підсистему або атрибут, 2) прапори цілей - дозволяють організувати цілеспрямований діалог, 3) прапори експертних груп - дозволяють керувати ходом консультації, роблять діалог лаконічним і осмисленим; 4) прапори анкет - організують форму діалогу ("анкета"). Математичний апарат виводу - логічний, заснований на обчисленні станів тверджень (таких станів може бути чотири: встановлене, спростоване, невизначене, конфліктне). ФИАКР реалізований на мові Паскаль.

ДИЭКС (розробник в [5] не вказаний). Система, орієнтована на створення ЕС діагностичного типу. Обмеження на застосування ДИЭКС накладаються не предметною областю, а характеристиками завдання. Система має два основні режими взаємодії з користувачами: 1) режим навчання системи експертами, 2) режим отримання консультації у системи. Факти в БЗ представляються у вигляді фреймів. Генератор баз правил орієнтований на реалізацію продукційних правил з елементами нечіткої логіки і байєсівського формалізму. У процесі логічного висновку система може задавати користувачеві питання чотирьох видів: чіткі («так», «ні», «не знаю»), нечіткі (із зазначенням суб'єктивного ступеня впевненості), числові (потрібне введення числових значень показника), вибору (потрібен вибір із запропонованих альтернатив). При взаємодії з користувачем в режимі пояснення можливий перегляд на екрані графічного зображення мережі прийняття рішення, з вибором цікавого для нас вузла і отриманням вичерпної інформації про нього. Результат консультації має вигляд звіту-списку досліджених гіпотез із зазначенням оцінок їх достовірності. Можливо протоколювання ходу консультації. Інтерфейс виконаний на основі вікон і меню. Система написана мовою Пролог, Сі і працює під управлінням операційних систем MS DOS, PC DOS.

ОЕС "Универсал" (НВО «Славутич», Київ). Засіб створення діагностичних

ЕС, що дозволяє створювати і експлуатувати в рамках одного програмного продукту практично необмежену кількість завдань і підзадач - причому завдання можуть належати різним проблемним областям.

Основні режими функціонування: консультування (діагностика), навчання (модернізація БЗ), пояснення (аргументація прийнятого рішення) і видача рекомендацій. Інформація про діагностичні ознаки, класи і значення відповідних умовних ймовірностей (при ймовірнісному механізмі міркувань) або ваг ознак (при детерміністському механізмі) зберігається в БЗ. Розробник може вибрати один з трьох механізмів виведення: байєсівську стратегія, зважене голосування, лінійне діагностичне правило. Результат діагностики представляється на моніторі у вербально-графічній формі, з кількісною оцінкою вірогідності альтернативних рішень, і може бути запроTOCOLьований для подальшого перегляду і редагування, а також використання в режимі навчання (для автоматичного оцінювання умовних ймовірностей ознак, з наступною модернізацій БД). Інтерфейс системи побудований на основі багатовіконної системи меню. Мова програмування - Clipper, операційне середовище MS DOS.

Більш докладно про ОЕС «Універсал» - у Розділі 9.

ОЕС МЕД і МEG. Оболонка МЕД - це спочатку система медичної спрямованості (наприклад, меню мають назви "Захворювання", "Лікування" і т.п.). Вузкість області докладання такої системи була подолана перейменуванням рівнів деревовидної схеми пошуку і вилученням "жорсткої" частини системи у вигляді інформаційно-довідкової системи і набору таблиць для рефлексодіагностики і рефлексотерапії: так була отримана оболонка МEG, яку, в принципі, можна застосовувати в самих різних проблемних областях.

Організація знань має вигляд ієрархічної системи, самий нижній рівень якої являє собою «інформаційний листок» - текстовий модуль, що містить найменування гіпотези, яка перевіряється (при вирішенні діагностичних завдань) чи рекомендації з лікування (при виборі лікувальних заходів). Механізм прийняття рішення надзвичайно простий і полягає в конкатенації вмісту обраних інформаційних листків, з наступним текстовим редагуванням отриманого таким чином документа (у медицині, наприклад, це висновок лікаря або програма лікувально-реабілітаційних заходів).

Більш докладно про системи МЕД і МEG - у Розділі 9.

5.3.3. Допоміжні засоби

Як приклад допоміжних засобів побудови ЕС наведемо спеціальну програму SINTEZ - програму для автоматизованої обробки протоколів, що містять результати функціонування ЕС ЭСИД. Ця програма допомагає витягати знання з протоколів, виробляючи їх статистичну обробку ("статистика прецедентів") і оцінюючи значення умовних ймовірностей $P(X_i/Y_j)$.

Результати функціонування програми SINTEZ потім можуть бути використані в програмі ЭСИД в режимі "Робота з БД": при виборі поєднання "ознака-

захворювання" програма запитує дозволу скористатися результатами статистичної обробки програми SINTEZ. Якщо такого дозволу немає, тоді користувач вводить в БД значення $P(X_i/Y_j)$, запозичене з літератури, у експертів, з особистого досвіду. Якщо дозвіл дається, тоді у відповідне місце БД заноситься обчислене програмою SINTEZ значення $P(X_i/Y_j)$.

Зауважимо, що допоміжний засіб такого роду може бути реалізовано не обов'язково у вигляді окремої допоміжної програми - це може бути і режим функціонування основної програми. Так, наприклад, проводиться обробка протокольної інформації в ОЕС "Універсал".

5.3.4. Засоби підтримки

Найважливішою частиною засобів підтримки програмування ЕС є редактор баз знань. При вдалій реалізації такого редактора користувач може отримати можливість прямого втручання у вміст баз знань, не вдаючись до допомоги професійних програмістів - це дуже важлива можливість, на неї часто вказують розробники відповідного інструментарію.

ЭСИД і ОЕС "Універсал" володіють таким редактором баз знань, що досягнуто завдяки двом моментам:

- база знань трактується як реляційна база даних з явно зазначеними зв'язками між подіями;
- максимально використовуються засоби редагування СУБД Clipper.

Іншою важливою частиною засобів підтримки програмування ЕС є засоби швидкої розробки процедур введення / виведення і пояснення. При розробці оболонки "Універсал" та спеціалізованої ЕС для іриодіагностики ЭСИД використані засоби введення / виведення, передбачені розробниками СУБД Clipper - цих коштів виявилось достатньо в силу відносної простоти розроблюваних ЕС, в яких передбачено два способи введення інформації в комп'ютер в режимі діалогу:

- введення даних з клавіатури;
- введення даних з протоколів, а також два способи виведення інформації:
- виведення результатів на екран дисплея;
- роздруківка результатів на принтері.

Узагальнюючи досвід розробки згаданих ЕС, а також інформаційно-довідкових систем МЕД і МEG, відзначимо бажаність наявності наступних засобів швидкої розробки процедур введення:

- організація деревовидної (ієрархічної) системи меню і підміню (такі кошти були б корисні при розробці ЭСИД, МЕД і МEG);
- організація багатовіконного інтерфейсу (що було б корисно при розробці програми "Універсал").

Що стосується бажаних коштів швидкої розробки процедур виведення, до таких віднесемо:

- засоби автоматичного формування ув'язнення з можливістю багаторівневого (ієрархічного) його структурування (наприклад, для цілей іри-

додіагностики розділи верхнього рівня - дані про пацієнта, результати конституціонального огляду, карта поглибленого огляду, оцінка стану систем організму, висновок і рекомендації; нижній рівень - оцінка стану підсистем та органів, групи рекомендацій і т.д.);

- кошти швидкого напівавтоматичного (з використанням заздалегідь заготовлених електронних довідників) редагування тексту висновку;
- засоби розробки та модернізації електронних довідників;
- засоби редагування тексту ув'язнення, типові для використовуваного середовища програмування.

Засоби пояснення у програмах ЕСІД та "Універсал" реалізовані схожим чином:

- в режимі "Діагноз" (при перегляді автоматично сформованого тексту висновку) перераховуються вибрані гіпотези і відповідні їм задані ознаки (із зазначенням ступеня вираженості ознак);
- разом з найменуванням виділеної гіпотези пропонується текст відповідних рекомендацій (цей текст заздалегідь заноситься в БД в режимі настрійки).

Короткі висновки

Вибір засобів розробки ЕС повинен відбуватися з урахуванням особливостей ЕС як програмного продукту. До таких особливостей відносяться:

- 1) компетентність (експертний рівень рішень, робастність);
- 2) здатність до символічних міркувань (подання і переформулювання знань в символічному вигляді);
- 3) глибина аналізу (важкі завдання, складні правила);
- 4) самосвідомість (здатність досліджувати свої міркування, пояснювати прийняті рішення).

Прийнято виділяти такі три групи критеріїв класифікації ЕС:

- 1) за характеристиками предметної області;
- 2) з організації спілкування з користувачем;
- 3) з реалізації внутрішньої структури.

До ЕС першої групи відносяться ЕС, вирішуючі завдання: інтерпретації даних; діагностики; прогнозування; проектування; планування; моніторингу; налагодження; навчання; управління.

Що стосується організації спілкування з користувачем, серед ЕС превалюють системи консультаційного типу, і медичні діагностичні системи відносяться саме до цього класу. Крім того, для спілкування з користувачем можуть використовуватися природні і штучні мови (меню, шаблони, строкове введення та ін.) Різними можуть і бути і способи пояснення прийнятих рішень.

За критерієм "програмна архітектура" розрізняють такі різновиди експертних систем: спеціалізовані; оболонки ЕС (порожні ЕС); розподілені ЕС (мережеві

ЕС) ; гібридні ЕС (поєднання моделей прийняття рішення); узагальнені ЕС (розподілені ЕС з сильним обчислювальним компонентом).

Інструментальні засоби побудови експертних систем поділяють на чотири категорії:

- 1) мови програмування;
- 2) мови інженерії знань;
- 3) допоміжні засоби;
- 4) засоби підтримки.

Серед мов програмування ЕС виділяються Лісп і Пролог, як найбільш популярні. Розрізняють такі види мов інженерії знань:

- 1) скелетні;
- 2) універсальні.

Скелетна мова інженерії знань є просто "роздягнутою" (або "порожньою") ЕС, тобто ЕС без спеціальних предметних знань. Скелетні системи зазвичай застосовуються до досить вузького класу проблем, що стримує можливості розробника ЕС .

Універсальна мова інженерії знань відрізняється від скелетної здатністю осягнути більш широке коло завдань - досягається це за рахунок відомого їх ускладнення в частині організації пошуку інформації, різноманітності механізмів виводу.

До універсальних мов можна віднести:

- 1) оболонки (shells) експертних систем;
- 2) інструментальні ЕС.

Література до Розділу 5

1. Уотермен Д. Руководство по экспертным системам. М., Мир, 1989. - 388с.
2. Экспертные системы. Принципы работы и примеры. Под ред. Р.Форсайта.- М.:Радио и связь, 1987.-224с.
3. Гриненко Л.М., Продеус А.М., Шукевич Ю.В., Якубов С.В. Прогнозування успішності навчання учнів за результатами психологічного тестування. – Педагогіка і психологія, 1996, N2, с.116-122.
4. Гаврилова Т.А., Червинская К.Р. Извлечение и структурирование знаний для экспертных систем.-М., Радио и связь, 1992.-197с.
5. Продеус А.Н., Сядро Т.А. Опыт разработки компьютерных систем поддержки ириодиагностики. Тезисы докладов международной конференции по физической и биомедицинской электронике. Киев, 1997.

Контрольні запитання до Розділу 5

1. Які особливості ЕС як програмного продукту повинні враховуватися при виборі засобів розробки ЕС?

2. Які групи критеріїв класифікації ЕС вам відомі?
3. Які інструментальні засоби побудови експертних систем вам відомі?
4. Які мови програмування, і чому, використовуються найбільш часто при розробці ЕС?
5. Що собою являють мови інженерії знань і які їхні різновиди вам відомі?
6. Що собою представляє «скелетна ЕС» і в чому полягає її головний недолік?
7. Що собою представляють «оболонки ЕС»?

Розділ 6 Розробка ЕС

В даному розділі розглядаються питання, пов'язані як з розробкою, так і з придбанням готової ЕС. При постановці завдання про розробку ЕС, перші питання, які повинні задаватися - це питання про доцільність і принципovu можливість такої розробки. І лише при позитивному вирішенні цих питань можна намічати етапи та рівні розробки, приступати до підбору експертів. При покупці готової ЕС проблема спрощується, проте як правильно це зробити - також важливе питання, докладно розглянутий у цьому розділі.

6.1. Об'єктивні передумови для розробки ЕС

Розглянемо наступні об'єктивні передумови для розробки ЕС:

- 1) низька ступінь структурованості знань чи неповнота інформації, необхідної для прийняття рішень у цій проблемній області;
- 2) значні обсяги інформації, якими необхідно оперувати при прийнятті рішення, у поєднанні з жорсткими вимогами до швидкості прийняття рішення;
- 3) висока відповідальність при прийнятті рішень;
- 4) стомлюваність осіб, котрі приймають рішення (ОПР);
- 5) різний рівень кваліфікації ОПР.

Низький ступінь структурованості знань або їх неповнота - це фактично визначення неалгоритмічності знань, оперування якими з використанням традиційних програмних засобів типу ППП практично неможливо.

Складність розв'язуваної задачі може в значній мірі визначатися великою кількістю даних, що підлягають обробці. У теорії розпізнавання образів цей ефект іноді іменують "прокляттям розмірності", маючи на увазі, що розпізнаваний об'єкт представляється точкою в багатовимірному просторі ознак - людині вельми скрутно змоглядно уявити собі такий простір, не кажучи вже про необхідність оперувати якимись багатовимірними розділяючими поверхнями для віднесення розпізнаваного (діагностуємого) об'єкта до того чи іншого класу об'єктів (у медицині - визначення найменування захворювання або групи захворювань, прогнозування стану).

Очевидно, що незважаючи на "погану якість" вихідних даних дуже часто потрібна висока вірогідність діагностики - це властиво ряду предметних областей і, особливо, медичній діагностиці. Від якості діагностики тут залежить правильність призначених лікувальних заходів і, в кінцевому рахунку, здоров'я і життя пацієнтів. Разом з тим, очевидно, що людина досить швидко стомлюється: за визнанням лікарів-діагностів, після п'ятого-шостого пацієнта вони вже "погано розуміють" - підтримка лікаря з використанням комп'ютерних засобів у даному випадку не здається зайвою. Нарешті, добре відомим фактом є і відмінність

кваліфікації ОПР, зокрема, медичних працівників. І якщо для висококваліфікованого фахівця комп'ютерна підтримка потрібна у меншій мірі, для початківця або недостатньо кваліфікованого фахівця вона повинна виявитися чи не необхідною підмогою.

6.2. Умови доцільності розробки експертної системи

Як вказується в літературі, вдаватися до розробки експертної системи слід тільки тоді, коли така розробка *можлива, виправдана і доречна*.

Умови можливості. Розробка ЕС вважається можливою при *спільному* виконанні наступного ряду умов:

- існують висококваліфіковані експерти, які згодні співпрацювати при розробці ЕС і здатні пояснити використовувані ними методи;
- думки експертів (якщо їх декілька) не суперечать один одному;
- розв'язувана задача містить компонент, що вимагає інтелектуальних навичок;
- розв'язувана задача не є новою (задача досить добре структурована, не потрібні значні дослідження шляхів її вирішення);
- розв'язуване завдання не занадто важке (піддається розбиттю на ряд простих підзадач);
- розв'язувана задача не базується на зверненнях до здорового глузду.

Умови виправданості. Виправданням розробки ЕС може слугувати хоча б одне з наступних умов:

- є ймовірність одержати великий дохід від експлуатації ЕС;
- для вирішення завдання потрібно багато висококваліфікованих експертів, чії послуги дорогі;
- можливий відхід наявних експертів створює загрозу вирішуваності важливого завдання (через втрату накопиченого досвіду);
- умови розв'язання задачі настільки несприятливі, що експертні рішення повинні приймати або неспеціалісти в даній проблемній області, або рішення взагалі повинні прийматися автоматично, без участі людини.

Умови доречності. Ключові фактори, що визначають доречність (розумність) розробки ЕС:

- наявність специфічних властивостей:
- маніпулювання символами і символічними структурами (обробка інформації, представлена якісно, тобто у вигляді шкал найменувань);
- можливе використання евристичних методів рішення задачі;
- завдання не повинне бути занадто легке (на навчання фахівця йде кілька років);
- розумна широта постановки задачі: задача повинна бути досить широкою, щоб представляти практичний інтерес, і досить вузькою, щоб впоратися з розробкою ЕС для її вирішення.

Приклад. Проілюструємо здійснимість наведених вище умов на прикладі розробки ЕС для іридодіагностики ЭСИД і ЭСИД -2.

Завдання іридодіагностики на момент розробки ЭСИД і тим більше ЭСИД -2 не була новою: існує наукова література, що містить значний обсяг результатів статистичних досліджень зарубіжних та вітчизняних авторів і відображає багаторічний досвід роботи фахівців в даній галузі медичної діагностики. У м. Києві існує іридологічна школа, тривалий час очолювала, В.В. Кривенко. Прихильники цієї школи, як правило, є висококваліфікованими експертами та мають досвід співпраці з інженерами і програмістами - розробниками автоматизованих систем іридодіагностики. Належність експертів одній школі служить певною гарантією того, що думки різних експертів не суперечитимуть одна одній.

Розв'язувана задача безумовно містить компонент, що вимагає інтелектуальних навичок - це завдання діагностики, що належить класу задач розпізнавання образів. Завдання такого типу досить добре піддаються формалізації як з використанням діагностичних таблиць, так і з використанням більш складних алгоритмів, аж до байесівської стратегії прийняття рішення.

Виправданням розробки ЭСИД і ЭСИД -2 служив успіх на Україні методу іридодіагностики: інтерес населення і лікарів до методу, здатному забезпечити доклінічну експрес-діагностику, призвела до значної кількості фахівців - іридодіагностів і, як наслідок, створив передумови для широкого розповсюдження іридодіагностичних ЕС серед таких фахівців - цьому одночасно сприяв і швидкий технічний прогрес у галузі створення і застосування відносно недорогих ПЕОМ.

Специфічними властивостями методу іридодіагностики є оперування з шкалами найменувань (найменування ознак і захворювань), тобто з інформацією якісного характеру, обробка якої цілком допускає використання як класичних, так і евристичних методів рішення задачі. Розв'язувана задача не є занадто легке - на навчання фахівця йде багато часу (мінімум два тижні на виклад теоретичних основ і мінімум рік практичної діяльності). Разом з тим, завдання задовольняє вимогу розумної широти постановки завдання: представляючи безсумнівний практичний інтерес, вона, разом з тим, досить добре визначена, що дозволяє бути впевненим в успішному завершенні розробки ЕС.

6.3. Рівні розробки ЕС

Основні принципи розробки ЕС:

- можливість розвитку системи;
- можливість переходу кількості в якість при розвитку системи.

Ці принципи реалізуються у створенні так званих прототипів ЕС, що відображають процес кількісної та якісної зміни ЕС в процесі їх розробки.

Розрізняють наступні прототипи ЕС:

- демонстраційна;
- дослідницька;

- досвідчена;
- промислова;
- комерційна.

Цілі створення демонстраційного прототипу:

- переконання органу, що фінансує, в необхідності створення ЕС;
- перевірка правильності проектних концепцій.

Дослідницький прототип являє собою "малу" ЕС, що володіє, в принципі, всіма необхідними функціями, але з базою знань відносно невеликих розмірів - мета розробки такого прототипу полягає в оптимізації прийнятих рішень щодо складу та взаємодії структурних компонентів ЕС між собою.

Досвідчена ЕС - система, що пройшла хорошу апробацію шляхом експлуатації в реальних умовах. Обсяг її варіює від середнього до великого.

Промислова ЕС - велика система, що пройшла інтенсивну перевірку в умовах дослідної експлуатації і, ймовірно, переписана на більш ефективній мові для прискорення функціонування та економії машинної пам'яті.

Лише дуже небагато з ЕС доведені до комерційного рівня - це промислові ЕС, що забезпечують дуже високу ефективність вирішення поставлених завдань і комерційно себе окупають.

6.4. Етапи розробки ЕС

Розробку ЕС можна вважати складається з п'яти взаємодіючих і перехреснюючихся етапів:

- ідентифікація;
- концептуалізація;
- формалізація;
- реалізація;
- тестування.

У поняття взаємодії та перекриття тут вкладається наступний зміст: хоча за часом ці етапи слідує послідовно, проте можливі повернення до попередніх етапів і паралельне виконання одночасно декількох етапів.

На етапі ідентифікації розробник і експерт визначають суттєві особливості завдання:

- тип завдання і широту її постановки;
- цілі створення ЕС;
- склад учасників розробки;
- необхідні тимчасові, фінансові та матеріально-технічні ресурси.

Типовий склад групи розробників зображений на рис.6.1.

На етапі концептуалізації інженер знань і експерт:

- вирішують, які поняття, відносини і механізми управління потрібні для опису вирішення завдань в обраній області;
- визначають підзадачі, стратегії та обмеження, властиві даній задачі;
- визначають ступінь подробиці, з якою потрібно представляти знання.

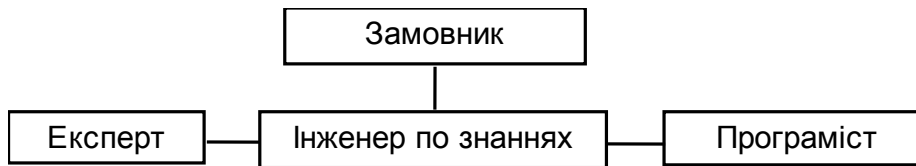


Рис. 6.1

Формалізацію здійснює інженер по знаннях:

- вибирає підходящий засіб розробки ЕС (мова програмування, мова інженерії знань);
- формує технічне завдання (ТЗ) - основний документ для програмістів.

У змісті ТЗ на ЕС входять:

- мета розробки (ЕС для практичної діяльності, для наукових досліджень, для навчання молодих фахівців, комбіновані ЕС ...);
- основні функції ЕС;
- бази знань (переліки симптомів і захворювань, встановлення відповідності між переліками);
- вимоги до інтерфейсу;
- вимоги до документації.

Проект ТЗ узгоджується з заказником, експертом і програмістами і видається програмістам.

На етапі реалізації розробники (інженер по знаннях, програмісти, експерт) перетворюють формалізовані знання в працюючу комп'ютерну програму. При цьому здійснюється контроль за виконанням програмістами вимог ТЗ.

Тестування здійснюється спільно розробниками і заказником і складається у визначенні правильності обраних рішень - в результаті можуть знадобитися неодноразові переробки ЕС. При цьому проводиться пошук відповідей на наступні питання:

- чи дозволяє використання системи приймати правильні рішення?
- чи не містить механізм виведення суперечливих правил?
- чи дозволяє стратегія управління системою працювати користувачеві в зручній для нього манері?
- чи адекватно пояснення системи описують, як і чому отримані ті чи інші висновки?
- чи охоплюють тестові завдання всю заплановану область компетентності системи?

Крім того, попутно оцінюється і корисність системи:

- чи допомагає використання системи користувачеві?
- чи є укладення системи добре структурованими і докладними в потрібному ступені?
- чи задовольняє користувача швидкість роботи системи?
- чи зручний для користувача інтерфейс системи?

Користувачі-практики вимагають більшого, ніж просто високоякісна робота системи:

- система повинна бути швидкою, надійною, легкою в обігу;
- повідомлення системи повинні бути легко зрозумілими;
- система повинна прощати користувачеві його помилки.

6.5. Вибір інструментального засобу побудови ЕС

Досить часто вибір інструментальних засобів роблять за такими принципами:

- інженер знань був уже добре знайомий з цим засобом;
- обраний засіб було найбільш ефективним і доступним з числа тих, які могли бути використані на обчислювальній техніці розробника.

Тим часом, корисно поглянути на завдання вибору інструментального засобу і більш об'єктивно, розбиваючи це завдання на етапи вибору засоби та оцінювання його можливостей. При цьому можна виділити наступні питання:

- чи задовольняє даний засіб бригаду розробників по наданих їм можливостях?
- чи дозволить даний інструментальний засіб завершити розробку ЕС в задані терміни?
- чи надійно вибирається засіб?
- підтримується і чи ведеться інструментальний засіб?

Вибір між мовою програмування та мовою інженерії знань слід робити, враховуючи терміни розробки, складність розробки і досвід розробників: використання мови програмування, як правило, подовжує терміни розробки, хоча мова програмування - більш гнучкий засіб у порівнянні з мовами інженерії знань; мова інженерії знань зручний тоді, коли терміни розроблення стиснуті або досвід розробників невеликий - при цьому втрачається гнучкість, властива мовам програмування.

Крім того, засіб розробки повинен мати ознаки, зумовленими характером завдань і способом застосування системи. Наприклад, при бажанні використовувати систему для науково-дослідної діяльності, в процесі якої передбачається збір значних обсягів фактичних даних з подальшою їх статистичною обробкою, дуже корисними можуть виявитися функції СУБД - за такою ознакою, наприклад, була вибрана мова програмування СУБД Clipper при розробці ЕС для іридодіагностики ЭСИД. А перенесення центру ваги з науково-дослідної на практичну діяльність привів до вибору іншої мови програмування - мови Visual C++ for Windows (як це було зроблено при розробці системи ЭСИД -2). Необхідність швидкої розробки ЕС для прогнозування успішності учнів за результатами їх психологічного тестування призвела до вибору інструментарію у вигляді ОЕС "Универсал".

Засіб розроблення повинен бути надійним і мати підтримку з боку розробників цього засобу (певною мірою це означає і те, що воно має бути придбано на законних підставах) - в іншому випадку розробка ЕС може призупинитися не з вини її розробників.

6.6. Робота з експертами (вилучення знань)

Предметні знання для експертної системи можуть бути почерпнуті з наукової літератури, звітів, досвідчених даних та особистого досвіду. Однак головне джерело знань - це фахівець-експерт.

Досвід спільної роботи інженерів по знаннях та експертів свідчить про відому парадоксальність знань експертів: чим вище компетентність експерта, тим менш здатний він описати ті знання, які він використовує для вирішення завдань. Пояснити цей парадокс можна, якщо прийняти за основу ієрархічну модель знань експертів: розмірність семантичного простору зменшується з підвищенням рівня професіоналізму - цей висновок узгоджується і з відомим положенням когнітивної психології про те, що процес пізнання супроводжується узагальненням.

Стосовно до ЕС медичного призначення подібна ієрархічна структура може мати, наприклад, вигляд, показаний на рис.6.2.

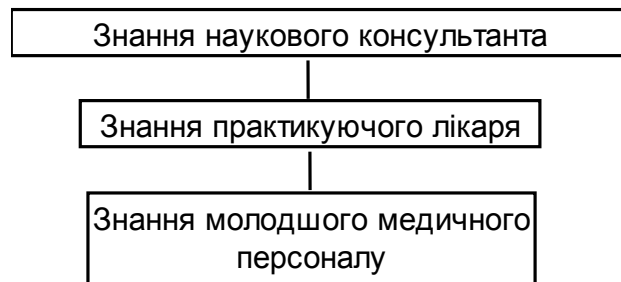


Рис. 6.2

Високий ступінь складності об'єктів, якими доводиться оперувати фахівцям «верхніх» поверхів зазначеної ієрархічної структури і становить головну проблему при співпраці експертів з інженерами знань - завдання математичного моделювання таких об'єктів надзвичайно складна і трудомістка.

Виділимо наступні складові процедури отримання інженером знань від експертів:

- спостереження (інженер по знаннях спостерігає, як експерт вирішує практичні завдання, піклуючись про те, щоб не вплинути на підхід експерта до вирішення завдання);
- обговорення (обговорення в неформальній вільній манері семантичного поля і відносин між елементами цього поля, особливостей даної задачі, типів вихідних даних і рішень для задачі, можливості розбиття задачі на підзадачі, критеріїв задовільності вирішення поставленого завдання);
- опис (одержання від експерта опису типової задачі);
- аналіз (інженер просить експерта вирішити кілька завдань, прагнучи осягнути стратегію вирішення, використовувану експертом);
- доведення системи (експерт дає інженеру завдання - той вирішує їх на

папері або на демонстраційному прототипі, намагаючись визначити працездатність системи);

- оцінювання системи (експерт перевіряє кожне правило і стратегію в цілому);
- перевірка (зіставлення рішень прикладів, отриманих системою та експертами - виробляють інші незалежні експерти).

6.7. Класифікація розробників ЕС

Розробники ЕС за кордоном діляться на три групи:

- університети;
- науково-дослідні центри;
- комерційні організації.

Зауважимо при цьому, що приблизно на такі ж групи діляться розробники і в країнах СНД, з тією лише відмінністю, що в часи існування СРСР третя група лише починала зароджуватися, а першим двом відповідали вузи і науково-дослідні інститути. У пострадянський час відсутність стабільного фінансування призвело до істотного скорочення обсягів розробок в НДІ і вузах, і лише в окремих комерційних організаціях розроблення тривають (головним чином - у спільних підприємствах).

За кордоном, і насамперед у США, багато досягнень в технології ШІ були стимульовані дослідженнями в університетах, звичайно за допомогою дисертацій. Хоча більшість великих університетів пропонують окремі курси або повні освітні програми з ШІ і ЕС, більша частина роботи в цій області була виконана всього в декількох з них: Станфорді, Карнегі-Меллоні і МТІ (Массачусетському технологічному інституті). Причому два з них, Станфорд і Карнегі-Меллон, з'явилися ініціаторами в розробці експертних систем та інженерії знань.

Причини такої неоднорідності, мабуть, полягають у наявності неординарних творчих особистостей, генеруючих наукові ідеї. Так, наприклад, в шістдесятих роках Аллан Ньюелл з Карнегі-Меллон висунув ключову ідею про представлення довготривалої пам'яті людини у вигляді послідовності правил типу "ситуація-дія", званих продукціями. Ідея Ньюелла потім була узагальнена в Станфорді. Результатом стало створення перших ЕС DENDRAL і MYCIN.

Дослідницькі центри США, що займаються розробкою ЕС - це як великі корпорації, що розташовані групою або відділом досліджень з ШІ, так і невеликі компанії, майже виключно займаються розробкою методів ШІ і додатків ЕС в певних областях. Наведемо кілька прикладів успішної діяльності таких дослідницьких центрів. Організація "Ренд Корпорейшн" (великий приватний некомерційний дослідницький центр, заснований в 1948р.), Що займається проблемами національної безпеки і соціального забезпечення, зосереджена на засобах побудови експертних систем і побудові самих ЕС - одним з напрямків такого роду є розробка мови інженерії знань ROSIE. Крім того, розроблені ЕС для юридичних міркувань в області відповідальності за випуск дефектної продукції,

що призвела до нещасних випадків або збитку для здоров'я споживачів.

Інший приклад - лабораторія ШІ в "Форд Аероспейс", що проводить дослідження в області ЕС для планування, складання розкладів і діагностики. Багато проектів, що розробляються цією лабораторією, пов'язані з програмою "Спейс Шаттл".

Нарешті, дослідженнями і розробками в галузі ШІ і ЕС займається безліч комерційних компаній, зокрема - банки та страхові компанії. Серед них можна виділити APЕХ, Syntelligence, CGI, Teknowledge, Intellicorp. На допомогу молекулярним біологам компанія Intellicorp розробила ЕС для моделювання та планування експериментів по рекомбінації ДНК, для аналізу нуклеотидних послідовностей, для управління базою біологічних даних і інших додатків генної інженерії. Компанія APЕХ розробляє ЕС для фінансової індустрії - для великих банків, компаній з кредитування, страхових компаній, бухгалтерських фірм. Її перший програмний продукт був впроваджений ще в 1983р. в різних місцях США, а більш пізні розроблення комерційно експлуатуються в даний час.

6.8. Труднощі розробки ЕС

Щоб розробити ЕС, потрібно, як вказується в літературі, «вкласти в це підприємство багато часу, грошей, енергії та ентузіазму». Якщо обрана проблема відповідає технології експертних систем (тобто має місце відповідність мети і засобів), тоді витрати на розробку окупляться багаторазово. Але навіть у цьому випадку розробка ЕС не обійдеться без напружених ситуацій. Деякі з них уникнути неможливо: штучний інтелект все ще дуже молода область. Але в ряді випадків їх можна уникнути, якщо попередньо знати про них.

Недолік ресурсів. На рис.6.3 показані ті труднощі, з якими можуть зіткнутися розробники при підготовці необхідних для роботи ресурсів.

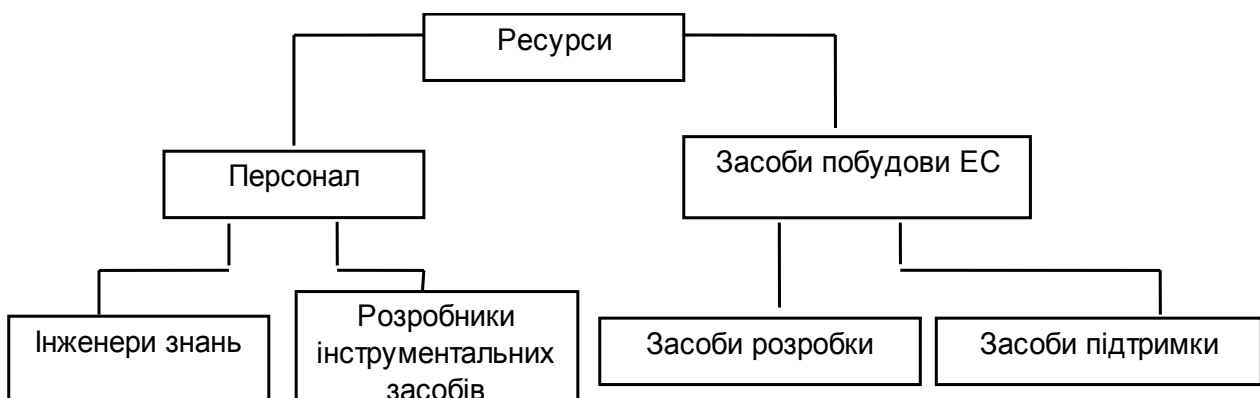


Рис. 6.3

Осіб, компетентних у проектуванні і створенні таких систем, дуже мало; ще менше коштів підтримки та мов високого рівня, які були б повністю доведені і надійні.

Є дві причини цієї нестачі ресурсів. По-перше, ЕС, як і що породила їх дис-

ципліна - штучний інтелект, - все ще нові і незнайомі більшості фахівців в області обчислювальної техніки та програмування і, отже, цим фахівцям важко розуміти і застосовувати їх. По-друге, попит на фахівців з штучного інтелекту і ЕС перевищує пропозицію - тому такі фахівці часто «створюються на ходу», з появою потреби в них, шляхом навчання на спеціальних курсах, у спеціальних школах і центрах.

Що стосується програмного забезпечення, то тут ситуація сприятливіша. Незважаючи на недостатню перевіреність і доведеність, багато високорівневі інструментальні засоби, спочатку розроблені в університетах і дослідницьких центрах, шляхом спрощення та доопрацювання перетворені в комерційні версії.

Обмеження, властиві експертним системам. Сучасним ЕС і інструментальним засобам притаманні обмеження, багато з яких поступово зникнуть по мірі розвитку досліджень в області ШІ:

- вельми складне уявлення тимчасових або просторових відносин - поки для цього типу уявлень необхідні величезні обсяги пам'яті;
- вельми складно побудова міркувань, заснованих на так званому «здоровому глузді»;
- виявляється складним навчити ЕС розпізнавати межі своєї компетентності - експертні системи демонструють досить невпевнену поведінку поблизу меж їх застосування, коли похибки даних виходять за певні межі або пропонуються завдання, відмінні від тих, для яких ЕС розроблялися;
- складно навчити ЕС працювати з помилковими або суперечливими знаннями - більшість ЕС покладаються на масу правил, але не вміють «покладатися» на основні принципи для розпізнавання невірних знань або для міркувань про протиріччях.

Інструментальні засоби для побудови ЕС також мають певні обмеження:

- нездатність виконувати набуття знань - витяг знань з експерта та втілення їх в базу знань являє собою виснажливий і повільний, важко піддається автоматизації, процес;
- недостатня здатність надати допомогу при уточненні та коригуванні БЗ - між тим, часом доводиться в кілька разів збільшувати обсяг бази знань для поліпшення якості роботи ЕС на десяток відсотків;
- недостатня універсальність і гнучкість - певні види знань, такі як тимчасові або просторові уявлення, не можуть бути ефективно представлені засобами фреймів і правил; крім того, багато мов не мають можливостями побудови потужних і зручних для користувача інтерфейсів.

Трудовитрати на розробку ЕС. Вартість розробки ЕС в істотній мірі визначається трудовитратами, тому часто замість питання вартості розглядають питання трудовитрат - одиницею виміру при цьому служать людино-роки.

Очевидно, вартість розробки залежить від її складності. Наведемо кілька прикладів для того, щоб оцінити порядок величини трудовитрат, необхідних для розробки досить складних ЕС.

Перша ЕС DENDRAL, чудова не тільки тим, що нею користуються вчені-хіміки повсюдно в США, але і тим, що стала джерелом багатьох ідей, що лежать в основі побудови ЕС, зажадала для свого створення значних зусиль - 30 людино-років.

Настільки ж трудомісткою виявилася розробка системи MACSYMA, робота над якою також була розпочата ще в середині шістдесятих років, а результати виявилися дуже високими: нею користуються повсюдно в США математики, інженери і вчені.

Слід зауважити, однак, що оцінка зусиль у людино-роках не дає точної картини необхідних трудових ресурсів і часу - 10 людино-років може означати як роботу 20 чоловік протягом півріччя, так і роботу однієї людини протягом 10 років. Обидві ці крайнощі не приведуть до успіху. Подвоєння персоналу не скоротить термін розробки вдвічі, оскільки процес створення ЕС - це послідовність кроків, тісно пов'язаних між собою безліччю зворотних зв'язків. Тому важко організувати роботу так, щоб окремі кроки виконувалися паралельно. Звичайно оптимальною називають групу розробників від 2 до 6 осіб: 2-4 людини слід використовувати в помірно важкій проблемній області; 3-5 осіб - у важкій; 4-6 людей - в дуже важкою. Недоцільно формувати групу з одного виконавця, який працює повний день - така «група» повільно набирає і насилу підтримує потрібний темп роботи. Коли розмір групи перевищує шість осіб, стає важким координувати їх діяльність.

Говорячи про час розроблення, відзначають, що при вказаних вище оптимальних розмірах групи помірно складне завдання може зажадати 0,5-1,5 року; важка - 1-3 роки; дуже важка - 3-5 років.

У середньому на розробку комерційної експертної системи йде приблизно 10 людино-років. У грошовому вираженні (на прикладі ЕС SureTrak американської фірми Primavera) вартість такої розробки близька \$ 10-11 млн.

Пастки при розробці ЕС. Говорячи про «пастки», фактично визнають, що розробка ЕС - не тільки наука, а й мистецтво. Такі пастки зустрічаються на всіх етапах розробки ЕС. Вкажемо ці пастки. При цьому, характеризуючи кожен пастку, спочатку опишемо її сутність, а потім приведемо необхідні рекомендації щодо її усунення.

Етап планування ЕС.

Вибір підходящого завдання:

1) задача настільки важка, що її не можна вирішити в рамках обмежень, що накладаються наявними ресурсами - в цьому випадку рекомендують *побудувати невелику прототипну систему* з тим, щоб з'ясувати, чи варто продовжувати повномасштабну розробку;

2) створення ЕС не зможе істотно полегшити труднощі, що зумовили створення ЕС - необхідно *заздалегідь проаналізувати очікуваний ефект* від експлуатації ЕС;

3) розв'язувана проблема вимагає непомірно великий БЗ, яку важко буде створити й експлуатувати (тривалість створення і повільність роботи) - необхідно *переглянути постановку задачі* на предмет обмеження її широти.

Ресурси, необхідні для розробки:

1) є можливість залучити значні ресурси, щоб укластися у встановлені терміни розробки ЕС - чисельність робочої групи слід вибирати, *керуючись рекомендаціями досвідчених інженерів знань*, оскільки подвоєння числа розробників зовсім не означає скорочення термінів розробки в два рази;

2) керівництво вважає, що ЕС є всього лише різновидом комп'ютерних програм, тому будь-який колектив програмістів може розробити ЕС - *необхідно залучити до розробки досвідченого інженера знань* для створення або консультування групи розробників.

Вибір інструментального засобу розробки:

1) при використанні обраного інструментального засобу виявляється його низька ефективність - вибираючи інструментальний засіб, необхідно *погодити його з властивостями розв'язуваної задачі*;

2) вибрано найбільш відомий інструментальний засіб, причому якщо завдання погано вирішується з його застосуванням, то здійснюється переформулювання завдання - вибираючи інструментальний засіб, необхідно *погодити його з властивостями розв'язуваної задачі*;

3) прийнято рішення створити ЕС на універсальній мові програмування типу Сі, Фортран і т.п. для підвищення швидкодії ЕС і для полегшення її переносимості на інші ЕОМ, що означає, однак, неприпустиме збільшення термінів розробки - *необхідно спочатку розробити ЕС на високорівневій мові побудови ЕС*, і лише потім при необхідності переходити на використання універсальних мов програмування;

4) обраний інструментальний засіб містить програмістські помилки, пошук і виправлення яких забирає багато часу - необхідно вибирати *сучасні перевірені і добре зарекомендували себе інструментальні засоби*.

Етап роботи з експертом.

Вибір експерта:

1) взаємодія з експертом вимагає значних зусиль і недостатньо ефективно - необхідно знайти, користуючись порадами та рекомендаціями, *висококваліфікованого експерта, яка вміє викласти свої знання і зацікавленого у співпраці*;

2) обраний експерт не має в своєму розпорядженні достатній час для участі в розробці - необхідно, щоб *експерт віддавав розробці не менше половини свого робочого часу*, в іншому випадку необхідно поміняти експерта.

Взаємодія з експертом:

1) експерт не відчуває впевненості, користуючись наявною базою правил для фіксації своїх знань - необхідно *піти назустріч експерту*, використовуючи його термінологію і полегшуючи механізм користування базою правил;

2) експерт формулює надто прості правила, які не гарантують високої ефективності створюваної ЕС - *необхідно допомогти експерту не уникати складних реальних ситуацій*;

3) експерт втрачає інтерес до роботи - *необхідно забезпечити регулярність контактів з експертом, дати йому можливість самому модернізувати створювану систему*;

4) експерт не знайомий з комп'ютерами і сумнівається в доцільності їх застосування - *необхідно нарощувати працездатність ЕС на робочому місці*;

5) залучено надто багато експертів, чиї думки важко поглиблено вивчити через нестачу часу - *починати розробку слід з залучення одного або двох експертів, залучаючи інших експертів на стадіях тестування і перегляду даного прототипу ЕС*.

Етап розробки ЕС.

Реалізація ЕС:

1) у процесі розробки знання експерта так переплітаються у програмі, що їх стає важко відокремити від знань для управління і побудови стратегій пошуку рішення - *необхідно попередньо ретельно відокремити предметні знання від інших видів знань*, що найлегше реалізувати з використанням правил продукції виду ЯКЦО-ТО ;

2) після тривалого періоду розробки бази знань виявляється відсутність багатьох основних понять - *необхідно постійно тестувати прийняті рішення і бути постійно готовим до перегляду концепцій і способів подання знань*;

3) при розробці ЕС виникають труднощі з реалізацією функції пояснень, оскільки використовуване інструментальний засіб не розташовує відповідною можливістю - *вибір інструментального засобу повинен бути продуманим, дуже бажано використовувати інструментальний засіб з вбудованим механізмом генерування пояснень*;

4) ЕС містить велику кількість різних специфічних правил, що ускладнює її експлуатацію - *необхідно мінімізувати кількість типів використовуваних правил*.

Тестування та оцінка ЕС:

1) у процесі тестування відповіді ЕС розчаровують як за рівнем компетентності, так і з точки зору практичної користі - *вже на етапі ідентифікації слід планувати методи оцінки системи*, що гарантують досягнення поставлених цілей;

2) користувачі скаржаться на труднощі діалогу з ЕС - *при розробці ЕС слід використовувати добре зарекомендовані методи ергономіки*;

3) при досягненні БЗ якоїсь «критичної маси» виправлення БЗ призводить до появи нових помилок - *необхідна наявність тестових завдань для перевірки несуперечності висновків ЕС, виробляючи таку перевірку кожного разу після проведених змін БЗ*.

6.9. Приклад розробки ЕС.

У п.8.1.2 наведені характеристики та перераховані основні труднощі розробки експертної системи для іридодіагностики ЭСИД-2. Тут ми, слідуючи наведеною в п.6.4 схемою поетапної розробки ЕС, наведемо технологічні характеристики процесу розробки системи ЭСИД -2.

Етап ідентифікації

Передумови розробки. Ідея розробки з'явилася після декількох років дослідної експлуатації іридодіагностичної ЕС ЭСИД (см.п.8.1.1), в процесі якої були виявлені наступні її «недоліки»:

1) система ЭСИД реалізована на мові програмування СУБД Clipper в середовищі MS DOS, що при значних розмірах баз даних, що містять інформацію про зв'язки декількох тисяч симптомів і сотень нозологічних форм, призводило до повільності роботи системи в режимі пошуку інформації - цей недолік передбачалося усунути шляхом переходу на більш сучасне середовище Windows, а також за допомогою використання мови програмування Сі;

2) мова програмування Clipper погано пристосована для операцій з графічними об'єктами типу кольорових зображень; між тим, хотілося підвищити комфортність роботи іридолога шляхом введення відеосигналу з виходу кольорової телевізійної камери в комп'ютер, з подальшим відображенням його на екрані монітора - цю можливість також передбачалося реалізувати з використанням мови Сі;

3) зазначена в п.2) особливість мови Clipper вельми ускладнювала аналіз топічних іридознаків з використанням вербально-графічних меню у вигляді відображуваних на моніторі різних схем проекційних зон райдужки ока - в ЭСИД-2 передбачалося реалізувати таку можливість, максимально наблизивши форму такого відображення до приводиться в літературі схемами проекційних зон;

4) система ЭСИД мала розвинені механізми прямого і зворотного логічного висновку, суттєво використовують елементи Байєсівської стратегії, що дозволяло кількісно оцінювати достовірність отримуваних висновків про наявність тих чи інших захворювань - це зручно для дослідницьких цілей, але незручно для практикуючих іридологів і пацієнтів, які бажають отримувати висновок лікаря у звичній вербальній формі;

5) модернізацію баз даних ЭСИД могли виробляти самі практикуючі іридологи, не звертаючись до професійних програмістам; це забезпечувалося досить простою структурою баз даних і можливістю статистичної обробки вмісту баз даних, з подальшим урахуванням результатів такої обробки при модернізації ЕС - однак ця можливість залишилася незатребуваною практикуючими іридологами, яких більше цікавила швидкість і достовірність формованого висновку лікаря, ніж дослідження, спрямовані на подальше удосконалення БЗ;

6) Система ЭСИД призначалася тільки для автономної експлуатації - до знову розроблювальної системи пред'являлася вимога можливості експлуатації як автономно, на одному комп'ютері, так і в локальній комп'ютерній мережі, в

комплексі з іншими комп'ютерними програмами медичного призначення (для проведення УЗД, аналізу ЕКГ і т. п.).

Призначення системи. Система призначалася в першу чергу для практикуючих іридологів. Тому в якості комплексного критерію якості системи прийнята корисність її для іридологів практично будь-якого рівня кваліфікації, з урахуванням відмінності вимог іридологів різних рівнів кваліфікації: для досвідченого лікаря - іридолога важливіше швидкість формування тексту висновку, тоді як для малодосвідченого - можливість швидкого пошуку в пам'яті комп'ютера достовірних даних про нозологічні форми для спостережуваних іридоознак.

Склад робочої групи. У робочу групу з розробки програми ЭСИД-2 входили 4 людини: інженер по знаннях, експерт і два програмісти (один з них - фахівець з системного програмування, інший - по функціональному програмуванню).

Тимчасові, фінансові та матеріально-технічні ресурси. Передбачалося завершити основний обсяг робіт, включаючи доведення системи до стану, придатного для дослідної експлуатації, за 9 місяців, що в кінцевому рахунку і було виконано. Оскільки програма ЭСИД-2 була складовою частиною апаратно-програмного комплексу АРМІР, що забезпечує виведення кольорового зображення райдужної оболонки на екрани монітора комп'ютера і допоміжного монітора, фінансові витрати на розробку ЭСИД -2 включали в себе і витрати на розробку підпрограми управління введенням-виведенням зображення.

Графік розробки ЭСИД-2 був складений таким чином, що підпрограма управління введенням-виведенням зображення могла бути написана паралельно з основним тілом програми, спрямованим на реалізацію обраного механізму логічного висновку.

Етап концептуалізації

Механізм логічного висновку. У силу практичної спрямованості знову розроблюваної ЕС було вирішено хоч і використовувати елементи байєсовської стратегії, однак робити це в значно меншому, в порівнянні з ЭСИД, ступеня (тільки при зворотному напрямку пошуку інформації "захворювання-ознака"). При прямому напрямку пошуку ("ознака-захворювання") вирішено було обмежитися лише підрахунком кількості ознак, "проголосували" за дане захворювання, з подальшим внутрішньосистемним (для кожної з систем організму людини) упорядкуванням (по числу "голосів") захворювань.

Таким чином, в ЭСИД-2 кодовано-числовій формі висновку надають перевагу вербальній формі, яка зручна як іридологу, так і пацієнту.

Дослідна експлуатація ЭСИД-2 згодом підтвердила правильність такого рішення. Більше того, виявилось зручним не відображати на екрані комп'ютера проміжний результат, що містить інформацію про результати «голосування» - зручніше виявилось відображати відразу вікно з автоматично формованим (при

введенні в комп'ютер інформації про спостережувані іридоознаки) висновком лікаря.

Структура висновку лікаря. Аналіз форм висновків численних прототипів системи ЭСИД-2 дозволив прийняти обґрунтоване рішення про доцільну структуру висновку лікаря - іридолога, що складається з двох компонентів:

- 1) результати експрес-оцінки конституціонального стану організму за методикою Е.С.Вельхова або В.В.Кривенко;
- 2) результати поглибленого огляду пацієнта.

У висновку можуть бути присутніми обидва компоненти або тільки один з них (залежно від умов обстеження).

При експрес-оцінці висновок містить перелік спостережуваних іридоознак і висновок про стан організму пацієнта в цілому (з градаціями «хороше», «задовільне», «незадовільне»).

При поглибленому огляді пацієнта висновок складається з двох розділів:

- карта огляду (перелік спостережуваних іридоознак);
- характеристика органів і систем організму пацієнта.

Причому другий розділ починається із загальної оцінки стану організму, і лише потім дається інформація про стан кожної системи окремо (див. приклади форми висновку в п.8.1.2).

Стан органів кожної системи дається у вигляді окремих фраз-пропозицій, що мають стандартну структуру:

генезис - найменування захворювання - ступінь (стадія) захворювання - етіологія захворювання

Наприклад: «схильність до виразкової хвороби шлунка і 12-палої кишки нервової етіології». Слово «схильність» тут використано в сенсі спадкової схильності до захворювання, а решта складові частини речення не потребують коментарів.

Інший приклад: «ознаки гастродуоденіту хронічного». Тут опущені частини «генезис» і «етіологія», але зате з'явилося слово «ознаки», яке несе в собі не стільки інформаційну, скільки етичне навантаження: по-перше, іридодіагностика як метод не претендує на постановку клінічного діагнозу, а по-друге, таким чином вдається піти від гіпердіагностики і зменшити ймовірність ятрогенії.

У результаті такої продуманості структури і форми укладення забезпечує не тільки достовірність відображення стану пацієнта, але і легко сприймається як лікарем, так і пацієнтом.

Технологія формування висновку іридолога. Завдяки плідній співпраці інженера по знаннях з експертом, обрана оптимальна стратегія обстежень, що забезпечує можливість швидкого формування більшої частини (70-80%) укладення вже в перші хвилини обстеження - досягається це тим, що в якості першого (за черговістю) оцінюваної ознаки використаний підтип конституції, що містить інформацію про генетичну схильність організму до тих чи інших захворювань.

Етап формалізації

Вибір відповідного засобу розробки ЕС. Такий вибір був здійснений спільно інженером по знаннях і програмістами - при цьому істотним моментом було добре володіння програмістами мовою програмування Сі (версія для середовища Windows).

Формування технічного завдання (ТЗ). Цьому моменту було приділено серйозну увагу, оскільки ТЗ є основним документом для програмістів, від глибини і якості якого в кінцевому рахунку залежать якість і терміни виконання роботи в цілому.

Тому в ТЗ, крім природних вказівок щодо мети розробки та основних функцій розроблюваної системи, увійшло найдокладніший опис баз знань:

- кодифікований перелік захворювань;
- ієрархічно організовані переліки ознак, із зазначенням відповідних захворювань;
- схеми проєкційних зон райдужної оболонки ока;
- ескізи вербально-графічних меню для таких ознак як зсув і сплющення зіниці.

Крім того, в ТЗ викладені основні вимоги до інтерфейсу. Вимоги до документації містили вказівки щодо структури та повноти «Керівництва користувача».

Проект ТЗ узгоджений з замовником, експертом і програмістами і виданий програмістам.

Етапи реалізації та тестування

На даних етапах розробки (інженер по знаннях, програмісти, експерт) перетворили формалізовані знання в працюючу комп'ютерну програму. При цьому здійснювався постійний контроль за виконанням програмістами вимог ТЗ.

На етапі реалізації розробки змушені були через зміну експерта замінити всю базу знань: замість структури та змісту БЗ, в основному запозичених з монографії М.С. Абрамова «Іридодіагностика з комп'ютером», були запропоновані нова структура і зміст БЗ, що представляли собою синтез інформації з більш «солідної» іридологічної літератури, запропонованої експертом інженеру по знаннях.

Крім того, функціональний програміст кілька разів змінював структуру баз даних, оптимізуючи її.

Нарешті, кілька разів змінювався інтерфейс з метою зробити його зручним для кінцевого користувача.

Неважко бачити, що етапи реалізації та тестування тісно перепліталися. Вони здійснювалися спільно розробниками і замовником, що в кінцевому рахунку призвело до задовільного рішення поставленого завдання - задовільного з точки зору як членів робочої групи, так і замовника.

Подальша дослідна експлуатація системи ЭСИД-2 підтвердила правильність

прийнятих рішень: система виявилася швидкою, надійною, легкою в експлуатації, її повідомлення легко зрозумілі, а прийнята форма укладення однаково задовільна як з точки зору лікаря будь-якої спеціалізації, так і з точки зору пацієнта.

6.10. Експертні системи та ринок

6.10.1. Плюси і мінуси розробки та придбання ЕС

Якісні результати порівняння розробки та придбання ЕС представлені в табл. 6.1.

Таблиця 6.1

Розробка ЕС	Придбання ЕС
Висока вартість	Невисока вартість
Тривалість періоду розробки	Можливість негайної експлуатації
Невеликий час освоєння ЕС	Значний час освоєння ЕС
Глибоке освоєння можливостей ЕС	Поверховість знань про можливості ЕС (наявність "секретів")
Максимальне задоволення вимог "замовника" ЕС	Неповне задоволення вимог "замовника" ЕС

На додаток до таблиці наведемо деякі кількісні характеристики. Головний недолік розробки ЕС - її висока вартість, на 3-4 порядки перевищує вартість покупної ЕС. Так, ринкова ціна готових ЕС досить високого ступеня складності орієнтовно становить \$ 500-5000, а вартість розробки ЕС - \$ 1-10 млн.

Тривалість розробки ЕС - питання взагалі особливий, оскільки часом буває важко прийняти рішення про завершеності системи. Наприклад, у 1995р. в газеті "Computer Week", # 27 (Москва), у статті Тома Давенпорта повідомлялося, що припинено розробку ЕС Xcon фірми Digital Equipment (ЕС для вибору конфігурації комп'ютерів) - однієї з найбільш зрілих і широко використовуваних комерційних ЕС. Перші версії Xcon почали експлуатуватися вже в 1980р., Оскільки розробники зрозуміли, що вони не повинні чекати того моменту, коли система буде повністю завершена.

Досвід експлуатації підтвердив їх правоту: трудомісткість як створення, так і розширення Xcon виявилася практично постійною і склала близько 4-х людино-років на рік (це відображає той факт, що обсяг знань, необхідних цій системі, зростає з постійною швидкістю). Виявилось також, що Xcon не тільки не забезпечує 100%-вої безпомилковості (відсутність помилок мало місце в 90-95% випадків), але, мабуть, буде робити помилки завжди, скільки б не розширювали її базу знань. Таким чином, розробники Xcon дійшли висновку, що побудова такої ЕС здається процесом, які не мають кінця і вимагає постійного нарощування системи протягом тривалого часу.

Що стосується часу освоєння ЕС, воно, безумовно, буде невеликим для роз-

робників, що експлуатують ЕС при вирішенні власних завдань, оскільки розробники знають її сильні і слабкі сторони краще, ніж будь-хто інший.

Крім того, ніяка документація не відображає повністю всіх особливостей складної системи - в результаті часто спостерігається ситуація, коли користувач організовує чи не цілу науково-дослідну роботу з метою з'ясування особливостей функціонування придбаної ЕС.

Нарешті, набуваючи ЕС, користувач повинен розуміти, що він може змінити базу знань, але не може змінити інтерфейс. Тому після невеликого періоду експлуатації придбаної ЕС його може чекати деяке розчарування, викликане природньою відмінністю його вимог до системи від вимог розробників.

6.10.2. Як придбати готову ЕС?

Основні критерії придбання:

- 1) корисність;
- 2) вартість;
- 3) якість;
- 4) наявність документації на продукт;
- 5) репутація розробника;
- 6) гарантійні умови;
- 7) авторський супровід.

Найпростішим комплексним критерієм є ставлення $\text{якість} / \text{вартість}$ (якість повинна бути високим при невисокій вартості).

Критеріями якості ЕС можуть бути:

- 1) відповідність ЕС цілям, для досягнення яких вона купується (за оцінкою фахівця, якому належить експлуатувати придбану ЕС, з можливим залученням для консультацій інших спеціалістів);
- 2) повнота переліку функцій ЕС;
- 3) відсутність збоїв при функціонуванні ЕС;
- 4) значимість видимих дефектів і можливість їх усунення;
- 5) повнота баз знань;
- 6) гнучкість - можливість самостійної, без залучення розробників, модернізації баз знань;
- 7) зручність експлуатації (простота інтерфейсу, наявність режиму підказок і т.п.);
- 8) якість документації ("Керівництво користувача", "Опис" та ін.)

Купуючи ЕС, обов'язково слід зібрати максимум інформації про розробників, в тому числі:

- 1) відгуки про розробників і постачальників ЕС (від організацій-користувачів);
- 2) солідність фірми-постачальника (час існування, ступінь участі в розробці, кількість і якість розробок, якість продажного і післяпродажного сервісу);
- 3) гарантійні умови, тривалість і якість авторського супроводу розробки (навчання, консультації, ремонт).

Короткі висновки

Об'єктивні передумови для розробки ЕС:

- 1) низька ступінь структурованості знань чи неповнота інформації, необхідної для прийняття рішень;
- 2) значні обсяги оброблюваної інформації в поєднанні з жорсткими вимогами до швидкості прийняття рішення;
- 3) висока відповідальність при прийнятті рішень;
- 4) стомлюваність осіб, котрі приймають рішення (ОПР);
- 5) різний рівень кваліфікації ОПР .

Умови можливості розробки ЕС :

- 1) існують висококваліфіковані експерти, які згодні співпрацювати і здатні пояснити використовувані методи;
- 2) думки експертів не суперечать один одному;
- 3) завдання, яке вирішується містить компонент, що вимагає інтелектуальних навичок;
- 4) завдання, яке вирішується не є новим;
- 5) завдання, яке вирішується не занадто важке;
- 6) завдання, яке вирішується не базується на зверненнях до здорового глузду.

Умови виправданості розробки ЕС:

- 1) є ймовірність отримати великий дохід від експлуатації ЕС;
- 2) для вирішення завдання потрібні висококваліфіковані експерти;
- 3) можливий відхід наявних експертів створює загрозу вирішення важливого завдання ;
- 4) умови розв'язання задачі настільки несприятливі, що експертні рішення повинні приймати або неспеціалісти в цій проблемній області, або рішення взагалі повинні прийматися автоматично, без участі людини.

Умови доречності розробки ЕС:

- 1) наявність специфічних властивостей;
- 2) маніпулювання символами і символічними структурами;
- 3) можливе використання евристичних методів рішення задачі;
- 4) завдання не повинно бути занадто легким;
- 5) розумна широта постановки завдання.

Основні принципи розробки ЕС :

- 1) можливість розвитку системи;
- 2) можливість переходу кількості в якість при розвитку системи.

Розрізняють такі прототипи ЕС:

- 1) демонстраційна;

- 2) дослідницька;
- 3) дослідна;
- 4) промислова;
- 5) комерційна.

Етапи розробки ЕС:

- 1) ідентифікація;
- 2) концептуалізація;
- 3) формалізація;
- 4) реалізація;
- 5) тестування.

Основні критерії придбання готової ЕС:

- 1) корисність;
- 2) вартість;
- 3) якість;
- 4) наявність документації на продукт;
- 5) репутація розробника;
- 6) гарантійні умови;
- 7) авторський супровід.

Література до Розділу 6

1. Уотермен Д. Руководство по экспертным системам. М., Мир, 1989. - 388с.
2. Экспертные системы. Принципы работы и примеры. Под ред. Р.Форсайта.- М.:Радио и связь, 1987. - 224с.
3. Продеус А.Н., Сядро Т.А., Стукалин В.А. Медицинская компьютерная информационно-справочная система: быстрое создание, эксплуатация и модернизация. Тезисы докладов на II Украинской научно-практической конференции по народной и нетрадиционной медицине. Днепропетровск, 1996г.
4. Сядро Т.А., Продеус А.Н., Стукалин В.А. Компьютерная поддержка иридолога в практической деятельности, при обучении и в научных исследованиях. Тезисы докладов на II Украинской научно-практической конференции по народной и нетрадиционной медицине. Днепропетровск, 1996г.,
5. Кричевич В.С. и др. Экспертные системы для персональных компьютеров. Методы, средства, реализации. Справочное пособие. – Минск, «Высшая школа», 1990. - 190с.
6. Лорьер Ж.-Л. Системы искусственного интеллекта. – М.,Мир, 1991. - 566с.
7. Довгялло А.М. Диалоговые системы. Современное состояние и перспективы развития. – К.,Наукова думка, 1987. - 245с.
8. Малпас Дж. Реляционный язык Пролог и его применение. – М., Наука, 1990. - 456с.
9. Гриненко Л.М., Продеус А.М., Шукевич Ю.В., Якубов С.В. Прогнозування

успішності навчання учнів за результатами психологічного тестування. – Педагогіка і психологія, 1996, №2, с.116-122; 1997, №2, с.114-123.

Контрольні запитання до Розділу 6

1. Які об'єктивні передумови для розробки ЕС?
2. Вкажіть умови можливості розробки ЕС?
3. Перерахуйте умови виправданості розробки ЕС.
4. У чому полягають умови доречності розробки ЕС?
5. Які види прототипів ЕС вам відомі?
6. Назвіть етапи розробки ЕС.
7. Вкажіть основні критерії придбання готової ЕС.

Частина 3

Приклади практичної реалізації медичних інформаційних систем

Розділ 7

Лікувально-діагностичні центри та експертні системи

Індивідуалізація процедури обстеження та лікування пацієнтів сьогодні вирішуються кількома шляхами. Один з них - підвищення технологічного рівня обслуговування, який передбачає використання найсучасніших систем зв'язку та оперативної обробки інформації, що надходить від різноманітних датчиків. Інший шлях - реалізація концепції «сімейний лікар», яка не відрізняється принциповою новизною, однак представляється вельми плідною.

7.1 Мікропроцесорні системи домашнього і віддаленого моніторингу

У медицині накопичений великий досвід створення мікропроцесорних медичних систем (кардіологія, органи дихання, кров'яний тиск, головний мозок і ін), суттєво розширили можливості діагностики, прогнозу, і вибору лікарської тактики, оцінки її ефективності, управління процесом лікування.

В основі функціонування мікропроцесорних медичних систем (ММС) лежить ідея оперативного двостороннього інформаційного обміну між системою і пацієнтом за рахунок максимального наближення мікропроцесорних систем до пацієнта, а також за рахунок побудови таких систем з використанням технологій штучного інтелекту взагалі і експертних систем зокрема.

Мініатюризація електронних приладів дозволяє реалізувати ці ідеї в різних формах, зокрема, у формі систем домашнього моніторингу.

Зокрема, відомі спроби повної автоматизації домашнього моніторингу на базі двох методів вимірів:

- вимірювання розподіленої температури тіла людини, що знаходиться у ліжку;
- вимірювання ЕКГ у ванні під час прийняття водних процедур.

У першому випадку 16 малих термометрів розташовуються в ліжку під простиралом на рівні грудей або стегна. Температура вимірюється з інтервалом 5 хвилин протягом декількох днів і заноситься в пам'ять комп'ютера і виводиться на його екран. У другому випадку 4 металевих електрода кріпляться до внутрішньої сторони ванни так, щоб 2 з них знаходилися поблизу плечей, а два інших - в області ніг (використовується електропровідність води). Електроди з'єднуються з реєстратором і комп'ютером.

Крім того, відомі успішні спроби вимірювань ЕКГ у ліжку з допомогою проводять простиралом без спеціальних електродів, а також про виміри у ліжку

інших характеристик: дихання, піт, кашель, циркуляція крові та ін Туалетне сидіння використовують для вимірювань ваги людини, калу, сечі, їх хімічного складу, параметрів шкіри стегна та ін.

Системи домашнього моніторингу спочатку створювалися з метою підвищення медичної безпеки самотньо живуть людей. Однак ця обставина, зрозуміло, ніяк не обмежує можливостей інших застосувань систем домашнього моніторингу. Одні й ті ж програмно-технічні засоби можуть бути використані для самих різних категорій людей:

- на різних тренажерах з контролем стану учня;
- для відпрацювання методик тренування спортсменів з динамічним контролем їх стану;
- для реабілітації інвалідів, у тому числі спортсменів, з контролем ефективності процесу і т.д.

Інший важливий напрямок застосування мікропроцесорних медичних систем - системи віддаленого моніторингу, створювані з використанням прогресивних інформаційних технологій:

- мережевих засобів, що розширюють можливості спілкування абонентів між собою (фахівці і пацієнти), використовуючи голос, текст, зображення;
- систем управління базами даних і базами знань, що дозволяють накопичувати, зберігати й обробляти великі обсяги інформації (за методами лікування, діагностики, лікам, пацієнтам та ін);
- інформаційно-довідкових систем різних рівнів складності, що дозволяють виробляти ефективний пошук та вилучення даних і знань з баз даних і знань;
- експертних систем, що дозволяють узагальнювати і ефективно використовувати важко формалізується досвід фахівців різних медичних профілів.

Прикладом мікропроцесорної системи віддаленого моніторингу є лікувально-діагностична система (ЛДС) з використанням так званих "електронних комбайнів" - спеціалізованих ЕОМ, що представляють собою своєрідний гібрид багатофункціонального телефону та персонального комп'ютера.

Мета створення таких ЛДС - підвищення якості лікування пацієнтів за рахунок впровадження в клінічну практику новітніх медичних та інформаційних технологій.

Призначення ЛДС:

- паралельне (швидке) обстеження груп пацієнтів на високому професійному рівні при мінімальній кількості медичного та технічного персоналу;
- архівування (тривале зберігання) інформації у вигляді пульсограм, температури тіла, ритму серця, ЕКГ, тиску крові та ін, отриманих засобами віддаленого моніторингу, для подальшого аналізу;
- медична, в тому числі психологічна, підтримка пацієнтів з використанням засобів віддаленого впливу, зокрема, за рахунок комплектування електронних комбайнів (ЕК) елементами мультимедійної технології.

Загальні завдання, які вирішуються з використанням такої ЛДС:

1. Профілактика, діагностика, лікування, реабілітація пацієнтів на всіх рівнях медичної практики:

- госпіталь (широкий набір методів лікування різних хвороб);
- спеціалізовані медичні центри, в першу чергу, для реабілітації інвалідів;
- медичне обслуговування пацієнтів в домашніх умовах.

2. Безпосередня допомога пацієнтові в екстремальних умовах (сигнал тривоги, рекомендації в критичних ситуаціях).

3. Розширення можливостей реабілітації людини після важких хвороб, у тому числі за рахунок зменшення ізольованості пацієнта від суспільного, культурного, спортивного життя (використання для діалогу пацієнта з електронним комбайном звуку, тексту, зображень не тільки розширює можливості спілкування із зовнішнім світом, а й зміцнює психологічну стійкість пацієнта, а наявність в ЕЛ криптографічних функцій і захист інформації від несанкціонованого доступу підвищують довіру пацієнта до ЛДС).

4. Можливість для пацієнтів з порушеннями функцій органів почуттів спілкуватися з інформаційними центрами і один з одним.

Розглянемо тепер принципи побудови і функціонування ЛДС з віддаленим моніторингом.

Структура ЛДС наведена на рис. 7.1. За ступенем глибини і повноти медичного обслуговування ЛДС являє собою трирівневу систему (Центр, госпіталь, домашні умови).

Нижній рівень (безпосередній моніторинг хворого в домашніх умовах) призначений для збору інформації та постійного відстеження стану пацієнта, а також для надання термінової медичної допомоги у критичних ситуаціях. Зберігання та первинна обробка інформації відбувається в ЕК, який пов'язаний через модем з комп'ютерними центрами відповідних лікувальних установ. Датчики, що вимірюють необхідні фізіологічні параметри пацієнта, через підсилювачі та фільтри (придушення перешкод) пов'язані з аналого-цифровим перетворювачем (АЦП), який переводить сигнал в цифрову форму для введення в ЕК. Необхідна для пацієнта інформація у вигляді тексту, образу (зображення) висвічується на табло ЕК, а звукова інформація подається на гучномовець ЕК. Спрощена клавіатура ЕК допомагає пацієнтові вести діалог з комп'ютерними центрами і отримувати рекомендації про лікувальних процедурах (стимулятор). Настроювання ЕК на конкретного пацієнта і облік його індивідуальних особливостей здійснюється записом відповідної інформації в пам'ять ЕК з комп'ютерних центрів. Також можлива заміна окремих чипів в пристрої ЕК.

Крім того, кожен пацієнт має магнітну картку, де записані основні дані про пацієнта з історії хвороби. Така картка допоможе лікарю швидше надати допомогу будь-якому пацієнту, який опинився в екстремальній ситуації.

На рівнях госпіталю та Центру також використовуються аналогічні системи моніторингу, але в більш розширеному варіанті. ЛДС госпіталю та Центру об'єднує моніторинг пацієнта і комп'ютерні центри в єдину локальну мережу.



Рис. 7.1

У комп'ютерних центрах зберігаються БД і БЗ з хвороб, пацієнтам, методам лікування і т.д. Доступ до БД виконується через електронні комбайни, які грають роль інтелектуальних терміналів. Завданням ЕК в цьому випадку є програмно-технічна підтримка діалогу пацієнта з сервером комп'ютерного центру, істотно використовує технології систем штучного інтелекту і, зокрема, експертних систем. Так, наприклад, допускається, що пацієнт може ставити питання комп'ютера некваліфіковано, тому останній, у свою чергу, може ставити уточнюючі і навідні запитання.

Крім того, передбачається спілкування ЕК з пацієнтами, що страждають порушеннями зору, слуху, мови.

Враховуючи, що БД і БЗ може користуватися не тільки лікар, а й пацієнт (в обмеженому обсязі), передбачається реалізація інформаційно-довідкових систем з використанням технології "гіпертекст", що полегшує доступ до БД і БЗ з тематики, рівню кваліфікації і т.д.

Блок-схема електронного комбайна представлена на рис.7.2.

Блок 1 центрального процесора виконує програму, що знаходиться в ОЗП і ПЗП (блоки 2 і 3, відповідно), в процесі виконання якої через порти введення-

виведення 6 читає і записує дані в ОЗП і ПЗП, індикує інформацію, виводячи її в блок 7 відображення інформації, приймає аналогові сигнали з блоку 8 зв'язку з телефонною лінією через блок 9 перетворення форми інформації (ЦАП і АЦП). Через фіксовані проміжки часу лічильник 5 видає сигнали переривання, що викликає переклад блоку 1 на виконання перериває програми.

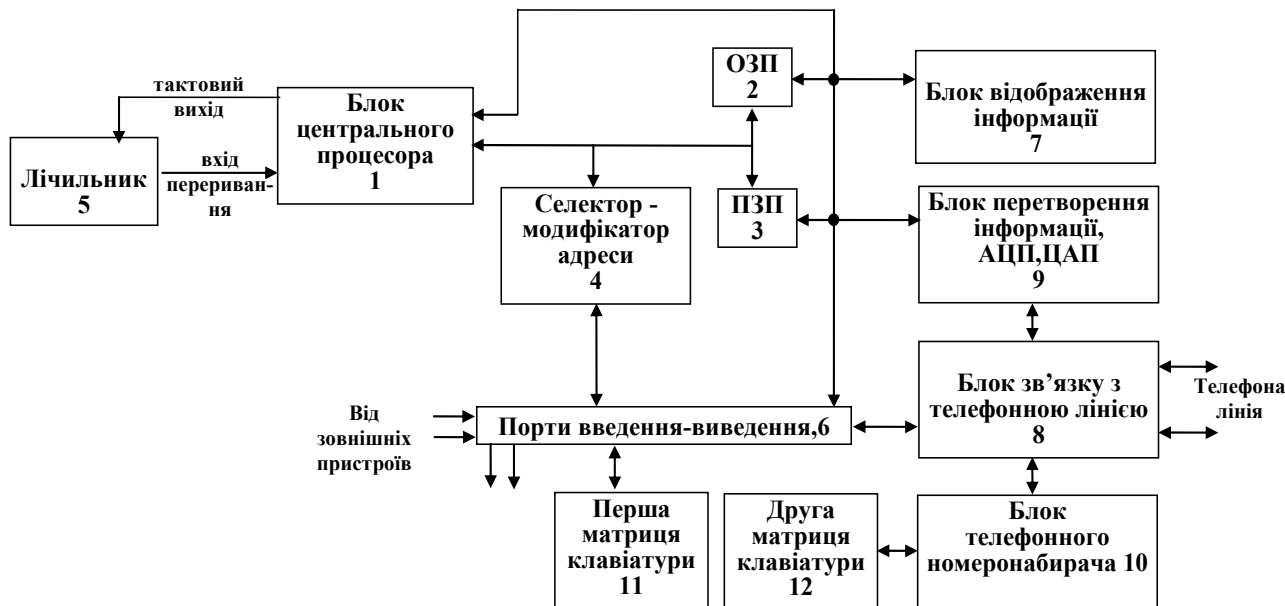


Рис. 7.2

Звернення блоку 1 до запам'ятовуючим вузлам 2 і 3 проводиться за допомогою селектора-модифікатора адреси 4. При цьому весь адресний діапазон блоку 1 розбитий на дві рівні частини, одна з яких відповідає ОЗП, а інша - ПЗП.

Інформаційний обсяг ПЗП в два рази перевищує половину простору, прямо адресується блоком 1. Програми роботи пристрою розміщені тільки в молодшій половині ПЗП. У старшій його частини розміщуються дані, які повинні зберігатися при відключенні живлення в мережі.

Введення інформації з матриці клавіатури 11 проводиться шляхом її програмного опитування. Набір номера абонента телефонної мережі може здійснюватися двома способами: за допомогою матриць клавіатури 11 або 12. У першому випадку блок центрального процесора отримує набирається номер через порти введення-виведення, через які управляє імпульсним і розмовною ключами. У другому випадку, який номер сприймається блоком 10 телефонного номеронабирача, який і керує імпульсним і розмовними ключами.

При реалізації функцій автовідповідача мовна інформація, що вводиться з мікрофону або з телефонної лінії, через блок 9 перетворення форми інформації записується блоком центрального процесора 1 в ПЗП, а інформація, що зберігається в ПЗП, через вузол 9 передається в телефонну лінію або на гучномовець.

Експлуатація електронного комбайна, з точки зору непрофесійного користувача, виглядає досить просто. Користувач задає на клавіатурі пристрою потрібну йому операцію, вибираючи її з набору меню, висвічується на індикаторі

комбайна, і операція тут же починає виконуватися. Якщо це телефонна функція, вона виконується без подальшого втручання користувача. Якщо це комуні-каційна функція (запит інформації з індикацією її на екрані або записом у пам'ять приладу, обмін повідомленнями за принципом електронної пошти), вона може виконуватися без подальшого втручання користувача або в діалоговому режимі (дистанційне виконання програм, доступ до віддалених баз даних і знань). У будь-якому випадку користувач може вважати, що всі його потреби задовольняються його власним ЕК, і йому немає ніякої потреби розбиратися в складній організації системи.

7.2. Лікувально-діагностичні системи з комплексуванням методів традиційної і нетрадиційної медицини

Концепція домашнього моніторингу тісно пов'язана з концепцією "сімейного лікаря" - обидві концепції спрямовані на підвищення якості медичного обслуговування за рахунок "наближення" медичної сфери обслуговування до пацієнта, підвищення уваги до потреб пацієнтів, усунення тієї знеосібки і стандартності підходу до обстежень, які дуже типові для сучасної поліклінічної системи. Багатогодинні черги - явище досить типове для поліклінік - негативно впливають на здоров'я не тільки пацієнтів, але і обслуговуючого персоналу поліклінік.

Російськими медичними фахівцями запропонована реалізація концепції "сімейний лікар" у вигляді трирівневої ієрархічної системи: сімейний лікар - спеціалізовані діагностичні центри - спеціалізовані лікувально-реабілітаційні центри.

У відповідності з цією концепцією, для роботи на рівні "сімейний лікар" у Росії необхідно 300 тисяч сімейних лікарів. За кожним з них закріплюється 100-150 сімей, в день відвідується 4-6 сімей. У цьому випадку кожному пацієнту приділяється 20-30 хвилин замість 3-5 хвилин в сучасних поліклініках.

Решта 350 тисяч нинішнього корпусу лікарів Росії повинні бути спрямовані в діагностичні центри та великі клініки. При цьому діагностичні центри, поперше, повинні бути спеціалізованими - кардіологічні, урологічні, гастроентерологічні і т.п., а по-друге, повинні бути оснащені сучасними апаратно-програмними засобами.

Досить подібною точки зору дотримуються фахівці багатьох сучасних медичних центрів України, що поєднують методи традиційної та нетрадиційної медицини. З тією лише відмінністю, що "сімейним лікарем", на їх думку, не може бути фахівець-одиначка - таким лікарем повинен бути або фахівець, тісно пов'язаний з медичним центром експрес-діагностики, або навіть весь медичний центр ("колективний сімейний лікар").

Пояснюється це тим, що одному фахівцю не під силу придбання сучасних складних і дорогих, часто комп'ютеризованих і портативних, засобів діагностики і лікування - витрати на придбання комплексів апаратури та програмного забезпечення під силу переважно медичним організаціям.

У цьому зв'язку розглянемо кілька підходів до впровадження прогресивних інформаційних технологій в медичних центрах, поєднують методи традиційної та нетрадиційної медицини при проведенні діагностичних, лікувальних і реабілітаційних заходів. Загальним для всіх розглянутих підходів є істотне використання ПЕОМ при реалізації технологічних процесів. Розрізняються підходи (аналіз цих відмінностей як раз і становить найбільший інтерес) багато в чому: точкою зору на вибір раціональної форми (технології) взаємодії пацієнта і лікаря (або медичного закладу), вибором поєднань методів традиційної і нетрадиційної медицини, вибором технічних рішень для реалізації заданих технологічних процесів.

Найбільшу увагу ми приділимо технічній (інформаційно-технологічній) стороні питання, не залишаючи, зрозуміло, осторонь і медичний аспект проблеми.

7.2.1. Технологія лікувально-діагностичного та реабілітаційного центру "ЛИДИР"

7.2.1.1. Технічне рішення на комп'ютерній платформі Macintosh

Дане технічне рішення являє собою можливий варіант реалізації технології "сімейний лікар", охоплюючи такі аспекти:

- 1) принципи загальносистемної організації в частині комп'ютерної підтримки діагностичної та лікувальної роботи фахівців;
- 2) номенклатуру устаткування (комп'ютери, адаптери, мережева підтримка);
- 3) склад і функції програмного забезпечення.

Рішення в цілому побудовано на комп'ютерній платформі Macintosh фірми Apple Computer з використанням сертифікованого програмного забезпечення для медичних установ, розробленого провідними програмістськими фірмами США. Необхідні в ряді випадків пристрою сполучення комп'ютерів з медичною апаратурою розроблені і сертифіковані спеціально для комп'ютерної платформи Macintosh американською фірмою National Instruments.

Концептуально рішення є комплексним, модульним, нарощуються (без зміни архітектури допускається виключення або введення одного або декількох АРМ'ів) і складається з двох частин - рішення для стаціонарної частини лікувально-діагностичного центру та рішення для його транспортного блоку, що встановлюється на мікроавтобусі. Стаціонарна і мобільна частини об'єднані за допомогою системи обміну інформацією, що використовує стандартний радіо-телефонний канал зв'язку (рис.7.3).

Стаціонарна частина включає в себе тринадцять автоматизованих робочих місць (АРМ) для фахівців центру: реєстратора, загальної діагностики, експрес-діагностики, спеціальної діагностики, додаткової діагностики, терапевтичної селекції, терапевтичного моніторингу та постановки діагнозу, лаборанта (лабораторний блок), рентгенолога, фармацевта (аптечний блок), трьох лікуючих лікарів - лікувального блоку, методів нетрадиційного лікування, гінеколога.

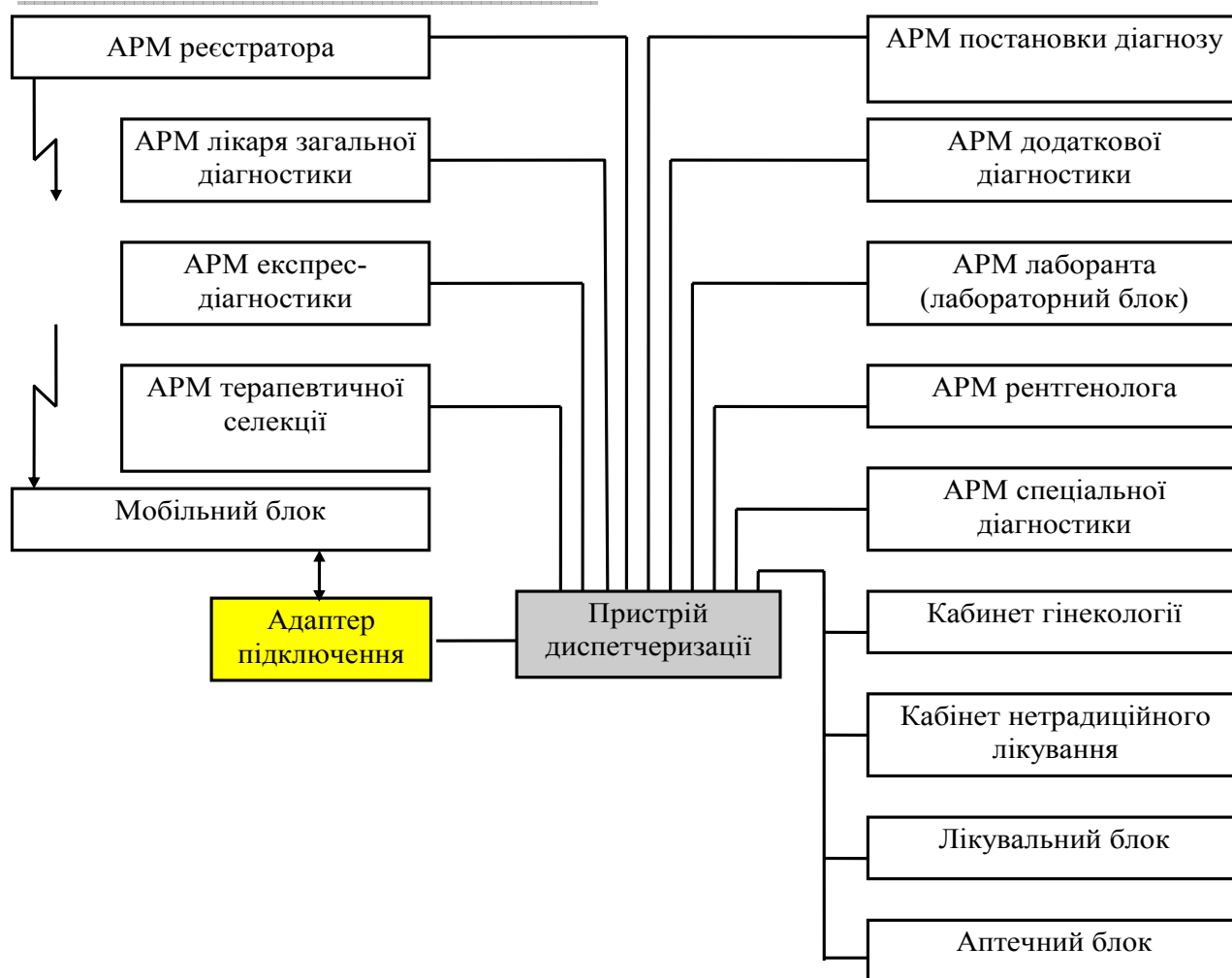


Рис. 7.3

Мобільна частина містить один АРМ лікаря, надаючи спеціалісту можливість загального діагностування, ведення карти пацієнта і оперативного обміну необхідною інформацією з базами даних знань, що зберігаються в комп'ютерах стаціонарної частини центру.

Основним структуруючим компонентом рішення є локальна комп'ютерна мережа (ЛКМ) типу Ethernet з топологією "зірка". Вибір даного типу ЛКМ та її топології обумовлений необхідністю забезпечення високошвидкісного обміну значними за обсягом масивами інформації (кольорові зображення, дані обстежень і т.п.) між кабінетами центру, а також між стаціонарною та мобільною частинами центру. Додатковим аргументом при виборі ЛКМ для центру є високий ступінь відмовостійкості. Диспетчеризацію роботи всіх користувачів ЛКМ забезпечує спеціальний контролер. З його допомогою можна здійснити спостереження за режимами роботи користувачів в мережі, встановити правила доступу до даних, швидко локалізувати місце виникнення збійних ситуацій.

Стаціонарна і мобільна частини пов'язані модемним зв'язком із застосуванням для бездротової передачі стандартного радіотелефону. Для забезпечення обміну великими масивами даних передбачається можливість прямого підключення

чення комп'ютера мобільної частини до ЛКМ Ethernet. Тим самим забезпечується також роздруківка даних, накопичених в мобільній частині, на високошвидкісному лазерному принтері стаціонарної частини.

Розглянемо конфігурацію АРМ'ів стаціонарної частини. Для більшості АРМ'ів обрана стандартна конфігурація: комп'ютер класу AV (Audio-Vision, з відеовходом, накопичувач на жорсткому диску ємністю, пристрій читання даних з лазерних дисків), дисплей, портативний принтер.

Вибір класу AV обумовлений необхідністю накопичення, обробки та високоякісного відображення даних про пацієнта, тобто текстової, візуальної, звукової та іншої інформації. При цьому значне місце займає саме візуально-графічна, а в ряді випадків - звукова і відеоінформація. Наявність на робочому місці портативного принтера дозволяє оперативно вивести текстову і графічну інформацію, необхідну для лікаря. У складі АРМ постановки діагнозу є кольоровий принтер для виведення на друк кольорових зображень, отриманих, наприклад, від тепловізора або апарату УЗД. Виведення зкомплексованих даних в повній стандартній формі забезпечується за допомогою високошвидкісного лазерного принтера, встановленого в реєстратурі. У лабораторному блоці встановлюється спрощена конфігурація - АРМ на основі комп'ютера без відеовходу, оскільки для роботи лаборантів він не потрібний.

На робочому місці реєстратора передбачається, на додаток до стандартної конфігурації, установка сканера, за допомогою якого забезпечується введення в систему даних про пацієнта, принесених їм від сторонніх лікувальних установ, рентгенівських знімків і т.п. Для зберігання великих масивів архівної інформації встановлюється пристрій зовнішньої пам'яті на змінних магніто - оптичних дисках.

В якості додаткового засобу підвищення якості обслуговування і діагностування пацієнтів на робочому місці лікаря іридіодіагностики може використовуватися цифровий фотоапарат для введення знятих знімків в комп'ютер для документування та подальшого аналізу.

До складу обладнання мобільної частини входять: комп'ютер типу notebook з вбудованим модемом, портативний принтер, пристрій сполучення з локальною мережею стаціонарної частини. Така конфігурація за своїми можливостями мало поступається стандартній і може забезпечити практично всі можливості введення, накопичення та первинної обробки даних про пацієнта.

За допомогою модему і спеціального програмного забезпечення комп'ютер мобільної частини може передати в стаціонарну частину для обробки будь-яку збережену в ньому інформацію, а також отримати за запитом результати цієї обробки або необхідні архівні дані.

Програмне забезпечення центру "ЛИДИР" доцільно будувати за функціонально-модульним принципом, з такими компонентами ПЗ:

- ПЗ комплексного призначення;
- ПЗ спеціального призначення (у складі АРМ'ів фахівців);
- ПЗ для обміну інформацією між стаціонарного та мобільного частинами.

До складу комплексного ПО повинні входити наступні програми:

- програма збору, обробки та аналізу загальної та спеціальної медичної інформації про пацієнта, включаючи графіку і відео;
- програма менеджменту медичного закладу.

До складу ПЗ АРМ фахівців входять програми, що забезпечують оперативну обробку інформації. У складі ПО зв'язку є програма дистанційного доступу до даних.

Істотним незручністю технічного рішення на комп'ютерній платформі Macintosh є та обставина, що ПЗ в цілому підлягає локалізації (перекладу та адаптації до умов країни-користувача). Враховуючи значний обсяг використаного ПЗ, а також високу вартість локалізації ПЗ, неважко прийти до висновку про економічну недоцільність такого технічного рішення.

7.2.1.2. Технічне рішення на базі ІВМ-сумісних комп'ютерів

Дане технічне рішення є альтернативним розглянутому вище (на комп'ютерній платформі Macintosh) і базується на наступних передумовах:

- рішення на платформі ІВМ-сумісних комп'ютерів представляється більш привабливим з причини більш низької вартості апаратного та програмного забезпечення, широкої поширеності цього забезпечення, "прозорості" його для користувача (можливість самостійного налаштування апаратно-програмного забезпечення);
- медичний центр повинен містити невелику (не більше 4 - 7) кількість АРМ'ів фахівців - в іншому випадку важко реалізувати індивідуальний підхід до кожного пацієнта, його швидке і одночасно високоякісне обстеження;
- при виникненні труднощів з організацією мобільної частини центру в повному обсязі таку можна реалізувати в усіченою формі шляхом транспортування пацієнта з дому (або з місця роботи) в медичний центр і назад транспортними засобами медичного центру.

Згідно з даним рішенням, ЛДЦ являє собою невеликий лікувально-діагностичний комплекс (діагностичні обстеження за трьома основними напрямками: іридодіагностика, УЗД, ЕКГ), обладнаний високоякісної медичної діагностичної апаратурою, замкнутої на локальну комп'ютерну мережу (тимчасову мережу Microsoft) з оригінальним (розробленим інженерним складом співробітників ЛИДИР) програмним забезпеченням у вигляді інформаційно - довідкових і експертних систем (рис.7.4).

При цьому в ЛДЦ вирішуються 7 завдань:

- 1) прийом відвідувачів;
- 2) реєстрація відвідувачів;
- 3) іридодіагностика;
- 4) УЗД;
- 5) ЕКГ - дослідження;

- 6) терапевтична селекція;
- 7) моніторинг (динамічне спостереження).

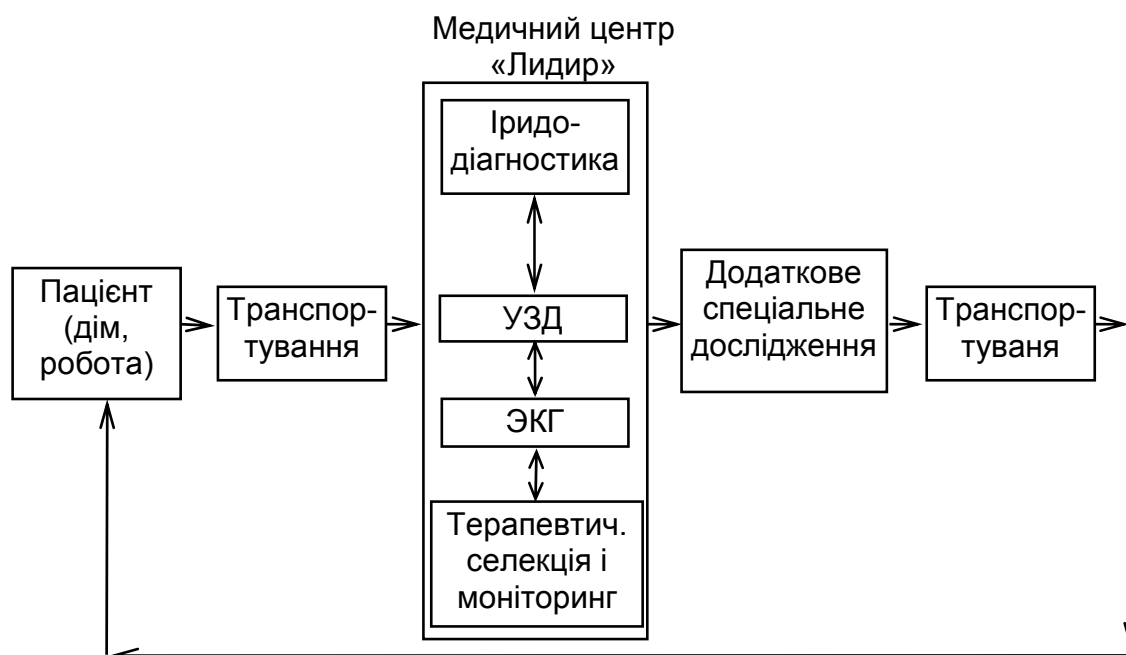


Рис. 7.4

Кількість АРМ'ів дорівнює 4, оскільки первинний прийом та реєстрація пацієнтів можуть проводитися на будь-якому АРМ'і, а терапевтична селекція і моніторинг реалізуються на одному АРМ'і.

Такий вибір конфігурації центру обумовлений прийнятою на озброєння головною концепцією діагностичних досліджень, яка складалася в проведенні експрес-діагностики для швидкого виявлення основних причин неблагополуччя в організмі людини. Ця концепція дозволяє мінімізувати обсяг спеціальних обстежень, аналізів та консультацій, і тим самим скоротити втрати часу на збір необхідних даних, знизити фінансові втрати пацієнтів.

Результатом експрес-діагностики є попередній діагноз і рекомендації по додатковому обстеженню у спеціалізованих медичних центрах і клініках (з метою уточнення діагнозу).

Всі види експрес-діагностики виробляються на базі стандартної медичної апаратури з організацією комп'ютерної програмно-апаратної підтримки, що забезпечує високий технологічний рівень медичних досліджень.

Програмне забезпечення центру побудовано за функціонально-модульним принципом з наступними компонентами ПЗ:

- ПЗ загального призначення;
- ПЗ АРМ'ів фахівців.

ПЗ загального призначення представляє собою оболонку у вигляді спеціалізованого мережевого пакету програм "MEDISERV" з наступними основними можливостями:

- ведення картотеки пацієнтів;
- автоматизоване формування текстів висновків, призначень і рекомендацій із застосуванням інформаційно-довідкової системи, яка може коригуватися і доповнюватися за бажанням користувача;
- роздруківку результатів огляду, включаючи дані інструментальних досліджень (УЗД- і іридо- зображення, криві ЕКГ), на принтері;
- запис результатів огляду в базу даних (архів) на жорсткому магнітному диску;
- перегляд, друк, сортування, видалення архівних записів.

ПЗ АРМ'ів фахівців являє собою програмні модулі, що ініціюються з оболонки "MEDISERV" та дублюючі всі її функції (крім ведення картотеки пацієнтів), а також реалізують:

- введення в пам'ять комп'ютера і автоматизований аналіз даних інструментальних досліджень (УЗД - зображень, ЕКГ, знімків райдужної оболонки очей);
- порівняння поточних і архівних результатів огляду пацієнтів з метою спостереження за динамікою патологічних процесів. Деталізуємо склад і функції програмно-апаратного забезпечення кожного з АРМ'ів.

Програмно-апаратна підтримка іридодіагностики полягає у виконанні таких функцій:

- 1) введення в комп'ютер кольорових зображень райдужних оболонок очей;
- 2) архівування в базі даних отриманих зображень;
- 3) автоматизація складання тексту висновку іридолога з використанням спеціальної інформаційно-довідкової системи ЭСИД-2.

При іридологічних дослідженнях використовується апаратно-програмний комплекс АРМІР, до складу якого входять:

- 1) щілинна лампа (типу ЩЛ-2Б), поєднана з фотокамерою;
- 2) комп'ютерна програма іридодіагностики ЭСИД-2, що забезпечує:
 - а) управління виводом зображень райдужної оболонки на екран комп'ютера або допоміжного монітора;
 - б) управління процесом запису зображень райдужної оболонки на магнітний носій, з можливістю подальшого відтворення цих записів на екрані комп'ютера;
 - в) швидкий набір і редагування тексту висновку, з використанням для цієї мети спеціальної інформаційно-довідкової системи.

Можливості програмно-апаратної підтримки УЗД та ЕКГ в чому схожі з такими для іридодіагностики, з тим лише природним відзнакою, що введення в комп'ютер інформації проводиться з виходів стандартної УЗД-апаратури та ЕКГ - апаратури.

Особливістю програмного забезпечення для УЗД-досліджень є можливість розцвічування УЗД-зображень (на моніторі комп'ютера) для підвищення ефективності візуального аналізу малоконтрастних УЗД-зображень (лише дорогі

моделі УЗД-апаратури володіють такою можливістю). Інший безумовно корисною можливістю є роздруківка УЗД-зображень на звичайному папері.

Особливістю програмного забезпечення електрокардіографічних досліджень (по 12 стандартним відведенням) є можливість виведення зображень ЕКГ - сигналів на екран комп'ютера, з подальшою "ручний" маркуванням цих зображень. Це дає можливість автоматизувати вимірювання параметрів ЕКГ - сигналів (результати вимірювань відображаються в спеціальній таблиці) в тих складних випадках, коли "інтелектуальна" ЕКГ - апаратура ставить явно помилковий діагноз через вплив перешкод на результати вимірювань або через недосконалість використовуваних алгоритмів вимірів (відсоток помилкових діагнозів при цьому вельми великий і може досягати 40-50%). Нарешті, як і при УЗД, можлива роздруківка кривих ЕКГ на звичайному папері.

При терапевтичній селекції та моніторингу використовуються практично всі описані вище можливості, надані пакетом програм "MEDISERV".

7.3. Технологія багатофункціонального комп'ютерного лікувально-діагностичного комплексу "АМЕКС"

Багаторічна експлуатація Київського багатофункціонального комп'ютерного лікувально-діагностичного комплексу (МКЛДК або, скорочено, ЛДК) "АМЕКС" показала досить високу ефективність реалізованої в ЛДК технології в рамках окремих медичних центрів (кабінетів), що дозволяє рекомендувати поширення і реалізацію даної технології в масштабах поліклінік, лікарень , санаторіїв, госпіталів та інших лікувально-діагностичних утворень системи медичного забезпечення (рис.7.5).



Рис. 7.5

Структурно-функціональна схема цієї технології наведена на рис.7.6.

Як слід з рис.7.6, комп'ютерна підтримка має багатофункціональний характер і реалізована у вигляді розподіленої експертної системи для вирішення завдань діагностики та управління лікуванням.

Спрямованість діагностичних заходів розкрита з використанням структурної схеми 7.7, де кожному з використовуваних методів діагностики поставлені у відповідність завдання, які вирішуються з використанням експертної системи.



Рис. 7.6

Короткі висновки

Індивідуалізація процедури обстеження та лікування пацієнтів сьогодні вирішуються кількома шляхами. Один з них - підвищення технологічного рівня обслуговування. Інший шлях - реалізація концепції «сімейний лікар».

Підвищення технологічного рівня обслуговування досягається використанням мікропроцесорних систем домашнього і віддаленого моніторингу, що дозволяють надати безпосередню допомогу пацієнтові в екстремальних умовах, розширити можливості реабілітації людини після важких хвороб.

Концепція домашнього моніторингу тісно пов'язана з концепцією "сімейного лікаря" - обидві концепції спрямовані на підвищення якості медичного обслуговування за рахунок "наближення" медичної сфери обслуговування до пацієнта, підвищення уваги до потреб пацієнтів, усунення тієї знеосібки і стандартності підходу до обстежень, які дуже типові для сучасної поліклінічної системи.



Рис. 7.7

Одна з форм реалізації концепції "сімейний лікар" - створення невеликих мобільних лікувально-діагностичних центрів (ЛДЦ), оснащених, крім медичного обладнання, сучасними засобами зв'язку та комп'ютерної обробки інформації. Це означає, в свою чергу, необхідність розробки відповідного ПЗ, яке забезпечує збір та організацію інформації про пацієнтів, оперативне формування і видачу висновків рекомендацій. Складовою частиною такого ПЗ є діагностичні ЕС, що дозволяють оптимізувати рішення поставленого завдання.

Література до Розділу 7

1. Микрокомпьютерные медицинские системы. Проектирование и применение, под ред. У.Томпкинса, Дж.Уэбстера. – М., "Мир", 1983.
2. Tatsuo Togand and others, Attempts of Fully Automated Home Health Monitoring Systems, Medical & Biological engineering & Computing, Vol. 34, Supplement 1, Part 1, 1996.
3. Juri Taskinen and others, Recording Methods for Home Monitoring, Medical & Biological engineering & Computing, Vol 34, Supplement 1, Part 1, 1996.

4. Лечебно-диагностическая система (ЛДС) с применением информационных технологий. Техничко-экономическое обоснование. – НАНУ, институт кибернетики им. В.М. Глушкова, 1998г.

5. Комплексное программно-техническое решение компьютерной поддержки центра "ЛИДИР". Пояснительная записка. – К., ГНВП "Аргус", 1995, 89 с.

6. Федоров С. Почему мы боеем и не излечиваемся? – "Известия" от 8.12.94г

Контрольні запитання до Розділу 7

1. Наведіть приклади відомих вам мікропроцесорних систем домашнього і віддаленого моніторингу?

2. Чому мікропроцесорні системи домашнього і віддаленого моніторингу дозволяють підвищити якість медичного обслуговування населення?

3. Чому прийнято вважати, що концепція домашнього моніторингу тісно пов'язана з концепцією "сімейного лікаря"?

4. Що собою представляє концепція "сімейного лікаря"?

5. Чи означає концепція "сімейного лікаря" те, що дільничний лікар повинен ставати власником дорогої медичної апаратури та іншого численного різноманітного обладнання?

6. Чому реалізація концепції "сімейний лікар" у вигляді створення невеликих мобільних лікувально-діагностичних центрів вважається раціональним вирішенням поставленого завдання?

Розділ 8

Експертні та інформаційно - довідкові системи

Рефлексодіагностика і іридодіагностика - методи, до певної міри об'єднуючі медичні культури Заходу і Сходу. Поява іридодіагностики, наприклад, пов'язують з ім'ям угорського лікаря Ігнаца Пекцелі (1826-1911). Своїми роботами в галузі теоретичної та практичної іридодіагностики відомі також фахівці США, Німеччини, Австралії, Росії, України (Б. Дженсен, М. Джонс, Б. Марч, Г. Озарко, Н.В. Вельховер, Ф.Н. Ромашов, У. Ф. Алексеев, Н.Б. Шульпина, В.В. Кривенко, Г.С. Лісовенко, Г.П. Потебня, Т.А. Сядро). Навпаки, методи сучасної електропунктурної рефлексодіагностики прийшли в медицину з Сходу (методи Р. Фолля і Й. Накатані). У цьому розділі розглянуто кілька експертних та інформаційно-довідкових систем для іридодіагностики і рефлексодіагностики, розгляд яких корисно, оскільки дозволяє скласти уявлення про реальні труднощі, з якими стикаються розробники і користувачі таких систем.

8.1. Іридодіагностичні комп'ютерні комплекси та програми

До теперішнього часу найбільш розвинені два основних напрямки автоматизації праці іридолог з використанням комп'ютерних засобів:

- 1) комп'ютерні програми;
- 2) програмно-апаратні комплекси.

У першому випадку мова йде про комп'ютер і комп'ютерній програмі, за допомогою яких іридолог здійснює збір та обробку даних з наступним автоматизованим формуванням уявлення. У другому випадку для вирішення тих же завдань, крім комп'ютера та комп'ютерної програми, використовуються ще й допоміжні технічні пристрої, що забезпечують введення зображення райдужки в комп'ютер з наступними обробкою і зберіганням цього зображення.

Порівнюючи між собою вказані два напрямки, відзначимо, що потенційні можливості програмно-апаратних комплексів вище, однак і вартість, в порівнянні з "чисто" програмними засобами, значно вище.

Що стосується програмного забезпечення для іридодіагностики, то його функціональні можливості, природно, головним чином визначаються фахівцями - іридолог, виходячи з розв'язуваних завдань (завдання практичної діагностики, навчання діагностуванню, науково-дослідні завдання). Разом з тим, можливості програмного забезпечення істотно залежать і від тих, хто займається безпосередньо його розробкою - в першу чергу, інженерів по знаннях, що визначають структуру і обсяг баз знань, механізми обробки даних у процесі діагностування. Нарешті, якість кінцевого програмного продукту і час його розробки істотно залежать від кваліфікації системних і функціональних програмістів. Тому порівняння готових програмних продуктів з найбільш зручному і потужному крите-

рієм "якість / ціна" - справа непроста, вимагає зваженого врахування багатьох додаткових показників.

Розглянемо характеристики деяких досить відомих засобів комп'ютерної підтримки іридолог, що дозволить скласти досить повне уявлення про їх достоїнства, недоліки і шляхи розвитку автоматизації іридодіагностики.

У 1985 р. в лабораторії проф. Є.С. Вельховера Університету дружби народів ім. П. Лумумби розроблено автоматизована система для іридодіагностичних досліджень. До складу системи входили відеокамера, апаратно-програмні засоби введення відеосигналу в комп'ютер, комп'ютер зі спеціальним програмним забезпеченням, допоміжний монітор.

Функції програмного забезпечення цієї системи: попередня обробка зображення райдужки, визначення азимутальних залежностей розмірів зіниць і райдужки, оцінка інформативних ознак (колір, просторова структура локальних знаків, оцінка статистичних характеристик іридологічних ознак, визначення координат іридологічних ознак та їх ототожнення зі знаками іридологічних схем.

У 1990-92 р.р. в літературі з'явилися повідомлення про розробку в тій же лабораторії (у співдружності з Московським НВО "Комета") комплексу "ИРИС" на базі ІВМ-сумісних комп'ютерів. Зовнішній вигляд комплексу "ИРИС" показаний на рис. 8.1.



Рис. 8.1

Призначення комплексу - проведення наукових досліджень і практична іридодіагностика. Наявністю дослідницької мети пояснюється розвиненість частини програмного забезпечення, спрямованої на обробку зображення райдужки: фільтрація зображень, масштабування фрагментів, поворот зображення, окантування зіниці і автономного кільця, псевдо розцвічування напівтонового зображення та ін. Для практичної іридодіагностики призначені режими топічної діагностики, системного аналізу, оцінки ступеня ризику розвитку певної пато-

логії (печінки, підшлункової залози, шлунка та ін), оцінки рейтингу старіння і спадкового статусу людини. Результати практичної діагностики друкуються на папері і складаються з двох частин: графічного зображення райдужки і текстового висновку лікаря про стан здоров'я пацієнта. Час діагностики на "ИРИС" складає 15-30 хв.

У літературі відзначені такі недоліки комплексу "ИРИС" як складність і громіздкість, яка перешкоджає його використанню навіть в поліклініках і лікарнях, не кажучи вже про використання в мобільному варіанті, наприклад, лікарями "швидкої допомоги", для експрес - діагностики. Ці недоліки частково можна пояснити перевантаженістю програмного забезпечення дослідними режимами обробки зображень (втрата "відчуття міри", можливо, обумовлена професійним інтересом програмістів до проблематики комп'ютерної обробки зображень). Нарешті, не бездоганна і форма видається укладення: карта огляду має вигляд послідовності кодів типу "2а", "17б" і т.д., що ускладнює розшифровку укладення іншими фахівцями.

Менш громіздким аналогом комплексу "ИРИС" була автоматизована система "ИРА"(розробка тієї ж лабораторії). Проте в літературі зазначалося, що "ИРА" все одно менш доступна лікувальним закладам, ніж іридографія (аналіз збільшених зображень райдужки, зафіксованих на кольорових слайдах) - причиною того, мабуть, була висока вартість комплексу.

Орієнтовно з 1993 р. здійснювалося комерційне поширення іридодіагностичного комплексу "АРМ лікаря - іридолога", розробленого в науково-виробничій фірмі "БЛИЦ" та АТ "РИНК" за участю Е.С.Вельховева. До складу цього АРМ'а входили: біокулярний іридоскоп з вбудованою кольоровою телевізійною камерою, персональний ІВМ-сумісний комп'ютер і відповідне програмне забезпечення. Призначення АРМ'а - автоматизація практичної іридодіагностики, навчання і перепідготовка іридодіагностів.

Програмне забезпечення комплексу дозволяло здійснювати автоматизоване виявлення патологій (приватна діагностика) в наступних системах органів: серцево - судинна, підшлункова залоза, легені, печінка, нирки, шлунково-кишковий тракт, молочна залоза, передміхурова залоза, портальна гіпертензія. Загалом комплекс забезпечує запис, зберігання, обробку і аналіз зображень райдужної оболонки ока, архівування та друк висновків, статистичних та реєстраційних даних огляду пацієнтів.

Переходячи до розгляду українських розробок, відзначимо, що наприкінці 80-х років Київське НВО "Квант" було одним з найбільших державних розробників технічних коштів іридодіагностики. Тому становить інтерес розгляд характеристик двох розробок цього підприємства - комплексів "ИРИС" і "ИРИС-ТМ".

Комплекс апаратури для іридодіагностики "ИРИС" розроблено у 1989 році за участю В.В.Кривенко - керівника спеціалізованої лабораторії нетрадиційних методів діагностики і терапії Інституту проблем онкології ім. Р.Є.Кавецького АН України.

Конструктивно комплекс виконаний у вигляді трьох окремих приладів: перетворення і передачі зображення, попередньої обробки і зберігання зображень, остаточної обробки зображення.

Схема функціонування комплексу коротко така. Після відображення на екрані монітора підсвіченої райдужки, її зображення поєднується з напівпрозорим, раніше вибраним, трафаретом ИРИС (схемою проєкційних зон), механічно закріпленим на екрані монітора. Поєднання досягається за допомогою рухомої платформи пристрою юстування, яка переміщається разом з телекамерою яка передає в трьох просторових координатах. У момент суміщення зображення запам'ятовується відеомагнітофоном, працювати в режимі "стоп-кадру".

В пристрої попередньої обробки і зберігання зображення реалізується можливості створення електронного трафарету і трансформації зображення райдужки (зміна масштабу зображення, виділення фрагментів зображення, накладення на зображення графічної та текстової інформації у вигляді графіка, гістограм, числових трафаретів, даних про пацієнта, дати обстеження, укладення лікаря - іридолога).

Апаратура обробки зображення райдужки складається з персонального комп'ютера, пристрою введення графічного зображення, накопичувача на магнітній стрічці. Призначення цієї програмно-апаратної системи полягало в розробці і налагодженні програмного забезпечення для напівавтоматичного і автоматичної постановки діагнозу.

Функції програмного забезпечення такі: визначення координат іридознаків, оцінювання ступеня децентралізації і сплюснення зіниць, побудова гістограм за кольором або яскравістю, змішування зображень, знаходження центра щільності в вибраному фрагменті зображення, корекція зображення, статистична обробка зображення, що вводяться зі слайдів.

На даний час, коли можливості апаратного і програмного забезпечення обчислювальної техніки істотно збільшились, очевидні громіздкість і моральна застарілість комплексу "ИРИС", початкова придатність його швидше для науково-технічних досліджень, ніж для практичної іридодіагностики.

Наступною і більш сучасною розробкою НВО "Квант" був програмно-технічний комплекс (ПТК) "ИРИС-ТМ", потрібний для проведення масових профілактичних обстежень пацієнтів і для навчання лікарів.

До складу ПТК "ИРИС -ТМ" входили кольорова телевізійна камера з оптичним пристроєм для макрозйомки, електронний блок для введення відеосигналу в комп'ютер, ІВМ-сумісний комп'ютер, відеомагнітофон, спеціальне програмне забезпечення.

Основні можливості програмного забезпечення: управління введенням відеосигналу в комп'ютер, виборами зони інтересу (вікна), робота із зоною інтересів (масштабування), збереження зображення на ГМД і жорсткому диску, введення зображення з ГМД або жорсткого диска, аналіз зображення (виділення контурів, робота з гістограмою яскравості, визначення геометричних характеристик зо-

браження), накладення трафаретів іридологічних карт на зображенні райдужної оболонки ока пацієнта, отримання довідкової іридодіагностичної інформації.

У 90-ті роки київськими інженерами були розроблені ще декілька комп'ютерних програм для комп'ютерної іридодіагностики. Особливістю даних програм є те, що вони представляли собою ЕС для іридодіагностики, названі ЭСИД і ЭСИД-2, при розробці яких істотно використані елементи Байєсівської стратегії.

Нижче представлена інформація про ці програми, а також про їх подальшому розвитку і створенні найбільш сучасного комплексу програм для іридодіагностики.

8.1.1. Експертна система для іридодіагностики ЭСИД

Комп'ютерна програма ЭСИД (MS DOS, мова програмування Clipper, обсяг програми 1,3 Мб, з яких 80% складають бази знань) розроблена в 1989-91гг. в Київському НВО "Славутич" і являє собою невелику, але разом з тим досить потужну і просту в експлуатації експертну систему продукційного типу, що дозволяє здійснювати загальну, диференціальну і приватну іридодіагностику з використанням елементів Байєсівської стратегії.

Призначення і можливості. Програма ЭСИД багатоцільова: вона може використовуватися для практичної іридодіагностики, наукових досліджень і навчання іридологів.

У перелік її можливостей входять:

- кількісна оцінка достовірності можливих захворювань;
- рішення завдань загальної, диференціальної та приватної діагностики з пошуком інформації за схемами «ознака - захворювання» і «захворювання - ознака»;
- використання, поряд з іридодіагностичними, додаткових ознак;
- аргументування діагнозу;
- автоматизоване формування висновку лікаря з рекомендаціями щодо лікування виявлених захворювань;
- протоколювання результатів обстежень, з можливістю подальшої автоматизованої статистичної обробки протоколів;
- модернізація бази знань без залучення професійних програмістів;
- можливість автоматизованого обліку результатів статистичної обробки протоколів при модернізації бази знань;
- кількість розпізнаваних нозологічних форм - близько 300;
- кількість використовуваних діагностичних ознак - порядку 1000;
- наявність режиму «Допомога» у вигляді кольорових графічних ілюстрацій і текстових коментарів.

База знань. Містить систему реляційних баз даних з переліками іридодіагностичних і додаткових (не іридодіагностичних) ознак, переліками нозологічних форм, інформацією про статистичні зв'язки між ознаками і захворюваннями, рекомендаціями щодо лікування.

Кількість використовуваних БД - 8, з них 7 постійних, використовуваних для пошуку інформації в прямому напрямку «ознака - захворювання», і одна тимчасова БД - для пошуку інформації в зворотному напрямку «захворювання - ознака» (див. рис.8.2).

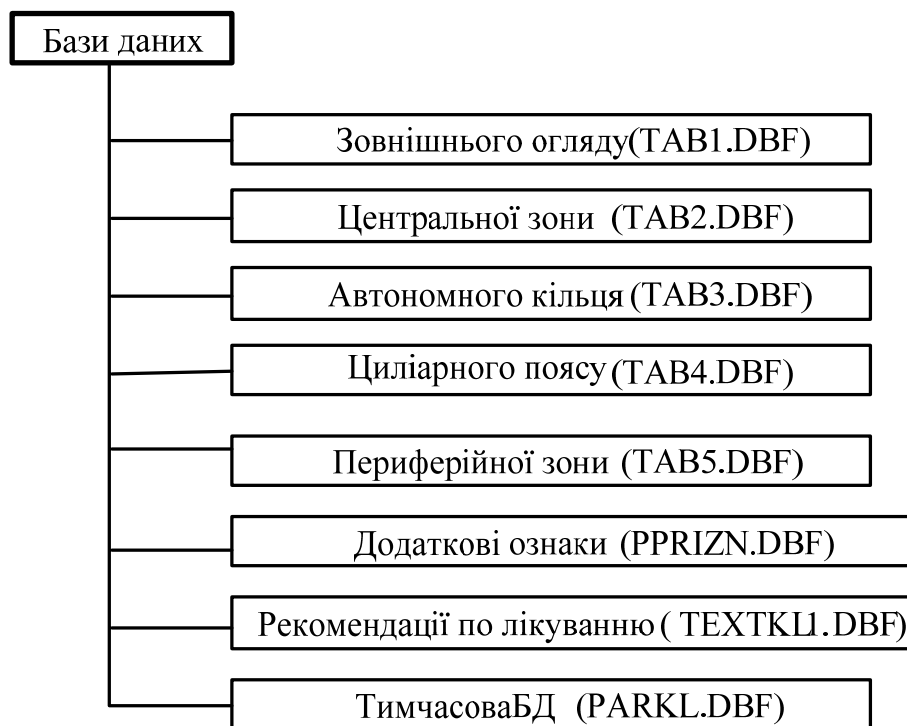


Рис. 8.2

Структура БД зовнішнього огляду, призначеної для зберігання інформації про конституціональні ознаки, має вигляд табл. 8.1. Зауважимо, що в цій таблиці зв'язку між ознаками і станами організму детерміновані - імовірнісний характер цих зв'язків відображено у вербальній формі в тексті описів цих станів (типу «Прогноз при важких захворюваннях задовільний").

Таблиця 8.1

Номер зони райдужної оболонки	Номер групи ознак	Ім`я групи ознак	Номер ознаки (всередині групи) i	Ім`я ознаки X_i	Прогноз стану здоров`я Y_{j1}	Генетичні особливості організму Y_{j2}
-------------------------------	-------------------	------------------	---------------------------------------	----------------------	------------------------------------	---

Структура БД TAB2-TAB5 має вигляд табл. 8.2. Тут імовірнісна залежність між захворюваннями і ознаками задана в числовій формі, у вигляді оцінок умовних ймовірностей $P(X_i / Y_j)$ - по суті, це традиційна «медична пам'ять» у формі, показаної в табл. 2.3.

Поля № № 6-7 повторені десятикратно з розрахунку, що одна ознака може відповідати максимум десяти хворобам.

Структура БД PPRIZN показана в табл. 8.3, а БД ТЕХТКЛ1 - в табл. 8.4.

Таблиця 8.2

Номер зони райдужної оболонки	Номер групи ознак	Ім'я групи ознак	Номер ознаки (всередині групи) i	Ім'я ознаки X_i	Номер (код) захворювання j	$P(X_i/Y_j)$
-------------------------------	-------------------	------------------	------------------------------------	-------------------	------------------------------	--------------

Таблиця 8.3

Номер (код) захворювання j	Ім'я захворювання Y_j	Номер ознаки i	Ім'я ознаки X_i	$P(X_i/Y_j)$	Індекс ознаки: (0–999)
------------------------------	-------------------------	------------------	-------------------	--------------	------------------------

Таблиця 8.4

Номер (код) захворювання j	Ім'я захворювання Y_j	Рекомендації по лікуванню
------------------------------	-------------------------	---------------------------

Процедура модернізації баз знань здійснюється самим користувачем, без допомоги професійних програмістів, і зводиться до редагування БД засобами вбудованої СУБД. При цьому редагування значень умовних ймовірностей $P(X_i / Y_j)$ може здійснюватися двома способами:

- 1) запозиченням з наукової літератури або з інших джерел інформації (думка експертів, особистий досвід та ін);
- 2) обчислюватися в спеціальному режимі функціонування програми ЭСИД на фактичному матеріалі, накопиченому шляхом протоколювання результатів обстежень пацієнтів (режим статистичного аналізу прецедентів) - такі обчислення проводяться з використанням спеціальної програми SINTEZ.

Можливості СУБД дозволяють користувачеві також змінювати переліки ознак і захворювань, редагувати тексти рекомендацій.

Вид графічного інтерфейсу програми ЭСИД в режимі введення значень діагностичних ознак зображений на рис. 8.3.

Механізм логічного виведення. Використовуваний в ЭСИД механізм логічного виведення базується на частковому використанні Байєсівської стратегії: в процесі загальної або диференціальної діагностики по заданим з клавіатури ПЕОМ значень діагностичних ознак (іридодіагностичних або додаткових) автоматично обчислюються і представляються у вербально-графічній формі значення оцінок апостеріорного розподілу ймовірностей $P(Y_j/X_i)$.

Основою алгоритму обчислень служать співвідношення:

$$P(Y_j / X_1, \dots, X_I) = \frac{I-1}{I} P(Y_j / X_1, \dots, X_{I-1}) + \frac{1}{I} P(Y_j / X_I)$$

де апостеріорна ймовірність $P(Y_j/X_i)$ захворювання Y_j для кожної ознаки X_i обчислюється за формулою:

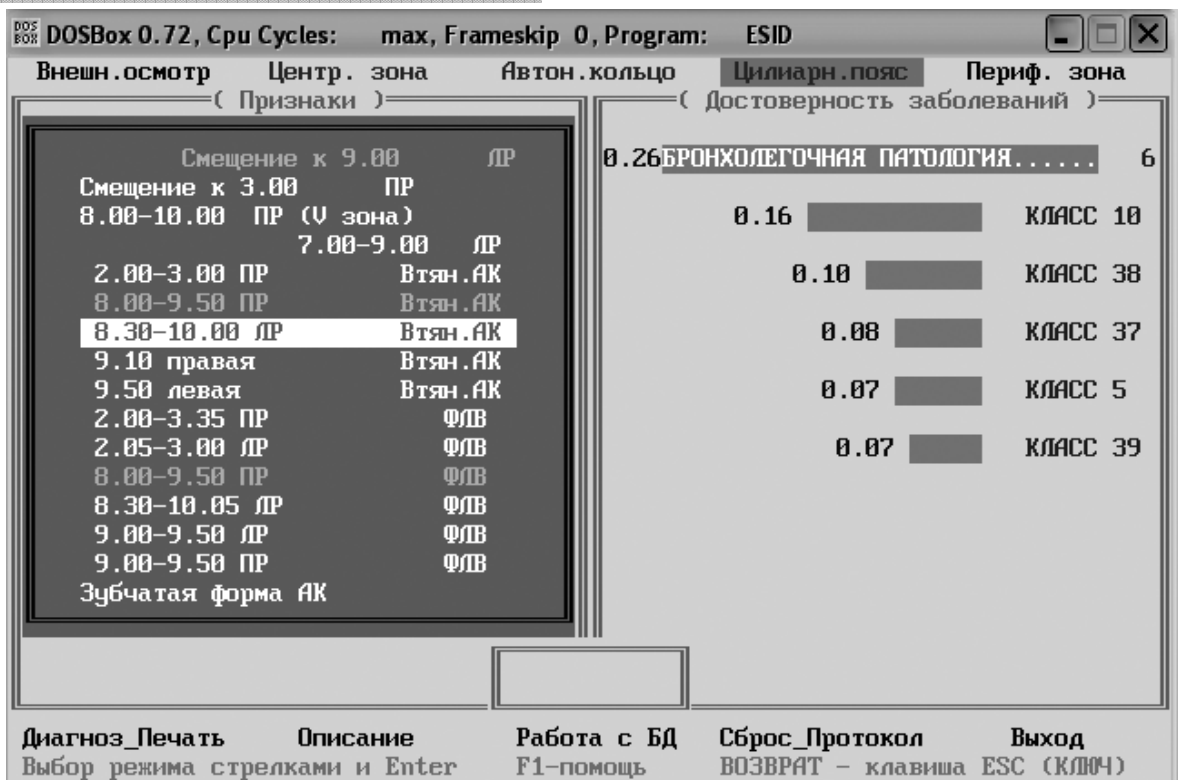


Рис. 8.3

$$P(Y_j / X_i) = \begin{cases} \frac{1}{J_i} \left(1 - \frac{r}{R}\right) + \frac{r}{R} \frac{P(X_i / Y_j)}{\sum_j P(X_i / Y_j)}, & J_i > 1, \\ \frac{r}{R} P(X_i / Y_j), & J_i = 1. \end{cases}$$

де r - числовий коефіцієнт ($r = 0 \dots R$), що характеризує вираженість спостережимої ознаки X_i .

Використання такого алгоритму дозволяє користувачеві спостерігати реакцію ЕС вже після введення в неї значення першої ознаки, а застосування коефіцієнта r пом'якшує жорсткість питання, який ЕС звертає до користувача: «Ознака X_i спостерігається? Так чи ні?». Так, значення $r = 0$ відповідає відповіді «Ні», $r = R$ - відповіді «Так», але значення $r = 3$ можна трактувати як проміжну відповідь типу «Ознака X_i спостерігається, але виражений досить помірно».

При приватній діагностиці в розгляд природно брати насамперед ті захворювання Y_j , яким відповідають найбільші значення розподілу $P(Y_j/X_1, \dots, X_I)$.

Технологія обстежень. Виявивши ознаки на райдужці ока, лікар вводить інформацію в комп'ютер, на підставі якої в комп'ютері здійснюється автоматичний пошук захворювань, відповідних заданій ознаці, та обчисленню їх рейтин-

гу. Результати відображаються на дисплеї у вигляді графіка достовірних можливих захворювань - значень розподілу $P(Y_j/X_1, \dots, X_l)$ (рис.8.3, 8.4).

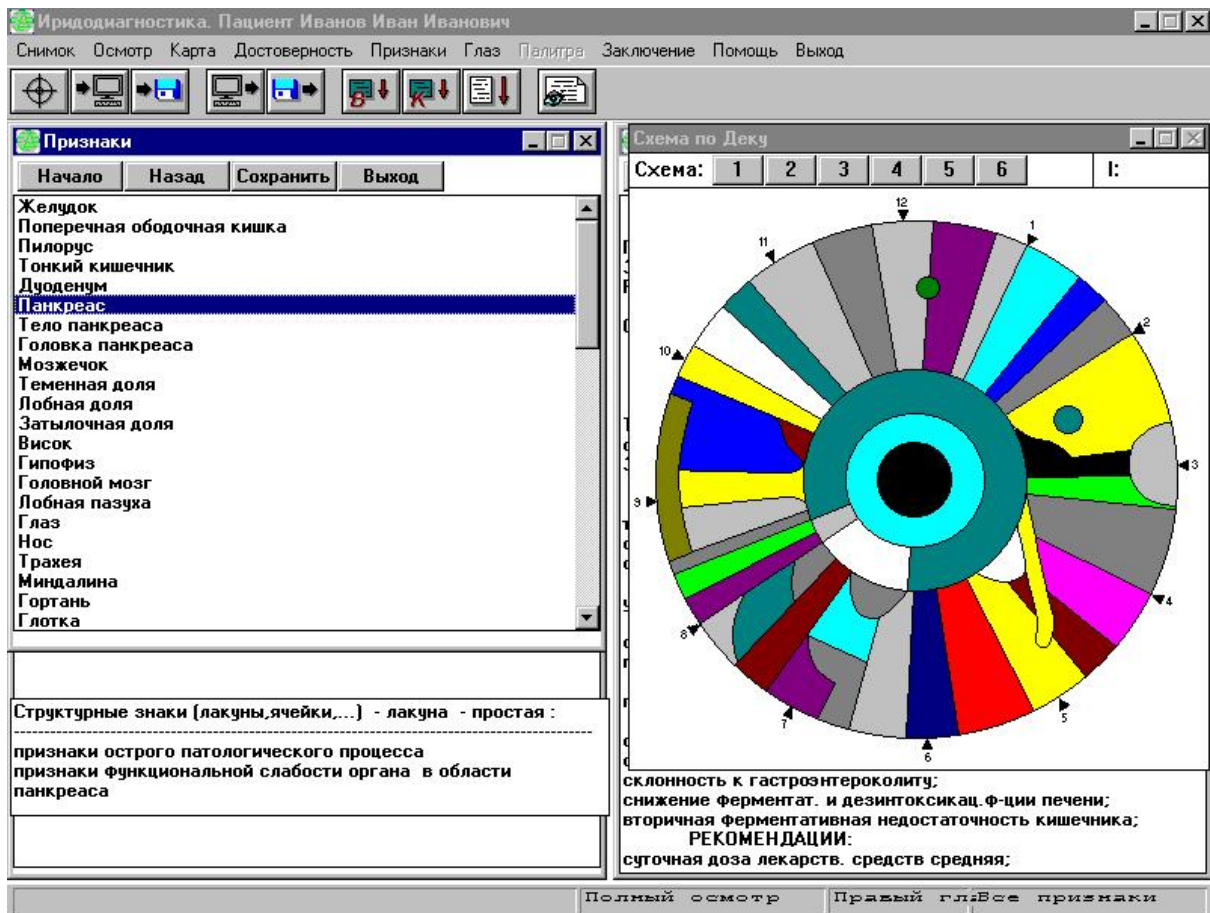


Рис. 8.4

Пошук можливих захворювань за схемою "ознака-захворювання" доцільний на початковій стадії діагностики (загальна або диференціальна діагностика), поки на шкалі рейтингів не проявиться тенденція до наявності одного або декількох, найбільш ймовірних, захворювань.

Потім можна приступати до приватної діагностики, виробляючи пошук у зворотному напрямку за схемою "захворювання-ознака". При цьому на дисплеї відобразиться повний перелік можливих ознак "підозрюваного" захворювання, що дозволяє більш цілеспрямовано повторити рух за схемою "ознака-захворювання" і остаточно підтвердити (або спростувати) попередній діагноз.

Технологія формування висновку іридолога. Для формування висновку іридолог позначає на графіку розподілу $P(Y_j/X_1, \dots, X_l)$ найменувань виявлених захворювань. За командою на дисплеї у вербальній формі відображається висновок іридолога, що містить:

- оцінку загального стану організму;
- оцінку стану відділів нервової системи;

- відносні рейтинги виявлених патологій та захворювань;
- аргументацію можливого діагнозу;
- рекомендації з додатковим цілеспрямованим обстеженням і лікуванню.

Висновок може бути виведено на друк за однією з кількох наперед заданих форм, що відрізняються ступенем подробиць виданого висновку.

Досвід експлуатації програми ЭСИД показав:

- 1) при автоматизованому формуванні укладення іридолога завжди існує небезпека гіпердіагностики, що призводить до необхідності обов'язкового "ручного" редагування тексту висновку;
- 2) використання великих баз даних на малопотужних комп'ютерах призводить до значних втрат часу при пошуку інформації;
- 3) значна кількість рівнів ієрархічної системи іридоознак призводить до помітних втрат часу при введенні даних навіть на потужних комп'ютерах.

8.1.2. Експертна система для іриодіагностики ЭСИД-2

Комп'ютерна програма ЭСИД-2 (середовище Windows, мова програмування Visual C ++, обсяг програми 1.1 Мб), розроблена в 1995-96гг. в Київському ГНВП "Аргус", є складовою частиною апаратно-програмного іриодіагностичного комплексу АРМИР і являє собою досить потужну експертну систему.

Вікно введення інформації програми ЭСИД-2 показано на рис. 8.4.

Призначення і можливості. Система призначена в першу чергу для практикуючих іридологів, хоча може використовуватися і в цілях навчання, і в науково-дослідних цілях.

В якості комплексного критерію якості системи прийнята корисність її для іридологів практично будь-якого рівня кваліфікації - в поняття корисності при цьому включені не тільки приватні вимоги високої достовірності та швидкості діагностики, придатності для масових профілактичних оглядів і глибоких індивідуальних обстежень, а й система наступних пріоритетів: для досвідченого лікаря - іридолога важливіше швидкість формування тексту висновку, тоді як для малодосвідченого - можливість швидкого пошуку в пам'яті комп'ютера надійної інформації про патології, відповідних піднаглядним іридоознакам.

Основні можливості системи:

- введення в пам'ять комп'ютера кольорових зображень РВ і їх відображення на екранах комп'ютера і допоміжного монітора в 256 кольорових градаціях;
- автоматизоване формування карти огляду та висновку іридолога;
- запис результатів обстеження до архіву на жорсткому магнітному диску;
- коректування наявних електронних довідників іридолога або створення нових довідників (модернізація баз даних здійснюється самим користувачем);
- перегляд, редагування архівних записів;
- друк результатів діагностики.

Відмінними, в порівнянні з ЭСИД, особливостями програми ЭСИД-2 є:

- програма призначена для функціонування в середовищі Windows, що відповідає більш сучасному рівню інформаційних технологій;
- програма може експлуатуватися як автономно, на одному комп'ютері, так і в локальній комп'ютерній мережі, в комплексі з іншими комп'ютерними програмами медичного призначення;
- є можливість аналізу введеного в комп'ютер кольорового зображення райдужки (з використанням відеокамери та спеціального пристрою введення);
- організація процедури формування висновку іридолога спрямована на задоволення вимог іридологів практично будь-якої кваліфікації, - від досвідченого до початківця іридолога;
- при аналізі топічних іридознаків забезпечена можливість використання вербально-графічних меню у вигляді відображуваних на моніторі різних схем проєкційних зон райдужки ока - форма такого відображення максимально наближена до наведених в літературі схем проєкційних зон.

Механізм логічного виводу. У силу комплексного характеру використовуваного критерію якості, механізм пошуку довідкової інформації хоча і використовує елементи байєсівської стратегії, проте ці елементи виражені меншою, в порівнянні з ЭСИД, ступеня: можливість відображення числового значення умовної ймовірності $P(X/Y_j)$ передбачена при зворотному напрямку пошуку інформації ("захворювання - ознака"), тоді як при прямому напрямку пошуку ("ознака - захворювання") проводиться лише підрахунок кількості ознак, "проголосували" за дане захворювання, з подальшим внутрішньосистемним (для кожної з систем організму людини) упорядкуванням (по числу "голосів") захворювань.

Такий підхід обумовлений насамперед практичною спрямованістю програми ЭСИД-2, відповідно до якої кодовано-числовій формі укладення надається перевага вербальна форма, зручна як іридологу, так і пацієнту. Що стосується можливих при цьому втрат у достовірності виводів, для зниження їх передбачено кілька вербальних градацій вираженості захворювань. У ЭСИД-2 таких градацій дві: "є схильність (тенденція) до ..." і "є ознаки ..." Як свідчить практика експлуатації програми "ЭСИД-2", такої кількості градацій достатньо, щоб уникнути гіпердіагностики і домогтися прийнятної (для методу іриодіагностики) точності формулювань висновку.

Технологія обстежень. Загальна та диференціальна іриодіагностика з використанням ЭСИД-2 реалізується в кілька етапів:

- 1) оцінка конституціонального стану організму за методикою Е.С.Вельховева або В.В.Кривенко;
- 2) напівавтоматичне формування фрагментів тексту укладення іридолога шляхом введення в комп'ютер найменувань спостережуваних іридологічних ознак;
- 3) редагування тексту висновку використанням простих інформаційно-довідкових систем ("електронних довідників") і засобів текстового редагування,

властивих середовищу Windows, за результатами збору анамнестичних даних (в процесі діалогу з пацієнтом).

Приватну іридіодіагностику з використанням ЭСИД-2 виробляють шляхом інформаційно-довідкового пошуку іридіологічних ознак, відповідних висунутим гіпотезам, - при цьому, на відміну від загальної іридіодіагностики, відсутній етап 1) і більш інтенсивно експлуатується програма на етапі 2) (з попереднім оцінюванням кількості іридіологічних ознак, "проголосували" за дане захворювання).

Технологія формування висновку іридіолога. Організована таким чином, що іридіолог вже після введення перших кількох ознак отримує на 70-80% готового висновку. Досягається цей ефект за рахунок наступних заходів:

- продуманість ієрархічної структури ознак, при якій обстеження починаються з спостереження комплексних, конституціональних, інформаційно найбільш насичених ознак (тип, підтип райдужної оболонки);
- використання вербальних градацій вираженості захворювань типу "є схильність (тенденція) до ..." і "є ознаки ..." замість числових оцінок;
- автоматичне посистемне (з систем організму людини) структурування висновку.

Після формування чорнового укладення іридіологу необхідно лише відредагувати текст з урахуванням анамнестичних даних, отриманих при діалозі з пацієнтом - це редагування виробляється з використанням «електронних довідників», як вбудованих, так і зовнішніх.

Час діагностики складає в середньому 15-20 хвилин.

Вікно редактора укладення показано на рис. 8.5.

8.1.3. Програмний комплекс для іридіодіагностики: ЭСИД-3, електронні атласи і графічний редактор GIDRA

Комплекс програм, до складу якого входять ЕС ЭСИД-3, графічний редактор GIDRA і два електронних атласи, розроблений в Національному технічному університеті України (КПІ) і призначений для експлуатації в ОС Windows всіх версій.

Експертна система ЭСИД-3

Призначення програми ЭСИД-3 - інтелектуальна підтримка іридіолога при проведенні експрес-діагностики або поглибленої діагностики, що завершується видачею висновку іридіолога. Узагальнена структура графічного інтерфейсу програми ЭСИД-3 наведена на рис.8.6, а на рис.8.7 показана сукупність вікон графічного інтерфейсу: головне вікно і вікна введення ознак при поглибленої діагностики. Безпосереднім прототипом програми ЭСИД-3 послужила її попередниця - програма ЭСИД-2, тому багато в чому ці програми схожі. Разом з тим, в програмі ЭСИД -3 враховано досвід розробки та експлуатації програми ЭСИД-2, а також досвід розробки та експлуатації аналогічних програмних про-

дуктів. Тому на сьогоднішній день програму ЭСИД-3 можна вважати однією з найбільш просунутих експертних систем такого роду.

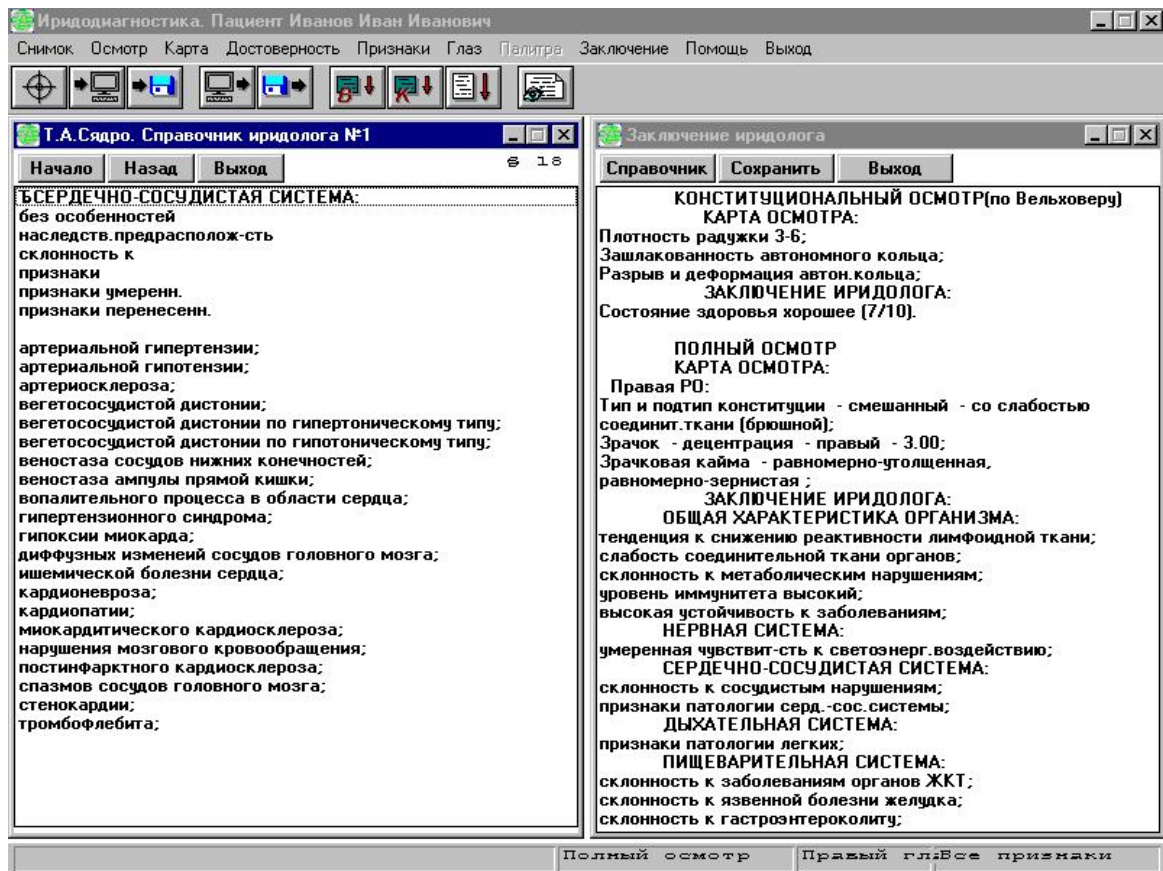


Рис. 8.5

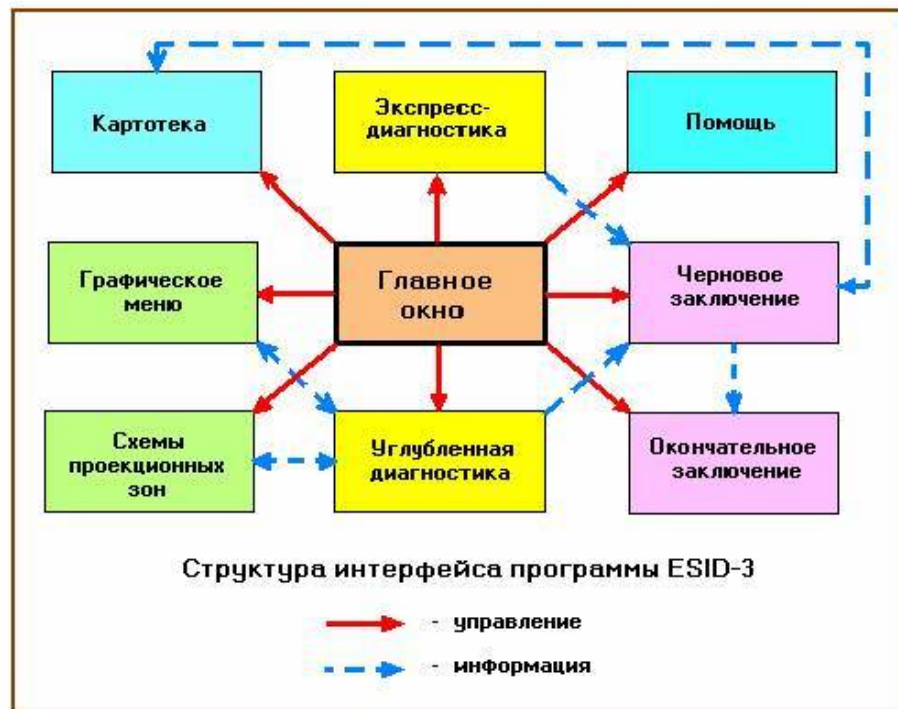


Рис. 8.6

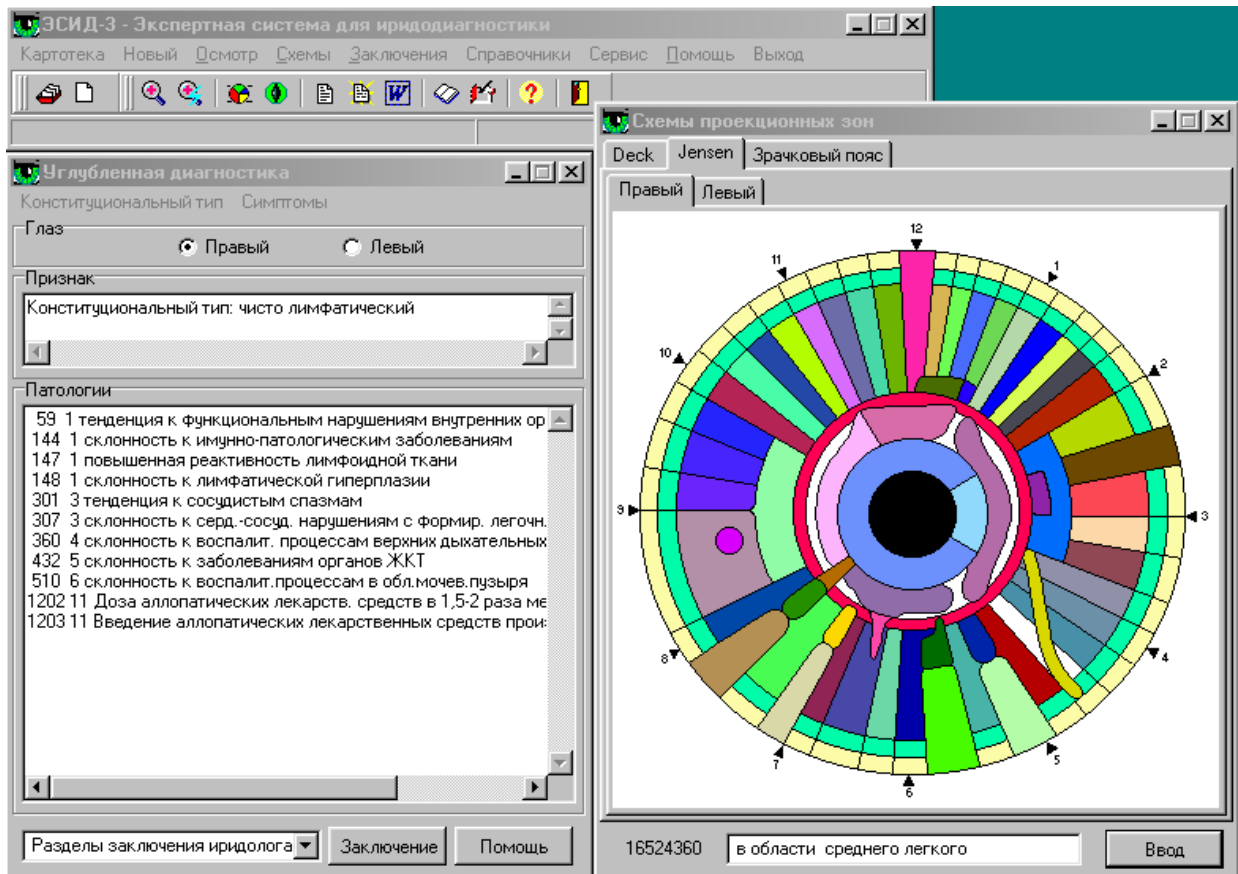


Рис. 8.7

У програмі ЭСИД-3 збережені режими експрес-діагностики та поглибленої діагностики. Режим поглибленої діагностики реалізований зручніше в тому сенсі, що занурення на нижній рівень розгалуженого меню здійснюється одним клацанням мишки, не вимагаючи багаторазового звернення до клавіатури. Вилучена можливість пошуку інформації в «зворотному» напрямку захворювання-ознака, оскільки експлуатація ЭСИД-2 показала, що такий пошук на практиці використовується рідко.

У програмі ЭСИД-3 додана можливість створення власних довідників іридолога (рис.8.8). Остаточне редагування ув'язнення може проводитися в середовищі MS Word.

Висновок іридолога формується поетапно. Спочатку, у міру введення в комп'ютер інформації про спостережуваних лікарем - іридологом іридоознак, автоматично формується "чорнове" висновок (рис.8.9).

Потім "чорновий" висновок редагують, в результаті чого отримують "чистовик" висновок, яке і виводять на друк. Для остаточного редагування висновок використовують або вбудований текстовий редактор (рис.8.10), або програму MS Word. В останньому випадку за допомогою спеціальної команди текст висновок копіюється в автоматично активізуючий додаток MS Word.

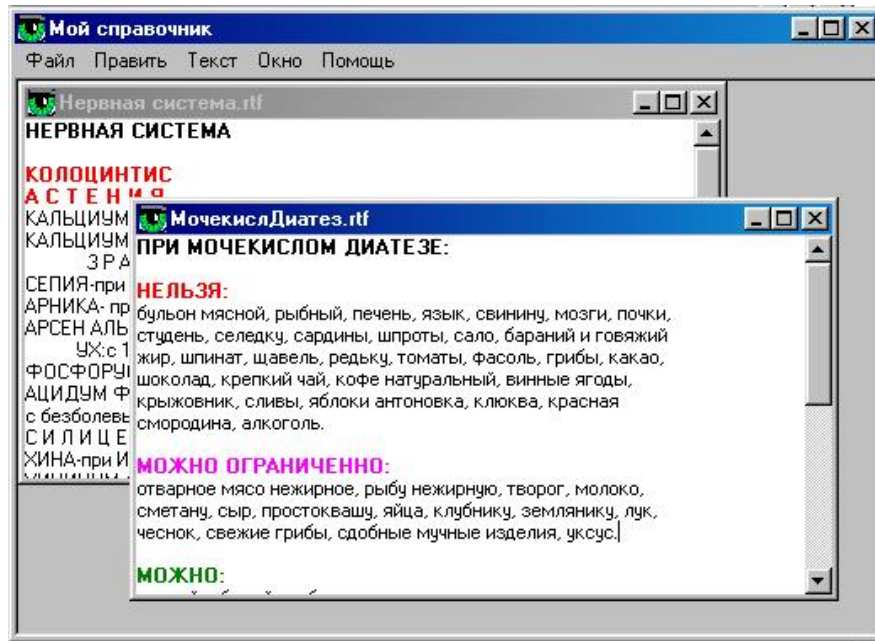


Рис. 8.8

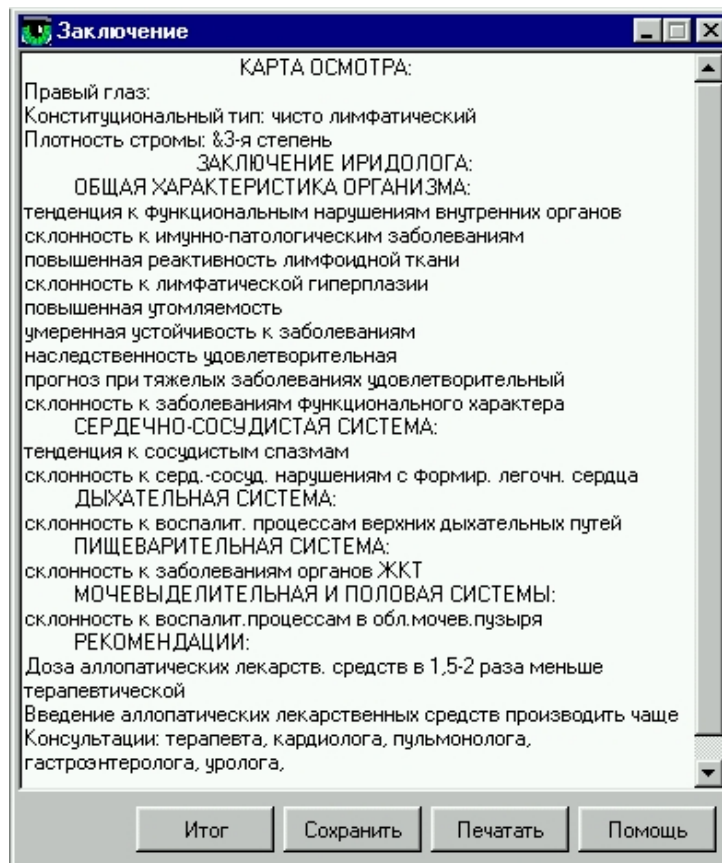


Рис. 8.9

При роботі з програмою, іридолог може звертатися як до вбудованого електронного довідника, так і до зовнішніх довідників, реалізованим у вигляді спеціальних додатків.

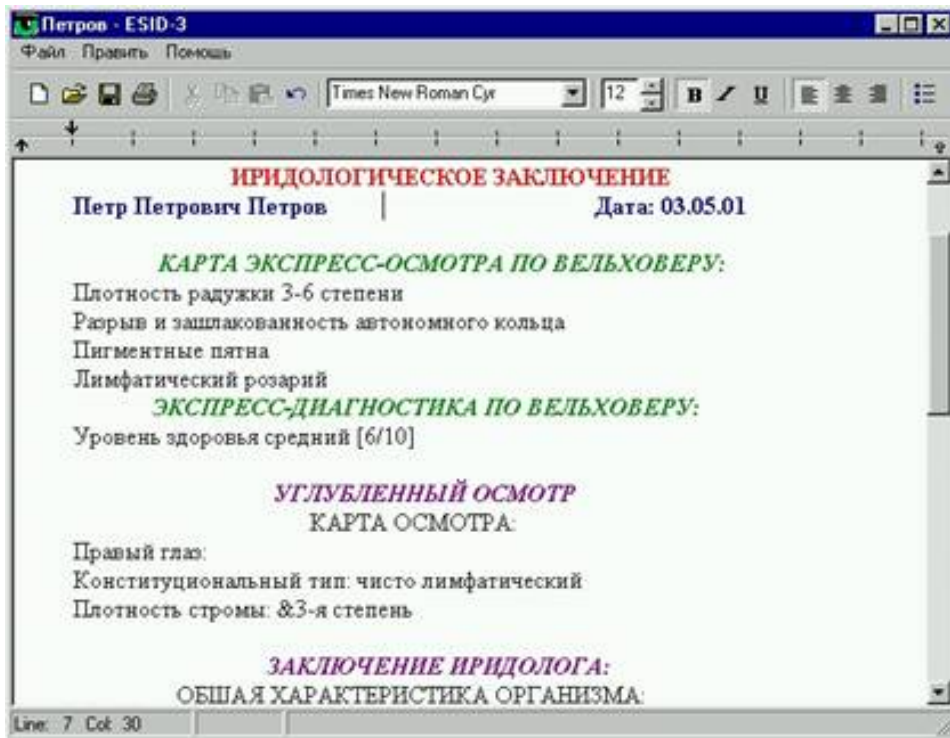


Рис. 8.10

Електронні атласи. «Електронний атлас іридолога. Частина 1. Іридодіагностичні синдроми» (розроблений у 1999 році, модернізований у 2002 році) представляє собою електронний довідник з великою базою даних про конституціональних типах райдужки (зображення з відповідними текстовими коментарями). Графічний інтерфейс даного Атласу-1 показаний на рис.8.11.

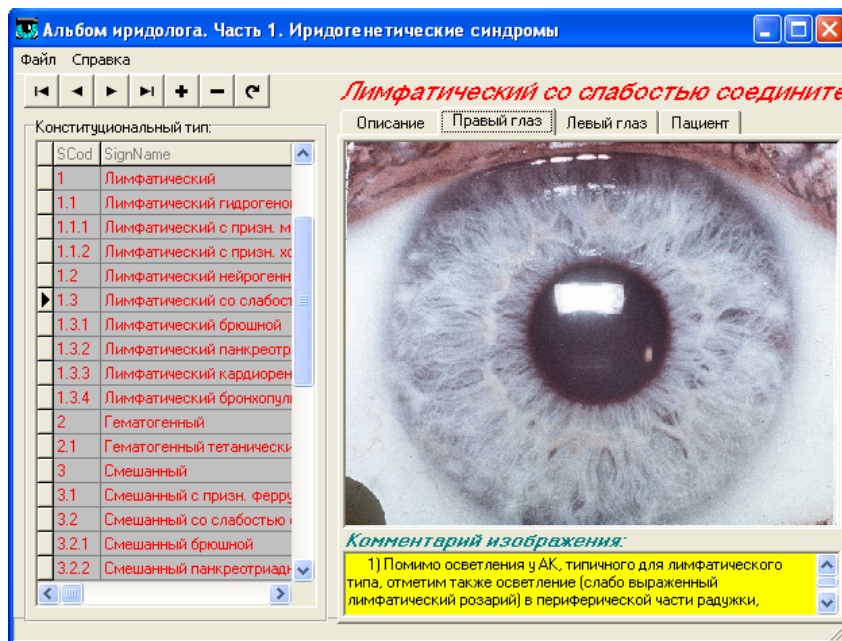


Рис. 8.11

«Електронний атлас іридолога. Частина 2. Іридологічні симптоми» (розроблений у 2004 році) - аналогічний електронний довідник, який зберігає текстову і графічну інформацію про різні іридологічні ознаки - плямах, лакунах і т.п. Графічний інтерфейс даного Атласу-2 зображений на рис.8.12.

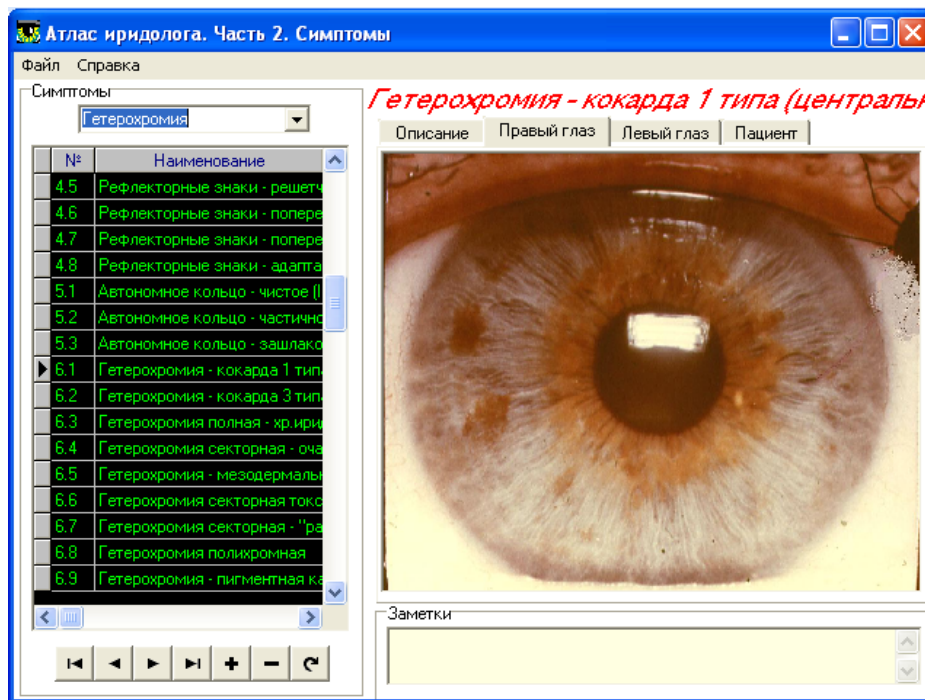


Рис. 8.12

Фрагменти текстової довідкової інформації можуть бути легко скопійовані з атласів в текст висновку, формованого програмою ESID-3.

Нарешті, для спільної роботи з програмою ЭСИД-3 призначений спеціально розроблений графічний редактор GIDRA, значно полегшує роботу з оцифрованими зображеннями райдужних оболонок.

Графічний редактор для іридологів GIDRA НТУУ (КПІ). Розроблена К.Н.Коваленко і А.Н. Продеусом в 2002-2004 роках на кафедрі акустики та акустоелектроніки НТУУ (КПІ), комп'ютерна програма GIDRA (рис.8.13) являє собою графічний редактор, що враховує вимоги іридологів різних категорій - практичних іридологів, іридологів - дослідників і іридологів - викладачів.

Програма дозволяє поєднувати зображення схеми проекційних зон із зображеннями райдужної оболонки. Можна вимірювати координати спостережуваних ознак у звичній для іридологів системі координат «годинниковий циферблат», а також вимірювати довжину, периметр і площу іридологічних знаків. Користувач може робити написи на зображеннях, малювати на них прямокутники, овали, лінії і стрілки, виділяти, вирізати або обрізати виділений фрагмент зображення і т.п. Вдало вирішена функція швидкого масштабування за допомогою елемента керування у вигляді повзунка. Можливі швидкі повороти зображення на 90 градусів, а також плавне обертання зображення. Можлива фільтра-

ція виділених ділянок зображення, що дозволяє підкреслити структурні особливості зображення.

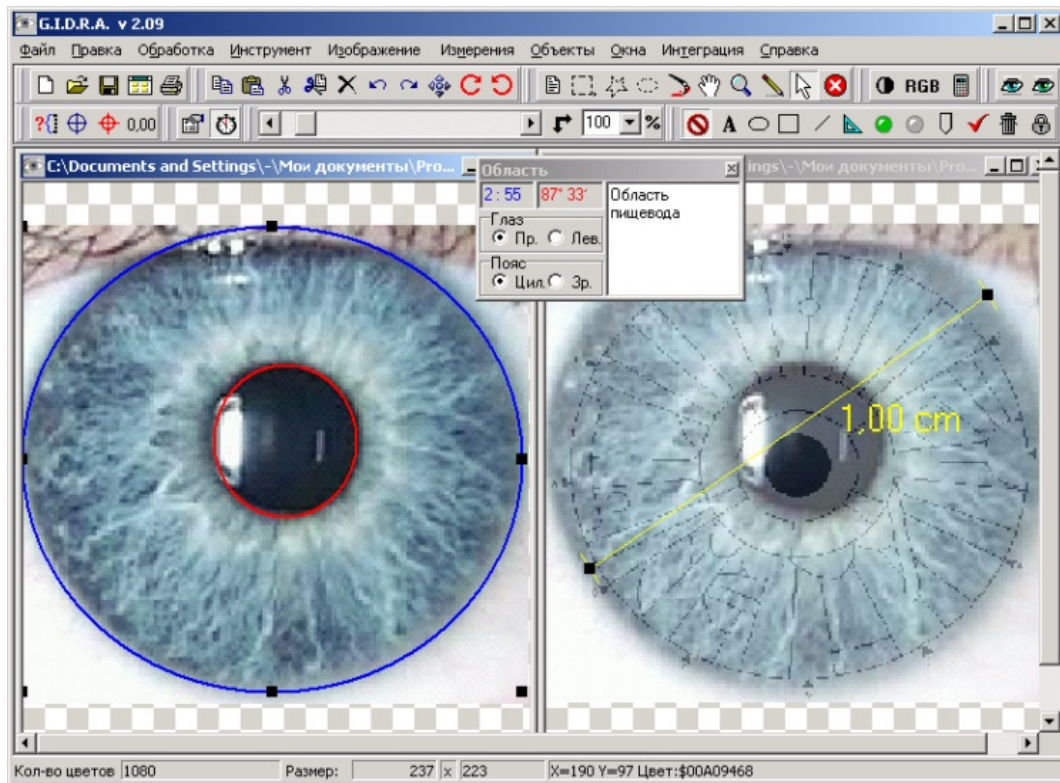


Рис. 8.13

8.1.4. Труднощі і перспективи комп'ютерної іридодіагностики

Аналіз особливостей функціонування сучасних іридодіагностичних систем, а також облік сучасного рівня інформаційних технологій дозволяють виробити узагальнені вимоги до системам що розробляються вперше:

- система повинна дозволяти проводити як масові профілактичні огляди, так і глибокі індивідуальні обстеження;
- організація процедури автоматизованого формування висновку іридолога повинна бути гнучкою, задовольняючи вимогам іридолога будь-якої кваліфікації;
- повинні бути можливими як автономна експлуатація системи, так і в локальній мережі, в комплексі з іншими аналогічними системами, що забезпечують, наприклад, комп'ютерну підтримку УЗД, аналізу ЕКГ і т.п.

Накопичений до теперішнього часу досвід розробок і експлуатації систем комп'ютерної іридодіагностики свідчить про наявність серйозних труднощів технічного, медичного та економічного характеру.

По-перше, підвищення якості і швидкості іридодіагностики за рахунок застосування комп'ютерних систем підтримки досягається ціною збільшення вар-

тості обстежень - провиною тому неминучі витрати на придбання або розробку комп'ютерних систем, тим більші, чим вище складність таких систем (наприклад, вартості іридологічних програм і програмно-апаратних комплексів співвідносяться як 1:50).

По-друге, якість одержуваного на моніторі комп'ютера кольорового зображення РВ зазвичай все ж не забезпечує тих можливостей, які забезпечує використання бінокулярних іридоскопії або щілинної лампи. Складовою частиною зазначеної вище проблеми є необхідність калібрування апаратних засобів, а також наявність таблиці відповідності зразків різних іридоознак і найменувань патологій та захворювань.

Практичний інтерес для іридологів представляє і принципова можливість роздруківки зображень РВ на кольорових принтерах, однак і тут актуальна проблема забезпечення якісної передачі кольору в поєднанні з прийнятною вартістю знімків.

Для молодосвідчених іридологів великий інтерес представляє завдання автоматизації розпізнавання діагностичних ознак за кольоровим зображенням РВ - однак складність аналізованого об'єкта в даному випадку настільки висока, що в найближчому майбутньому можна сподіватися на автоматизацію розпізнавання лише найпростіших іридоознак.

Таким чином, сучасний рівень розвитку науково-технічних методів і засобів розпізнавання образів дозволяє задовільно вирішити лише кілька приватних завдань автоматизації іридодіагностики:

- моделювання процесу логічного висновку іридолога в задачах загальної та приватної діагностики;
- автоматизація процесу формування висновку іридолог.

Ці завдання сьогодні вирішуються шляхом створення експертних систем та інформаційно-довідкових систем. Зауважимо однак, що рішення і цих приватних завдань пов'язане з рядом суттєвих проблем:

- достовірність висновку іридолога, сформованого з використанням ЕС, в значній мірі залежить від ефективності алгоритмів обробки інформації - між тим, застосуванню потенційно найбільш ефективної Байєсівської стратегії перешкоджає ряд факторів, які важко усунути (статистична залежність ознак, неповнота апріорних даних, неповнота і неоднорідність іридологічної інформації, наявність заважаючих факторів); подолати деякі складності можна при більш уважному ставленні іридологів до організації досліджень і опублікуванню їх результатів;
- уявлення про раціональну форму укладення іридолога істотно відрізняються у різних користувачів - між тим, укладення іридолога, по суті, є цільовою функцією, яка в значній мірі визначає структуру і функції комп'ютерних систем підтримки; допомогти тут може стандартизація вимог до форми висновку;
- швидкість формування висновку іридолога істотно знижується через втрати часу на маніпулювання з клавіатурою комп'ютера при введенні

інформації про наявність тієї чи іншої ознаки - глибина занурення на нижній рівень ієрархічно організованого переліку іридологічних ознак може досягати 5-7 рівнів; можливий вихід з цієї ситуації, що базується на головному управлінні комп'ютерною програмою, також не представляється привабливим чинності недостатньо високої достовірності результатів роботи систем автоматичного розпізнавання мови.

Перераховані проблеми частково подолані при розробці програми ЭСИД-3, що базується на таких передумовах:

- "вершиною" діяльності лікаря - іридолога є висновок лікаря, тому мета комп'ютерної, з використанням ЕС, підтримки іридолога полягає в забезпеченні швидкого автоматизованого формування добре структурованого укладення з гарантовано високим рівнем достовірності висновків;
- ступінь автоматизації процедури формування висновку іридолога не повинна бути надмірно високою в тому сенсі, що не повинна виключатися можливість "ручного" коригування висновку - це вимога пояснюється, з одного боку, труднощами забезпечення високої достовірності автоматично формованих висновків, а з іншого боку, відповідальністю лікаря за якість виданого висновку;
- консультаційні ЕС продукційного типу (тобто базуються на використанні правил "якщо ... - то ...") перспективні, в плані забезпечення комп'ютерної підтримки іридолога, практично при будь-якому вигляді діагностики;
- байєсівська стратегія, трактується як форма реалізації правила продукції, є потужним і зручним механізмом трансформації шкал ознак в шкали захворювань, однак при побудові систем, призначених для іридолог - практиків, більш кращий більш простий і наочний механізм конкатенації (логічного об'єднання) приватних висновків.

Перераховані вище положення можна розглядати як рекомендації різним категоріям розробників іридодіагностичних ЕС - інженерам по знаннях, лікарям-експертам, програмістам.

8.2. Інформаційно-довідкові системи для гомеопатії

Сьогодні відомо безліч простих програм-довідників для лікарів-гомеопатів. Одним із прикладів довідників такого роду є комп'ютерна програма фірми "Арника Монтана" (м. Тамбов, Росія), що представляє собою інформаційно-пошукову систему (ІПС), пошук інформації в якій проводиться по деревовидній (ієрархічній) схемі, показаної на рис.8.14.

Знайшовши потрібну інформацію, лікар або знайомиться з нею візуально (читаючи текст на екрані монітора), або роздруковує на принтері. При цьому може бути роздрукований тільки один інформаційний "листок" - той, який відображений в даний момент часу на екрані монітора.

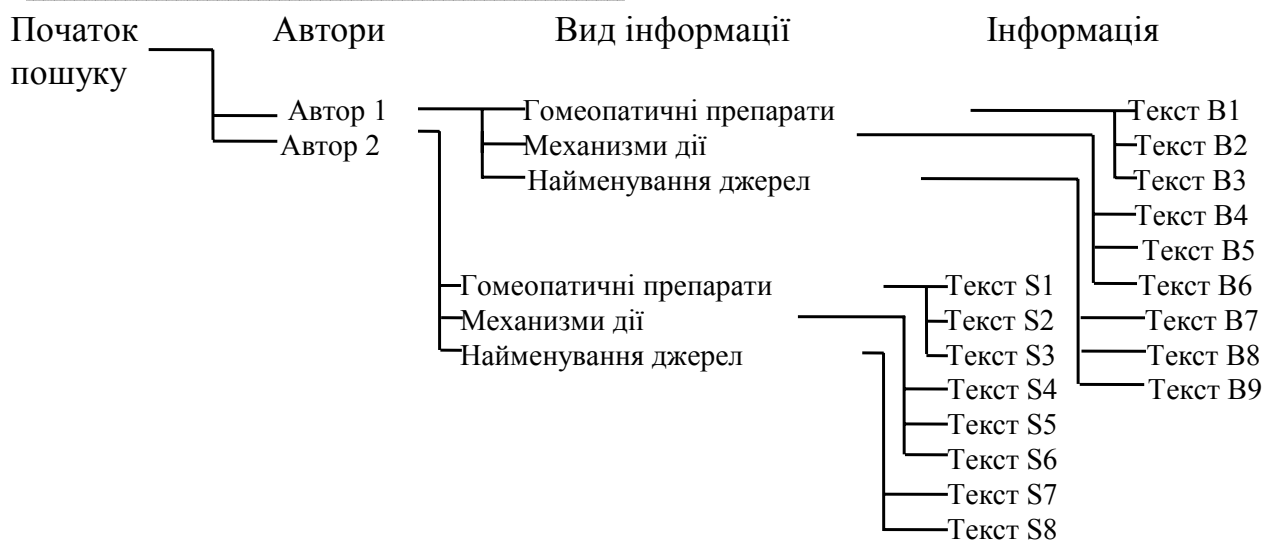


Рис. 8.14

Зразок змісту такого інформаційного листка має, наприклад, такий вигляд (автор = Verice, вид інформації - "гомеопатичні препарати = KALIUM ARSENICUM):

KALIUM ARSENICUM

ЗАГАЛЬНИЙ ОПИС:

Хворий типу **МИШЯКОВОКИСЛОГО КАЛІЯ** має схильність до малігнізації свого хворобливого стану, до переходу шкірних захворювань в застарілі форми. Він стурбований, хвилюється, анемічен.

ШКІРА:

Нестерпний свербіж, який погіршується при роздяганні. Шкіра **СУХА**, луската, поблякла. Акне; пустули, погіршуються під час менструації. Хронічна екзема. Свербіж погіршується **ВІД ТЕПЛА**, при ходьбі, при роздяганні. **ПСОРІАЗ**, **ЛИШАЙ**, фagedонічні виразки. Тріщина в ліктьових та колінних згинах. Подагричні вузли: з погіршенням при змінні погоди. Шкірний рак, при якому раптово без передуючих зовнішніх проявів, раптом розвивається ураження тканин, яке вселяє тривогу. Численні дрібні вузли під шкірою.

ШЛУНОК:

Відрижка з запахом і присмаком зіпсованості яйця. **ЗАВЖДИ ДУЖЕ ГОЛОДНИЙ, МАЄ ХОЧ - НЕБУТЬ З'ІДАТИ СЕРЕД НОЧІ.** Нудота, блювота вагітних.

Біль в шлунку після їжі.

СТІЛЕЦЬ:

Слизистий, **КРОВЯНИСТИЙ, ВКРАЙ СМЕРДЮЧИЙ, РІДКИЙ, ТЕМНИЙ.** Твердий, важко виштовхуючий, з пекучою біллю, із виділенням крові з прямої

кишки, пекучі гемороїдальними вузлами. ЗАПОР (У ДІТЕЙ - ГРУДНИКІВ і у більш старших – хворобливих скрофульоз них дітей).

ВІДНОШЕННЯ:
РАДІЙ

МОДАЛЬНОСТІ:

Погіршення від кофе, хворий «типу Псорика» не отримує полегшення від кофе. Погіршення при змінні погоди, на сонячній спеці, від холоду. БОЇТЬСЯ НАЙМЕНШОГО ХОЛОДУ АБО ПРОТЯГУ. ПОЛЕГШЕННЯ від спеки, тепло-го одягу – навіть влітку.

ДОЗУВАННЯ:

Розведення від III до XXX.

Відмітна особливість таких програм-довідників - робота з інформацією полягає лише в "простому" пошуку і роздруківці потрібних фрагментів інформації. У поняття "простий пошук" при цьому вкладається наступний зміст: процедура пошуку полягає в перегляді вмісту баз даних і виборі потрібного інформаційного листка на підставі аналізу його семантики, виробленого користувачем без допомоги інтелектуального інтерфейсу (іншими словами, інтерфейс не забезпечує які-небудь обчислення, спрямовані на полегшення вибору потрібної інформації).

8.2.1. Комп'ютерна програма «REPER» для підбору гомеопатичних препаратів

Програма REPER, розроблена фахівцями Київського НВО "Славутич" і 25-ї поліклініки м. Києва, містить більш складний, порівняно з наведеним вище прикладом, механізм роботи з інформацією і призначена не стільки для заповнення прогалин у знаннях лікаря-гомеопата, скільки для швидкого підбору мінімальної кількості найбільш ефективних (для даного пацієнта) гомеопатичних препаратів.

8.2.2. Методика призначення гомеопатичних препаратів за допомогою програми REPER

Дана методика полягає в наступному:

1) пацієнт заповнює особисту карточку, фіксуючи в ній скарги, перенесені операції і захворювання, хвороби батьків, а найголовніше - перераховуючи свої симптоми і відчуття за допомогою тест - карт трьох видів: А, В і С (форма особистої карточки показана на рис. 8.15);

2) лікар за допомогою програми REPER вводить в пам'ять комп'ютера коди симптомів, перерахованих в особистій карточці пацієнта (графічний інтерфейс програми REPER зображений на рис.8.16);

ОСОБИСТА КАРТОЧКА

Прізвище.....Ім'я.....По-батькові..... Дата.....

СКАРГИ
(перерахувати скарги за останні 2-3 неділі)

.....

An vitae

(перерахувати перенесені захворювання та операції в хронологічному порядку.
Приклад: 1978р.- Видалення ж/п; 1980р. – гип.б-нь та ін.)

.....

Сімейний анамнез (перерахувати хвороби батьків)

ВІДПОВІДІ НА ТЕСТ-КАРТУ
(Приклад: А: 1.2.10 / 2.3.15 / 4.8.10 і т.і., В: 1.2.1 / 1.2.2 / 2.2.10 і т.і.,
С: 2.1.3 / 2.2.4 / 3.1.1 і т.і.)

Рис. 8.15

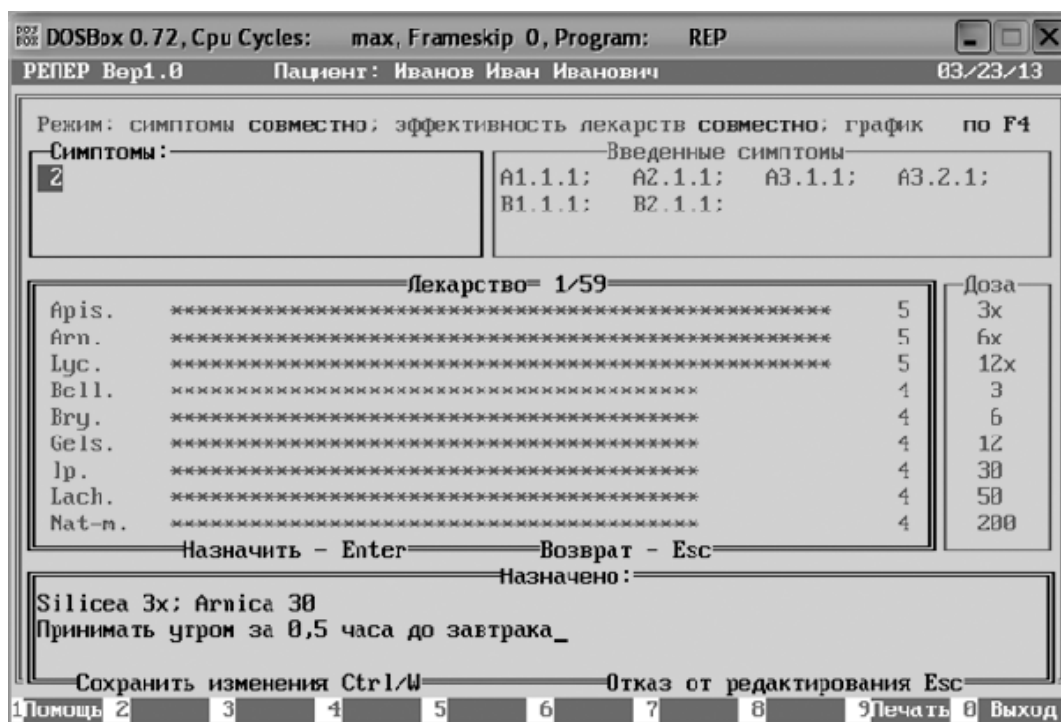


Рис. 8.16

3) після закінчення введення даних лікар за допомогою програми REPER отримує графік частоти зустрічання препаратів, відповідних введеним симптомам, і при виборі препаратів намагається мінімізувати їх кількість, орієнтуючись на препарати, які найбільш частіше зустрічаються - при цьому лікар враховує додаткові відомості про пацієнта, які перераховані в особистій карточці і

уточнені при особистому спілкуванні з пацієнтом, а також враховує активність призначуваних препаратів;

4) лікар виписує рецепт або видає пацієнтові готові препарати, орієнтуючись на результати функціонування програми REPER.

Існує три види тест - карт: А, В і С.

Тест-карта А, іменована "Загальні симптоми", відображає особливості всього організму, тому є найбільш цінною – відомості, які в ній містяться, допоможуть лікарю підібрати гомеопатичні препарати будь-якій людині при будь-якій патології. Рубрика всередині тест-карти розташована в порядку спадання їх значимості для життєдіяльності певного організму. Серед них: риси характеру і інтелект, сновидіння і ставлення до сексу, конгестии і пітливість, характеристика шкірних покривів та ін. Кожна рубрика розбита на кілька різних позицій: відчуття, модальності, етіологія і ін. Це зроблено для того, щоб була можливість вивчити всі індивідуальні відтінки кожного симптому.

Фрагменти тест-карти А:

1. Психологічні симптоми

1.1. Бажання, схильності, звички

1.1.1. Ображати

1.1.2. До перебільшення

1.1.3. До критики, осудженню

1.1.4. Докоряти себе

1.1.5. Докоряти інших

.....

1.1.14. Стогнати, стенати

1.1.15. Сомнамбулізм

1.1.16. Плювати в лице

1.2. Риси характеру

1.2.1. Прихильність, ніжність

1.2.2. Гнівливність, злість

1.2.3. Очікування неприємностей

1.2.4. Тривога, занепокоєння

.....

1.2.44. Вибуховість

1.2.45. Підозрілість

1.2.46. Прострація розумова

І т.ін.

Аналогічно виглядають тест-карти В і С.

Тест-карта В представляє характеристику різних частин людського тіла, а також функціонування органів почуття. Тут і конституційні особливості людей (зовнішній вигляд, вираженню обличчя і ін.), і симптоми ЛОР - патології, офтальмології, ревматології, стоматології та ін.

Тест-карта С містить найбільш зрозумілі для лікарів традиційної школи симптоми і діагнози. Вона цікава для терапевтів широкого профілю, гастроентерологів, пульмонологів, урологів та ін.

Інформація в тест - картці В і, особливо, в С, має другорядне значення при виборі препаратів.

8.2.3. Облік сили впливу гомеопатичних препаратів

Проводиться таким чином. Оскільки використовувані гомеопатичні препарати відносяться до висококонституційним, вони фігурують в більшості рубрик і розрізняються тільки своїми коефіцієнтами (балами):

- бал "3" свідчить про безумовну доцільність призначення;
- бал "2" - сила впливу препарату трохи нижче;
- бал "1" або відсутність балу - засіб може бути призначено для лікування даного, окремо взятого, симптому, але без великої надії на успіх.

Важливою особливістю програми REPER, що вигідно відрізняє її від "паперового" довідника, є можливість швидкого (протягом декількох хвилин) вибору з великої кількості препаратів тільки тих, котрим відповідає один і той самий бал ("3", "2" або "1") - причому одержувані підмножини препаратів автоматично сортуються в порядку зменшення частоти зустрічаємості.

Ще однією важливою особливістю програми REPER є можливість автоматичної або напівавтоматичного дозування призначуваних препаратів. Автоматичне дозування проводиться по правилам, які зазначені у "паперовому" довіднику - аналозі.

При напівавтоматичному дозуванні використовується наступна шкала дозувань:

- 3× – максимальна доза (мінімальне розведення)
- 6×
- 12×
- 3
- 6
- 12
- 30
- 50
- 200
- 500
- 1000
- 10000 – мінімальна доза (максимальне розведення)

8.2.4. Інтерфейс - коротка характеристика

Інтерфейс програми REPER, показаний на рис. 8.16, багатовіконний. Активізація вікон здійснюється натисканням визначених функціональних клавіш. Активізоване вікно містить деяке "меню", строки котрого слід заповнити або просто вибрати для подальшого використання.

Послідовність роботи з вікнами зазвичай така:

- 1) введення симптомів - у вікні "Симптоми";
- 2) перегляд введених симптомів - у вікні "Введені симптоми";
- 3) перегляд повного списку гомеопатичних препаратів, відповідних введеним симптомам, - у вікні " Ліки ";
- 4) виборами призначуваних препаратів з автоматичним їх дозуванням - також у вікні "Ліки";
- 5) напівавтоматичне дозування призначуваних препаратів - у вікнах "Доза" і "Ліки";
- 6) редагування списку призначених препаратів перед виведенням його на принтер - у вікні "Призначення".

Послідовність роботи з вікнами можна змінювати, виходячи з потреб та уподобань користувача. Наприклад, можна ввести спочатку всі симптоми тест-карти А і обрати відповідні препарати, потім ввести симптоми тест-карти В і розширити список призначуваних препаратів, потім аналогічно врахувати симптоми тест-карти С. А можна відразу ввести симптоми для всіх тест - карт, і тільки після цього приступити до обрання і призначення препаратів (цей спосіб найбільш доцільний і тому організовується "за умовчанням", тобто автоматично).

Короткі висновки

Розгляд конкретних експертних та інформаційно - довідкових систем для іридіодіагностики і рефлексодіагностики корисний, оскільки дозволяє скласти уявлення про реальні труднощі, з якими стикаються розробники і користувачі подібних систем.

До теперішнього часу найбільш розвинені два основних напрямки автоматизації праці іридіологів і рефлексодіагностів з використанням комп'ютерних засобів :

- 1) комп'ютерні програми;
- 2) програмно - апаратні комплекси.

У першому випадку мова йде про комп'ютер і комп'ютерну програму, за допомогою яких фахівець здійснює збір та обробку даних з наступним автоматизованим формуванням висновку. У другому випадку для вирішення тих же завдань, крім комп'ютера та комп'ютерної програми, використовуються ще й допоміжні технічні пристрої, що забезпечують введення інформації в комп'ютер з наступними обробкою і зберіганням цієї інформації.

Порівнюючи між собою вказані два напрямки, слід зазначити, що потенційні можливості програмно - апаратних комплексів вище, однак і вартість, в порівнянні з "чисто" програмними засобами, значно вище.

Можливості програмного забезпечення істотно залежать і від тих, хто займається безпосередньою його розробкою - в першу чергу, експертів-лікарів і інженерів по знаннях, що визначають структуру і обсяг баз знань, механізми обробки даних у процесі діагностування. Нарешті, якість кінцевого програмного продукту і час його розробки істотно залежать від кваліфікації системних і функціональних програмістів.

Література до Розділу 8

1. Абрамов М.С. Иридодиагностика с компьютером. – Ташкент, изд. им. Ибн Сины, 1991. - 194 с.
2. Верба А.І., Уланов М.М. Свідоцтво про державну реєстрацію прав автора на твір, ПА#31. Лікувально-оздоровчий реабілітаційний центр «ЛІДІР» та особливості технології надання медичних послуг, яка реалізується в ньому. Дата реєстрації 6 грудня 1995 р.
3. Вельховер Е.С. Клиническая иридология. М., "Орбита", 1992. - 431с.
4. Кривенко В.В., Лисовенко Г.С., Потебня Г.П., Сядро Т.А. Иридодиагностика. Справочник. Киев, "Украинская Советская Энциклопедия" им. М.П. Бажана, 1991. - 133с.
5. Петенко О.В., Гречишникова Н.И. Практическое пособие по иридодиагностике. – Душанбе: Изд-во "ИРФОН", 1991.
6. Продеус А.Н., Сядро Т.А. Опыт разработки компьютерных систем поддержки иридодиагностики. Электроника и связь, вып.2, часть II, 1997. - с.307-311.
7. Продеус А.Н., Сядро Т.А. Проблемы реализации байесовской стратегии при автоматизации иридодиагностики. Электроника и связь, вып.2, часть II, 1997. - с. 312-316.
8. Продеус А.Н. Компьютерная программа "ESID-2" для иридодиагностических исследований. – Народна та нетрадиційна медицина України, Вип.1(2), Київ, 1997. - с. 9-11.
9. Сядро Т.А., Продеус А.Н., Стукалин В.А. Компьютерная поддержка иридолога в практической деятельности, при обучении и в научных исследованиях. Тезисы докладов на II Украинской научно-практической конференции по народной и нетрадиционной медицине. Днепропетровск, 1996 г. - с. 58-59.
10. Продеус А.М., Сядро Т.А., Верба А.І., Зафіріді О.Г., Мартиненко В.О. Свідоцтво про державну реєстрацію прав автора на твір, ПА#262. Комп'ютерна програма "ESID-2". Дата реєстрації 14 жовтня 1996 р.

Контрольні запитання до Розділу 8

1. Вкажіть два основних напрямки автоматизації праці іридологів і рефлексодіагностів.
2. Що являють собою програмно-апаратні комплекси?
3. У чому принципова відмінність програмно-апаратних комплексів від прикладних діагностичних програм?
4. Перерахуйте переваги і недоліки програмно-апаратних комплексів і прикладних діагностичних програм.

Розділ 9

Оболонки експертних та інформаційно-довідкових систем

Зручність «програмних» оболонок як інструментальних засобів важко переоцінити. По суті, з їх допомогою розробник, в ролі якого може виступати навіть кінцевий користувач, може швидко створити такий програмний продукт, який найкраще відповідає його вимогам. У даній главі наведено приклади декількох інструментальних оболонок, що представляють в розпорядження розробника вельми зручний графічний інтерфейс, що дозволяє вирішувати порівняно нескладні завдання діагностичного плану.

9.1. Оболонка експертних систем "Універсал"

ОЕС "Універсал" - досить проста в експлуатації програма для ІВМ-сумісних комп'ютерів, адресована користувачам з різноманітним рівнем теоретичної підготовки, - від наукових робітників до студентів, - і призначена для рішення завдань діагностики, прогнозування, аналізу гіпотез, прийняття рішень в різноманітних проблемних областях, у тому числі в медицині та психології.

При проведенні наукових досліджень, що базуються на накопиченні значних обсягів фактичного матеріалу, оцінюванні статистичних характеристик, моделюванні різних вирішальних (діагностичних) правил, у тому числі з використанням статистики прецедентів, ОЕС дозволяє як протоколювати процес і результати прийняття рішення, так і згодом використовувати результати протоколювання для модернізації баз знань. Приклади використання ОЕС для вирішення науково-дослідних завдань в галузі соціальної та педагогічної психології наведені далі.

Як навчальний засіб, ОЕС «Універсал» дозволяє швидко і на конкретних прикладах освоїти елементи технології розробки ЕС.

ОЕС реалізована мовою програмування СУБД Clipper, експлуатується в середовищі MS DOS.

9.1.1. Можливості та режими функціонування

Перелік основних можливостей ОЕС "Універсал":

1) при створенні баз знань можуть використовуватися як знання людини - експерта, так і фактичний матеріал (базовану на протоколах, створюваних за результатами вирішення конкретних діагностичних завдань);

2) база правил дозволяє здійснити кількісну оцінку достовірно аналізованих гіпотез трьома способами - з використанням Байєсівської стратегії, методом зваженого голосування, а також методом підрахунку балів (лінійне діагностичне правило);

3) можуть бути задані граничні умови для будь-якої класифікаційної ознаки чи групи ознак;

4) при оцінюванні значення ознак можна вказати ступінь достовірності цієї оцінки (ступінь впевненості, ступінь виразності);

5) кількісний аналіз достовірностей гіпотез може бути супроводжений аргументацією та видачею рекомендацій;

6) можливе використання графічного та музичних ілюстрацій і заставок, що дозволяє надати інтерфейсу розроблюваної експертної системи необхідну естетичну форму.

ОЕС «Універсал» має два основних режими функціонування:

- настройка (навчання, модернізація БЗ);
- обчислення (діагностика, прийняття рішення).

В режимі настроювання здійснюється генерація баз правил, що описуються зошитами виду "умова – ознака – параметр – клас" або тріадами "ознака – параметр – клас", із зазначенням значень умовних імовірностей ознак обраних класів, що необхідно для подальших обчислень достовірностей класів (гіпотез, що перевіряються) за методом Байєса або за методом зваженого голосування. Якщо для обчислень достовірностей класів використовується метод підрахунку балів (лінійне діагностичне правило), тоді замість умовних ймовірностей вказується кількість балів (вага), відповідне тому або іншому значенню ознаки.

В режимі обчислень проводиться оцінка достовірностей класів відповідно до одного з трьох можливих математичних апаратів: метод Байєса, метод зваженого голосування і метод підрахунку балів. Метод Байєса забезпечує коректність результатів при умові статистичної незалежності ознак. Інші два методи допускають залежність ознак.

Окрім 2-х основних режимів є також чотири допоміжних режими:

- 1) діагноз;
- 2) графіка - музика;
- 3) скидання;
- 4) опис.

Режим "Діагноз" потрібний для видачі розгорнутого діагнозу (з аргументацією висновків) і рекомендацій.

У режимі "Графіка - Музика" проводиться настроювання інтерфейсу для надання йому необхідної естетичної форми за допомогою графічних і музичних заставок - ілюстрацій.

Приведення системи в робочий стан після закінчення її розробки або модернізації здійснюється в режимі "скидання".

Режим "Опис" містить стислий опис оболонки "Універсал".

Механізми логічного виводу. При використанні методу Байєса достовірність класів оцінюється відповідно до виразу:

$$P(Y_j / X_i, R_i) = P(Y_j / X_{i-1}, R_{i-1}) \left\{ 1 + \frac{R_i}{R_{\max}} \left[\frac{P(X_i / Y_j)}{\sum_{j=1}^J P(Y_j / X_{i-1}, R_{i-1}) P(X_i / Y_j)} - 1 \right] \right\}, \quad j = 1, \dots, J$$

де $P(Y_j/X_i, R_i)$ – апостеріорна ймовірність j -го класу при умові виміру i -ої ознаки; R_i – коефіцієнт, що характеризує ступінь впевненості користувача в якості вимірів (або ступеня вираженості) i -ої ознаки; R_{max} – верхня межа шкали ступеня впевненості; $P(X_i/Y_j)$ – умовна ймовірність i -ої ознаки при наявності j -го класу.

При використанні методу голосування достовірність класів оцінюється відповідно до виразу:

$$P(Y_j / X_i, R_i) = \frac{i-1}{i} P(Y_j / X_{i-1}, R_{i-1}) + \frac{1}{i} P(Y_j) \left\{ 1 + \frac{R_i}{R_{max}} \left[\frac{P(X_i / Y_j)}{\sum_{j=1}^J P(Y_j) P(X_i / Y_j)} - 1 \right] \right\}, j = 1, \dots, J,$$

де $P(Y_j)$ – апостеріорна ймовірність j -го класу.

При використанні методу підрахунку балів достовірність класів оцінюється відповідно до виразу:

$$S_i = S_{i-1} + B_i,$$

де S_i – сума балів при завданні i -ої ознаки ($S_0 = 0$);

B_i – кількість балів, відповідне i -ій ознаці.

Отримувані в результаті обчислень оцінки достовірностей класів відображаються на екрані у вербально-графічній та числовій формі.

Настроювання оболонки "Універсал". Режим настроювання потрібний для створення і модернізації бази знань. В процесі настроювання відбувається навчання майбутньої ЕС:

- 1) вибирається механізм логічного виводу (один з трьох можливих);
- 2) формується шкала ступеня впевненості (задається R_{max});
- 3) присвоюється ім'я майбутньої ЕС (найменування розв'язуваної задачі);
- 4) присвоюються імена підзадач, на які розбивається розв'язувана задача;
- 5) для кожної підзадачі формується список ознак, використовуваних при діагностиці;
- 6) формується шкали значень для кожного з ознак;
- 7) зазначаються обмежувальні умови вимірювань кожної ознаки або групи ознак (виконання котрих гарантує ефективне використання ознак);
- 8) формується список класів (гіпотез);
- 9) формується правила типу "якщо ... то ..." (продукційні правила), що описують зв'язок між ознакою і класом на рівні байєсівського формалізму або на рівні простого підрахунку балів.

Експлуатація оболонки "Універсал". У режимі обчислень здійснюється визначення ступеня достовірності класів за оцінками ознак, що вводиться користувачем. Взаємодія користувача з ЕС має вигляд діалогу: користувач відповідає на аналогічні питання системи, користуючись сукупністю меню багатомовного інтерфейсу. Черговість використання меню близька черговості процедур настроювання ОЕС (див. вище).

У режимі обчислень здійснюється як "пряме" направлення міркувань за схемою "ознака-класи", так і "зворотне" напавлення за схемою "клас-ознаки".

Особливістю зворотного направлення міркувань є можливість зібрати в один список всі ознаки різноманітних підзадач, які характеризують вибраний клас.

"Прямий" напрямок міркувань доцільно використовувати в початковій стадії діагностування, коли необхідно виявити основний напрямок пошуку. Достовірності гіпотез, що перевіряються, при цьому оцінюються грубо. "Зворотній" напрямок міркувань - це допоміжний прийом, що дозволяє сконцентруватися лише на ознаках найбільш "підозрілої" гіпотези. Як тільки список таких ознак сформований, доцільно знову повернутися до "прямого" напрямку міркувань і остаточно підтвердити (або спростувати) справедливості гіпотези, що перевіряється.

Зразковий вид екрану дисплея при функціонуванні ОЕС "Універсал" в режимі "Обчислення" зображений на рис.9.1 (фрагмент рішення задачі орієнтованого прогнозування стану здоров'я новонародженого за результатами оцінювання передпологового стану матері і з урахуванням особливостей протікання пологів).

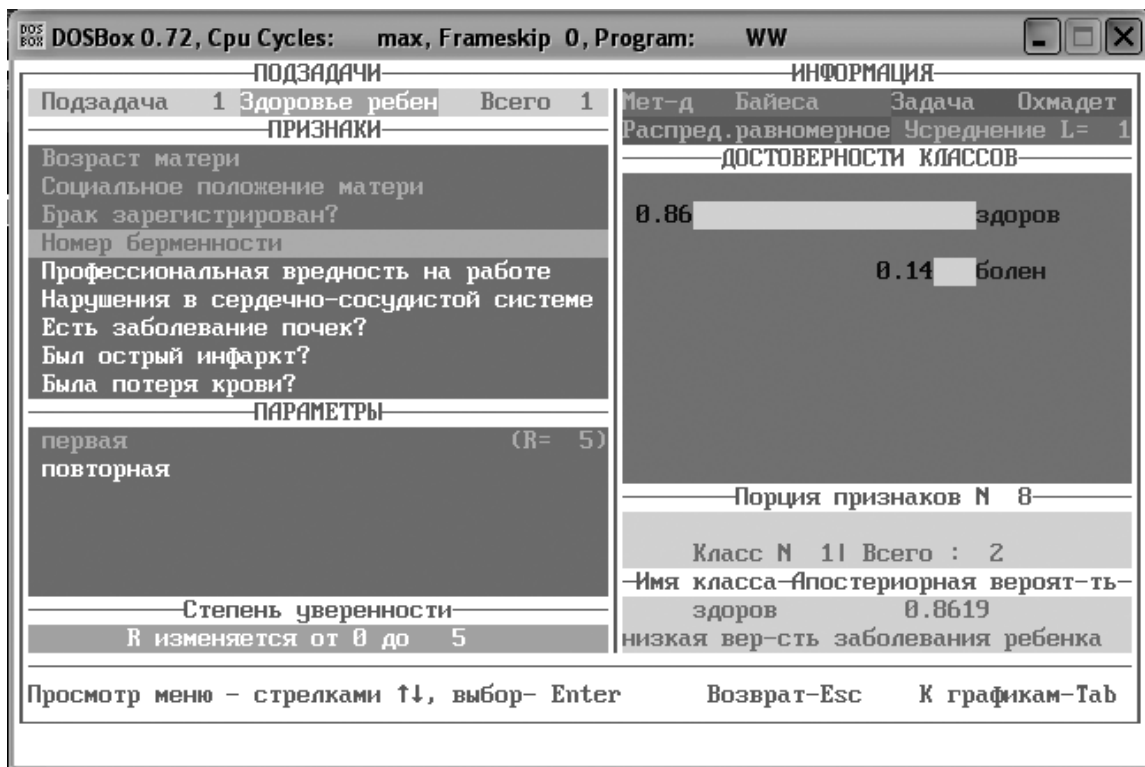


Рис. 9.1

Управління законом апріорного розподілу ймовірностей. Таке управління дозволяє задати значення функції $P(Y_j)$ або, що те ж, функції $P(Y_j/X_i, R_i)$ при $i = 0$. Така можливість важлива, оскільки дозволяє, в принципі, скоротити кількість врахованих діагностичних ознак без зниження достовірності одержуваних висновків, або підвищити достовірність висновків при тій же кількості ознак.

Управління здійснюється в режимі "Налаштування" по запиту програми. У разі відмови від завдання функції $P(Y_j)$ її значення формуються автоматично -

при цьому вибирається рівномірний закон розподілу апіорних ймовірностей.

Режим "Діагноз". Цей режим забезпечує виконання таких функцій:

1) уявлення розгорнутого діагнозу на екрані, а також виведення його на друк;
 2) виведення на друк повних списків ознак, параметрів, умов і класів, що використовувалися при вирішенні даної задачі;

3) редагування класифікації рекомендацій, що видаються в результаті.

Зразковий вид дисплея комп'ютера в режимі "Діагноз" зображений на рис.9.2.

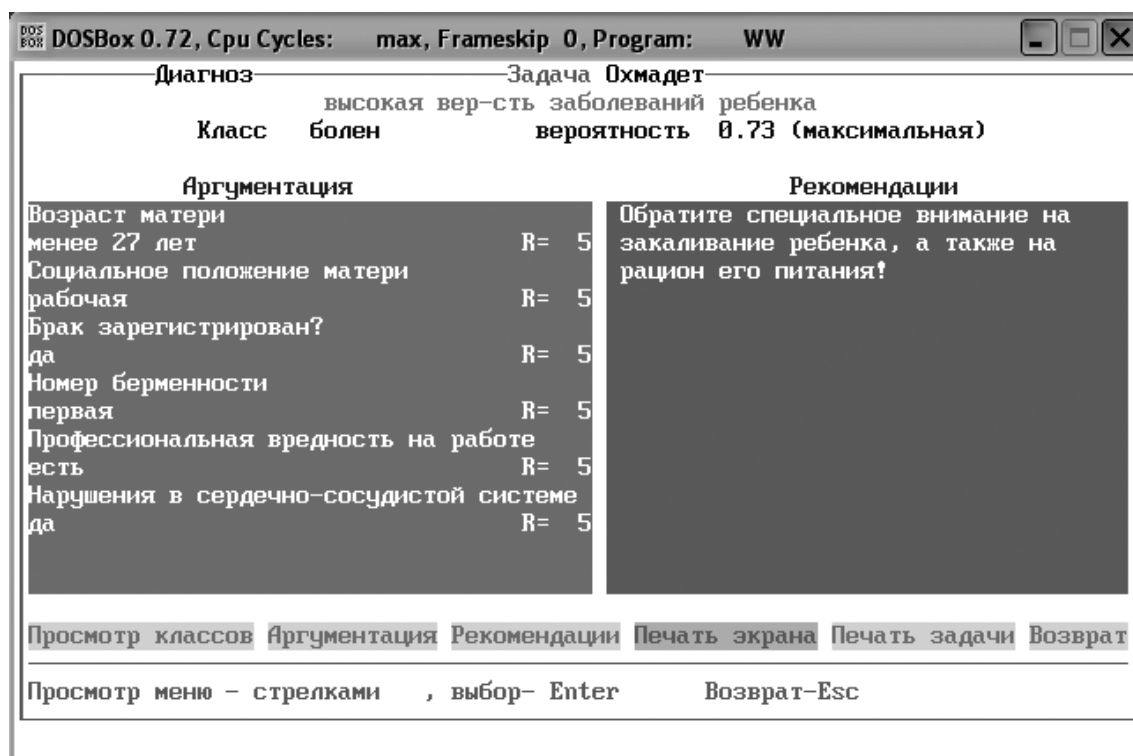


Рис.9.2

Робота з протоколами фактів. Мета протоколювання - накопичення в базі даних, що іменується протоколом, сукупностей ознак і їх значень, що спостерігалися при вирішенні деякої задачі діагностики. Протокол складається з порцій - сукупностей ознак, їх значень і відповідних результатів, що спостерігалися при вирішенні задачі класифікації з новими вихідними даними (наприклад, порцією ознак є результати обстеження одного пацієнта, результати обміру параметрів одного пристрою, що перевіряється і т.п.).

Сформовані протоколи фактів можуть піддаватися огляду і редагуванню, завдяки чому користувач може:

1) перевірити правильність реєстрації фактів і подій і, у разі необхідності, відредагувати протокол;

2) познайомитися з ознаками і фактами деякої проблемної області - у цьому випадку ЕС "Універсал" є "вчителем", який навчає на прецедентах, тобто "випадках з життя";

3) осягти різні механізми "розкладання знань по полицках", що базується на трьох закладених в ЕС "Універсал" стратегіях прийняття рішення, найчастіше використовуваних при розпізнаванні образів: байєсівської, зваженого голосування та підрахунку балів;

4) осягти механізми прийняття рішення як би "граючи", що робить ЕС "Універсал" доступною особам з рівнем освіти, що вельми розрізняється: від вчених-дослідників до студентів і навіть школярів.

Зібрані в протокол факти можуть піддаватися також статистичній обробці, завдяки чому користувач може оцінити частоту зустрічальності того чи іншого признаку для вибраного класу, користуючись фактами з даного протоколу.

В результаті такої процедури:

1) користувач може зрівняти статистичні дані, властиві різноманітним протоколам;

2) користувач може навчити або довчити ЕС, використовуючи статистичні дані з протоколів;

3) користувач може просто задовольнятися отриманими статистичними даними, розглядаючи їх як самостійний цінний результат.

Модернізація раніше створеної експертної системи. Здійснюється в режимі "Налаштування" і мало чим відрізняється від створення нової експертної системи (див. вище) – лише з тою природною відмінністю, що обсяг роботи, як правило, менше (багато ознак і шкали їх значень вже задані).

При цьому "стара" ЕС може бути збережена; нова, модернізована ЕС, організується шляхом автоматичного копіювання баз даних "старої" ЕС - користувачеві для цього потрібно лише привласнити нову назву модернізованої ЕС (у відповідь на відповідний запит програми "Універсал").

Перспективність оболонок типу "Універсал". За допомогою ОЕС "Універсал" було вирішено кілька практично корисних завдань медичного і психологічного прогнозування та діагностики:

1) медична діагностика: прогнозування стану здоров'я новонародженого за результатами оцінювання стану здоров'я матері в період вагітності (див. рис.9.1 - 9.2);

2) психологічна діагностика: прогнозування успішності учнів середньої школи за результатами їх психологічного тестування (рішення цього завдання див. п.9.1.2);

3) соціологічна діагностика: оцінювання ступеня задоволеності життям за методикою американського психолога М. Аргайла.

Досвід практичної експлуатації ОЕС "Універсал" свідчить про перспективність розвитку ідеї створення оболонок такого типу. Враховуючи вдосконалення комп'ютерних технологій, представляється доцільною розробка аналогічних ОЕС для експлуатації в середовищі Windows, на базі СУБД типу Paradox, Delphy тощо, що дозволить істотно розширити можливості таких ОЕС.

9.1.2. Приклад рішення задачі прогнозування з використанням ОЕС "Універсал"

Зазвичай метою тестування учнів (студентів, школярів) є пряме оцінювання рівня їхніх знань. Відмінна особливість психологічного тестування полягає в тому, що оцінюється не рівень знань, а здатність ефективно їх обробляти: сприймати, запам'ятовувати, перетворювати і передавати.

Цілком природно припустити, що рівень знань учнів вельми тісно пов'язаний з їх здатністю обробляти знання. Існує ряд робіт, в яких наведені результати таких досліджень, в ході яких отримано кількісні дані, що підтверджують справедливість цього припущення і дозволяють вирішувати корисні практичні завдання. Зокрема - прогнозувати успішність учнів за результатами їх психологічного тестування. Викладемо коротко ці результати. При цьому звернемо спеціальну увагу на два моменти:

- місце і роль ОЕС «Універсал» у вирішенні даної задачі прогнозування;
- які компоненти природного інтелекту є визначальними для досягнення високої успішності.

Досить великий обсяг результатів психологічного тестування учнів (102 учні) був отриманий при прийомі їх в 9-й клас технічного ліцею Шевченківського району м.Києва восени 1992 і 1993 рр.. - При цьому результати були добре організовані завдяки використанню локальної мережі ПВМ "Корвет" (12 комп'ютерів) з програмним забезпеченням у вигляді пакету "АНІМА" (розробник пакету - психолого-педагогічний центр "АНІМА", м. Мінськ, 1991р.).

Статистична обробка результатів тестування проводилася на ІВМ - сумісному комп'ютері з програмним забезпеченням у вигляді ОЕС "Універсал".

Схема організації збору та обробки експериментальних даних представлена на рис.9.3.



Рис. 9.3

З психологічних тестів пакету "АНІМА" використано 2 групи тестів: "Особливості мислення" і "Тип характеру". У першій групі були тести "Виключення понять", "Аналіз відносин понять", "Складні асоціації", "Відшукування закономірностей" і "Відшукування чисел". Друга група складалася з тестів "Екстраверсія", "Ригідність", "Збуджуваність", "Темп реакцій", "Активність" і "Щирість".

Для статистичної обробки даних були сформовані дві вибірки результатів психологічного тестування: навчальна містила результати тестування 76 учнів, а контрольна - 26 учнів. Серед учнів обох вибірок були представники всіх "класів": "відмінники", "четвірочники", "трієчники" і "двієчники". Належність учнів до того чи іншого класу визначалася за допомогою експертів, роль яких виконували три викладача: інформатики, фізики та хімії.

Розглянемо тепер результати експериментальних досліджень.

Оцінювання ступенів статистичної залежності тестів пакету "АНІМА" шляхом обчислення коефіцієнтів регресії і коефіцієнтів кореляції показало, що результати тестування між собою практично незалежні (коефіцієнт кореляції не перевищує 0.28) і, як наслідок, використання рекурентної форми формули Байєса є коректним.

Зробити висновки щодо принципової можливості розв'язання поставленої задачі можна, аналізуючи результати навчання ОЕЕС "Універсал" по навчальній вибірці, представлені в табл. 9.1 у вигляді оцінок розподілів ймовірностей $P(X_i/Y_j)$, де $Y_j = 2,3,4,5$ - рівні успішності учнів, згрупованих у чотири групи з номерами $j = 2,3,4,5$, відповідно; X_i - досліджувані фактори зі значеннями рівнів "низький", "середній", "високий", причому $i = 1,2,\dots,5$ для групи тестів "Особливості мислення" і $i = 6,7,\dots,11$ для групи тестів "Тип характеру".

Дані табл. 9.1 дозволяють розглядати регресію кожного з факторів на шкалу класів і таким чином, по-перше, зіставити фактори по інформативності, а по-друге, оцінити характер зв'язку між факторами і "класами" і скласти психологічні "портрети" представників кожного з "класів". Так, наприклад, у групі тестів "Особливості мислення" найбільш інформативні результати тестів "Аналіз відносин понять" і "Складні асоціації", а в групі тестів "Тип характеру" - результати тестів "Екстраверсія" і "Збуджуваність". Найменш інформативними представляються результати тестів "Активність" і "Щирість".

Що стосується характеру зв'язку між факторами і класами, неважко бачити, що криві регресії факторів групи "Особливості мислення" досить добре апроксимуються чітко нахиленими до осі "класів" прямими лініями, тоді як для факторів групи "Тип характеру" маємо або істотно нелінійні залежності ("збудливість", "Екстраверсія"), або близькі до лінійних, практично паралельні осі "класів".

Таким чином, отримані результати свідчать про принципові можливості розв'язання задачі прогнозування успішності учнів за результатами їх психологічного тестування, а також свідчать про меншу інформативність групи тестів "Тип характеру" (порівняно з тестами групи "Особливості мислення").

Виходячи з отриманих залежностей, маємо, наприклад, такі психологічні портрети "відмінників" і "двієчників". Для "відмінників" характерні високий рівень комбінаційних особливостей, здібностей до абстрагування, структурування і класифікації, встановлення відносин між кількома закономірностями, добре розвинене асоціативне мислення. Характер "відмінників" відрізняється непоганими гнучкістю, швидкістю реакцій, самозаглибленість, середнім рівнем збудливості.

Таблиця 9.1

Номер теста i	Найменування теста (фактора X_i)	Рівень по шкалі значень теста	$P(X_i/Y_1)$	$P(X_i/Y_4)$	$P(X_i/Y_3)$	$P(X_i/Y_2)$
			«відмін.»	«четв.»	«трієчн.»	«двієчн.»
1	Виключення понять	Низький	0.23	0.10	0.49	0.86
		Середній	0.50	0.80	0.39	0.14
		Високий	0.27	0.10	0.12	0.0
2	Аналіз відносин понять	Низький	0.05	0.30	0.67	1.0
		Середній	0.23	0.10	0.24	0.0
		Високий	0.72	0.60	0.09	0.0
3	Важкі асоціації	Низький	0.05	0.0	0.18	0.79
		Середній	0.32	0.60	0.55	0.21
		Високий	0.63	0.40	0.27	0.0
4	Відшукування закономірностей	Низький	0.41	0.50	0.55	0.93
		Середній	0.41	0.50	0.36	0.07
		Високий	0.18	0.0	0.09	0.0
5	Відшукування чисел	Низький	0.41	0.40	0.49	0.71
		Середній	0.54	0.60	0.48	0.29
		Високий	0.05	0.0	0.03	0.0
6	Екстраверсія	Низький	0.29	0.0	0.11	0.17
		Середній	0.53	0.57	0.48	0.33
		Високий	0.18	0.43	0.41	0.50
7	Ригідність	Низький	0.35	0.12	0.15	0.17
		Середній	0.35	0.50	0.44	0.50
		Високий	0.30	0.38	0.41	0.33
8	Збуджуваність	Низький	0.29	0.63	0.34	0.17
		Середній	0.42	0.25	0.44	0.33
		Високий	0.29	0.12	0.22	0.50
9	Темп реакції	Низький	0.12	0.0	0.07	0.09
		Середній	0.53	0.50	0.63	0.66
		Високий	0.35	0.50	0.30	0.25
10	Активність	Низький	0.12	0.0	0.03	0.16
		Середній	0.18	0.37	0.34	0.09
		Високий	0.70	0.63	0.63	0.75
11	Щирість	Низький	0.06	0.25	0.18	0.0
		Середній	0.41	0.37	0.41	0.42
		Високий	0.53	0.38	0.41	0.58

Психологічний портрет типового "двієчника", як і слід було очікувати, в значній мірі є антиподом портрета "відмінника" - дані табл. 9.1, характеризують їх розумові здібності ($i = 1, \dots, 5$), підтверджують цей висновок. Разом з тим розгляд даних, що відповідають значенням $i = 6, 7, \dots, 11$, свідчать, що за характером відрізнити "двієчника" від "відмінника" вельми непросто: серед "двієчників"

зустрічається досить багато цілеспрямованих (33%), щирих (58%) і дуже багато активних (75%). Дещо полегшує завдання той факт, що "відмінники" краще збалансовані за показниками "екстраверсія" та "збудливість", так як у більшості своїй відносяться до амбавертам (53%) з середнім рівнем збудливості (42%), тоді як 50% "двієчників" є екстравертами з високим рівнем збудливості.

Була проведена експериментальна оцінка можливої ефективності прогнозування з використанням ОЕС "Універсал", шляхом обробки контрольної вибірки (26 учнів) по п'яти алгоритмам (перші три з них реалізовані з використанням ОЕС "Універсал"):

- 1) байєсова стратегія;
- 2) зважене голосування;
- 3) підсумовування балів;
- 4) вибір найближчого еталону в метриці простору L_1 :

$$\rho_{1j} = \sum_{i=1}^{11} |X_i - E_{ij}|$$

де E_{ij} - вектори-еталони;

- 5) вибір найближчого еталону в метриці простору L_2 :

$$\rho_{2j} = \left[\sum_{i=1}^{11} (X_i - E_{ij})^2 \right]^{1/2}$$

Підсумком проведених досліджень є наступні висновки:

1) завдання прогнозування успішності учнів за результатами їх психологічного тестування можна вважати вирішеною: при використанні байєсівської стратегії правильний прогноз мав місце в 54-65% випадків, причому більш високу якість прогнозу досягалося при завданні істотно нерівномірного апріорного розподілу апріорних ймовірностей (у припущенні, що «четвірочки» і «трієчники» складають переважну більшість учнів);

2) використання байєсівської стратегії призводить до найкращих результатів прогнозування; алгоритм складання балів може забезпечити непогану якість прогнозування, однак за умови обмеження тестами групи "Особливості мислення"; порівняно непогані результати прогнозування забезпечує використання алгоритмів, які базуються на виборі найближчого еталону в метриці простору L_1 або L_2 (завдяки відносній простоті вони можуть використовуватися навіть для ручних розрахунків);

3) використання тестів групи "Тип характеру" хоча і менш ефективно для прогнозування, в порівнянні з тестами групи "Особливості мислення", однак воно помітно сприяє збільшенню ефективності прогнозування;

4) у групі тестів "Особливості мислення" найбільш інформативними виявилися результати тестів "Аналіз ставлення понять" і "Складні асоціації", а в групі тестів "Тип характеру" - результати тестів "Збуджуваність" і "Екстраверсія".

Наведені вище результати свідчать, що ОЕС "Універсал" є зручним і ефективним інструментом для експериментальних психологічних досліджень.

9.1.3. Валеоаналіз уроку з використанням ОЕС «Універсал»

Перш ніж розглянути завдання створення ЕС для валеоаналіза занять у школі, зробимо кілька попередніх зауважень.

Зазначимо існування двох категорій користувачів ОЕС «Універсал»:

- випробувачі (дослідники);
- випробовувані (обстежувані особи).

У першому випадку ОЕС «Універсал» використовують як генератор вербальних медичних тестів, як засіб протоколювання результатів обстежень пацієнтів з подальшою статистичною їх обробкою, як інструмент формування баз знань.

У другому випадку «Універсал» використовують як базу знань з інтерфейсом у вигляді безлічі вербальних тестів, що дозволяє отримувати відповіді на питання типу «що, якщо ...» (наприклад, якщо мова йде про прогнозування інсульту, пацієнт може отримати відповідь на питання, наскільки знизиться ризик інсульту, якщо йому вдасться знизити систоличний або діастолічний тиск, усунути психотравмуючі ситуації, відмовитися від сигарет і алкоголю, і т.ін.).

Побудова консультаційних центрів з використанням ОЕС «Універсал» - завдання, що не вимагає значних матеріальних витрат і, в принципі, легко здійсненне в силу наступних причин. По-перше, експериментально доведені надійність і ефективність ОЕС «Універсал» - вимоги до програмно-апаратного забезпечення при цьому мінімальні (практично будь-який сумісний комп'ютер). По-друге, методики діагностики та прогнозування досить численні, досить добре, як правило, верифіковані і цілком доступні (завдяки мережі бібліотек).

Прикладами методик такого роду є роботи з валеоаналізу уроку, різні методики використання валеологічних знань при самоорганізації здорового способу життя, методика ранньої діагностики судинних захворювань мозку, включаючи методики прогнозування мозкових інсультів, і т.ін.

Таким чином, ще одна область застосування ОЕС «Універсал» - валеологія, яку можна розглядати як продукт взаємодії медицини та обов'язкової середньої освіти. Тут «Універсал» можна використовувати як генератор валеологічних тестів, з подальшим практичним їх використанням адміністративно-викладацьким складом шкіл, учнями.

А тепер розглянемо невеликий приклад, що демонструє принципову можливість побудови таких, порівняно простих, експертних систем - розглянемо завдання побудови експертної системи, що допомагає оцінити, з погляду валеології, якість організації занять у школі.

Очевидно, навчально-виховний процес у школі повинен будуватися на основі загальнопедагогічного принципу дбайливого ставлення вчителів до здоров'я учнів. Ступінь реалізації цього відношення в навчально-виховному процесі можна визначити на підставі таких критеріїв як порядок і дисципліна в класі, психологічна обстановка, умови і організація навчального процесу і т.ін. Пере-

лік цих критеріїв є, по суті, переліком діагностичних ознак, фрагмент якого наведено в табл. 9.2 (всього таких ознак 15).

Таблиця 9.2

Ознаки	Значення ознак	Бали
1. Організація класу	• порядок, дисципліна, вчитель володіє ситуацією	+3
	• неорганізованість початку уроку, але вчитель володіє ситуацією	+1
	• незначні порушення дисципліни	0
	• порушення порядку і дисципліни, шум	-1
	• постійне порушення дисципліни, шум, вчитель не володіє ситуацією	-2
2. Психологічний настрій учнів на початку уроку	• є	+3
	• нема	-2
3. Доброзичливість вчителя	• є	+3
	• нема	-2
4. Тепловий режим	• витримується	+3
	• порушується	-1
5. Повітряний режим	• клас досить провітрюється	+3
	• задушливе повітря	-3

Оскільки спочатку оцінку значень ознак передбачено проводити в балах, при налаштуванні ОЕС «Універсал» доцільно обрати механізм логічного висновку у вигляді підсумовування балів. Рішення про якість уроку (валеологічну оцінку уроку) приймають, налаштовуючи «Універсал» на перелік класів (видів рішень) відповідно до табл. 9.3.

Таблиця 9.3

Сума балів	Клас	Рекомендації
38-45	Відміна валеооцінка	Продовжувати в тому ж дусі
31-37	Хороший урок	Чи не заспокоюватися на досягнутому
23-30	Задовільна валеооцінка	Необхідно звернути увагу на недоліки
16-22	Незадовільний урок	Організація уроку погано впливає на здоров'я школярів, необхідно обов'язково поліпшити умови проведення уроку
менше 15	Дуже незадовільний урок	Організація уроку дуже погано впливає на здоров'я школярів. Необхідно терміново покращити умови проведення уроку

9.1.4. Використання ОЕС «Універсал» для самоорганізації здорового способу життя

Шкільна програма з валеології передбачає прогнозування учнями, які вивчають цей предмет, стану свого здоров'я на довгі роки. ОЕС «Універсал» може бути використана для вирішення цього завдання.

При цьому застосування ОЕС «Універсал» дозволяє:

- дати учням знання з подолання факторів ризику;
- сформулювати переконання, що жоден лікар не може зробити для людини того, що він може і повинен зробити для себе сам;
- навчитися планувати своє здоров'я на довгі роки.

Особливий інтерес у школярів викликає самостійна дослідницька робота, проведення якої з використанням ОЕС «Універсал» дозволяє, до того ж, водночас ознайомитися і з застосуванням сучасних комп'ютерних технологій. Як приклад такої роботи можна навести дослідження ризику виникнення інфаркту міокарда у родичів і знайомих. В основу цієї роботи може бути покладений американський тест (по Шенкману), фрагмент якого наведено в таблиці 9.4 (всього в тесті 9 ознак).

Таблиця 9.4

Ознаки	Значення ознак	Бали
1. Вік (років)	20-30	0
	31-40	2
	41-50	3
	50-60	4
	більше 60	5
2. Стать	жіночий	1
	чоловічий	2
3. Стрес: чи має спосіб життя стресовий характер?	ні	0
	частково	4
	так	8
4. Спадкові фактори:	відсутність кровних родичів з інфарктом	0
	кривний родич з інфарктом після 60	1
	один кровний родич з інфарктом до 60 років	2
	два кровних родича з інфарктом до 60 років	3
	три кровних родича з інфарктом до 60 років	8
5. Куріння	не курить	0
	10 сигарет на день	2
	20 сигарет на день	4
	30 сигарет на день	6
	40 сигарет на день	8

У даному прикладі, як і в попередньому, спочатку передбачено оцінку значень ознак здійснювати в балах, тому при налаштуванні ОЕС «Універсал» також доцільно обрати механізм логічного виводу у вигляді підсумовування балів. Рішення про ступінь ризику приймають, налаштовуючи «Універсал» на перелік класів (видів рішень) відповідно до табл. 9.5.

Таблиця 9.5

Сума балів	Клас	Рекомендації
до 10	Ризик відсутній	Немає підстав для занепокоєння
11-18	Є тенденція	Дослідити фактори ризику
19-25	Явний ризик	Зменшити фактори ризику
26-32	Серйозний ризик	Консультація кардіолога
більше 32	Дуже серйозний ризик	Спостереження у кардіолога

9.1.5. Рання діагностика судинних захворювань мозку з використанням ОЕС «Універсал»

Мозковий інсульт є однією з найбільш частих причин загибелі та інвалідації сучасних людей. Якщо в минулі роки смертність, обумовлена мозковим інсультом, займала третє місце після ішемічної хвороби серця і онкологічних захворювань, то останнім часом вона перемістилася на друге місце.

При цьому особливо висока смертність серед людей з артеріальною гіпертензією та атеросклерозом - це дозволяє деяким дослідникам і організаторам охорони здоров'я вважати, що в боротьбі з мозковими інсультами основну увагу слід приділяти первинній профілактиці, насамперед профілактиці артеріальної гіпертензії та атеросклерозу. Дотримуючись цієї стратегії, профілактику артеріальної гіпертензії пропонують починати з юнацького, а атеросклерозу - з раннього дитячого віку. При цьому, однак, виникає ціла зв'язка питань. По-перше, кого піддавати такій профілактиці? Якщо всіх, тоді виникає друге питання - яким чином це можна забезпечити? Якщо не всіх, то як знайти осіб, особливо схильних до дії факторів ризику? І нарешті, якщо такі особи знайдені, то які саме профілактичні заходи мають бути вжиті?

Разом з тим, добре відомо, що крім профілактики захворювань (первинної профілактики) існує профілактика їх ускладнень (вторинна профілактика). Оскільки вже тепер десятки мільйонів людей мають атеросклероз, артеріальну гіпертензію, або і те й інше, виникає питання раціональної організації вторинної профілактики судинних захворювань мозку.

Рішення якщо і не всіх, то, принаймні, частини перерахованих вище питань може бути полегшено при використанні ОЕС «Універсал», налаштованої на ранню діагностику таких захворювань та їх ускладнень з урахуванням існуючих верифікованих і досить ефективних методик.

Одна з найпростіших методик подібного роду базується на використанні прогностичної картки, що містить всього 11 факторів ризику - фрагмент цієї карти, призначеної для прогнозування виникнення мозкових інсультів і створеної на основі методу експертних оцінок, наведено в таблиці 9.6.

Таблиця 9.6

Ознаки	Значення ознак	Бали
1. Склероз судин мозку	помірно	2
	значно	3
2. Систолічний тиск	151-160 мм рт.ст.	1
	вище 160 мм рт.ст.	3
3. Діастолічний тиск	100 мм рт.ст. і вище	3
4. Біль в області серця	Частіше 1 раз в неділю	1
5. Зміни судин очного дна	Значні (звуження артерій, їх звивистість і т.ін.)	3

Суть методу експертних оцінок при цьому полягає в наступному. Спочатку експерти складають перелік основних факторів ризику, кожному з яких присвоюють числове значення (кількість балів). Потім на групі навчання верифікують складений перелік, прагнучи досягти 90%-ої точності прогнозу. Остаточні висновки про ефективність методики роблять за результатами апробації переліку на контрольній групі - якщо точність прогнозу перевищує 70%, карта визнається придатною для практичного застосування.

Діагностичне рішення приймають такий спосіб. Бали спостережуваних ознак підсумовують. Якщо у віці 30-50 років сума дорівнює 7, у віці 51-60 років -10, а у віці 61 рік і старше - 12 балам, то таких обстежуваних необхідно віднести до групи підвищеного ризику.

Деяка інша методика використана при розробці іншої прогностичної карти - для прогнозування судинних захворювань мозку і гострих інфарктів у людей з нормальним артеріальним тиском. Числові значення ознак при цьому визначалися як значення оцінок умовних ймовірностей $P(Y_j/X_i)$. Згодом ці значення були «зіпсовані» шляхом заміни цілочисельними значеннями - балами (фрагмент такої карти наведений у табл. 9.7; в цілому карта містить 19 ознак). Однак при бажанні можна повернутися до оцінок умовних ймовірностей $P(Y_j/X_i)$, тоді при налаштуванні ОЕС «Універсал» замість методу підсумовування балів можна задати метод зваженого голосування.

Діагностичне рішення приймають аналогічним чином. Бали спостережуваних ознак підсумовують і констатують велику ймовірність виникнення мозкового інсульту або інфаркту міокарда, якщо сума балів дорівнює 7 і більше.

Значний інтерес представляють висновки фахівців про те, що оскільки причини виникнення мозкових інсультів і гострих інфарктів міокарда ідентичні, то розглянуті вище системи прогнозування майже з однаковою точністю проорокують і те й інше. У результаті немає необхідності створювати самостійну протиінсультну

службу - досить об'єднати зусилля невропатологів і кардіологів у рамках існуючої служби з профілактики хвороб органів кровообігу та їх ускладнень.

Таблиця 9.7

Ознаки	Значення озгпк	Бали
1. Атеросклероз судин головного мозку	I стадія	1
	II стадія	2
	III стадія	4
2. Зміни судин очного дна	У вигляді звуження і звивистості артерій, «Салюс I, II або III», симптоми «мідного дроту», «срібного дроту», ангіопатія сітківки	2
3. Перехідні порушення мозкового кровообігу	У вертебро-базиллярної басейні	2
	У каротидному басейні	3
4. Зміни серця за даними ЕКГ	значні	1
	грубі зміни	3
5. Стенокардія	Почастішання або поява нападів стискаючих болів	2

Додамо до цього, що застосування систем автоматизації прийняття рішення, подібних ОЕС «Універсал», значно підвищує надійність ранньої діагностики мозкових інсультів і гострих інфарктів міокарда. Налаштування ж ОЕС «Універсал» на реалізацію якої нової методики - завдання, легко розв'язувана самими користувачами протягом декількох годин.

9.2. Оболонки інформаційно-довідкових систем МЕД і МЕГ

Проблема ефективного запам'ятовування і подальшого відтворення великих обсягів інформації добре знайома фахівцям, особливо початківцям, в області методів нетрадиційної медицини. Особливо актуальна вона для фахівців, які намагаються опанувати відразу декількома взаємодоповнюючими методами діагностики і лікування, що типово і природно, оскільки зростання кількості незалежних джерел інформації про стан організму пацієнта призводить до збільшення достовірності діагностування, а раціональне поєднання різних методів лікування підвищує його ефективність.

Для вирішення зазначеної проблеми призначена комп'ютерна програма МЕД (середовище MS DOS, мова програмування Clipper), що представляє собою оболонку інформаційно-довідкових систем, призначену для швидкого створення, експлуатації та модернізації інформаційно-довідкових систем практично будь-якого медичного напрямку.

Поряд з основним "гнучким" режимом створення та експлуатації власних довідкових систем, оболонка МЕД містить додаткові "жорсткі" режими "Таблиці" і "Діагностика по Riodoraku", що забезпечують комп'ютерну підтримку

рефлексотерапевта при постановці діагнозу і проведенні відповідного курсу лікування. Такий вибір "жорсткої" частини оболонки обумовлений широкою поширеністю останнім часом методів рефлексотерапії. Вид графічного інтерфейсу в режимі "Діагностика по Riodoraku" зображений на рис. 9.4.

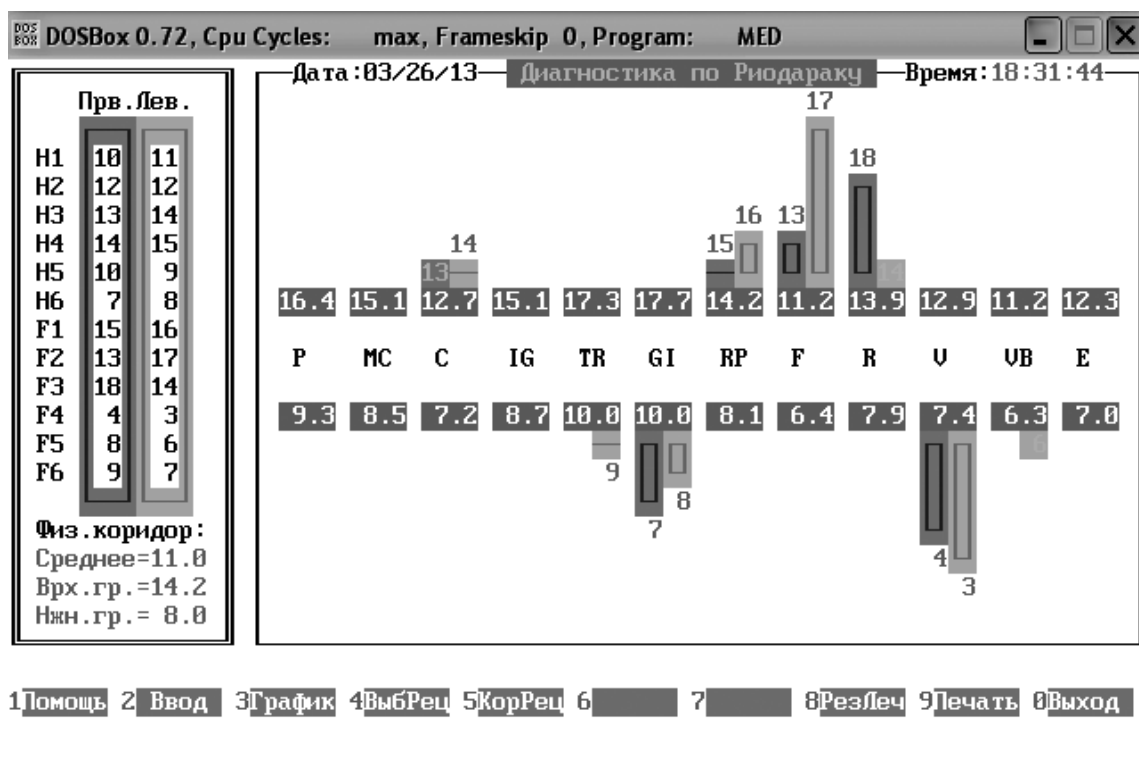


Рис. 9.4

Інша комп'ютерна програма MEG представляє собою оболонку інформаційно-пошукових систем, призначену для створення та експлуатації інформаційно-пошукових систем не тільки медичного, а й інших призначень (насамперед такого роду оболонка потрібна психологам при формуванні висновків та рекомендацій).

Формально оболонка MEG являє собою урізану версію оболонки MED: відсутні режим підтримки діагностування "по Riodoraku" і режим "Таблиці" (інша природне відмінність - змінені імена системи меню і підміню, що пояснюється більшою універсальністю, в плані можливих додатків, оболонки MEG).

Механізм обробки інформації в обох оболонках схематично представлений на рис.9.5, а графічний інтерфейс оболонок в режимі «Робота» показаний на рис. 9.6.

Таким чином, особливістю програм MED і MEG є наявність двох видів обробки:

- 1) об'єднання ("конкатенація") текстів знайдених "інформаційних листків" в один текстовий документ, що являє собою заготовіть висновку лікаря;
- 2) редагування тексту висновку - виділення текстових блоків, переміщення їх, копіювання, видалення, набір з клавіатури нового тексту.

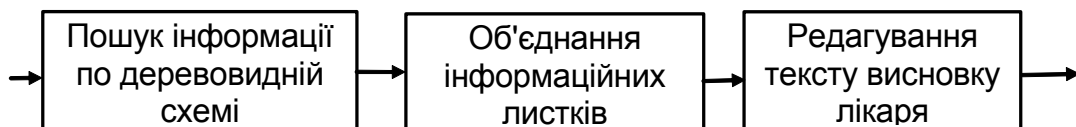


Рис. 9.5

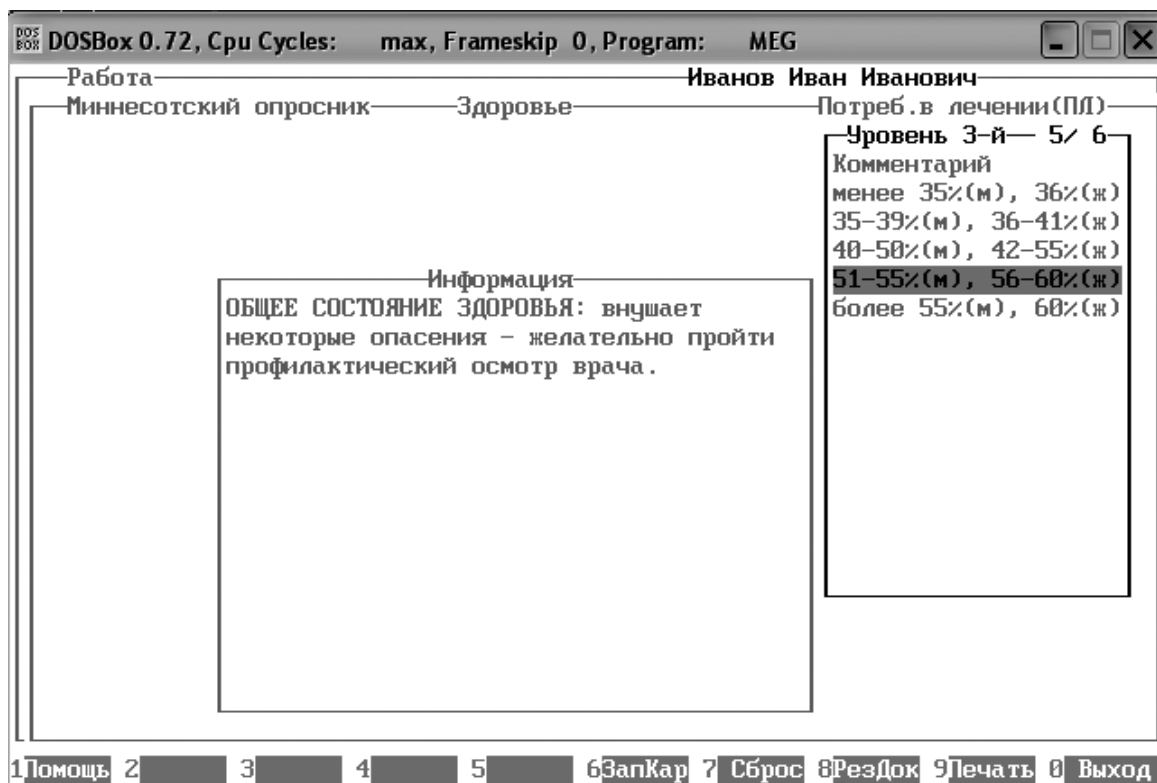


Рис. 9.6

9.2.1. Основні характеристики оболонок

Програму МЕД і МЕГ забезпечують:

- швидкі створення, модернізацію та експлуатацію інформаційно-пошукових систем на базі ІВМ-сумісних комп'ютерів;
- ефективне ведення електронної картотеки з інформацією про результати виробленого пошуку;
- текстове редагування зібраної інформації;
- розпечатку на папері результатів інформаційного пошуку.

Основні режими функціонування:

- робочий (називається "Лікування" в МЕД і "Робота" - у МЕГ);
- налаштувальний ("Налаштування" - в обох оболонках).

У робочому режимі здійснюється експлуатація раніше створеної інформаційно-пошукової системи, а в налаштувальному режимі - створюється нова або модернізується наявна система.

У робочому режимі користувач має можливість зібрати воедино всю його інформацію, здійснюючи пошук підходящої інформації за деревовидною схе-

мою. У першій версії оболонки МЕД це рівні:

- система;
- підсистема;
- захворювання;
- спосіб лікування;
- варіант симптоматики;
- симптоматика;
- лікування;
- рецепт.

В більш пізній, другій версії оболонки МЕД глибина пошуку зменшена: вилучені рівні 4) -6), що сприяло прискоренню пошуку інформації та спрощенню експлуатації самої системи.

Якщо довідник призначений для пошуку рекомендацій з лікування виявлених захворювань, перший рівень може бути заповнений переліком методів лікування ("Рефлексотерапія", "Апітерапія", "Фітотерапія" і т.ін.), другий - переліком систем і органів людського організму, третій - переліком захворювань систем і органів, четвертий - переліком лікувальних заходів, п'ятий - інформацією, яка розкриває сутність цих заходів.

Методика експлуатації такого довідника проста і складається з наступних етапів:

- автоматично, при включенні режиму "Робота", викликається електронна картотека, куди заносять основні дані про пацієнта (П.І.Б., адреса і т.ін.);
- проводять пошук потрібної інформації шляхом послідовного спуску по рівнях деревовидної структури інформаційно-довідкової системи; при досягненні передостаннього рівня на моніторі одночасно відображаються перелік лікувальних заходів та опис цих заходів;
- перекладом курсора на потрібну позицію переліку лікувальних заходів добиваються відображення на моніторі опису заходів;
- натисканням клавіші "Виконання" на клавіатурі комп'ютера "скидають" текст опису обраного лікувального заходу в заключний документ "Рекомендації з лікування";
- при необхідності повторюють пошук відсутньої інформації, розширюючи зміст заключного документа;
- здійснюють редагування заключного документа;
- роздруковують на папері сформовані рекомендації з проведення лікувальних заходів.

Оболонка МEG за кількістю рівнів пошуку збігається з цією останньою версією і відрізняється від неї, як уже зазначалося вище, відсутністю режиму підтримки рефлексодіагностики і рефлексотерапії.

Найменування рівнів у МEG:

- рівень 0 (відповідає рівню "система" в МЕД);
- рівень 1 ("підсистема");

- рівень 2 ("захворювання");
- рівень 3 ("лікування");
- інформація ("рецепт").

У режимі "настройки" користувач має можливість наповнити скелет майбутньої інформаційно-пошукової системи потрібною інформацією. Це фіксована схема, користувач не може змінити її і повинен структурувати інформацію в дозволених рамках. У цьому ж режимі здійснюється корегування (модернізація) вже наявної інформаційно-пошукової системи.

Коли користувач вважатиме систему "готовою до використання", йому достатньо перейти до робочого режиму, щоб змусити її працювати.

9.2.2. Бази даних та знань

Система МЕД являє собою сукупність інтерфейсу і баз даних. Бази даних можна об'єднати в чотири групи:

- 1) системні бази даних "гнучкої" частини оболонки МЕД;
- 2) бази даних "жорсткої" частини оболонки МЕД, призначеної для підтримки рефлексодіагностики "по Riodoraku" та відповідних рекомендацій з лікування (рефлексотерапія);
- 3) бази даних користувача, створювані в процесі настройки і що містять дані по підсистемах і захворювань, іменовані в цілому базою знань;
- 4) бази даних, що містять результати обстежень пацієнтів і представляють собою електронні файли - картотеки.

9.2.3. Особливості інтерфейсу

Робота з оболонками організовано максимально просто: на моніторі комп'ютера по мірі просування по рівнях схеми пошуку інформації відображаються що вкладаються один в одного "вікна", написи на яких вказують на глибину пошуку. Кожне вікно представляє собою якесь "меню", рядки якого слід вибрати для подальшого використання (в режимі "Робота") або заповнити текстом (в режимі "настройки").

Послідовність роботи з вікнами зазвичай така:

- 1) вибір режиму роботи (робочий або настроювальний);
- 2) занурення на $n + 1$ -й рівень пошуку з вікна з n -тим рівнем;
- 3) перегляд знайденого інформаційного листка;
- 4) копіювання вмісту потрібного інформаційного листка в результуючий документ, призначений для виведення на друк;

Якщо не потрібно довго (кілька днів, місяців ...) зберігати атрибути споживача інформації (П.І.Б., адреса, телефон) і результуючий документ із зібраною інформацією, тоді при включенні програми електронна картотека не активується, у користувача при цьому запитуються лише атрибути споживача інформації (їх можна і не вводити).

Якщо тривале зберігання інформації необхідно, тоді потрібна і електронна картотека, при активізації якої доведеться відповісти на кілька запитань комп'ютера, що стосуються місця розташування та імені файлів-розділів картотеки і повідомити відомості про споживача інформації (П.І.Б., адреса і т.п.). Після введення цієї інформації в правій верхній частині екрану відобразяться П.І.Б. споживача, і програма буде готова до виконання основної своєї функції - пошуку інформації.

Зауважимо, що робота з електронною картотекою доцільна не тільки при бажанні зберегти надовго інформацію про результати виробленого інформаційного пошуку, а й при поетапному формуванні результуючого документа.

Текст результуючого документа можна редагувати - для цього є простий текстовий редактор, що дозволяє:

- прибрати одну чи кілька рядків тексту;
- додати в тексті кілька порожніх рядків і вписати в них текст;
- відредагувати текст в обраному рядку;
- записати весь текст у "кишеню" (буфер пам'яті);
- прочитати текст з "кишені".

Створювати і модернізувати інформаційно-пошукову систему дуже легко - робота в режимі "настройки" мало чим відрізняється від роботи в режимі "Робота".

Вид графічного інтерфейсу в режимі «Висновок» показаний на рис. 9.7.

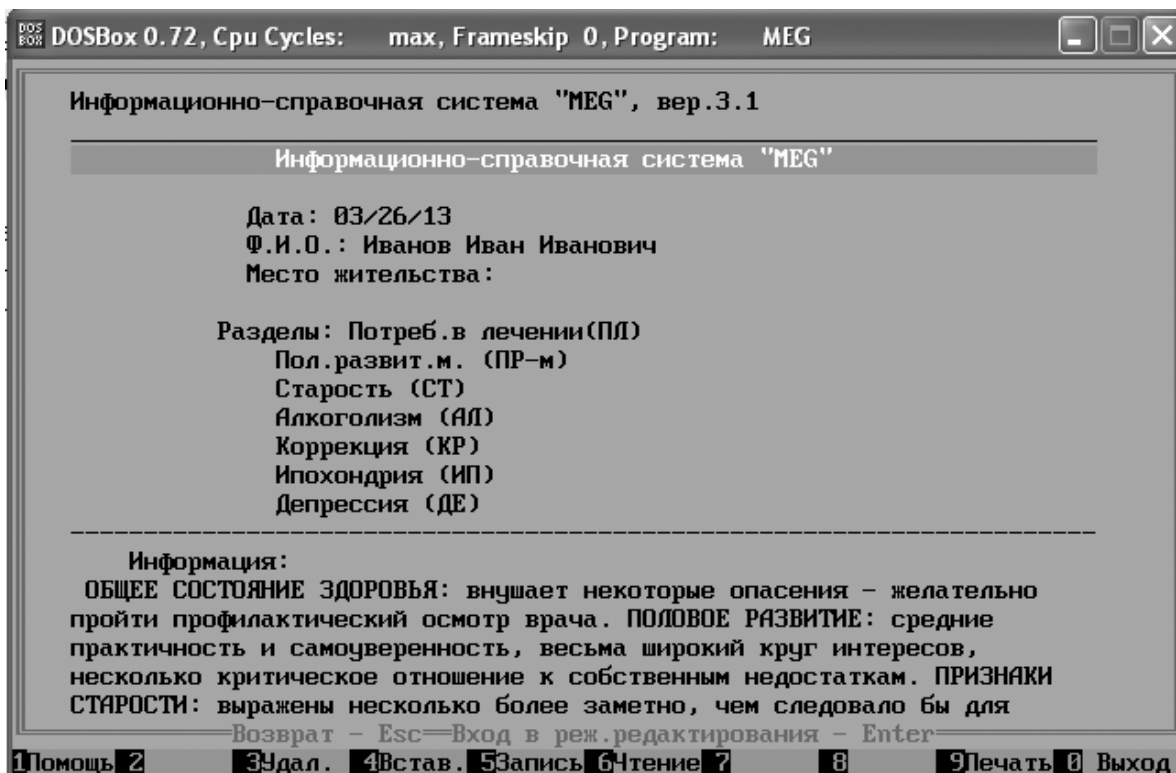


Рис. 9.7

9.2.4. Принципи формування висновку

Цих принципів небагато і вони дуже прості:

- висновок має бути високопрофесійним;
- висновок повинен формуватися швидко;
- процедура формування висновку повинна бути простою і гнучкою, зручною в експлуатації.

Реалізація цих принципів досягається за рахунок таких прийомів:

- заповнення баз даних інформаційно-довідкової системи здійснюється висококваліфікованим фахівцем (зазвичай - самим лікарем-користувачем);
- швидкість пошуку інформації забезпечується деревовидною структурою процесу інформаційного пошуку, а швидкість формування висновку забезпечується можливістю копіювання в текст висновку блоків знайденої текстової інформації (обсяг таких блоків може бути значним);
- гнучкість процедури формування висновку забезпечується поєднанням режиму копіювання готових фрагментів тексту, що зберігаються в базах даних, з режимом введення і редагування текстової інформації за допомогою клавіатури комп'ютера - в останньому випадку можна говорити про наявність простого текстового редактора, що дозволяє реалізувати редагування тексту висновку;
- зручність в експлуатації досягається за рахунок мінімізації обсягу функцій редактора, а також наявністю спеціальної функції "Збереження в буфері", що дозволяє створювати архівну копію висновку на випадок аварійного збою комп'ютера (при раптовому виключенні мережевого живлення, наприклад).

9.2.5. Досвід експлуатації та перспективи використання

Оболонка медичних інформаційно - довідкових систем "МЕД" експлуатується в наступних організаціях МОЗ України:

- Фонд милосердя України;
- 25-а поліклініка м.Києва, кабінет методів нетрадиційної медицини.

Оболонка інформаційно - довідкових систем "MEG" знайшла застосування в таких організаціях:

- Київське ТОВ "Макрос" (для цілей психологічної діагностики);
- ліцей бізнесу м. Києва (для цілей психологічної діагностики);
- 25-а поліклініка м. Києва, кабінет методів нетрадиційної медицини (для іридодіагностики, рефлексодіагностики, апітерапії та інших методів діагностики і лікування).

Досвід експлуатації оболонок типу МЕД і MEG свідчить про корисність і перспективність такого роду систем у різних проблемних областях і в першу чергу - в медицині та психології.

Короткі висновки

Зручність «програмних» оболонок експертних систем як інструментальних засобів полягає в тому, що з їх допомогою розробник, в ролі якого може виступати навіть кінцевий користувач, може швидко створити такий програмний продукт, який найкраще відповідає його вимогам.

Оболонки ЕС мають два основних режиму функціонування :

- 1) налаштування (навчання , модернізація БЗ) ;
- 2) обчислення (діагностика, Прийняття рішення) .

Режим налаштування призначений для створення і модернізації бази знань.

В процесі налаштування відбувається навчання майбутньої ЕС :

- 1) вибирається механізм логічного висновку ;
- 2) формується шкала ступеня впевненості ;
- 3) присвоюється ім'я майбутньої ЕС;
- 4) присвоюються імена підзадач, на які розбивається розв'язувана задача;
- 5) для кожної підзадачі формується список ознак, використовуваних при діагностиці ;
- 6) формуються шкали значень для кожної з ознак;
- 7) зазначаються обмежувальні умови вимірювань кожної ознаки або групи ознак;
- 8) формується список класів (гіпотез) ;
- 9) формуються правила, що описують зв'язок між ознакою і класом на рівні того чи іншого формалізму.

У режимі обчислень здійснюється визначення ступеню достовірності класів за оцінками ознак , що вводиться користувачем.

Може існувати ряд додаткових режимів, спрямованих на підвищення зручності експлуатації ЕС. Серед таких режимів першочергове місце займає режим швидкого формування і редагування висновку, від якості реалізації якого істотно залежить правильність наступних лікувально - реабілітаційних заходів.

Література до Розділу 9

1. Мисюк Н.С. Системы прогнозирования мозговых инсультов. (Методическое пособие). Минский госуд.мед.институт, 1983. - 59с.
2. Трошин Г.И. Методические указания по диагностике заболеваний мозга при массовых обследованиях. (Методическое пособие). Горьковский госуд. мед. университет, 1989. - 89с.
3. Шахненко В.І. Валеоаналіз уроку. – Валеологія (науково-практичний журнал), 1(2)'97 . - с. 22.
4. Волкова С.С. Методика використання валеологічних знань при самоорганізації здорового способу життя. – Валеологія (науково-практичний журнал), 1(2)'97 . - с. 11-14.

5. Гриненко Л.М., Продеус А.М., Шукевич Ю.В., Якубов С.В. Прогнозування успішності навчання учнів за результатами психологічного тестування. – Педагогіка і психологія, 1996, №2, с. 116-122; 1997, №2. - с. 114-123.

6. Продеус А.Н., Сядро Т.А., Стукалин В.А. Медицинская компьютерная информационно-справочная система: быстрое создание, эксплуатация и модернизация. Тезисы докладов на II Украинской научно-практической конференции по народной и нетрадиционной медицине. Днепропетровск, 1996 г.

7. Стукалин В.А., Сядро Т.А., Продеус А.Н. Подбор гомеопатических препаратов с использованием компьютерной поддержки. Тезисы докладов на II Украинской научно-практической конференции по народной и нетрадиционной медицине. Днепропетровск, 1996 г.

Контрольні запитання до Розділу 9

1. У чому полягає головна зручність оболонок ЕС як інструментального засобу?
2. Які основні режими функціонування мають оболонки ЕС?
3. Який сенс вкладається в поняття «навчання ЕС»? З яких етапів складається таке навчання?
4. Які функції виконує режим обчислень в оболонці ЕС?
5. Які додаткові режими можуть існувати у оболонки ЕС, і для чого вони призначені?

Розділ 10 Вибрані питання теорії експертних систем

Теоретичним основам розробки експертних систем були присвячені Розділи 1-4 першої частини цієї книжки, що дозволило скласти цілісне і досить повне уявлення про такий непростий об'єкт як експертні системи. Разом з тим, деякі важливі, з прикладної точки зору, питання не отримали належного освітлення.

У цьому розділі зазначений пропуск певною мірою компенсовано. При цьому особливу увагу приділено таким важливим і складним проблемам як вилучення знань у експертів і обробка експертних оцінок. Крім того, враховуючи широку поширеність байєсівського підходу до створення механізму логічного висновку, наведено ряд додаткових матеріалів, прикладів і завдань, які допомагають краще зрозуміти особливості байєсівського підходу.

10.1. Вилучення знань у експертів

10.1.1. Експертне оцінювання як процес вимірювання

Ефективність початкових етапів розробки ЕС (етапів ідентифікації та концептуалізації) багато в чому визначається успішним формуванням авторитетної групи експертів і отриманням від них якісних знань, що складають основу будь-якої ЕС.

Суть процесу виявлення знань полягає в організації проведення експертами інтуїтивно-логічного аналізу проблемної області з кількісною оцінкою формульованих ними суджень. На цьому етапі експерти:

- формують об'єкти і поняття предметної області (мету, рішення, альтернативні ситуації тощо);
- виробляють вимірювання характеристик (ймовірності звершення подій, коефіцієнти значимості цілей, перевагу рішень і т.д.).

Експертне оцінювання являє собою процес вимірювання, який можна визначити як процедуру порівняння об'єктів за обраними показниками (ознаками). У цьому визначенні фігурують три поняття: об'єкт, показник (ознака), і процедура порівняння.

Об'єктами можуть бути предмети, явища, рішення. В якості показників порівняння можуть використовуватися просторово-часові, фізичні, психічні та інші властивості і характеристики об'єктів. Процедура порівняння включає в себе:

- визначення причинно-наслідкового зв'язку між об'єктами;
- встановлення ступеня впливу одних об'єктів на інші.

Остання обставина вимагає проведення порівняння об'єктів, що визначають будь-який результат, за ступенем їх впливу на нього. Введення конкретних показників порівняння дозволяє, експертам встановлювати відносини між об'єк-

тами, наприклад, «велике», «краще», «більш ніж», «гірше», «однакові», «переважніше» і т.д.

Для формального опису множини об'єктів і відношення між ними вводиться поняття системи з співвідношенням

$$M = \langle O; R \rangle,$$

де $O = (O_1, O_2, \dots, O_n)$ - множина об'єктів предметної області; $R = (R_1, R_2, \dots, R_n)$ - множина відношень між ними.

Запис $O_i R_k O_j$ означає, що об'єкт O_i знаходиться в співвідношенні R_k до об'єкту O_j . Таке відношення називається бінарним (двомісним), оскільки зв'язую два об'єкта. Відношення можуть бути тримісними і т.д. В загальному випадку k - місними, по числу об'єктів, які вони зв'язують. Визначимо основні властивості відношень:

Відношення R *рефлексивне*, якщо $O_i R O_j$ - істинно;

Відношення R *симетричне*, якщо із $O_i R O_j$ слідує $O_j R O_i$;

Відношення R *транзитивне*, якщо із $O_i R O_j$ і $O_j R O_k$ слідує $O_i R O_k$.

Відношення, яке володіє властивостями рефлексивності, симетричності і транзитивності називається відношенням *еквівалентності*:

$$O_i \sim O_j.$$

При експертному оцінюванні, крім відношення еквівалентності, використовується відношення *порядку*. Це відношення може означати, наприклад: «раніше ніж», «більше ніж», «більше ніж», «переважніше ніж» і т.д. Відношення порядку *антірефлексивне*, *транзитивне*, позначається:

$$O_i \succ O_j$$

для випадку, коли O_i переважніше O_j .

Експертне оцінювання в процесі виявлення знань від експертів, використовує поряд з емпіричною, також числову систему з відношенням

$$H = \langle N; S \rangle,$$

де N - множина дійсних чисел, а S - множина відношень між числами. Саме ця система дозволяє не тільки встановлювати причинно-наслідкові зв'язки між об'єктами предметної області, але і визначати їх взаємозв'язок і вплив.

10.1.2. Методи вимірювання ступеня впливу об'єктів

До найбільш часто використовуваних при експертному оцінюванні методів належать: ранжування, парне порівняння, безпосередня оцінка. При описі кожного з перерахованих методів будемо вважати, що є кінцеве число вимірюваних

об'єктів і сформульована одна або кілька ознак порівняння, по яких вивчається ступінь впливу об'єктів на результат.

Отже, методи вимірювання будуть відрізнятися лише процедурою порівняння об'єктів. Ця процедура включає:

- побудова відносин між об'єктами емпіричної системи;
- вибір функції f , що відображає об'єкти емпіричної системи на числову систему;
- визначення шкали вимірювань.

Розглянемо докладніше всі ці питання, що виникають при використанні кожного з методів вимірювань.

10.1.2.1. Метод ранжування

Ранжування - це процедура впорядкування об'єктів за ступенем їх впливу на результат, виконується екпертом в процесі виявлення його знань. На основі своїх знань і досвіду експерт розташовує об'єкти в порядку переваги, керуючись одним або декількома показниками порівняння. Залежно від виду відносин між об'єктами можливі різні варіанти упорядкування об'єктів.

Нехай серед об'єктів немає *еквівалентних* за ступенем впливу на результат. У цьому випадку між об'єктами існує відношення строгого порядку, що володіє властивостями:

- несиметричності (якщо $O_i > O_j$, то $O_i < O_j$);
- транзитивності (якщо $O_i > O_j$, $O_j > O_k$, то $O_i > O_k$);
- і зв'язності (для будь-яких двох об'єктів, або $O_i > O_j$, або $O_i < O_j$).

У результаті порівняння всіх об'єктів експерт складає упорядковану послідовність:

$$O_1 > O_2 > \dots > O_n,$$

де об'єкт з номером один є найбільш переважним з усіх об'єктів, об'єкт з номером два менш переважний, ніж перший, але переважніше всіх інших і т.д.

Отримана система з відношенням порядку $<O>$ утворює *серію*. Для серії доведено існування числової системи, елементами якої є числа, а відношення порядку $>$ є відношенням «більше ніж», «краще ніж».

Це означає, що існує числове представлення $f(O_i)$ таке, що послідовності відповідає послідовність чисел

$$f(O_1) > f(O_2) > \dots > f(O_n).$$

На практиці експертного ранжування найчастіше використовується послідовність натуральних чисел

$$r_1 = f(O_1) = 1; r_2 = f(O_2) = 2; \dots; r_n = f(O_n) = n.$$

Числа r_1, r_2, \dots, r_n називаються рангами. Найбільш переважачому присвоюється ранг 1, іншому - ранг 2 і т.д. На практиці, серед об'єктів можуть бути і ек-

вівалентні за ступенем їх впливу на результат. Наприклад, упорядкування може мати вигляд

$$O_1 > O_2 > O_3 \sim O_4 \sim O_5 > \dots > O_{n-1} \sim O_n.$$

У цій послідовності об'єкти O_3 , O_4 і O_5 еквівалентні між собою, а O_{n-1} і O_n - між собою. Для еквівалентних об'єктів прийнято призначати однакові ранги, рівні середньому арифметичному значенню рангів, призначених однаковим об'єктам. Такі ранги отримали назву зв'язаних рангів. Для прикладу впорядкування в разі $n = 10$ ранги об'єктів O_3 , O_4 і O_5 будуть однаковими :

$$r_3 = r_4 = r_5 = (3+4+5) / 3 = 4$$

$$r_9 = r_{10} = (9+10) / 2 = 9.5 .$$

Як видно з прикладу, зв'язані ранги можуть бути дробовими. Зручність використання зв'язаних рангів полягає в тому, що сума рангів n - об'єктів дорівнює сумі натуральних чисел від 1 до n . При цьому будь-які комбінації зв'язаних рангів не змінюють цю суму. Ця обставина істотно спрощує обробку результатів ранжування при груповій експертній оцінці.

10.1.2.2 . Метод парних порівнянь

Парне порівняння являє собою процедуру встановлення переваги об'єктів при порівнянні всіх можливих пар. На відміну від ранжування, при якому здійснюється впорядкування об'єктів відразу, парне порівняння представляє для експертів більш просту задачу. При порівнянні кожної пари об'єктів можливе відношення або порядку, або еквівалентності. Парне порівняння є вимірювання у шкалі порядку.

У результаті порівняння кожної пари об'єктів O_i , O_j експерт має впорядкувати цю пару, визначивши, що:

$$\text{або } O_i > O_j, \text{ або } O_j > O_i, \text{ або } O_i \sim O_j .$$

Перехід від емпіричної системи до числової системи з відношенням здійснюється за допомогою такої функції f , що :

$$\text{якщо } O_i > O_j, \text{ то } f(O_i) > f(O_j)$$

$$\text{якщо } O_j > O_i, \text{ то } f(O_j) > f(O_i)$$

Нарешті якщо об'єкти еквівалентні, то природно припустити, що $f(O_i) \sim f(O_j)$. Найчастіше в практиці експертного оцінювання використовують наступні чисельні представлення:

Таблиця 10.1

Емпірична система	Представлення 1		Представлення 2		Представлення 3	
	$f(O_i)$	$f(O_j)$	$f(O_i)$	$f(O_j)$	$f(O_i)$	$f(O_j)$
$O_i > O_j$	2	0	1	-1	1	0
$O_i \sim O_j$	1	1	0	0	0.5	0.5

Результати порівняння експертом усіх пар об'єктів зручно представити у вигляді таблиці, стовпці і рядки якої складають об'єкти, а в комірках таблиці проставляються числові значення.

Приклад: Як приклад розглянемо табличне відображення результатів проведеного парного порівняння п'яти об'єктів при використанні числового представлення 1.

Таблиця 10.2

O_i	O_1	O_2	O_3	O_4	O_5
O_1	1	2	2	1	2
O_2	0	1	2	1	0
O_3	0	0	1	0	1
O_4	1	1	2	1	0
O_5	0	2	1	2	1

З цієї таблиці випливає, що об'єкт O_1 переважніше об'єктів O_2, O_3, O_5 і еквівалентний O_4 . Об'єкт O_2 переважніше O_3 , еквівалентний O_4 і менш кращий, ніж O_1 і O_5 . Порівняння об'єктів у всіх можливих парах не дає повного упорядкування всіх об'єктів. Тому виникає задача про ранжування об'єктів на основі парного порівняння.

10.1.2.3 . Метод безпосередньої оцінки

Безпосередня оцінка являє собою процедуру приписування об'єктам числових значень в шкалі інтервалів. Ці значення відповідають ступеню впливу того чи іншого об'єкта на спостережуваний результат.

У процесі виявлення значень експерт повинен поставити у відповідність кожному об'єкту точку на безперервній числовій осі, наприклад, на відрізьку $[0,1]$. Природно вимагати, щоб еквівалентним по впливу об'єктам приписувалося б одне і теж число.

Вимірювання уподобання в шкалі інтервалів можна виконати з високим ступенем довіри тільки при гарній інформованості експертів про властивості об'єктів і предметної області.

У ряді випадків, з метою ослаблення цих умов, але, природно, за рахунок зменшення точності вимірювання замість безперервної числової осі розглядають велику оцінку, яка використовує 5, 10, 100 - бальні шкали.

Однак безпосередня оцінка не завжди повинна використовувати числові шкали. Наприклад, колір об'єкта неможливо уявити у вигляді якого-небудь числового значення, а перехід до значень частот спектру в багатьох випадках досить скрутний для експертів.

10.1.3 . Формування групи експертів

При формуванні групи експертів на стадії виявлення знань необхідно враховувати такі характеристики експертів як:

- компетентність - ступінь кваліфікації експертів в даній галузі знань;
- креативність - здатність вирішувати творчі завдання;
- відношення до експертизи - негативне чи пасивне ставлення, чи зайнятість істотно впливає на якість роботи експерта в групі;
- конформізм - схильність впливу авторитетів, при якому думка авторитету може пригнічувати осіб, що володіють більш високою компетентністю;
- колективізм і самокритичність.

Розглянемо один з можливих шляхів кількісного опису характеристик експерта, заснований на обчисленні відносних коефіцієнтів компетентності за результатами висловлювання фахівців про склад експертної групи.

Суть методики зводиться до того, що ряду фахівців пропонується висловити думку про обліковий склад експертної групи. Якщо в цьому списку з'являються особи, що не увійшли до вихідного списку, їм теж пропонується назвати фахівців для участі в експертизі. Після кількох етапів буде отриманий досить повний список кандидатів у групу.

За результатами опитування складається матриця (табл.10.3), по рядках і стовпцям якої записуються прізвища експертів, а елементами таблиці є змінні

$$x_{ij} = \begin{cases} 1, & \text{якщо } j\text{-ий експерт назвав } i\text{-го;} \\ 0, & \text{якщо } j\text{-ий експерт не назвав } i\text{-го.} \end{cases} \quad (10.1)$$

При цьому експерт може включати чи не включати себе в експертну групу (тобто $x_{jj}=0$ або $x_{jj}=1$). По даній таблиці можна обчислити відносні коефіцієнти компетентності, використовуючи алгоритм вирішення завдань про лідера. Введемо відносні коефіцієнти компетентності h -порядку для кожного експерта

$$K_i^h = \frac{\sum_{j=1}^m x_{ij} k_j^{h-1}}{\sum_{i=1}^m \sum_{j=1}^m x_{ij} k_j^{h-1}}, \quad (i=1, \dots, m; h=1, 2, \dots) \quad (10.2)$$

де m - число експертів у списку (розмірність матриці $\|x_{ij}\|$), h - номер порядку коефіцієнта компетентності. Коефіцієнти компетентності нормовані так, що їх сума дорівнює одиниці:

$$\sum_{i=1}^m k_i^h = 1, \quad h=1, 2, \dots \quad (10.3)$$

За формулою (10.2) можна обчислити значення компетентності для різних порядків, починаючи з першого. При $h=1$ вираз (10.2) буде мати вигляд:

$$K_i^1 = \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^m \sum_{j=1}^m x_{ij}}, \quad i = 1, 2, \dots, m \quad (10.4)$$

Сенс цієї формули в тому, що підраховується кількість голосів, поданих за i -го експерта, і ділиться на загальну кількість голосів, поданих за всіх експертів. Таким чином, коефіцієнт компетентності першого порядку - це відносне число експертів, що висловилися за включення i -го експерта в групу.

Відносний коефіцієнт компетентності другого порядку отримують з (10.2) для $h = 2$ за умови, що $k_j^1 (j=1, 2 \dots m)$ визначені за (1.4) :

$$K_i^1 = \frac{\sum_{j=1}^m x_{ij} k_j^1}{\sum_{i=1}^m \sum_{j=1}^m x_{ij} k_j^1}, \quad i = 1, 2, \dots, m \quad (10.5)$$

Коефіцієнти другого порядку являють собою відносну кількість голосів, врахованих за допомогою коефіцієнта компетентності першого порядку.

Послідовно обчислюючи відносні коефіцієнти компетентності більш високого порядку, можна переконатися, що процес швидко сходиться після 3-4 обчислень, тобто відносні коефіцієнти швидко стабілізуються. У загальному випадку коефіцієнти відносної компетентності визначаються як:

$$k_i = \lim_{h \rightarrow \infty} k_i^h, \quad \sum_{i=1}^m k_i = 1 \quad (10.6)$$

Приклад: У результаті опитування трьох експертів про склад експертної групи отримані дані (x_{ij}) про думку кожного з них по вхід експертів у робочу групу. Ці дані зведені в табл.10.3 .

Таблиця 10.3

	Думка експертів		
	Експерт 1 (А)	Експерт 2 (В)	Експерт 3 (С)
Експерт 1 (А)	1	1	1
Експерт 2 (В)	0	1	0
Експерт 3 (С)	1	0	1

Результати послідовної обробки отриманих даних за описаним вище алгоритмом матимуть вигляд.

На першому кроці, вважаючи рівну компетентність всіх експертів, приймаємо $k^0 = [1 \ 1 \ 1]^T$ і обчислюємо коефіцієнти відносної компетентності першого порядку:

$$y = \sum_{i=1}^2 x_{ij} = 1+1+1+0+1+0+1+0+1 = 6$$

$$k_A^1 = k_1^1 = \frac{1}{y} \sum_{j=1}^2 x_{1j} k_j^0 = \frac{1}{6} (1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1) = \frac{3}{6} = 0.5$$

$$k_B^1 = k_2^1 = \frac{1}{y} \sum_{j=1}^2 x_{2j} k_j^0 = \frac{1}{6} (0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1) = \frac{1}{6} = 0.167$$

$$k_C^1 = k_2^1 = \frac{1}{y} \sum_{j=1}^2 x_{2j} k_j^0 = \frac{1}{6} (1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1) = \frac{2}{6} \approx 0.333$$

На другому кроці, використовуючи отримані значення, обчислюємо коефіцієнти відносної компетентності другого порядку:

$$y = \sum_{i=1}^2 \sum_{j=1}^2 x_{ij} k_j^1 = 1 \cdot \frac{3}{6} + 1 \cdot \frac{1}{6} + 1 \cdot \frac{2}{6} + 0 \cdot \frac{3}{6} + 1 \cdot \frac{1}{6} + 0 \cdot \frac{2}{6} + 1 \cdot \frac{3}{6} + 1 \cdot \frac{3}{6} + 0 \cdot \frac{1}{6} + 1 \cdot \frac{2}{6} = \frac{12}{6} = 2$$

$$k_A^2 = k_1^2 = \frac{1}{y} \sum_{i=1}^2 x_{1i} k_i^1 = \frac{1}{2} \left(1 \cdot \frac{3}{6} + 1 \cdot \frac{1}{6} + 1 \cdot \frac{2}{6} \right) = \frac{16}{26} = 0.5$$

$$k_B^2 = k_2^2 = \frac{1}{y} \sum_{i=1}^2 x_{2i} k_i^1 = \frac{1}{2} \left(0 \cdot \frac{3}{6} + 1 \cdot \frac{1}{6} + 0 \cdot \frac{2}{6} \right) = \frac{11}{26} \approx 0.083$$

$$k_C^2 = k_2^2 = \frac{1}{y} \sum_{i=1}^2 x_{2i} k_i^1 = \frac{1}{2} \left(1 \cdot \frac{3}{6} + 0 \cdot \frac{1}{6} + 1 \cdot \frac{2}{6} \right) = \frac{15}{26} \approx 0.417$$

Продовжуючи аналогічні обчислення до тих пір, доки не буде відрізнятися від k_i^{h-1} з точністю 0.01, отримаємо

$$k^3 = [0.5 \quad 0.042 \quad 0.458]^T$$

$$k^4 = [0.5 \quad 0.02 \quad 0.48]^T$$

$$k^5 = [0.5 \quad 0.01 \quad 0.49]^T.$$

$$\text{При } h \rightarrow \infty \quad k^h \rightarrow [0.5 \quad 0.0 \quad 0.5]^T$$

Можна показати, що граничні значення коефіцієнтів компетентності являють собою компоненти власного вектора для максимального власного числа матриці $X = \|x_{ij}\|$. Власні числа матриці X визначаються як корені алгебраїчного рівняння

$$|X - \lambda \times E| = 0,$$

де λ - вектор власних чисел матриці голосування $|X|$, E - єдина матриця. Власний вектор матриці, відповідний максимальному власному числу, обчислюється із системи $m + 1$ порядку лінійних алгебраїчних рівнянь

$$XK = \lambda_0 K, \quad \sum_{i=1}^m k_i = 1,$$

де $K = [k_1, k_2, \dots, k_m]$ - вектор компетентності, що є власним вектором матриці для максимального власного числа.

Існують підходи до оцінки компетентності, засновані на обліку апостеріорних даних, тобто результатів експертного оцінювання.

10.2. Обробка експертних оцінок

10.2.1. Завдання обробки

Залежно від цілей експертного оцінювання та методу обліку експертних оцінок виникають такі основні завдання:

- 1) побудова узагальненої оцінки понять і об'єктів на основі індивідуальних оцінок експертів;
- 2) побудова узагальненої оцінки на основі парного порівняння об'єктів кожним з експертів;
- 3) визначення відносного взаємозв'язку об'єктів;
- 4) визначення залежностей між ранжуванням;
- 5) визначення узгодженості думок експертів;
- 6) оцінка надійності обробки результатів.

При вирішенні багатьох завдань недостатньо упорядкування об'єктів по одному або групі показників. Необхідно мати числові значення для кожного об'єкта, що визначають його перевагу перед іншими об'єктами. Наявність таких оцінок дозволить визначити узагальнену оцінку для всієї групи експертів.

Визначення узгодженості думок експертів створюється шляхом обчислення числової міри, що характеризує ступінь схожості індивідуальних думок. Аналіз значень міри узгодження сприяє виробленню та виявленню угруповань думок експертів.

Обробка експертних оцінок дозволяє розкрити зв'язані показники порівняння і здійснити угруповання за ступенем зв'язку. Так, наприклад, якщо показники порівняння - засоби досягнення цих цілей, то встановлення взаємозв'язку між ранжуванням, впорядковуваними витрати з точки зору досягнення цілей, дозволяє обґрунтовано відповісти на питання: " в якій мірі досягнення однієї мети по даних засобах сприяє досягненню інших цілей " (тобто встановити причинно-наслідковий зв'язок).

Оцінки, одержувані на основі обробки, являють собою випадкові об'єкти, тому одним з найважливіших завдань процедури обробки є визначення їх надійності.

Існує безліч підходів до вирішення даної задачі. З метою ілюстрації розглянемо один з найпростіших. Нехай m експертів провели оцінку n об'єктів по l показникам. Результати оцінювання представлені величинами x_{ij}^h , де i - номер

об'єкт, j - номер експерта, h - номер показника. Величини, x_{ij}^h отримані методом безпосереднього оцінювання, являють собою числа з деякого відрізка числової осі, або бали.

В якості групової оцінки для кожного з об'єктів можна прийняти середнє зважене значення його оцінки

$$x_i = \sum_{h=1}^l \sum_{j=1}^m q_h x_{ij}^h k_j, (i = 1, 2, \dots, n)$$

де q_h - коефіцієнти значення показників порівняння об'єктів, k_j - коефіцієнти компетентності експертів. Величини q_h і k_j є нормованими, тобто

$$\sum_{h=1}^l q_h = 1, \sum_{j=1}^m k_j = 1$$

Коефіцієнти q_h можуть бути визначені експертним шляхом, як середній коефіцієнт ваги h - го показника по всім експертам, тобто

$$q_h = \sum_{j=1}^m q_{hj} k_k$$

Можливість отримання групової експертної оцінки шляхом підсумовування індивідуальних оцінок з значенням компетентності і важливості ґрунтується на виконанні :

- аксіом теорії корисності фон Неймана-Моргенштерна для індивідуальних і групових оцінок;
- умов нерозрізненості об'єктів в груповому відношенні, якщо вони невідрізнні у всіх індивідуальних оцінках (частковий принцип Парето) .

Коефіцієнти компетентності експертів можна обчислити за апостеріорними даними, тобто за результатами оцінки об'єктів. Основною ідеєю цього обчислення є припущення про те, що компетентність експерта повинна оцінюватися за ступенем узгодженості його оцінок з груповою оцінкою об'єктів.

Для спрощення подальшого викладення, обмежимося розглядом випадку $h = 1$. Тобто коли групове оцінювання об'єктів проводиться на основі тільки одного показника. Алгоритм обчислення групових оцінок і коефіцієнтів компетентності експертів для цього випадку має вигляд:

а) початкові умови при $t = 0$

$$k_j^0 = \frac{1}{m} (j = \overline{1, m}),$$

тобто початкове значення коефіцієнтів компетентності для всіх експертів приймається однаковим і рівним.

б) рекурентні співвідношення для $t = 1, 2, 3 \dots$

$$x_i^t = \sum_{j=1}^m x_{ij} k_j^{t-1}, (i = \overline{1, n}) - \text{групова оцінка для } i - \text{го об'єкта на } t - \text{му кроці на основі}$$

індивідуальних оцінок x_{ij} ;

$$\lambda^t = \sum_{j=1}^n \sum_{i=1}^m x_i^t x_{ij} - \text{нормований коефіцієнт};$$

$$k_j^t = \frac{1}{\lambda^t} \sum_{i=1}^n x_{ij} x_i^t, (j = \overline{1, m-1}) - \text{коефіцієнти компетентності } j - \text{ого експерта на } t - \text{му кроці};$$

$$k_m^t = 1 - \sum_{j=1}^{m-1} k_j^t - \text{коефіцієнти компетентності } m - \text{ого експерта із умови нормування};$$

в) ознака закінчення ітераційного процесу

$$\max(|x_i^t - x_i^{t-1}|) < E$$

Збіжність даної ітераційної процедури доведена в літературі для випадку, коли індивідуальні оцінки невід'ємні, а експерти і об'єкти не розпадаються на окремі групи (тобто кожна група експертів не оцінює об'єкти своєї групи). У більшості практичних завдань ці умови виконуються, що доводить збіжність алгоритму.

Приклад: Три експерти ($m = 3$) оцінили значення двох заходів ($n = 2$) за ступенем їх впливу на вирішення однієї з проблем ($l = 1$). Результатами експертизи з'явилися нормовані оцінки заходів $x_{1j} + x_{2j} = 1, j = 1, 2, 3$.

x_{ij}	Експерт 1	Експерт 2	Експерт 3
Захід 1	0,3	0,5	0,2
Захід 2	0,7	0,5	0,8

Обчислимо групові оцінки заходів, що приводять до вирішення проблеми і коефіцієнти компетентності кожного з експертів. Для цього скористаємося наведеним вище алгоритмом, задавши точністю обчислення $E = 0,001$.

Середні оцінки об'єктів першого наближення (при $t=1$) будуть рівні:

$$x_1^1 = \frac{1}{3}(0,3 + 0,5 + 0,2) = 0,333 \quad ; \quad x_2^1 = \frac{1}{3}(0,3 + 0,5 + 0,8) = 0,667$$

$$x^1 = (0,333; 0,667)$$

Обчислимо нормований коефіцієнт λ^1 :

$$\lambda^1 = \sum \sum x_{ij} x_i^1 = x_1^1(0,3 + 0,5 + 0,2) + x_2^1(0,7 + 0,5 + 0,8) = 0,333 * 1 + 0,667 * 2 = 1,665$$

Значення коефіцієнтів компетентності першого наближення приймуть значення:

$$k_1^1 = \frac{1}{1,665}(0,3 * 0,333 + 0,7 * 0,667) = 0,34$$

$$k_2^1 = \frac{1}{1,665}(0,3 * 0,333 + 0,5 * 0,667) = 0,3$$

$$k_3^1 = 1 - (0,34 + 0,3) = 0,36 \text{ і тоді } k^1 = (0,34; 0,30; 0,36)$$

Обчислюючи групові оцінки другого і т.д. наближення, отримаємо:

$$x^2 = (0,324; 0,676) \quad x^3 = (0,3235; 0,6765)$$

$$\lambda^2 = 1,676 \quad \lambda^3 = 1,6765$$

$$k^2 = (0,341; 0,298; 0,361) \quad k^3 = (0,341; 0,298; 0,361)$$

Результат третього кроку задовольняє умову закінчення ітераційного процесу і за значення групової оцінки приймається $x \approx x^3 = (0,3235; 0,6765)$.

10.2.2. Обробка парних порівнянь

При встановленні причинно-наслідкових залежностей між об'єктами предметної області, експертам у ряді випадків складно висловити їх чисельно. Тобто важко встановити кількісно ступінь впливу тієї чи іншої причини (об'єкта) на конкретне слідство. Особливо психологічно це складно, якщо таких об'єктів багато.

Разом з тим, експерти порівняно легко вирішують завдання парного порівняння. Це завдання полягає в тому, що експерт встановлює уподобання об'єктів при порівнянні всіх можливих пар. Тобто експерт, розглядаючи всі можливі пари об'єктів, в кожній з них встановлює ту причину, яка, на його думку, надає більший вплив на слідство. Виникає питання, як отримати оцінку всієї сукупності об'єктів на основі результатів парного порівняння, виконаного групою експертів.

Нехай кожен з експертів проводить оцінку впливу на результат всіх пар об'єктів, даючи числову оцінку

$$r_{ij}^h = \begin{cases} 1 & , \text{ якщо об'єкт } O_i \text{ більш значущий, ніж } O_j \\ 0,5 & , \text{ якщо об'єкти } O_i \text{ і } O_j \text{ рівноправні} \\ 0 & , \text{ якщо об'єкт } O_i \text{ менш значущий, ніж } O_j \end{cases}$$

де $h=1,2,\dots,m$ - номер експерта, $i,j=1,2,\dots,n$ - номери об'єктів, досліджуваних при експертизі. Тобто за результатами експертизи маємо m таблиць (матриць) виду, що показано на рис. 10.1.

Як слід з рис. 10.1 послідовність обробки парних порівнянь полягає в тому, що на підставі таблиць парних порівнянь експертів будується матриця математичних очікувань оцінок всіх пар об'єктів. Потім по цій матриці обчислюється вектор коефіцієнтів відносної важливості об'єктів.

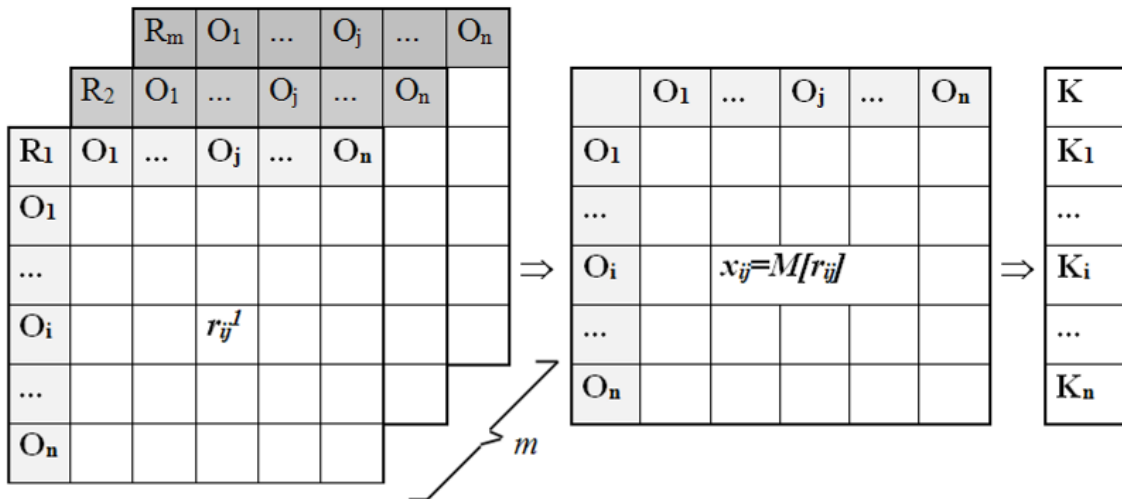


Рис. 10.1. Послідовність обробки парних порівнянь

Якщо при оцінці пари O_{ij} із загальної кількості експертів m_i висловилися на користь O_i , m_j , на користь O_j , а m_p вважає ці об'єкти рівноправними, то оцінка математичного очікування дискретної випадкової величини буде дорівнювати:

$$x_{ij} = M[r_{ij}^h] = 1 * \frac{m_i}{m} + 0,5 * \frac{m_p}{m} + 0 * \frac{m_j}{m}, h = \overline{1, m}$$

Так як загальна кількість експертів $m = m_i + m_p + m_j$, то визначаючи звідси m_p і підставляючи його в вищенаведений вираз отримаємо

$$x_{ij} = \frac{m_i}{m} + 0,5 \left(\frac{m - m_i - m_j}{m} \right) = \frac{1}{2} + \frac{m_i - m_j}{2m}$$

Очевидно, що $x_{ij} + x_{ji} = 1$. Сукупність величин x_{ij} утворюють матрицю $X = \|x_{ij}\|$ розмірності $n \times n$, на основі якої можна побудувати ранжування всіх об'єктів і визначити коефіцієнти відносної важливості об'єктів, тобто вектор

$$k = [k_1, k_2, \dots, k_n]^T$$

Одним із способів визначення значень елементів вектора є ітераційний алгоритм виду:

а) початкова умова $t = 0$

$$k^0 = [k_1, k_2, \dots, k_n]^T$$

б) рекурентні відношення

$$k^t = \frac{1}{\lambda^t} * X * k^{t-1};$$

$$\lambda^t = [1 \ 1 \ 1 \ \dots \ 1] * X * k^{t-1}, t = (1, 2, \dots, n),$$

де X - матриця математичних очікувань оцінок пар об'єктів, k^t - вектор коефіцієнтів відносної важливості об'єктів порядку t .

$$\sum_{i=1}^n k_i^t = 1 \text{ - умова нормування;}$$

в) ознака закінчення $\| k^t - k^{t-1} \| < E$.

Якщо матриця ненегативна і нерозкладна (тобто шляхом перестановки рядків і стовпців її не можна привести до трикутного вигляду), то при збільшенні порядку $t \rightarrow \infty$ величина λ^t сходиться до максимального власного числа матриці X , тобто

$$k = \lim_{x \rightarrow \infty} k^t, \sum_{i=1}^n k_i = 1.$$

Це твердження випливає з теореми Перрона-Фробеніуса та доводить збіжність наведеного вище алгоритму.

Приклад: Припустимо, що в результаті опитування трьох ($m = 3$) експертів про ступінь впливу на результат трьох ($n = 3$) різних факторів (об'єктів) отримані наступні таблиці парних порівнянь:

Експерт 1 (R_1)

	O_1	O_2	O_3
O_1	0.5	1	1
O_2	0	0.5	0
O_3	0	1	0.5

Експерт 2 (R_2)

	O_1	O_2	O_3
O_1	0.5	0.5	0.5
O_2	0.5	0.5	0.5
O_3	0.5	0.5	0.5

Експерт 3 (R_3)

	O_1	O_2	O_3
O_1	0.5	1	0.5
O_2	0	0.5	0
O_3	0.5	1	0.5

Для отримання групової оцінки ступеня впливу кожного з об'єктів на результат, побудуємо матрицю математичних очікувань оцінок кожної з пар об'єктів, яка для розглянутого прикладу буде мати вигляд:

	O_1	O_2	O_3
O_1	$3/6$	$5/6$	$4/6$
O_2	$1/6$	$3/6$	$1/6$
O_3	$2/6$	$5/6$	$3/6$

Значення елементів цієї матриці отримані з таких виразів:

$$x_{11} = \frac{1}{2} + \frac{0 - 0}{2 \cdot 3} = \frac{1}{2}; \quad x_{11} = x_{22} = x_{33} = \frac{1}{2};$$

$$x_{12} = \frac{1}{2} + \frac{2 - 0}{2 \cdot 3} = \frac{5}{6}; \quad x_{21} = 1 - x_{12} = \frac{1}{6};$$

$$x_{13} = \frac{1}{2} + \frac{1 - 0}{2 \cdot 3} = \frac{4}{6}; \quad x_{31} = 1 - x_{13} = \frac{2}{6};$$

$$x_{23} = \frac{1}{2} + \frac{0 - 2}{2 \cdot 3} = \frac{1}{6}; \quad x_{32} = 1 - x_{23} = \frac{5}{6}.$$

Скористаємося вищеописаним алгоритмом для отримання вектора відносної важливості об'єктів. Для наочності, кожен з кроків представимо у вигляді:

крок 0:

$$K^0 = [1 \ 1 \ 1]^T;$$

крок 1:

$$Y^1 = X \times K^0 = \frac{1}{6} \times \begin{bmatrix} 3 & 5 & 4 \\ 1 & 3 & 1 \\ 2 & 5 & 3 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{6} \times \begin{bmatrix} 12 \\ 5 \\ 10 \end{bmatrix}$$

$$\lambda^1 = [1 \ 1 \ 1] \times Y^1 = [1 \ 1 \ 1] \times \frac{1}{6} \times \begin{bmatrix} 12 \\ 5 \\ 10 \end{bmatrix} = \frac{27}{6}$$

$$K^1 = \frac{1}{\lambda^1} \times Y^1 = \frac{6}{27} \times \frac{1}{6} \times \begin{bmatrix} 12 \\ 5 \\ 10 \end{bmatrix} = \begin{bmatrix} 0.444 \\ 0.185 \\ 0.370 \end{bmatrix}$$

крок 2:

$$Y^2 = X \times K^1 = \frac{1}{6} \times \begin{bmatrix} 3 & 5 & 4 \\ 1 & 3 & 1 \\ 2 & 5 & 3 \end{bmatrix} \times \frac{1}{27} \times \begin{bmatrix} 12 \\ 5 \\ 10 \end{bmatrix} = \frac{1}{6 \cdot 27} \times \begin{bmatrix} 101 \\ 37 \\ 79 \end{bmatrix}$$

$$\lambda^2 = [1 \ 1 \ 1] \times Y^2 = [1 \ 1 \ 1] \times \frac{1}{6 \cdot 27} \times \begin{bmatrix} 101 \\ 37 \\ 79 \end{bmatrix} = \frac{217}{6 \cdot 27}$$

$$K^2 = \frac{1}{\lambda^2} \times Y^2 = \frac{6 \cdot 27}{217} \times \frac{1}{6 \cdot 27} \times \begin{bmatrix} 101 \\ 37 \\ 79 \end{bmatrix} = \begin{bmatrix} 0.465 \\ 0.171 \\ 0.364 \end{bmatrix}$$

Продовжуючи ітераційний процес до тих пір, доки норма оцінки не буде менше заданої ($\max(|K_i^t - K_i^{t-1}|) < 0.001$) отримаємо

$$K^3 = [0.468 \ 0.169 \ 0.363]^T$$

$$K^4 = [0.468 \ 0.169 \ 0.363]^T$$

На четвертому кроці виконується умова виходу, що дозволяє за групову оцінку ступеня впливу на результат прийняти вектор коефіцієнтів відносної важливості об'єктів виду:

$$K = [0.500 \ 0.349 \ 0.151]^T$$

10.2.3. Визначення узагальнених ранжировок

При груповій експертній оцінці кожному i -ому об'єкту кожен з j -х експертів присвоює r_{ij} . У результаті проведення експертного оцінювання виходить матриця рангів $\|r_{ij}\|$ розмірності $n \times m$, де n - кількість об'єктів, а m - кількість експертів ($j = \overline{1, m}$).

Найпростіший спосіб отримання узагальненої ранжировки полягає в ранжуванні об'єктів за величиною сум рангів, отриманих кожним об'єктом від усіх експертів. У цьому випадку для матриці ранжировок $\|r_{ij}\|$ обчислюються суми:

$$r_i = \sum_{j=1}^m r_{ij} \quad (i = 1, 2, \dots, n).$$

Далі об'єкти упорядковуються по ланцюжку нерівностей $r_k < r_l < \dots < r_q$, де $r_k = \min_i(r_i)$, $r_l = \min_{i, i \neq k}(r_i)$, ..., $r_q = \max_i(r_i)$. Звідси випливає узагальнена ранжировка об'єктів:

$$O_k \succ O_l \succ \dots \succ O_q.$$

Для обліку компетентності експертів досить помножити i - у ранжировку на коефіцієнти компетентності j - го експерта $0 \leq k \leq 1$. У цьому випадку обчислення суми рангів для i - го об'єкта здійснюватися за формулою:

$$r_i = \sum_{j=1}^m k_j r_{ij},$$

що дозволяє впорядкувати об'єкти по ланцюжку нерівностей. Слід зазначити, що побудова таких узагальнених ранжировок є коректною процедурою тільки в тому випадку, якщо ранги призначаються як місця об'єктів у вигляді натуральних чисел $1, 2, \dots, n$.

Однак ранги об'єктів визначають тільки порядок розташування об'єктів за показниками порівняння. Ранги як числа не дають можливості зробити висновок про те, на скільки або в скільки разів кращий один об'єкт в порівнянні з іншим. Якщо ранги 3, то звідси не слід робити висновки про те, що об'єкт, з рангом 1, в три рази краще, ніж об'єкт, що має ранг, що дорівнює трьом.

Разом з тим, для використання в ЕС знань, отриманих від експертів, необхідно не тільки упорядкування або ранжування об'єктів за ступенем їх впливу чи на який-небудь результат, але й визначення кількісної оцінки ступеня впливу кожного з об'єктів на результат.

Найпростішим методом для реалізації цього завдання є підхід, заснований на побудові узагальненої ранжировки шляхом переходу від матриці ранжировок до матриці парних порівнянь. Для цього на основі матриці $\|r_{ij}\|$ будується m матриць парних порівнянь R_j ($j=1, 2, \dots, m$), де m - число експертів. Елементи цих матриць визначаються таким чином:

$$R_j = \|r_{ik}^j\| = \begin{cases} 1 & , \text{якщо } O_i^j \succ O_k^j, \text{ тобто } r_{ij} < r_{kj} \\ 0,5 & , \text{якщо } O_i^j \sim O_k^j, \text{ тобто } r_{ij} = r_{kj} \\ 0 & , \text{якщо } O_i^j \prec O_k^j, \text{ тобто } r_{ij} > r_{kj} \end{cases}$$

де j - номер експерта, i та k - номери порівнюваних об'єктів.

Потім до отриманих матриць парних порівнянь всіх експертів застосовується розглянутий раніше метод обробки парних порівнянь. Його ітераційна процедура дозволяє отримати коефіцієнти відносної важливості об'єктів за ступенем їх впливу на результат. Проілюструємо застосування цього підходу на прикладі.

Приклад. Нехай три експерта ($m = 3$) провели ранжування трьох об'єктів ($n=3$) за ступенем їх впливу на який-небудь результат і таблиця ранжировок має вигляд:

Об'єкт O_i	Експерт 1	Експерт 2	Експерт 3
O_1	1	1	2
O_2	2	3	1
O_3	3	2	3

На основі цієї таблиці матриця парних порівнянь для першого експерта буде мати вигляд:

$$R_1 = \|R_{kj}^1\| = \begin{vmatrix} 0.5 & 1 & 1 \\ 0 & 0.5 & 1 \\ 0 & 0 & 0.5 \end{vmatrix}$$

Аналогічні матриці парних порівнянь для другого і третього експерта будуть мати вигляд:

$$R_2 = \begin{vmatrix} 0.5 & 1 & 1 \\ 0 & 0.5 & 0 \\ 0 & 1 & 0.5 \end{vmatrix}; \quad R_3 = \begin{vmatrix} 0.5 & 0 & 1 \\ 1 & 0.5 & 1 \\ 0 & 0 & 0.5 \end{vmatrix}$$

Використовуючи метод обробки парних порівнянь отримаємо послідовність векторів коефіцієнтів відносної важливості об'єктів:

Крок	K_1	K_2	K_3
0	1,0	1,0	1,0
1	0,481	0,333	0,185
2	0,489	0,346	0,156
3	0,5	0,348	0,152
4	0,5	0,349	0,151

Ітераційна процедура із заданою точністю ($E = 0,001$) є збіжною на четвертому кроці до значень:

$$K = [0.500 \ 0.349 \ 0.151]^T$$

Що дозволяє оцінити кількісно ступінь впливу кожного об'єкта на результат, отриманий на основі вихідного ранжування експертів.

10.3 . Експертні системи з невизначеними знаннями

10.3.1 . Проблеми , властиві ЕС з невизначеними знаннями

На практиці часто доводиться оцінювати гіпотези для яких є неповна або недостатня інформація. Іноді важко зробити точні оцінки, але, незважаючи на невизначеність, ми приймаємо розумні рішення. Щоб ЕС були корисними, вони теж повинні вміти це робити. Класичним прикладом цього завдання є медична діагностика. Завжди існують деякі сумніви в чіткості прояве симптомів того чи іншого захворювання. Сумніви в наявності у пацієнта конкретного захворювання зберігаються навіть у тому випадку, коли всі його симптоми чітко виражені.

Як же виявляється і враховується невизначеність в експертних системах? Розглянемо найпростішу ситуацію. Нехай використовується правило:

якщо (A), то (B)

і, припустимо, ніякі інші правила і посилки не мають відношення до ситуації, що розглядається. Де ж виникає невизначеність? В ЕС вона може бути двох типів:

- невизначеність в істинності самої посилки (наприклад , якщо ступінь впевненості в тому , що A істинно становить 90 % , то які значення прийме B);
- невизначеність самого правила (наприклад ми можемо сказати, що в більшості випадків , але не завжди якщо є A, то є також і B).

Ще більш складна ситуація виникає у разі , якщо правило має вигляд:

якщо (A і B) , то (C),

де ми можемо з деякою мірою бути впевнені як в істинності кожної з посилок (A, B), а тим більше їх спільного прояву, так і в істинності самого аргументу.

При проектуванні і створенні ЕС з невизначеними знаннями виникають чотири важливі проблеми :

- як кількісно виразити ступінь визначеності при встановленні істинності (або хибності) деякої частини даних?
- як виразити ступінь підтримки висновку конкретною посилкою?
- як використовувати спільно дві (або більше) посилок, незалежно впливаючих на висновок ?
- як бути в ситуації, коли потрібно обговорити ланцюжок висновку для підтвердження в умовах невизначеності?

Насамперед розглянемо можливості використання, теорії ймовірності при введенні в умовах невизначеності.

10.3.2 . Суб'єктивні ймовірності

Основне поняття настільки природно, що воно відіграє значну роль у повсякденному житті. Розмови, що стосуються ймовірності дощу або хорошого урожаю в городі часто зустрічаються в нашому житті. Поняття ймовірності було розроблено декілька століть тому. Але вже тисячі років людина використовує такі слова , як « може бути » , « шанс » , « успіх » або інші їх еквівалентні в розмовній мові .

Математична теорія ймовірностей була сформульована відносно недавно (близько 1660 року). Ймовірність події класично визначається як відношення випадків, в яких дана подія відбувається до загального числа спостережень .

Однак можливі й інші визначення. У теперішній час існують кілька інтерпретацій теорії ймовірностей. Розглянемо три найбільш домінуючі погляди.

Об'єктивний погляд. Цей підхід заснований на законі великих чисел, що гарантує те, що при наявності досить великої кількості спостережень частота результатів, що цікавить події буде прямувати до об'єктивної ймовірності .

Персоніфікований, суб'єктивний чи заснований на думках погляд. Полягає в тому, що ймовірнісна міра розглядається як ступінь довіри того, як окрема особистість судить про істинність деякого висловлювання. Цей погляд говорить, що дана особистість має в деякому сенсі ставлення до цієї події. Термін «байєсовські» часто використовується як синонім суб'єктивної ймовірності .

Необхідний або логічний погляд. Характеризується тим, що ймовірнісна міра поширюється на безліч тверджень, що мають логічний зв'язок, такий, що істинність одного з них може залежати від іншого.

Ці ймовірнісні інтерпретації використовують і різні схеми виведення. Однак існують всього дві школи ймовірнісних розрахунків: школа Паскаля (або загальноприйнята), школа Бекона (або індуктивна). Розрахунки по Паскалю використовують байєсовські правила для перевірки та обробки границі довіри. Обчислення по Бекону використовують правила логіки для доведення або спростування гіпотез. Об'єктивний і суб'єктивний погляди використовують розрахунки по Паскалю. Ті, хто підтримують логічні висновки, використовують розрахунки по Бекону.

Існують ЕС, побудовані на обох напрямках. Проте в ЕС бази знань накопичують людські знання, тому для представлення знань експертів з урахуванням ймовірностей найбільш придатними є інтерпретації на основі суб'єктивних довір. У результаті чого і більшість сучасних ЕС, що використовують теорію ймовірностей, є «байєсівськими».

10.3.3 . Байєсівське оцінювання

Перед тим, як ввести теорему Байєса, розглянемо деякі фундаментальні поняття теорії ймовірностей. Нехай A є деяка подія реального світу. Сукупність всіх елементарних подій називається вибірковою просторою або простір події (Ω). Ймовірність події A , позначається $p(A)$ і кожна ймовірнісна функція p повинна задовольняти три аксіоми:

1) Ймовірність будь-якої події A є невід'ємною, тобто

$$p(A) \geq 0 \text{ для } \forall A \in \Omega;$$

2) Ймовірність всіх подій вибіркового простору дорівнює 1, тобто

$$p(\Omega) = 1.$$

3) Якщо k подій A_1, A_2, \dots, A_k є взаємно незалежними (тобто не можуть відбутися одночасно), то ймовірність, принаймні, однієї з цих подій дорівнює сумі окремих ймовірностей, або

$$p(A_1 \cup A_2 \cup \dots \cup A_k) = \sum_{i=1}^k p(A_i)$$

Аксіоми 1 і 2 можна об'єднати, що дає

$$1 \geq p(A) \geq 0 \text{ для } \forall A \in \Omega.$$

Це твердження показує, ймовірність будь-якої події знаходиться між 0 і 1. За визначенням, коли $p(A) = 0$, то подія A ніколи не станеться. У тому випадку і коли $p(A) \geq 1$, то подія A має статися обов'язково.

Доповнення до A , позначається ($\neg A$), містить сукупність всіх подій Ω за виключаємо A . Так як A і $\neg A$ є взаємно незалежними (тобто $A \cap \neg A = \emptyset$), то з аксіоми 3 випливає

$$p(A) + p(\neg A) = p(A \cup \neg A) = p(\Omega) = 1.$$

Перепишучи цю рівність у вигляді $p(\neg A) = 1 - p(A)$, ми отримуємо шлях для отримання $p(\neg A)$ из $p(A)$.

Припустимо тепер, що $B \in \Omega$ деяка інша подія. Тоді ймовірність того, що відбудеться A за умови, що сталося B записується у вигляді $p(A|B)$ і називається умовною ймовірністю події A при заданій події B .

Ймовірність того, що обидві події A і B відбудуться $p(A \cap B)$ називається *спільною ймовірністю подій A і B* . Умовна ймовірність $p(A|B)$ дорівнює відношенню спільної ймовірності $p(A \cap B)$ до ймовірності події B , за умови, що вона не дорівнює 0, тобто

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

Аналогічно умовна ймовірність події B за умови A , що позначається $p(B|A)$ дорівнює:

$$p(B|A) = \frac{p(B \cap A)}{p(A)}$$

і таким чином

$$p(B \cap A) = p(B|A) \times p(A).$$

Так, як спільна ймовірність комутативна (тобто від перестановки місць сума не змінюється), то

$$p(A \cap B) = p(B \cap A) = p(B|A) \times p(A).$$

Підставляючи цю рівність у раніше отриманий вираз для умовної ймовірності $p(B|A)$ отримуємо правило Байєса

$$p(A|B) = \frac{p(B|A) \times p(A)}{p(B)}$$

У ряді випадків наше знання того, що відбулася подія B , не впливає на ймовірність події A (або навпаки A і B). Іншими словами, ймовірність події A не залежить від того, сталося чи ні подія B , так що

$$p(A|B) = p(A) \text{ и } p(B|A) = p(B).$$

У цьому випадку говорять, що подія A і B є незалежними.

10.4 . Ймовірнісний логічний висновок

10.4.1 . Найпростіший логічний висновок

Розглянемо випадок, коли всі правила в експертній системі відображаються у формі :

Якщо $\langle H \text{ є істинною} \rangle$ то $\langle E \text{ спостерігається з ймовірністю } p \rangle$.

Очевидно, якщо H сталося, то це правило говорить про те, що подія E відбувається з ймовірністю p . Але що буде, якщо стан H невідомо, а E відбулося?

Використані теореми Байєса дозволяють обчислити вірогідність того, що H істинно. Заміна « A » і « B » на « H » і « E » не істотна для формули Байєса, але з її допомогою ми можемо покинути загальну теорію ймовірності і перейти до аналізу ймовірнісних обчислень в ЕС. У цьому контексті:

- H - подія, що полягає в тому, що дана гіпотеза вірна;
- E - подія, що полягає в тому, що є доказ (свідчення), який може підтвердити правильність вказаної гіпотези.

Перепишучи формулу Байєса в термінах гіпотез і свідчень, отримуємо:

$$p(H | E) = \frac{p(E | H) \times p(H)}{p(E | H) \times p(H) + p(E | \neg H) \times p(\neg H)}.$$

Ця рівність встановлює зв'язок гіпотези зі свідченням і, в той же час, свідчення, що спостерігається з поки ще не підтвердженою гіпотезою. Ця інтерпретація припускає також визначення апіорної ймовірності гіпотези $p(H)$ до спостереження або отримання деякого факту.

В ЕС ймовірності, необхідні для вирішення деякої проблеми, забезпечуються експертами і запам'ятовуються в базі знань. Ці ймовірності включають:

- апіорні ймовірності всіх гіпотез $p(H)$ (хвороба);
- умовні ймовірності виникнення доказів за умови існування кожної з гіпотез $p(E | H)$.

Так, наприклад, у медичній діагностиці експерт повинен задати апіорні ймовірності всіх можливих хвороб в деякій медичній галузі.

Крім того, повинні бути визначені умовні ймовірності прояву тих чи інших симптомів при кожній з хвороб. Умовні ймовірності повинні бути отримані для всіх симптомів і хвороб, припускаючи, що всі симптоми незалежні в рамках однієї хвороби.

Дві події E_1 і E_2 є умовно незалежними, якщо їх спільна ймовірність за умови деякої гіпотези H дорівнює добутку умовних ймовірностей цих подій за умови H , тобто

$$p(E_1 E_2 | H) = p(E_1 | H) \cdot p(E_2 | H).$$

Користувачі дають ЕС інформацію про спостереження (наявності певних симптомів) $p(H_i | E_j \dots E_k)$ для всіх гіпотез (H_1, \dots, H_m) у світлі пред'явлених симптомів (E_j, \dots, E_k) і ймовірностей, збережених в БЗ (без знань).

Ймовірність $p(H_i | E_j \dots E_k)$ називається *апостеріорною ймовірністю* гіпотез за спостереженнями (E_j, \dots, E_k) . Ці ймовірності дають порівняльне ранжування всіх можливих гіпотез, тобто гіпотез з нульовими апостеріорними ймовірностями.

Однак, наведена вище формула Байєса обмежена в тому, що кожне свідчення впливає на одну гіпотезу. Можна узагальнити цей вираз на випадок множинних

гіпотез (H_1, \dots, H_m) і множинних доказів (E_j, \dots, E_k) . Ймовірності кожної з гіпотез за умови виникнення деякого конкретного доказу E можна визначити з виразу:

$$p(H_i | E) = \frac{p(E | H_i) \times p(H_i)}{\sum_{k=1}^m p(E | H_k) \times p(H_k)}, \quad i = \overline{1, m}.$$

10.4.2. Діагностичний алгоритм із застосуванням формули Байєса

Мається ціла група алгоритмів для обчислювальної діагностики, яка заснована на ймовірнісних критеріях. У медичних довідниках по діагностиці захворювань зазвичай симптоми та ознаки захворювань супроводжуються вказівками на частоту їх появи при даній хворобі. При цьому використовується такі терміни, як «часто», «дуже часто», «рідко» і т.п.

Діагностичний алгоритм, який визначає порядок переробки медичної інформації для вирішення завдання постановки діагнозу включає в себе збір, аналіз і оцінку інформації. Алгоритми аналізу інформації в експертних системах (ЕС) ґрунтуються на кількісних операціях. В алгоритмі максимальної оцінки інформації, основну роль грають логічні операції. У цьому випадку здійснюється постійне порівняння інформації від пацієнта з інформацією напрацьованою експертами. Найпростіший діагностичний алгоритм такої ЕС зображений на рис. 10.2.

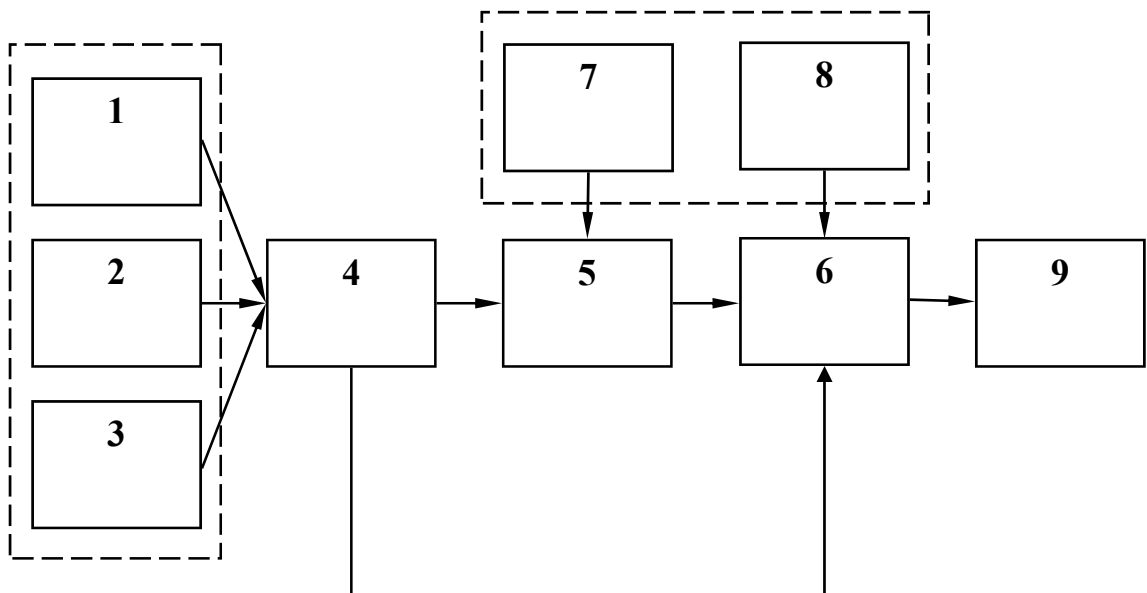


Рис. 10.2 Діагностичний алгоритм для ЕС.

Тільки останнім часом починають в окремих випадках проводити статичні характеристики симптомокомплексів, що дозволяє оцінювати їх вірогідність. У

той час, як в медицині основою діагностики є встановлення залежності симптомів від наявності захворювань, машинна діагностика ґрунтується на вивченні ймовірності тих чи інших захворювань за наявності певного набору симптомів.

Блоки на рис. 10.2 означають наступне: 1 - дані опитування та огляду; 2 - лабораторно-клінічні дані; 3 - результати фізіологічних досліджень; 4 - оперативне накопичення інформації; 5 - порівняння накопиченої інформації з константами різних медичних показників; 6 - порівняння виділених симптомів і відхилень з симптомокомплексами, відбір симптомів (пам'ять ЕОМ); 7 - запам'ятовування констант і меж «норми»; 8 - запам'ятовування симптомокомплексів; 9 - видача висновків (діагнозів).

Завдання ймовірнісного діагнозу може бути сформульоване як задача визначення ймовірності даної хвороби по заданому набору симптомів, якщо відомі ймовірні наявності окремих симптомів при певних захворюваннях. У зв'язку з цим завданням в обчислювальній діагностиці отримала застосування формула Байєса, яка для даної ЕС може бути записана у вигляді:

$$P(H_i / E) = \frac{P(E / H_i) \cdot P(H_i)}{\sum_{k=1}^m P(E / H_k) \cdot P(H_k)},$$

де $i = \overline{1, m}$, $P(H_i / E)$ - ймовірність того, що за наявності комплексу симптомів E_j є хвороба H_i ; $P(H_i)$ - ймовірність захворювання при випадковому виборі; $P(E_j / H_i)$ - відома з медичного досвіду ймовірність наявності симптомів комплексу E_j при хворобі H_i ; $\sum P(E / H_k) \cdot P(H_k)$ - сума добутків ймовірностей кожної з розглянутих хвороб на ймовірність даного комплексу симптомів при кожній з цих хвороб.

Прокоментуємо цю формулу: $P(H_i / E_j)$ - шукана ймовірність хвороби по даному набору симптомів $P(E_j / H_i)$. Це той медичний досвід, який накопичено в даний час і виражений кількісно у вигляді ймовірності даного симптомокомплексу при даному захворюванні. Ці дані, як правило, визначаються експертами за допомогою методики, наведеної вище.

Ця ймовірність може бути виражена числом від 0 до 1,0. Хоча в навчальних посібниках та монографіях ще немає точних розрахунків таких ймовірностей, вони можуть бути виражені приблизно таким чином: патогномонічні симптомокомплекси мають ймовірність, рівну одиниці, що часто зустрічні симптомокомплекси 0,7-0,9; середня частота симптомокомплексів 0,5-0,6; рідко зустрічаються симптомокомплекси 0,3-0,4; дуже рідко 0,1-0,2; нуль означає, що симптомокомплекс при даному захворюванні ніколи не зустрічається. $P(H_i)$ - це ймовірність даного захворювання у групі випадково відібраних пацієнтів; такою групою може бути контингент даної лікарні, даного району, даного міста. Сенс введення в діагностику величини $P(H_i)$ полягає в тому, що вона непостійна і залежить від географічних, сезонних, епідеміологічних та інших факторів, які лікар повинен враховувати при постановці діагнозу.

Так, наприклад, під час епідемії грипу симптомокомплекс: головний біль, слабкість, підвищення температури тіла з ймовірністю 0,95-1,0 відповідає діагнозу грипу; в тропічних країнах цей же симптомокомплекс, в відсутність епідеміологічних факторів, має велику ймовірність для діагнозу теплового удару, гіпертермії.

Сума добутоків ймовірностей кожної з розглянутих хвороб на ймовірність обраного (досліджуваного) симптомокомплексу при кожному даному захворюванні являє собою умовну величину, що показує порівняльну значимість даного симптомокомплексу при даному захворюванні у порівнянні з іншими захворюваннями. По суті це показник патогномонічності. Чим більше відрізняється знаменник від чисельника, тим менш патогномонічний симптомокомплекс.

В якості прикладу розглянемо таблицю 10.4, в якій є умовні та безумовні ймовірності для 7 симптомокомплексів. Припустимо, що у нашого хворого є симптомокомплекс $M_4 (E_4)$, при якому можливі хвороби № 1-4 (H_1, H_2, H_3, H_4). За формулою Байєса знаходимо значення $P(H_i/E_j)$, тобто ймовірності кожної з 4 хвороб за наявності симптому E_4 .

$$P(H_1 / E_4) = \frac{0,6 \cdot 0,333}{0,6 \cdot 0,333 + 0,15 \cdot 0,067 + 0,05 \cdot 0,3 + 0,005 \cdot 0,2} = 0,855$$

$$P(H_2 / E_4) = 0,044$$

$$P(H_4 / E_4) = 0,067$$

$$P(H_6 / E_4) = 0,004$$

Таблиця 10.4

Симптомокомплекси E_i	Хвороба-комплекси H_i							
	H_0	H_1	H_2	H_3	H_4	H_5	H_6	H_7
E_0	0	0	0	0	0	0	0	0
E_1	0	0	0,167	0	0,1	0	0	0
E_2	0	1,167	0,067	0	0	0	0	0
E_3	0	0	0,2	0	0	0	0	0
E_4	0	0,33	0,06	0	0,3	0	0,2	0
E_5	0	0	0,16	0,33	0,6	0	0,2	0,2
E_6	0	0,5	0,16	0	0	0	0,2	0
E_7	0	0	0,26	0,66	0	0	0,1	0,8
Хвороба-комплекс	0	0,6	0,15	0,15	0,05	0,04	0	0

таким чином найбільш ймовірна хвороба H_i .

У разі множинних доказів:

$$p(H_i | E_1 E_2 \dots E_n) = \frac{p(E_1 E_2 \dots E_n | H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1 E_2 \dots E_n | H_k) \times p(H_k)}, i = \overline{1, m}.$$

На жаль, даний вираз має ряд недоліків. Так, знаменник вимагає від нас знання умовних ймовірностей всіх можливих комбінацій доказів і гіпотез, що робить правило Байєса малоприматним для ряду додатків. Однак у тих випадках, коли можливо припустити умовну незалежність свідчень, правило Байєса можна привести до простішого вигляду:

$$p(H_i | E_1 E_2 \dots E_n) = \frac{p(E_1 | H_i) \times p(E_2 | H_i) \times \dots \times p(E_n | H_i) \times p(H_i)}{\sum_{k=1}^m p(E_1 | H_k) \times p(E_2 | H_k) \times \dots \times p(E_n | H_k) \times p(H_k)}, i = \overline{1, m}.$$

Разом з тим припущення про незалежність подій у ряді випадків пригнічують точності суджень і доказів в ЕС.

10.4.3. Поширення ймовірностей в ЕС

Ймовірності подій поширюються по БЗ експертної системи на основі правила Байєса для обчислення всіх апостеріорних ймовірностей гіпотез за умови спостережуваних доказів. Ці апостеріорні ймовірності дають ранжовану інформацію про потенційно істинну гіпотезу. Розглянемо приклад, який ілюструє цей процес.

Приклад. Припустимо, що в деякій БЗ є всього три взаємно незалежних гіпотези: H_1, H_2, H_3 , які мають апіорні ймовірності: $p(H_1), p(H_2), p(H_3)$, відповідно. Правила БЗ містять два умовно незалежних доказів, які підтримують вихідні гіпотези в різного ступеня. Апіорні і умовні ймовірності всіх гіпотез і доказів цього прикладу позначені таким чином. $p(H_i | E)$ - бажана ймовірність хвороби по даному набору симптомів; $p(H_i)$ - ймовірність захворювання H при випадковому виборі (апіорна); $p(E | H_i)$ - це той медичний досвід, який накопичено останнім часом; $\sum p(E | H_k) \times p(H_k)$ - сума добутків ймовірностей кожної з розглянутих хвороб на ймовірність даного одного з симптомів при кожній хворобі.

$p(i)$	1	2	3
$p(H_i)$	0,5	0,3	0,2
$p(E_1 H_i)$	0,4	0,8	0,3
$p(E_2 H_i)$	0,7	0,9	0,0

При цьому вихідні гіпотези характеризує подія, пов'язана визначенням надійності деякої медичної фірми:

H_1 - «середня надійність лікувального закладу»

H_2 - «висока надійність лікувального закладу»

H_3 - «низька надійність лікувального закладу»

Подіями, які є умовно незалежними доказами, що підтримують вихідні гіпотези є: E_1 - "наявність помилок" і E_2 - "наявність довідкового матеріалу".

У процесі збору фактів ймовірності гіпотез будуть підвищуватися, якщо факти підтримують їх або зменшуватися, якщо спростовують їх. Припустимо, що ми маємо тільки одне свідчення E_1 (тобто з ймовірністю одиниця настав факт E_1). Спостерігаючи E_1 , ми обчислюємо апостеріорні ймовірності для гіпотез згідно з формулою Байєса для одного доказу:

$$p(H_i | E_1) = \frac{p(E_1 | H_i) \cdot p(H_i)}{\sum_{k=1}^3 p(E_1 | H_k) \cdot p(H_k)}, \quad i=1,2,3.$$

Таким чином,

$$p(H_1 | E_1) = \frac{0.4 \cdot 0.5}{0.4 \cdot 0.5 + 0.8 \cdot 0.3 + 0.3 \cdot 0.2} = 0.40$$

$$p(H_2 | E_1) = \frac{0.8 \cdot 0.3}{0.4 \cdot 0.5 + 0.8 \cdot 0.3 + 0.3 \cdot 0.2} = 0.48$$

$$p(H_3 | E_1) = \frac{0.3 \cdot 0.2}{0.4 \cdot 0.5 + 0.8 \cdot 0.3 + 0.3 \cdot 0.2} = 0.12$$

Після того як E_1 сталося, довіра до гіпотез H_1 і H_3 знизилася, в той час як довіра до H_2 зросла. У тих випадках, коли є факти, які підтверджують як подія E_1 і E_2 , то апостеріорні ймовірності вихідних гіпотез також можуть бути обчислені за правилом Байєса:

$$p(H_i | E_1 E_2) = \frac{p(E_1 E_2 | H_i) \times p(H_i)}{\sum_{k=1}^3 p(E_1 E_2 | H_k) \times p(H_k)}, \quad i=1,2,3.$$

Оскільки події E_1 і E_2 умовно незалежні при даних гіпотезах H_i , формулу Байєса можна переписати у вигляді:

$$p(H_i | E_1 E_2) = \frac{p(E_1 | H_i) \times p(E_2 | H_i) \times p(H_i)}{\sum_{k=1}^3 p(E_1 | H_k) \times p(E_2 | H_k) \times p(H_k)}, i = 1, 2, 3.$$

Звідки

$$p(H_1 | E_1 E_2) = \frac{0,4 \times 0,7 \times 0,5}{0,4 \times 0,7 \times 0,5 + 0,8 \times 0,9 \times 0,3 + 0,3 \times 0,0 \times 0,2} = 0,393,$$

$$p(H_2 | E_1 E_2) = \frac{0,8 \times 0,9 \times 0,3}{0,4 \times 0,7 \times 0,5 + 0,8 \times 0,9 \times 0,3 + 0,3 \times 0,0 \times 0,2} = 0,607,$$

$$p(H_3 | E_1 E_2) = \frac{0,3 \times 0,0 \times 0,2}{0,4 \times 0,7 \times 0,5 + 0,8 \times 0,9 \times 0,3 + 0,3 \times 0,0 \times 0,2} = 0,0.$$

Це означає, що тільки H_1 і H_2 залишилися після отримання доказів E_1 і E_2 . При цьому H_1 , менш ймовірно, ніж H_2 . На цьому прикладі ми розглянули процес поширення ймовірностей по елементам при вступі до неї тих чи інших доказів.

10.4.4. Послідовне поширення ймовірностей

Однак реально поширення ймовірностей відбувається поетапно, підсумовуванням окремих доказів та їх впливу на умовну ймовірність, по мірі надходження окремих E_i . Це можна зробити, використовуючи апіорні і апостеріорні ймовірності, наступним чином:

1. Задаємо $p(H_i)$ - апіорну ймовірність подій H_i .
2. Для отриманих доказів E_j записуємо $p(E_j | H_i)$.
3. З урахуванням теореми Байєса підраховуємо $p(E_j | H_i)$ залежно від результату E_j , тобто обчислюємо апостеріорну ймовірність події H_i .
4. Тепер можна не звертати уваги на всі наставші E_j і перепозначивши поточну апостеріорну ймовірність події H_i , як нову апіорну ймовірність H_i . Отже, нехай $p(H_i)$ дорівнює $p(H_i | E_j)$ залежно від значення E_j .
5. Потім виберемо нове свідчення для розгляду і перейдемо до п. 2.

Проілюструємо цю послідовність на наведеному вище прикладі в положенні, що спочатку надійшов доказ E_2 . тоді:

$$p(H_1 | E_2) = \frac{0,7 \times 0,5}{0,7 \times 0,5 + 0,9 \times 0,3 + 0,0 \times 0,2} = 0,565,$$

$$p(H_2 | E_2) = \frac{0,9 \times 0,3}{0,7 \times 0,5 + 0,9 \times 0,3 + 0,0 \times 0,2} = 0,435,$$

$$p(H_3 | E_2) = \frac{0,0 \times 0,2}{0,7 \times 0,5 + 0,9 \times 0,3 + 0,0 \times 0,2} = 0,0.$$

Отримані ймовірності можна прийняти за нові апостеріорні ймовірності гіпотез H_1 , H_2 і H_3 , тобто:

$$p(\tilde{H}_1) = 0,565 \quad , \quad p(\tilde{H}_2) = 0,435 \quad , \quad p(\tilde{H}_3) = 0,0$$

І якщо тепер додатково надійде доказ E_1 , то нові апостеріорні ймовірності гіпотез можуть бути обчислені тільки на основі знову обчисленого його доказу:

$$p(H_1 | E_1 E_2) = \frac{0,4 \times 0,565}{0,4 \times 0,565 + 0,8 \times 0,435 + 0,3 \times 0,0} = 0,393,$$

$$p(H_2 | E_1 E_2) = \frac{0,8 \times 0,435}{0,4 \times 0,565 + 0,8 \times 0,435 + 0,3 \times 0,0} = 0,607,$$

$$p(H_3 | E_1 E_2) = \frac{0,3 \times 0,0}{0,4 \times 0,565 + 0,8 \times 0,435 + 0,3 \times 0,0} = 0,0.$$

З наведеного прикладу видно, що ітераційна процедура послідовного розподілу ймовірностей в міру надходження доказів дозволяє отримати результат, аналогічні безпосередньому застосуванню правила Байеса для випадку одночасно двох надійшовших доказів .

Короткі висновки

Суть процесу виявлення знань полягає в організації проведення експертами інтуїтивно - логічного аналізу проблемної області з кількісною оцінкою формульованих ними суджень. На цьому етапі експерти:

- 1) формують об'єкти і поняття предметної області;
- 2) проводять виміри характеристик.

До найбільш часто використовуваних при експертному оцінюванні методів належать: ранжування, парне порівняння, безпосередня оцінка.

Ранжування - це процедура впорядкування об'єктів за ступенем їх впливу на результат, виконується екпертом в процесі виявлення його знань .

Парне порівняння являє собою процедуру встановлення переваги об'єктів при порівнянні всіх можливих пар.

Безпосередня оцінка являє собою процедуру приписування об'єктам числових значень в шкалі інтервалів. Ці значення відповідають ступеню впливу того чи іншого об'єкта на спостережуваний результат.

Ранжування і парне порівняння - це вимірювання в шкалі порядку. Безпосередня оцінка - це вимірювання в шкалі інтервалів.

При формуванні групи експертів на стадії виявлення знань необхідно врахувати такі характеристики експертів як:

- 1) компетентність
- 2) креативність;
- 3) ставлення до експертизи
- 4) конформізм;
- 5) колективізм і самокритичність

Залежно від цілей експертного оцінювання та методу обліку експертних оцінок виникають такі основні завдання:

- 1) побудова узагальненої оцінки понять і об'єктів на основі індивідуальних оцінок експертів;
- 2) побудова узагальненої оцінки на основі парного порівняння об'єктів кожним з експертів;
- 3) визначення відносного взаємозв'язку об'єктів;
- 4) визначення залежностей між результатами ранжування;
- 5) визначення узгодженості думок експертів;
- 6) оцінка надійності обробки результатів.

Можливість отримання групової експертної оцінки шляхом підсумовування індивідуальних оцінок з врахуванням компетентності і важливості ґрунтується на виконанні :

- 1) аксіом теорії корисності фон Неймана-Моргенштерна для індивідуальних і групових оцінок;
- 2) умов непомітності об'єктів в груповому відношенні, якщо вони не відрізняються у всіх індивідуальних оцінках (частковий принцип Парето).

При проектуванні і створенні ЕС з невизначеними знаннями виникають чотири важливі проблеми:

- 1) як кількісно виразити ступінь визначеності при встановленні істинності або хибності) деякої частини даних?
- 2) як висловити ступінь підтримки висновку конкретною посилкою?
- 3) як використовувати спільно дві (або більше) посилок, незалежно впливаючих на висновок?

4) як бути в ситуації, коли потрібно обговорити порядок висновка для підтвердження в умовах невизначеності?

Існують ЕС, побудовані на об'єктивному і суб'єктивному поглядах на поняття ймовірності події. У базах знань ЕС накопичуються людські знання, тому для представлення знань експертів з урахуванням ймовірностей найбільш придатними є інтерпретації на основі суб'єктивних довір. У результаті чого і більшість сучасних ЕС що використовують теорію ймовірностей, є « байєсівськими ».

Використання байєсівської стратегії в ЕС реалізують з використанням Байєсівської формули "зворотних ймовірностей", тобто оцінки умовних ймовірностей гіпотез. За наявності декількох ознак (симптомів) такі обчислення просто реалізувати в припущенні статистичної незалежності ознак, що далеко не завжди відповідає дійсності. Разом з тим, практика показує, що такий підхід, незважаючи на його очевидну математичну некоректність, цілком застосовний, оскільки зазвичай приводить до правильних висновків.

Література до Розділу 10

1. Хабаров С.П. Экспертные системы. Конспект лекций. – С-Пб: 2008. [Электронный ресурс] – Режим доступа: <http://www.twirpx.com/file/61731/> (03.04.2013).
2. Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование - М.: Вильямс. – 2007. – 1152 с.
3. Частиков А.П., Гаврилова Т.А., Белов Д.Л. Разработка экспертных систем. Среда CLIPS. - СПб.: БХВ-Петербург. - 2003. – 608 с.
4. Ковалев В.З., Татевосян А.С.. Интеллектуальные информационные системы: Конспект лекций. - Омск: Изд-во ОмГТУ, 2005.
5. Vizureanu P. Expert SHstems – Vukovar, Croatia.: Intech – 2010. – 246 p.
6. Продеус А.Н., Захарова Е.Н. Экспертные системы в медицине. – К.: «ВЕК+», 1998. – 320 с.
7. Новиков Ф., Яценко А. Microsoft Office'97 в целом. – «ВНУ–Санкт-Петербург», 1998. – 624 с.

Контрольні запитання до Розділу 10

1. У чому полягає суть процесу виявлення знань?
2. Що собою являє ранжування і як воно використовується при виявленні знань?
3. Що собою являє парне порівняння і як воно використовується при виявленні знань?
4. Що собою являє безпосередня оцінка і як вона використовується при виявленні знань?

5. У яких шкалах проводяться вимірювання при реалізації процедур ранжирування, парного порівняння та безпосередній оцінці?
6. Які характеристики експертів необхідно враховувати при формуванні групи експертів на стадії виявлення знань?
7. Які основні завдання виникають при експертному оцінюванні?
8. На чому ґрунтується можливість отримання групової експертної оцінки?
9. Назвіть чотири важливі проблеми, що виникають при проектуванні і створенні ЕС з невизначеними знаннями.
10. Чому більшість сучасних ЕС, що використовують теорію ймовірностей, є «байєсівськими»?
11. Яке припущення, що не відповідає дійсності, зазвичай використовують розробники «байєсівських» ЕС?

Додатки

У Додатках наведені завдання до лабораторних робіт з дисципліни «Експертні системи в медицині», а також теоретичний матеріал. Всі лабораторні роботи виконуються на ПК, тому можуть одночасно розглядатися як комп'ютерні практикуми.

Додаток 1

Лабораторна робота 1 Оцінювання інформативності діагностичних ознак

Основні цілі роботи:

1. Практично закріпити знання про існування 2-х груп показників інформативності діагностичних ознак: класової і міжкласової інформативності.
2. Практично освоїти способи оцінювання інформативності діагностичних ознак.

Додаткові цілі роботи :

Освоїти елементи статистичної обробки результатів спостережень (оцінювання математичного сподівання (МО), середньоквадратичного відхилення (СКВ), регресії, лінійного та нелінійного тренду, показників інформативності) в середовищі Excel.

Завдання:

1 . На аркуші «Регресія» вивчіть представлені вихідні дані та результати їх регресійного аналізу. Зробіть висновки про можливість та ефективність використання регресійного аналізу для оцінки інформативності діагностичних ознак. Використовуючи Excel, самостійно побудуйте лінію регресії E_i на H_j , а також її поліноміальну апроксимацію (лінію тренду) для одного з 11 представлених тестів.

2 . На аркуші «Відстані» вивчіть результати обчислень відстаней Бхатачарія. Завершіть обчислення відстаней Махаланобіса. Поясніть природу можливого «поділу на нуль» при обчисленні відстані Махаланобіса. Порівняйте ці два способи обчислень міжкласових відстаней і виберіть кращий, на вашу думку, спосіб. Аргументуйте свій вибір. Організуйте обчислення дивергенції Кульбака. Зробіть висновки щодо зручності цього показника.

3 . На аркуші «Інформація» вивчіть алгоритм і результати обчислень міжкласової інформативності ознаки. Зіставте ці результати з отриманими раніше результатами. Вкажіть переваги і недоліки даного показника.

4 . На аркуші «Порівняння» наведені результати обчислення коефіцієнта кореляції між коефіцієнтами регресії і показником міжкласової інформативності.

Перевірте правильність цих результатів і зробіть висновки про переваги і недоліки порівнюваних показників.

5. На аркуші «Ентропія» вивчіть алгоритм і результати обчислень класової ентропії. На цьому ж аркуші обчисліть інформаційну значимість симптому. Порівняйте результати і зробіть висновки про переваги і недоліки кожного з показників.

Зміст звіту:

Уявіть і прокоментуйте результати ваших спостережень і обчислень.

Контрольні питання:

1. Лист «Регресія»: які 2 способи тут використані для побудови ліній тренда?
2. Лист «Відстані»: відстані між чим тут обчислюються?
3. Лист «Інформація»: чим зручніше «інформація» у порівнянні з «відстанями»?
4. Лист «Порівняння»: у чому схожість і в чому відмінність «інформації» та коефіцієнту регресії?
5. Лист «Ентропія»: чому у вихідних даних 0 замінений на дуже мале число?

Теоретичні відомості

Розрізняють 2 види показників інформативності - класові і міжкласові. Для діагностики більший інтерес представляють міжкласові показники. Класові показники цікаві з точки зору мінімізації опису об'єктів одного обраного класу.

Міжкласові показники

Про інформативності ознаки можна судити по кривих регресії x_i на Y_j - чим сильніше виражена статистична залежність, тим інформативніше ознака. Якщо значення ознак і захворювань представлені шкалами порядку, тоді в якості міри залежності, а отже, і заходи інформативності, можна використовувати тангенс кута нахилу лінійного компонента регресійної залежності (коефіцієнт кореляції Пірсона).

Інші показники:

1) відстань Махаланобіса:

$$d_{ijt} = \frac{|m_{ij} - m_{it}|}{\sqrt{\sigma_{ij}\sigma_{it}}}$$

2) відстань Бхаттачарія:

$$\mu_{ijt} = -\ln \int_{x_i} [P(x_i | H_j)P(x_i | H_t)]^{1/2} dx_i$$

3) дивергенція Кульбака:

$$K_{ijt} = \sum_k [P(E_{ik} / H_j) - P(E_{ik} / H_t)] \log_2 \frac{P(E_{ik} / H_j)}{P(E_{ik} / H_t)}$$

4) кількість інформації про межкласову відмінність:

$$I_i = \sum_j \sum_k \log_2 \frac{P(E_{ik}, H_j)}{P(E_{ik})P(H_j)}$$

Класові показники

1) інформаційна значущість симптому:

$$I_{ij} = \sum_k \log_2 \frac{P(E_i / H_j)}{P(E_i)} = \sum_k \log_2 \frac{P(H_j / E_i)}{P(H_j)}$$

де $P(E_i / H_j)$ - ймовірність наявності ознаки E_i при наявності захворювань H_j ; $P(E_i)$ - ймовірність спостереження ознаки E_i взагалі; $P(H_j / E_i)$ - ймовірність захворювання H_j при спостереженні ознаки E_i ; $P(H_j)$ - ймовірність захворювання H_j .

2) класова ентропія:

$$h_{ij} = - \sum_k P(E_{ik} | H_j) \log_3 P(E_{ik} | H_j)$$

Коментарі до програми, реалізованої в середовищі Excel

Програма реалізована на 5-ти аркушах (рис. П1.1-П1.5): "Регресія", "Відстані", "Інформація", "Порівняння" і "Ентропія".

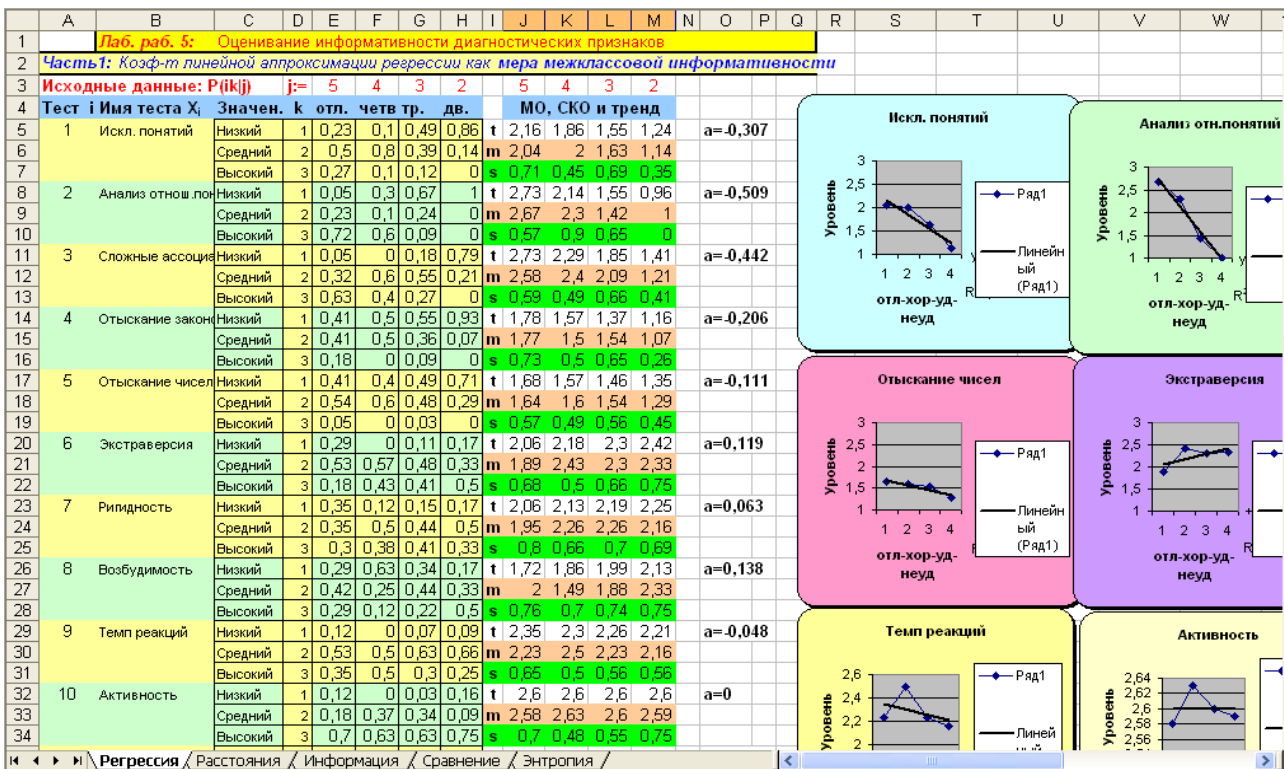


Рис. П1.1

1. Лист "Регресія" (рис. П1.1): Вихідні дані - отримані в результаті тестування школярів оцінки умовної ймовірності $P(X_i / Y_j)$ - частота результатів психологічного тестування, що зустрічається для різних категорій ("відмінники", "хорошисти" і т.д.) школярів. Результати - графіки ліній регресії, що характеризують залежність результатів психологічного тестування для різних категорій школярів.
2. Лист "Відстані" (рис. П1.2): результати обчислень відстаней Бхатачарія і Махаланобіса.
3. Лист "Інформація" (рис. П1.3): результати обчислень міжкласової інформативності.
4. Лист "Порівняння" (рис. П1.4): порівняння результатів регресійного аналізу і обчислень міжкласової інформації.

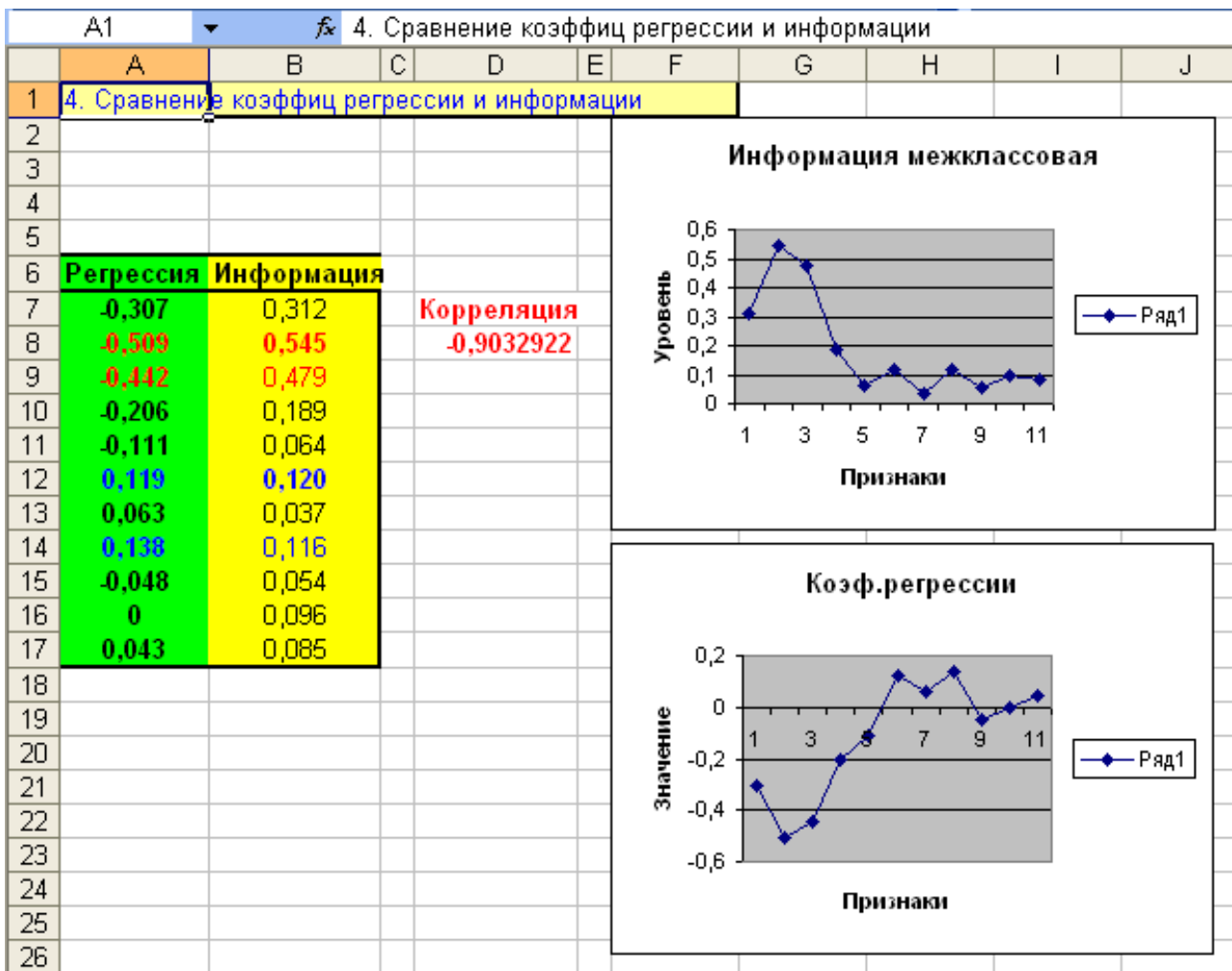


Рис. П1.4

5. Лист "Ентропія" (рис. П1.5): обчислення класової ентропії.

Q9		0,630929755667361																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	5. Внутрикласовая энтропия																				
2											Энт. ij-того приз.(формула)			Внутрикласовая энтропия							
3	Исходные данные: P(i k)																				
4											$h_{ij} = -\sum_k \{P(i k) \log_3 P(i k)\}$										
5	Тест i	Имя теста X _i	Значен. k	отл.	четв.	тр.	дв.	P(x _i Y)	лгрфмP(x _i Y)	произвед.	Тест i	Имя теста X _i	отл.	четв.	тр.	дв.	Сумма:				
6	1	Искл. понятий	Низкий	1	0,23	0,1	0,49	0,86	1E-10	-20,9590327	-2,1E-09	1	Искл. понятий	0,94	0,582	0,9	0,4	2,779			
7			Средний	2	0,5	0,8	0,39	0,14	0,42	-0,78963305	-0,33165	2	Анализ отнош.л	0,66	0,817	0,8	0	2,230			
8			Высокий	3	0,27	0,1	0,12	0	0,58	-0,49583204	-0,28758	3	Сложные ассоци	0,73	0,613	0,9	0,5	2,716			
9	2	Анализ отнош. пон	Низкий	1	0,05	0,3	0,67	1				4	Отыскание зако	0,95	0,631	0,8	0,2	2,640			
10			Средний	2	0,23	0,1	0,24	0				5	Отыскание чисе	0,77	0,613	0,7	0,5	2,667			
11			Высокий	3	0,72	0,6	0,09	0				6	Экстраверсия	0,91	0,622	0,9	0,9	3,333			
12	3	Сложные ассоци	Низкий	1	0,05	0	0,18	0,79	Примечание:			7	Рипидность	1	0,882	0,9	0,9	3,723			
13			Средний	2	0,32	0,6	0,55	0,21	так помечено число 1*10 ⁻¹⁰			8	Возбудимость	0,99	0,812	1	0,9	3,686			
14			Высокий	3	0,63	0,4	0,27	0				9	Темп реакций	0,87	0,631	0,8	0,8	3,029			
15	4	Отыскание закон	Низкий	1	0,41	0,5	0,55	0,93				10	Активность	0,74	0,6	0,7	0,7	2,695			
16			Средний	2	0,41	0,5	0,36	0,07				11	Искренность	0,79	0,985	0,9	0,6	3,343			
17			Высокий	3	0,18	0	0,09	0				Сумма:									
18	5	Отыскание чисел	Низкий	1	0,41	0,4	0,49	0,71				9,36	7,787	9,3	6,4						
19			Средний	2	0,54	0,6	0,48	0,29													
20			Высокий	3	0,05	0	0,03	0													
21	6	Экстраверсия	Низкий	1	0,29	0	0,11	0,17													
22			Средний	2	0,53	0,57	0,48	0,33													
23			Высокий	3	0,18	0,43	0,41	0,5													
24	7	Рипидность	Низкий	1	0,35	0,12	0,15	0,17													
25			Средний	2	0,35	0,5	0,44	0,5													
26			Высокий	3	0,3	0,38	0,41	0,33													
27	8	Возбудимость	Низкий	1	0,29	0,63	0,34	0,17													
28			Средний	2	0,42	0,25	0,44	0,33													
29			Высокий	3	0,29	0,12	0,22	0,5													
30	9	Темп реакций	Низкий	1	0,12	0	0,07	0,09													

Рис. П1.5

Література

1. Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование - М.: Вильямс. – 2007. – 1152 с.
2. Продеус А.Н., Захарова Е.Н. Экспертные системы в медицине. – К., «ВЕК+», 1998. – 320 с.
3. Частиков А.П., Гаврилова Т.А., Белов Д.Л. Разработка экспертных систем. Среда CLIPS. - СПб.: БХВ-Петербург. - 2003. – 608 с.
4. В.З. Ковалев, А.С. Татевосян. Интеллектуальные информационные системы: Конспект лекций. - Омск: Изд-во ОмГТУ, 2005.

Лабораторна робота 2 Рангова кореляція діагностичних ознак

Основні цілі роботи:

1. На прикладі аналізу реальних даних (психодіагностика і іриодіагностика) ознайомитися з 2-ма видами оцінок коефіцієнта рангової кореляції: Спірмена і Кендела.
2. Практично освоїти метод перевірки суттєвості оцінки коефіцієнта Кендела.

Додаткові цілі роботи:

Освоїти елементи статистичної обробки результатів спостережень (ранги, коефіцієнт кореляції) в середовищі Excel і елементи програмування на Visual Basic for Applications (VBA).

Завдання:

1. Використовуючи вихідні дані з аркуша "11А + Б_98г", обчисліть коефіцієнт кореляції Пірсона для однієї з ситуацій, помічених "діркою" в таблиці 1.1 на аркуші "Корел" - для обчислень скористайтесь "Майстром функцій" програми Excel.
2. Аналогічним чином обчисліть коефіцієнти рангової кореляції для "дірок" в таблицях 1.2-1.3, 2.2-2.3 - для обчислень скористайтесь листом "РанКорСпір". При цьому на аркуші "РанКорСпір": X_i і Y_i - масиви вихідних даних; кнопки «Ранговая кореляція по Спірмену» і «Ранговая кореляція по Кенделу» - обчислення коефіцієнтів рангової кореляції; r_{sa} і комірка I2, r_{sb} і комірка K2 - коефіцієнти Спірмена (некоректний і коректний, відповідно); T_a і комірка K5, T_b і комірка J2 - коефіцієнти Кендела (некоректний і коректний, відповідно).
3. Зіставте ваші результати з правильними відповідями на аркуші "Відповіді".
4. Обчисліть S/sig Кендела для однієї з "дірок" таблиці 3.3 (порівняйте з правильним результатом). Обчисліть межі довірчого інтервалу для даного випадку.

Зміст звіту:

1. Уявіть і прокоментуйте результати Ваших обчислень.
2. Зіставте коефіцієнти Пірсона, Спірмена і Кендела по вмісту таблиць на аркуші «Відповіді».

Самостійна робота:

Розробіть програму обчислення рангів за формулами (П2.1) - (П2.3). Перевірте, як впливає вибір визначення рангу на результати обчислень коефіцієнтів рангової кореляції. Спробуйте самостійно отримати графік, аналогічний наявному на аркуші «Графік».

Теоретичні відомості

На сьогоднішній день найбільш відомими є два підходи до визначення коефіцієнта рангової кореляції - Спірмена і Кендела. При практичних обчисленнях КРК Спірмена найбільш часто користуються формулою:

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}, \quad (\text{П2.1})$$

де, $d_i = p_i - q_i$; p_i і q_i - послідовності рангів; n - кількість елементів у кожній з цих послідовностей. Відповідно, КРК Кендела зазвичай обчислюють у відповідності з виразом:

$$\tau = \frac{S}{\frac{1}{2}n(n-1)}, \quad (\text{П2.2})$$

де S - показник узгодженості пар рангів, який зручно представити у вигляді:

$$S = \frac{1}{2} \sum_i \sum_j \text{sgn}(p_j - p_i) \text{sgn}(q_j - q_i). \quad (\text{П2.3})$$

У програмі Excel застосовується наступний алгоритм обчислення рангів:

$$R_i = \frac{1}{2} \sum_{k=1}^n \text{sgn}_-(x_i - x_k) + \frac{n+2}{2}, \quad (\text{П2.4})$$

де

$$\text{sgn}_-(x) = \begin{cases} 1, & x > 0 \\ -1, & x \leq 0 \end{cases}$$

За наявності пов'язаних рангів формули (П2.1)-(П2.2) призводять до значної похибки (наприклад, коефіцієнт автокореляції Кендела може як завгодно близько наближатись до нуля!) У цьому зв'язку замість формул (П2.1)-(П2.2) доцільно застосовувати вираз для узагальненого коефіцієнта кореляції :

$$\Gamma = \frac{\sum_i \sum_j a_{ij} b_{ij}}{\sqrt{\sum_i \sum_j a_{ij}^2 \sum_i \sum_j b_{ij}^2}}, \quad (\text{П2.5})$$

де кожній парі елементів першої послідовності, скажемо i і j , приписана якась величина a_{ij} , і аналогічно кожній парі другої послідовності приписана величина b_{ij} (при цьому постулюється $a_{ii} = b_{ii} = 0$ для всіх $i=1, \dots, n$). Тоді алгоритм коректного обчислення КРК Кендела виходить з (П2.5) шляхом підстановки

$$a_{ij} = \text{sgn}(p_j - p_i), \quad b_{ij} = \text{sgn}(q_j - q_i),$$

$$\text{де } \text{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}.$$

Аналогічно, КРК Спірмена коректно обчислюється, якщо в (П2.5) прийняти $a_{ij} = p_j - p_i$, $b_{ij} = q_j - q_i$. До речі, "звичайний" коефіцієнт кореляції Пірсона виходить з (П2.3), якщо прийняти $a_{ij} = x_j - x_i$, $b_{ij} = y_j - y_i$. Остання обставина дозволяє стверджувати, що КРК Спірмена є ні що інше як коефіцієнт кореляції Пірсона, розрахований на рангах.

Достовірність обчислень КРК Кендела можна оцінювати по-різному. По-перше, обчислюючи відношення S/σ_s , де σ_s - СКВ величини S . Тоді наприклад, враховуючи, що розподіл S при $n = 30$ близький до нормального, при $|S/\sigma_s| > 1,96$ впливає, що з імовірністю більше 0,95 справедливий висновок про статистичну незалежність. При $|S/\sigma_s| > 1,96$ ця ймовірність перевищує 0,99, а при $|S/\sigma_s| > 3,3$ - перевищує 0,999. За наявності зв'язаних рангів σ_s^2 обчислюється по досить великій формулі:

$$\sigma_s^2 = 1/18 * (c1 - a1 - b1) + c2 * a2 * b2 + c3 * a3 * b3, \quad (\text{П2.6})$$

де

$$c1 = n * (n-1) * (2 * n + 5), \quad a1 = \sum_t [t * (t-1) * (2 * t + 5)],$$

$$b1 = \sum_u [u * (u-1) * (2 * u + 5)], \quad c2 = 1 / (9 * n * (n-1) * (n-2)),$$

$$a2 = \sum_t [t * (t-1) * (t-2)], \quad b2 = \sum_u [u * (u-1) * (u-2)],$$

$$c3 = 1 / (2 * n * (n-1)), \quad a3 = \sum_t [t * (t-1)],$$

$$b3 = \sum_u [u * (u-1)], \quad \sum_t [] = \sum_t [].$$

По-друге, можна обчислити довірчий інтервал:

$$\tau = (t \pm x * (2 * (1 + 2 * x^2 / n - t^2) / n)^{(1/2)}) / (1 + 2 * x^2 / n), \quad (\text{П2.7})$$

де x - нормоване відхилення (тобто $x = S/\sigma_s$), відповідне заданій довірчій ймовірності.

Коментарі до програми, реалізованої в середовищі Excel:

Програма реалізована на 5-ти аркушах.

1. Аркуш "11A + Б_98г" (рис. П2.1): результати тестування та обстеження.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		Таб.раб.№4: коефіцієнт рангової кореляції																
2	Перед Вами таблиця результатів психологічної та іридологічної діагностики -																	
3	результати представлені по 5-бальній шкалі: 3-умерено, 4-повищено,																	
4	5 суттєво підвищено, 2- понижено, 1- суттєво понижено																	
5																		
6		Психологічне тестування							Іридологія					Вспомогательные				
7	Фамилия	экстр.эмоц.	темп.	истц.эмоц.	безбт.	смел.	напр.	трабе.	нрв.к	зрчк	ЗП	парас	симп	парас	симп			
8	Войтенко	5	5	5	3	3	4	3	3	5	2	4	4	4	2		4	2
9	Воробец	5	3	4	1	4	4	3	3	5	1	5	5	4	4		4	4
10	Горбань	5	3	4	3	3	5	4	3	3	3	2	4	2	4		2	4
11	Грошева	2	4	3	4	3	3	2	3	4	4	4	3	1	4		1	4
12	Дегтярева	4	5	4	3	3	4	2	4	5	2	5	4	1	2		1	2
13	Карась	4	2	4	3	3	4	3	3	4	4	4	4	2	3		2	3
14	Кур	4	2	5	3	5	4	5	2	2	5	3	4	2	3		2	3
15	Кучерявый	5	2	3	2	5	4	5	2	2	3	2	3	2	2		2	2
16	Павлов	4	1	5	3	5	5	4	1	2	5	2	3	3	2		3	2
17	Пневская	5	4	5	3	2	4	3	3	5	2	4	4	4	4		4	4
18	Скрынник	5	2	5	3	3	5	3	3	2	3	3	4	2	4		2	4
19	Тулиголовец	5	1	4	3	3	4	4	2	4	2	3	3	3	1		3	1
20	Чумак	5	3	5	3	4	4	4	3	2	3	3	3	4	4		4	4
21	Яворский	5	4	5	3	2	3	3	3	4	2	4	3	5	2		5	2
22	Бойченко	5	3	5	3	4	5	4	2	3	1	3	4	2	5		2	5
23	Верникова	5	4	4	3	3	4	4	3	2	3	3	3	2	2		2	2
24	Врублевская	5	2	5	5	4	5	5	3	5	4	3	3	2	4		2	4
25	Горук	5	2	4	1	3	4	3	3	5	4	4	4	3	4		3	4
26	Гриценко	3	4	4	3	3	4	3	3	3	3	3	4	4	2		4	2
27	Климов	2	3	3	3	2	2	3	4	4	3	3	4	2	2		2	2
28	Т...	2	4	2	2	2	2	2	4	1	4	2	2	1	2		1	2

Рис. П2.1

2. Аркуш "Корел" (рис. П2.2): результати оцінювання коефіцієнта кореляції і його достовірності.

3. Аркуш "РанКорСпір" (рис. П2.3): обчислення коефіцієнтів рангової кореляції та їх достовірності.

4. Аркуш "Графік" (рис. П2.4): графік коефіцієнта автокореляції Кендела для формули (П2.2) при наявності пов'язаних рангів.

5. Аркуш "Відповіді" (рис. П2.5): те ж, що Аркуш "Корел", але без "пропусків" в таблицях.

Экспертні системи в медицині

X26 Иридодиагностика (П) - S/sig Кендала															
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1. Не учитывается наличие связанных рангов										2. Учитывается наличие связанных рангов					
(ранги вычисляются по алгоритму Excel)										(ранги вычисляются по алгоритму Excel)					
Табл.1.1															
Иридодиагностика - коэфф. корреляции Пирсона															
Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.									
экстрав.		-0,233	-0,097	0,027	0,445	0,096									
эмоц.возб.	0,250		0,498	0,180	-0,084	0,113									
темп реакц.	0,131	-0,049		0,032		0,056									
истощ.НС	0,143	0,082	0,231	0,158	-0,141	0,246									
эмоц.зрел.		0,390	-0,529		-0,128	-0,061									
беззаботн.	-0,041	0,041	-0,237	0,170	0,109	0,111									
смелость	-0,202	0,104	-0,518	-0,154	0,137	-0,006									
напряженн.	0,398	-0,150		0,150	-0,211										
Табл.1.2															
Иридодиагностика - ранг. коэф. корр-ции Спирмена (r _{sa})										Иридодиагностика - ранг. коэф. корр-ции Спирмена (r _{sb})					
Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.	Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.		
экстрав.		0,007	0,172	0,254	0,539	0,499	экстрав.		-0,411	-0,066	0,092	0,092	0,092	0,092	0,092
эмоц.возб.	0,377		0,402	0,075	-0,080	0,092	эмоц.возб.	0,348		0,457	0,113	0,113	0,113	0,113	0,113
темп реакц.	0,273	0,141		0,154		0,376	темп реакц.	0,087	-0,034		0,056	0,056	0,056	0,056	0,056
истощ.НС	0,196	0,269	0,167	-0,133	-0,196	0,432	истощ.НС	0,105	0,203	0,147	-0,061	-0,061	-0,061	-0,061	-0,061
эмоц.зрел.		0,243	-0,511		-0,145	0,141	эмоц.зрел.		0,291	-0,361					
беззаботн.	0,050	0,150	-0,213	0,098	0,022	0,433	беззаботн.	-0,127	0,033	-0,289	0,111	0,111	0,111	0,111	0,111
смелость	-0,334	0,124	-0,644	-0,389	-0,220	0,148	смелость	-0,333	0,162	-0,501	-0,006	-0,006	-0,006	-0,006	-0,006
напряженн.	0,385	-0,040		0,109	-0,281		напряженн.	0,340	-0,093						
Табл.1.3															
Иридодиагностика - ранг коэф. корр-ции Кендала (T _a)										Иридодиагностика - ранг коэф. корр-ции Кендала (T _b)					
Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.	Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.		

Рис. П2.2

J1 T _b															
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Ранговая корреляция по Спирмену						Ранг корр-ция по Кендалу						Sigsig			
						Sum di ²	n	r _{sa}	T _b	r _{sb}	S	sig	s/sig	D	
						2105	30,00	0,53	0,57	0,5553	180	49,48834	3,637221		
	Xi	Yi	rangX	rangY	di	di ²	t, u								
Номер	эмоц.вз	напржн.													
1	5	3	27	9	18	324	0								
2	3	3	14	9	5	25	7	0	T _a	Z1	Z2				
3	3	3	14	9	5	25	0	0	0,4138	0,898274	0,808006				
4	4	3	21	9	12	144	0	0	0	0					
5	5	4	27	25	2	4	0	1	1	1	0				
6	2	3	5	9	-4	16	0	1	0	0	0	0			
7	2	2	5	2	3	9	0	0	1	1	1	1	1	1	
8	2	2	5	2	3	9	0	0	0	1	1	1	1	1	
9	1	1	1	1	0	0	16	1	1	1	1	1	1	1	
10	4	3	21	9	12	144	0	1	1	1	0	1	0	0	
11	2	3	5	9	-4	16	0	0	1	0	0	0	0	1	
12	1	2	1	2	-1	1	0	1	1	0	0	0	0	1	
13	3	3	14	9	5	25	0	1	0	0	1	1	1	1	
14	4	3	21	9	12	144	0	0	1	0	0	1	1	1	
15	3	2	14	2	12	144	0	1	0	0	-1	1	1	1	
16	4	3	21	9	12	144	0	1	0	0	1	0	0	0	
17	2	3	5	9	-4	16	0	0	-1	0	0	1	0	0	
18	2	3	5	9	-4	16	0	0	0	-1	0	0	0	1	
19	4	3	21	9	12	144	0	0	0	0	1	0	0	0	
20	3	4	14	25	-11	121	0	-1	1	1	-1	0	0	-1	
21	4	4	21	25	-4	16	0	0	0	1	1	0	0	1	
22	5	3	27	9	18	324	0	-1	-1	0	0	0	0	0	
23	3	4	14	25	-11	121	0	-1	0	0	-1	1	1	1	

Рис. П2.3

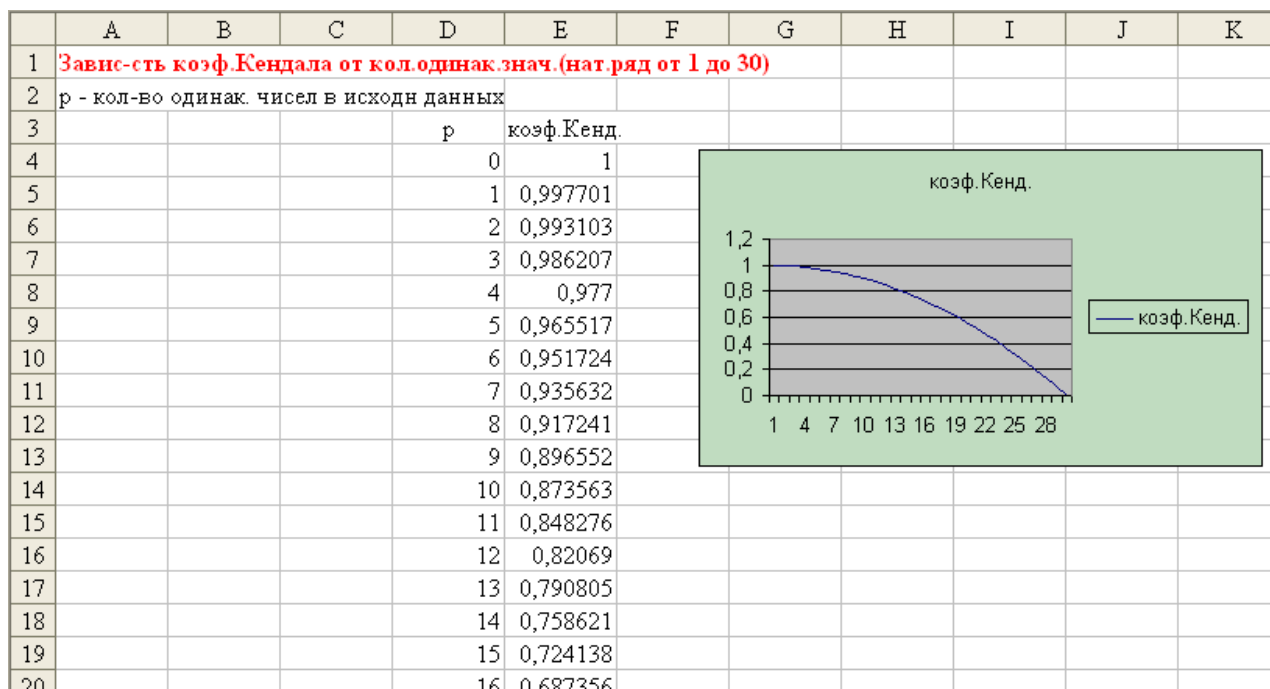


Рис. П2.4

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	1. Не учитывается наличие связанных рангов										2. Учитывается наличие связанных рангов					
2	(ранги вычисляются по алгоритму Excel)										(ранги вычисляются по алгоритму)					
3																
4	Приодиагностика - коэф. корреляции Пирсона															
5	Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.									
6	экстрав.	0,142	-0,233	-0,097	0,027	0,445	0,096									
7	эмоц.возб.	0,250	-0,199	0,498	0,180	-0,084	0,113									
8	темп реакц.	0,131	-0,049	-0,082	0,032	0,496	0,056									
9	истоц.НС	0,143	0,082	0,231	0,158	-0,141	0,246									
10	эмоц.зрел.	-0,517	0,390	-0,529	-0,150	-0,128	-0,061									
11	беззаботн.	-0,041	0,041	-0,237	0,170	0,109	0,111									
12	смелость	-0,202	0,104	-0,518	-0,154	0,137	-0,006									
13	напряженн.	0,398	-0,150	0,612	0,150	-0,211	0,227									
14																
15	Приодиагностика - ранг. коэф. корр-ции Спирмена (r_s)															
16	Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.									
17	экстрав.	0,433	0,007	0,172	0,254	0,539	0,499									
18	эмоц.возб.	0,377	-0,275	0,402	0,075	-0,080	0,092									
19	темп реакц.	0,273	0,141	0,052	0,154	0,417	0,376									
20	истоц.НС	0,196	0,269	0,167	-0,133	-0,196	0,432									
21	эмоц.зрел.	-0,407	0,243	-0,511	-0,206	-0,145	0,141									
22	беззаботн.	0,050	0,150	-0,213	0,098	0,022	0,433									
23	смелость	-0,334	0,124	-0,644	-0,389	-0,220	0,148									
24	напряженн.	0,385	-0,040	0,420	0,109	-0,281	0,232									
25																
26	Приодиагностика - ранг. коэф. корр-ции Кендала (T_k)															
27	Психодиаг	трабекул	нервн.кол	зрачок	ЗП сужен	парасим.	сипатич.									
28	экстрав.	0,183	-0,411	-0,066	0,000											
29	эмоц.возб.	0,348	-0,286	0,457	0,111											
30	темп реакц.	0,087	-0,034	-0,051	0,000											
31	истоц.НС	0,105	0,203	0,147	-0,000											
32	эмоц.зрел.	-0,386	0,291	-0,361	-0,000											
33	беззаботн.	-0,127	0,033	-0,289	0,000											
34	смелость	-0,333	0,162	-0,501	-0,200											
35	напряженн.	0,340	-0,093	0,426	0,100											

Рис. П2.5

Література

1. Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование - М.: Вильямс. – 2007. – 1152 с.
2. Продеус А.Н., Захарова Е.Н. Экспертные системы в медицине. – К.: «ВЕК+», 1998. – 320 с.
3. Кендэл М. Ранговые корреляции. – М.: Статистика, 1975. – 213с.
4. Хеттманспергер Т. Статистические выводы, основанные на рангах. –М.: Финансы и статистика, 1987. – 333 с.

Лабораторна робота 3 Рекурентна форма формули Байєса

Основні цілі роботи:

1. Закріпити наявні знання про рекурентну форму формули Байєса.
2. Оцінити можливості створення демонстраційних версій діагностичних систем в середовищі Excel.

Додаткові цілі роботи:

1. Освоїти техніку створення демонстраційних додатків в середовищі Excel
2. Закріпити (освоїти) техніку редагування і написання програм в середовищі Excel

Завдання:

1. Скопіювати файл даної лабораторної роботи в робочу директорію. Переіменувати цей файл і запустити його на виконання. Натисканням графічної кнопки «Клацни тут перед початком роботи» привести систему в робочий стан.

2. Натисненням однієї з кнопок «Есть1», «Слабо1», «нет1» задати значення ознаки «Ознака 1». Зафіксувати у своєму блокноті результати автоматичних обчислень, представлені таблично і графічно. Проробити вручну викладки, що підтверджують правильність автоматичних обчислень.

3 . Повторити дії за пунктом 2, але вже для ознаки « Ознака 2».

4 . Зробити скидання результатів обчислень (кнопка « Скидання »).

5 . Змінити вміст комірок таблиці БЗ1 і повторити дії з пп.2 і 3.

6 . Змінити вміст комірок таблиці Бз2 і повторити дії з пп.2 і 3.

7 . Включити редактор Visual Basic (сервіс-макрос-редактор Visual Basic), у вікні "Проект" знайти "Модуль1" і подвійним клацанням лівої клавіші миші активізувати режим перегляду і редагування тексту програми. Вивчити текст програми і намалювати її збільшену блок-схему (кожен блок - окрема підпрограма - макрос).

8 . Вивчити текст підпрограми (макросу) " есть_1 . Намалювати блок-схему підпрограми "єсть_1". Порівняти цю підпрограму з іншими аналогічними підпрограмами.

9 . Сформулювати пропозиції щодо скорочення обсягу програми.

10 . Зробити спробу модернізувати програму, додавши "Ознака 3". Зафіксувати в блокноті труднощі, які зустрілися при такій модернізації, а також проаналізувати труднощі, які можуть зустрітися при подальшому збільшенні кількості ознак. Сформулювати пропозиції щодо подолання (хай і часткового) цих труднощів.

Зміст звіту

1. Представити і прокоментувати результати автоматичних обчислень і результати ручної їх перевірки.
2. Представити дві блок-схеми: збільшену блок-схему всієї програми і блок-схему підпрограми "єсть_1".
3. Викласти свої міркування про недоліки наявної діагностичної програми.
4. Представити свої пропозиції з удосконалення наявної програми.

Теоретичні відомості

При обчисленнях за формулою Байєса в якості міри вірогідності висновку про наявність тієї чи іншої патології використовується умовна ймовірність $P(H_j / E_i)$:

$$P(H_j / E_i) = \frac{P(H_j)P(E_i / H_j)}{P(E_i)} \quad (ПЗ.1)$$

Для безлічі ознак $E = \{E_1, \dots, E_I\}$ формула Байєса набуває вигляду:

$$P(H_j / E_1, \dots, E_I) = \frac{P(H_j)P(E_1, \dots, E_I / H_j)}{P(E_1, \dots, E_I)} \quad (ПЗ.2)$$

або, для статистично незалежних ознак,

$$P(H_j / E_1, \dots, E_I) = P(H_j) \prod_{i=1}^I \frac{P(E_i / H_j)}{P(E_i)} \quad (ПЗ.3)$$

Вирішальне правило при використанні формули Байєса полягає у пошуку максимуму функції $P(H_j / E_1, \dots, E_I)$.

У додатках зручніше рекурентний варіант формули Байєса:

$$P(H_j / E_1, \dots, E_I) = P(H_j / E_1, \dots, E_{I-1}) \frac{P(E_I / H_j)}{P(E_I)} \quad (ПЗ.4)$$

що дозволяє робити обчислення по мірі оцінювання нових симптомів, не чекаючи моменту, коли будуть оцінені всі I симптомів. Цей принцип дозволяє припинити облік нових симптомів, якщо лікар вважатиме оцінку ймовірності аналізованої гіпотези досить високою.

Коментарі до діагностичної системи, реалізованої в середовищі Excel :

Діагностична система реалізована на аркуші "Байес_рекур" (рис. ПЗ.1). У лівому верхньому кутку цього аркуша розташована таблиця Б31 значень апіорного розподілу ймовірностей. Під нею розміщено таблицю Б32 значень умовних розподілів ймовірностей $P(E_i / H_j)$. Під Б32 розташовані таблиці з резуль-

татами обчислень $P(H_j / E_1, \dots, E_D)$. Кількість таблиць дорівнює кількості заданих ознак.

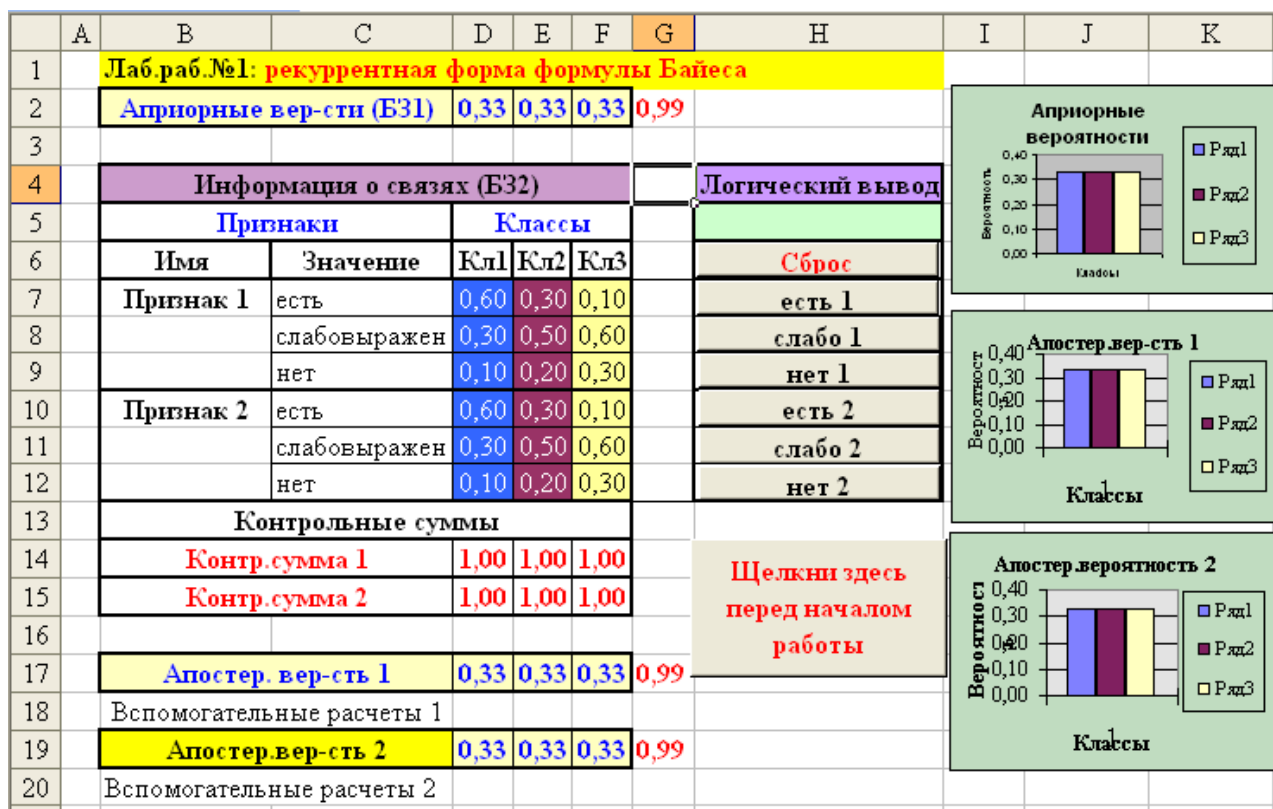


Рис. ПЗ.1

У правій частині аркуша знаходяться автоматично оновлювані графіки $P(H_j / E_1, \dots, E_D)$.

У центрі аркуша розташований ряд кнопок, натискання яких дозволяє задавати те або інше значення ознак і робити обчислення за формулою (ПЗ.4) в автоматичному режимі.

Самостійна робота:

1. Виробити пропозиції щодо вдосконалення наявної демо - програми.
2. Реалізувати вироблені пропозиції шляхом модернізації наявної демо - програми.

Література

1. Продеус А.Н., Захарова Е.Н. Экспертные системы в медицине. – К.: «ВЕК+», 1998. - 320 с.

Лабораторна робота 4 Адаптивне вимірювання розподілу апіорних ймовірностей

Основні цілі роботи:

Провести моделювання процесу адаптивного вимірювання розподілу апіорних ймовірностей $P(H_j)$, використовуючи обчислювальні можливості середовища Excel.

Додаткові цілі роботи:

Освоїти елементи програмування на Visual Basic for Applications (VBA).

Завдання:

1. Скопіювати файл лабораторної роботи в робочу директорію. Перейменувати цей файл (наприклад, Lab_4_volodia). Запустити файл.

2. Натисканням графічної кнопки «Клацни тут перед початком роботи» привести систему в робочий стан.

3. Натисканням кнопки «N=0» зробити скидання результатів попередніх обчислень (якщо таке скидання потрібне).

4. Натисканням кнопки «N = N + 1» збільшити вміст лічильника на 1 - це відповідає реєстрації знову прийшовшого пацієнта. Зафіксувати у своєму блокноті результати автоматичних обчислень, представлених в комірках "N=", "(N-1)/N=", "1/N=". Проробити вручну викладки, що підтверджують правильність автоматичних обчислень .

5 . Натисненням однієї з кнопок «Клас 1», «Клас 2» або «Клас 3» задати результати об'єктивного діагнозу даного пацієнта (взяти їх з таблиці на аркуші "Результати діагностики"). Зафіксувати у своєму блокноті результати автоматичних обчислень, представлених в комірках "функція -діагноз", "архівна апіорна ймовірність", "оновлена апіорна ймовірність". Проробити вручну викладки, що підтверджують правильність автоматичних обчислень.

6 . Повторити дії по пунктах 4-5 до вичерпання даних таблиці на аркуші "Результати діагностики".

7 . Побудувати графіки функцій "функція -діагноз", "архівна апіорна ймовірність", "оновлена апіорна ймовірність" (для цього клацнути по графіку правою клавішею миші і використовувати позицію "вихідні дані ..." меню, що з'явилося).

8 . По закінченні обчислень зробити скидання результатів обчислень і очищення графіків (графік "Оновлена апр. ймовірність" попередньо скопіювати на аркуш "Архів").

9 . Включити редактор Visual Basic (сервіс-макрос-редактор Visual Basic), у вікні "Проект" знайти " Модуль 2" і подвійним клацанням лівої клавіші миші

активізувати режим перегляду і редагування тексту програми Вивчити текст програми і намалювати її збільшену блок-схему (кожен блок - окрема підпрограма - макрос).

10 . Вивчити і забезпечити коментарями текст підпрограм (макросів) "Сброс_N", "Увел_на_1_і_вчислення" і "Класс1".

11. Подумати над можливою ситуацією: отримана в результаті обчислень оцінка $P(H_j)$ містить хоча б одне нульове значення. Як це відіб'ється на обчисленні за формулою Байєса? Виробити (і обґрунтувати) пропозиції щодо усунення даної ситуації. Спробувати модернізувати програму, усунувши помічений недолік. Вказати труднощі, які зустрілися при такій модернізації.

Зміст звіту:

1. Представити і прокоментувати результати обчислень і ручної їх перевірки.
2. Уявити збільшену блок-схему програми та тексти підпрограм з введеними коментарями.
3. Викласти свої міркування про недоліки наявної програми і свої пропозиції з удосконалення наявної програми.

Самостійна робота:

1. Виробити пропозиції щодо вдосконалення наявної програми.
2. Реалізувати вироблені пропозиції шляхом модернізації наявної програми.

Теоретичні відомості

При обчисленнях за формулою Байєса

$$P(H_j / E_i) = \frac{P(H_j)P(E_i / H_j)}{P(E_i)} \tag{П4.1}$$

або по рекурентному варіанту формули Байєса:

$$P(H_j / E_1, \dots, E_l) = P(H_j / E_1, \dots, E_{l-1}) \frac{P(E_l / H_j)}{P(E_l)}, \tag{П4.2}$$

необхідно мати інформацію про розподіл апіорних ймовірностей $P(H_j)$.

В разі відсутності такої інформації перед початком обчислень вважають

$$P(H_j) = 1/J, \quad j = 1, \dots, J,$$

а в процесі обчислень оцінюють $P(H_j)$ по мірі надходження нової інформації.

Алгоритм обчислень має вигляд:

$$P_{oN}(H_j) = 1 / N \sum_{n=1}^N \delta(j - j_{on}), \tag{П4.3}$$

де

$$\delta(j) = \begin{cases} 1, & j = 1 \\ 0, & \text{інш. } j \end{cases}$$

j_{on} - код рішення за результатами об'єктивних обстежень, прийнятого для n -того пацієнта.

Рекурентна форма алгоритму (П4.3):

$$P_{oN}(H_j) = \frac{N-1}{N} P_{o(N-1)}(H_j) + \frac{1}{N} \delta(j - j_{oN}) \quad (П4.4)$$

Формула (П4.4) дозволяє робити обчислення $P(H_j)$ у міру обстеження пацієнтів, не чекаючи моменту, коли будуть зібрані результати обстежень N пацієнтів. Цей принцип фактично означає додаткове навчання діагностичної системи в процесі її практичної експлуатації.

Коментарі до програми, реалізованої в середовищі Excel:

Програма реалізована на 3-х аркушах: "Апріор_ймовірність", "Рез-ти діагностики" і "Архів".

1. **Аркуш "Апріор_ймовірність"** (рис. П4.1): зліва - зона скидання в 0 лічильника пацієнтів і створення номера N нового пацієнта; посередині - зона задання функції-діагнозу, описувана одиничною дельта-функцією, а також двох масивів відліків оцінки апіорного розподілу ймовірностей $P(H_j)$ - архівного (для пацієнта з номером $N-1$) і оновленого (для пацієнта з номером N); праворуч - заготовки для графіків $P(H_j)$ (вихідні дані для графіків слід вказувати вручну).

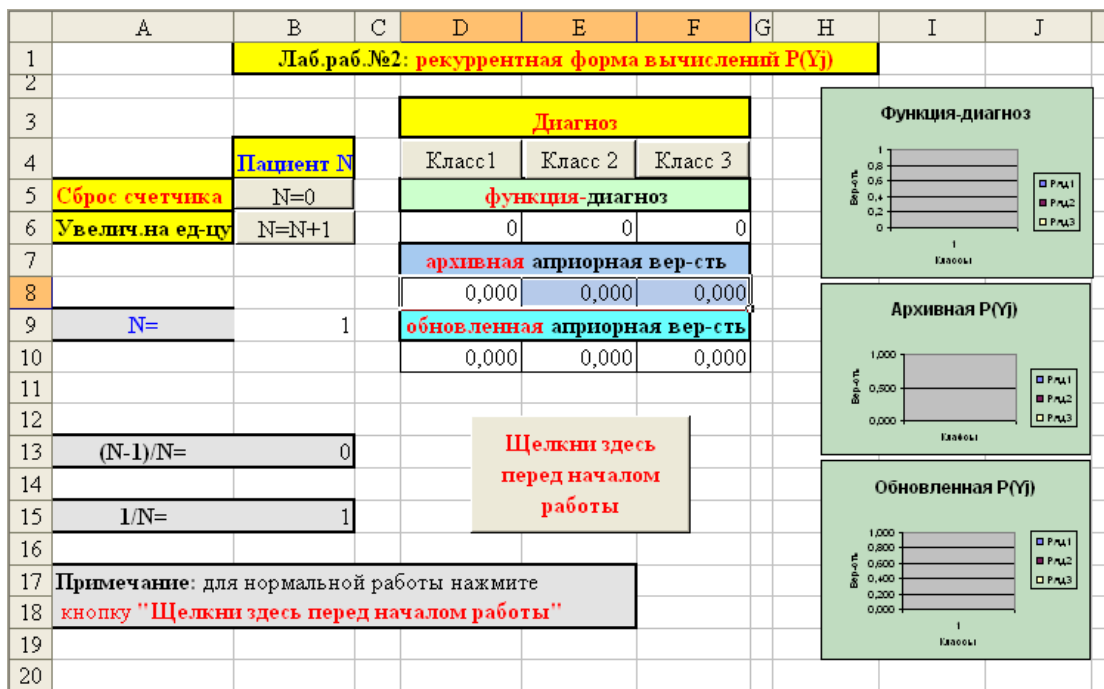


Рис. П4.1

2. Аркуш "Апріор_ймовірність" (рис. П4.2) таблиця результатів "об'єктивної діагностики".

	A	B	C	D	E	F	G	H	I	
1		Результаты диагностики								
2										
3	N	1	2	3	4	5	6	7	8	
4	Вар.1	1	1	2	3	3	3	2	1	
5	Вар.2	2	2	1	3	2	1	3	3	
6	Вар.3	3	3	2	1	1	1	2	2	
7	Вар.4	1	1	2	1	1	3	3	2	
8	Вар.5	2	2	3	3	1	1	3	2	
9	Вар.6	1	1	2	2	3	3	1	1	
10	Вар.7	1	1	1	2	2	3	3	3	
11	Вар.8	3	3	2	1	1	1	2	2	
12	Вар.9	1	1	2	2	1	3	3	3	
13	Вар.10	2	2	1	3	3	1	3	2	
14	Вар.11	1	2	2	1	3	3	1	3	
15										
16	Примечание: в желтой зоне таблицы - номера классов (коды заболеваний)									
17										

Рис. П4.2

3. Аркуш "Архів" - сховище графіків.

Література

1. Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование - М.: Вильямс. – 2007. – 1152 с.
2. Продеус А.Н., Захрабова Е.Н. Экспертные системы в медицине. – К.: «ВЕК+», 1998. – 320 с.

Лабораторна робота 5 Адаптивне вимірювання розподілу умовних ймовірностей

Основні цілі роботи:

Провести моделювання процесу адаптивного вимірювання умовного розподілу ймовірностей $P(E_i/H_j)$, використовуючи обчислювальні можливості середовища Excel.

Додаткові цілі роботи:

Освоїти елементи програмування на Visual Basic for Applications (VBA).

Завдання:

1. Скопіювати файл лабораторної роботи в робочу папку і запустити програму.
2. Натисканням кнопки «Клацни тут перед початком роботи» привести систему в робочий стан (режим покрокових обчислень).
3. Натисканням кнопки «N=0» зробити скидання результатів попередніх обчислень (якщо таке скидання потрібне).
4. Натисканням кнопки «N = N +1» збільшити вміст лічильника на 1 - це відповідає реєстрації знову прийшовшого пацієнта.
5. Провести обчислення умовної ймовірності $P(E_i/H_j)$, користуючись вихідними на аркуші "Вихідні дані" (номер варіанта відповідає номеру робочого місця), для чого:
 - а) натисненням однієї з кнопок «Клас 1», «Клас 2» або «Клас 3» задати результат об'єктивного діагнозу даного пацієнта;
 - б) натисненням розташованих праворуч кнопок «є 1», «слабо 1» і т.д. задати значення відповідних ознак;
 - в) клацнути курсором по кнопці "3 -дельта - функція";
 - г) клацнути курсором по кнопці "Оновл. $P(E_i/H_j)$ " - в результаті в нижчій зоні з'являться результати обчислень $P(E_i/H_j)$ за формулою (П5.4).
6. Повторити дії по пунктах 4-5 до вичерпання даних таблиці на аркуші "Исходн.данние".
7. Користуючись Майстром діаграм, побудувати графік функції $P(X_i/H_j)$ (на аркуші "Графіки") і замалювати його для звіту.
8. По закінченні обчислень зробити скидання результатів обчислень (N= 0).
9. Включити редактор Visual Basic (сервіс-макрос-редактор Visual Basic), у вікні "Проект" знайти " Модуль 1" і подвійним клацанням лівої клавiшi мишi активiзувати режим перегляду i редагування тексту програми. Вивчити текст програми, спробувати знайти в ній недоліки.

Зміст звіту:

1. Описати постановку задачі і спосіб її вирішення з застосуванням Excel.
2. Представити і прокоментувати результати обчислень, вказавши, в тому числі, виявлені в програмі недоліки.

Самостійна робота:

1. Виробити пропозиції щодо вдосконалення наявної програми.
2. Реалізувати вироблені пропозиції шляхом модернізації наявної програми

Контрольні питання:

1. Що буде, якщо отримана в результаті обчислень оцінка $P(E_i/H_j)$ містить хоча б одне нульове значення? Як це відіб'ється на обчисленнях за формулою Байєса? Спробуйте модернізувати програму, усунувши помічений недолік. Вкажіть труднощі, які вам зустрілися при такій модернізації.
2. Наведена програма дає правильні результати при виконанні 2-х умов: а) немає пропусків ознак, б) діагностується одне і те ж захворювання. Чому? Спробуйте удосконалити програму так, щоб вона завжди працювала вірно.

Теоретичні відомості

При обчисленнях за формулою Байєса

$$P(H_j / E_i) = \frac{P(H_j)P(E_i / H_j)}{P(E_i)} \tag{П5.1}$$

або по рекурентному варіанту формули Байєса:

$$P(H_j / E_1, \dots, E_l) = P(H_j / E_1, \dots, E_{l-1}) \frac{P(E_l / H_j)}{P(E_l)}, \tag{П5.2}$$

необхідно мати інформацію про умовний розподіл ймовірностей $P(E_i/H_j)$.

В разі відсутності такої інформації перед початком обчислень вважають

$$P(H_j) = 1/J, \quad j = 1, \dots, J,$$

де K_i - число градацій за шкалою вимірювань i -тої ознаки, і оцінюють $P(E_i/H_j)$ у процесі обчислень, по мірі надходження нової інформації.

Алгоритм обчислень має вигляд:

$$P_{oN}(x_{ik} | H_j) = 1 / N_{ij} \sum_{n=1}^{N_{ij}} \delta(i - i_{on}, k - k_{on}, j - j_{on}), \tag{П5.3}$$

де

$$\delta(j) = \begin{cases} 1, & i = 0, k = 0, j = 1 \\ 0, & \text{інш.} \quad j \end{cases},$$

i_{0n} - код заданої ознаки, k_{0n} - код заданого значення-того ознаки; i_{0n} - код рішення за результатами об'єктивних обстежень, прийнятого для i -того пацієнта.

Рекурентна форма алгоритму (П5.3):

$$P_{oN_{ij}}(H_j) = \frac{N_{ij} - 1}{N_{ij}} P_{o(N_{ij}-1)}(H_j) + \frac{1}{N_{ij}} \delta(i - i_{0N}, k - k_{0N}, j - j_{0N}) \quad (П5.4)$$

Формула (П5.4) дозволяє робити обчислення $P(H_j)$ у міру обстеження пацієнтів, не чекаючи моменту, коли будуть зібрані результати обстежень N пацієнтів. Цей принцип фактично означає вдосконалення діагностичної системи в процесі її практичної експлуатації.

Коментарі до програми, реалізованої в середовищі Excel:

Програма реалізована на 3-х аркушах: "Обчислення", "Вихідні дані", "Графіки".

1. **Аркуш "Обчислення"** (рис. П5.1): зліва - зона скидання в 0 лічильника пацієнтів і створення номера N нового пацієнта; посередині - 1) зона завдання функції-діагнозу, описуваної одиничною дельта-функцією, 2) зона формування 3-вимірної дельта-функції, що містить інформацію про задані ознаки і відповідного діагнозу, 3)зони архівних і поточних результатів обчислень $P(E_i/H_j)$; праворуч - кнопки для завдання значень діагностичних ознак і таблиця типу "медична пам'ять".

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1		Лаб.раб.№3: рекуррентная форма вычисления P(X Y)																	
3		Диагноз				Ошибка 1: N не зависит от i (пропуск признаков не учтен)													
4		Пациент N	Класс	Класс	Класс	Ошибка 2: N не зависит от j (при разных j некорректно вычисляется P(X Y))													
5	Сброс счет	N=0	функция-диагноз																
6	Увелич.ед	N=N+1																	
7			3-дельта-функция		Обновл.P(X)														
8																			
9	N=	0																	
10																			
11																			
12																			
13	(N-1)/N	0																	
14																			
15	1/N	0																	
16	Щелкни здесь перед началом работы																		
17																			
18																			
19																			
20																			

Рис. П5.1

	A	B	C	D	E	F	G	H	I	J
1		Результаты диагностики								
2										
3	N	1	2	3	4	5	6	7	8	
4	Вар.1	1(e_1,e_2)	1(e_1,c_2)	1(n_1,c_2)	1(e_1,c_2)	1(e_1,n_2)	1(c_1,c_2)	1(e_1,e_2)	1(c_1,n_2)	
5	Вар.2	2(e_1,n_2)	2(c_1,n_2)	2(c_1,e_2)	2(c_1,c_2)	2(n_1,e_2)	2(e_1,e_2)	2(e_1,n_2)	2(n_1,n_2)	
6	Вар.3	3(c_1,n_2)	3(e_1,n_2)	3(n_1,c_2)	3(e_1,c_2)	3(n_1,e_2)	3(e_1,e_2)	3(e_1,e_2)	3(c_1,n_2)	
7	Вар.4	1(e_1,n_2)	1(c_1,c_2)	1(e_1,e_2)	1(c_1,n_2)	1(e_1,e_2)	1(e_1,c_2)	1(n_1,c_2)	1(e_1,c_2)	
8	Вар.5	2(n_1,e_2)	2(e_1,e_2)	2(e_1,n_2)	2(n_1,n_2)	2(e_1,n_2)	2(c_1,n_2)	2(c_1,e_2)	2(c_1,c_2)	
9	Вар.6	3(n_1,e_2)	3(e_1,e_2)	3(e_1,e_2)	3(c_1,n_2)	3(c_1,n_2)	3(e_1,n_2)	3(n_1,c_2)	3(e_1,c_2)	
10	Вар.7	1(e_1,e_2)	1(e_1,c_2)	1(e_1,e_2)	1(c_1,n_2)	1(n_1,c_2)	1(e_1,c_2)	1(e_1,n_2)	1(c_1,c_2)	
11	Вар.8	2(c_1,e_2)	2(c_1,c_2)	2(n_1,e_2)	2(e_1,e_2)	2(e_1,n_2)	2(c_1,n_2)	2(e_1,n_2)	2(n_1,n_2)	
12	Вар.9	3(n_1,c_2)	3(e_1,c_2)	3(n_1,e_2)	3(e_1,e_2)	3(c_1,n_2)	3(e_1,n_2)	3(e_1,e_2)	3(c_1,n_2)	
13	Вар.10	1(n_1,c_2)	1(e_1,c_2)	1(e_1,n_2)	1(c_1,c_2)	1(e_1,e_2)	1(e_1,c_2)	1(e_1,e_2)	1(c_1,n_2)	
14	Вар.11	2(c_1,n_2)	2(c_1,e_2)	2(c_1,c_2)	2(n_1,e_2)	2(e_1,n_2)	2(e_1,e_2)	2(e_1,n_2)	2(n_1,n_2)	
15										
16	Примечание: в желтой зоне таблицы - номера классов (коды заболеваний)									
17		скобках - соответствующие значения признаков								
18										
19										

Рис. П5.2

2. Аркуш "Вихідні данні" (рис. П5.2) таблиця результатів "об'єктивної діагностики" та відповідних значень діагностичних ознак.

3. Аркуш "Графіки" з результатами побудови графіків.

Література

1. Джарратано Дж., Райли Г. Экспертные системы: принципы разработки и программирование - М.: Вильямс. – 2007. – 1152 с.

2. Продеус А.Н., Захрабова Е.Н. Экспертные системы в медицине. – К.: «ВЕК+», 1998. – 320 с.

Лабораторна робота 6 Методи нечіткої логіки в задачах підтримки прийняття рішення

Основні цілі роботи:

1. Ознайомитися з основними ідеями методів нечіткої логіки
2. Побудувати засобами Excel модель системи прийняття нечітких рішень

Завдання:

1. Вибрати параметр (наприклад, вага, зріст, вік тощо) і задати для нього лінгвістичні змінні.
2. Побудувати таблицю можливих значень і провести їх оцінку за аналогією з прикладом, наведеним у розділі «Теоретичні відомості».
3. Побудувати графіки залежності власних функцій лінгвістичних змінних від значень числової шкали параметра.
4. Дослідити, як змінюються значення лінгвістичних змінних при зміні експертних оцінок.
5. По поєднанню параметрів побудувати кон'юнктивну або диз'юнктивну модель для оцінки властивостей об'єкта.

Зміст звіту:

1. Короткі теоретичні відомості про алгоритми обробки даних з використанням нечіткої логіки.
2. Опис розробленого прикладу.
3. Роздрукована програма і результатів дослідження.

Теоретичні відомості

За визначенням розпливчата множина \tilde{A} множини Q - це безліч пар:

$$\tilde{A} = \{ \langle \mu_A(q), q \rangle \mid q \in Q, \mu_A \in [0,1] \},$$

де $\mu_A(q)$ - функція належності розпливчастої множини \tilde{A} ; Q - базова множина або шкала.

З точки зору теорії нечітких множин, кожному параметру q , визначеному на числовій шкалі, необхідно поставити у відповідність деякий набір лінгвістичних змінних. Побудувавши залежності $Q = f(\mu_A)$, можна для кожного значення параметра q визначити значення лінгвістичних змінних на інтервалі 0-1. Далі аналіз об'єкта або ситуації по його параметрам зводиться до логічних операцій над власними функціями.

Модель прийняття рішення будується за формулою:

ЯКЩО "Ситуація", ТО "Рішення"

"Ситуація" визначається за поєднанням параметрів, об'єднаних в нечітку модель. Для побудови нечіткої моделі використовуються логічні операції над власними функціями розпливчастих множин, наприклад:

доповнення: $\neg \tilde{A} = 1 - \mu_A(q)$;

перетин: $\tilde{A} \wedge \tilde{B} = \min(\mu_A(q), \mu_B(q))$;

об'єднання: $\tilde{A} \vee \tilde{B} = \max(\mu_A(q), \mu_B(q))$;

Базова шкала виражається в одиницях виміру кількісних характеристик об'єкта, наприклад, для оцінки стану організму це можуть бути такі параметри як температура, тиск, результати аналізів. Оцінка параметрів організму виробляється експертами, в результаті чого формуються значення лінгвістичних змінних.

Розглянемо як параметр температуру, фрагмент базової кількісної шкали для якого

$$T = \{35; 36; 37; 38; 39; 40\}$$

Лінгвістичні змінні, відповідні параметру "Температура" можуть мати значення "низька" (н), "норма (базова)" (б), "висока" (в). Після оцінювання базової шкали чотирма експертами (оцінки обрані умовно) отримуємо матрицю визначення розпливчастої підмножини для параметра "температура" і відповідні власні функції:

$$\tilde{A}_T = \begin{vmatrix} м & н & н & б & б & б \\ н & н & н & н & б & б \\ м & н & н & н & н & б \\ м & н & н & н & б & б \end{vmatrix}$$

$$\mu_m = \left\{ \frac{3}{4}, 0, 0, 0, 0, 0 \right\},$$

$$\mu_n = \left\{ \frac{1}{4}, 1, 1, \frac{3}{4}, \frac{1}{4}, 0 \right\},$$

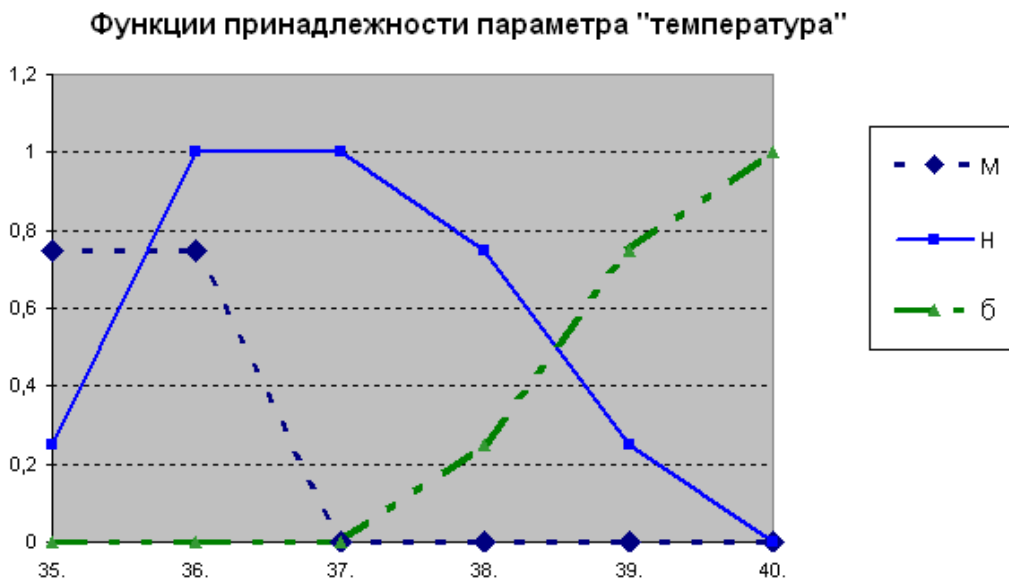
$$\mu_b = \left\{ 0, 0, 0, \frac{1}{4}, \frac{3}{4}, 1 \right\}.$$

Нижче наведено приклад побудови лінгвістичних змінних "мала" (м), "норма" (н), "велика" (б) для параметра «Температура» з використанням засобів табличного процесора Excel.

	A	B	C	D	E	F	G
1	Всього експертів:						
2	4						
3	Експертні	Шкала параметра "Температура"					
4	оцінки	35	36	37	38	39	40
5	1-й експерт	<i>м</i>	<i>н</i>	<i>н</i>	<i>б</i>	<i>б</i>	<i>б</i>
6	2-й експерт	<i>н</i>	<i>н</i>	<i>н</i>	<i>н</i>	<i>б</i>	<i>б</i>
7	3-й експерт	<i>м</i>	<i>н</i>	<i>н</i>	<i>н</i>	<i>н</i>	<i>б</i>
8	4-й експерт	<i>м</i>	<i>м</i>	<i>н</i>	<i>н</i>	<i>б</i>	<i>б</i>
9							
10							
11	Лінгвістичні змінні	35	36	37	38	39	40
12	м	0,75	0,75	0	0	0	0
13	н	0,25	0,75	1	0,75	0,25	0
14	б	0	0	0	0,25	0,75	1

36
=РАХУНОКЯКЩО(C5:C8;A12)/A2
=РАХУНОКЯКЩО (C5:C8;A13)/A2
=РАХУНОКЯКЩОС5:C8;A14)/A2

Для обчислення значень функцій належності в клітинах b12:g14 використовується функція "рахунок, якщо", віднесена до числа експертів - клітина A2. Фрагмент формул наведено для значення температури 36⁰C і графіки для отриманих значень власних функцій лінгвістичних змінних.



Побудувавши залежності $Q = f(\mu_A)$, отримаємо можливість для кожного значення параметра q визначити значення лінгвістичних змінних на інтервалі 0 - 1.

Література

1. Zadeh L.A. Fuzzy sets // Inform. Contr. -1965- vol. 8.- P. 338- 353.
2. Кофман А. Введение в теорию нечетких множеств. - М.: Радио, 1986. -350с.
3. Нечеткие множества в моделях управления и искусственного интеллекта. /Под редакцией Д.А. Поспелова. - М.: Наука, 1986.- 312 с.
4. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – Санкт-Петербург: БХВ-Петербург, 2003. – 719 с.

Лабораторна робота 7 Моделювання систем нечіткої логіки в середовищі Matlab

Основні цілі роботи:

1. Ознайомлення з інструментарієм середовища Matlab, призначеним для моделювання систем нечіткої логіки
2. Побудова засобами Matlab моделі системи прийняття нечітких рішень

Завдання:

1. Активізувати редактор систем нечіткого висновку FIS (редактор FIS) командою `fuzzy` (попередньо активізувавши середовище Matlab).
2. Активізувати демонстраційний приклад «Чайові в ресторані». Для цього в меню вікна редактора FIS вибрати пункт **File> Import> From File**, після чого провести пошук папки `toolbox\fuzzy\fuzdemos` (зазвичай знаходиться на диску C: всередині папки **Program Files\MATLAB**) і файлу з ім'ям `tipper.fis`.
3. Практично ознайомитися зі схемою системи нечіткого висновку на ім'я **tipper**.
4. Розробити власну систему нечіткого виведення медичного призначення для двох вхідних лінгвістичних змінних і однієї вихідної логічної змінної, використовуючи для цього інформацію про захворювання та супроводжуючих їм діагностичних ознаках, наявних в даній книзі.

Зміст звіту:

1. Опис розробленої системи нечіткого виведення.
2. Обґрунтування рішень, прийнятих при розробці системи нечіткого виведення.

Теоретичні відомості

Системи нечіткого виведення призначені для перетворення значень вхідних змінних в значення вихідних змінних на основі використання нечітких правил продукції.

Основними етапами нечіткого виведення є:

- формування бази правил;
- фазифікації вхідних змінних;
- агрегування підумови в нечітких правилах продукції;
- активізація або композиція висновку;
- акумулювання висновків в нечітких правилах продукції;

- дефазифікація вихідних змінних.

Фазифікацією називається процес або процедура знаходження значень функцій належності нечітких множин (термів) на основі звичайних («чітких») вихідних даних. Метою фазифікації є встановлення відповідності між конкретним (зазвичай - чисельним) значенням вхідної змінної і значенням власної функції відповідного їй терма вхідної лінгвістичної змінної.

Агрегування являє собою процедуру визначення ступеня істинності умов по кожному з правил системи нечіткого виведення.

Активізація являє собою процедуру або процес знаходження ступеня істинності кожного з висновку правил нечіткої продукції .

Акумуляція (акумулювання) являє собою процедуру або процес об'єднання всіх ступенів істинності висновків для знаходження власних функції для кожної з вихідних лінгвістичних змінних.

Дефазифікація являє собою процедуру або процес знаходження звичайного «чіткого» значення для кожної з вихідних лінгвістичних змінних.

Етапи нечіткого виведення можуть бути реалізовані різними способами. Вибір поєднання конкретних способів реалізації різних етапів означає формування конкретного алгоритму нечіткого виведення.

Одним з найбільш популярних є алгоритм Мамдані, етапи якого володіють наступними особливостями .

1 . Проводять фазифікації вхідних змінних.

2 .При агрегуванні підумови використовують парні нечіткі логічні операції. Правила, ступінь істинності яких відмінний від нуля, вважаються активними і використовуються для розрахунків.

3 .Активізацію висновків здійснюють за формулою

$$\mu'(y) = \min \{c_i, \mu(y)\} ,$$

при цьому для скорочення часу виведення враховують тільки активні правила.

4 . Акумуляцію висновків здійснюють за формулою

$$\mu_D(x) = \max \{ \mu_A(x), \mu_B(x) \}, (\forall x \in \mathbf{X}) ,$$

для об'єднання нечітких множин.

5 . Дефазифікації вихідних змінних здійснюють методом центру тяжіння або методом центру площі.

Література

1. Леоненков А. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – Санкт-Петербург: БХВ-Петербург, 2003. – 719 с.