



# AKADEMIA OBRONY NARODOWEJ

## CENTRUM INFORMATYKI

### SYSTEMY EKSPERTOWE WSPOMAGANIA DECYZJI

Plk dr hab.inż. Józef ZAPIÓR

## WPROWADZENIE DO INŻYNIERII WIEDZY I SYSTEMÓW EKSPERTOWYCH

1-2

Studium teoretyczne  
pod kryptonimem E

~~Biblioteka Główna  
Akademii Obrony Narodowej~~

~~S/1124~~

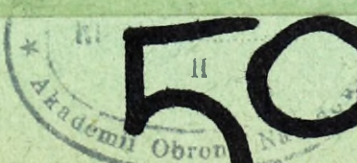


~~05-001924-001-0~~

Warszawa

1992

59776





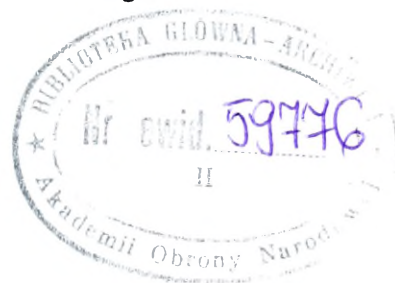
# AKADEMIA OBRONY NARODOWEJ

---

## CENTRUM INFORMATYKI

### SYSTEMY EKSPERTOWE WSPOMAGANIA DECYZJI

Plk dr hab.inż. Józef ZAPIÓR



## WPROWADZENIE DO INŻYNIERII WIEDZY I SYSTEMÓW EKSPERTOWYCH

1-2

Studium teoretyczne  
pod kryptonimem EKSPERT-1



---

Warszawa

1992

Redakcja techniczna: J. Zapiór

1

Edytor tekstu: WordPerfect 5.1  
Drukarka: Hewlett Packard LaserJet III  
Druk ukończono: 22-12-1992 r.

Objętość 5.2 ark. aut.

~~ASO/te~~

## SPIS TREŚCI

Przedmowa . . . . .	5
<b>1. WSTĘP . . . . .</b>	<b>7</b>
1.1. Co to jest przetwarzanie wiedzy ? . . . . .	7
1.2. W poszukiwaniu nowych systemów przetwarzania informacji . . . . .	8
1.3. Rozwiązywanie problemów za pomocą współczesnych systemów liczących . . . . .	10
1.4. Możliwości przetwarzania wiedzy . . . . .	12
1.5. Nowe stosunki w układzie "człowiek - maszyna" . . . . .	15
1.6. Krótka historia rozwoju sztucznej inteligencji I systemów ekspertowych . . . . .	17
<b>2. PODSTAWY INŻYNIERII WIEDZY I SYSTEMÓW EKSPERTOWYCH . . . . .</b>	<b>23</b>
2.1. Obszar problemowy i metody poszukiwania . . . . .	24
2.2. Przykład prostego problemu . . . . .	24
2.3. Użyteczność wiedzy . . . . .	27
2.4. Inżynieria wiedzy . . . . .	33
2.4.1. Proces rozwoju inżynierii wiedzy . . . . .	34
2.4.2. Zasady pracy systemu ekspertowego . . . . .	35
2.5. Przedstawienie wiedzy I wnioski w systemach ekspertowych . . . . .	38
2.5.1. Reguły produkcji . . . . .	39
2.5.1.1. Przedstawienie wiedzy za pomocą reguł produkcji . . . . .	39
2.5.1.2. Wnioski oparte na regułach produkcji . . . . .	40
2.5.1.3. Przykłady systemów produkcyjnych . . . . .	42
2.5.2. Sieci semantyczne . . . . .	44
2.5.2.2. Wnioski w sieciach semantycznych. . . . .	45
2.5.2.3. Dekompozycja opisywanego świata . . . . .	47

2.5.3. Logika predykatów . . . . .	48
2.5.3.1. Przedstawienie logiki predykatów . . . . .	48
2.5.3.2. Wnioskowanie w logice predykatów . . . . .	49
2.5.4. Model tablicy ogłoszeń . . . . .	50
2.5.4.1. Przedstawienie modelu tablicy ogłoszeń . . . . .	50
2.5.4.2. Wnioskowanie I przetwarzanie w modelu tablicy ogłoszeń . . . . .	51
2.5.4.3. Sterowanie wnioskami . . . . .	53
2.5.5. Systemy frejmowe (wiedza proceduralna) . . . . .	55
2.5.5.1. Teoria frejmów i systemy frejmowe . . . . .	55
2.5.5.2. Formy przedstawienia systemu frejmów . . . . .	56
<b>2.6. Właściwości rozmaitych sposobów przedstawienie wiedzy . . . . .</b>	<b>62</b>
2.6.1. Porównanie współczesnych sposobów przedstawienia wiedzy . . . . .	63
2.6.1.1. Reguły produkcji . . . . .	64
2.6.1.2. Sieci semantyczne . . . . .	64
2.6.1.3. Logika predykatów . . . . .	65
2.6.1.4. Model tablicy ogłoszeń . . . . .	66
2.6.1.5. Systemy oparte na frejmach . . . . .	67
2.6.2. Zbieżności i rozbieżności w sposobach przedstawienia wiedzy . . . . .	69
<b>2.7. Problemy współczesnych systemów inżynierii wiedzy i nowe cele jej przetwarzania . . . . .</b>	<b>73</b>
<b>Spis literatury . . . . .</b>	<b>78</b>

## Przedmowa

Niniejsza praca zawiera niektóre wyniki badań literatury przedmiotu na temat inżynierii wiedzy i systemów ekspertowych. Wyniki te posłużą w przyszłości do opracowania materiałów dydaktycznych dla studentów Akademii Obrony Narodowej.

Potrzeba tworzenia takiej podstawy do tworzenia materiałów dydaktycznych wynika z kilku względów praktycznych wynikających ze specyfiki kształcenia dowódców i oficerów sztabu w Akademii Obrony Narodowej.

Pierwszym względem praktycznym jest burzliwy rozwój inżynierii wiedzy objawiający się w gwałtownym wzroście liczby i jakości systemów ekspertowych, w tym o zastosowaniach wojskowych. Potrzeba dopasowania wymogów kształcenia do realiów współczesności wymaga podejmowania badań w tej dziedzinie.

Drugim jest potrzeba rozwiązywania w praktyce wojskowej, głównie w dowodzeniu wojskami i pracy sztabowej, problemów trudno poddających się jakiegokolwiek formalizacji. Uważa się powszechnie, że takie problemy mogą być rozwiązywane z wykorzystaniem odpowiednich systemów ekspertowych.

Trzecim względem praktycznym jest fakt, że absolwenci Akademii Obrony Narodowej nie są i nie będą profesjonalnymi informatykami, mimo to powinni umieć w przyszłości posługiwać się różnymi programami komputerowymi, w tym również systemami ekspertowymi. Zatem potrzebne jest dostarczenie im w procesie kształcenia pewnego minimum wiedzy i umiejętności umożliwiających zrozumienie istoty funkcjonowania systemów ekspertowych.

Wreszcie czwarty względ praktyczny to potrzeba budowy w przyszłości rozmaitych systemów ekspertowych. Absolwenci Akademii Obrony Narodowej mogą ewentualnie tu występować w roli ekspertów. Stąd konieczność dostarczenia im pewnego minimum wiadomości z zakresu inżynierii wiedzy, co powinno umożliwić lub polepszyć w

przyszłości ich efektywną współpracę z konstruktorami systemów ekspertowych.

Niniejsza praca otwiera cykl dalszych prac na temat inżynierii wiedzy i systemów ekspertowych.

J.Z.

## 1. WSTĘP

W niniejszej pracy omawiana jest budowa i funkcjonowanie systemów intelektualnych, wykorzystujących osiągnięcia nauki i techniki w dziedzinie sztucznej inteligencji. Systemy te skupiają w chwili obecnej powszechną uwagę bowiem z ich opracowaniem i użytkowaniem wiąże się nadzieje na znacznie lepsze niż dotychczas rozwiązywanie różnych problemów praktycznych. Systemy te powstają w wyniku wdrażania nowych technologii przetwarzania informacji, które rozwinęły się szczególnie intensywnie w ostatnich latach w dziedzinie sztucznej inteligencji.

W systemach intelektualnych występuje rodzaj informacji nazywanej wiedza, rozumianą jako zbiór wiadomości z określonej dziedziny przedmiotowej. Zatem, systemy intelektualne są równocześnie systemami przetwarzania wiedzy. W tym sensie, w niniejszej pracy, przetwarzanie informacji będzie nazywane przetwarzaniem wiedzy. W literaturze przedmiotu przetwarzanie wiedzy bywa również nazywane przetwarzaniem intelektualnym lub przetwarzaniem informacji intelektualnej lub jeszcze inaczej, jednak wszystkie te nazwy wskazują na jedną i tę samą technologię. Zanim przejdziemy do szczegółów, rozpatrzmy istotę przetwarzania wiedzy z punktu widzenia w celów i możliwości przetwarzania wiedzy oraz różnic w odniesieniu do tradycyjnych metod przetwarzania informacji.

### 1.1. Co to jest przetwarzanie wiedzy ?

Przyjmujemy, że wiedzą będziemy nazywać informację przechowywaną w komputerze, sformalizowaną zgodnie z określonymi zasadami strukturalnymi, którą można wykorzystywać podczas rozwiązywania różnych problemów praktycznych według algorytmów opartych na wnioskowaniu logicznym. Głównym celem tej pracy jest próba konkretyzacji określenia wiedzy i ułatwienie jego zrozumienia, z punktu widzenia przydatności do rozwiązywania konkretnych problemów.

W ogólnie przyjętym sensie termin "wiedza" jest bardzo popularny. W "Słowniku języka polskiego" wiedza to: 1. "ogół wiadomości zdobytych dzięki uczeniu się; zasób wiadomości z jakiejś dziedziny, gałąź nauki". 2. "znajomość czegoś; uświadomienie sobie czegoś". Często jednak ten termin, zwłaszcza termin "wiedza naukowa" zawiera elementy filozoficzne. Jeśli trzymać się dosłownego znaczenia tego terminu rozpatrywanego na przykład w naukoznawstwie, to jego ogólne koncepcje mogą być niezbyt jasne i trudno zrozumiałe nie tylko dla użytkowników komputerowych systemów intelektualnych, ale również dla specjalistów. Stąd w niniejszej pracy rozpatrującej przedstawianie wiedzy w komputerze, ma się na myśli wspomniany wyżej ograniczony sens. Innymi słowami, przez przetwarzanie wiedzy będziemy rozumieli przekształcanie jej treści za pomocą reguł do postaci umożliwiającej jej zapis w pamięciach komputera oraz wykonywanie na nich różnorodnych operacji, głównie logicznych. Zatem, przy przetwarzaniu wiedzy najbardziej fundamentalnym i ważnym problemem jest przede wszystkim opis sensu treści problemów w jak najszerszym zakresie oraz zastosowanie takiej formy opisu, która zagwarantuje, prawidłowe przetwarzanie ich treści w komputerze za pomocą sformalizowanych zasad i metod przetwarzania. Problem ten zwykle nazywa się problemem przedstawienia wiedzy. Wiedzę opisuje się za pomocą określonych formalizmów, lecz powstaje przy tym mnóstwo problemów dotyczących jej przetwarzania, na przykład w jaki sposób użyć wiedzy do osiągnięcia postawionego celu (rozwiązania problemu) lub w jaki sposób wprowadzać wiedzę do komputera. Celem nauki o przetwarzaniu wiedzy, nazywanej również inżynierią wiedzy, jest rozwiązywanie wszystkich tych problemów, dalszy postęp w rozwijaniu nowego stylu przetwarzania informacji oraz badanie i ulepszanie związków między użytkownikiem a komputerowym systemem intelektualnym.

## **1.2. W poszukiwaniu nowych systemów przetwarzania informacji**

Przetwarzanie wiedzy jest jedną z dziedzin przetwarzania informacji. Komputerom przetwarzającym informacje, niezależnie od sposobu ich organizacji, zadaje się wartości liczbowe. Ich zadawanie i otrzymywanie jest elementem procesu przetwarzania, w których bierze udział człowiek -

użytkownik. Dlatego podczas projektowania systemów przetwarzania informacji rozwiązuje się zwykle również problemy podziału pracy między przyszłym użytkownikiem systemu a komputerem (programem). Wykorzystuje się przy tym - między innymi - dorobek psychologii inżynierskiej i ergonomii pod kątem problematyki współdziałania między człowiekiem i maszyną.

Jak wiadomo, współczesne komputery opierają się na architekturze von Neumanna, wypracowanej 40 lat temu. Realizują one metodę obliczeń bazującą na maszynie Turinga. W przeciągu owych 40 lat, erze błyskawicznego postępu, doskonalono zarówno środki techniczne jak i oprogramowanie współczesnych komputerów, jednak podstawowe zasady ich funkcjonowania nie zmieniały się. Według poglądów japońskich wyrażonych w jednej z prac<sup>1)</sup> taka organizacja komputerów nie jest odpowiednia do zastosowań we współczesnej technologii informacyjnej. Systemy przetwarzania wiedzy powinny przetwarzać informację według zupełnie innych zasad oraz umożliwiać rozwiązywanie takich problemów, których przedtem nie można było rozwiązywać za pomocą metod tradycyjnych. Spróbujemy dalej analizować poglądy japończyków z których narodziła się idea komputerów V generacji, przy tym kiedy będzie mowa o tradycyjnych typach komputerów, ich historii lub przyszłości, to będziemy mieli na myśli komputery z architekturą Turinga - von Neumanna.

Mówiąc o różnicach w metodach przetwarzania informacji, trzeba wyjaśnić co leży u ich podstaw oraz jak wpływają one na system przetwarzania informacji. Najbardziej ogólnymi charakterystykami systemu przetwarzania informacji są: "język opisu" do przedstawiania obiektów przetwarzania oraz "mechanizm, który operuje przedstawieniami obiektów opisanych w tym języku". Każdy system przetwarzania można opisać wcześniej określonym dla niego językiem i będzie on najbardziej efektywny do rozwiązywania zadań, które zorientowane są na mechanizmy, zbudowane na bazie tego języka. We współczesnych komputerach wykorzystuje się języki typu proceduralnego, a jako mechanizmy operacyjne - mechanizmy sterujące i urządzenia arytmetyczne, wypełniające ściśle program przedstawiony w tym języku. W odróżnieniu od tego, w systemach intelektualnych mamy do czynienia z

---

<sup>1)</sup> Osuga. S. Obrabotka znanij. Wyd. "Omsja", Tokio, 1986. (Tłum. na j. ros. Wyd. "Mir", 1989).

językami przedstawienia wiedzy, a jako mechanizmy operacyjne wykorzystuje się rozpatrywane w dalszej części mechanizmy wnioskowania. Zatem właściwości obydwu typów systemów staną się bardziej jasne przy porównaniu ich języków i mechanizmów operacyjnych. Dalej, ponieważ mechanizmy operacyjne komputera są przedstawiane w jakichś językach, to można powiedzieć, że właściwości odpowiednich systemów są skoncentrowane w samych językach. Dlatego w dalszej części tej pracy skupimy uwagę na przetwarzaniu wiedzy jako takiej, a następnie na opisach systemów przetwarzania.

### **1.3. Rozwiązywanie problemów za pomocą współczesnych systemów liczących**

W świecie rzeczywistym powstają różnorodne problemy lecz brak jest podstaw do twierdzenia, że można je wszystkie rozwiązywać pomyślnie za pomocą współczesnych komputerów. Do kategorii problemów najlepiej rozwiązywanych przez skonstruowane do chwili obecnej komputery, należy zaliczyć problemy z ustalonym i stosunkowo prostym cyklem realizacji, w którym problem najpierw jest opisywany, następnie formalizowany i analizowany a na samym końcu, wykorzystując do spreparowanego w ten sposób problemu środki przetwarzania, otrzymuje się pożądaną wynik. Do chwili powstania komputerów cały ten proces był wykonywany przez człowieka i jego automatyzacja była dawniej jedynie nierealizowalnym marzeniem. Z pojawieniem się komputerów typu Turinga - von Neumanna stało się możliwe maszynowe wykonywanie tylko ostatniego etapu tego procesu. W wyniku takiego sposobu rozwiązywania problemu otrzymywało się bardzo dużą ilość różnorodnych danych, przeważnie w postaci wielkich kolumn cyfr na wielostronowych wydrukach, bardzo trudnych do analizowania przez ludzi. Powstała konieczność przystosowywania wyników obliczeń do możliwości ludzi - użytkowników.

Pojawił się zatem nowy etap polegający na przystosowywaniu rozwiązań zorientowanych maszynowo do rozwiązań uwzględniających różnorodne, a przede wszystkim percepcyjne możliwości człowieka, przy minimalizacji kosztów takiego przekształcenia (w szerokim sensie - włączając także czas). Jakość dotychczas ręcznie wykonywanych prac,

które rozpoczynały się od opisu problemu, przeprowadzenia jego analizy aż do "ręcznego" rozwiązania, istotnie różniła się od jakości prac z przetwarzaniem za pomocą komputera w końcowym etapie. Taki sposób rozwiązywania problemów, kiedy dopiero w końcowym etapie wykorzystuje się komputer był i nadal jest uważany jako jak najbardziej naturalny. Ten styl przetwarzania informacji, u podstaw którego leży koncepcja tradycyjnego wykorzystania komputera, rzutuje również na rozwiązywanie problemów programowania. Styl ten do tej pory dobrze zaspokajał wymagania przetwarzania informacji dotyczące współdziałania człowieka i maszyny. W tym ergonomicznym układzie człowiek-maszyna (człowiek-komputer) rozwiązywano pomyślnie wiele problemów takich jak budowa baz danych, systemów operacyjnych i pamięci, które tworzone w celu bardziej efektywnego wykorzystania mechanizmów komputera. Wyobrażenia te miały również wpływ na konstrukcję urządzeń peryferyjnych. W miarę rozszerzania się sfery działania komputerów, obciążenie człowieka, który musiał sprostać takiemu stylowi pracy było jednak tak duże, że zaczęto już mówić o pojawiającym się kryzysie w oprogramowaniu.

Z drugiej strony problemy powstające w świecie rzeczywistym są wielorakie i złożone. Nie zawsze przedstawiane są one wyraźnie i nie zawsze stopień trudności ich przedstawienia jest zgodny z rytmem pojawiania się kolejnych generacji komputerów z architekturą von Neumanna. Oprócz tego same właściwości problemów często różnią się zupełnie od tradycyjnych i z tego względu w wielu przypadkach konieczny jest inny ich opis formalny. Do nich zalicza się problemy projektowania w rozmaitych dziedzinach technicznych takie na przykład jak projektowanie maszyn i mechanizmów, projektowanie układów elektronicznych o wielkiej skali integracji, wytwarzanie wielkomolekularnych związków chemicznych, projektowanie stopów metali, problemy podejmowania decyzji w sferze politycznej i ekonomicznej, problemy sterowania złożonymi systemami, w tym ekonomicznymi, militarnymi i technologicznymi, które na pewno jeszcze bardziej rozwiną się w przyszłości, problemy kształcenia, problemy przesyłania informacji na wielką skalę, w tym rozumienie języków naturalnych i rozumienie ilustracji oraz te problemy z zakresu samego przetwarzania informacji, które bezpośrednio stykają się z problematyką programowania. Wszystko to są nadzwyczaj skomplikowane zadania i możliwości ich rozwiązania wychodzą daleko poza granice, które są zakreślone w ramach tradycyjnych form technologicznych. Jeśli nawet można zapewnić środki, to

znaczy odpowiednią klasę sprzętu oraz narzędzi programistycznych, odpowiednio do stojących zadań, to okaże się, że albo nie ma wysoko wyspecjalizowanych programistów, albo praca programistów będzie nieefektywna. Oprócz tego trudno zapewnić odpowiedniość systemów przetwarzania do tych zadań, pozostając w granicach istniejącego stylu przetwarzania Turinga - von Neumann'a. Przykłady problemów, których nie można rozwiązywać albo rozwiązywać jest je bardzo trudno, pozostając w granicach koncepcji właściwych dla takiego rodzaju współczesnych systemów przetwarzania, będą opisane dalej.

Nakreśliliśmy granice funkcjonalnych możliwości współczesnych komputerów i bazujących na nich systemów przetwarzania informacji. Aby przekroczyć te granice wypracowano w Japonii w końcu lat 80-tych ideę komputerów V-tej generacji. Miały one realizować nowe funkcje, których zadaniem byłoby rozwiązywanie problemów intelektualnych w oparciu o nowe związki między użytkownikiem a komputerem.

#### **1.4. Możliwości przetwarzania wiedzy**

W wyniku odpowiedniego sformułowania omawianych tu problemów, w ostatnich latach ogromnie wzrosło zainteresowanie uparcie rozwijającą się dziedziną - sztuczną inteligencją, zwłaszcza zaś inżynieria wiedzy, w której zaczęto przechodzić do etapu praktycznych zastosowań. Opracowano wiele systemów charakteryzujących się całkowicie lub częściowo specyficznymi cechami przetwarzania wiedzy. W większości tych systemów, które otrzymały nazwę systemów ekspertowych, osiągnano cel formalizacji i gromadzenia wiedzy ekspertów oraz możliwości wykorzystania tej wiedzy przez nie specjalistów. Jednakże systemy ekspertowe tworzone narzędziami inżynierii wiedzy, będące wynikiem badań w dziedzinie sztucznej inteligencji i znajdujące zastosowanie w dzisiejszych czasach, są na ogół zwyczajnymi programami komputerowymi, w których funkcje wnioskowania logicznego realizowane są za pomocą algorytmów. Wykorzystują one stosunkowo proste mechanizmy sterowania pamięcią w celu przechowywania bazy wiedzy i dialogowego wprowadzania i wyprowadzania danych, lecz samo przedstawienie wiedzy w tych systemach jest dość proste. W porównaniu ze zwykłymi programami systemy przetwarzania wiedzy wyróżniają się takimi

zaletami jak możliwość stosunkowo prostej realizacji, łatwej adaptacji do zmian w otoczeniu systemowym, prostotą przyswojenia i dlatego przy właściwym wyborze problemu, mieszczącego się w koncepcji systemów przetwarzania wiedzy, można żywić nadzieję na otrzymanie dobrych wyników.

Systemy wykorzystujące tylko podstawowe funkcje, w których główną rolę odgrywają wnioskowanie dedukcyjne i baza wiedzy, mają jednak bardzo uproszczoną postać. Z tego powodu zakres rozwiązywanych przez nie problemów może być ograniczony właśnie przez ich prostotę.

Stworzone do dzisiejszego dnia systemy przetwarzania wiedzy nie otrzymały zatem jakiegoś specyficznego dla nich kształtu. Ponieważ są to "zwyczajne" programy komputerowe, to nie ma gwarancji, że przy narastaniu rozpiętości wiedzy będą one zdolne wykonywać bardziej złożone funkcje. Co prawda, we współczesnych komputerach z proceduralnym opisem problemów możliwe jest także budowanie dużych systemów oprogramowania przez łączenie względnie niewielkich modułów programowych. Nie jest to jednak proste. Jednak dziś kiedy poważnie bada się wszystkie sposoby, które można wykorzystać w przewyciężeniu kryzysu tworzenia oprogramowania, taka konstrukcja znajduje się w zasięgu możliwości człowieka. Aby jednak zbudowanie skomplikowanego systemu stało się rzeczywiście możliwe, należy wypracować takie metody przedstawienia informacji i sterowania przetwarzaniem, żeby system funkcjonował niezależnie (w sensie logicznym) od rozmiaru krążącej w nim informacji. W dzisiejszych systemach inżynierii wiedzy w zasadzie możliwe jest nieograniczone narastanie rozmiaru wiedzy, lecz w praktyce z tego powodu powstają znaczne trudności w sterowaniu wiedzą. Z tego powodu istnieje niebezpieczeństwo przekroczenia pewnego progu trudności związanych z tym, co nazywamy kryzysem oprogramowania. Tym nie mniej możliwości tające się w technologii przetwarzania wiedzy, nie mają granic. Aby budować solidne podstawy tej technologii trzeba chyba zgodzić się z poglądami japończyków nawołujących do analizy tej technologii od samych podstaw, a następnie tworzyć nowe, niekonwencjonalne style przetwarzania informacji oraz nowe środki techniczne, znajdujące swój wyraz w idei komputerów V-tej generacji. Potrzebny jest skok jakościowy w konstrukcji oprogramowania podobny do tego jaki kilkadziesiąt lat temu został dokonany w technologii baz danych. Jak pamiętamy były

trudności ze sterowaniem informacją w bazach danych ponieważ były one "zaszywane" w programach sterujących. Dziś mamy taką sytuację, że bazy danych funkcjonują zupełnie niezależnie od programów sterujących ich zawartością. Opracowywanie tych programów sterujących staje się coraz skuteczniejsze dzięki istnieniu odpowiednich języków manipulacji danymi oraz wypracowaniu odpowiednich metod projektowania takich na przykład jak metoda CASE w systemach ORACLE lub INFORMIX. Miejmy nadzieję, że już w niedalekiej przyszłości podobna sytuacja wystąpi w sterowaniu bazami wiedzy i mechanizmy wprowadzania wiedzy i wnioskowania zostaną od nich zupełnie oddzielone.

Jak na razie przyjęte w tradycyjnym programowaniu rozgraniczenie programu na "dane" i sam "program" w programowaniu intelektualnym jest często niemożliwe. Dla przykładu translatory i interpretatory, to programy, które traktują inne programy jako dane. W zależności od celu jedne i te same dane mogą w każdym konkretnym przypadku być przedstawiane lub interpretowane w rozmaity sposób, jako deklaratywne, albo jako proceduralne.

W komputerach tradycyjnych przedstawienie danych i komend na poziomie maszynowym jest jednakowe. W językach wyższego poziomu (C, Pascal, Basic itp.) przeciwnie, program i dane zwykle są od siebie wyraźnie oddzielone. Program, procedura lub funkcja napisana w tych językach, to zamknięte swojego rodzaju "światy" operujące na pasywnych danych. Oddziaływanie na nie możliwe jest jedynie za pomocą parametrów.

W programowaniu logicznym w przedstawieniu programu i danych obserwuje się inną skrajność. Uważa się, że w programowaniu logicznym nie ma konieczności oddzielania danych od programu. Program można traktować jako zbiór reguł logicznych.

W podrozdziale 1.2 stwierdziliśmy, że właściwości odpowiednich systemów są skoncentrowane w samych językach. W rzeczywistości każdy język programowania posiada swoisty punkt widzenia na problematykę przedstawienia programu i danych. Na przykład autorzy jednej z prac na temat języka Lisp uważają<sup>2)</sup>, że we wzajemnych stosunkach między programem a danymi osiągnięto pewien kompromis i traktują to

---

<sup>2)</sup> Hyvönen E., Seppänen J. 2 LISP-MAAILMA. Ohjelmointimenetelmat ja -jarjestelmat. "Kirjayhtymä", Helsinki, 1987. (tłum. na j.ros. Mir Lispa. Metody i sistemi programirowanija. Tom 2. "Mir", Moskwa, 1990).

jako zaletę tego języka. Forma przedstawienia programu i danych jest jednolita. Program można traktować dość swobodnie jako dane, a dane można interpretować i stosować jako program. W konstrukcjach tego języka istnieją dwie interpretacje, które zależą od ich wykorzystania. Sposób wykorzystania i interpretacji wyrażeń określa sam programista. Środki i mechanizmy niezbędne do wywołania interpretatora i wykorzystania obiektów (EVAL, APPLY, FUNCALL, QUOTE, FUNCTION itp.) tak samo, jak środki wykorzystywane do realizacji obliczeń, znajdują się w dyspozycji programisty. Programista decyduje o przekształceniu programu w dane i danych w program.

Nie mniej jednak autorzy wyżej wspomnianej pracy zauważają, że w niektórych systemach lispowych możliwe jest programowanie zorientowane obiektowo, tak jak jest to realizowane w systemie SmallTalk, w którym dane i procedury są od siebie oddzielone, mimo że stanowią one jednolity obiekt.

### **1.5. Nowe stosunki w układzie "człowiek - maszyna"**

Technologia przetwarzania informacji w ciągu jej całej 30-letniej historii, aż do chwili obecnej jak już wspomnieliśmy, była prowadzona prawie bez wyjątków z wykorzystaniem mechanizmów liczących opartych na idei maszyny Turinga - von Neumanna. Systemy ekspertowe, będące praktyczną realizacją idei przetwarzania wiedzy, w swojej istocie idą tą samą drogą, śladami tych samych metod, a realizacja ich funkcji bazuje na tych samych mechanizmach. Jednakże przy analizie wspomnianych problemów sprawą najważniejszą powinno być samo rozwiązywanie zadań praktycznych występujących w świecie rzeczywistym. Główne trudności jakie tu występują to: po pierwsze - pożądana metoda tworzenia struktury systemu przetwarzania wiedzy, która posiadałaby możliwości rozwiązywania złożonych problemów nie jest jeszcze doskonała; po drugie - w intelektualnym przetwarzaniu informacji cele powinno się osiągać w układzie "człowiek - komputer" zupełnie inaczej, aniżeli w maszynie Turinga - von Neumanna. Według poglądów japońskich przyjęcie nowego stylu przetwarzania informacji wymaga najpierw przeanalizowania zachowania się człowieka podczas osiągania przez niego celu w drodze rozwiązywania określonego problemu, a następnie

rozpatrzenia takich pytań jak: jaką część procesu rozwiązywania zadania można automatyzować za pomocą środków przetwarzania wiedzy, na ile efektywnie można wspomagać za pomocą tych środków rozwiązywanie problemów i co będzie dawał system, który udało się zrealizować i wdrożyć do zastosowań praktycznych?

Obecnie systemy przetwarzania wiedzy, które osiągnęły w swoim rozwoju poziom eksperymentalny i komercyjny zostały zrealizowane wyłącznie na komputerach typu tradycyjnego. Przedstawienie problemów zorientowane na przetwarzanie wiedzy, jest obecnie prowadzone tylko po ich przekształceniu i przedstawieniu przyjeźtemu w komputerach tradycyjnych. Jeśli można byłoby taki mechanizm przetwarzania utworzyć na bazie zasad Turinga - von Neumanna, to również na tych zasadach można byłoby realizować i przetwarzanie wiedzy. Więcej, rzeczywiste wprowadzenie do praktyki przetwarzania maszynowego, operującego na przedstawionej informacji, zorientowanego na przetwarzanie wiedzy, możliwe jest z wykorzystaniem innych środków różniących się od von neumanowskich. Fakt ten, dotyczy nie tylko klasyfikacji, przy której dokonywano by przedstawienia koncepcji przetwarzania wiedzy w koncepcji Turinga - von Neumanna. Sens jego polega na wdrożeniu nowych zasad przetwarzania informacji. Inaczej mówiąc, wyróżniającą cechą przetwarzania wiedzy jest możliwość zmiany współdziałania w układzie "człowiek - komputer" i zgodnie z tym - stworzenia nowego stylu rozwiązywania problemów. Kiedy tworzy się system przetwarzania wiedzy, należy na początku wyraźnie wyjaśnić ten fakt jako taki, a następnie wybrać najbardziej przydatną architekturę (włączając tę część systemu, która dokonuje rozwiązania zorientowanego na przetwarzanie wiedzy). Powstaje tu niebezpieczeństwo, że użycie tradycyjnych systemów przetwarzania wiedzy pociągnie za sobą błędne rozwiązania. Jak zobaczymy dalej, błędne jest samo porównanie systemu, w którym problem opisywany jest w języku typu proceduralnego z systemem, w którym jest on opisany w języku typu deklaratywnego.

## 1.6. Krótka historia rozwoju sztucznej inteligencji I systemów ekspertowych

Od najdawniejszych czasów zawsze marzono o tym aby budować urządzenia lub mechanizmy, które naśladowałyby działania ludzi po to, aby zadziwiały albo wykonywały różne rzeczy od najprostszych, aż do takich, które umożliwiły by ich właścicielowi panowanie nad światem. Stąd powstawały pewne legendy.

Jedną z nich jest opowieść o tym jak w XIII wieku uczony, filozof i teolog, Albert von Bollstädt zwany Albertem Wielkim, pokazał swojemu uczniowi zbudowaną przez siebie maszynę o ludzkiej postaci, która potrafiła chodzić i mówić. Uczniem tym był sam Tomasz z Akwinu. Przyjął on tę kukłę za diabelski wytwór i zniszczył ją laską, którą zwykle nosił po ulicach Kolonii.

Inna legenda opowiada o tym jak w średniowiecznej Pradze kabalista, astronom i alchemik cesarza Rudolfa II, rabbi Liwa (Löw) ben Becalel, stworzył Golema - postać ludzką ulepioną z gliny. Golem miał zaspokajać zachcianki i wykonywać różne prace dla kłótlivej żony rabbiiego, co robił bardzo dobrze. Miał tylko jedno ograniczenie - jak twór prawowierne go żyda nie mógł pracować w sobotę. Dlatego rabbi zawsze w piątek wieczorem wyjmował mu szem - klucz uruchamiający Golema. Pewnego razu zapomniał tego zrobić. W sobotę Golem zaczął szaleć. Nie słuchał żadnych rozkazów, niszczył wszystko co napotkał na swojej drodze i tylko z wielkim trudem udało się rabbiemu wyjąć szem, po czym Golem przewrócił się i rozpadł.

Jeszcze jedna legenda wywodzi się z idei możliwości stworzenia sztucznego człowieka zwanego homunkulusem za pomocą odpowiednich zabiegów magicznych i alchemicznych. Powstała w XVI wieku i jej prekursorem był Paracelsus (Theophrastus Bombastus von Hohenheim) lekarz, alchemik i filozof urodzony w Szwajcarii. Idea ta długi czas zaprzętała myśli alchemików i uczonych. Idea ta została wykorzystana dotąd jedynie w literaturze i sztuce. Homunkulus prowadzi Fausta i Mefista na Klasyczną Noc Walpurgii. Można wymienić wiele postaci stworzonych sztucznie w literaturze science fiction, obdarzonych nieprzeciętną inteligencją. Dla nas interesujące jest szerokie rozpowszechnienie wiary w możliwość stworzenia istot podobnych do człowieka w inny sposób, aniżeli czyni to przyroda.

W XVIII wieku rozwinęła się bardzo sztuka zegarmistrzowska. Wielu słynnych zegarmistrzów próbowało z powodzeniem budować mechanizmy człowiekopodobne, wykonujące pewne czynności podobnie jak ludzie.

Powyższe przykłady świadczą o dążeniach ludzi do zbudowania istot na swoje podobieństwo wykonujących pewne celowe działania. Dążenia te mogły być zrealizowane dopiero w miarę rozwoju cybernetyki, zwłaszcza zdefiniowania i zrozumienia zjawiska sprzężenia zwrotnego. Poważniejsze prace w dziedzinie sztucznego intelektu mogły być jednak realizowane dopiero w erze komputerów.

W 1954 roku A. Newell postanowił zbudować program do gry w szachy. Przedtem C.E. Shannon, ojciec teorii informacji, zaproponował odpowiednią do tego metodę. A.M. Turing, jeden z prekursorów informatyki, udokładnił tę metodę i zbudował odpowiedni model. W amerykańskiej kompanii Rand Corporation J.C. Shaw i H.A. Simon zespolili wysiłki w pracy nad projektem Newell'a. Wspierał ich zespół psychologów z Amsterdamu pod kierownictwem De Groot'a <sup>3)</sup>, który studiował styl gry arcymistrzów szachowych. Język programowania, wytworzony specjalnie przez tę grupę był przeznaczony do łatwej manipulacji informacją w symbolicznej postaci maszynowej, mógł pracować z systemem wskaźników i przetwarzać listy. Był to język programowania IPL1 (1956), poprzednik języka Lisp (1960), którego twórcą był J. McCarthy <sup>4)</sup>. Pierwszym programem w dziedzinie sztucznej inteligencji był program o nazwie "Logik - Teoretyk", przeznaczony do dowodzenia twierdzeń i rachunku zdań. Program ten został po raz pierwszy zademonstrowany 9 września 1956 r. <sup>5)</sup>

Program do gry w szachy NSS (Newell, Shaw, Simon) został opracowany w roku 1957. Jego struktura i struktura programu "Logik - Teoretyk", wyobrażenie o "pożądanych sytuacjach" i "heurystykach" (regułach, które pozwalają dokonać wyboru przy braku ścisłych podstaw teoretycznych) doprowadziły później do koncepcji "Uniwersalnego Rozwiązywacza Problemów". Program ten analizując różnice między

---

<sup>3)</sup> De Groot. Thought and choice in chess. Amsterdam, Mouton 1965.

<sup>4)</sup> McCarthy J. Lisp 1.5 Programmers Manual. MIT Press, Cambridge, Massachusetts, 1963.

<sup>5)</sup> Lauriere J.-L. Intelligence artificielle. Eyrolles, Paris 1987.

sytuacjami i konstruuując cele, pomyślnie rozwiązuje łamigłówki typu "Wieża Hanoi".

Specjaliści w dziedzinie informatyki rozpoczynają następnie interesować się bezpośrednio sztuczną inteligencją i niektórzy z nich, tacy jak J. McCarthy, M. Minsky i H.A. Simon piszą wspaniale i często cytowane prace. Powstają nowe programy. H. Gelernter<sup>6)</sup> pokazuje, że jego program dowodzenia twierdzeń ze szkolnej geometrii może pracować lepiej aniżeli jej twórca! Aby udowodnić, że trójkąt ABC posiadający równe sobie dwa kąty u podstawy (w wierzchołkach B i C), jest równoramiennym, program zamiast klasycznego dowodu przytaczanego w podręcznikach szkolnych, polegającego na budowie wysokości opuszczanej z wierzchołka A na podstawę, po prostu stosuje twierdzenie o równości trójkątów ABC i ACB. Wynik jest oczywisty.

Powstaje również program EPAM (Elementary Perceiving and Memorizing Program - program prostego odbioru i zapamiętywania) opracowany przez E. Feigenbaum'a<sup>7)</sup>, przeznaczony do modelowania sytuacji psychologicznych.

Programy posługujące się pytaniami sformułowanymi w języku naturalnym, stworzono już dawno. Znalazły one zastosowanie przy poszukiwaniu informacji w bazach danych. Na przykład program BASEBALL<sup>8)</sup> odpowiadał na pytania dotyczące wyników rozgrywek bejsbolowych jakie odbyły się w przeszłości, a program STUDENT<sup>9)</sup> potrafił rozwiązywać zadania algebraiczne, sformułowane w języku angielskim.

Bardzo duże nadzieje pokładali badacze w pracach w dziedzinie maszynowego tłumaczenia języków naturalnych. W tej sferze kontynuują prace duże zespoły badaczy. Prace te zorientowane są przede wszystkim

---

<sup>6)</sup> Gelernter H. Realisation of a geometry theorem-proving machine. Proc. Inter. Conf. Inform. Proc., UNESCO House, Paris 1960.

<sup>7)</sup> Feigenbaum E., and Feldman J. (Eds.) Computers and Thought. McGraw-Hill, New York 1963.

<sup>8)</sup> Green B.F., Wolf A.K., Chomsky C., Laughery K. Baseball: an automatic question answerer (w pracy Feigenbauma, 1963).

<sup>9)</sup> Bobrow. D.G., Natural Language Input for a Computer Problem Solving System. M.Minsky (ed.) Semantic Information Processing. Cambridge. MIT, 1968.

na wykorzystania analizy syntaktycznej i informacji otrzymanej ze słowników (tzw. metoda słów kluczowych)<sup>10)</sup><sup>11)</sup>, tym nie mniej badacze stracili wiele lat na to, aby przekonać się, że automatyczne tłumaczenie języków nie jest problemem izolowanym i do pomyślnego jego rozwiązania niezbędne jest zrealizowanie etapu dotyczącego rozumienia mowy.

Nowe podejście do logiki formalnej, oparte na doprowadzeniu wnioskowania do sprzeczności, pojawiło się w roku 1965<sup>12)</sup>. Podejście to pozwala formalizować wiele zadań i przeprowadzać ich interpretację maszynową. Podejście to było wykorzystywane z powodzeniem do dowodzenia twierdzeń<sup>13)</sup> i weryfikacji programów<sup>14)</sup>. To samo podejście posłużyło jako punkt wyjścia przy tworzeniu oryginalnego języka programowania - Prolog, który dysponuje mocą logiki predykatów pierwszego rzędu. Jest on produktem europejskim, powstał w roku 1971 na uniwersytecie w Marsylii. Opracowany został przez zespół pod kierownictwem A. Colmerauer'a<sup>15)</sup>.

Badaniom w dziedzinie sztucznej inteligencji towarzyszą opracowania języków programowania nowych generacji i tworzenie bardziej wyrafinowanych systemów programowania. Daje to możliwość podczas opracowania programów komputerowych wykorzystywać metody rozumowania w

---

<sup>10)</sup> Dreyfus H.L. What computers cannot do. Revised ed. Harper (1972).

<sup>11)</sup> Lightill J., "Artificial Intelligence: A General Survey", Artificial Intelligence a Paper Symposium. Science Research Council, 1973.

<sup>12)</sup> Robinson J.A., A machine-oriented logic based on the resolution principle. JACM, 12(1)- 1965.

<sup>13)</sup> 1) Slagle J.R. Artificial Intelligence: The Heuristic Programming Approach. New York: McGraw-Hill, 1971, 2) Green C. Application of Theorem Proving to Problem Solving. In IJCAI-1, 1969, 3) Kowalski R. Predicate Logic as a Programming Language. Information Processing 74. North-Holland, Amsterdam 1974.

<sup>14)</sup> Waldinger R.J. and Levitt, K.N. Reasoning about Programs. Artificial Intelligence, 5(3) 1974.

<sup>15)</sup> Colmerauer A., Kanoul H., Pasero R., Roussel P. Un système de communication homme-machine en français. In Rapport Groupe d'Intelligence Artificielle, Université d'Aix-Marseille, Luminy, France. 1973.

oparciu o zwykły zapas słów. Więcej, języki programowania Lisp, Prolog, PLANNER, QA4 (wymieniono tu tylko najważniejsze) pozwalają za pomocą koncepcji celu i twierdzenia, modelować i formalizować logiczny wniosek podczas rozwiązywania zadań; języki MACSYMA i REDUCE pozwalają przeprowadzać formalne manipulacje z wyrażeniami matematycznymi; język TMS pozwala sterować przy niepewnych wiadomościach i sprawdzać ich zgodność między sobą.

Do roku 1968 w dziedzinie sztucznej inteligencji pracowano w zasadzie nad oddzielnymi "mikroprzestrzeniami". Tworzono systemy przydatne do takich specyficznych i ograniczonych sfer zastosowań jak gry, geometria euklidesowa, rachunek całkowy, przetwarzanie krótkich fraz języka naturalnego z niewielkimi zbiorami słów. Prawie we wszystkich tych systemach wykorzystywano jedno i to samo podejście - uproszczoną kombinatorykę, bazującą na ograniczonym wyborze alternatyw, na podstawie zdrowego rozsądku, użyciu liczbowych funkcji ocenowych oraz rozmaitych heurystykach. Obecnie powstało wiele doskonałych programów wykorzystujących dorobek w dziedzinie sztucznej inteligencji. Na przykład, opracowano wiele łatwo dostępnych programów do gry w szachy o takiej jakości, że niewielu graczy ma szansę wygrania partii z "przeciwnikiem" zrealizowanym w programie komputerowym.

Na początku lat 70-tych nastąpił skok jakościowy w badaniach nad sztuczną inteligencją. Tłumaczy się to dwoma przyczynami:

Po pierwsze, wszyscy badacze stopniowo uświadamiali sobie, że wszystkie dotychczas opracowane programy w dziedzinie sztucznej inteligencji, nie dysponują najważniejszym - głęboką wiedzą w odpowiedniej dziedzinie. Taką głęboką wiedzę, w postaci nagromadzonych w przeciągu wielu lat doświadczeń, posiadają tylko eksperci w uprawianej przez siebie dziedzinie. Dlatego aby istotnie polepszyć wyniki pracy jakiegokolwiek programu sztucznej inteligencji trzeba nie po prostu doskonalić heurystyki lub jakiejkolwiek współczynniki liczbowe, na podstawie których pracuje program, a odwrotnie - wykorzystywać metody wnioskowania logicznego oraz wiedzę nagromadzoną w doświadczeniu, przedstawioną w formie symbolicznej.

Abym uczynić program intelektualnym, trzeba go wyposażyć w wysokiej jakości zasób wiedzy specjalistycznej w określonej dziedzinie przedmiotowej.

Po drugie, Powstaje konkretny problem jak przekazać tę wiedzę programowi, jeśli jego twórca jej nie posiada. Odpowiedź jest jasna - sam program powinien je wydzielić z danych otrzymywanych od eksperta. Badacze zetknęli się z koniecznością wyposażenia systemów sztucznej inteligencji w możliwości, których nie ma w zwykłych językach programowania, zwłaszcza, że programy sztucznej inteligencji powinny umieć same zbierać informację, przechowywać i wykorzystywać ją tylko przy wystąpieniu dostatecznych podstaw. W danym przypadku istnieje granica między wnioskiem o jakimś fakcie, a wykorzystaniem tego faktu. W przeciwieństwie do tego, zwykły język programowania pozwala wyrażać tylko możliwe do wykonania zadania lub instrukcje. Właściwość ta jest bardzo ważna, ponieważ ekspert wyposaża system w oddzielne, izolowane fakty, nie będąc uprzedzony o tym, w jakim momencie program weźmie je pod uwagę.

Powstał zatem problem odpowiedniego przedstawienia wiedzy wyrażonej w języku naturalnym. Doprowadziło to do rozwoju specjalizowanych programów komputerowych, z których każdy jest ekspertem w określonej dziedzinie. Programy te otrzymały nazwę systemów ekspertowych.

Na początku lat 70-ych powstało kilka pierwszych systemów ekspertowych o takich nazwach jak DENDRAL, MYCIN, PROSPECTOR, PUFF itp., których opis, można znaleźć prawie w każdej szanującej się pracy na temat systemów ekspertowych. Stały się one zatem pewnymi przykładami klasycznymi, a wypracowane podczas ich projektowania takie elementy jak szkielety baz wiedzy oraz mechanizmy wnioskowania posłużyły do budowy następnych, jeszcze bardziej doskonałych, systemów ekspertowych. Obecnie można kupić gotowe oprogramowanie w postaci tak zwanych "shell'ów" zawierających szkielety baz wiedzy oraz mechanizmy wnioskowania, wyposażone w odpowiednie specjalistyczne języki programowania. Za ich pomocą można budować własne systemy ekspertowe.

Początkowo projektowanie i budowę systemów ekspertowych postrzegano bardziej jako sztukę programowania, aniżeli jako pracę naukową. Obecnie jednak proces ten stał się bardziej zrozumiały i wyraźniej określony, czego dowodem jest pojawienie się na zachodzie całego szeregu prac ujmujących całościowo tę problematykę. Jako przykład

z  
-  
d  
w  
h  
y  
o  
e  
a  
c

takiej pracy można wymienić prace<sup>16)</sup>, które były dostępne w Polsce, w tłumaczeniu na język rosyjski.

Inżynieria wiedzy to technologia budowy systemów ekspertowych. W procesie budowy systemu ekspertowego wymaga się specyficznej formy współdziałania między twórcą systemu ekspertowego, którego nazywa się inżynierem wiedzy z jednym lub kilkoma ekspertami w określonej dziedzinie. Inżynier wiedzy pozyskuje od ekspertów procedury, strategie, reguły empiryczne, które wykorzystują oni do rozwiązywania zadań.

## **2. PODSTAWY INŻYNIERII WIEDZY I SYSTEMÓW EKSPERTOWYCH**

Jak powiedziano w rozdziale 1, w dziedzinie sztucznej inteligencji istnieje potrzeba wykorzystania możliwości jakie daje nowa technologia przetwarzania informacji, jednak obecne realizacje sprowadzają się tylko do rozwiązań cząstkowych. Co trzeba zrobić aby wykorzystać ten ukryty potencjał? Dla lepszego zrozumienia tej okoliczności spróbujemy w sposób uproszczony rozpatrzyć proces rozwoju i podstawy sztucznej inteligencji. Ponieważ celem tej książki jest danie wyobrażenia o systemach przetwarzania informacji zupełnie innego typu, to uwagę czytelnika skupia się na istocie sztucznej inteligencji, rozpatrywanej właśnie z tej pozycji. Co się tyczy współczesnej technologii sztucznej inteligencji i jego zastosowania, to w tym celu trzeba sięgnąć do innej literatury.

### **2.1. Obszar problemowy i metody poszukiwania**

Podczas rozwijania sztucznej inteligencji powstało szereg nowych punktów widzenia, dano podstawy nowej technologii przetwarzania informacji. Przy tym nowa technologia znalazła zastosowanie w najbardziej różnorodnych zastosowaniach, na przykład w rozumieniu języka naturalnego, rozpoznawaniu obrazów i systemach ekspertowych. Przydała się również w projektowaniu systemów intelektualnych. W rozdziale tym

---

<sup>16)</sup> Hayes-Roth F., Waterman D.A., Lenat D. Building Expert Systems. Addison-Weseley, 1983 oraz Waterman D.A. A Guide to Expert Systems. Addison-Weseley, 1986.

zwrócono uwagę na sposób myślenia i podejścia leżące u podstaw zastosowań metod sztucznej inteligencji.

We wczesnym etapie rozwoju zastosowań praktycznych sztucznej inteligencji badano metody wyszukiwania. W okresie kiedy badania były prowadzone faktycznie po ciemku, nie wdawano się w systematyzację metod. Ludzie jak zawsze mieli skłonność do rozwiązywania na komputerach różnych gier (zagadek), które jak i dziś były bodźcem rozwoju intelektualnego człowieka. Programy do rozwiązywania zagadek i prowadzenia gier odróżniały się od zwykłych programów, ponieważ wymagały przede wszystkim tworzenia nowych algorytmów. Algorytmy te, nazywane dziś metodami poszukiwania (wyszukiwania), prawie wszystkie weszły do programów intelektualnych, jako fundamentalny kierunek sztucznej inteligencji. Stąd na początku rozpatrzmy prosty przykład ilustrujący metodę wyszukiwania.

## 2.2. Przykład prostego problemu

Wiele zastosowań sztucznej inteligencji związane jest z określeniem kolejności wykonania serii operacji. Jako przykład przytoczymy rozwiązywanie gry w "8" podane w pracy N.J. Nilsona <sup>17)</sup>. Zadania tego typu, można spotkać podczas rozwiązywania problemów związanych ze sterowaniem ruchami robota, automatycznym programowaniem i innymi.

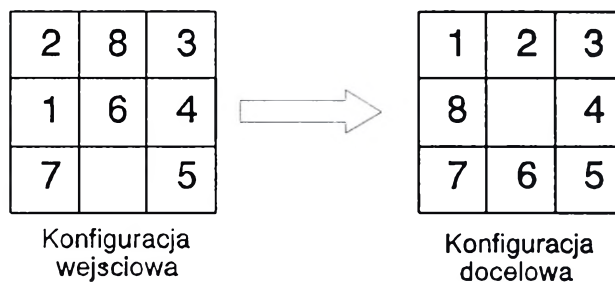
Gra w "8" jest prostym zadaniem ilustrującym tego rodzaju problemy. Wykorzystuje się w niej planszę składającą się z  $3 \times 3 = 9$  pól zawierających zbiór cyfr od 1 do 8 i jedno pole puste. Na rysunku 1 przedstawiono położenie cyfr wejściowe i docelowe. W położeniu wejściowym cyfry na planszy rozmieszczone są w sposób przypadkowy, przy tym przedstawiono tu tylko jedną z 362880 (tzn.  $9!$ ) możliwych konfiguracji 8-miu cyfr i pustego pola.<sup>18)</sup> Rozwiązanie gry polega na ustawieniu cyfr tak, jak przedstawiono to na planszy docelowej, przy zachowaniu tylko jednej reguły gry. Reguła ta polega na tym, że wolno przesuwac jedną

---

<sup>17)</sup> Nilsson N.J. Principles of Artificial Intelligence. Tiog Publishing Company, Palo Alto. 1980.

<sup>18)</sup> Z tej liczby tylko z połowy stanów wejściowych można dojść do stanu docelowego. Druga połowa to stany nierozwiązywalne.

cyfrę na wolne miejsce tylko ruchami w kierunku pionowym lub poziomym (prostokątnym) dotąd, aż otrzyma się stan docelowy. Jako rozwiązanie należy podać prawidłową kolejność ruchów, na przykład: przesunąć cyfrę 6 w dół, przesunąć cyfrę 8 w dół itd., aż do otrzymania roz-



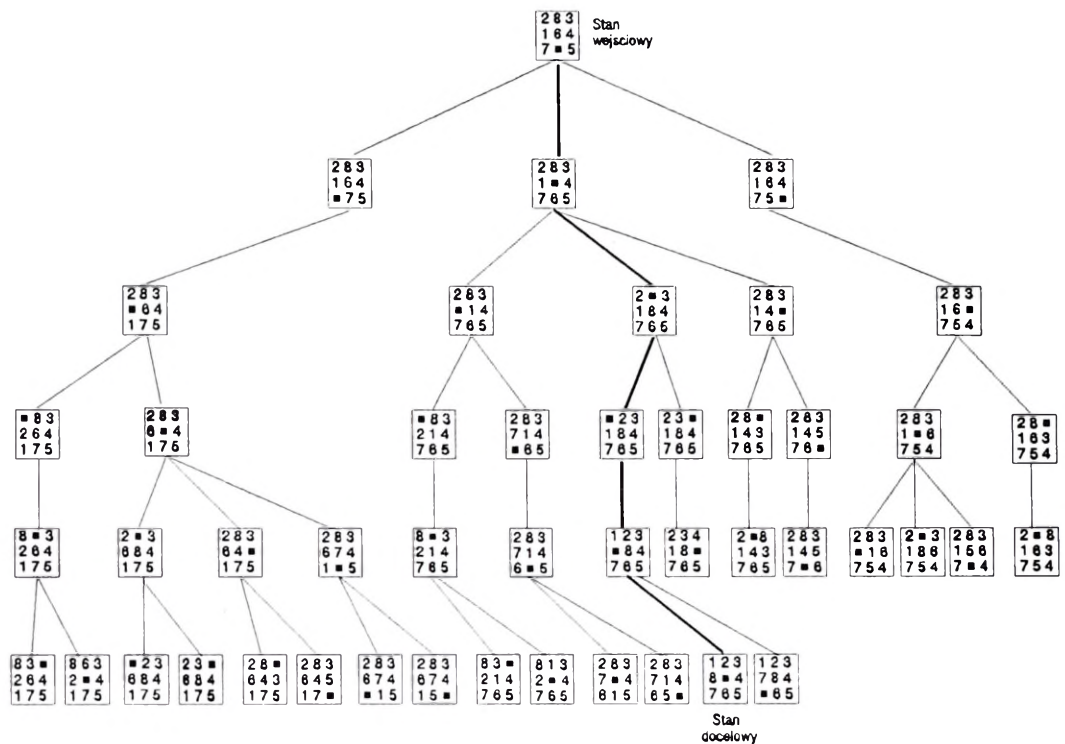
Rys. 1. Konfiguracja wejściowa i docelowa w grze w "8".

mieszczenia cyfr na planszy docelowej. Można również zadać inną konfigurację położenia cyfr na planszy docelowej z tym, że nie każda kombinacja z wyżej wymienionych 362880 da się sprowadzić do położenia docelowego.

Na planszy wejściowej mamy zatem określony stan zadania, który możemy zapisać w bazie danych za pomocą tablicy o wymiarach 3x3. Dla innego typu zadań wejściowa baza danych może mieć inną postać ale będzie zawsze zawierała opis początkowego (wejściowego) stanu zadania. W praktyce może to być dowolna struktura danych opisująca stan zadania. Mogą to być ciągi symboli, wektory, zbiory, macierze, tablice, drzewa, listy itp., w zależności od rodzaju zadania. W naszym przykładzie gry w "8" struktura danych odzwierciedla fizyczne położenie cyfr na planszy lecz w zadaniach innego typu można odzwierciedlać co innego.

Każdy ruch na planszy przekształca jeden stan zadania w następny. Wygodnie jest również interpretować grę w "8" jako przesuwanie pustego pola do następnego w czterech kierunkach: w dół, w lewo, w górę lub w prawo. Na rys. 2 przedstawiono drzewo zawierające pełny zbiór stanów jakie mogą wystąpić podczas przesuwania pustego pola o jedną pozycję, z których tylko jeden jest stanem docelowym. Pogrubionymi liniami pokazano drogę najlepszego sposobu osiągnięcia stanu docelowego, do czego potrzebne jest wykonanie 5 przesunięć cyfr, przy stanie wejściowym określonym na rys. 1.

Oczywiście jeśli stan wyjściowy lub (i) docelowy będzie inny, to zarówno kształt drzewa jak i minimalna liczba przesunięć będą inne.



Rys. 2. Drzewo stanów gry w "8".

### 2.3. Użyteczność wiedzy

Powróćmy do opisu gry w "8". Opisany algorytm wyszukiwania rozwiązania można realizować w postaci programu komputerowego. Wynik wydawany przez ten program, będzie przedstawiał drogę (kolejność ruchów cyframi) od stanu wejściowego do docelowego. Załóżmy jednak, że nie dysponujemy komputerem, nie mamy pojęcia o efektywnych sposobach rozwiązywania, heurystykach itd., a chcemy jednak rozwiązać to zadanie. Prawdopodobnie będziemy próbować przesuwać poszczególne cyfry metodą prób i błędów dotąd, aż otrzymamy stan docelowy. Jeśli wykonamy kilka ruchów cyframi i nie otrzymamy stanu docelowego to będziemy musieli wracać w górę drzewa aby ponownie wykonać inną serię ruchów i tak dalej aż do skutku. Jeśli

nie będziemy zapamiętywać lub zapisywać wykonywanych czynności, to prawdopodobnie spostrzeżemy się, że niektóre serie niewłaściwych ruchów wykonujemy powtórnie. Będziemy jednak zawsze w trakcie rozwiązywania zadania ciągle porównywali aktualny stan na planszy ze stanem docelowym i na tej podstawie podejmowali decyzje co do dalszych ruchów. Będziemy zatem powtarzać cyklicznie pewien ciąg następujących czynności:

1. Porównać stan wejściowy (aktualny) z docelowym.
2. Wykonać ruch, który naszym zdaniem przybliży aktualny stan do stanu docelowego.
3. Powtarzać czynności 1 i 2 dotąd aż uzyska się stan docelowy.

W miarę rozwiązywania większej liczby zadań przy różnych stanach wejściowych będziemy nabierali pewnych umiejętności, które pozwolą nam coraz sprawniej uzyskiwać stan docelowy.

Jeśli dysponujemy komputerem to możemy zbudować algorytm oraz napisać odpowiedni program, który wykona za nas całą pracę przy rozwiązywaniu zadania, to znaczy określi jak kolejno przesuwac cyfry na planszy aby otrzymać optymalną drogę dojścia do stanu docelowego.

Algorytm, na podstawie którego można będzie napisać program w języku proceduralnym (Basic, Pascal, C lub jakimkolwiek innym) musi być zbudowany bardziej szczegółowo aniżeli uprzednio przytoczony ciąg czynności wykonywanych podczas rozwiązywania tego zadania przez człowieka. Na pewno program powinien zapamiętać dwa stany na planszy - wejściowy i docelowy, ponieważ są one po prostu danymi wejściowymi. Ponadto aby opracowywany przez nas algorytm był realizowalny to musi przewidywać działania na modelu samego urządzenia do gry, to znaczy planszy i cyfr. W przypadku rozwiązywania zadania przez człowieka mamy do czynienia z prawdziwą planszą i klatkami z cyframi, czyli rzeczywistym urządzeniem do gry a nie jego modelem.

W każdej nawet najprostszej grze obowiązują jakieś reguły. Również w przypadku naszej gry w "8" mamy do czynienia z taką regułą. Dotyczy ona ruchów cyframi i można ją sformułować następująco:

*Cyfrę można przesuwać na wolne miejsce planszy tylko w kierunkach pionowym lub poziomym (nie wolno przesuwać cyfry po przekątnych).*

Reguła jest częścią składową elektronicznego modelu planszy do gry. Zatem w komputerowym modelu gry w "8" mamy do czynienia z odwzorowaniami:

1. 9-ciu pól, które można zapisać w postaci tablicy 3x3.
2. Definicji 8-miu cyfr lub innych dowolnych symboli, zapisanych w postaci 8 liczb. Oczywiście cyfry nie są jednakowe i mogą być zawsze uporządkowane w zadany sposób. Będzie to również dotyczyć każdych innych 8 symboli, które zdecydujemy się użyć zamiast cyfr.
3. Reguly gry, oparte na zasadzie, która mówi, że klatka z cyfrą można przesunąć tylko w kierunku prostopadłym i na puste miejsce. Oczywiście nie można jej przesunąć poza granice planszy co musi być odzwierciedlone w regułach gry. Powyższa zasada gry powinna być zamodelowana bardziej szczegółowo, to znaczy powinna zawierać opis szczegółowych reguł dotyczących ruchów jakie można wykonać z każdej pozycji na planszy. Reguly te mogą być zapisane w postaci JEŚLI, TO, a więc:

Reguła 1 JEŚLI klatka znajduje się w górnym lewym rogu,  
 TO można go przesunąć: JEŚLI z prawej puste miejsce  
 TO wprawo LUB JEŚLI poniżej puste miejsce TO w  
 dół;

.....

Reguła 4 JEŚLI klatka znajduje się w środku planszy (nie  
 dotyka do krawędzi planszy)  
 TO można go przesunąć: JEŚLI powyżej puste miejsce  
 TO w górę LUB JEŚLI z prawej puste miejsce TO w  
 prawo LUB JEŚLI z dołu puste miejsce TO w dół  
 LUB JEŚLI z lewej puste miejsce TO w lewo;

.....

Reguła 9 JEŚLI klatka znajduje się w dolnym prawym rogu,  
 TO można go przesunąć: JEŚLI z lewej puste miejsce TO  
 w lewo LUB JEŚLI powyżej puste miejsce TO w  
 górę.

W naszym wypadku mamy do czynienia tylko z dziewięcioma polami na planszy, zatem mamy jedynie 9 reguł dotyczących modelu użycia urządzenia do gry w "8". Gra w "8" przekształca się w grę w "15" jeśli użyjemy planszy 4x4 pola. Gra w "24" wymagać będzie planszy o

wymiarach 5x5 pól. Liczba reguł będzie stała i będzie zawsze wynosić 9.

Rozwiązywanie każdej gry za pomocą komputera wymaga w pierwszym etapie zamodelowania urządzenia do tej gry oraz opisu reguł jakie w niej obowiązują. Zarówno urządzenia do gry jak i jej reguły mogą mieć różny stopień złożoności. Oprócz opisanej tu gry w "8" istnieje cały szereg innych znanych gier i zagadek takich jak na przykład: wilk koza i kapusta, zadanie komiwojażera, kostka Rubika, warcaby, szachy i wiele innych. Za pomocą właściwie zbudowanego modelu gry możemy wprowadzać różne jej stany początkowe. Przy tym liczba prawidłowych stanów początkowych, to znaczy takich, z których można dojść do stanu docelowego jest mniejsza aniżeli maksymalna liczba stanów gry jaką można uzyskać z obliczeń kombinatorycznych. Przekonali się o tym z pewnością ci, którym rozsypała się kostka Rubika. Jeśli ją źle złożyć (to znaczy wprowadzić stan nieprawidłowy), to nie da się uzyskać takiego rozwiązania, w którym każda ściana kostki będzie miała jeden kolor. Byłoby rzeczą interesującą udowodnić, że w grach bardziej złożonych, na przykład szachach, istnieją stany nieprawidłowe, to znaczy takie, do których nie można dojść z położenia wejściowego przesuwając figury zgodnie z regułami gry. Świadczą o tym chyba niektóre bardzo dziwne sytuacje na szachownicy publikowane w niektórych czasopismach jako zadania szachowe.

Drugi etap budowy programu do rozwiązywania gry za pomocą komputera polega na wprowadzeniu strategii gry, to znaczy zasady postępowania, która doprowadzi do rozwiązania zagadki lub rozwiązania gry.

Rozpatrzmy to na przykładzie gry w "8". Strategią, którą można byłoby nazwać "brakiem strategii" jest działanie w sposób zupełnie przypadkowy, a więc traktowanie każdego ruchu jako równorzędnego. Jeśli w stanie wejściowym przedstawionym na rys. 1 każemy w projektowanym algorytmie losować pierwszy ruch do pustej klatki cyfrą 5, 6 lub 7 z jednakowym prawdopodobieństwem, z następnym porównaniem nowej konfiguracji z docelową to jest to początek wspomnianego "braku strategii". Jeśli wszystkie ruchy będziemy wykonywać w taki właśnie sposób, to za którymś z kolei losowaniem i wykonaniem ruchu uda się wreszcie osiągnąć konfigurację docelową. Będzie to jednak algorytm bardzo nieefektywny, te same ruchy będą często powtarzane i czas rozwiązania zadania będzie bardzo długi.

Na tym prostym przykładzie widzimy, że możemy zbudować nieefektywny algorytm rozwiązywania gry w "8", to znaczy działający w sposób przypadkowy. Nazwijmy taki algorytm nieinteligentnym. Algorytm inteligentny to taki, który będzie działał racjonalnie. Racjonalizacja postępowania będzie tu polegać na tym aby:

1. Niedopuszczać do powtórnego wchodzenia do tych samych gałęzi. W tym celu trzeba likwidować już wykonywane wybory, czyli ograniczać liczbę wariantów wyboru ruchów w rozgałęzieniach.
2. Po wejściu do niewłaściwej gałęzi wykonywać powrót do rozgałęzienia wyjściowego i wchodzić do istniejących (niezlikwidowanych) gałęzi.

Wynikiem postępowania zgodnie z opracowanym algorytmem będzie ustalenie kolejności ruchów cyframi jakie należy wykonać aby rozwiązać zadanie.

Jeśli zamiast rozpatrywania kolejności stanów gry, rozpatrzymy samą drogę jako kolejność ruchów cyframi, to znaczy opiszemy ją jako kolejność działań, które zmieniają stany, to wynik będzie prawie równoważny programowi współczesnego komputera, to znaczy równoważny przedstawieniu problemu przez opisanie metody jego rozwiązania za pośrednictwem języka typu proceduralnego. Innymi słowy proceduralne przedstawienie problemu jest równoważne jego przestawieniu, otrzymywanemu jako wynik procesu rozwiązania problemu wyszukiwania. Różnice tych dwu metod rozwiązania problemu, to jest tradycyjnej metody programowania algorytmu i metody wyszukiwania, pojawiają się w następującej postaci. W tradycyjnych metodach programowania algorytmem otrzymania rozwiązania jest opis problemu i rozwiązanie jego całości zależy od samego problemu jako takiego. Stąd wynika, że jeśli w rozpatrywanym przykładzie zaszły jakiegokolwiek zmiany, to znaczy zaszły zmiany w stanie początkowym, stanom terminalowym, grupach stanów, zasadach przejść, warunkami ocen, to program rozwiązania powinien być napisany od nowa. Z drugiej strony według metody wyszukiwania problem przedstawiony jest nie proceduralnie i jedna i ta sama metoda może być użyta w innym zadaniu do pewnego zakresu problemów, to znaczy do problemów z określoną formą przedstawienia stanów lub idąc dalej do zakresu problemów posiadających zadane struktury danych zasady przejść i kryteriów oceny stanów. Jeśli nawet warunki budowy problemu zmieniają się w tym zakresie, to wszystko jedno

metoda rozwiązania problemu pozostanie taka sama, a zgodnie ze zmianą problemu zmieni się samo rozwiązanie.

Omówiony w tym przykładzie uproszczony sposób rozwiązywania problemów oparty na wyszukiwaniu, był charakterystyczny dla wczesnego etapu rozwoju sztucznej inteligencji. Wszystkie stany zwykle były znane uprzednio i można było wybrać drogę poszukiwania wychodząc z zadanego stanu początkowego. Było to możliwe ze względu na stosunkowo niewielkie liczby stanów. Stany te można było określić w sposób sformalizowany i jednocześnie dostępny do zrozumienia.

Jak widać z powyższego przykładu tego rodzaju proste problemy są dalekie od tych jakie zwykle spotyka się w życiu. Już była mowa, że jeśli liczba stanów rośnie, to odpowiednio rośnie czas niezbędny do otrzymania rozwiązania. W celu przewycięzenia tego stosuje się metodę zadawania oceny każdemu stanowi, o którym była mowa uprzednio w odniesieniu do problemu gry. Jednakże to podejście nie zawsze jest uniwersalne dlatego, że taką funkcję oceniającą wprowadza się w warunkach, jednorodności przedstawienia stanów i możliwość badania jednej i tej samej funkcji dla wszystkich stanów. Takie warunki wskazują na to, że krąg problemów jest ograniczony do problemów sformalizowanych. I rzecz nie polega na tym, że to nieuchronnie prowadzi do uniwersalnych algorytmów, które wykorzystuje się do rozwiązywania w pełni sformalizowanych problemów. Faktycznie w wielu przypadkach dla takiego rodzaju problemów stosuje się rozmaite metody. Na przykład gra w szachy. Ta gra w istocie polega również na poszukiwaniu, jednak tu przy określaniu sposobu następnego ruchu wykorzystuje się zasady jako metodę zawężenia obszaru poszukiwania. Zawężenie obszaru poszukiwania oznacza, że istnieje informacja równoważna zasadom, która określa wszystkie warianty poszukiwania i warunki ich sprawdzenia. Ściślej mówiąc przy nieprawidłowych ruchach, może powstać sytuacja taka, że informacja wynikająca z rozłożenia figur na desce nie jest odpowiednia do zasad, lecz jest to zupełnie inny problem, tu nie rozpatrywany. Jest jednak jasne, że zasady gry same jako takie nie są realizacją procesu poszukiwania. Przedstawiają one sobą przechowywaną w formie nieproceduralnej informację, a nie poszukiwanie jako takie. Jednakże w innej terminologii zasady uważa się jako wiadomości (składające się na wiedzę). W szczególności w programie komputerowym stworzonym do gry w szachy, zasady gry są przechowywane w określonej formie przedstawienia wykorzystywanej przez program. Dlatego w sytuacji,

kiedy zasady wykorzystuje się w celu wyboru ruchów, staje się rzeczywiście zrozumiałą sens definicji wiedzy podanej na początku książki jako "przechowywanej w komputerze informacji, sformalizowanej zgodnie z określonymi zasadami strukturalnymi, którą komputer może wykorzystać autonomicznie podczas rozwiązywania problemów według takich algorytmów jako wnioskowanie logiczne". Świadczy to o tym, że wiedza (wiadomości) odgrywa nadzwyczajną rolę w rozwiązywaniu złożonych problemów. Podobną sytuację obserwuje się w wielu innych złożonych grach. Aby stać się lepszym graczem, trzeba opanować dużą liczbę zasad gry, to znaczy wiedzy.

Chociaż do osiągnięcia pewnego celu gry konieczne są wiadomości (wiedza) i techniczne mistrzostwo, przestrzeń problemowa tych gier jest ograniczona istniejącymi na planszy pozycjami, gry stają się problemami zamkniętymi z prostą strukturą. Ponieważ gra opisywana jest przez konfigurację figur na planszy, liczba takich sytuacji będzie wzrastać w sposób kombinatoryczny, jednakże przedstawienie każdej sytuacji zachowuje regularność, którą wykorzystuje się do przedstawienia wiedzy. Innymi słowami, przy formalizacji gry w koncepcjach rozwiązania problemu najpierw ustala się przestrzeń stanów, a następnie w tej przestrzeni ustala się położenie konkretnego każdego stanu. W koncepcji problemu gry wchodzi zasady ruchów, sposoby rozmieszczenia i sposoby przesuwania figur, opisywane pewną liczbą parametrów. Każdy oddzielny stan w grze jest kombinacją jednych i tych samych parametrów, zatem wszystkie stany opisywane są w jednolitej formie przedstawienia (struktura danych). Oznacza to, że "przedstawienie każdego stanu zachowuje regularność".

Jasne jest, że wiedza (wiadomości) odgrywają ważną rolę w rozwiązywaniu rozmaitych problemów powstających w realnym świecie, i ta ich ważność nie ma żadnego porównania z grami. Problemy rzeczywiste zwykle są bardzo złożone i nie posiadają takich jasno określonych ram jak w grach. Zatem, w wielu sytuacjach nie można dostatecznie jasno określić samego pojęcia stanu, ponieważ określenie to nie jest takie proste jak w grach. Powstaje problem jak najlepiej przedstawić wiedzę w podobnych okolicznościach? Problem ten stał się w połowie lat 60-tych ważnym tematem badań w dziedzinie sztucznej inteligencji. Ostatnio wielu badaczy zaproponowało najrozmaitsze sposoby przedstawienia wiedzy. Jednakże zagadnienie jakim powinno być przedstawienie wiedzy na razie jest otwarte. Wyjaśnienie tego jest jednym z zadań niniejszej

pracy, lecz przed tym omówimy krótko istniejącą obecnie sytuację w dziedzinie przedstawiania wiedzy, w której dziś znajdują się podstawowe sfery użytkowe sztucznej inteligencji.

#### 2.4. Inżynieria wiedzy

Inżynieria wiedzy to zbiór technicznych metod tworzenia systemów, przeznaczonych do rozwiązywania problemów, z wykorzystaniem zbiorów wiadomości z określonych dziedzin przedmiotowych<sup>19)</sup>. W odniesieniu do systemów ekspertowych to ogół metod, technik i technologii wykorzystywanych przy ich budowaniu i użytkowaniu.

Wiedza wykorzystywana w komputerowych systemach intelektualnych jak to przedstawiono we wstępie, jest informacją odwzorowującą problemy rzeczywiste z pewnymi ograniczeniami. Z punktu widzenia praktycznej użyteczności konieczne jest aby wiedza -jako zbiór wiadomości - posiadała taką formę, która umożliwiała by pewną swobodę osiągania postawionego celu. Pytanie, w jakim właśnie stopniu dopuszczalna jest ta swoboda, lub mówiąc bardziej jasno, jakim warunkom powinny odpowiadać wiadomości, włączając tu również możliwości ich opisu, posiada różne odpowiedzi w zależności od obszaru zastosowania. Na przykład w sferze zastosowań technicznych wykorzystuje się jak najbardziej różnorodne środowisko przedstawienia i oprócz opisu językowego wykorzystuje się także rysunki, wzory matematyczne itp.

W zastosowaniach technicznych w przedstawieniach językowych wykorzystuje względnie proste formalizmy. Związane jest to ze specyficznym charakterem obszaru techniki, ponieważ wszystko w niej jest określone faktami i obiektywną rzeczywistością. Jednakże, jeśli dana metoda opisu wprowadza się w tych obszarach stosowanych, gdzie w większej części nacisk kładzie się na językowe przedstawienie, na przykład, w takiej dziedzinie jak prawodawstwo, to ze wszech miar potrzebna jest znacznie większa zdolność opisu. Dalej, opis językowy, wymagany w obszarach stosowanych informacji (włączając w to zarówno

---

<sup>19)</sup> Osuga. S., Przetwarzanie wiedzy. Wyd. "Omsja", Tokio, 1986. (praca przetłumaczona na j. Rosyjski w roku 1989, wyd. "MIR")

język w szerokim jego pojęciu jak i grafikę), będziemy nazywać językiem przedstawienia wiedzy. W celu wykorzystania takiej informacji w postaci wiedzy potrzebne są funkcje intelektualne, przewyższające możliwości współczesnych komputerów. Wszystko to - przedstawienie wiedzy i ich wykorzystanie, rozpatrywane w odniesieniu do konkretnego obszaru stosowanego, jest przedmiotem inżynierii wiedzy.

#### 2.4.1. Proces rozwoju inżynierii wiedzy

Inżynieria wiedzy jako technologia użytkowania wiedzy wyszła ze źródeł sztucznej inteligencji i zaczęła się intensywnie rozwijać w ostatnich latach. Istotą sztucznej inteligencji jest naukowa analiza i automatyzacja funkcji intelektualnych człowieka. Jednakże dla większości problemów, z wyjątkiem takich względnie łatwo poddających się algorytmizacji jak omówione uprzednio problemy wyszukiwania (poszukiwanie), bardzo złożoną rzeczywistością jest trudność ich realizacji maszynowej.

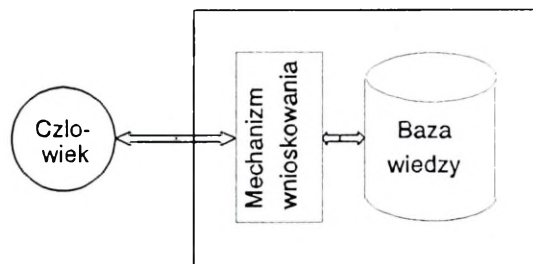
Wszystkie możliwe badania w dziedzinie sztucznej inteligencji pozwoliły nieodwracalnie upewnić się w poglądzie, że prawdziwie niezbędną dla rozwiązania problemów jest wiedza ekspertów. Zatem jeśli stworzyć system, zdolny do zapamiętania i wykorzystania wiedzy ekspertów, to znajdzie on zastosowanie w działalności praktycznej. Kiedy badacze w dziedzinie sztucznej inteligencji rzeczywiście zbudowali takiego rodzaju systemy, wszystkie te poglądy stały się usprawiedliwione. Takie systemy jak DENDRAL, MYCIN i im podobne, natychmiast stały się znane na całym świecie. Systemy te gromadzą w pamięci komputera wiedzę ekspertów i wykorzystują tę wiedzę do rozwiązywania problemów, wyciągając ją w razie konieczności z pamięci, to otrzymały one nazwę ekspertowych, a profesor E. Feigenbaum, jeden z twórców systemów ekspertowych, wprowadził dla tego obszaru zastosowań technicznych termin "inżynieria wiedzy". Wszystkie metody i podejścia inżynierii wiedzy znalazły ogólne uznanie i poparcie badaczy. Następnie w USA opracowano dużą liczbę systemów ekspertowych, które znalazły bardzo szerokie zastosowanie w innych krajach.

Chociaż jest tu mowa o systemach ekspertowych, jako o czymś jednolitym, to posiadają one w konkretnych realizacjach, jak będzie pokazane dalej na kilku przykładach, bardzo różnorodne formy. Tym nie

mniej wszystkie systemy ekspertowe mają wspólną cechę: przetwarzanie informacji na komputerze wykonywane jest zupełnie inaczej aniżeli tradycyjne przetwarzanie Turinga - von Neumanna. Jest rzeczą naturalną, że z tymi systemami wiąże się duże nadzieje jako na środki wspomaganie działalności intelektualnej, włączając tu także i projektowanie. Obecnie istnieje jednak wiele problemów jeszcze nie rozwiązywanych przez systemy ekspertowe na poziomie praktycznej realizacji. Stąd spróbujemy omówić podstawowe zasady systemów ekspertowych, pozostawiając inne problemy do rozpatrzenia w dalszej treści.

#### 2.4.2. Zasady pracy systemu ekspertowego

System ekspertowy jak pokazano na rysunku 3 jest to system "człowiek - maszyna", zbudowany z takich składników jak baza wiedzy - zbiór wiedzy (wiadomości), opisanych z wykorzystaniem wybranej formy i przedstawienia - i mechanizm wnioskowania, który zapewnia manipulację wiedzą podczas rozwiązywania problemów użytkowych.



Rys.3. System ekspertowy jako układ "człowiek-maszyna".

W systemach ekspertowych nie tylko wybrano określony sposób do przedstawienia wiedzy. Jak będzie powiedziane dalej, stosuje się w nich najrozmaitsze podejścia i na razie jeszcze nie wiadomo, które z nich są najlepsze. Sytuacja taka wytworzyła się dlatego, że nasze doświadczenia w użytkowaniu wiedzy jako informacji są bardzo płytkie i nie wystarczy sformułować poglądu o

tym, jakim warunkom powinno odpowiadać przedstawienie wiedzy. Zagadnienie to będzie omawiane dalej, a najpierw rozpatrzmy zasady pracy systemu ekspertowego na przykładzie reguł produkcji - najprostszego sposobu przedstawienia wiedzy. W takim systemie ekspertowym całą wiedzę przedstawia się za pomocą opisów typu "JEŚLI - TO". Część reguły "JEŚLI" nazywa się założeniem, a "TO" - wnioskiem lub działaniem. W ogólnej postaci regułę będziemy zapisywać jako "JEŚLI A1, A2, ..., An, TO B". Taki zapis oznacza, że jeśli "wszystkie warunki od A1 do An są prawdziwe, to należy wykonać działanie B".

Rozpatrzmy regułę:

JEŚLI (1) y jest ojcem x (A1)  
(2) z jest bratem y (A2) (2.1)  
TO z jest wujkiem x (B)

W tym przypadku,  $n = 2$ ; przypadek kiedy  $n = 0$  opisuje zjawisko, składające się tylko z wniosku, to znaczy prosty fakt. Przykładem takiej wiedzy jest fakt "ciężar atomowy żelaza wynosi 55,847".

Mechanizm wnioskowania łączy wiedzę pojedynczo, a następnie wyprowadza z kolejnych wiadomości wniosek. Załóżmy na przykład, że w bazie wiedzy razem z powyższymi opisami wiedzy istnieją i takie wiadomości:

JEŚLI (1) z jest ojcem x  
(2) z jest ojcem y (2.2)  
(3) x i y nie są jedną  
i tą samą osobą  
TO x i y są braćmi

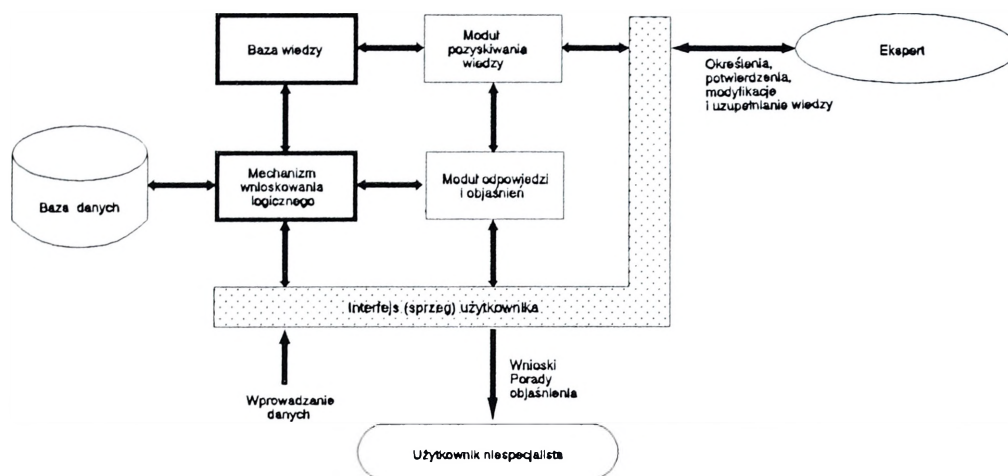
Kuba jest ojcem Jasia (2.3a)

Kuba jest ojcem Maciusia (2.3b)

Jaś jest ojcem Piotrusia (2.3c)

Tutaj  $x, y, z$  - zmienne. Taki zapis pokazuje, że przy warunku podstawienia jednakowych wartości do jednakowych zmiennych, będą one uogólnieniem dla różnych wiadomości. W regule (2.1) i regule (2.2) wykorzystuje się jednakowe symbole zmiennych  $x, y, z$  jednakże jakikolwiek związek między tymi zmiennymi nie istnieje, ponieważ wszystkie wliczone wiadomości są niezależne od siebie.

Z tych wiadomości można wyciągnąć wniosek formalny, że "Maciuś jest wujkiem Piotrusia". To znaczy w  $z, x, y$  reguł (2.2) podstawia się odpowiednio "Kuba", "Jaś", "Maciuś", z otrzymanego wyniku i twierdzeń (2.3a) i (2.3b) wyprowadza się wniosek "Jaś i Maciuś są braćmi". Podstawia się następnie do  $x, y, z$  zasady (2.1) odpowiednio "Jaś", "Piotruś", "Kuba", można wyprowadzić wniosek (2.3c) i otrzymany wcześniej wynik.



Rys.4. Uproszczony schemat systemu ekspertowego.

Ta w pełni sformalizowana procedura, wykorzystująca takie metody jak dopasowanie wzorców (przy których ustala się, czy są zgodne ze sobą dwie formy przedstawiania, włączając tu otrzymanie takiego podstawienia zmiennych, przy których formy są zgodne), Poszukiwanie w bazie danych, powrót (powrót do stanu wyjściowego przy nieudanej próbie rozwiązywania), reprezentuje właśnie mechanizm wnioskowania. W chwili obecnej mechanizmy wnioskowania realizowane są środkami programowymi, jednakże japończycy w swojej koncepcji komputerów piątej generacji przewidują ich maszynową realizację.

Planują oni w pewien określony sposób zadać formalizmu przedstawienia wiedzy i zbudować określony na bazie tego przedstawienia mechanizm wnioskowania, analogiczny do powyższego, to przez sformułowanie wiedzy w formie dopuszczalnej w takim przedstawieniu co ma umożliwić wyprowadzanie wniosków w sposób automatyczny. Wszystko to różni się od tradycyjnych procedur (programów) i standardowych danych (baz danych) i pozwala realizować w komputerze przedstawienie, które można wykorzystywać bardziej elastycznie aniżeli dotychczasowe procedury. Faktycznie jeśli dodawać nowe wiadomości to z reguły (jeśli tylko w systemie są kolejno wprowadzone zasady przetwarzania wiedzy) te wiadomości będą wykorzystywane dokładnie w taki sam sposób jak do momentu ich wprowadzenia. Na przykład jeśli w rozpatrzonym przykładzie do bazy wiedzy wprowadzić nowe twierdzenie "Kuba jest ojcem Bonifacego", to od razu będą wprowadzone takie

wnioski jak "Jaś i Bonifacy są braćmi" i "Bonifacy jest wujkiem Piotrusia".

W ten sposób, wnioski (a dokładniej wnioski dedukcyjne) są określone jako sposób użycia bazy wiedzy, jednakże wszystko co powiedziano wyżej odnosi się i do wniosków, bazujących na regułach produkcji, i jeśli przedstawienie wiedzy będzie posiadać inną formę to będzie i inna forma wniosków. Wychodząc z tej okoliczności, zaznajomimy się z charakterystycznymi sposobami przedstawienia wiedzy, wykorzystywanymi w chwili obecnej i z wnioskami wynikającymi z tych sposobów.

## **2.5. Przedstawienie wiedzy I wnioski w systemach ekspertowych**

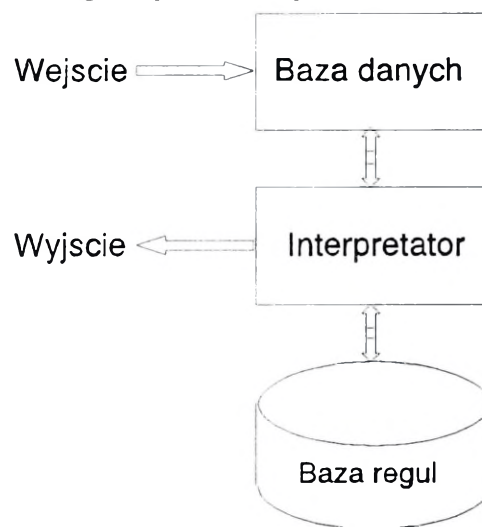
Charakterystyczną cechą systemów ekspertowych, wyróżniającą je spośród tradycyjnych systemów przetwarzania informacji, jest wykorzystanie nowego rodzaju informacji, nazywanego wiedzą. Formalizmu opisu takiego rodzaju wiedzy określa się jako przedstawienie wiedzy a składnikiem, który wykorzystuje do rozwiązania problemu wiedzę ekspertów, opisaną we wcześniej wybranych dla niej formach przedstawienia, jest mechanizm wnioskowania. W systemach z bazami wiedzy, w tej liczbie i systemów ekspertowych, przedstawienie wiedzy jest pojęciem fundamentalnym, a decyzja o wyborze sposobu przedstawienia wiedzy ma ogromny wpływ na każdą ich część składową. Można nawet powiedzieć, że przez sposób przedstawienia wiedzy określa się możliwości systemu bazy wiedzy. I na odwrót, aby system przetwarzania wiedzy odpowiadał określonym potrzebom użytkowym, powinno się stworzyć odpowiednie przedstawienie wiedzy. Jednakże w chwili obecnej nie można jeszcze powiedzieć, że wymagania w odniesieniu do przedstawiania wiedzy z tego punktu widzenia sformułowane są niesprzecznie. Możliwość wykorzystania wiedzy powstała równocześnie z momentem pojawienia się mechanizmów wnioskowania, a mechanizm wnioskowania był określony w swoim czasie razem ze sformulowaniem sposobów przedstawienia wiedzy. Ponieważ przedstawienie wiedzy jest środkiem opisu wiedzy człowieka, to pożądane jest aby jego możliwości opisu były możliwie największe; z drugiej strony jeśli forma przedstawienia jest zbyt skomplikowana, to komplikuje się mechanizm wnioskowania, przy tym

znacznie trudniejsze staje się projektowanie systemu ekspertowego i powstaje niebezpieczeństwo utraty wiarygodności wykonywanych przez niego działań. W ostatecznym rozrachunku projektowanie przedstawienia wiedzy przewiduje wypracowanie wszystkich tych warunków, a następnie i wybór rozwiązania na podstawie pewnego kompromisu między nimi. Problem ten będzie jeszcze omawiany w rozdz. 5, tu tylko krótko zaznajomimy się z niektórymi charakterystycznymi sposobami przedstawienia wiedzy, wykorzystywanymi we współczesnych systemach ekspertowych.

### 2.5.1. Reguły produkcji

#### 2.5.1.1. Przedstawienie wiedzy za pomocą reguł produkcji

Reguły produkcji, jak już było powiedziane, opisują wiedzę w postaci JEŚLI - TO. Prostota i obrazowość tego sposobu spowodowała, że zastosowano go w wielu systemach. Systemy przetwarzania, wykorzystujące przedstawienie wiedzy za pomocą reguł produkcji, otrzymały nazwę systemów produkcji. W skład systemu produkcji wchodzi baza reguł, globalna baza danych i interpretator reguł. Baza reguł - to obszar pamięci, w który utrzymuje się baza wiedzy - zbiór wiadomości, przedstawionych w postaci JEŚLI - TO; globalna baza danych - to obszar pamięci zawierający faktyczne dane (fakty), które opisują wprowadzane dane i stany systemu. Bazy danych w różnych systemach mają różną formę, jednakże wszystkie one mogą być opisane jako grupa danych, zawierających nazwy danych, atrybuty i wartości atrybutów. Interpretator jest to mechanizm wyprowadzania i jest



Rys.5. Rola Interpretatora w systemie ekspertowym

on tym składnikiem systemu, który formułuje wnioski, wykorzystując bazę reguł i bazę danych (rys.5).

Na rys. 6 pokazano przykład przedstawienia wiedzy w MYCIN - systemie ekspertowym, wykorzystującym reguły produkcji. System MYCIN jest napisany w Lisp'ie. Na rysunku tym uwidoczniło w dole - wewnętrzne przedstawienie reguły, a na górze - wyjaśniające jego zewnętrzne przedstawienie. Charakterystyczną właściwością przedstawienia wiedzy w MYCIN jest istnienie narzędzia pozwalającego bez szczególnych trudności wykorzystywać w danych i zasadach informacje rozmyte, za pomocą wyliczenia oceny, nazywanej współczynnikiem wiarygodności.

#### 2.5.1.2. Wnioski oparte na regułach produkcji

Mechanizm wnioskowania realizowany we współczesnych systemach produkcji, w zasadzie nie jest skomplikowany. Posiada on funkcje wyszukiwania w bazie wiedzy, kolejnego wykonywania operacji nad wiadomościami i otrzymywaniem wniosków; przy tym istnieją dwa sposoby przeprowadzenia takich wniosków - wnioskowanie bezpośrednie i odwrotne. We wnioskowaniu bezpośrednim wykonuje się przechodzenie do postawionego celu z kolejnym użyciem reguł do danych (faktów), które przyjmuje się jako punkt wyjścia. We wnioskowaniu bezpośrednim wybiera się jeden z elementów zawartych w bazie danych i jeśli przy konfrontacji element ten jest zgodny z założeniem reguły, przeważnie wyprowadza się odpowiedni wniosek i umieszcza się go w bazie danych lub wykorzystuje się działanie, określone regułą i w odpowiedni sposób zmienia się zawartość bazy danych. Często takie wnioski określa się jako wnioski sterowane danymi lub wnioski wstępującymi, kiedy kolejno wyprowadza się nowe wyniki, rozpoczynając z już znanych danych. Wnioski, przy których proces przesuwa się w kierunku od postawionego celu do punktu wyjścia, są odwrotnymi. Nazywane są również zstępującymi lub wnioskami zorientowanymi na cel. Proces wnioskowania zstępującego rozpoczyna się od postawienia celu. Jeśli ten cel jest zgodny z konsekwentem (wnioskiem) reguły, to założenie reguły jest przyjmowane za podcel, lub hipotezę, i ten proces prowadzi się do tej pory

dopóki nie otrzyma się uzgodnienia podceli lub hipotez z otrzymanymi danymi. Nie można jednoznacznie odpowiedzieć na pytanie, który z tych sposobów jest lepszy, ponieważ zależy to od tego problemu, dla którego jest on wykorzystywany. W systemach, którym stawia się wymaganie wysokiej uniwersalności, konieczne jest istnienie obydwu tych sposobów wnioskowania. W MYCIN i EMYCIN wykorzystuje się wnioskowanie odwrotne, w OPS-5 - bezpośrednie.

Zwykle w wiedzy (wiadomościach) przedstawianych za pomocą reguł proceduralnych jednocześnie występuje kilka warunków w założeniu, jednakże istnieje jedyny wniosek (jeśli wniosków jest więcej jak 2, to wiadomość zwykle rozdziela się i przytacza ją w postaci z jednym "TO"). Wymagania do mechanizmów wnioskowania wynikają z warunku dopasowania wiedzy, wykorzystywanej dla wnioskowania, przy tym przetwarzanie staje się szczególnie łatwe, jeśli dla wyciągnięcia wniosku wystarczy jednokrotne dopasowanie. W tym sensie wnioski zstępujące są prostymi, jednakże istnieją takie zadania, dla których nie jest tak prosto dać wyraźne opisanie celu. Na przykład, w systemach diagnostyki medycznej, kiedy stawia się diagnozę według symptomów, trudno otrzymać wniosek zstępujący, jeśli cel jest nieznan.

#### 2.5.1.3. Przykłady systemów produkcyjnych

Systemy produkcyjne w zasadzie są bardzo proste. Powodzenie pierwszych systemów ekspertowych DENDRAL I MYCIN, wykorzystujących reguły produkcji, wniosło olbrzymi wkład w tworzenie inżynierii wiedzy. Na przykładzie tych systemów zaznajomimy się z systemami produkcyjnymi.

(1) DENDRAL. Badania nad systemem ekspertowym DENDRAL rozpoczęły się dość wcześnie - w 1965 r., i ten system można nazwać pionierem inżynierii wiedzy. Pierwszoplanowym celem jego tworzenia była ocena struktury nieznanymi związków chemicznych na podstawie wyników eksperymentów w dziedzinie spektroskopii (później uwzględniono również wyniki jądrowego rezonansu magnetycznego). DENDRAL realizuje procedurę generacji wszystkich chemicznych struktur - kandydatów, które potencjalnie łączą się z otrzymanymi danymi wejściowymi i wyboru ich ze struktury, dokładnie dobranej do tych danych. W takim

procesie z wykorzystaniem wiedzy ekspertów, możliwe jest wykluczanie na podstawie analizy nieodpowiednich struktur i szybkie otrzymywanie właściwych wyników. Przy tym postawiony cel może być osiągnięty w sposób tym bardziej efektywny, im wiedza którą rozporządza system będzie miała pokaźniejszą objętość. Z tego powodu w stadium przygotowania bazy wiedzy DENDRAL wykorzystuje się system METADENDRAL, który pozwala wprowadzać do bazy wiedzy wiadomości otrzymywane przy analizie danych. System DENDRAL był wysoko oceniony z punktu widzenia przydatności do tworzenia inżynierii wiedzy, a przyjęta w nim metoda jednoczesnego wykonywania operacji generacji-testowania jest identyczna do algorytmów, wykonywanych przez człowieka podczas projektowania. Jest bardzo ważny jako sposób tworzenia systemów, których celem jest wykorzystywanie wiedzy.

(2) MYCIN i jego analogi. W pierwszych systemach ekspertowych przeważały systemy diagnostyki medycznej. Wynika to stąd, że wiedza z zakresu diagnostyki klinicznej jest w znacznym stopniu zdeterminowana osądami praktycznego doświadczenia lekarzy i chociaż wiadomości te są typu proceduralnego, to nie są one jednak opisywane metodami formalnymi stąd niezbędny sposób ich przedstawienia, są bliższe przedstawianiu w języku naturalnym. Uwzględniając to, diagnostykę medyczną rozpatrywano jako obiekt bardzo nadający się do opracowania w formie systemu ekspertowego.

MYCIN jest charakterystycznym przykładem, systemu, wykorzystującym reguły produkcji. Przedstawienie wiedzy w MYCIN jest proste pod względem formy i łatwo zrozumiałe, jednak opisanie wszystkich wyrażeń językowych w formie JEŚLI - TO nie zawsze jest proste. Poglądowość tej formy przedstawienia wiedzy w MYCIN znalazła odbicie u wielu następców, jednakże i to stało się jego ograniczeniem. Nie tylko w MYCIN lecz w wielu innych systemach nie zawsze jest łatwo wyrazić w jednej i tej samej formie wiedzę, którą posiada człowiek. Zwykle zarówno w systemach z przetwarzaniem takiej wiedzy jak i w dziedzinie zawodowej, gdzie stosuje się te systemy, niezbędny jest pośrednik, dostatecznie dobrze rozumiejący wszystkie strony, jednak pozyskać albo przygotować takiego pośrednika jest trudno. To właśnie było przyczyną pojawienia się systemu, który otrzymał nazwę TEIRESIAS, którego forma realizacji pochodziła od systemu MYCIN. TEIRESIAS jest systemem ekspertowym, który podtrzymuje przeniesienie wiedzy od eksperta do bazy wiedzy nawet przy braku pośrednika. System ten jest

przeznaczony do przysposobienia wiedzy w dialogu eksperta z systemem ekspertowym, wykrycia błędów w wiedzy i ich korekcji.

W systemie MYCIN wnioski realizowane są w formie zależnej od treści wiedzy w dziedzinie przedmiotowej, a w EMYCIN mechanizm wnioskowania jest już oddzielony od zawartości bazy wiedzy i jest składnikiem samodzielnym. System EMYCIN pracuje analogicznie jak MYCIN, jednak cechuje go uniwersalność, dzięki której jest on w określonym stopniu niezależny od dziedziny przedmiotowej. Zatem, jeśli wiedza z dziedziny przedmiotowej jest opisana w formie przyjętej w MYCIN, i wprowadzić ją do bazy wiedzy to można będzie utworzyć system ekspertowy, zorientowany na daną dziedzinę przedmiotową.

W ten sposób MYCIN wniósł swój istotny wkład w rozwój inżynierii wiedzy oraz sprzyjał utworzeniu szeregu nowych systemów, np. PUFF<sup>20)</sup>. W innym systemie ekspertowym PROSPECTOR zapisano wiedzę o pożytecznych kopalinach. Jest on przeznaczony do przetwarzania danych geologicznych - rozpoznania złóż geologicznych. Istnieją również programy uniwersalne pomocne we wspomaganie budowy systemów ekspertowych, które pozwoliły zbudować systemy od RITA do ROSIE, stanowiące jednolity z MYCIN szereg diagnostycznych systemów ekspertowych.

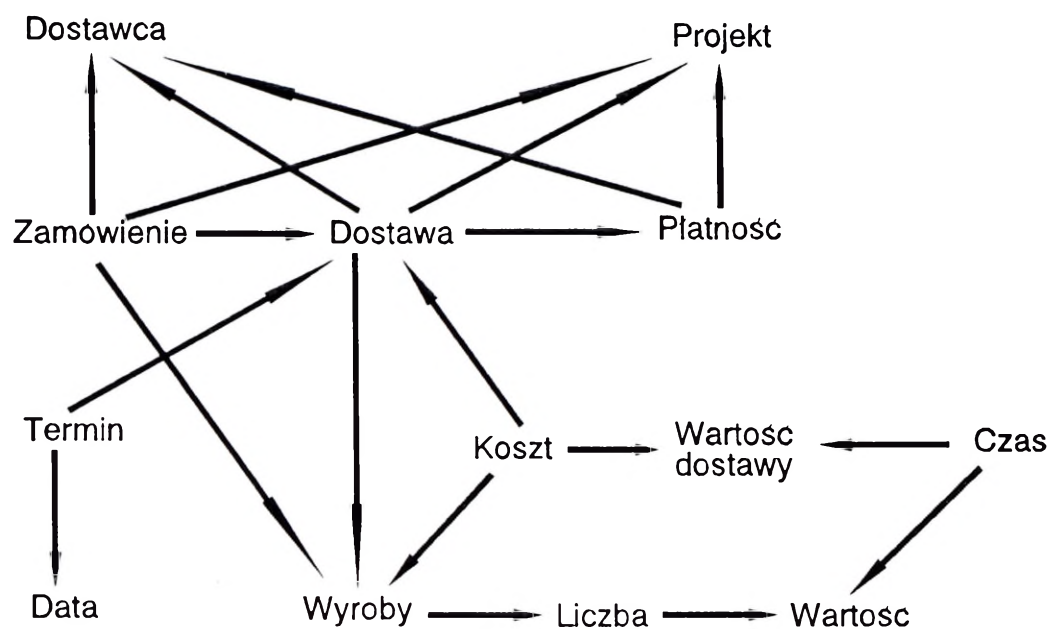
## 2.5.2. Sieci semantyczne

### 2.5.2.1. Przedstawienie wiedzy za pomocą sieci semantycznych

Wiadomościami można nazywać opisy związków między wyabstrahowanymi pojęciami i właściwościami odwzorowujących konkretne obiekty świata rzeczywistego. Pojęcia i związki między nimi można opisać za pomocą sieci składającej się z węzłów i luków. Węzły w takiej sieci wyrażają właściwości i pojęcia a luki są opisem ich wzajemnych zależności. Wszystkie węzły i luki mogą być wyposażone w etykiety,

---

<sup>20)</sup> System diagnostyczny chorób płuc.



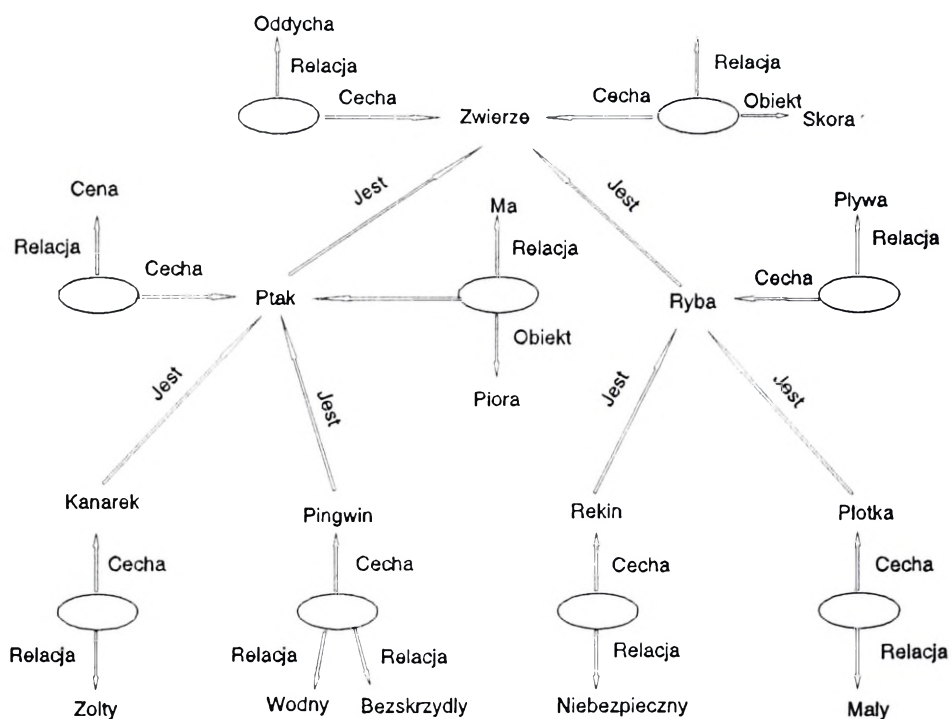
Rys. 6. Przykład sieci semantycznej.

które pokazują co właściwie one opisują. Tego rodzaju formalizmu opisu przedstawienia wiedzy otrzymał nazwę sieci semantycznej. Na rys. 2.8 przedstawiony został przykład takiej metody przedstawienia wiedzy. Rysunek ten opisuje sytuację, kiedy do pewnego projektu zamawia się części i trzeba za nie zapłacić dostawcy.

Nazwa "sieć semantyczna" odzwierciedla nie tylko możliwość otrzymywania określonego formalizmu, jak to ma miejsce w systemach produkcji, lecz obejmuje ona wszystko to, co opisuje się jak omówione wyżej sieci związków, właściwości i pojęć. Następnie metody realizacji konkretnego systemu dzielą się w zależności od tego, w jaki sposób wykorzystuje się w nich przedstawienie wiedzy.

Specyficzne właściwości sieci semantycznych zdecydowały o ich dużej popularności wśród badaczy jako środka wnioskowania a wśród struktur danych, które znalazły zastosowanie w komputerach, wykorzystuje się struktury takiego samego rodzaju jak sieci semantyczne. Trudno obecnie ustalić, kto pierwszy posłużył się sieciami, jednakże dobrze wiadomo, że były wykorzystywane one semantykę terminów do opisu związków conceptów (pojęć) kiedy w najwcześniejszym okresie badań

nad stosowaną sztuczną inteligencją sieci służyły do ściśle określonego celu - przedstawieniu semantyki albo wiedzy. Wychodząc z opisu semantyki jednych terminów za pomocą innych, przedstawiono je w postaci sieci związków konceptów z odpowiadającymi im terminami. Na rys.7 pokazano taki przykład. Kulingham postawił sobie zadanie sieciowego przedstawienia i zmniejszenia objętości wiedzy

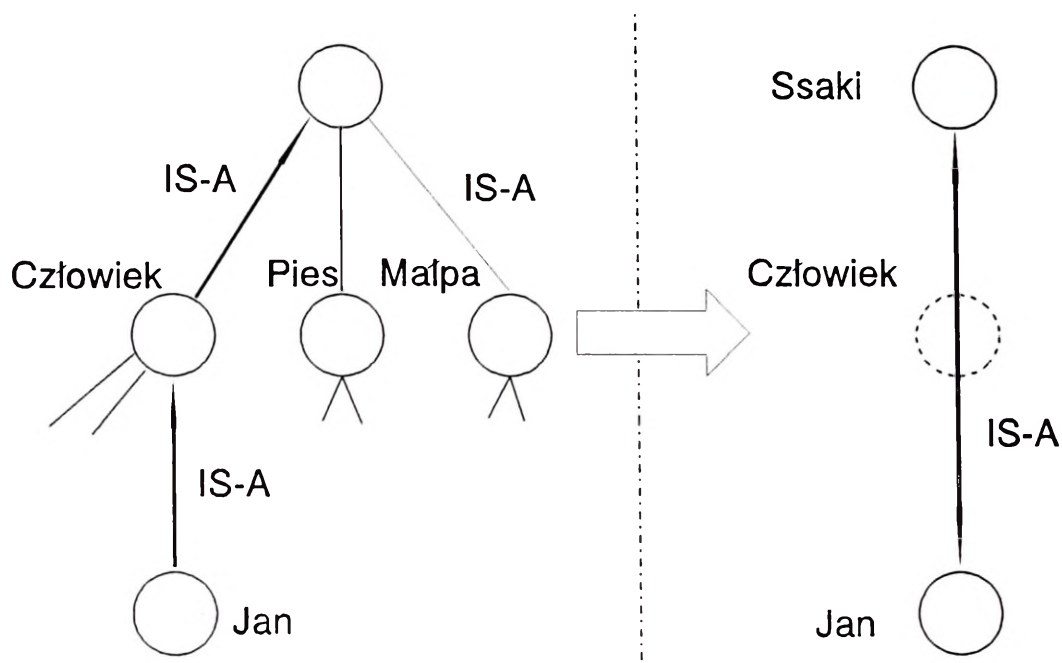


Rys. 7. Sieć semantyczna Kulinghama.

przy wykorzystaniu sieci w celu rozwiązywania problemów. Wychodząc z tego faktu, że jeśli w sieci zadany jest jeden koncept, to bezpośrednio otrzymuje się i inny, znajdujący się z nim w ścisłym związku, były przeprowadzone badania (głównie w dziedzinie psychologii) w celu zastosowania sieci semantycznych w charakterze modeli asocjacyjnych.

#### 2.5.2.2. Wnioski w sieciach semantycznych.

Równoległe z badaniem ogólnych właściwości sieci semantycznych prowadzono badania nad przedstawieniem wiedzy za pomocą sieci

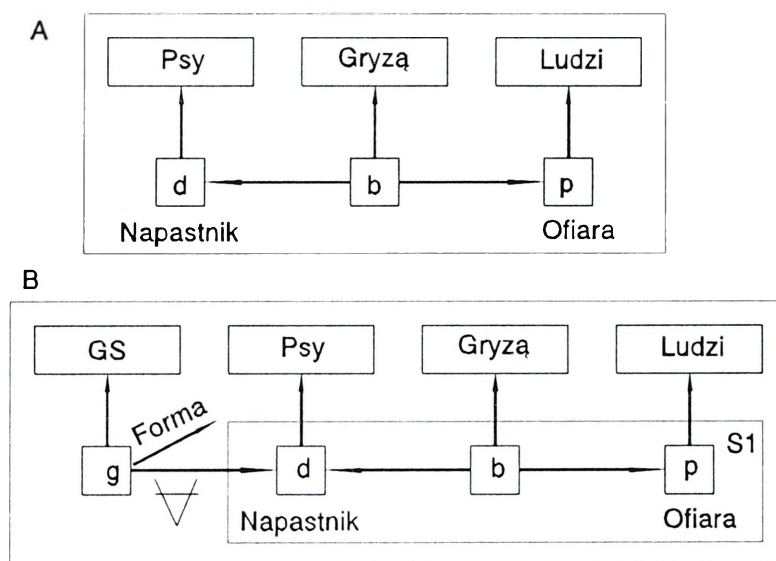


Rys. 8. Wniosek w sieci semantycznej.

semantycznych i budową sposobu wnioskowania dla tego przedstawienia. Wnioski z sieci semantycznych określa się za pomocą stosunków między zbiorem łuków posiadających wspólne węzły. Najprostszym rodzajem wniosku jest wniosek jaki można sformułować między zadanymi łukami. Na przykład, jak pokazano na rys. 8, kiedy zależnością "JEST" (IS-A), opisującą związki implikacji istoty i pojęć, wiąże się "Jan JEST człowiekiem" ("Jan IS-A człowiek") i "człowiek JEST ssakiem", (człowiek IS-A ssak) do ogniwa "Jan - człowiek - ssak", to z tego czyni się wniosek, że "Jan JEST ssakiem" (Jan IS-A ssak). W ostatnich czasach przeprowadzono badania mające na celu określenie wniosków o bardziej ogólnej postaci dla fragmentów sieci, jednakże nie udało się otrzymać zadowalających wyników z powodu trudności w unifikacji wniosków z sieci semantycznych. Prawdopodobnie dobry wynik pozwala otrzymać takie podejście, w którym wykorzystuje się kombinację sieci semantycznych z innymi metodami przedstawienia. Daje to możliwość pomyślnego wcielenia w życie wszystkich zalet przedstawienia wiedzy tylko za pomocą sieci semantycznych.

### 2.5.2.3. Dekompozycja opisywanego świata

Innym problemem sieci semantycznych jest przesunięcie grup wiedzy, odnoszących się do zasadniczo różnych sytuacji podczas definiowania luków między węzłami, co czyni sieć trudną do zrozumienia. Problem



A i B przedstawiają: "(pewien) pies ugryzł człowieka", "wszkie psy gryzą ludzi", S1 opisuje granice kwantyfikatora  $\forall$  w przestrzeni,

Rys. 9. Sieć Hendrixa.

ten jest istotny nie tylko dla sieci semantycznych lecz także dla systemów produkcji. W celu rozwiązania tego problemu zaproponowano takie metody jak dekompozycję zależną od rodzaju sieci (H. Hendrix, rys. 9), organizację hierarchiczną (R. Schubert) i inne. Jednakże w sieciach semantycznych istnieją pewne ograniczenia, z tego powodu trzeba się dobrać do sedna sprawy. Wychodząc z tej okoliczności, należy nie tylko dla sieci semantycznych lecz także dla wszystkich wykorzystywanych obecnie sposobów przedstawienia wiedzy, dokładnie badać problem ich stosowalności do rozwiązywania wielkich i złożonych problemów.

### 2.5.3. Logika predykatów

#### 2.5.3.1. Przedstawienie logiki predykatów

Logika predykatów od dawna była przedmiotem badania logiki jako nauki. W chwili obecnej jest to dziedzina nauki nazywana logiką matematyczną (ściślej mówiąc jest to rozdział logiki matematycznej), lub logiki symbolicznej. W szczególności logika predykatów była badana przez wielu matematyków, ponieważ stanowiła ona matematyczny fundament, stanowiący podstawę do formalnego opisu systemów.

Logika predykatów jako system jest konstruowana z języka, nazywanego językiem rachunku predykatów pierwszego rzędu. Twierdzenia przedstawione w terminach tego języka opisują stosunki leżące u podstaw logicznie pełnego systemu i reguł wnioskowania, które z zadanej grupy wyrażeń wyprowadzają zależności różniące się od wszystkich wyrażeń tej grupy. Podstawowym formalizmem przedstawienia za pomocą predykatów jest "term", ustanawiający odpowiedniość symboli znakowych do opisywanego obiektu oraz predykat do opisu zależności w postaci formuły relacyjnej obejmującej termy. Na przykład, Jan, Piotr - to termy. Kiedy między nimi istnieje zależność (związek?) "ojciec" i "dziecko", to stosunek ten opisuje się jako OJCIEC(Jan, Piotr). Predykat, którego wszystkie termy, są termami-konstantami, nazywa się wyrażeniem. Kiedy mówi się "predykat" to zwykle ma się na myśli, że wchodzi do niego term-zmienna. Na przykład takim predykatem jest OJCIEC(x,y). Jeśli w predykacie zapewniono możliwość użycia języka naturalnego to odpowiada on zdaniu w najprostszej jego formie, tak zwanemu zdaniu prostemu. Wychodząc z tej formy, zadaje się specjalne symbole i reguły syntaktyczne w celu określenia wzorów predykatowych odpowiadających zwykłemu zdaniu. Na przykład, takim zdaniem jest OJCIEC(Jan, Piotr) & MATKA(Maria, Piotr) (ojcem Piotr jest Jan i matką Piotr jest Maria lub (każdy człowiek x)[CZŁOWIEK(x) - > ŚMIERTELNY(x)] (jeśli podmiot jest człowiekiem, to on jest śmiertelny -

wszyscy ludzie umierają). Zwróćcie uwagę, że ostatni przykład jest podobny do reguły produkcji.

#### 2.5.3.2. Wnioskowanie w logice predykatów

Badania zastosowania logiki predykatów do przedstawienia wiedzy nabrały szerokiego rozmachu od roku 1965, kiedy J. Robinson zaproponował sposób wnioskowania maszynowego, nazywaną zasadą rezolucji. Po tym powszechne uznanie uzyskał pogląd, że z punktu widzenia efektywności, metody obróbki maszynowej zorientowane na określony obszar przedmiotowy i wykorzystujące właściwości samego problemu, są bardziej przydatne aniżeli metody ogólne. W ostatnich latach kiedy wypracowano języki programowania typu logicznego, na przykład Prolog, poglądy na logikę predykatów znacznie się zmieniły i od nowa bardzo aktywnie rozpoczęto badania w tej dziedzinie. Przyczyną tej aktywizacji były plany wykorzystanie Prologu jako podstawowego języka do tworzenia komputerów 5-tej generacji. Jak wiadomo idea komputerów 5-tej generacji powstała w Japonii w wyniku pojawienia się możliwości realizacji uniwersalnych metod przetwarzania informacji wynikającej z burzliwego rozwoju mikroprocesorów.

Przedstawienie wiedzy z wykorzystaniem logiki predykatów jest w niektórych sytuacjach zbieżne z regułami produkcji. Wynika to z tego, że w logice predykatów także rozwiązuje się wzory, które zadają odpowiednie reprezentacje w postaci JEŚLI-TO, dopuszczając kilka opisów w składzie warunku JEŚLI i wniosku TO, a wzory z podobnym uporządkowaniem opisu również są obiektem logiki predykatów. Istnieją jednak między tymi przedstawieniami także istotne różnice. W odniesieniu do przedstawienia wiedzy istotne są dwa przeciwstawne warunki: niezależność wiedzy jednej od drugiej i spójność wiedzy jako jednej całości. W regułach produkcji z przedstawieniem oddzielnych wiadomości związana jest pełna niezależność, a właściwości wiedzy jak jednej całości nie są rozpatrywane. Stwarza to możliwość dużych komplikacji z pozycji praktycznej realizacji sposobu przedstawienia wiedzy. W logice predykatów pod uwagę bierze się zachowanie właściwości wiedzy jako jednej całości. Tym właśnie uwarunkowane są różnice w systemach

przetwarzania wiedzy, które przedstawiają przechowywaną w nich informację dużej liczbie użytkowników.

Logika predykatów jako jeden z języków przedstawienia wiedzy, wykorzystywana jest również w charakterze teoretycznej podstawy do opisu samego systemu przedstawienia wiedzy.

#### 2.5.4. Model tablicy ogłoszeń

##### 2.5.4.1. Przedstawienie modelu tablicy ogłoszeń

W systemie produkcyjnym gromadzi się w bazie wiedzy reguły jednakowe pod względem formy, stąd jeśli dopuścić, że w jednym i tym samym systemie rozwiązywane są problemy z kilku różnych dziedzin praktycznych, to może dojść do pomieszania reguł z różnymi właściwościami, istniejącymi w tych dziedzinach, co będzie przyczyną istotnego obniżenia efektywności przetwarzania. Zatem operowanie różnymi problemami w jednym systemie produkcyjnym jest niepożądane. Bywa jednak w niektórych przypadkach tak, że jeden problem jest budowany z kilku problemów częściowych z różnymi właściwościami. Do tego problemy te nie są wspólne i zachowują niezależność. Na przykład, nie można rozdzielić problemów szczegółowych przy istnieniu takiego słabego związku, kiedy w wyniku rozwiązania jednego problemu szczegółowego inicjuje się początek rozwiązania problemu następnego.

Jednym z takich problemów jest problem rozpoznawania mowy. Wynikiem jego rozwiązania jest rozumienie mowy. Składa się z kilku różnych co do właściwości podproblemów, rozpoczynając z etapu wykrycia sygnałów przez detektory dźwięku aż do rozumienia na wysokim poziomie abstrakcji, kiedy kolejno dokonuje się wydzielania fonemów, rozpoznawania słów, budowy zdań i rozumienia zdania.

Istnieje system HEARSAY-II, opracowany w amerykańskim uniwersytecie Carnegie-Mellon i zorientowany na tego rodzaju problemy. Z charakterystycznych właściwości jego budowy, rozpatrywanych dalej, trzeba wymienić zastosowanie modelu tablicy ogłoszeń (Ścisłej rzecz biorąc, model tablicy ogłoszeń jest sposobem organizacji i wykorzystania,

a nie przedstawienia wiedzy.). W tym modelu dla każdego oddzielnego problemu, które razem stanowią jedną całość, istnieje odpowiedni zbiór danych, przy tym wszystkie zbiory przez wspólny roboczy obszar pamięci, nazywany tablicą ogłoszeń, są tak sterowane, że cała wiedza jest wykorzystywana spójnie, jako jedna całość. Takie oddzielne zbiory wiedzy nazywają się zwykle źródłami wiedzy (Z). Każdy Z sam jest budowany jako system produkcyjny.

#### 2.5.4.2. Wnioskowanie I przetwarzanie w modelu tablicy ogłoszeń

Na rys. 10 jako przykład modelu tablicy ogłoszeń pokazano strukturę systemu HEARSAY-II.

Za pomocą tego modelu zapewnia się dopasowanie pracy rozmaitych źródeł wiedzy, które odpowiadają rozmaitym poziomom hierarchicznym przedstawienia problemu. W systemie HEARSAY-II tablica ogłoszeń składa się z 7 hierarchicznych poziomów i dane na każdym poziomie są utrzymywane do tej pory, aż przetwarzanie nie osiągnie tego poziomu. Dane przetwarzane są kolejno rozpoczynając od przetwarzania parametrów akustycznych na najniższym poziomie, dalej przechodzą one poziom wydzielenia fonemów, znajdujący się nad nim poziom fraz zbudowanych z fonemów, następnie wyższy poziom oddzielnych słów oraz poziom jednostek syntaktycznych. Przetwarzanie na każdym z tych poziomów odbywa się nie w sposób zdeterminowany; odpowiednie wyniki pokazują tylko jedną z możliwości rozumienia i w tym sensie są one hipotezami. Kiedy wykorzystując odpowiednią hipotezę, tworzy się hipotezę na wyższym poziomie to przenosi się ona na poziom, na którym odbywa się również przetwarzanie danych. Na przykład, jeśli uda się z fonemów ukształtować oddzielne słowo mające sens, to hipoteza ta w dalszym przetwarzaniu liczy się jako wiarygodna. Niektóre źródła wiedzy, posiadające odniesienie do hipotezy odpowiedniego poziomu, przyjmują hipotezę niższego poziomu jako wskazówkę wykorzystywaną do formowania mechanizmu wnioskowania wstępującego generującego hipotezę dla poziomu wyższego. Po przyjęciu hipotezy wyższego poziomu jako wniosku, źródło wiedzy przeprowadza wnioskowanie zstępujące i sprawdza wynik przetwarzania, przekazując go jako sprzężenie zwrotne na poziom niższy.

przetwarzania wiedzy, które przedstawiają przechowywaną w nich informację dużej liczbie użytkowników.

Logika predykatów jako jeden z języków przedstawienia wiedzy, wykorzystywana jest również w charakterze teoretycznej podstawy do opisu samego systemu przedstawienia wiedzy.

#### 2.5.4. Model tablicy ogłoszeń

##### 2.5.4.1. Przedstawienie modelu tablicy ogłoszeń

W systemie produkcyjnym gromadzi się w bazie wiedzy reguły jednakowe pod względem formy, stąd jeśli dopuścić, że w jednym i tym samym systemie rozwiązywane są problemy z kilku różnych dziedzin praktycznych, to może dojść do pomieszania reguł z różnymi właściwościami, istniejącymi w tych dziedzinach, co będzie przyczyną istotnego obniżenia efektywności przetwarzania. Zatem operowanie różnymi problemami w jednym systemie produkcyjnym jest niepożądane. Bywa jednak w niektórych przypadkach tak, że jeden problem jest budowany z kilku problemów częściowych z różnymi właściwościami. Do tego problemy te nie są wspólne i zachowują niezależność. Na przykład, nie można rozdzielić problemów szczegółowych przy istnieniu takiego słabego związku, kiedy w wyniku rozwiązania jednego problemu szczegółowego inicjuje się początek rozwiązania problemu następnego.

Jednym z takich problemów jest problem rozpoznawania mowy. Wynikiem jego rozwiązania jest rozumienie mowy. Składa się z kilku różnych co do właściwości podproblemów, rozpoczynając z etapu wykrycia sygnałów przez detektory dźwięku aż do rozumienia na wysokim poziomie abstrakcji, kiedy kolejno dokonuje się wydzielania fonemów, rozpoznawania słów, budowy zdań i rozumienia zdania.

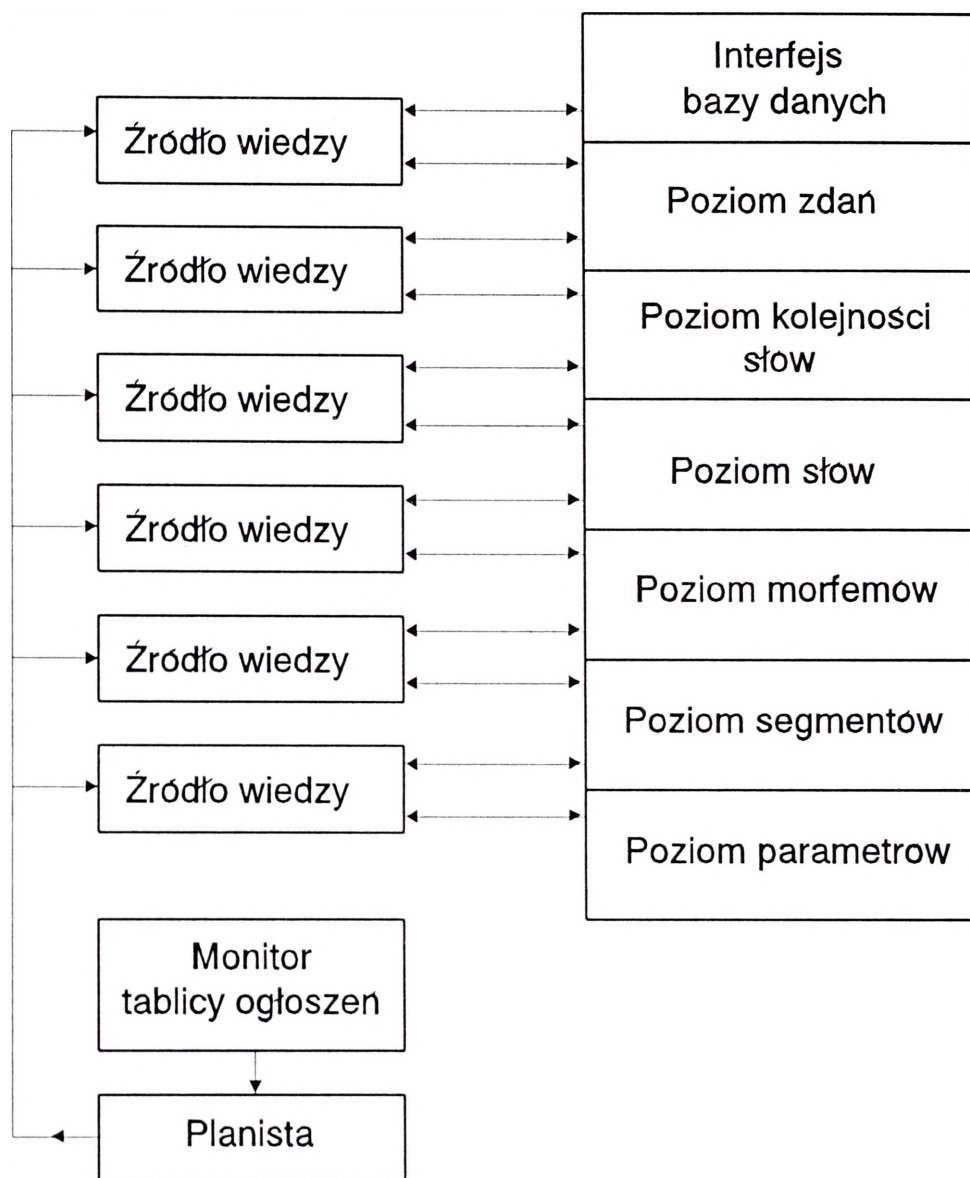
Istnieje system HEARSAY-II, opracowany w amerykańskim uniwersytecie Carnegie-Mellon i zorientowany na tego rodzaju problemy. Z charakterystycznych właściwości jego budowy, rozpatrywanych dalej, trzeba wymienić zastosowanie modelu tablicy ogłoszeń (Ścisłej rzecz biorąc, model tablicy ogłoszeń jest sposobem organizacji i wykorzystania,

a nie przedstawienia wiedzy.). W tym modelu dla każdego oddzielnego problemu, które razem stanowią jedną całość, istnieje odpowiedni zbiór danych, przy tym wszystkie zbiory przez wspólny roboczy obszar pamięci, nazywany tablicą ogłoszeń, są tak sterowane, że cała wiedza jest wykorzystywana spójnie, jako jedna całość. Takie oddzielne zbiory wiedzy nazywają się zwykle źródłami wiedzy (Z). Każdy Z sam jest budowany jako system produkcyjny.

#### 2.5.4.2. Wnioskowanie I przetwarzanie w modelu tablicy ogłoszeń

Na rys. 10 jako przykład modelu tablicy ogłoszeń pokazano strukturę systemu HEARSAY-II.

Za pomocą tego modelu zapewnia się dopasowanie pracy rozmaitych źródeł wiedzy, które odpowiadają rozmaitym poziomom hierarchicznym przedstawienia problemu. W systemie HEARSAY-II tablica ogłoszeń składa się z 7 hierarchicznych poziomów i dane na każdym poziomie są utrzymywane do tej pory, aż przetwarzanie nie osiągnie tego poziomu. Dane przetwarzane są kolejno rozpoczynając od przetwarzania parametrów akustycznych na najniższym poziomie, dalej przechodzą one poziom wydzielenia fonemów, znajdujący się nad nim poziom fraz zbudowanych z fonemów, następnie wyższy poziom oddzielnych słów oraz poziom jednostek syntaktycznych. Przetwarzanie na każdym z tych poziomów odbywa się nie w sposób zdeterminowany; odpowiednie wyniki pokazują tylko jedną z możliwości rozumienia i w tym sensie są one hipotezami. Kiedy wykorzystując odpowiednią hipotezę, tworzy się hipotezę na wyższym poziomie to przenosi się ona na poziom, na którym odbywa się również przetwarzanie danych. Na przykład, jeśli uda się z fonemów ukształtować oddzielne słowo mające sens, to hipoteza ta w dalszym przetwarzaniu liczy się jako wiarygodna. Niektóre źródła wiedzy, posiadające odniesienie do hipotezy odpowiedniego poziomu, przyjmują hipotezę niższego poziomu jako wskazówkę wykorzystywaną do formowania mechanizmu wnioskowania wstępującego generującego hipotezę dla poziomu wyższego. Po przyjęciu hipotezy wyższego poziomu jako wniosku, źródło wiedzy przeprowadza wnioskowanie zstępujące i sprawdza wynik przetwarzania, przekazując go jako sprzężenie zwrotne na poziom niższy.



Rys. 10. Model tablicy ogłoszeń

#### 2.5.4.3. Sterowanie wnioskami

Wszystkie źródła wiedzy mogą pracować niezależnie, asynchronicznie i równolegle. W celu realizacji efektywnego procesu przetwarzania wiedzy w całości system zawiera mechanizm sterowania uruchamianiem źródeł wiedzy zgodnie z bieżącą sytuacją. W modelu tablicy ogłoszeń działanie mechanizmu sterowania jest najważniejsze. Analizuje on gromadzoną na tablicy ogłoszeń sytuację, wyciąga wnioski o tym, jaki poziom przetwarzania powinien posiadać najwyższy priorytet i realizuje planowanie uruchomienia źródeł wiedzy. Uważa się, że w systemie HEARSAY-II część sterująca odpowiada możliwościom intelektualnym i sposobom myślenia człowieka. Jednakże w systemach z takiego rodzaju organizacją powstaje problem, powiązania sposobu realizacji mechanizmu sterowania z algorytmem sterowania. Innymi słowami, algorytm sterowania realizuje się jako mechanizm sterowania, będący bezpośrednio wprowadzony do części sterującej systemem, albo mechanizm sterowania realizuje się w postaci uniwersalnej a algorytm sterowania formułuje się w innej, oddzielonej od własnego mechanizmu sterowania środowiskiem, które inicjuje ten mechanizm. Przy takim wyborze mechanizm wnioskowania w systemach produkcyjnych, na przykład w MYCIN, określa się w zależności od wiedzy, a w EMYCIN należy niestety zwrócić uwagę na to, że wiedza i mechanizm wnioskowania są zupełnie rozdzielone. W systemach ekspertowych typu MYCIN wiedza jest bezpośrednio związana z problemami obszaru przedmiotowego, jednakże wiedza dotycząca problemu obiektowego nie zależy od mechanizmu wnioskowania i jest przedstawieniem stosunku wiedzy o metodach wykorzystania wiedzy obszaru przedmiotowego i mechanizmu, który wykorzystuje tę wiedzę.

W ten sposób, złożona struktura, którą posiada system HEARSAY-II, odpowiada złożoności problemów. Do takich problemów można zaliczyć nie tylko rozumienie mowy, lecz i automatyczne programowanie, planowanie, planowanie eksperymentu, rozumienie tekstów, przetwarzanie obrazów. Wszystkie te dziedziny przedmiotowe wymagają jednakże jeszcze bardziej złożonych koncepcji aniżeli system HEARSAY-II. Wszystkie te problemy nie są na tyle proste żeby ograniczać przetwarzanie do ram systemów produkcyjnych, które buduje się obecnie z kilku różniących się między sobą jednostek operacyjnych, co daje o wiele bardziej złożoną strukturę aniżeli w HEARSAY-II. Zatem, do tworzenia

systemów bardziej uniwersalnych pod względem funkcji trzeba stosować metody dopasowania wielu źródeł wiedzy, i rozstać się z ideą budowy systemu w formie zależnej od problemu, jak to ma miejsce w systemie HEARSAY-II. W szczególności, dla tablicy ogłoszeń należy stosować strukturę danych o postaci ogólnej, niezależnej od rodzaju problemu. Może to uczynić jeszcze bardziej złożonym całe sterowanie, w wyniku czego będzie ono nałożone na samego użytkownika. Takie systemy, jak AGE, ART, ESHELL, oparte na wykorzystaniu idei HEARSAY-II, są niewątpliwie jeszcze bardziej uniwersalnymi systemami w wyniku orientacji właśnie na taki kierunek. To właśnie doprowadziło do odejścia od źródła pierwotnego i wzmocnieniu w nich elementów proceduralnych.

#### 2.5.5. Systemy frejmowe<sup>21)</sup> (wiedza proceduralna)

##### 2.5.5.1. Teoria frejmów i systemy frejmowe

Frejm posiada wszystkie właściwości istotne dla języka przedstawienia wiedzy, a jednocześnie jest on sam w sobie nowym sposobem przetwarzania informacji. W chwili obecnej do systemów frejmów wniesiono takie właściwości języka programowania, które pozwoliły na bazie teorii frejmów programować zadania. Frejm jest metodą przedstawienia wiedzy, opartą na teorii frejmów, opublikowanej przez M. Minskiego w roku 1975., jako jednego z naukowych podejść do opisu wiedzy, przydatnej dla zrozumienia scen i mowy potocznej. Frejm jako taki jest pojęciem rozpatrywanym z pozycji zastosowania niektórych zadań. Aby przedstawić sobie bardziej konkretny obraz frejmu założymy, że niezbędna jest, inna aniżeli procedury lub predykaty wielowymiarowa forma, która posiada pewne właściwości, różne od innych form. Jeśli można tak się wyrazić, frejm - to forma przedstawienia wiedzy, która opisuje ramy rozpatrywanego świata i wykorzystuje do tego opisu

---

<sup>21)</sup> Świadomie używam tu tytułem eksperymentu terminu "frejm" jako powszechnie stosowanego w literaturze światowej, zamiast przyjętego u nas terminu "rama" lub "ramka".

systemy produkcyjne lub predykaty. W wyniku tego semantyka wiedzy określana jest zgodnie z otaczającym środowiskiem, a sfera działania tego środowiska będzie określana zgodnie z zakresem interesów człowieka.

Przeanalizujemy myślenie człowieka, który wraca do domu z pracy. Najpierw obmyśla drogę powrotu, pozostając przy podzakresie rozmyślań o określonym terytorium, wyobrażając go sobie według schematu tras przejazdu i punktów przesiadek. Następnie w drodze na przystanek aż do powrotu do domu rozmyśla w bardziej wąskim podzakresie. Kiedy przychodzi do domu, to podzakres jego myślenia ogranicza się teraz do pomieszczeń własnego domu, lecz jeśli wejdzie do swojego pokoju, to myślenie jest ograniczone jego ścianami. Jeśli przyjąć, że analogicznymi podzakresami myślenia są frejmy, to każdy frejm ma określoną swoją własną granicę przedstawienia i działania.

W teorii frejmów prawie nie mówi się niczego konkretnego na temat metod realizacji. W ślad za powstaniem teorii frejmów zaczęły pojawiać się i języki, posiadające możliwości opisu procesów formalnych, kiedy w postaci programu opisuje się wszystkie działania, które powinny być wykonane dla każdego świata złożonego z obiektów (frejmu). Programy wywołuje się z określonego frejmu, a przy wymianie informacji między różnymi frejmami odbywa się albo wymiana informacji między frejmami, albo przekazywanie sterowania. Jest to realizacja systemu frejmów, jednakże nie posiada ona jeszcze bezpośredniego odniesienia do teorii frejmów, a przedstawia sobą program komputerowy, którego przesłanką stała się teoria frejmów. Metodę frejmową można przedstawiać i do przedstawiania wiedzy, lecz zadania przy tym będą opisywane w formie proceduralnej i dlatego przy porównaniu rozmaitych metod, które rozpatrywaliśmy uprzednio, trzeba mówić o metodzie proceduralnej. Znane są również języki zorientowane na obiekty i posiadające zbieżną z systemem frejmów strukturę sterowania. Dalej krótko opiszemy przedstawienie wiedzy i sterowanie wnioskowaniem za pomocą frejmów.

#### 2.5.5.2. Formy przedstawienia systemu frejmów

Jednostką przedstawienia w systemie frejmów jest obiekt, nazywany frejmem. Jest on formą przedstawienia jakiejś sytuacji, którą można (lub jest celowo) opisywać pewnym zbiorem pojęć i prawidłowości. Jako

identyfikatora frejmu używa się jego nazwy. Imię to powinno być unikalne w całym systemie frejmów. Frejm posiada określoną strukturę wewnętrzną, składającą się ze zbioru elementów, nazywanymi słotami<sup>22)</sup>, którym również nadaje się nazwy. Każdy slot z kolej jest przedstawiany jako określona struktura danych.

Rodzaj sytuacji opisywanej w postaci frejmu określa użytkownik. Pewna prawidłowość, na przykład pewna grupa ludzi, z której każdy jest

<u>Freim : Nazwa</u>	
Klasa	; Zwierzę
Element struktury	; Głowa, szyja, ręce, nogi,...
Wzrost	; 30 - 220 cm
Masa	; 1 - 200 kg
Ogon	; Brak
Język	; Polski /angielski/rosyjski
Frejm analogii	; Małpa

Rys. 11. Prosty przykład frejmu opisującego "człowieka".

poddany wszelkiego rodzaju opisom, tworzy jedną określoną sytuację i dlatego jest opisywana jako frejm. Jako wartość slotu przedstawia się konkretną informację odnoszącą się do obiektu opisywanego za pomocą frejmu. Na rys. 11 na prostym przykładzie pokazano frejm opisujący człowieka.

Zalóżmy, że tego rodzaju frejmy są gromadzone w postaci wiedzy i chcemy rozpoznać pewien obiekt. W tym celu rozpatrzemy kilka zadanych we frejmie "człowiek" slotów i przypisanym im atrybutom.

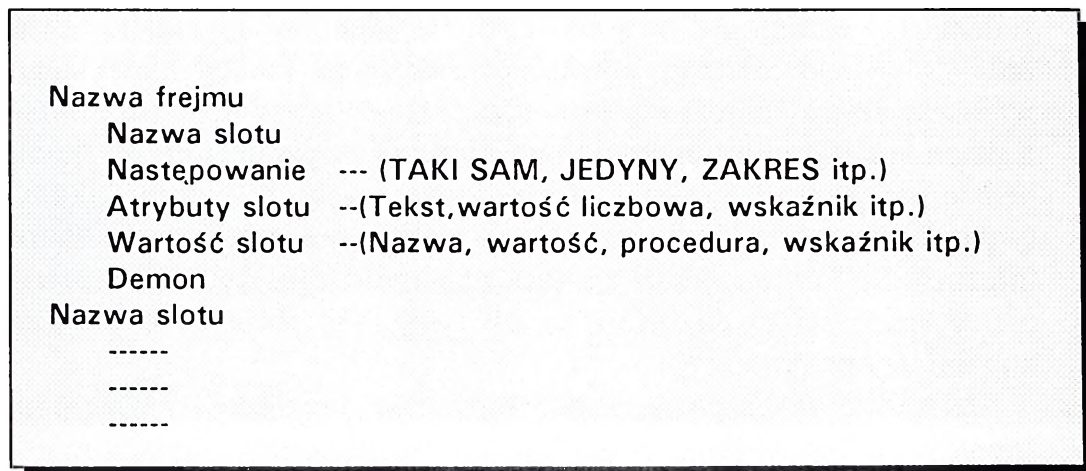
<sup>22)</sup> Używam tu świadomie pojęcia "slot" zamiast przyjętych u nas nazw "klatka", "szczelina" itp.

Rozpoczynamy procedurę dopasowania wywoływaną w tym celu z pamięci frejm "człowiek". Kiedy otrzyma się dane, które spełnią wszystkie warunki zawarte w slotach, obiekt zostanie zidentyfikowany jako człowiek. Jeśli dane wejściowe nie będą zgodne z warunkami zadanymi we frejmie "człowiek". na przykład będzie zadane, że masa obiektu wynosi 300 kg, obiekt posiada ogon, to będzie wyciągnięty wniosek, że taki obiekt nie jest człowiekiem. Po tym, posłużwszy się wskaźnikiem frejmu analogii i wyzwawszy z pamięci frejm "małpa", będzie przeprowadzone analogiczne dopasowanie. \* - wartość domyślna (wykorzystuje się, kiedy dla danego punktu nie otrzymano konkretnej informacji). Ten przykład pozwala zrozumieć sens sytuacji nawet przy nie w pełni zadanych informacjach, jednakże stwarza on możliwość popełniania błędów.

Jedną z charakterystycznych właściwości przedstawienia za pomocą frejmów jest określenie następowania (następowania atrybutów): jeśli obiekt, który jest opisany pewną grupą frejmów, znajduje się w związku konceptualnym z górnymi i dolnymi poziomami (stosunek IS-A lub PART OF), to odpowiadający jemu frejm można skonstruować na bazie jego powiązań hierarchicznych i przy tym można zapewnić następowanie atrybutów slotów z jednakowymi imionami wewnątrz frejmu górnego poziomu. W charakterze informacji opisującej następowanie wewnątrz frejmu można wskazać sposób następowania, na przykład "ten sam", "jedyny", "zakres". Załóżmy, że wskazaniem następowania "ogon" we frejmie "człowiek" jest "ten sam", wówczas wartość slotu "ogon" frejmu "mężczyzna" w dolnym poziomie frejmu "człowiek" będzie mu bezpośrednio podporządkowana (tzn. przyjmie wartość "nie") Dalej, jeśli następowanie jest wskazane jako "zakres", i do wartości slotu "masa ciała" podstawią się "1 - 200kg" (patrz rys. 2.13), to wartość masy w jednakowych slotach wszystkich frejmów, które następują po tym atrybucie, powinny znajdować się wewnątrz takiego zakresu. Następowanie pozwala wyeliminować dublowanie informacji, jego efekt przejawia się również w eliminowaniu przeciwstawnych wiadomości.

Do tej pory była mowa o wartościach slotów, jednakże dla każdego slotu są określone rozmaite typy danych takie na przykład, jak wartości liczbowe i ich odwołania się do innych frejmów (innymi słowami nazwy innych frejmów). Takim określeniem wskazywane są atrybuty slotów, dopuszcza się także zadawanie funkcji. Jeśli są one zadawane tak jak w Lispie, to wartość slotu będzie nazywać się funkcją Lispu (tzn. dodawana

procedura) (rys. 12). Tym samym w systemie frejmów dopuszczalna jest możliwość zamiany wiedzy przez procedury.

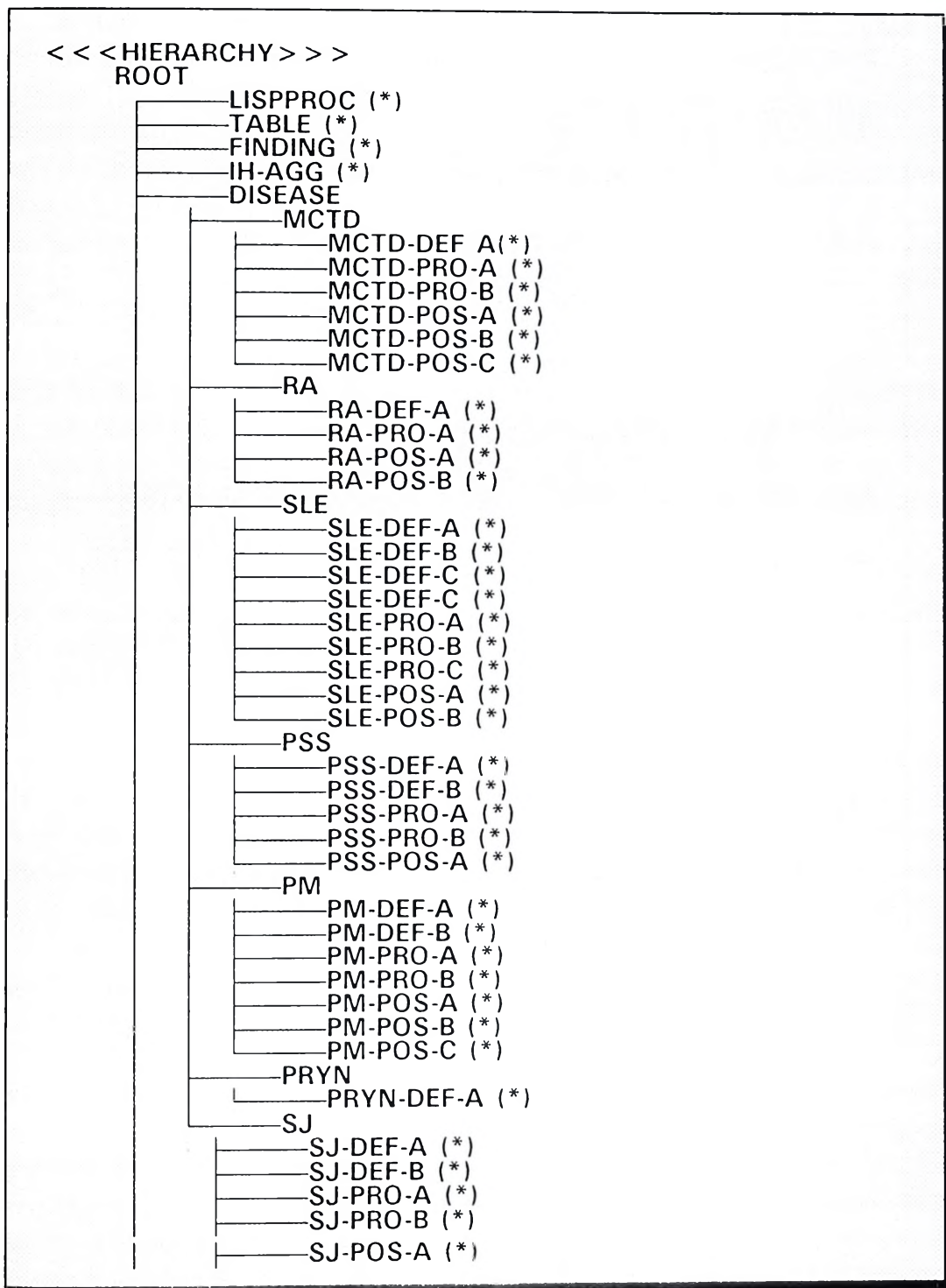


Rys. 12. Struktura frejmu.

### 2.5.5.3. Wnioskowanie w systemie frejmów

Wnioski w systemie frejmów wykorzystuje się do wymiany wiadomości między frejmami. Dla przykładu, rozpatrzmy zastosowanie systemu frejmów do celów diagnostyki medycznej. Na rys. 13 pokazano przykład systemu ekspertowego do diagnostyki reumatyzmu, wytworzonego na bazie systemu frejmów.

Na rysunku tym przedstawiono całkowitą strukturę sformułowaną z większej liczby frejmów. Na najwyższym poziomie istnieje frejm - korzeń <<ROOT>>, dalej formułuje się niezbędną hierarchię frejmów, opisującą formy bardziej konkretne. Frejmy poziomu najniższego oznaczono symbolem \* i nazywa się je frejmami następowania. Frejmy pośrednie nazywają się frejmami klas, a frejmy leżące niżej od ich klas opisywane są przez ogólne dla klas pojęcia. Na rys. 14 pokazano wewnętrzną strukturę frejmu. <<I. Role>> opisuje wskaźnik następowania, część <<From>> opisuje położenie w hierarchii frejmów, gdzie jest określony ten slot.



Rys. 13. Frejm ze strukturą hierarchiczną  
 (system diagnostyczny reumatyzmu, wykorzystujący język FMS).

Frame-name: PM-POS-B Frame-type: INSTANCE

Slot NameL.	Role	From	Type	Value
A-Kind-Of	(U)	ROOT	FRAME	PM
D.Deacendnants	(U)	ROOT	FLIST	NIL
Description	(U)	ROOT	TEXT	"Polymyosltis,Possible-B"
CreatedBy	(U)	ROOT	STRING	("G00126""83.12.29")
ModifiedBy	(U)	ROOT	STRING	("G00126""83.12.30")
MajorElements	(S)	PM	LIST	(AMES MBXS PMWS...)
MinorElements	(S)	PM	LIST	(HELI RRES AMEM...)
MajorCriteria	(U)	PM	LIST	(00 NIL)
MinorCriteria	(U)	PM	LIST	(00 NIL)
SerologyCr	(U)	PM	ATOM	NIL
Exclusion	(U)	PM	ATOM	PMXX
Certainty	(U)	PM	ATOM	Possible
Conclusion	(U)	DISEASE	LIST	*NOT-DEFINED*
Logic	(U)	DISEASE	LIST	CASELOGIC
UnknowLogic	(U)	DISEASE	LIST	UNKNOWCASELOGIC

*CASELOGIC i UNKNOWNCASELOGIC sa procedurami przyłączanymi.*

Rys. 14. Struktura frejmu PM-POS-B.

Kiedy system frejmów wykorzystuje się do diagnostyki, trzeba określić nazwę choroby podłączoną do slotu <<LOGIC>> tego frejmu. Za pomocą frejmu <<DISEASE>>, inicjuje się procedura przyłączona do jego slotu <<LOGIC>>. Frejm <<DISEASE>> w celu przeprowadzenia diagnostyki według każdego kryterium diagnostycznego posyła jedną za drugą wiadomości z <<PM-DEF-A>>, będącego frejmem jego niższego poziomu, do frejmu <<PM-POS-C>> i zapewnia uruchomienie procedury <<CASEOLOGIC>>, przyłączonej do slotu <<LOGIC>>. Dla wszystkich tych frejmów zadaje się informację o chorobie, za pomocą której jest prowadzona diagnostyka, oparta na przeprowadzenia kryteriów diagnostycznych, wynik jest wracany do frejmu <<PM>> jako odpowiedź na wiadomość. Frejm <<PM>> opuszcza na swoim dolnym poziomie wniosek diagnostyczny, oparte na rozmaitych odnoszących się do <<PM>> kryteriach i zwraca jego do frejmu <<DISEASE>>, który jest źródłem generacji tej wiadomości. Tym samym frejm <<DISEASE>> powtarza uogólniony wniosek diagnostyczny jako odpowiedź do źródła generacji wiadomości do frejmu <<DISEASE>>.

Specyficzną funkcją we frejmach jest demon. Demon zadaje automatycznie uruchamianą procedurę, kiedy do jego slotu podstawia się pewną wartość lub kiedy przeprowadza się konfrontację wartości. Przypuśćmy, że wprowadzamy wskazanie <<IF-NEEDED>> i jednocześnie z tym wprowadzamy nazwę procedury ASK, jeśli tylko wartości jeszcze nie mogą być podstawione. Program ten zadaje pacjentowi pytanie (na przykład, <<czy boli głowa?>>), podstawia jego odpowiedź <<tak>> lub <<nie>> do slotu i jednocześnie z tym wykonuje powrót do początku konfrontacji. Demon wypełnia działania zapewniające elastyczność w tego rodzaju wnioskowaniu.

W ten sposób, w systemie frejmów każdemu z frejmów zadaje się określone obowiązki i zapewnia się uzgodnione rozwiązanie problemów za pomocą wniosków we wcześniej przydzielonym zakresie obowiązków. Odpowiedzialność za usystematyzowane prowadzenie przetwarzania wiedzy podczas rozwiązywania problemu jest nakładana na użytkownika, który, wychodząc z własnego wyobrażenia o całości problemu powinien zapisać system frejmów. W istocie nie różni się to prawie od programowania na tradycyjnym komputerze i dlatego tę metodę trzeba nazywać proceduralną.

## 2.6. Właściwości rozmaitych sposobów przedstawienie wiedzy

Rozmaite sposoby przedstawienie wiedzy rozpatrzone w rozdziale dotyczącym metody proceduralnej, były omawiane tak, jak gdyby nie były one w ogóle związane między sobą, co oczywiście nie jest prawdą. Odzwierciedla to tylko ten fakt, że wszystkie prace prowadzono według różnych koncepcji. Jednakże, zawsze istniał wspólny cel - przedstawienie wiedzy i ich wykorzystanie do rozwiązywania problemów, stąd można powiedzieć, że wzajemne rozróżnianie sposobów przedstawienia jest zawarte bardziej w formie zewnętrznej, a nie zmienia się istota. Oczywiście sposób osiągnięcia celu będzie inny dla każdego odpowiedniego sposobu przedstawienia, kiedy w płaszczyźnie metodologicznej bardzo jest potrzebne łatwe osiągnięcie celu podczas wnioskowania. Oznacza to jednak, że jeśli posłużyć się pytaniem - jaki cel powinien osiągać system przedstawienia wiedzy, - to odpowiedź nie zawsze będzie jasna. Kiedyś przed inżynierią wiedzy stawiano zadanie "pomyślnie posiłkować się wiedzą". O współczesnych systemach inżynierii wiedzy można także powiedzieć, że ich praktyczna realizacja osiągnęła na razie tylko etap eksperymentalny. Aby zapewnić pełnowartościowy rozwój technologii przetwarzania wiedzy, należy na samym początku postawić cel, który powinna osiągać ta technologia.

*Zadaniem przetwarzania wiedzy jest wspomaganie działalności intelektualnej człowieka, stąd tworzenie systemu przetwarzania wiedzy należy rozpocząć od analizy metod rozwiązywania problemów przez człowieka.*

Omówienie tego problemu jako jednego z ważniejszych w tej pracy, będzie omówione w dalszych rozdziałach.

Tu jedynie w ramach wytworzonej sytuacji rozpatrzemy tylko specyficzne właściwości, przejawiające się w zewnętrznej postaci rozmaitych sposobów przedstawienia wiedzy i to wspólne co jest w nich istotne.

### 2.6.1. Porównanie współczesnych sposobów przedstawienia wiedzy

Różne sposoby przedstawienia wiedzy posiadają istotne tylko dla nich cechy charakterystyczne te, które omawialiśmy poprzednio oraz inne. Na początku omówimy ich najbardziej oczywiste właściwości.

#### 2.6.1.1. Reguły produkcji

Przedstawienie wiedzy za pomocą reguł produkcji jest proste a wnioski, które wyciąga się na podstawie formalizmu JEŚLI-TO, łatwo zrozumiałe i analogiczne do tych wszystkich sylogizmów, które przyswoiliśmy sobie jeszcze w wieku dziecięcym. Oprócz tego jasno przedstawiona modułowość reguł pozwala zadawać nowe wartości nie wdając się w sens innych wiadomości. Dzięki temu oraz dzięki możliwości budowy wniosków w jednorodnej formie otrzymywane systemy są proste i łatwe do zrozumienia. Z drugiej strony nietrudno zauważyć, że prostota ta, jako wynik niezbyt wielobarwnego przedstawienia wszystkich opisów, nakierowana jest na rozwiązywanie prostych oraz jednorodnych pod względem właściwości zadań i prowadzi do wyraźnego obniżenia efektywności rozwiązywania takich problemów, które składają się z kilku różnorodnych zadań. Lecz nawet jeśli problem jest prosty, to nie udaje się zbudować systemu do sterowania wiedzą jako jedną całością, stąd proces sterowania powinien być realizowany przez człowieka. Jeśli problem staje się trochę bardziej złożony to i sterowanie staje się trudniejsze. W szczególności właściwość systemu wiedzy polega na tym, że człowiek może w ciągu dłuższego czasu gromadzić lub starannie dopracowywać wiedzę. Jednakże w tym przypadku możliwe jest, że z powodu niektórych przyczyn albo źródeł tych przyczyn, spowodowanych przez ludzi lub brak czasu, powstanie wiedza przeciwstawna do otrzymanej wcześniej. Człowiek posiada pewne ograniczenia związane z sterowaniem takim procesem, co następnie nakłada ograniczenia co do skali i różnorodności rozwiązywanych problemów.

#### 2.6.1.2. Sieci semantyczne

Sieci semantyczne to ogólna nazwa metod opisu wykorzystujących sieci do opisu wiedzy. charakterystyczną właściwością sieci semantycznych jest dobra poglądowość wiedzy jako systemu. Każda oddzielna wiadomość jest rozpatrywana jako pewien stosunek między istotą a pojęciami i formalnie, tak jak i w systemach produkcji, można gromadzić określone wiadomości dodając je do już istniejących wewnątrz systemu, zachowując ich modułowość. Jednocześnie cała wiedza odnosząca się do jednakowych treści i pojęć może być przedstawiona w postaci związków między różnymi węzłami, opisującymi te treści i to daje podstawę do sądu o łatwości rozumienia takiego przedstawienia. Na podstawie tych sieci przeprowadza się wnioskowanie, jednakże potrzebne są tu specjalne algorytmy wnioskowania. W systemach produkcji wnioski określa się zgodnie z ograniczonym formalizmem Jeśli - TO, dlatego algorytmy takiego wnioskowania są również sformalizowane, jednocześnie łącznie ze swoją prostotą są one wystarczająco dokładne. Ponieważ sieci semantyczne są zbiorczą nazwą systemów przedstawienia wiedzy wykorzystujących sieci, nie ma sensu określać specjalnie dla nich odrębnych algorytmów wnioskowania, stąd bardzo trudno porównywać sieci z innymi sposobami przedstawienia wiedzy. Jednakże z tego powodu, że forma przedstawienia za pomocą sieci nie jest ustalana dla konkretnego formalizmu, będą potrzebne odrębne własne reguły wnioskowania, stąd silniejszy jest element dowolności wnoszony przez człowieka. Wnioski, które nie są dostatecznie starannie sprawdzone, stwarzają niebezpieczeństwo powstawania przeciwstawności. Zatem, w sieciach semantycznych trzeba bardziej aniżeli w systemach produkcji zwracać uwagę na takie okoliczności jak usuwanie przeciwstawności. Sam system nie dysponuje takimi możliwościami i dlatego w wielu przypadkach funkcja ta powierzana jest człowiekowi. Przeglądając wszystkie wiadomości człowiek jest w stanie sterować ich nieprzeciwstawnością, jednakże jeśli objętość wiadomości będzie się zwiększała, to ich przedstawienie będzie się znacznie utrudniać, co ograniczy krąg rozwiązywanych zagadnień do stosunkowo niewielkich problemów.

### 2.6.1.3. Logika predykatów

Logikę predykatów cechuje wysoki poziom modułowości wiedzy i jednocześnie pozwala ona otrzymywać jednolity system reprezentacji, w którym logicznie objaśniane są właściwości wiedzy jako jednej całości. Zatem z pomocą logiki predykatów można, określając w dowolny sposób wiadomości, wyjaśnić czy istnieją lub brak jest przeciwstawności między nowymi i już istniejącymi wiadomościami. Z drugiej strony, logikę predykatów cechują, w następstwie zachowania cechy całościowości, takie niedostatki jak nadmierny poziom formalizacji przedstawienia wiedzy, trudność ich odczytu, nie za bardzo wysoka wydajność przetwarzania. Oprócz tego, w logice predykatów wszystkie relacje opisywane są predykatami, stąd przy komputerowym przetwarzaniu te zalety, które posiada struktura danych, nie mogą się uwydatnić w pełnej mierze. To także pogarsza efektywność przetwarzania.

### 2.6.1.4. Model tablicy ogłoszeń

Model tablicy ogłoszeń jest jedną z rozwiniętych modyfikacji systemu produkcji. Celem opracowania modelu tablicy ogłoszeń było stworzenie takiego systemu, który kompensowałby ograniczoność monotonicznych wniosków. Może ona być użyta do rozwiązywania złożonych problemów w drodze wykorzystania mechanizmu podziału źródeł wiadomości i ich dopasowanych działań we wspólnej pamięci roboczej, przy tym z wykorzystaniem takiej cechy dodatniej systemu produkcyjnego, jak modułowość wiedzy. Jeśli obiekt przetwarzania jest zbiorem rozmaitych prac to mieszanie ze sobą różnorodnych wiadomości nie prowadzi do niczego dobrego. Kiedy powstaje skomplikowany problem z wyrażeniem wyrażoną niezależnością oddzielnych prac, to nie ma szczególnych przeszkód na drodze do praktycznej realizacji systemu, jednakże będzie ona różnić się od systemu produkcyjnego. Jeśli oddzielne jednostki operacyjne, zgodnie z wynikłą sytuacją, będą znajdować w bardziej ścisłych związkach wzajemnych, to problemu nie będzie można rozwiązać za pomocą prostego mechanizmu sterowania tablicą ogłoszeń.

Inaczej mówiąc, najważniejszym momentem w pomyślnej pracy systemu wnioskowania jest sterowanie tablicą ogłoszeń, jednakże w praktyce komplikacje te wywołują także trudności przedstawienia metawiedzy i sposobów jej wykorzystania. W przypadku bardziej ogólnym sterowanie przekazywane jest samemu użytkownikowi. Wymaga to aby posiadał on wysoki poziom wiedzy uwzględniający również dużą indywidualność systemu, co komplikuje posługiwanie się systemem. Oczywiście powstaje tu również, tak jak i w przypadku sieci semantycznych, problem przeciwstawności wiedzy.

#### 2.6.1.5. Systemy oparte na frejmach

Systemy oparte na frejmach określające tak jak systemy produkcji formę przedstawienia wiedzy i różnią się od innych systemów tym, że pozwalają użytkownikowi na większą swobodę. System oparty na frejmach nie tylko opisuje wiedzę, lecz może także być wykorzystywany przez użytkownika do napisania algorytmów wnioskowania. Z punktu widzenia użytkownika, który bardzo dobrze zna przetwarzanie wiedzy, wygląda to na zaletę, jednakże dla zwykłego użytkownika jest to wielkie obciążenie. Można powiedzieć, że system frejmów jest rozszerzeniem systemów tradycyjnych typu proceduralnego. Dzięki takim właściwościom można formalnie budować systemy frejmów w bardzo szerokim zakresie, dlatego i oceny otrzymywane za pomocą systemów opartych na frejmach powinny być, prawdopodobnie bardzo różnorodne.

Frejm jako struktura opisuje jedną z jednostek przetwarzania, posiadającą w pewnym stopniu niezależność i może dostarczać środków wiążących te jednostki strukturalne. Zatem, jeśli porównać system oparty na frejmach z tradycyjnym systemem oprogramowania, gdzie cały sens przetwarzania powinien być opisany grupą kolejno łączonych operacji, to można mówić o częściowym zdjęciu ograniczeń kolejnego procesu przetwarzania z powodu tego, że poddane temu przetwarzaniu jednostki są opisywane w postaci frejmów, które (w skrajnym przypadku, dla ich części) nie koniecznie następują jedna po drugiej. Jednakże ta okoliczność, że wiedza zadawana jest tu w swej istocie procedurami komplikuje w porównaniu z innymi metodami przysposabianie wiedzy i zubaża możliwości dynamicznej adaptacji do zmian otoczenia. Oprócz tego,

zakłada się, że w złożonych problemach relacje między frejmami również stają się bardziej złożone.

W systemach opartych na frejmach istnieje jeszcze jeden istotny problem, związany z ich strukturą. Jako taki, frejm jest zbiorem opisów (slotów), będących między sobą w pewnych wzajemnych zależnościach. Ta okoliczność, że sloty związane są z procedurami oraz to, że przekazują one między sobą sterowanie za pośrednictwem wymiany wiadomości, oznacza, że struktura systemu frejmów opisuje strukturę sterowania przetwarzaniem proceduralnym. Jednakże w wielu przypadkach opis pewnej sytuacji rozpatruje się jako jednolitą całość i dlatego często frejm jest przedstawieniem tej sytuacji, analogicznie do frejmu "człowiek" z poprzedniego przykładu. Kiedy wszystkie zbieżne między sobą sytuacje zebrane są w jedną całość i zbiór ten jest opisany w postaci frejmu jako pewnej sytuacji, to między otrzymanym frejmem i frejmami każdej oddzielnej sytuacji będą sformułowane związki strukturalne. Na bazie tej struktury, która może być rozpatrywana jako struktura danych będzie realizowane następowanie. W końcowym wyniku system frejmów jest doprowadzany do formy przedstawienia jednostek struktur sterujących i jednostek struktur danych we frejmach ogólnych, dlatego czasem niezbędne jest posługiwanie się dowolnie wymyślonymi przykładami; przy tym pogarsza się lekkość rozumienia intuicyjnego, będącego właściwością tego systemu. Analogiczne ograniczenie dopuszczalne jest dla problemów o stosunkowo niewielkiej skali, lecz jeśli tylko problem staje się trochę większy, to i sterowanie w systemie opartym na frejmach jest bardziej złożone aniżeli w systemach tradycyjnego przedstawienia proceduralnego, stąd trudno o nim mówić jako o uniwersalnym systemie przetwarzania wiedzy.

W ten sposób dla większości wykorzystywanych dziś systemów przedstawienia wiedzy i wytworzonych na jej podstawie systemów charakterystyczne jest nagle skomplikowanie się opisu problemu przy rozrastaniu się jego skali. Tę niedogodność trudno będzie usunąć nie pozbawiając się jednocześnie zalet odpowiednich metod. Jedynym wyjątkiem jest logika predykatów. Posiada ona oczywiście również ten niedostatek, lecz jeśli we wszystkich innych metodach przy skomplikowaniu problemu do pewnego stopnia faktycznie zanika możliwość sterowania wiadomościami (otrzymany wynik nie jest gwarancją jego wiarygodności), to logika predykatów pozwala gwarantować otrzymanie niezawodnych wyników nawet przy obniżeniu efektywności przedstawienia.

Zatem, logika predykatów daje jednolity system, którego zastosowanie pozwala określać wielki problem, krok po kroku zadając informacje wejściową.

Kiedy w jakiej by to nie było dziedzinie, włączając w to i systemy sztucznej inteligencji, jest proponowana nowa metodologia, bardzo ważne jest aby ustanowić, czy jest ona przydatna do rozwiązywania złożonych problemów w tym stopniu, jak to wymaga się w rzeczywistości. Kiedy mowa jest o systemach sztucznej inteligencji, to metodologia systemowego projektowania, która nadawałaby się do problemów o wielkiej skali, jest bardzo daleka od idei prostego ilościowego rozszerzenia metodologii wytworzonej dla niewielkich systemów i w tym przejawia się istotna różnica między takimi metodologiami.

Do oceny przydatności systemów ekspertowych lub innych systemów inżynierii wiedzy do rozwiązania problemów rzeczywistych dochodzono w drodze ilościowego wzrostu technologii i algorytmów wykorzystywanych w tych systemach. W wyniku tego okazało się, że w pewnych etapach narastania lub na samym jego końcu są one rzeczywiście pożyteczne w rozwiązywaniu problemów i tym wyrównają się od dowolny innych wykorzystywanych dziś metod.

#### 2.6.2. Zbieżności i rozbieżności w sposobach przedstawienia wiedzy

Chociaż wszystkie opisywane uprzednio metody różnią się na pierwszy rzut oka to w istocie posiadają one wiele wspólnego i różnice między nimi posiadają charakter konceptualny. Systematyzacja metod i objaśnienie związków wzajemnych między podstawowymi pojęciami i funkcjami są przydatne podczas tworzenia nowych sposobów przedstawienia wiedzy.

Rozpatrzmy najpierw zbieżności tych metod. Jak już o tym mówiliśmy w rozdziale 2.4.2, zdania logiki predykatów posiadają wiele wspólnego z regułami produkcji. W istocie, interpretując zasadę produkcji jako zdanie logiki predykatów, można bez jakichkolwiek trudności zamienić go na predykaty. Przy takich działaniach wyrażenie OJCIEC(WŁADYSŁAW, JAN), które opisano za pomocą formuły predykatów, odpowiada stosunkowi WŁADYSŁAW - OJCIEC - JAN w sieci semantycznej. W ogólnym przypadku każdy luk sieci semantycznej

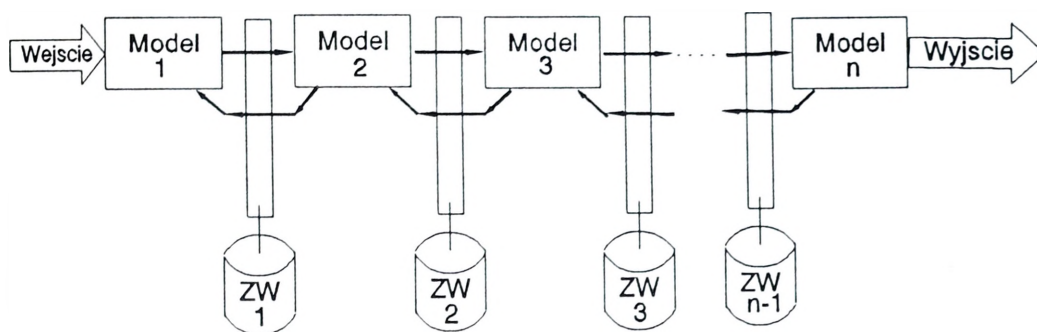
można zadać predykatami jako relację między prawidłowościami (regułami), które opisują węzły na końcach łuków. Możliwość opisu pewnego wyrażenia otrzymanego jednym ze sposobów za pomocą sposobu drugiego oznacza, że wyrażenia te są równoznaczne. Jednak ta parzysta równoznaczność relacji w ogóle nie oznacza, że takie sposoby przedstawienia jak reguły produkcji, logika predykatów lub sieci semantyczne pozwalają otrzymywać równoznaczne sposoby przedstawienia. Tłumaczy to się tym, że często to co można przedstawić za pomocą jednego wzoru nie można wyrazić za pomocą innego wzoru oraz tym, że różne są same idee i systemy leżące u podstaw tych wzorów i to wywiera wpływ na metody wnioskowania. Z ogólnych wyobrażeń wynika, że teoretyczny fundament reguł produkcyjnych i sieci semantycznych jest dosyć słaby w porównaniu z logiką predykatów. Z tego, że jest on słaby nie wynika konieczność uporczywego trzymania się przyjętych w nich formalizmów i że można zmieniać ich postać. Jednym z takich przykładów jest pełnoprawne włączenie do opisu tych pojęć, które określają ilościową ocenę tego opisu typu WP (współczynnika pewności)<sup>23)</sup> w MYCIN, a w logice predykatów nie jest to takie łatwe do zrobienia i powinny być stworzone do tego odpowiednie podstawy teoretyczne. Tym nie mniej opisująca moc logiki predykatów jako jednolitego systemu formalizmów przedstawienie wiedzy jest wyższa aniżeli w innych systemach. Tłumaczy się to tym, że w logice predykatów można ze wzoru swobodnie wyłączyć zmienne a w innych metodach jest to albo zupełnie niemożliwe albo możliwe w stopniu bardzo ograniczonym. Lecz jeśli nawet przy istnieniu takich rozbieżności wszystkie te metody są do siebie podobne pod tym względem, że przedstawione za pomocą ich formalizmów wiadomości nie różnią się zbyt między sobą i jeśli problem rozrasta się i komplikuje to sposoby te nie mogą jej adekwatnie odpowiadać. Duży problem to nie proste narastanie jednakowych formuł lecz jest to zbiór wielu różniących się co do właściwości dopasowanych i wzajemnie powiązanych problemów szczegółowych. Dla każdego problemu szczegółowego wykorzystuje się tylko istotne dla niego wiadomości. Przy tym najlepiej przeprowadzić dekompozycję zbioru wiedzy i przygotować źródła wiadomości dla każdego problemu szczegółowego. Takie podejście jest realizowane nie tylko w tablicy

---

<sup>23)</sup> Znane są także nazwy "współczynnik wiarygodności" i "wskaźnik określoności."

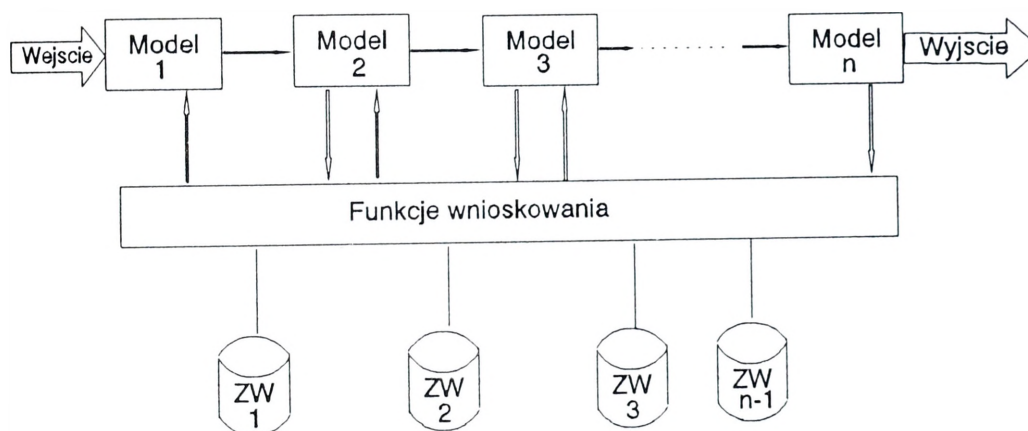
ogłoszeń. Analogiczna do niego jest również dekompozycja sieci semantycznej. Lecz jeśli tablica ogłoszeń jest realizowana za pomocą reguł produkcji, to sieci semantyczne są opisywane identycznymi dla nich w przytoczonym wyżej sensie przedstawieniami sieciowymi. Chociaż z pozycji logiki formalnej istnieje tu analogia (podobieństwo), w praktyce systemy produkcji są prostsze aniżeli sieci semantyczne i dlatego model tablicy ogłoszeń można jak widać, rozpatrywać jako praktyczną realizację produkcji. Oprócz tego struktura taka jest powszechnie używana w systemie frejmów. Frejmy rozpatrywane jako zbiór wiedzy z silnymi związkami wzajemnymi, odpowiadają źródłom wiadomości. Istotna różnica frejmów od metody tablicy ogłoszeń i rozdzielonych semantycznie sieci polega na przesłance wejściowej, w sieciach i tablicy ogłoszeń sterowanie dużą liczbą źródeł wiedzy wykonuje sam system, w tym czasie jak w systemie frejmów to początkowo jest wykonywane przez samego użytkownika w drodze wprowadzenia mechanizmów wymiany informacji. Jest to różnica nadzwyczaj ważna. Z tego właśnie powodu system frejmów można rozpatrywać jako język typu proceduralnego.

Jednakże w rzeczywistości przykłady pomyślnego funkcjonowania modelu tablicy ogłoszeń są ograniczone przypadkami bardzo niewielkiego współdziałania problemów składowych i prostego sterowania źródłami wiedzy. W pozostałych przypadkach sterowanie staje się skomplikowane i jego algorytmy często nawet trudno jest wyraźnie przedstawić. Spróbujemy przeanalizować tę okoliczność, przeprowadzając porównanie z bardziej ogólnymi modelami rozwiązywania problemów w systemie HEARSAY-II, wykorzystującej model tablicy ogłoszeń. Na rys. 2.17 pokazano organizację rozwiązania problemu w systemie HEARSAY-II. W tym wypadku każdy problem składowy (szczegółowy, częściowy) jest związany tylko z problemami stycznymi. Reguły przekształcenia problemów szczegółowych są przedstawione w formie reguł produkcji JEŚLI-TO, źródła wiedzy (ZW) są położone tak, jak pokazano na rysunku. Wybór inicjowanego ZW odbywa się stosunkowo łatwo z rozdziału danych, znajdujących się na tablicy ogłoszeń. Sytuacja kiedy inicjuje się i-te ZW i dane w przesłaniach jego reguł na tablicy ogłoszeń nie występuje, jest niedopuszczalna; trzeba jej zapobiec na poziomie modelu poprzedniego i-1. Trzeba tu pomyśleć i o optymalizacji procesu rozwiązania, jednakże jeśli porównać z wypadkiem, kiedy Z cechuje się skomplikowanymi związkami wzajemnymi, to sterowanie tym procesem



ZW - zdroj wiedzy

Rys.2.17. Model tablicy ogłoszeń



ZW - zdroj wiedzy

Rys.2.18. Model ogólny.

za pomocą tego procesu jest dość łatwe. Na rys 2.18 w charakterze bardziej ogólnej struktury pokazano schemat, w którym problemy lokalne są związane między sobą (nie liniowo, jak na rys, 2.17) w strukturze sieciowej. Taka struktura dobrze zdaje egzamin w takich problemach, jak projektowanie, jednak trudno jest dla nich zbudować algorytmy sterowania. W wyniku tego pojawia się tendencja kolejnego zbliżenia systemów, opartych na model tablicy ogłoszeń, z systemami typu frejmowego. Tendencja ta posiada ważne znaczenie. Celem systemu sztucznej inteligencji jest danie możliwości każdemu człowiekowi łatwego

rozwiązywania problemów na komputerze, które uważa się za trudne przy tradycyjnym opisie proceduralnym. Wspomniana tendencja oznacza, że szerokie rozpowszechnienie systemów przetwarzania wiedzy prowadzi w końcowym efekcie do proceduralnego typu opisu problemów. Choć tendencja ta rzeczywiście ma miejsce, to nie należy z tego powodu popadać w pesymizm, jak również nie należy twierdzić, że stworzone do chwili obecnej metody rozwiązywania problemów wyczerpały swoje możliwości. Bardziej jest to związane z tym, że oddzielne składniki systemów, które pojawiły się do chwili obecnej, zbudowane są na stosunkowo prostych koncepcjach. Dalej rozpatrzmy te przesłanki wejściowe, których gorliwym zwolennikiem powinien stać się każdy kto chce robić postępy w tworzeniu systemów intelektualnych. Po pierwsze, to modułowość wiedzy; po drugie - dalszy rozwój w kierunku systemów o wielkiej skali; po trzecie - to systemy z różnorodnymi źródłami wiedzy. Realizacja tych przesłanek napotyka na wiele trudności, dla ich przezwyciężenia potrzebna jest nowa metodologia. Za nim przystąpi się do rozwiązywania tego problemu, jeszcze raz wrócimy do niedostatków współczesnych systemów inżynierii wiedzy. W określonym sensie jest to ważniejsze od tego wszystkiego, co powiedzieliśmy do tej pory.

## **2.7. Problemy współczesnych systemów inżynierii wiedzy i nowe cele jej przetwarzania**

Wszystkie wykorzystywane w chwili obecnej przedstawienia wiedzy i zbudowane na ich podstawie systemy inżynierii wiedzy cechuje moc opisująca, nie ograniczona wąskimi problemami takimi na przykład, jak gry, rozpatrywane na początku rozdziału. Stało się możliwe stosowanie tych systemów, bez zawężania sfery ich działania w określony sposób, w konkretnych dziedzinach przedmiotowych. Spróbujemy przeanalizować te problemy z trochę innego punktu widzenia.

Na zewnątrz między problemem gry i praktycznie rozwiązywanymi problemami systemów inżynierii wiedzy istnieje określona zbieżność, lecz istnieją między nimi i wielorakie różnice. Jak pokazano w rozdziale poprzednim, znane są bardzo różnorodne formy systemów inżynierii wiedzy, lecz wszystkie one rozpatrywane są jako właśnie forma realizacji

systemu, dlatego dalej spróbujemy wydzielić ich istotę nie wnikając w szczegółowe przypadki. Wspólnym dla tych systemów jest chyba istnienie elementów poszukiwania, izomorfizm formowanych drzew wyszukiwania i wynikające z nich zawile algorytmy poszukiwania. Na przykładzie systemu produkcji pokażemy, że system inżynierii wiedzy posiada takie drzewo poszukiwania, jak na rys 2.3, gdzie zbudowano je dla gry.

Powróćmy do systemu przeprowadzającego wnioskowanie odwrotne. Niech w tym systemie istnieje grupa reguł, zawierających jednakowy wniosek A, któremu odpowiadają różne przesłanki (przyczyny):

$B_1 \rightarrow A$

$B_2 \rightarrow A$

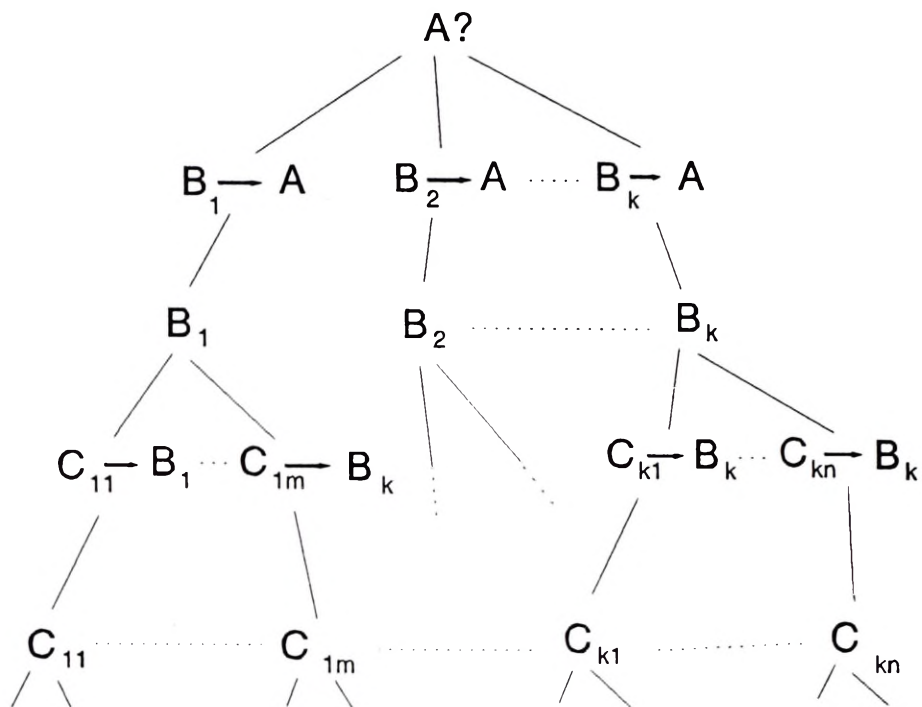
.

.

.

$B_k \rightarrow A$

Załóżmy, że otrzymano pytanie "A?", W tym przypadku zbiór reguł zapewnia osiągnięcie wyniku i dlatego wszystkie one są kandydatami do wniosku a ich przesłanki stają się kluczami następnego poszukiwania. Dalej niech zostanie zadane kilka reguł  $C_{ij} \rightarrow B_i$ ,  $i = 1, 2, \dots, k$ ,  $j = 1, 2, \dots, m$ , posiadających w konsekwencji (wyniku) każde z tych  $B_i$ ,  $i = 1, 2, \dots, k$ . W takim wypadku drzewo poszukiwania które rozpoczyna się z pytania A, przyjmie postać pokazaną na rys 2.19. Pod względem struktury jest ono identyczne to tego, które przedstawiono na rys. 2.3, i metoda poszukiwania będzie taka sama. Istotna różnica między grą a systemem inżynierii wiedzy polega na tym, że przedstawienie wiedzy w tej ostatniej przybiera możliwości opisujące przez odrzucanie pojęcia stanu, zawartego w pojęciu gry. W grze wszystkim stanom zadaje się warunki ocenowe; każdemu stanowi, który osiąga się w procesie rozwiązywania problemu przypisuje się pogląd ocenowy; wyjaśnia się dopuszczalność drogi, znalezionej w wyniku bieżącej próby i jeśli ta droga jest dopuszczalna, to po przejrzeniu jego trasy przebytej do tego stanu wykonuje się kolejny krok, a jeśli trasa jest niedopuszczalna, to osiągnięty stan zostaje wyrzucony z rozpatrywania i wnosi się odpowiednie poprawki. W grze możliwości takiego procesu są określone przez możliwość jego opisu w prostej i jednakowej formie, co oznacza prostotę koncepcji problemu i przedstawienie wszystkich stanów kombinacji niewielkiej liczby parametrów. Systemy inżynierii wiedzy odrzucają to i unikają tym samym podobnych ograniczeń w przedstawieniu wiedzy,



Rys.2.19. Drzewo poszukiwania powstające podczas pracy systemu produkcji

lecz jednocześnie z tym czynią niemożliwym i zastosowanie metody rozwiązania problemu gry. Na przykład, do przesłanek reguł systemu produkcji zakłada się niezbędne z pozycji wykonania działania konsekwentów i posiadające minimalne ograniczenia już istniejące warunki. Przy tym opisu wiadomości powinien dokonać człowiek i jeśli w celu wyprowadzenia określonego wniosku niezbędne są jakieś tam inne, mimo wprowadzonych, warunki, lecz człowiekowi o tym nie wiadomo, to takie przeoczenie odbije się bezpośrednio na regułach. Jedną z zalet reguł w przedstawieniu wiedzy jest ich absolutna elastyczność, lecz nie wykluczona jest przy tym możliwość, że wnioskowanie przeprowadzone przy powtórным wykorzystaniu takich wiadomości będzie prawidłowe, a w systemie brak jest środków, którymi możnaby to sprawdzić. Innymi słowami w systemie takiego typu człowiek sam powinien upewnić się w tym, że wiarygodność każdej oddzielnej wiadomości zapewnia wszechstronne przekonanie o prawidłowości wyciągniętego wniosku albo powinien on sprawdzić prawidłowość otrzymanego wniosku, przez prześledzenie całej drogi wnioskowania. Różnica w formalizacji problemu gry i systemu inżynierii wiedzy polega na tym, że w pierw-

szym wypadku mamy do czynienia z systemem, który "przeprowadza ocenę według wyników", osiągniętych przez wykorzystanie wniosków, a system inżynierii wiedzy wykonuje tylko "ocenę wiadomościami" lub "ocenę według procesu wnioskowania". Okoliczność tę można łatwo wyjaśnić przez porównanie rysunków 2.3 i 2.19. Na rys 2.3 pokazano stany dla każdego węzła pośredniego, i tylko na jego podstawie wyciąga się przypuszczenie o dopuszczalności tego węzła, innymi słowami - czy dopuszczalna jest droga na której osiągnięto ten węzeł. Węzły pokazane na rys 2.19. są przesłankami reguł i reguły te to istota subiektywnych wyobrażeń człowieka, który je wytworzył. Zatem jedynymi tylko regułami nie wolno osiągać oceny.

W przypadku kiedy można "oceniać według wyników", jest nam wszystko jedno na jakie drodze osiągnięto ten wynik. Przy tym wiadomości wykorzystuje się jako środek szybszego osiągnięcia prawidłowego wniosku. Jeśli te wiadomości są nieprawidłowe, to będzie dokonane przejście do stanu niedopuszczalnego, jednakże fakt ten będzie następnie wykryty i dlatego od razu można będzie wrócić do początku wnioskowania. Tym samym efektywność trochę się pogorszy, jednakże nie ma potrzeby niepokoić się o wpływ błędu na wynik końcowy. Oprócz tego jeśli pewna wiadomość jest bardzo dobra to taki fakt będzie zarejestrowany i okaże to swój pozytywny efekt na dalsze zastosowania systemu. W ten sposób ważną zaletą "oceny według wyników" jest możliwość rozwiązania problemu, przez wykorzystanie samej wiedzy. Pogląd o dużej mocy opisującej systemu przetwarzania wiedzy jest wynikiem "oceny według wyników".

Z drugiej strony, problemy grove rozporządzają pełną formalizacją wewnętrznego świata problemowego lub obiektu badań, zatem dokładnie opisują problem aż do najmniejszych detali jego struktury. O systemach inżynierii wiedzy również można powiedzieć, że rozporządzają one jasną dla rozumienia metodą przedstawienia, jeśli tylko nie wdawać się przy tym w detale opisu obiektów. Takie wiadomości przyjęto nazywać powierzchniowymi, ponieważ otrzymywane one są z opisu (w postaci wiedzy) wyników obserwacji zewnętrznego zachowania się obiektów. Na przykład, jeśli do pewnego obiektu przychodzi z zewnątrz oddziaływanie A, to według obserwowalnego odzewu B na to oddziaływanie można, nawet jeśli wewnętrzna struktura obiektu jest nieznana, przedstawić odzew w postaci "JEŚLI A, TO B". Stąd wynika, że nawet przy istnieniu istotnych zmian wewnątrz obiektów lub wewnątrz zadanego świata

problemowego, które są ściśle związane z relacjami opisywanymi jako wiadomości, nie odbije się to na wiadomościach. Metoda systematyzacji takich wiadomości zależy od właściwości dziedziny przedmiotowej, która określa przeznaczenie całego systemu wiedzy. W przypadku kiedy nasza wiedza o dziedzinie przedmiotowej jest nadzwyczaj skąpa i możemy otrzymywać tylko bardzo nieznaczące płytkie wiadomości, detalizacja obiektu aż do jego struktury jest niemożliwa, dlatego trudno będzie negować zastosowanie współczesnych metod. Na przykład, w diagnostyce medycznej - dziedzinie przedmiotowej większości zbudowanych do tej pory systemów przetwarzania wiedzy - obiektem badania jest człowiek, lecz wiedza diagnostyczna nie zawsze była pozyskiwana w drodze wychodzenia od najdrobniejszych wyobrażeń o niej. Większość wiadomości było wynikiem osobistego doświadczenia lekarzy i logiczna podstawa tych wiadomości w niektórych przypadkach może być dość uboga. Innymi słowami, prawie wszystkie systemy diagnostyczne ograniczane były faktem wykorzystania płytkich wiadomości. Jednakże w znacznej części sfery technicznej zastosowanie systemów intelektualnych przyczyniło się to do uzyskania większej wiedzy aniżeli w sferze diagnostyki. Przykładem tego może być budowa maszyn, elektronika, przemysł lotniczy oraz chemia molekularna, metalurgia i inne dziedziny stosowane. Jeśli obiekty reprezentują sobą twory sztuczne, to powinna być znana ich struktura. We wszystkich tych dziedzinach wymagania do rozwiązania zadań nie mogą być zaspokojone w drodze jedynie tylko formalizowania powierzchownych wiadomości - niezbędne są dane, które byłyby związane ze strukturą obiektów. Dlatego na pierwszy rzut oka wydaje się, że system przetwarzania wiedzy powinien przede wszystkim być prawidłowo wykorzystywany zgodnie z jego dziedziną przedmiotową. Jednakże też nie jest to tak. W każdej dziedzinie stosowanej zwykle istnieją tendencje progresywne: na nieznaną przelewa się świat nauki i granice poznania rozszerzają się. Na przykład, w medycynie olśniewającego postępu w ostatnich latach doznały takie przyrządy diagnostyczne jak skanujące tomografy komputerowe, które pozwalają wykryć wewnętrzne związki między stanem organizmu ludzkiego i chorobami. W następstwie tego wiedza medyczna także wychodzi poza ramy wiedzy powierzchownej (płytkiej). Jeśli system przetwarzania wiedzy może przybierać tylko wiedzę płytką, to jest ona niestety pozbawiona sensu istnienia. Niezbędna jest elastyczna adaptacja takiego systemu do zmian świata obiektów i jego otoczenia. Problem polega na

tym, czy ograniczać się płytkim podejściem do wykorzystania systemów czy też wnikać w istotę aż do struktury obiektu, dotyczy metod zastosowania systemu, określanych w zależności od poziomu wiedzy o obiekcie. System powinien być projektowany na podstawie tych wyobrażeń, aby przy jego wykorzystaniu w największej mierze odpowiadał wytworzonej sytuacji. Gry i współczesne systemy inżynierii wiedzy znajdują się na przeciwstawnych biegunach pod względem rozumienia swojego świata problemowego, jednakże wszystkim im brakuje elastyczności. W rozdziale 5 bardziej szczegółowo będzie mowa o osiągniętym postępie w rozumieniu istoty obiektów. Zatem, pierwszym celem przy tworzeniu systemu przetwarzania wiedzy jest narastanie i rozwijanie wiadomości o dziedzinie przedmiotowej, budowa systemu, rozporządzającego wiadomościami o dużej mocy i bogatymi możliwościami opisu metod ich przedstawienia.



## Spis literatury

- [1] Andrew A.M. Artificial Intelligence. "Abacus Press", 1983.
- [2] Bubnicki Z.: Wstęp do systemów ekspertowych. PWN. Warszawa. 1990.
- [3] Hayes-Roth F., Waterman D.A., Lenat D.B. Building Expert Systems. Addison-Wesley Pub.Co. 1983.
- [4] Laurière J.-L. Intelligence Artificielle. "Eyrolles", Paris, 1987, (Tłum. na j.ros. Sistiemy isskustwiennogo intieliekta. "Mir", Moskwa, 1991.)
- [5] Malpas J. A relational Language and Its Applications. "Prentice Hall", Englewood Cliffs, 1987.
- [6] Nilsson N.J. Principles of Artificial Intelligence. Tioga Publ.Co, Paolo Alta, California, 1980.
- [7] Osuga S. Obrabotka znaniy, Wyd. "Omsja", Tokio, 1986 (Tłumaczenie na j. ros. Wyd. "Mir", 1989).
- [8] Osuga S., Saeki Y. Priobrietienije znaniy, Wyd. "Omsja", Tokio, 1987 (Tłumaczenie na j. ros. Wyd. "Mir", 1989).
- [9] Owoc L. (red.). Elementy systemów ekspertowych. Skrypt AE. Wrocław 1991.
- [10] Pospielow G.S. Isskustwiennyj inteillekt - osnowa nowoj informacjonnoj tiechnologii. Izd. "Nauka", 1988.
- [11] Ueho H., Isidzuka M (red.). Priedstawilienije i ispolzowanije znaniy. Wyd. "Omsja", Tokio, 1987 (Tłumaczenie na j. ros. Wyd. "Mir", 1989).
- [12] Waterman D.A.: A Guide to Expert Systems. Addison-Wesley Pub. Comp. 1986.

