

## Podstawowe wiadomości o systemach ekspertowych

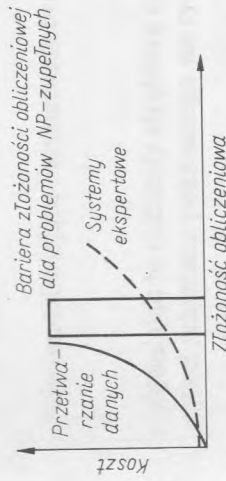
### 2.1. Podstawowe koncepcje systemów ekspertowych

System ekspertowy jest programem komputerowym, który wykonuje złożone zadania o dużych wymaganiach intelektualnych i robi to tak dobrze jak człowiek, będący ekspertem w tej dziedzinie. Określenie „system ekspertowy” może być zastosowane do dowolnego programu komputerowego, który na podstawie szczególnej wiedzy może wyciągać wnioski i podejmować decyzje, działając w sposób zbliżony do procesu rozumowania człowieka. W niniejszej książce terminu „system ekspertowy” używamy w szerszym sensie. W ten sposób określamy system, który niekoniecznie zastępuje eksperta – człowieka. W wielu bowiem sytuacjach, np. podczas podejmowania decyzji w siłowniach energetycznych, podczas sterowania skomplikowanych urządzeń i maszyn, na pokładzie statków kosmicznych itd., człowiek nie mógłby swymi zmysłami ogarnąć całej sytuacji. Są więc potrzebne systemy, które pracują w czasie rzeczywistym i wykonują swoje funkcje lepiej niż człowiek. Z tego względu w dalszej części przez system ekspertowy rozumiemy program komputerowy przeznaczony do rozwiązywania specjalistycznych problemów, które wymagają profesjonalnej ekspertyzy.

Systemy ekspertowe można podzielić na trzy ogólne kategorie: *doradcze* (advisory), podejmujące decyzje bez kontroli człowieka (dictatorial), *krytykujące* (criticizing). Systemy doradcze prezentują rozwiązania dla użytkownika, który jest w stanie ocenić ich jakość. Użytkownik może odrzucić rozwiązanie oferowane przez system i zająć inne rozwiązanie. Systemy podejmujące decyzje bez kontroli człowieka są same dla siebie końcowym autorytetem. Są używane np. do sterowania różnymi obiektami, gdzie udział człowieka jest utrudniony lub wręcz niemożliwy. Z kolei dla systemów krytykujących jest przedstawiany problem jakież jego rozwiązanie. System dokonuje w tym przypadku analizy i komentuje uzyskane rozwiązania.

Systemy ekspertowe stosuje się w wielu przypadkach w takich dziedzinach, gdzie informacja (wiedza) o danej dziedzinie jest niepewna. To samo dotyczy

dziejzin, które nie są w sposób jednoznaczny sformalizowane (nie istnieje model matematyczny algorytmów rozwiązyjących zadania w danej dziedzinie) [11, 35, 59]. Inną dziedziną jest rozwiązywanie problemów NP-zupełnych, co zilustrowano na rys. 2.1. Dla problemów tych mogą być opracowane teorie, lecz jest niemożliwością ogarnąć za pomocą algorytmów wszystkich teoretycznie możliwych przypadków w rozsądnym czasie. Dla problemów NP-zupełnych złożoność obliczeniowa (czyli liczba kroków koniecznych do ich rozwiązania) jest funkcją wykładniczą ich rozmiarów. Innymi słowy, przy liniowym wzroście wymiaru problemu czas potrzebny na jego rozwiązanie wzrasta wykładniczo. Zjawisko takie jest określane terminem *eksplozji kombinatorycznej*.



Rys. 2.1. Przetwarzanie wiedzy umożliwia pokonanie bariery złożoności obliczeniowej problemu

Utworzenie systemu opartego na bazie wiedzy wymaga pozyskania (akwizycji) wiedzy eksperta, który często znajduje rozwiązanie na podstawie informacji o problemie oraz na swoim doświadczeniu. Ponieważ pozyskanie i strukturalizacja wiedzy eksperta jest zajęciem pracochłonnym, więc wysiłek włożony w utworzenie systemu ekspertowego jest uzasadniony wówczas, gdy będzie on wykorzystywany w długim odstępie czasu przez wystarczająco dużą liczbę użytkowników [31, 93, 99, 111].

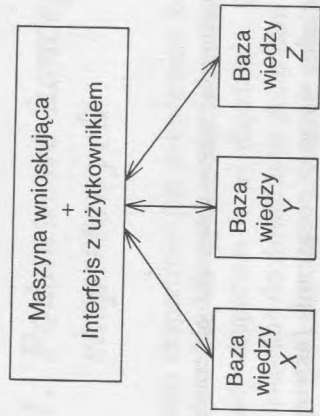
System ekspertowy, dysponując zapisaną wiedzą eksperta z wybranej dziedziny, może jej używać wielokrotnie w sposób ekonomicznie efektywny, gdyż nie jest przy tym wymagana obecność eksperta. Umożliwia to ekspertowi jednoczesne uwolnienie się od powtarzania analogicznych ekspertyz i zajęcie się bardziej twórczymi zadaniami. Szczególną zaletą takich systemów jest możliwość rozwiązywania określonych zadań bez bezpośredniego udziału eksperta, jak również możliwość agregacji wiedzy w jednym systemie licznego zespołu ekspertów.

Pod koniec lat siedemdziesiątych dokonano istotnego spostrzeżenia. Okazało się, że moc programu ekspertowego (w zakresie rozwiązywania danego problemu) tkwi w zakodowanej w nim wiedzy, a nie w formalizmie i schematach wnioskowania, których ten program używa. Możemy to wyrazić w postaci stwierdzenia „more knowledge less search”, co oznacza: im pełniejsza jest wiedza, tym szybciej uzyskuje się rozwiązanie.

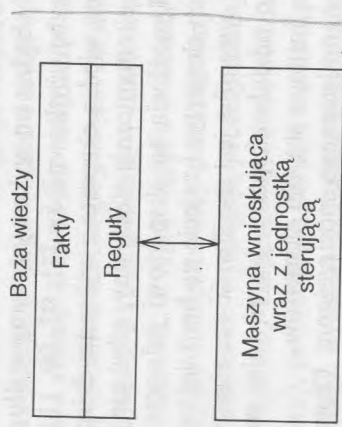
Tak więc problem z posiadaniem pełnej wiedzy tkwi w bazie wiedzy, a nie w sposobie realizacji procesu wnioskowania systemu ekspertowego. Oznacza to, że aby zbudować inteligentny program, należy go wyposażać w dużą ilość dobrej jakości, specyficznej wiedzy o danym przedmiocie.

Proces konstruowania systemów ekspertowych należy do zagadnień tzw. *inżynierii wiedzy* (knowledge engineering). Zasadniczym celem inżynierii wiedzy jest pozyskiwanie wiedzy, jej strukturalizacja, przetwarzanie. Zajmuje się ona także rozwijaniem metodologii i narzędzi budowy systemów ekspertowych. Obejmuje ona w szczególności tworzenie metod programowania dla:

- a) pozyskiwania i strukturalizacji wiedzy ekspertów;
- b) dopasowania i wyboru odpowiednich metod wnioskowania i wyjaśniania rozwiązywanych problemów;
- c) projektowania odpowiednich interfejsów (układów pośredniczących) między komputerem a użytkownikiem.



Rys. 2.2. Ogólna struktura systemu ekspertowego



Rys. 2.3. Podstawowe bloki systemu ekspertowego

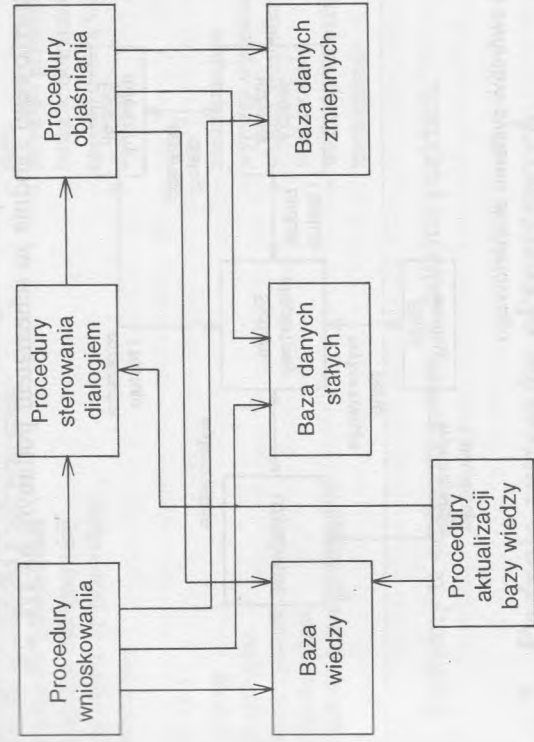
Systemy ekspertowe jako programy są w zasadzie zorganizowane w ten sposób, że wiedza dotycząca danej dziedziny, zwana *bazą wiedzy* (knowledge base), jest odseparowana od reszty systemu, zawierającej m.in. sterowanie wiedzą, do czego stosuje się odpowiednie mechanizmy wnioskowania. Ta część systemu jest często nazywana *maszyną wnioskującą* (inference engine). Tak więc system ekspertowy ma dwa główne bloki, przedstawione na rys. 2.2 oraz 2.3 [2, 12].

### 2.1.1. Struktura systemów ekspertowych

Rozważając bardziej szczegółowo strukturę systemów ekspertowych możemy w niej wyróżnić następujące podstawowe elementy:

- baza wiedzy (np. zbiór reguł);
- baza danych (np. dane o obiekcie, wyniki pomiarów, hipotezy);
- procedury wnioskowania – maszyna wnioskująca;
- procedury objaśniania – objaśniają strategię wnioskowania;
- procedury sterowania dialogiem – procedury wejścia/wyjścia umożliwiające formułowanie zadań przez użytkownika i przekazywanie rozwiązań przez program;
- procedury umożliwiający rozszerzanie oraz modyfikację wiedzy – pozyskiwanie wiedzy.

Uwzględniając te elementy strukturę systemu ekspertowego można przedstawić w postaci schematu pokazanego na rys. 2.4.

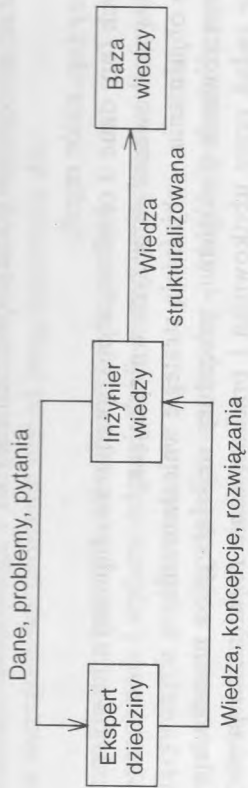


Rys. 2.4. Główne elementy systemu ekspertowego

Współpraca użytkownika z systemem jest nazywana *konsultacją* albo *sesją*. Odbywa się przez interfejs z użytkownikiem. Wąskim gardłem interfejsu jest klawiatura. Przewiduje się, że w przyszłości komunikacja z systemem będzie odbywać się za pomocą głosu.

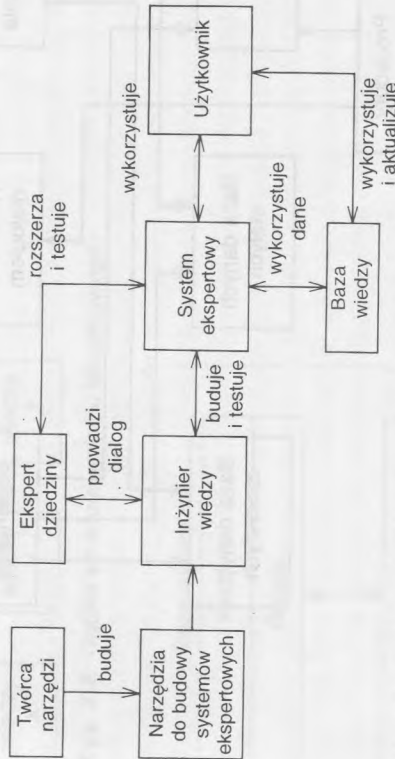
Systemy z bazami wiedzy wydzielnymi od pozostałych modułów programu nazywamy *systemami opartymi na bazie wiedzy* (knowledge based systems). Systemy takie często są wykorzystywane jako systemy wspomagające podejmowanie decyzji (decision support systems). System ekspertowy podczas swego działania

korzysta z bazy wiedzy. Należy jednak zaznaczyć, że nie zawsze system mający odseparowaną bazę jest systemem ekspertowym. Na przykład programy niektórych gier nie są zaliczane do systemów ekspertowych.



Rys. 2.5. Typowy proces pozyskiwania wiedzy

Wiedza z danej dziedziny w postaci strukturalizowanej i sformalizowanej może być wprowadzona do bazy wiedzy systemu ekspertowego różnymi sposobami, na przykład przez *inżyniera wiedzy* (knowledge engineer). Jedną z metod pozyskiwania wiedzy jest prowadzenie dialogu z ekspertem, co ilustruje rys. 2.5. Systemy ekspertowe mogą być budowane z różnych narzędzi. Prowadzi to do specjalizacji twórców systemu ekspertowego zgodnie ze schematem podanym na rys. 2.6.



Rys. 2.6. Role twórców systemu ekspertowego

W tablicy 2.1 dokonano porównania konwencjonalnego przetwarzania danych z inżynierią wiedzy, a w tabl. 2.2 zestawiono wady i zalety ekspertyz naturalnej i sztucznej.

Podstawowe zagadnienia w systemach opartych na bazie wiedzy są następujące:

- reprezentacja wiedzy,
- pozyskiwanie wiedzy,
- sposób użycia wiedzy,
- objaśnianie i uczenie się.

Tablica 2.1. Porównanie konwencjonalnego przetwarzania z inżynierią wiedzy

Konwencjonalne przetwarzanie danych	Inżynieria wiedzy
Programista analityk systemów	Inżynier wiedzy
Program	System ekspertowy
Baza danych	Baza wiedzy
Reprezentacja i użycie danych	Reprezentacja i użycie wiedzy
Algorytmy	Heurystyki
Efektywna manipulacja dużymi bazami danych	Efektywna manipulacja dużymi bazami wiedzy

Tablica 2.2. Porównanie ekspertyz naturalnej z ekspertyzą sztuczną

Ekspertyza naturalna, wykonana przez człowieka	Ekspertyza sztuczna
<b>W a d y:</b>	<b>Z a l e t y:</b>
- tracąca na wartości z upływem czasu	- stała
- trudna do przeniesienia	- łatwa do przeniesienia
- trudna w dokumentacji	- łatwa w dokumentacji
- nie dająca się przewidzieć	- zgodna z bazą wiedzy
- kosztowna	- dostępna
<b>Z a l e t y:</b>	<b>W a d y:</b>
- twórcza	- nie inspirująca
- adaptacyjna	- wymaga wprowadzenia wiedzy
- wykorzystanie zmysłów	- wejścia symboliczne
- szeroki zakres	- wąski zakres
- wiedza zdrowego rozsądku	- wiedza przetwarzana w sposób mechaniczny

Problemy te omówimy szerzej w dalszych punktach.

## 2.1.2. Rodzaje systemów ekspertowych

Aby ocenić system ekspertowy, należy najpierw odpowiedzieć na pytania, co otrzymujemy na jego wyjściu. Otóż wyniki te możemy podzielić na trzy zasadnicze grupy: diagnozy, prognozy oraz plany. *Diagnoza* jest to ocena stanu istniejącego na podstawie posiadanych danych. Do tej grupy zaliczamy programy: MYCIN [113], CADUCEUS [98], EDORA [6], DENDRAL [74], PROSPECTOR [42], HEARSAY [107], PUFF [4] i CASNET [7] (diagnoza w potocznym rozumieniu tego słowa czy rozpoznawanie zmian stanu (monitoring)). Zagadnienie to jest blisko związane z rozpoznawaniem wzorców, choć zwykle programów takiego rozpoznawania nie zalicza się do systemów ekspertowych.

Prognoza jest to przewidywanie stanu przyszłego na podstawie istniejących danych. Niektóre programy medyczne, np. CASNET [7], stawiają prognozę stanu (w tym przypadku zdrowia pacjenta) na podstawie własnej diagnozy.

Plan wreszcie jest rozumiany jako opis pewnego stanu, do którego należy dążyć. Jednym z przykładów jest planowanie konfiguracji komputerów systemu VAX wykonane przez system R1 [80], znany także jako XCON, XSEL (projektowanie). Innym przykładem jest planowanie działań robota przemysłowego w zmieniającym się otoczeniu. Jeszcze innym przykładem może być system DELTA/CATS [28], którego przeznaczeniem jest utrzymanie zdolności eksploatacyjnej lokomotyw.

Łączenie ze sobą tych trzech typów pracy systemu ekspertowego jest oczywiście możliwe. Dzięki połączeniu rozpoznawania wzorca, diagnozy i planowania uzyskujemy systemy sterujące, np. system sterowania statkiem kosmicznym. Inną grupę stanowią systemy diagnozowania i planowania na przykład systemów uczących. Pod kątem tak określonych zadań należy zatem porównywać i oceniać różne systemy ekspertowe. W tablicy 2.3 zestawiono różne rodzaje systemów ekspertowych w zależności od realizowanych przez te systemy zadań. Z kolei w tablicy 2.4 przedstawiono podstawowe obszary zastosowań systemów ekspertowych.

Tablica 2.3. Rodzaje systemów ekspertowych

Kategoria	Zadania realizowane przez systemy ekspertowe
Interpretacyjne	dedukują opisy sytuacji z obserwacji lub stanu czujników, np. rozpoznawanie mowy, obrazów, struktur danych
Przyjęte	wnioskuje o przyszłości na podstawie danej sytuacji, np. prognoza pogody, rozwój choroby
Diagnostyczne	określają wady systemu na podstawie obserwacji, np. medycyna, elektronika, mechanika
Kompletowania	konfigurują obiekty w warunkach ograniczeń, np. konfigurowanie systemu komputerowego
Planowania	podjęmują działania, aby osiągnąć cel, np. ruchy robota
Monitorowania	porównują obserwacje z ograniczeniami, np. w elektrowniach atomowych, medycynie, w ruchu ulicznym
Sterowania	kierują zachowaniem systemu; obejmują interpretowanie, predykcję, naprawę i monitorowanie zachowania się obiektu
Poprawiania	podają sposób postępowania w przypadku złego funkcjonowania obiektu, którego te systemy dotyczą
Naprawy	harmonogramują czynności przy dokonywaniu napraw uszkodzonych obiektów
Instruowania	systemy doskonalenia zawodowego dla studentów

Tablica 2.4. Podstawowe obszary zastosowań systemów ekspertowych

Wektor zastosowania	Bankowość i ubezpieczenia	Przemysł	Handel i usługi	Sektor publiczny i inne
Monitorowanie i sterowanie	obserwowanie trendów	nadzorowanie procesów, sterowanie procesami, raportowanie specjalnych sytuacji	obserwowanie trendów	monitorowanie reaktorów jądrowych oraz dużych sieci (gazowe, wodne)
Projektowanie		projektowanie zakładów i produktów, komputerów	wyбір asortymentów, doradztwo dla rolnictwa	sieci (pocztowe, energetyczne)
Diagnostyka	kredyty, pożyczki na nieruchomości, analiza ryzyka, przetwarzanie skarg	wykrywanie uszkożeń, utrzymywanie zdolności produkcyjnej	kredyty, analiza ryzyka	diagnoza medyczna, diagnoza techniczna
Planowanie	analiza ryzyka, planowanie inwestycji	projektowanie funkcji logicznych, planowanie projektu	analiza ryzyka, analiza rynku	planowanie inwestycji, planowanie na wypadek klęski, planowanie dystrybucji

Ze względu na sposoby realizacji systemy ekspertowe możemy podzielić na dwie grupy: systemy dedykowane oraz systemy szkieletowe. Systemy dedykowane są tworzone od podstaw przez inżyniera wiedzy współpracującego z informatykiem. Systemy szkieletowe (shells) są to systemy z pustą bazą wiedzy. Proces tworzenia finalnego systemu ekspertowego jest w tym przypadku krótszy niż w pierwszym przypadku, gdyż jest wymagane tylko pozyskanie wiedzy i jej odowiednia implementacja w systemie. W Dodatku 1 przedstawiono podstawowe właściwości niektórych znanych dedykowanych systemów ekspertowych.

Ze względu na metodę prowadzenia procesu wnioskowania systemy ekspertowe dzieli się na:

- z logiką dwuwartościową (Boole'a),
- z logiką wielowartościową,
- z logiką rozmytą.

Ze względu na rodzaj przetwarzanej informacji systemy ekspertowe dzieli się na dwie grupy: systemy z wiedzą pewną, czyli zdefiniowaną, oraz systemy z wiedzą niepewną, w przetwarzaniu której wykorzystuje się przede wszystkim aparat probabilistyczny.

Charakterystyczną cechą większości znanych systemów ekspertowych jest przedstawienie informacji w postaci symbolicznej. Umożliwia to wykorzystanie w systemie wszelkich istniejących układów przetwarzania symboli. Możemy przyjąć, że właśnie posługiwanie się symbolami wyróżnia klasyczne systemy ekspertowe spośród innych programów. Istnieją jednak systemy ekspertowe oparte na niesymbolicznych metodach przetwarzania informacji, np. sieci neuronowe, algorytmy genetyczne itp. Wśród technik opartych na symbolicznym przetwarzaniu informacji kryterium podziału systemów ekspertowych jest dokonywane ze względu na rodzaj reprezentacji wiedzy. Możemy tu wyróżnić takie metody reprezentacji wiedzy jak: reguły, ramy, sieci semantyczne, predykaty itp. (o czym mówimy w rozdz. 3).

## 2.2. Właściwości systemów ekspertowych

### 2.2.1. Poprawność systemu

System ekspertowy powinien zapewnić wysoki poziom wydawanych ekspertyz. W tym sensie możemy mówić o poprawności systemu, jeśli daje on dobre rezultaty, rozwiązuje zadania w czasie dopuszczalnym i dysponuje strategiami umożliwiającymi imitowanie wiedzy i intuicji eksperta, uzyskanej w wyniku wieloletniego doświadczenia.

Jakość pracy możemy ocenić porównując wyniki działania systemu z rezultatami pracy człowieka. Porównanie takie przeprowadzono dla systemu DENDRAL. Okazało się, że system ten lepiej określa konfigurację cząsteczki – to znaczy jest mniej pomyłek w długiej serii testów – niż u znakomitej większości wykonujących to samo zadanie chemików. W punktach dystrybucji do konfigurowania zestawów komputerów systemu VAX używa się systemu ekspertowego RI. Okazał się on wyraźnie lepszy od wykonujących wcześniej tę pracę ludzi.

Dopuszczalny czas rozwiązania zadania jest pojęciem relatywnym względem samego zadania. Niektóre systemy odznaczają się czasem odpowiedzi mierzonym w sekundach, inne zaś mogą pracować w ciągu długiego czasu (tygodniami – chociażby PROSPECTOR). Nie zawsze jednak wydłużanie czasu odpowiedzi poprawia w istotny sposób wyniki działania. Dla systemów ekspertowych wykorzystujących technikę przeszukiwania przestrzeni rozwiązań w postaci struktur drzewiastych czas rozwiązania wzrasta wykładniczo. Jeśli zatem czas w jakikolwiek sposób ogranicza możliwość oczekiwania na odpowiedź, to istnieje konieczność kompromisu między wymogami jakości i szybkości działania. Ocena tych dwu czynników łącznie jest nader skomplikowana.

Ekspert-człowiek często opiera się na intuicji i nie potrafi dobrze uzasadnić swoich rozstrzygnięć, choć są one poprawne (a przynajmniej dopuszczalne). W tej sytuacji stworzenie odpowiednich reguł wnioskowania jest bardzo skomplikowane.

### 2.2.2. Uniwersalność

Ważną cechą charakterystyczną systemu ekspertowego jest jego zdolność do rozwiązywania obszernej klasy zadań z danej dziedziny. Aby zdolność tę przejawiać, system nie powinien zawierać wielu sztywnych, wcześniej przygotowanych rozwiązań, lecz dużą liczbę reguł obejmujących dostatecznie szeroki zakres heurystyk z dziedziny problemowej. Powinny one nie tylko umożliwić przeszukiwanie przestrzeni rozwiązań dobrze postawionego zadania, lecz także coś w rodzaju eks-trapolacji w przypadku zadań będących na granicy wiedzy znanej systemowi. Dostateczna liczba uniwersalnych reguł powinna zapewnić płynne (nie skokowe) pogarszanie się jakości pracy systemu w sytuacji, w której posiadane reguły wnioskowania okazują się niewystarczające. Wspomniana tu uniwersalność jest jednak rozumiana wewnątrz dziedziny wiedzy systemu.

Uniwersalność rozumiana jako możliwość rozwiązania zadań z różnych dziedzin wiedzy na podstawie strukturalnego podobieństwa reguł wnioskowania jest jeszcze nieosiągalna. Tworzenie systemów zdolnych do takiego działania – można je nazywać *metasystemami* – jest jednak nieuniknionym kierunkiem rozwoju badań nad sztuczną inteligencją.

### 2.2.3. Złożoność

Stopień komplikacji systemu ekspertowego jest w naturalny sposób określony przez dziedzinę, dla której jest wykonany. Występuje tu paradoksalne zjawisko, że problem niezbyt skomplikowany, o krótkim i mało rozgałęzionym drzewie przeszukiwań czyni konstrukcję systemu nieoptymalną, ponieważ problem taki łatwiej jest rozwiązać używając zwykłych programów. Podobnie rzecz się ma w przypadku istnienia ścisłych algorytmów rozwiązania. Tak więc nie nazwiemy systemem ekspertowym ani programu wyboru jednego z czterech możliwych przedmiotów, ani algorytmu szybkiego przekształcenia Fouriera.

Niekiedy sytuacja jest inna. Istnieje złożony system ekspertowy w dziedzinie, która bardzo trudno poddaje się opisowi, człowiek natomiast radzi sobie z nią dobrze. Przykładem może być tutaj sterowanie procesem technologicznym o nieznanym modelu matematycznym. Wtedy system może służyć jako pomoc osobom mało zaznajomionym z daną dziedziną, jeśli starannie dobierze się sposób uzasadniania podejmowanych decyzji.

Ogólnie ocena złożoności systemu jest w pewien sposób możliwa – na przykład przez liczbę reguł wnioskowania, wielkość bazy danych itp. Systemy ekspertowe są klasyfikowane ze względu na liczbę reguł na trzy grupy:

małe	(100 – 300 reguł),
średnie	(300 – 2000 reguł),
duże	(ponad 2000 reguł).

Ze względu na trudności w przeszukiwaniu dużych baz wiedzy w procesie wnioskowania nie stosuje się obecnie więcej niż 10 000 reguł.