

УДК 004.75:004.272.26:004.032.26:519.1

О.В. ГАЛКІН

Київський національний університет імені Тараса Шевченка

ПОБУДОВА ЕКСПЕРТНИХ СИСТЕМ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ

У статті розглядається представлення та організація знань в експертній системі Гомеопат, а також засоби і управління представленням знань. Описані основні етапи побудови ЕС Гомеопат, а також алгоритм функціонування системи в цілому. Розглянуті основні інтерфейсні можливості користувача.

Ключові слова: інтелектуальні системи, експертні системи, нейронні мережі.

Протягом останніх десятиліть у рамках досліджень зі штучного інтелекту сформувався самостійний напрямок – експертні системи (ЕС). Величезний інтерес до ЕС з боку користувачів викликаний наступними причинами:

- вони орієнтовані на розв’язання широкого кола завдань у неформалізованих областях, які донедавна вважалися недосяжними для обчислювальної техніки;
- за допомогою (ЕС) фахівці, які не знають програмування можуть самостійно розробляти додатки, що дає змогу розширити сферу використання ЕОМ;
- експертні системи при розв’язанні практичних завдань досягають результатів, що не поступаються, а іноді переважають можливості людей – експертів.
- Розробка ЕС зазвичай включає шість наступних етапів: ідентифікація, концептуалізація, формалізація, виконання, тестування, дослідну експлуатацію:
 - на етапі *ідентифікації* – визначаються завдання, які підлягають рішенню, виявляються цілі розробки, визначаються експерти і типи користувачів;
 - на етапі *концептуалізації* – проводиться змістовний аналіз проблемної області, виявляються використовувані поняття і їх взаємозв'язки, визначаються методи вирішення завдань;
 - на етапі *формалізації* – вибираються інформаційні системи (ІС) і визначаються способи представлення всіх видів знань, формалізуються основні поняття, визначаються способи інтерпретації знань, моделюється робота системи, оцінюється адекватність цілям системи зафіксованих понять, методів рішень, засобів вистави і маніпулювання знаннями;
 - на етапі *виконання* – здійснюється наповнення експертом бази знань (БЗ). У зв'язку з тим, що основою ЕС є знання, даний етап є найбільш важливим і найбільш трудомістким етапом розробки ЕС. Процес придбання знань розділяють на витягання знань з експерта, організацію знань, що забезпечує ефективну роботу системи, і представлення знань у вигляді, зрозумілому ЕС. [1,2]

У даній статті розглядається представлення знань та архітектура ЕС Гомеопат, яка призначена для визначення діагнозу за введеними лікарем симптомами.

Представлення знань в ЕС Гомеопат

Основою ЕС Гомеопат є нейронні мережі. Добре відомо, що нейронні мережі працюють з даними представленими у числовому вигляді [3], але з іншого боку лікар отримує інформацію про симптоми хвороби у звичайному вигляді, наприклад, пацієнт може повідомити, що в нього сильна нежить. Тому експертна система повинна формалізувати симптоми (тобто відобразити їх у числовому

вигляді), які повідомив пацієнт, а потім передати їх на вхід спеціалізованої нейронної мережі [6] (СНМ) (тобто мережі, яка була навчена на прикладах із певної галузі, наприклад з кардіології). Таким чином, при розробці ЕС Гомеопат постало питання представлення знань у числовому вигляді.

Основні питання, які треба вирішити при представленні знань, – це питання визначення структури та питання представлення знань у вибраному формалізмі. Ці дві проблеми не є незалежними. Також слід зазначити що, обраний формалізм може виявитися непридатним в принципі або неефективним для відображення деяких знань. Розглянемо простий приклад представлення знань в ЕС Гомеопат. Вектор (табл. 1), що подається на вхід СНМ може мати наступний вигляд:

Таблиця 1. Приклад вектора

Симптом1	Симптом2	Симптом3
1	0	2

У цьому випадку значення вхідного вектора свідчать про те, що у пацієнта є симптом1, відсутній симптом2, та присутній симптом3, але виражений сильніше ніж, наприклад, симптом1. Звичайно така шкала відображення симптомів у числову форму є доволі груба, про створення більш точної шкали ідеться нижче.

При проектуванні ЕС Гомеопат вирішувались наступні питання про представлення знань: визначення складу знань, що представляються; організація знань; представлення знань, тобто визначення моделі подання. Склад знань ЕС Гомеопат визначався наступними факторами: проблемним середовищем; архітектурою ЕС; потребами та цілями користувачів; мовою спілкування.

Організація знань в базі даних ЕС Гомеопат

Наступною проблемою після представлення знань, яка вирішувалась при створенні ЕС Гомеопат була проблема організації знань в БД. Відомо, що показником інтелектуальності системи, з точки зору представлення знань, вважається здатність системи використовувати в потрібний момент необхідні (релевантні) знання. Системи, що не мають засобів для визначення релевантних знань, неминуче стикаються з проблемою «комбінаторного вибуху» [2]. Таким чином, організація та зв'язність знань в ЕС є дуже важливим аспектом проектування. Зв'язність або агрегація знань є також основним способом, що забезпечує прискорення пошуку релевантних знань. У ЕС Гомеопат знання організовані навколо найбільш важливих об'єктів (сутностей) предметної області, а саме симптоми, діагноз, препарат для лікування. Якщо казати на мові реляційних БД, знання в БД ЕС Гомеопат організовані у вигляді наступних п'яти зв'язних інформаційних таблиць:

- 1) таблиця «діагноз» – описує можливі діагнози та посилання на препарати для лікування даного діагнозу;
- 2) таблиця «галузь захворювання» – подає перелік галузей до яких належать, ті чи інші захворювання;
- 3) таблиця «Пацієнт» – наводить перелік пацієнтів та встановлені для них системою діагнози;
- 4) таблиця «Препарат» – це перелік препаратів та їх опис;
- 5) таблиця «Симптоми» – це перелік можливих симптомів.

Варто також зауважити, що БД ЕС Гомеопат насправді складається з декількох баз даних. По-перше, це БД Гомеопат, яка складається з п'яти інформаційних таблиць (вони були описані вище). Крім основної БД Гомеопат в ЕС Гомеопат є службові БД, кожна з яких складається з одинадцяти таблиць які використовуються для роботи спеціалізованих нейронних мереж та комутаційного елемента (КЕ), а

також ці БД необхідні в процесі навчання відповідних НМ. Всі знання, що характеризують деяку сутність, зв'язуються і представляються у вигляді окремого об'єкта. При подібній організації знань, якщо системі потрібна інформація про деяку сутність, наприклад, ліки якими можна вилікувати хворобу пацієнта, то вона шукає об'єкт (тобто запис з таблиці Діагноз) , що описує цю сутність, а потім вже всередині об'єкта відшукує необхідну інформацію. Наприклад, якщо пацієнту було встановлено діагноз – гострий тонзиліт, тоді система в таблиці діагноз знаходить посилання на препарати, що описані в таблиці Препарат. У об'єктах ЕС Гомеопат виділяють два типи зв'язків між елементами: зовнішні і внутрішні. Внутрішні зв'язки об'єднують елементи в єдиний об'єкт і призначені для відображення структури об'єкта. Зовнішні зв'язки відображають взаємозалежності, що існують між об'єктами в області експертизи. При роботі з великою базою знань ЕС Гомеопат постала також проблема пошуку знань, релевантних розв'язуваній задачі. У зв'язку з тим, у даних, що оброблюються може не міститися явних вказівок на необхідні значення. Тому загальна задача пошуку релевантного діагнозу, в ЕС Гомеопат, розбивалась на більш прості задачі. По-перше, вирішувалось до якої галузі належить захворювання пацієнта, яке описується введеними симптомами (цю задачу вирішував так званий КЕ). По-друге, після встановлення галузі захворювання, симптоми передавались на вхід спеціалізованій НМ, яка і встановлювала відповідний діагноз. Для того щоб ЕС Гомеопат могла керувати процесом пошуку рішення, була здатна набувати нові знання і пояснювати свої дії, вона повинна вміти не тільки використовувати свої знання, а й мати здатність розуміти і досліджувати їх, тобто ЕС повинна мати знання про те, як представлені її знання про проблемне середовище [4]. Якщо знання про проблемне середовище назвати знаннями нульового рівня подання, то перший рівень уявлення містить метазнання, тобто знання про те, як представлені у внутрішньому світі системи знання нульового рівня. Перший рівень містить знання про те, які засоби використовуються для представлення знань нульового рівня. Знання першого рівня грають істотну роль при управлінні процесом рішення, при придбанні та поясненні дій системи. У зв'язку з тим, що знання першого рівня не містять посилань на знання нульового рівня, знання першого рівня незалежні від проблемного середовища. Кількість рівнів подання може бути більше двох. Другий рівень представлення містить відомості про знання першого рівня, тобто знання про подання базових понять першого рівня. Поділ знань за рівнями подання забезпечує розширення області застосування системи.

Виділення рівнів детальності дозволяє розглядати знання з різним ступенем деталізації. Кількість рівнів детальності визначається специфікою завдань, що розв'язуються, обсягом і способом їх подання. Як правило, виділяється не менше трьох рівнів детальності, що відбивають відповідно загальну, логічну і фізичну організацію знань. Введення декількох рівнів детальності забезпечує додатковий ступінь гнучкості системи, тому що дозволяє робити зміни на одному рівні, не зачіпаючи інші. Зміни на одному рівні деталізації можуть призводити до додаткових змін на цьому ж рівні, що виявляється необхідним для забезпечення узгодженості структур даних і програм. Проте наявність різних рівнів перешкоджає поширенню змін з одного рівня на інший.

Таким чином, організація БД ЕС Гомеопат, яка була описана вище та вимоги до керування процесом пошуку рішення привела до того, що в ЕС Гомеопат був використаний КЕ. Більш того, використання КЕ дало можливість розбити одну велику задачу на певну кількість більш простих задач. Інакше для рішення задачі встановлення діагнозу потрібно використовувати доволі довгий процес

навчання НМ методом оберненого розповсюдження похибки [3]. Але добре відомо, що цей метод має досить суттєві вади, а саме [3]:

- «Параліч» мережі. В процесі навчання, ваги можуть стати дуже великими. Це може призвести до того, що більшість нейронів будуть працювати при дуже великих вихідних значеннях, в зоні, де похідна стискувальної функції дуже мала. Як результат процес навчання може припинитися.
- Локальні мінімуми. Мережа може потрапити в локальний мінімум, в той час коли поряд є більш глибокий мінімум. В точці локального мінімуму усі напрямки ведуть в гору, тому мережа не може з нього вийти.
- Розмір шага навчання. Очевидно, що розмір шага повинен бути скінченним, але якщо його вибрати досить великим, то може виникнути «параліч», або нестійкість. Тобто існує проблема оптимального розміру шага навчання.

З огляду на вище описані проблеми, які виникають при використанні довгого процесу навчання, розбиття початкової задачі на декілька більш простих задач, за допомогою КЕ виправдано.

Зауважимо, що для доступу до БД ЕС Гомеопат та пошуку в ній була використана технологія ADO.NET. Технологія ADO.NET – це частина Microsoft .NET framework. ADO.NET відрізняється від інших подібних технологій доступу до даних тим, що дозволяє взаємодіяти з БД автономно, тобто за допомогою кешу даних. Така автономна взаємодія з БД дуже важлива при роботі з НМ, особливо при навчанні НМ (тобто при роботі ЕС Гомеопат в режимі експерта, в якому відбувається навчання НМ). Важливим елементом автономного доступу до даних є контейнер для табличних даних. Такий автономний контейнер в технології ADO.NET реалізований на основі класу DataSet. Цей клас був використаний в ЕС Гомеопат для доступу до даних.

Інтерфейс з кінцевим користувачем

ЕС Гомеопат може функціонувати в двох режимах: в режимі навчання (придбання знань) і в робочому режимі (діагностування). У першому режимі експерт предметної області засобами діалогових вікон вводить в БЗ опис конкретних навчаючих вибірок, а також іншу службову інформаційну складову у термінах мови представлення знань ЕС Гомеопат.

Після завантаження системи лікар-користувач має можливість обрати область захворювання. Ця область визначається лікарем при зверненні пацієнта з відповідним до цієї області захворюванням. Також це діалогове вікно дозволяє лікарю-користувачу за вимогою обрати один з режимів роботи з ЕС Гомеопат, а саме режим «Експерта» та режим «Лікаря».

Режим «Експерта» (рис. 1) надає можливість роботи з БД «Діагноз», «Симптоми», «Перепарат». Як зазначено вище ці таблиці беруть участь у навчанні спеціалізованої НМ та комутаційного елемента.

Номер	Номер діагнозу	Симптом	вектор симптома 1	вектор симптома 2	вектор симптома 3	вектор симптома 4
5	1	головна біль	1	0	1	1
6	1	фронтальної очей	0	1	1	0
7	1	нервове в горлі	0	1	0	0
8	1	сухобрихат...	0	1	0	0
9	1	притупленість...	0	1	2	0
10	1	двоїтвене лям.	0	1	2	2
11	1	испанська крові	0	0	1	0
12	1	випередки	0	0	1	0
13	1	неповноцінності	0	0	1	0
14	1	нервовість	0	0	0	1
15	1	супернативності	0	0	0	1

Рис. 1. Режим «Експерта»

Кнопка «Обучение специализированной нейронной сети» (рис.2) використовується для навчання спеціалізованих нейронних мереж, приклади для навчання беруться з відповідних службових баз даних(кожної спеціалізованої нейронної мережі відповідає службова база даних). Перед запуском процесу навчання експерт має задати кількість ітерацій. Навчання закінчується, або якщо було досягнуто необхідної точності, або якщо було проведено задану кількість ітерацій. Є також можливість «скинути» стан мережі. Для цього призначена кнопка «Сделать сеть не обученной».

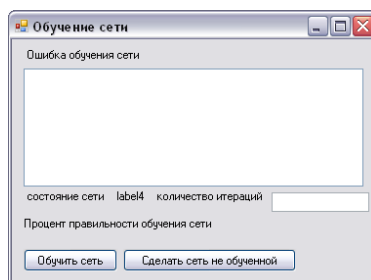


Рис. 2. Навчання мережі

Кнопка «Обучение терапевта» (рис.3а) призначена для навчання нейронної мережі комутаційного елемента. Як вже було зазначено вище комутаційний елемент визначає до якої галузи належать симптоми. В даному випадку експерт має можливість сформувати приклади для навчання. Для цього існує кнопка «сформировать таблицу выборки». Також експерт має вести кількість ітерацій. Як і у випадку з спеціалізованими нейронними мережами, експерт має можливість відмінити навчання нейронної мережі.

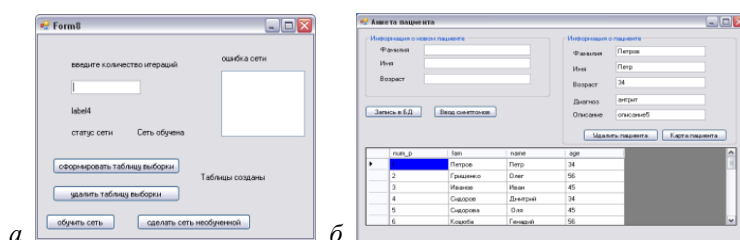


Рис. 3. а– навчання «терапевта» (комутаційного елемента), б– режим «Лікар»

У режимі «Лікаря» (рис. 3 б) відбувається безпосередня діагностика пацієнта, а саме: введення інформації про нового пацієнта, перегляд карти пацієнта (наявна якщо пацієнт раніше користувався ЕС Гомеопат), видалення пацієнта з БД ЕС Гомеопат, введення симптомів пацієнтів для можливості діагностування. Після введення інформації про пацієнта лікар-експерт співставляє у відповідність зі скаргам пацієнта симптоми БД ЕС Гомеопат та натисканням відповідної кнопки встановлює попередній діагноз пацієнту (рис. 5).

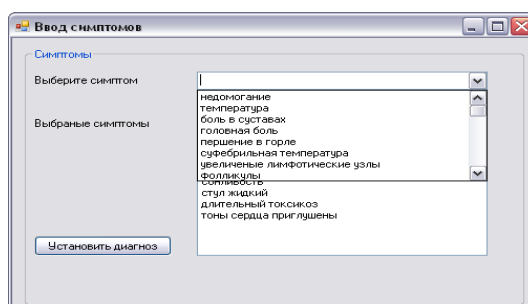


Рис. 5. Введення симптомів

При встановленні таким чином діагнозу виконує свої функції відповідна спеціалізована нейронна мережа. Діагноз є попереднім оскільки на наступному етапі лікар-експерт може скорегувати результат, оскільки, як відомо, при навчанні НМ існує похибка навчання і відповідно є ймовірність того, що діагноз може бути не коректно встановлений. Якщо таке трапилось, на цьому етапі лікар-експерт, змінює вихідні дані і система «перенавчається» з урахуванням зауважень лікаря-експерта.

Висновки

В статті було продемонстровано можливість розділити задачу встановлення діагнозу та вибір засобів лікування на декілька більш простих задач. Тобто спочатку на основі введених лікарем симптомів та за допомогою комутаційного елемента встановлюється галузь до якої належать дані симптоми, а потім вже за допомогою спеціалізованої нейронної мережі встановлюється діагноз пацієнта та визначаються гомеопатичні лікарські засоби для лікування відповідного захворювання. При розробці системи Гомеопат для роботи з базами даних та реалізації нечіткої специфікації логічного виведення була використана технологія ADO.NET. Використання цієї технології обумовлено тим, що ця технологія має декілька переваг, а саме: – ADO.NET дозволяє працювати з різними джерелами даних; наявність автономних об'єктів дозволяє істотно підвищити ефективність та знизити навантаження на СУБД; інтеграція з XML. Що дозволяє досить легко масштабувати ЕС Гомеопат.

Список використаної літератури

1. Элти Дж., Кумбс М. Экспертные системы: концепции и примеры. М.: Финансы и статистика, 1987. – 191 с.
2. Уинстон П. Искусственный интеллект. – М.: Мир, 1980. – 519с.
3. Рутковская Д., Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. – М: Горячая линия – Телеком, 2006. – 452с.
4. Андон Ф.И., Балабанова А.С. Выявление знаний и изыскания в базах данных: подходы, модели, методы и системы //Проблемы программирования. –2000. – N1-2. – С.513–526.
5. Круглов В. В., Дли М. И., Голунов Р. Ю. Нечеткая логика и искусственные нейронные сети: Учеб. пособие. – М.: Издательство физико-математической литературы, 2001. – 224 с.
6. Jacek Leski. Systemy neuronowo-rozmyte. – Warszawa: Wydawnictwo Naukowo-Techniczne, 2008. – 686 s.

Стаття надійшла до редакції 01.11.2012

Построение экспертных систем с использованием нейронных сетей

Галкин О.В.

Киевский национальный университет имени Тараса Шевченко

В статье рассматривается представление и организация знаний в ЕС Гомеопат, а также средства и управление представлением знаний. Описаны основные этапы построения ЕС Гомеопат, а также алгоритм функционирования системы, в целом. Рассмотрены основные интерфейсные возможности пользователя.

Ключевые слова: интеллектуальные системы, экспертные системы, нейронные сети.

Build expert system using neural networks

O. Galkin

Kyiv National Taras Shevchenko University

In the present paper we show the representation and the organization of the knowledge in N-Homeopat. We describe the main stages of construction of ES the Homeopath, and also algorithm of functioning of system, as a whole. The main features of the user interface are also considered.

Keywords: intellectual systems, expert systems, neural networks.