

Elżbieta Artowicz

**REPREZENTACJA
WIEDZY
W SYSTEMIE
INFORMACYJNO-
-WYSZUKIWAWCZYM**

WYDAWNICTWO

SBP



NAUKA-DYDAKTYKA-PRAKTYKA

**REPREZENTACJA WIEDZY W SYSTEMIE
INFORMACYJNO-
-WYSZUKIWAWCZYM**

Polish Librarians Association
SCIENCE-DIDACTICS-PRACTICE

Elżbieta Artowicz

**REPRESENTATION OF
KNOWLEDGE IN
INFORMATION RETRIEVAL
SYSTEMS**

Problems of relevance

**WYDAWNICTWO
SBP**



Warsaw 1997

Stowarzyszenie Bibliotekarzy Polskich
NAUKA-DYDAKTYKA-PRAKTYKA

Elżbieta Artowicz

**REPREZENTACJA WIEDZY
W SYSTEMIE
INFORMACYJNO-
-WYSZUKIWAWCZYM**
Zagadnienia relewancji

WYDAWNICTWO

SBP



Warszawa 1997

Komitet Redakcyjny serii wydawniczej
<<NAUKA — DYDAKTYKA — PRAKTYKA>>

Marcin DRZEWIECKI (przewodniczący), Stanisław CZAJKA, Zofia GACA-
-DĄBROWSKA, Janusz KAPUŚCIK, Danuta KONIECZNA, Krzysztof MIGOŃ,
Mieczysław MURASZKIEWICZ, Janusz NOWICKI (sekretarz), Wanda PINDŁOWA,
Maria PRÓCHNICKA, Barbara STEFANIAK, Hanna TADEUSIEWICZ,
Zbigniew ŻMIGRODZKI

**Książka wydana staraniem Stowarzyszenia Bibliotekarzy Polskich
przy finansowym wsparciu Komitetu Badań Naukowych**

Recenzenci:

prof. dr hab. inż. Mieczysław MURASZKIEWICZ
prof. dr hab. Teresa GIERMAK-ZIELIŃSKA

Redakcja techniczna i korekta:
Anna LIS

© Copyright by Stowarzyszenie Bibliotekarzy Polskich

ISBN 83-87629-01-4

CIP — Biblioteka Narodowa

Artowicz Elżbieta

Reprezentacja wiedzy w systemie informacyjno-wyszukiwawczym: zagadnienia rele-
wancji / Elżbieta Artowicz ; Stowarzyszenie Bibliotekarzy Polskich. - Warszawa :
Wydaw. SBP, 1997.- (Nauka, Dydaktyka, Praktyka ; 26)

Wydawnictwo SBP. Warszawa 1997. Wydanie I.
Ark. wyd. 15,47. Ark. druk. 16,75. Skład i łamanie: Ryszard LIPNICKI
Druk i oprawa: Zakład „PRIMUM” ul. Mokronoskich 7a,
05-825 Grodzisk Maz., tel. 755-68-58

SPIS TREŚCI

PRZEDMOWA	9
1. CECHY DYSTYNKTYWNE SYSTEMU INFORMACYJNO-WYSZUKI- WAWCZEGO	13
1.1. Kryteria podziału systemów informacyjno-wyszukiwawczych: struktura zbioru informacji	13
2. TEORIE I INTERPRETACJE RELEWANCJI	22
2.1. Problemy relewancji w filozofii	23
2.2. Relewanca w socjologii fenomenologicznej	26
2.2.1. Relewanca w nauce o prawdopodobieństwie sceptyka Karneadesa z Cyreny	27
2.3. Interpretacje relewancji w logice	31
2.3.1. Relewanca w teorii indukcji i prawdopodobieństwa Rudolfa Carnapa	31
2.3.2. Problemy implikacji we współczesnej logice formalnej: entailment i logiki relewancji	37
2.4. Teoria relewancji Sperbera i Wilson	52
2.5. Interpretacja relewancji w informacji naukowej	65
2.5.1. Relewanca jako relacja między elementami systemu informacyjno-wyszukiwawczego	69
2.5.2. Relewanca jako relacja między zbiorami systemu informacyjno-wyszukiwawczego i jego otoczeniem	76
2.6. Zasady relewancji w systemie informacji naukowej	82
3. REPREZENTACJA WIEDZY W SYSTEMIE INFORMACYJNO-WYSZUKI- WAWCZYM	87
3.1. Wiedza reprezentowana w strukturze języka informacyjno-wyszukiwawczego	87
3.1.1. Relacje paradygmatyczne jako wykładnik językowych kryteriów relewancji	91
3.1.2. Kategoryzacja leksyki JIW jako wykładnik pozajęzykowych kryteriów relewancji	94
3.2. Reprezentacja wiedzy o języku informacyjno-wyszukiwawczym	95
3.2.1. Reprezentacja znaczenia w języku informacyjno-wyszukiwawczym	96
3.2.2. Wykładniki wiedzy o języku i rzeczywistości w gramatyce języka informacyjno-wyszukiwawczego	101
3.3. Reprezentacja wiedzy w językach informacyjno-wyszukiwawczych systemów dokumentacyjnych	103
3.3.1. Wiedza reprezentowana w paradygmatyce języków informacyj- no-wyszukiwawczych	103
3.3.1.1. Wiedza o rzeczywistości w uniwersalnych językach informacyj- no-wyszukiwawczych	104
3.3.1.2. Wiedza o rzeczywistości w języku haseł przedmiotowych (jhp)	107

3.3.1.3. Wiedza o rzeczywistości w specjalistycznych JIW	117
3.3.2. Reprezentacja wiedzy w gramatyce języków informacyjno-wyszukiwawczych	124
3.3.2.1. Reprezentacja wiedzy w gramatyce Uniwersalnej Klasyfikacji Dziesiątej	124
3.3.2.2. Odwzorowanie wiedzy w gramatyce Klasyfikacji Dwukropkowej Ranganathana	127
3.3.2.3. Odwzorowanie wiedzy w gramatyce języka haseł przedmiotowych ..	129
3.3.2.4. Wiedza w gramatyce Międzynarodowej Klasyfikacji Patentowej	131
3.3.2.5. Odwzorowanie wiedzy w gramatyce języków deskryptorowych	134
3.3.2.6. Wiedza w gramatyce kodów semantycznych	137
3.4. Reprezentacja wiedzy w systemach faktograficznych	138
3.4.1. Kognitywne modele przetwarzania informacji	138
3.4.2. Sposoby reprezentacji wiedzy	143
3.4.2.1. Systemy logiczne	146
3.4.2.2. Sieci semantyczne	148
3.4.2.3. Elementarne cechy semantyczne	151
3.4.2.4. Reprezentacje proceduralne	154
3.4.2.5. Reguły produkcji	156
3.4.2.6. Teoria ram i skryptów	158
3.4.2.7. Sieci neuronowe	167
4. DOSTĘP DO WIEDZY W SYSTEMIE INFORMACYJNO-WYSZUKIWAWCZYM	173
4.1. Strategia wyszukiwawcza	175
4.2. Wyszukiwanie informacji bez transformacji semantycznej	185
4.2.1. Strategia wyszukiwawcza w systemie dialogowym	186
4.2.2. Prezentacja kryteriów relewancji w systemie dialogowym	189
4.2.3. Wiedza o zbiorach informacyjnych systemu	193
4.2.4. Prezentacja wiedzy o języku	196
4.2.4.1. Prezentacja leksyki JIW	198
4.2.4.2. Prezentacja reguł gramatyki	203
4.2.4.3. Operatory tekstowe	210
4.2.5. Wiedza o rzeczywistości w systemie online	213
4.2.6. Problemy komunikacji użytkownika z systemem	214
4.3. Wyszukiwanie informacji z transformacją semantyczną	215
4.3.1. Dostęp do wiedzy w języku naturalnym	216
4.3.2. Mechanizmy wnioskowania	222
4.3.3. Wnioskowanie w inteligentnych systemach faktograficznych	222
PODSUMOWANIE	238
ZAŁĄCZNIKI	240
LITERATURA	245
SKOROWIDZ RZECZOWY	255
INDEKS NAZWISK	264

CONTENTS

FOREWORD	9
1. THE DISTINCTIVE FEATURES OF INFORMATION SYSTEM	13
2. THEORIES AND INTERPRETATIONS OF RELEVANCE	22
2.1. Problems of relevance in philosophy	23
2.2. Relevance in phenomenological sociology	26
2.2.1. Relevance and sceptic Carneades probability science	27
2.3. Interpretations of relevance in logics	31
2.3.1. Relevance in R.Carnap theory of induction and probability	31
2.3.2. Entailment and the logics of relevance	32
2.4. D.Sperber and D.Wilson theory of relevance	52
2.5. Interpretations of relevance in information science	65
2.5.1. Relevance as the relationship between the elements of information retrieval system	69
2.5.2. Relevance as the relationship between information retrieval system and its environment	76
2.6. Principles of relevance in information retrieval systems	82
3. REPRESENTATION OF KNOWLEDGE IN INFORMATION RETRIEVAL SYSTEM	87
3.1. Knowledge represented in the structure of information retrieval system	87
3.1.1. Paradigmatic relationships as the exponent of linguistic criteria of relevance	91
3.1.2. Categorization of vocabulary as the exponent of extralinguistic criteria of relevance	94
3.2. Representation of knowledge of information retrieval language	95
3.2.1. Representation of meaning in information retrieval language	96
3.2.2. Exponents of linguistic and extralinguistic knowledge in the grammar of information retrieval language	101
3.3. Representation of knowledge in languages of bibliographic information retrieval systems	103
3.3.1. Knowledge represented through paradigmatic relationships of information retrieval languages	103
3.3.1.1. Knowledge of world in universal information retrieval languages ...	104
3.3.1.2. Knowledge of world in specialised information retrieval languages ..	107
3.3.1.3. Knowledge of reality in professional information retrieval languages ..	117
3.3.2. Representation of knowledge in the grammar of information retrieval languages	124
3.3.2.1 Knowledge in the grammar of the Universal Decimal Classificatio ...	124
3.3.2.2. Knowledge in the grammar of Ranganathan Colon Classification ...	127
3.3.2.3. Knowledge in the grammar of the subjects headings language	129
3.3.2.4. Knowledge in the grammar of the International Patent Classification .	131

3.3.2.5. Knowledge in the grammar of descriptor languages	134
3.3.2.6. Knowledge in the grammar of semantic codes	137
3.4. Representation of knowledge in non-bibliographic information retrieval systems	138
3.4.1. The cognitive models of information processing	138
3.4.2. Means of knowledge representation	143
3.4.2.1. Logical systems	146
3.4.2.2. Semantic networks	148
3.4.2.3. Semantic primitives	151
3.4.2.4. Procedural representations	154
3.4.2.5. Production rules	156
3.4.2.6. Frames and scripts theory	158
3.4.2.7. Neural networks	167
4. ACCESSING KNOWLEDGE IN INFORMATION RETRIEVAL SYSTEMS	173
4.1. Searching strategy	175
4.2. Information retrieval without semantic transformation	185
4.2.1. Searching strategy in dialogue system	186
4.2.2. Presenting the criteria of relevance in dialogue system	189
4.2.3. Knowledge of information system's files	193
4.2.4. Presenting the knowledge of language	196
4.2.4.1. Presenting the information retrieval language's vocabulary	198
4.2.4.2. Presenting the grammatical rules	203
4.2.4.3. Text operators	210
4.2.5. Knowledge of world in online system	213
4.2.6. Problems of user – system communication	214
4.3. Information retrieval with semantic transformation	215
4.3.1. Accessing knowledge in natural language	216
4.3.2. Inference mechanisms	222
4.3.3. Inferencing in intelligent non-bibliographic systems	222
SUMMARY	238
APPENDICES	240
BIBLIOGRAPHY	245
SUBJECT INDEX	255
INDEX OF NAMES	264

PRZEDMOWA

Zagadnienie reprezentacji wiedzy można rozpatrywać w wielu aspektach, zważywszy zróżnicowanie systemów informacyjnych pod względem ich struktury, funkcji, technicznych środków przetwarzania informacji oraz obszaru zastosowań. Zróżnicowanie to może stanowić źródło nieporozumień, z których najczęstszym jest przeciwstawianie *informacji, danych i wiedzy*, dość typowe w obszernej literaturze powstałej w ostatnich dwudziestu paru latach poświęconej sztucznej inteligencji, inżynierii wiedzy oraz informatycznym technologiom przetwarzania informacji. Lektura prac z tego zakresu prowadzi często do złudnego wrażenia jakoby przetwarzanie informacji było czymś z gruntu innym niż reprezentacja i przetwarzanie wiedzy, które mają miejsce tylko w określonych systemach komputerowych z wyrafinowanym oprogramowaniem, mających określone obszary zastosowań, na przykład zarządzanie i wspomaganie procesów decyzyjnych. Wyjaśnienie podobnych stereotypów jest jednym z zamierzeń niniejszej pracy. *Celem głównym* jest przedstawienie językowych (lingwistycznych) sposobów reprezentacji wiedzy szeroko rozumianej, poczynając od tej gromadzonej i przetwarzanej w tradycyjnych systemach bibliotecznych, kończąc na inteligentnych systemach eksperckich.

Układ pracy wyznacza teza mówiąca, że zakres i sposób reprezentowania wiedzy w systemie informacyjnym jest zdeterminowany przez przyjęte dla tego systemu *kryteria relewancji* oraz sposób jej udostępniania użytkownikowi. Z rozwinięciem tej tezy wiążą się trzy założenia określające zakres pracy i kolejność prezentacji omawianych zagadnień.

Założenie pierwsze dotyczy definicji systemu informacyjno-wyszukiwawczego, umożliwiającej wyodrębnienie tych jego elementów, które są istotne z punktu widzenia podjętej problematyki. Nie jest to zadanie proste, istnieje bowiem wiele definicji znanych z literatury przedmiotu, często obciążonych jednak typowymi błędami logicznymi. Za punkt wyjścia przyjęto definicję, w świetle której cechami dystynktywnymi systemu informacyjnego, niezbędnymi i wystarczającymi do odróżnienia od innych systemów, są:

- zbiór informacyjny o określonej strukturze,
- język informacyjno-wyszukiwawczy oraz
- transformacja (przetwarzanie) informacji. [Słownik encyklopedyczny, 1993]

O rodzajach transformacji w systemie decydują przyjęte kryteria relewancji oraz — w znacznej mierze — instrumenty informacyjne, przede wszy-

stkim komputer z oprogramowaniem. Techniczny opis narzędzi informatycznych stosowanych w różnych systemach informacyjno-wyszukiwawczych wykracza poza ramy tej pracy.

Założenie drugie wiąże się z omówieniem różnych interpretacji *relewanccji* jako nadrzędnego kryterium, stanowiącego punkt odniesienia w charakterystykach sposobu reprezentacji informacji (wiedzy), jej selekcjonowania i udostępniania. Relewanccja jako tradycyjna, podstawowa miara sprawności funkcjonowania systemu informacyjnego nie jest homogeniczną kategorią wyłącznie informacyjną. Stanowi przedmiot badań dyscyplin pokrewnych — filozofii, w której rozważania nad relewanccją mają długą tradycję sięgającą starożytności, a także logiki, językoznawstwa, psychologii oraz teorii komunikacji. Do dorobku tych dyscyplin nawiązuje się obecnie w ramach badań kognitywnych (*cognitive science*) oraz w sztucznej inteligencji, w związku z próbami budowy psychologicznie uzasadnionych modeli przetwarzania informacji konceptualnej w pamięci ludzkiej. Do stopniowego różnicowania interpretacji relewanccji w informacji naukowej, początkowo definiowanej jako kategoria absolutna, przyczyniało się z jednej strony wykorzystanie aparatu logiki, na przykład teorii indukcji i prawdopodobieństwa R. Carnapa, z drugiej zaś — rozwój techniki, telekomunikacji i informatyki umożliwiającących stworzenie nowych technik selekcjonowania i udostępniania informacji w systemach online.

Założenie trzecie, wyznaczające sposób prezentacji materiału polega na uznaniu wpływu instrumentów informacyjnych (technicznych środków przetwarzania informacji) na odwzorowanie wiedzy w systemie. Wpływ ten dotyczy reprezentacji wiedzy o języku oraz o rzeczywistości pozajęzykowej i przejawia się w różnych obszarach systemu:

- strukturze zbioru informacyjnego,
- języku informacyjno-wyszukiwawczym,
- procesach udostępniania informacji i jej selekcjonowania.

Rozróżnienie tych obszarów odwzorowania wiedzy ma służyć wykazaniu związku między tendencjami w ewolucji środków językowych w systemach informacyjnych i rozwojem technologii przetwarzania informacji. Wydaje się to istotne dla projektowania języków i systemów informacyjno-wyszukiwawczych, zarówno z teoretycznego jak i pragmatycznego punktu widzenia. Zagadnień tych nie podejmowano dotąd w sposób syntetyzujący, chociaż istnieje bogata literatura w zakresie języków informacyjno-wyszukiwawczych poświęcona cechom strukturalnym języka oraz wiele prac o profilu informatycznym poświęconych organizacji zbiorów informacyjnych systemów (baz danych) i technicznemu aspektom selekcjonowania informacji. Analizowane w pracy przykłady baz danych, których jest już wiele tysięcy służą jedynie ukazaniu typowych cech systemów i języków, a nie ich szczegółowej charakterystyce przedmiotowej czy technicznej.

Praca składa się z czterech rozdziałów.

W rozdziale pierwszym podjęto ogólne zagadnienia reprezentacji i lokalizacji wiedzy w systemie, wprowadzono rozróżnienie wiedzy językowej i po-

zajęzykowej oraz definicje systemów informacji dokumentacyjnej i faktograficznej. Chodziło tu o wskazanie, że odmiennosc denotacji wyrazów języków informacyjnych w tych dwu rodzajach systemów pozostaje w ścisłej zależności z zakresem wiedzy pozajęzykowej i jej miejscem w systemie.

W systemie dokumentacyjnym wiedza o rzeczywistości reprezentowana jest pośrednio, za pomocą języka informacyjno-wyszukiwawczego opisującego cechy dokumentów pierwotnych (tekstów języka naturalnego), zewnętrznych wobec systemu. W systemie faktograficznym na zbiór informacyjny składają się wyrażenia opisujące bezpośrednio obiekty rzeczywistości pozajęzykowej, nazywane faktami. W rozdziale omawiającym odwzorowanie wiedzy w językach systemów dokumentacyjnych (uniwersalnych i specjalistycznych) rozróżniono wiedzę reprezentowaną na osi paradygmatycznej (w słowniku) i syntagmatycznej (w tekstach) języka.

Rozdział drugi jest poświęcony przeglądowi różnych ujęć i teorii relewancji, powstałych na gruncie filozofii, socjologii, logiki, z uwzględnieniem systemów *entailment* i logik relewancji, a także teorii D. Sperbera i D. Wilson reprezentującej kognitywne ujęcie procesów komunikacji. Zwłaszcza ta ostatnia zasługuje na szczególną uwagę projektantów języków i systemów informacyjnych jako reprezentująca wieloaspektowe ujęcie uniwersalnych mechanizmów przetwarzania informacji przez człowieka na poziomie werbalnym (językowym) i mentalnym (konceptualnym). Zamknięciem rozdziału jest przegląd interpretacji w informacji naukowej formułowanych w różny sposób w miarę rozwoju zautomatyzowanych (skomputeryzowanych) systemów informacyjno-wyszukiwawczych i stosowanych w nich technologii przetwarzania informacji. Na tej podstawie sformułowano zasady relewancji informacji w systemie.

Rozdział trzeci zawiera charakterystykę sposobów odwzorowania wiedzy w językach systemów dokumentacyjnych i faktograficznych. W podrozdziałach omawiających zakres wiedzy reprezentowanej w słownikach (na osi paradygmatycznej) języków stosowanych w systemach dokumentacyjnych wyróżniono klasę języków uniwersalnych, z założenia obejmujących całe universum wiedzy oraz języków specjalistycznych, o ograniczonym polu tematycznym i semantycznym. Oddzielnie scharakteryzowano wiedzę wyrażaną za pomocą gramatyki (na osi syntagmatycznej) języków uniwersalnych i specjalistycznych. W podrozdziałach poświęconych wiedzy o obiektach rzeczywistości pozajęzykowej, przetwarzanej w systemach faktograficznych uwzględniono zarówno systemy proste, jak i skomplikowane systemy inteligentne, funkcjonujące dzięki wykorzystaniu formalizmów logicznych i semantycznych.

Rozdział czwarty traktuje o udostępnianiu wiedzy w systemach dokumentacyjnych i faktograficznych. W tej części pracy podstawową kategorią jest strategia wyszukiwawcza, której głównym przeznaczeniem jest odwzorowanie w zapytaniu informacyjnym cech informacji relewancyjnej z punktu widzenia użytkownika. Opis zasad udostępniania informacji ze zbioru systemu opiera

się na rozróżnieniu dwu rodzajów systemów, w których proces wyszukiwania polega na:

— prostej selekcji wyrażen według koincydencji znaków planu wyrażania użytych w opisie obiektów dokumentacyjnych (językowych) i/lub pozadokumentacyjnych (pozajęzykowych),

— generowaniu nowej informacji w wyniku transformacji semantycznej wyrażen należących do bazy wiedzy.

Jako materiał ilustrujący udostępnianie informacji bez transformacji semantycznej wykorzystano przede wszystkim dostępną dokumentację instruktażową systemów online o zasięgu międzynarodowym. Relewancja informacji wyselekcjonowanej przez system interpretowana jest w nich jako wartość komparatywna i zmieniająca się, oceniana w trakcie wyszukiwania w trybie interakcyjnym.

Zasady udostępniania wiedzy w systemach dokonujących transformacji semantycznej informacji, a więc w systemach inteligentnych przedstawiono z uwzględnieniem:

— rozwoju modułów komunikacji użytkownika z systemem,

— wykorzystania systemów logicznych (*logik entailment* i relewancji) jako podstawy transformacji informacji w procesach wnioskowania.

Przyjmując naszkicowane wyżej założenia konstrukcji pracy, autorka adresuje ją do Czytelnika zainteresowanego odpowiedzią na następujące pytania:

- Co znaczy informacja relewantna?
- Jaki rodzaj i zakres wiedzy można wyrazić w danym języku informacyjno-wyszukiwawczym?
- Jakie możliwości i ograniczenia dla wyboru języka wynikają ze stosowanej techniki selekcjonowania i udostępniania informacji?

1. CECHY DYSTYNKTYWNE SYSTEMU INFORMACYJNO-WYSZUKIWAWCZEGO

W literaturze poświęconej systemom informacyjno-wyszukiwawczym często rozróżnia się *przetwarzanie informacji* i *przetwarzanie wiedzy*. Termin pierwszy używany jest głównie w kontekście opisu systemów informacji dokumentacyjnej, natomiast drugi upowszechnił się w związku z rozwojem pewnej klasy systemów faktograficznych nazywanych systemami eksperckimi (ekspertowymi), systemami doradczymi, systemami z bazą wiedzy — systemami inteligentnymi [Popowska H., 1995, 22-30].

Zróznicowanie terminologii staje się źródłem sugestii, jakoby systemy te różniły się zasadniczo przedmiotem przetwarzania (informacja – wiedza), z czym trudno się zgodzić, przedmiotem przetwarzania każdego systemu są bowiem teksty pewnego języka z semantyką, a więc informacja. Istotną natomiast różnicą jest odmienna w opisujących te systemy językach denotacja ich wyrażeń: obiektami denotowanymi przez wyrażenia języka systemu dokumentacyjnego są dokumenty (które zdefiniujemy jako informację utrwaloną na pewnym nośniku), natomiast wyrażenia języka systemu faktograficznego denotują obiekty rzeczywistości pozajęzykowej. Uściślenie to nie jest jednak wystarczające do scharakteryzowania wyżej wymienionych klas systemów, ponieważ za systemy przetwarzające wiedzę uznaje się jedynie pewną podklasę systemów faktograficznych, tzw. systemy inteligentne. Proste systemy faktograficzne typu obiekt-cecha traktowane są jako systemy przetwarzające informację, nie zaś wiedzę.

W związku z tym przeprowadzimy krytyczny przegląd definicji i podziałów systemów przyjmowanych w literaturze przedmiotu.

1.1. KRYTERIA PODZIAŁU SYSTEMÓW INFORMACYJNO-WYSZUKIWAWCZYCH: STRUKTURA ZBIORU INFORMACJI

Terminy *system informacyjno-wyszukiwawczy*, *system informacyjny*, *system wyszukiwania informacji*, *system informacji* są używane do określenia struktur zbudowanych z różnych elementów i relacji, co wynika z zasady względnej niepodzielności elementów systemu. Można je bowiem traktować jako podsystemy (systemów) lub wręcz jako elementy otoczenia systemu.

Jako dość charakterystyczny przykład można zacytować definicję ze *Słownika terminologicznego informacji naukowej*, w którym: system informacyjno-wyszukiwawczy został określony jako „zespół środków umożliwiających wyszukiwanie i przekazywanie informacji”; należą tu język informacyjny, zasady opracowywania i wyszukiwania informacji oraz środki techniczne zapewniające realizację tych procesów, natomiast system informacyjny (system informacji) jako: „system organizacyjny, którego elementami są użytkownicy informacji, pracownicy informacji, zbiory informacyjne oraz zespół metod i środków służących do realizacji procesu informacyjnego” [*Słownik terminologiczny*, 1979]. Odrębne definiowanie tych terminów sugeruje odrębność ich desygnatów, z czym trudno się zgodzić. Zauważmy, że obu definicji nie sposób uznać za poprawne — w definicji pierwszej pominięto tak istotny element systemu jak zbiór informacyjny i jego tworzenie, w definicji drugiej nie uwzględniono między innymi języka jako instrumentu realizacji procesów informacyjnych w systemie i komunikowania się systemu z otoczeniem. W obu definicjach zabrakło tak istotnej cechy SIW jak kryterium relewancji, stanowiącego główne kryterium wyszukiwania (selekcjonowania) informacji, o którym będzie mowa w dalszej części pracy.

Z kolei w sformułowanych w tym samym źródle definicjach terminu wyszukiwanie informacji:

1) „wybieranie ze zbioru informacyjnego tych dokumentów pochodnych i/lub pierwotnych, których charakterystyki wyszukiwawcze odpowiadają instrukcji wyszukiwawczej dla danej kwerendy”,

2) (informat.) „metody i procedury zautomatyzowanego wyszukiwania określonych danych”

pojawia się bądź błąd definicji za wąskiej (definicja pierwsza, w której ograniczenie wyszukiwania do dokumentów eliminuje systemy informacji faktograficznej), bądź błędne koło, jak to ma miejsce w definicji drugiej.

Krytyczna analiza zasygnalizowanych tu definicji systemu informacyjno-wyszukiwawczego i terminów z nim związanych mogłaby posłużyć jako ilustracja problemów interdyscyplinarnej komunikacji w nauce i częstego zjawiska patologii w terminologii naukowej. Problem ten został podjęty w odrębnych pracach [Noniewicz U., 1988].

Definicja systemu informacyjno-wyszukiwawczego, którą przyjmujemy dla celów dalszej analizy i uściśleń terminologicznych oparta jest na konotacyjnym rozumieniu znaczenia, czyli uwzględnia te cechy systemu, które przysługują wszystkim denotatom terminu. Cechą konotacyjną przysługującą wszystkim systemom jest *transformacja informacji* i związane z nią reguły odwzorowania. Definicja taka umożliwia ponadto wyjaśnienie skomplikowanego zagadnienia związanego z typologią systemów informacyjno-wyszukiwawczych (dość oczywistego przy nieostrych definicjach obiektu), w szczególności tradycyjnego podziału na systemy informacji dokumentacyjnej i faktograficznej oraz

nowszej podziału na systemy przetwarzające wiedzę i systemy przetwarzające informację.

Systemem informacyjno-wyszukiwawczym (SIW), systemem informacyjnym, systemem wyszukiwania informacji nazwiemy system przetwarzający, a więc transformujący informację, który z informacji wejściowych (obecnie często w postaci dokumentów) tworzy odpowiednio ustrukturalizowany zbiór informacyjny poprzez odpowiednie transformacje. Na zbiorze informacyjnym dokonuje dalszych transformacji, spośród których najważniejsze jest selekcjonowanie informacji na podstawie określonej dla danego systemu relacji relewancji. Relacja ta określa podobieństwo elementów zbioru informacyjnego do instrukcji wyszukiwawczej, powstającej w wyniku poddania przez system odpowiedniej transformacji zapytania informacyjnego skierowanego przez użytkownika systemu. Informacje spełniające warunek relewancji tworzą zbiór informacji wyjściowej i udostępniane są użytkownikom systemu.

Zależnie od rodzaju informacji wyjściowej, a więc i rodzaju transformacji tworzących z informacji wejściowej zbiór informacyjny, wyróżnimy:

— *systemy informacji dokumentacyjnej* — gdy na podstawie informacji wejściowej tworzy się zbiór metainformacji w postaci charakterystyk wyszukiwawczych. Powstają one w wyniku transformacji streszczania, wyboru słów kluczowych i indeksowania przy użyciu odpowiedniego (używanego w tym systemie) języka informacyjno-wyszukiwawczego, (przy czym streszczanie jest transformacją fakultatywną). System taki daje na wyjściu metainformację;

— *systemy informacji faktograficznej* — gdy na wyjściu system udostępnia odpowiednio wyselekcjonowane informacje wejściowe, będące wyrażeniami denotującymi bezpośrednio obiekty rzeczywistości pozajęzykowej. W wypadku systemów inteligentnych informacje te poddane są odpowiednim transformacjom treściowym (semantycznym) za pomocą operacji wnioskowania i/lub wyjaśniania.

Przy pewnym rozumieniu informacji dokumentacyjnej, jako podzbioru informacji faktograficznej, wszystkie systemy informacyjno-wyszukiwawcze można uznać za systemy informacji faktograficznej. W wypadku, gdy wszystkie transformacje wykonywane są w systemie w sposób automatyczny, pracownicy obsługujący system należą do otoczenia systemu.

Tak więc, za kryterium wyróżnienia w powyższej definicji systemów informacji dokumentacyjnej i faktograficznej przyjmujemy *rodzaj informacji wyjściowej*, udostępnianej użytkownikowi po transformacjach dokonywanych w systemie, którą jest:

— *metainformacja* (informacja o dokumentach): o adresie dokumentu, jego charakterystyce formalnej i treściowej lub

— *wszelka informacja o rzeczywistości nie dotycząca dokumentów*. Przy rozumieniu dokumentu jako specyficznego faktu, systemy informacji dokumentacyjnej można traktować jako podklasę systemów informacji faktograficznej. W większości opisów systemów informacji faktograficznej termin

„faktograficzny” nie jest używany (zwłaszcza w literaturze angielskojęzycznej, gdzie wprowadza się częściej opozycję „bibliographic — non-bibliographic” lub używany jest rzadko (*fact, factual system*). Z kolei w opisach systemów informacyjno-wyszukiwawczych dostarczających metainformacji, ich autorzy na ogół kwalifikują charakteryzowane obiekty w sposób wyraźny jako dokumentacyjne, wyróżniając klasy i podklasy zależnie od przyjętych kryteriów podziału.

Spośród kryteriów najczęściej stosowanych w typologii SIW, oprócz wspomnianego kryterium denotacji informacji wyjściowej, wymienia się [Popowska, 1985, 1988]:

- zakres tematyczny zbiorów, według którego wyróżnia się systemy jednodziedzinowe i wielodziedzinowe (np. system informacji nauk społecznych, biologicznych, itp.), zaś wśród wielodziedzinowych także systemy *transdyscyplinowe*, gdzie w zbiorach przeważa informacja jednej głównej dziedziny (np. medycyny), uzupełniana informacją z dziedzin komplementarnych (w wypadku medycyny np. z chemii, biologii);
- zasięg systemu, według którego wyróżnia się systemy międzynarodowe, krajowe/narodowe, regionalne, lokalne;
- rodzaj nośnika informacji w *systemie*, a więc proporcje udziału w procesach informacyjnych systemu człowieka i instrumentów informacyjnych, głównie komputerowych. Wyróżnia się tu tzw. *systemy manualne* (np. karty dokumentacyjne, karty przezierne, karty dziurkowane, informatory, czasopisma abstraktowe, bibliografie, katalogi), *systemy małej mechanizacji*, w których instrumenty informacyjne systemu nie ograniczają się tylko do przechowywania informacji, lecz także służą do wykonywania prostych transformacji (np. systemy mikrofilmowe, karty perforowane wraz z selektorami), *systemy skomputeryzowane* (historycznie nazywane systemami dużej mechanizacji), w których informacja utrwalona jest na nośniku maszynowym (taśmach, dyskach), a narzędziem dokonującym transformacji informacji jest komputer wraz z oprogramowaniem¹; przy czym systemy skomputeryzowane mogą różnić się stopniem automatyzacji poszczególnych operacji transformowania informacji;
- zasady organizacji zbiorów informacji w systemie. Wyróżnia się organizację prostą, przy której informacje porządkowane są w układzie kreślonym przez jedną cechę formalną lub treściową (adres/sygnatura, autor, temat w haśle przedmiotowym, symbol klasyfikacyjny) lub organizację inwersyjną, przy której informacje są porządkowane w układzie określonym przez więcej niż jedną cechę formalną i/lub treściową (np. unitermy, temat i każdy z jego określników w haśle przedmiotowym, nazwiska autorów, symbole klasyfikacyjne, tytuły czasopism, język czasopism).

¹ W opracowaniach informatycznych poświęconych oprogramowaniu służącego do transformacji informacji w systemie informacyjno-wyszukiwawczym systemem informacji lub systemem informatycznym nazywany jest często pakiet programów (oprogramowanie).

W skomputeryzowanych systemach informacji do opisu struktury zbiorów informacji stosuje się tzw. modele danych, charakteryzujące powiązania między wyrażeniami (tekstami) składowymi zbioru. Modele te dzieli się na: *hierarchiczne*, w których relacje organizujące zbiór ujęte są w postaci drzewa, *sieciowe*, w których wyrażenie (tekst) może być przyporządkowane równocześnie kilku wyrażeniom (tekstom) nadrzędnym oraz *relacyjne*, w których związki między wyrażeniami mają postać co najmniej relacji dwuargumentowych.

O podziale systemów może ponadto decydować:

- rodzaj transformacji na zbiorze informacji wejściowej systemu przy tworzeniu charakterystyk wyszukiwawczych dokumentów (wraz ze stosowanym do tego celu językiem), na przykład streszczanie przy tworzeniu serwisów abstraktowych, wybór słów kluczowych, formułowanie opisu deskryptorowego, wybór dokumentów cytujących dany dokument lub w nim cytowanych przy opracowywaniu indeksu cytowań;
- zakres realizacji funkcji wyszukiwawczych w charakterystykach wyszukiwawczych dokumentów. Według tego kryterium wyróżnia się *systemy pełnotekstowe* zawierające w swoich zbiorach teksty dokumentów pierwotnych, a więc właściwie systemy faktograficzne; systemami pełnotekstowymi nazywane są niekiedy systemy bibliograficzne, w których możliwe jest selekcjonowanie informacji w pełnym tekście charakterystyki wyszukiwawczej, bez ograniczania jej pól składowych (tzw. wyszukiwanie pełnotekstowe);
- sposób przetworzenia informacji udostępnianych na wyjściu systemu, określonej przez reguły odwzorowania zbioru informacyjnego w podzbiór informacji wyjściowej, spełniającej warunek relewancji. W systemach dokumentacyjnych oraz w prostych systemach faktograficznych system dokonuje jedynie wyselekcjonowania tych informacji, które spełniają warunek podobieństwa do instrukcji wyszukiwawczej, natomiast w inteligentnych systemach faktograficznych dokonuje ponadto transformacji treściowej. Systemy dokonujące semantycznej transformacji informacji, a więc generujące na wyjściu nową informację nie zawartą *explicite* w swoim zbiorze, nazywane są systemami eksperckimi (ekspertowymi, doradczymi, systemami z wiedzą) i przeciwstawiane systemom jedynie selekcjonującym informację;
- wpływ użytkownika informacji na przebieg realizacji procesów informacyjnych w systemach skomputeryzowanych. Według tego kryterium wyróżnia się systemy funkcjonujące w trybie interakcyjnym (dialogowym, online, konwersacyjnym) w opozycji do systemów funkcjonujących w trybie wsadowym (partiovym, offline). W tych pierwszych użytkownik ma możliwość ingerowania w przebieg danego procesu informacyjnego, na przykład dokonywania zmiany przyjętej strategii wyszukiwawczej, w tych drugich nie ma takiej możliwości, a ocena relewancji wyselekcjonowanej informacji odbywa się na podstawie końcowych wyników wyszukiwania;

- sposób aktualizacji zbiorów informacyjnych systemu, który w wydawnictwach informacyjnych zwartych (bibliografiach, informatorach) ma charakter skokowy, przeciwstawiany aktualizacji ciągłej.

W starszej literaturze informatycznej wyróżnia się ponadto inne rodzaje trybów realizacji procesów informacyjnych, odpowiadające technicznemu aspektowi funkcjonowania komputera, na przykład w zależności od liczby realizowanych równocześnie programów:

- tryb jedno- i wieloprogramowy.

Zagadnienia te nie należą do przedmiotu naszych rozważań.

Spśród wymienionych kryteriów podziału systemów informacyjno-wyszukiwawczych za relewantne z punktu widzenia przedmiotu naszych uznamy te, które pozwalają wyróżnić:

a) systemy dokumentacyjne i faktograficzne, a więc udostępniające na wyjściu odpowiednio *metainformację* i *informację o obiektach pozatekstowych*;

b) systemy dokonujące jedynie selekcji przechowywanych w nich informacji bez transformacji semantycznej (niededukcyjne) i systemy dokonujące transformacji semantycznej (faktograficzne systemy inteligentne, dedukcyjne);

c) systemy przetwarzające informację w trybie interakcyjnym (dialogowe) i w trybie wsadowym (systemy wsadowe określane są często jako systemy tradycyjne).

Wbrew potocznym intuicjom rozróżnienie „system tradycyjny — system dialogowy” nie jest równoważne opozycji „system nieskomputeryzowany — system skomputeryzowany”. To pierwsze odwołuje się bowiem do trybu komunikowania się użytkownika z systemem i metod dostępu do informacji. Problemowi temu poświęcono wiele uwagi w okresie, gdy selekcjonowanie informacji w systemach realizowano wyłącznie w trybie wsadowym. Na pewien aspekt tego zagadnienia zwróciła uwagę E. Chmielewska-Gorczyca w artykule poświęconym nieostrym interpretacjom tzw. pre- i postkoordynacji w teorii języków informacyjno-wyszukiwawczych [Chmielewska-Gorczyca E., 1981], gdzie prekoordynacją nazywa się:

- koordynację (łączenie) jednostek leksykalnych języka informacyjnego na etapie budowy języka lub
- koordynację (łączenie) jednostek leksykalnych języka informacyjno-wyszukiwawczego na etapie tworzenia charakterystyk wyszukiwawczych dokumentów (indeksowania),

natomiast postkoordynacją — łączenie jednostek leksykalnych na etapie wyszukiwania informacji lub indeksowania.

Autorka ta udowodniła, odwołując się do przykładów z różnych języków systemów informacji dokumentacyjnej, że problem pre- i postkoordynacji sprowadza się w istocie do dostępnych w systemie metod wyszukiwania i możliwości formułowania strategii wyszukiwawczej, dotyczy więc systemu nie zaś języka, przy czym w praktyce podział na systemy pre- i postkoordynowane jest podziałem nieostrym.

W świetle przeprowadzonych przez cytowaną autorkę wywodów kryterium uznania systemu za względnie pre- lub postkoordynowany stanowi (oprócz metod wyszukiwania) odpowiednia organizacja zbiorów systemu, która może zapewniać selekcjonowanie informacji według niewielu wybranych cech (niekiedy tylko według jednej stanowiącej pierwsze wyrażenie charakterystyki wyszukiwawczej dokumentu — na przykład według tematu hasła przedmiotowego, symbolu prostego Uniwersalnej Klasyfikacji Dziesiętnej, autora) lub też wielu cech (a więc także np. tytułu czasopisma, roku publikacji, języka publikacji, wydawcy, afiliacji autora, określników w hasle przedmiotowym i in.). Należy zauważyć, że zakres cech informacji podlegającej selekcji we wczesnych systemach skomputeryzowanych był również ograniczony, podobnie jak w systemach manualnych czy w systemach małej mechanizacji. Rozszerzenie repertuaru cech, według których selekcjonowana jest informacja poprzez odpowiednią organizację zbiorów informacji (inwersyjną) możliwe jest w systemach wykorzystujących różne nośniki informacji i techniczne środki wyszukiwania (ręcznych, małej mechanizacji i skomputeryzowanych). Powstaje wówczas problem nieekonomiczności przetwarzania, co stanowi odrębny aspekt zagadnienia. Uwagi te można odnieść także do systemów informacji faktograficznej, w których możliwości wyszukiwania informacji o cechach obiektów (atrybutach) mogą być zróżnicowane, na przykład dostęp do informacji o osobie w systemie ewidencji ludności można uzyskać tylko przez nazwisko i imię w kartotece o organizacji prostej lub też według innych cech w zbiorze o organizacji inwersyjnej, np. daty urodzenia, zawodu itp.

W świetle powyższych uwag natury terminologicznej oraz przyjętych zasad podziału systemów przedmiotem dalszych rozważań będą:

- interpretacja terminu relewancja,
- relacje między wiedzą o rzeczywistości i zbiorami informacyjnymi systemu,
- sposoby odwzorowania tej wiedzy oraz wiedzy o języku służącym do tworzenia zbioru systemu i jego udostępniania w wyróżnionych klasach systemów. W rozważaniach tych terminami kluczowymi są *strategia wyszukiwawcza* oraz *relewancja*, które omawiamy w dalszych częściach pracy (rozdz. 2 oraz rozdz. 4).

Przyjęte dotychczas ustalenia definicyjne dotyczące poszczególnych rodzajów SIW wykorzystamy obecnie do określenia relacji między zbiorami informacyjnymi systemu a wiedzą o rzeczywistości. W pierwszej kolejności odwołamy się do tych cech, które pozwalają rozróżnić systemy dokumentacyjne i faktograficzne. Odnotujemy równocześnie, że definicje tych systemów, a zwłaszcza systemów faktograficznych, budziły wiele kontrowersji, których źródłem było przede wszystkim używanie w funkcji definiensów nieostrych wyrażen typu „fakt”, „obiekt z atrybutami przyjmującymi wartości liczbowe lub graficzne” (z konotacją „element rzeczywistości pozajęzykowej”) [Bielicka L., 1984]. Przyczyniły się do tego zamętu terminologicznego także prace z zakresu informatyki, w których koncentrowano się na technicznych aspektach manipulowania znakami (planem wyrażania języka) w pamięci komputera, traktując jako nieistotne

rozdzielenie obiektów denotowanych przez wyrażenia języka, w którym kodowana jest informacja.

Problem wyraźnej definicji systemów faktograficznych został sformułowany *explicite* w związku z budową dla nich języków informacyjno-wyszukiwawczych i oceną przydatności języków stosowanych w systemach dokumentacyjnych. Jako cechy odróżniające systemy faktograficzne od systemów dokumentacyjnych wymienia się między innymi sposób opracowania informacji wejściowej w systemie, konieczność jej syntezy i parametryzacji oraz formę odpowiedzi dostarczanej użytkownikowi przez system, zawierającej wartości atrybutów obiektów pozajęzykowych. Cechy te wynikają z relacji między rzeczywistością pozajęzykową i zbiorami informacyjnymi systemu.

Jak stwierdziliśmy wyżej, w systemie dokumentacyjnym wyrażenia języka, w którym zakodowana jest informacja, denotują dokumenty, a więc teksty zawierające informację o rzeczywistości pozatekstowej. Zbiór wiedzy o rzeczywistości stanowi więc bazę danych zewnętrzną w stosunku do zbiorów systemu, zawierających metainformację (należy do otoczenia systemu).

W systemie faktograficznym wiedza o obiektach pozajęzykowych zawarta jest w bazie danych, stanowiącej element systemu. Wyrażenia języka, w którym kodowana jest informacja, denotują obiekty pozajęzykowe, odwzorowując ich pewne relewantne cechy. Sposób pozyskiwania tej wiedzy (najczęściej na podstawie dokumentów) możemy uznać w naszych rozważaniach za problem drugorzędny, istotna jest bowiem denotacja wyrażen języka tego systemu.

W systemie dokumentacyjnym rzeczywistość pozajęzykowa jest odwzorowywana jedynie pośrednio, w stopniu takim, jaki zakłada język pełniący funkcje metainformacyjne (zakres i sposoby odwzorowania wiedzy w językach systemów dokumentacyjnych omawiamy w rozdz. 3).

W systemie faktograficznym rzeczywistość pozajęzykowa reprezentowana jest bezpośrednio za pomocą języka, którego główną funkcją jest funkcja informacyjna. Polega ona na wskazywaniu obiektów zewnętrznych wobec systemu, nie będących dokumentami oraz na określeniu ich cech (atrybutów) stałych i/lub sytuacyjnych oraz związków z innymi obiektami (stałych lub sytuacyjnych — schemat 1).

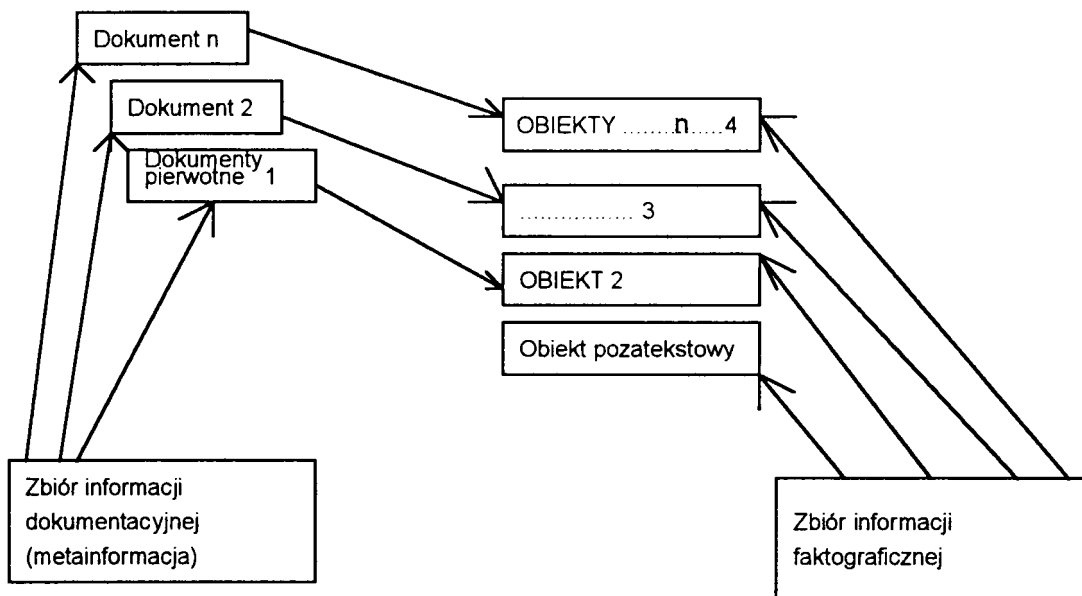
Lokalizacja wiedzy o rzeczywistości — zewnętrzna wobec systemu dokumentacyjnego i wewnętrzna w systemie faktograficznym — wyznacza różne kryteria relewancji informacji udostępnianej użytkownikowi przez system, odwzorowywane w języku informacyjno-wyszukiwawczym.

Pełniący metainformacyjne funkcje język systemów dokumentacyjnych musi wskazywać cechy formalne i treściowe dokumentów. W wielkim uproszczeniu można powiedzieć, że w systemie dokumentacyjnym język ma umożliwiać uzyskanie odpowiedzi na pytanie: *Gdzie, kto pisał o problemie/obiekcie X.*

Język systemów faktograficznych określa jakościowe i ilościowe cechy fragmentów rzeczywistości, a więc ma umożliwiać uzyskanie odpowiedzi na pytanie: **Co wiemy o problemie/obiekcie X.**

Mimo że charakterystyki wyszukiwawcze dokumentów oraz charakterystyki wyszukiwawcze obiektów w prostych systemach faktograficznych (nie-dedukcyjnych) typu obiekt-atrybut mogą różnić się doborem odwzorowywanych w nich cech uznanych za relewantne, to informacja zawarta w takich zbiorach ma charakter statyczny, gdyż przetwarzające ją systemy informacyjne nie dokonują transformacji semantycznej zawartości bazy danych.

Podejście to określane jest w literaturze z zakresu sztucznej inteligencji jako *deklaratywne*. Inaczej dzieje się w systemach inteligentnych (dedukcyjnych), w których informacje o cechach obiektów pozajęzykowych zawarte w faktograficznej bazie systemu stanowią tylko zbiór przesłanek przekształconych w informację wyjściową, za pomocą przyjętych w danym systemie reguł transformacji semantycznej w procesach wnioskowania. Informacja zawarta w bazie danych takich systemów informacyjnych ma więc charakter dynamiczny. Zagadnienia te omawiamy w rozdz. 3 pracy.



Schemat 1. Denotacja zbioru informacji dokumentacyjnej i faktograficznej

2. TEORIE I INTERPRETACJE RELEWANCJI

W świetle definicji słowników języków naturalnych terminy *relewancja*, *relewantry* traktowane są aksjomatycznie i intuicyjnie zrozumiale. Relewancja to relacja:

- *relevant* — ‘connected with what is happening or being discussed’ (związane z tym, co się dzieje lub o czym jest mowa)¹;
- *relevant* — ‘correct or suitable for a particular purpose’ (poprawny lub stosowny do określonego celu)²;
- *relevant* — ‘bearing upon or connected with the matter in hand; to the purpose pertinent’ (odnoszący się do lub związany ze sprawą/rzeczą, o którą chodzi; właściwy dla danego celu)³;
- *relewancja* — związek — zwłaszcza logiczny z rzeczą, o którą chodzi; relewantny — pozostający w (logicznym) związku z, odnoszący się do, dostarczający dowodu słuszności albo niesłuszności sprawy; istotny dla⁴;
- *relewantry* (jęz.) — cechy, opozycje istotne dla funkcji komunikacyjnej języka⁵.

W logice i filozofii relewancja stanowi jedną z najważniejszych kategorii poznawczych. Zmagania z relewancją w logice sięgają — zdaniem niektórych autorów — zarania jej dziejów, jako że wiążą się z teoriami i systemami opisującymi relację wynikania jednych zdań z innych i orzekania o ich prawdziwości w reprezentacjach rzeczywistości pozajęzykowej. W filozofii relewancja uważana jest za podstawową zasadę organizacji wiedzy odwzorowanej w języku naturalnym oraz na poziomie mentalnym. Trudno jest przy tym wskazać wyraźną granicę między interpretacjami logicznymi i filozoficznymi, trudno rozstrzygalny jest bowiem problem relacji między logiką i filozofią.

Przedstawione w rozdz. 2.3. zasady eksplikacji semantycznej przyjęte w teorii indukcji i prawdopodobieństwa R. Carnapa łączą się w sferze problemowej z teorią prawdopodobieństwa sceptyka Karneadesa, do której z kolei nawiązał w latach czterdziestych amerykański filozof A. Schutz, próbujący

¹ *Cambridge International Dictionary of English* (1995), 1197.

² *Oxford Advanced Learners Dictionary of Current English*, (1981), vol 2., 711.

³ *Webster's Encyclopedic Unabridged Dictionary of the English Language* (1989), New York: Portland House, 1212.

⁴ Kopaliński W. (1978), 831.

⁵ *Słownik języka polskiego* (red.) M.Szymczak, (1978), t.3, 42.

określić zasady organizacji wiedzy potocznej w umyśle człowieka. Jego rozważania pozostają z kolei pod wpływem tradycji filozofii fenomenologicznej.

Do wzbogacenia sfery interpretacji relewancji, jako naczelnej zasady organizacji wiedzy człowieka i procesów komunikacji międzyludzkiej, przyczyniły się badania psychologiczne oraz rozwój pragmatyki w językoznawstwie, stawiającej jako problem podstawowy ustalenie warunków efektywnej komunikacji.

Problem ten ma wymiar teoretyczny i aplikacyjny w inteligentnych systemach faktograficznych, albowiem o efektywności funkcjonowania tych systemów decyduje sposób odwzorowania cech obiektów pozajęzykowych oraz reguły użycia tych odwzorowań w procesach inferencyjnych systemu, umożliwiających treściową transformację informacji prowadzącą do dostarczenia użytkownikowi odpowiedzi relewantnej i użytecznej z punktu widzenia jego indywidualnych potrzeb.

2.1. PROBLEMY RELEWANCJI W FILOZOFII

Chcąc zilustrować złożoność problemu relewancji w informacji naukowej i wskazać jego aspekty filozoficzne, można posłużyć się parafrazą „w poszukiwaniu zdrowego rozsądku i uniwersalnej wizji świata, czyli o kryteriach relewancji w systemie informacyjno-wyszukiwawczym”. Do użycia tego sformułowania zainspirowała autorkę praca T. Hołówki poświęcona zdrowemu rozsądkowi i myśleniu potocznemu jako powszechnie przyjętemu zbiorowi elementarnych poglądów na rzeczywistość, tworzących uniwersalną wizję świata odwzorowaną w języku potocznym [Hołówka T., 1986]. Paralelizm problemów i poglądów filozofów opowiadających się za istnieniem zdrowego rozsądku jako „pierwszych zasad danych przez naturę” odwzorowanych w języku (lub też przeciw nim) oraz zadanie dostarczenia przez SIW informacji relewantnej wydają się tylko pozornie odległe. Intuicyjne rozumienie relewancji w języku potocznym zastosowane w odniesieniu do pewnego rodzaju automatów, jakimi są SIW, oznacza przyjęcie kryteriów „zdroworozsądkowych” do wartościowania efektów przetwarzania informacji. Oznacza to założenie uniwersalizmu w prezentacji wiedzy w systemie, umożliwiające wyszukanie informacji relewantnej, tj. o satysfakcjonującej wartości poznawczej. Termin relewancja nie został użyty w omawianej pracy ani razu, ale w kontekście naszych rozważań analogie rysują się dość wyraźnie.

Pytanie o relewancję informacji w początkowym okresie rozwoju SIW, w latach pięćdziesiątych (omawiane dalej w rozdziale 2.5.) można porównać z próbami stosowania kryterium zdroworozsądkowego dla ustalenia podstawowej wizji świata Szarego Człowieka w języku potocznym. Analizując spory filozofów, a w szczególności argumenty i kontrargumenty dotyczące poglądów G. E. Moore'a, przedstawicieli szkockiej filozofii zdrowego rozsądku, szkoły oksfordzkiej i innych, oraz posługując się bogatym materiałem językowym, T. Hołówka

udowodniła tezę swojej pracy, iż zdrowy rozsądek, z kwalifikacją „heterogeniczny”, podlega zasadniczym ograniczeniom, nie pozwala na ukształtowanie systemu wiedzy o dostatecznej wartości poznawczej z racji braków metodologicznych, stanowi przy tym kryterium zmienne czasowo i przestrzennie: konsensus osiągnięty w wyniku dysput filozoficznych wyraża się w stwierdzeniu „pluralizmu zdrowych rozsądków” [Hołówka T., 1986, 91]. Do podobnych wniosków prowadziły — jak to wykażemy — dyskusje nad problemem relewancji w informacji naukowej.

Problem relewancji w filozofii, a właściwie na jej pograniczu z socjologią wiedzy, został *explicite* sformułowany stosunkowo niedawno, z niewielkim wyprzedzeniem w stosunku do informacji naukowej, w pracach A. Schutza [Schutz A., 1970]. Należy jednak zauważyć, że w ujęciu informacyjnym klasyczne problemy filozofii można interpretować jako problemy kryteriów i metod pozyskiwania przez użytkownika relewantnej informacji. Stąd też:

— problem źródeł poznania i spór natywiwistów z empirystami genetycznymi w ujęciu informacyjnym dotyczyłby genezy elementarnej wiedzy przetwarzanej przez człowieka;

— problemy epistemologii w poszczególnych odmianach „empiryzmu — racjonalizmu — irracjonalizmu” można ująć w kategoriach relacji między źródłem informacji i jej użytkownikiem, w kategoriach metody pozyskiwania przez człowieka relewantnej informacji;

— o wymiarze informacyjnym ontologii stanowią konwencje ontologiczne (kryteria relewancji) przyjmowane za podstawę różnych teorii znaczenia. Różnie definiowane kategorie ontologiczne decydują o zaliczeniu znaczenia do sfery zjawisk psychicznych, przedmiotów fizycznych, przedmiotów idealnych, obiektów językowych lub też do sfery relacji między użytkownikiem języka, językiem i opisywaną rzeczywistością [Schutz A., 1970].

W dyskusjach filozofów nad reprezentatywnością i uniwersalizmem wiedzy odwzorowanej w języku potocznym, zaostzonych w związku z hipotezą B. Whorfa o dyktaturze języka w interpretacji świata, sformułowano problem istotny dla naszych rozważań — relacji między zdroworozsądkowymi wizjami rzeczywistości oraz między wiedzą potoczną i naukową. Poglądy formułowane w tych kwestiach można ująć w formie następujących przeciwstawnych konstatacji:

a) wiedza potoczna:

— jest silnie nacechowana empirycznie i behawioralnie, wyklucza z realnego istnienia wszystko, co nie jest doświadczalne sensorycznie [Manheim K., 1952; Nagel E. 1970, Ryle G., 1974];

— wyrasta z elementarnych odruchów poznawczych gatunku ludzkiego, odwzorowuje konstytutywne i niezienne rysy rzeczywistości, tworząc najbardziej skonsolidowaną jej wizję [James W., 1952; Peirce Ch. S., 1960];

b) natomiast w zakresie relacji między wiedzą potoczną, nauką i filozofią:

— istnieje konflikt między myśleniem teoretycznym i potocznym; wiedza teoretyczna zaczyna się tam, gdzie kończy się wiedza zdroworozsądkowa,

myślenie teoretyczne angażuje bowiem nieustanne wyjaśnianie, systematyzację i weryfikację sądów, podczas gdy myślenie potoczne ma charakter nierefleksyjny [Nagel E., 1970; Scheler M., 1973];

— nie ma konfliktu między nauką i filozofią a wiedzą zdroworozsądkową, myślenie naukowe jest ujętym w karby myśleniem potocznym [Korner S., 1974], tworzącym przesłanki dla myślenia teoretycznego i wyrażającym potrzebę porządkowania rzeczywistości [Manheim K., 1974; Polanyi M., 1958];

c) w zakresie struktury wiedzy potocznej:

— z pluralizmem wiedzy zdroworozsądkowej złożonej z wiązek spontanicznych wrażeń łączy się jej uwarstwienie, zróżnicowanie geograficzne i ewoluowanie w czasie. Największą warstwę wiedzy potocznej tworzą sądy przejmowane z poszczególnych kręgów kulturowych w postaci presupozycji ukrytych w kategoriach językowych, rytuałach i instytucjach społecznych, warstwę najbardziej powierzchowną stanowią sądy reprezentowane przez niewielkie grupy (np. społeczność wioski), choć bywa i tak, że tzw. „zdanie ogółu” reprezentuje poglądy kilku osób [Hołówka T., 1986, 94];

— istnieją pewne ramy interpretacyjne (schematy kognitywne, „idiomy”), dzięki którym możliwe jest ukazanie własności strukturalnych wiedzy zdroworozsądkowej i ustalenia tego, co konstytutywne („naturalne, oczywiste”) dla poszczególnych wizji rzeczywistości i które powodują, że rzeczywistość odbieramy jako „gotową, uporządkowaną” [Manheim K., 1974; Polanyi M., 1958];

— różne odwzorowania rzeczywistości w wiedzy potocznej tworzą ciąg perspektyw, wzajemnie się korygujących, wytwarzanych przez poszczególne zbiorowości [Manheim K., 1974;].

O ile można mówić o osiągnięciu w połowie wieku konsensusu w sprawie pluralizmu wiedzy potocznej odwzorowanej w języku, o tyle nie rozwiązany pozostawał problem ustalenia cech konstytutywnych tworzących „jądro, trzon” światopoglądów właściwych poszczególnym kulturom oraz codziennego nastawienia poznawczego człowieka. Chodziło o ustalenie elementarnych jednostek opisu w socjologii wiedzy.

Jako grunt epistemologiczny dla tego nurtu badań przyjęto zasady analizy fenomenologicznej, za pomocą której zamierzano ujawnić cechy strukturalne myślenia potocznego, dane prejęzykowo i funkcjonujące jako korelat języka. Próby określenia zasad selekcjonowania cech rzeczywistości odwzorowanych w języku potocznym podejmował M. Scheler, posługując się nieostrymi kategoriami „składników względnie naturalnych” i „względnie sztucznych”, odpowiadającym pierwiastkom poznania zdroworozsądkowego i naukowego w różnych kulturach [Scheler M., 1973].

Analizy paremiologiczne, zmierzające do wykrycia cech myślenia potocznego w przysłowiach, doprowadziły do stwierdzenia, iż cechuje je niespójność i stanowią raczej narzędzie interakcji, zaspokajające pozapoznawcze potrzeby człowieka [Hołówka T., 1986, 135]. Koncepcje Schelera poddał krytyce Mannheim, wskazując wspomniane niedostatki metodologiczne i trudności

ci z dyskursywnym ujęciem różnic światopoglądowych w różnych kulturach ze względu na ich wzajemne uzależnienie [Manheim K., 1974]⁶.

2.2. RELEWANCJA W SOCJOLOGII FENOMENOLOGICZNEJ

Kolejną próbę ustalenia zasad organizacji wiedzy potocznej tworzącej różne wizje rzeczywistości podjął amerykański filozof Albert Schutz, nawiązujący do tradycji metody redukcji fenomenologicznej Husserla, uważany za twórcę szkoły *socjologii życia codziennego* (socjologii fenomenologicznej, etnometodologii). Przyjmując założenie o pluralizmie odwzorowań wiedzy potocznej, określił swój program badawczy za pomocą pytania „What makes the social world tick”? (Co powoduje, że rzeczywistość społeczną odbieramy jako coś uporządkowanego, gotowego?). W codziennej rzeczywistości nie ma niczego — twierdzi Schutz — czemu nie sprostaby doraźna wiedza zrelatywizowana do konkretnych celów [Schutz A., 1964; 1970]. Decyduje o tym relewancja, „podstawowa kategoria filozoficzna, zasada strukturalizacji rzeczywistości”. Rozważania w tej kwestii zawarte w opublikowanych pośmiertnie pracach Schutza nie spełniają warunków mocnej teorii relewancji. Można je uznać za dokładne postawienie problemu wymagającego rozwiązania w związku z dwiema głównymi tezami autora, zgodnie z którymi:

- nasza wiedza zdroworozsądkowa na temat codziennej rzeczywistości (*the natural attitude*) jest systemem konstruktów tego, co typowe,
- codzienna rzeczywistość (*the lifeworld*) stanowi splot intersubiektywnych „działań aktorów na scenie społecznej”.

Zakres problemów wyznaczany przez te tezy ma bezpośrednio odniesienia do badań nad percepcją i strukturą doświadczenia ludzkiego w psychologii, psychoanalizie, filozofii, a także do teorii znaczenia, których celem jest ustalenie zasad rządzących umysłem w tworzeniu różnych projekcji rzeczywistości, nazywanych przez Schutza „sferami znaczenia” (*provinces of meaning*) i decydujących, iż w danym momencie jedna z projekcji staje się dominująca (stanowi „accent reality”). Pozostając w kręgu tradycji pragmatyzmu i analizy fenomenologicznej Husserla, Schutz podejmuje polemikę z Jamesem, opowiadającym się za interpretacją rzeczywistości jako relacji między emocjami i działaniem człowieka, która to relacja wyznacza różne porządki rzeczywistości (*sub-univers*), a więc przedmiotów fizycznych, religii, teorii naukowych, stanów emocjonalnych. Za uproszczenie uważa psychologiczne ujmowanie struktury rzeczywistości w kategoriach wiary — braku wiary (*belief — disbelief*), twierdząc równocześnie, iż to struktura naszego doświadczenia, nie zaś emocje

⁶ Interesującą ilustrację trudności w tworzeniu modelu poznawczego różnych kultur w antropologii przedstawił E.T.Hall w pracy *Bezgłośny język*, w której równocześnie zaproponował pewien metajęzyk opisu „bezgłośnego języka kultur”, opartego na metodach strukturalizmu. Zob. [Hall E.T., 1987].

ani struktura ontologiczna przedmiotów (jak utrzymywał Husserl), determinuje nasze projekcje rzeczywistości. Związki metodologiczne prac Schutza z fenomenologią wyrażają się w poszukiwaniu jednostek sensu (*sense unities*), przy czym za konieczne uważał Schutz także określenie sposobu konstytuowania się owych jednostek w społecznej interakcji i komunikacji⁷.

Naszą codzienną wizję rzeczywistości — twierdzi Schutz — tworzy wiedza pochodząca z percepcji i doświadczania świata zewnętrznego (zbiór wiedzy podręcznej — *stock of knowledge at hand*) oraz wiedza socjokulturowa, przekazywana przez tradycję (*world handed down to me*). Jest ona zorganizowana w hierarchiczne sfery relewancji poprzez funkcję naszych zainteresowań, określające pewien styl kognitywny, archetyp doświadczenia. Każda ze sfer w projekcji rzeczywistości jest zdeterminowana przez zasadę relewancji. Sfery te (*provinces of meaning*) tworzą w sumie heterogeniczny system różnych relewancji (*the whole systems of relevances*); Schutz nie sformułował definicji relewancji, co stanowi o niepełności jego wywodów. Podjęta przez niego próba wyjaśnienia zasady relewancji w sferze percepcji i zachowań społecznych człowieka jest jednak inspirująca.

2.2.1. Relewancja w nauce o prawdopodobieństwie sceptyka Karneadesa z Cyreny

Schutz odwołuje się do poglądów przedstawiciela sceptyków greckich Karneadesa z Cyreny (około 213-129 p.n.e.) twórcy tzw. trzeciej Akademii, zreferowanych w pismach jego ucznia Klejtomacha oraz stoika Sekstusa Empiryka [*Filozofia starożytna*, 1968]. Karneades, zwalczający dogmatyzm stoików w sferze epistemologii, teologii i etyki, sformułował własną teorię dotyczącą kryteriów fałszu i prawdy, którą przeciwstawiał wszelkim wcześniejszym teoriom filozoficznym. Twierdził, że nie ma absolutnych kryteriów prawdy wobec zawodności rozumu, zmysłów i wyobrażeń. „Tylko takie wyobrażenie stanowi kryterium prawdy, które wydaje się prawdziwe i wystarczająco jasno jawi się jako prawdziwe (...) niekiedy narzuca się również wyobrażenie fałszywe, tak że niekiedy z konieczności trzeba posługiwać się wspólnym wyobrażeniem i fałszu i prawdy” [*Filozofia starożytna*, 1968, 277].

Karneades sformułował własną naukę o prawdopodobieństwie (gr. *pithanon*); dla tego terminu za najbliższy znaczeniowo Schutz proponuje ang. *plausible* — wiarygodny, możliwy do przyjęcia, posiadający pozory prawdopodobieństwa), w której polemizował ze stoikami (Chryzyposem) w kwestii klasyfikacji wyobrażeń (reprezentacji postrzeżeń). Punkt wyjścia stanowiło stwierdzenie konieczności rozróżnienia tego, co jest nam nieznanie od tego, co nie-

⁷ Zdaniem K. Ajdukiewicza: „...to, co fenomenologowie nazywają wglądem w istotę, można też nazwać uważnym wzywaniem się w znaczenie wyrazów” [Ajdukiewicz K., 1985, 71]. Dlatego też „poszukiwanie jednostek sensu” można interpretować jako próbę ustalenia elementarnych cech semantycznych odwzorowujących cechy dystynktywne obiektów rzeczywistości.

pewne. Nie ma prawdy jako takiej, istnieją jedynie prawdy problematyczne: tworzone dla nas, w nas i przez nas prawdy ludzkie. Jako istota mądra człowiek przyjmuje najmniej ryzykowną postawę poznawczą, polegającą na powstrzymaniu się od sądów o rzeczach (termin ten ma inne znaczenie niż to przyjęto później w fenomenologii). Nieosiągalne są reprezentacje postrzeżeń zmysłowych rozumiane w sensie przyjmowanym przez stoików jako obrazujące rzeczy takimi, jakimi są. Stopień przekonania o prawdziwości naszych postrzeżeń jest określony przez motywy i przyczyny działania. Pewność i niepewność nie mają charakteru kontradycji, istnieje bowiem między nimi cała gama stanów pośrednich spowodowana własnościami umysłu, sprawiającymi, iż nie ma w nim „reprezentacji czystych”.

Karneades wyróżnia trzy stopnie wiarygodnego orzekania o względnej prawdziwości wyobrażeń (reprezentacji postrzeżeń) mogących służyć jako motywy lub przyczyny działania człowieka:

1) prawdopodobieństwo zwykle (gr. *pithanai*), występujące w przypadku rozwiązywania problemu przez proste rozstrzygnięcie „tak — nie”;

2) prawdopodobieństwo większe niż zwykle, poparte konkretną reprezentacją sytuacji i połączoną z niedowierzaniem (gr. *pithanai kai aperispastoi* — prawdopodobne i niesprzeczne), np. Herkules przynoszący Admetowi z podziemia żywą Alkestis;

3) prawdopodobieństwo wszechstronnie zweryfikowane przy uwzględnieniu całego zespołu cech postrzeganego przedmiotu (gr. *pithanai kai aperispastoi kai periodeuomena* — prawdopodobne, niesprzeczne i wszechstronnie zbadane).

Jako przykład ilustrujący najwyższy stopień prawdopodobieństwa reprezentacji Karneades podaje kłęb zwinięty w kącie w mrocznej izbie zimą. Przedmiot ten może być uznany przez wchodzącego za kłęb powrozu lub za węża przybierającego zimą kolor szaro-brązowy przypominający powróż. Wiarygodność wyników wszechstronnego badania reprezentacji przedmiotu (procesu „*periodeusis*”) oraz empirycznego ustalania pewności (gr. *diexodos*) decyduje o działaniach podejmowanych przez postrzegającego przedmiot człowieka.

Schutz podkreśla wyraźne analogie między teorią doświadczenia Husserla i poglądami Karneadesa w zakresie określania prawdopodobieństwa prawdziwości sądów przez percepcje i motywacje działań. Elementem wspólnym w poglądach obu filozofów jest założenie, iż w danej sferze naszej świadomości konkurują różne konfiguracje reprezentacji perceptualnych lub imaginacyjnych przeznaczonych do interpretacji problemów („*problematic possibilities*” w terminologii Husserla; *peripastos* — *aperipastos* według Karneadesa). Wszelka nowa wiedza (*actually experienced*) jest porównywana w procesie tzw. biernej syntezy (*passive synthesis*) z wiedzą posiadaną (*already experienced*), z reprezentacjami uprzednich doświadczeń, stanowiących ramę, tło odniesienia (*field of unproblematic horizon*) [Husserl, 1967]. Reprezentacje przechowywane stanowią według Husserla sferę nieustrukturalizowaną (*field of open possibilities*).

Interpretacja (*passive synthesis*) należąca — zdaniem Husserla — do sfery pre-predykcji opiera się na relacjach tożsamości, podobieństwa i in. W wyniku porównania, a więc dokonania wyboru określonej konfiguracji reprezentacji, nowa wiedza może zostać wyróżniona, stać się „theme” (według Husserla) lub wtopić się w tło (*horizon*) jako element tła („field of unproblematic”). W psychologii postaci (*Gestalt*) również zakłada się istnienie w świadomości nieustrukturalizowanej sfery, która jest organizowana w procesie interpretacji dzięki selektywnym własnościom umysłu w przedstawienie czegoś (*Gestalt*) i jego tło. Zdaniem Schutza zasadnicze ograniczenie przedstawień wyraża się w możliwości wyboru tylko jednej interpretacji (konfiguracji) z wyłączeniem interpretacji wieloznacznych. Tym, co powoduje o wyborze określonej konfiguracji interpretacyjnej dla nowej reprezentacji (kłąb powrozu — wąż) jest relewancja uwarunkowana przez posiadaną sytuacyjną „wiedzę podręczną” (*stock knowledge at hand*) oraz historyczną, socjokulturową (*world handed down to me*), które w sumie wyznaczają nieostrą ramę interpretacyjną możliwych doświadczeń (np. pomieszczenia, izby z kłębem „wąż — powróż”). O uznaniu nowej reprezentacji jako dostatecznie relewantnej dla wyróżnienia czegoś jako nowego problemu decydują cechy typowe wspólne dla wiedzy nowej i posiadanej, stanowiące warunki przekładu tego, co nieznanne na znane.

W ustalaniu stopnia nowości reprezentacji rzeczywistości i organizacji pól znaczenia w potocznej wizji świata pierwszoplanową funkcję przypisuje Schutz relacji podobieństwa (*familiarity*) w sferze percepcji, interakcji społecznej oraz motywacji działań ludzkich. Stąd też, przyjmując zasadę typowości jako podstawę strukturalizacji wiedzy, sformułował koncepcję trzech wzajemnie zależnych (*interdependent*) systemów relewancji: *relewancji przedmiotowej*, *relewancji interpretacyjnej* oraz *relewancji motywacyjnej*.

Relewancja przedmiotowa (*topical relevance*) dotyczy wiedzy pochodzącej z percepcji świata zewnętrznego i pozwala wyróżnić to, co nieznanne (*unfamiliar, problematic*) w nieustrukturalizowanym polu wiedzy posiadanej (*unstructuralized field of unproblematic familiarity*). Istnieje praktycznie nieskończona liczba relewancji przedmiotowych ze względu na możliwość organizowania reprezentacji świata według różnych cech obiektów (*theme*), różnych ich konfiguracji oraz hierarchizacji cech i konfiguracji (*subthematization*). Relewancję przedmiotową związaną z hierarchizacją reprezentacji nazywa *wewnętrzną relewancją przedmiotową* (*intrinsic topical relevance*).

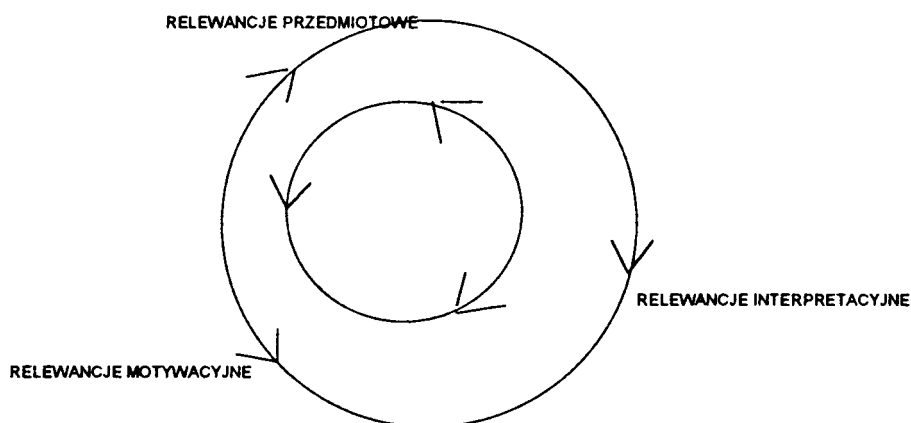
Ze względu na sposób doświadczania rzeczywistości i identyfikowania w niej cech nieznanych — spontaniczny lub kontrolowany — rozróżnione zostały: relewancja wymuszona (*imposed relevance*) — przez szok, zmianę w czasie, przestrzeni, interakcję społeczną oraz niewymuszona, spowodowana dobrowolną selekcją percypowanych obiektów.

Relewancja interpretacyjna (*interpretative relevance, periodeusis* u Karneadesa) określa kwalifikowanie, ocenę postrzeganych obiektów w kontekście rzeczywistości, a więc ich interpretację semantyczną i lokalizację tej interpre-

tacji w określonym polu znaczeniowym (np. interpretacja cech obiektu „wąż — powróż” [Schutz A., 1970, 36]. Podobnie jak w przypadku relewancji przedmiotowych należałoby mówić o relewancjach interpretacyjnych, których system spełnia warunek spójności [Schutz A., 1970, 44].

Relewancja motywacyjna (motivational relevance) określa istotne cechy sposobu postępowania (działania) ustalone w wyniku pewnej decyzji interpretacyjnej. Na przykład, zinterpretowanie kłęba w mrocznym kącie jako węża powoduje, że zbliżający się do tego przedmiotu człowiek zaopatrzy się w kij do ewentualnej obrony, odwołując się do wcześniejszych doświadczeń. Relewancje motywacyjne tworzą wzajemnie powiązany łańcuch, oparty na doświadczeniach motywujących i motywowanych (*in-order-to and because-of motives*), mogą mieć charakter wymuszony lub spontaniczny.

Wyróżnienie trzech rodzajów systemów relewancji jest — jak podkreśla Schutz — uproszczeniem, są one bowiem dynamiczne i współdziałające. Motywacja działań wymaga często dodatkowych interpretacji, te z kolei oznaczają poszukiwania uzupełniającej relewancji przedmiotowej (schemat 2).



Schemat 2. Sfery relewancji

Dynamizm systemów relewancji — twierdzi Schutz — powoduje, że wiedza pozyskiwana ze świata zewnętrznego w procesach interakcji społecznej nie jest homogeniczna ani zintegrowana. Jej elementy nie muszą być spójne. O jej strukturalizacji decyduje:

- stopień wiarygodności (*plausibility*) — od przekonania o pewności, poprzez jej modyfikację, po postawę indyferentną,
- zróżnicowane czasowo i przestrzennie procesy zdobywania nowych doświadczeń,
- procesy inferencyjne, które można opisać w kategoriach logiki,
- sposób odwzorowania procesów interakcji społecznej w poszczególnych systemach relewancji [Schutz A., 1970, 90], przy czym w odwzorowaniu tym znaczenie podstawowe przypisuje się kategoriom wiarygodności (*plausibility*) i podobieństwa (*familiarity*).

Struktura znaczeniowa wiedzy posiadanej (*the meaning — context*) ..., wyznacza zakres wszelkich możliwych problemów relewantnych interpretacyjnie dla rozpatrywanego przedmiotu (*the topic in hand*); zakres ten staje się relewantny motywacyjnie do tego, aby oczywisty dotąd stan naszej wiedzy traktować jako niejasny i problematyczny, wymagający aktualizacji i organizacji” [Schutz A., 1970, 157]. Tak więc, w świetle wywodów Schutza, zasada różnych relewancji określa organizację zbioru wiedzy człowieka, a zakres wiedzy poszukiwanej stanowi projekcję struktury semantycznej tego zbioru.

2.3. INTERPRETACJE RELEWANCJI W LOGICE

Problem relewancji w logice rozważany jest od czasów starożytnych, poczynając od sylogistyki Arystotelesa, i wiąże się z budową systemów logicznych, opisujących metasystemową relację wynikania (inferencji) jednych zdań z innych i przekonującego orzekania o ich prawdziwości jako reprezentacji świata zewnętrznego (por. rozdz. 2.2.1.).

Wyróżnia się dwie interpretacje prawdopodobnego orzekania (prawdopodobieństwa):

- interpretacje subiektywistyczne, relatywizujące prawdopodobieństwo do osoby orzekającej, określane także jako „osobiste, psychologiczne” (*personal probability*), oparte na tzw. funkcji przekonania uwarunkowanej behavioralnie oraz
- interpretacje obiektywistyczne [*Filozofia a nauka*, 1987, 513-518].

2.3.1. Relewancja w teorii indukcji i prawdopodobieństwa Rudolfa Carnapa

Logiczna interpretacja relewancji Rudolfa Carnapa, a w szczególności ujęcia zaprezentowane w tych jego pracach, w których podjął problem relewancji w kontekście prawdopodobieństwa i natury wnioskowania indukcyjnego w naukach empirycznych, wywarły wpływ na późniejsze sformułowania definicji relewancji w informacji naukowej. (zob. rozdz. 2.5.)

Koncepcja prawdopodobieństwa R. Carnapa sformułowana w *Logical Foundations of Probability* [Carnap R., 1962, a także koncepcje przedstawione w pracach jego poprzedników np. J. Keynesa [Keynes J. M., 1921] oraz jego kontynuatora J. Hintikki [Hintikka J., 1966], reprezentują obiektywistyczne interpretacje prawdopodobieństwa, wśród których rozróżnia się interpretacje logiczne i interpretacje statystyczne. Uznając równoprawność metod prawdopodobieństwa statystycznego w nauce, Carnap stwierdzał równoczesną ich nieadekwatność wobec potrzeb szeroko rozumianej logiki indukcji (logiki niededukcyjnej, *non-demonstrative*) niezbędnej w badaniach empirycznych do oceny stopnia potwierdzenia hipotezy lub wniosku (*h*) na podstawie przesłanek lub pewnego świadectwa (*e*-evidence). Sformułowany przezeń system logiki indukcji z prawdopodobieństwem logicznym jako podstawową kategorią określają następujące założenia:

- wszystkie zasady i twierdzenia logiki indukcji są analityczne;
- prawdopodobieństwo logiczne jest relacją między zdaniami (*statements or propositions*), określającą status hipotez naukowych poprzez tzw. funkcję konfirmacji (*confirmation function, c-function*), czyli stopień potwierdzenia pewnych zdań przez inne zdania);
- logiczny charakter funkcji konfirmacji można wyjaśnić przez analogię z relacją implikacji w logice dedukcji między hipotezą (*h*) i przesłanką/świadcstwem (*e*), traktując relację implikacji⁸ w indukcji jako konwers implikacji dedukcyjnej, określane także terminami *entailment, necessary implication*. Rozwiązywanie problemów implikacji zarówno w logice dedukcji, jak i indukcji należy do sfery semantyki logicznej;
- prawdopodobieństwo zdania zależy od języka, do którego ono należy i wyraża się w stosunku do wszelkich możliwych zdań w danym języku;
- wyrażeniami elementarnymi języka są opisy stanów (*state descriptions*) (ζ), odwzorowujące możliwe własności i stany obiektów (*possible worlds* — możliwe światy). Tworzą one pewne klasy (*ranges of sentences*) (\mathfrak{R});
- podstawowe pojęcia logiki (*L-concepts*) traktowane są jako elementarne semantyczne opisy stanów:

Prawda (*L-true, logical truth, necessary truth*) — zdanie atomowe (*i*) jest *L-true* (*f-i*), jeśli dotyczy każdego opisu stanu (?), zatem jeśli (\mathfrak{R}_i) ma zakres uniwersalny.

Falsz (*L-false, logical impossibility, self-contradiction*) — zdanie atomowe (*i*) jest *L-false* (*fvi*), jeśli nie dotyczy żadnego opisu stanu (ζ_i), zatem (\mathfrak{R}_i) jest zerowy.

Implikacja (*L-implication, logical or necessary implication, entailment*) — zdanie atomowe (*i*) implikuje logicznie zdanie (*j*) jeśli dotyczy każdego opisu stanu (ζ), którego dotyczy (*i*), zatem jeśli (\mathfrak{R}_i) jest podklasą (\mathfrak{R}_j). Jest to relacja

⁸ **Wynikanie.** Implikacja w klasycznej logice ujmowana jest jako jeden z rodzajów zdań złożonych różważanych w rachunku zdań, przypominających pod pewnymi względami okres warunkowy w językach naturalnych. Symbol implikacji \rightarrow „*p* \rightarrow *q*” odczytuje się przy pomocy spójnika „jeśli ... to” – „jeśli *p*... to *q*”. Wyróżnia się *Implikację materialną, formalną i ścisłą*, przy czym za podstawową przyjmuje się implikację materialną, jako służącą do zdefiniowania pozostałych. *Implikacja materialna* „*p* \rightarrow *q*” jest prawdziwa wówczas, gdy nie jest tak, że pierwszy człon (poprzednik) jest prawdą, a drugi (następnik) fałszem. *Implikacja formalna* zachodzi między funkcjami zdaniowymi (zawierającymi zmienne wolne) wg schematu $\wedge x[f(x) \rightarrow g(x)]$. Jest prawdziwa, gdy nie jest tak, że spełniony jest poprzednik, a nie jest spełniony następnik. *Implikacją ścisłą* nazywana jest implikacja, w której następnik wynika w pewien sposób z poprzednika, a więc nie może być tak, by poprzednik był prawdą, a następnik fałszem. [*Mala encyklopedia logiki*, 1970, 81-82]. Implikację ścisłą opisują omawiane w rozdz. 2.3.2. logiki modalne Lewisa.

W językoznawstwie podejmowany jest problem *implikacji semantycznej* rozumianej jako własność wyrażen predykatywnych (wyrażen o niepelnym sensie). Własność ta polega na tym, że sensory reprezentowane przez te wyrażenia wymagają uzupełnienia, czyli zamknięcia przez sensory innych wyrażen – np. sensory *mądry* (*x*), *logik*(*x*) implikują po jednym argumente, *żona*, *zna(c)* – po dwie pozycje argumentowe *żona* (*x,y*), *zna(c)*(*x,y*). Szerzej zob. [*Encyklopedia językoznawstwa ogólnego*, 1993, 220-222].

łącząca zdanie (*i*) i (*j*) wówczas, gdy nie jest możliwe, że prawdziwe jest zdanie (*i*) i nieprawdziwe (*j*), czyli gdy (*j*) dotyczy tych przypadków, których dotyczy (*i*).

Ekwiwalencja (*L-equivalence, necessary equivalence, mutual entailment*) — jest definiowana przez identyczność zakresową. Jest to inaczej wzajemna implikacja zdań j_1, j_2, \dots, j_n przy $n \geq 2$.

Alternatywność (*L-disjunctness*) — gdy na mocy konieczności logicznej przynajmniej jedno ze zdań atomowych j_1, j_2, \dots, j_n jest *L-true*; znaczy to, iż alternatywa zdań jest prawdziwa logicznie.

Wykluczanie (*L-exclusion, logical incompatibility, logical impossibility of joint truth*) — relacja łącząca zdania atomowe (*i*) oraz (*j*), kiedy nie istnieje stan którego oba dotyczą. Jest to równoznaczne z koniunkcją zdań fałszywych logicznie [Carnap R., 1962, 83].

Pojęcie⁹ funkcji confirmacji zostało wprowadzone jako nazwa rodzajowa dla pewnej nieskończonej klasy funkcji przyporządkowujących dowolnej parze zdań języka liczbę rzeczywistą z przedziału 0-1 (od fałszu do prawdy), reprezentującą osłabione prawdopodobieństwo wynikania logicznego. W klasie tej wyróżniono pewną podklasę tzw. symetrycznych funkcji confirmacji (*symmetrical c-functions*), odwzorowujących zmianę stopnia confirmacji hipotezy (*h*) na skutek wprowadzenia do korpusu dowodowego (*body of evidence* — (*c*)) nowej informacji (*i*). Stopień ten może się zwiększyć, zmniejszyć lub pozostać niezmienny, zaś nowa informacja nazywana jest odpowiednio: relewantną pozytywnie, relewantną negatywnie lub irrelewantną.

W takim ujęciu **relewantcję** można interpretować jako funkcję definiowaną na podstawie *n*-wartości określającą siłę związku logicznego (semantycznego) między wyrażeniami (zdaniami) pewnego języka. Interpretując relewantcję informacji jako funkcję liczbową, Carnap podjął próbę parametryzacji zmian stopnia confirmacji hipotez pod wpływem nowej informacji za pomocą miar relewantcji. Wykazał przy tym, iż:

— (*i*) jest pozytywna/negatywna lub irrelewantna dla (*h*) przy (*e*), jeśli $r(i, h, e)$ jest pozytywna/negatywna lub irrelewantna;

— przy zastąpieniu (*i*) przez $\sim(i)$ lub (*h*) przez $\sim(h)$ relewantcja zmienia wartość na przeciwną;

— (*r*) ma charakter addytywny w przypadku rozszerzenia korpusu (*e*) o dwie nowe obserwacje (*i, j*) dla hipotezy (*h*) bądź też w przypadku rozważania wpływu (*i*) na dwie hipotezy (*h, k*) przy (*e*).

Na tej podstawie wyróżnił Carnap relewantcję ekstremalnie pozytywną (*extremely positive*), stanowiącą sumę wartości relewantcji informacji (*i*) dla alternatywnych opisów stanu (ξ) będących składnikami szeregu (\mathfrak{R}) wówczas, gdy wartość (*r*) nie jest negatywna dla żadnego ze składników oraz relewantcję całkowicie pozytywną (*completely positive*) dla koniunkcji składników (*conjunctive*), gdy (*r*) nie jest negatywna dla żadnego z elementów składowych.

⁹ Termin pojęcie (*concept*) jest dla Carnapa — jak sam zaznacza — synonimem znaczenia (*meaning*). Jego użycie pozwala odróżnić semantykę terminów od ich planu wyrażania.

System indukcji Carnapa i jego próba logicznej interpretacji prawdopodobieństwa był krytykowany ze względu na ograniczenie do języków bardzo prostych, pozwalających opisywać jedynie cechy obiektów bez uwzględniania relacji między nimi [Filozofia a nauka, 1987, 517-518]. Zarzut ten, z pewnością zasadny z punktu widzenia ciągle nie rozwiązanych problemów mocnej logiki indukcji i prawdopodobieństwa logicznego, z punktu widzenia potrzeb teorii języków informacyjno-wyszukiwawczych wydaje się osłabiony ze względu na wartość metodologiczną teorii Carnapa, co postaramy się dalej uzasadnić. Trzeba też zasygnalizować inspirującą rolę jego metody w rozwoju teorii prawdopodobieństwa, kontynuowanej w systemie Hintikki, różniącym się od systemu Carnapa tym, że zdanie atomowe (opisy stanów) zastąpiono tzw. konstytuentami (zdaniami zawierającymi kwantyfikator), co z jednej strony pozwala uzyskać funkcje konfirmacji Carnapa, z drugiej zaś daje się zastosować do zdań ogólnych (*h*), mających w systemie Carnapa wartości bliskie zeru [Filozofia a nauka, 1987].

Inna koncepcja prawdopodobieństwa logicznego, koncepcja H. Kyburga, nazwana epistemologiczną, oparta jest na relacji semantycznej między pewnym zdaniem a zbiorem zdań stanowiących wiedzę, spełniających przy tym warunek odpowiedniej względnej częstości występowania w zbiorze wiedzy. Wartości funkcji prawdopodobieństwa logicznego stanowią pewne przedziały liczbowe (zamiast liczb) ustalane drogą analizy semantycznej (nie zaś empirycznie) [Kyburg H., 1970].

Formułując zasady prawdopodobieństwa logicznego niezbędne do ustalenia funkcji konfirmacji dla zdań typu: „...eksperyment ponownie potwierdza teorię *T*”, „... dostarcza nowych argumentów za”, teoria jest potwierdzona w stopniu znacznie wyższym przez dane z eksperymentu znane obecnie niż przez dostępne przed dwudziestu laty”, Carnap poświęcił wiele uwagi metodologicznej dyscyplinie swojej teorii, eksponując przy tym rolę semantyki języka nauki (reguł interpretacji), warunkującej poprawność wszelkich procesów inferencyjnych „W momencie gdy wykraczamy poza matematykę formalną przechodząc do wiedzy o faktach z dziedziny przyrody, innymi słowy do nauk empirycznych, do których należy też matematyka stosowana, potrzebujemy czegoś więcej niż rachunku lub systemu aksjomatów; system musi być uzupełniony regułami interpretacji” [Carnap, 1962, 18].

Pomiar relewancji informacji ze względu na funkcję konfirmacji staje się — w świetle wywodów Carnapa — bezpośrednio zależny od interpretacji semantycznej terminów wyjaśnianych w nauce, a zaczerpniętych z języka potocznego (tzw. *prescientific terms*) lub też wprowadzanych do języka nauki (teorii) niezależnie. W szczególności wprowadzanie pojęć nowych powinno odbywać się w dwu etapach:

- 1) formalizacji (aksjomatyzacji), czyli konstrukcji systemu aksjomatycznego,
- 2) interpretacji systemu aksjomatycznego polegającej na określeniu znaczeń elementarnych terminów aksjomatycznych (primitive axiomatic terms),

a więc na ustaleniu reguł semantyki, które niekiedy nazywane są definicjami korelacyjnymi (*correlative definitions*) lub korelacjami epistemicznymi (*epistemic correlations*) [Carnap, 1962, 16].

Jako przykład ilustrujący różnicę między obydwoma etapami podaje Carnap słynny system aksjomatyczny G. Peano dla arytmetyki liczb naturalnych, w którym terminy elementarne pozostają niezinterpretowane (zero, liczba, następnik liczby jest liczbą) oraz system Fregego umożliwiający transformację zdań typu „liczba palców mojej ręki wynosi pięć” w postać nie zawierającą terminów arytmetycznych [Frege G., 1884].

Proces transformacji pojęć (tj. znaczeń) niewyraźnych (*inexact*) na wyraźne (dokładne, *exact*) nazywa Carnap *eksplikacją*, zaś znaczenie terminu wyjaśnianego i wyjaśniającego odpowiednio *explicandum* i *explicatum*. Terminy te są — jak zaznacza — zakresowo szersze niż *definiendum* — *definiens*, choć analogiczne pod względem struktury. Eksplikacja polega na dostarczeniu definicji terminu poprzez skorelowanie jego znaczenia ze znaczeniami innych terminów danego języka, jest więc formą interpretacji. (Interpretacja może przybierać także formę definicji *explicite*, która stanowi wówczas regułę semantyki języka, prezentującą znaczenie terminu jako znaczenie wyrażenia złożonego z terminów semantycznie zinterpretowanych).

Z semantycznym ujmowaniem konfirmacji hipotez naukowych wiążą się w teorii Carnapa cztery zasady niezbędne dla poprawnej interpretacji semantycznej wyrażen pełniących funkcję *explicatum*, a mianowicie:

1) podobieństwa do *explicandum* (*similarity to the explicandum*) „*Explicatum* powinno być podobne do *explicandum* w sposób umożliwiający zamienne użycie; nie wymaga się bliskiego podobieństwa, dopuszcza znaczne różnice” — np. potocznie *fish* (zwierzę żyjące w wodzie) i zoolog. *piscis*;

2) dokładności/precyzji (*exactness*) „Charakterystyka *explicatum*, to znaczy reguły jego użycia (na przykład w formie definicji), powinna być podana w formie dokładnej, umożliwiającej wprowadzenie *explicatum* do spójnego systemu pojęć naukowych”;

3) *produktywności* (*fruitfulness*) „*Explicatum* powinno być semantycznie produktywne (dosłownie pojęciem płodnym), to znaczy użytecznym do formułowania wypowiedzi ogólnych” (...) „pojęcie naukowe jest tym bardziej produktywne, w im więcej wchodzi związków z innymi pojęciami na podstawie obserwowanych faktów — tym większa jest jego użyteczność w formułowaniu praw;

4) *prostoty* (*simplicity*) „*Explicatum* powinno być tak proste na ile to możliwe, to znaczy w stopniu takim, na jaki pozwalają bardziej istotne zasady 1-3” [Carnap R., 1962, 5-7].

W świetle powyższych wymagań ustalenie stopnia konfirmacji hipotezy i miary relewancji informacji z tą hipotezą związanej jest w pierwszym rzędzie kwestią analizy semantycznej wyrażen języka oraz określenia związków między nimi, a dopiero w dalszym planie tworzenia odpowiedniego aparatu for-

malnego i technicznego do badań eksperymentalnych. Przygotowaniu struktury semantycznej języka służy kategoryzacja wyrażeń elementarnych semantycznie (concepts), na podstawie których można orzekać o sile związku logicznego między hipotezą (h) i świadectwem (e). Kategoryzacja ta odwołuje się do tych cech znaczeniowych (elementarnych predykatów), które — zdaniem Carnapa — są w naukach empirycznych szczególnie istotne. Są to:

1) **Cechy klasyfikacyjne** (*classificatory concepts*), stanowiące warunek niezbędny i wystarczający do uznania czegoś za element klasy, służące podziałowi obiektów lub stanów na dwie lub więcej rozłącznych klas. Na przykład w podziale: *metale* — *niemetale*, *metale* — *żelazo*, *miedź*, *srebro* — jako cechy klasyfikacyjne wykorzystuje się własności przedmiotów, natomiast w zdaniu *The person X is acquainted with the field of science Y* (Osoba X jest zaznajomiona z dziedziną Y) cechą tą jest relacja, którą można traktować jako własność pary uporządkowanej.

2) **Cechy ilościowe** (*quantitative concepts*), a więc metryczne, liczbowe, funkcje służące charakterystyce obiektów, zdarzeń lub cech tychże przez przypisanie wartości liczbowych pochodzących z pomiaru lub wyliczenia — na przykład długość, czas, szybkość, masa, siła, temperatura, cena itp. W wielu przypadkach cechy ilościowe charakteryzują własności klasyfikacyjne: temperaturę charakteryzuje własność *ciepłota*, odległość 5 mil — własność *bliskość* itp.

3) **Cechy porównawcze** (*comparative concepts*), nazywane także typologicznymi, porządkowymi, reprezentujące kategorię pośrednią między cechami klasyfikacyjnymi i ilościowymi, zawsze wyrażają pewne relacje — najczęściej dwuargumentową (gdy porównywane są cechy klasyfikacyjne — *ciepło*), niekiedy trójargumentowe — *Peter's knowledge in physics* (Wiedza Piotra w zakresie fizyki) i *Jack's knowledge in history* (Wiedza Jacka w zakresie fizyki) bądź też czteroargumentowe *X is better acquainted with (the field) Y than Z with V* (X jest lepiej zaznajomiony z dziedziną Y niż Z z dziedziną V). Znaczna część cech porównawczych jest odwzorowana w leksyce języka potocznego: *hot*, *warm*, *luke-warm*, *cold*, *more warm* (gorący, ciepły, letni, zimny, cieplejszy) itp.

Przyjmując wymienione cechy semantyczne wyrażeń stanowiących hipotezę lub jej przesłanki, Carnap rozróżnia odpowiednio:

1) konfirmację klasyfikacyjną (*classificatory concept of confirmation*), reprezentującą pewne relacje między hipotezą i przesłanką w aspekcie własności obiektów, zwykle wyrażaną zdaniami „ h jest potwierdzane przez e ”; „ e dostarcza (pozytywnych) dowodów dla h , e jest świadectwem uzasadniającym (wzmacniającym) założenie h ”;

2) konfirmację porównawczą (*comparative concept of confirmation*), odwzorowującą pewne wieloargumentowe relacje między (h) oraz (e) i (h') oraz (e'), wyrażaną zdaniami typu: „ h jest mocniej potwierdzana (wspierana, uzasadniana) przez (e) niż (h') przez (e')”;

3) konfirmację ilościową (*quantitative or metrical concept of confirmation*) (degree of confirmation), reprezentowaną liczbowo przez funkcję $c(h,e)=q$ w przedziale $0-1$.

Formułując zasady scharakteryzowanej tu fragmentarycznie teorii prawdopodobieństwa logicznego, Carnap skoncentrował się na ilościowej interpretacji funkcji konfirmacji, przyjmując ją za naczelną explicatum w logice indukcji, choć równocześnie przyznawał, że jest to interpretacja najbardziej kontrowersyjna [Carnap R., 1962, 23]. Z punktu widzenia różnych interpretacji relewancji w dyscyplinach związanych z teorią języków i systemów informacyjnych, szczególną wartość w systemie Carnapa reprezentują założenia metodologiczne, wiążące konfirmację i relewancję z semantyką języka opisującego pewną rzeczywistość, a ściślej z organizacją jego pola semantycznego, warunkującą wynikanie pewnych zdań z innych, na mocy relacji *L-implikacji* (implikacji logicznej lub *entailment*). Problem metasystemowej relacji wynikania (implikacji) i jej związku z relewancją rozpatrywany jest w systemach dedukcji współczesnej logiki formalnej, które omówimy w następnym rozdziale.

2.3.2. Problemy implikacji we współczesnej logice formalnej: entailment i logiki relewancji

Współczesna logika formalna rozwiązuje problemy podejmowane w logice tradycyjnej, konstruowanej na wzór logiki Arystotelesa, a także w logice klasycznej w systemach bliskich ujęciu *Principia Mathematica* Bertranda Russella. Poszukuje odpowiedzi na pytania dotyczące poprawności rozumowań, przyjmując pewne zasady metodologiczne matematyki: metajęzyk, obiektywizm — polegający na eliminowaniu wszelkich czynników psychologicznych oraz formalizm.

Dotychczasowe sukcesy logiki formalnej w rozwiązywaniu szeregu tradycyjnych problemów filozoficznych (paradoksu klas Russella, definicji prawdy Tarskiego, twierdzenia Goedla o niezupełności) wynikają z precyzji jej języka, a ponadto powodują, że dyscyplina ta jest równocześnie instrumentem i częścią filozofii, badającą najogólniejsze własności wszelkich obiektów [Bocheński J. M., 1988, 25].

Klasycznym problemem podejmowanym przez współczesną logikę formalną (nie-klasyczną) jest ustalanie relacji między treścią i formą wnioskowania, określane jako „paradoks implikacji”. W niesformalizowanych systemach logicznych implikacja ujmowana jest czysto ekstensjonalnie. Przy dwuwartościowej semantyce logiki klasycznej dla uznania prawdziwości zdania złożonego „ p implikuje q ” ($p \rightarrow q$) jest niezbędne i wystarczające, aby nie było możliwe przypisanie na tych samych warunkach wyrażeniu p znaczenia prawdy i wyrażeniu q znaczenia fałszu [Marciszewski W., 1977]. Nieadekwatność tego

ujęcia wykazuje się na przykładzie zdania typu *Jeżeli polecę na księżyc, to pójdę do kina* [Marciszewski W., 1987, 81]¹⁰.

Potrzebę formalnego opisu relacji wynikania wskazał E.G. Moore, pisząc o odwrotności relacji między zdaniem p i q polegającej na tym, iż możemy mówić o q , że jest dedukowalne z p . Zaproponował termin *entailment* jako nazwę tej relacji — p entails q znaczy tyle, co q jest dedukowalne z p , pozostaje z nim w pewnym związku [Moore E. G., 1920]. Termin ten jest obecnie stosowany w trzech znaczeniach:

- jako nazwa spójnika logicznego;
- jako nazwa systemów logicznych konstruowanych metodami współczesnej logiki formalnej i opisujących relację entailment;
- jako nazwa kierunku we współczesnej logice formalnej związanego z systemami typu entailment [Marciszewski W., 1987].

W charakterystykach systemów opisujących relację wynikania terminem o znaczeniu podstawowym jest *relewancja*. Zdaniem autorów klasycznej, choć nie pionierskiej, pracy poświęconej relacji entailment Alana Andersona i Nuela D. Belnapa, klasyczna logika matematyczna zaniedbała problem implikacji materialnej, pozostawiając nierozwiązane zagadnienie relewancji logicznej, często sygnalizowane w aneksach podręczników logiki hasłem „niepowodzenia relewancji” (fallacies of relevance) [Anderson A. R., Belnap N. D., 1975]¹¹.

Systemy logiczne konstruowane z zamierzeniem sformalizowanego opisu warunków, jakie powinny być spełnione, by zachodziła relacja entail-

¹⁰ Problem implikacji jako reguły prawomocnego dowodzenia uzyskał dodatkowy wymiar w związku z potrzebą wyjaśnienia związku między zdaniami języka naturalnego i prowadzonymi na ten temat rozważaniami w filozofii języka. Jest to istotne w sytuacji, gdy związek ten nie jest komunikowany za pomocą wykładników formalnych, odpowiadających logicznemu spójnikowi *jeśli ... to* bądź też, gdy sygnalizowany jest za pomocą różnych wyrażen, takich jak *dlatego, stąd, stąd wynika, w konsekwencji*.

¹¹ Jako przykład ilustrujący istotę problemu implikacji i relewancji przytaczają dyskusję między matematykiem, autorem artykułu poświęconego zagadnieniu przestrzeni Banacha i redaktorem czasopisma. Autor opatruje końcową hipotezę przypisem stwierdzając ...,hipoteza ta dotyczy też innych dziedzin matematyki, czego czytelnik może nie dostrzec bezpośrednio. Na przykład, jeśli nasze przypuszczenie jest prawdziwe, wówczas rachunek funkcjonalny pierwszego rzędu jest zupełny, jeśli jest fałszywe, to implikuje, że ostatnia hipoteza Fermata jest poprawna”. Redaktor akceptuje artykuł, kwestionując przypis jako wprowadzający zamieszanie. Nie widzi związku między końcową hipotezą i innymi dziedzinami matematyki, nie wymienionymi w przypisie. Autor replikuje „Posłużyłem się *if ... then* zgodnie z tradycją logiczną, rachunek funkcjonalny pierwszego rzędu jest zupełny, a zatem cokolwiek implikuje ten fakt jest konieczne; jeśli domniemanie jest fałszywe, to przypuszczalnie jest niemożliwe, zatem implikuje cokolwiek. Skoro kwestionuje pan to użycie, to po prostu nie rozumie pan technicznego sensu *if ... then* pieczołowicie wypracowanego dla nas przez logików”. Argumentuje redaktor: „Rozumiem aspekty techniczne, ale jest to po prostu niepoprawne. Pomimo tego, co mówią logicy, zasady naszego czasopisma wymagają, aby poprzednik wyrażenia *if ... then* był relewantny dla wyciągniętego wniosku. Pan natomiast nie udowodnił, że domniemanie na temat przestrzeni Banacha jest relewantne dla zupełności twierdzenia bądź dla hipotezy Fermata” [Anderson A. R., Belnap J. D., 1975., 17-18].

ment, nazywane są *logikami relewancji* (relevance logics)¹². Wymienia się dwa rodzaje systemów, z których wywodzą się logiki relewancji:

- logiki modalne,
- systemy nawiązujące do koncepcji analityczności I. Kanta, według której zdanie implikacyjne jest zdaniem analitycznym wówczas, gdy znaczenie poprzednika zawiera znaczenie następnika [Marciszewski W., 1987].

Logiki relewancji wykorzystujące kryterium modalności scharakteryzował obszernie N. D. Belnap, przyjmując za punkt odniesienia pięć działów przedmiotowych logiki klasycznej: rachunek zdań, rachunek kwantyfikatorów, arytmetykę związaną z twierdzeniem Goedla, teorię mnogości oraz teorię funkcji rekurencyjnych [Belnap N. D., 1981]. Przyjęty przez tego autora układ prezentacji logik relewancji jest odmienny niż przyjęty w polskiej literaturze [Marciszewski W., 1987]¹³.

Charakterystykę prób rozwiązania paradoksów implikacji materialnej w ramach modalnego rachunku zdań rozpoczynają systemy Lewisa rozwijane przez tego autora wspólnie z Langfordem. Zasadniczym punktem w koncepcji Lewisa było wprowadzenie terminu implikacja ścisła (*strict implication*), którego znaczenie charakteryzowano przez zbiór aksjomatów: koniunkcji — *jest konieczne, że; jest możliwe, że* oraz znaków interpunkcyjnych [Marciszewski, 1987, 311-319].

Interpretację modalnej formuły implikacji według Lewisa, będącej uściśleniem implikacji materialnej, można ująć następująco: *p implikuje ściśle q* znaczy tyle, że nie jest tak, aby *p* było *prawdą* i *q* *falszem*. Początkowy zestaw aksjomatów był następnie modyfikowany. Jako alternatywne odmiany implikacji ścisłej przedstawiono uporządkowane systemy oznaczone symbolami S1 — S5. Na ich podstawie konstruowano systemy nowe drogą wymiany aksjomatów (*axiom-chopping*) oraz modyfikację reguł syntaktycznych — na przykład w systemie T. Feyesa, M. von Wrighta oraz McKinseya-Tarskiego

¹² Termin „relevance logics” został niefortunnie przetłumaczony na język polski jako *logiki relewantne*. Sugeruje to relewantność logik zamiast relewancję jako ich przedmiot. W pracy stosujemy termin logiki relewancji.

¹³ Termin logiki modalne jest różnie interpretowany w literaturze logicznej ze względu na wielowiekową tradycję (por. [Marciszewski W., 1987, 311-319]). Marciszewski wyróżnia modalność aletyczną, deontyczną i epistemiczną. W znaczeniu węższym (modalności aletycznej) uwzględnia się jedynie jednoargumentowe funkcury łącznie z pojęciami konieczności, możliwości i niemożliwości, przy czym nie określono czy chodzi o konieczność logiczną, fizyczną czy też metafizyczną.

Kabziński wyróżnia trzy rodzaje systemów określaných jako logiki relewancji: w znaczeniu węższym, szerszym i najszerszym. Jako podstawową logikę relewancji wskazuje system R skonstruowany niezależnie przez Moha w 1950 roku i Churcha w 1951 roku oraz nadlogiki tego systemu. Do logik relewancji w znaczeniu szerszym zalicza logiki typu R oraz systemy entailment (E) reprezentowane przez system N. R. Andersona i N. D. Belnapa, związane z logikami modalnymi. Jako logiki relewancji w znaczeniu najszerszym wymienia wszelkie systemy logiczne realizujące postulat relewancji, a więc typu R.E. oraz systemy implikacji analitycznej (PAI, DAI) [Marciszewski W., 1987].

[Belnap N. D., 1981]. Równocześnie toczyły się dyskusje na temat teoretycznych i aplikacyjnych wartości logik modalnych. Jak stwierdza N. D. Belnap, wyrażane opinie oscylowały między akceptacją (Quine, Russell) a odrzuceniem (A.Prior). Podejmowano także prace nad formalizacją logik modalnych metodami algebraicznymi oraz nad ich opisem semantycznym (m. in. K. Gödel, McKinsey, A. Tarski). Metodę *semantycznej interpretacji logik modalnych*, stosowaną także w innych dziedzinach wiedzy, przedstawił S. Kripke [Kripke S. A., 1959]. Charakteryzują ją następujące założenia:

- prawdziwość zdania nie jest kategorią absolutną; jest zrelatywizowana do pewnego zbioru wskaźników (indeksów) nazwanych możliwym światem (*possible world*);
- jako parametr struktury semantycznej języka stosuje się pewną relację binarną Rxy , interpretowaną jako możliwość zrelatywizowaną do światów.

Prawdziwość zdania *konieczne A w świecie (w)* powinna być określana poprzez jego prawdziwość w każdym (w'), który można zrelatywizować do (w). Po nałożeniu pewnych ograniczeń na relację R (np. symetryczności, przechodności) Kripke przeprowadził analizę semantyki różnych systemów modalnych, wskazując że prawdziwość zdania jest zdeterminowana przez (porównywalne) reguły semantyczne różnych języków.

Prezentację rozwiązań implikacji na gruncie modalnego rachunku zdań ograniczymy do wybranych zagadnień dość elementarnych, niezbędnych do scharakteryzowania logik relewancji w kontekście ich użyteczności w systemach informacyjnych. Jak zauważa Belnap, konstruowanie logik modalnych stanowiących nadbudowę klasycznych systemów Lewisa i Langforda, a także struktur Kripkego zaczęto w pewnym momencie traktować jako swoisty sport polegający na wprowadzaniu drobnych zmian i alternatywnych technik.

Rozwój logik relewancji opartych na rachunku zdań i modalności wynikał z niedoskonałości modalnego rachunku zdań. Koncepcja ścisłej implikacji pozwalała częściowo eliminować paradoksy implikacji materialnej, nie była jednak wolna od paradoksów typu *kontradykcja implikuje wszystko* lub *cokolwiek implikuje konieczną prawdę* — *Nie jest możliwe, że Ziemia jest okrągła i nie jest tak, że Ziemia jest okrągła, i nie jest tak, że $2 + 2 = 4$* . [Marciszewski W., 323, 340].

Autorem inspirującej reinterpretacji implikacji ścisłej, umożliwiającej częściowe eliminowanie paradoksów, był R. Ackerman. Jego koncepcja tzw. *mocnej implikacji* polegała na takim zawężeniu ścisłego implikowania, aby okres warunkowy: jeżeli p to q był prawdziwy jedynie wówczas, gdy p i q są związane w sposób konieczny i relewantny (*necessarily and relevantly*). Problemem dyskusyjnym był jednak brak matematycznej interpretacji relewancji. Za niesatysfakcjonujące uznawano opisowe wyjaśnienie, że p implikuje relewantnie q wówczas, gdy może być użyte w dyskusji, z której wnioskiem jest q [Ackerman R., 1956].

Chcąc wyjaśnić nieostry termin „użycie” Ackerman posłużył się relacją *entailment*. W skonstruowanym systemie **P'** zdanie zależne w dowodzie hipotez przy dedukcji naturalnej opatrywał wskaźnikami pozwalającymi stwierdzić, które z kolejnych kroków dowodu są relewantnie zależne od danej hipotezy bądź też umożliwiają śledzenie wykorzystania danej hipotezy do osiągnięcia danego kroku dowodu [Belnap N. D., 1981]. W ten sposób modalną „konieczność” zdefiniowano za pomocą relacji *entailment*, określającej konieczność występowania w poprzedniku i następniku okresu warunkowego *Jeżeli ... to* wspólnej zmiennej w postaci spójnika *entailment (entailment connective)*, zapewniającego minimalny warunek relewancji [Marciszewski W., 1987, 341].

Autorzy prac omawiających logiki relewancji podkreślają, że koncepcja spójnika *entailment* wykracza poza logiki modalne w sensie Lewisa i właśnie dzięki relacji *entailment* możliwe jest zdefiniowanie klasycznego funktora konieczności modalnej. Wykazano, że relacja ta ma charakter semantyczny nie może bowiem być tak, aby pewna nierelacyjna własność *A* połączona z pewną nierelacyjną własnością *B* była wystarczająca dla formuły *A entails B*.

Koncepcję mocnej implikacji Ackermana wraz z rachunkiem **P'** wykorzystali Anderson i Belnap, konstruując system **E** (*rachunek entailment*), którego założenia określili następująco [Anderson N. R., Belnap N. D., 1975]:

a) regułami wnioskowania w systemie **E** (*that ... entails*) są reguły dedukcji naturalnej, a więc takiej, w której proces dowodzenia twierdzeń polega na stosowaniu reguł z pominięciem aksjomatów. Dla każdego spójnika relacji *entailment* istnieje jedna reguła uzasadniająca jego wprowadzenie (*I — entailment introduction*) lub jego użycie (*entailment elimination — „ $\rightarrow E$ ”, modus ponens). Interpretując $A \rightarrow B$ jako „*A entails B*” lub *B jest dedukowalne z A*, zakłada się asercję $A \rightarrow B$ wszędzie, gdzie ma miejsce dedukcja *B z A*;*

b) podstawową, archeotypową formą wnioskowania jest $A \rightarrow A$ (prawo identyczności) należące do czterech aksjomatów schematu:

$A \rightarrow A$ (*identity — identyczność*)

$A \rightarrow B \rightarrow B \rightarrow C \rightarrow A \rightarrow C$ (*transitivity — przechodniość*)

$(A \rightarrow B \rightarrow C) \rightarrow B \rightarrow A \rightarrow C$ (*permutation — permutacja*)

$(A \rightarrow B \rightarrow C) \rightarrow A \rightarrow B \rightarrow A \rightarrow C$ (*self-distribution — samodystribucja*).

Wspólną cechą systemów spokrewnionych z systemem **E** jest warunek wspólności znaczenia poprzednika i następnika formuły implikacyjnej. Formuła $A \rightarrow B$ może być uznana za twierdzenie wówczas, gdy *A* i *B* pozostają w związku znaczeniowym przynajmniej w zakresie minimalnym. Brak tego związku powoduje niepowodzenia relewancji (*fallacies of relevance*).

Warunek minimalnego związku znaczeniowego stanowi podstawową zasadę w systemie **R**, skonstruowanym niezależnie przez Moha w 1950 r. i Churcha w 1951 r. Rachunek **R** uważany jest za podstawową logikę relewancji, różniącą się od systemu **E** tym, iż stosuje się w nim kategorie modalności. Został nazwany przez Churcha pozytywnym rachunkiem słabej implikacji (*weak posi-*

ive implication calculus) i scharakteryzowany następująco [Anderson N. R., Belnap N. D., 1975]:

..., Jeśli istnieje dowód dla B przy hipotezach A_1, \dots, A_n , w którym wszystkie A_1, \dots, A_n są stosowane przy dochodzeniu do B , wówczas istnieje dowód $A_n B$ z A_1, \dots, A_{n-1} spełniający ten sam warunek [Anderson N. R., Belnap N. D., 1975, 21].

System **R** opiera się na czterech aksjomatach:

$R \rightarrow 1$ $A \rightarrow A$ (*identity – identyczności*)

$R \rightarrow 2$ $A \rightarrow B \rightarrow C \rightarrow A \rightarrow C \rightarrow B$ (*transivity – przechodności*)

$R \rightarrow 3$ $(A \rightarrow B \rightarrow C) \rightarrow B \rightarrow A \rightarrow C$ (*permutation – permutacji*)

$R \rightarrow 4$ $(A \rightarrow A \rightarrow B) \rightarrow A \rightarrow B$ (*contraction – kontrakcji, ściągnięcia*).

Jako reguły dedukcji stosuje się reguły odrywania oraz adiunkcji (koniunkcji) [Anderson N. R., Belnap N. D., 1975, 20].

Żaden z wymienionych systemów logik relewancji nie jest systemem zupełnym (*the full relevance logics*). Zdaniem Belnapa przyczyn tego stanu rzeczy należy upatrywać w złudnym założeniu, że można zbudować system rozstrzygalny (*decidable*) na poziomie rachunku zdań, posługując się kategorią relewancji, która jest relacją semantyczną zawsze prowadzącą do problemu podejmowania decyzji o wyborze znaczeń wyrażeń występujących w poprzedniku i następniku formuły implikacyjnej.

Trafność tego spostrzeżenia potwierdziły badania R. K. Meyera, który opracował wspólnie z R. Routleyem system nazwany *semantyką relacyjną* (*relational semantics*), nawiązujący do koncepcji semantyki Kripkego [Meyer R. K., Routley R., 1972].

W semantyce Meyera-Routleya prawdziwość zdania implikacyjnego jest również zrelatywizowana do pewnego zbioru wskaźników (*indeksu*). Zrezygnowano jednak z kontrowersyjnej terminologii Kripkego, na przykład z podziału na światy „możliwe — niemożliwe”, „normalne — nienormalne”. W opisie indeksów służących do oceny prawdziwości formuł zdaniowych wykorzystywane są pewne układy (*set-ups*). Układy te mogą być niedookreślone i niepełne (*underdetermined, incomplete*), co znaczy, że nie muszą zawierać zdania A ani jego negacji, bądź też mogą być nadokreślone (*overdetermined*) i niespójne (*inconsistent*), czyli mogą zawierać zarówno zdanie A jak i jego negację. Ten rodzaj układów, wyróżniony obok układów normalnych (*normal setups*) odwzorowujących świat rzeczywisty (*real world*), może być wykorzystany do opisu sytuacji niezupełnie określonych lub niespójnych, pozwalając przy tym na wyeliminowanie ze zdań matematyki czynników psychologicznych¹⁴. Orzekanie o logicznej prawomocności zdań dotyczy jedynie „układów normalnych”.

¹⁴ Można na przykład wyobrazić sobie, że twierdzenie „czterech barw” jest fałszywe, jakkolwiek udowodniono za pomocą „bezprzykładowych obliczeń komputerowych”, że jest prawdziwe [Anderson N. R., Belnap N. D., 1975].

W analizie implikacji relewantnej w systemie Meyera-Routleya stosowana jest — zamiast binarnej relacji Rxy jak to ma miejsce w semantyce Kripkego — relacja trójczłonowa (*three-termed*) $Rxyz$. Jest interpretowana następująco: x i y są współmożliwe ze względu na z (W systemie Kripkego $Rxy - y$ jest możliwe ze względu na x).

Zdanie implikacyjne $A \rightarrow B$ jest prawdziwe przy zbiorze wskaźników (indeksie) (x) wówczas, gdy dla każdego (y) oraz (z) takiego, że $Rxyz$, jeśli A jest prawdziwe przy y , to B jest prawdziwe przy z [Belnap N. D., 1981, 143]. Meyer wykazał, że zmieniając postulaty dotyczące układu reprezentowanego przez trójczłonową relację współmożliwości, można opisać każdą z dotychczas skonstruowanych logik relewancji.

Logika relewancji nawiązująca do koncepcji analityczności I. Kanta reprezentuje system implikacji analitycznej W. Parry'ego. Punktem wyjścia było słynne zdanie „...,związek orzeczenia z podmiotem wyraża się przez identyczność” [Kant I., 1781], w myśl którego analityczność wyjaśniano „ S jest P jest zdaniem analitycznym w przypadku, gdy podmiot jest w pewnym sensie częściowo identyczny z predykatem”. Zdanie to utrzymało swą prawomocność do dziś.

Sformułowanie Kanta wykorzystał W. Parry, budując system PAI, w którym podstawowym warunkiem implikacji jest zawieranie się znaczenia następnika w znaczeniu poprzednika formuły zdaniowej. Jego założenia są następujące (według wersji podanej przez Andersona i Belnapa [Parry W. T., 1933]).
Aksjomaty:

$\rightarrow, \&, \sim$, oraz uzupełnione przez Andersona i Belnapa $\Leftrightarrow V$

Postulaty :

$\&$:

- 1) $(A \& B) \rightarrow (B \& A)$
- 2) $(A \rightarrow (B \& C)) \rightarrow (A \rightarrow B)$
- 3) $(A \rightarrow B) \& (A \rightarrow C) \rightarrow A (B \& C)$

\sim :

4) $A \rightarrow \sim \sim A$

oraz $\&$:

- 5) $(A \rightarrow C) \& (B \rightarrow C) \rightarrow (A \vee B) \rightarrow C$
- 6) $(A \& (B \vee C)) \rightarrow (A \& B) \vee (A \& C)$

oraz \sim :

7) $A \rightarrow B \sim A \vee B$

$\&$, oraz \sim :

8) $(A \vee (B \& \sim B)) \rightarrow A$

Formuły arbitralne (... A ...) zawierające A

9) (... A ...) $\textcircled{R} A \textcircled{R} A$

10) $((A \Leftrightarrow B) \& (... A ...)) \rightarrow (... B ...)$

Reguły :

$\rightarrow E$ oraz $\& I$.

Podstawowe twierdzenie systemu **PAI** brzmi: Jeśli $A \rightarrow B$ jest dowodliwe w systemie, to wszystkie zmienne zawarte w B występują także w A .

System ten wywołał wiele dyskusji. Goedel podnosił problem semantycznej interpretacji składni, pisząc „... p implikuje analitycznie q ” można przypuszczalnie interpretować następująco: q jest wywodliwe z p i z aksjomatów logicznych i nie może zawierać żadnych innych pojęć niż p , a to oznaczałoby, gdyby sprecyzować tę definicję, zmierzanie do całkowitej dowodliwości aksjomatu Parry’ego w sensie, że wszystkie zdania prawomocne w powyższej interpretacji są wyprowadzalne” [149]¹⁵.

Przyczyn krytycznych uwag wobec systemu **PAI** należałoby dopatrywać się w niejasnym sformułowaniu Kanta o „częściowej identyczności”, z którym wiąże się problem kryteriów podobieństwa znaczeniowego wyrażeń stanowiących poprzednik i następnik formuły implikacyjnej. Jest on nierozstrzygalny przy ekstensjonalnej interpretacji znaczenia. W związku z tym J. M. Dunn, autor systemu **DAI**, będącego modyfikacją systemu Parry’ego, twierdził, że system **PAI** jest faktycznie systemem modalnym, równoważnym systemowi **S3** Lewisa.

Dunn zaproponował „demodalizację” **PAI** poprzez wprowadzenie dodatkowych aksjomatów:

$$11) A \rightarrow A \rightarrow A \rightarrow A \text{ (lub, równoważne: } A \rightarrow A \rightarrow A)$$

dzięki czemu system staje się częściowo podobny do systemu **E** oraz **R**, a ponadto uzupełniając:

$$12) A \& B \rightarrow A \rightarrow B$$

$$\text{(lub równoważne } (A \rightarrow A \& B \rightarrow B) \rightarrow A \rightarrow B \rightarrow A \rightarrow B)$$

Na tej podstawie sformułował twierdzenie dedukcji analitycznej (analytic deduction theorem): ...Jeśli istnieje w **DAI** dowód dla B w sensie formalnym na podstawie hipotez C_1, \dots, C_n, A oraz jeżeli każda zmienna występująca w B występuje także w A , wówczas w **DAI** istnieje dowód dla $A \rightarrow B$ z C_1, \dots, C_n [Anderson A. R., Belnap N. D., 1975, 433].

Definicja ta zbliża się do ujęcia właściwego semantyce relacyjnej Meyera-Routleya, a w szczególności do określenia „współmożliwe powiązanie” (compossible relative). Anderson i Belnap wykazali, że wprowadzenie do systemu **R** spójnika „ \cdot ” — „współutrzymanie” (*co-tenability*) pozwala zdefiniować formułę $A \rightarrow B$ (implikacji koniecznej) jako $(A \rightarrow B) \cdot (B \rightarrow A)$, a więc relację *co-entailment* (wzajemnej dedukowalności). Spójnik „ \cdot ” (oraz „ $+$ ”) w systemie **R** nie zawierającym funktorów modalnych oraz w semantyce relacyjnej Meyera-Routleya, jest interpretowany jako intensjonalny odpowiednik koniunkcji i dysjunkcji, jakkolwiek jego użycie nie oznacza, że A wyklucza B . Innym pro-

¹⁵ „... p impliziert q analytisch, kann man vielleicht so interpretieren: „ q ist aus p und den logischen Axiomen ableitbar und enthält keine anderen Begriffe als p ” und es wäre, nachdem diese Definition genauer präzisiert hat, ein Vollständigkeitsbeweis für die Parryschen Axiome zu erstreben, in dem Sinn, das alle Satze, welche für die obige Interpretation von \rightarrow gelten, ableitbar sind”. [Parry, 1933, 5-6]

ponowanym określeniem dla tego spójnika jest „spójność relewancyjna” (relevant consistency). Ze względu na teoretyczne znaczenie tej interpretacji dla języków informacyjno-wyszukiwawczych należy dodać, że funktor ten rozważano także w kontekście modalnym systemów ścisłej implikacji.

W systemach implikacji ścisłej, wywodzących się z poglądów Diodora Krokosa, poszukiwano m.in. rozwiązania problemu spójników logicznych umożliwiających sformalizowany opis mocnego związku między członami zdania implikacyjnego. Chryzyp z Soloi pisał ... „A ci, którzy wprowadzają pojęcie powiązania twierdzą, że okres warunkowy jest mocny, gdy przeciwieństwo jego następnika jest niespójne z jego poprzednikiem”... [Sextus Empiricus, 1929]. Na tej podstawie Anderson i Belnap stwierdzają, że spójnik „•” użyty w formule $A \bullet B$ oznacza, że A jest kompatybilne z B i pozwala uzyskać definicję Chryzopa

- 1) $A \rightarrow B =df \downarrow A \bullet B$, jeśli A implikuje B , to A jest kompatybilne z B ;
- 2) $A \rightarrow B \rightarrow A \bullet B$, a następnie posługując się klasyczną negacją;
- 3) $A \rightarrow B \rightarrow A \rightarrow B$.

Uwagi te przytaczamy ze względu na możliwość wykorzystania interpretacji spójnika *co-tenability* w teorii języków informacyjno-wyszukiwawczych, na przykład do wyjaśnienia problemu kompatybilności języków. Równocześnie należy zauważyć, że terminologiczne uwagi poczynione w związku z tym spójnikiem dostarczają dodatkowych argumentów za koniecznością intensjonalnej interpretacji znaczenia przy budowie języków informacyjno-wyszukiwawczych.

Koncepcję pewnego modelu implikacji dla rachunku zdań przedstawił R. L. Epstein, zakładając, że przyjęta w tym modelu semantyka może okazać się przydatna do ustalania przedmiotów treści tekstów języka naturalnego [Epstein R. L., 1979]. Punkt wyjścia stanowi stwierdzenie nieadekwatności klasycznego rachunku zdań do wyjaśnienia związku wynikania według formuły *jeżeli p to q*, ponieważ rachunek ten nie dostarcza środków umożliwiających odwzorowanie relacji między dziedzinami znaczenia zmiennych indywidualnych. Autor zakłada, że semantyczny opis relacji implikacji powinien uwzględniać zależności między zdaniem, określone terminem *relatedness*. Termin ten traktuje się jako równoważny terminowi „relewantny”, który ma ustalony sens techniczny w filozofii logiki, określony między innymi przez Andersona i Belnapa.

Jako właściwy w opisie języków naturalnych oraz w teorii zdarzeń Epstein proponuje termin związany (*related*): związany treściowo (*content related*) oraz związany czasowo i przestrzennie (*space-time related*). Termin ten jest interpretowany jako elementarny semantycznie, podobnie jak prawda i fałsz. W opisie relewancji semantycznej, a więc związków treściowych między zdaniem, prawdziwość formuły implikacyjnej $p \rightarrow q$ interesuje nas jedynie wówczas, gdy p jest w jakiś sposób związane znaczeniowo z q . W przeciwnym wypadku relacja nie jest spełniona.

W podstawowej wersji modelu Epsteina chodzi zatem o opis pewnej relacji binarnej (predykatu) $r(p,q)$. W wersji rozwiniętej rozważane są także relacje trójczłonowe, a faktycznie n -członowe, definiowane przy użyciu następującej symboliki:

Język modelu	Znaczenie w języku angielskim	Inne symbole niesformalizowanego metajęzyka
\sim	not	\neg
\rightarrow	implies	\Rightarrow
$\langle \rangle$	if and only if	\Rightarrow iff
	and	\wedge
	or	\vee
$p_0, p_1, p_2 \dots$	propositions	$p, q, q_0, q_1 \dots$
		s, t
$(,)$	parantheses	$(,) \dots\dots$
	well formed formula	
	(wff) sentence	$..f..$
	set of wffs	

Za jedyną własność binarnej relacji r konieczną w budowie modeli tekstów języka naturalnego uznawana jest zwrotność (reflexivity). Każde zdanie p jest związane ze sobą samym $r(p,p)$. Przy modelowaniu działań (p,q) warunkiem wstępnym relacji p powoduje q (p causes q) jest pozostawianie p i q w bliskości czasowo-przestrzennej. Przy ustalaniu relacji powiązania poprawnie zbudowanych formuł (wffs) A, B konieczne jest spełnienie warunku, aby każde A było związane ze samym sobą (A, A). Oznacza to rozszerzenie relacji r na wszystkie wffs.

Regułą wnioskowania jest modus ponens, zgodnie z którą, jeśli prawdziwe jest A , prawdziwe jest $A \rightarrow B$, to prawdziwe jest także B wówczas, gdy A jest powiązane z B , co oznacza konieczność zachodzenia relacji w A dla pewnego p , a także w B dla pewnego q ; $r(p,q)$. Oznacza to więc konieczność spełnienia minimalnego warunku relewancji. Przy spełnieniu tego warunku można stwierdzić, że $(A, B \rightarrow C) \Leftrightarrow (A, B)$ lub $\mathcal{R}(A, C)$. Spójniki \wedge, \vee definiowane są za pomocą negacji i implikacji.

$$\mathcal{R}(A, B) \Leftrightarrow \mathcal{R}(A, \neg B)$$

$$\mathcal{R}(A, B \wedge C) \Leftrightarrow \mathcal{R}(A, B) \text{ lub } \mathcal{R}(A, C)$$

$$\mathcal{R}(A, B \vee C) \Leftrightarrow \mathcal{R}(A, B) \text{ lub } \mathcal{R}(A, C)$$

Do orzekania o prawdziwości zdań v (validity), gdy $v(p) \in T$ lub F , stosuje się matryce identyczne jak w rachunku zdań. Po rozszerzeniu v do $V, V(p) = v(p)$ dla każdego p otrzymuje się: $V(A) = T \Leftrightarrow V(B) = F$ oraz $V(A) = F \Leftrightarrow V(B) = T$. Stosowalność relacji powiązania zapewniana jest przez warunek, że $A \rightarrow B$ nie jest

prawdziwe, gdy $\sim \mathcal{R}(A, B)$. Udowodniono następnie, że zbiór poprawnie zbudowanych formuł $A(A_1, \dots, A_n)$ jest prawdziwy wówczas, gdy każde A_i należy do tego zbioru lub można je uzyskać na podstawie modus ponens. O zupełności systemu \mathcal{R} i możliwości przeprowadzenia wnioskowania $\vdash A \rightarrow C$ z $\vdash A \rightarrow B$ decydują reguły podstawiania. W tym celu wprowadzono konstrukty nazwane modelami i ramami M (models and frames). Model jest parą (r, v) . Poprawnie zbudowana formuła A jest prawdziwa w M ($M \models A$) wtt, gdy $V(A) = T$. Zdanie jest prawomocne (valid) wtt, gdy jest możliwe w każdym modelu. Jeśli każda poprawnie sformułowana formuła jest prawdziwa w M , to jest on modelem zbioru formuł S . M_r jest ramą dla określonej klasy r , a więc zbiorem wszystkich modeli z relacją r , natomiast M_v jest zbiorem modeli z oceną prawdziwości v . Poprawnie zbudowana formuła A obowiązuje w $M_r(M_v)$ oznaczane jako $M_r \models A(M_v \models A)$ wtt gdy $M_{r,v} \models A$ wszystkich v względem r .

Przedstawiono dowód, że wszelkie spójniki zdaniowe, których ocena prawdziwości zależy jedynie od prawdziwości zdań oraz od sposobu ich powiązania można konsekwentnie definiować w kategoriach negacji i implikacji:

$A \wedge B$ (niepowiązane *i – and*) jest prawdziwe w klasycznym rachunku zdań wtt, gdy prawdziwe są A i B , niezależnie od ich związku semantycznego:

$$\neg A \rightarrow \{B \otimes (A \rightarrow B)(A \rightarrow B)\}.$$

Przy $R(A, B)$ jako charakterystyce syntaktycznej $\mathcal{R}(A, B)$ umożliwiającej przekład semantycznych warunków powiązania na aksjomaty prezentowanego języka, $R(A, B)$ jest prawdziwe w modelu wtt, gdy A jest w tym modelu powiązane z B , na przykład $R(A, B) \Leftrightarrow R(B, A)$ jest ramą relacji symetrycznej;

— $A \vee B$ (klasyczna *lub – or*), jest prawdziwe wówczas, gdy prawdziwe jest A lub B definiowane jako $\neg(\neg A \wedge \neg B)$,

— $A \vee B$ (zależne relacyjnie *lub – relatedness dependent – or*) prawdziwe jest wówczas, gdy prawdziwe jest A lub B oraz gdy są one powiązane, definiowane jako $(A \vee B) \wedge (R(A, B) \vee R(B, A))$;

— $A \& B$ (zależne relacyjnie *i – and*) definiowane jako $A \wedge \{B \wedge (R(A, B) \vee R(B, A))\}$;

— $A \supset B$ (*implikacja materialna, implikacje niepowiązane – unrelated implies*), kiedy nie jest tak, że $A=T, B=F$ niezależnie od tego, czy A jest powiązane z B , definiowane jako:

$$\neg(A \wedge \neg B) \text{ lub } (A \leftrightarrow B) \text{ czyli } (B \rightarrow A) \wedge (A \rightarrow B).$$

Problem spójników zdaniowych rozwiązano, traktując je jako matryce logiczne dla n zdań, takie jak \neg, \rightarrow . Przyjmując Q_1, Q_2, \dots, Q_n za zmienne dla poprawnie zbudowanych formuł, stwierdzono, że spójnik zależy jedynie od tego, czy $V(Q_i)=T$ lub F oraz czy zachodzi relacja $r(Q_i, Q_j)$. Na tej podstawie definiowane są wszelkie spójniki oraz wyjaśniany jest problem przechodniości relacji r (*transitivity – r_T*):

$$r_T(p, q) \Leftrightarrow \text{dla pewnego } n, \text{ istnieją } q_1 \dots q_n \text{ takie, że } r(p, q_1) \text{ oraz } r(q_1, q_2) \text{ oraz } \dots \text{ oraz } r(q_n, q).$$

Nieprzechodnia relacja r (*non-transistive*) odwzorowuje ściśle powiązania pierwszego rzędu (first-order relatedness), natomiast przechodnia r_T oznacza powiązanie odległe (distantly related), przy czym r nieprzechodnie jest relacją rozstrzygalną, zaś r_T nie jest. Do ustalenia odległych powiązań p i q można dojść w określonej liczbie kroków (m), badając różne ramy M_{r_m} oraz porównując je z M_r i z M_T :

niech r_2 będzie (p, q) pewne $q_1, r(p, q_1)$

oraz r_m będzie (p, q) istnieją $q_1 \dots q_{m-1}, q$

takie, że $r(p, q_1)$ oraz (q_1, q_2) oraz ... oraz $r(q_{m-1}, q)$,

gdy relacja r (z założenia zwrotna) jest równocześnie symetryczna i przechodnia, wówczas stanowi regułę ekwiwalencji, otwierając w ten sposób miejsce dla reguł podstawiania. W modelu M_r możliwe jest wówczas podstawienie B zamiast C oraz D będącej częścią A , w dowolnej poprawnie zbudowanej formule $A(\text{wff})$, jeśli B i C mają tę samą wartość prawdziwościową:

$\mathcal{H}(B, D) \Leftrightarrow \mathcal{H}(D, C)$ oraz $\mathcal{H}(D, B) \Leftrightarrow \mathcal{H}(D, C)$.

Reguła ta staje się bardzo prosta, umożliwia bowiem podstawienie B za C wtt $M_r \models B \Leftrightarrow C$. Relacja ekwiwalencji dzieli zdania na klasy ekwiwalencji. Rama określonej klasy ekwiwalencji M_r składa się ze zdań uniwersalnie prawomocnych (universally valid sentences) oraz poszczególnych rzeczywistości (worlds) zawierających powiązane zdania. Każda z rzeczywistości denotowanych przez zdania jest modelem klasycznego rachunku zdań. Poprawnie zbudowane formuły zawierające zdania z klas ekwiwalencji są prawomocne w M_r wtt, gdy są prawomocne w rachunku zdań [Epstein R. L., 1979, 154].



Schemat 3. Zdania klasy ekwiwalencji - ujęcie 1

Scharakteryzowany model powiązania znaczeniowego (relatedness) został opracowany z myślą o jego zastosowaniu do opisu semantyki języka naturalnego. Założono, że zdania języka naturalnego są powiązane znaczeniowo wtt, gdy dotyczą tego samego przedmiotu. Zasady stosowania modelu określono następująco:

a) punktem wyjścia jest arbitralnie ustalony zbiór przedmiotów treści (topics), oznaczany T ,

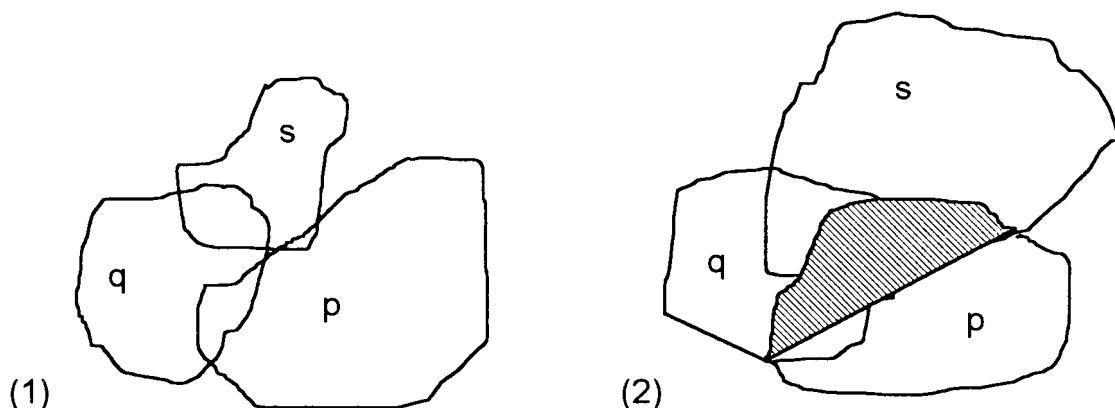
b) każde zdanie p wyznacza podzbiór T , oznaczany r . Podzbiór ten może dotyczyć jednego przedmiotu treści lub wszystkich elementów T ,

c) zdania p i q są powiązane wówczas, gdy przedmioty ich treści krzyżują się, a więc $r(p, q)$ wtt, gdy $r \cap q \neq \emptyset$. Oznacza to relację zwrotną, która powinna być symetryczna: $r \cap q = q \cap r$.

Cecha przechodniości stanowiłaby tu przeszkodę: $r(p, q)$ wtt $p \subseteq q$;

d) relację $r(p, q)$ rozszerza się na $\mathcal{R}(A, B)$ na podstawie zdań q_1, \dots, q_n występujących w poprawnie zbudowanych formułach A i B , a zatem $\mathcal{A}(A, B)$ $r(p, q)$ wtt $r \cap q \neq \emptyset$. Przedmiotem treści wyrażen złożonych są więc kombinacje przedmiotów treści wyrażen składowych;

e) ustalenie przedmiotu treści zdań języka naturalnego sprowadza się *de facto* do ustalania relacji powiązania (*relatedness*). Wiąże się to z ustalaniem nasilenia związku między wieloma przedmiotami treści. Miarę tego związku przedmiotowego między zdaniami p, q, s można przedstawić dwoma sposobami:



Schemat 4. Zdania klasy ekwiwalencji - ujęcie 2

Stanowi to o ograniczeniu modelu wynikania logicznego Epsteina, jakkolwiek zaletą tego modelu jest prostota. Do opisu rodziny relacji złożonych niezbędny jest — jak stwierdza w konkluzji autor — język bogatszy lub co najmniej dorównujący siłą teorii mnogości.

Potrzebę skonstruowania logik relewancji wyższego rzędu stwierdza także Belnap, podkreślając, że dotychczasowe systemy pierwszego rzędu są nieadekwatne wobec złożoności relacji w języku naturalnym. Próby skonstruowania skwantyfikowanej wersji logik relewancji przez dołączenie do rachunku zdań dodatkowych aksjomatów oraz reguł wnioskowania związanych z kwantyfikatorami (*każde, niektóre*) nie doprowadziły do pozytywnych rezultatów. W logikach relewancji opartych na rachunku zdań nie rozstrzygnięty pozostał problem podejmowania decyzji o wyborze najbardziej charakterystycznego przedmiotu treści (według Epsteina). Podobnie przy próbie kwantyfikacji pow-

staje problem kompletności i konieczności odwoływania się do pozalogicznej wiedzy o zakresie znaczenia zmiennych indywidualnych.

Jako kierunek rokujący nadzieje na skonstruowanie logik relewancji wyższego rzędu wskazuje się analizę funkcji relewancji (nazywanej także strict function). Funkcja ta powinna umożliwić reprezentację rodziny funkcji rzeczywiście zależnych od ich argumentów, umożliwiającą konsekwentne stosowanie dowodu poprzednika w przeprowadzaniu dowodu dla następnika [Belnap N. D., 1981].

Problem ten podejmowali uczniowie Belnapa i Andersona, kontynuując kierunek entailment. Przedstawiono między innymi kryteria podziału funkcji na relewantne i irrelewantne, stwierdzając, że dana formuła jest dowodliwa w systemie \mathbf{R} w przypadku, gdy typ reprezentowany przez tę formułę (na przykład: $A \rightarrow B$ jako reprezentujący typ funkcji relewantnej (z A na B) stanowi typ terminu syntaktycznie relewantnego) [Belnap N. D., 1981]. Na tej podstawie wykazano spójność logicznej i matematycznej koncepcji relewancji, co powinno doprowadzić do sformułowania teorii relewantnej rekurencji (relevant recursion theory).

W uzupełnieniu prezentacji systemów logicznych stanowiących próbę rozwiązania problemu relewantnej implikacji zasygnalizujemy koncepcję A. Zinowjewa [Zinowjew A., 1976]. Jej założenia różnią się od wyżej scharakteryzowanych tym, iż autor uważa za błędne traktowanie zdań implikacyjnych jako wypowiedzi złożonych z X wynika Y , zapisywane jako $X|-Y$. Zinowjew twierdzi, że formułę tę należy traktować jako wypowiedź elementarną, której częściami nie są wypowiedzi X i Y lecz terminy $[X]$, $[Y]$ reprezentujące te wypowiedzi. Symbol $|-$ proponuje interpretować jako dwuargumentowy predykat z wypowiedzi pierwszej wynika wypowiedź druga, zapisywana jako:

$$([X], [Y]) \leftarrow (|-).$$

Symbol $|-$ oznacza bowiem, nie same przedmioty X i Y lecz zachodzące między nimi związki. Podkreśla także konieczność rozróżniania formuł odpowiadających okresowi warunkowemu (\rightarrow) od metawypowiedzi o wynikaniu, jakkolwiek pewne ich własności mogą być zbieżne. Podstawową zasadę dedukcji traktuje jedynie jako wstępny warunek ustalenia reguł wynikania logicznego, na mocy których można stwierdzić, że z X wynika Y , co oznacza uznanie Y przy założeniu, że X jest czymś, co spełniło już swoje zadanie. Definiowanie wynikania logicznego ma polegać na wyliczaniu reguł wynikania przy uwzględnieniu związków między wartościami logicznymi wypowiedzi. Związki te określa podstawowa zasada dedukcji. Jej przestrzeganie jest warunkiem niezbędnym lecz nie jest wystarczające. Nie powinna przy tym stanowić problemu, ponieważ w każdym dziale logiki reguły wynikania są ustalane dla określonych struktur o zdefiniowanych wartościach logicznych. Za nieistotną uważa liczbę tych wartości, ważne jest bowiem jedynie uzyskanie prawdziwych następstw z prawdziwych przesłanek.

Zinowjew zakłada, że nie istnieje „naturalne”, „intuicyjne”, „prawdziwe” wynikanie logiczne. Możliwa jest jedynie eksplikacja rozumienia niektórych

form wynikania intuicyjnego wraz z ustaleniem różnic między tymi formami i ich powiązań. Przedstawiona teoria mocnego (wąskiego) wynikania S^M charakteryzuje się tym, że do wniosku Y nie wchodzi wyrażenia elementarne nie występujące w przesłance X , a ponadto we wszystkich twierdzeniach systemu występuje co najmniej jedno wyrażenie złożone z co najmniej dwu terminów (znaków o ustalonym znaczeniu) oraz z co najmniej jednego operatora logicznego definiującego własności wypowiedzi. Jest to więc koncepcja zbliżona do interpretacji W. Parry'ego nawiązującej do ujęcia analityczności według Kanta.

Przytaczamy podstawową charakterystykę systemu S [Zinowjew A., 1976, 166]:

Alfabet :

- 1) litery p, q, r, z z indeksami lub bez indeksów — wypowiedzi elementarne,
- 2) \cdot, \vee, \sim — operatory logiczne (*i, lub*) nierozłączne, *nie, przy czym* „ \cdot ” może być pomijany),
- 3) \vdash - predykat wynikania logicznego,
- 4) nawiasy i przecinki.

Wypowiedzi :

- 1) wypowiedzi elementarne są wypowiedziami,
- 2) jeśli X jest wypowiedzią, to $\sim X$ jest wypowiedzią,
- 3) jeśli $X^1 \dots X^n$ ($n \geq 2$) są wypowiedziami, to (X^1, X^2, \dots, X^n) oraz $(X^1 \vee X^2 \vee \dots \vee X^n)$ są wypowiedziami,
- 4) wypowiedzią jest tylko to, co definiują punkty 1 - 3,
- 5) $([X], [Y]) \leftarrow (\vdash)$ jest wypowiedzią o wynikaniu logicznym wtt, gdy X i Y są wypowiedziami; zapis ten może być uproszczony do postaci $X \vdash Y$,
- 6) znak \vdash może występować tylko raz, zgodnie z założeniem, że w ogólnej teorii wynikania logicznego nie rozpatruje się budowy wypowiedzi.

Aksjomaty S^M :

- 1) $\sim \sim X \vdash X$
- 2) $XY \vdash X$
- 3) $XY \vdash YX$
- 4) $X^1, X^2, \dots, X^n \vdash Y$, gdzie Y różni się od X^1, X^2, \dots, X^n tylko rozkładem nawiasów,
- 5) $\sim (XY \vdash \sim X \vee \sim Y)$
 $(X^1, X^2, \dots, X^n) \vdash \sim X^1 \vee \sim X^2, \dots \vee \sim X^n$
- 6) $(X \vee Y)Z \vdash XZ \vee YZ$
- 7) $(XZ \vee YZ) \vdash (X \vee Y)Z$
- 8) $(XY \vee Z) \vdash (XY \vee Z) (Y \vee Y)$.

Reguły wyprowadzania teorematów z aksjomatów S^M :

- 1) jeśli $X \vdash Y$ i $Y \vdash Z$, to $X \vdash Z$
- 2) jeśli $X \vdash X$ i $Y \vdash Z$, to $X \vdash YZ$
- 3) jeśli $X^1 \vdash X^2$, to $Y^1 \vdash Y^2$, gdzie Y^2 jest utworzone z Y^1 przez zamianę którejś z wypowiedzi X^1 na wypowiedź X^2 .

Omówione przykładowo systemy opisujące relację implikacji, wywodzące się z tradycji badań nad modalnością oraz nad analitycznością, traktowano

jako próby sformułowania uniwersalnych teorii wynikania logicznego. Niezależnie od różnic w założeniach wyjściowych i stosowanej notacji wykazują one jedno zasadnicze ograniczenie, wyrażające się w konieczności odwoływania się do semantycznych stosunków zachodzących między wyrażeniami zdania implikacyjnego, są więc zrelatywizowane do wiedzy odwzorowanej w języku opisującym rzeczywistość. Wiedza ta wyznacza bowiem znaczenie wyrażań. Znaczenie musi być interpretowane intensjonalnie, ponieważ jest to podstawowy sposób na wykazanie zachodzenia relacji entailment i spełnienia minimalnego warunku relewancji logicznej. Spełnienie tego warunku stanowi z kolei podstawę procesu wnioskowania.

2.4. TEORIA RELEWANCJI SPERBERA I WILSON

Punktem wyjścia w teorii relewancji Dana Sperbera i Deirde Wilson, reprezentującej kognitywny nurt badań w językoznawstwie, jest rozróżnienie dwu sposobów opisu komunikacji, adekwatnych dla różnych form przekazu informacji (np. werbalnej, ostensywnej, percepcyjnej itp.) [Sperber D., Wilson D. 1986]:

- *ujęcia semiotycznego*, reprezentowanego przez wszystkie teorie od Arystotelesa po F. de Saussure'a i C. Shannona, zgodnie z którym proces rozumienia i komunikowania wypowiedzi polega na *kodowaniu* i *dekodowaniu* (znaczenia) sygnału językowego,
- *ujęcia inferencyjnego*, zgodnie z którym rozumienie i komunikowanie odbywa się poprzez *interpretację* wypowiedzi w kontekście.

Jednym z głównych celów teorii jest podważenie poglądu, dość powszechnego w tradycji zachodnioeuropejskiej humanistyki, o jednorodnym charakterze procesów komunikacji, sprowadzanych do procesów semantycznego kodowania i dekodowania. Autorzy ci twierdzą, że interpretacja wypowiedzi stanowi interakcję między strukturą językową, informacją pozajęzykową oraz strukturami informacji mentalnej, tworzącymi zasadniczy kontekst wypowiedzenia (utterance). Interakcja ta możliwa jest dzięki procesom *wnioskowania*, wypełniającym lukę między (zakodowaną) reprezentacją semantyczną i mentalną wypowiedzenia. Różnica między dekodowaniem i wnioskowaniem wyjaśniana jest następująco:

— dekodowanie zaczyna się od sygnału i kończy na odkryciu informacji (message) związanej z sygnałem przez podstawowy kod (język);

— wnioskowanie zaczyna się od zbioru przesłanek i reguł, a kończy zbiorem wniosków, które nie są bezpośrednio związane z przesłankami kodem; proces ten może stanowić część procesu dekodowania.

Jeśli nadawca i odbiorca wypowiedzi stosują identyczne reguły wnioskowania odnoszące się do tego samego zbioru przesłanek, dochodzą do tego samego wniosku — identycznie interpretują tę wypowiedź. Te same przesłanki mogą dopuszczać wiele lub nieskończenie wiele wniosków (interpretacji), lecz tylko jedna lub niektóre okazują się najczęstsze, np. najczęstszą interpretacją wieloznacznego zdania:

(1) Kowalski kupił „Przegląd Sportowy”

(a) Kowalski kupił egzemplarz „Przeglądu Sportowego”

(b) Kowalski kupił przedsiębiorstwo, które wydaje „Przegląd Sportowy”

jest interpretacja (a), podczas gdy zdanie:

(1) Fibak kupił „Przegląd Sportowy”

(a) Fibak kupił egzemplarz „Przeglądu Sportowego”

(b) Fibak kupił przedsiębiorstwo, które wydaje „Przegląd Sportowy”

może mieć częstszą interpretację (b). Pytanie stawiane w związku z tym w pragmatyce językoznawczej i teorii komunikacji brzmi: *jak to się dzieje, że interpretując wypowiedź odbiorca wyciąga z danej przesłanki wniosek zakładany przez nadawcę?*

Jedna z prób wyjaśnienia tego problemu odwołuje się do tzw. *hipotezy wspólnej wiedzy uczestników komunikacji (mutual knowledge hypothesis, shared knowledge hypothesis)*, stanowiącej kontekst, w którym przetwarzany jest zbiór przesłanek. Na kontekst ten, traktowany jako konstrukt psychologiczny, składa się pewien zbiór założeń o rzeczywistości (*assumptions*), automatycznie kształtowany przez uczestników procesu komunikacji. O podobieństwie zbioru założeń członków jednej społeczności komunikacyjnej ma stanowić wspólny kod (język) i zdolności wnioskowania, a o zróżnicowaniu — szeroko rozumiane odmienne doświadczenia, pochodzące z percepcji; np. relacje świadków wypadku ulicznego zwykle różnią się. Zauważa się, że odmienne doświadczenia uczestników komunikacji są neutralizowane przez struktury gramatyczne wypowiedzi, np. odpowiedź na pytanie: *Napijesz się kawy?* — *Kawa działa pobudzająco* — można interpretować zarówno jako akceptację (chcę być bardziej rześki), jak i odmowę (chcę spać). Udowodnienie prawdziwości hipotezy o wspólnej wiedzy jako gwarancji skutecznej komunikacji jest bardzo trudne, jeśli w ogóle możliwe.

Na odwzorowanie rzeczywistości w strukturach mentalnych, nazywane środowiskiem poznawczym człowieka (*cognitive environment, shared cognitive environment*), składają się informacje pochodzące z różnych źródeł: z procesów komunikacji werbalnej, ostensywnej oraz z percepcji, zindywidualizowane w stopniu znacznie wyższym, niż się przypuszcza. Złożoność i zindywidualizowanie struktur informacji mentalnej spowodowane są: złożonością środowiska poznawczego, zróżnicowaniem możliwości poznawczych człowieka (percepcyjnych, inferencyjnych, językowych, pamięciowych) oraz posiadaną wiedzą teoretyczną. Proces komunikacji między nadawcą i odbiorcą informacji, a zwłaszcza interpretacja wypowiedzi, możliwy jest w zakresie wyznaczonym przez skrzyżowanie odwzorowań mentalnych nadawcy i odbiorcy informacji.

Inspirację i punkt wyjścia opisu procesu interpretacji wypowiedzenia i jego relewancji w teorii Sperbera i Wilson stanowił pragmatyczny model komunikacji Paula Grice'a, w szczególności inferencyjna interpretacja znaczenia

oraz zasady (postulaty) kooperacji konwersacyjnej¹⁶ [Grice H. P., 1957, 377-388] [Grice H. P., 1969, 147-177].

Według Grice'a stwierdzenie: *[S] rozumie coś przez X*, oznacza *[S] chce, aby wypowiedzenie X wywarło określony efekt na odbiorcy przez uznanie intencji [S]*. Ujęcie to zmodyfikował P. Strawson, 'charakteryzując komunikację inferencyjną następująco:

- wypowiadając *X [S]* chce spowodować reakcję *r* u odbiorcy *[A]*,
- *[A]* uznaje intencję (*a*) nadawcy *[S]*,
- uznanie przez *[A]* intencji (*a*) jest przynajmniej częścią jego reakcji, dowodem uznania intencji nadawcy *[S]*.

Wprowadzenie kategorii *intencji informacyjnej* oznacza uznanie komunikacji za proces niejednorodny, w którym kod (język), rozumiany jako znany uczestnikom zbiór określonych konwencji, jest jednym z elementów obok kontekstu i reguł wnioskowania, służebnych wobec procesu inferencji. Zauważa się przy tym, że posługiwanie się kodem nie jest obligatoryjne w komunikacji, ponieważ przekazanie (umożliwienie wywnioskowania) informacji jest możliwe również przez ostensywne zachowanie nie kodowane, np. w odpowiedzi na pytanie *Co słyhać?* — pokazanie opakowania aspiryny (jestem przeziębiony).

Identyfikację intencji informacyjnej nadawcy wypowiedzi poprzedza akceptacja ze strony odbiorcy tzw. *intencji komunikacyjnej nadawcy* [Strawson P., 1964, 439-460]. Brak tej akceptacji oznacza niepowodzenie komunikacji, np. odpowiedź kancelarii prezydenta na propozycję osobistego spotkania z ministrem X w pilnej sprawie — „prezydent przyjmuje od X tylko informacje przesłane pocztą” można uznać za demonstrację braku intencji komunikacyjnej. Demonstracją intencji komunikacyjnej w dialogowych systemach informacji są przewidziane w protokołach dialogu formuły witania użytkownika. (Por. rozdz. 4.2.2.) W komunikacji między ludźmi rozpoznawanie intencji komunikacyjnej odbywa się na podstawie obserwacji zachowania, wiedzy o ludziach oraz wnioskowania, które z możliwych zachowań jest przewidziane i pożądane. W komunikacji bezpośredniej (*face-to-face*) intencja komunikacyjna jest realizowana konwencjonalnie, np. za pomocą formuł pozdrowień.

¹⁶ Zajmując się ogólnymi zasadami i prawidłowościami leżącymi u podstaw komunikacji międzyludzkiej, Grice sformułował tzw. zasadę kooperacji (współdziałania) — cooperative principle, rozwiniętą w dziewięciu maksymach, ujętych w cztery kategorie. **(I) Maksyma ilości:** (1) *Uczyn swoją wypowiedź na tyle informatywną, ile to niezbędne*, (2) *Nie czyn swojej wypowiedzi bardziej informatywnej niż to niezbędne*; **(II) Maksyma jakości:** (1) *Nie mów tego, co uważasz za nieprawdę*; (2) *Nie mów tego, na co nie masz wystarczających dowodów*; **(III) Maksyma związku:** (1) *Bądź relewantny*; **(IV) Maksymy sposobu:** (1) *Unikaj mętności wypowiedzi*; (2) *Unikaj wieloznaczności*; (3) *Bądź zwięzły*; (4) *Bądź uporządkowany*. [Sperber D., Wilson D., 1986, 33] Wprowadzając kategorię relewancji Grice nie sformułował jej definicji. W późniejszych pracach z zakresu pragmatyki lingwistycznej zasady kooperacji Grice'a uzupełnił Leech, formułując tzw. *zasady grzeczności*, złożone z sześciu maksym: *taktu, wielkoduszności, aprobaty (uznania), skromności, zgodności oraz sympatii i zrozumienia* [Leech G.N., Principles of Pragmatics, London].

W komunikacji niebezpośredniej (nie odwzajemnianej), np. między autorem książki i czytelnikiem — za pomocą środków językowych mających nadać ujawnionym intencjom informacyjnym cechy wiarygodności, w formie uzasadnienia tez czy wywodów.

W pragmatycznych ujęciach procesów komunikacji podkreśla się znaczenie *ostensji* (postrzegalnego zachowania uczestników) w dostarczaniu dowodów intencji, poprzedzających ujawnienie przez nadawcę podstawowej informacji (znaczenia), np. demonstracyjne pociągnięcie nosem po wejściu do mieszkania.

W odróżnieniu od klasycznego ujęcia semiotycznego, zakładającego, że komunikacja polega na (de)kodowaniu pewnego zestawu znaczeń (konotacji), pragmatycy są zdania, że komunikowanie znaczenia nie należy do wyraźnych klas zjawisk, np. — *Co będziesz dziś robić? — Okropnie boli mnie głowa.* — (a) Nic nie będę robić; (b) Zrobię, co będzie możliwe; (c) Nie wiem, co będę robić.

Poprzez intencję informacyjną nadawca ujawnia pewien zbiór założeń (*assumptions*), których treść jest reprezentowana mentalnie. Ujawnione intencje komunikacyjna i informacyjna wraz z założeniami wykorzystywane są przez odbiorcę w inferencyjnym procesie interpretacji wypowiedzenia, przy czym celem procesu komunikacji jest bezpośrednia modyfikacja środowiska poznawczego odbiorcy informacji.

Założenia — prawdziwe lub fałszywe — uznaje się za podstawową (choć nie elementarną) jednostkę języka informacji mentalnej (konceptualnej). Przyjmują postać formuł logicznych, które muszą spełniać jedynie warunek bycia poprawnie sformułowanych (*well-formed*). Założenia te mogą być kompletne semantycznie (lub nie), stanowiąc wówczas tzw. założenia faktualne (*factual assumptions*). Zbiór założeń faktualnych tworzy informację encyklopedyczną. Ponadto jest podłożem dla spontanicznego wnioskowania (*non-demonstrative inference*) pod wpływem informacji pochodzącej ze środowiska poznawczego. Formuły niekompletne semantycznie, a więc schematy założeń pełnią ważną funkcję w procesach poznawczych, ponieważ są uzupełniane na podstawie wiedzy pozajęzykowej. Stanowią podstawową jednostkę formatu pamięci, pełniąc równocześnie funkcje metajęzykowe (założenia o założeniach), czego dowodzi zdolność niespontanicznego myślenia o tym, co wiemy i zajmowania pewnych postaw. Przypuszcza się, że podstawowe formuły założeń mają postać typu:

(*Temperatura wynosi ... °C*)

(„... jest (*kawalerem/ żonaty/ rozwiedziony/ wdowcem*)”)

Zródłem założeń faktualnych jest:

— percepcja: *To jest jabłko* (wzrok); *Dzwoni telefon* (słuch); *To jest gorzkie* (smak);

— dekodowanie językowe: *Piotr mówi, że boli go głowa*;

— procesy wnioskowania (dedukcja).

Założenia faktualne składające się na konceptualne odwzorowanie rzeczywistości nie są równoważnościowe. Różnią się mocą (strength of assumptions), odpowiadającą stopniu przekonania o ich prawdziwości. Moc założeń faktualnych wyznaczana jest przez funkcję confirmacji i zależy od¹⁷:

a) historii przetwarzania, wpływającej ponadto na ich dostępność; założenia częściej przetwarzane są łatwiej dostępne, np.: *Kair jest obecnie stolicą Egiptu* jest łatwiej dostępne niż *Teby były stolicą Egiptu za panowania dwudziestej dynastii*;

b) sposobu zdobywania — mocniejsze są założenia oparte na doświadczeniu percepcyjnym, np. *Zaopatrzenie sklepów w żywność w Polsce na początku lat osiemdziesiątych było fatalne* — ma moc większą dla osoby wówczas dorosłej, która robiła zakupy niż dla osoby młodej, która była wówczas dzieckiem, i nie doświadczała w takim stopniu uciążliwości zaopatrzenia.

c) stopnia zaufania odbiorcy do nadawcy informacji językowej.

Wyjściowa moc założeń zależy od sposobu ich zdobywania: mocne są założenia pochodzące z percepcji, słabsze — tworzone na podstawie akceptacji czyjejś wypowiedzi, a więc uwarunkowane stopniem zaufania do nadawcy informacji. Moc ta zmienia się po każdorazowym wykorzystaniu założenia do przetwarzania nowej informacji. Wydaje się mało prawdopodobne, aby człowiek posiadał odrębny system przetwarzania i odwzorowania mocy założeń, która reprezentuje raczej pewną wartość porównawczą w skali heurystycznej niż zaś numerycznej (por. rozdz. 2.3.1. — teoria indukcji i prawdopodobieństwa R. Carnapa). Zmiana ta przypuszczalnie nie jest rezultatem odrębnego wyspecjalizowanego przetwarzania logicznego, lecz „produktem ubocznym” różnych procesów inferencyjnych (dedukcyjnych i niededukcyjnych).

Transformacja zbioru założeń (ich zasobu i mocy) pod wpływem nowej informacji pochodzącej ze środowiska poznawczego dokonywana jest na podstawie reguł dedukcji (implikacji) interpretowanych semantycznie i spontanicznie wykorzystywanych przez wyspecjalizowany ośrodek inferencyjny (deductive device):

If P then Q; If (not P) then (not Q); If Q then P (because) P;

Spontaniczne wnioskowanie jest możliwe dzięki ustrukturalizowaniu założeń w formuły zdaniowe, złożone z etykiety (adresu) pełniącego dwie różne funkcje — „nagłówek” (heading) i składnika formuły logicznej. Podstawowe znaczenie w procesie inferencji przypisuje się relacji entailment (zob. rozdz. 2.3.2.).

Konieczność przeprowadzania równoczesnej semantycznej i syntaktycznej interpretacji, dedukcji i implikacji wyjaśniana jest następującymi uwarunkowaniami:

— punktem wyjścia do przetworzenia zbioru założeń jest ich formuła logiczna,

¹⁷ Por. rozdz. 2.3.1. — Teoria indukcji i prawdopodobieństwa R. Carnapa.

— relacja implikacji semantycznej zachodzi między założeniami (P) i (Q) wtt, gdy odwołujemy się do stanu rzeczy, które te założenia reprezentują.

Według autorów omawianej teorii, zdolność człowieka do spontanicznej dedukcji decyduje o ekonomii przetwarzania informacji, dokładności wniosków wyciąganych z przesłanek oraz o precyzji odwzorowania rzeczywistości na poziomie mentalnym.

Spontaniczne wnioskowanie dedukcyjne, uwzględniające semantyczne własności założeń, możliwe jest dzięki ich strukturze, dotychczas charakteryzowanej ogólnie ze względu na funkcje w odwzorowaniu rzeczywistości. Przyjmuje się, że elementarną jednostką założenia jako formuły zdaniowej jest pojęcie (*concept*) traktowane jako pewien konstrukt psychologiczny. Wyróżniane jest za pomocą etykiety (*label*), pełniącej równocześnie dwie różne, choć komplementarne funkcje: adresu w pamięci oraz hasła (*heading*), pod którym przechowywane są założenia zawierające dane pojęcie.

Wyróżniono trzy rodzaje informacji konceptualnej organizowanej za pomocą etykiet pojęć:

a) Informacja logiczna określa reguły dedukcji związane z danym pojęciem. Są one najprawdopodobniej ograniczone do reguły eliminacji (*and-elimination*) (*Input P&Q; Output P; Input P&Q; Output Q*), *modus ponendo ponens* i *modus tollendo ponens*. Reguła eliminacji ogranicza wnioskowanie tylko do tych zbiorów założeń (przesłanek), w których występuje dane pojęcie i które dostarczają wniosków nie zawierających danego pojęcia, co pozwala uniknąć jej nieskończonej replikacji.

b) Informacja encyklopedyczna wskazuje ekstensję (denotację) danego pojęcia (obiekty, zdarzenia, osoby).

c) *Informacja leksykalna*, wskazuje odpowiedniki danego pojęcia w języku naturalnym na poziomie fonologicznym i syntaktycznym (frazy lub zdania), np. *pet (ulubione zwierzątko) — kotek, piesek*.

Rozróżnienie wejść encyklopedycznych i logicznych dla elementarnych jednostek informacji konceptualnej odpowiada tradycyjnemu rozróżnieniu prawd (zdań) *syntetycznych i analitycznych*. Jest to podział formalny służący podkreśleniu, że ta sama informacja może być przechowywana w dwóch formach i funkcjonować dwoma sposobami. Podział ten odpowiada rozróżnieniu reprezentacji (*representation*) i przetwarzania (*computation*) w psychologii kognitywnej, uznawanych za komplementarne i niezbędne w procesie rozumienia wypowiedzi.

Problemem nie wyjaśnionym i trudno weryfikowalnym jest sposób reprezentacji znaczenia słów (wyrazów). Przypuszcza się, że jest odwzorowywane przez układy powiązanych (kojarzonych) pojęć, dostarczanych w różnych formatach wraz z regułami eliminacji wskazującymi warunki niezbędne i wystarczające do użycia danego słowa, np. *matka — rodzic żeński; Input X matka Y Output X rodzic żeński Y*.

Przywołuje się tu kognitywną teorię prototypów, według której znaczenie słowa nie jest zdeterminowane przez własności logiczne słowa, lecz przez reprezentację mentalną rzeczy, którą denotuje dane słowo — np. znaczeniem słowa *żyrafa* jest prototypowy model żyrafy. Zdaniem J. A. Fodora znaczenia większości słów nie da się zdekomponować na elementarne cechy semantyczne (*semantic primitives*), np. *kolor*, *sól*, *zwierzę*, są one bowiem nieredukowalne, choć mają własności logiczne [Fodor J. A., Garret M., Walker E., Parkes C., 1980, 263-367], [Fodor J.A., 1983]. Autor ten uznaje za prawdopodobne reprezentowanie znaczenia przez tzw. postulaty znaczeniowe związane z regułami eliminacji, np.

SÓL — reguła eliminacji

Input (*X-sól-Y*)

Output (*X-substancja pewnego rodzaju Y*)

ŻYRAFA — reguła eliminacji

Input (*X-żyrafa-Y*)

Output (*X-zwierzę pewnego gatunku Y*)

ŻÓŁTY reguła eliminacji

Input (*X-żółty-Y*)

Output (*X-kolor pewnego rodzaju Y*)

Mechanizm transformacji informacji (założeń/przesłanek) polega na analizowaniu struktury konceptualnej założeń oraz na manipulowaniu nimi na podstawie wskazanych reguł wprowadzania (*AND Introduction*, *OR Introduction*, *DOUBLE-NEGATION*) i reguł eliminacji przypisanych wejściom logicznym. Dedukcyjna transformacja założeń może obejmować tylko implikacje (*entailments*) nietrywialne umieszczone w kontekście innych implikacji, również nietrywialnych. Implikacjami trywialnymi są takie, które na wyjściu zawierają pojęcie (założenie) wejściowe i co najmniej jedno nowe. Na przykład, w wyniku zastosowania reguł wprowadzania i eliminacji:

(1) *Input* : (i) *P*

(ii) *Q*

Output (*P and Q*) — koniunkcja jako wniosek

(2) *Input* : (i) *P*

(ii) *Q*

Output (*P or Q*) — dysjunkcja jako wniosek

(3) *Input* : (i) *P*

(ii) *Q*

Output (*not(not P)*) — podwójna negacja jako wniosek

możliwe są następujące implikacje nietrywialne (*entailments*):

(a) *P and P*; (b) *P or Q*; (c) (*not (not P)*)

(d) (*If (not P) then Q*); (e) (*If Q then P*).

Oznacza to, że żaden nadawca ani odbiorca nie wyciągnie następujących wniosków na podstawie trywialnych wypowiedzi typu:

- (1) *Premier podał się do dymisji.*
- (1a) *Premier podał się do dymisji i premier podał się do dymisji.*
- (1b) *Albo premier podał się do dymisji albo dziś jest trochę ciepiej.*
- (1c) *Nieprawda, że premier nie podał się do dymisji.*
- (1d) *Jeśli premier nie podał się do dymisji, tygrysy wyginą.*
- (1e) *Jeśli są urodziny królowej, premier podał się do dymisji.*

W spontanicznym wnioskowaniu rolę odgrywają jedynie implikacje nietrywialne, a nietrywialność gwarantują właśnie reguły eliminacji działające przy analizie (wyjaśnianiu) treści założeń dostarczonych do ośrodka inferencyjnego w mózgu. Powoduje to docelową modyfikację (doskonalenie) odwzorowania rzeczywistości. O zmianie tego odwzorowania decydują dwie odmiany implikacji — analityczne i syntetyczne.

• Implikacja analityczna — zbiór założeń $\{P\}$ implikuje analitycznie założenie Q *wtt*, gdy Q jest jedną z tez końcowych dedukcji, w której tezami wyjściowymi są $\{P\}$, i w których stosuje się jedynie reguły analityczne (określające znaczenie wyrażeń języka), wystarczające do rozumienia zbioru założeń, np.

(1) *Jeśli kolej strajkuje i samochód zepsuł się, nie można dostać się do pracy.*

- (1a) *Kolej strajkuje.*
- (1b) *Samochód zepsuł się.*
- (1c) *Nie można dostać się do pracy.*

Uzyskanie wniosku (1c) umożliwiają zastosowane do wejściowych założeń reguły wprowadzania *AND* oraz *modus ponens* (koniunkcja lub dysjunkcja).

• Implikacje syntetyczne — zbiór założeń $\{P\}$ implikuje syntetycznie założenie Q *wtt*, gdy Q jest jedną z tez końcowych w dedukcji, w której tezami wyjściowymi są $\{P\}$, a Q nie jest implikacją analityczną $\{P\}$. W praktyce implikacja syntetyczna jest rezultatem derywacji, w której stosuje się przynajmniej jedną regułę syntetyczną, a więc wymagającą nie tylko językowego zrozumienia wypowiedzi, lecz także odwołania się do wiedzy pozajęzykowej (encyklopedycznej). Na przykład, do zrozumienia następujących zdań (1a-c) i wywnioskowania z nich zdań (2a-c):

- (1a) *Bilet jest w portfelu.*
- (1b) *Portfel jest w walizce*
- (1c) *Walizka jest w samochodzie.*
- (2a) *Bilet jest w walizce*
- (2b) *Bilet jest w samochodzie.*
- (2c) *Portfel jest w samochodzie.*

oprócz wychwycenia implikacji analitycznych — (*Bilet jest gdzieś; Coś jest w portfelu; Portfel jest gdzieś; Coś jest w walizce; Walizka jest gdzieś; Coś jest w samochodzie*) — konieczna jest znajomość syntetycznej reguły zawierania (containment rule), czyli wiedzy, o tym że: *portfel, walizka, samochód* są rodzajem pojemnika.

Ilustracja różnicy w funkcjonowaniu założeń (i implikacji) analitycznych oraz syntetycznych jest istotna z punktu widzenia sposobu reprezentacji i przetwarzania wiedzy w inteligentnych systemach eksperckich, omawianych w rozdziale 3.

Doskonalenie mentalnego odwzorowania świata wyraża się w zainteresowaniu ludzkiego organizmu wykryciem wśród założeń aktualnie przetwarzanych możliwie dużej liczby implikacji syntetycznych, przed ich rozbiciem na mniejsze jednostki (pojęcia). Wymaga tego zasada efektywności przetwarzania informacji przez człowieka. *Efektywność* można definiować w kategoriach celów absolutnych (np. celny rzut piłki do kosza) lub względnych (długofalowych lub krótkofalowych) — np. mnożenie potomstwa, doskonalenie uderzenia w grze w tenisa, rozwijanie wiedzy. O efektywności przetwarzania założeń faktualnych (encyklopedycznych) decyduje ich oddziaływanie na zbiór założeń już posiadanych. Warunkiem efektywnego oddziaływania nowych założeń (nowej informacji) na założenia posiadane (informację starą), tworzącą kontekst przetwarzania jest ich treściowe powiązanie, a więc łączne używanie jako przesłanek w procesach inferencyjnych, powodujących efekt zwielokrotnienia wiedzy. Nazywa się to **zasadą maksymalizacji przetwarzania informacji**, stanowiącą decydujący czynnik w interakcji ludzi i ich komunikowaniu się.

Nowe założenie (*P*) wpływające do pamięci oddziałuje w pewien sposób na organizację zbioru przechowywanych założeń (*C*). Proces ten określany jest jako *kontekstualizacja (P) w kontekście (C)* lub jako *implikacje kontekstowe* (analityczne lub syntetyczne). Polega na formułowaniu wniosków, których uzyskanie nie jest możliwe na podstawie samego (*P*) lub (*C*). Powstanie implikacji kontekstowych jest możliwe *wtt*, gdy:

- połączenie $\{P\}$ i $\{C\}$ nietrywialnie implikuje Q ,
- $\{P\}$ nie implikuje nietrywialnie Q oraz
- $\{C\}$ nie implikuje nietrywialnie Q .

Implikacje kontekstowe powstają automatycznie i spontanicznie, będąc źródłem tzw. efektów kontekstowych. Mogą one polegać na wzmocnieniu, osłabieniu lub odrzuceniu przechowywanych założeń, np.:

(1) *Jeśli Kowalscy, Malinowscy i Nowakowie przyszli na imieniny, to imieniny były udane.*

(a) *Kowalscy, Malinowscy i Nowakowie przyszli na imieniny.*

(b) *Imieniny były udane.*

Zdaniu (1b) przypisywana jest wartość confirmacji, która jest tu najwyższa, wyższa niż w sytuacji, gdy:

2 (a) *Kowalscy i Nowakowie przyszli na imieniny,*

(b) *Imieniny były udane.* (Nie było Malinowskich — osłabienie wniosku).

3 (a) *Kowalscy i Malinowscy przyszli na imieniny.*

(b) *Imieniny były udane.* (Nie było Nowaków — osłabienie wniosku).

4 (a) *Kowalscy przyszli na imieniny.*

(b) *Imieniny były udane.* (Nie było Nowaków ani Malinowskich — większe osłabienie wniosku).

5 (a) *Kowalscy, Malinowscy i Nowakowie nie przyszli na imieniny.* -

W zdaniu (5a) następuje odrzucenie wniosku, ponieważ dolna granica konfirmacji nie może być mniejsza od wartości koniunkcji przesłanek.

Ocena wartości konfirmacji zależy od wartości elementów koniunkcji przesłanek, przy czym:

- we wnioskowaniu analitycznym siła wniosku równa jest sile przesłanek,
- we wnioskowaniu syntetycznym są trzy możliwości:

- (a) gdy obie przesłanki są pewne — wniosek jest pewny,
- (b) gdy jedna z przesłanek nie jest pewna — wniosek też nie jest pewny,
- (c) gdy obie przesłanki nie są pewne; siła wniosku jest mniejsza niż słabszej z przesłanek; wniosek jest wówczas niewyraźny (słabo potwierdzony).

W przypadku efektów kontekstowych powstających w wyniku implikacji syntetycznych, każda przesłanka jest argumentem dla wniosku lub jego dowodem w kontekście innych przesłanek, a zatem może wzmacniać wniosek wspólnie implikowany. Poziom wzmocnienia jest funkcją jej własnej siły. Kontekstowe wzmocnianie siły założeń może być zależne od przesłanki $\{P\}$ i założeń tworzących kontekst $\{C\}$ lub niezależne. Wzmocnienie niezależne ma miejsce wówczas, gdy jeden wniosek jest niezależnie implikowany przez dwa różne zbiory przesłanek, np.:

- (1) *Jeśli Kowalscy, Malinowscy i Nowakowie przyszli na imieniny, to imieniny były udane.*
 - (a) *Kowalscy, Malinowscy i Nowakowie przyszli na imieniny.*
 - (b) *Imieniny były udane.*
- (2) *Jeśli imieniny skończyły się późno, to były udane.*
 - (a) *Imieniny skończyły się późno.*
 - (b) *Imieniny były udane.*

Innym rodzajem wzmocnienia przesłanek jest *wzmocnienie przez kombinację* prowadzącą do dodania innych implikacji, wzmocnienia posiadanych założeń, eliminacji założeń fałszywych oraz usuwania założeń sprzecznych z uwzględnieniem ich siły. Usuwanie założeń sprzecznych obejmuje również te, które implikują je analitycznie jako przesłanki oraz słabsze pary założeń syntetycznych. Jest to procedura rekursywna.

Wzmocnianie retroaktywne zachodzi wówczas, gdy założenie efektywnie używane w kontekstualizacji ulega wzmocnieniu przez fakt, że kontekstualizacja przyniosła oczekiwane rezultaty, np. w komunikacji werbalnej, w której podstawą siły założeń są gwarancje (autorytet) nadawcy informacji. Oznacza to, że odbiorca O zakłada, że nadawca N włączy do wypowiedzi relewantną informację $\{P\}$ i uważa N za prawdomównego (wiarygodnego). Gdy $\{P\}$ okazuje się relewantne w oczekiwany sposób, do kontekstu wprowadzane jest wzmocnione założenie Q .

Zasada kontekstualizacji odwzorowywanej informacji o rzeczywistości oraz powstawanie efektów kontekstowych w wyniku mechanizmów inferencyjnych ma podstawowe znaczenie w charakterystyce procesu rozumienia oraz w definiowaniu relewancji jako psychologicznej cechy procesów mentalnych.

W świetle scharakteryzowanego modelu przetwarzania informacji mentalnej **relewancja jest komparatywną i stopniowalną cechą informacji**, określaną jako relacja między założeniem w kontekście, w którym jest przetwarzana, przy uwzględnieniu biologicznego wysiłku organizmu.

Def. 1. Założenie jest relewantne w kontekście w tym zakresie, że jego efekty kontekstowe dla tego kontekstu są duże.

Def. 2. Założenie jest relewantne w tym zakresie, że wysiłek wymagany do przetworzenia w tym kontekście jest mały.

Informacja jest nierelewantna wówczas, gdy nie powoduje efektów kontekstowych, a więc nie dochodzi do jakościowej i ilościowej transformacji informacji, co ma miejsce wówczas, gdy:

a) nowa informacja (założenie) nie jest powiązana semantycznie z informacją tworzącą kontekst, np. *5 maja 1881 r. był w Kabulu dniem słonecznym*;

b) założenie jest już obecne w kontekście, nie zmienia się jego siła, np. ktoś stwierdza *Czytasz teraz książkę* w momencie, gdy jesteśmy zajęci czytaniem książki;

c) założenie jest niespójne z kontekstem, np. *Jesteś senny* — w momencie, gdy tak nie jest.

Relewancję założeń można oceniać tylko w kontekście. Jako ilustrację przytoczymy podany przez Sperbera i Wilson następujący dialog między przechodniem i sprzedawcą chorażówek: [Sperber D., Wilson D., 1986, 121-122]

S: Czy zechciałby Pan kupić chorażówkę na rzecz Royal National Lifeboat Institution?

P: Nie, dziękuję. Zawsze spędzam wakacje z siostrą w Birmingham.

Aby dostrzec relewancję odpowiedzi przechodnia, trzeba znać przesłanki (założenia) użyte przez przechodnia i wygenerować pewne syntetyczne implikacje kontekstowe:

(1) *Birmingham jest położony w głębi kraju.*

(2) *Royal National Lifeboat Institution jest instytucją charytatywną.*

(3) *Kupienie chorażówki jest formą składki na działalność charytatywną.*

(4) *Ktoś, kto spędza wakacje w głębi kraju nie potrzebuje korzystać z pomocy Royal National Lifeboat Institution.*

(5) *Od kogoś, kto nie potrzebuje korzystać z pomocy instytucji charytatywnej, nie należy spodziewać się składek.*

(6) **Implikacja : przechodzień nie jest zainteresowany propozycją.**

Istotnym problemem w przetwarzaniu nowej informacji, a więc nowych założeń komunikowanych werbalnie lub ostensywnie, jest dobór podzbioru ze zbioru informacji przechowywanej w pamięci, tworzącego kontekst. Czy kontekst przetwarzania założeń jest dany czy wybierany?

W literaturze poświęconej ustnej komunikacji werbalnej uważa się, że kontekst niezbędny do zrozumienia wypowiedzi bądź jest dany, a wyrażone explicite założenie jest łącznikiem między kontekstem obecnym w umyśle odbiorcy na początku rozmowy, bądź jest określony przez poprzednie wypo-

wiedzi. Autorzy prezentowanej teorii relewancji reprezentują pogląd, że nie jest możliwe arbitralne określenie, który z podzbiorów założeń może stać się kontekstem przetwarzania, ponieważ w naturze kontekstu nie ma niczego takiego, co wykluczałoby możliwość tworzenia nowego kontekstu lub jego zmiany w trakcie procesu rozumienia (interpretacji) wypowiedzi. Ilustrują to warianty możliwego przebiegu dialogu, zależnie od założeń wyrażanych lub implikowanych przez wypowiedź [Sperber D., Wilson D., 1986, 138-140]:

(1) Wszystkie założenia są wyrażone *explicite*:

P(eter): *Jestem zmęczony.*

Impl.: *Peter chce, żeby Mary przygotowała posiłek.*

M(ary): *Skoro jesteś zmęczony, przygotuję posiłek.*

P: *Jestem zmęczony.*

M: *Przygotuję posiłek.*

(2) Założenia są implikowane:

(2.1.)

P: *Jestem zmęczony.*

M: *Deser jest gotowy. Przygotuję główne danie.*

Implikacja syntetyczna: *Posiłek składa się z co najmniej dania głównego i deseru. Mary przygotowuje posiłek.*

(2.2.)

P: *Jestem zmęczony.*

M: *Deser jest gotowy. Przygotuję osso-buco¹⁸.*

Implikacje syntetyczne: *Ossobuco jest daniem głównym. Posiłek składa się z co najmniej dania głównego i deseru. Mary przygotowuje posiłek.*

Założenia wyrażone i implikowane przez poprzednie wypowiedzi oraz wiedzę encyklopedyczną (założenia faktualne) mogą być związane z którymkolwiek pojęciem użytym w nowej wypowiedzi. Jednym ze wstępnych etapów rozumienia wypowiedzi jest identyfikowanie pojęć i dodanie do kontekstu związanej z nimi, już posiadanej wiedzy encyklopedycznej:

(2.3.)

P: *Jestem zmęczony.*

M: *Deser jest gotowy. Przygotuję specjalność restauracji Capri.*

Implikacje syntetyczne:

Specjalnością restauracji Capri jest osso-buco. Ossobuco jest daniem głównym. Posiłek składa się co najmniej z dania głównego i deseru. Mary przygotowuje posiłek.

Automatyczne wypełnianie kontekstu w rezultacie implikacji dużą ilością informacji encyklopedycznej, która wzmaga efekty kontekstowe, powoduje zwiększenie wysiłku na przetworzenie, np.:

(2.4.)

P: *Jestem zmęczony.*

M: *Deser jest gotowy. Przygotuję specjalność tej restauracji, obok której mieszka John.*

¹⁸ Potrawa — noga cielęca z kością szpikową.

Implikacje syntetyczne:

John mieszka obok restauracji Capri. Specjalnością restauracji Capri jest osso-buco. Osso-buco jest daniem głównym. Posiłek składa się co najmniej z dania głównego i deseru. Mary przygotowuje posiłek.

Dla każdego elementu nowej informacji za kontekst może posłużyć wiele różnych zbiorów założeń z różnych ośrodków (pamięci długofalowej, średniofalowej, doraźnej percepcji itp.), co pozwala sądzić, że nie da się określić arbitralnego kontekstu przetwarzania informacji przez człowieka, a klasę potencjalnych kontekstów ogranicza organizacja długofalowej pamięci encyklopedycznej. Jej modelowaniem zajmuje się m.in. sztuczna inteligencja i stosowane w niej sposoby reprezentacji wiedzy (ramy, prototypy, scenariusze itp. — zob. rozdz. 3). Przyjmuje się, że wiedza encyklopedyczna jest przechowywana w postaci zgrupowanej — klocków (*chunks*), a nie w formie pojedynczych założeń, i nawet najmniejsze pobierane porcje informacji są klockami. Nawiązując do podanych wyżej przykładów można przypuszczać, że informacja o „osso-buco” jest przechowywana z informacją o „restauracji Capri” i „winem Valpolicella” tam serwowanym jako specjalność.

Brak dostatecznie uzasadnionej teorii wyszukiwania informacji konceptualnej nie pozwala uwzględnić w charakteryzowanej teorii czynnika dostępności założeń jako wpływającego na ocenę relewancji. Można jedynie spekulatywnie zakładać, że łatwiej dostępne są założenia już ocenione w procesach inferencyjnych, przy czym należałoby rozróżnić dostęp w jednym kroku lub w wielu, co oznacza rozszerzenie kontekstu w zakresie często nieopłacalnym z punktu widzenia wysiłku organizmu. Niekiedy informacja niezbędna do przetwarzania jest łatwiej dostępna w otoczeniu niż w pamięci, np. ustalenie daty urodzin postaci historycznej przez sprawdzenie w encyklopedii zwykle wymaga mniej wysiłku niż jej wyszukanie w pamięci. Za prawdopodobne uważa się równoczesne przetwarzanie nowych założeń w wielu kontekstach ze względu na dążenie do ustalenia różnych kryteriów relewancji.

Na zakończenie każdego procesu transformacji informacji (dedukcji) człowiek dysponuje pewnym zbiorem kontekstów, częściowo uporządkowanym przez relację inkluzji, która determinuje (na poziomie psychologicznym) kolejność dostępu do poszczególnych podzbiorów, stanowiąc — jak wspomniano — jeden z istotnych czynników w ocenie relewancji. Spośród innych czynników ważna także jest organizacja pamięci encyklopedycznej, z której korzystanie przez ośrodek dedukcji wpływa z kolei na wzbogacanie jej zasobów. Dotyczy to w szczególności założeń odwzorowywanych *explicite*, istotnych w implikacji analitycznej.

Przetwarzanie informacji konceptualnej nie ogranicza się w świetle teorii relewancji do prostego wychwycenia założenia zawartego w wypowiedzeniu nadawcy informacji i przeprowadzenia dedukcji na podstawie informacji encyklopedycznej. Istotne znaczenie mają także implikacje syntetyczne, których derywacja angażuje dwa zbiory założeń: służących jako kontekst i przetwarzane w danym

kontekście, co można wykorzystać w psychologicznym wyjaśnieniu różnicy między informacją „będącą w centrum uwagi” i informacją „pozostającą w tle”.

Tak więc, w ujęciu teorii Sperbera i Wilsona relewancja stanowi niereprezentowalną własność procesów mentalnych (*non-representational dimension*), a jeśli nawet stanowi, to reprezentacja ta przyjmuje postać osądów porównawczych lub zgrubnie absolutnych (nierelewantny, bardzo relewantny, mało relewantny), nie zaś ocen kwantytatywnych i jest silnie uzależniona od wiedzy encyklopedycznej i językowej człowieka.

2.5. INTERPRETACJA RELEWANCJI W INFORMACJI NAUKOWEJ

Dyskusje nad definicją relewancji w informacji naukowej ilustrują dość charakterystyczne cechy procesów komunikacji w nauce. Z jednej strony reprezentują typowe etapy rozwoju terminologii nowej dyscypliny (od sformułowania problemu i prób przedstawienia pierwszych definicji poprzez ich kolejne weryfikacje po próby stawiania hipotez), z drugiej zaś wykazują swoisty izolacjonizm wyrażający się we fragmentarycznym uwzględnianiu dorobku innych dyscyplin w rozwiązywaniu tego samego problemu.

Pierwsze użycie w informacji naukowej terminu relewancja (relewantny) w znaczeniu obecnie przyjmowanym przypisuje się S. C. Bradfordowi, który w latach trzydziestych posłużył się nim dla określenia pewnych mechanizmów selekcji informacji zawartej w czasopismach naukowych, wprowadzając podział artykułów na „relewantne i nierelewantne”. Miało to miejsce w sytuacji zachwiania równowagi w „ekologii komunikacji naukowej” spowodowanej:

- wzrostem liczby publikacji naukowych związanym z ilościowym rozwojem kadr naukowych,

- brakiem lub raczej niedostatkami kryteriów selekcji powiększających się zbiorów informacji bibliotecznej,

- równoczesnymi procesami dezintegracji (specjalizacji) i integracji w nauce (powstawania dziedzin interdyscyplinarnych) [Bradford S. C., 1932].

Początków kryzysu komunikacyjnego w nauce należałoby upatrywać — zdaniem niektórych — w wieku XVII, od momentu powstania pierwszych czasopism naukowych, stanowiących podstawowy element mechanizmu przekazywania wiedzy cząstkowej dotyczącej pewnych problemów (systematycznego rozpowszechniania informacji) w ramach pewnych sieci służących równocześnie jej integracji [Price D. J. de S., 1965; Saracevic T., Rees A. M., 1963]. Środkiem służącym wyróżnieniu wiedzy (informacji) relewantnej i nierelewantnej są najpierw klasyfikacje biblioteczne, następnie metody indeksowania i dokumentowania publikacji.

Dyskusje nad metodami efektywnej selekcji informacji z relewancją jako głównym kryterium efektywności rozpoczynają się w latach czterdziestych i pięćdziesiątych XX wieku, kiedy to powstają systemy informacyjno-wyszuki-

wawcze jako wyspecjalizowane mechanizmy selekcjonowania informacji, komplementarne wobec niedostatecznej sprawności bibliotek w komunikowaniu wiedzy. Przy ogólnym consensusie w sprawie zasadności przyjęcia kryterium relewancji w wartościowaniu informacji dostarczanej użytkownikowi przez system oraz akceptacji interpretacji terminu jako pewnej relacji między różnie ujmowanymi elementami procesu komunikacyjnego podejmowane są próby sformułowania „dobrej” definicji relewancji, przyjmujące formę hipotez wyliczających możliwe do uwzględnienia argumenty relacji według formuły:

RELEWANCJA JEST TO :

COSĆ (A) — miara, stopień, wymiar, ocena, relacja

CZEGOŚĆ (B) — odpowiedniości, użyteczności, związku, usatysfakcjonowania, dostosowania, zbieżności

ISTNIEJĄCE MIĘDZY CZYMSĆ (C) — dokumentem, artykułem, formą tekstu, informacją, faktem

ORAZ CZYMSĆ (D) — zapytaniem informacyjnym, instrukcją wyszukiwawczą, wykorzystywaną informacją, potrzebą informacyjną, punktem widzenia

OKREŚLONEGO PRZEZ PEWNEGO (E) — osobę, użytkownika, autora sądu, autora zapytania, specjalistę w dziedzinie informacji. [Saracevic T., Rees A. M., 1963, 328]

Interpretacja znaczenia tego podstawowego dla mechanizmów selekcji informacji terminu obciążona była wieloma błędami. W definicjach stosowano jako definiensy terminy semantycznie złożone bądź nieostre wyrażenia języka potocznego. Nie wykazywano związków semantycznych między terminami używanymi jako definiensy, denotującymi różne elementy procesu komunikacji użytkownika z systemem. Interpretując relewancję jako relację wieloargumentową, traktowano ją w kategoriach celu absolutnego, o którego osiągnięciu rozstrzyga się binarnie: dokument wyszukany powinien być relewantny, niewyszukany — nierelwantny.

Przyczyn takiego uwikłania definicji relewancji w błędne koło należałoby upatrywać nie tylko w nieporadności metodologicznej prac publikowanych w latach pięćdziesiątych, lecz także w warunkach dialogu między użytkownikiem i systemem informacyjno-wyszukiwawczym. Dyskusje nad pomiarami efektywności SIW na podstawie potocznego kryterium relewancji słabną lub przenoszą się w sferę rozważań teoretycznych praktycznie z początkiem lat siedemdziesiątych. Nie prowadzi się eksperymentów w zakresie oceny relewancji. Rok 1970 można umownie traktować jako cezurę w rozwoju systemów informacyjnych, określonych w naszych rozważaniach jako tradycyjne i dialogowe [Saracevic T., 1975]. W systemach tradycyjnych statyczny dialog użytkownika z systemem miał doprowadzić do uzyskania odpowiedzi relewantnej we wszelkich możliwych aspektach. W systemie dialogowym (interakcyjnym) pojawiają się możliwości korygowania przyjętych kryteriów relewancji lub

wprowadzenia innych kryteriów. Znamienne jest również to, że wprowadzenie pewnego ładu interpretacyjnego w sporach nad znaczeniem terminu pierwotnego (primitive) w teorii systemów informacyjnych związane jest z wykorzystaniem logicznej interpretacji relewancji, praktycznie ignorowanej przez dwa dziesięciolecia mimo znacznego dorobku logiki w tym zakresie. Próby formułowania teorii relewancji na gruncie teorii komunikacji nawiązują już do ujęć relewancji w logice, filozofii, a także psychologii.

Źródłem sporów nad relewancją w informacji naukowej jest złożoność procesów komunikacji, które można opisać w aspekcie aplikacyjnym lub teoretycznym, przyjmując jako definiendum cechy dystynktywne elementów tych procesów należące do systemu informacji lub też do jego otoczenia. Ujęcie aplikacyjne dominuje w początkowym okresie dyskusji, chociaż wprowadzeniu terminu relewancja przez Bradforda przyświecał cel teoretyczny, a mianowicie sformułowanie kwantytatywnych praw komunikacji naukowej wiążące się z powstawaniem bibliometrii i naukometrii.

Przegląd różnych ujęć metodologicznych i definicji relewancji w informacji naukowej przeprowadził T. Saracevic, do którego pracy odwołamy się jako ilustrującej problem w stopniu najbardziej reprezentatywnym, wprowadzając niezbędne uzupełnienia na podstawie innych źródeł [Saracevic T., Rees A. M., 1963, 328]. Jako cechą znaną dla poszczególnych postaw metodologicznych wymienia on swoisty izolacjonizm, przejawiający się w ignorowaniu istniejących już interpretacji, oraz jednostronność, często połączoną z nadmiernym pragmatyzmem. Sam uznaje kompleksowe, teoretyczne podejście do relewancji za problem zasadniczy w rozwoju komunikacji w nauce, w tym systemów informacyjnych, stwierdzając: „Historia nauki dostarcza wielu dowodów na to, że nie ma nic bardziej praktycznego niż dobra teoria. (...) Brak teorii prowadzi najczęściej do zgadywanek z niewielkim prawdopodobieństwem pozytywnego wpływu” [Saracevic T., 1975, 337-338].

Saracevic podjął także próbę klasyfikacji, a właściwie typologii interpretacji, przyjmując za punkt wyjścia dwa pytania:

1) jakie cechy dystynktywne (elementy, czynniki) przyjmowano za konieczne dla opisu znaczenia terminu?

2) jakie relacje specyfikowane są w definicjach tego terminu?

Analiza definicji i poglądów ich autorów na relewancję pozwoliła ustalić następujące wyrażenia konotowane przez ten termin w odpowiedzi na pytanie (1) wymieniane jako argumenty „relacji relewancji”:

— wiedza przedmiotowa¹⁹,

¹⁹ Rozróżnienie znaczeń terminów *wiedza* i *informacja* nie jest zbyt precyzyjne w rozważaniach Saracevica. Wiedzy przypisuje jako cechę dystynktywną organizację odwzorowującą m.in. proces jej nawarstwienia (sedimentation). Również związki między wiedzą i dokumentami pierwotnymi (subject literature) są określone enigmatycznie: „... są powiązane, lecz także dość odmienne”... różnią się strukturą, dynamiką, łączy je cecha redundancji [Saracevic T., 1975, 332].

- literatura przedmiotowa (subject literature), tj. ich treść ujmowana przedmiotowo,
- języki służące reprezentacji wiedzy,
- zbiory informacyjne systemu,
- adresaci (użytkownicy) informacji, ze szczególnym uwzględnieniem ich zasobu wiedzy (informacji),
- systemy informacyjne,
- otoczenie (environment) użytkownika, tj. realia wykorzystywania informacji,
- systemy wartości stanowiące punkt odniesienia dla orzekania o relewancji.

Na tej podstawie Saracevic wyróżnił kilka ujęć relewancji, uwzględniając częściowo chronologię ich powstawania:

1. Aspekt „wiedzy przedmiotowej” — relewancja jako relacja między wiedzą o przedmiocie lub na temat przedmiotu (*on or about the subject*) i pytaniem/problemem (*question/topic*) dotyczącym przedmiotu.

2. Aspekt „literatury przedmiotowej” — relacja między literaturą (dokumentami pierwotnymi) i pytaniem/problemem dotyczącym przedmiotu.

3. Aspekt logiczny — relacja między wnioskiem i przesłanką w ujęciu dedukcyjnym lub probabilistycznym.

4. Aspekt systemowy — relacja między systemem informacyjnym (jego zbiorami i procesami) oraz pytaniem/problemem lub literaturą przedmiotową lub użytkownikiem.

5. Aspekt odbiorcy informacji — ocena (*judgement* — sąd o) relacji między dokumentem i zapytaniem wydawana przez człowieka.

6. Aspekt pertynencji, aspekt odbiorcy informacji — relacja między zbiorem wiedzy użytkownika (*stock of knowledge at hand*) i wiedzą przedmiotową lub literaturą przedmiotową.

7. Aspekt pragmatyczny — relacja między problemem rozwiązywanym bezpośrednio przez użytkownika oraz dostarczoną informacją, przy uwzględnieniu kryterium użyteczności i preferencji jako podstawy wnioskowania.

Kryteria wyróżnienia powyższych ujęć relewancji nie wydają się dostatecznie ostre. W szczególności wątpliwości budzi niejasne rozróżnienie kryteriów: wiedza — informacja; literatura — dokument — zbiory informacyjne systemu — system informacyjny. Ponadto trudno zgodzić się z Saracevicem, że chronologia interpretacji relewancji odwzorowuje precyzję kryteriów służących jako *explicata*. Dlatego też wśród eksplikacji relewancji wyróżnimy:

- interpretacje umieszczające relewancję w sferze relacji między elementami systemu informacyjnego (zbiorami informacji i wyrażeniami języka),
- interpretacje umieszczające relewancję w sferze relacji między systemem informacyjnym i elementami jego otoczenia.

2.5.1. Relewanca jako relacja między elementami systemu informacyjno-wyszukiwawczego

W historii rozwoju skomputeryzowanych systemów informacyjnych najwcześniej powstała binarna interpretacja relewancji, zgodnie z którą każda odpowiedź systemu spełniająca warunek minimalnej zgodności z instrukcją wyszukiwawczą (wydana przez system) jest relewantna. W ujęciu tym, obecnie określanymi terminami *relewanca techniczna*, *relewanca formalna*, *relewanca inżynierska* traktuje się relewancję jako relację między wyrażeniami języka, składającymi się na zbiory systemu i wyrażeniami zawartymi w instrukcji wyszukiwawczej. Skala oceny relewancji jest binarna.

Kryterium relewancji Bradforda, nie poparte co prawda definicją explicite, stosowane w bibliometrii i językoznawstwie statystycznym polega na interpretacji relewancji jako funkcji przyporządkowującej na podstawie wykładników językowych (terminów języka naturalnego) cechę bycia przedmiotem treści (tematem) zbiorom dokumentów (pierwotnych artykułów), stanowiących zbiór informacyjny systemu nauki. Skala oceny relewancji jest tu wielostopniowa i pozwala wyznaczyć w zbiorze informacji pewne podzbiory (strefy) odpowiadające pewnym przedziałom funkcji. Podobne stanowisko wobec relewancji zajmowano w badaniach kwantytatywnych, na przykład w pracach Lotki nad dystrybucją częstości prac różnych autorów w poszczególnych dyscyplinach [Lotka A. J., 1926], Price'a nad paradygmatem (sieciami cytowań) w literaturze [Price D. J. de S., 1965], Urquarta nad wykorzystaniem czasopism w bibliotekach naukowych [Marulli L., Koenig M. E. D., 1979 [Urquart D. J., 1959].

Modele statystyczne języka naturalnego skonstruowane przez Zipfa wraz ze sformułowaną przez niego tezą dotyczącą ekonomii wysiłku w komunikacji językowej popartą analizą częstości wyrażeń języka w tekstach i słowniku można uznać za ilustrację ilościowo ujmowanej zasady relewancji, określanej przez wyraźne kryteria: relację między częstością wyrazów i ich długością, częstością wyrazów i liczbą ich znaczeń, wiekiem i pochodzeniem oraz między częstością wyrazu i jego rangą [Sambor J., 1972].

Próby powiązania z relewancją prawidłowości ustalonych przez bibliometrię dla zbiorów informacji pierwotnej podejmowali Kozakchov oraz Saracevic, który wysunął hipotezę o zdeterminowaniu przez mechanizm relewancji struktury zbiorów informacji w nauce w przekroju diachronicznym [Kozakchov L. S., 1969, 3-11] [Saracevic T., Rees A. M., 1963].

Wśród poglądów na relewancję wiążących ją ze sferą relacji między zbiorami informacji i wyrażeniami językowymi wyróżnia się podgrupa poglądów autorów postulujących konieczność interpretacji tego terminu w informacji naukowej w języku logiki i/lub matematyki.

Interesującą analizę problemów związanych z definiowaniem relewancji przedstawił J. D. Hillman [Hillman J. D., 1977, 23-30]. Odwołał się do materia-

łów z Międzynarodowej Konferencji nt. Informacji Naukowej w Waszyngtonie w 1958 r., w których zreferowano dyskusję nad relewancją i powiązaniem informacji w systemie (*relatedness*). Swoje uwagi dotyczące dwu nurtów w dyskusji połączył następnie z próbą sformułowania definicji relewancji na podstawie ekstensjonalnej interpretacji znaczenia oraz przedstawionej przez Carnapa metody opisu własności elementów klas przez wykorzystanie relacji podobieństwa (*similarity*). Metoda ta stanowi część omówionych w podrozdziale 2.3. zasad eksplicacji semantycznej wyrażen w nauce.

W dyskusji nad relewancją jako systemową relacją określającą związek zapytania informacyjnego z klasami dokumentów w systemie B. C. Vickery reprezentował pogląd, iż relewancja jest cechą systemu ustalaną przez jego projektanta według następujących kryteriów [Hillman J. D., 1977, 23-30]:

- dokument (charakterystyka wyszukiwawcza) jest (lub nie jest) zakwalifikowany jako dotyczący określonego przedmiotu (treści),
- dokumenty zakwalifikowane są (lub nie są) jako dotyczące pewnych przedmiotów treści, w których zawarty jest dany przedmiot (1),
- dokumenty zakwalifikowane są (lub nie są) jako dotyczące pewnych przedmiotów zawartych w przedmiocie (1),
- dokumenty zakwalifikowane są jako dotyczące pewnych przedmiotów współrzędnych z (1).

Zdaniem Bar-Hillela takie ujęcie relewancji jako cechy systemu sprowadza problem odwzorowania związków między klasami dokumentów w zbiorze systemu do formuły „*A* jest związane z *B* wtedy i tylko wtedy, jeśli *A* i *B* występują razem w pewnym *C*, a ponadto uwzględnia się wszystkie wywodzące się z nich klasy w pewnym łańcuchu określonej długości” [Hillman J. D., 1977, 23-30]. Taka formuła relewancji stanowi podstawę klasyfikacji hierarchicznych w rozwoju systemów informacji.

Opowiadając się za stanowiskiem Bar-Hillela, Hillman stwierdza, iż kluczem do wyjaśnienia problemu relewancji jako relacji między zapytaniem użytkownika i zbiorem dokumentów w systemie jest zinterpretowanie znaczenia (*concept*) oraz związków semantycznych (*conceptual relatedness* — por. rozdz. 2.3.2.). Konieczne jest skonstruowanie teorii relewancji opisującej w sposób formalny relację podobieństwa sądów wyrażonych przez użytkowników o wynikach wyszukiwania informacji. Na tej podstawie byłoby możliwe określenie uniwersalnej zasady ustalania klas dokumentów „podobnych semantycznie” (*conceptually similar*), z uwzględnieniem zmienności i wielowartościowości relewancji. Sformułowanie teorii relewancji wymaga — zdaniem Hillmana — rozwiązania dwu kolejnych problemów:

- zdefiniowania klas ekwiwalencji znaczeniowej, odpowiadających klasom denotatów, przyjmując jedynie za podstawę podobne sądy uporządkowane parami wyrażone, przez użytkownika;
- zdefiniowania związku znaczeniowego (*relatedness of concepts*) przez określenie, jakie klasy można utworzyć na podstawie szerokiej relacji podobieństwa.

Metodę, której przydatność do budowy teorii relewancji analizuje Hillman, przedstawił Carnap w dwu wersjach. Wersja pierwsza zakłada, iż dziedzinę relacji stanowi wykaz par jednostek elementarnych (*ground element pairs*), natomiast relacja podobieństwa rozumiana jest jako symetryczna, zwrotna relacja częściowej identyczności (*part-identity*); w przypadku systemu informacyjnego — częściowej identyczności charakterystyk wyszukiwawczych dokumentów i instrukcji wyszukiwawczej. Warunek związku znaczeniowego spełniony byłby wówczas, gdy:

- 1) każda badana para z klasy znaczeniowej będzie częściowo identyczna,
- 2) klasę znaczeń powinny stanowić największe spośród spełniających warunek (1):

$$(x) (y) [(x \in a, y \in a) R x, y]. \bullet (x) (\exists y) [(y \in a \supset R x, y) x \in a].$$

Ograniczenia tej metody nazwał Hillman „problemem towarzyszenia” (*companionship difficulty*), powstającym w przypadku klas znaczeniowo niezależnych oraz „niedoskonałej wspólnoty” (*imperfect community*), gdy każda para zbioru charakterystyk wyszukiwawczych posiada wspólne cechy znaczeniowe, lecz nie posiada ich cały zbiór, np. 1.abc 2.bcd 3.cde 4 def. [Hillman J. D., 1977, 23-30].

W drugiej wersji metody Carnapa konstrukcja klas znaczeniowo związanych oparta jest na relacji częściowego podobieństwa *Ae* (*part-similarity*) zastępującej relację częściowej identyczności. Relacja ta zachodzi między każdą z dwu jednostek elementarnych (ChWD), które bądź posiadają wspólną własność bądź posiadają własności takie, że własność jednej jest podobna do własności drugiej jednostki. Na podstawie relacji *Ae* budowane są tzw. kręgi podobieństwa (*similarity circles*). Klasa własności *b* powinna spełniać następujące warunki:

- 1) w całości należeć do każdego z kręgów podobieństwa, który zawiera co najmniej połowę *b* oraz

- 2) dla każdej jednostki elementarnej *x* należącej do *b* istnieje przynajmniej jeden krąg podobieństwa, do którego nie należy *x*, lecz który zawiera w całości *b*.

Podobieństwo własności (*similarity between qualities*) *Aq* zdefiniowano jako:

$$Aq = Df \alpha\beta \{ \alpha \bullet \beta \in qual. \alpha \subset \beta Ae$$

— relacja *Aq* jest klasą wszystkich α oraz klasą wszystkich β , gdzie α i β są klasami podobnymi takimi, że każdy element α jest częściowo podobny (*Ae*) do każdego elementu β .

W konkluzji Hillman stwierdza, że przy ekstensjonalnej interpretacji znaczenia nie jest możliwe zdefiniowanie podobieństwa (powiązania znaczeniowego) dokumentów na podstawie metody Carnapa także w wersji drugiej, ponieważ:

- 1) nie dysponujemy teorią umożliwiającą organizację i strukturalizację zapytań użytkowników dla wszystkich relewantnych dokumentów (ChWD) ze względu na cechę „przedmiot dokumentu”;

- 2) udowodniono, iż nie jest możliwe utworzenie klas obiektów podobnych na podstawie posiadania wspólnych predykatów (*predicate-sharing*): „Liczba

predykatów wspólnych dla dowolnej pary rozróżnialnych obiektów jest stała. Zatem, jeśli podobieństwo ocenia się na podstawie wspólności predykatów, żadne dwa obiekty nie są bardziej podobne niż jakakolwiek inna para obiektów. Liczba predykatów wspólnych dla dowolnego zbioru różnych obiektów jest stała, jakkolwiek zbiór by to był” [Satosi Watanabe, 1966].

Po wykazaniu nieadekwatności zasady podobieństwa znaczeniowego klas charakterystyk wyszukiwawczych do sformułowania uniwersalnej teorii relewancji informacji, Hillman wyraził przypuszczenie, iż bardziej obiecujące byłoby rozważenie podstaw, na jakich opiera się formułowanie przez użytkowników orzeczeń o relewancji, zapowiadając próbę sformułowania innej wersji teorii relewancji. Zapowiedzi tej nie zrealizował dotąd, co nie jest zaskakujące zważywszy, że oznaczałoby to przeniesienie interpretacji relewancji do sfery psychologii i intencjonalnego ujęcia znaczenia.

Autorem, którego poglądy na relewancję wywarły inspirujący wpływ na interpretację terminu oraz na rozwój kryteriów oceny sprawności systemów informacyjnych był W. S. Cooper [Cooper W. S., 1971]. W artykule *A definition of Relevance for Information Retrieval* poddał krytyce dość szeroko akceptowaną definicję Cuadry i Kattera (odmawiając jej statusu definicji), którą prezentujemy wśród poglądów łączących relewancję informacji z różnymi elementami otoczenia systemu (pkt. 2.5.2. niniejszego rozdziału). Podobnie jak Hillman, podstaw definicji relewancji w informacji naukowej upatrywał w logice. Krytykując niewyraźne określenia relewancji nawiązujące do kategorii psychologicznych, przyjął następujące założenia i rozróżnienia:

- informacja jest zawsze reprezentowana w formie językowej,
- relewancja jest relacją definiowaną dla reprezentacji językowych,
- podstawową jednostką języka w reprezentacji informacji jest zdanie,
- dane zawarte w zbiorze informacji systemu informacyjnego można traktować jako zbiór zdań oznajmujących języka poprawnie zbudowanego (*well-formed*), na przykład języka predykatów pierwszego rzędu,
- podstawowym konstruktem logicznym niezbędnym do zdefiniowania relewancji w systemie informacyjnym jest relacja wynikania (*entailment*),
- wyszukiwanie informacji w SIW ma charakter inferencyjny, polega na dedukowaniu bezpośrednich odpowiedzi dla pytań (*premises*).

Za źródło nieporozumień w dyskusjach nad relewancją uważa Cooper nierozróżnianie poszczególnych aspektów wyszukiwania informacji, a mianowicie:

— utożsamianie „potrzeby informacyjnej” użytkownika odnoszącej się do sfery stanów psychicznych, niedostępnych bezpośrednio obserwacji z zapytaniem informacyjnym (*information query*), będącym jedynie wstępnym, przybliżonym odwzorowaniem potrzeby informacyjnej,

— utożsamianie zapytania informacyjnego z instrukcją wyszukiwawczą (*the request*), które tylko niekiedy są identyczne,

— utożsamianie „logicznej relewancji informacji” z „użytecznością informacji”.

Użyteczność informacji jest zdeterminowana przez różnorodne czynniki pozasystemowe, do których należą: nowość, łatwość orzekania o relewancji przez system (*humanly discernible relevance*), wartość przypisywana zdaniom języka reprezentującego informację oraz stopień wiarygodności informacji (*credibility*); użytkownik z zasady niedowierza dokładności systemu. Kryterium użyteczności traktowane jest przy tym jako neutralne — celem systemu jest dostarczenie użytkownikowi informacji użytecznej, nie zaś zaledwie relewantnej.

Relewancję logiczną wyszukiwanej informacji Cooper interpretuje jako dwuargumentową relację między zdaniami zbioru informacyjnego systemu i zdaniem (zdaniami) odwzorowującymi potrzebę informacyjną użytkownika: „Zdanie ze zbioru informacji jest logicznie relewantne dla odwzorowania potrzeby informacyjnej wtedy i tylko wtedy, gdy jest elementem minimalnego zbioru przesłanek pewnego zdania składowego z odwzorowania potrzeby”, „(...) minimalny zbiór przesłanek dla zdania składowego (instrukcji wyszukiwawczej) rozumiany jest jako najmniejszy z możliwych, to znaczy taki, że usunięcie któregośkolwiek z jego elementów spowoduje, że powstałe zdania przestaną być logiczną konsekwencją zbioru przesłanek” [Cooper W. S., 1971, 20].

Powyższe definicje opatrzone zostały zastrzeżeniami i komentarzem, który podsumujemy następująco, traktując poglądy autora jako istotne dla dalszych rozważań, a ponadto uproszczone w pracy Saracevica.

1. Definicji relewancji sformułowanej w języku naturalnym można jedynie przyznać status definicji „w zasadzie”. Faktycznie sprowadza się do zastąpienia nie zdefiniowanego terminu relewancja innym równie nieostrym terminem „implikacja” (wynikanie), nie zdefiniowanym dla języka naturalnego. Do sformułowania precyzyjnej definicji niezbędny jest odpowiedni język formalny.

2. Istnieje różnica między mechanizmami wyszukiwania w systemach informacji faktograficznej i dokumentacyjnej. W przetwarzanie informacji w systemach faktograficznych zaangażowany jest tylko zbiór informacji systemu oraz przyjęte w nim reguły dedukcji, stosowane przez sam system. Wyszukiwanie informacji w systemie dokumentacyjnym jest logicznie bardziej podobne do wnioskowania probabilistycznego niż dedukcyjnego. Proces ten angażuje dwa zbiory informacji: systemowy oraz wiedzę użytkownika, program wyszukiwawczy oraz mózg użytkownika jako urządzenie inferencyjne. Na początku procesu wszystkie zdania zbioru informacji są — z punktu widzenia systemu — elementami zbioru przesłanek, posiadają jednakowe prawdopodobieństwo bycia relewantnymi. Prawdopodobieństwo to zmienia się w miarę wyszukiwania, pewne zdania zostają uznane za prawdopodobnie relewantne w stopniu wyższym niż inne na mocy kryteriów relewancji arbitralnie ustalonych w systemie. Dla każdego zdania składowego instrukcji wyszukiwawczej musi być znaleziony minimalny zbiór przesłanek. Probabilistyczny z natury proces wyszukiwania i proces dedukcji są — twierdzi Cooper — nieroz-

różnialne w systemie dokumentacyjnym, wyrażając przy tym wątpliwość, aby możliwa była budowa doskonałego systemu (programu) wyszukiwawczego.

3. Definicja relewancji ma swój wymiar logiczno-matematyczny i filozoficzny, jeśli uwzględnić różnice w zasadach wyszukiwania informacji. W wersji podstawowej umożliwia udzielanie przez system odpowiedzi binarnych w dedukcyjnych systemach faktograficznych. W wersji zmodyfikowanej uwzględniającej konieczność odwzorowania relacji między zdaniami składowymi instrukcji wyszukiwawczej tworzącymi pewne drzewo zależności (*component statement tree*), przedstawia się następująco:

„Zdanie zbioru informacji (w pamięci systemu lub jego użytkownika) jest logicznie relewantne dla odwzorowania potrzeby informacyjnej wtedy i tylko wtedy, gdy:

a) jest elementem minimalnego zbioru przesłanek dla zdania składowego w drzewie zależności odwzorującym potrzebę informacyjną oraz

b) dla wszystkich zdań składowych w tym drzewie (z wyjątkiem wierzchołka), będących poprzednikami tego zdania istnieją zbiory przesłanek [Cooper W. S., 1971, 37].

4. „Relewancja logiczna dla potrzeby informacyjnej jest ważnym czynnikiem, lecz nie jedynym w ustalaniu użyteczności informacji ze względu na zaspokojenie tej potrzeby” — brzmi wniosek końcowy [Cooper W. S., 1971, 37].

Znaczenie logicznej interpretacji relewancji tego autora jest wieloaspektowe:

- sformułowana została definicja relewancji precyzyjna w stopniu wyznaczonym przez aparat logiki,
- wyraźnie określono dziedzinę relacji relewancji, sygnalizując jej zdeterminowanie przez sieć relacji odwzorowanych w języku,
- podjęto próbę scharakteryzowania procesu wyszukiwania informacji w systemie w kategoriach logicznych procesów wnioskowania dedukcyjnego i probabilistycznego, wywodzącego się z logiki indukcji,
- explicite sformułowano tezę o istnieniu różnych rodzajów relewancji.

Probabilistyczne ujmowanie relewancji w systemie informacyjnym reprezentuje jedna z pierwszych prac w informacji naukowej, podejmujących teoretyczne zagadnienia wyszukiwania informacji, praca Marona i Kuhna [Maron M. E., Kuhn J. L., 1960, 216-244]. Relewancji przypisali oni znaczenie podstawowe w stopniu takim samym, jak ilości informacji w teorii komunikacji Shannona. Znaczenie terminu relewancja w sensie probabilistycznym zdefiniowano jako wieloargumentową relację między: zapytaniem użytkownika (*user's request*), zakresem tematycznym (*subject area*), instrukcją wyszukiwawczą oraz dokumentem dostarczonym przez system.

Podstawę wyszukiwania miały stanowić, arbitralnie ustalone, wzorce odpowiedzi prawdopodobnie relewantnych. Zakładano przypisywanie charakterystykom wyszukiwawczym dokumentów wartości p reprezentującej systemową ocenę prawdopodobieństwa relewancji, tzw. *status wyszukiwawczy* (re-

retrieval status value) lub *wskaznik relewancji (relevance number)*, dla którego ustala się pewną wartość progową P . Wyszukiwane są te dokumenty, dla których ustalony wskaźnik relewancji p przekracza wartość progową P . Kryterium relewancji Marona i Kuhna ma więc interpretację binarną. Oznaczało to równocześnie zakładanie prawdopodobieństwa zawodności systemu. Model wyszukiwania informacji uwzględniał koszty błędów.

Wiele analogii z ujęciem Marona i Kuhna wykazuje *model wyszukiwania informacji Swetsa*, w którym wskaźnik relewancji (*relevance number*) nazwano wskaźnikiem pertynencji (*pertinence number*), ustalonym na podstawie funkcji odwzorowującej dystrybucję dokumentów uznanych za relewantne i nierелеwantne [Swets J. A., 1963, 245-250].

Z binarną koncepcją relewancji informacji w systemie Marona i Kuhna oraz Swetsa polemizował Bookstein [Bookstein A., 1979, 269-272]. Przedmiotem jego krytyki było także przyjęcie jako semantycznego kryterium binarnej interpretacji informacji cechy „bycia przedmiotem treści” (*topicality*). Wykazując matematycznie nieadekwatność obu modeli, Bookstein zaproponował tzw. *operacyjną interpretację relewancji (operational interpretation of relevance)* i odróżnienie relewancji systemowej od relewancji dla użytkownika oraz wprowadzenie stopni relewancji, ustalonych w pewnym continuum, zależnie od tego, kto podejmuje decyzje co do relewancji informacji — system czy użytkownik. Koncepcja ta nie wnosi w zasadzie nowych pomysłów interpretacyjnych poza interpretacją relewancji jako relacji między użytkownikiem a systemem (elementem otoczenia systemu), co odpowiada rozróżnieniu relewancji i pertynencji informacji omawianej w dalszej części rozdziału (pkt 2.5.2.).

Interesującą koncepcję relewancji, wywodzącą się z nurtu logiczno-probabilistycznego, przedstawili Goffman i Nevill w tzw. epidemicznej teorii komunikacji, nawiązującej nazwą do zasady rozprzestrzeniania się chorób przez efektywny kontakt [Goffman W., Nevill A., 316-334]. Do sformułowania założeń teorii, w której relewancja jest interpretowana jako relacja ekwiwalencji umożliwiająca podział zbioru informacyjnego na klasy ekwiwalencji, wykorzystano matematyczną teorię miar, a ściślej cztery aksjomaty, określające cechy niezbędne i wystarczające miary:

- 1) „bycie wartością rzeczywistą i nienegatywną”,
- 2) zupełna addytywność,
- 3) uporządkowanie,
- 4) posiadanie zera absolutnego.

Aksjomaty te wykorzystano następnie do wykazania, że z zasady probabilistyczny pomiar relewancji powinien uwzględniać także wewnętrzną dynamikę zbioru, tj. wzajemne związki między jednostkami zbioru informacyjnego systemu. Pomiar taki nie może ograniczać się do porównywania instrukcji wyszukiwawczej z każdą charakterystyką wyszukiwawczą dokumentu oddzielnie, ponieważ wówczas nie jest spełniony warunek addytywności. Wyszukiwanie informacji jest dynamicznym podprocesem komunikacyjnym, opisywanym za

pomocą procedur interrogacji. Relewanca stanowi wówczas miarę efektywności kontaktu procedury ze zbiorem informacyjnym. Probabilistyczny charakter relewancji jako miary powoduje, że nie jest ona związana z jednym podzbiorem informacji w systemie. Procedury interrogacyjne mogą wpływać na zmianę stopnia relewancji poszczególnych klas charakterystyk wyszukiwawczych dokumentów lub obiektów. Na podstawie twierdzeń prawdopodobieństwa statystycznego, określającego proces zmiany w czasie danej populacji, wskazano różne warunki, w których zbiór informacyjny ujawnia różne powiązania — relacja relewancji wyznacza różne podzbiory informacji zależnie od procedury interrogacyjnej. Zaproponowano wprowadzenie miary warunkowego prawdopodobieństwa relewancji zdeterminowanego przez aktualne związki w zbiorze informacji i procedurę interrogacyjną, dowodząc, iż „jakakolwiek miara efektywności dostarczanej informacji musi zależeć od tego, co dotychczas wiadomo i stąd konieczne jest wprowadzenie pojęcia warunkowego prawdopodobieństwa informacji” [Goffman W., 1964].

Ujęcie relewancji warunkowo prawdopodobnej, wielowartościowej i stopniowalnej, orzekanej i mierzony przy uwzględnieniu relacji między elementami zbioru informacji w danym czasie, jest pierwszą teorią łączącą aspekt interakcji w procesie komunikacji ze statycznym ujęciem relewancji w dystrybucji zbiorów informacji w nauce przez Bradforda. Stanowi kontrpropozycję wobec binarnej miary efektywności SIW. Traktowanie wyszukanej przez system informacji jako warunkowo relewantnej przy danej procedurze interrogacyjnej (instrukcji wyszukiwawczej) określa zasadę działania dialogowych systemów informacji, które omawiamy w rozdz. 4.2.1.

Wiele analogii z komunikacyjnym, interakcyjnym ujęciem relewancji warunkowo prawdopodobnej wykazuje uwzględniająca aspekt wyszukiwania *definicja Saracevica*: „Komunikowanie wiedzy jest efektywne wówczas, jeśli informacja przekazywana z jednego zbioru powoduje zmiany w innym zbiorze. Relewanca jest miarą tych zmian” [Saracevic, 1975, 325].

2.5.2. Relewanca jako relacja między zbiorami systemu informacyjno-wyszukiwawczego i jego otoczeniem

Jako typowy przykład interpretacji relewancji jako relacji między zbiorami systemu informacji a elementami otoczenia systemu, a więc między wyrażeniami języka i ich użytkownikiem, interpretacji, bardzo rozpowszechnionej w literaturze lat sześćdziesiątych, przytoczymy sformułowanie Cuadry i Kattera: „Relewanca jest zgodnością w kontekście między zapytaniem informacyjnym i artykułem, to jest zakresem w jakim artykuł obejmuje materiał odpowiedni dla zapytania informacyjnego” (...) „Relewanca jest terminem niewyjaśnionym” [Cuadra C. A., 1964].

Określenie Cuadry i Kattera, nie spełniające warunków poprawnej definicji jest symptomatyczne dla pragmatycznego, psychologicznego nurtu w dyskusjach nad relewancją, w ramach którego utożsamia się — jak wykazał Cooper — ocenę

relewancji z jej interpretacją. Spory nad ustaleniem czynników wpływających na orzeczenie o relewancji informacji wyszukanej przez system były szczególnie intensywne po opublikowaniu raportu z tzw. eksperymentu cranfieldskiego [Cleverdon C. W., Mils J., Keen E. M., 1966]. Niektórzy z autorów (m.in. Taube) kwestionowali w ogóle kryterium relewancji jako subiektywne i efemeryczne [Dooye L. B., 1966] [Taube M. A., 1965].

Za czynniki wpływające na orzeczenie o relewancji Cuadra i Katter uznają:

- 1) cechy badanego dokumentu: tematyka, stopień trudności, zwięzłości, styl,
- 2) warunki formułowania orzeczenia: czas, organizacja zbioru dokumentów, wielkość zbioru, sposób sformułowania zadania dla orzekającego,
- 3) wyrażenia określające wymagane informacje, które decydują o sposobie orzekania o relewancji,
- 4) cechy orzekającego: doświadczenie, wykształcenie, nastawienie [Cuadra C. A., Katter R. V., 1967].

Jak zauważa Salton, różnica zdań wśród orzekających o relewancji pojawiała się głównie w tzw. przypadkach granicznych, którym przypisywano „niski współczynnik relewancji”. W przypadku dużego zbioru charakterystyk wyszukiwawczych binarne decyzje o podziale na podzbiory informacji „relewantnej i nierelwantnej” budziły znacznie mniej kontrowersji [Salton G., 1975].

Próby ustalenia aplikacyjnych kryteriów orzekania o relewancji przedstawiane w formie przytoczonych na początku rozdziału hipotez wyliczających możliwe argumenty relacji między informacją wyszukaną przez system a elementami jego otoczenia, z których najważniejszy był użytkownik informacji, uwikłane były w problem zdefiniowania innego, równie nieprecyzyjnego terminu potrzeba informacyjna (*information need*). Praktyka serwisów informacyjnych zawsze wykazywała kłopoty z artikulacją potrzeb informacyjnych przez użytkowników. Włączenie aspektu psychologicznego i socjologicznego do orzekania o relewancji prowadziło do jej interpretacji jako relacji między informacją wydaną przez system a domniemanym stanem umysłu użytkownika uświadamiającego sobie potrzebę bycia poinformowanym, a więc definicji *ignotum per ignotum*.

Do wprowadzenia pewnego ładu interpretacyjnego i równoczesnego uznania możliwości istnienia różnych odmian relewancji przyczyniło się opublikowanie artykułu Coopera, wyraźnie odróżniającego relewancję informacji od jej użyteczności jako innego równoprawnego kryterium oceny, a ponadto inne prace, również wskazujące konieczność odróżnienia potrzeby informacyjnej — jako stanu umysłu — od językowego odwzorowania tego stanu [O'Connor J., 1967, 1968].

Wprowadzenie, już w latach sześćdziesiątych, kryterium „pertynencji” jako relacji wiążącej wyszukaną informację z psychologiczną potrzebą informacyjną, miało pewien walor porządkujący w dyskusji, przybliżało bowiem wyraźniejsze ujmowanie relewancji jako relacji między tekstami języka, jednak nie

wyjaśniało jeszcze kategorii użyteczności informacji istotnej w ustalaniu pertynencji.

„Relewancja jest własnością przyporządkowującą pewne elementy zbioru (informacji) np. dokumenty zapytaniu”... [Rees A. M., Saracevic T., 1963].

„Pertynencja jest własnością przyporządkowującą te elementy potrzebie informacyjnej” [Goffman W., Nevill V. A., 1966].

W propozycjach autorów próbujących uściślić, w celach aplikacyjnych, kryteria relewancji, a przede wszystkim pertynencji, niewyraźna kategoria potrzeba informacyjna zostaje zastąpiona przez zbiór wiedzy, którym to terminem określa się zbiór informacji stanowiący podstawę oceny przez użytkownika informacji wyszukanej przez system. Użycie terminu wiedza, nie zaś informacja nie jest przypadkowe, wyznacza bowiem dodatkowy wymiar opisu relacji wiążących zbiór informacji wskazany przez system jako relewantny w sferze relacji językowych z innymi zbiorami.

Foskett wprowadza opozycję wiedza społecznie akceptowana — wiedza jednostki (*public knowledge* — *private knowledge*). Włącza w ten sposób do sfery rozważań nad dziedzinami relacji sumaryczne zbiory informacji w danej dziedzinie nauki oraz algorytmny jej konceptualnego przetwarzania, będące wynikiem pewnego konsensusu specjalistów (np. paradygmat Kuhna), które tworzą ową wiedzę społecznie akceptowaną (*public knowledge*). Wiedzę jednostki tworzy pewien podzbiór wiedzy społecznej różniący się nie tylko wielkością z powodu ograniczenia pojemności pamięci człowieka, lecz także, i przede wszystkim, indywidualnym sposobem jej przetwarzania konceptualnego. „*Relewancja* jakiegokolwiek elementu wiedzy oznacza, że można go traktować jako spójny z ogólnym paradygmatem w szerszej dziedzinie, jako uznany i uznawalny element konsensusu między specjalistami w tej dziedzinie”. *Pertynencja* oznacza, że jest on zgodny ze szczególnym paradygmatem, jaki jednostka próbuje konstruować we własnym umyśle. Relewancja jest tym, co determinuje zawartość wszelkich serwisów informacyjnych; na przykład służby informacyjne selekcjonują materiały według pewnego konsensusu, nie zaś według kryteriów jednostkowych. *Pertynencja* natomiast jest tym, co determinuje dobór materiałów przez bibliotekarza lub pracownika służby informacyjnej dla konkretnego użytkownika. Dobór ten jest bardziej subiektywny, mniej zgodny z konsensem. Wkład jednostki do rozwoju wiedzy można oceniać w zakresie takim, w jakim udaje się jej transformować to, co dla niej pertynentne w to, co jest relewantne względem istniejącego konsensusu [Foskett D. J., 1972; 1981].

Ujęcie relewancji i pertynencji w wersji sformułowanej przez Fosketta, oparte na rozróżnieniu wiedzy akceptowanej społecznie i wiedzy jednostki rozwija Kemp, wskazując ich cechy dystynktywne niezbędne także do opisu znaczeń innych par terminów *public-ness* (bycie społecznym, wspólnym) i *private-ness* (bycie indywidualnym, jednostkowym). Przykładami takich par terminów z dziedzin pokrewnych są [Kemp D. A., 1974]:

Dziedzina	Spoleczne	Jednostkowe
Informacja naukowa	Relewanca	Pertynencja
Filozofia nauki	Wiedza społeczna	Wiedza jednostki
Psychologia	Denotacja	Konotacja
Lingwistyka	Semantyka	Pragmatyka
Socjologia wiedzy	Komunikacja formalna	Komunikacja nieformalna

W interpretacji relewancji jako relacji między odwzorowaniem w systemie informacyjnym zbioru wiedzy społecznie akceptowanej i zapytaniem informacyjnym, za cechę semantycznie prymarną uważa się „bycie przedmiotem treści” i ewentualnie relacje między przedmiotami treści (relewancji przedmiotowej — *topical relevance*). Z kolei w interpretacji pertynencji jako relacji między zbiorem wiedzy społecznie akceptowanej i zbiorem wiedzy jednostki (powstałej w wyniku selekcji i nawarstwiania się elementów wiedzy społecznej) istotne znaczenie ma wartościowanie/osądzanie (*judgement*) — relewanca interpretacyjna (*interpretational relevance*) — por. rozdz. 2.2.

Stosunkowo słabe rozpoznanie psychologicznych zasad wydawania przez człowieka sądów o użyteczności informacji było powodem nieadekwatności ujęć relewancji i pertynencji w kategoriach relacji między dwoma zbiorami wiedzy i jej oceną psychologiczną. Na nieuchwytność kryterium „osądzania” (*judgement*) zwrócił uwagę P. Wilson, formułując koncepcję tzw. relewancji sytuacyjnej, w której do określenia relacji między odpowiedzią systemu informacyjnego a zbiorem wiedzy użytkownika włączone zostaje pozasystemowe kryterium użyteczności informacji wskazane przez Coopera [Wilson P., 1973, 457-471].

Punktem wyjścia była tu definicja relewancji logicznej Coopera, wobec której Wilson wysunął zarzut, iż jest zbyt wąska, bo uwzględnia tylko jeden z rodzajów relewancji logicznie związanej z systemami dedukcyjnymi. Zarzut ten wydaje się niezupełnie słuszny, ponieważ, przedstawiając definicję relewancji opartą na implikacji dedukcyjnej, Cooper zaznaczał, że jest to próba sformułowania mocnej definicji dla systemów faktograficznych. Równocześnie, ukazując różnice w probabilistycznie zorientowanym wyszukiwaniu informacji dokumentacyjnej, podkreślił nieadekwatność obecnego stanu teorii indukcji wobec potrzeby mocnej, indukcyjnej definicji relewancji. Twierdził równocześnie, że wnioskowanie probabilistyczne można interpretować dedukcyjnie, opisując relację implikacji jako stopień prawdopodobieństwa, dzięki czemu możliwe staje się wprowadzenie stopni relewancji.

Poglądy Wilsona na koncepcję relewancji sytuacyjnej dotyczą dwu aspektów — jej podstaw logicznych oraz użyteczności. Polemizując z Cooperem w kwestii reguł wnioskowania, na jakich opiera się wyszukiwanie informacji Wilson twierdzi, że procesy inferencyjne człowieka mają charakter niededukcyjny (non-deductive, non-demonstrative). Nawet przy braku dobrej ogólnej teorii

indukcji, najmocniejsze ugruntowanie znajduje relewancja w logice konfirmacji, której przedmiotem są relacje między przesłankami i wnioskiem w sytuacji, gdy przesłanka podtrzymuje wniosek, lecz nie zachodzi między nimi relacja *entailment* (por. rozdz. 2.3.1. — poglądy Carnapa dotyczące funkcji konfirmacji). Tę odmianę relewancji logicznej nazywa Wilson relewancją ewidencyjną (*evidential relevance*), twierdząc równocześnie, że jej wyjaśnienie jest możliwe przez odwołanie się do prawdopodobieństwa (empirycznego potwierdzenia) wniosku w stosunku do przesłanek. Wilson przytacza definicję relewancji Keynesa: „element informacji *I* jest relewantny dla wniosku (*h*) ze względu na przesłanki (*e*) jeśli stopień konfirmacji /prawdopodobieństwa/ (*h*) przy świadectwie (*e*) oraz *I* jest większy lub mniejszy niż stopień konfirmacji (prawdopodobieństwa) przy samym (*e*). [Keynes J. M., 1921]. Za pojęcie centralne w teorii systemów informacyjno-wyszukiwawczych uważa relację „podtrzymywania/poświadczenia” (*support*) jako wyraźną w stopniu nie mniejszym niż konfirmacja wniosków przy danych przesłankach.

Definiując relewancję sytuacyjną (i jej odmiany) Wilson wyróżnił te elementy systemu informacyjnego i jego otoczenia, które są — jego zdaniem — niezbędne, aby sprecyzować kryterium użyteczności informacji i równocześnie wyeliminować nieostre kryteria psychologiczne:

1) zainteresowanie użytkownika (*concern*) wyznaczane przez zbiór atrybutów sytuacji, w jakiej się znajduje. Zbiór ten nie jest określany obiektywnie lecz subiektywnie;

2) preferencje dotyczące alternatyw związanych z zainteresowaniami (zbiorem atrybutów sytuacji). Nie zakłada się hierarchizacji preferencji tak jak w teoriach ekonomicznych, przyjmując że mają charakter „holistyczny” nie zaś „atomistyczny”;

3) zbiór wiedzy użytkownika, przy czym termin wiedza traktowany jest jako synonim informacji, bez uwzględniania kategorii psychologicznej wiarygodności (*belief*). W zbiorze tym wyróżnione są podzbiory wiedzy ogólnej i wiedzy sytuacyjnej. Wiedzę sytuacyjną wyznaczają atrybuty sytuacji w danym czasie;

4) zapytanie informacyjne odwzorowujące pewne atrybuty sytuacji/zainteresowania użytkownika (*concern question*);

5) zbiór możliwych odpowiedzi na zapytanie (*concern set statements*).

Ponadto autor ten przyjął następujące założenia:

- istnieje wiele rodzajów relewancji;
- należy odróżniać relewancję psychologiczną, określającą efekty wykorzystania informacji, od relewancji logicznej (*evidential*), będącej relacją między elementem informacji, indywidualnym odwzorowaniem rzeczywistości (zbiorem wiedzy użytkownika) oraz sytuacją użytkownika w tej rzeczywistości;
- zbiór wiedzy użytkownika pozostaje w pewnym związku ze zbiorem potencjalnych odpowiedzi,

- zdania tworzące zbiór potencjalnych odpowiedzi powinny spełniać warunek rozłączności i adekwatności, tak aby dokładnie jeden element zbioru odpowiedzi był elementem opisu sytuacji; spełnienie tego warunku powinien zapewniać poprawnie skonstruowany język (*well-formed*) — por. [Cooper W. S., 1971] [Carnap R. 1962].
- warunkiem relewancji sytuacyjnej jest relewancja logiczna.

Oznaczając I – zbiór wiedzy użytkownika w danym czasie, Wilson stwierdza:

„Element I jest relewantny sytuacyjnie jeśli łącznie z innymi elementami zbioru I jest relewantny logicznie dla pewnego zapytania (*concern question*), to jest jeśli jest elementem minimalnego zbioru przesłanek w I , z którego wynika logicznie jedna z możliwych odpowiedzi na zapytanie” [Wilson P., 1973, 465];

„Element I jest relewantny ewidencyjnie dla pytania jeśli istnieje możliwa odpowiedź, której wiarygodność (*credibility*) zostanie zmniejszona lub zwiększona po jego rozważeniu, to znaczy jeśli istnieje pewien podzbiór elementów I nie zawierający I , którego użycie z I spowoduje, że wiarygodność odpowiedzi będzie mniejsza lub większa niż przy użyciu jego samego (I). Jeśli potencjalne odpowiedzi na zapytanie tworzą rodzinę wyczerpujących i rozłącznych hipotez, to nowa informacja (wyrażenie) będzie relewantna ewidencyjnie, jeśli spowoduje zmiany w rozkładzie prawdopodobieństwa wśród elementów tej rodziny, przy czym nie można zakładać z góry warunku rozłączności i adekwatności.

Relewancja sytuacyjna jest bezpośrednia (*direct*) wówczas, gdy element zbioru wiedzy użytkownika (I) jest elementem zbioru atrybutów wyznaczających zainteresowania, natomiast pośrednia (*indirect*), gdy element zbioru I nie jest elementem zbioru zainteresowań. Jako warunek relewancji sytuacyjnej wymieniana jest psychologiczna akceptacja relewancji logicznej, zdeterminowana przez zbiór preferencji. Na tej podstawie można — jak twierdzi Wilson — wytłumaczyć fakt, że informacja relewantna logicznie w danej sytuacji jest kwalifikowana jako „interesująca — nieinteresująca, nudna”. Preferencje wpływają na motywację, chęć dowiedzenia się czegoś.

Inną odmianą relewancji sytuacyjnej wyróżnionej przez Wilsona jest tzw. relewancja praktyczna, zdeterminowana przez cele i plany działania (*concern, goals*). Interpretacje i odmiany relewancji scharakteryzowane przez tego autora wykazują wiele podobieństw z ujęciem Schutza w filozofii (por. rozdz. 2.2.).

Do autorów podejmujących próbę wyjaśnienia zasady funkcjonowania kryterium użyteczności informacji w systemie należy M.Kochen, który przedstawił sformalizowany opis tzw. funkcji użyteczności i informacji (*utility function*). Punktem wyjścia do zdefiniowania „użyteczności” na podstawie pewnych preferencji było przyjęcie za Cooperem rozróżnienia użyteczności i relewancji. Relewancję umieścił on w sferze relacji językowych (*relation between propositions*), natomiast użyteczność w sferze relacji pozajęzykowych między informacją a systemem preferencji, których zmiana powoduje zmiany funkcji

użyteczności. Definicja funkcji użyteczności u została sformułowana na podstawie czterech aksjomatów teorii użyteczności²⁰.

A1. Dla każdego z dwu dokumentów d i d' użytkownik U powinien być zdolny jednoznacznie stwierdzić, który z nich preferuje lub zajmuje wobec niego postawę indyferentną.

A2. Jeśli użytkownik U preferuje d zamiast d' oraz d' zamiast d'' , zatem preferuje d zamiast d'' dla każdego d, d', d'' .

A3. Jeśli użytkownik U preferuje d zamiast d' oraz d' zamiast d'' , to potrafi określić prawdopodobieństwo P takie, że dla wszystkich d, d', d'' zajmuje postawę indyferentną między opcją pewności dla d' a opcją polegającą na zastosowaniu metody, za pomocą której wyszuka d z prawdopodobieństwem P oraz d'' z prawdopodobieństwem $1-P$.

A4. Jeśli U zajmuje postawę indyferentną wobec d i d^* , zaś d^* zostaje zastąpione przez d w pewnej strategii wyszukiwawczej (d z prawdopodobieństwem P lub d'' z prawdopodobieństwem $(1-P)$, tak jak w A3.), wówczas U zajmuje postawę indyferentną wobec tej strategii oraz d^* z prawdopodobieństwem P lub d'' z prawdopodobieństwem $(1-P)$.

Funkcja użyteczności $u(d; U, Q)$ wyznacza dla każdego dokumentu (d) wartość przy danym U i Q tak, że jeśli U preferuje d zamiast d' , to $u(d)$ przy danym $u(d')$, a spodziewana użyteczność strategii jest spodziewaną użytecznością dokumentów w tej strategii.

Zasadniczą wadą definicji funkcji użyteczności jest odwoływanie się do niejasnego kryterium „jest zdolny”, co przyznaje pośrednio sam autor zaznaczając, że zdobywanie wiedzy przez użytkownika może zmienić porządek preferencji, a zatem wartość funkcji u jest pośrednio zdeterminowana przez zmiany w organizacji zbioru informacji (wiedzy) użytkownika.

2.6. ZASADY RELEWANCJI W SYSTEMIE INFORMACJI NAUKOWEJ

Cechą wspólną omówionych prób zdefiniowania relewancji jest traktowanie jej jako relacji, której jednym z argumentów są wyrażenia pewnego języka, w którym zakodowana jest informacja (por. tab.1 i 2). Jednakże tylko nieliczni spośród cytowanych autorów konsekwentnie odróżniają reprezentację informacji w języku naturalnym od metainformacji utrwalonej w języku informacyjno-wyszukiwawczym. [Cooper W. S., 1971] [Goffman W., 1964] [Kochen M., 1974].

Dla większości definicji symptomatyczna jest nieostra interpretacja terminu „dokument” polegająca na utożsamianiu charakterystyki wyszukiwawczej z dokumentem pierwotnym, a także zapytania informacyjnego z instrukcją

²⁰ Skwantyfikowaną definicję użyteczności wprowadzili do teorii gier J. Neuman i O. Morgenstern. Zob. Pszczołowski T., *Mała encyklopedia prakseologii i teorii organizacji*, Warszawa 1978, Zakład Narodowy im. Ossolińskich.

wyszukiwawczą odwzorowującą to zapytanie. Ma to miejsce także wówczas, gdy zakłada się konsekwentne umieszczanie relewancji w sferze relacji językowych, zachodzących między wyrażeniami JIW należącymi do zbioru informacyjnego SIW i instrukcji wyszukiwawczej i odróżnia tę dziedzinę relacji od relacji między wyrażeniami języka informacyjnego (języka naturalnego) i ich użytkownikiem w definicjach pertynencji (relewancji dla użytkownika/relewancji sytuacyjnej). Stanowi to źródło trudności przy ustalaniu odwzorowywanych w języku informacyjnym cech desygnatów jego wyrażen, a więc określaniu reguł semantyki wyznaczających kryteria relewancji w SIW. Konieczność interpretacji relewancji na gruncie poprawnie sformułowanego języka (*well-formed*) podkreślali Carnap i Cooper.

Tabela 1

Relewancja jako relacja między elementami systemu informacyjno-wyszukiwawczego

Autor	Pierwszy argument relacji	Drugi argument relacji	Skala oceny
Bradford	Teksty dokumentów pierwotnych	Językowe wykładniki cechy <i>bycia przedmiotem treści</i>	Wielostopniowa
Price, Urquart i in.	Teksty dokumentów pierwotnych	Językowe wykładniki cechy <i>bycia elementem opisu bibliograficznego</i>	Wielostopniowa
Vickery	Zbiory dokumentów SIW	Hierarchia wyrażen JIW	Wielostopniowa
Hillman	Zbiory dokumentów SIW	Relacje semantyczne między wyrażeniami JIW odwzorowującymi sądy użytkowników	Wielostopniowa
Cooper	Zbiór informacyjny SIW (zdania języka predykatów)	Zbiór informacyjny SIW (zdania języka predykatów)	Wielostopniowa (Relacja implikacji)
Maron i Kuhn	Zbiór ChWD	Treść zapytania; Instrukcja wyszukiwawcza	Binarna
Goffman i Nevill	Podzbiory zbioru informacji SIW	Procedury interrogacyjne	Funkcja konfirmacji
Foskett	Zbiór informacyjny SIW (wiedza społecznie akceptowana)	Odpowiedź na zapytanie wydana przez SIW	Wielostopniowa
Bookstein	Podzbiór Informacji wydany przez system	Instrukcja wyszukiwawcza	Funkcja
Kochen	Zdania zbioru informacyjnego SIW	Zdania JIW	Wielostopniowa

Kolejnym zagadnieniem rozważanym w dyskusjach nad definicjami i kryteriami relewancji, skłaniającym do refleksji metodologicznej, jest używanie nieprecyzyjnego terminu *potrzeba informacyjna*, zastąpionego w nowszych pracach

Relacje między systemem informacyjno-wyszukiwawczym i jego otoczeniem

Autor	Nazwa relacji	Pierwszy argument relacji	Drugi argument relacji	Trzeci argument relacji
Cuadra i Katter	Relevancja	Dokument	Potrzeba informacyjna	Zapytanie
Rees, Saracevic	Pertynencja	Elementy zbioru informacyjnego SIW	Potrzeba informacyjna	
Bookstein	Relevancja dla użytkownika	Informacja wydana przez SIW	Potrzeba informacyjna	
Foskett	Relevancja	Informacja wydana przez SIW	Wiedza społecznie akceptowana i paradygmat jej przetwarzania	
	Pertynencja	Informacja wydana przez SIW	Wiedza użytkownika i indywidualny paradygmat jej przetwarzania	
Kemp	Relevancja	Informacja wydana przez SIW	Wiedza społeczna	
	Pertynencja	Informacja wydana przez SIW	Wiedza użytkownika	
Wilson	Relevancja sytuacyjna - pośrednia - bezpośrednia - praktyczna	Informacja wydana przez SIW	Wiedza użytkownika	Zbiór atrybutów sytuacji (zainteresowania użytkownika) Zbiór atrybutów sytuacji oraz cele działania
	Relevancja ewidencyjna (logiczna)	Informacja wydana przez SIW	Wiedza użytkownika	Zbiór preferencji
Kochen	Funkcja użyteczności	Informacja wydana przez SIW (zbiór zdań)	Wiedza użytkownika	System preferencji
Saracevic	Relevancja	Zbiór informacyjny SIW	Wiedza użytkownika	

określeniami wiedza użytkownika, wiedza jednostki (*private knowledge*), przeciwstawianymi wiedzy ogólnej, społecznie akceptowanej (*public knowledge*). Zmiana terminologii wydaje się wnosić porządek pozorny, ponieważ zarówno potrzeba jak i wiedza jednostki denotują pewien stan mózgu użytkownika wraz z informacją zakodowaną w pewnym języku, przypuszczalnie innym niż język naturalny — *lingua mentalis*²⁰.

Jej wykorzystanie w procesie komunikacji użytkownika z SIW zawsze oznacza konieczność transformacji (przekładu) z „wewnętrznego” *lingua mentalis* na „zewnątrzny” język naturalny, tego zaś na właściwy język sztuczny — JIW lub inny język właściwy danej dyscyplinie naukowej (dziedzinie wiedzy).

Tak więc, uwzględnianie kolejnych niezbędnych odwzorowań wiedzy w językach odpowiadających różnym poziomom przetwarzania informacji przez człowieka w procesach komunikacji uznamy za *pierwsze założenie niezbędne do poprawnego zdefiniowania relewancji*. Założenie to pozwala — jak sądzimy — wyjaśnić na płaszczyźnie językowej rozróżnienie *relewancji i pertynencji* (relewancji dla użytkownika, relewancji sytuacyjnej) oraz wytłumaczyć rozbieżności w ich ocenach.

Założenie drugie przyjmujemy za Carnapem i Cooperem, uznając za podstawę definicji relewancji w SIW relewancję logiczną, z uwzględnieniem wskazanych wyżej różnic w procesach inferencyjnych, właściwych procesom wyszukiwania informacji w systemach faktograficznych (dedukcyjnych) i dokumentacyjnych (probabilistycznych).

Założenie trzecie dotyczy przyjęcia planu treści języka (jego semantyki), stanowiącego przedmiot naszych rozważań jako dziedziny relacji relewancji. Uznajemy przy tym plan wyrażania za wtórny wobec planu treści, odwzorowujący powierzchwniowo relacje semantyczne. Założenie to pozwala wyjaśnić znaczenie terminów *relewancja techniczna, relewancja inżynierska, relewancja formalna* jako relacji wyznaczającej w zbiorze informacyjnym SIW taki podzbiór informacji, który spełnia warunek minimalnego podobieństwa (ustalony dla danego SIW) z instrukcją wyszukiwawczą w planie wyrażania JIW. Ustalanie podobieństwa planów wyrażania tekstów JIW, tj. instrukcji wyszukiwawczej i charakterystyk wyszukiwawczych jest elementem procesu wyszukiwania informacji relewantnej semantycznie.

Założenie czwarte nawiązuje do zasad eksplikacji semantycznej terminów języka nauki sformułowanych przez Carnapa, których przestrzeganie warunkuje poprawność procesów inferencyjnych, a więc także procesów wyszukiwania informacji.

Spośród scharakteryzowanych wyżej czterech zasad jedynie zasada podobieństwa klas została wykorzystana w definiowaniu relewancji przez Hillmana, który przyjął równocześnie ekstensjonalną interpretację znaczenia. Rozważania Hillmana, poparte dowodem matematycznym, doprowadziły do stwierdzenia ko-

²⁰ Zagadnienie sposobu odwzorowania informacji konceptualnej w mózgu człowieka w świetle wyników badań kognitywnych poruszam w rozdz. 3.4.1.

nieczności intensjonalnego interpretowania znaczenia przy definiowaniu relewancji w SIW, łączonej z cechą określaną terminem *relatedness* (powiązania znaczeniowego). Uznając komplementarność zasad Carnapa (podobieństwa, dokładności, produktywności i prostoty), przyjmujemy, że relewancję informacji w systemie wyznacza sieć relacji odwzorowanych w polu semantycznym JIW (paradygmatycznych i syntagmatycznych), reprezentujących związki między denotowanymi przez wyrażenia języka obiektami rzeczywistości dokumentacyjnej i/lub pozadokumentacyjnej. Relacje te, tworzące według terminologii Carnapa definicje korelacyjne wyrażen językowych (korelacje epistemiczne), decydują o kryteriach relewancji w systemie informacyjnym.

O precyzji kryteriów relewancji w danym JIW decyduje przede wszystkim *zasada produktywności (fruitfulness)*, zgodnie z którą stopień prawdopodobieństwa logicznego wniosków formułowanych na podstawie pewnych hipotez, przy danym korpusie dowodowym, pozostaje w bezpośredniej zależności od skorelowania znaczeń wyrażen stosowanego języka. Zasady podobieństwa i dokładności należałoby uznać za podporządkowane zasadzie produktywności. Sposób semantycznego skorelowania wyrażen JIW wyznacza zasady reprezentacji wiedzy w systemie informacyjno-wyszukiwawczym

3. REPREZENTACJA WIEDZY W SYSTEMIE INFORMACYJNO-WYSZUKIWAWCZYM

3.1. WIEDZA REPREZENTOWANA W STRUKTURZE JĘZYKA INFORMACYJNO-WYSZUKIWAWCZEGO

Przyznanie zasadzie relewancji nadrzędnej roli w procesie wyszukiwania informacji sprawiło, że projektanci systemów informacyjno-wyszukiwawczych, a w szczególności języków informacyjno-wyszukiwawczych, dążyli do zapewnienia takiej struktury systemu, aby możliwe było wyszukanie informacji nie tylko możliwie kompletnej, ale także relewantnej dla użytkownika wiedzy o obiektach opisywanych w zbiorach informacji. Podstawę oceny użyteczności informacji dostarczanej użytkownikowi tworzy — jak wskazywaliśmy w poprzednim rozdziale — sieć relacji semantycznych JIW, wyznaczających kryteria relewancji, co wykazali Carnap, a następnie Cooper. W nowszej literaturze informatycznej, poświęconej budowie systemów informacji z dostępem w języku naturalnym, problem odwzorowania w zbiorach informacyjnych systemu relewantnej dla użytkownika tego systemu informacji określany jest jako problem *reprezentacji wiedzy*. Termin reprezentacja wiedzy definiowany jest różnie, jako:

- „... formalizm zapisu informacji posiadanych przez system wraz ze sposobami odwzorowania”. [Bolc L., Cichy M., Różańska L., 1982];
- „... sposób odwzorowania fragmentu świata w określonym czasie oraz rodzaje transformacji dokonywane przez system (wyszukiwanie informacji, wnioskowanie)” [Mc Dermott D., 1978, 206-221].

Według McDermotta konieczne jest rozróżnienie reprezentacji wiedzy w programie komputerowym od modelu tej reprezentacji, określającego sposób odwzorowania konkretnej wiedzy oraz rodzaje operacji dokonywanych na reprezentacjach zgodnych z pewnym modelem. Elementami obligatoryjnymi reprezentacji rozumianej jako model są notacja, semantyka notacji dla określonej dziedziny wiedzy oraz algorytmy wyszukiwania i aktualizacji danych wraz z regułami wnioskowania [Mc Dermott D., 1978, 206-221].

Powyższe definicje, jakkolwiek sformułowane dla pewnej klasy SIW, są dostatecznie ogólne, aby stwierdzić (przy pewnych niezbędnych uściśleniach), że problem reprezentacji wiedzy w SIW sprowadza się przede wszystkim do struktury języka, a zwłaszcza struktury jego pola semantycznego, odwzorowanej

w planie wyrażania, a następnie do transformacji dokonywanych na jego tekstach. W systemach informacyjnych z dostępem w języku naturalnym ekspozowany jest problem transformacji wyrażen niezbędnych do „rozumienia” przez system komunikatów użytkownika i generowania odpowiedzi. Należy jednak zauważyć, że wszelka informacja utrwalona na pewnym nośniku (w języku naturalnym lub sztucznym) stanowi odwzorowanie wiedzy. Reprezentacją wiedzy są więc zarówno nowoczesne systemy dialogowe, jak i tradycyjne katalogi biblioteczne, serwisy bibliograficzne i faktograficzne [Sosińska B., 1985, 19-36].

Wadą przytoczonych definicji jest to, iż nie pozwalają wyraźnie wskazać tych elementów języka w systemie, które są wykładnikami cech relewantnych, a więc wystarczających do rozwiązania zadania wyszukania pożądanej informacji lub jej wygenerowania. Elementy te najczęściej wymieniane są w ciągu, nie uwzględniającym hierarchii ich ważności (na przykład, McDermott wymienia na pierwszym miejscu notację stosowaną do opisu modelu reprezentacji) [Mc Dermott D., 1978, 206-221]. W omówionych definicjach relewancji logicznej podkreślono znaczenie kryteriów ustalania wyrażen elementarnych języka (zdań rachunku predykatów pierwszego rzędu) oraz określenia relacji (drzewa zależności) zachodzących między zdaniami składowymi zapytania informacyjnego, warunkujących wyselekcjonowanie ze zbiorów systemu tekstów odpowiadających formule implikacyjnej *entailment*, świadczącej o istnieniu związku znaczeniowego między zapytaniem informacyjnym a wyszukany wyrażeniem (por. zasada produktywności Carnapa — rozdz. 2.3.1.).

W związku z tym na plan pierwszy w opisie struktury JIW wysuwa się semantyka, a więc reguły przyporządkowania wyrażeniom tego języka elementów opisywanej przezeń rzeczywistości. Notacja odwzorowująca przyjęte reguły semantyki ma w naszych rozważaniach znaczenie drugoplanowe.

Jako adekwatną podstawę opisu struktury JIW w planach treści i wyrażania przyjmujemy dla potrzeb dalszych rozważań ujęcie B. Sosińskiej [Sosińska B., 1987]. Autorka ta przyjęła następujące założenia teoretyczne:

- za punkt wyjścia posłużyła dualna definicja znaku F. de Saussure’a (rozdzielenie „signifié — signifiant”) [Saussure de F., 1971];
- na podstawie teorii glossematycznej L. Hjelmsleva wyróżniono w strukturze języka dwa wzajemnie zależne plany: treści i wyrażania, ze wskazaniem w każdym z nich poziomów substancji i formy. W teorii tej istotne było założenie o funkcjonowaniu na poziomie substancji planu treści elementarnych cech semantycznych (glossemów), tworzących znaczenie na mocy reguł semantycznych poszczególnych języków [Hjelmslev L., 1979];
- w opisie kategorii planów JIW wykorzystano teorię budowy klasyfikacji S. R. Ranganathana, wykazującą pewien paralelizm z teoriami F. de Saussure’a oraz Hjelmsleva. Wyróżnione przez Ranganathana trzy plany językowe (plan treści dokumentów, plan werbalny oraz plan notacyjny JIW) świadczą o zasadności uzależniania kategoryzacji planu werbalnego (w uproszczeniu planu

wyrażania języka naturalnego) od kategoryzacji jednostek planu treści JIW, nie zaś odwrotnie [Ungurian O., 1975];

- interpretacja znaczenia w JIW (denotacyjna i konotacyjna) nawiązuje do definicji treści i zakresu znaków języka K. Ajdukiewicza oraz teorii ekstensji i intensji wyrażen R. Carnapa [Ajdukiewicz K., 1985, 102-136] [Carnap R., 1956];
- w charakterystyce funkcji metainformacyjnej JIW decydującej o interpretacji znaczenia odwołano się do definicji metainformacji H. Greniewskiego oraz do definicji językowej funkcji metainformacyjnej B. Bojar [Bojar B., 1979].

W rezultacie rozważań teoretycznych i krytycznej analizy koncepcji Ranganathana w opisie B. Sosińskiej wyróżniono cztery poziomy znaków JIW:

1) poziom figur treści (elementarnych cech semantycznych) i poziom figur wyrażania (elementarnych figur graficznych),

2) poziomy znaczenia i wyrażania elementarnych jednostek leksykalnych,

3) poziomy znaczenia i wyrażania zdań JIW,

4) poziomy znaczenia i wyrażania tekstów JIW.

Kolejnym ustaleniem metodologicznym w pracy B. Sosińskiej, do którego odwołamy się obecnie, jest *interpretacja znaczenia* w JIW. Jest to zagadnienie istotne z różnych względów. Z opisu funkcji metainformacyjnej JIW wynika, że wyrażenia JIW desygnują równocześnie dwa typy obiektów: bezpośrednio — dokumenty poprzez opis ich treści i/lub formy oraz pośrednio — obiekty rzeczywistości pozadokumentacyjnej, informacja o których składa się na treść dokumentów. Rozróżnienie desygnatów pośrednich i bezpośrednich jest podstawą podziału informacji na dokumentacyjną i faktograficzną (i odpowiednio systemów informacji) — por. rozdz. 1. Konkluzje sformułowane przez B. Sosińską w wyniku obszernie egzemplifikowanych rozważań na temat interpretacji znaczenia wyrażen JIW można ująć następująco:

1) Istnieje ścisła korelacja między bezpośrednimi i pośrednimi desygnatami wyrażen JIW, polegająca na konieczności odwoływania się do rzeczywistości językowej (dokumentacyjnej) w celu uzyskania informacji o obiekcie rzeczywistości pozajęzykowej za pośrednictwem SIW. Z kolei możliwość informowania przez system o obiektach rzeczywistości pozadokumentacyjnej jest uzależniona od uprzedniego odwzorowania w jego zbiorach pewnej wiedzy językowej i pozajęzykowej o tym obiekcie, poczynając od kodowania jego nazwy po uporządkowanie zbiorów informacyjnych według określonych kryteriów (treściowych lub formalnych), tj. dystynktywnych cech semantycznych wyznaczających kryteria relewancji.

2) Projektanci JIW odwołują się do denotacji bądź konotacji wyrażen, zależnie od potrzeby odwzorowania w języku określonego rodzaju wiedzy:

a) do denotacji bezpośredniej nawiązują najczęściej reguły syntaktyczne JIW systemów dokumentacyjnych jako oparte na operacjach teoriomnogościowych sumy i iloczynu określonych na zbiorach desygnatów bezpośrednich (zbiorach dokumentów);

b) denotacja pośrednia jest wykorzystywana w pewnym zakresie przy budowie systemów leksykalnych JIW, tj. do określenia relacji mereologicznych oraz części relacji kojarzeniowych, z wyjątkiem relacji kojarzeniowych opartych na współwystępowaniu wyrażen w tekstach.

3) Denotacyjna i konotacyjna interpretacja znaczenia mogą stanowić podstawę klasyfikacji JIW ze względu na semantyczne umotywowanie (lub nieumotywowanie) znaków językowych. Języki, których wyrażenia są semantycznie motywowane, odwzorowują w polu semantycznym cechy dystynktywne desygnatów pośrednich (obiektów rzeczywistości pozadokumentacyjnej) oraz relacje między tymi desygnatami. U podstaw organizacji pola semantycznego leży więc konotacyjne rozumienie znaczenia (zbiór cech desygnatów). Reprezentantami tej klasy są kody semantyczne, klasyfikacje fasetowe i pewne typy klasyfikacji monohierarchicznych. Języki, których wyrażenia są semantycznie arbitralne (niemotywowane) odwołują się do relacji między desygnatami bezpośrednimi (zbiorami dokumentów), np. języki deskryptorowe, języki hasel przedmiotowych, pewne typy klasyfikacji. Organizacja pól semantycznych tych języków opiera się na denotacyjnej interpretacji znaczenia. Relacje językowe, odwzorowywane w obu typach języków, związane są ze strukturą języka nie zaś z desygnatami wyrażen, pozostają poza problemem referencji znaków językowych.

4) O miejscu wyrażenia w strukturze JIW, a więc o relacjach, jakie tworzy z innymi wyrażeniami, decyduje konotacja bezpośrednia i/lub pośrednia, czyli zespół cech semantycznie niezbędnych i wystarczających (tzw. konotacja charakterystyczna) do identyfikacji desygnatów wyrażen wynikających z postulatów znaczeniowych danego języka. Jest to szczególnie istotne przy ustalaniu konotacji jednostek leksykalnych typu predykatów pierwszego stopnia.

5) Denotacja bezpośrednia i pośrednia oraz odpowiednio konotacja, zdefiniowane przez B. Sosińską w sensie ekstensji i intensji znaków językowych Carnapa, mają charakter uniwersalny, ponieważ są niezależne od istnienia dokumentów i obiektów rzeczywistości pozadokumentacyjnej, spełniających warunki określone przez wyrażenia JIW.

Przytoczone konkluzje pozwalają sformułować kolejne założenia związane z przedmiotem naszych rozważań:

- dla określenia rodzaju wiedzy odwzorowanej w JIW ważne znaczenie ma przyjęcie denotacyjnej lub konotacyjnej koncepcji znaczenia jako podstawy organizacji pola semantycznego, niezależnie od tego, czy dany język jest stosowany w systemie informacji dokumentacyjnej czy faktograficznej;
- wszystkie wyrażenia JIW mają charakter predykatów. Jednoargumentowe predykaty pierwszego stopnia (tzw. leksyka podstawowa) wskazują obiekty rzeczywistości dokumentacyjnej lub pozadokumentacyjnej; ich konotację (intensję) tworzą zespoły własności desygnowanych obiektów. Jednoargumentowe predykaty drugiego stopnia desygnują pewne wyrażenia języka (leksykę podstawową); ich konotację tworzą własności tych wyrażen w systemie języka.

Dwuargumentowe predykaty drugiego stopnia desygnują (podobnie jak predykaty jednoargumentowe) wyrażenia języka; na ich konotację składają się związki między tymi wyrażeniami w tekstach JIW. B. Sosińska zwróciła uwagę na heterogeniczny charakter denotacji niektórych jednostek leksykalnych JIW, określających relacje lub działania (np. WOJNA), dla których liczba argumentów jest zmienna, zależnie od dziedziny denotacji: w odniesieniu do rzeczywistości dokumentacyjnej są jednoargumentowe, w odniesieniu do rzeczywistości pozadokumentacyjnej są wieloargumentowe. Wydaje się, że tę rozbieżność denotacji w przypadku systemów faktograficznych można by uznać za pozorną, przyjmując, że na etapie budowy JIW desygnatem bezpośrednim wyrażenia tego języka jest wyrażenie języka naturalnego (pozostaje więc predykatem jednoargumentowym pierwszego stopnia), natomiast desygnatami pośrednimi są uporządkowane n -tki obiektów rzeczywistości pozajęzykowej, pozostające w określonych relacjach lub stanie;

- o zakresie wiedzy odwzorowanej w polu semantycznym JIW decyduje rozległość sieci relacji między desygnatami bezpośrednimi i/lub pośrednimi wyrażen, a więc definicje korelacyjne według Carnapa. Miejsce wyrażenia JIW (elementarnej jednostki leksykalnej) stanowi równocześnie o jego epistemicznej korelacji z innymi wyrażeniami. W ten sposób realizowana jest cecha kompleksowości komunikowania znaczenia w tzw. podstawowej leksyce JIW.

3.1.1. Relacje paradygmatyczne jako wykładnik językowych kryteriów relewancji

Związki semantyczne zachodzące między jednostkami systemu JIW w planie treści, odwzorowywane w planie wyrażenia, są komunikowane za pomocą wykładników:

- a) relacji paradygmatycznych w słownikach JIW,
- b) kategorii semantycznych [Sosińska B., 1987].

Relacje paradygmatyczne, wykazywane w słownikach JIW jako pewne kompleksy podstawowych relacji znanych z teorii mnogości, sprowadzane są do pięciu rodzajów związków:

- 1) równoważności (synonimii),
- 2) nadrzędności (hiperonimii),
- 3) podrzędności (hiponimii),
- 4) krzyżowania,

5) wykluczania, w tym przeciwstawności (antonimii), [Rasiowa H., 1975] [Szejder J. A., 1975].

Analiza relacji paradygmatycznych, wyznaczających porządek pionowy JIW, doczekała się obszernej literatury jako zagadnienie o znaczeniu podstawowym w projektowaniu języka informacyjno-wyszukiwawczego i określaniu dla systemu kryteriów relewancji semantycznej [Wereszczyńska-Cisło B., 1984]. Wszechstronną specyfikację relacji paradygmatycznych w JIW przeprowadziła Sosińska, definiując je jako związki organizujące system językowy na każdym

jego poziomie. Wyróżniając pięć grup relacji: synonimii, przeciwstawności znaczenia, hierarchiczne, współrzędności oraz asocjacyjne, scharakteryzowała:

• Jako odmiany relacji synonimii:

— odpowiedniość wyszukiwawczą, definiowaną jako wszelkie związki semantyczne z wyjątkiem całkowitego wykluczenia, zachodzące między językiem naturalnym i JIW, na poziomie paradygmatyki i syntaktyki;

— synonimię wyszukiwawczą, definiowaną jako relację językową między wyrażeniami JIW, spełniającymi warunek identyczności denotacji i konotacji bezpośredniej i pośredniej;

• Jako odmiany relacji przeciwstawności znaczenia:

— antonimię, definiowaną jako nie pełniący funkcji porządkującej w leksyce związek między wyrażeniami JIW, których cechą charakterystyczną jest to, iż stwierdzenie jednego z nich implikuje zaprzeczenie drugiego, jakkolwiek zaprzeczenie jednego z nich nie implikuje stwierdzenia drugiego;

— komplementarność, zdefiniowaną jako związek między parami wyrażeń, których jeden człon jest tożsamy znaczeniowo z zaprzeczeniem członu drugiego, przy czym pary takie tworzą wyrażenia oznaczające własności stopniowalne na skali binarnej (*ludzie chorzy* — *ludzie zdrowi*);

— konwersywność, zdefiniowaną jako relacja między wyrażeniami z odwróconymi rolami syntaktycznymi, umożliwiającą tworzenie synonimów wyszukiwawczych na poziomie syntaktycznym (np. *import węgla z Polski do Holandii* — *eksport węgla z Polski do Holandii*).

• Jako odmiany relacji hierarchicznych:

— relację generyczną, zdefiniowaną jako związek hierarchiczny zachodzący między nazwą gatunkową i rodzajową (podrzędności generycznej) lub rodzajową i gatunkową (nadrzędności generycznej), pośredni lub bezpośredni (relacja bezpośredniej podrzędności/nadrzędności generycznej, relacja pośredniej podrzędności/nadrzędności generycznej). Za jej pomocą określana jest zależność semantyczna wyrażeń JIW ze względu na przynależność desygnatów pośrednich i bezpośrednich wyrażenia podrzędnego do zbioru desygnatów bezpośrednich lub pośrednich wyrażenia nadrzędnego;

— relację mereologiczną zdefiniowaną jako związek zachodzący między nazwą pewnej całości i nazwą pewnej jej części lub odwrotnie. Jest ona realizowana w dwu wariantach — jako relacja mereologiczna stała (właściwa), na przykład *drzewo* - *konar* lub relacja mereologiczna sytuacyjna (niewłaściwa), na przykład *samochód* - *silnik elektryczny* [209]. Relacja ta jest często interpretowana w literaturze przedmiotu jako odwzorowująca pozajęzykowe związki między desygnatami wyrażeń. Dyskusyjność tego ujęcia słusznie zauważa Sosińska, ze względu na to, iż rzeczywistość opisywana w języku jest wobec niego zewnętrzna, a zatem zewnętrzne są także związki między obiektami rzeczywistości pozajęzykowej;

— relację hierarchii tematycznej (hierarchii dokumentacyjnej, *topic inclusion*) [Marciszewski W., 1973, 35-52], zdefiniowaną jako związek między wyra-

zeniami reprezentującymi pewne dziedziny wiedzy (dyscypliny naukowe), stanowiącymi nazwy pewnych szerszych tematów badań i nazwaniami pewnych problemów (tematów) węższych lub też obiektów badań, na przykład *socjologia rodziny — rozwody*. Relacja ta stwarza analogiczne problemy interpretacyjne jak relacja mereologiczna z punktu widzenia językowego/pozajęzykowego charakteru związku;

— jako relację współrzędności czyli równoważności, określanej ze względu na pozostawanie w tej samej zależności hierarchicznej wobec wyrażenia semantycznie nadrzędnego dla argumentów relacji, którego denotacja jest równa sumie denotacji wzajemnie współrzędnych wyrażen hiponimicznych, zaś konotacja stanowi iloczyn teoriomnogościowy konotacji wyrażen hiponimicznych. Wyróżnione zostały relacje współrzędności: generycznej (*drzewo iglaste* - <*świerk - jodła*>), współrzędności mereologicznej (*drzewo* - <*konar - pień*>) oraz współrzędności tematycznej (*biologia* - <*zoologia - botanika*>).

• Relacje asocjacyjne (kojarzeniowe, skojarzeniowe, pokrewieństwa) — interpretowane w językoznawstwie jako wszelkie systemowe związki semantyczne między wyrażeniami języka [Saussure de F., 1961], zostały zdefiniowane jako wszelkie relacje oznaczane w słownikach JIW, z wyjątkiem synonimii wyszukiwawczej i zależności generycznych, — są to więc relacje antonimii, komplementarności, konwersywności, mereologiczne, hierarchii tematycznej. Podobnie jak w przypadku relacji mereologicznych, charakter tych relacji jest interpretowany ze względu na zależności istniejące między obiektami rzeczywistości pozajęzykowej, a więc w odniesieniu do desygnatów pośrednich JIW. Analogicznie, pewne typy relacji asocjacyjnych też mają charakter stały (systemowy) i z tego względu odwzorowywany w słownikach JIW (np. *skutek - przyczyna, proces - produkt, obiekt - proces, proces - własność*). Jak wykazały analizy przeprowadzone w pracach omawiających problem szczegółowo, wyodrębnienie relacji asocjacyjnych jest ściśle uzależnione od dziedzin wiedzy i opisywanych w nich obiektów badań [Aitchinson J., Gilchrist A., 1972; Nikitina S. E., 1978]. Z tego względu specyfikacja relacji asocjacyjnych jest szczególnie istotna w tych JIW (systemów faktograficznych i/lub dokumentacyjnych), których funkcja polega na odwzorowaniu jako cech relewantnych w wyszukiwaniu informacji własności obiektów pozajęzykowych oraz zależności zachodzących w układach tych obiektów, dzięki czemu możliwe jest różnicowanie charakterystyk wyszukiwawczych dokumentów i instrukcji wyszukiwawczych, a więc precyzyjna specyfikacja kryteriów relewancji w SIW. Na przykład Wereszczyńska-Cisło wyróżniła dla dziedziny technologii żywności 186 relacji asocjacyjnych, zredukowanych do 122 i tworzących 32 grupy [Wereszczyńska-Cisło B., 1984]. W teaurusie lingwistyki formalnej Nikitiny wyspecyfikowano 28 typów związków skojarzeniowych [Nikitina S. E., 1978].

Odwzorowaniu korelacji epistemicznych w systemie JIW, a więc ustaleniu jego porządku pionowego, służą przede wszystkim relacje hierarchiczne oraz uznane za stałe (właściwe) relacje asocjacyjne. W literaturze z zakresu teorii JIW

najczęściej przyjmuje się, że relacje mereologiczne i asocjacyjne są środkami korelacyjnego definiowania opartego na wiedzy o obiektach pozajęzykowych denotowanych pośrednio lub bezpośrednio przez wyrażenia języka.

Relacje organizujące pole semantyczne języka tworzą dwa rodzaje struktur: struktury monorelacyjne i polirelacyjne, przy czym faktycznie wszystkie JIW cechuje polirelacyjna organizacja pola semantycznego [Sosińska B., 1986, 29-40]. Równocześnie wiadomo, że rzeczywistość pozajęzykowa i językowa opisywana w zbiorach informacyjnych systemu jest zbyt złożona, aby odwzorować za pomocą wykładników wyspecyfikowanych relacji złożone związki zachodzące w układach elementów rzeczywistości językowej (dokumentacyjnej) i pozajęzykowej, a także by odzwierciedlić przy tym różnorodne aspekty tych układów odpowiadające różnym ich cechom.

3.1.2. Kategoryzacja leksyki JIW jako wykładnik pozajęzykowych kryteriów relewancji

Niewystarczalność zbioru relacji paradygmatycznych do odwzorowania relacji semantycznych, zachodzących między wyrażeniami JIW, była powodem wprowadzenia kategoryzacji leksyki języka. Pozwoliło to częściowo rozwiązać problem reprezentacji semantycznej wyrażeń wieloznacznych drogą podziału leksyki na pewne klasy według określonych cech dystynktywnych semantycznie. Proces kategoryzacji różnie rozumiany w zależności od przyjętych za podstawę podziału cech, np. gramatycznych, semantycznych lub leksykalnych, jest niezwykle trudny. Praktyka leksykograficzna języka naturalnego wykazuje bowiem, że nie budzącymi w zasadzie zastrzeżeń podziałami semantycznymi terminów są podziały na rzeczy i procesy lub też na nazwy konkretne i abstrakcyjne, co nie rozwiązuje problemu precyzyjnego określenia kryteriów relewancji [Sosińska B., 1987].

Kategoryzacja semantyczna słownictwa wymaga zawsze odwołania się do wiedzy pozajęzykowej (specjalisty z danej dziedziny oraz leksykografa) umożliwiającej wskazanie:

- w języku naturalnym — najbardziej istotnych dla użytkownika terminologii cech desygnatów, które powinny być odwzorowane w strukturze morfologicznej terminu i w jego definicji [Jadacka H., 1979];
- w JIW — cech konotacyjnych wyrażeń uznanych za relewantne w procesie wyszukiwania informacji.

Rolę kategoryzacji leksyki JIW w budowie definicji korelacyjnych organizujących pole semantyczne wykazały badania prowadzone przez Erica de Groliera, B. Vickeriego oraz D. J. Fosketta [Foskett D. J. 1981], [Gardin J. C., de Grollier E., Levery F., 1964], a także omawiane w dalszej części pracy koncepcje S. R. Ranganathana, oraz B. Bojar i O. A. Wojtasiewicza. Jak zauważa B. Sosińska w obszernej charakterystyce konkretnych uniwersalnych i specjalistycznych kategoryzacji leksyki JIW, specyficzną cechą rozwiązań przyjmowanych w różnych JIW jest niemożność rozgraniczenia kategorii semantycznych i syntaktycznych.

W niektórych JIW, zarówno uniwersalnych jak i specjalistycznych, przeprowadza się najpierw kategoryzację syntaktyczną leksyki, wyróżniając klasy wyrażen *auto-syntaktycznych* i *synsyntaktycznych*, a następnie — zależnie od przyjętych kryteriów relewancji semantycznej — ustala klasy wyrażen reprezentujących kategorie semantyczne (por. UKD, AWION, języki hasel przedmiotowych).

Zasadom semantycznej kategoryzacji leksyki JIW poświęcone są prace teoretyczne O. Unguriana [Ungurian O., 1982, 21-40], D. J. Fosketta [Foskett D. J., 1981], J. Hutchinsa [Hutchins W. J., 1978], CRG [*A Classification...*, 1975], S. R. Ranganathana [Ungurian O., 1975]. Teoretycy klasyfikacji podkreślają analogie w określaniu kategorii ontologicznych w filozofii i kategorii semantycznych w JIW. Dotyczy to w szczególności definicji kategorii w ujęciu S. Ranganathana [Sosińska B., 1987] [Ungurian O., 1982, 21-40]. W ujęciu CRG (Classification Research Group) kategoria jest interpretowana jako pewna ogólna klasa znaczeniowa wyodrębniona na podstawie cech dziedzinowych /dyscyplinowych (np. bibliotekoznawstwa, nauki o glebie, sportu i wychowania fizycznego). W pracach z zakresu teorii JIW kategorie semantyczne są często utożsamiane z fasetami. Różnice w interpretacji znaczeniowej obu terminów w ujęciu Ranganathana oraz CRG scharakteryzowała Sosińska, wykazując, że:

— w teorii Ranganathana faseta traktowana jest jako „przejaw kategorii w klasie”, powodujący wyodrębnienie w ramach jednej kategorii podzbioru jednostek leksykalnych (izolat) należących do określonej klasy dziedzinowej,

— w ujęciu CRG następuje zawężenie uniwersalnego pola semantycznego JIW do pól branżowych, podyktowane względami pragmatycznymi [*A Classification...*, 1975].

3.2 REPREZENTACJA WIEDZY O JĘZYKU INFORMACYJNO-WYSZUKIWAWCZYM

Definiując relewancję jako sieć relacji organizujących pole semantyczne danego JIW, tworzących w ten sposób definicje korelacyjne jego wyrażen, podkreśliliśmy pierwszorzędne znaczenie zasad organizacji planu treści języka i sposobów eksplikacji znaczenia. Strukturę powierzchniową traktujemy więc w naszych rozważaniach jako pochodną struktury głębokiej. Nie oznacza to deprecjacji organizacji planu wyrażania, a wręcz przeciwnie, chodzi o podkreślenie konieczności przestrzegania zasady współzależności planu treści i planu wyrażania, nie zawsze uwzględnionej przez projektantów systemów informacyjnych, niekiedy wręcz kierujących się błędnym przekonaniem, że wprowadzenie sztucznej notacji hierarchicznej automatycznie zapewnia spójność organizacyjną planu treści.

Obszerną charakterystykę zasad i środków organizacji planu wyrażania JIW przedstawiła B. Sosińska, omawiając zewnętrzne i wewnętrzne kryteria doboru notacji oraz prezentując różne rodzaje systemów notacyjnych stosowanych do odwzorowania struktury paradygmatycznej języka [Sosińska B., 1987]. Zagadnienie to podjął także E. Scibor w ramach typologii strukturalnej języków infor-

macyjnych, omawiając cechy JIW związane z prezentacją słownictwa i wyróżniając wśród nich dwie grupy: cechy tzw. uporządkowania zasadniczego (obligatoryjnego) oraz pomocniczego [Ścibor E., 1982].

Sposób prezentacji znaczenia w języku jest wykładnikiem kryteriów relewancji na jego osi paradygmatycznej (słowniku) i syntagmatycznej (w tekście), decyduje o łatwości dostępu do informacji w systemie, możliwościach formułowania strategii wyszukiwawczej przez użytkownika, determinuje więc sposób realizacji funkcji wyszukiwawczej i porządkującej języka.

Optymalny język informacyjny powinien posiadać organizację pola semantycznego przystosowaną do odwzorowania możliwie bogatego repertuaru kryteriów relewancji. Własność tę określa się jako cechę wielowymiarowości pola semantycznego. Liczba proponowanych wymiarów dla systemów dokumentacyjnych wynosi od dwu — przedmiotu i treści (Ranganathan) do pięciu — cech formalnych, treści, przedmiotu, czasu i przestrzeni [Ungurian O., 1982]. Uniwersalne określenie wymiarów pola semantycznego dla języków systemów faktograficznych nie jest możliwe, ponieważ przy ekstensjonalnej interpretacji znaczenia wymagałoby to zdefiniowania cech relewantnych wszelkich obiektów pozajęzykowych i ich cech. Sieć relacji będących wykładnikami relewancji determinuje układ (porządek pionowy, uporządkowanie) słownika JIW oraz zbioru informacyjnego systemu (katalogu, kartoteki, bibliografii, indeksu itp.). Pozostaje ponadto w ścisłej zależności z syntagmatyczną organizacją wyrażen złożonych w charakterystykach wyszukiwawczych dokumentów (ChWD) i/lub charakterystykach obiektów pozajęzykowych (ChWO).

3.2.1. Reprezentacja znaczenia w języku informacyjno-wyszukiwawczym

Charakterystyka metod prezentacji znaczenia przeprowadzana była bądź w celach pragmatycznych, w ramach formułowania zasad konstrukcji słowników JIW, bądź teoretycznych, jakkolwiek, na przykład w przypadku Ranganathana, teoria budowy klasyfikacji i określania ustroju pola semantycznego często związana była z powstaniem konkretnego języka [Dobrowolski Z., 1956] [Ungurian O., 1982, 21-40; 1975].

O repertuarze kryteriów relewancji wbudowanych w system paradygmatyczny języka decyduje jego *struktura relacyjna*. Najprostszym typem jest struktura płaska, której elementy nie są powiązane żadną relacją hierarchiczną. Przykładem języka o strukturze płaskiej może być język słów kluczowych lub wczesne typy języków deskryptorowych, w słownikach których oznaczano jedynie relację odpowiedniości wyszukiwawczej, z równoważnymi zakresowo wyrażeniami języka naturalnego. Zbiór jednostek leksykalnych tego typu języków tworzy szereg, o którego porządku pionowym decyduje kolejność liter alfabetu języka naturalnego (układ alfabetyczny).

Pozostałe języki informacyjne posiadają strukturę, w której elementy powiązane są różnego typu relacjami, tworzą więc strukturę polirelacyjną. Struktury monorelacyjne, w których wszystkie elementy byłyby powiązane relacją jednego

typu, w praktyce nie są spotykane. Możliwości prezentacji znaczenia za pomocą struktur polirelacyjnych są wykorzystywane w różnym zakresie w poszczególnych językach. W każdym JIW jednostki słownikowe pozostają wobec siebie co najmniej w relacji współrzędności i/lub równorzędności, tworząc szereg (np. *sosna - dąb, drzewo iglaste - drzewo liściaste*). Zasady ustalania porządku pionowego w szeregu stanowiły przedmiot zainteresowań zarówno teoretyków klasyfikacji (Ranganathana, Fosketa), jak i specjalistów zajmujących się językami o słownictwie paranaturalnym [Ścibor E., 1982; Ungurian O., 1982].

Ungurian scharakteryzował sformułowane przez Ranganathana dwadzieścia trzy zasady użytecznego następstwa (*helpful sequence*), wyróżniając wśród nich pięć grup odwzorowujących cechy desygnatów wyrażen (tu izolat), takie jak:

- czas („późniejszy w czasie”, „późniejszy w rozwoju ewolucyjnym”),
- lokalizacja (przestrzenna): od dołu do góry, od lewej ku prawej lub odwrotnie, zgodnie z ruchem wskazówek zegara, od peryferii ku środkowi lub odwrotnie, wzdłuż dowolnej linii określonego punktu,
- miara jednostek: zwiększająca lub zmniejszająca się, np. punkt, linia, płaszczyzna, bryła, przestrzeń czterowymiarowa; biblioteka regionalna, biblioteka miejska, biblioteka dzielnicowa,
- kompleksowość, np. głoska, sylaba, słowo, zdanie,
- inne obejmujące trzy zasady: następstwa kanonicznego, liczebności publikacji i kolejności alfabetycznej.

Z zasadami Ranganathana zbieżne są zasady sformułowane przez D. J. Fosketa: chronologiczna, ewolucyjna, wzrastającej złożoności, wzrastającego lub zmniejszającego się wymiaru, przestrzenna, preferowanej kategorii, kanoniczna i alfabetyczna [Hutchins W. J., 1978] [Ścibor E., 1982].

Cytowani autorzy podkreślają pragmatyczny charakter wymienionych zasad oraz ich niekonsekwencję, widoczną zwłaszcza w tych przypadkach, kiedy możliwe jest stosowanie do ustalania porządku pionowego identycznych szeregów dwu różnych zasad. Chodzi w szczególności o zasadę preferowanej kategorii Fosketa i liczebności publikacji Ranganathana. Niekonsekwencja ta ilustruje równocześnie złożoność problemu prezentacji znaczenia w JIW na podstawie pozajęzykowej wiedzy o desygnatach pośrednich wyrażen JIW. Stąd też w JIW, w których cały zasób leksykalny tworzący szereg (np. w języku haseł przedmiotowych lub w językach deskryptorowych z tezaurem, zawierającym jedynie część alfabetyczną i/lub alfabetyczno-hierarchiczną), częstym kryterium relewancji wyznaczającym porządek pionowy słownika są cechy planu wyrażania języka, tj. kolejność znaków w bazie notacyjnej. Ma to miejsce także w UKD, gdzie podstawową formą prezentacji znaczenia jest łańcuch klasyfikacyjny [Ścibor E., 1982, 147].

Wykładnikami związków znaczeniowych w zbiorze leksyki języka tworzącym co najmniej jeden szereg są odsyłacze, interpretowane w pracy Ścibora jako wykładniki uporządkowania pomocniczego. Określenie to wydaje się niezbyt fortunate, ponieważ może sugerować drugorzędną funkcję tego sposobu wskazywa-

nia związków znaczeniowych w słowniku JIW, zwłaszcza w językach o notacji paranaturalnej. Te metajęzykowe wyrażenia umieszczone w słowniku pełnią funkcję wykładnika relacji niespecyfikujących lub specyfikujących (często relacji odpowiedniości wyszukiwawczej), przybierając różne formy graficzne:

- odsyłacz całkowity (ogólny, pusty), najczęściej występujący w formie *zob.*, *U* (używaj), *stos.*, *>* jest wykładnikiem międzyjęzykowej relacji odpowiedniości wyszukiwawczej, np. *Lingwistyka zob. Językoznawstwo* (jhp), *Pies stos. Psy*, *ONZ > Organizacja Narodów Zjednoczonych*, *Brain drain U Drenaż mózgow:*
- odsyłacz systematyczny (odsyłacz uzupełniający), najczęściej występujący w formie *zob. też*, *por.*, *SD*, *SDR*, *SDC*, *WD*, *WDC*, *WDR* jest wykładnikiem językowej lub pozajęzykowej relacji hierarchicznej, np. *Broń palna zob. też Pistolet*; *Zboże WD/WDR Jęczmień, Pszenica, Żyto*;
- odsyłacz uzupełniający (odsyłacz pełny), w formie *zob. też.*, *por.*, *SD*, *WD*, *KD* jest wykładnikiem relacji hierarchicznej lub skojarzeniowej;
- odsyłacz porównawczy (odsyłacz uzupełniający) w formie *zob. też*, *por.*, *por. też*, *KD* — jest wykładnikiem relacji skojarzeniowej, innej niż relacja hierarchiczna, np. *Szkoły zob. też Nauczanie*; *Mróz KD Odmrozenia*;
- odsyłacz orientacyjny (odsyłacz tekstowy) w formie *zob. też* wskazuje związek wyrażenia z ogólnie określoną kategorią semantyczną lub wskazuje zalecaną strukturę syntagmatyczną, np. *Bibliografia narodowa zob. też Poszczególne bibliografie narodowe*, jak *Bibliografia czeska*, *Bibliografia polska*, *Bibliografia rosyjska*; *Pamiętniki zob. też Pamiętniki dotyczące poszczególnych dziedzin, krajów*, *Bezrobocie - pamiętniki*, *Chłopi - pamiętniki*, *Medycyna - pamiętniki*, *Polska - pamiętniki*.

Znaczenie odsyłaczy jako wykładników kryteriów relewancji ilustrują zwłaszcza te JIW, w których związki semantyczne między wyrażeniami oparte są na cechach obiektów rzeczywistości pozajęzykowej — na przykład w Międzynarodowej Klasyfikacji Patentowej (por. „reguła pierwszeństwa” — rozdz. 3.3.1.2.1.).

Ważnymi sposobami eksplikacji znaczenia wyrażen w językach o notacji paranaturalnej i szeregowym porządku pionowym są symbole pomocnicze (wyrażenia metajęzykowe) służące do definiowania znaczeń i eliminacji homonimii, wynikającej z równości wyrażen JIW z wyrażeniami języka naturalnego. Nazwy tych symboli, funkcjonalnie równoważnych, są różne, co wynika z tradycji rozwoju poszczególnych rodzajów języków i praktyki ich stosowania przez służby biblioteczno-informacyjne. Do symboli stosowanych w tej funkcji należą:

- dopowiedzenie w języku haseł przedmiotowych, precyzujące znaczenie tematu, np. *Hel*, *gaz* i *Hel*, *wieś* (podanie znaczenia w języku naturalnym explicite);
- *scope note*, stosowany najczęściej w języku deskryptorowym przez podanie części znaczenia explicite, tj. wskazanie kategorii znaczeniowej, np. *Odra (rzeka)*, *Odra (choroba)* lub wskazanie nazwy dziedziny, w której jest używane, *Róża (medycyna)*, *Róża (botanika)*, *Róża (ogrodnictwo)*;

- kwalifikatory — wyrażane analogicznie jak dopowiedzenie i scope note. Ich wprowadzenie do słownika wymaga odwoływania się do wiedzy o znaczeniu wyrażen języka naturalnego oraz często do wiedzy pozajęzykowej w przypadku definiowania znaczenia przez wskazanie dziedziny stosowania wyrażenia.

Kategoryzacja semantyczna i/lub syntaktyczna stanowi — jak to wykazemy omawiając prezentację wiedzy w poszczególnych językach — podstawowy sposób zapewnienia wielowymiarowego pola semantycznego i komunikowania uniwersalnych cech obiektów rzeczywistości pozajęzykowej drogą kombinacji wyrażen w zdaniach JIW. W połączeniu z wykładnikami relacji hierarchicznych i skojarzeniowych reprezentuje metodę odwzorowania stosunkowo mocnych kryteriów relewancji, umożliwiając spełnienie warunku produktywności określonego w zasadach eksplikacji Carnapa. W obszernej literaturze z zakresu teorii JIW podkreśla się znaczenie kategoryzacji dla zapewnienia wyrażeniom językowym jednoznaczności i wyeliminowania polisemii właściwej językowi naturalnemu [Foskett D. J., 1981] [Robowski J., 1974] [Sosińska B., 1987] [Ungurian O., 1982, 21-40], co z kolei umożliwia spełnienie zasady precyzji eksplikacji terminów w ujęciu Carnapa.

Sposoby komunikowania znaczenia wyrażen w języku, którego równorzędne jednostki leksykalne powiązane są relacjami hierarchicznymi oraz skojarzeniowymi, są zróżnicowane stosownie do rodzaju notacji odwzorowującej w planie wyrażania cechy relewantne planu treści [por. Sosińska B., 1987].

Wykładnikami kategorii i/lub faset¹ są litery bazy notacyjnej, najczęściej reprezentujące notację mieszaną. Kategoryzacja wyrażen (uporządkowanie fasetowe) powoduje wyodrębnienie w zbiorze wyrażen JIW kilku podzbiorów słownictwa, pozostających w stosunku do siebie w relacji równorzędności semantycznej, choć czasami różniących się funkcją syntagmatyczną — przykładem mogą tu być modyfikatory w tezaurusach fasetowych, wykładniki relacji syntagmatycznych (wskaźniki roli) w kodach semantycznych, tzw. kody indeksowe w Międzynarodowej Klasyfikacji Patentowej.

W językach klasyfikacyjnych (klasyfikacjach) wyróżnienie kategorii i faset prowadzi do polihierarchicznej struktury pola semantycznego, przeciwstawianej strukturze monohierarchicznej. Podstawową jednostką organizacji pola semantycznego w klasyfikacjach monohierarchicznych jest łańcuch klasyfikacyjny, a więc część klasyfikacji obejmująca ciąg klas. Definicje łańcucha klasyfikacyjnego omawiają m.in. Ungurian [Ungurian O., 1976] i Ścibor [Ścibor E., 1982]. Kryteria relewancji odwzorowywane za pomocą łańcucha klasyfikacyjnego (i klasyfikacji) stanowią tzw. słabą interpretację relewancji (por. rozdz. 2.1.3.), określaną przez

¹ Podkategoria; termin wprowadzony przez S. R. Ranganathana w jego teorii klasyfikacji, zapożyczony z terminologii jubilerskiej, gdzie oznacza ścięcie graniastej krawędzi naroża szlifowanego kamienia, potęgujące efekty barwne wskutek załamywania i rozszczepiania promieni świetlnych. Metaforyczne użycie nazwy w teorii klasyfikacji miało podkreślić wieloaspektową analizę rzeczywistości, której elementy podlegają klasyfikacji. [Słownik encyklopedyczny terminologii języków i systemów informacyjno-wyszukiwawczych, 1993, 38].

formułę „*A jest związane z B wtt jeśli występują razem w pewnym C, a ponadto uwzględnia się wszystkie wywodzące się z nich klasy*” [Hillman D. J., 1964].

Relacje zachodzące między wyrażeniami będącymi nazwami klas w łańcuchu stanowią wszelkiego rodzaju związki hierarchiczne językowe i pozajęzykowe, a więc generyczne, mereologiczne, hierarchii tematycznej, co powoduje, że na znaczenie wyrażeń JIW składają się zarówno cechy obiektów językowych jak i pozajęzykowych, odwzorowane w notacji, np. w UKD [Ścibor E., 1982]:

631 Rolnictwo w ogólności. Gospodarstwo wiejskie (Dziedzina)

631.3 Maszyny. Narzędzia i urządzenia rolnicze (Obiekty)

631.31 Maszyny i narzędzia do uprawy roli. Maszyny i narzędzia pomocnicze (Rodzaj obiektów 1)

631.312 Pługi (Rodzaj obiektów 1.1.)

631.312.8 Pługi talerzowe (Rodzaj obiektów 1.1.1)

Ten sposób prezentacji znaczenia jednostek leksykalnych Ścibor nazywa uporządkowaniem formalno-semantycznym, uznając je za niepoprawne z racji odwzorowania w łańcuchu różnych relacji językowych i pozajęzykowych. Właśność ta jest jednak typowa dla wszelkich klasyfikacji, których kryteria relewancji oparte są głównie na specjalistycznej wiedzy pozajęzykowej.

Zasady prezentacji znaczenia wyrażeń języków klasyfikacyjnych pozostających w relacji współrzędności, a więc tworzących szereg klasyfikacyjny, omówiono wyżej (Ranganathan, Foskett) jako ilustrację względności kryteriów relewancji opartych na kryteriach pozajęzykowych, co widoczne jest zwłaszcza w tych przypadkach, kiedy układ wyrażeń w szeregu w zasadzie uznaje się za oparty na kryteriach formalnych, np. (za: [Ścibor E., 1982]):

633.811 Kwiaty

633.812 Ziola

633.814 Drewno

633.816 Rośliny balsamiczne. Kadzidło

633.819 Inne. Kemia piżmowa, piżmian właściwy. Hibiscus abelmaschus.

Wykorzystanie struktury dendrytowej z łańcuchem klasyfikacyjnym w organizacji pola semantycznego języków o notacji paranaturalnej, na przykład języków deskryptorowych, polega na budowie części systematycznej tezauryusa, komplementarnej wobec części alfabetycznej i/lub alfabetyczno hierarchicznej, zawierającej deskryptory uporządkowane w układzie odwzorowującym klasyfikację przedmiotów treści dokumentów. Należy zauważyć, że ten sposób prezentacji pola semantycznego języków deskryptorowych traktowany jest w świetle dokumentów różnych krajów (Polska Norma, American National Standard) jako fakultatywny.

Eksplikacja znaczeń wyrażeń JIW w ramach struktur łańcuchowych powoduje z jednej strony kompleksowość sposobu prezentacji, ponieważ dane zawarte w poszczególnych elementach symbolu są rozwinięciem danych z symboli poprzedzających, z drugiej strony wymaga od posługującego się językiem znajo-

mości całej struktury paradygmatycznej języka w celu odnalezienia wyrażenia niezbędnego w opracowywanej ChWD.

Stąd też coraz szersze zastosowanie jako pomoc w prezentacji zasobu leksykalnego słowników JIW (o notacji sztucznej i paranaturalnej) znajdują indeksy, czyli spisy pomocnicze do zbioru informacji (tu leksyki), wyliczające elementy tego zbioru w układzie innym niż przyjęty dla tego zbioru, powstałym w wyniku odpowiednich transformacji, np. indeks przedmiotowy do UKD opisuje znaczenie symboli klasyfikacyjnych w języku haseł przedmiotowych, indeks permutacyjny do tezaury umożliwia wybór deskryptorów równokształtnych z wielowyrzowymi wyrażeniami języka naturalnego według różnych ciągów liter [Chmielewska-Gorczyca E., 1977]. Rosnące znaczenie indeksów w funkcji metajęzykowej, oprócz ich tradycyjnej funkcji metainformacyjnej, jako wskaźników adresowych wyrażenń relewantnych w zbiorze informacyjnym (katalogu, bibliografii itp.), widoczne jest w związku z powiększaniem się zbiorów informacji w systemach oraz stosowaniem w nich różnych języków informacyjnych. Tendencję tę, ilustrującą znaczenie sposobu prezentacji wiedzy o języku, prześledzimy w rozdziale poświęconym systemom informacji online.

3.2.2. Wykładniki wiedzy o języku i rzeczywistości w gramatyce języka informacyjno-wyszukiwawczego

Relacje syntagmatyczne w JIW stanowiły przedmiot rozważań w szeregu prac ukierunkowanych teoretycznie i aplikacyjnie. Ocena sprawności wyszukiwania informacji wykazuje, że precyzja odwzorowania kryteriów relewancji na osi syntagmatycznej języka wpływa na zdolność selekcji informacji w stopniu nie mniejszym niż spójność organizacyjna pola semantycznego na osi paradygmatycznej [Buchanan B., 1979] [Foskett D. J., 1981] [Sosińska B., 1983, 111-114].

Typologii relacji syntagmatycznych w JIW poświęcona jest praca E. Chmielewskiej-Gorczycy, w której uwzględniono JIW systemów dokumentacyjnych, a wśród nich tych, których zbiory informacyjne tworzą ChWD [Chmielewska-Gorczyca E., 1983, 144-151]. Spośród najważniejszych ustaleń, postulatów i wniosków sformułowanych przez tę autorkę na podstawie analizy różnych JIW wymienimy:

- wykazanie istnienia zależności organizacji syntagmatycznej języka od jego organizacji paradygmatycznej;
- wykazanie istnienia związku między strukturą paradygmatyczną JIW, instrumentami informacyjnymi systemu („możliwościami technicznymi”) oraz potrzebami użytkowników;
- przedstawienie projektu optymalnej struktury zapisu indeksowego opartego na wyróżnieniu w paradygmatyce języka dwu kryteriów relewancji: kategorii przedmiotu treści dokumentu, a więc kryterium relewancji przedmiotowej, oraz kategorii sposobu ujęcia przedmiotu treści, a więc kryterium relewancji interpretacyjnej;

- scharakteryzowanie sposobów wzbogacenia repertuaru wykładników kryteriów relewancji w języku i ich rozszerzenia na wszystkie elementy ChWD. Polegają one na wykorzystaniu reguł transformacji tekstów ChWD umożliwiających przeprowadzanie operacji permutacji, rotacji, obcinania, zanurzania i przetaczania.

Typologia reguł gramatycznych JIW stanowi zwykle element rozważań w pracach poświęconych typologii JIW [Ścibor E., 1982]. Ścibor wyróżnił cztery rodzaje gramatyk stosowanych w JIW: gramatykę zerową, pozycyjną, częściowo pozycyjną oraz niepozycyjną.

Typologia przedstawiona przez E. Chmielewską-Gorczycę oparta jest na siedmiu kryteriach: liczbie argumentów relacji, rodzaju wykładników, funkcji, poziomie języka, typie tekstu, formie wykładników oraz obligatoryjności stosowania wykładników relacji. Za najistotniejsze uznała autorka podziały:

- według kryterium funkcji, na podstawie którego wyróżniła *relacje łączące i specyfikujące*,
- według kryterium formy (postaci) wykładników, na podstawie którego wyróżniła pozycyjne i niepozycyjne sposoby wyrażania funkcji elementarnych jednostek leksykalnych w tekście JIW, zaś wśród sposobów niepozycyjnych wyróżniono z kolei wykładniki izolowane, pełniące funkcje analogiczne jak morfemy swobodne w języku naturalnym oraz nieizolowane wbudowane (jako afiksy) w jednostkę leksykalną.

Jako pierwszy typ relacji syntagmatycznych (nie specyfikujących) wyróżniono, analogicznie jak w innych pracach z omawianego zakresu, prostą koordynację jednostek leksykalnych, będącą faktycznie relacją współwystępowania nazw obiektów denotowanych bezpośrednio i/lub bezpośrednio w SIW. Ten typ relacji jest stosowany najczęściej w językach słów kluczowych, językach deskryptorowych oraz w niektórych klasyfikacjach biblioteczno-bibliograficznych. Por. [Bielicka L., 1985] Również relacje nie specyfikujące reprezentują te języki, w których stosowane są szeroko rozumiane wskaźniki więzi (nie zawężone jedynie do języków deskryptorowych).

Ze względu na postać wykładników stosowanych do odwzorowania relacji specyfikujących wyróżnia się:

- 1) wskaźniki roli określające funkcję desygnatu pośredniego jednostki leksykalnej w danym kontekście;
- 2) systemy relatorów, wykorzystujące zarówno symbolikę sztucznie skonstruowaną, jak i wyrażenia języka naturalnego;
- 3) modyfikatory, pełniące funkcję wykładników kategoryalnych (akcydenalnych cech funkcjonalnych lub kategorii ontologicznych);
- 4) reguły gramatyki pozycyjnej, a wśród nich reguły gramatyki pozycyjnej oraz formalnej;
- 5) kategoryzację semantyczną przeprowadzaną na osi paradygmatycznej języka, natomiast funkcjonalnie podobną do wskaźników roli i różniącą się od nich jedynie formą; wskaźniki roli są porównywalne z morfemami izolowanymi

w języku naturalnym, podczas gdy wykładniki kategorii semantycznych można porównać na płaszczyźnie syntagmatycznej z morfemami nieizolowanymi (związanymi), stanowiącymi integralną część jednostki leksykalnej.

Najczęściej przyjmuje się, że relacje syntagmatyczne są relacjami językowymi zachodzącymi między elementami charakterystyki wyszukiwawczej dokumentu, charakteryzowanymi w literaturze, głównie na płaszczyźnie syntaktycznej. Niektórzy autorzy (Ranganathan, Foskett, Buchanan) przeprowadzają analizę struktur syntaktycznych na poziomie semantyki składniowej, opisując funkcje semantyczne jednostek leksykalnych w tekście (por. [Sosińska B., 1983, 111-144]). Opis ten polega, mówiąc najogólniej, na odwzorowaniu w lub w charakterystyce wyszukiwawczej obiektu systemu faktograficznego funkcji bezpośrednich i pośrednich desygnatów wyrażen języka w określonej sytuacji. Wyrażenia te są równocześnie scharakteryzowane w strukturze paradygmatycznej JIW, która odwzorowuje pewne cechy rzeczywistości językowej (dokumentacyjnej) i/lub pozadokumentacyjnej (pozajęzykowej). Stąd też w tych językach, w których struktura paradygmatyczna odzwierciedla zależności między obiektami świata zewnętrznego, wykładniki relacji tekstowych również służą do odwzorowania wiedzy pozajęzykowej poprzez określenie funkcji i/lub cech obiektów w danej sytuacji. Stwierdzenie to postaramy się zweryfikować, dokonując syntetycznego przeglądu sposobów odwzorowania wiedzy w syntagmatyce omawianych JIW (rozdz.3.3.2.).

3.3 REPREZENTACJA WIEDZY W JĘZYKACH INFORMACYJNO-WYSZUKIWAWCZYCH SYSTEMÓW DOKUMENTACYJNYCH

W charakterystyce sposobów reprezentacji wiedzy w różnych językach informacyjno-wyszukiwawczych uwzględniamy dwa aspekty określające układ rozdziału:

- oddzielny opis wiedzy odwzorowywanej odpowiednio na osi *paradygmatycznej* (słowniku) i *syntagmatycznej* (tekście) języka,
- podział języków informacyjno-wyszukiwawczych na *uniwersalne* i *specjalistyczne*, tradycyjnie przyjmowany w literaturze ze względu na zakres tematyczny odwzorowywanej w nich wiedzy.

3.3.1. Wiedza reprezentowana w paradygmatyce języków informacyjno-wyszukiwawczych

Komunikowanie cech obiektów rzeczywistości pozajęzykowej jako cech relewantnych w systemie informacyjnym wymaga odwołania się do tych własności desygnatów pośrednich, które są odwzorowywane jako wspólny element znaczenia danej kategorii semantycznej leksyki JIW. Precyzyjne określenie kryteriów relewancji w SIW wymaga, jak to udowodnili Hillman i Sosińska, przyjęcia konotacyjnej interpretacji znaczenia, na podstawie której określa się zespół cech nie-

zbędnych i wystarczających do identyfikacji opisu desygnatu w zbiorze informacyjnym systemu. Wyrażenia JIW mogą reprezentować, z punktu widzenia ich konotacji, następujące kategorie [Sosińska B., 1987]:

- 1) predykaty proste, tj. własności elementarnych obiektów,
- 2) predykaty złożone, tj. konkatenacje predykatów prostych,
- 3) predykaty złożone, których intensję stanowi zespół własności ich desygnatów wskazywany za pomocą odpowiednich postulatów znaczeniowych.

Cechy semantycznie proste desygnatów wyrażen JIW opisane w języku predykatów odwzorowują zależności między tymi desygnatami i jednostkami języka, są więc semantycznie motywowane w odróżnieniu od tych cech, które odwołują się wyłącznie do struktury systemu językowego i zakresu znaczeniowego jego wyrażen bez względu na to, co desygnują w rzeczywistości pozajęzykowej.

Wykładnikami elementarnych semantycznie cech obiektów pozajęzykowych w paradygmatyce języka są *kategorie semantyczne* w klasyfikacjach o notacji numerycznej lub alfanumerycznej, *relacje mereologiczne* oraz część *relacji asocjacyjnych*, odwzorowujących strukturę rzeczywistości, a więc układy obiektów i ich własności akcydentalne.

3.3.1.1. Wiedza o rzeczywistości w uniwersalnych językach informacyjno-wyszukiwawczych

Wyróżniając klasy języków uniwersalnych i specjalistycznych kierujemy się tradycją opartą na kryterium zakresu tematycznego wiedzy w nich odwzorowanej. Jak wykażemy w dalszych rozważaniach ta kwalifikacja nie zawsze jest uzasadniona.

3.3.1.1.1. Wiedza o rzeczywistości w Uniwersalnej Klasyfikacji Dziesiętnej. Uniwersalna Klasyfikacja Dziesiętna jest językiem, w którym przyjęto „uniwersalne” kryteria relewancji typowe dla arbitralnej reprezentacji wiedzy, w której dominuje dążenie do odwzorowania w zbiorze informacyjnym systemu cech przedmiotów treści dokumentów (relewancji przedmiotowej) i podawanie gotowych symboli dla wszystkich zagadnień. Wiedza o języku i wiedza o rzeczywistości są w UKD ściśle ze sobą powiązane — związek ten wyraża się w wyodrębnieniu dwu kategorii syntaktycznych wyrażen:

- autosyntaktycznych, zawartych w tablicach głównych, obejmujących dziesięć klas wiedzy, podlegających dalszemu podziałowi do kilkunastu stopni,
- synsyntaktycznych, umieszczonych w tablicach pomocniczych, w których z kolei wyróżniono poddziały wspólne oraz poddziały analityczne [Ungurian O., 1978].

Kategoryzacja semantyczna wyrażen elementarnych, jakkolwiek nie założona *explicite*, została wprowadzona w tablicach pomocniczych dla wyrażen synsyntaktycznych. Polega ona na wyróżnieniu dziewięciu kategorii semantycznych odwzorowujących następujące cechy desygnatów:

1) W poddziałach pomocniczych (formalnych):

— cechy języka tekstu dokumentu, uporządkowane następnie częściowo za pomocą hierarchicznej relacji generycznej, a więc odwzorowującej pewną wiedzę językową, na przykład: =089 *języki sztuczne*, =089,2 *esperanto*, =30 *język niemiecki* oraz =393 *język holenderski, język flamandzki, Afrikaans*;

— cechy formy (wydawniczej lub piśmienniczej) dokumentu, również zorganizowane następnie za pomocą relacji generycznej, np. (047) *raporty sprawozdania, zawiadomienia* i (047.2) *sprawozdania z podróży*.

2) W poddziałach wspólnych kategorii treściowych:

— cechy czasu (chronologii i czasu trwania zdarzenia), służące określeniu dat, periodyczności, czasu geologicznego, chronologicznego, kulturowego. Cechy te również organizowane są za pomocą językowej relacji generycznej;

— cechy lokalizacyjne (miejsca) przedmiotu, zdarzenia lub zjawiska charakteryzowanego w dokumencie, zorganizowane według formuły fasetowej, tj. w ramach czterech faset: miejsca, i przestrzeni w ogólności, fizyko-geograficznej, podziału politycznego (współczesnego). Następnie, przy użyciu relacji generycznej zorganizowano siedem podkategorii pomocniczych dla poddziałów wspólnych miejsca (strefy, orientacji, jednostek administracyjno-politycznych, wspólnot związkowych państw niepodległych, pola działania inicjatywy prywatnej lub kolektywnej, rozmieszczenia, przebiegu, pochodzenia, przeznaczenia i podziału regionalnego). Kraje świata współczesnego uporządkowane są według pozajęzykowej relacji mereologicznej;

— cechy etniczne (narodowości, rasy) zagadnień odwzorowywanych symbolem głównym (autosyntaktycznym);

— cechy punktu widzenia (aspektu), a więc odwzorowujące kryteria relewancji interpretacyjnej, np. .001 *punkt widzenia teoretyczny*, .003 *punkt widzenia ekonomiczny*;

— cechy metodologiczne (interpretacyjne) sposobu ujęcia przedmiotu treści dokumentu, np. *ateistyczny, marksistowski, parapsychologiczny*. Cechy te są odwzorowywane przez modyfikację symboli głównych poprzedzonych znakiem .000;

— cechy strukturalne (zawartościowe) materiału będącego przedmiotem treści dokumentu, uporządkowane następnie w siedmiu fasetach za pomocą relacji generycznej, np. -034 *metale*, -035 *materiały pochodzenia organicznego*;

— cechy osobowe, zorganizowane w siedmiu fasetach, np. -053 *osoby według kategorii wieku*, -054 *osoby według charakterystyk etnicznych, narodowościowych ub obywatelskich*.

3) W poddziałach wspólnych analitycznych:

— cechy aspektu treści dokumentów charakterystycznych dla danej dziedziny wiedzy, np. w dziale 82 — .01 *estetyka literatury*, .0: *interpretacje, przekłady*,

— cechy odwzorowujące stałe lub ruchome obiekty lub ich rodzaje specyficzne dla poszczególnych dziedzin wiedzy, np. w dziale 62 *Inżynieria -- -1 ogólny podział maszyn, -111 maszyny proste, -15 turbiny*.

Organizacja pola semantycznego leksyki tablic głównych odwzorowuje w pierwszej kolejności *kryteria relewancji przedmiotowej* „bycia przedmiotem treści dokumentu”, „bycia o przedmiocie treści” w układzie dziedzinowym, w którym wyliczono zorganizowane za pomocą wszystkich trzech rodzajów relacji hierarchicznych (językowych i pozajęzykowych), wyrażenia desygnujące szczegółowe dyscypliny, przedmioty, metody i problemy badawcze. Jak wykazała Sosińska w pracach poświęconych problemom denotacji i konotacji wyrażzeń JIW [Sosińska B., 1986], w zasadzie monohierarchiczna struktura leksyki tablic głównych UKD odwzorowuje własności desygnatów pośrednich (obiektów rzeczywistości pozajęzykowej) oraz pozajęzykowe układy (zależności) tych obiektów, np. *być o rybach* — także *być o mięsie* — *być o produkcji pochodzenia zwierzęcego* — *być o produkcji leśnictwa, rolnictwa*. Równocześnie wykazywane są zależności paradygmatyczne między wyrażeniami języka komunikowane za pomocą językowych relacji generycznych.

Na podstawie powyższej charakterystyki możemy zatem stwierdzić, że UKD projektowana (podobnie jak i inne klasyfikacje pokrewne strukturalnie) jako język podporządkowany uniwersalnym kryteriom relewancji ma wbudowaną w polu semantycznym całą wiedzę o rzeczywistości językowej, tworzącą swoisty amalgamat z wiedzą językową komunikowaną za pomocą relacji generycznych, organizujących część leksyki zarówno autosyntaktycznej, jak i synsyntaktycznej.

3.3.1.1.2. Wiedza o rzeczywistości w Klasyfikacji Dwukropkowej Ranganathana. Klasyfikacja Dwukropkowa Ranganathana jest językiem informacyjnym, w którym zastosowano konsekwentnie paralelne odwzorowanie wiedzy o rzeczywistości w trzech planach językowych (w planie treści, werbalnym oraz notacyjnym) drogą kategoryzacji leksyki według dwu kryteriów epistemicznych:

- kryterium dziedzinowej systematyzacji przedmiotów i aspektów treści dokumentów oraz obiektów i metod badań, odwzorowanego explicite w porządku pionowym języka,
- kryterium pięciu kategorii ontologicznych (interpretowanych odmiennie niż w tradycji filozofii europejskiej), odwzorowanego implicite w klasach w postaci faset, czyli tzw. przejawów kategorii, zorganizowanych z kolei za pomocą konsekwentnie stosowanych (językowych) relacji generycznych. Kategoryzacja jako jawne odwzorowanie wiedzy pozajęzykowej ma miejsce na poziomie syntagmatycznym języka w postaci tzw. formuł fasetowych [Ungurian O., 1975] [Sosińska B. 1983].

Szczegółowe omówienie kategoryzacji w Klasyfikacji Ranganathana zawierają prace Unguriana i Sosińskiej. Ungurian przeprowadził krytyczną analizę kategoryzacji przyjętej przez Ranganathana oraz zaproponował własną, uwzględniającą pięciowymiarowy ustrój pola semantycznego JIW. Kategoryzacja Ranganathana obejmuje wyrażenia, których desygnatami pośrednimi są:

— w kategorii Czas (*Time*) — zjawiska i jednostki związane z określeniem trwania,

— w kategorii Przestrzeń (*Space*) — jednostki przestrzeni pozaziemskiej, atmosfery, powierzchni i wnętrza Ziemi,

— w kategorii Energia (*Energy*) — wszelkie działania dokonywane na wszelkiego rodzaju bytach lub też wykonywane przez wszelkiego rodzaju byty,

— w kategorii Materia (*Matter*) — jednostki materiałów, surowców, katalizatorów, produktów, usług, właściwości i ich wartości,

— w kategorii Indywidualność (*Personality*) — wszelkie desygnaty nie należące do pozostałych kategorii.

Dzięki wyróżnieniu trzech planów JIW w Klasyfikacji Ranganathana wiedza o rzeczywistości jest wyraźnie oddzielona od wiedzy językowej z zachowaniem jednojednoznaczności odwzorowania jednostek wiedzy o rzeczywistości wyróżnionych w planie treści, a następnie w planach werbalnym i notacyjnym. Jednostkom planu treści, klasom i izolatom odpowiadają w planie werbalnym tematy (nazwy klas) i nazwy izolat, w planie notacyjnym symbole klas i izolat.

Dalsza szczegółowa kategoryzacja klas i izolat jest wykazywana w systemie paradygmatycznym klasyfikacji (por. [Sosińska 1983, 11-114]).

Wskazane cechy strukturalne Klasyfikacji Dwukropkowej świadczą o tym, że organizacja jej pola semantycznego oparta została na konotacyjnej interpretacji znaczenia, to zaś z kolei dowodzi, że jedynie przy tej interpretacji możliwe jest konsekwentne odróżnienie w JIW wiedzy o rzeczywistości i wiedzy o języku (desygnatach pośrednich i bezpośrednich wyrażeniach). Decyduje to o przydatności języka zarówno w systemach informacji dokumentacyjnej jak i faktograficznej.

3.3.1.2. Wiedza o rzeczywistości w języku hasel przedmiotowych (jhp)

Kryteria relewancji przyjęte w języku hasel przedmiotowych odwołują się do cech języka oraz do cech desygnowanych pośrednio obiektów, będących przedmiotami treści dokumentów oraz bezpośrednio cech dokumentów. Podobnie jak w UKD podstawowym wykładnikiem cech relewantnych w tym języku jest *kategoryzacja syntaktyczna* leksyki, polegająca na rozróżnieniu wyrażeni samodzielnych składniowo (autosyntaktycznych tematów) oraz niesamodzielnych składniowo (syntaktycznych określników). Kategoryzacja semantyczna przeprowadzana jest w ramach dwu wyróżnionych kategorii syntaktycznych.

Kategoryzacja tematów, interpretowanych niekiedy jako przedmioty treści dokumentu, przeprowadzana w polskiej literaturze przedmiotu według dwu kryteriów, odwołuje się do:

1) cech desygnatów bezpośrednich (cech dokumentów), na podstawie czego wyróżnia się:

— tematy przedmiotowe oraz tematy klasowe, służące odwzorowaniu cech treściowych dokumentów,

— tematy formalne odwzorowujące cechy istotne w opisie formy dokumentu;

2) cech desygnatów pośrednich będących przedmiotami treści dokumentów, na podstawie których wyróżnia się tematy ogólne i jednostkowe (indywidualne). Tematy jednostkowe podlegają dalszej kategoryzacji na: *tematy osobowe, językowo-etniczne, geograficzne, chronologiczne oraz rzeczowe* [Ćwiekowa J., 1974].

Cechy obiektów pozajęzykowych desygnowanych przez poszczególne rodzaje tematów obejmują:

a) dla tematów przedmiotowych: nazwy i własności obiektów, osób, zjawisk, stanów, czynności (np. *Pojazdy, Lekarze, Uczucia, Rodzina, Nauczanie*),

b) dla tematów klasowych: dziedziny wiedzy naukowej i praktycznej (*Socjologia, Stolarstwo, Elektronika*),

c) dla tematów formalnych: cechy formy wydawniczej (piśmienniczej dokumentów).

Kategorie tematów ogólnych odwołują się do pewnych wspólnych cech obiektów jednakowych (podobnych), desygnowanych przez tematy klasowe i formalne, choć także do cech nie przysługujących wszystkim desygnatom nazw (*Psychologia, Ptaki, Czasopisma*).

Kategoria tematów jednostkowych (indywidualnych) obejmuje podkategorie, określane za pomocą wyrażen będących nazwami:

— osób, dynastii, rodów (*Adam Mickiewicz, Jagiellonowie, Estreicherowie*),

— języków naturalnych (*Język polski*), grup językowych (*Języki słowiańskie*), a także zjawisk związanych z przynależnością do określonego narodu (*Literatura węgierska, Malarstwo francuskie, Czasowniki polskie*),

— jednostek geograficznych (państw, regionów, kontynentów, wysp, pasm górskich, oceanów i mórz, rzek, jezior itp.),

— jednostek chronologicznych, uściślających cechy wydarzeń za pomocą dat (*Wojna 1914-1918 I Światowa, Powstanie 1863 Styczniowe*),

— obiektów nieożywionych reprezentowanych przez nazwy instytucji, wytworów myśli ludzkiej, pierwiastków i związków chemicznych oraz ożywionych, na przykład gatunków zwierząt i roślin.

Kategoryzacja tematów w języku haseł przedmiotowych odwzorowuje cechy desygnatów pośrednich i bezpośrednich wyrażen odpowiadających kryterium przedmiotu (treści) dokumentu jako kryterium relewancji, co — nawiązując do naszych wcześniejszych rozważań — można uznać za kryterium relewancji przedmiotowej. Przedstawione przykłady kategorii semantycznych, określających kryteria relewancji w jhp są reprezentatywne dla polskiej praktyki katalogowania przedmiotowego zbiorów dokumentów bibliotek uniwersalnych. Instrukcje opracowania przedmiotowego zbiorów innych krajów (USA, ZSRR), bądź zbiorów specjalistycznych poszczególnych dziedzin wiedzy, świadczą o zróżnicowaniu kryteriów ustalania kategorii przedmiotu dokumentu, co w praktyce uniemożliwia wskazanie ogólnych, uniwersalnych kryteriów relewancji przed-

miotowej w omawianym języku. Są one bowiem w znacznym stopniu zdeterminowane przez strukturę zbiorów informacyjnych systemu (biblioteki). Właściwość ta różni jhp od omówionych wyżej uniwersalnych klasyfikacji.

W pracach teoretycznych i metodycznych poświęconych kategoryzacji semantycznej wyrażen synsyntaktycznych jhp, a więc określników podkreśla się, że ich funkcja polega na modyfikacji, zawężeniu zakresu znaczeniowego tematu w zdaniu jhp, tj. haśle przedmiotowym, poprzez wskazywanie aspektu ujęcia przedmiotu dokumentu, a więc aspektu opisu bezpośredniego lub pośredniego desygnatu wyrażenia językowego [Ćwiekowa J. 1974] [Sadowska J., 1982; 1983, 53-68]. Kategoryzacja określników odwzorowuje kryteria relewancji interpretacyjnej ze względu na określony tematem przedmiot. W polskiej praktyce katalogowania przedmiotowego wyróżnia się trzy podstawowe kategorie określników:

- określniki szczegółowe służące wskazaniu indywidualnych cech przedmiotów treści lub formy dokumentu,
- określniki ogólne, służące wskazaniu aspektu interpretacji tematu lub ujęcia formy dokumentu,
- określniki lokalizujące, służące wskazaniu miejsca przedmiotu (formy) dokumentu w czasie lub przestrzeni.

Dalszy podział wymienionych kategorii prowadzi do wyróżnienia w kategorii określników szczegółowych tych, które wskazują cechy istotne w interpretacji tematu (przedmiotu / formy):

— cechy jednostkowe (określniki jednostkowe), np. dzieła twórcy (*Sienkiewicz Henryk — Krzyżacy*), dzieło wyróżnione w zbiorze innych dzieł (*Literatura polska — Bogurodzica*), wydarzenia związanego z konkretnym miejscem (*Grunwald — 1410*);

— cechy typowe obiektów pozadokumentacyjnych, desygnowanych pośrednio przez którykolwiek z tematów (określniki specjalne), np. *Pieniądz — obieg*;

— cechy rodzajowe desygnatów pośrednich tematów (określniki gatunkowe), np. *Muzyka organowa — kościelna*;

— cechy językowo-etniczne desygnatów pośrednich (określniki językowo-etniczne, np. *Folklor — polski*).

Dla kategorii określników ogólnych podkategoriami odwzorowującymi dalsze cechy relewantne interpretacyjnie są:

— aspekt prezentacji przedmiotu (tematu) dokumentu (określniki klasowe), np. *Językoznawstwo — historia*,

— cechy formalne prezentacji przedmiotu (określniki formalne), np. *Fizyka — encyklopedia*, *Literatura węgierska — bibliografia*.

W ramach kategorii określników lokalizujących wyróżnia się podkategoria określników geograficznych i chronologicznych, wskazujących miejsce desygnowanego przez temat obiektu w przestrzeni (*Elektrownie atomowe — Związek Radziecki*) lub w czasie (*Literatura polska — XIX wiek*). Kategoryzacje określników przyjmowane w poszczególnych jhp narodowych i/lub specjalistycznych

wykazują podobne zróżnicowanie kryteriów relewancji interpretacyjnej jak to ma miejsce w przypadku tematów.

3.3.1.2.1. Wiedza o rzeczywistości w kodach semantycznych. Wbudowanie rozległej wiedzy o rzeczywistości pozajęzykowej i jej powiązanie (niekiedy wręcz przemieszanie) z wiedzą o związkach między jednostkami języka w postaci hierarchii cech obiektów, jak ma to miejsce w UKD, można wyjaśnić odwołując się do funkcji realizowanych przez scharakteryzowane języki w systemie informacyjnym. Sosińska wyróżnia cztery funkcje: pierwotne — metainformacyjną oraz impresywną, znajdującą wykładniki formalne jedynie w systemach zautomatyzowanych, oraz wtórne — organizacji zbiorów i określania ich zakresów tematycznych [Sosińska B., 1987]. Tradycyjnie eksponowaną funkcją JIW jest funkcja metainformacyjna i związana z nią funkcja wyszukiwawcza.

W odróżnieniu od tradycyjnych języków, zwłaszcza klasyfikacji oraz jhp. pełniących funkcję metainformacyjną, a także wtórnie organizacyjną oraz funkcję określania zakresu tematycznego zbiorów, kody semantyczne budowane są do realizacji funkcji informacyjnej (opisowej, semantycznej, przedstawiającej, komunikatywnej, symbolicznej), właściwej językowi naturalnemu, a ponadto do realizacji funkcji metainformacyjnej (i wyszukiwawczej), analogicznie jak tzw. metajęzyki semantyczne stosowane w pracach lingwistycznych do opisu znaczeń wyrażen języka naturalnego [Apresjan J. D., 1971] [Bojar B., 1979] [Wierzbicka A., 1972]. Z opisowej (informacyjnej) funkcji kodów semantycznych wynika konieczność odwzorowania w ich polu semantycznym relewantnych cech obiektów rzeczywistości pozajęzykowej, składających się na konotację konstytutywną jednostek leksykalnych projektowanego języka.

Najbardziej znanymi JIW typu kodu semantycznego są: amerykański Kod Semantyczny W.Perry'ego i A.Kenta oraz ukraiński język *RX*-kodów E. F. Skorochocki [Perry W. J., Kent A., 1956], [Skorochocki E. F., Pšeničnaja L. E., Karjalajne I. N. 1968]. W Polsce prace nad budową kodu semantycznego prowadzili w latach siedemdziesiątych B. Bojar i O. A. Wojtasiewicz [Wojtasiewicz O. A., Bojar B., 1978]. Z funkcją informacyjną kodów semantycznych łączy się przyjęcie jako podstawowej zasady odwzorowania kryteriów relewancji w języku kategoryzacji obiektów desygnowanych przez jednostki leksykalne. Zakresy pól semantycznych poszczególnych kodów, a więc wielkość fragmentów rzeczywistości przez nie opisywanej, są zróżnicowane. I tak np. w kodzie Semantycznym SINTO przeprowadzono kategoryzację obiektów rzeczywistości należących do całego uniwersum wiedzy, natomiast pozostałe z wymienionych kodów były projektowane dla określonych dziedzin wiedzy: Kod Perry'ego i Kenta dla metalurgii, kod Skorochocki dla elektrotechniki, SYNTOL J. C. Gardina dla antropologii, psychologii, filozofii i etnografii [Gardin J. C., 1975]. Szczegółową charakterystykę kategorii semantycznych odwzorowujących cechy relewantne opisywanych obiektów przedstawiła Sosińska [Sosińska B., 1987], tu wskażemy

rozwiązania najbardziej reprezentatywne dla sposobu komunikowania wiedzy o rzeczywistości w językach omawianego typu.

3.3.1.2.1.1. Wiedza o rzeczywistości w kodzie semantycznym SINTO i jego wariantach. Kod semantyczny dla Systemu Informacji Naukowej Technicznej i Organizacyjnej (SINTO) oparty został na tych samych założeniach metodologicznych, co opis znaczeń polskich czasowników ruchu dyscyplin sportowych oraz operacji transportowania [Wojtasiewicz O. A., Bojar B., 1978]. Zrealizowana tam metoda opisu nazwana metodą pojęć elementarnych, polega na:

- ustaleniu zbioru elementarnych relewantnych cech dystynktywnych obiektów rzeczywistości pozajęzykowej,
- określeniu reguł odwzorowywania tych cech w pewnym metajęzyku oraz reguł posługiwania się wyrażeniami tego metajęzyka (reguł semantyki),
- określeniu reguł konstruowania jednoznacznych wyrażeń złożonych semantycznie, służących do opisu fragmentów rzeczywistości pozajęzykowej i/lub znaczenia wyrażeń języka naturalnego [Bojar B., 1979, 6.].

Ustalenie semantycznych dystynktywnych cech w kodzie SINTO opiera się na szczegółowej kategoryzacji jednostek leksykalnych, desygnujących następujące rodzaje elementów rzeczywistości:

- obiekty, wśród których wyróżniono trzy podkategorie obiektów prymarnych: *Informacja (I)*, *Energia (E)* oraz *Substancja (S)*,
- obiekty sekundarne, wśród których wyróżniono dwie podkategorie: *Instrument (Inst)* i *Miejsce (Loc)*, podlegające dalszej kategoryzacji;
- operacje na obiektach, wśród których wyróżniono podkategorie: *Transport (T)*, *Produkcja (P)*, *Konsumpcja (C)*. Są one traktowane jako funktory, których jednym argumentem musi być nazwa obiektu. Kategoria ta jest ściśle związana z kategoryzacją obiektów;
- operacje na operacjach (kategoria dotychczas nie rozwinięta). Założono, że będą to operacje, których argumentami są inne operacje, na przykład *blokowanie*, *hamowanie*, *rozpoczynanie*, *kończenie* i ich modyfikacje;
- charakterystyki obiektów, dla których wyróżniono dwie podkategorie: *Środowiska fizycznego* oraz *Skali zjawisk*, poddawane dalszej kategoryzacji;
- charakterystyki operacji, z dwiema podkategoriami jak dla charakterystyki obiektów;
- role semantyczne, na które składa się siedemnaście rodzajów wyrażeń służących do opisu funkcji sytuacyjnych obiektów i operacji: *agens*, *akcens*, *pacjens*, *aktant*, *surowiec*, *substancja pomocnicza*, *instrument*, *produkt*, *cel*, *rezultat*, *metoda działania*, *środowisko*, *okoliczność*, *miejsce* (lokalizacja stała, początkowa, końcowa, trasy), *czas* (daty, termin początkowy, końcowy, okres trwania), *miara*, *proces* [Bojar B., 1979] [Wojtasiewicz O. A., Bojar B., 1978].

Kategoryzacja semantyczna przyjęta w kodzie do opisu dyscyplin sportowych obejmuje cechy relewantne w opisie istoty zwycięstwa w poszczególnych dyscyplinach sportu, tworzące opozycje w ramach następujących ośmiu kategorii:

- rodzaj poruszającego się ciała (zawodnik - niezawodnik),
- instrument ruchu (brak, obecny przez cały czas trwania ruchu, obecny tylko w momencie początkowym ruchu),
- ośrodek ruchu (powierzchnia ciała stałego, woda, powietrze),
- grawitacja (nierелеwantna, ruch przy wykorzystaniu siły grawitacji, ruch w kierunku przeciwnym do działania siły grawitacji),
- kształt toru ruchu ciała (nierелеwantny, relewantny),
- dystans (relewantny, nierелеwantny),
- punkt docelowy ruchu (nierелеwantny, wyznaczone miejsce przestrzeni, odległość od wyznaczonego punktu docelowego),
- czas trwania ruchu (relewantny, nierелеwantny).

Przedstawioną kategoryzację zastosowano następnie do opracowania matrycy dystynktywnych cech dyscyplin sportowych, odwzorowującej równocześnie znaczenie terminów będących nazwami poszczególnych dyscyplin. Tablica może być rozbudowana przez wprowadzenie dodatkowych cech dystynktywnych, odwzorowujących własności sytuacji w zawodach — na przykład cechy *waga poruszającego się ciała* [Wojtasiewicz O. A., Bojar B., 1978., 224].

Rozbudowa cech semantycznie dystynktywnych w opisie sytuacji rzeczywistości pozajęzykowej, spójna z ogólną, ontologiczną kategoryzacją przyjętą w kodzie SINTO, została zastosowana w *opisie polskich czasowników ruchu*, a więc tych czasowników, które komunikują zmianę relacji przestrzennych jakiegoś ciała [Bojar B., 1979]. Wykładnikami kryteriów relewancji są następujące cechy:

1) informacje o poruszającym się ciele (szczegółowo wyspecyfikowany rodzaj, kształt, pozycja w momencie zakończenia ruchu, liczba poruszających się ciał);

2) informacje o sile sprawczej ruchu (lokalizacja siły, kontakt między ciałem poruszającym się a powodującym ruch, przemieszczenie się ciała powodującego ruch, ochrona przed spadaniem ciała zmieniającego lokalizację przez ciało powodujące tę zmianę, czas działania siły sprawczej);

3) informacje o przestrzeni ruchu (zmiana ośrodka w czasie trwania ruchu, jednoczesne ośrodki ruchu, kontakt między powierzchnią ruchu a powierzchnią poruszającego się ciała, ściana kontaktująca, rodzaj ośrodka, opór ośrodka, kształt toru, zagęszczenie przestrzeni ruchu);

4) informacje o rodzaju ruchu (zmiana lokalizacji w jego wyniku, zmiana zwrotu, wielkość zmiany, powtarzalność elementarnych ruchów składowych, częstotliwość zmian, lokalizacja osi obrotu);

5) informacje o kierunku ruchu, stosunek do wcześniejszej lokalizacji poruszającego się ciała;

6) informacje o prędkości ruchu (wartość prędkości, zmiana prędkości, harmonijność przebiegu zmiany prędkości);

7) informacje o zjawiskach akustycznych towarzyszących ruchowi (brak, obecność);

8) relacje między poruszającym się ciałem a innymi obiektami (obecność innych obiektów w przestrzeni ruchu, stosunek między kierunkami poruszających się ciał, kierunek ruchu ciała poruszającego się w stosunku do ciała współwystępującego w tej samej przestrzeni, komunikowanie lokalizacji poruszającego się ciała w momencie początkowym i końcowym ruchu, odległość poruszającego się ciała od lokalizatora w relewantnym momencie ruchu, zmiana odległości między ciałami leżącymi we wspólnej przestrzeni, kolizyjność toru ruchu, usytuowanie w momencie t poruszającego się ciała w stosunku do ciała, którego ruch jest implikowany, zakładanie usytuowania w momencie zakończenia ruchu;

9) możliwość określania czasownikiem przebiegu drogi;

10) możliwość określania czasownikiem relacji bycie początkowym punktem ruchu.

11) wartość ekspresywna czasownika określana na podstawie stosunku nadawcy informacji do poruszającego się ciała lub wynikająca ze stosunku nadawcy informacji do odbiorcy informacji.

Trzy ostatnie cechy mają charakter metajęzykowy. W pracach z zakresu teorii JIW, uwzględniających przedstawioną kategoryzację semantyczną (wykorzystaną częściowo w kodzie SINTO) raczej eksponuje się jej wartość z punktu widzenia informacji naukowej [Sosińska B., 1987]. Warto jednak zwrócić uwagę na to, iż scharakteryzowana metoda odwzorowania w języku sztucznym relewantnych cech rzeczywistości pozajęzykowej wykazuje uniwersalizm przy jej zastosowaniu metajęzykowym do opisu języka naturalnego, co stanowi o ewentualnej przydatności w systemach informacyjnych z dostępem w języku naturalnym. Ponadto omówiony metajęzyk semantyczny spełnia warunek produktywności, o czym świadczą zilustrowane możliwości rozbudowy ogólnej kategoryzacji wyjściowej. Zasadą komunikowania w języku cech relewantnych jest fasetowy charakter ich organizacji, co umożliwi niezależne kodowanie cech składających się na znaczenie terminu i dokładną ich specyfikację. Na przykład cechy dyscypliny sportowej „koszykówka” kodowane są w języku Bojar następująco: (*bOcaOOaO*)

— *b* przemieszczanie pewnego przedmiotu (*cecha rodzaj poruszającego się ciała — niezawodnik*),

— *O* bez posługiwania się dodatkowym instrumentem (*instrument ruchu*),

— *c* w powietrzu (*cecha ośrodek ruchu*),

— *a* wykorzystywana jest siła grawitacji (*cecha grawitacja*),

— *O* kształt toru ruchu jest nerelewantny (*cecha kształt toru ruchu*),

— *O* przemierzany dystans jest nerelewantny (*cecha dystans*),

— *a* przedmiot umieszczany w wyznaczonym punkcie (*cecha punkt docelowy ruchu*),

— *O* czas trwania ruchu nerelewantny, limitowany (*cecha czas trwania*).

Dla porównania, w równoważnym zakresowo symbolu UKD.(796.323) odwzorowano, uznane za relewantne, następujące cechy:

796 — gry sportowe, atletyka, gimnastyka

796.3 — gry z użyciem piłki

796.32 — piłka rzucona w górę

796.323 — koszykówka

Jak więc wynika z tego porównania, w UKD hierarchizacja odwzorowanych w jego wyrażeniach cech desygnowanych w rzeczywistości pozatekstowej elementów (sytuacji) przeprowadzana jest na każdym poziomie według różnych kryteriów, co powoduje łączne odwzorowanie w systemie paradygmatycznym wiedzy o relacjach językowych (por. 796 i 796.3) oraz relacjach między obiektami pozajęzykowymi (796.3, 796.32). Dowodzi to niejednorodności kryteriów relewancji w konstrukcji pola semantycznego tego języka.

3.3.1.1.4.2. Wiedza w Kodzie Semantycznym J. W. Perry'ego i A. Kenta. Kod semantyczny *ASM-WRU* (*American Society for Metals*), zbudowany w latach pięćdziesiątych przez W. Perry'ego i A. Kenta, różni się od scharakteryzowanego kodu SINTO zakresem pola semantycznego, ograniczonego w praktyce do metalurgii, oraz zakładaną funkcją — metainformacyjną. Zasady strukturalizacji wiedzy przyjęte dla tego języka pozwalają jednak umieścić go wśród języków przeznaczonych do uniwersalnej reprezentacji wiedzy [Perry J. W., Kent A., 1958].

W języku tym kryteria relewancji w odwzorowaniu wiedzy o obiektach rzeczywistości, które mogą być przedmiotem treści dokumentów, zostały wykorzystane do kategoryzacji elementarnych cech semantycznych, reprezentowanych przez zbiór 214 tzw. mnożników semantycznych. W planie wyrażenia cechy te reprezentowane są przez trzy znaki, ustalone według kryterium mnemoniki rdzeniowej wyrażeń języka angielskiego, np. *M-CH* (machine), *F-IH* (ryba), *M-TL* (metal). Miejsce wolne w ciągu znaków jest przeznaczone dla infiksów (wykładników relacji paradygmatycznych). Wyróżniono sześć kategorii semantycznych, poddawanych dalszemu podziałowi:

Pojęcia ogólne (General concepts):

- Idee (z teoriami)
- Stosunki ogólne
- Własności
- Pole badań
- Siły
- Klasyfikacje

Stosunki (Relationships):

- Ekonomiczne
- Fizyczne
- Społeczne

Stany (States):

- Psychologiczne
- Wyzwalające działanie
- Fizyczne

Procesy (Processes):

- Fizyczne
- Ogólne i główne

- Materialne
- Społeczne i przemysłowe
- Substancje (Substances):
- Ogólne
- Szczególne (organiczne, nieorganiczne)
- Preparaty
- Przedmioty (Objects):
- Ogólne (naturalne, artefakty)
- Specjalne

Elementarnymi jednostkami leksykalnymi o funkcji autosyntaktycznej są tzw. deskryptory proste będące wyrażeniami złożonymi z owych mnożników semantycznych, uzupełnionych jednym z wykładników dziesięciu relacji paradygmatacznych, tzw. relacji analitycznych, którymi są litery *A, E, I, O, U, Q, W, X, Y, Z*. Ich funkcją jest precyzowanie znaczenia przez wskazanie rodzaju związku między oznaczanym przedmiotem a kategorią, do której został zaliczony. Wykładniki te są wyodrębnione jako niezależna kategoria leksykalna nie podlegająca dalszemu podziałowi, wyznaczająca strukturę paradygmatyczną języka. A oto lista tych wykładników:

- A - być typowym reprezentantem,
- M-CH - być urządzeniem,
- MACH - być maszyną,
- E - być składnikiem,
- I - być częścią,
- O - być złożonym z kilku elementów danej klasy,
- U - być producentem,
- Q - być użytkownikiem czegoś, być zdeterminowanym przez coś,
- W - być środkiem, przedmiotem działania,
- X - nie posiadać jakiejś cechy,
- y - być istotną cechą obiektu,
- Z - posiadać pewne cechy, nie reprezentowane przez mnożnik.

Pozostałe cechy relewantne przedmiotu wyrażane są przez sufiksy cyfrowe dołączane do deskryptorów prostych, będące równocześnie numerami wyrażen związanych relacją generyczną z reprezentującym przedmiot deskryptorem, na przykład *FZSH.17 — delfin*. Sufiksy te są więc wykładnikami wiedzy o zakresach znaczeniowych wyrażen. Inną metodą zawężania znaczenia (precyzowania cech dystynktywnych przedmiotu) jest konkatencja deskryptorów prostych.

W kodzie ASM-WRU przeprowadzono także kategoryzację wyrażen reprezentujących relacje syntagmatyczne (syntetyczne) niezbędne w opisie pozajęzykowych sytuacji, w których występują desygnaty deskryptorów prostych lub w komunikowaniu aspektu ujęcia tych desygnatów w dokumencie. Pełniąc funkcję podobną jak określniki w jhp, są wykładnikami relewancji interpretacyjnej. Wyróżniono cztery klasy wykładników relacji syntetycznych, zależnie od ich łą-

czliwości syntaktycznej z deskryptorami należącymi do określonych kategorii ogólnych, a więc łączące się z:

- reprezentacjami materiałów,
- reprezentacjami procesów,
- reprezentacjami własności,
- reprezentacjami warunków.

W planie wyrażania są one reprezentowane przez trzyliterowe symbole, zaczynające się literą *K* i kończące kropką (*KOV.*, *KAP.*, *KAI.*). Ich zasób pierwotnie wynosił 24, następnie został ograniczony do 14. (Na nierozłączność klasyfikacji wykładników relacji syntetycznych, czyli syntagmatycznych, nazywanych także wskaźnikami roli, zwróciła uwagę Sosińska) [Sosińska B., 1987].

Tak więc, kod Perry'ego i Kenta jest językiem, w którym odwzorowywana wiedza strukturalizowana jest według kryteriów własności obiektów pozajęzykowych (kategorii), a więc *relewancki przedmiotowej i sytuacyjnej*, a także według kryteriów relewancki językowej.

3.3.1.2.1.2. Wiedza w języku *RX*-kodów. Zaletą kodów semantycznych jest umotywowanie semantyczne ich wyrażen, służące odwzorowaniu istotnych cech desygnatów, a co się z tym wiąże konotacyjna interpretacja znaczenia, gwarantująca w stopniu wyższym niż interpretacja denotacyjna precyzję kryteriów relewancki w systemie. Określenie kryteriów relewancki przy odwzorowywaniu w nich wiedzy pozajęzykowej jest punktem wyjścia do kategoryzacji wyrażen. Kategoryzację taką przeprowadzono także w języku *RX*-kodów, stosowanym w systemie BIT [Skorochod'ko E. F., Pšeničnaja L. E., Karjalajne I. N., 1968]. Kategorie ustalone dla tego języka obejmują:

- terminy (*X*) — wyrażenia desygnujące przedmioty i pojęcia, reprezentowane w formie wykazu / listy,
- relatemy (*R*) — wykładniki cech (przedmiotów) i relacji binarnych.

Podobnie jak w kodzie Perry'ego i Kenta, terminy desygnujące przedmioty i procesy różnią się stopniem złożoności semantycznej, co odwzorowane jest w planie wyrażania języka za pomocą indeksów dolnych i górnych (*X*) oraz w podziale na terminy podstawowe (elementarne), nazywane *termami*, oraz na *terminy złożone*. Odpowiednio, *relatemy* dzielone są na: *podstawowe* (pierwotne), których liczba wynosi 52, oraz *pochodne*. Obiekty pozajęzykowe, jako przedmioty treści dokumentów, reprezentowane są w charakterystykach wyszukiwawczych w formie ciągów wyrażen elementarnych lub pochodnych, na przykład:

$$R_{001}X^0_{306}R_{253}X^0_{305} = X^1_{168} \text{ — zwiększenie mocy,}$$

gdzie: X^0_{305} = moc, R_{001} = być elementem klasy (obligatoryjny relatem każdego terminu), X^0_{305} = zwiększenie, R_{253} = mieć obiekt.

W odróżnieniu od scharakteryzowanych wyżej kodów, w języku *RX* kryteria relewancki odnoszące się do wyodrębnienia obiektów pozajęzykowych są ustalone arbitralnie, bez dalszej kategoryzacji semantycznej (podział na pierwotne terminy i pochodne terminy uwzględnia jedynie stopień złożoności semantycznej). Ta sama uwaga dotyczy relatemów przeznaczonych do odwzorowania relacji sytu-

cyjnych (syntagmatycznych) zachodzących między obiektami reprezentowanymi przez terminy. Ograniczenie kategoryzacji desygnatów terminów w języku *RX-kodów* można wyjaśnić częściowo metainformacyjną funkcją tego języka (charakteryzowanie treści dokumentów), a ponadto budową systemu leksykalnego i przyjętą indukcyjną metodą gromadzenia słownictwa, która jednak nie gwarantuje jego kompletności.

3.3.1.3. Wiedza o rzeczywistości w specjalistycznych JIW

Wyróżnianie wśród języków informacyjno-wyszukiwawczych języków specjalistycznych i uniwersalnych jest zabiegiem dyskusyjnym, ukształtowanym przez pewną tradycję. Źródeł tego podziału należałoby przypuszczalnie upatrywać w sugerowaniu się zakresem dziedzinowym wiedzy reprezentowanej w słownikach poszczególnych języków. Tymczasem wydaje się, że o faktycznej uniwersalności (lub wyspecjalizowaniu) języka decyduje repertuar kategorii semantycznych (faset) organizujących plan treści języka i odwzorowujących przy różnym stopniu uniwersalności cechy dystynktywne obiektów/fragmentów rzeczywistości denotowanych przez wyrażenia języka. Tak więc, za języki reprezentujące wiedzę w sposób uniwersalny należałoby uznać kody semantyczne Bojar i Wojtasiewicza, w pewnym zakresie kod ASM-WRU Perry'ego i Kenta, Klasyfikację Dwukropkową Ranganathana, kategoryzację przedmiotów treści dokumentów Unguriana, SYNTOL Gardina. Natomiast Uniwersalna Klasyfikacja Dziesiętna wydaje się, wbrew pozorom, językiem specjalistycznym, o czym decydują różne kryteria podziału denotowanych pośrednio obiektów pozajęzykowych na poszczególnych poziomach klasyfikacji. Można stwierdzić, że to metoda ustalania cech dystynktywnych obiektów pozajęzykowych, nie zaś zakres pola tematycznego języka, decyduje o jego uniwersalności bądź wyspecjalizowaniu [Gardin J. C., de Grollier E., Levery F., 1964] [Sosińska B., 1985 19-36] [Ungurian O., 1982- 21-40].

3.3.1.3.1. Wiedza odwzorowana w Międzynarodowej Klasyfikacji Patentowej. Charakterystycznym przykładem ilustrującym powyższe stwierdzenie, iż o uniwersalności (wyspecjalizowaniu) JIW decyduje metoda ustalania kryteriów relewancji, a więc cech dystynktywnych desygnatów pośrednich jego wyrażen odzworowywanych w polu semantycznym, jest *Międzynarodowa Klasyfikacja Patentowa*, którą przyjęto uważać za język uniwersalny, gdyż jej zakres tematyczny obejmuje wszelkie dziedziny, w których możliwe jest dokonywanie wynalazków [Kwapisz A., 1981] [Moryk A., 1981].

Międzynarodowa Klasyfikacja Patentowa należy do języków informacyjno-wyszukiwawczych o stosunkowo długiej tradycji powstawania i rozwoju. Jej pierwsze projekty opracowano w latach 1904-1926. Decyzję o wznowieniu prac podjęto w 1951 r., ogólne zasady budowy klasyfikacji ustalono w 1952 r. Opracowanie I wersji tablic MPK, oficjalnie opublikowanych w 1968 r., rozpoczęto w drugiej połowie lat pięćdziesiątych. W Polsce MPK obowiązuje od 1.01.1969 r. i jest aktualizowana co pięć lat. Dotychczas opracowano cztery edycje MPK, ostatnią czwartą (INT CL 4) w 1985 roku (obowiązuje w Polsce od 1.01.1986 r.).

W materiałach metodycznych poświęconych zasadom stosowania MKP wymienione są równoległe różne funkcje przewidziane dla tego języka:

- systematyzacja dokumentów patentowych,
- obsługa informacyjna użytkowników (wyszukiwanie informacji),
- badanie stanu techniki w określonych dziedzinach,
- prowadzenie badań statystycznych w zakresie ochrony własności przemysłowej i intelektualnej [Kwapisz A., 1981].

Sformułowanie takie sugeruje, że główną funkcją tego języka jest funkcja metainformacyjna oraz, wtórnice, funkcja organizowania zbiorów informacji patentowej, z czym trudno się zgodzić. Symbole MPK są bowiem obligatoryjnym elementem opisu patentowego, a więc dokumentu przyznającego prawo do wyłącznego korzystania z wynalazku w celach zarobkowych, handlowych lub w produkcji przemysłowej. „Wynalazek” jest obiektem rzeczywistości denotowanym przez wyrażenie MKP. Stąd też język ten pełni funkcję analogiczną jak kody semantyczne przeznaczone do opisywania rzeczywistości pozajęzykowej, służąc równocześnie do organizacji zbiorów informacyjnych (opisów patentowych, abstraktów tych opisów).

Wynalazki stanowiące obiekty opisywane w tym języku podlegają kategoryzacji. Wyróżnia się trzy podstawowe kategorie wynalazków (przedmiotów):

Sposoby, na przykład: *fermentacja, polimeryzacja, rozdzielanie, kształtowanie, przemieszczanie, obróbka, przesyłanie i przekształcanie energii, budownictwo, sposoby obsługi maszyn, przetwarzanie i przekazywanie informacji*. Odwołując się do kategoryzacji przyjętej w kodzie SINTO możemy stwierdzić, że kategoria ta odpowiada wymienionym przez autorów kodu SINTO kategoriom operacje na obiektach i operacje na operacjach.

Urządzenia, na przykład: *instalacje związane z określonymi Sposobami, przyrządy, aparatura, maszyny, aparaty do urządzeń roboczych*. Kategoria ta odpowiada wyróżnionej w kodzie SINTO w ramach kategorii obiektów sekundarnych podkategorii *Instrumenty*.

Wytwory, na przykład: *związki chemiczne, kompozycje, wytwory procesów technologicznych*. Odpowiada to podkategorii *Substancja* w kategorii obiektów primarnych kodu SINTO.

Charakterystyka obiektów reprezentujących jedną z trzech wymienionych kategorii dokonywana jest według dwu kryteriów (cech dystynktywnych):

— według kryterium funkcji w rzeczywistości niezależnej od dziedziny zastosowania, na przykład:

- rozdzielanie, oddzielanie (symbol *B01D*)
- zawory, kurki, zawory kurkowe, urządzenia odpowietrzające (*F16K*)
- układy sterowania (*G05B*)
- według kryterium funkcji zależnej od dziedziny zastosowania, np. zastosowanie związku chemicznego jako nawozu lub detergentu.

Specyfikacja cech relewantnych przedmiotu wynalazku, określanych w materiałach metodycznych niezbyt precyzyjnie jako niezbędnych do wskazania „istoty wynalazku”, odwzorowana jest w strukturze paradygmatycznej języka.

Ogólny schemat zakresu pola semantycznego MPK wyznacza podział na 8 działów, odpowiadających pewnym sferom wiedzy technicznej, których część można sprowadzić do pewnych ogólniejszych kategorii semantycznych:

- A — Podstawowe potrzeby ludzkie
- B — Różne procesy przemysłowe. Transport
- C — Chemia i metalurgia
- D — Włókiennictwo i papiernictwo
- F — Budownictwo. Górnictwo. Konstrukcje zespolone
- F — Mechanika. Oświetlenie. Ogrzewanie. Uzbrojenie. Technika minerska
- G — Fizyka
- H — Elektrotechnika

Dalsza pionowa rozbudowa słownika MKP (tablic klasyfikacyjnych) odwzorowuje podział przedmiotów wynalazku na:

klasy (oznaczone w planie notacyjnym dwucyfrową liczbą, np. *Rolnictwo. Leśnictwo. Hodowla zwierząt. Lowiectwo. Zakładanie sidel. Rybołówstwo*);

podklasy, oznaczone w planie notacyjnym dużą literą alfabetu łacińskiego umieszczaną po symbolu klasy, np. *A01B — Uprawa gleby w rolnictwie, Części, elementy lub osprzęt maszyn i narzędzi rolniczych ogólnie*;

grupa główna, oznaczana w planie notacyjnym ciągiem cyfr rozdzielonych kreską ukośną (pierwsza część ciągu może zawierać od 1 do 3 cyfr, po kresce umieszczane są dwa zera), np. *A01B/00 — Narzędzia ręczne, A63F1/00 — gry w karty*;

podgrupa, oznaczana w planie notacyjnym dwoma cyframi umieszczonymi po kresce ukośnej w symbolu grupy głównej, różnymi od 00, np. *A01B3/01 — Pługi z nieruchomymi korpusami płuźnymi, A01B3/02 — z ludzką siłą pociągową, A01B3/04, ze zwierzęcą siłą pociągową*. Wykładnikami relacji paradygmatycznych na dalszych poziomach podziału podgrupy są kropki, których liczba oznacza poziom podziału (poprzedzają one odpowiednik słowny symbolu klasyfikacyjnego). Dopuszcza się sześć stopni podziału. Wyrażenia należące do podgrup organizowane są najczęściej za pomocą relacji generycznych i/lub me-reologicznych, na przykład:

A01B3/04 Pługi ze zwierzęcą siłą pociągową

A01B3/06 nienawrotne, bez możliwości tworzenia przyległej skiby w czasie

A01B3/06 Pługi bezkoleśne

A63H3/00 Lalki

A63H3/36 Elementy. Urządzenia i wyposażenie dodatkowe

A63H3/38 Oczy dla lalek

A63H3/40 Ruchome

W MKP wprowadzano także kategoryzację syntaktyczną wyrażeń (w III edycji) uzupełniając niektóre podklasy, np. *C12R*, wyrażeniami synsyntaktycznymi nazwanymi kodami indeksowymi. Tworzą one odrębną podklasę zawierającą wykaz mikroorganizmów uczestniczących w niektórych procesach biochemicznych, istotne składniki kompozycji lub mieszaniny, elementy konstrukcji (*B29K9 : 06*). Wyrażenia te wykazują znaczne podobieństwo ze wskaźnikami

roli w kodzie Perry'ego i Kenta, służą bowiem odwzorowaniu cech sytuacyjnych obiektów, reprezentowanych przez jednostki leksykalne (symbole klasyfikacyjne).

3.3.1.3.2. Wiedza o rzeczywistości w językach deskryptorowych. W literaturze z zakresu teorii języków informacyjno-wyszukiwawczych dość powszechny jest pogląd, że języki deskryptorowe należą do grupy JIW odwołujących się do denotacyjnej interpretacji znaczenia, ich jednostki leksykalne (deskryptory) są znakami arbitralnymi, zaś relacje między nimi odwzorowują relacje między zakresami znaczeniowymi, stąd też odwzorowują relacje między zbiorami dokumentów nie zaś między obiektami rzeczywistości i są konstruowane do wyszukiwania informacji o dokumentach [Sosińska B., 1981, 41-60]. Przyjmuje się ponadto, że w języku deskryptorowym wykładnikami wiedzy o strukturze rzeczywistości są:

- relacje mereologiczne oraz część relacji skojarzeniowych, z wyłączeniem tej ich części, która ustalana jest na podstawie częstości współwystępowania w tekstach [Bielicka L., Ścibor E., 1982],
- kategoryzacja leksyki (deskryptorów), wyznaczająca strukturę fasetową tezauryusa w jego części systematycznej [Bielicka L., Ścibor E., 1982] [Ungurian O., 1976].

Zagadnienie kategoryzacji semantycznej w językach deskryptorowych doczekało się obszernej literatury przedmiotu. Kategoryzacja jest efektywnym sposobem zapewnienia wielowymiarowej, polirelacyjnej struktury pola semantycznego języka. Przy projektowaniu języka informacyjno-wyszukiwawczego przeznaczonego dla konkretnego systemu specjalistycznego podejmuje się często ocenę możliwości wykorzystania kategoryzacji uniwersalnych lub pokrewnych specjalistycznych w konstruowanym języku [Dao Thi Qui, 1989] [Foskett D. J., 1981] [Hutchins W. J., 1978] [Ungurian O., 1982] [Wereszczyńska-Cisło B., 1984] [Ivankin V. J., 1973, 14-19]. Przeglądowi kategoryzacji w językach deskryptorowych projektowanych dla specjalistycznych SIW poświęcone są prace wielu autorów (Ungurian, Sosińska, Dao Thi Qui, Hutchins, Foskett), omawiające propozycje różnych zestawów kategorii semantycznych:

- dla nauki o glebie Vickery'ego,
- dla polskiego nazewnictwa naukowo-technicznego Unguriana,
- dla naukoznawstwa i polityki naukowej w systemie AWION [Artowicz E., 1978],
- dla dziedziny higieny i bezpieczeństwa pracy Fosketta [Foskett D. J., 1981].

dla przemysłu naftowego, fizyki ciała stałego, technologii żywności i in. [Sosińska B., 1987].

W dyskusjach nad optymalną liczbą i szczegółowością ustalanych kategorii podkreśla się, że ich dobór nie pozostaje w zależności wprost proporcjonalnej do zakresu tematycznego systemu i stopnia specjalizacji dziedziny wiedzy [Hoyle W. G., 1965, 1-6]. Według Bielickiej i Ścibora do najczęściej wyróżnianych kategorii w językach deskryptorowych należą: dziedziny i/lub gałęzie

gospodarki narodowej, przedmioty, osoby, procesy, zjawiska, zdarzenia, relacje, własności i wielkości, prawa, miejsce, czas [Bielicka L., Ścibor E., 1982].

Obecny rozwój SIW zdaje się podważać tradycyjny pogląd o ograniczeniu zastosowania języków deskryptorowych do systemów dokumentacyjnych z racji odwzorowywanych w nich relacji między zakresami znaczeniowymi.

Analizy JIW stosowanych w tzw. prostych systemach faktograficznych (niededukcyjnych) typu obiekt-cecha wykazują, że środkiem służącym do identyfikacji obiektów, ich atrybutów i wartości w SIW jest często język deskryptorowy ze słownikiem ustrukturalizowanym podobnie do tezausa oraz rozbudowanymi wykładnikami relacji syntagmatycznych (wskaźnikami roli i więzi) [Bielicka L., 1984].

W nielicznych pracach teoretycznych podejmujących zagadnienie struktury pola semantycznego języków dla systemów informacji faktograficznej wyróżnia się następujące klasy semantyczne jednostek leksykalnych, niezbędnych do opisu obiektów pozajęzykowych i wyszukania informacji o nich:

- nazwy obiektów,
- nazwy cech obiektów,
- nazwy wartości cech obiektów,
- nazwy grup cech obiektów [Bielicka L., 1984].

Semantyczną kategoryzację leksyki języka deskryptorowego w zakresie technologii preparatów enzymatycznych, przewidzianego zarówno dla systemów dokumentacyjnych, jak i faktograficznych, wprowadziła Dao Thi Quy [Dao Thi Qui, 1989]. W języku tym wyodrębniono trzy szerokie kategorie semantyczne niezbędne w opisie tematyki dokumentów:

- I – technologie
- II – badanie (własności, ocena wyników badań, ocena efektywności ekonomicznej)
- III – zastosowanie.

Kategoryzacja obiektów pozajęzykowych, będących przedmiotem treści dokumentów w dziedzinie technologii, obejmuje 14 kategorii:

- A** – Czynniki biochemiczne biosyntezy i aktywności enzymów,
- B** – Katalizatory
- C** – Obiekty ogólne
- D** – Odpady z produkcji preparatów enzymatycznych
- E** – Preparaty badawcze i użytkowe
- F** – Procesy i zjawiska
- G** – Metody
- H** – Producenci enzymów oraz ich składniki i wydzieliny
- I** – Substraty i produkty reakcji enzymatycznych
- J** – Surowce i materiały pomocnicze
- K** – Urządzenia (narzędzia, elementy urządzeń, aparaty)
- L** – Własności
- M** – Zastosowanie preparatów enzymatycznych
- O** – Cechy formalne dokumentu

Cechy wyróżnionych obiektów podzielono na dwie kategorie:

- własności (np. aktywność enzymatyczna, wartość *PH*, masa cząsteczkowa, punkt izoelektryczny, postać fizyczna itd.)
- zastosowania, a w niej dwie podkategorie:
 - wykorzystanie preparatów enzymatycznych *jako co* lub *w czym*,
 - produkty otrzymywane przy użyciu preparatów enzymatycznych.

Wśród przytoczonych kategorii i podkategorii niezależnie wyróżniono dwa rodzaje typów ich podziału:

- odwzorowującego relacje hierarchiczne,
- odwzorowującego niestale (sytuacyjne) własności obiektów ze względu na wykorzystanie *jako co* lub *w czym*.

Powyższy schemat kategoryzacji został zastosowany w budowie części systematycznej tezausa technologii preparatów enzymatycznych, w którym deskryptory usystematyzowano zgodnie ze wskazanymi zasadami podziału, przy wykorzystaniu relacji generycznych i mereologicznych. W podziale wykorzystano różne cechy obiektów wyznaczające przynależność deskryptorów do różnych faset, na przykład *ENZYMY* dzielono według:

- zdolności katalizowania określonych procesów chemicznych,
- zdolności do optymalnej aktywności w określonym *PH*,
- pochodzenia.

Zaprojektowany przez Dao Thi Quy JIW stanowi pewną propozycję modelową nie ograniczoną technologicznymi warunkami implementacji w konkretnym systemie informacyjnym, na przykład oprogramowaniem. Zakłada ponadto możliwość modyfikacji w zakresie zasobu leksykalnego tezausa i odwzorowanych w nim relacji między deskryptorami stosownie do potrzeb użytkowników systemu.

Przykładem języka deskryptorowego projektowanego dla systemu dokumentacyjno-faktograficznego implementowanego przy konkretnym oprogramowaniu (CDS-ISIS) oraz przy uwzględnieniu potrzeb konkretnych użytkowników jest język Systemu Informacyjnego o Metodach i Technikach Organizatorskich w Sterowaniu Jakością (SIMTOJ). System jest budowany przez zespół specjalistów w Instytucie Organizacji Systemów Produkcyjnych Politechniki Warszawskiej [Artowicz E., Szomański B., 1989, 103-124].

W systemie tym, zgodnie z życzeniem użytkowników, wyróżniono:

- blok informacji dotyczących dokumentów źródłowych, których przedmiotem są obiekty — Metody lub Techniki Organizatorskie (MTO)².
- blok informacji o cechach (atrybutach) obiektów „kryteriów identyfikacji MTO”.

Blok informacji dokumentacyjnej zawiera zwięzły opis obiektu (tzw. opis encyklopedyczny), uwzględniający cechy relewantne z punktu widzenia użyt-

² Analiza różnych definicji terminów metoda i technika wykazała znaczną dowolność i niekonsekwencje w ich stosowaniu przez różnych autorów. Niemożliwe okazało się ustalenie cech dystyngtywnych działań składających się na pewien „wzorzec postępowania ze względu na cel” nazywany metodą lub techniką [Pszczółowski T., 1978].

kownika sfery produkcyjnej, a zwłaszcza pracowników komórek sterowania jakością w przedsiębiorstwach przemysłu elektromaszynowego. W bloku tym wyróżniono następujące grupy pól opisu dokumentacji obiektu:

- nazwa obiektu (nazwa systemowa MTO)
- askryptory
- nazwy obiektów nadrzędne zakresowo
- nazwy obiektów podrzędne zakresowo
- nazwy obiektów związanych relacją mereologiczną
- nazwy obiektów związanych relacją asocjacyjną
- opis dokumentacyjny obiektu (tekst o objętości do 1650 znaków) zawierający wykaz działań składowych, sprzężeń między nimi, zasad oraz ewentualnie informację graficzną (schematy, wykresy)
- opis dokumentacji wyników wdrożenia MTO w konkretnym przedsiębiorstwie
 - nazwa przedsiębiorstwa
 - adres
 - realizator wdrożenia
 - opis dokumentacyjny wdrożenia
 - cechy formalne dokumentów źródłowych, których przedmiotem

jest MTO

- autor
- tytuł
- miejsce wydania
- wydawnictwo
- rok
- rodzaj dokumentu
- cechy identyfikacyjne autora charakterystyki wyszukiwawczej
- nazwisko i imię
- stopień naukowy
- tytuł zawodowy
- adres.

Deskryptory stanowiące nazwy obiektów są grupowane w Słowniku (docelowym teaurusie MTO) przy wykorzystaniu relacji generycznych i mereologicznych, istnieją bowiem MTO, które są częścią składową innych MTO, na przykład *ANALIZA TECHNOLOGICZNOŚCI* jest nadrzędna mereologicznie w stosunku do:

- *ANALIZA TECHNOLOGICZNOŚCI WYKONANIA, podobnie*
- *METODA TAKSONOMICZNA*
- *STANDARYZACJA CECH*
- *ANALIZA KORELACYJNA*
- *NORMOWANIE WSKAŹNIKÓW WAŻNOŚCI*

Cechy (atrybuty) pozajęzykowych obiektów systemu są odwzorowane w podzbiorze słownika JIW nazwanym kryteria identyfikacji MTO. Ich ustalenie

okazuje się zadaniem bardzo trudnym ze względu na złożoność obiektów systemu. W pierwszej wersji Słownika uwzględniono cechy MTO, relewantne z punktu widzenia użytkowników systemu, związane z:

- czynnikami procesu produkcji
- warunkami stosowania MTO.

Wśród cech MTO niezbędnych do określania ich przydatności ze względu na różnorodne czynniki i elementy procesów produkcji uwzględniono m.in.:

- podmiot (komórki przedsiębiorstwa),
- przedmiot, o którym uzyskuje się informację, np. informację o ludziach, narzędziach pracy itp.

- obszar stosowania MTO według sfer, faz i etapów produkcji:

- sfery przedprodukcyjnej,
- sfery produkcyjnej,
- sfery poprodukcyjnej,
- sfery pozaprodukcyjnej,
- funkcje realizowane za pomocą MTO w systemie sterowania jakością,
- skala produkcji,
- wielkość przedsiębiorstwa,
- problemy jakości wyrobów,
- etapy rozwiązywania problemów.

Spośród cech określających warunki użytkowania obiektu (MTO) uwzględniono:

- znajomość języka (naturalnego lub sztucznego),
- wykształcenie,
- rodzaj wykształcenia,
- doświadczenie zawodowe,
- koszty (orientacyjne),
- środki organizacyjno-techniczne,
- stopień trudności (skala trójstopniowa),
- ograniczenia stosowalności w poszczególnych działach przemysłu.

3.3.2. Reprezentacja wiedzy w gramatyce języków informacyjno-wyszukiwawczych

3.3.2.1. Reprezentacja wiedzy w gramatyce Uniwersalnej Klasyfikacji Dziesiątej

Kryteriami relewancji determinującymi strukturę paradygmatyczną UKD są — jak wskazaliśmy w rozdz. 3.3.1.1.1. — cechy obiektów rzeczywistości pozajęzykowej, należących do różnych układów. Świadczy o tym m.in. stwierdzana nieadekwatność odpowiednika słownego symbolu przy ustalaniu jego zakresu znaczeniowego. W materiałach metodycznych formuluje się zalecenie sprawdzania tego zakresu przez odwołanie się do znaczenia symbolu nadrzędnego, np. poprawne znaczenie symbolu 299.26 *Fenicjanie* można ustalić na podstawie

znaczenia symbolu nadrzędnego 299.2 *Religie Semitów (Religia Fenicjan)* (por. [Bielicka L. Ścibor E., 1982]).

Funkcje semantyczne wyrażen w tekstach UKD są dość złożone i niekonsekwentnie zróżnicowane ze względu na fakt, iż desygnaty wyrażen wyodrębnione są zarówno według kryteriów językowych, jak i pozajęzykowych. Scharakteryzowane w rozdz. 3.3.1.1.1. kategorie semantyczne wyrażen niesamodzielnych syntaktycznie (poddziałów wspólnych) służą do specyfikacji cech obiektów dokumentacyjnych: *formy, języka oraz pozajęzykowych przedmiotów treści dokumentów: miejsca i czasu wydarzeń, punktu widzenia (interpretacji), wykorzystania materiału, osób jako przedmiotów lub podmiotów działań w sytuacji charakteryzowanej w dokumencie*. Od 1980 r. dopuszcza się autosyntaktyczne użycie wyrażen należących do kategorii języka, formy, miejsca, grupy etnicznej oraz czasu, co świadczy o możliwości zmiany ról semantycznych argumentów relacji w tekście, zależnie od funkcji przypisywanej wyrażeniu reprezentującemu daną kategorię w sytuacji charakteryzowanej w dokumencie.

Kryteria relewancji odwzorowane w regułach konstruowania zdań UKD z elementów leksyki zawartej w tablicach głównych i pomocniczych możemy przedstawić w pewnym uproszczeniu, posługując się wymienioną typologią relacji E. Chmielewskiej-Gorczyicy.

Wykładnikami niespecyfikującymi relacji są symbole proste nie zawierające innych wyrażen oprócz jednostek leksykalnych autosyntaktycznych (symboli tablic głównych). Odwzorowana w ten sposób wiedza ogranicza się do wskazania jednego lub więcej obiektu denotowanego przez wyrażenie języka. Użycie w zdaniu (symbolu złożonym) symbolu łączącego + w funkcji wskaźnika więzi służy m.in. do zasygnalizowania istnienia związku sytuacyjnego między desygnatami wyrażen autosyntaktycznych, zajmujących różne miejsca w systemie paradygmatycznym języka, np. 54 + 66 *Chemia i przemysł spożywczy*.

Symbol „/” umożliwia scharakteryzowanie sytuacji, będącej przedmiotem treści dokumentu przez wskazanie uczestniczących w niej desygnatów wyrażen, należących do wspólnego obszaru systemu paradygmatycznego języka. Służy więc do łączenia symboli kolejnych, np. 548/549 *Krytalografia i mineralogia*.

Stosunkowo rozbudowany repertuar wykładników relacji pełni funkcje dwojakie: służy do odwzorowania cech sytuacyjnych obiektów, przedmiotów treści dokumentów oraz do tworzenia nazw obiektów nowo powstałych i/lub nie uwzględnionych dotychczas w zasobie leksykalnym.

Do odwzorowania cech sytuacyjnych obiektów służą wyrażenia synsyntaktyczne umieszczone w systemie paradygmatycznym w ramach podziałów wspólnych, wprowadzane za pomocą symboli:

- (1/9) *poddziały wspólne miejsca*
- (= ...) *poddziały wspólne grupy etnicznej*
- = ... *poddziały wspólne języka*
- „ ... ” *poddziały wspólne czasu*

- (O...) *podziały wspólne formy*
- .OO... *podziały wspólne punktu widzenia*
- O *podziały wspólne charakterystyki ogólnej:*
- O3 *materiału*
- O5 *osoby*

O tym, że symbole podziałów specjalnych (-1/-9, .01/.09, znak apostrof) są związane funkcjonalnie z sytuacyjną wiedzą pozajęzykową, świadczy ich akcydentalne wprowadzenie jedynie w niektórych działach UKD jako środka modyfikującego relacje hierarchiczne (zasady podziału) w pewnych określonych sytuacjach, na przykład wskazania jako cechy relewantnej partii politycznej jej statusu w państwie, podczas gdy cechą decydującą o miejscu partii w strukturze hierarchicznej są jej założenia polityczne, np:

- 329.11 *Założenia konserwatywne*
- 329.051 *Partie rządzące*
- 329.052 *Partie opozycyjne*
- 329.11.051 *Partie konserwatywne rządzące.*

Bielicka i Ścibor zauważają, że w ten sposób wprowadzany jest w języku element polihierarchii, co oznacza przyjęcie pewnej kategoryzacji semantycznej na podstawie kryteriów pozajęzykowych [Bielicka L., Ścibor E., 1982].

Analogiczne uwagi dotyczą tworzenia tzw. symboli syntetycznych za pomocą apostrofu przewidzianego dla niektórych działów, umożliwiającego charakteryzowanie obiektu poprzez wskazanie więcej niż jednej cechy, wynikającej z miejsca reprezentującego go symbolu w strukturze paradygmatycznej, np:

- 329.21 *Założenia monarchistyczne*
- 329.11'21 *Partie o założeniach konserwatywno-monarchistycznych lub*
- 666.113 *Klasyfikacja szkieł według tlenków składowych*
- 666.113.41'31'28 *Szkoło alkaliczno-wapniowo-krzemionkowe*

Znak apostrofu służy tu do wskazania pozajęzykowej relacji sytuacyjnej część — całość (obiekt — składniki).

O ile funkcja symboli podziałów analitycznych polega na modyfikacji struktury paradygmatycznej języka, o tyle kolejne wykładniki relacji specyfikujących (dwukropek, dwukropek podwójny, nawias kwadratowy) służą do odwzorowania relacji semantycznych na osi tekstu poprzez wskazanie, np.

- relacji oddziaływania pewnego obiektu na inny obiekt:
- 338:37.014 *Wpływ sytuacji gospodarczej na politykę oświatową*
- operacji dokonywanych na obiektach lub operacjach:
- 622.28::693.625 *Bielenie nakrapianiem obudowy kopalnianej*
- 684.4::658.5 *Organizacja produkcji w przemyśle meblarskim*

Różnicę między funkcjami semantycznymi dwukropka i dwukropka podwójnego określa się zazwyczaj w sposób uproszczony, przez sprowadzenie do możliwości inwersji symboli składowych w zbiorze informacyjnym. Wydaje się, że chodzi tu faktycznie o zmianę funkcji wyrażenia umieszczanego po

dwukropku podwójnym jako wykładnika pewnej sytuacyjnej cechy obiektu, reprezentowanego przez wyrażenie pierwsze, a więc pełniącego funkcję modyfikatora.

W sytuacjach złożonych z wielu obiektów, między którymi zachodzą różne relacje wymagające wyróżnienia, stosowany jest nawias kwadratowy, umożliwiający tworzenie grup tematycznych w zdaniu rozwiniętym UKD (symbolu złożonym), np.:

62[669.14]0.174 *Technologia stali, próby zginania*

620.1[669.14]74 *Badanie materiałów. Stal. Próby zginania*

620.17[669.14]4 *Badanie właściwości mechanicznych stali. Próby zginania* [Bielicka L., Ścibor E., 1982].

Zalecenia metodyczne UKD określające formułę zdaniową dla symboli rozwiniętych (symbol główny, symbole syntetyczne wprowadzane za pomocą apostrofu, symbole podziałów analitycznych, symbole podziałów wspólnych), reprezentują pewien układ wartościujący cechy sytuacyjne przedmiotu treści dokumentu. Cechy dokumentu (język i forma) zajmują w tym układzie miejsce peryferyjne.

Uzupełnieniem omówionych wykładników relacji syntagmatycznych były systemy relatorów: pierwszy D. Kerveganta i późniejszy I. Perraulta [Kervegant D., 1962] [Perrault I., 1965, 136-144] [Ungurian O., 1978], służące do specyfikacji związków zachodzących między obiektami rzeczywistości pozajęzykowej, odwzorowywanych w ChWD. D. Kervegant wyróżnił pięć rodzajów związków, oznaczanych za pomocą wskaźników zależności (> lub <):

- 1) przynależność (implikacje, właściwości, całość — część),
- 2) proces (czynności, współdziałanie, działanie, produkty itp.),
- 3) zależność (stosunek, więź, połączenie),
- 4) ukierunkowanie (aspekt, zastosowanie, wykorzystanie),
- 5) porównanie (analogia, równość, różnice).

System Perraulta, w którym wykładnikami relacji są kombinacje dziewięciu liter alfabetu łacińskiego od *a* do *i*, obejmuje podobne relacje sytuacyjne: część — całość (*dec — dea*), współrzędność lub współdziałanie (*ebgf*), klasa — odmiana (*ffdcffda*), zastosowanie (*eche*). Obydwa systemy potwierdzają pogląd, że zarówno organizacja paradygmatyczna jak i syntagmatyczna tego „uniwersalnego” języka oparte są faktycznie na specjalistycznej wiedzy pozajęzykowej, w której zakłada się typową dla języków klasyfikacyjnych hierarchizację cech desygnatów pośrednich wyrażen, powodującą arbitralne wbudowanie konotacyjnej definicji znaczenia symbolu językowego (por. [Sosińska B., 1981, 41-60]).

3.3.2.2. Odwzorowanie wiedzy w gramatyce Klasyfikacji Dwukropkowej Ranganathana.

Zasady odwzorowywania informacji w strukturach syntaktycznych Klasyfikacji Dwukropkowej określa w teorii klasyfikacji Ranganathana jego koncepcja analizy informacyjnej dokumentów (analizy fasetowo-fazowej), prowa-

dzącej do ustalenia dwuargumentowych relacji tekstowych (relacji fazowych), odwzorowujących powiązania między bytami stanowiącymi przedmiot treści dokumentu w sytuacji w nim charakteryzowanej [Ungurian O., 1975]. Związki te są komplementarne wobec stałych relacji paradygmatycznych, odwzorowanych w systemie języka w ramach przyjętej kategoryzacji ontologicznej bytów. Krytyczną charakterystykę koncepcji analizy fasetowo-fazowej przeprowadziła Sosińska, zwracając równocześnie uwagę na pewną jej niekonsekwencję, polegającą na porównywaniu trzech różnych planów językowych: treści i wyrażania JIW oraz planu wyrażania języka naturalnego [Sosińska B., 1983, 111-114]. Tu przytoczymy podstawowe założenia reprezentacji wiedzy w planie syntagmatycznym tego JIW.

Podstawowymi jednostkami planu treści tekstów Klasyfikacji Dwukropkowej są *klasy* (koincydencje bytów rzeczywistości pozajęzykowej), reprezentowane w planie werbalnym języka naturalnego przez *tematy* złożone z *izolat* (reprezentacji bytów jednostkowych). Indywidualny temat może tworzyć jedynie klasa, a więc pewien układ bytów, ewentualnie z izolatą jako elementem składowym. Układy bytów, będących przedmiotami treści dokumentów, są reprezentowane przez następujące rodzaje tematów:

- tematy główne, odpowiadające nazwom dziedzin wiedzy naukowej i praktycznej, a wśród nich tematy naczelne i nienaczelne (kanoniczne, specjalistyczne i systemowe). W połączeniu z izolatami (reprezentacjami bytów jednostkowych) tworzą cztery rodzaje układów, reprezentowanych przez:
- tematy główne proste, złożone z jednego tematu naczelnego i jednego z tematów nienaczelných;
- tematy główne rozwinięte, złożone z jednego tematu naczelnego i dwu lub więcej tematów nienaczelných;
- tematy rozwinięte, złożone z tematu głównego i jednej lub więcej izolat;
- tematy złożone, reprezentowane przez więcej niż jeden temat główny.

Między układami bytów zachodzą pewne *relacje sytuacyjne*, odwzorowywane za pomocą piętnastu rodzajów *relatorów*, nazwanych przez Ranganathana *relacjami fazowymi*. Funkcjonują one wewnątrz szeregów klasyfikacyjnych, łącząc izolaty wewnątrz faset oraz między niewspółrzednymi klasami jako wykładniki pięciu rodzajów relacji semantycznych.

Relacja nieokreślona jest relacją niespecyfikującą, równoważną funkcjonalnie wskaźnikom więzi, oznaczaną zależnie od poziomu symbolami t, j, a . Relacje porównania, różnicowania i wpływu, oznaczane odpowiednio — zależnie od poziomu — symbolami (v, m, c) , (w, n, d) , (y, r, g) są specyfikującymi relacjami pozajęzykowymi. Relacja ukierunkowania różni się od trzech wymienionych tym, iż nie zachodzi między bytami stanowiącymi przedmiot dokumentu, lecz między dokumentem (informacją) a jego intencjonalnym odbiorcą, np. *elementy astronomii dla podróżników*.

Gramatyka Klasyfikacji Dwukropkowej jest gramatyką pozycyjną, symbol relacji jest poprzedzany znakiem (ampersand), np.:

Różnice w inteligencji między ludnością wiejską i miejską

Symbol ,31 w5:72

- Socjologia

,31 — Socjologia mieszkańców wsi

w — relacja wewnątrzszeregową różnicowania

,31 w5:7 — Różnice w osobowości między ludnością wiejską i miejską

,31 w5:72 — Różnice w inteligencji między ludnością wiejską i miejską

Technologia paliw w przemyśle szklarskim

Symbol : F,55 b M6

F — Technologia

F5 — Technologia organiczna

F55 — Technologia paliwa

b — Relacja międzyklasowa ukierunkowania

M6 — przemysł szklarski

Wpływ buddyzmu na chrześcijaństwo

Symbol : Q,6 r,4

Q — Religia

Q,6 — Buddyzm

r — relacja wewnątrzszeregową wpływu

Q,4 — Chrześcijaństwo.

3.3.2.3. Odzworowanie wiedzy w gramatyce języka hasel przedmiotowych

Strukturę zdania języka hasła przedmiotowego określają reguły gramatyki pozycyjnej, przypisujące auto- i synsyntaktycznym jednostkom leksykalnym tego języka stałe miejsca. Dotyczy to w szczególności pozycji poszczególnych rodzajów okreśników, ponieważ zgodnie z regułą tej gramatyki temat reprezentuje określoną kategorię semantyczną (por. rozdz. 3.3.1.1.3.). Formuła zdaniowa języka hasel przedmiotowych, ustalona w polskiej tradycji katalogowania przedmiotowego przez Łysakowskiego, przewiduje następującą kolejność okreśników, reprezentujących scharakteryzowane wyżej kategorie semantyczne (za [Sadowska J., 1982]):

- 1) Okreśniki szczegółowe
 - a) jednostkowe
 - b) gatunkowe, językowe, specjalne
- 2) Okreśniki ogólne
 - a) klasowe i inne rzeczowe
 - b) formalne
- 3) Okreśniki lokalizujące
 - a) geograficzne
 - b) chronologiczne.

Jak zauważa J.Sadowska w swojej pracy poświęconej semantyczno-syntaktycznym aspektom okreśników w jhp, ustalony tradycyjnie szereg okreśników

w haśle przedmiotowym „staje się uciążliwy w SIW”, powodując wieloznaczność interpretacyjną haseł i rozdzielenie materiału jednorodnego [Sadowska J., 1982, 60], narusza więc zasady pragmatyki organizacji zbiorów informacji w systemie, co ilustruje następujący przykład:

Kobieta — socjologia — Polska

Kobieta — socjologia — podręcznik — Polska.

Autorka sugeruje celowość restrukturyzacji tradycyjnej formuły zdaniowej, polegającej na przeniesieniu określników formalnych na ostatnie miejsce i utworzenie dwu skupień informacyjnych dotyczących treści dokumentu i jego formy. Zasygnalizowany w ten sposób problem polega, jak się wydaje, na przemieszaniu w tradycyjnej formule zdaniowej jhp wiedzy o rzeczywistości dokumentacyjnej (określniki formalne, częściowo językowe) oraz o rzeczywistości pozadokumentacyjnej, co staje się źródłem trudności przy selekcjonowaniu informacji w systemie według językowych lub pozajęzykowych kryteriów relewancji. Przeprowadzona przez J.Sadowską analiza określników jhp wykazuje przewagę kryteriów relewancji pozajęzykowej, odwzorowywanej w strukturze hasła przedmiotowego, co widoczne jest w przypadku określników jednostkowych, np. hasło: *WARSZAWA — Uniwersytet — Biblioteka — Zakład Rycin* zawiera określnik geograficzny wbudowany w temat (*Warszawa*).

Znaczenie wiedzy pozajęzykowej w rozstrzygnięciu o szyku określników, i tym samym o strukturalizacji zbioru informacyjnego systemu, ilustrują przykłady haseł zawierających określniki tego samego rodzaju (np. rzeczowe lub inne klasowe):

Koszty — budownictwo — rachunkowość

Koszty — gospodarka — rachunkowość

Koszty — rachunkowość — budownictwo

Koszty — rachunkowość — przemysł

Koszty — transport — rachunkowość,

gdzie pozycja określnika wyznacza kryteria relewancji związane ze strukturą gospodarki lub z pewnym aspektem tematu (interpretacją).

Analiza tzw. łączliwości określników (jednostkowych, klasowych) również wykazuje ich bezpośrednią zależność od tematu lub ich wzajemną zależność wynikającą z tzw. znajomości realiów, np.: *Warszawa — Biblioteka Narodowa — Instytut Bibliograficzny — Zakład Dokumentacji Księgoznawczej.*

Świadczy to o tym, że kryteria relewancji na osi syntagmatycznej jhp w tradycyjnych katalogach przedmiotowych odwołują się przede wszystkim do wiedzy pozajęzykowej, rozstrzygającej o szyku określników w haśle przedmiotowym. Język ten wykazuje wiele analogii z językami klasyfikacyjnymi, zwłaszcza z poddziałami wspólnymi UKD oraz częścią izolat Klasyfikacji Dwukropkowej Ranganathana (por. [Sadowska J., 1982, 50-53]).

3.3.2.4. Wiedza w gramatyce Międzynarodowej Klasyfikacji Patentowej

Kryteria relewancji przyjęte przy projektowaniu struktury paradygmatycznej Międzynarodowej Klasyfikacji Patentowej dotyczą — jak wskazywaliśmy w rozdz. 3.3.1.2.1. — cech nowo powstałych obiektów rzeczywistości pozajęzykowej. W zasadach opisu tych obiektów założono wartościowanie i systematyzację cech w celu odróżnienia tego ich podzbioru, który stanowi przedmiot ochrony patentowej, od cech nie objętych zastrzeżeniem patentowym, uzupełniających informację o danym rozwiązaniu. Wykładnikiem rozgraniczającym tzw. informację o wynalazku (relewantną) od informacji dodatkowej (nierelewantnej) jest znak // (podwójna kreska ukośna).

Charakterystyka wyszukiwawcza obiektu chronionego prawem patentowym (sposobu, wytworu lub urządzenia) powinna odwzorowywać cechy składające się na jego funkcje:

— niezależnie od dziedziny zastosowań, np.

BO1D Rozdzielanie i oddzielanie

F16K Zawory, Kurki, Zawory kurkowe, Urządzenia odpowietrzające

GO5B Układy sterowania

HO3K Technika impulsowa

CO7 Chemia organiczna — sposób wytwarzania

— związane z konkretnym zastosowaniem, np.: zastosowanie związku chemicznego jako nawozu lub detergentu, bądź też w medycynie lub w procesach przemysłowych:

A61K47/00, gdzie

A61 Medycyna lub weterynaria; Higiena

K preparaty do celów farmaceutycznych, dentystycznych lub kosmicznych 47/00 preparaty znamienne składnikami nieczynnymi.

B29035/00, gdzie

B29 Procesy przemysłowe — Mechaniczna obróbka metali w zasadzie bez ubytku materiału

C Wytwarzanie blach, drutu, prętów metalowych inaczej niż przez walcowanie

35/00 Odbieranie wyrobów lub usuwanie resztek materiału z pras wyciskowych.

Cechy obiektu związane z jego zastosowaniem mogą należeć do informacji dodatkowej nie objętej ochroną patentową.

Każdy obiekt charakteryzowany jest w tym języku przez wskazanie wszystkich jego cech relewantnych. Jeśli cechy te odwzorowane są w słowniku na poziomie grup, wymagana jest ich specyfikacja przy uwzględnieniu dopowiedzeń i odsyłaczy wbudowanych w strukturę paradygmatyczną języka.

Gramatyka MPK przewiduje proste i rozwinięte formuły zdaniowe. Zdanie proste składa się z jednego symbolu działu, klasy, podklasy, a także ewentualnie

grupy, podgrupy, np.: *A62D9'00 Zastosowanie w pożarnictwie kompozycji substancji chemicznych w aparatach do oddychania.*

W przypadku, gdy cechy obiektu tworzą kombinację będącą podstawą umieszczenia rozwiązania w różnych grupach współrzędnych na osi paradygmatycznej (na przykład „sposoby wielostopniowe”, „urządzenia do mieszania znamienne częściami składowymi”), zalecane jest indeksowanie symbolem reprezentującym wyższy poziom hierarchii, o ile nie rozstrzyga inaczej „reguła pierwszeństwa”. Gdy cechy obiektu uwzględniane są w kilku różnych grupach głównych, obowiązuje rozwinięta formuła zdaniowa, złożona z dwu lub więcej symboli.

Zdania rozwinięte mogą zawierać następujące struktury:

1. Symbole działu, klasy, podklasy, grup i podgrup oddzielone przecinkami, przy uwzględnieniu ewentualnej „reguły pierwszeństwa”, która wyznacza pozycję wyrażenia o wyższym stopniu relewancji, na przykład spośród symboli:

EOIB27/04 Maszyny służące do wyrównywania podsypu

EOIB27/02 Wyrównywanie podsypu — ma pierwszeństwo,

a więc poprawne zdanie ma postać: *EOIB27/02,27/04;*

2. Symbole działu, klasy, podklasy, grup i podgrup połączone z symbolami synsyntaktycznymi, tzw. kodami indeksowymi, wprowadzonymi jako wykładniki faset w niektórych klasach w celu wyróżnienia elementów niektórych procesów, np. mikroorganizmów, składników kompozycji lub mieszanin, elementów metod lub konstrukcji. Wykładnikiem tej relacji syntagmatycznej wyróżniony składnik obiektu, jest dwukropek, np.: *B29K23 : 06.*

Kod indeksowy może być dołączany do wyrażen składowych zdania rozwiniętego typu (1) przez umieszczenie po symbolu grupy głównej lub podgrupy, np. *CO87210/16, 214 : 06.* Jest on interpretowany wówczas jako wykładnik składnika desygnatów obu wyrażen, w celu wskazania wyrażenia, z którym jest związany wprowadzany jest nawias pojedynczy ;

3. Formuła zdaniowa składa się z dwu części, rozdzielonych znakiem //, w pozycji pierwszej, przed kreskami, odwzorowywane są cechy obiektu chronione patentem, przedstawione w postaci zdania prostego lub rozwiniętego, po kresce umieszczone są informacje dodatkowe, a więc informacje o cechach nie podlegających ochronie patentowej, np.

CO8F210/16,255/04 // A6IK47/00,B29035/00,CO9J3/14 — Kopolimery etylenu z alfaoleinami szczepionymi na kopolimerach etyleno-propylenowych,

gdzie:

C Chemia, Metalurgia: chemia, wodór, wodorki, gaz

CO8 Organiczne związki wielocząsteczkowe; ich wytwarzanie lub obróbka chemiczna; Mieszanki na ich podstawie

CO8F Związki wielocząsteczkowe utrzymywane w reakcjach, w których biorą udział tylko wiązania nienasycone węgiel — węgiel

CO8F210 Kopolimery nienasyconych węglowodorów alifatycznych, zawierających tylko jedno wiązanie podwójne węgiel — węgiel

CO8F255 Związki wielocząsteczkowe otrzymywane przez polimeryzację monomerów na polimerach węglowodorów, określonych w grupie 10/00. Homopolimery lub kopopolimery nienasyconych węglowodorów alifatycznych, zawierających tylko jedno wiązanie podwójne węgiel — węgiel

A Podstawowe potrzeby ludzkie

A61 Medycyna lub weterynaria. Higiena

A6IK Preparaty do celów farmaceutycznych, dentystycznych lub kosmicznych

A6IK47/00 Preparaty medyczne znamienne składnikami nieczynnymi

B Różne procesy przemysłowe

B29 Mechaniczna obróbka metali zasadniczo bez ubytku materiału

B29C Wytwarzanie blach, drutu, prętów metalowych inaczej niż przez walcowanie

B29C35/00 Odbieranie materiałów lub usuwanie resztek materiału z pras wyciskowych

C Chemia, metalurgia

CO9 Barwniki, farby, środki nadające połysk, żywice naturalne, środki klejące, mieszaniny różnego rodzaju

CO9J Zastosowanie jako środków klejących materiałów innych niż klej

CO9J3/4 Środki klejące inne niż klej ze składników organicznych.

W sytuacji, gdy w słowniku (tablicach) MPK brak jest wyrażenia odwzorującego cechy nowo powstałego obiektu, stosowany jest symbol *X*, dołączany do symbolu podklasy, grupy lub podgrupy. Charakterystyki wyszukiwawcze obiektów, zawierające ten symbol, stanowią podstawę aktualizacji tablic MPK.

Możliwe związki sytuacyjne obiektu odwzorowywane są w przypadku, gdy do jego cech relewantnych należą cechy określające dziedzinę zastosowania wynalazku, niezależnie od tego, czy podlega ochronie prawnej. Dziedzina ta jest również charakteryzowana za pomocą wyrażen specyfikujących obiekty, w stosunku do których przedmiot wynalazku może pełnić określone funkcje (składnika, narzędzia itp.)

W rozdziale poświęconym organizacji paradygmatycznej (słownika) JIW wskazywaliśmy, że MPK pełni faktycznie funkcje kodu semantycznego, służącego do opisu obiektów rzeczywistości pozajęzykowej, w którym przyjęte kryteria relewancji odwołują się do własności strukturalnych lub funkcjonalnych tych obiektów. Pogląd ten potwierdza analiza struktury syntagmatycznej języka, podporządkowanej zasadzie hierarchizacji najbardziej relewantnych cech obiektów zgodnie z regułami gramatyki pozycyjnej, dokonywanej w niektórych podpolach przez wyróżnienie wykładników istotnych elementów procesów lub konstrukcji za pomocą wskaźników, pełniących funkcje podobne do wskaźników roli.

3.3.2.5. Odwzorowanie wiedzy w gramatyce języków deskryptorowych

Metodom reprezentacji związków syntagmatycznych w językach o słownictwie paranaturalnym, a zwłaszcza w językach deskryptorowych, poświęcona jest obszerna literatura przedmiotu. Zagadnienie specyfikacji związków semantycznych między wyrażeniami współwyznaczającymi sytuację odwzorowaną w dokumencie pierwotnym w strukturze powierzchniowej tekstów ChWD absorbowoło uwagę teoretyków i praktyków JIW w latach sześćdziesiątych i siedemdziesiątych, a więc w okresie, kiedy większość systemów informacyjnych funkcjonowała w trybie wsadowym, zaś powszechnie stosowane metody indeksowania współrzędnego nie gwarantowały oczekiwanej (intuicyjnie) przez użytkowników sprawności wyszukiwania informacji, prowadząc do powstawania tzw. fałszywej koordynacji wyrażen. W systemach informacji online obserwuje się — jak postaramy się wykazać w rozdziale następnym — tendencję do ograniczania formalnych wykładników relacji semantycznych na osi tekstu lub do zupełnej z nich rezygnacji, co rekompensuje rozbudowa operatorów tekstowych wyspecjalizowanych we wskazywaniu warunków, jakie powinny spełniać na płaszczyźnie powierzchniowej wyrażenia relewantne JIW.

Zainteresowanie teoretyków JIW związkami wyrażen w tekstach języków deskryptorowych podyktowane jest różnymi względami:

- istnieniem ścisłej zależności między strukturą pola semantycznego tezaury (zasobem leksyki, strukturą paradygmatyczną, metodami prezentacji znaczenia) oraz gramatyką języka deskryptorowego [Chmielewska-Gorczyca E., 1983, 141-151];
- koniecznością specyfikacji relacji syntagmatycznych w systemach specjalistycznej informacji dokumentacyjnej, a także w systemach faktograficznych [Dao Thi Qui, 1989];
- dążeniem do rozszerzenia repertuaru tzw. kluczy wyszukiwawczych w SIW drogą takiej transformacji tekstów ChWD, aby możliwe było wykorzystanie w procesie wyszukiwania każdego znaczącego wyrażenia składowego bez zmiany znaczeń wyrażen złożonych. Doprowadziło to do zainteresowania tzw. relacjami fazowymi [Calkins M. L., 1980, 53-67].

Zakres wiedzy odwzorowywanej na osi paradygmatycznej języka deskryptorowego pozostaje — jak wskazywaliśmy wyżej za autorami szczegółowo rozważającymi omawiany problem — w ścisłej zależności z organizacją syntagmatyczną. Stosowana w językach deskryptorowych kategoryzacja wyrażen obejmuje niekiedy kategorię wyrażen synsyntaktycznych, a więc modyfikatorów (deskryptorów aspektowych, deskryptorów uściślających), będących wykładnikami wiedzy sytuacyjnej o obiekcie desygnowanym przez deskryptory auto-syntaktyczne (naczelne, podstawowe, zasadnicze) [Artowicz E., 1978, 35-56] [Bielicka L., Ścibor E., 1982]. Do odwzorowania wiedzy o dokumencie służy kategoria tzw. deskryptorów formalnych [Bielicka L., Ścibor E., 1982].

W charakterystykach wyszukiwawczych dokumentów (ChWD), powstałych w wyniku indeksowania współrzędnego, prezentowana wiedza sprowadza się do podania nazw obiektów desygnowanych pośrednio i bezpośrednio. Jako odmiana indeksowania współrzędnego traktowane jest indeksowanie hierarchiczne, polegające na wprowadzaniu do ChWD dodatkowo wyrażen JIW pozostających w relacji nadrzędności w stosunku do tych, które zostały wytypowane na podstawie tekstu indeksowanego dokumentu.

Wskaźniki więzi są wykładnikami relacji niespecyfikujących, pełnią więc przede wszystkim funkcję komunikowania wiedzy o języku poprzez wskazanie, które z wyrażen ChWD pozostają w pewnej relacji. Wśród autorów analizujących wpływ wskaźników więzi na zwiększenie sprawności wyszukiwania obserwuje się rozbieżne opinie na temat celowości stosowania tego sposobu wyrażania kryteriów relewancji językowej [*The „Mixed Term-links”*, 1975, 288-293] i eliminowania fałszywej interpretacji wyrażen złożonych. Zarówno oceny negatywne jak i pozytywne poparte są często wynikami wyszukiwania w konkretnych SIW (np. w systemie INIS w byłej NRD). O ograniczonej przydatności decyduje niemożność specyfikacji relacji między wyrażeniami w tekście, a więc niewielka przydatność w prezentacji wiedzy o pozajęzykowych sytuacjach przedstawionych w ChWD.

Wskaźniki roli — jak stwierdziła E.Chmielewska-Gorczyca — są funkcjonalnie podobne w prezentacji wiedzy do kategorii semantycznych organizujących pole semantyczne języka deskryptorowego na osi paradygmatycznej, służą bowiem określeniu funkcji desygnatu wyrażenia językowego w opisywanej sytuacji. Tradycyjnie traktowane są jako wykładniki wiedzy pozajęzykowej w tekstach języków deskryptorowych, co nie jest słuszne zważywszy ich zastosowanie np. w kodach semantycznych. Budzą one wiele kontrowersji co najmniej z dwu powodów [Robowski J., 1974]:

- wymagania wysokich kwalifikacji lingwistycznych od indeksatorów;
- trudności z opracowaniem ich uniwersalnego zestawu przydatnego w językach o różnych zakresach tematycznych, które to trudności należałoby uznać za wręcz „oczywiste”, ponieważ zrealizowanie takiego zadania byłoby równoznaczne ze wskazaniem uniwersalnego zbioru funkcji sytuacyjnych dla wszelkich obiektów rzeczywistości pozajęzykowej.

Zainteresowanie wskaźnikami roli jako wykładnikami kryteriów relewancji w ChWD jest charakterystyczne dla rozwoju SIW w latach sześćdziesiątych, a więc w okresie, kiedy warunki dialogu użytkownika z SIW zmuszały projektantów systemów do poszukiwania uniwersalnych, arbitralnych metod prezentacji wiedzy w systemie [Wereszczyńska-Cisło B., 1984]. Późniejszy rozwój systemów informacji dokumentacyjnej dowiódł celowości ich stosowania w systemach specjalistycznych, zwłaszcza w zakresie chemii, metalurgii itp. [Bielińska L., 1985].

Szeroką interpretację wskaźników roli przedstawiła W. Ogórkiewicz w pracy doktorskiej poświęconej analizie wpływu tego rodzaju wykładników relacji

syntagmatycznych na wyszukiwanie informacji w systemie w dziedzinie technologii żywności [Ogórkiewicz, W., 1985, 71-94]. Autorka zdefiniowała wskaźniki roli jako „Wykładniki sytuacyjnej funkcji pełnionej przez dany obiekt w określonym kontekście językowym lub w określonej sytuacji rzeczywistości pozajęzykowej”, wyróżniając trzy grupy:

- 1) pozajęzykowo-tematyczne, służące określeniu kategorii tematycznej dokumentów,
- 2) językowo-tekstowe, służące określeniu funkcji wyrażen w kontekście,
- 3) pozajęzykowo-tekstowe, służące określeniu funkcji sytuacyjnych desygnatów terminów, będących słowami kluczowymi dokumentów.

W tym ujęciu pełnią one funkcje prezentowania zarówno wiedzy językowej (wskaźniki językowe-tekstowe) jak i pozajęzykowej (określenie kategorii tematycznych dokumentu oraz roli desygnatu wyrażenia w rzeczywistości pozajęzykowej), co pozwala zakładać ich przypuszczalną przydatność w systemach informacji dokumentacyjnej i faktograficznej. Potwierdzają to przypuszczenie wyniki pracy nad projektem JIW systemu dokumentacyjno-faktograficznego w dziedzinie preparatów enzymatycznych przedstawione przez Dao Thi Qui [Dao Thi Qui, 1989].

Wskaźniki wagi (wagi, współczynniki wagowe) są de facto explicite wyrażoną systemową oceną relewancji wyrażen języka deskryptorowego przez przypisanie tym wyrażeniom pewnych symboli pomocniczych, odzwierciedlających ich przydatność w pełnieniu funkcji wyszukiwawczej. Najprostszy sposób ważenia deskryptorów ze względu na ich zakładaną relewancję polegał na wprowadzeniu skali binarnej oceny, np. 0 mniejszy stopień relewancji, brak wskaźnika — wyższy stopień relewancji [*New Kind of „Weight-relators”*, 1975, 294-296]. Najczęściej stosowane są skale trój-, cztero- oraz ośmiostopniowe [Bielicka L., 1985].

Ocena przydatności wag jako wykładników relewancji dyskutowana jest w obszernej literaturze przedmiotu [Chmielewska-Gorczyca E., 1981, 22-26], głównie ze względu na subiektywność kryteriów relewancji przyjmowanych przez indeksatorów przy tzw. metodach manualnych. W ostatnich latach zainteresowanie wykorzystaniem wag wzrosło w związku z wykorzystaniem wyników badań statystycznych, dotyczących częstości występowania w dokumentach wyrażen równokształtnych z deskryptorami lub ich ekwiwalentów wyszukiwawczych. W polskiej literaturze z zakresu JIW zagadnienie to podejmował K. Choroś [Choroś K., 1981, 22-26]. Oprócz metod statystycznych wykorzystywany jest także aparat probabilistyki w powiązaniu z oceną potrzeb użytkowników (por. probabilistyczne ujęcie relewancji Marona i Kuhna — rozdz. 2.5.1.).

Prezentowane w literaturze jako odrębny sposób odwzorowania relacji syntagmatycznych w językach deskryptorowych *indeksowanie ankietowe* (fasetowe, pozycyjne, rastrowe) polega właściwie na wykorzystaniu wiedzy o cechach obiektów pozajęzykowych (kategorii semantycznych) odwzorowanych w strukturze paradygmatycznej języka oraz reguł gramatyki pozycyjnej do tworzenia

ChWD. np. przez wymienienie w formularzu do indeksowania takich cech obiektów, jak dziedzina, do której należy badany obiekt, nazwa obiektu, metody badania obiektu, instrumenty badania obiektu, realizator badań, miejsce i czas badania, cel badania, modyfikatory. Zastosowanie ankietowej metody prezentacji wiedzy o obiektach systemów dokumentacyjnych traktowane uniwersalistycznie nie pozwoliło osiągnąć zakładanej sprawności wyszukiwania informacji [Ivankin V. J., 1973, 14-19]. Taki sposób prezentacji wiedzy w tekstach JIW staje się popularny w specjalistycznych systemach informacji faktograficzno-dokumentacyjnej, ponieważ umożliwia jednoznaczne wskazanie cech obiektów pozajęzykowych w ramach przyjętej formuły zdaniowej z równoczesnym wskazaniem dokumentów opisujących te obiekty.

3.3.2.6. Wiedza w gramatyce kodów semantycznych

Przykłady kodów semantycznych, scharakteryzowane w rozdz. 3.3.1.1.4. z punktu widzenia struktury paradygmatycznej, prezentowane są w literaturze jako języki z rozbudowanymi środkami wyrażania relacji syntagmatycznych [Chmielewska-Gorczyca E., 1983, 144-151] [Czerny A. I., 1978]. Przyjęty w literaturze dotyczącej języków informacyjno-wyszukiwawczych sposób prezentacji formy wykładników ról semantycznych jednostek leksykalnych w tekstach sprawia wrażenie, iż każdy z tych języków posiada odrębny, niepowtarzalny repertuar środków wyrażania wiedzy o funkcji desygnatów jednostek leksykalnych w opisywanych sytuacjach.

Z punktu widzenia semantycznej specyfikacji relacji najbardziej szczegółowe ujęcie reprezentuje kod SINTO, w którym przewidziano siedemnaście rodzajów ról semantycznych, jakie mogą pełnić obiekty rzeczywistości pozajęzykowej. Role te komunikowane są za pomocą reguł gramatyki pozycyjnej.

W kodzie ASM-WRU relacje syntagmatyczne (stosunki syntetyczne) są reprezentowane przez czternaście rodzajów związków odwzorowywanych za pomocą wskaźników roli, trzyliterowych kodów zaczynających się literą *K* i kończących kropką [Perry J. W., Kent., A., 1958].

W języku *SYNTOL* przewidziano cztery typy relacji, służących do odwzorowania związków zachodzących między wyrażeniami składowymi syntagm: koordynacji, konsekcji, przyczyna — skutek, asocjacji oraz predykcji.

W języku *RX-kodów* wykładnikami relacji między obiektami reprezentowanymi przez terminy są relatywy, których zestaw obejmuje 52 jednostki, z uwzględnieniem cechy ukierunkowania relacji oraz jego negacji. Mając na uwadze wielokrotne omawianie tego zagadnienia w literaturze przedmiotu, możemy stwierdzić, że wykładnikami wiedzy na osi syntagmatycznej kodów semantycznych są wskaźniki roli, których zestaw zasadzie można sprowadzić do wykazu tzw. przypadków głębokich Fillmore'a (por rozdz. 4.3.1), najbliższego repertuariowi ról semantycznych w kodzie SINTO.

Znamiona kodów semantycznych, ze względu na specyfikację związków sytuacyjnych między obiektami w tekstach JIW, noszą także: indeksowanie relacyjne Farradane'a oraz tzw. relacje fazowe A. C. Fosketa, reprezentujące różne

pojęcia metodologiczne, lecz nie wykraczające poza repertuar ról semantycznych w kodzie SINTO. Nie związany z konkretnym językiem informacyjnym system relatorów Farradane'a obejmuje dziewięć rodzajów związków syntagmatycznych (zgodności, odpowiedności, odmienności, samoczynności, wymiaru, działania, skojarzenia, przynależności i zależności funkcjonalnej) [Chmielewska-Gorczyca E., 1981, 83-97]. Foskett wyróżnił cztery typy relacji, stanowiących modyfikację typologii Ranganathana: temat — odbiorca, wpływ, zależności ekspozycyjne i porównania [Sosińska B., 1983, 111-114].

3.4. REPREZENTACJA WIEDZY W SYSTEMACH FAKTOGRAFICZNYCH

Natura wiedzy i inteligencji należy do klasycznych problemów nauki absorbujących filozofów, psychologów, językoznawców i socjologów. Na gruncie sztucznej inteligencji zainteresowanie wiedzą wynikało z celów pragmatycznych, do których początkowo należało konstruowanie inteligentnych artefaktów, później algorytmizacja procesów myślowych człowieka i komputerowa realizacja programów symulujących myślenie. W badaniach nad reprezentacją wiedzy, stanowiących jeden z bardziej aktywnych nurtów sztucznej inteligencji, podkreśla się jednak ich pragmatyczny charakter uwarunkowany potrzebami systemów komputerowych.

Najczęstsze określenie „wiedzy” komputera, jako struktur danych i procedur interpretacyjnych w zakresie akceptowalnym przez użytkownika oraz procedur manipulacyjnych, wiąże się nie tylko z eksponowaniem technologicznego ujęcia zagadnienia. Chodzi także o podkreślenie rezygnacji z referencjalnej interpretacji znaczenia, nie poddającej się komputerowej implementacji. Komputer nie musi (i nie może) rozumieć całego języka naturalnego, dlatego też z pragmatycznego punktu widzenia informatyki, niezbędne jest przede wszystkim ustalenie zakresu zadań realizowanych przez system, określających stopień formalizacji języka oraz reguł transformacji tekstu, przede wszystkim zaś reguł wnioskowania zgodnie z nawykami inferencyjnymi użytkownika oraz granic akceptowalności konwersacji prowadzonej przez system. Pytanie, czym jest wiedza uzyskuje tu więc inny wymiar niż na gruncie języka naturalnego.

3.4.1. Kognitywne modele przetwarzania informacji

Charakterystyczne dla teorii baz danych systemów inteligentnych sprowadzanie reprezentacji wiedzy do problemu struktury danych i operacji manipulowania nimi zawiera pewien paradoks, związany z dążeniem do zapewnienia użytkownikowi zindywidualizowanego sposobu korzystania z systemu faktograficznego w celu uzyskania relewantnej informacji. Zagadnienie to ma szerszy wymiar metodologiczny związany z różnymi dyscyplinami naukowymi.

W aspekcie filozoficznym zindywidualizowana reprezentacja wiedzy oznacza opozycję wobec klasycznej arystotelesowskiej teorii poznania oraz obiektywizmu jako jej naczelnego założenia, przyjmowanego za podstawę tradycyjnej psychologii i językoznawstwa strukturalnego. Opozycja ta wyraża się w zakwestionowaniu teorii kategoryzacji Arystotelesa, w myśl której świat składa się z bytów o ustalonych cechach, które to byty pozostają między sobą w pewnych relacjach, obiektywnych i niezależnych od umysłu człowieka (równocześnie kwestionowana jest jej logiczna podstawa — teoria mnogości). Przyczyniły się do tego m.in. badania psychologiczne Eleonory Rosch oraz *Dociekania filozoficzne* Wittgensteina [Rosch E., 1977] [Wittgenstein L., 1972]. Z badań E. Rosch wynikało, że nie wszystkie okazy kategorii są jednakowo reprezentatywne, można je bowiem ułożyć w tzw. łańcuchy podobieństwa, w których część okazów zajmuje miejsce centralne, część zaś peryferyjne.

W aspekcie psychologicznym modelowanie procesów przetwarzania informacji przez człowieka polega na wykorzystaniu struktur języka naturalnego do badania procesów poznawczych i percepcji. Zakłada się nierozdzielność problemów psychologii i semantyki języka naturalnego w teorii poznawania, co oznacza nawiązanie do niearystotelesowskich tradycji w filozofii starożytnej. Charakterystyczne dla psychologii kognitywnej jest traktowanie mózgu ludzkiego jako procesora informacji, w którym wyróżnia się kanały kodowania, przechowywania i przekazu informacji oraz próbuje określić formalne własności procesów, jakim podlega informacja. Autorzy jednego z bardziej znanych podręczników psychologii kognitywnej stwierdzają: „Gdy usiłujemy znaleźć zasady organizacji mózgu ludzkiego celowe jest przeanalizowanie zasad organizacji i struktur wszelkich systemów informacji. Nacisk kładziemy na słowo zasady (...) nie ma znaczenia, czy system zbudowany jest z neuronów, elektronicznych obwodów scalonych czy też z przekładni i dźwigni — heurystyka we wszystkich przypadkach jest taka sama” [Lindsay P. H., Norman D. A., 1984, 520].

Akcentuje się więc dynamizm i proceduralny charakter reprezentacji rzeczywistości w mózgu człowieka oraz podleganie ciągłym transformacjom na skutek zmiany stanów pamięci. Problemy modelowania pamięci, istotne w opracowaniu zasad organizacji zbiorów informacyjnych systemów, stanowią bardzo obszerne zagadnienie, wykraczające poza ramy rozprawy. Większość autorów jest w zasadzie zgodna w wyróżnieniu trzech rodzajów pamięci: sensorycznej, długotrwałej oraz krótkotrwałej (*sensory, long-term, short-term memory*) [Norman D., 1970]. Szczegółowe koncepcje zależą w dużej mierze od celów, do jakich przeznaczony jest dany model, np. zastosowania teorii uczenia się (dla systemów samouczących), symulacji zachowań czy też przetwarzania informacji [*Computer Simulation of Cognitive Process*, 1963]. Teorię pamięci dynamicznej R. C. Schanka, związaną bezpośrednio z reprezentacją wiedzy w inteligentnym systemie informacji faktograficznej omawiamy w pkt. rozdz. 4.3.5.

W aspekcie językoznawczym metody kognitywne obejmują m.in. próby sformułowania nowej filozofii, traktowanej jako opozycja wobec zasad języko-

znawstwa strukturalnego (arbitralności znaku językowego, rozróżnienia synchronicznego i diachronicznego opisu języka). Formułowany jest postulat „wzbogacenia teorii semantyki języka naturalnego przez badanie związków między językiem, percepcją i myśleniem, dzięki którym możemy mówić o tym, co postrzegamy i co robimy” [Jackendoff R., 1983, 283]. Chodzi więc o opracowanie takich modeli przetwarzania informacji językowej przez człowieka, które stanowiłyby łącznik między teorią poznania, psychologią i teoriami semantycznymi. W myśl hipotezy sformułowanej przez Jackendoffa, w mózgu istnieją tzw. konceptualne struktury informacji, w ramach których informacja pochodząca z komunikacji językowej staje się kompatybilna z informacją pochodzącą z systemów peryferyjnych (wizji, słuchu, węchu, kinestezji), dzięki czemu człowiek potrafi komunikować o swoich doznaniach. Struktury te można scharakteryzować za pomocą skończonego zbioru uniwersalnych i wrodzonych reguł, tzw. reguł poprawności konceptualnej (*conceptual well-formedness rules*) reprezentujących wzmocnioną wersję teorii rozwoju Piageta. Jackendoff wskazuje dwa możliwe sposoby powiązania struktury konceptualnej mózgu z systemem (podsystemem) informacji językowej:

1) struktura ta może stanowić odrębny poziom, związany ze strukturami informacji semantycznej poprzez reguły pragmatyki (*rule component*), określające relacje między znaczeniem językowym i rzeczywistością pozajęzykową [Katz J. J., Fodor J. A.];

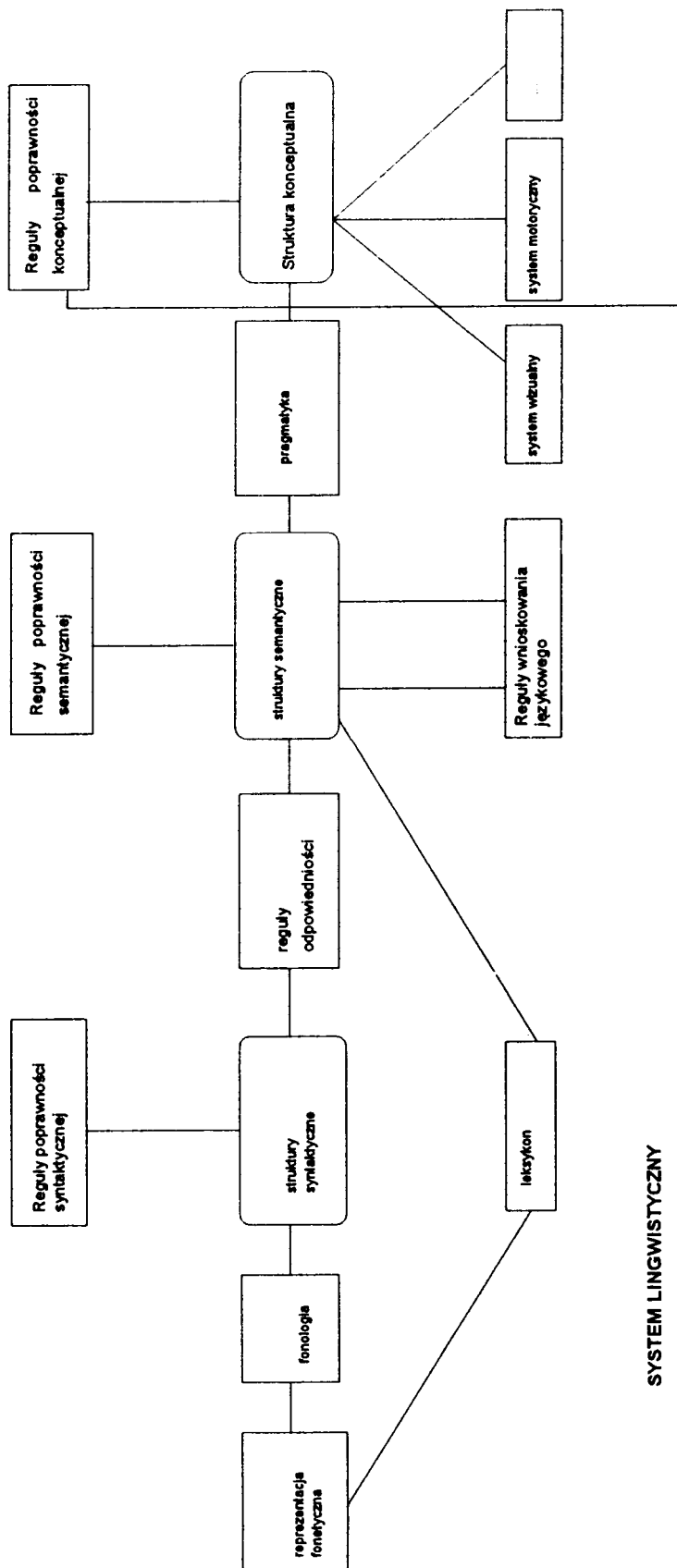
2) struktury semantyczne stanowią podzbiór struktur konceptualnych, zaś łącznikiem między nimi są reguły odpowiedniości (*correspondance rules*) wraz z regułami pragmatyki i regułami wnioskowania (por. schematy 5 i 6).

Ujęcie drugie jest typowe dla organizacji wiedzy w sztucznej inteligencji, gdzie tzw. bazę danych konceptualnych (*conceptual data base*) stanowią wyrażenia języka opisującego obiekty systemu, transformowane na mocy przyjętych reguł inferencji, zaś kontrola poprawności wnioskowania realizowana jest za pomocą przyjętych zasad heurystyki (pragmatyki systemu).

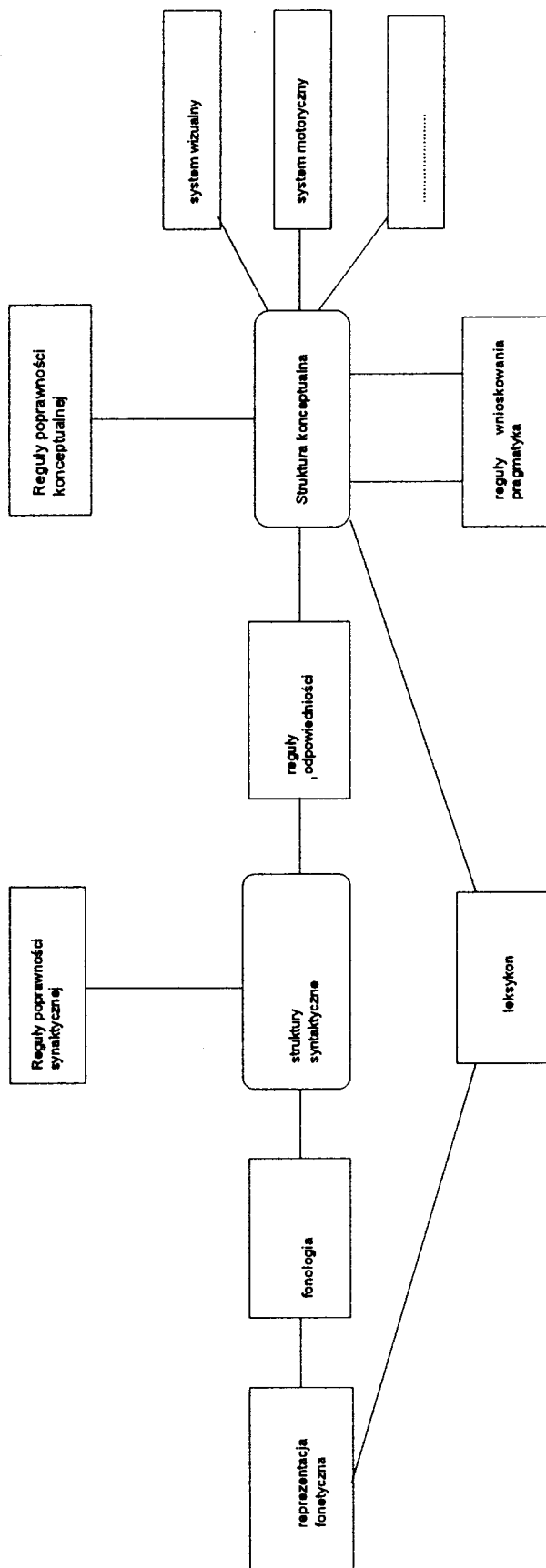
W interpretacji kognitywnej idealne modele konceptualne odnoszą się do całej wiedzy człowieka o świecie i wyodrębniane są na podstawie określonych modeli kulturowych, światopoglądowych, poznawczych oraz stereotypów związanych z przesądami czy uprzedzeniami. Na przykład, na model znaczenia wyrazu *matka* składają się stereotypowe cechy różnych modeli:

- urodzenia (*kobieta, która urodziła*),
- genetyczny (*samica*),
- edukacyjny (*kobieta, która wychowuje*),
- matrymonialny (*żona ojca*),
- genealogiczny (*najbliższa krewna wstępna*).

Idealny model realizuje wszystkie cechy, inne modele mogą być niepełne (*matka pracująca*), bądź rozszerzone (*dobra matka, choć nie umie prowadzić domu*). [Encyklopedia językoznawstwa ogólnego, 1993, 310] Cechy ujęte w kognitywnym modelu znaczeniowym mogą być źródłem metafor — *komisja-matka* = *rodzica*, oraz metonimii odnoszących się do funkcji *matka* — *najważniejszy*



Schemat 5. Powiązanie struktur konceptualnych z systemem informacji językowej. Ujęcie 1.



Schemat 6. Powiązanie struktur konceptualnych z systemem informacji językowej. Ujęcie 2.

gracz w drużynie (najważniejsza osoba). Kognitywiści uznają metonimizację i metaforyzację za główną treść języka, co w konsekwencji prowadzi do odrzucenia podziału na wiedzę encyklopedyczną (pozajęzykową) i językową, ponieważ — jak twierdzą — człowiek posługuje się językiem w kontekście innych, nierozdzielnych rodzajów wiedzy. Metody tej jeszcze niedostatecznie ukształtowanej metody opisu języka wykorzystuje — oprócz sztucznej inteligencji — translatoryka [Langacker R., 1987] [Lakoff G., Johnson M., 1988] [Tabakowska E., 1990].

3.4.2 Sposoby reprezentacji wiedzy

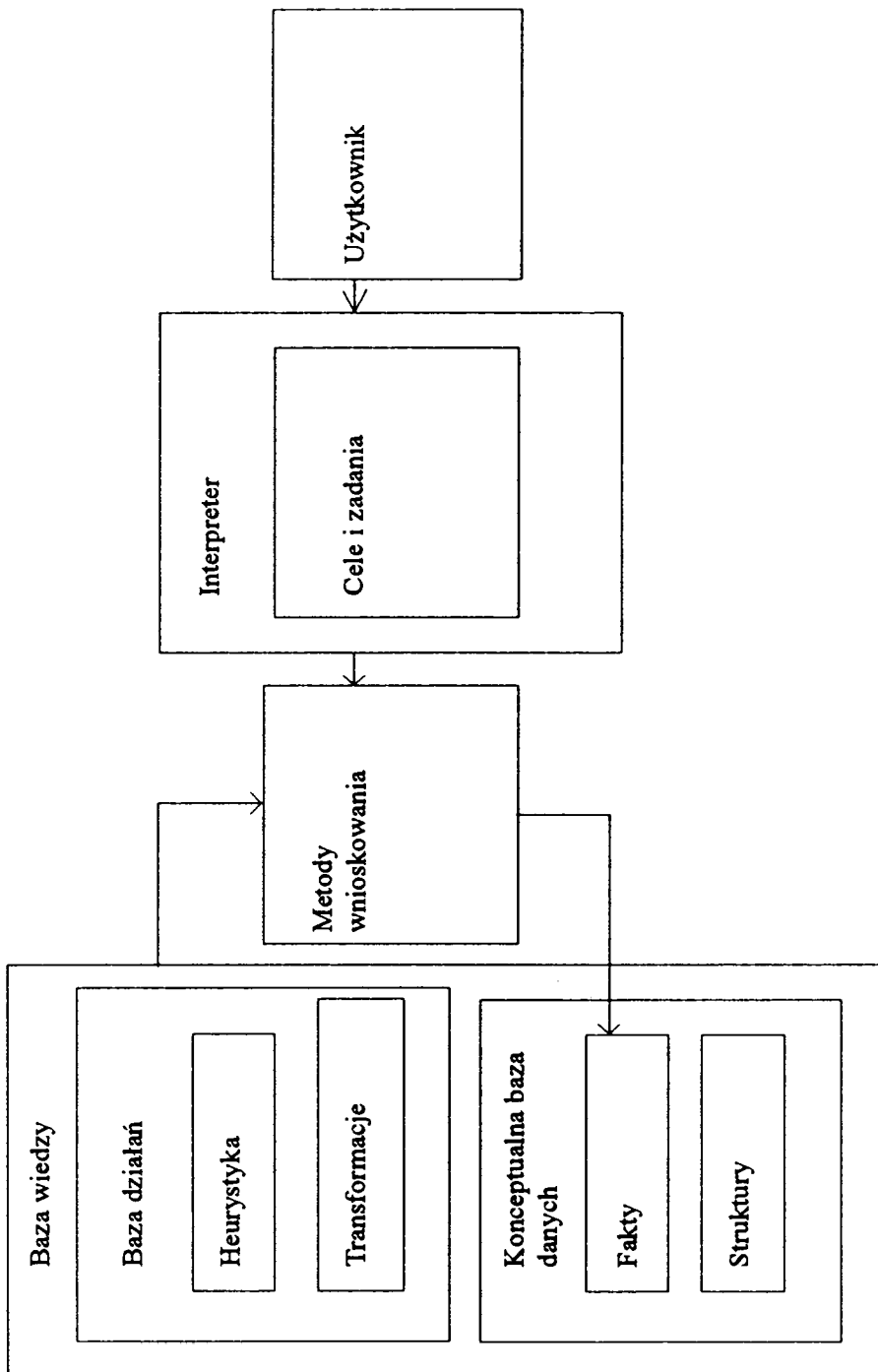
Problem reprezentacji wiedzy w systemach informacyjnych podnoszony jest jako zagadnienie teoretyczne i aplikacyjne od stosunkowo niedawna, w związku z rozwojem *systemów inteligentnych* (eksperckich/ekspertowych, doradczych — expert, advisory systems), a więc generujących ze zbioru tekstów zawierających informację faktograficzną, na zasadzie transformowania ich treści, informacje nie będące podzbiorami tych tekstów. Podstawą generowania nowej informacji są operacje *wnioskowania* i *wyjaśniania*. W ramach badań nad procesami kognitywnymi człowieka powstają, co prawda, atrakcyjne modele przetwarzania informacji mentalnej, jak na przykład przedstawiony w rozdziale 2.4 model Sperbera i Wilson, czy zasygnalizowane wyżej ujęcie Langackera, mają one jednak w znacznej mierze charakter spekulatywny, niesatysfakcjonujący z punktu widzenia technologii systemów.

Z punktu widzenia pragmatyki epistemologicznej wiedzę systemu faktograficznego charakteryzuje się poprzez rozróżnienie dwu podzbiorów informacyjnych:

- podzbioru cech odwzorowywanych za pomocą pewnego języka w tzw. konceptualnej bazie danych (conceptual data basis), z uwzględnieniem wzajemnych powiązań między elementami bazy;
- podzbioru reguł transformacji informacji bazy danych niezbędnych do gromadzenia informacji, jej wyszukiwania oraz derywowania nowej na mocy przyjętych reguł wnioskowania [Barr A., Fegenbaum E. A., 1982].

Wskazanie tych dwóch zasadniczych podzbiorów w zbiorze wiedzy systemu jest oczywiście dużym uproszczeniem, pełniącym funkcję porządkującą. Zarówno w przypadku podzbioru informacji faktograficznej, jak i inferencyjnej niezbędne jest wyposażenie systemu w wiedzę na poziomie *meta-*, opisującą z jednej strony strukturę danych i warunki ich poprawności, z drugiej zaś mechanizmy heurystyczne, kontrolujące poprawność procesów wnioskowania (heuristics, meta-reasoning). Strukturę systemu faktograficznego (eksperckiego), w którym możliwa jest transformacja informacji ilustruje schemat 7.

Obecnie zajmujemy się zasadami reprezentacji wiedzy w zbiorach systemu, a więc w konceptualnej bazie danych. Charakterystykę zasad wyszukiwania informacji z uwzględnieniem reguł semantycznej transformacji informacji omawiamy w dalszej części pracy poświęconej metodom dostępu do wiedzy w syste-



Schemat 7. Struktura inteligentnego systemu informacji faktograficznej

- *reprezentacji modularnej* łatwo przyswajalnej dla użytkownika, ze względu na łatwość dekompozycji struktur danych, przeciwstawianej reprezentacji niemodularnej (proceduralnej), w której występuje silne uzależnienie zawartości bazy (procedur) od kontekstu, w którym przetwarzana jest informacja;
- *reprezentacji deklaratywnej* przeciwstawianej *reprezentacji proceduralnej*. Jako zalety reprezentacji deklaratywnej wymienia się elastyczność związaną z łatwością modyfikacji bazy danych, natomiast reprezentacji proceduralnej przypisuje się cechę przejrzystości wnioskowania, przystosowanego do heurystyki danej dziedziny wiedzy. Różnicę między ujęciem deklaratywnym i proceduralnym może zilustrować przykład prostego programu do sortowania alfabetycznego, w którym wiedza proceduralna jest odwzorowana *implicite* w postaci: „*A poprzedza B, B poprzedza C*”, co odzwierciedlają odpowiednie kody liter; *explicite* określana jest w procedurach sortowania.

Reprezentacje deklaratywne są historycznie wcześniejsze. Projektanci systemów inteligentnych koncentrowali się początkowo na ustaleniu tych rodzajów wiedzy, którą można odwzorować w systemie za pomocą języka sieci semantycznych oraz systemów logicznych (rachunku zdań, rachunku predykatów), traktując problemy transformacji jako zagadnienie drugoplanowe. Refleksje na temat równorzędności znaczenia statycznych reprezentacji wiedzy oraz sposobów jej transformacji pojawiły się połowie lat siedemdziesiątych. Przedmiotem rozważań stał się także problem relewancji i użyteczności wiedzy [Barr A., Feigenbaum E.A., 1982].

Wspólną cechą wszystkich metod reprezentacji wiedzy w systemach faktograficznych jest ich ograniczenie do określonych dyscyplin. Najczęściej zakłada się, że *apriori* znane są fakty (obiekty, zdarzenia etc.), których cechy są odwzorowywane w bazie danych, natomiast zadaniem projektanta systemu jest określenie sposobu kodowania informacji w ramach przyjętych struktur oraz sposobu transformowania wyrażen.

Do najczęściej stosowanych metod odwzorowania wiedzy należą:

- 1) typowe dla ujęcia deklaratywnego:
 - a) systemy logiczne,
 - b) sieci semantyczne,
 - c) elementarne cechy semantyczne,
- 2) typowe dla ujęcia proceduralnego:
 - a) reprezentacje proceduralne,
 - b) reguły produkcji;
- 3) łącząca cechy ujęcia deklaratywnego i proceduralnego — teoria ram.

3.4.2.1. Systemy logiczne

We wczesnych systemach inteligentnych językami reprezentacji wiedzy były systemy klasycznej logiki formalnej. Zadecydowały o tym dwa czynniki: łatwość definiowania aksjomatów systemu dla danej dziedziny semantycznej oraz wyraźnie określona składnia i reguły wnioskowania dedukcyjnego. Naj-

częściej stosowanym językiem logiki formalnej jest rachunek predykatów pierwszego rzędu. Rachunek zdań jest zbyt ubogi ze względu na konieczność odwzorowania w bazie systemu nie tylko wiedzy o prawdziwości lub fałszywości zdań, lecz także o obiektach pozajęzykowych, relacjach między nimi i ich klasami. Rachunek predykatów jest jednak językiem bardzo ogólnym i często niewygodnym aplikacyjnie. Dlatego też uzupełniany jest funkcjami (operatorami), których argumentami może być zmienna, stała lub inna funkcja, oraz znakiem równoważności predykatów (equals) $XY \text{ wtt}$ dla wszelkich predykatów $P, P(X), P(Y)$ i funkcji $F(X)=F(Y)$.

Podstawową jednostką zbioru informacji zakodowanej w języku rachunku predykatów jest poprawnie zbudowana formuła (*well-formed formula* — *wff*), tworzona przy użyciu spójników zdaniowych, predykatów, stałych, zmiennych i kwantyfikatorów oraz operatorów funkcyjnych. Interpretacja semantyczna formuły odbywa się na podstawie wiedzy o obiektach danej dziedziny wiedzy pozajęzykowej oraz relacji między nimi zachodzących. Stanowią więc one modele semantyczne danej dziedziny, w których:

- każdemu symbolowi stałemu przyporządkowany jest pewien obiekt dziedziny,

- każdemu symbolowi funkcyjnemu przyporządkowana jest funkcja określona na dziedzinie,

- każdemu symbolowi predykatowemu przyporządkowane są relacje między obiektami [Madalińska A., Bugaj E., 1983, 85-102].

Każda formuła może być prawdziwa lub fałszywa w danym modelu. Jest uznawana za prawidłową jedynie wówczas, gdy jest prawdziwa (spełniona) dla wszystkich modeli pewnej dziedziny semantycznej i stanowi wówczas twierdzenie systemu logicznego. Stwierdzenie prawidłowości (prawdziwości) budowy wszystkich formuł systemu najczęściej nie jest możliwe, i dlatego do ustalenia twierdzeń danej dziedziny stosuje się formalne procedury dowodzenia (tzw. procedury decyzyjne). Powinny one spełniać warunek zgodności (formuły przez nie ustanowione powinny być prawidłowe, a więc prawdziwe) i kompletności (powinny ustanawiać twierdzenia). Regułami wnioskowania, a więc transformacji informacji, są reguły dedukcji naturalnej.

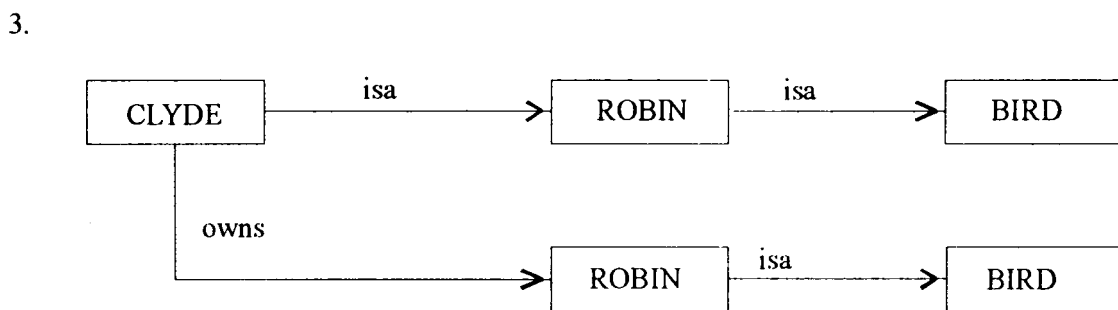
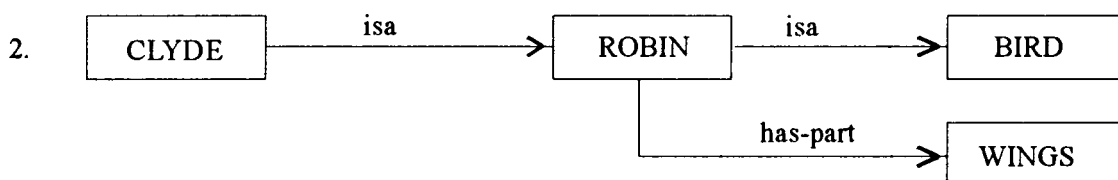
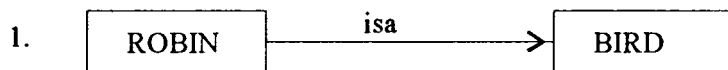
Rachunek predykatów znajduje zastosowanie w tych systemach, w których o dostarczeniu użytkownikowi informacji relewantnej decydują reguły dowodzenia i wnioskowanie dedukcyjne (najczęściej metody dedukcji naturalnej) — w matematyce, chemii, sterowaniu robotami oraz automatycznego programowania.

O powszechności wykorzystania rachunku predykatów jako języka reprezentacji wiedzy decyduje jego precyzja w definiowaniu cech obiektów i relacji między nimi, elastyczność umożliwiająca zatomizowaną reprezentację rzeczywistości w postaci niezależnych formuł oraz modularność ułatwiająca rozbudowę bazy asercji. Jako zasadniczą wadę wymienia się rozdział wiedzy o statycznych cechach obiektów oraz wiedzy na temat metod transformacji wiedzy asertorycznej, a więc rozdział epistemologicznej i heurystycznej bazy danych.

Próby rozbudowy rachunku predykatów przez dołączenie do formuł zdaniowych reguł ich użycia podejmowane są w systemach opartych na reprezentacjach ramowych (por. rozdz. 3.4.2.5.).

3.4.2.2. Sieci semantyczne

Terminem sieci semantyczne określa się różne języki reprezentacji wiedzy, których jedyną wspólną cechą jest podobieństwo strukturalne notacji. Bazę notacji stanowią węzły (nodes), graficznie oznaczane jako kropki, kółka lub prostokąty oraz łuki (arcs), graficznie oznaczane strzałkami. Zarówno węzły jak i łuki są opatrywane etykietami (labels), pełniącymi funkcję wykładników znaczenia wyrażen składowych sieci. Etykiety węzłów reprezentują wyrażenia elementarne języka desygnujące obiekty, stany, zdarzenia itp., natomiast łuki odwzorowują relację między faktami desygnowanymi przez wyrażenia elementarne [Barr A., Fegenbaum E. A., 1982].



W literaturze przedmiotu podkreśla się, że poza podobieństwem notacji trudno jest wskazać uniwersalne zasady interpretacji semantycznej sieci. Ich charakterystyka jako sposobu reprezentacji wiedzy ogranicza się do prezentacji rozwiązań już historycznych, bądź najczęściej stosowanych zasad organizacji pola semantycznego języka.

Jako autora koncepcji wykorzystania sieci semantycznych wymienia się Rossa Quilliana [Quillian R., 1968, 227-270], który zaprojektował sieć pomyślaną jako model pamięci asocjacyjnej człowieka. Sieć ta składała się z grup węzłów nazwanych *planes*, z których każdy reprezentował wiedzę o znaczeniu wyrażenia (*concepts*). Wnioskowanie w sieci odbywało się drogą tzw. rozszerzonej aktywizacji (*spreading activation*), polegającej na odnajdywaniu powiązań między węzłami. Proces ten rozpoczynał się od dwu węzłów, powodując powstawanie tzw. sfery aktywizacji (*sphere of activation*). Powiązanie odnajdywano w momencie, gdy węzeł reprezentujący „pojęcie” był aktywizowany z dwu stron. W późniejszej wersji modelu Quillian podjął próbę rozwiązania problemu bezpośredniego ustalania relacji między węzłami, który stanowił znaczne utrudnienie w wyszukiwaniu informacji o pojęciach powiązanych, niezbędnych w procesach wnioskowania.

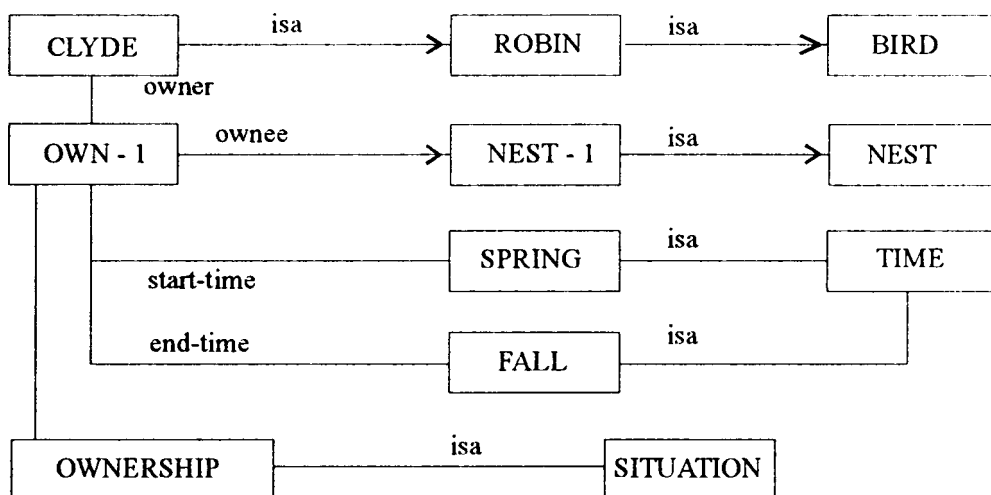
Zaletą sieci semantycznych jest możliwość opisu obiektów za pomocą stosunkowo niewielkiego zbioru wyrażenia elementarnych semantycznie, reprezentowanych w węzłach, np. *FLYING is a MOTION* (Latanie jest ruchem). Jako wady wymienia się ograniczenie możliwości odwzorowania relacji do relacji dwuargumentowych, brak możliwości rozróżnienia cech obiektów indywidualnych i cech klas obiektów oraz kłopoty z odwzorowaniem relacji sytuacyjnych.

Najprostszym sposobem organizacji semantycznej sieci jest wykorzystanie relacji generycznych, ustalanych na podstawie klasyfikacji obiektów (faktów), a także relacji mereologicznych, które wykorzystywane są w procesach wnioskowania dedukcyjnego na mocy zasady dziedziczenia cech przez wyrażenia umieszczone w węzłach nadrzędnych. Cechy przypisane wyrażeniom nadrędnym interpretowane są jako asercje wszystkich wyrażenia podrzędnych zakresowo lub mereologicznie. Na przykład: *Robin is a bird — Drozd jest ptakiem(1)*; *Clyde is a robin — Clyde jest drozdem(2)*; *Bird has (part) wings — Ptak ma (część) skrzydła(3)*;

Ograniczenie sieci widoczne staje się w momencie potrzeby odwzorowania relacji sytuacyjnych. Zdanie *Clyde owns a nest* musi być reprezentowane przy rozróżnieniu obiektu jednostkowego (gniazda, które posiada) i klasy obiektów (gniazd).

Jednakże przy tej strukturze sieci nie jest możliwe odwzorowanie zdania *Robin owns a nest from spring until fall* (Drozd ma gniazdo od wiosny do jesieni) ze względu na ograniczenie reprezentacji do relacji binarnych. Rozwiązanie tego problemu zaproponowali Simmons i Slocum [Winograd T., 1980, 209-241], dostosowując strukturę sieci do prezentowania wieloargumentowych predykatów sytuacyjnych; zaproponowany przez nich model nosi nazwę ramy sytuacyjnej (*case frames*):

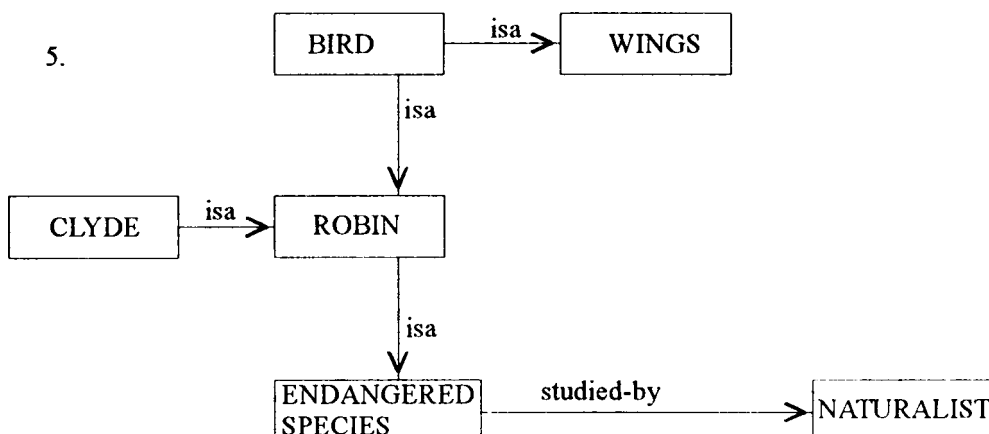
4.



Problemy z interpretacją semantyczną zdań opisujących sytuację powstają także w związku z trudno odwzorowalnymi różnicami między cechami obiektów jednostkowych i klasami jednostek, przy równoczesnym obowiązywaniu zasady dziedziczenia cech wyrażeń nadrzędnych. Na przykład interpretując zdanie: system może wydedukować zdanie: *Naturalists study endangered species* (Przyrodnicy badają zagrożone gatunki) oraz *Naturalists study Clyde* (Przyrodnicy badają Clyde'a), które mogą być prawdziwe lub fałszywe. Możliwość powstania takich interpretacji spowodowana jest z jednej strony brakiem sformalizowanej interpretacji semantycznej struktur sieciowych, z drugiej zaś zróżnicowaniem mechanizmów inferencji w poszczególnych systemach.

W nowszych pracach na temat sieci semantycznych obserwuje się tendencję do łączenia sieci w złożone struktury nazywane ramami, scenariuszami (*frames, scenarios*), w których wykorzystuje się także proceduralne metody reprezentacji wiedzy. Dotychczasowe zastosowanie sieci semantycznych wiąże się przede wszystkim z systemami przetwarzającymi teksty języka naturalnego w połączeniu z programami analizy syntaktycznej i morfologicznej (parser),

5.



na przykład ATN (Augmented Transition Network) [Woods W. A., 1975, 11-154]. Tak więc, *sieci semantyczne służą de facto reprezentacji wiedzy o języku, a jedynie pośrednio — wiedzy o rzeczywistości pozajęzykowej.*

3.4.2.3. Elementarne cechy semantyczne

Budowa każdego języka reprezentacji wiedzy obejmuje tworzenie jego słownika — na przykład tezaurusu języka deskryptorowego, słownika predykatów przy formalizmach logicznych itp. Słowniki elementarnych cech semantycznych (semantic primitives) definiuje się najczęściej jako słowniki wyrażen reprezentujących elementy znaczenia innych wyrażen. Z literatury językoznawczej doskonale znane są prace Apresjana, Melčuka i Żółkowskiego nad budową metajęzyka semantycznego w ramach modelu sens — tekst [Apresjan J.D., 1971; 1980], podobnie jak prace A. Wierzbickiej [Melčuk I., Zholkovsky A., 1984; Wierzbicka A., 1972]. Niestety nie podejmowano prób ich wykorzystania w konkretnych systemach.

Jednym z wcześniejszych systemów, w których wykorzystano ten sposób reprezentacji wiedzy, był opracowany przez Y. Wilksa system zautomatyzowanego przekładu z języka angielskiego na francuski [Wilks Y., 1973]. Teksty języka angielskiego były transformowane do postaci sieci semantycznych, w których węzły etykietowano za pomocą, a następnie na tej podstawie generowano teksty języka francuskiego.

Koncepcję słownika prymitywów semantycznych sformułowaną przez Wilksa określają zasady tzw. semantyki preferencyjnej (*preference semantics*), polegające na ustaleniu cech dystynktywnych, składających się na znaczenie pewnego wyrażenia i wskazaniu wśród nich cechy preferowanej, niezbędnej i wystarczającej do identyfikacji znaczeń innych wyrażen (ich sensów), np. *drink* — *fluid object* (wino, benzyna). W przypadku gdy zbiór preferowanych sensów okazywał się niewystarczający do rozróżnienia znaczeń, wprowadzano opis pełnej interpretacji wyrażenia. Wilks sformułował warunki, jakie powinien spełniać poprawnie zbudowany słownik elementarnych cech semantycznych:

- 1) skończoności (*finitude*) — liczba cech powinna być skończona i mniejsza niż liczba wyrażen definiowanych za ich pomocą;
- 2) adekwatności (*comprehensiveness*) — zbiór powinien być adekwatny w wyrażaniu sensów wyrażen i ich rozróżnianiu;
- 3) niezależności (*independence*) — żadne z wyrażen elementarnych nie może być definiowalne w kategoriach innych wyrażen elementarnych;
- 4) nie-cyrkularności (*noncircularity*) — dwa wyrażenia nie mogą tworzyć błędnego koła definicyjnego;
- 5) elementarności (*primitiveness*) — żaden z podzbiorów cech nie może być zastępowalny przez inny podzbiór.

Słownik systemu pełnił faktycznie funkcję metajęzykową, służył bowiem do opisu i odróżniania znaczeń wyrażen dwu języków naturalnych, a w szczegól-

ności wyrażen wieloznacznych. Zawierał wykaz 80 cech dystynktywnych, podzielonych na pięć kategorii: JEDNOSTKI (*Entities*), DZIAŁANIA (*Actions*), PRZYPADKI/SYTUACJE (*Cases*), KWALIFIKATORY (*Qualifiers*), WSKAŹNIKI TYPU (*Type indicators*). Ponadto wprowadzono wyrażenia o funkcji wyrażen podstawowych, choć nie elementarnych, pozostające w relacji nadrzędności generycznej z wyrażeniami elementarnymi, wyróżnione znakiem *, np.: *ANI — MAN. FOLK (a human group — grupa ludzi), BEAST (a nonhuman animal — zwierzę nie będące człowiekiem):

ENTITIES (jednostki)

MAN (człowiek)

STUFF (substancja)

PART (część jednostki)

ACTIONS (działania)

CAUSE (powodowanie czegoś)

BE (być — bycie jako równoważność lub orzeczenie)

FLOW (pływać — poruszać, przemieszczać się jak płyny)

CASES TO (przypadki — kierunek ku czemuś)

IN (zawieranie)

QUALIFIERS (kwalifikatory)

GOOD (dobry — słuszny moralnie lub akceptowany)

MUCH (dużo — w odniesieniu do substancji)

TYPE INDICATORS (wskaźniki typu)

HOW (jak -bycie typem działania dla konstrukcji przysłówkowych)

KIND (rodzaj — bycie jakością dla konstrukcji przymiotnikowych)

A oto przykład definicji słowa BREAK (złamać)

(BREAK: (*HUM SUBJ)

(*PHYSOB OBJE)

((((NOTWHOLE KIND) BE) CAUSE) GOAL)

(THING INSTR)

(STRIK)

— co oznacza uderzenie (*STRIK*ing) realizowane raczej przez *HUM*an *SUBJ*ect za pomocą pewnego instrumentu (*INSTR*) to jest rzeczy (*THING*) w celu (*GOAL*) spowodowania (*CAUS*ing), aby obiekt fizyczny (*PHYS*ical *OBJE*ct) nie był całością (*BE NOT WHOLE*).

Zdania przedstawiono za pomocą formuł (templates), których podstawową jednostkę stanowiły trójki (triples) — reprezentujące agensa, przedmiot działania i działanie, np. *Small men sometimes father big sons* (Mali mężczyźni są czasem ojcami wielkich synów) odwzorowywano następująco:

[men]	↔	[father]	↔	[sons]
↑		↑		↑
[small]		[sometimes]		[big]

W odróżnieniu od aplikacyjnego rozwiązania w systemie Wilksa, teoria zależności konceptualnej R. Schanka (*Conceptual Dependence Theory*)

miała stanowić przede wszystkim propozycję intuicyjnego modelu przetwarzania języka naturalnego przez człowieka, niezależnego od konkretnych aplikacji komputerowych, jakkolwiek przewidywano jej wykorzystanie do parafrazowania tekstów, wnioskowania oraz konwersacji z użytkownikiem [Schank R., 1975].

Podstawą reprezentacji tekstów jest słownik złożony z jedenastu jednostek nazwanych *AKTAMI* (*ACTs*) oraz kategorii pomocniczych niezbędnych do jednoznacznego odwzorowania cech sytuacji opisywanych w zdaniach wieloznacznych, np. *I saw the Grand Canyon flying to New York* (Widziałem Grand Canyon, lecąc do Nowego Jorku — Widziałem Grand Canyon lecący do Nowego Jorku) lub też bliskoznacznych, np. *I want a book* (Chcę książkę) — *I want to get a book* (Chcę dostać książkę) — *I want to have a book* (Chcę mieć książkę). W zbiorze jedenastu działań elementarnych (*ACTs*) wyróżniono cztery kategorie:

PHYSICAL ACTS (Działania fizyczne)

PROPEL (poruszać — stosować siłę wobec obiektu fizycznego)

MOVE (ruszać — przemieszczać część ciała)

INGEST (polykać — umieścić coś wewnątrz obiektu ożywionego)

EXPEL (wydalać — zmusić coś do wydostania się z wnętrza obiektu ożywionego)

GRASP (chwycić obiekt fizyczny)

ACTS CHARACTERIZED BY RESULTING STATE CHANGES (Działania charakteryzowane przez zmiany stanów)

PTRANS (zmienić lokalizację obiektu fizycznego)

ATRANS (zmienić relację abstrakcyjną np. posiadanie ze względu na pewien obiekt)

ACTS USED MAINLY AS INSTRUMENTS FOR OTHER ACTS (Działania pełniące funkcję instrumentalną dla innych działań)

SPEAK (mówić — wytwarzać dźwięk)

ATTEND (kierować narząd zmysłu ku bodźcowi)

MENTAL ACTS (działania mentalne)

MTRANS (przekazywać informację)

MBUILD (tworzyć nową informację ze starej).

Zbiór kategorii pomocniczych zawiera:

PICTURE PRODUCERS (*PPs*) (obiekty fizyczne, a wśród nich siły natury, takie jak np. wiatr)

PICTURE AIDERS (*PAs*) (cechy obiektów)

TIMES (czasy)

LOCATIONS (lokalizacje)

ACTION AIDERS (*AA.s*) (atrybuty działań (*ACTs*))

Charakterystyka sytuacji prezentowanych w tekstach polega na połączeniu wyrażeń kategorii elementarnych i pomocniczych za pomocą relacji instrumentalizacji (*instrumentality*) lub kauzatywnych (*causation*), Tworzenie opisów sy-

tuacji określane jest jako proces konceptualizacji (conceptualization) odwzorowujący:

- 1) działającego (actor) — *PP* (picture produce) jako realizatora aktu elementarnego (*ACT*),
- 2) działający obiekt (actor) — *PP* oraz opis jego stanu, a więc cechy obiektu (*PA*).

Zbiór elementarnych działań miał odpowiadać — zdaniem Schanka — raczej mentalnej reprezentacji sytuacji w umyśle człowieka niż reprezentacji znaczeń tekstów, w których część informacji może być zawarta implicite. Uzyskanie tej informacji umożliwia operacje wnioskowania realizowane na zbiorze prymitywów semantycznych w połączeniu z cechami obiektów i zdarzeń odwzorowujących daną sytuację. Każdy z nich implikuje bowiem odpowiedni zbiór rozumowań (inferences). Równocześnie pozwala to zaoszczędzić czas przetwarzania oraz przestrzeń w pamięci. Na przykład elementarne działanie *MTRANS* (przekaz informacji) służy reprezentacji sytuacji *wiedzieć* (to know) wówczas, gdy: coś słyszymy, widzimy, czytamy, a więc gdy zachodzi przekaz informacji z otoczenia do pamięci systemu (komputera).

Dość szybko stwierdzono nieadekwatność pierwotnych założeń teorii zależności konceptualnej i małą przydatność w systemach rozumiejących, a ściślej interpretujących język naturalny. Ograniczenia dotyczyły m.in. :

- braku powiązania informacji przetwarzanej z wiedzą już posiadaną przez człowieka (kontekstu),
- możliwości opisu sytuacji odwzorowujących działania jedynie w kategoriach zdarzeń fizycznych, np. całować (to kiss) — *to MOVE lips to lips* (przemieszczać usta do ust).
- niemożności interpretacji wyrażen metaforycznych oraz dokonywania kwantyfikacji.

Założenia teorii zależności konceptualnej rozwinął Schank w teorii ram łączącej — jak wspomnieliśmy wyżej — cechy ujęcia deklaratywnego i proceduralnego, którą omawiamy w dalszej części rozdziału. Teoria ta zastosowana była w systemie *SAM* (Story Comprehension Program) [Riesbeck K., 1978].

3.4.2.4. Reprezentacje proceduralne

Wprowadzenie metod proceduralnej reprezentacji wiedzy w końcu lat sześćdziesiątych było właściwie zabiegiem komplementarnym wobec metod deklaratywnych [Barr A., Fegenbaum E. A., 1982]. Integralną część zbioru informacyjnego systemu stanowiła w dalszym ciągu faktograficzna baza danych, które poddawano transformacji w procesie wyszukiwania przy użyciu procedur, ustrukturalizowanych odpowiednio dzięki nowym językom programowania, np. LISP, PLANER. Procedury określające kierunek rozumowania zostały dołączone do danych faktograficznych, tak aby system korzystał w procesie wyszukiwania jedynie z wiedzy relewantnej przy danym celu, nie zaś z wszelkiej wiedzy posiadanej. Na przykład do potwierdzenia, że pewien obiekt potrafi latać może być potrzebna wiedza, że obiekt ten jest ptakiem, reprezentowana następująco:

(IF-NEEDED FLIES(X)
TRY-SHOWING BIRD(X)).

Zgodnie z taką heurystyką, jeśli chcemy stwierdzić, że *Clyde* (drozd) potrafi latać system musi udowodnić, że jest on ptakiem. Równocześnie mogą wystąpić ograniczenia kierunku rozumowania: ze zdania „*All birds can fly*” (Wszystkie ptaki potrafią latać) i *Fred is a bird* (Fred jest ptakiem) system nie jest w stanie wydedukować zdania *Fred can fly* (Fred może latać). Określenie dopuszczalnego kierunku wnioskowania, podobnie jak wskazanie danych bazy faktograficznej relewantnych przy danym zapytaniu, jest zadaniem języka programowania. Charakterystyka języków programowania jest zagadnieniem wykraczającym poza zakres niniejszej pracy. Zasygnalizujemy jedynie, że kierowanie wnioskowaniem (inference guiding) umożliwiają specjalne procedury nazwane w języku *PLANNER* *filtrami* (filters), rekomendujące nie tylko elementy danych relewantne w danej sytuacji i służące jako twierdzenia (theorem) w procesie wnioskowania, lecz także całe klasy twierdzeń relewantnych w danej dziedzinie i przypuszczalnie w danej sytuacji, np. w celu udowodnienia, że pewien obiekt jest ptakiem, system zaleca użycie filtru kierującego do zbioru twierdzeń z zakresu zoologii:

(GOAL BIRD (FILTER ABOUT-ZOOLOGY)).

Jako jedno z rozwiązań w zakresie proceduralnej reprezentacji wiedzy wymienia się wnioskowanie domyślne (default reasoning) stanowiące uzupełnienie klasycznego wnioskowania dedukcyjnego (extended logical inference). Na przykład w języku *PLANNER* ten sposób wnioskowania możliwy jest dzięki wyrażeniu *THNOT*.

(*THNOT OSTRICH (X) ASSUME FLIES (X)*),

które interpretowane jest następująco: o ile nie można udowodnić, że *X* (wszystkie ptaki) nie jest strusiem, to zakładamy, że *X* potrafi latać.

Jako zaletę proceduralnej reprezentacji wiedzy podaje się jej przydatność w organizowaniu zbiorów informacji i korzystny wpływ na ukierunkowanie i bezpośredniość procesów wnioskowania (*directedness*), przyporządkowanych zakładanym celom. Jako ograniczenie wymienia się zdarzającą się niemożność spełnienia warunku kompletności i spójności przetwarzanej wiedzy. Problem niekompletności powstaje wówczas, gdy mimo wyszukania wszelkich danych faktograficznych w bazie, system nie potrafi przeprowadzić oczekiwanego wnioskowania. Problem niespójności związany jest z wnioskowaniem domyślnym, na przykład wówczas, gdy ze zdania *Fred is a bird* (Fred jest ptakiem) system wnioskuje: *Fred can fly* (Fred może latać) i *Fred is an ostrich* (Fred jest strusiem). Rozwiązania tych problemów upatruje się w ścisłym wiązaniu struktur danych faktograficznych z procedurami, składającymi się na program realizujący odpowiednie transformacje danych (procedural attachment) [Cichy M., 1983, 103-136].

3.4.2.5. Reguły produkcji

Reguły produkcji (production rules, productions) są odmianą *proceduralnej reprezentacji wiedzy*, przydatną jako mechanizm kontroli interakcji między wyrażeniami zbioru *wiedzy deklaratywnej* (faktograficznej) i *proceduralnej* (heurystycznej). Ich scharakteryzowanie wiąże się z problemami podobnymi do wspomnianych przy opisie sieci semantycznych, istnieje bowiem szereg różnych rozwiązań aplikacyjnych przystosowanych do architektury konkretnych systemów o odmiennej interpretacji semantycznej [Newell A., 1973, 463-526].

Jako elementy typowe dla tej metody reprezentacji wiedzy wyróżnia się:

- zbiór reguł określających warunki realizowania określonych działań (*rule base*),
- zbiór wyrażen bazy danych tworzących kontekst, który powinien spełniać warunki niezbędne i wystarczające do realizowania danego zadania w określonym momencie,
- interpreter sterujący działaniami systemu (jak w każdym komputerze), a w szczególności podejmujący decyzje o kolejności działań.

Każda reguła ma postać formuły logicznej *If ... then*, interpretowanej jako: *Jeśli spełniony jest dany warunek, to działanie jest właściwe*. W formule tej wyróżnia się tzw. część warunkową czyli lewostronną (*condition part, left-side part*) oraz część działaniową czyli prawostronną (*action part, right-side part*). Po stwierdzeniu spełnienia (lub nie) warunków określonych w części *If ...*, system wnioskuje o relewancji działań i dokonuje transformacji informacji.

Zasadniczą rolę w podejmowaniu decyzji o relewancji danego działania pełni zbiór wyrażen tworzących tzw. aktualny kontekst przetwarzania, budowany doraźnie przez interpreter na podstawie analizy zapytania i obrazu bazy wiedzy systemu. Kontekst może zmieniać się pod wpływem działań już podjętych na mocy pewnych reguł lub pod wpływem uzupełnień dokonywanych w bazie faktograficznej. (por. tzw. efekty kontekstowe w teorii relewancji Sperbera i Wilson — rozdz. 2.4) Stąd też reguły produkcji znajdują zastosowanie w systemach samouczących, w których niezbędna jest struktura zbiorów informacyjnych zapewniająca łatwość ich modyfikacji. Wyrażenia tworzące aktualny kontekst mogą być prezentowane w postaci prostego wykazu lub w formie ustrukturalizowanej.

Przykład prostego systemu, w którym reguły produkcji służą do identyfikacji produktów żywnościowych, podaje Barr [Barr A., Fegenbaum E. A., 1982]. Kontekst tworzy tu wykaz (*CL* — context list). Identyfikowany obiekt oznaczono symbolem *X*; wyrażenie *On-CL X* oznacza przetwarzanie informacji o danym obiekcie w kontekście. System ten posiada bazę reguł oraz interpreter sterujący wyszukiwaniem informacji niezbędnej do zidentyfikowania produktu:

P1. *IF On-CL green THEN Put-On-CL produce*

(JEŚLI w kontekście zielony to umieść na liście wytwór)

P2. *IF On-CL packed in small container THEN Put-On-CL delicacy*

(JEŚLI w kontekście *zapakowany w mały pojemnik* umieść na liście *delicates*)

P3. *IF On-CL refrigerated OR On-CL produce THEN Put-On-CL perishable*

(JEŚLI w kontekście *chłodzony* lub *wytwór TO* umieść na liście *łatwo psujący się*)

P4. *IF On-CL weighs 15 lbs AND On-CL inexpensive AND NOT On-CL perishable THEN Put-On-CL staple*

(JEŚLI w kontekście *waży 15 funtów* i w kontekście *niedrogi* I *NIE psujący się TO* umieść na liście *podstawowy*)

P5. *IF On-CL perishable AND On-CL weighs 15 lbs THEN Put-On-CL turkey*

(JEŚLI w kontekście *łatwo psujący się* I *waży 15 funtów TO* umieść na liście *indyk*)

P6. *IF On-CL weighs 15 lbs AND On-CL produce THEN Put-On-CL watermelon*

(JEŚLI w kontekście *waży 15 funtów* I *wytwór TO* umieść na liście *arbuz*)

Zadaniem interpretera jest wyszukanie tych wszystkich formuł (reguł), których część warunkowa (lewostronna) odpowiada kryterium prawdziwości. Kryterium to sprawdzane jest przez interpreter metodą podobną do pytań quizowych. Potwierdzenie spełnienia warunku określonego w lewej części formuły implikacyjnej oznacza spełnienie kryterium prawdziwości w procesie wnioskowania i tym samym stwarza podstawę podjęcia działania (action) wskazanego w prawej części formuły. W tym momencie system podejmuje działania związane z wyszukiwaniem i transformacją informacji, które omawiamy w rozdziale 4. Często okazuje się, że warunek prawdziwości (na gruncie wiedzy systemu) spełnia więcej niż jedna formuła. Dzieje się tak w przypadku, gdy cechy obiektów systemu odwzorowane są w postaci prostej, nieustrukturalizowanej listy reguł, w której pełnią funkcję poprzednika zdania warunkowego *If ... then*.

Reguły produkcji są często stosowane w systemach inteligentnych ze względu na stosunkową prostotę reprezentacji wiedzy o obiektach i łatwość manipulowania nimi w tzw. aktualnym kontekście. Cecha modularności w organizacji wiedzy ułatwia z jednej strony rozbudowę bazy reguł (względnie niezależnych), z drugiej zaś stwarza problemy z konstruowaniem przez system aktualnego kontekstu pytania w przypadku, gdy spełnienie warunku określonego w lewostronnej części reguły dotyczy pewnego większego zbioru reguł wyznaczających ciągi (serie) wzajemnie powiązanych działań. Powstają przez to trudności z odwzorowaniem w programie pełnego algorytmu transformacji informacji niezbędnych do rozwiązania określonego problemu.

Zalety i ograniczenia reguł produkcji decydują o ich przydatności jedynie w pewnych dziedzinach, które cechuje względna niezależność naturalnego sposobu opisu obiektów i ich cech będących przedmiotem badań [Nillson N., 1973].

Do dziedzin tych należą:

a) medycyna kliniczna, w której wiedzę tworzy zbiór rozproszonych faktów (deklaratywnie ujmowanych cech obiektów), w przeciwieństwie do fizyki reprezentowanej w spójnych teoriach;

b) systemy diagnostyki medycznej, w których wiedza jest reprezentowana w postaci zbioru względnie niezależnych działań, w przeciwieństwie na przykład do systemów operacji finansowych, powiązanych we wzajemnie zależne podprocesy;

c) biologia (taksonomia), w której stosunkowo łatwo jest oddzielić wiedzę o obiektach badań od sposobu jej użycia (tj. reguł transformacji), przeciwstawiana na przykład wiedzy farmaceutycznej.

Przykładem systemu, w którym wykorzystano reguły produkcji jest *MYCIN* stosowany w diagnostyce medycznej w zakresie doboru terapii dla pacjentów z bakteremią (bacteremia) lub zapaleniem opon mózgowych. Rozbudowa bazy reguł w systemie *MYCIN* realizowana jest za pomocą podsystemu *TEIRESIAS*, wyspecjalizowanego w modyfikacji bazy danych deklaracyjnych i proceduralnych [Shortlife E., 1976].

3.4.2.6. Teoria ram i skryptów

Sformułowane w połowie lat siedemdziesiątych przez M.Minsky'ego założenia teorii ram stanowiły propozycję metody reprezentacji wiedzy łączącej, z punktu widzenia technologii przetwarzania, cechy reprezentacji deklaracyjnych i proceduralnych, zaś z teoretycznego punktu widzenia uwzględniały wyniki badań nad poznaniem (kognitywnych) prowadzonych w psychologii i lingwistyce. [Minsky M. A., 1975] Podobnie jak nazwy metod omówionych wyżej, terminy skrypt, rama (*frame, script*) stosowane są w literaturze z dość dużą dowolnością do oznaczenia bardzo różnych rozwiązań aplikacyjnych. Bezpośrednią inspiracją do sformułowania koncepcji ram i skryptów było stwierdzenie nieadekwatności metod dotychczas stosowanych (rachunku predykatów, sieci semantycznych) w systemach komunikujących się z użytkownikiem w języku naturalnym. W szczególności zauważono, że nie sposób rozwiązać problemu uczenia komputerów rozumienia tekstów przy równoczesnym pomijaniu problemu organizacji pamięci człowieka zawierającej informację służącą jako kontekst, w którym przetwarzana jest nowa informacja. Teoria semantyczna języka powinna odwoływać się do teorii pamięci lub wręcz być jej częścią. Argumentów za takim podejściem dostarczały ponadto eksperymenty psychologiczne w zakresie procesów rozumienia tekstów.

Założenia teorii Minsky'ego, rozwijane m.in. przez Schanka i Abelsona określały więc potrzebę strukturalizacji wiedzy w bazie danych systemu tak, aby w interpretacji sytuacji odwzorowanych w tekście języka naturalnego zaangażowana była wiedza pochodząca z wcześniejszych doświadczeń, zorganizowana w różnych strukturach pamięciowych i przez nie sterowana. Początkowo przypuszczano, że rozwiązaniem problemu będzie stworzenie modeli reprezentacji sytuacji w pamięci stanowiących predykcje dotyczące stereotypowych obiektów w danej sytuacji oraz stereotypowych sekwencji zdarzeń,

które wyznaczają cechy dystynktywne (typowe) kontekstu, w którym interpretowana jest nowa informacja. [Schank R. C., Abelson R. P., 1975]

Teoria ram i skryptów jako podstawowych schematów reprezentacji wiedzy sytuacyjnej, została przedstawiona w dwu wersjach. Wersja wcześniejsza została sformułowana w latach siedemdziesiątych. Wersja zmodyfikowana, uwzględniająca dynamiczną teorię pamięci, została przedstawiona przez R. Schanka w pierwszej połowie lat osiemdziesiątych [179, 181]. Przedstawimy je więc w układzie chronologicznym, jako zasady reprezentacji wiedzy o obiektach i sytuacjach. [Schank R. C., Abelson R. P., 1975; Schank R. C., 1984]

W wersji wcześniejszej zakładano, że rama jest pewną ogólną strukturą, umożliwiającą odwzorowanie cech dystynktywnych obiektów pozajęzykowych oraz sytuacji, których elementami są pewne obiekty. Obiekt reprezentowany jest przez nazwę ramy, zaś jego atrybuty przez tzw. klatki (*slots*) powiązane ze znaczeniem nazwy ramy relacjami paradygmatycznymi (np. relacją generyczną) i/lub skojarzeniowymi. Ramy reprezentujące klasy pewnych obiektów również mogą być zorganizowane za pomocą relacji paradygmatycznych (generycznych), które umożliwiają przeprowadzenie wnioskowania uogólniającego, na mocy dziedziczenia cech dystynktywnych, często niezbędnego do interpretacji opisu nowych obiektów. W literaturze w zakresie sztucznej inteligencji zasadę przypisywania cech dystynktywnych obiektom, reprezentowanym przez ramy związane relacją generyczną, określa się jako hierarchię dziedziczenia (*generalization/inheritance hierarchy*) odpowiadającą relacji prototyp — egzemplarz prototypu.

Nazwom ram (obiektów) oraz ich atrybutów (klatkom), interpretowanym jako zmienne przybierające określone wartości w charakterystykach wyszukiwawczych obiektów nazywanych deskrypcjami (*description*), przypisane są reguły wnioskowania, a więc elementy wiedzy proceduralnej. Dzięki temu cechy dystynktywne pewnych obiektów mogą być automatycznie przypisywane (przez przeniesienie) opisom innych obiektów, należących do tej samej klasy. Przykład zawartości ramy prototypowej oraz ramy egzemplarza prototypu ilustrują deskrypcje ram *Chair* (krzesło) i *John's chair* (krzesło Johna):

CHAIR Frame

<i>Specialization-of:</i>	<i>FURNITURE</i>	(Odmiana: MEBLA)
<i>Number-of-legs:</i>	<i>an integer</i> (<i>DEFAULT=4</i>)	(Liczba nóg: liczba całkowita) (DOMYŚLNA = 4)
<i>Style-of-back:</i>	<i>straight,</i> <i>cushioned</i>	(Rodzaj oparcia: proste, poduszkowe)
<i>Number-of-arms:</i>	<i>0,1 or 2</i>	(Liczba podparć: 0, 1 lub 2)

JOHN'S CHAIR Frame

<i>Specialization:</i>	<i>CHAIR</i>	(Odmiana: MEBLA)
<i>Number-of-legs:</i>	<i>4</i>	(Liczba nóg: 4)
<i>Style-of-back:</i>	<i>cushioned</i>	(Rodzaj oparcia: poduszkowe)
<i>Number-of-arms:</i>	<i>0</i>	(Liczba podparć: 0)

Reprezentacja sytuacji interpretowanych na poziomie proceduralnym jako sekwencja zdarzeń różni się od reprezentacji obiektów lub ich klas tym, że w ramie sytuacyjnej wyróżnia się na ogół więcej niż jeden obiekt z odpowiednimi atrybutami, zaś relacje między tymi obiektami są relacjami czasowo-przestrzennymi i/lub przyczynowo-skutkowymi, zachodzącymi między obiektami w wyniku realizacji pewnych działań, również scharakteryzowanych za pomocą określonych cech dystynktywnych. Wartości atrybutów obiektów i sytuacji mogą być reprezentowane przez wartości stałe lub domyślne. Odwzorowanie cech obiektów i działań w sytuacji prototypowej lub podrzędnej generycznie powiązane jest z odwzorowaniem cech (atrybutów) działań, reprezentowanych przez odpowiednie klatki — tzw. klatki kolejności zdarzeń (*Event-Sequence Slot*). Klatkom odwzorowującym atrybuty zdarzeń przypisane są odpowiednie reguły wnioskowania. Tworzą one w sumie uporządkowane czasowo i przestrzennie charakterystyki wyszukiwawcze sytuacji nazywane skryptami (*scripts*). Jako klasyczny przykład skryptu cytowany jest w literaturze opis ramy *RESTAURANT (1)*, powiązanej ze skryptem *EAT-AT-RESTAURANT (Jedzenie w restauracji) (2)* [Schank R. C., Abelson R. P., 1984]:

1. *RESTAURANT Frame*

<i>Specialization-of:</i>	<i>Business-Establishment</i>
(Odmiana):	(Instytucja handlowa)
<i>Types:</i>	(Typy)
<i>Range:</i>	<i>(Cafeteria, Seat-Yourself, Wait-To-Be-Seated)</i>
(Zakres):	(Bufet, Miejsce dowolnie wybrane, Miejsce wskazane)
<i>Default:</i>	<i>Wait-To-Be-Seated</i>
(Domyślny):	Miejsce wskazane
<i>If-needed:</i>	<i>IF plastic-orange-counter, THEN Fast-Food</i> <i>IF stack-of-trays THEN Cafeteria</i> <i>IF wait-for-waitress-sign or reservation-made</i> <i>THEN Wait-To-Be-Seated, OTHERWISE Seat-Yourself</i>
(Jeśli-warunki):	(JEŚLI pomarańczowa-plastikowa lada TO Bar szybkiej obsługi JEŚLI sterta tac TO Bufet JEŚLI znak „oczekuj kelnerki” lub „rezerwacja” TO Oczekuj wskazania miejsca, W POZOSTAŁYCH PRZYPADKACH Wybierz miejsce sam)
<i>Location:</i>	(Lokalizacja)
<i>Range:</i>	<i>an ADDRESS</i>
(Zakres):	(Adres)
<i>Name:</i>	(Nazwa)
<i>If-needed:</i>	<i>(Look-at the menu)</i>

(Jeśli-warunki):	(Sprawdź jadłospis)
<i>Food-Style</i>	(Typ kuchni)
<i>Range:</i>	<i>(Burgers, Chinese, American, Seafood, French)</i>
(Zakres):	(Niemiecki, Chiński, Amerykański, Frutti di mare, Francuski)
<i>Default:</i>	<i>American</i>
(Domyślny):	(Amerykański)
<i>If-needed:</i>	<i>(Update Alternatives of Restaurant)</i>
(Jeśli-warunki):	(Zaktualizuj alternatywy restauracji)
<i>Times-Operation:</i>	(Godziny działania)
<i>Range:</i>	<i>A Time-of-Day</i>
(Zakres):	(Dzień)
<i>Default:</i>	<i>Open evenings except Mondays</i>
(Domyślny):	(Otwarte wieczorem z wyjątkiem poniedziałków)
<i>Payment-Form:</i>	(Sposób płacenia):
<i>Range:</i>	<i>Cash, CreditCard, Check, Washing-Dishes-Script</i>
(Zakres):	(Gotówka, Karta kredytowa, Czek, Skrypt Transakcje)
<i>Event-Sequence:</i>	(Kolejność zdarzeń) :
<i>Default:</i>	<i>Eat-at-Restaurant Script</i>
(Domyślny):	(Skrypt: Jedzenie w restauracji)
<i>Alternatives:</i>	(Alternatywy):
<i>Range:</i>	<i>all reataurants with the same FoodStyle</i>
(Zakres):	wszystkie restauracje o tym samym stylu kuchni)
<i>If-needed:</i>	<i>(Find all Restaurants with the same FoodStyle)</i>
(Jeśli-warunki):	(Wyszukaj wszystkie restauracje o tym samym stylu kuchni)

2. **EAT-AT-RESTAURANT SCRIPT**

<i>Props:</i>	<i>(Restaurant, Money, Food, Menu, Tables, Chairs)</i>
(Obiekty):	(Restauracja, pieniądze, jedzenie, jadłospis, stoły, krzesła)
<i>Roles:</i>	<i>(Hungry-Persons, Wait-Persons, Chef-Persons)</i>
(Role):	(Osoby głodne, osoby oczekujące, osoby obsługi)
<i>Point of view:</i>	<i>Hungry-Persons</i>
(Punkt widzenia):	(Osoby głodne)
<i>Time-of Occurence:</i>	<i>(Times-of-Operation of Restaurant)</i>
(Czas Zdarzenia):	(Pora funkcjonowania restauracji)
<i>Place-of-Occurence:</i>	<i>(Location of Restaurant)</i>
(Miejsce Zdarzenia):	(Lokalizacja restauracji)
<i>Event-Sequence:</i>	(Sekwencja zdarzeń):
<i>first:</i>	<i>Enter the Restaurant Script</i>
(najpierw):	(Rozpocznij od skryptu restauracja)
<i>then:</i>	<i>if (Wait-To-Be-Seated-Sign or Reservations)</i>

tation Language) Winograda i Bobrowa [Bobrov D. G., Winograd T., 1977, 213-222; Winograd T., 1980, 209-241].

Zasadniczym mankamentem scharakteryzowanej pierwotnej wersji ram i skryptów było — jak stwierdzają jej autorzy — statyczne ujmowanie struktury pamięci systemu informacyjnego i pomijanie aspektu ich rozwoju w procesach wyszukiwania i transformacji informacji. Charakterystyka struktur pamięci w tej wersji sprowadzała się w zasadzie do wyróżnienia w niej warstw chronologicznych: długo-, krótkotrwałej oraz pośredniej. Stwierdzano ponadto nieadekwatność skryptów jako sposobu odwzorowania typowej wiedzy sytuacyjnej. Kłopoty z ustaleniem typowych cech zdarzeń powstawały przy próbie opracowania nadrzędnego *skryptu wypadek samochodowy*, jakkolwiek możliwe było wskazanie cech relewantnych obiektów i działań dla związanej z ogólną sytuacją, np. *Wezwanie karetki pogotowia, Zawiadomienie policji, Proces sądowy* itp.

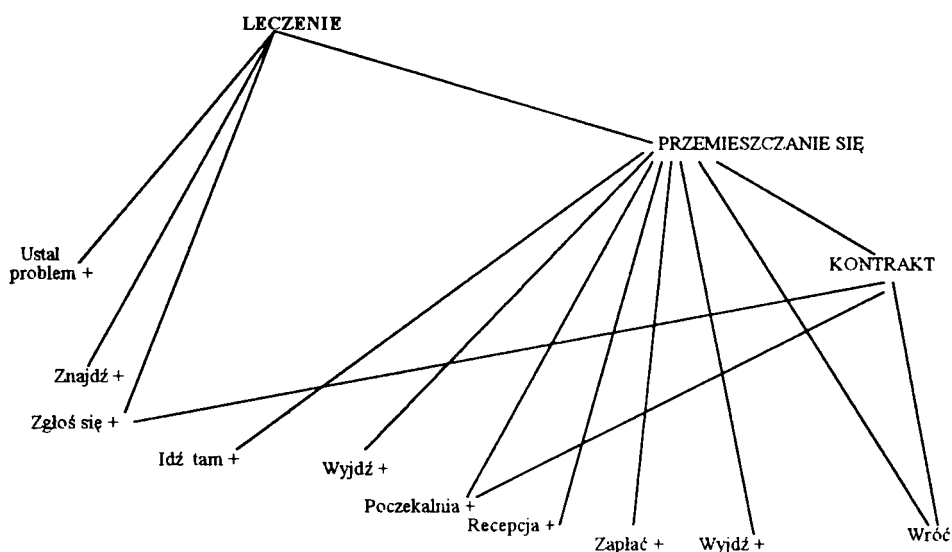
Druga zmodyfikowana wersja reprezentacji wiedzy deklaratywnej i proceduralnej opierała się na założeniu, że skrypty i ramy są elementami struktur *pamięci dynamicznej*, posiadającej różne poziomy i reorganizującej się pod wpływem przetwarzania informacji: przypominania, zapamiętywania i generowania nowej informacji [Schank R. C., 1984]. O sposobie reorganizacji struktur pamięciowych decyduje zasada *relatedness*, a więc powiązania zapamiętanych, relewantnych w danej sytuacji obiektów i ich cech oraz działań.

Reorganizacja pamięci następuje w dwu przypadkach: generalizacji (uogólnienia) doświadczeń językowych lub pozajęzykowych oraz uczenia się systemu. Istotne znaczenie ma tu proces przypominania, w którym cechy obiektów, zdarzeń (epizodów) stanowią podstawę interpretacji nowego doświadczenia, a więc rozumienia tekstu i/lub odwzorowanej w nim sytuacji pozajęzykowej. Struktury ramowo-klatkowe okazują się tu zbyt sztywną i uproszczoną strukturą pamięci, nie pozwalają bowiem na odwzorowanie wielowymiarowych powiązań między obiektami i sytuacjami. Wykorzystanie relacji paradygmatycznych między nazwami obiektów i ich cech sprowadzało się faktycznie do ich klasyfikacji monohierarchicznej z rozbudowanym aparatem odsyłaczy.

Odwzorowanie cech obiektów w dynamicznych strukturach pamięci polega najogólniej ujmując, na ich kategoryzacji połączonej ponadto z oznaczeniem ich pochodzenia z różnych typów doświadczenia: językowego, np. związanego z użyciem wyrażen spoza słownictwa potocznego oraz pozajęzykowego, związanego z percepcją wzrokową, realizacją celów i planów działania wraz z formułowanymi przy tych okazjach wnioskami i towarzyszącymi im przekonania-*mi (beliefs)*. Stwierdzono, że muszą istnieć struktury pamięci umożliwiające zapamiętywanie i wyszukiwanie informacji o obiektach należących do różnych kategorii. Podważano więc zasadę blokowej (klockowej) organizacji wiedzy, charakterystyczną dla wczesnych systemów o strukturze modularnej. Zmodyfikowane założenia możemy scharakteryzować następująco.

W opisie sytuacji reprezentowanych wcześniej przez skrypty i ramy wyróżnia się mniejsze jednostki odpowiadające układom obiektów i działań towa-

rzyszających — sceny (*scenes*); są one podporządkowane celom, określanym instrumentalnie. Tworzą pewne sekwencje odwzorowane w jednostkach pamięci nazwanych kieszeniami (*memory pockets*), różniące się od skryptów zasobem i ogólnością wiedzy przechowywanej. Kieszenie zawierają więc odwzorowania relevantnych cech obiektów i działań stanowiących predykcje w interpretacji nowych doświadczeń. Cechy relevantne różnych sekwencji działań mogą być reprezentowane w wielu wariantach, a więc w różnych kombinacjach. Wyróżnia się kieszenie uniwersalne (*UM — universal memory pocket*) i szczegółowe, oznaczane (*M*). Elementy sytuacji (sekwencji działań) uznane za relevantne mogą być wykorzystane w różnych kieszeniach, a więc różnych podzbiorach wiedzy angażowanej w przetwarzanie nowej informacji. Są ponadto skorelowane chronologicznie. Informacja nowo wprowadzana poddawana jest dekompozycji przy wykorzystywaniu wiedzy przechowywanej w różnych miejscach pamięci. Ilustrację stanowi kieszeń *leczenie* (*M-HEALTH PROTECTION*), powiązana z kieszeniami *przemieszczania się* (*M-POV*) i *kontrakt* (*M-CONTRACT*) — zob. wykres 5).



Schemat 8. Kieszeń "LECZENIE" (M-HEALTH PROTECTION)

Działania relevantne występujące w wielu kieszeniach wyróżniane są ze względu na cel, stanowiący według Schanka centrum struktur pamięciowych — na przykład związanych z reprezentacją cech sytuacji fizycznych, społecznych i jednostkowych (tj. relevantnych z punktu widzenia zachowania się danej osoby). Przechowywanie informacji o relacjach między wyróżnionymi elementami poszczególnych sytuacji umożliwiają struktury (kieszenie) poziomu wyższego — *Meta-Memory Organization Pockets (Meta-MOPs)*, sterujące ponadto wykorzystaniem zawartości niższych poziomów. Ich zasadnicza funkcja polega na uogólnianiu wiedzy.

Tworzeniu nowych pojemników wiedzy, ich koordynacji na różnych poziomach oraz ocenie relewancji poszczególnych elementów sytuacji w różnych kombinacjach mają służyć w charakteryzowanym modelu reprezentacji wiedzy w pamięci tzw. punkty organizacji tematycznej (*TOPs* — *Theme Organization Points*). Ich funkcja polega na tworzeniu rodzaju mapy wskazującej sposób poszukiwania elementów sytuacji (ram) przydatnych dla określonych celów i planów — na przykład *posiadanie, usatysfakcjonowanie kogoś przez znalezienie nowego rozwiązania problemu, dokonanie wyboru rozwiązania, udoskonalenie czegoś, zlikwidowanie sytuacji kryzysowej*. Punkty organizacji tematycznej (raczej problemowo-celowej) stanowią więc środek reprezentacji metainformacji, niezależnej od dziedziny (*domain independent*). Zdaniem autora modelu decydują one o inteligencji człowieka, zawierają bowiem wiedzę proceduralną, określającą sposób użycia wiedzy deklaratywnej, zawartej w kieszeniach (szczegółowych i uogólniających). Wyszukanie informacji o obiektach, ich cechach i działaniach za pośrednictwem punktów sterujących (*TOPs*) możliwe jest dzięki wieloaspektowemu indeksowaniu, wskaźnikom pełniącym funkcję adresów.

Scharakteryzowany model pamięci dynamicznej Schanka, stanowiący modyfikację pierwotnej wersji ram i skryptów wykazuje wiele cech wspólnych z modelem organizacji informacji mentalnej przedstawionym w ramach teorii relewancji D. Sperbera i D. Wilson. Różni się konwencjami terminologicznymi i aplikacyjnym przeznaczeniem. Zakłada wieloaspektowość odwzorowania cech obiektów pozajęzykowych, różne poziomy organizacji informacji ze względu na jej ogólność oraz niezbędność odwzorowania metainformacji na poziomie deklaratywnym i proceduralnym. Wyznacza więc różne sfery relewancji wiedzy. Model ten był wykorzystywany w konkretnych systemach interpretujących teksty języka naturalnego — w systemie *IPP (Integrated Partial Parser)*, przetwarzającym teksty agencji UPI oraz eksperymentalnym systemie *CYRUS*, interpretującym teksty prasowe o bieżących wydarzeniach politycznych (dyplomatycznych) [Schank R. C., 1984]. Przykład reprezentacji wiedzy o spotkaniach dyplomatycznych zawiera schemat 6.

Przyjęto tu zasadę „kieszeniowej” organizacji pamięci systemu w opisie zdarzeń (*episodic MOPs*) oraz drzewiastą strukturę reprezentacji zdarzeń poprzez specyfikację różnic między cechami relewantnymi (*NORMS*): uczestnicy, lokalizacja, temat, związki z innymi zdarzeniami (*DIFFERENCES*). Niektóre cechy mogą stanowić przesłanki w procesach wnioskowania uogólniającego (są predyktywne — *predictive*), na przykład na temat roli politycznej danego uczestnika. Cechy te są opatrywane specjalnymi indeksami służącymi jako podstawa generalizacji wiedzy umożliwiającej uczenie się systemu.

Przegląd najczęściej stosowanych zasad reprezentacji wiedzy w inteligentnych systemach faktograficznych prowadzi do wniosków związanych z różnymi aspektami podjętych problemów:

- zbieżności różnych interpretacji relewancji w ujęciu chronologicznym i dyscyplinowym,

- różnic w konwencjach terminologicznych, związanych z charakterystyką języków reprezentacji wiedzy w systemach dokumentacyjnych i inteligentnych faktograficznych,
- zakresem wiedzy odwzorowanej w językach systemów faktograficznych.

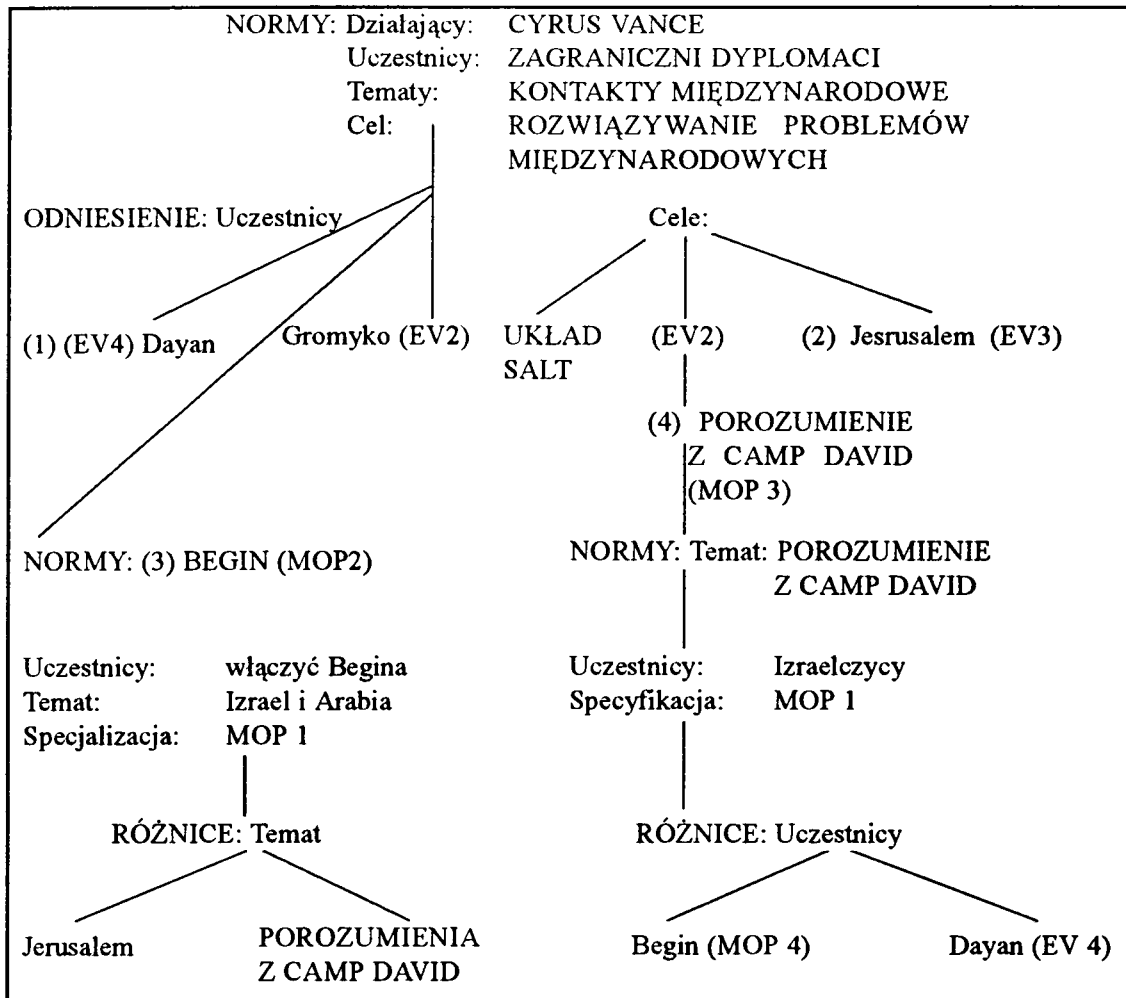
Porównanie wywodów na temat relewancji w ujęciu fenomenologicznym A. Schutza, inspirowanego poglądami sceptyka Karneadesa, prowadzi do konstatacji znacznego podobieństwa wniosków formułowanych na gruncie filozofii z wnioskami prezentowanymi w wyniku interdyscyplinarnych badań kognitywnych w teorii relewancji D. Sperbera i D. Wilson oraz z wnioskami, do których dochodzą aplikacyjnie myślący autorzy modeli i systemów faktograficznych, podlegających ewolucji na gruncie sztucznej inteligencji. Stwierdza się mianowicie wielowymiarowość i relatywność oceny relewancji informacji odwzorowanej w pewnym języku (mentalnym, naturalnym lub sztucznym) systemu komunikacyjnego.

Rozwój poglądów na temat sposobów odwzorowania relewantnych cech obiektów rzeczywistości pozajęzykowej wykazuje swoisty izolacjonizm, przejawiający się w warstwie terminologicznej prac z zakresu filozofii, teorii komunikacji i sztucznej inteligencji (por. sfery relewancji A. Schutza i teoria prawdopodobieństwa Karneadesa, zasada relewancji Sperbera i Wilson oraz teoria wielowymiarowej pamięci dynamicznej R. Schanka).

Zakres wiedzy odwzorowywanej w języku systemów faktograficznych może być praktycznie nieograniczony, język informuje bowiem o ekstensjach. Rodzaj wiedzy relewantnej sprowadza się do opisu sensów wyrażen językowych — intensji. Można je faktycznie sprowadzić do wiedzy o obiektach wyróżnianych za pomocą ich nazw oraz relacji między nimi, ujmowanymi statycznie lub dynamicznie (w sytuacjach, w których występują pewne działania).

Repertuar relacji opisujących cechy sytuacji jest identyczny jak stosowany tradycyjnie w językach systemów dokumentacyjnych, sprowadza się do relacji paradygmatycznych i skojarzeniowych, zmienia się jedynie zakres ich wykorzystania. Relacje paradygmatyczne (generyczne) są przydatne w odwzorowaniu związków między klasami obiektów i klasami ich cech wyróżnianych klasycznymi metodami językoznawstwa strukturalnego, relacje skojarzeniowe stosowane są w znacznie większym zakresie niż w systemach dokumentacyjnych w opisie sytuacji denotowanych przez teksty języka naturalnego. Zasada eksponowaną w stopniu wyższym niż w systemach dokumentacyjnych jest kojarzenie cech obiektów i działań ze względu na cel (por. punkty organizacji tematycznej Schanka).

Innym rodzajem wiedzy jest wiedza proceduralna, a więc określająca reguły transformacji semantycznej wyrażen w systemie służąca udostępnianiu użytkownikowi informacji relewantnej ze względu na założony cel. Wiedza proceduralna staje się coraz ściślej powiązana z wiedzą deklaratywną, warunkuje bowiem relewancję informacji. Omawiamy ją w rozdziale poświęconym sposobom udostępniania informacji w systemach informacyjno-wyszukiwawczych.



Schemat 9. Reprezentacja wiedzy o spotkaniach dyplomatycznych

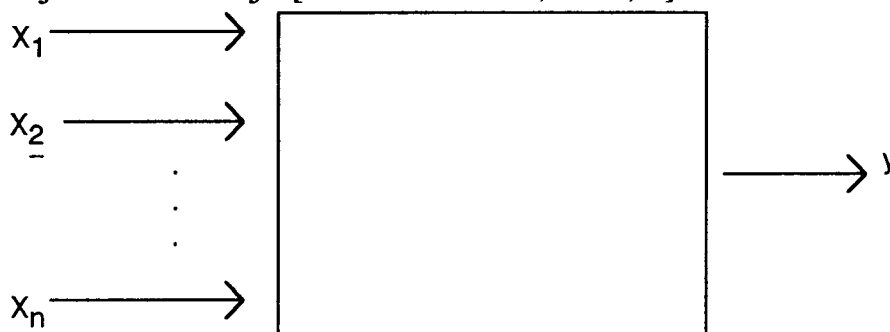
3.4.2.7. Sieci neuronowe

Sieci neuronowe są metodą reprezentacji, a przede wszystkim *transformowania i strukturalizowania informacji* opracowywaną niezależnie od metod stosowanych w systemach faktograficznych, scharakteryzowanych w poprzednich rozdziałach, zdeterminowanych przez technologiczne możliwości „klasycznej” informatyki i elektroniki. Uzupełnienie opisu tych „tradycyjnych” metod szkicowym omówieniem odrębnego nurtu badań związanych z przetwarzaniem informacji wydaje się celowe zarówno ze względu na ich genezę, jak i spodziewany (przełomowy) wpływ na rozwój technologii informatycznych (instrumentów informacyjnych SIW), wyznaczających — jak podkreślaliśmy — repertuar kryteriów relewancji informacji udostępnianej przez system.

Koncepcja sieci neuronowych powstała przed ponad pięćdziesięcioma laty na gruncie interdyscyplinarnym: najpierw neurofizjologii i matematyki,

uzupełnianych następnie przez bionikę i biocybernetykę. Obecnie tę dziedzinę badań i zastosowań sieci neuronowych nazywa się *neurokomputeringiem*. Wyjaśnienie, czym są sieci neuronowe Ryszard Tadeusiewicz, autor syntetyzującej monografii w języku polskim³ sprowadza do lapidarnego stwierdzenia — „bardzo uproszczonym modelem mózgu” [Tadeusiewicz R., 1993, 13]. Jako pierwszą, źródłową pracę, która zainspirowała powstawanie neurokomputeringu wymienia się sformułowanie przez McCullocha w 1943 r. matematycznego modelu komórki nerwowej [McCulloch W. S., 1943], do którego nawiązywały prace opublikowane w latach pięćdziesiątych i sześćdziesiątych, podejmujące aspekty techniczne i biologiczne modelowania systemu nerwowego (i/lub jego elementów), łącznie z budową *perceptronu* — pierwszego, elektromechanicznego układu funkcjonującego na zasadach neuropodobnych. Był on przeznaczony do rozpoznawania znaków alfanumerycznych i równoczesnego uczenia się [Neumann von J., 1958], [Taylor W. K., 1960].

W literaturze przedmiotu podkreśla się, że główną ideą wykorzystania własności neuronów i ich struktur, jako naturalnych systemów przetwarzających informację, było przełamanie podstawowego ograniczenia wyrażającego w sposobie przetwarzania informacji przez komputer — przetwarzania szeregowego i prostego przechowywania informacji. Za cel postawiono stworzenie teoretycznych (matematycznych) i technicznych środków przetwarzania równoległego przy równoczesnym uczeniu się systemu, a więc transformowaniu i przyswajaniu informacji. [Tadeusiewicz R., 1993, 8]



Źródło: Tadeusiewicz R., (1993), Sieci neuronowe, Warszawa: Akademicka Oficyna Wydawnicza RM 27

Punktem wyjścia pionierskich teoretycznych i aplikacyjnych modeli struktur systemu nerwowego, a nawet całego mózgu, powstałych w latach sześćdziesiątych, były obserwacje funkcjonowania prostych organizmów żywych — np. mózgu ośmiornicy, systemu wzrokowego kraba. W tym też okresie na Uniwersytecie Stanforda w USA podjęto próbę budowy pierwszego neurokomputera,

³ Wykorzystana jako podstawa niniejszej, skrótowej prezentacji praca R. Tadeusiewicza zawiera wykaz kilkuset pozycji literatury zagadnienia.

wytwarzanego przez ponad 20 lat dla celów komercyjnych i stosowanego w radarach, modemach i urządzeniach telekomunikacyjnych do tzw. adaptacyjnego przetwarzania sygnałów⁴

Niezależnie od technologii stosowanych do wytworzenia (historycznie i współcześnie) fizycznych modeli komórek nerwowych i złożoności tych modeli, sieć neuronową można najogólniej określić jako:

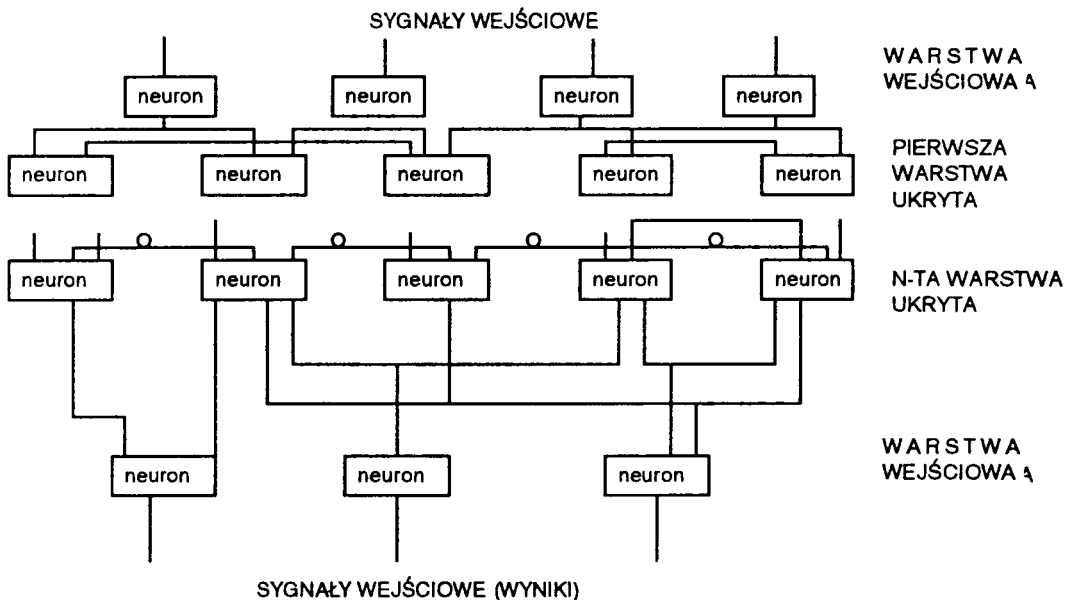
- pewien automat odbierający i transformujący sygnały przynoszące informację,
- relację lub funkcję w sensie matematycznym przyporządkowującą wielu sygnałom wejściowym jeden wynik.

W podstawowej, klasycznej wersji funkcja ta ma następujący zapis:

$$y = f(x_1, x_2, \dots, x_n)$$

gdzie: x_1, x_2, \dots, x_n — sygnały wejściowe, y — wynik na wyjściu, f — funkcja;

Podstawowa struktura neuronu przedstawiana jest graficznie w postaci:



Źródło: Tadeusiewicz R. (1993), 13.

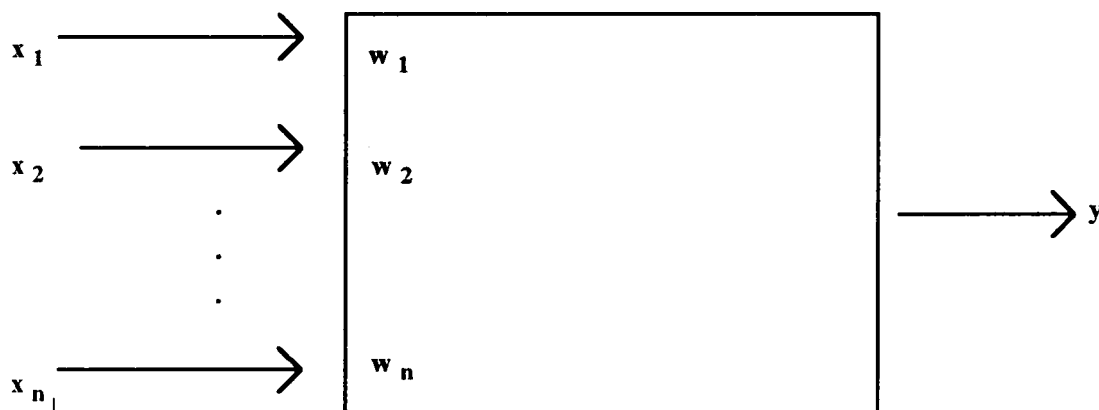
Charakterystyczną cechą struktur złożonych z podobnych elementów (sieci) jest ich wzajemne powiązanie oraz warstwowość. Wyróżnia się tzw. warstwy wejściowe, wyjściowe oraz ukryte. Ta własność struktury sieci neuronowych decyduje o możliwości równoległego i współzależnego interpretowania i transformowania sygnałów.

Interpretacja i późniejsza transformacja sygnałów (x_1, x_2, \dots, x_n) odbywa się dzięki przypisaniu im wartości nazywanych wagami synaptycznymi (analo-

⁴ W niniejszym skrótowym omówieniu sieci neuronowych abstrahujemy od technicznych aspektów rozwiązań, która to problematyka przekracza kompetencje autorki. Czytelniczka zainteresowanego technicznymi rozwiązaniami odsyłamy do obszernej polsko- i angielskojęzycznej literatury przedmiotu.

gia z naturalnymi neuronami) (w_1, w_2, \dots, w_n) wykorzystywanymi także w procesie uczenia się systemu.

Ocena i porównanie wag przeprowadzana jest w kategoriach wektorowych, a więc parametrów położenia sygnału wejściowego x w n -wymiarowej przestrzeni. Rozpoznawanie sygnałów wejściowych polega na ustaleniu podobieństwa ich wektorów do wektora określonego zestawu wag (wzorcowego obiektu). O sile wynikowego sygnału y w neuronie decyduje stopień podobieństwa wejściowego sygnału x_j do zadanego wzorca, na który reaguje dany neuron. Tak można w wielkim uproszczeniu streścić zasadę rozpoznawania różnego typu sygnałów (optycznych, akustycznych itp.). Neurony wyspecjalizowane w klasyfikacji sygnałów według zadanych wzorców nazywane są „grandmother cells”. Modelowanie skomplikowanych zależności między wagami sygnałów i strukturami neuronów w warstwach i sieciach, a więc ustalanie relacji między cechami dystynktywnymi informacji o obiektach napływającej do systemu, a następnie ich klasyfikowanie, przeprowadzane jest za pomocą matematycznego aparatu równań liniowych i nieliniowych. Naszkicowane zdolności rozpoznawania sygnałów powodują, że system jest w stanie odpowiadać na pytania typu co to jest (obiekt, cecha, proces, zjawisko itp.).



Źródło: Tadeusiewicz R., (1993), 13.

Zdolność uczenia się sieci neuronowych uzyskano poprzez wyposażenie neuronów w dwa mechanizmy nazywane procesorem zmiany wag i detektorem błędów, dzięki czemu możliwe jest korygowanie przez system wektorów wag i realizowanie postawionych zadań obliczeniowych według zadanej zależności (funkcji). Neurony mające te własności określane są *ADALINE* (*ADaptive LINear Element*).

Podstawowe mechanizmy rozpoznawania sygnałów i uczenia się przez sieci neuronowe, opracowane w latach sześćdziesiątych dla modeli złożonych zaledwie z 8 neuronów o 128 połączeniach (*ADALINE*, *MADALINE*) wykorzystano w rozwijaniu późniejszych, coraz bardziej złożonych sieci rezonansowych, sieci ze sprzężeniem zwrotnym i in. (najnowsze dysponują $1.5 \cdot 10^6$ połączeń). Zamierzeniem docelowym jest stworzenie dwukierunkowej pamięci-

ci skojarzeniowej (*BAM* — *Bidirectional Associative Memory*), której koncepcję sformułowano na początku lat osiemdziesiątych. Jej funkcjonowanie można najogólniej ująć w następujących zasadach:

- wzajemnego oddziaływania sygnałów wejściowych (x) i wyjściowych (y) w zamkniętej pętli sprzężenia zwrotnego;
- wprowadzenia dwu typów wektorów sygnałów wejściowych w_i oraz w_j ,
- rejestrowania (uczenia się) zależności między wejściem i wyjściem, co oznacza tworzenie wewnętrznych obiegów sygnałów w sieciach, umożliwiających „przypominanie” zapamiętanej informacji.

Właśnie z tym ostatnim kierunkiem prac badawczo-wdrożeniowych związane są oczekiwania przelomu w technologii informatycznej, ze szczególnym odniesieniem do sztucznej inteligencji⁵.

Zakres zastosowania skrótowo przedstawionych zasad strukturalizacji i transformacji informacji wynika z ograniczonej interpretacji sygnałów, która sprowadza się do rozpoznawania i przekształcania parametrów numerycznych reprezentujących obiekty fizyczne w rzeczywistości. W związku z tym techniczne rozwiązania funkcjonujące w praktyce dotyczyły w pierwszej kolejności przestrzennej identyfikacji obiektów wyraźnie zdefiniowanych: graficznych (rozpoznawanie znaków pisma), akustycznych (synteza i rozpoznawanie sygnałów mowy), przedmiotów o określonych cechach (kontrola bagażu na lotniskach ze względu na ew. zawartość materiałów wybuchowych), ruchu (sterowanie, koordynacja i planowanie ruchów robotów w automatyce).

Wykorzystanie sieci neuronowych jako metody reprezentacji i przekształcania informacji numerycznej (danych odwzorowujących ilościowe cechy obiektów w rzeczywistości) poza techniką obejmuje w pierwszej kolejności sparametryzowane dane ekonomiczne (statystyczne), związane z funkcjonowaniem gospodarki, rynku i finansów, niezbędne na przykład dla celów takich jak:

— predykcja: ocena i prognozowanie rynków, tendencji w gospodarce, ocena zdolności kredytowej przedsiębiorstw;

— identyfikacja i klasyfikowanie obiektów gospodarczych na podstawie danych bilansowych połączone z oceną ich atrakcyjności inwestycyjnej;

— analiza i kojarzenie danych przez ich powiązanie z procesami wnioskowania, prowadzące do wykrycia pewnych zależności lub ustalania przyczyn niepowodzeń gospodarczych, niemożliwe przy zastosowaniu tradycyjnego programowania;

⁵ Warto odnotować podkreślany w literaturze przedmiotu fakt znacznego zahamowania prac nad sieciami neuronowymi w latach siedemdziesiątych pod wpływem krytycznej oceny możliwości zastosowania sieci jednowarstwowych, opublikowanej przez M. Minsky'ego cytowanego w poprzednich rozdziałach, jednego z czołowych twórców sztucznej inteligencji. Przelamanie tego impasu (wraz z renesansem problematyki) miało nastąpić dopiero w latach osiemdziesiątych po opublikowaniu pracy [Anderson J.A., Rosenfeld E. (1988), *Neurocomputing — Foundations of Research*, Cambridge: MIT Press]; Por. [Tadeusiewicz R. (1993), 11]

— filtracja sygnałów, analogiczna jak w telekomunikacji czy automatyce, polegająca na eliminowaniu szumów informacyjnych i pozwalająca usuwać tzw. klasyczne zakłócenia (szумы) o charakterze losowym występujące w statystyce, utrudniające procesy decyzyjne;

— optymalizacja procesów decyzyjnych w sytuacjach gospodarczych, a zwłaszcza tych, które wiążą się z trudnym wyborem optymalnego wariantu w warunkach eksplozji kombinatorycznej. Jako klasyczny przykład w tej sferze wymieniane jest rozwiązanie tzw. problemu komiwojażera (*travelling salesman problem*) — ustalenie optymalnej trasy objazdu n miast przez akwizytora i maksymalne oszczędzenie kosztów podróży. Problem ten rozwiązywany jest przy zastosowaniu tzw. sieci Hopfielda, poprzez ocenę parametrów możliwych kombinacji n odległości (d), wynoszących $D=n!/2n$; np. dla $n=60$, liczba ta wynosi $D=69,34155 \cdot 10^{78}$ [Tadeusiewicz R., 1993, 94]

W literaturze podkreśla się, że o przewadze sieci neuronowych nad metodami reprezentacji i transformacji wiedzy w „tradycyjnych” systemach faktograficznych decyduje zdolność uczenia się systemu i wynikająca stąd możliwość przewidywania sygnałów wyjściowych (wyników) bez konieczności formułowania apriori hipotez o związku między przesłanką i możliwym wnioskiem (por. rozdz. 3.4.2.6 — Teoria ram i skryptów).

Trudno obecnie wskazać przykłady wykorzystania sieci w przetwarzaniu tekstów języka naturalnego połączonym z interpretacją semantyczną wyrażen sensu stricto. Teoretycznie, nie wydają się istnieć przeszkody w wykorzystaniu w systemie semantycznej identyfikacji obiektów językowych na podstawie zdolności wykrywania sieci relacji generycznych, mereologicznych i skojarzeniowych (scharakteryzowanych w rozdziale 1) oraz rozbudowy tych relacji na podstawie własności kontekstowych i statystycznych wyrażen. Ten kierunek zastosowań można przypuszczalnie łączyć z zamierzeniami budowy wskazanej wyżej dwukierunkowej pamięci skojarzeniowej (BAM) i technicznego doskonalenia instrumentów informacyjnych (konstruowania neurokomputerów). Oznaczałoby to z jednej strony zbliżenie sposobów odwzorowania, transformowania i udostępniania informacji w systemach do „naturalnych” sposobów, zgodnie z postulatami badań kognitywnych (por. teoria relewancji Sperbera i Wilson), z drugiej zaś — pozwoliłoby przełamać ograniczenia w gromadzeniu i przetwarzaniu wiedzy przyjętych w sztucznej inteligencji (systemach eksperckich) dzięki wspieraniu badań nad sieciami przez obserwacje neurologiczne dotyczące struktury i własności procesów uczenia się.

4. DOSTĘP DO WIEDZY W SYSTEMIE INFORMACYJNO-WYSZUKIWAWCZYM

W dotychczasowych rozważaniach koncentrowaliśmy się na zagadnieniu sposobów reprezentacji wiedzy w zbiorach systemu informacyjno-wyszukiwawczego, charakteryzując języki służące do odwzorowania wiedzy o dokumentach oraz o obiektach pozadokumentacyjnych (pozajęzykowych) zgodnie z pewnymi kryteriami relewancji (rozd. 1, 2). W ten sposób uwzględniony został jedynie ten aspekt funkcjonowania systemu, który wiąże się z gromadzeniem wiedzy, jej kodowaniem, organizacją oraz przechowywaniem. Gromadzenie i strukturalizacja wiedzy według odwzorowanych w języku informacyjnym kryteriów relewancji należą do tych procesów informacyjnych, które poprzedzają udostępnianie informacji przechowywanej w systemie drogą jej transformacji, a więc przetwarzania.

Potrzebę wyjaśniania różnych znaczeń terminów „*udostępnianie*” (korzystanie z) wiedzy systemu sformułowano explicite dopiero w pracach związanych z rozwojem systemów inteligentnych, wskazując dwa sposoby używania zbiorów systemu:

- wyszukanie informacji relewantnej, a więc jej wyselekcjonowanie,
- wygenerowanie informacji relewantnej w procesach wnioskowania, na podstawie wiedzy deklaratywnej i reguł transformacji prawdziwościowej (wiedzy proceduralnej) [Barr A., Fegenbaum E. A., 1981].

Rozróżnienie to przyjmujemy za podstawę prezentacji zagadnień w tej części pracy, pozwala bowiem wyraźnie wskazać te systemy, w których:

1) udostępnianie informacji drogą selekcjonowania określane jest jedynie przez relację podobieństwa charakterystyki wyszukiwawczej dokumentu lub obiektu pozadokumentacyjnego do instrukcji wyszukiwawczej (systemy dokumentacyjne, proste systemy faktograficzne typu obiekt-cecha) — *systemy udostępniające wiedzę bez transformacji (nie-dedukcyjne)*;

2) udostępnianie informacji polega na realizowaniu przez system transformacji semantycznej wyrażen z zbioru systemu, będącej równocześnie transformacją prawdziwościową, a więc taką, że z wyrażen wyjściowych (przesłanek) uznawanych w systemie za prawdziwe otrzymuje się inne wyrażenia (wnioski) prawdziwe w stopniu takim samym, jaki przysługuje przesłankom (wnioskowanie logiczne) lub mniejszym (wnioskowanie uprawdopodobniające) — *systemy udostępniające wiedzę po jej transformacji (dedukcyjne)*. Odmianą transformacji semantycznej i prawdziwościowej realizowanej w niektórych syste-

mach inteligentnych jest wyjaśnianie, polegające na dobraniu do danego wyrażenia uznanego za prawdziwe takiego wyrażenia ze zbioru systemu, że mogłoby być przesłanką w procesie wnioskowania.

W literaturze z zakresu języków i systemów informacyjnych termin transformacja stosowany jest często w znaczeniu bliższym selekcjonowaniu informacji — na przykład streszczanie tekstów dokumentów pierwotnych interpretowane jest jako transformacja tekstu (połączona ze zmianą języka naturalnego na inny język naturalny), w wyniku której powstaje inny tekst krótszy, zawierający wyrażenia będące wykładnikami relewancji. Transformacje wyrażen językowych dokonywane w procesie przekładu zapytania informacyjnego na instrukcję wyszukiwawczą i formułowania strategii wyszukiwawczej w ujęciu teorii pytań (Ajdukiewicza, Marciszewskiego), które charakteryzujemy w dalszej części tego rozdziału, są transformacjami semantycznymi i prawdziwościami zarazem, wyznaczanymi przez funkcję presupozycji.

Sprowadzenie różnic w sposobach udostępniania wiedzy przez system do dokonywania (lub nie) transformacji prawdziwościowej informacji w procesach wnioskowania byłoby jednak dużym uproszczeniem prowadzącym do nieporozumień. Należy zauważyć, że chodzi tu właściwie o moment, w którym odbywa się transformacja oraz o jej realizatora.

W systemach tradycyjnych, w których ocena relewancji i pertynencji informacji wyselekcjonowanej przez system odbywa się na podstawie końcowych wyników wyszukiwania, proces prawdziwościowej transformacji informacji poprzedza realizację strategii wyszukiwawczej. Dążenie do zapewnienia maksymalnej kompletności i relewancji wyników wyszukiwania powoduje, że cała wiedza o rzeczywistości pozajęzykowej jest wbudowana w język informacyjno-wyszukiwawczy, łącznie z wiedzą językową.

W systemach dialogowych online (dokumentacyjnych i faktograficznych), w których użytkownik komunikuje się z systemem według zasady *akcja-reakcja*, zależnie od oceny pertynencji wyszukanej informacji transformacja odbywa się w trakcie sesji wyszukiwawczej i polega na zmianie zbioru wyrażen składowych instrukcji wyszukiwawczej, odwzorowujących pytanie informacyjne użytkownika i wskazane kryteria relewancji. Wpływa to na zakres wykładników relewancji wbudowanych w język informacyjno-wyszukiwawczy, odwołujących się do wiedzy językowej i pozajęzykowej.

W systemach inteligentnych faktograficznych transformacja semantyczna i prawdziwościowa informacji zawartej w pytaniu, polegająca na wykorzystaniu jego wyrażen składowych jako przesłanek w procesie wnioskowania, prowadzącego do wygenerowania odpowiedzi relewantnej (najczęściej ze względu na cel), realizowana jest przez system formułujący strategię wyszukiwawczą na podstawie własnej interpretacji pytania. System określa więc samodzielnie algorytm realizacji funkcji wyszukiwawczej na podstawie wiedzy proceduralnej oraz wiedzy deklaratywnej o obiektach pozajęzykowych ich cechach oraz sytuacjach, w których są użyte.

Stopniowe wbudowywanie w system środków umożliwiających transformację prawdziwościową informacji i formułowanie strategii wyszukiwawczej wpływa także na sposób komunikowania się z użytkownikiem. W systemach tradycyjnych funkcję metainformacyjną i wyszukiwawczą pełni w zasadzie jeden język (np. UKD, język haseł przedmiotowych w katalogu bibliotecznym). W systemach online wyodrębnia się język kwerend, stosunkowo niezależny od języków informacyjnych baz danych. W systemach inteligentnych następuje oddzielenie się modułu komunikacji w języku zbliżonym do naturalnego. Wiedza o języku i o rzeczywistości pozajęzykowej prezentowane są względnie niezależnie.

4.1. STRATEGIA WYSZUKIWAWCZA

Interakcyjny lub nieinterakcyjny tryb przesyłania komunikatów (instrukcji wyszukiwawczej i odpowiedzi o wynikach wyszukiwania) między systemem i użytkownikiem wyznacza skalę możliwości prezentacji kryteriów relewancji oraz transformacji niezbędnych do wyszukania informacji relewantnej, a więc określenia algorytmu selekcjonowania i/lub generowania odpowiedzi.

Przechodzimy w ten sposób do kolejnego istotnego problemu w teorii systemów i języków informacyjno-wyszukiwawczych, jakim jest formułowanie strategii wyszukiwawczej. Poświęcono mu wiele uwagi, zwłaszcza w tym okresie, gdy wyszukiwanie informacji realizowane było w trybie nieinterakcyjnym (wsadowym), stanowiąc źródło kłopotów natury technicznej i merytorycznej przy określaniu relewantnych cech informacji o dokumentach. W kontekście rozważań nad metodami formułowania optymalnej strategii podnoszono problem arbitralności reprezentacji wiedzy w systemie i jej nieadekwatności w stosunku do potrzeb użytkowników.

W charakterystykach procesu wyszukiwania informacji w systemie używa się terminów *strategia wyszukiwawcza*, *metody wyszukiwania informacji* oraz *instrukcja wyszukiwawcza*, *klucz wyszukiwawczy*, które nie zawsze definiowane są w sposób wyraźny. Szczególne problemy stwarza rozróżnienie strategii wyszukiwawczej oraz metod wyszukiwania odnoszących się do różnych poziomów opisu procesu selekcjonowania informacji.

Meadow określa strategię wyszukiwawczą jako „plan układu i kolejności stawiania pytań przez przeszukującego w trakcie realizacji określonego zapotrzebowania na informację”. Odróżnia strategię wyszukiwawczą od technicznych metod selekcjonowania informacji utrwalonej na określonym nośniku, które nazywa planem wyszukiwania („metoda przeszukiwania kartoteki”). Podkreśla przy tym, że metoda wyszukiwania na poziomie realizacji technicznej jest stałym elementem strategii wyszukiwawczej w systemie, podczas gdy strategia zmienia się z każdym pytaniem. Metody wyszukiwania w planie realizacji technicznej definiuje jako „funkcję organizacji zbioru informacji oraz logiki, jaką dopuszcza dane pytanie informacyjne” [Meadow Ch. T., 1972].

Charakterystyka strategii wyszukiwawczej (*search algorithms*) podana przez Saltona stanowi właściwie opis technicznych metod przeszukiwania zbioru informacji według kryterium określonego przez tzw. *klucz wyszukiwawczy* [Salton G., 1975]. Jest nim wyrażenie pełniące funkcję metatekstową, wskazujące wyrażenia relewantne w danym zbiorze informacyjnym — na przykład rodzaj tekstu utrwalonego w zbiorze, elementarną jednostkę leksykalną w charakterystyce wyszukiwawczej lub zapytaniu informacyjnym.

W ujęciu Lancastera strategię wyszukiwawczą tworzą operacje logiczne, określone w instrukcji wyszukiwawczej. Interpretacja ta zawęży z kolei strategię do reguł gramatyki języka informacyjno-wyszukiwawczego [Lancaster F W., 1968].

Ungurian określa strategię wyszukiwawczą jako *schemat wyszukiwania*, wymieniając jej elementy obligatoryjne: reguły przekształcania charakterystyki wyszukiwawczej i instrukcji wyszukiwawczej w zapis formalny, metody przeszukiwania zbioru charakterystyk wyszukiwawczych i zbioru leksyki języka informacyjno-wyszukiwawczego, metody mierzenia stopnia relewancji, ustalenia dotyczące tzw. funkcji decyzji stanowiące kryterium uznania dokumentu za relewantny lub nierelwantny w systemie (może to być decyzja binarna lub graniczna, względna) oraz reguły modyfikacji instrukcji wyszukiwawczej — tzw. reguły sprzężenia zwrotnego. Podana przez tego autora charakterystyka elementów składowych strategii obejmuje zarówno metody w sensie technicznym, jak i elementy strategii w rozumieniu Meadowa, przy czym nie jest jasne sformułowanie mówiące o regułach przekształcania charakterystyki wyszukiwawczej w zapis formalny, ponieważ są one — jako teksty języka sztucznego (informacyjno-wyszukiwawczego) — sformalizowane [Ungurian O., 1976].

Czerny definiuje strategię wyszukiwawczą (strategia poszuka) jako „... przewidziane w instrukcji wyszukiwawczej operacje logiczne dokonywane w procesie wyszukiwania, z uwzględnieniem kolejności ich przeprowadzania — w celu osiągnięcia optymalnych rezultatów...” [Czerny A. I., 1978].

W literaturze poświęconej wyszukiwaniu informacji w systemach dialogowych strategia wyszukiwawcza (*search strategy*) definiowana jest najczęściej jako zbiór zaplanowanych instrukcji wyszukiwawczych (*search statement*), określających sposób wyszukiwania w systemie informacji, stanowiącej odpowiedź na zapytanie użytkownika [Guide to Dialog Searching, 1978].

W *Słowniku terminologicznym informacji naukowej* brak jest hasła strategia wyszukiwawcza, zdefiniowano jedynie instrukcję wyszukiwawczą jako treść kwerendy (zapytania informacyjnego) wyrażoną w języku informacyjnym [Słownik terminologiczny..., 1979].

W związku z tym, iż przytoczone definicje niezbyt wyraźnie określają relacje między znaczeniami analizowanych terminów, konieczne jest wprowadzenie pewnych uściśleń. Pierwsze dotyczy potrzeby rozróżnienia metod wyszukiwania informacji w sensie technicznym oraz strategii wyszukiwawczej. Przykładów takich rozróżnień dostarczają w systemach skomputeryzo-

wanych języki programowania, wśród których mamy wyspecjalizowane w funkcji opisu struktury informacji (struktury danych) — tzw. języki opisu danych (*data description language*) oraz wyspecjalizowane w transformacji informacji na poziomie realizacji technicznej — języki manipulowania danymi (*data manipulation language*). Techniczne (informatyczne) metody realizacji procesów wyszukiwania informacji wykraczają poza ramy tej pracy.

Charakterystyka strategii wyszukiwawczej związana jest z opisem procesu wyszukiwania informacji z punktu widzenia użytkownika systemu i projektanta języka informacyjno-wyszukiwawczego. Mając na uwadze to rozróżnienie *strategię wyszukiwawczą* można zdefiniować jako algorytm realizacji funkcji wyszukiwawczej w systemie informacyjno-wyszukiwawczym w celu zapewnienia optymalnych wyników wyszukiwania informacji, określonych przez kryteria relewancji, kompletności i pertynencji.

W przytoczonej definicji Czernego instrukcja wyszukiwawcza i strategia zostały w zasadzie sprowadzone do jednego poziomu, ponieważ instrukcja jest traktowana jako odwzorowanie strategii, co nie wydaje się słuszne.

Przyjmując drugie rozróżnienie między strategią wyszukiwawczą i instrukcją wyszukiwawczą, zdefiniujemy *instrukcję* jako tekst języka informacyjno-wyszukiwawczego odwzorowujący treść zapytania informacyjnego. Definicje podawane przez cytowanych autorów są w tym względzie zbieżne. W przypadku najprostszym strategia może być odwzorowana tylko w jednej instrukcji — na przykład w przypadku zapytania o dokumenty tylko jednego autora w systemie informacji dokumentacyjnej lub też zapytania o wykaz osób, które urodziły się 1.01.1950 r., w systemie informacji faktograficznej. Najczęściej strategia, rozumiana jako algorytm, a więc uporządkowany zbiór transformacji opisanych w pewnym języku, określających sposób selekcjonowania informacji w zbiorze, jest odwzorowana w więcej niż jednej instrukcji wyszukiwawczej. Zależy to w dużej mierze od instrumentów informacyjnych systemu. Kryterium, któremu podporządkowane są kolejne transformacje umożliwiające wyselekcjonowanie pożądanej informacji jest określone przez wskazany w tekście instrukcji klucz wyszukiwawczy. Należy przy tym zaznaczyć, że termin klucz wyszukiwawczy jest stosowany także na poziomie informatycznego opisu struktury zbioru informacji i technicznych metod jej wyszukiwania, w znaczeniu wyrażenia pełniącego funkcję adresu danego ciągu znaków, lokalizującego ten ciąg w zbiorze znaków.

Podstawą do określania strategii wyszukiwawczej i sformułowania instrukcji, stanowiącej element tej strategii, jest *zapytanie informacyjne* (*kwerenda*, *pytanie informacyjne*). Jest to pytanie skierowane do systemu, określające cechy dystynktywne informacji, będące dla danego użytkownika informacji cechami relewantnymi — na przykład cechy formalne i/lub treściowe dokumentu lub cechy obiektów nie będących dokumentami. Jest ono najczęściej sformułowane w języku naturalnym. Konieczność precyzyjnego określenia na jego podstawie tych cech informacji, które powinny być odwzorowane przy

wyborze strategii wyszukiwawczej i formułowaniu instrukcji wyszukiwawczej, powoduje, że formułowanie instrukcji wyszukiwawczej poprzedzone jest pośrednią transformacją zapytania w tzw. *instrukcję słowną*, sformułowaną w języku słów kluczowych. Instrukcja słowna jest następnie transformowana w instrukcję wyszukiwawczą, stosownie do przyjętej strategii. Wiąże się to z częstą wśród użytkowników systemów informacyjno-wyszukiwawczych nieumiejętnością artykułowania swoich potrzeb informacyjnych i w konsekwencji z nieprecyzyjnym formułowaniem zapytań informacyjnych, często różniących się od rzeczywistych potrzeb [Computer Readable Databases, 1985].

Teoretyczne podstawy określania strategii wyszukiwawczej i formułowania instrukcji wyszukiwawczej nie zostały jak dotąd wyczerpująco opracowane. W literaturze z zakresu JIW systemów dokumentacyjnych, problemy interpretacji i transformacji zapytań informacyjnych traktuje się bądź wąsko pragmatycznie, bądź ogólnikowo, ograniczając problem do wskazania etapów transformacji zapytania w instrukcję wyszukiwawczą [Lancaster F. W., 1968] [Ungurian O., 1976].

Próbie sprecyzowania teoretycznych podstaw transformacji zapytań informacyjnych kierowanych do SIW podjęły B. Wereszczyńska-Cisło i W. Ogórkiewicz, wykorzystując w tym celu pracę K. Ajdukiewicza *Zdania pytajne* oraz elementy teorii pytań W. Marciszewskiego [Ajdukiewicz K., 1986, 278-286], [Marciszewski W., 1977], [Wereszczyńska-Cisło B., Ogórkiewicz W., 1985, 41-60]. Teoria ta została opracowana z punktu widzenia metodologii formułowania pytań i rekonstrukcji przekształceń problemów badawczych, w ramach logicznej analizy tekstu naukowego. Jak zaznaczają autorki — cel, jakiemu ma służyć jest nieco inny niż wyszukiwanie informacji w systemie, rozumiane jako selekcjonowanie informacji już utrwalonej. Celem badania naukowego jest bowiem generowanie nowej informacji, co implikuje inne rodzaje transformacji. Jeśli jednak uwzględnić najnowsze systemy informacyjne, tzw. systemy inteligentne — na przykład symulujące w swoich efektach pracę eksperta, a więc generujące nową informację, to różnice te stają się niewyraźne.

W ujęciu Marciszewskiego możliwe odpowiedzi na dane pytanie tworzą pewien zbiór alternatywnych zdań oznajmujących. Elementami pytania są: założenie pytania (alternatywa wszystkich odpowiedzi) oraz niewiadoma, łącząca dane pytanie z możliwymi schematami odpowiedzi (*datum questionis*). Wyróżnia się dwa rodzaje pytań:

- *pytanie rozstrzygnięcia*, formułowane za pomocą partykuły *czy* i będące z założenia tautologiami, na przykład *czy zbiór dokumentów systemu zawiera dokumenty dotyczące X*,
- *pytanie uzupełnienia* (dopełnienia), formułowane za pomocą wyrażen innych niż partykuła *czy* (*gdzie, jak, dlaczego, który*), nie będące tautologiami, na przykład *które z dokumentów systemu dotyczą X*.

Pytanie rozstrzygnięcia wyraża kompletną niewiedzę pytającego, odpowiedź na nie może być binarna „tak — nie” już po próbie wyszukania jednego dokumentu.

Pytanie uzupełnienia zakłada pewną wiedzę użytkownika i udzielenie na nie odpowiedzi wymaga wyszukania w systemie wszystkich dokumentów relevantnych. Opis założeń pytań uzupełnienia opiera się w teorii Marciszewskiego na języku rachunku kwantyfikatorów. Założenie traktowane jest jako zdanie egzystencjalne, a więc uogólniona alternatywa, złożona z części twierdzącej (założenia pozytywnego) oraz przeczącej (założenia negatywnego). Konjunkcja części twierdzącej i przeczącej (założenie pełne) umożliwia ocenę zawartości informacyjnej pytania uzupełnienia dotyczącego określonego zbioru obiektów poprzez porównywanie liczby niezależnych logicznie członów alternatywy, będących założeniami pozytywnymi. Istnieje odwrotnie proporcjonalna zależność między liczbą członów alternatywy a informacją o odpowiedzi zawartą w danym pytaniu [Marciszewski W., 1977]. Założenie pytania powinno spełniać warunek prawdziwości i efektywności. Prawdziwość jest zapewniana przez fakt, iż założenie jest zdaniem egzystencjalnym, a więc najogólniejszym. Efektywność założenia wiąże się z koniecznością ograniczenia liczby członów alternatywy, które należy sprawdzić, aby ustalić elementy zbioru posiadające wskazaną w pytaniu cechę. Ogólność i efektywność założenia są więc warunkami przeciwstawnymi. Poszukiwanie optymalnego sposobu udzielania odpowiedzi na pytanie polega więc na zachowaniu proporcji między efektywnością (ograniczoną liczbą założeń) a ogólnością (prawdopodobieństwem niepominięcia właściwej odpowiedzi).

Rozwiązaniem problemu jest — w świetle dotychczasowych ustaleń — przyjęcie ogólnego założenia pytania przy równoczesnym dokładnym określeniu obiektów z tym założeniem związanych. Dokładność określenia zależy od wykorzystania wiedzy o obiekcie pytania przy formułowaniu tego pytania. Podstawą do orzekania o cechach obiektu wskazanego w pytaniu są jego wyrażenia składowe — nazwy i predykaty, z którymi wiąże się wiedza językowa (znajomość znaczenia i reguł języka, w którym formułowane jest pytanie) oraz często pewna niezbędna wiedza empiryczna (pozajęzykowa). Nazwy i predykaty pytania wyjściowego są punktem wyjścia do generowania pytań pochodnych, tworzących w sumie łańcuch zależności, który przyjmuje postać drzewa, np. pytanie *czym jest dany obiekt* generuje pytanie *z czego składa się obiekt; z jakich elementów się składa; jakie są relacje między tymi elementami*. Generowanie pytań pochodnych (pomocniczych) przyczynia się do zwiększenia ogólności założenia pytania, a więc do zwiększenia prawdopodobieństwa niepominięcia właściwej odpowiedzi. Oprócz tego stosuje się inny rodzaj modyfikacji członów założenia: poprzez eliminowanie wyrażen nieostrych, tj. powodujących powstanie zależnie od interpretacji alternatyw złożonych lub wyrażen wieloznacznych, co przyczynia się do spełnienia warunku efektywności (ograniczanie członów alternatywy). Zwiększeniu prawdopodobieństwa

niepominięcia właściwej odpowiedzi może służyć także operacja dołączania nowych członów. Oba rodzaje modyfikacji odbywają się na gruncie danej wiedzy językowej lub pozajęzykowej oraz dyrektyw zawartych w pytaniu. Ostateczne ustalenie zbioru założeń (członów alternatywy) prowadzi do określenia schematu dopuszczalnych odpowiedzi na pytanie.

Różnice między transformacjami pytań stanowiących rekonstrukcję problemów badawczych i transformacjami zapytań informacyjnych są — jak wspomniano — nieostre. Dotyczy to nie tylko celów transformacji. Zapytania informacyjne najczęściej nie są pytaniami w sensie gramatycznym, a raczej hasłowymi sformułowaniami problemów badawczych, nad którymi pracują użytkownicy, lub wyliczeniami nazw przedmiotów badań, cech tych przedmiotów, czynności lub procesów. Wymagają przekształcenia w pytanie uzupełnienia, a następnie niezbędnej modyfikacji (rozszerzenia, zawężenia) oraz przekładu na język informacyjno-wyszukiwawczy danego systemu i sformułowania instrukcji wyszukiwawczej. Operacje transformacji można sprowadzić do kilku charakterystycznych etapów:

1. Formułowanie pytania uzupełnienia według formuły (A): — *jaki podzbiór informacji (które dokumenty) należące do zbioru SIW dotyczą X.*

2. Generowanie pytania komplementarnego według formuły (B): — *które z opisów sytuacji rzeczywistości zawartych w przeszukiwanych podzbiorach SIW opisują X.*

3. Formułowanie charakterystyki słownej dla X w języku słów kluczowych — ustalenie listy wyrażen reprezentujących X .

4. Generowanie pytań pochodnych dla każdego wyrażenia (Y) z wyjściowej listy słów na podstawie wiedzy językowej lub pozajęzykowej, na przykład:

- a) czym jest Y w danym kontekście,
- b) do jakich zbiorów należy Y ,
- c) z jakich podzbiorów składa się Y ,
- d) z jakich indywiduów składa się Y ,
- e) z jakich obiektów składają się desygnaty Y ,
- f) częściami jakich obiektów są desygnaty Y ,
- g) jakie własności sytuacyjne posiadają desygnaty Y ,
- h) jakie inne desygnaty współwystępują z Y w danej sytuacji,
- i) jakie role pełnią desygnaty Y wobec innych desygnatów w danej sytuacji.

5. Budowa założenia pytania B na podstawie charakterystyki słownej pytań pomocniczych.

6. Modyfikacja założenia pytania B (członów alternatywy) przez dołączenie członów nowych lub eliminowanie istniejących w celu zwiększenia jego efektywności i prawdopodobieństwa.

7. Przekład słów kluczowych składających się na założenie pytania B na niezależne syntaktycznie wyrażenie IJW danego systemu — formułowanie charakterystyki leksykalnej X jako podstawy instrukcji wyszukiwawczej.

8. Rozbudowa charakterystyki leksykalnej przez wprowadzenie wyrażeń nowych, związanych z danym zbiorem leksyki JIW relacjami językowymi lub pozajęzykowymi, wyspecyfikowanymi w słowniku JIW lub na podstawie relacji współwystępowania w tekstach dokumentów.

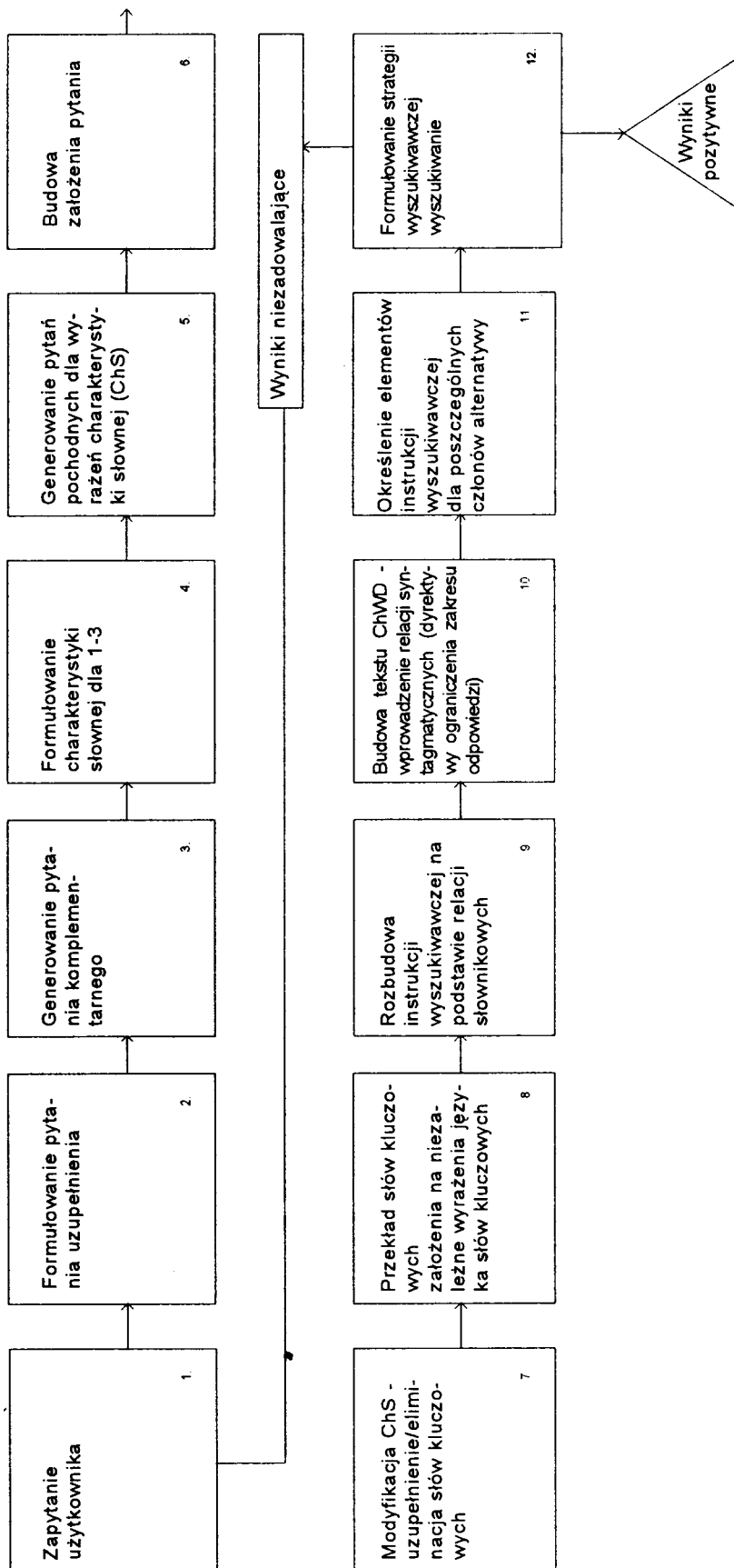
9. Wprowadzenie wykładników relacji syntagmatycznych między wyrażeniami JIW w celu ograniczenia zakresu treściowego i/lub formalnego odpowiedzi w postaci odpowiednich dyrektyw wyrażanych przez te relacje — formułowanie drugiej wersji charakterystyki wyszukiwawczej.

10. Precyzowanie instrukcji wyszukiwawczej zgodnie z zasadą, iż powinna ona zawierać wyrażenia JIW odpowiadające jednemu członowi alternatywy poprzez określenie pojedynczych elementów instrukcji lub ich zestawów reprezentujących poszczególne człony alternatywy dopuszczalnych odpowiedzi.

11. Sformułowanie na podstawie tekstów JIW, reprezentujących poszczególne człony alternatywy, strategii wyszukiwawczej, stanowiącej schemat wszystkich dopuszczalnych odpowiedzi na pytanie, a więc odwzorowującej wszystkie niezbędne w procesie wyszukiwania transformacje. Relacje między tekstami instrukcji odpowiadającymi członom alternatywy mogą być reprezentowane przez operatory logiczne (np. sumę, iloczyn, różnicę zbiorów).

W tradycyjnych SIW, w których użytkownik nie ma możliwości wglądu i ingerencji w przebieg procesu wyszukiwania informacji, precyzja transformacji prowadzących do sformułowania instrukcji wyszukiwawczych, składających się na przyjętą strategię, decyduje o efektywności wyszukiwania. W przypadku wyników niezadowolających możliwa jest modyfikacja instrukcji i strategii poprzez ograniczanie lub rozbudowę wyrażeń JIW, reprezentujących człony alternatywy dla zapytania informacyjnego, jest to jednak uzależnione od przyjętych w systemie reguł sprzężenia zwrotnego. Na przykład w systemach Selektywnej Dystrybucji Informacji, instrukcja wyszukiwawcza reprezentująca tzw. profil użytkownika, służąca do wyszukiwania w aktualizowanych okresowo zbiorach informacyjnych systemu, jest na ogół weryfikowana z częstością paromiesięczną. Każda modyfikacja instrukcji oznacza zatem konieczność rewizji wszystkich etapów transformacji zapytania.

Przy tak scharakteryzowanych warunkach dialogu użytkownika z systemem, instrumentem determinującym proces wyszukiwania jest język informacyjno-wyszukiwawczy. Scharakteryzowane wyżej operacje transformacji zapytania informacyjnego w instrukcję wyszukiwawczą, a w szczególności związane ze specyfikowaniem cech dokumentów lub obiektów pozadokumentacyjnych w języku słów kluczowych i w JIW świadczą o tym, iż o efektywności procesu wyszukiwania decyduje wiedza odwzorowana w języku, a także zbieżność z potrzebami użytkownika sposobu odwzorowania tej wiedzy w systemie. Operacje transformacji językowych, składające się na proces indeksowania dokumentów lub obiektów pozadokumentacyjnych, są w zasadzie te same, co dla instrukcji wyszukiwawczej. Punktem wyjścia jest pytanie: *czego dotyczy dokument X / czym jest obiekt Y*, stanowiące podstawę generowania



Schemat 10. Dialog użytkownika z systemem tradycyjnym

pytań pochodnych i sformułowania charakterystyki wyszukiwawczej jako zbioru dopuszczalnych alternatyw dla pytań użytkownika, spełniających warunek prawdopodobieństwa i efektywności w sensie wyżej scharakteryzowanym. Jest to problem bardzo trudny, jeśli uwzględnić sposób myślenia i postępowania użytkownika systemu i indeksatora opracowującego charakterystykę wyszukiwawczą i instrukcję wyszukiwawczą, przy czym najczęściej są to dwaj różni indeksatorzy. Użytkownik zakłada zwykle, że indeksator powinien posiadać tę samą co on wiedzę merytoryczną z danej dziedziny, więc formułuje pytanie hasłowo. Indeksator, dysponujący ograniczoną wiedzą specjalistyczną, jest ograniczony w swoim działaniu przez przyjęte w systemie reguły odwzorowania wiedzy.

Jako rozwiązanie problemu dostosowania struktury języka informacyjnego w systemie do potrzeb użytkowników przyjmowano uniwersalistyczne podejście przy ustalaniu modelu języka. Zgodnie z nim zakładano, że charakteryzowanie treści dokumentów poprzez odwzorowanie głównego ich przedmiotu stanowi prawdopodobną i efektywną alternatywę wszelkich zapytań informacyjnych. Podejście takie, ukształtowane przez tradycję systemów bibliotecznych oraz potrzeby informacyjne przede wszystkim nauk ścisłych, ma istotne wady z punktu widzenia zróżnicowania potrzeb informacyjnych w różnych dyscyplinach naukowych, w różnych etapach kariery naukowej użytkowników, w różnych etapach realizacji tematów badawczych, a także ze względu na niejednolity stopień przygotowania użytkowników do korzystania z różnych serwisów informacyjnych [Adam R., 1982] [Swift D. F., 1977, 90-94].

Uniwersalistyczne, przedmiotowe podejście w kształtowaniu modelu JIW umożliwia eksponowanie w charakterystyce wyszukiwawczej, a więc uznawanie za relewantną, jednej z trzech kategorii informacji wykorzystywanej przez pracowników nauki, a mianowicie informacji empirycznej. Tę kategorię informacji tworzą podstawowe dane do badań, najłatwiejsze do zidentyfikowania i zlokalizowania w tekstach dokumentów źródłowych, stanowiące przedmiot zainteresowania głównie *przedstawicieli nauk przyrodniczych*. Dwie pozostałe kategorie informacji — *informacja metodologiczna* (o procedurach i technikach badań) oraz *teoretyczno-koncepcyjna* (o nowych teoriach, hipotezach i ideach) interesują przedstawicieli poszczególnych dziedzin w stopniu zróżnicowanym. Informacja teoretyczno-koncepcyjna stanowi niezbędne uzupełnienie informacji empirycznej w naukach ścisłych, natomiast informacja metodologiczna często uznawana jest za najwartościowszą (relewantną) dla uprawiających nauki społeczne [Adam R., 1982], [Wojtasiewicz O. A., Sosińska B., 1982, 3-20].

Zróżnicowanie potrzeb informacyjnych użytkowników informacji z punktu widzenia specyfiki poszczególnych dziedzin wiedzy nie pokrywa się ze zróżnicowaniem wynikającym z układu chronologicznego kariery pracownika nauki czy też prac nad konkretnym tematem badawczym. Niezależnie od specyfiki poszczególnych dziedzin wiedzy wyróżnia się etapy pracy naukowej, w których dominuje wyszukiwanie informacji wszystkich trzech kategorii lub jej rozpowszechnia-

nie, traktowane naturalnie z dużym uproszczeniem jako procesy komplementarne. Wyszukiwanie informacji metodologicznej, teoretyczno-koncepcyjnej i empirycznej jest charakterystyczne dla początków kariery naukowej (od osiągnięcia stopnia doktora), a także dla wstępnych etapów prac nad nowym tematem badawczym. Rozpowszechnianie informacji w ramach kategorii specyficznych dla poszczególnych dziedzin cechuje etap samodzielnej pracy naukowej oraz końcowy etap badań w ramach konkretnego tematu. Pozyskiwanie i dystrybucja informacji za pośrednictwem systemów informacyjno-wyszukiwawczych jest jedną z form komunikacji w nauce, w najlepszym razie równorzędną obok kanałów nieformalnych [Adam R., 1982], [Adler H., 1981, 13-28].

Uniwersalistyczne podejście w doborze modeli języków informacyjno-wyszukiwawczych w systemach tradycyjnych porównuje się w literaturze poświęconej problemom badań informacyjnych użytkowników z tzw. arbitralnymi i zindywidualizowanymi modelami nauczania [Ford N., 1983, 209-222], [Sosińska B., 1985, 19-36]. Arbitralność w traktowaniu potrzeb informacyjnych polega na modelowaniu systemu i języka według jednolitych kryteriów oceny wartości informacji pod względem stopnia szczegółowości opisu obiektów systemu (dokumentów lub niedokumentów) i wyróżniania jednej kategorii informacji przedmiotowej (empirycznej) jako podstawy kluczy wyszukiwawczych. Na ogół nie uwzględnia się tak istotnych z punktu widzenia różnych stopni przygotowania merytorycznego użytkowników cech informacji, jak wieloaspektowość opisu (relewanca interpretacyjna) oraz różne poziomy uogólnienia (abstrakcji). Problem ten jest szczególnie dotkliwie odczuwany wśród użytkowników systemów informacji nauk społecznych, w których te same reprezentacje zjawisk uzyskują, poprzez różne konwencje terminologiczne różne interpretacje na gruncie różnych szkół i koncepcji opisu rzeczywistości, czego rezultatem jest zapotrzebowanie na informację metodologiczną i teoretyczno-koncepcyjną. W systemach tradycyjnych próby znalezienia rozwiązań w zakresie indywidualizacji prezentacji wiedzy systemu i wyrażania potrzeb użytkowników przybierają różne formy, co znajduje odzwierciedlenie w poszczególnych rodzajach języków.

Arbitralne kształtowanie obrazu rzeczywistości (dokumentacyjnej lub pozadokumentacyjnej) w systemach tradycyjnych zasadnie krytykowano z punktu widzenia potrzeb użytkowników informacji. Podejście to ma jednak specyficzne „usprawiedliwienie”, jeśli zważyć sytuację komunikacyjną, w jakiej znajduje się system i użytkownik porozumiewający się za pomocą języka informacyjno-wyszukiwawczego.

Język ten pełni dwie funkcje podstawowe: metainformacyjną i wyszukiwawczą (impresywną, heurystyczną) oraz dwie funkcje wtórne: organizacyjną i określania zakresu tematycznego zbiorów informacji, pochodne funkcji metainformacyjnej.

Funkcja metainformacyjna polega na ograniczaniu funkcji semantycznej języka naturalnego do opisu informacji — np. prezentacji cech treściowych

i/lub formalnych dokumentów czy też prezentacji dystynktywnych cech obiektów pozadokumentacyjnych.

Funkcja wyszukiwawcza polega na sterowaniu procesem wyszukiwania informacji poprzez wyrażenia charakteryzujące warunki, jakie powinien spełniać wyszukany podzbiór informacji. O możliwościach realizacji funkcji wyszukiwawczej decyduje w znacznej mierze funkcja organizacyjna, ponieważ określa ona miejsce wyrażen syntaktycznie prostych języka w zbiorze informacji systemu poprzez formę tych wyrażen (plan wyrażenia).

Istnieje odwrotnie proporcjonalna zależność między miejscem wyrażenia danego JIW w organizacji zbioru leksyki języka i organizacji zbioru informacyjnego (w tzw. porządku pionowym) a jego miejscem w organizacji syntagmatycznej (w tzw. porządku poziomym) [Sosińska B., 1987].

Przy realizacji funkcji wyszukiwawczej obiektem sterowania jest zarówno odbiorca komunikatów (tekstów) języka informacyjno-wyszukiwawczego: system informacyjno-wyszukiwawczy, odbierający komunikaty użytkownika (instrukcje wyszukiwawcze) jak i użytkownik odbierający odpowiedzi systemu o własnościach formalnych lub treściowych dokumentów. Język informacyjny umożliwia zatem formułowanie tylko dwu typów wypowiedzi. Należy jednak zauważyć, że użytkownik systemu tradycyjnego po uzyskaniu odpowiedzi niezadowolającej po przeprowadzonym wyszukiwaniu może sformułować nową instrukcję i powtórzyć wyszukiwanie na podstawie modyfikacji swojej informacji introspektywnej, spowodowanej wprowadzeniem wiedzy przekazanej przez system. Wiedza systemu i jego możliwości komunikowania się pozostają niezmienione. System stanowi więc stronę pasywną w dialogu z użytkownikiem, a przy ograniczonej interakcji użytkownika z systemem efektywność wyszukiwania informacji zapewnia arbitralność w sposobie prezentowania informacji poprzez język.

Nowe podejście do formułowania strategii wyszukiwawczej, wraz z określaniem warunków relewancji informacji i jej oceną, pojawia się w momencie powstania systemów udostępniających informację w trybie dialogowym, zapoczątkowując równocześnie dyskusje nad indywidualizacją metod reprezentacji wiedzy w systemie i metod komunikowania się z użytkownikiem.

4.2. WYSZUKIWANIE INFORMACJI BEZ TRANSFORMACJI SEMANTYCZNEJ

Udostępnianie wiedzy systemu bez jej transformacji semantyczno-prawdziwościowej ilustrujemy przykładem systemów dialogowych. Ich funkcjonowanie charakteryzujemy przede wszystkim z punktu widzenia wpływu sposobu realizacji funkcji wyszukiwawczej (zmienianej przez użytkownika strategii wyszukiwania) na prezentację wykładników cech relewantnych. Należy przy tym podkreślić, że powstanie systemów dialogowych możliwe było dzięki rozwojowi technicznych instrumentów informacyjnych:

- powstaniu komputerów z pamięciami dyskowymi, umożliwiającymi szybkie wyszukanie informacji, działających w trybie wielodostępnym,

- rozwojowi sieci telekomunikacyjnych, umożliwiających szybki przekaz informacji między komputerami przy wykorzystaniu kanałów łączności satelitarnej, linii telefonicznych itp.,
- opracowaniu oprogramowania (*software*) dostosowanego do organizacji informacji na nowych nośnikach (pamięciach dyskowych) oraz do przetwarzania informacji w trybie interakcyjnym na podstawie procedur interrogacyjnych.

Czynnikiem pośrednio wpływającym na powstanie systemów dialogowych był liczebny wzrost zbiorów informacji na nośnikach maszynowych (baz danych). Pierwsze wydanie katalogu opracowywanego od lat siedemdziesiątych na Uniwersytecie Illinois (USA), uznawanego za jedną z najbardziej reprezentatywnych „baz danych o bazach danych” zawierało charakterystykę około trzystu skomputeryzowanych zbiorów informacji (dane opublikowane w 1976 r.). Wydanie czwarte z 1985 r. podaje informacje o około trzech tysiącach baz (2805) [*Computer Readable Databases*, 1985].

4.2.1. Strategia wyszukiwawcza w systemie dialogowym

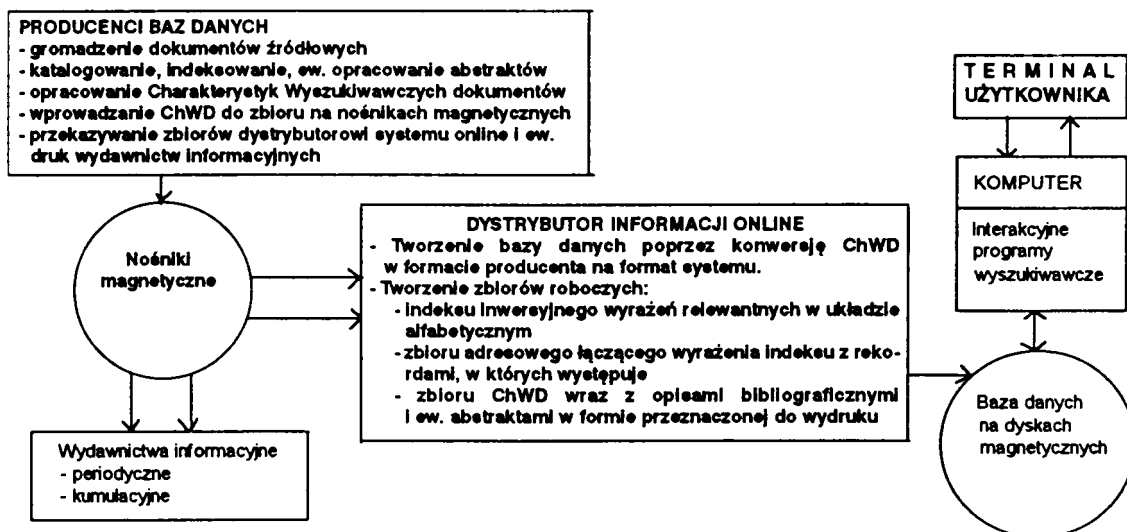
Przyjęcie zasady indywidualnego, szybkiego komunikowania się użytkownika z systemem, dysponującym zbiorami informacji na dyskach zawierających od kilkudziesięciu do kilkuset tysięcy rekordów z dowolnej dziedziny wiedzy, wpłynęło na zmianę jego struktury funkcjonalnej. Zmiana ta polegała na rozróżnieniu ról producenta informacji (*data base producer, supplier*) i dystrybutora informacji, nazywanego także operatorem systemu (odp. *vendor, operator*). Zmiany te zapoczątkowały w latach sześćdziesiątych w USA firmy Lockheed oraz System Development Corporation (SDC) bez ingerowania w zasady opracowania zbiorów informacyjnych systemów udostępnianych tradycyjnie przez producentów baz danych w postaci wydawnictw abstraktowych, katalogów bibliotecznych, indeksów przedmiotowych i innych, a więc przez biblioteki naukowe, służby abstraktowo-bibliograficzne, instytucje naukowe¹.

¹ Prace nad budową dwu najstarszych i największych pod względem zbiorów informacji oraz zasięgu działania systemów online Lockheed-DIALOG oraz SDC-ORBIT rozpoczęto w latach sześćdziesiątych. W roku 1963 firma Lockheed utworzyła Pracownię Informacji Naukowej (Information Sciences Laboratory). W roku 1966 wygrała konkurs na budowę systemu informacji bibliograficznej online dla National Aeronautics and Space Administration (NASA), podejmując opracowanie oprogramowania programów wyszukiwawczych nazwanego RECON (REmote CONsole), za pomocą którego pierwotnie wyszukiwano informację w zbiorach *International Aerospace Abstracts, Scientific and Technical Aerospace Reports*, tworzonych w NASA. Programy RECON zaczęto następnie wykorzystywać w US Atomic Energy Commission, the US Office of Education oraz National Technical Information Service (NTIS) ESA. Od 1972 r. firma Lockheed stosuje zmodyfikowaną wersję RECON we własnym systemie on-line, udostępnianym użytkownikom na zasadach komercyjnych pod nazwą DIALOG. System zaczął funkcjonować z trzema bazami danych, których liczba powiększa się systematycznie. W 1975 r. system posiadał 24 bazy, w 1978 r. — ponad 70, w 1980 r. — ponad 100, a w połowie lat osiemdziesiątych — około 500.

Budowę systemu (programów wyszukiwawczych) SDC ORBIT podjęto w roku 1963 na zlecenie US Department of Defence Advanced Research Projects Agency (ARPA), dysponującego wówczas zbiorem informacji zawierającym ponad 200 tys. rekordów bibliograficznych (*ORBIT-Online Ret-*

Do zadań dystrybutorów baz danych w systemie online (instytucji i organizacji komercyjnych, dużych bibliotek naukowych o zasięgu międzynarodowym typu National Library of Medicine, National Agricultural Library, Library of Congress, Ohio College Library Center, British Library, międzynarodowych organizacji naukowych typu European Space Agency) należy:

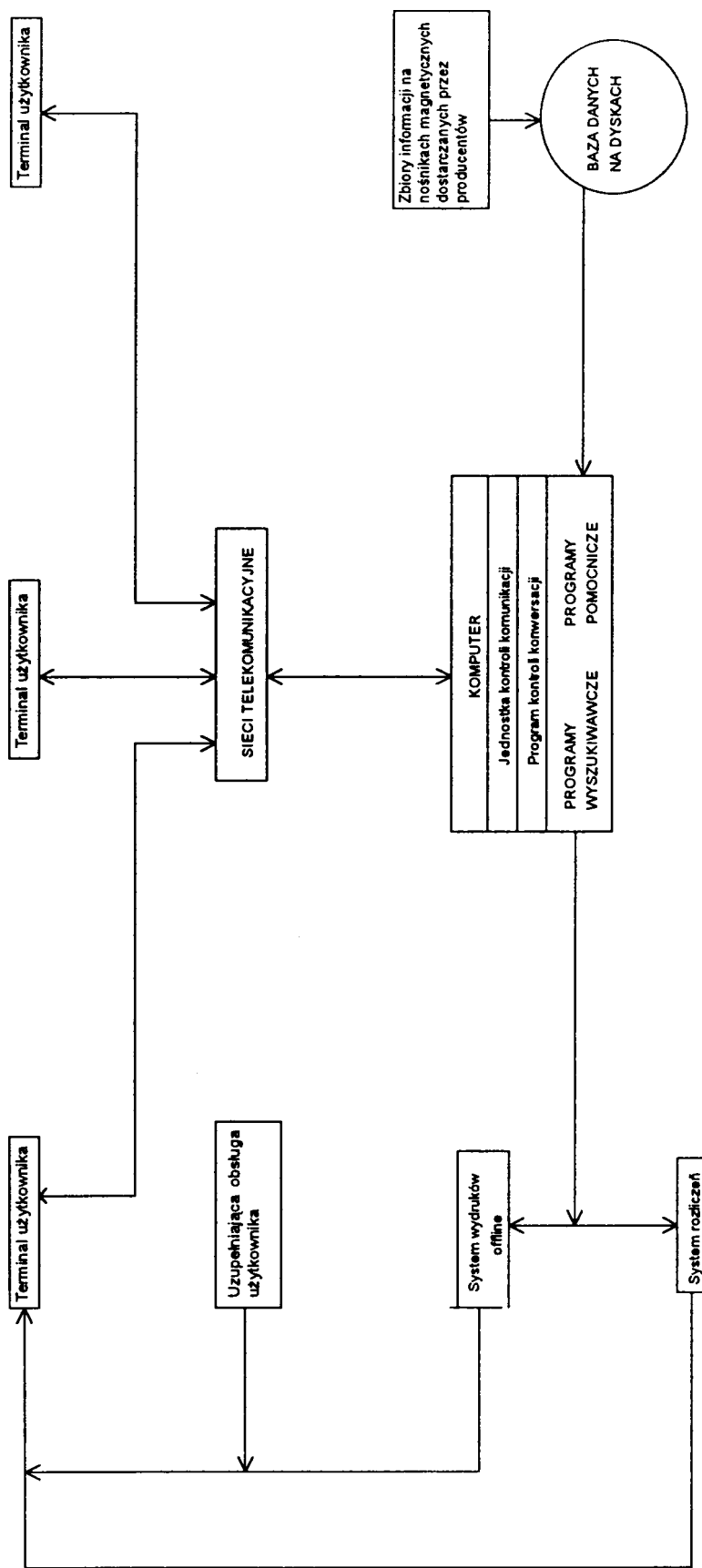
- dokonywanie konwersji indywidualnych formatów danych, przyjętych w zbiorach informacji poszczególnych producentów, na format systemu, a więc transformacja struktur tekstów ChWD w planie wyrażania w strukturę obowiązującą w systemie;
- stworzenie i aktualizacja bazy danych systemu o organizacji informacji ułatwiającej selekcję informacji według różnych kluczy wyszukiwawczych, a więc wyrażen pełniących funkcję metatekstową, wskazujących wyrażenia relewantne w procesie wyszukiwania;



Schemat 11. Funkcje uczestników systemu informacyjno-wyszukiwawczego online

rievial of Bibliographic Information and Text). W roku 1968 firma SDC rozpoczęła prace na zlecenie amerykańskiej National Library of Medicine nad oprogramowaniem do interakcyjnego wyszukiwania informacji w zbiorach wydawnictwa abstraktowego *Index Medicus*. Programy te wraz z językiem wyszukiwawczym ELHILL (Lister Hill National Centre for Biomedical Communication) zaczęto stosować od 1970 r. w systemie MEDLINE. Od 1973 r. SDC uruchomiła własny system SDC ORBIT, funkcjonujący na zasadach komercyjnych, posiadający w momencie uruchomienia trzy bazy danych, w 1980 r. — około 60. Kolejne systemy online o zasięgu międzynarodowym powstawały w drugiej połowie lat siedemdziesiątych, wśród nich m.in.:

- *Bibliographic Retrieval Service (BRS)* w State University of New York w 1976 r., funkcjonujące ze zmodyfikowaną wersją oprogramowania STAIRS, — US New York Times Information Bank,
- *Infoline* w zakresie nauk biomedycznych (Wielka Brytania) — od 1979 r.;
- *BLAISE* (British Library Automated Information Service — od 1977 r.;
- *AUSINET* w Australii,
- *EURONET-DIANE* (Europejska Wspólnota Gospodarcza — od 1980 r.



Schemat 12. Instrumenty informacyjne w systemie online

- opracowanie języka wyspecjalizowanego w funkcji wyszukiwawczej, umożliwiającego użytkownikowi prowadzenie dialogu z systemem kurtuazyjnie reagującym na zapytania;
- dostarczanie użytkownikowi informacji ze zbiorów systemu, przygotowanej w trybie wsadowym (offline) i ewentualnie kopii dokumentów źródłowych. Zasady podziału funkcji w systemie on-line oraz układ ich instrumentów informacyjnych ilustrują następujące schematy (za [Hoover R. E., 1980]):

4.2.2. Prezentacja kryteriów relewancji w systemie dialogowym

Podstawą selekcjonowania informacji w systemie jest — jak zaznaczono w schemacie 11 — indeks wyrażen relewantnych (indeks podstawowy — *basic index*), nazywanych potocznie terminami wyszukiwawczymi (*search terms*), uporządkowanymi alfabetycznie, w stosunku do którego zbiór adresowy oraz zbiór charakterystyk wyszukiwawczych dokumentów w systemie dokumentacyjnym pełnią funkcję pomocniczą, związaną z technicznymi (informatycznymi) aspektami przetwarzania informacji. Świadczy to o przyjęciu innych niż w systemach tradycyjnych założeń określających minimalny warunek relewancji systemowej — każde wyrażenie umieszczone w indeksie inwersyjnym, niezależnie od jego miejsca w strukturze paradygmatycznej JIW, może być relewantne. Ewentualną rozbieżność między systemową oceną relewancji i oceną użytkownika (pertynencji) rekompensuje możliwość zmiany strategii wyszukiwawczej w trakcie sesji wyszukiwawczej dzięki szybkiej reakcji systemu na przekazywane przez użytkownika komendy sformułowane w języku wyszukiwawczym systemu — języku kwerend (*query language*).

Każdy z wymienionych przykładowo systemów posiada własny schemat dialogu z użytkownikiem, określony w tzw. protokołach konwersacji, niezależnych od języków informacyjnych stosowanych w funkcji metainformacyjnej przez producentów baz danych. Protokoły te i związane z nimi języki kwerend, oparte najczęściej na alternatywnej notacji paranaturalnej i sztucznej, wskazują rodzaje operacji wykonywanych przez komputer w celu wyszukania informacji. Stanowią zasadniczą część dokumentacji instruktażowej każdego systemu online. Różnią się w szczególności pod względem repertuaru operacji i ich nazw, co stwarza kłopoty użytkownikom korzystającym równolegle z kilku systemów. Odpowiadają podstawowym etapom realizacji procesu wyszukiwania (por. tab. 3).

W dokumentacji instruktażowej poszczególnych systemów można wyróżnić następujące grupy operacji niezbędnych do przeprowadzenia sesji wyszukiwawczej:

1. Operacje telekomunikacyjne (podejmowania i kończenia dialogu z systemem — *Logon and Logoff*). do których należy:

Słowniki języków wyszukiwawczych niektórych systemów online
(nazwy niektórych operacji selekcjonowania operacji)

Tabela 3

Rodzaj operacji	DIALOG Nazwa komendy - Skróć	SDC-ORBIT	BRS	DIANE	ESA-IRS
Uzyskiwanie (meta-) metainformacji o zbiorach systemu	?FILESn ?FILESUM	FILES?		INFO	?FILES
Uzyskiwanie metainformacji o JW	?EXPLAIN	EXPLAINitem (EXP.?)		HELP	?QUEST ?SESAME
Uzyskiwanie metainformacji o strukturze ChWD	?FIELDn				?FILDs
Uzyskiwanie informacji o bieżących zmianach w systemie	?NEWS ?WHATsNEW	NEWS	NEWS	NEWS	?NEWS
Uzyskiwanie informacji o aktualizacji zbiorów w bieżącym tygodniu i w poprzednim	?UPDATE				?UPDATE
Wybór bazy danych do wyszukiwania informacji	?BEGIN	FILE	CHANGE	BASE	BEGINn
Przegląd indeksu wyrazów relewantnych	EXPAND	NEIGHBOR NBR	ROOT*	DISPLAY	EXPAND
Polecenie wyszukania ChWD/ ChWO według wskazanej instrukcji wyszukiwawczej	SELECT SUPERSELECT	FIND FD'	SEARCH	FIND	SELECT
Modyfikacja (ograniczenie) zakresu przeszukiwanego zbioru	LIMIT	FIND FD'	LIMIT	L	LIMIT
Przegląd wyników stosowanej strategii wyszukiwawczej	DISPLAY SETS	HISTORY HIS	DISPLAY ALL	BACK	DISPLAYSETS DS
Wydruk wyszukiwanej informacji - online -offline	TYPE PRINT	PRINT PRINT OFFLINE PRT OFF	PRINT PRINTOFF PO	SHOW PRINT	TYPE PRINT

- a) uruchomienie sprzętu: terminala, modemu, drukarki,
- b) uzyskanie połączenia z siecią telekomunikacyjną², na przykład poprzez linię telefoniczną, po przekazaniu identyfikatora terminala, następnie hasła (*password*), otwierającego dostęp do sieci oraz kolejnego hasła (*password*), otwierającego dostęp do systemu online,
- c) reagowanie na informacje przekazywane z sieci lub systemu w czasie zakłóceń w łączności lub przy niesprawności sprzętu dystrybutora,
- d) ortograficzna korekta tekstów pisanych na ekranie przez użytkownika,
- e) uzyskanie informacji o kosztach wyszukiwania.

2. Operacje selekcjonowania informacji za pomocą języka kwerend, wśród których można wyróżnić następujące grupy procedur interrogacyjnych, umożliwiających:

- a) wybór baz danych (zbioru informacyjnego), w których ma być wyszukiwana informacja,
- b) przegląd indeksów baz danych, zawierających wyrażenia uznane za wykładniki relewancji w systemie,
- c) formułowanie strategii wyszukiwawczej — wprowadzenie wyrażen składających się na instrukcję wyszukiwawczą (*search statements*),
- d) modyfikację instrukcji wyszukiwawczej, np. przez ograniczenie liczby dokumentów spełniających warunek relewancji technicznej,
- e) przegląd instrukcji wyszukiwawczych stosowanych w trakcie sesji wraz z wynikami wyszukiwania uzyskanymi w poszczególnych etapach,
- f) wydruk w trybie (*online* lub *offline*) wyselekcjonowanej informacji,
- g) uzyskanie informacji pomocniczej w procedurach interrogacyjnych dopuszczalnych w ramach stosowanego języka wyszukiwawczego (języka kwerend),
- h) uzyskanie informacji o bieżących zmianach w zakresie organizacji zbiorów w systemie oraz ich zasobach.

Spośród wymienionych dwu grup operacji, pozostających w dyspozycji użytkownika prowadzącego dialog z systemem, przedmiotem naszego zainteresowania są operacje grupy drugiej, służące do selekcjonowania informacji. Procedury telekomunikacyjne i inne związane są z technicznymi i czysto pragmatycznymi aspektami funkcjonowania systemu, stanowią w naszych rozważaniach zagadnienie marginalne.

Realizacja operacji selekcjonowania informacji możliwa jest — jak wspomnieliśmy — dzięki językowi kwerend, którego wyspecjalizowaną funkcją jest funkcja impresywna sterowania operacjami komputera. Język ten jest funkcjonalnie komplementarny wobec języków informacyjnych, stosowanych w poszczególnych bazach danych w funkcji metainformacyjnej. Wyszukanie informacji w systemie online nie jest możliwe wyłącznie za pomocą języka służącego do

² Systemy amerykańskie udostępniane są m.in. za pośrednictwem dwu sieci TELENET i TYMNET. Istnieją ponadto sieci regionalne, np. SCANNET, EURONET AUSINET, międzybiblioteczne — OCLC [Sosińska B., 1983, 99-114].

opracowania charakterystyk wyszukiwawczych dokumentów lub charakterystyk obiektów pozajęzykowych w systemach faktograficznych. Funkcji tej nie można także zrealizować wyłącznie przy użyciu samego języka kwerend. Tak więc, nowy rodzaj instrumentu informacyjnego w systemie (interakcyjne programy wyszukiwawcze) powoduje wyodrębnienie w JIW podzbiorów jego wyrażeń wyspecjalizowanych w funkcjach: metainformacyjnej, wyszukiwawczej i impresywnej. Z realizacją funkcji wyszukiwawczej i impresywnej wiąże się stosowanie operatorów logicznych oraz operatorów tekstowych, które omawiamy w dalszej części pracy. Interesującą cechą niektórych systemów online jest możliwość zmiany języka kwerend przy wyszukiwaniu informacji w tym samym zbiorze, udostępnianym za pośrednictwem różnych sieci teleinformatycznych. Na przykład w systemie ESA-IRS (European Space Agency — Information Retrieval System) możliwa jest zmiana języka QUEST podobnego do Lockheed-DIALOG na język SESANE (EUROLANGUAGE), jeden z języków sieci EURONET przez podanie odpowiedniej komendy *SET SESANE ON* lub *SET EUROLANGUAGE ON*.

Użytkownik systemu online, dysponujący często dostępem równocześnie do kilku systemów teleinformatycznych (o zasięgu światowym lub regionalnym) staje nie tylko przed problemem metod selekcji informacji w zbiorach liczących miliony rekordów, lecz także przed problemem wyboru bazy danych, a więc jednego z podzbiorów systemu. Pojawia się pytanie *gdzie wyszukiwać informację*, poprzedzające problem *jak wyszukiwać*. W ten sposób w systemie online zostaje wskazane nowe kryterium relewancji, nie nazywane *explicite* w systemach tradycyjnych, mianowicie informacji o zasobach informacyjnych systemu, czyli o bazach danych bibliograficznych i faktograficznych³. Należy zauważyć, że ten istotny dla użytkownika rodzaj relewancji, oznacza wyróżnienie kolejnego poziomu metainformacji, gdyż jest to:

- informacja o metainformacyjnych bazach danych dokumentacyjnych (bibliograficznych), a więc *meta-metainformacja*, a także
- informacja o informacyjnych bazach danych faktograficznych, a więc *metainformacja*.

Prezentacji kryteriów relewancji odwzorowujących rodzaj informacji w podzbiorach systemu służą komendy przeznaczone do:

- wykazania zasobów informacyjnych systemu,
- wyboru jednej lub więcej baz danych, w których ma być wyszukiwana informacja (por. tabela 3) zawierająca zestawienie komend przeznaczonych do realizacji poszczególnych rodzajów operacji w niektórych systemach online).

Ten rodzaj relewancji odwzorowany jest także w indeksie przedmiotowym wspomnianego katalogu baz danych [*Computer Readable Databases*, 1985].

³ W systemach amerykańskich operuje się terminami *baza danych bibliograficznych* (*bibliographic data base*) i *baza danych nie-bibliograficznych* (*non-bibliographic data base*). Ten dytomiczny podział zbiorów informacyjnych systemu pozwala (pragmatycznie) rozwiązać definicyjne dylematy związane z terminem *fakt* (faktograficzny).

Scharakteryzowanie zakresu tematycznego baz danych jest istotne dla użytkownika ze względów merytorycznych i ekonomicznych, ponieważ do zjawisk powszechnych należy częściowe pokrywanie się zasobów informacyjnych poszczególnych baz tworzonych przy wykorzystaniu różnych JIW. Oprogramowanie większości systemów online umożliwia obecnie przeprowadzenie wyszukiwania równoległe w kilku bazach danych na podstawie tej samej instrukcji wyszukiwawczej, co z jednej strony decyduje o obniżeniu kosztów wyszukiwania, z drugiej zaś pozwala porównać rezultaty wyszukiwania w bazach o podobnym zakresie tematycznym, np. *COMPENDEX (Computerized Engineering Index)*, *ISMEC (Information Services for Mechanical Engineering)* i *INSPEC* zawierają wspólne pole tematyczne „fizyka”. Analogicznie w systemach udostępniających informację katalogową, np. w systemie *SOCRATES*, Stanford University USA, możliwe jest wyszukiwanie informacji równocześnie w katalogach wydawnictw zwartych, ciągłych, filmów, map, archiwów, raportów z badań lub w jednym z nich, wskazanym przez użytkownika [*Reference Guide to Socrates*, 1985].

4.2.3. Wiedza o zbiorach informacyjnych systemu

Meta-metainformacyjny opis zasobów informacyjnych systemu dostarczany użytkownikowi w bardzo obszernej, zalgorytmizowanej dokumentacji instruktażowej⁴, w układzie podporządkowanym operacjom realizowanym za pomocą instrukcji języka wyszukiwawczego, stanowi jeden ze sposobów „humanizacji” procesu komunikowania się użytkownika z systemem. Służy temu wymieniona w tabeli komenda *NEWS*, występująca w repertuarze instrukcji, jak też instrukcje umożliwiające uzyskiwanie informacji o zasobach informacyjnych systemu, w postaci *HELP*, *FILES*, *INFO*, *INSTRUCT* oraz *EXPLAIN* związana z prezentacją wiedzy o JIW.

Wykorzystanie tych komend może dotyczyć różnych poziomów organizacji informacji w zbiorach. Na przykład w systemie *DIALOG* — Instrukcja:

— *:?EXPLAIN* umożliwia przekazanie przez system informacji o zasobie słownika języka kwerend;

— *:?FILES* powoduje wydruk listy wszystkich baz danych dostępnych w momencie wyszukiwania dla użytkownika legitymującego się określonym hasłem (Zbiory poufne „confidential files” nie są udostępniane wszystkim użytkownikom).

Możliwe jest ograniczenie wykazu baz danych przez dołączenie numeru bazy, na przykład:

— *FILES 90* spowoduje wydruk listy baz, których numer jest wyższy niż 90.

⁴ Dokumentacja instruktażowa systemu *Lockheed-DIALOG* wydana w 1980 r. składała się z czterech kilkusetstronicowych tomów, z których pierwszy zawierał opis procedur telekomunikacyjnych oraz języka wyszukiwawczego, natomiast trzy pozostałe — poświęcone były charakterystyce poszczególnych baz danych.

— ?FILE *n* — powoduje przekazanie syntetycznego opisu bazy danych, typu:

?FILE 47: MAGAZINE INDEX PROVIDES COVERAGE OF OVER 350 POPULAR AMERICAN MAGAZINES TO SUPPLY INFORMATION ON A BROAD RANGE OF TOPICS INCLUDING CURRENT AFFAIRS, CONSUMER PRODUCT EVALUATION, SPORTS AND RECREATION, BUSSINESS, ENVIRONMENT, SCIENCE AND TECHNOLOGY.

THE FILE CONTAINS RECORDS FROM 1977 TO THE PRESENT. APPROXIMATELY 5000.

NEW RECORDS ARE ADDED MONTHLY. RETROSPECTIVE COVERAGE FOR 1976 IS CURRENTLY IN PROCESS. AS OF SEPTEMBER 1978 THE FILE CONTAINED OVER 100.000 RECORDS.

COST OF SEARCHING IS \$ 45/CONNECT HOUR AND \$10/FULL RECORD PRINTED OFF-LINE.SEE ALSO ?FIELD 47⁵ [Guide to Dialog Searching, 1979].

?FIELD — służy do prezentacji tych wszystkich cech dokumentu lub obiektu pozajęzykowego, których cechy są odwzorowywane w charakterystyce wyszukiwawczej, powodując wydruk wszystkich pól wyszukiwawczych rekordu danej bazy. W wykazie pól wyróżnia się pola odpowiadające cechom formalnym dokumentu (tzw. *prefix-coded fields*) oraz pola tekstowe (*full-text fields*), np.:

?FIELD 75 (baza dokumentacyjna)

FIELD 75: MANAGEMENT CONTENTS

PREFIX-CODED FIELDS	(EXPAND OR SELECT)
FIELD NAME	EXAMPLE
AUTHOR	AU=ZUCKER,S.
DESCRIPTOR CODE	DC=0933
PUBLICATION DATE	DT=76/08
JOURNAL CODE	JC=NAB
JOURNAL NAME	JN=NATIONS BUSSINESS
PUBLICATION YEAR	PY=1976
UPDATE	UD=9999
FULL-TEXT FIELDS	(SELECT ONLY)
ABSTRACT	ADVERSING (F) MODEL/AB — gdzie (F) operator tekstowy
DESCRIPTOR	MANAGEMENT STYLE

⁵ Zbiór 47 *Magazine Index* pokrywa ponad 350 popularnych czasopism amerykańskich w celu dostarczenia informacji o szerokim zakresie tematycznym obejmującym bieżące wydarzenia, konsumencką ocenę wyrobów, sport i wypoczynek, biznes, ochronę środowiska naukę i technikę.

Zbiór zawiera rekordy poczynając od 1977 do chwili obecnej. Około 5000. Nowe rekordy są wprowadzane co miesiąc. W trakcie opracowania są informacje retrospektywne za 1976 rok. We wrześniu 1978 r. zbiór zawierał ponad 100 000 rekordów.

Koszt wyszukiwania wynosił \$45 za godzinę połączenia i \$10 za wydruk kompletnych rekordów offline. Zob. także ?Field 47.

TITLE

MEDIA (F) RESPONSE/TI

FIELD 106: TRADE OPPORTUNITIES WEEKLY (baza faktograficzna)

PREFIX-CODED FIELDS

EXAMPLE

BUSSINESS CODE

BC=A

BUSSINESS NAME

BN=AGENT

COUNTRY CODE

CC=241

COUNTRY NAME

CN=BRAZIL

DATE OF INFORMATION DI

=790618

PRODUCT AMOUNT

PA=0050

PRODUCT CODE

PC=991

TYPE OF OPPORTUNITY

CODE

TC=204

TYPES OF OPPORTUNITY TO

=DIRECT SALE

UPDATE

UD=999

FULL-TEXT FIELDS

ABSTRACT

DESIGN PLANNING/AB

COMPANY

PORT(W)SAID(F)HOUSING/CO

W przytoczonych przykładach pól wyszukiwawczych baz danych dokumentacyjnej i faktograficznej równocześnie załączono uwagę o możliwości zastosowania wyrażen języka kwerend związanych z selekcjonowaniem informacji. Każdy z systemów online posiada wykaz zbiorczy symboli pól wyszukiwawczych, załączany w dokumentacji instruktażowej. Jego zawartość uzależniona jest od liczby i rodzajów baz danych (por. Zał. 1 i 2 „Pola wyszukiwawcze w systemach DIALOG i SDC-ORBIT). Znacznie bogatszy jest repertuar pól wyszukiwawczych systemu DIALOG, w którym udostępnia się kilkadziesiąt baz danych faktograficznych, głównie zawierających informację handlową i statystyczną, np. bazy zaliczane do grup BUSINESS COMPANIES, PREDICAST — MARKET RESEARCH, BUSINESS STATISTICS, funkcjonujące komplementarnie z bazami danych bibliograficznych, BUSINESS BIBLIOGRAPHIC, LAW GOVERNMENT, PUBLIC AFFAIRS i inne. Symbole stosowane w funkcji etykiet pól wyszukiwawczych, pełniące rolę wykładników kryteriów relewancji dla obiektów danej bazy mają często przyporządkowanych kilka znaczeń. Wieloznaczność ta jest automatycznie eliminowana przez wybór zbioru informacji do wyszukiwania. Podany przykładowo sposób prezentacji wiedzy o zasobach informacyjnych całego systemu oraz poszczególnych baz danych wraz ze strukturą rekordów bazy, a więc strukturą tekstów spójnych języka informacyjnego, może stanowić podstawę jednej z dwu metod wyszukiwania, tzw. metody jadłospisowej (menu driven), zalecanej dla użytkowników początkujących w korzystaniu z systemów online. Posługując się podanym przez system wykazem pól wyszukiwawczych, nazywanych opcjami wyszukiwawczymi, użytkownik dokonuje wyboru pól odpowiadających odpowiednim cechom formalnym lub treściowym obiektu i wypełnia je wyrażeniami, które powinny zawierać charakterystyki wyszukiwawcze. Formułowanie instrukcji wyszukiwawczej odbywa się wówczas metodą indek-

sowania ankietowego, szczególnie często wykorzystywanego w bibliotecznych katalogach online, obsługujących wielu użytkowników, na przykład studentów bibliotek akademickich we wspomnianym systemie *SOCRATES* oraz w National Library of Medicine (USA). Ten sposób prezentacji wiedzy o cechach relevantnych zbioru informacyjnego decyduje o tym, iż system jest przyjazny dla użytkownika (user cordial, user friendly), a ponadto uwalnia użytkownika od obowiązku znajomości często uciążliwej, bo zróżnicowanej w poszczególnych zbiorach, interpunkcji. Przykład wyszukiwania w katalogu autorskim w National Library of Medicine:

Nazwisko	Imię	Imię
FREUD	SIGMUND	

Informacja systemu:

„...Proszę nacisnąć klawisz *ENTER* po wypełnieniu poszczególnych podpól... ”
 „...Proszę nacisnąć klawisz *TAB* w celu przeniesienia się do kolejnego podpola... ”
 „...Pytanie zostało przyjęte. Proszę czekać ... ”

W systemie *SOCRATES* menu katalogów sieci bibliotecznej zawiera wykaz następujących pól wyszukiwawczych:

<i>A</i>	<i>AUTHOR, personal name</i>	np.	<i>J. D. Salinger</i>
<i>O</i>	<i>ORGANIZATION, as author</i>		<i>Smithsonian</i>
<i>TP</i>	<i>TITLE, Initial phrase</i>		<i>Portrait of the Artist as ...</i>
<i>T</i>	<i>TITLE, any words</i>		<i>Portrait Artist</i>
<i>S</i>	<i>Subject, topical words</i>		<i>Genetic Engineering</i>
<i>SN</i>	<i>SUBJECT, personal name</i>		<i>Martin Luther King</i>

Po wyborze opcji system prosi użytkownika o podanie wyrażen relevantnych, tzw. wartości wyszukiwawczych (search values), na podstawie których selekcjonuje informację po wprowadzeniu komendy *FIND* [*Reference Guide to SOCRATES Search*, 1985].

Zaletą wykorzystania cech strukturalnych zbioru informacji w wyszukiwaniu jądłospisowym jest łatwość posługiwania się systemem. Oznacza to równocześnie znaczne ograniczenie zasady indywidualizacji komunikowania się z systemem oraz schematyzację w formułowaniu strategii wyszukiwawczej, której sprecyzowanie i umiejętność modyfikacji wymaga pogłębionej wiedzy o języku systemu.

4.2.4. Prezentacja wiedzy o języku

Selekcjonowanie informacji według sformułowanej przez użytkownika strategii wyszukiwawczej odbywa się po wskazaniu, za pomocą komendy oznaczającej rozpoczęcie wyszukiwania (np. *BEGIN*, *CHANGE*, *BASE*):

- jednej bazy danych: np. *?BEGIN* lub *B3* (nr bazy *INSPEC* w systemie *DIALOG*)
- kilku baz danych:

a) przez wskazanie ich numerów:

?SELECT FILE lub SF 2, 3, 4, 8, 32, 34, 94, gdzie:

FILE2: Chemical Abstract Search 67-71

FILE3: Chemical Abstract Search 72-76

FILE4: Chemical Abstract Search 77-80

FILE6 COMPENDEX 70-80

FILE32 METADEX 66-80

FILE34 SCISEARCH 78-80

FILE94 SCISEARCH 74-77

b) przez polecenie wyboru baz według kryteriów przedmiotowych:

?SF ENERGY, ENVIRONMENT

Odpowiedź systemu:

FILE6 NTIS 64-80

FILE8 COMPENDEX 70-80/OCT

FILE12 INSPEC 69-77/OCT

FILE13 INSPEC 78-80/O

FILE28 OCEANIC ABS 64-80/OCT

FILE29 MET/GEOASTRO ABS 70-79/DEC

FILE40 ENVIROLINE 71-80/AUG

FILE41 POLLUTION ABS 70-80/SEP

FILE44 AQUATIC SCI ABS 78-80/JUL

FILE45 APTIC 66-78/OCT

FILE68 ENVIRONMENTAL BIBLIOGRAPHY 74-80/OCT

FILE69 ENERGYLINE 71-80/AUG

Podstawą realizacji procesu wyszukiwania jest — jak wspominaliśmy — tzw. indeks podstawowy (basic index) każdej z baz danych, który stanowi słownik wyrażeń relewantnych w systemie, odwzorowujących cechy obiektów według omówionych wyżej pól (opcji) wyszukiwawczych. Do cech tych należą:

- cechy formalne dokumentów,
- cechy treściowe dokumentów,
- cechy obiektów pozajęzykowych.

Repertuar cech formalnych dokumentów jest bardzo obszerny i wykorzystywany w różnym zakresie w poszczególnych bazach (por. Zał. 1 i 2 *Wykaz pól wyszukiwawczych*). Uogólniając można stwierdzić, że jako opcja wyszukiwawcza nie występuje w nim jedynie cecha paginacji.

Cechy treściowe dokumentów w bibliograficznych bazach danych najczęściej odwzorowywane są w słowniku bazy za pomocą wyrażeń należących do języka naturalnego oraz języka informacyjno-wyszukiwawczego. Na podstawie przytoczonych wykazów pól wyszukiwawczych systemów *DIALOG* i *SDC-ORBIT* możemy wymienić następujące JIW, najczęściej stosowane w funkcji meta-informacyjnej:

- język słów kluczowych,
- język haseł przedmiotowych,

- język deskryptorowy,
- specjalistyczne kategoryzacje,
- klasyfikacje patentowe (narodowe i międzynarodowa).

Należy przy tym zastrzec, że na skutek zasady nieingerowania przez dysytributora informacji w metody charakterystyki treściowej dokumentów, obserwuje się dużą dowolność i nieprecyzyjność w terminologii stosowanej w charakterystykach struktury rekordów w systemie i zawartości ich pól wyszukiwawczych. Częstym zjawiskiem jest używanie terminów deskryptor, termin, hasło, fraza, pojęcie (*concept*). Termin identyfikator definiowany w tradycji europejskich służb informacyjnych jako deskryptor równokształtny z nazwą własną, w systemach amerykańskich oznacza słowo kluczowe ustalone w ramach swobodnego indeksowania dokumentów. Dokładna charakterystyka JIW w bazach danych online wymagałaby dostępu do szczegółowej dokumentacji instruktażowej każdej z nich i zapoznania się z ich kontrolowanymi słownikami, co nie jest możliwe. Słowniki nazywane tezaurusami są niekiedy kontrolowanymi wykazami słów kluczowych z zaznaczoną relacją odpowiedniości wyszukiwawczej. Dla użytkownika nie posiadającego doświadczenia w korzystaniu z systemów online myląca może okazać się informacja o stosowaniu kodów w charakterystykach wyszukiwawczych. W językach o notacji paranaturalnej, także deskryptorowych, stosuje się często równoległą notację sztuczną, np. cyfrową DC=0508 (baza danych MANAGEMENT CONTENT), która pozwala oszczędzać czas i koszty wyszukiwania.

Kolejną interesującą cechą w systemach online związaną z zakresem tradycyjnych JIW jest minimalny udział klasyfikacji w charakteryzowaniu treści dokumentów. W niektórych amerykańskich bazach danych stosowana jest *Klasyfikacja Dziesiąta Dewey'a* (np. Government Printing Office — Monthly Catalogue). Ponadto w bazach BIOSIS z zakresu nauk przyrodniczych BIOSIS PREVIEWS równoległe z językiem deskryptorowym wykorzystuje się język klasyfikacyjny — tzw. biosystematic code. *Międzynarodowa Klasyfikacja Patentowa* oraz klasyfikacje narodowe, pełniące — jak już wskazywaliśmy — funkcję kodów semantycznych, znajdują zastosowanie w tych bazach danych, które zawierają charakterystyki opisów patentowych (INPADOC — o zasięgu międzynarodowym, amerykańskich — CLAIMS/CHEM 50-70, CLAIMS/US PAT 71-77, CLAIMS/US PAT 78-, CLAIMS/WEEKLY) [Guide to DIALOG Searching, 1979].

4.2.4.1. Prezentacja leksyki JIW

Prezentacja cech obiektów pozadokumentacyjnych w faktograficznych bazach danych oparta jest na kontrolowanych słownikach nazw obiektów usystematyzowanych w ramach pragmatycznie interpretowanych kategorii, nazywanych niekiedy zdarzeniami (*events*). Cechy obiektów specyfikowane są bądź ze względu na ich lokalizację, bądź parametry ilościowe, zwłaszcza w dominujących w zbiorach DIALOGU-u bazach informacji rynkowej, przemysłowej, czy też w zakresie finansowania i sprawozdawczości z prac naukowo-badawczych

i rozwojowych. Słowniki te, z równoległą notacją paranaturalną i sztuczną (product, name, event name, phase name — product code, event code, phase code), udostępniane są użytkownikom uprawnionym do korzystania z baz. Na przykład obiekty bazy *EIS INDUSTRIAL PLANTS* (Economic Information Systems — US industrial activity), charakteryzowane są za pomocą następujących cech [*Guide to DIALOG Searching*, 1979]:

BC = Branch city
BN = Branch name
BS = Branch state
BZ = Branch zip code
CC = Company code
CN = Company number
DE = Descriptor
EX = Employee size
HC = Headquarters city
HN = Headquarters name
HS = Headquarters state
HX = Headquarters code
HZ = Headquarters zip code
PC = Product code
PN = Product name
SD = Sales in milion \$
SF = Special feature
SM = Share of market

Tak zwane nie-bibliograficzne bazy danych łączą faktycznie informacje o dokumentach z informacją o innych obiektach, co odwzorowywane jest w zestawie kategorii, w ramach których umieszczone są nazwy (i kody) obiektów i ich cech. Na przykład w bazie *PTS FEDERAL INDEX* (Predicasts FI) przyjęto następującą kategoryzację obiektów:

00 Annotation and Status Codes (Federal Index)
10 Organization and institutions
20 Management procedures
30 Products and processes
40 Resources and resource use
50 People
60 Market information
70 Unit costs and prices
80 Financial data
90 Government and Society.

Drugim źródłem wyrażeń indeksu bazy danych, oprócz słowników JIW, są pola charakterystyki wyszukiwanego dokumentu lub obiektu pozadokumentacyjnego, które określane są jako pola pełnotekstowe (*full-text fields*), zawierające informacje w języku naturalnym, a więc *TYTUŁ, ABSTRACT, ADNOTACJA,*

a także w niektórych bazach faktograficznych, nazwy własne instytucji lub przedsiębiorstwa. Selekcja wyrażeń relewantnych z tych pól dokonywana jest automatycznie przez zastosowanie tzw. *stop lists*, a więc drogą eliminacji tych wyrażeń, które nie mogą pełnić funkcji wyszukiwawczej, tradycyjnie uznawanych za „asemantyczne” — spójników, przedimków, poimków, przyimków. Wyrażenia relewantne są umieszczane w indeksie bazy w postaci unitermów. Odróżnienie wyrażeń JIW od unitermów (słów kluczowych) selekcjonowanych automatycznie umożliwia sposób ich prezentacji w indeksie, w którym zaznacza się niepodzielność deskryptorów równokształtnych z wielowyrzowymi wyrażeniami języka naturalnego, np. [*Guide to DIALOG Searching*, 1979]:

<i>E1 NUCLEAR</i>	17564
<i>E2 NUCLEAR/DE</i>	5333
<i>E3 NUCLEAR/ID</i>	7288
<i>E4 NUCLEAR FISSION/DE</i>	4792
<i>E5 NUCLEAR POWER/DE</i>	4298
<i>E6 NUCLEAR REACTOR</i>	6736

Zapoznanie się z indeksem bazy jako słownikiem wyrażeń relewantnych, charakteryzujących przedmioty treści dokumentów lub obiekty pozadokumentacyjne umożliwia komenda: *EXPAND* lub *E* (DIALOG), *NEIGHBOR* lub *NBR* (SDC), *ROOT* (BRS) lub *DISPLAY* (DIANE). System automatycznie przypisuje pozycjom ukazywanego na ekranie indeksu numery, które można wykorzystać w instrukcji wyszukiwawczej zamiast pełnych wyrażeń, umieszczając ponadto po ostatniej pozycji pytanie o kontynuację, np. w systemie *DIALOG-2* — *more*. Potwierdzenie przez naciśnięcie klawisza enter powoduje dalsze wyświetlenie indeksu bez potrzeby powtarzania komendy.

W niektórych bazach danych zasób leksykalny JIW udostępniany jest w postaci odrębnego podzbioru — słownika (dialist) lub tezaury online. Informację o możliwości korzystania ze słownika JIW online podaje się w charakterystyce bazy danych, np. w *ERIC*, *PSYCHOLOGICAL ABSTRACTS*, *GEOARCHIVE*, *US POLITICAL SCIENCE ABS*, *BIOSIS*, *MESH* (Medical Subject Headings) i innych.

Szczególnie rozbudowaną informacją słownikową odznaczają się bazy informacji chemicznej (*CA* — *Chemical Abstracts*), liczące kilka milionów rekordów. Dla baz tych (*CA*, *CA Condensates*, *CASIA*, *CA PATents*) zbudowano słownik konkordancyjny, funkcjonujący jako niezależna baza danych *CHEMNAME*.

Nazwy substancji chemicznych są reprezentowane za pomocą wyrażeń języka naturalnego oraz symboli chemicznych, dzięki czemu możliwe jest wyszukiwanie informacji według cech strukturalnych związków oraz według kryteriów przedmiotowych (subject search). Często aktualizacja zbiorów informacji, rozbudowa terminologii chemicznej oraz złożoność zależności między substancjami (strukturalna i aplikacyjna) wpłynęły na szczegółowość specyfikacji relacji w słowniku. Odwzorowano w nim następujące cechy wyrażeń i ich desygnatów:

— nazwa w języku deskryptorowym, obejmującym nazwy substancji zawarte w *CA Index Names* (IDE), np. *ACID, AZULENE, BENZ, CARBOXYLIC, DERIV, DI, GIBBAN*;

— wykaz pierwiastków składowych (*element count*) oparty na formule molekularnej, umożliwiający dostęp do informacji o wszystkich atomach formuły (EC);

— słowa kluczowe (*ID — identifiers*), wyselekcjonowane z pola synonimy (*SY*);

— nazwa nadrzędna (*Heading Parent*) (*HP=*) w *CA Index Names*, np. *HP=AZULEN*;

— zastępniki (*Substituent*) (*SB=*), np.; *SB=METHYLEN*;

— modyfikacja nazwy (*Name Modification*) (*NM=*) w *CA Index Names*;

— nazwa podstawowa (*name match*) (*CN=*), umożliwiająca wyszukanie dokumentów stanowiących pierwotne publikacje o danej substancji z pominięciem późniejszych pochodnych;

— synonimy (*SA=*) nazw umieszczonych w *CA Index Names*, np. *SY=GIBBERELLIN*;

— formuła molekularna (*Molecular Formula*) (*MF=*), np. *MF=C19H244-04*, według systemu Hilla;

— wzór strukturalny (*Stereochemical Descriptor*) (*ST=*), odwzorowujący układ przestrzenny substancji;

— numer grupy (*Group Number*) w schemacie klasyfikacyjnym CA, np. *GN=A1*;

— termin indeksu periodycznego (*Periodical Index Term*) (*PI=*).

Prezentacja leksyki języków w postaci indeksów poszczególnych baz danych, ewentualnie słowników (tezaursów, słowników jhp) online, przy częstym pokrywaniu się zakresów tematycznych zbiorów stwarza użytkownikom systemów trudności spowodowane faktycznym ograniczeniem opisu jednostek leksykalnych do planu wyrażania. Stanowi to także problem dla służb informacyjnych, opracowujących zbiory informacji, lecz równocześnie współdziałających w zakresie wymiany informacji w ramach sieci biblioteczno-informacyjnych. Specjaliści zajmujący się globalnymi problemami rozwoju systemów online coraz częściej piszą o swoistej „dżungli językowej” w bazach danych, z których korzystanie staje się coraz trudniejsze i coraz bardziej kosztowne [Lewis D. A., 1980, 59-74]. Jako propozycję rozwiązania problemów semantycznej porównywalności i przekładalności języków poszczególnych baz danych przedstawia się koncepcje lingwistycznych (terminologicznych) banków danych wbudowanych w architekturę systemu jako samodzielny moduł (podsystem). Zakłada się, że moduł taki spełniałby pewne funkcje:

a) wewnątrzsystemowe, w tym:

— kontrola i korekta zapytań niepoprawnych leksykalnie lub ortograficznie,

— integracja językowa zbiorów systemu,

- pomoc terminologiczna w formułowaniu strategii wyszukiwawczej;
- b) zewnętrzne (międzysystemowe), w tym:

- ułatwianie dostępu do baz danych w ramach lokalnych sieci teleinformatycznych,

- podstawa terminologiczna w budowie JIW nowych baz danych i aktualizacji istniejących.

Koncepcję budowy zintegrowanego leksykonu baz danych (authority file — kartoteki wzorcowej), skupiającego słownictwo języków deskryptorowych, języków słów kluczowych, języków haseł przedmiotowych, zaczęła realizować w latach siedemdziesiątych Biblioteka Kongresu USA, współpracująca z Government Printing Office USA, bibliotekami uniwersyteckimi (Northwestern, Texas, Wisconsin), National Library of Medicine, dysponująca indeksem haseł przedmiotowych *Medical Subject Headings (MESH)*, oraz National Agricultural Library [Colombo D. S., Niehoff R. T., 1977].

Jako inny przykład prac nad scalaniem zasobów leksykalnych słowników JIW baz danych udostępnianych w systemach online może posłużyć tzw. *Vocabulary Switching System (VSS)*, tworzony w Battelle Columbus Laboratories, USA. Początkowo zasoby jego obejmowały kilka słowników JIW, m.in. tezaurus *TEST*, tezaurus *NASA*, tezaurus *INSPEC*, indeksy przedmiotowe; na początku lat osiemdziesiątych ich liczba sięgała dwudziestu o łącznych zasobach około 500 tys. jednostek leksykalnych, tworzących terminologiczny bank danych. Zbiory tego leksykonu zgrupowano w czterech modułach tematyczno-dziedzinowych [Hillman D. J., 1977]:

- 1) nauk technicznych i ścisłych, zawierającym wyrażenia indeksów przedmiotowych *Chemical Abstracts*, deskryptory *Tezaurusu Departamentu Energetyki USA*, *Tezaurusu TEST. INSPEC, NASA* oraz trójjęzycznego *Tezaurusu Żelaza i Stali*;

- 2) nauk biomedycznych, zawierającym wyrażenia indeksu przedmiotowego bazy *Biosis Previews, MESH* oraz część indeksu *CA*;

- 3) nauk społecznych, zawierającym m.in. deskryptory tezaurusów *ERIC* i *Psychological Abstracts*,

- 4) businessu, administracji i zarządzania, zawierającym leksykę tezaurusów stosowanych w bazach *ABSTRACTS BUSINESS inc. INFORMATION, MANAGEMENT CONTENT* oraz *ECONOMIC ABSTRACTS INTERNATIONAL*.

Zasoby leksykalne integrowanych słowników prezentowane są w układzie alfabetycznym (*form file*), z uwzględnieniem wykazu relacji paradygmatycznych każdego wyrażenia, a więc wykładników relacji generycznych, skojarzeniowych, odsyłaczy związanych z międzyjęzykową relacją ekwiwalencji wyszukiwawczej, typu: *używaj — używaj zamiast, zob. — zob. też* oraz precyzujących znaczenie *scope notes*.

Funkcje indeksów pomocniczych, ułatwiających wyszukanie wyrażeń (terminów) w zbiorach leksykonu, pełnią słowniki inwersyjne opracowane przy wykorzystaniu operacji permutacji i maskowania (obcinania):

- indeks unitermów (*word file*) grupujący wyrazy mogące pełnić samodzielnie funkcję wyszukiwawczą;
- indeks rdzeni wyrazowych (*stem file*), tworzony za pomocą algorytmów obcinania końcówek fleksyjnych wyrazów i części morfemów słowotwórczych w elementach składowych terminów;
- indeks rdzeni terminów (*stem phrase file*), powstały analogicznie jak indeks wyrazów (unitermów).

Wykazaniu podobieństwa znaczeniowego terminów należących do leksyki różnych JIW, a tym samym wskazaniu ich synonimii wyszukiwawczej, służy tzw. zbiór pojęć (*concept file*), w którym odwzorowane są także relacje paradygmatyczne, organizujące pola semantyczne poszczególnych języków.

Podobne znaczenie mają zintegrowane leksykony, pełniące funkcję makrotezaurusów dla systemów specjalistycznych, np. system *SCORPIO* (Subject Oriented Retriever for Processing Information Online), stosowany przy wyszukiwaniu informacji do obsługi informacyjnej członków Kongresu USA [76]. Ich powstanie ilustruje interesującą tendencję rozbudowy w systemach online wiedzy o strukturze paradygmatycznej JIW oraz o relacjach międzyjęzykowej ekwiwalencji wyszukiwawczej, wobec ograniczenia zakresu wiedzy o rzeczywistości charakterystycznej dla języków systemów tradycyjnych.

4.2.4.2. Prezentacja reguł gramatyki

Charakterystyczną cechą języka wyszukiwawczego w systemach online jest — jak wspomnieliśmy — wzbogacanie jego funkcji związanych z instrumentami przetwarzania informacji o funkcję impresywną. Realizacja procesu selekcionowania informacji w zbiorach systemu po:

- uzyskaniu meta-metainformacji o ich zasobach,
- wyborze bazy/baz danych,
- zapoznaniu się z indeksem wyrażenń relewantnych, reprezentujących cechy formalne i/lub treściowe dokumentów bądź obiektów pozadokumentacyjnych, odbywa się za pomocą komendy o znaczeniu wyselekcjonuj informację o obiektach według wskazanych cech. Komenda ta może mieć postać np.: *SELECT (DIALOG, ESA-IRS)*, *FIND (SDC ORBIT, DIANE)*, *SEARCH (BRS)*. Z jej użyciem wiąże się specyfikacja reguł pozycyjnej gramatyki języka.

Wyrażenie wskazane przez użytkownika jako wykładnik relewancji musi zawsze wystąpić po pełnej lub skrótowej nazwie operacji.

Najprostsza formuła zdaniowa komendy *SELECT* (w systemie *DIALOG*) miała postać:

?SELECT lub *?S WYRAŻENIE* jednowyrazowe — *?S CANCER*

Odpowiedź: numer wyszukanego podzbioru, zawierającego wskazane wyrażenie, liczba zawierających je ChWD, wyrażenie:

1 3655 *CANCER*.

Ten sposób formułowania instrukcji interpretowany jest przez system jako polecenie wyszukania poprzez indeks bazy charakterystyk wyszukiwawczych, które zawierają w dowolnym polu tekstowym wskazane wyrażenie. Do specyfikacji cech obiektów systemu odwzorowanych w przewidzianym dla danej bazy zestawie pól wyszukiwawczych służą wymienione wyżej etykiety (por. załączniki 1, 2), posiadające stałe miejsce w formule zdaniowej języka kwerend.

W systemie *DIALOG* etykieta (symbol) pola, które powinno zawierać wskazane wyrażenie, umieszczana jest prepozycyjnie lub postpozycyjnie w stosunku do tego wyrażenia, zależnie od przyjętego ogólnego kryterium selekcji (formalnych cech dokumentu, cech obiektu pozadokumentacyjnego lub treściowych cech dokumentu).

Komenda sformułowana w celu wyselekcjonowania informacji o dokumentach według cech formalnych powinna odpowiadać wzorcowi:

?S symbol pola tzw. prefix ze znakiem równości, wykładnik relewancji

?SAU = JONES C⁶

2 422 AU = JONES C

lub (przykład opcji wyszukiwawczych dla pól oznaczanych w zdaniu prepozycyjnie z bazy *LANGUAGE AND LANGUAGE BEHAVIOR ABSTRACTS*):

?SAU=FAISS, KLAUS

?SCA=CHOMSKY (CA= Cited author)

?SCO=LECI-4 (CO= Coden)

?SDT=REVIEW (DT= Document type)

?SJO=LINGUISTICS (JO= Journal)

?SLA=GER (LA= Language)

?SPD=1975 (PD= Publication Date)

?SSC=110 (SC= Section code)

?SUD=76 (UP= Update)

Gdy liczba pozycji związanych z niedokładnie określoną cechą jest duża, system przekazuje informację, np.:

?SAU=SAWYER

3 > 1 000 TERMS; RESPECIFY

W układzie prepozycyjnym wskazywane są również cechy obiektów pozadokumentacyjnych, np. w bazie *FOUNDATION GRANTS INDEX (USA)*:

?SAM=00020 (AM= Grant amount, \$ thousands)

?SAG=00020 (AG= Greater than or equal)

?SAL=00030 (AL= Less than)

?SCC=3.3 (CC= Category code; tu: History]

?SDF=F (DF= Domestic or foreign recipient; tu: Francja)

?SFS=NC (FS= Foundation state)

?SLO=ENGLAND (LO= Sites of activity)

⁶ Ogromnym praktycznym problemem, np. przy wyszukiwaniu ChWD według nazwisk autorów dokumentów, jest interpunkcja, np. kropka lub przecinek umieszczane po inicjałach imienia/imion. Obserwuje się tendencję do eliminacji interpunkcji i zastępowanie jej spacjami.

?S RC=RALEIGH (RC= Recipient city)
 ?S RS=NC (RS= Recipient state)
 ?S YA=73 (YA= Year authorized)

Cecha obiektu może być również wskazana przy wykorzystaniu numerów pozycji wyrażenia w indeksie bazy, np.: ?S E9, ?# E8

Selekcjonowanie informacji według cech treściowych tekstów wymaga w systemie *DIALOG* postpozycyjnego umieszczenia etykiety określającej pola wyszukiwawcze, poprzedzanej kreską ukośną /, np.: /TI, /AB, /DE i in.

Przykłady selekcjonowania informacji w cytowanych bazach *LANGUAGE AND LANGUAGE BEHAVIOR ABS* oraz *FOUNDATION GRANTS INDEX*:

FOUNDATION GRANTS INDEX

?S ACQUISITION/AB — /AB abstrakt

?S NORTH CAROLINA HISTORY/DE — /DE deskryptor

?S NEW(W)YORK (F)RESIDENTS/LM

— /LM ograniczenie geograficzne subwencji, tu: do mieszkańców Nowego Jorku

— (W), (F) — operatory tekstu, omawiane dalej

?S Z(F)REYNOLDS/FN — nazwa fundacji

?S ART(W)CULTURE(W)CAROLINA/RN — /RN — recipient name

?S GOVERNMENTAL/RT — /RT Recipient type, według przyjętej kategoryzacji, np. *SCHOOLS, HOSPITALS*

LANGUAGE AND LANGUAGE BEHAVIOR ABS

?S PHONOLOGY /AB

?S BOCHUM /CS — /CS Corporate Source, źródło korporatywne

?S GENERATIVE PHONOLOGY/DE

?S BOOK(W)REVIEW/ID — /ID — słowa kluczowe, amer. „identifier”

?S SYLLABE(W)STRUCTURE/TI — /TI — tytuł

W wielu systemach pozycja etykiety pola wyszukiwawczego zależy od rodzaju cech relewantnych, określanych w zdaniu instrukcji wyszukiwawczej, nie jest rozróżniana. Dotyczy to np. systemów *SDC ORBIT, BLAISE, BRS*, w których symbol pola wyszukiwawczego jest zawsze umieszczany postpozycyjnie, przy zastosowaniu różnych znaków delimitacji.

W bazach *BRS* przy wyszukiwaniu pełnotekstowym nie jest wymagana kwalifikacja pól wyszukiwawczych za pomocą etykiet dołączanych do wyrażen jako sufiksy. W systemie *DIALOG* możliwość taką zapewnia wersja komendy *SELECT — SUPERSELECT (SS)*.

<i>SDC ORBIT</i> — baza <i>INSPEC</i>	<i>BRS</i> — baza <i>INSPEC</i>
<i>FIND SEISMOTECTONIC/TI</i>	<i>SEARCH SEISMOTECTONIC.TI</i>
<i>FIND NUCLEAR POWER STATIONS/IT</i>	
<i>FIND WEBER,C/AU SEARCH</i>	(<i>WEBER ADJ C</i>). <i>AU</i> <i>ADJ</i> — operator tekstu
<i>FIND GERMAN/LA</i>	<i>SEARCH GE.LG</i>

Stosowanie podstawowej komendy, umożliwiającej selekcjonowanie informacji ze wskazaniem wyrażenia określającego kryterium relewancji dla użytkownika, związane jest z istotnym ograniczeniem w specyfikacji cech obiektów bazy danych. Ograniczenie to polega na wyszukiwaniu informacji według jednej cechy relewantnej, na przykład według cech treści dokumentów lub ich własności formalnych, bez specyfikacji relacji zachodzących między tymi cechami w tekstach ChW.

We wszystkich systemach online podstawowym środkiem językowym, służącym budowie złożonych zdań instrukcji wyszukiwawczej i zawężania zakresu podzbiorów wyselekcjonowanych w wyniku działania komendy *SELECT*, są operatory algebry Boole'a, wyrażające działania na zbiorach (iloczyn, suma różnica i ew. uzupełnienie zbioru). Wykładnikiem operacji — *AND*, *OR*, *NOT* na zbiorach jest (najczęściej) komenda *COMBINE* (w skrócie *C*), która może być stosowana łącznie z numerami wyszukanych podzbiorów lub z wyrażeniami JIW. Łącznie liczba zbiorów poddawanych operacjom logicznym nie może przekroczyć 29. Na przykład:

SCIENCE CITATION INDEX — DIALOG:

1) ?S CANCER

1 3655 CANCER

?S CARCINOMA

2 2597 CARCINOMA

?S LUNG

3 2190

?C (1 OR 2) AND (3 OR 4)

5 422(1 OR 2) AND (3 OR 4)

2) ?S PAPER(W)MILL OR PAPERMILL

1 177 PAPER(W)MILL OR PAPERMILL

?S AU= SAWYER S

2 498 SAWYER S

?CIAND2

?S AU= JONES C

4 422 AU=JONES C

?CIAND4

5 0 IAND4

?SS AU=SAWYER S AND (PERMILL/TI OR WATER/TI)

6 498 AU=SAWYER S

7 2755 PAPERMILL/TI

8 21505 WATER/TI

9 7 6AND (7 OR 8)

Po uzyskaniu odpowiedzi systemu (podzbioru 9) użytkownik zwykle zapoznaje się z charakterystykami wyszukiwawczymi, przekazując komendę ich wydruku na ekranie (całego zbioru lub wybranych pozycji) w wybranym formacie:

?T9/2/1-3, gdzie:

T — wydruk na ekranie (wydruk na papierze wymaga komendy *P print*, sformułowanej według tych samych zasad),

/2/ — numer formatu⁷

1-3 — numery pozycji ChWD w wyszukanym podzbiorze.

Do modyfikacji zakresu wyszukanego podzbioru, przez ograniczenie liczby charakterystyk wyszukiwawczych dokumentów wskazanych przez system, służy także komenda *LIMIT (L)*; najczęściej przydatna do ograniczenia wyszukiwania informacji o dokumentach według ich cech formalnych (chronologicznych, językowych, lokalizacyjnych itp.), np.

?L 5/00550133-0669809 — wskazanie podzbioru według numerów systemowych ChWD

?L 2/ART — ograniczenie wyszukiwania tylko do ChWD artykułów w Science Citation Index

?L 21/NAR — ograniczenie wyszukiwania tylko do ChWD nie będących artykułami (SCI)

?L 9/ENG — ograniczenie wyszukiwania tylko do publikacji angielskojęzycznych

?L 10/FOR — ograniczenie wyszukiwania tylko do publikacji obcojęzycznych

?L %/AVAIL — ograniczenie wyszukiwania do dokumentów,

?L 6/UNAVAIL — ogólnie dostępnych (lub niedostępnych).

Możliwości zawężania zakresu podzbiorów są zróżnicowane w poszczególnych bazach. Najbardziej uniwersalnym kryterium jest numer akcesyjny charakterystyki wyszukiwawczej dokumentu w systemie.

Funkcję ograniczania zakresu przeszukiwanego zbioru według cech chronologicznych dokumentu może spełniać także komenda *SELECT* wraz ze znakiem „:” (dwukropek), np.: *?S PY=1976:PY=1978*

Symbol ten jest szczególnie przydatny w wyszukiwaniu w zbiorach, w których JIW są klasyfikacje, w celu wskazania jako wykładników relewancji symboli należących do jednego łańcucha klasyfikacyjnego, np.: *?S CC=64072:CC 078*.

Sesja wyszukiwawcza w systemie przeznaczona na wyselekcjonowanie informacji na jedno zapytanie użytkownika polega zazwyczaj na realizacji strategii złożonej z instrukcji wyszukiwawczych o zmiennych wykładnikach relewancji zależnie od wyników etapowych wyszukiwania.

⁷ Formaty prezentacji informacji wyszukanej w systemie są zróżnicowane w poszczególnych bazach. W systemie DIALOG użytkownik dysponuje ośmioma opcjami: *Format 1* — DIALOG — numer akcesyjny, *Format 2* — pełny rekord (ChWD) z wyłączeniem abstraktu, *Format 3* — Cytacja bibliograficzna, *Format 4* — Tytuł dokumentu i abstrakt, *Format 5* — Pełny tekst ChWD, *Format 6* — Numer akcesyjny DIALOG oraz tytuł dokumentu, *Format 7* — Cytacja bibliograficzna i Abstrakt, *Format 8* — Tytuł i ChWD w języku informacyjnym. Do sortowania informacji według cech innych niż przewidziane służy komenda *SORT* i instrukcja zbudowana według formuły *?SORT* numer podzbioru, numer ChWD, kryterium selekcji, np. *SORT8/1-48/AU TI*; *SORT10/1-54/SD, DT*, a następnie *P/11/5/AM, D*.

? DS
 SET ITEMS DESCRIPTION
 1 31784 AT
 2 13733 S
 3 1413 ARTIFICIAL(W)INTELLIGENCE
 4 1 INFORMATION(W)SYSTEM(W)LANGUA
 5 27 INTELLIGEN?(W)SYSTEM?
 6 10196 LANGUAGE?
 7 4 5 AND 6
 8 448 NATURAL(W)LANGUAGE?
 9 0 4AND8
 10 1 5AND8
 11 115 TEXT(W)PROCESS?
 12 0 5AND11
 ? S INFORMATION((W)SYSTEM?
 13 8968 INFORMATION(W)SYSTEM?
 ? C8AND13
 14 38 8AND13
 ? C3AND14
 15 20 3AND14
 ? T15/2/1-20

Przyczyną niepowodzeń w wyszukiwaniu, niezależnie od nietrafnego doboru wyrażen w instrukcji wyszukiwawczej, bywa rozbieżność ortograficzna wyrażen instrukcji i zbioru informacyjnego systemu lub supletywność rdzeni wyrazowych. Źródłem poważnych trudności dla użytkowników europejskich są różnice w ortografii brytyjskiej i amerykańskiej, np. *CENTER* — *CENTRE*, *THEATER* — *THEATRE*, łączna lub rozdzielna pisownia wyrażen złożonych (*PAPERMILL* — *PAPER MILL*). Środkiem systemowo włączanym do reguł języka wyszukiwawczego jest operacja maskowania i obcinania (*truncation*), polegająca na neutralizacji co najmniej jednej cechy dystynktywnej wyrażenia składowego instrukcji wyszukiwawczej, najczęściej litery. Wykładnikiem tej operacji jest symbol „?”. W systemie *DIALOG* możliwości wyszukania wyrażen maskowanych ograniczają się do jednostek indeksowych (indeksu podstawowego lub indeksów innych pól wyszukiwawczych). Jest ona stosowana w zdaniach rozpoczynających się komendą *EXPAND*, *SELECT* lub *SUPERSELECT*, na przykład:

?S ON?LINE — neutralizacja znaku „-”

?S TRANSFORMER? — neutralizacja końcówki liczby mnogiej lub innych morfemów, dzięki czemu możliwe jest wyszukanie znacznej liczby wyrażen indeksu, należących zarówno do tezaury jak i słów kluczowych :

1) ?E TRANSFORME?

E 1 TRANSFORMED

E 2 TRANSFORMEDTO

E 3 TRANSFORMEE

E 4 TRANSFORMEER
E 5 TRANSFORMEES
E 6 TRANSFORMER
E 7 TRANSFORMER BUBBLING
E 8 TRANSFORMER CAPACITANCE MODELS
E 9 TRANSFORMER CHARGING
E10 TRANSFORMER COILS
E11 TRANSFORMER CORE
E12 TRANSFORMER FEEDBACK
E13 TRANSFORMER INDUCTANCE MODELS
E14 TRANSFORMER LIFE
E15 TRANSFORMER LOADING POLICIES
E16 TRANSFORMER OIL
E17 TRANSFORMER SHEET
E18 TRANSFORMER SHEETS
E19 TRANSFORMER STEEL
E20 TRANSFORMER SUBSTANTIIONS
E21 TRANSFORMER WINDINGS
E22 TRANSFORMER-RECTIFIERS
E23 TRANSFORMER — TERMINATED LINES
E24 TRANSFORMER'S
E25 TRANSFORMERETTES
E26 TRANSFORMERLESS
E27 TRANSFORMERS
E28 TRANSFORMERS, PULSE
E29 TRANSFORMERSAND
E30 TRANSFORMES
E31 TRANSFORMESS
E32 TRANSFORMHOLOGRAMS
E33 TRANSFORMIERENDE

2) ?S AU=SAWYER?

1 498 AU=SAWYER

3) ?S JN=SCIENTIFIC AMER?

2 615 JN=SCIENTIFIC AMER?

4) ?S PY=197?

3 40252389 PY=197?

5) ?S AGGRESS?/TI, DE

4 780 AGGRESS?/TI, DE

W przytoczonych przykładach liczba neutralizowanych znaków nie jest specyfikowana. System wyszukuje wszelkie wyrażenia złożone ze znaków następujących po symbolu operacji. W rozwiniętej wersji operacja maskowania może być realizowana następującymi sposobami:

a) ze specyfikacją maksymalnej liczby znaków neutralizowanych — na końcu wyrażenia:

?S SELECT HORSE??

?S THEAT???

w wyniku czego system wyszukuje wyrażenia o określonej długości;

b) przez wbudowanie symbolu maskowania między znaki składowe wyrażenia (*embedded*):

?S WOM?N

?S ORGANI?ATION/TI, DE

c) przez kombinację sposobu a) oraz b):

?S WORKM?N?.

4.2.4.3. Operatory tekstowe

Rozszerzenie opcji wyszukiwawczych w systemie online do w zasadzie tych wszystkich cech dystynktywnych obiektów, które są wyróżnione jako pola wyszukiwawcze reprezentowane przez wyrażenia odpowiednich indeksów bazy danych, powoduje — jak wskazywaliśmy — rozbudowę wiedzy o planie wyrażania języków informacyjnych o notacji paranaturalnej dzięki transformacjom wyrażen złożonych. Wiedza o rzeczywistości pozajęzykowej jest ograniczona do tych wykładników, które są tradycyjnie stosowane w słownikach języków deskryptorowych (kategorii semantycznych, wykładników relacji mereologicznych i skojarzeniowych). Na poziomie syntagmatycznym *nie stosuje się w praktyce wskaźników roli* jako wykładników sytuacyjnych związków między desygnatami wyrażen. Możemy stwierdzić, że obowiązek znajomości relacji między wyrażeniami języka naturalnego i JIW pełniących funkcję kluczy wyszukiwawczych i obiektami rzeczywistości pozajęzykowej zostaje przeniesiony na użytkownika systemu (pracownika służb informacyjnych lub bezpośrednio zainteresowanego specjalistę). W porównaniu z systemami tradycyjnymi zmienia się także sposób, w jaki system pomaga użytkownikowi wskazać związki zachodzące w tekstach charakterystyk wyszukiwawczych między wyrażeniami składowymi. Sposób ten jest pochodną operacji dokonywanych na zbiorach bazy danych. W literaturze przedmiotu często jest określany nieprecyzyjnym terminem wyszukiwanie pełnotekstowe (*full-text searching*).

Wyszukiwaniem pełnotekstowym nazywa się różne operacje selekcjonowania informacji w systemie online, najczęściej interpretując je jako:

a) wyszukiwanie w systemie, którego zbiór tworzą teksty dokumentów pierwotnych, a więc wyszukiwanie odbywa się w „pełnych tekstach” dokumentów (pierwotnych). Do systemów takich, funkcjonujących w skali pełnej eksploatacji, należy komercyjny system informacji legislacyjnej *LEXIS* w USA [Grabowska M., 1986, 89-102];

b) wyszukiwanie za pomocą operatorów tekstowych w pełnym tekście, na przykład w streszczeniu, charakterystyce wyszukiwawczej dokumentu;

c) wyszukiwanie w dowolnych wskazanych ciągach wyrażen charakterystyki wyszukiwawczej dokumentu.

W omawianych przykładowo systemach online dostępnych w skali międzynarodowej nie stosuje się wyszukiwania pełnotekstowego w znaczeniu pierw-

szym (a), ponieważ zbiory informacyjne tych systemów tworzą bądź charakterystyki wyszukiwawcze dokumentów, bądź charakterystyki wyszukiwawcze obiektów pozajęzykowych (zbiory faktograficzne). Jako cechę dystynktywną wyszukiwania pełnotekstowego w znaczeniu (b) i (c) podaje się możliwość stosowania *operatorów tekstowych*, interpretowanych jako wyrażenia języka wyszukiwawczego wyspecjalizowane w funkcji metatekstowej, a więc w określaniu warunków, jakie powinny spełniać wyrażenia wskazane w instrukcji wyszukiwawczej, aby mogły być uznane za relewantne. Do warunków tych należą:

- odległość wyrażen (pojedynczych wyrazów) w tekście, określana za pomocą tzw. operatorów sąsiedztwa (*proximity operators*);
- rodzaj kontekstu, w którym powinno wystąpić wyrażenie w charakterystyce wyszukiwawczej dokumentu.

Możliwość łącznego stosowania wykładników operacji maskowania ze wskaźnikami położenia wyrażenia w tekście jest różna w poszczególnych systemach. Należy podkreślić, że w odróżnieniu od wykładników relacji semantycznych w zdaniu JIW w systemach tradycyjnych, operatory tekstowe systemów online są jedynie wykładnikami relewancji technicznej, a więc określającej zgodność wyrażen ChWD z wyrażeniami instrukcji wyszukiwawczej w planie wyrażenia.

Do wskazania kryteriów relewancji przez specyfikację odległości wyrażen w tekście służą najczęściej dwa symbole, interpretowane jako „przyległy, sąsiadujący” (*adjacent*):

(*W*) — zawsze umieszczany w nawiasie, pełniącym funkcję delimitacyjną (stosowany np. w systemie *DIALOG*, *ESA-IRS*);

[*ADJ*] — również umieszczany w nawiasie, stosowany w systemie *BRS*;

(*NBR*) — w systemie *SDC ORBIT*, *BLAISE*.

Umieszczenie tego operatora między dwoma wyrażeniami oznacza, że wyrażenie stojące po nim powinno następować bezpośrednio po wyrażeniu stojącym przed nim, np.:

?*S SOLAR(W)ENERGY = SOLAR ENERGY*

?*S DEEP(ADJ)SPEEN = DEEP SPEEN*

?*S NUCLEAR(ADJ)POWER(ADJ) (PLANTS OR STATIONS)*

Operator sąsiedztwa można rozbudować przez dodanie przed nim wskaźnika liczbowego *n*, np. (*1W*), (*3W*). Precyzowanie liczby wyrażen rozdzielających wyrażenia instrukcji jest istotne ze względu na możliwość wystąpienia między wyrażeniami relewantnymi wyrażen zakazanych tzw. *stop words*, automatycznie pomijanych przez system przy tworzeniu indeksów pól tekstowych na podstawie wspomnianej *stop listy*. W ten sposób operator tekstowy pełni funkcję ukrytego wskaźnika wagi dla wyrażen niewspółrzędnych w ChWD. Na przykład, w celu wyszukania tytułu powieści *Przeminęło z wiatrem* (*Gone with the wind*) w języku angielskim konieczne jest użycie operatora w postaci: ?*S GONE(2W)WIND*.

Określanie rodzaju kontekstu, w którym powinno wystąpić wyrażenie bez względu na jego odległość od innych wyrażen i ich kolejność, umożliwiają op-

ratory wskazujące jednostki strukturalne tekstu spójnego, a więc rekordu, pola (*field*), podpola (*subfield*), cytacji (*citation*). Rozbudowany system operatorów kontekstowych, stosowanych łącznie z operatorami maskowania oraz operatorami logicznymi, opracowano dla systemu *DIALOG* (umieszczane w nawiasach):

(*F*) — wyrażenia powinny wystąpić w tym samym polu: *SOLAR(F)ENERGY*;

(*C*) — wyrażenia powinny wystąpić w tej samej cytacji, w dowolnym polu: *SOLAR(C)ENERGY*,

(*L*) — (*link*) stosowany tylko w tych bazach, w których używany jest język deskryptorowy, umożliwiający wyszukiwanie hierarchiczne, a więc w tych ChWD, które zawierają w polu opisu deskryptorowego wyrażenia pozostające w relacji paradygmatycznej z wyrażeniami (deskryptorami) instrukcji wyszukiwawczej: *SOLAR(L)ENERGY*; *ANTENNA?(L)TELEVISION*,

(*S*) — stosowany w tych bazach, w których w strukturze rekordu wyróżnia się podpola wyszukiwawcze; wskazane wyrażenia instrukcji powinny wystąpić w podpolach w określonej kolejności: *SOLAR(S)ENERGY*).

Użycie operatorów kontekstowych jest niekiedy funkcjonalnie równoważne użyciu komendy *SUPERSELECT*, oznaczającej polecenie wyszukania ChWD według dowolnych pól, w połączeniu z operatorami logicznymi, np. (*DIALOG*, *Guide to searching*):

?S ARTHRITIS(C)ACUPUNCTURE

1 41 ARTHRITIS(C)ACUPUNCTURE

?SS ARTHRITIS AND ACUPUNCTURE

19555 ARTHRITIS

875 ACUPUNCTURE

2 41 ARTHRITIS AND ACUPUNCTURE

Zastosowanie operatorów odległości i operatorów kontekstowych nie jest ograniczone do pól tekstowych, jakkolwiek jest w nich najczęstsze. Okazuje się pomocne w wyszukiwaniu:

?S CS=JOHNS?(C)CS=HOPKINS — CS=Cited source (źródło cytowane)

?S CA=JOHNS?(C)CS=HOPKINS

Kombinacje przedstawionych operatorów dają różne możliwości precyzowania instrukcji wyszukiwawczej, które przytaczamy w porządku wzrastającej precyzji [*Guide to DIALOG Searching*, 1979]:

?S ANTENNA?(C)MOBILE

?S ANTENNA?(C)LA=FRENCH

?S ANTENNA?(F)MOBILE

?S ANTENNA?(F)CS=WESTERN

?S ANTENNA?(F)MOBILE/DE

?S ANTENNA?(F)MOBILE/DE,/TI

?S MOBILE(*nW*)ANTENNA?

?S CS=LIBRARY(*nW*)CS=CONGRESS

?S MOBILE(*W*)ANTENNA?

?S CS=NEW(W)CS=JERSEY
?S MOBILE(W)ANTENNA?(F)MICROWAVE
?S CS=NATIONAL(W)CS=LIBRARY(F)CS=MEDICINE
?S ANTENNA(L)TELEVISION>

W literaturze od lat prezentowane są różne oceny wyników wyszukiwania pełnotekstowego oraz kontrolowanego, ograniczonego do opisów deskryptorowych lub haseł przedmiotowych. Ocena efektywności różnych metod formułowania strategii stanowi bardzo obszerne zagadnienie, wykraczające poza ramy naszych rozważań, wymagałaby ponadto dysponowania dostępem do któregoś ze wspomnianych systemów przez dłuższy czas. Opinie użytkowników profesjonalnych i nieprofesjonalnych, prezentowane na łamach czasopism wyspecjalizowanych w problematyce systemów online (*Database, Online Review*), formułowane są zwykle na podstawie doświadczeń związanych z określoną bazą danych lub grupą baz pokrewnych tematycznie, zebranych przy obsłudze określonej grupy użytkowników. Spory o relewancję w systemie online kończą się zwykle wnioskiem o uzależnieniu sposobu formułowania strategii od potrzeb użytkowników, celu wyszukiwania, środków finansowych oraz kosztów wyszukiwania w poszczególnych bazach. Są de facto sporami o ocenę relewancji systemowej i pertynencji, przy różnych strategiach wyszukiwawczych. Z eksperymentalnych badań (case studies) nad wynikami wyszukiwania informacji w zakresie ochrony środowiska w USA (US ENVIRONMENTAL PROTECTION AGENCY, Ohio) wynika, że wbrew potocznym opiniom o efektywności wyszukiwania pełnotekstowego, lepszą trafność osiągnęto posługując się kontrolowanymi słownikami JIW do budowy instrukcji wyszukiwawczych, zawierających od dwóch do czterech wyrażań, reprezentujących terminy nadrzędne w systemie paradygmatycznym. Wyszukiwanie pełnotekstowe oceniano jako bardziej efektywne przy szerokich profilach tematycznych, określanych w celu zebrania np. kompletnej bibliografii [Calkins M. L., 1980, 53-67].

4.2.5. Wiedza o rzeczywistości w systemie online

Możliwość dokonywania transformacji tekstów charakterystyk wyszukiwawczych dokumentów i stosowania różnych kluczy wyszukiwawczych w systemie online powoduje ograniczanie wiedzy pozajęzykowej w systemie, przy równoczesnej rozbudowie wiedzy meta-metainformacyjnej (informacji o bazach danych) oraz wiedzy o języku komunikowania się z systemem. Charakterystyczne jest ograniczenie wykorzystania w funkcji metainformacyjnej i wyszukiwawczej tradycyjnych klasyfikacji bibliotecznych. W nielicznych komercyjnie udostępnianych bazach danych katalogowych bibliotek amerykańskich stosowana jest Klasyfikacja Dziesiąta Deweya. Wyjątek stanowią klasyfikacje patentowe, pełniące *de facto* (jak wcześniej wykazano) funkcję kodów semantycznych w opisie nowo powstałych obiektów rzeczywistości pozajęzykowej, a także klasyfikacje (systematyzacje) obiektów w naukach przyrodniczych, np. *BIOSIS PREVIEW*S.

Spośród sposobów odwzorowania wiedzy pozajęzykowej należy wymienić przede wszystkim te, które są właściwe językom deskryptorowym na ich osi paradygmatycznej:

- kategoryzację obiektów pozajęzykowych, zróżnicowaną w tezaursach poszczególnych baz danych,
- wykładniki relacji mereologicznych oraz skojarzeniowych.

Kategoryzacja obiektów, jako zabieg organizujący pole semantyczne języka, jest często stosowana w faktograficznych bazach danych, zwłaszcza zawierających informację handlową i polityczną, których słowniki obejmują nazwy tysięcy obiektów (np. PREDICASTS). Przyporządkowanie obiektom odpowiednich atrybutów i spełnienie warunku rozłączności ich zbiorów wymaga narzucenia JIW pewnego „szkieletu organizacyjnego” w postaci kategoryzacji. Potwierdza się przy tej okazji uniwersalność języków deskryptorowych jako języków systemów dokumentacyjnych i faktograficznych. Należy przy tym zauważyć umowność podziału systemów (baz danych) na faktograficzne i dokumentacyjne, wynikające m.in. z tego, że zarówno charakterystyki wyszukiwawcze obiektów w bazach faktograficznych, jak i ich słowniki budowane są na podstawie analizy dokumentów: prasy codziennej (np., *NEWS-SEARCH — DIALOG*), wydawnictw artystycznych, biuletynów agend rządowych i międzynarodowych (np. *PREDICASTS FEDERAL INDEX, PTS, PTS FEDERAL INDEX WEEKLY, PTS Funk and Scott Indexes, PTS INTERNATIONAL STATISTICAL ABSTRACTS, PTS PROMT, PTS US ANNUAL TIME SERIES, PTS US STATISTICAL ABSTRACTS, PTS WEEKLY* i in.).

4.2.6. Problemy komunikacji użytkownika z systemem

Budowa sieci teleinformatycznych, umożliwiających szybkie przesyłanie informacji między systemem i użytkownikiem, oraz opracowanie interakcyjnych programów wyszukiwawczych, dokonujących transformacji zbiorów informacji w sposób rozszerzający repertuar kluczy wyszukiwawczych w systemie (a więc zbiór wykładników relewancji), wpłynęły na zmianę proporcji w zakresie wiedzy wbudowanej w JIW, powiększając zakres wiedzy o języku. Oznacza to równocześnie, że użytkownik korzystający z systemu rozstrzyga o relewancji wyselekcjonowanej informacji na podstawie własnej wiedzy specjalistycznej. Sposób reprezentacji wiedzy w systemie, głównie w postaci zbiorów terminów indeksów uporządkowanych alfabetycznie — w odróżnieniu od „kompedium wiedzy”, jakimi są tradycyjne klasyfikacje typu UKD, stanowi źródło zarzutów kierowanych pod adresem systemów online, iż ich zbiory są zbiorami nie powiązanych logicznie elementów informacji (rekordów), które nie odzwierciedlają sposobu myślenia człowieka i nie są przystosowane do zindywidualizowanego korzystania z nich — są „zbiorami informacji lecz nie są zbiorami wiedzy”.

Jako drugą zasadniczą wadę systemów online, funkcjonujących w latach siedemdziesiątych, wymienia się sztywność języka wyszukiwawczego (języka kwerend), przystosowanego do protokołów konwersacji w poszczególnych systemach i dającego *de facto* niewielkie możliwości indywidualnego określania

potrzeb informacyjnych w procesie komunikacji z systemem, a więc szczegółowej specyfikacji kryteriów relewancji. Wiąże się to z postulatem budowy odrębnych modułów komunikacji (*interface*), stanowiących element podsystemu lingwistycznego wraz ze zintegrowanymi leksykonami JIW (Vocabulary switching systems, authority files — kartoteki wzorcowe). Prace w tym zakresie prowadzono w ośrodkach uniwersyteckich oraz jednostkach badawczych systemów komercyjnych (Lockheed) — np. nad językiem CITE (*Current Information Transfer in English*), IIDA (*Individualized Instruction for Data Access*) dla systemu DIALOG [Goldstein Ch., M., Ford W., 1978, 270-278] [Gregory N., 1979].

Jako pożądane cechy języka komunikacji z systemem i odpowiedniego oprogramowania wymienia się:

- możliwość wykorzystania dokumentów w języku naturalnym do automatycznej budowy baz danych faktograficznych,
- możliwość formułowania instrukcji wyszukiwawczych w języku zbliżonym do naturalnego i „rozumienia” tych instrukcji przez system, a także generowania odpowiedzi.

Budowa tzw. psychologicznych baz danych z dostępem w języku naturalnym, przetwarzających „wiedzę” (przeciwstawianą informacji) stanowi odrębne zagadnienie, posiadające obszerną literaturę [Bolc L., (red.), 1983]. Z ich powstaniem wiąże się nowa interpretacja relewancji, uwzględniająca dynamizm procesów komunikacji oraz indywidualizację reprezentacji wiedzy.

4.3. WYSZUKIWANIE INFORMACJI Z TRANSFORMACJĄ SEMANTYCZNĄ

Udostępnianie wiedzy w systemach inteligentnych (eksperckich/ekspertowych, doradczych, systemach z bazą wiedzy) charakteryzuje dokonywanie transformacji informacji. Transformacji dokonuje się w dwu etapach:

- komunikowania się z użytkownikiem, na co składa się interpretacja zapytania oraz przekazywanie odpowiedzi,
- generowania odpowiedzi na podstawie zgromadzonej w bazie wiedzy faktograficznej (deklaratywnej i proceduralnej).

Użytkownicy SIW uważają, że optymalnym sposobem komunikowania się z systemem jest język zbliżony do naturalnego, eliminujący niedogodności języków kwerend: skomplikowaną składnię oraz konieczność znajomości organizacji informacji w faktograficznej bazie danych. Wyposażenie systemu w moduł komunikacji w języku naturalnym polega w praktyce na wzbogaceniu funkcji interpretera (por. rozdz. 3.4. schemat 11), działającego w każdym programie komputerowym, o reguły transformacji informacji umożliwiające:

- przeprowadzanie analizy gramatycznej tekstów zapytań (przez tzw. *parser*),
- interpretację semantyczną tych tekstów,
- generowanie instrukcji wyszukiwawczej w języku reprezentacji pozajęzykowej wiedzy, przyjętym w określonej bazie danych,

- generowanie odpowiedzi w języku naturalnym po jej uprzednim sformułowaniu w języku bazy danych faktograficznych, w wyniku realizacji procesów wnioskowania pod kontrolą reguł heurystyki systemu.

Moduł komunikacji z użytkownikiem dokonuje więc transformacji informacji na wejściu i wyjściu systemu, polegającej na zmianie języka reprezentacji wiedzy: naturalny - sztuczny - naturalny.

Transformacje wewnątrzsystemowe prowadzące do wygenerowania odpowiedzi relewantnej polegają na wyselekcjonowaniu informacji o obiektach i sytuacjach z bazy deklaratywnej i przeprowadzeniu wnioskowania — na mocy wiedzy proceduralnej — dostarczającego nowej informacji, nie występującej w zbiorach systemu.

Udostępnianie wiedzy przez system polega więc na wykorzystaniu tej samej wiedzy pozajęzykowej, w celu zinterpretowania zapytania i wywnioskowania odpowiedzi, oraz wiedzy o języku naturalnym, niezbędnej do przekładu zapytania na reprezentację wewnątrzsystemową, a następnie do przekładu odpowiedzi w języku wewnętrznej reprezentacji na język naturalny. Wiedza o języku i o rzeczywistości pozajęzykowej jest reprezentowana oddzielnie, jakkolwiek różni się jedynie ukierunkowaniem jej wykorzystania w różnych cyklach działania systemu [Cichy M., 1983].

4.3.1. Dostęp do wiedzy w języku naturalnym

Przetwarzanie informacji w języku naturalnym w systemach „rozumiejących” ten język i poprawnie posługujących się nim w generowanych odpowiedziach ma już ponad czterdziestoletnią tradycję, związaną z różnym wykorzystaniem prac, których zasadniczym celem było odwzorowanie w sposób sformalizowany struktury syntaktycznej i semantycznej tekstów. Tworzenie modułów komunikacji z bazami danych reprezentuje jeden z nowszych kierunków zastosowań badań w tym zakresie, prowadzonych w latach siedemdziesiątych i stale kontynuowanych⁸. Prace wcześniejsze wiązały się z:

- a) przekładem komputerowym,
- b) modelowaniem zachowań językowych człowieka.
- c) próbami generowania wypowiedzi w systemach konwersacyjnych umożliwiających wychwytywanie i eksponowanie niekonsekwencji w wypowiedziach użytkownika analizowanych w celach diagnozy psychiatrycznej (na przykład w klasycznym już systemie *ELIZA J. Weizenbauma*),
- d) transformacją tekstów języka naturalnego prowadzącą do tworzenia baz danych faktograficznych, różniącą się od procedur automatycznego indeksowania w systemach dokumentacyjnych stopniem strukturalizacji wiedzy o obiektach denotowanych przez wyrażenia językowe.

⁸ Z ankiety przeprowadzonej w roku 1982 wynika, że funkcjonowało wówczas ponad sto systemów przetwarzających język naturalny. Jak stwierdza M. Wallace [Wallace M., 1984] prace wcześniejsze można było podsumować w jednym referacie.

Literatura na temat systemów przetwarzających język naturalny i sposobów reprezentacji odwzorowanej w nim wiedzy jest bardzo obszerna [Bolc L., Cichy M., Różańska L., 1982]. Scharakteryzujemy jedynie założenia metodologiczne właściwe podejściom stosowanym we wczesnym okresie oraz rozwiązaniom nowszym. Addis stwierdza, że obecnie można wyróżnić dwie generacje systemów rozumiejących język naturalny (*NLU — natural language understanding*) [Addis T. R., 1985].

Generacja pierwsza, do której należą systemy lat pięćdziesiątych i sześćdziesiątych (w roku 1965 wymieniano 15-16 systemów) związana była z próbami automatyzacji przekładu opartego na klasyfikacji wyrażen w języku źródłowym i docelowym oraz wnioskowaniu statystycznym⁹, a także z budową systemów konwersacyjnych, w których rozumienie tekstu pytania użytkownika polegało na wyszukaniu w nim słów kluczowych, którym przypisana była standardowa odpowiedź, traktowana jako przypuszczalna ich implikacja, np.:

Użytkownik: *I do not enjoy rain in July* (Nie lubię deszczu w lipcu).

System: *Well, we don't usually have rainy weather in July, so you will probably not be disappointed.* (Zazwyczaj nie mamy deszczowej pogody w lipcu, więc przypuszczalnie nie będziesz zawiedziony) [Addis T. R., 1985].

W tej metodzie reprezentacji wiedzy zakładano, że podstawową jednostką znaczenia są słowa (words), łączone w odwzorowaniach zapytań za pomocą spójników logicznych. Nie dysponowano sformalizowanymi modelami gramatyk języka naturalnego, ograniczając się do analizy frazowej umożliwiającej eliminowanie pewnych struktur wieloznacznych.

Druga generacja, rozwijana od połowy lat sześćdziesiątych charakteryzuje się zastosowaniem formalnych metod definiowania struktur syntaktycznych (frazowych), z uwzględnieniem zależności kontekstowych. Metody inspirowane w dużym stopniu teorią gramatyk generatywno-transformacyjnych Chomsky'ego są bardzo zróżnicowane w poszczególnych systemach. W obszernej literaturze zagadnienia wymienia się najczęściej: formalizmy logiczne, gramatyki transformacyjne, rekursywne gramatyki sieciowe [Woods W. A., 1975, 111-154] oraz rozszerzone modele gramatyk sieciowych [Bolc L., Cichy M., Różańska L., 1982]. Jeden z pierwszych programów analizy gramatycznej Thoma (1968) przydzielał słowom w tekście (kontrolowanym przez słownik) etykiety syntaktyczne. Wyróżniano w ten sposób *dwie klasy słów: otwartą* (open), zawierający rzeczowniki, których rolę syntaktyczną można określić bez znajomości ich znaczenia oraz *zamkniętą* (closed) zawierającą przyimki, istotne w definiowaniu relacji syntaktycznych w zdaniu.

M. Cichy podkreśla, że charakterystyczne dla rozwoju sformalizowanych metod analizy tekstów języka było:

— w początkowym okresie (a więc w latach sześćdziesiątych) — modularne, „poziome” traktowanie struktury tekstu, polegające na rozdzieleniu kolejnych etapów analizy morfologicznej, syntaktycznej i semantycznej,

⁹ Pierwszy przekład maszynowy zrealizowano dla paru zdań języka rosyjskiego w roku 1954 w Georgetown University, przy słowniku zawierającym około 250 wyrażen.

— w systemach nowszych ściśle wiązanie analizy morfologiczno-syntaktycznej z analizą semantyczną, wprowadzaną już w fazie wczesnej analizy syntaktycznej.

Zasadniczy przełom w metodach analizy syntaktycznej i semantycznej tekstów w systemach prowadzących dialog z użytkownikiem wiąże się z teorią Katza i Fodora [Katz J. J., Fodor J. A., *Language*, 52], w której wyjaśniono różnice między składnią i semantyką, postulując analizę zdania w kategoriach semantyki. Od tego momentu rozpoczyna się — jak podkreślają autorzy prac przeglądowych — „era semantyki i sieci semantycznych”, których pierwotna koncepcja łączona jest z Quillianem (por. rozdz. 3.4.2.2.) i jego modelem pamięci konceptualnej. Wiedza pozajęzykowa o obiektach, i zachodzących między nimi relacjach sytuacyjnych, okazuje się niezbędna do interpretacji zdania, wyznacza bowiem relacje między jego składnikami. Analiza wyrażen składowych komunikatów kierowanych do systemu przez użytkownika odwołuje się do zakładanego poziomu wiedzy o pewnym fragmencie rzeczywistości, zakresu realizowanych zadań oraz procedur interpretacji tej wiedzy. Przegląd metod reprezentacji wiedzy pozajęzykowej w faktograficznej bazie danych, poczynając od systemów logiki pierwszego rzędu, poprzez sieci semantyczne i złożonej struktury ram i skryptów, przedstawiliśmy w rozdz. 3.4.2.5.

Oprócz teorii Katza i Fodora, dla kształtowania się koncepcji opisu struktury semantyczno-syntaktycznej zdania w systemach *NLU*, duże znaczenie miała teoria przypadków głębokich Fillmore'a, opisująca zdanie (*S* — *sentence*) jako pewną konstrukcję złożoną ze zdania w sensie logicznym (*P* — *preposition*), na którą nałożona jest pewna rama modalna (*M* — *modality*) definiowana przez negację, czas, tryb, aspekt ($S \rightarrow M + P$). Głównym przedmiotem opisu struktury zdania były relacje semantyczne, odwzorowujące związki między obiektami rzeczywistości opisywanej w zdaniu (*P*). Relacje te, reprezentowane przez tzw. *funkcje przypadku* (*Kasus functions*), są konotowane przez czasownik (predykat): $P \rightarrow V + C_1 + C_2 + C_3 + \dots + C_n$. Czasowniki klasyfikowane są według tzw. ram przypadków (*case frames*), konotujących struktury z użyciem określonych klas rzeczowników.

Zasadniczym ustaleniem wynikającym z teorii Fillmore'a było stwierdzenie, że istnieje skończony zbiór relacji opisujący rzeczywistość w sposób wystarczający do wykazania podobieństwa semantycznego zdań różnych syntaktycznie¹⁰. Z kolei, do ustalenia zakresu generowanych zdań sensownych za konieczne uznano uwzględnienie ograniczeń wyznaczanych na podstawie klasyfikacji semantycznej czasowników (*case frames*). W ten sposób strukturę semantyczną

¹⁰ Fillmore wyróżnił 17 przypadków. Do najczęściej wymienianych należą: *agentivus* — wyrażający agensa; *benefactivus* — oznaczający istotę żywą, ze względu na którą odbywa się czynność nazwana czasownikiem *Janek zrobił to dla Pawła*; *comitativus* — oznaczający stosunek towarzyszenia *Janek z Pawłem poszli do szkoły*; *dativus* — istotę żywą, która nie jest agensem, a której dotyczy czynność lub stan *Janek wierzył, że*; *factitivus* — przedmiot będący rezultatem czynności *Janek zrobił stół*; *instrumentalis* — narzędzie — *Janek uderzył Pawła kijem*; *locativus* — miejsce *W Chicago jest mroźnie*; *pacjens* — odbiorca czynności *Wiatr otworzył drzwi, Drzwi się otworzyły*. [Encyklopedia językoznawstwa ogólnego, 1993, 190]

zdania można opisać dla potrzeb analizy i syntezy przez ustalenie dla zawartych w nich czasowników konotowanych przez nie przypadków głębokich, wypełnianych przez nazwy obiektów i ich atrybutów zaczerpniętych z faktograficznej bazy danych.

Do teorii przypadków głębokich i konotacji czasownika lub innych części mowy nawiązywano w pracach nad automatycznym indeksowaniem realizowanych dla firmy Lockheed na początku lat siedemdziesiątych [Addis T. R., 1985]; w ramach tych prac zbudowano tablice konotacji dla około 8000 słów języka angielskiego, opisywanych w kategoriach rządzenia przypadkami (word government). W rezultacie powstał jednak czysto syntaktyczny (walencyjny) słownik, uwzględniający tzw. wyjątki (około 1000) dla części mowy konotujących inne konstrukcje składniowe niż założone w systemie.

Teoria konotacji składniowej pozwalała częściowo wyjaśnić problem relacji między semantyką i składnią zdania, prowadziła bowiem do stwierdzenia, że podstawową zasadą konotacji jest jej semantyczny charakter. Prace poszczególnych autorów związanych z tym kierunkiem badań różnią się przyjmowanym zakresem konotacji składniowej:

— w gramatyce przypadków Fillmore'a własność tę przypisuje się tylko czasownikom,

— w teorii zależności konceptualnej Schanka oraz teorii elementarnych cech semantycznych Wilksa własność tę przypisuje się wszelkim wyrażeniom elementarnym używanym odpowiednio w ramach i skryptach (Schanka) oraz trójkach (Wilksa) (por. rozdz. 3.4.2.3., 3.4.2.6.),

— Winograd posługuje się metodą konotacji w systemie SHRDLU w opisie twierdzeń specyfikujących znaczenie wyrażen [Winograd T., 1980, 209-241].

Zauważmy wreszcie, że wszystkie teorie pamięci konceptualnej operujące kategoriami prototyp obiektu, stereotypowe działanie lub stereotypowa sytuacja (np. w teorii Schanka) odwołują się do konotacji semantycznej.

Wszystkie systemy interpretują i generują zdania jako podstawowe jednostki wypowiedzenia (ewentualnie skupione w paragrafy), których opis musi uwzględniać analizę morfologiczną na poziomie paradygmatyki i syntagmatyki języka. Budowa każdego systemu zapewniającego dostęp do bazy danych faktograficznych związana jest z pracą nad słownikiem specyfikującym systemowe relacje między jego jednostkami drogą kategoryzacji i ustalania związków hierarchicznych i skojarzeniowych.

Udzielanie przez system odpowiedzi w języku naturalnym łączy się z zadaniem określenia sposobów interpretacji semantycznej zdań wygenerowanych w wyniku transformacji wiedzy zawartej w bazie danych (przeprowadzonego wnioskowania) — por. [Mykowiecka A., 1987; 1992]. Każdy z systemów pod tym względem jest ograniczony, ponieważ nie umożliwia interpretacji zdań niepoprawnych gramatycznie, chociaż zrozumiałych dla ludzi, a przy tym może generować zdania poprawne, lecz niezrozumiałe z powodu ograniczenia wiedzy pozajęzykowej.

Jako główny powód zróżnicowania systemów rozumiejących język naturalny wskazuje się rozdzielanie wiedzy niezbędnej w przetwarzaniu języka na wiedzę deklaratywną i proceduralną, co widoczne jest w różnych mechanizmach inferencyjnych. Żaden ze zbudowanych dotychczas systemów nie może stanowić przykładu uniwersalnej metody generowania relewantnej informacji [Addis R., 1985].

W pracach poświęconych przeglądowi systemów rozumiejących język naturalny podkreśla się znaczenie wiedzy faktograficznej dla tzw. kompletności systemu. W systemach wczesnych bazę danych faktograficznych traktowano jako dodatek do implementacji rozbudowanego modułu konwersacyjnego, koncentrując się na analizie i syntezie struktur syntaktycznych i morfologicznych wypowiedzi [Cichy M., 1983, 103-136]. Wspomniany przełom w metodach opisu składni języka naturalnego łączony z pracami Fodora, Katza i Fillmore'a oznaczał w pragmatyce systemów zmianę podejścia do wiedzy pozajęzykowej, co polegało na traktowaniu wiedzy o świecie (tzw. wiedzy ogólnej i wiedzy szczegółowej) oraz wiedzy o języku jako komplementarnych, jakkolwiek oddzielnie reprezentowanych w systemie. Umożliwia to wykorzystywanie modułów konwersacyjnych niezależnie od realizacji dialogu ze specjalistycznymi bazami danych faktograficznych.

Mykowiecka przedstawiła charakterystykę kilkunastu systemów (modułów) generujących odpowiedzi w języku naturalnym [Mykowiecka A., 1987]. Jako cechy dystynktywne wymienia:

- względną niezależność bazy wiedzy o świecie zewnętrznym, reprezentowanej najczęściej za pomocą sieci semantycznych i elementarnych cech semantycznych języka predykatów (listy obiektów i ich cech), a także ram;
- prezentowanie wiedzy o języku, z podziałem na słownik i gramatykę lub też łącznie we wspólnych strukturach, zależnie od przyjętego modelu gramatyki;
- repertuar struktur językowych dostosowanych do ich funkcji w dialogu z użytkownikiem (na przykład polecenia, przekazania komunikatu, pytania, zadania identyfikacji obiektu w bazie danych) oraz zasady łączenia zdań w dłuższe odpowiedzi.

Na przykład system *TEXT*, opracowany w uniwersytecie stanu Pensylwania USA, udostępnia informacje o stanie jednostek pływających floty amerykańskiej, odpowiadając na pytania typu: *co wiesz o <obiekt>*, *co to jest <obiekt>*, *jaka jest różnica między <obiekt> i <obiekt>*. Wiedza o rzeczywistości jest tu reprezentowana w hierarchicznie zorganizowanych sieciach semantycznych zawierających nazwy obiektów (statków) i innych obiektów z nimi związanych, identyfikowane za pomocą klas atrybutów odpowiadających cechom wyglądu i parametrom technicznym. W organizacji relacji między atrybutami wykorzystuje się związki paradygmatyczne.

Udzielenie odpowiedzi polega na wykorzystaniu schematów retorycznych, do których utworzenia posłużyła kategoryzacja zdań określająca ich zakładane

funkcje w wypowiedzi. Wyróżniono kategorie zdań stosowane następnie do ustalenia tzw. schematów przebiegu dialogu:

- 1) **opis atrybutu** — *Maja ma żółtą sukienkę;*
- 2) **zrównoważenie** — *Wina określane jako wspaniałe to wina z dobrych winnic;*
- 3) **uszczerbowienie faktu** — *Wczoraj było gorąco. Termometr wskazywał 35° C;*
- 4) **wyjaśnienie** — *Wróciłem do domu, ponieważ ...;*
- 5) **potwierdzenie faktu** — *Publiczność zauważyła różnicę. Już po pierwszych kadrach filmu wybuchł śmiech;*
- 6) **analogia** — *Przyrządź to tak samo jak poprzednio, ale dodaj wina;*
- 7) **podanie cechy wyróżniającej** — *Żyrafy wyróżniają się długą szyją;*
- 8) **wprowadzenie opisu cech bądź ich podklas** — *Łódź była duża, biała z biało-niebieskim żaglem;*
- 9) **określenie możliwego ciągu zdarzeń** — *Jeśli Piotr wróci wcześniej, to nam wszystko opowie;*
- 10) **alternatywa** — *Możemy pójść do X lub do Y;*
- 11) **skutek** — *Naciśnięcie przycisku spowodowało wybuch;*
- 12) **przypuszczenie** — *To było chyba w maju;*
- 13) **wniosek** — *Nie przyjechałaś, więc Krzyś się obraził;*
- 14) **wniosek końcowy** — *Tak więc możesz liczyć na sukces.*

Na podstawie wymienionych kategorii zdań wyróżniono następnie kilka typów wypowiedzi, których funkcją jest: porównanie, wprowadzenie do głównego tematu, podanie przykładu ogólnego lub szczegółowego, podkreślenie cechy, przeciwstawienie cech oraz sformułowanie wniosku końcowego.

Przyjęte schematy przebiegu dialogu przewidują: identyfikację obiektu (*co to jest statek?*), opis obiektu na podstawie części składowych, opis obiektu na podstawie jego cech, opis porównawczo — różnicujący. Generowane teksty są sekwencjami zdań reprezentujących określone kategorie, na przykład:

(Analogia / Opis atrybutów / Atrybut / Przemianowanie / Przykład szczegółowy / Potwierdzenie),*

(Podkreślenie / Analogia / Opis atrybutu / Przykład szczegółowy / Potwierdzenie).*

Statek jest to jednostka pływająca po powierzchni. Możliwości przewozowe opisywane są przez atrybuty bazy danych: *Wyporność* i *Załoga*. Inne atrybuty opisujące statek to: *maksymalna szybkość*, *napęd*, *paliwo*, *pojemność paliwa*, *rodzaj paliwa*. Na przykład: *DOWNES ma maksymalną prędkość 20, napęd — STMTURGD.*

Wybór schematu retorycznego do generowania odpowiedzi jest zdeterminowany przez typ pytania. Pytanie o definicję powoduje wygenerowanie odpowiedzi według schematu identyfikacji statku lub schematu opisu jego części składowych. Pytanie o różnice między obiektami odwołuje się do schematów porównawczo-różnicujących. Schematy mogą też być reprezentowane wariantowo,

o wyborze wariantu decyduje temat przetwarzanego aktualnie zdania pytającego, według którego system selekcjonuje listy obiektów przewidywanych jako temat kolejnego zdania, zapamiętując zdania zawierające tematy już wykorzystane (tworzony jest stos). Dzięki temu możliwe jest kontynuowanie tematu, rozpoczęcie nowego wprowadzonego w ostatnim zdaniu bądź powrót do tematów poprzednich [Mykowiecka A., 1987].

4.3.2. Mechanizmy wnioskowania

Generowanie przez system odpowiedzi na zinterpretowane składniowo i semantycznie zapytanie użytkownika musi być poprzedzone kolejną transformacją informacji polegającą na wyselekcjonowaniu z bazy faktograficznej systemu wyrażen stanowiących przesłanki w procesach wnioskowania.

Podstawowy mechanizm wnioskowania w systemach inteligentnych określony jest przez reguły dedukcji i implikacji logicznej. W pragmatycznych ujęciach podręczników sztucznej inteligencji teoretyczne aspekty prawomocnego wnioskowania dedukcyjnego ograniczają się najczęściej do stwierdzenia, że mają one długą tradycję wywodzącą się z czasów starożytnych. Równocześnie wymienia się najbardziej przydatne sposoby wnioskowania: wnioskowanie formalne, wnioskowanie przez uogólnienie, wnioskowanie przez analogię, wnioskowanie przez domyślanie się (default reasoning).

Z badań nad procesami przetwarzania informacji przez człowieka, a w tym nad organizacją informacji konceptualnej, przekonująco przedstawionych w teorii D. Sperbera i D. Wilson, wynika że o relewancji informacji decyduje nie tylko sposób strukturalizacji wiedzy statycznie odwzorowującej rzeczywistość lecz także, w stopniu co najmniej równym, reguły jej logicznej transformacji, której podstawę stanowi relacja wynikania (implikacji). Relacja ta jest przedmiotem badań współczesnej logiki formalnej.

Mając na uwadze znaczenie wyników tych badań teoretycznych dla rozwoju inteligentnych systemów faktograficznych, omówiliśmy je dość szczegółowo w rozdz. 2.3.2, gdyż problem ten traktowany jest w aplikacyjnych pracach z zakresu sztucznej inteligencji raczej fragmentarycznie. Ograniczenia w formalnym definiowaniu relacji wynikania logicznego pozwalają zrozumieć doraźny charakter rozwiązań przyjmowanych w konkretnych systemach.

4.3.3. Wnioskowanie w inteligentnych systemach faktograficznych

Semantyczna transformacja informacji, w wyniku której system dostarcza użytkownikowi relewantnej wiedzy, angażuje wykorzystanie wiedzy deklaratywnej i proceduralnej. Warunkiem wygenerowania nowej informacji jest wyselekcjonowanie z bazy systemu danych służących jako przesłanki we wnioskowaniu. Uważa się, że podstawowym elementem procesu wnioskowania jest rozpoznanie obiektów wskazanych w zapytaniu drogą ich porównania z odpowiednimi opisami w bazie danych (matching) [Bobrov D. G., Winograd T., 1977, 213-222]. Jest

to możliwe poprzez powiązanie wiedzy proceduralnej ze strukturami danych deklaracyjnych (procedural attachment), dzięki czemu zostaje wskazany zakres stosowalności danej reguły do określonego zbioru wyrażań, a tym samym wyznaczony zakres, w którym zachodzi relacja *entailment*. Przypisywanie strukturom danych dopuszczalnych reguł wnioskowania jest więc najefektywniejszym obecnie sposobem gwarantowania relewancji informacji w systemie.

Należy jednak zauważyć, że przyjmowane rozwiązania mają charakter do-
rażny. Nie korzysta się jeszcze z dorobku pragmatyki w językoznawstwie, a zwa-
szcza z teorii presupozycji na poziomie zdań i wyrazów jako najczęstszych
wykładników presupozycji [Austin J.L., 1962], [Grice H.P., *Philosophical Review*
78]. Na ten stan rzeczy wpływa przypuszczalnie także nieustabilizowanie tej
dziedziny badań i związane z nią kontrowersje dotyczące samego terminu presu-
pozycji [Bogusławski A., 1978, 7-28].

W literaturze przedmiotu stosunkowo obszernie traktowane jest zastoso-
wanie procedur dedukcji w systemach, w których językiem reprezentacji wiedzy
jest język predykatów pierwszego rzędu, charakterystyczny dla wczesnego etapu
rozwoju sztucznej inteligencji oraz dla tych dziedzin, w których przydatne jest
wnioskowanie formalne — na przykład w matematyce, automatycznym pro-
gramowaniu (por. rozdz. 3.4.2.1.). Przegląd tych procedur z punktu widzenia pro-
gramowania przedstawiła E. Madalińska-Bugaj, podkreślając, że udzielanie przez
system odpowiedzi polega na ocenie prawdziwości lub fałszywości formuł pre-
dykatowych na podstawie danych z bazy faktograficznej, pełniących funkcje
twierdzeń systemu. Autorka ta wyróżniła trzy grupy procedur dedukcyjnych
[Madalińska-Bugaj E., 1983, 85-102]:

- związane z tzw. podejściem bezpośrednim, przy którym relacje dwuargumen-
towe przechowywane są w słowniku jako pary (atrybut, wartość), natomiast
regułami wnioskowania są *modus ponens* i *podstawianie*. Generowanie odpo-
wiedzi na pytanie umożliwiają dwie funkcje *FINDR* i *VFIND*. Pierwsza służy
do wyszukania w bazie obiektów spełniających warunki określone w pytaniu.
Funkcja *VFIND* po użyciu *FINDR* i stwierdzeniu, że informacja nie jest za-
warta w bazie, powoduje wyszukanie niezbędnego aksjomatu i podjęcie próby
wydedukowania odpowiedzi;
- oparte na tzw. zasadzie rezolucji, przy których dowolne pytanie traktowane
jest jako twierdzenie do udowodnienia: *znaleźć takie x , że $P(x)$ jest prawdzi-
we*, gdzie $P(x)$ jest wyspecyfikowanym predykatem, który może być sformu-
łowany na wejściu w postaci $P(x)$. Na przykład, każde wyrażenie pytania jest
sprowadzane do tzw. postaci klauzuli (formuły atomowej) w standardowej
notacji logicznej, stosownie do zasady rezolucji. Zbiór klauzul tworzy listę,
wykorzystywaną w procesie dowodzenia według pewnych zasad. Zasada
pierwsza polega na zanegowaniu zdań, które mają być dowiedzione i próbie
skonstruowania dowodu. Zasada druga określa kolejność pobierania z pa-
mięci dodatkowych klauzul uzupełniających zanegowane wyrażenia wejś-

ciowe w procesie dowodzenia. Zasady doboru relewantnych klauzul są określone oddzielnie;

- oparte na zastosowaniu tzw. generatora funkcji w języku programowania LISP, przy którym informację wejściową stanowią deklaracje relacji i funkcji rachunku predykatów, formuły atomowe wyrażen bazy danych, pytania reprezentowane w postaci klauzul i traktowane jako przedmiot dowodzenia oraz reguły będące faktycznie wyrażeniami funkcyjnymi, do których system odwołuje się udzielając odpowiedzi (dedukcja wstecz).

Interesującego przykładu wnioskowania i transformacji wiedzy dostarcza opis opracowanego w połowie lat siedemdziesiątych systemu *POLITICS*, w którym wiedza o rzeczywistości reprezentowana jest za pomocą skryptów, a wyrażenia tego języka denotują przekonania ludzi o określonych wydarzeniach politycznych [Carbonell J. G. Jr., 1979, 27-51]. System ten konstruowano jako próbę symulacji procesu interpretacji zdarzeń i integracji różnych poziomów wiedzy, a więc różnych rodzajów relewancji (przedmiotowej i interpretacyjnej). Wnioskowanie w tym systemie oparte jest na sterowaniu zadaniami (*goal-directed*) przypisanymi poszczególnym skryptom.

Wiedza odwzorowywana w systemie uwzględnia trzy aspekty:

a) funkcję ideologii w interpretacji zdarzenia, jakkolwiek nie zakładano budowy precyzyjnych modeli ideologii,

b) reguły i strategie przyjmowane przez uczestników zdarzeń (państwa) w celu przeciwdziałania osiągnięciu celów przez stronę przeciwną,

c) łączenie w skryptach wiedzy o świecie zewnętrznym z zależnymi kontekstowo regułami wnioskowania.

Ideologie reprezentowane są w postaci drzew celów przypisanych uczestnikom zdarzeń (*actors*). Cele zorganizowane są więc za pomocą relacji paradygmatycznych. Drzewa celów, stworzone do odwzorowania celów Stanów Zjednoczonych, występują w dwu wersjach: konserwatywnej i liberalnej. Interpretacja zdarzenia w ramach każdej ideologii oparta jest na teorii zależności konceptualnej Schanka (por. rozdz. 3.4.2.4.). Konwersacja z systemem odbywa się w języku angielskim [Riesbeck Ch. K., 1978]. W przytaczanych przykładach transformacji wyrażen stosowana jest notacja protokołu konwersacji systemu:

— wyrażenie reprezentuje jedną z piętnastu klas obiektów, na przykład *PERSON, ORG(anization), COUNTRY, WEAPON*

* — wyrażenie reprezentuje prototyp w pamięci systemu,

§ — wyrażenie jest nazwą skryptu.

Transformacji informacji angażującej wnioskowanie dokonuje się dwukrotnie.

Po raz pierwszy na etapie przekładu tekstu wejściowego (niekoniecznie za pytania) na język reprezentacji wiedzy w bazie danych, na przykład:

***INTERPRET US-CONSERVATIVE**

INPUT TEXT:

The United States Congress voted to fund the Trident submarine project.

PARSING ... COMPLETED

((ACTOR (#ORG TYPE (*GOVT*) NAME (CONGRESS) PART (*US*))REF
(DEF))

< = > (*ATRAN*)

OBJ (*MONEY*) AMOUNT (SUFF)) TO(#ORG INVOLVING (#WEAPON
TYPE (*SUBMARINES*) NAME (TRIDENT))

REF (INDEF)

TIME ((AFTER NOW X))

MODE ((INTENTIONAL))).

Można to sparafrazować następująco: „Pewna agenda rządu Stanów Zjednoczonych nazywana Kongresem ma zamiar przekazać pieniądze (z nieokreślonego źródła) do innej agendy, która ma coś wspólnego z bronią należącą do kategorii łodzie podwodne o nazwie <<Trident>>. Odnotowano, że suma przekazanych pieniędzy jest wystarczająca (choć nie wyspecyfikowana).

Wnioskowanie związane z interpretacją zdania wejściowego dotyczy instytucji (agendy) otrzymującej pieniądze. System uruchamia w tym celu skrypt \$FUND, przy użyciu którego odpowiada na pytania: jaka agenda otrzyma pieniądze, na co będą wykorzystane, skąd pochodzą, jaki związek istnieje między łodziami Trident i agencją uzyskującą pieniądze, jakie implikacje ma to działanie dla celów USA wyższego rządu (jeśli ma). Uruchomienie skryptu \$FUNDING powoduje wywnioskowanie kontekstu, w którym odbywa się dalsza transformacja:

INSTANTIATING SCRIPT : \$FUNDING

INFERENCE: (ARMEDFORCES PART (*US*) RANK (COMMAND) TO
RECEIVE (*MONEY* AMOUNT (SUFF))

INFERENCE: (ARMEDFORCES PART (*US* RANK (COMMAND)) TO
BUILD (WEAPON TYPE (*SUBMARINES*) NAME(TRIDENT

Uruchomienie skryptu \$FUND jest wynikiem porównania (*matching*) reprezentacji tekstu wejściowego wygenerowanej przez parser (moduł analizy języka naturalnego) z wiedzą systemu i stwierdzenia występowania wspólnej cechy relewantnej „przekazanie pieniędzy przez agencję rządową najwyższego szczebla” (*ATRANS MONEY*). Przekazanie pieniędzy jest jedną z ról semantycznych wyróżnionych jako klatki (slots) w skrypcie, które z kolei powodują realizację określonych procesów wnioskowania. (Rolami semantycznymi w skrypcie \$FUND są ponadto: *FUNDEE*, *FUNDER*, *MONEY-GIVEN*, *FUND-CON* — przedmiot finansowania *MONEY-NEEDED*.) Podajemy przykłady reguł wnioskowania specyfikującego w systemie *POLITICS* związanego z interpretacją tekstu wejściowego przy użyciu skryptu \$FUND:

RULE #F3

IF X delegates (AUTHORITY, RESPONSABILITY, or MONEY) to Y

THEN Y is N levels below X in the Governmental structure

REFINEMENT:

IF <that which is delegated> is of high (MILITARY, ECONOMIC or JUDICIAL) significance

THEN probably N = 1.

RULE #F5

*IF the GOAL of the Funding is of a MILITARY nature
THEN expect a MILITARY agency to receive the AUTHORITY to spend the money (i.e. CONTROL of the money)*

(APPLYING RULE #F3 FROM \$FUND)

INFERENCE:

*(GOV-ORG PART (*US*) RANG (LEV VAL (1))) TO RECEIVE (*MONEY* AMOUNT (SUFF))*

SPECIFIED:

*&FUNDEE = (#GOV-ORG PART(*US*) RANK (LEV VAL(1)))*

(APPLYING RULE #F5 FROM \$FUND)

INFERENCE:

*(#ARMEDFORCES PART (*US*) RANK (COMMAND TO RECEIVE (*MONEY* AMOUNT (SUFF)))*

SPECIFIED:

*&FUNDEE = (# ARMEDFORCES PART (*US*) RANK (COMMAND))*

RULE #F6

*IF &FUND-CON involves promoting X (e.g. helping bring X into existence)
AND &FUNDER does not already have X*

AND &MONEY-GIVEN is grater than or =&MONEY-NEEDED

THEN expect &FOUNDEE to do the necessary actions to bring X about.

(APPLYING RULE #F6 FROM \$FUND)

SPECIFIED:

*&FUND-CON= ((CON ((ACTOR (*MONEY* AMOUNT (SUFF) IS (CONTROL VAL (#ARMEDFORCES PART (*US*) RANK (COMMAND))))))*

ENABLE

*(ACTOR (ARMEDFORCES PART (*US*) RANK (COMMAND)) < = > (*DO*))*

RESULT

*((ACTOR (#WEAPON TYPE (*SUBMARINES*) NAME (TRIDENT)) IS (CONTROL VAL- (#ARMEDFORCES PART (*USA*) RANK (COMMAND))))))*

RULE #F8

IF X (X same as in RULE #6) has a physical realization

AND &FOUNDEE controls the creation of X

THEN&FUND-CON = (&FUNDEE will build X)

(APPLYING RULE FROM \$FUND)

INFERENCE:

*(#ARMEDFORCES PART (*US*) RANK (COMAND)) TO BUILD (WEAPON TYPE (*SUBMARINES*) NAME(TRIDENT))*

SPECIFIED:

*&FUND-CON ((ACTOR (#ARMEDFORCES PART (*US*) RANK (COMMAND))*

< = > (\$CONSTRUCT ENABLEMENT: (*MONEY* AMOUNT)) POBJ:
(#WEAPON TYPE (*SUBMARINES*) NAME(TRIDENT))
MODE ((INTENTIONAL)).

Po raz drugi system realizuje wnioskowanie, określając cele, jakie usiłuje osiągnąć działający (*goal-directed inference* — wnioskowanie sterowane celami). Zdarzenia wyrażone w tekście wejściowym explicite lub wywnioskowane traktowane są jako instrumenty w osiągnięciu celów założonych jako cechy relewantne ideologii. Reguły wnioskowania w wersji ideologii konserwatywnej USA, której celem nadrzędnym jest zdobycie potęgi militarnej scharakteryzowano następująco [Carbonell J. G., Jr., 1979, 37]:

RULE #4

IF a new weapon system is built by some agency of the armed forces of country X

THEN the military power of X will increase

REFINEMENT: INVOKE (RULE #4.1 RULE 4.2)

RULE #4.1

IF increase in military power is a high level goal of X

THEN expect to do any scripts enabled by the increase in power which do not violate other goals of X

RULE #4.2

IF increase in military power is a low goal of X

THEN expect to be working on some goal for which military strength is a precondition

RULE #5

IF Country X Achieves increase in military power

THEN look for country such that either is open hostilely to X, or Y poses a treat to X

(APPLYING RULE #4 FROM GOAL DETERMINATION)

INFERENCE:

**US* INCREASE IN MILITARY POWER*

ACTIVATE: \$CONFRONTATION

(APPLYING RULE #5 FROM GOAL DETERMINATION)

*(PENDING DEMON (SEARCH THREAT *US*¹¹).*

We wnioskowaniu w wersji ideologii liberalnej USA przy zastosowaniu reguły #4 powstają problemy, ponieważ zdobycie potęgi militarnej nie występuje w skryptach aktywowanych do interpretacji zdarzeń, co jest niezbędne do wyjaśnienia, dlaczego USA potrzebują więcej broni. System uruchamia wówczas skrypt wyścig zbrojeń (\$ARMSRACE).

W konsekwencji zastosowania reguły #5 automatycznie wywołującej skrypt *Konfrontacja* system dochodzi do następujących wniosków, reprezentowanych w języku bazy wiedzy w kategoriach relewantnych cech celów:

¹¹ „Demon” jest nazwą procedury uruchamianej automatycznie [Madalińska-Bugaj E., 1983, 85-102].

INFERENCE:

*GOAL OF *US* IS (#COUNTRY TYPE (*COMMUNIST*) SIZE (LARGE) TO FEAR *US**

INSTANTIATING SCRIPT: *\$CONFRONTATION*

INFERENCE:

*(#COUNTRY TYPE (*COMMUNIST*) SIZE (LARGE) TO STOP EXPANDING POLITICAL CONTROL*

INTERPRETATION COMPLETED, READY TO ACCEPT QUESTIONS.

W module komunikacji z użytkownikiem zastosowano zasady analizy tekstu systemu SAM [Riesbeck Ch. K., 1978]. Na podstawie scharakteryzowanych transformacji tekstu wejściowego system udziela odpowiedzi na następujące pytania dotyczące budowy łodzi Trident:

Wersja interpretacji konserwatywnej

Pyt. 1: *What did the U.S. Congress do this for?*

Odp.: *THE UNITED STATES CONGRESS WANTS THE UNITED STATES ARMED FORCES TO BE STRONGER*

Pyt. 2: *Why should the U.S. be stronger?*

Odp.: *THE UNITED STATES ARMED FORCES SHOULD BE STRONGER TO STOP COMMUNIST EXPANSION*

Pyt. 3: *What should the U.S. do if Russia builds nuclear submarines?*

Odp.: *THE UNITED STATES SHOULD BUILD MORE SUBMARINES*

Wersja interpretacji liberalnej

Pyt. 1: *What did the U.S. Congress do this for?*

Odp.: *THE UNITED STATES CONGRESS FEARED FALLING BEHIND IN THE ARMS RACE*

Pyt. 2: *What will Russia do next?*

Odp.: *RUSSIA WILL PROBABLY BUILD MORE WEAPONS*

Pyt. 3: *What should the U.S. do if Russia builds submarines?*

Odp.: *THE UNITED STATES SHOULD NEGOTIATE WITH RUSSIA TO STOP THE ARMS RACE.*

Różnica w odpowiedziach na trzecie pytanie wynika z użycia we wnioskowaniu jako wykładników cech relewantnych skryptu *\$CONFRONTATION* w wersji konserwatywnej oraz skryptu *\$ARMSRACE*, charakteryzujących cele polityki USA. Możliwość wyjaśnienia przyczyn angażowania się w wyścig zbrojeń uwzględnione są w skrypcie *\$ARMSRACE* — uniknięcie konsekwencji rezygnacji z wyścigu zbrojeń, które są jednym z celów szczegółowych związanych z pokojem na świecie.

Wygenerowanie odpowiedzi na pytanie trzecie nie jest możliwe drogą transformacji cech relewantnych celów odwzorowanych w skryptach. Do wnioskowania zostaje włączony odrębny moduł wnioskowania operacyjnego niezbędnego do określenia konfliktu celów uczestników działań oraz do wskazania działań w danej sytuacji niezbędnych. Zagadnienie planowania działań niezbędnych do rozumienia tekstów podejmowali Schank i Abelson, Schmidt i Sridharan [Schank R. C., Abelson R. P., 1975] [Schmidt C., Shridharan N., 1977]. W systemie *POLI-*

TICS we wnioskowaniu istotne jest planowanie działań przeciwnych (counterplanning). Polega ono na doborze zbioru strategii określających przyczyny, miejsce i czas konfliktu celów (lub planów) oraz sposób osiągnięcia celów jednej strony mimo zaistnienia konfliktu. Scharakteryzowano strategie działań dla uczestników w wersji prewencyjnej oraz w wersji aktywnej:

Ogólna strategia dla uczestnika A działającego prewencyjnie

1. Ustal jaki plan X realizuje A , aby osiągnąć cel $G(X)$.
2. Gdy zawodzi (1) zdecyduj, jakiego planu użyłby A , gdyby był w sytuacji X -a. Załóż, że X postępuje według tego planu.
3. Ustal najslabsze powiązanie w planie, najłatwiejsze do zablokowania.
4. Zastosuj jedną z reguł kontrplanowania (nizej wymienione) do zablokowania planu w najslabszym punkcie.
5. Oczekuj, że X zmodyfikuje swój plan w celu ominięcia zablokowanego powiązania lub ponownie go sformułuje.
6. Zablokuj plan nowy lub zmodyfikowany.
7. Jeśli $G(X)$ jest planem o najwyższym priorytecie, oczekuj, że X może realizować równocześnie kilka planów do osiągnięcia $G(X)$. Zablokuj wszystkie plany.
8. Jeśli w którymkolwiek punkcie koszt blokowania $G(X)$ przekracza koszty związane z X i osiaganiem $G(X)$, zrezygnuj z kontrplanowania.

Ogólna strategia dla uczestnika A osiagania celu mimo przeciwności

1. Zastosuj normalną strategię planowania, aby znaleźć plan osiągnięcia $G(X)$.
2. Jeśli X blokuje plan, wyszukaj jakikolwiek plan alternatywny lub:
 - a) omiń zablokowane ogniwo (modyfikując plan A),
 - b) zablokuj niektóre plany X -a i staraj się spowodować wzajemne blokowanie waszych planów,
 - c) czasowo zrezygnuj z $G(A)$ do momentu, gdy X nie będzie blokował istotnych ogniów.
3. Jeśli $G(A)$ jest dużej wagi, realizuj równocześnie różne plany osiągnięcia $G(A)$.
4. Jeśli $G(A)$ jest celem szczegółowym wśród celów wyższych $G'(A)$, znajdź dla $G'(A)$ nowy plan, który nie wymaga osiągnięcia $G(A)$.
5. Utrzymuj plan w tajemnicy przed X i przed każdym, kto może komunikować się z X .
6. Jeśli X potrzebuje osiągnąć pewien $G(X)$ w celu zablokowania ogniwa w planie dla $G(A)$, równocześnie blokuj X dla $G(X)$ i realizuj $G(A)$.

Założono także konieczność wnioskowania w sytuacji, kiedy strategie są stosowane przez więcej niż dwu uczestników. Dobór strategii, i w konsekwencji wskazanie reguł wnioskowania na etapie interpretacji zdarzenia przedstawionego w tekście wejściowym oraz na etapie udzielania odpowiedzi przez system, jest określany przez reguły bardziej ogólne niż te, które są przypisane pewnym cechom relewantnym zdarzeń w skryptach. Chodzi tu o odwzorowanie pewnych cech wspólnych różnym ideologiom w polityce międzynarodowej, na przykład

w sytuacji, gdy do przejęcia kontroli politycznej nad jakimś państwem dąży równocześnie kilka stron. Przewidziano tu następujący zbiór reguł:

RULE#C1:

To stop actor X from accomplishing his goal G(X) see if there is any goal G(A) which is mutually exclusive with G(X). If so, give high priority to accomplishing G(A)

RULE#C2:

If A wishes to accomplish G(A) and G(A) is being blocked by X working on a mutually exclusive G(X) then try to thwart G(X) by some other means

RULE#C3:

To prevent from achieving G(X) threaten a higher level goal of X, say G'(X). Thus expect to divert its efforts to G'(X)

RULE#C4:

To prevent X from achieving G(X) prevent X from achieving a necessary subgoal of G(X)

RULE#C5:

To achieve G(A) if X is blocking a subgoal G(A) try to find an alternative plan (or script) for G(A)

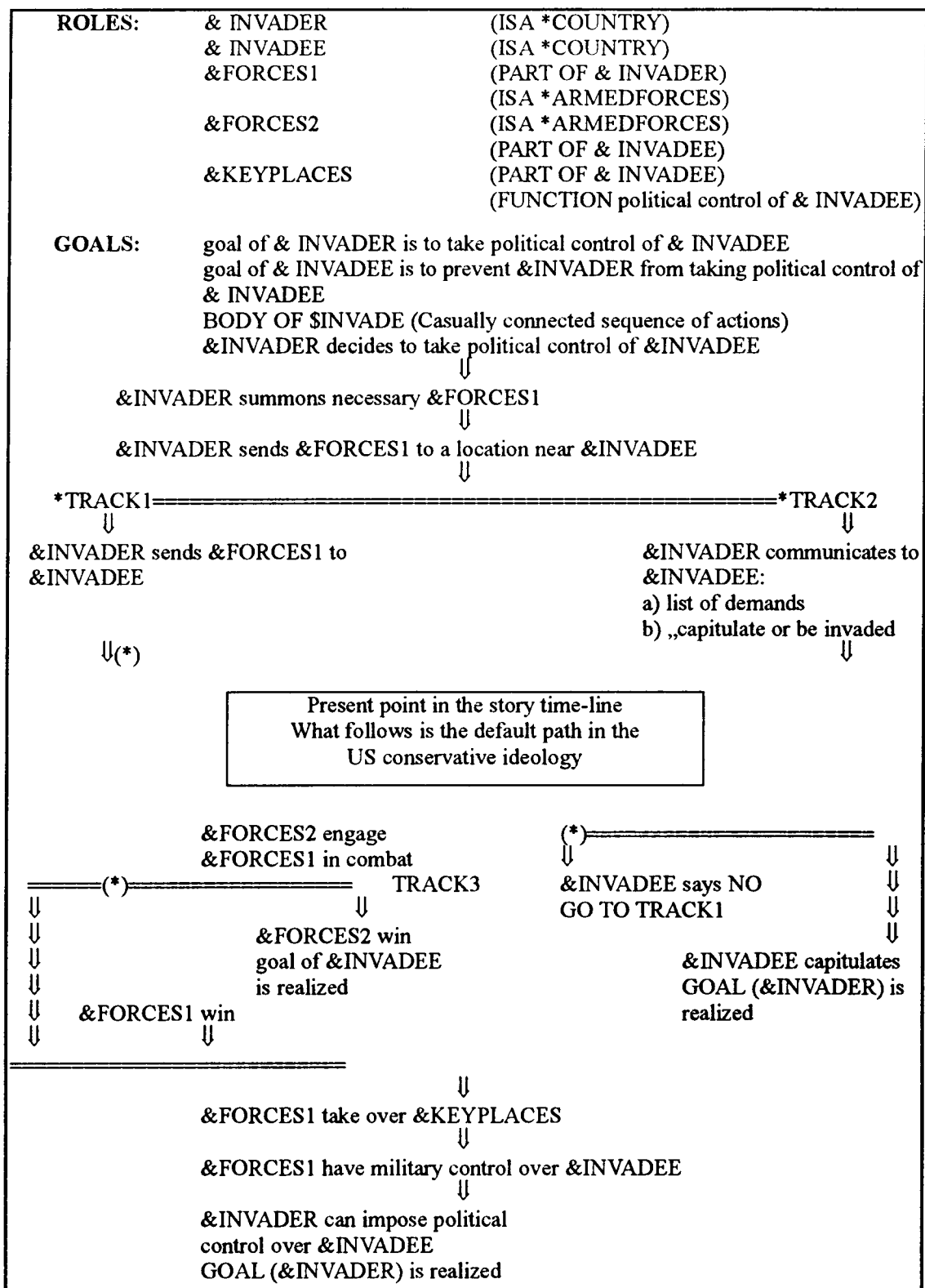
RULE#C6:

If G(A) and G(X) are in conflict and there exists G'(A) = G'(X) where the G' goal is higher level than A should try mediating with X to mutually achieve G' and abandon their respective G goals. (Metaphorically speaking: Why fight over a molehill when together we can conquer a mountain).

Przytoczone reguły są *de facto* regułami heurystyki, w wyniku zastosowania których wyszukiwane są reguły szczegółowe wnioskowania przypisane poszczególnym skryptom. Ich użycie w systemie wymaga odwołania się do wcześniej zgromadzonej wiedzy — tu reprezentowanej jako sekwencje działań i celów w odpowiednich wariantach skryptów. Warianty takie opracowano dla skryptu inwazji *SINVADE* interpretującego przebieg wydarzeń w Czechosłowacji w 1968 r. według możliwych wersji sekwencji działań, tzw. ścieżek (*tracks*), a także przewidującego przebiegu wydarzeń (schemat 13).

Wybór jednego z wariantów skryptu jest procesem skomplikowanym, z decyzyjnego punktu widzenia wymaga bowiem:

- porównania reprezentacji tekstu wejściowego z reprezentacją ról i celów w skrypcie,
- uruchomienia skryptu,
- w przypadku rozbieżności reprezentacji wiedzy w tekście wejściowym i skrypcie — zastąpienia cech relewantnych zdarzeń innymi, ekwiwalentnymi, na zasadzie przyporządkowania tym samym celom,
- wskazania domyślnej ścieżki w skrypcie, której cele są najbardziej zbieżne z celami określonych uczestników działań,



Schemat 13. Domyślne sekwencje działań skryptu INVADE w wersji konserwatywnej

INSTANTIATED SCRIPT ROLES

&INVADEE← *RUSSIA*
&INVADEE← *CZECHOSLOVAKIA*
&FORCES1← (*TROOPS* PART OF (*RUSSIA*))
&FORCES2 uninstantiated
&KEYPLACES uninstantiated

INSTANTIATED GOALS:

GOAL OF RUSSIA: <Take political control of Czechoslovakia>
SUBGOAL: <Take military control of key places in Czechoslovakia>

GOAL OF CZECHOSLOVAKIA <Preventing Russia from taking political control>
SUBGOAL: <Preventing from taking military control of key places

INSTANTIATED PATH IN SCRIPT

RUSSIA decides to take political control of *CZECHOSLOVAKIA*
↓
RUSSIA summons necessary *TROOPS*
↓
RUSSIA sends TROOPS to the border between *RUSSIA* and *CZECHOSLOVAKIA*
↓
RUSSIA sends TROOPS into *CZECHOSLOVAKIA*
↓
Russian TROOPS defeat Czech TROOPS in combat
↓
Russian TROOPS take over Czech &KEYPLACES
↓
Russian TROOPS HAVE MILITARY CONTROL OVER *CZECHOSLOVAKIA*
↓
RUSSIA imposes political control over *CZECHOSLOVAKIA*, achieving its goal

Schemat 14. Realizacja skryptu INVADE według ścieżki (1)

- aktualizacji wiedzy w bazie danych, a więc modyfikacji modelu rzeczywistości w systemie na podstawie nowej wiedzy z tekstu wejściowego.

Interpretując zdanie *Russia massed troops on the Chech border* (*Rosja skoncentrowała wojska na granicy Czechosłowacji*) system porównuje cechy odwzorowane w skrypcie *INVADE* i odnajduje jedyną wspólną cechę *koncentrację sił*, która może być przypisana różnym celom szczegółowym działającego. Niezbędne jest tu włączenie reguł pragmatyki skryptu (tzw. trigger — zapadka — por. rozdz. 3.4.2.5.), które powodują wygenerowanie dwu wniosków: *Rosja zamierza dokonać inwazji na Czechosłowację lub Rosja chce pomóc militarnie Czechosłowacji*. Dalsze działanie tego mechanizmu wnioskowania prowadzi z jednej

strony do stwierdzenia, że Czechosłowacja nie jest zaangażowana w żaden konflikt militarny i nie potrzebuje pomocy, z drugiej zaś, że uzyskanie kontroli nad małymi narodami jest możliwe w wyniku inwazji. Takie wnioskowanie sterowane celami ilustruje poniższy przykład [Carbonell J. G., Jr., 1979, 47]:

Let X be a small country, Y be a large country. (Niech X będzie małym krajem, Y - dużym)

RULE #13

(TRIGGER: Y moves troops near or inside of X) (Y przemieszcza wojska w pobliże lub do X)

IF X is involved in a military conflict and needs assistance (X jest wciągnięty w konflikt zbrojny i potrzebuje pomocy)

AND having X win the conflict helps Y to achieve one of its goals (A pokonanie, X pomaga Y osiągnąć jeden ze swoich celów)

THEN Y will have the goal of military aiding X (ZATEM celem Y będzie udzielenie pomocy militarnej X)

ELSE Y expanded its political control as a higher level goal

THEN Y has the goal of military take over of X.

Przytoczony przykład wydaje się interesujący jako ilustracja spełnienia przez zdania minimalnego warunku relewancji (zawierania w poprzedniku i następniku formuły implikacyjnej wspólnej zmiennej), który jest niezbędny — lecz nie jest wystarczający — do uzyskania relewantnego wniosku. Umożliwia jego uzyskanie klauzula *ELSE*. Autorzy systemu zauważają, że przy tym samym mechanizmie wnioskowania interpretacja interwencji amerykańskiej w Wietnamie w wersji ideologii konserwatywnej byłaby zdefiniowana w systemie jako pomoc, ponieważ ówczesny Wietnam Południowy był zaangażowany w konflikt zbrojny z Wietnamem Północnym, a ponadto byłoby to zbieżne z celami nadrzędnymi USA — powstrzymaniem ekspansji komunizmu.

Dialog z użytkownikiem po zinterpretowaniu przytoczonego zdania ma następujący przebieg:

Wersja ideologii konserwatywnej

INPUT TEXT: Russia massed troops on the Czech border. (Rosja skoncentrowała wojska na granicy Czechosłowacji) INTERPRETATION COMPLETED, READY TO ACCEPT QUESTIONS

Pyt.: *Why did Russia do this ? (Dlaczego Rosja tak postąpiła?)*

Odp.: *BECAUSE RUSSIA THOUGHT THAT IT COULD TAKE POLITICAL CONTROL OF CZECHOSLOVAKIA BY SENDING TROOPS. (PONIEWAŻ ROSJA SĄDZI, ŻE MOŻE PRZEJĄĆ KONTROLĘ NAD CZECHOSŁOWACJĄ POPRZEZ WYSŁANIE WOJSK)*

Pyt.: *What will Russia do next? (Co następnie zrobi Rosja?)*

Odp.: *RUSSIA MAY WAIT FOR CZECHOSLOVAKIA TO GIVE IN TO RUSSIAN DEMANDS (ROSJA MOŻE OCZEKIWAĆ, ŻE CZECHOSŁOWACJA PODPORZĄDKUJE SIĘ JEJ ROZKAZOM)*

Pyt.: *What happens if Czechoslovakia does not give in? (Co stanie się, jeśli Czechosłowacja nie podda się?)*

Odp.: *RUSSIA MAY DECIDE TO START THE INVASION (ROSJA MOŻE ROZPOCZĄĆ INWAZJĘ)*

Pyt.: *What should the United States do? (Co powinny zrobić Stany Zjednoczone?)*

Odp.: *THE UNITED STATES SHOULD INTERVENE MILITARY (STANY ZJEDNOCZONE POWINNY ZBROJNIE INTERWENIOWAĆ)*

Pyt.: *What else can the United States do? (Co jeszcze mogą zrobić?)*

Odp.: *THE UNITED STATES CAN DO NOTHING, OR IT CAN INTERVENE DIPLOMATICALLY BY CONFERRING WITH RUSSIA ABOUT CZECHOSLOVAKIA, THESE ALTERNATIVES ARE BAD FOR THE UNITED STATES. (STANY ZJEDNOCZONE MOGĄ NIC NIE ROBIĆ LUB INTERWENIOWAĆ DYPLOMATYCZNIE PERTRAKTUJĄC W SPRAWIE CZECHOSŁOWACJI. ALTERNATYWA TA JEST NIEKORZYSTNA DLA STANÓW ZJEDNOCZONYCH.)*

Wersja ideologii liberalnej (różnice w odpowiedziach)

Pyt.: *What should the United States do? (Co powinny zrobić Stany Zjednoczone?)*

Odp.: *THE US SHOULD DENOUNCE THE RUSSIAN ACTION TO THE UNITED NATIONS. (USA POWINNY ZDEMASKOWAĆ DZIAŁANIA ROSJI PRZED ONZ)*

Nowy tekst:

The US denounces the Russian treat to the UN. Russia rejects U.S criticism. The UN resolves to condemn the Russian action. Russian troops move into Czechoslovakia. (USA demaskują zagrożenie rosyjskie przed ONZ. Rosja odrzuca oskarżenie amerykańskie. ONZ postanawia potępić działania Rosji. Wojska rosyjskie wkraczają do Czechosłowacji)

Pyt.: *What should the US do now ? (Co teraz powinny uczynić USA?)*

Odp. *THE US SHOULD APPEAL TO RUSSIA TO STOP THEIR INVASION (USA POWINNY ZAAPELOWAĆ DO ROSJI O POWSTRZYMANIE INWAZJI)*

W sytuacji, gdy działanie środkami dyplomatycznymi w układach międzynarodowych okazuje się zawodne, ponieważ nie są skuteczne w oddziaływaniu na władze na szczeblu lokalnym (USA), możliwości kontynuacji dialogu kończą się. System nie dysponuje bowiem wiedzą na temat stosunków międzynarodowych oraz sposobów realizacji celów środkami politycznymi, odwzorowanych w odpowiednich regulacjach wnioskowania.

Interesującą cechą systemu *POLITICS* jest sposób aktualizacji bazy wiedzy o świecie zewnętrznym w wyniku interpretacji tekstów wejściowych zawierających wyrażenia, których „nie rozumie”. Próby interpretacji wyrażen nieznanych

ilustrują więc proces uczenia się systemu, zgodnie z postulatami teorii pamięci dynamicznej. Przytaczamy przykład podany przez autorów [Carbonell J. G., Jr., 1979]:

TEKST WEJŚCIOWY:

Russia sends massive arms shipments to the MPLA in Angola. (Rosja wysyła duże transporty wojska dla MPLA w Angoli)

PARSING ... (UNKNOWN WORD: MPLA) (... NIEZNANE SŁOWO: MPLA)

(SYNTACTIC EXPECTATION: NOUN) (OCZEKIWANY SYNTAKTYCZNIE: RZECZOWNIK)

(SEMANTIC EXPECTATION): (FRAME: (ATRANS PTRANS)

SLOT: RECIP REQ: (LOC ACTOR)))

Po przeprowadzeniu analizy syntaktycznej system próbuje zidentyfikować nieznane wyrażenie *MPLA*, przekazując komunikat dotyczący aspektów językowych i interpretacji semantycznej opartej na wykorzystaniu we wnioskowaniu ram i klatek (por. rozdz. 3.4.2.5.). Prawdopodobne znaczenie rzeczownika *MPLA*: klatka *otrzymujący* (recipient) przy zmianie lokalizacji przedmiotu i zmianie relacji posiadania, przedmiot fizyczny lub lokalizacja:

RUSSIA ⇔ PTRANS ← *ARMS ←	(MPLA)
	RECIP1
	RUSSIA

ARMS ⇔ SIZE val (> NORM)

Na podstawie wiedzy o tym, że Angola jest krajem w stanie wojny, potrzebuje zatem pomocy militarnej, oraz o tym, że Rosja jest krajem komunistycznym, system wnioskuje w dwu kierunkach: MPLA chce przejąć kontrolę nad Angolą oraz Rosja wykorzystuje MPLA do przejęcia kontroli nad Angolą:

INFERENCE: *MPLA* MAY BE A POLITICAL FACTION OF *ANGOLA*

INFERENCE: *RUSSIA* ATRANS *ARMS* TO *MPLA*

INFERENCE: *MPLA* IS PROBABLY *COMMUNIST*

INFERENCE: GOAL OF *MPLA* IS TO TAKE OVER *ANGOLA*

INSTANTIATING SCRIPT: SAIDNMF

INFERENCE:

RUSSIA MAY WANT TO CONTROL *ANGOLA* THROUGH *MPLA*

INTERPRETATION COMPLETED, READY TO ACCEPT QUESTION

Dialog w wersji konserwatywnej

Pyt.: *Why did Russia do this ? (Dlaczego Rosja tak postąpiła?)*

QUESTION TYPE IS: **WHY** (TYP PYTANIA: **DLACZEGO**)

SEARCHING INFERENCE

Odp.: *RUSSIA WANTS TO CONTROL ANGOLA THROUGH THE MPLA*

(ROSJA CHCE KONTROLOWAĆ ANGOŁĘ POPRZEZ MPLA)

Pyt.: What will the MPLA use arms shipments for? (Do czego wykorzystała MPLA wojska?)

QUESTION TYPE IS WFR (what for) (TYP PYTANIA: DO CZEGO)

SEARCHING GOLDWATER IDEOLOGY

SEARCHING SCRIPT: \$AIDMF

Odp.: THE MPLA WANTS TO TAKE OVER ANGOLA USING THE WEAPONS. (MPLA CHCE PRZEJĄĆ WŁADZĘ NAD ANGOLA)

Pyt.: What will other factions in Angola do? (Co zrobią inne ugrupowania w Angoli?)

Odp.: THE OTHER FACTIONS MAY ASK SOME OTHER COUNTRY FOR ARMS. (INNE UGRUPOWANIA MOGĄ PROSIĆ INNY KRAJ O WOJSKO)

Pyt.: Which country? (Który kraj?)

*ATTEMPTING FURTHER SPECIFICATION OF *COUNTRY* IN CURRENT CONTEXT (PRÓBA SPECYFIKACJI *KRAJU* W BIEŻĄCYM KONTEKŚCIE)*

Odp.: PROBABLY THE UNITED STATES> (PRAWDOPODOBNIESTANY ZJEDNOCZONE)

Pyt.: Should the United States do anything about it ? (Czy Stany Zjednoczone mają nic nie robić?)

Odp.: YES, THE UNITED STATES SHOULD SEND ARMS TO THE OTHER FACTIONS IN ANGOLA. (O TAK, STANY ZJEDNOCZONE POWINNY WYSLAĆ WOJSKA DLA INNYCH UGRUPOWAŃ W ANGOLI)

Transformacja informacji w bazie danych realizowana jest przy użyciu analogicznych reguł heurystyki, jak w przykładzie „Czechosłowacja”, oraz wiedzy zawartej w skryptach dotyczących konfliktów zbrojnych i roli partii (stron) politycznych w konflikcie.

Dokładne przedstawienie zasad wnioskowania i semantycznej transformacji informacji w systemie POLITICS wydawało się interesujące z kilku powodów:

a) zastosowano tu zasadę relewancji interpretacyjnej związanej z relewancją motywacyjną (celami); relewancji interpretacyjnej odpowiadają „konserwatywna” i „liberalna” wersja opisu zdarzeń w procesie jej przyswajania przez system i udostępniania w procesie dialogu;

b) jest to przykład reprezentacji wiedzy łączącej statyczne i dynamiczne aspekty odwzorowania rzeczywistości (deklaratywne i proceduralne) za pomocą skryptów (rozwiniętej w stosunku do pierwszej wersji Minsky’ego, Schanka i Abelsona [Minsky M. A., 1975]) zastosowanej w systemie SAM (Story Acquiring Module), interpretującym teksty doniesień gazetowych o wypadkach ulicznych i wizytach dyplomatycznych;

c) użyte w systemie reguły wnioskowania sterowanego celami i ograniczającego kontekstem stanowią ilustrację wnioskowania gwarantującego spełnienie minimalnego warunku relewancji logicznej, a tym samym prawdziwością transformację informacji zawartej w bazie wiedzy systemu; wystąpienie przynajmniej jednej wspólnej cechy relewantnej w poprzedniku i następniku formuły implikacyjnej tekstu wejściowego jest podstawą aktywizacji określonych reguł wnioskowania, zaś zmiana tej cechy może wpływać na zmianę kierunku wnioskowania i interpretację semantyczną wyrażenia;

d) system opracowano przed kilkunastoma laty, na jego wiedzę (traktowaną asertorycznie) składają się cechy relewantne zdarzeń politycznych, odwzorowywane w kategoriach celów poszczególnych ideologii. Zmiana układów rzeczywistości powoduje degradację wiedzy pozajęzykowej. Obecnie relewancja informacji o rzeczywistości i jej przydatność do interpretacji i predykcji zdarzeń politycznych wydaje się mocno problematyczna, nawet przy zakładanej konieczności uproszczeń i naiwności generowanych wniosków. Nie uległa tej degradacji wiedza o języku odwzorowana we względnie niezależnym module komunikacji z bazą wiedzy.

PODSUMOWANIE

Rozważania przeprowadzone na podstawie literatury przedmiotu oraz wyniki analizy systemów informacyjno-wyszukiwawczych z punktu widzenia wpływu instrumentów informacyjnych na zakres i sposób odwzorowania wiedzy w systemie pozwalają sformułować następujące konkluzje:

1. Podstawową zasadą określającą sposób odwzorowania wiedzy w systemie jest zasada relewancji, interpretowana różnie w informacji naukowej oraz w dyscyplinach, których przedmiotem są procesy poznania i reprezentacji wiedzy: filozofii, logice, językoznawstwie, teorii komunikacji oraz ostatnio w sztucznej inteligencji. Analiza propozycji uniwersalnych, przedmiotowych definicji relewancji wykazuje ich nieadekwatność i nieostrość. Porównanie rozważań w tym zakresie w filozofii (A. Schutza), inspirowanych poglądami sceptyka Karneadesa, w logice oraz w teorii komunikacji Sperbera i Wilson pozwala stwierdzić zbieżność wniosków o istnieniu różnych sfer interpretacji znaczenia tego terminu: *sfery przedmiotowej, interpretacyjnej i operacyjnej*. Skłania to do rozumienia relewancji jako intensjonalnych własności obiektów odwzorowanych w języku za pomocą sieci relacji semantycznych (paradygmatycznych i syntagmatycznych).

2. Repertuar relewantnych cech informacji w systemie jest zdeterminowany przez trzy czynniki:

- lokalizację zbioru wiedzy o rzeczywistości zewnątrz systemu (dokumentacyjnego) lub wewnątrz (faktograficznego),
- instrument informacyjny umożliwiający interakcyjny lub nie-interakcyjny tryb komunikowania się z użytkownikiem;
- dokonywanie (lub nie) semantycznej transformacji informacji w celu jej udostępnienia w systemach dedukcyjnych (inteligentnych) i nie-dedukcyjnych.

Czynniki te wyznaczają sposób reprezentacji wiedzy o rzeczywistości: pośredni za pomocą języka informacyjno-wyszukiwawczego wyspecjalizowanego w funkcji metainformacyjnej w systemie dokumentacyjnym oraz bezpośredni za pomocą języka reprezentacji wiedzy w systemie faktograficznym. *Uniwersalnymi wykładnikami cech relewantnych obiektów pozajęzykowych* we wszystkich typach systemów są relacje mereologiczne i część relacji skojarzeniowych, odwzorowywanych w słowniku i/lub tekstach języka systemu. *Wykładnikami relewantnych cech obiektów językowych* są relacje hierarchii zakresowej wyrażen oraz część relacji skojarzeniowych, ustalanych na podstawie możliwości współwystępowania tych wyrażen w tekście.

3. Instrument informacyjny systemu (komputer z oprogramowaniem) wyznaczający tryb udostępniania wiedzy relewantnej oraz możliwość jej wygenerowania w procesie wnioskowania wpływa na sposób odwzorowania wiedzy o języku i o rzeczywistości:

- w systemach tradycyjnych, w których ocena relewancji wyselekcjonowanej informacji odbywa się na podstawie końcowych wyników wyszukiwania, cała wiedza o rzeczywistości jest wbudowana w język, co skłania do weryfikacji rozpowszechnionego podziału języków informacyjnych na „uniwersalne” (klasyfikacje biblioteczne) i „specjalistyczne”;
- w systemach dialogowych (online), w których ocena relewancji odbywa się w trakcie procesu wyszukiwania, przy możliwości zmiany kryteriów relewancji, zwiększa się zakres wiedzy o języku, reprezentowanej głównie za pomocą operatorów tekstowych;
- w dialogowych systemach inteligentnych, w których ocena relewancji dokonywana jest przez system na podstawie faktograficznej bazy danych oraz własnej interpretacji zapytań, realizowanej w procesie wnioskowania, następuje rozdzielenie reprezentacji wiedzy o świecie zewnętrznym (bazy danych) i wiedzy o języku (modułu komunikacyjnego), jakkolwiek pozostają komplementarne funkcjonalnie.

4. Relewancja informacji udostępnianej przez inteligentny system faktograficzny jest zdeterminowana przez reguły implikacji. Prawomocność użycia pewnych reguł i spełnienie warunku wynikania koniecznego zapewnia się poprzez ich przypisanie tylko tym wyrażeniom w bazie wiedzy, które reprezentują cechy obiektów relewantne operacyjnie (ze względu na cel) lub interpretacyjnie. Połączenie tzw. deklaratywnego i proceduralnego sposobu odwzorowania wiedzy w systemie odpowiada możliwościom opisu implikacji relewantnej za pomocą dostępnego aparatu logiki, systemów entailment i logik relewancji.

POLA WYSZUKIWAWCZE W BAZACH DANYCH SDC-ORBIT

(Układ alfabetyczny)

Symbol	Zawartość pola
AAB	Analytical abstract
AB	Abstract
AC	Accession number
AN	Accession Number
AOS	Author organisation source
AU	Author
AV	Availability
AW	Awardee
AY	Accession number year
BI	Basic Index
BN	International Standard Book Number
CA	Cited Author
CALL	Library of Congress call number
CC	Category code name
CEN	Century
CG	Contract/grant numbers
CHAN	Clearing house accession number
CI	City
CJ	Cited journal
CL	Derwent classes
CN	NAL call number
CNO	Congressional Record number
COCO	Committee code
CORD	Co-ordinates
CP	Cited page
CS	Congress and session number
CV	Cited volume
CY	Cited year
DD	Deadline data, Dewey Decimal number
DEG	Degree
DT	Document type
EC	Event code
EY	Entry year
FR	Foreign
FS	File segment
FUND	Funding
HE	Heading note
IC	International patent classification
ICC	Issue category code
IN	Principal/associate investigators
IS	Issue
IT	Index terms
IW	Index words
JC	Journal citation, Journal coden
LA	Language
LC	Library of Congress card number
LO	Location, Location country
ME	Main entry (author)
NO	Notes
NU	Document/report/bill numbers
ON	Order number

OS	Organisational source
OSCO	Organisational source code
PA	Patent countries, Patent assignee
PC	Patent classification, Patentee/company code
PCC	Primary category code
PD	Publication date
PL	Patent location
PN	Patent number, Project number
PR	Priority country
PY	Publication year, Priority number
RN	Report number
RNO	Number of rings
RPR	Rings present
RSD	Ring system description
SCC	Secondary category code
SE	State
SF	Substituent fragment
SI	Superintendent of document / item number
SN	Source name
SO	Source
SOCO	Source code
SPO	Sponsoring organisation
ST	Supplementary terms
STAT	Citations to statutes at large
STF	Stereochemistry fragment
TA	Title annotation
TC	Taxonomic codes
TI	Title words
TN	Task number
TT	Title terms
UP	Update code
VOL	Volume
WT	Witness terms
XR	Cross- eference index terms

POLA WYSZUKIWAWCZE W BAZACH DANYCH LOCKHEED-DIALOG

Układ alfabetyczny

Symbol	Zawartość pola
AA	Asset amount, Author affiliation
AB	Abstract
AC	Legislative authority code, Country of patent application, Area code, Activity code, Assignee code
AG	Asset greater than, Grant greater than
AI	Alloys index
AL	Asset less than, Grant less than
AM	Grant amount, Contract amount
AN	Assignee name, Accession number, Agency name, Abstract number
AP	Approach
AR	Authority record
AS	Agency state, Agency state code, Agency/Service abbreviation
AT	Asset type, Article type
AU	Personal author, Inventor
AV	Availability on microfiche, Availability
BC	Biosystematic code, Branch city
BN	Biosystematic name, Branch name, Bureau number, Book number, SBN
BS	Branch state
BZ	Branch zip code
CA	Card alert, Cited author, Call number
CC	Concept code, Corporate source, Category code, Country code, Contractor/ /Company code, Class code, Geographic code CF Cosati field
CG	Contributions greater than, Corporate source location
CH	Clearinghouse code
CL	Patent classification number, Classification code, Contribution less than, Con- ference location
CM	Community code
CN	Contract/grant number, Concept name, Country name, Geographic name, Class name, Corporate source name, Company number
CO	Coden, Company, Corporate source, Contractor company name
CP	Contract number prefix, Country of publication
CR	Cited reference, CA reference, Contributions received
CS	Corporate source, Contractor state, Contractor state code
CT	Conference title
CY	Conference year
DC	Descriptor code, UDC number
DD	Dewey Decimal Number
DE	Descriptor
DF	Single-word descriptor, Domestic or foreign recipient
DG	Degree year
DN	Document number
DO	Donor
DS	Division/station code
DT	Document type, Date, Date of contract, Publication date
EC	Event code, Element count
ED	Edition, Errata data
EG	Expenditure greater than
EI	Internal revenue service number
EL	Expenditure less than
EN	Event name
ES	Establishment data, Element symbol
EV	Errata volume

EX	Employee size, Expenditure, Exchange
FC	Foundation city
FG	Founding greater than
FL	Founding less than
FC	Foreign patent number, Foundation name
FR	Federal region
FS	Foundation state, Field of science
FT	Foundation type
FU	Actual funding
FY	Fiscal year
GA	Grant amount
GC	Group code, Auxiliary descriptor code
GE	Geographic descriptor
GG	Grant greater than
GI	Gallery index
GL	Grant less than, Geographic location
GN	Grant number, Number of grants, Group number
GP GPO	(Government Printing Office) Number
GR	Growth rate, Grade
GS	Tag, Auxiliary descriptor, Geographic descriptor
HC	Headquarters city, Hardware capability code
HE	Historical period ending date
HI	Highest grant
HN	Headquarters name, Hardware capability name, Section heading
HP	CA (Chemical Abstracts) parent index name, Historical period
HS	Headquarters state, Historical period starting date
HX	Headquarters code
HZ	Headquarters zip code
ID	Identifier
IS	Issue number
IN	Investigator name
IV	Investigator
JA	Journal announcement
JC	Journal coverage, Journal code
JN	Title of publication, Journal name, Country of granted patent
JO	Journal name, Journal announcement, Journal citation, Journal
LA	Language, Loan amount
LC	Licensing country, Library of Congress card number
LM	Limitation
LN	Number of loans
LO	Location, Lowest grant, Location code
MA	Matching gifts
MC	Measure code
MF	Molecular formula
MN	Number of matching gifts, Measurement name, Manufacturer's name, Meeting number
NA	Name of program, Named person
NM	CA index name modification
NN	NICSEM (National Information Centre for Special Education Materials) accession number
NO	NAL Call number
NP	Publisher
NT	Descriptive note
NU	Number of program
OB	Objectives
OC	OCLC number
ON	Officer's name
PA	Patent assignee, Purpose and activities, Program amount
PB	Publisher, Publications
PC	Product code, Phase code

PD	Publication date, Process date
PG	Performing organisation place
PI	Country of patent, Periodic index term
PN	Bureau/project number, Patent number, Phase name, Product name, Number of programs
PO	Country of original patent, Performing organisation, Performing organisation code
PP	Place of publisher, Progress period
PR	Progress report
PS	Program sponsor, institution
PT	Phase type
PU	Publisher
PX	PACS code
PY	Publication year
PZ	Performing organization zip code
QC	Quick code
RC	Recipient city
RI	Title
RN	Report number, ID number, Abstract number, Recipient name, Record number, CAS registry number
RP	Research problem area code
RS	Recipient state
RT	Recipient type, Registration type
SA	Sponsoring agency, Scholarship amount
SB	CA substituent index name, Source publication
SC	Corporate source code, Section code, Subject category, System category code, CA section, Sponsor code, Source code, Subject code, Section heading code
SD	Search date, Sales in million \$, Start date
SE	Series note, Series, Series statement
SF	Special feature, Subfile
SG	Sponsoring organisation place
SH	Section heading
SL	Summary language
SM	Search month, Share of market, Subcommodity code
SN	Number of scholarship, Subject name, ISSN, SSIE number
SO	Source
SP	Sponsoring program, Sponsor
ST	Stereochemical descriptor, System category name, State/country abbreviation, Section heading
SY	Synonyms, Search year
SZ	Sponsoring organisation zip code
TC	Tag code
TD	Termination date
TI	Title
TN	Trade name
TX	Taxonomic descriptor
UC	Universal decimal code
UD	Update
VL	EC volume number
WT	Work type
YA	Year authorised
YR	Year, Year of publication, Degree year, Journal year
ZP	Foundation zip code

LITERATURA

- ACKERMAN R. (1956), „Begrundung einer strengen Implikation” *Journal of Symbolic Logic* No 2, 113-128.
- A *CLASSIFICATION of Library and Information Science* (1975), By R. Daniel, J. Mills with the assistance of R. Selwood and P. Elliot for the Classification Research Group, London: The Library Association.
- ADAM R. (1982), „Społeczne podejście do informacji w naukach społecznych” *Zagadnienia Informacji Naukowej* nr 2(41), 3-20.
- ADDIS T. R. (1985), *Designing Knowledge-Based Systems*. London: Kogan Page Ltd.
- ADLER H. (1981), „Zapotrzebowanie na informację biblioteczną w zakresie nauk biologicznych, doświadczalnych” *Zagadnienia Informacji Naukowej* nr 1(38), 13-28.
- AGAZZI E. (ed.), (1981), *Modern Logic — A Survey. Historical, Philosophical and Mathematical Aspects of Modern Logic and its Applications*, Dordrecht Holland: D. Reidel Publishing Company.
- AITCHINSON J., GILCHRIST A. (1972), *Thesaurus Construction. A Practical Manual*, London: Aslib.
- AJDUKIEWICZ K. (1985), „O znaczeniu wyrażeń”, [W: *Język i poznanie*]. Warszawa: Państwowe Wydawnictwa Naukowe, 102-136.
- AJDUKIEWICZ K. (1949), *Zagadnienia i kierunki filozofii. Teoria poznania — Metafizyka*, Warszawa: Czytelnik.
- AJDUKIEWICZ K. (1985), „Zdania pytajne” [W: *Język i poznanie*]. Warszawa: Państwowe Wydawnictwo Naukowe, 278-286.
- AMERICAN National Standard. *Guidelines for Thesaurus Structure, Construction and Use*, Z 39.19-1974.
- ANDERSON, A. R., BELNAP J. D. (1975), *Entailment: the Logic of Relevance and Necessity*. vol.1, Princeton — New York: Princeton University Press.
- ANDERSON, A. R., BELNAP J. D. Jr., DUNN J. M. (1992), *Entailment: the Logic of Relevance and Necessity*. vol.2, Princeton — New York: Princeton University Press.
- ANDERSON J. A., ROSENFELD E. (1988) *Neurocomputing — Foundations of Research*, Cambridge: MIT Press.
- APRESJAN J. (1971), *Koncepcje i metody współczesnej lingwistyki strukturalnej*, Warszawa: Państwowy Instytut Wydawniczy.
- APRESJAN J. D. (1980), *Semantyka leksykalna. Synonimiczne środki języka*, Wrocław: Zakł. Narodowy im. Ossolińskich.
- ARTOWICZ E. (1978), „Założenia języka informacyjno-wyszukiwawczego w systemie AWION” *Zagadnienia Informacji Naukowej* nr 2(33), 35-56.
- ARTOWICZ E. (1983), „Zastosowanie sztucznej inteligencji w systemach przetwarzania wiedzy” *Zagadnienia Informacji Naukowej* nr 2(43), 158-165.
- ARTOWICZ E., SZOMAŃSKI B. (1989), „System informacji o metodach i technikach organizatorskich w sterowaniu jakością” *Zagadnienia Informacji Naukowej* nr 1(54), 103-124.
- AUSTIN J. L. (1962), *How to do Things with Words*, Oxford University Press.
- BACH and HARMS H. (eds.), (1968), *Universal Linguistic Theory*, Rinehart and Winston. New York.
- BARR A., FEGENBAUM E. A. (1982), *The Handbook of Artificial Intelligence*,. vol. 1-2, Heuris Tech Press.

- BELNAP N. D. (1981), *Modal and Relevance Logics*. [W: AGAZZI E. (ed.), (1981)], 131-150.
- BIELICKA L. (1984), „Języki informacyjno-wyszukiwawcze typu obiekt-atrybut systemów faktograficznych” *Prace IINTE* 49.
- BIELICKA L. (1985), *Zasady gramatyczne w językach deskryptorowych*, Warszawa: IINTE.
- BIELICKA L., ŚCIBOR E. (1982), *Języki informacyjne. Rodzaje i zastosowanie w działalności dokumentacyjnej*, Warszawa: CIINTE.
- BOBROV D. G. (1975), *Dimension of Representation*, [W: BOBROV D.G., COLLINS A., 1975], 1-33.
- BOBROV D. G., COLLINS A. (1975), *Representation and Understanding*, New York: Academic Press.
- BOBROV D. G., WINOGRAD T. (1977), „An Overview of KRL: A Knowledge Representation Language” *Cognitive Science* vol. 1, 213-222.
- BOCHENSKI J. M. (1988), „Co logika dała filozofii” *Studia Filozoficzne* nr 6-7, 7-13.
- BOGUSŁAWSKI A. (1980), *Termin „presupozycja” i „ponjatijnyj aparat teorii teksta”*, [W: MA-YENOWA M. R. (red.)], 7-28.
- BOJAR B. (1979), *Opis semantyczny polskich czasowników ruchu i pojęć związanych z ruchem*, Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- BOJAR B. (1991), *Zarys językoznawstwa dla studentów bibliotekoznawstwa i informacji naukowej*, Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- BOLC L., CICHY M., ROŻAŃSKA L. (1982), *Przetwarzanie języka naturalnego*, Warszawa: Wydawnictwa Naukowo-Techniczne.
- BOLC L. (red.), (1983), *Przetwarzanie informacji reprezentowanej w postaci naturalnej*, Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- BOOKSTEIN A. (1979), „Relevance” *Journal of American Society for Information Science* vol. 29, 269-272.
- BRADFORD S. C. (1932), *The Documentation Chaos*. [W: *Documentation*]. London: Lockwood, 106-121.
- BUCHANAN B. (1979), *Theory of Library Classification*, London: Clive Bingley.
- CALKINS M. L. (1980), „Free Text or Controlled Vocabulary? A Case History Step by Step Analysis, plus other Aspects of Search Strategy” *Database*, June, 53-67.
- CAMBRIDGE *International Dictionary of English* (1995), Cambridge University Press.
- CARBONELL J. G. Jr. (1979), „POLITICS: Automated Ideological Reasoning” *Cognitive Science* vol. 2, 27-51.
- CARNAP R. (1962), *Logical Foundations of Probability*, The University of Chicago Press, 613.
- CARNAP R. (1956), *Meaning and Necessity. A Study in Semantics and Modal Logic*, 2nd ed. Chicago: The University of Chicago Press.
- CARNAP R. (1971), *The Logical Syntax of Language*, New York: Harcourt Brace.
- CHASE W. (1973), *Visual Information Processing*, New York: Academic Press.
- CHMIELEWSKA-GORCZYCA E. (1977), *Indeksy permutacyjne*, Warszawa: CIINTE.
- CHMIELEWSKA-GORCZYCA E. (1979), „Problem synonimii w teorii języków informacyjno-wyszukiwawczych” *Zagadnienia Informacji Naukowej* nr 2(35), 79-98.
- CHMIELEWSKA-GORCZYCA E. (1981), „Metoda indeksowania relacyjnego Farradane’a” *Zagadnienia Informacji Naukowej* nr 1(38), 83-97.
- CHMIELEWSKA-GORCZYCA E. (1981), „Organizacja zbioru informacji jako element języka systemu informacyjno-wyszukiwawczego” *Zagadnienia Informacji Naukowej* nr 2(39), 83-97.
- CHMIELEWSKA-GORCZYCA E. (1983), „Relacje syntagmatyczne w językach informacyjno-wyszukiwawczych” *Zagadnienia Informacji Naukowej* nr 2(43), 144-151.
- CHOROŚ K. (1981), „Indeksowanie dokumentów deskryptorami ważonymi” *Aktualne Problemy Informacji i Dokumentacji* nr 5-6, 22-26.
- CICHY M. (1983), *Reprezentacja wiedzy w systemach wyszukiwania informacji z dostępem w języku naturalnym*. [W: BOLC L., 1983], 103-136.
- CLEVERTON C. W., MILLS J., KEEN E. M. (1966), *Factors Determining the Performance of Indexing System. Test Results*. vol. 2., Aslib — Cranfield Research Project. Cranfield.

- COLE P., MORGAN J. (eds.), (1975), *Syntax and Semantics 3: Speech Acts*, New York: Academic Press.
- COLOMBO D. S., NIEHOFF R. T. (1977), *Improved Access to Scientific and Technical Information through Automated Vocabulary Switching*, Final Report Battele Columbus Labs.
- COMPUTER Readable Databases. A Directory and Data Sourcebook. vol. 1 Science, Technology, Medicine, vol. 2 Business, Law, Humanities, Social Sciences (1985), Williams M. E., (ed.), North Holland, Amsterdam: New York.
- COMPUTER Simulation of Cognitive Process (1963), [W: LUCE R. D., BUSH R. R., GALANTER E. (eds.), (1963)]
- COOPER W. S. (1971), „A Definition of Relevance for Information Retrieval” *Information Storage and Retrieval* No 7, 19-39.
- CUADRA C. A. (1964), *On the Utility of the Relevance Concept*, SP-1595 Systems Development Corporation. Santa Monica, California.
- CUADRA C. A., KATTER R. V. (1967), *Experimental Studies of Relevance Judgement. Final Report*, Report TM-3520. vol. 1 Project Summary, vol. 2 Description of Individual Studies, Systems Development Corporation, Santa Monica, California, June .
- CZERNY A. I. (1978), *Wstęp do teorii wyszukiwania informacji*, Warszawa: Ośrodek Informacji Naukowej PAN.
- ĆWIEKOWA J. (1974), *Opracowanie przedmiotowe piśmiennictwa*, Warszawa: Ośrodek Informacji Naukowej PAN.
- DAO THI QUI (1989), *Metodyka budowy języka informacyjno- wyszukiwawczego dla systemu informacji specjalistycznej. (Na przykładzie informacji o technologii preparatów enzymatycznych)*. Praca doktorska wykonana pod kierunkiem doc. dr hab. B. Bojar, Warszawa: Uniwersytet Warszawski. Wydział Neofilologii.
- DAVIS R., KING J. J. (1977), *An Overview of Production Systems*, [W: ELCOCK E., MITCHIE D. (1977)], 300-332.
- DAVIS R., LENAT B. D. (1982), *Knowledge-based Systems in Artificial Intelligence*, New York: McGraw-Hill International Book Company.
- DOBROWOLSKI Z. (1956), *Budowa Klasyfikacji*, Warszawa: Państwowe Wydawnictwa Techniczne.
- DOOYLE L. B. (1966), *Is Relevance an Adequate Criterion in Retrieval System Evaluation?*, Proceedings of the American Documentation Institute. Part 2, Washington, 199-200.
- ELCOCK E., MITCHIE D. (1977), *Machine Intelligence 8*, Chichester, England: Ellis Horwood.
- ENCYKLOPEDIA językoznawstwa ogólnego Polański K. (red.), (1993), Warszawa: Zakład Narodowy im. Ossolińskich.
- ENCYKLOPEDIA współczesnego bibliotekarstwa polskiego (1976), Warszawa: Zakład Narodowy im. Ossolińskich.
- EPSTEIN R. L. (1979), „Relatedness and Implication” *Philosophical Studies* nr 36, 137-173.
- FILLMORE Ch. (1968), *The Case for Case*, [W: BACH and HARMS H. (eds.), Rinehart and Winston. New York], 1-90.
- FILOZOFIA a nauka. Zarys encyklopedyczny (1987), Warszawa: Zakład Narodowy im. Ossolińskich, 513-518.
- FILOZOFIA starożytna. Wybrane teksty z historii filozofii (1968), Warszawa: Państwowe Wydawnictwo Naukowe, 275-281.
- FODOR J. A., GARRET M., WALKER E., PARKES C. (1980), „Against definitions”, *Cognition* 8.3, 263-367.
- FODOR J. A. (1983), *The modularity of mind*, Cambridge Mass.: MIT Press.
- FORD N. (1983), „Knowledge Structure in Human and Machine Information Processing — their Representation and Interaction”, *Social Science Information Studies* vol. 4, 209-222.
- FOSKETT D. J. (1977), „Classification and Indexing in Social Sciences”, *Aslib Proc.* vol. 29: No 22, 90-101.
- FOSKETT D. J. (1972), „A Note on the Concept of Relevance”, *Information Storage and Retrieval* No 8(2), 457-471.
- FOSKETT D. J. (1981), *The Subject Approach to Information*, 4th ed., London: Clive Bingley.

- FREGE G. (1884), *Die Grundlagen der Arithmetik: eine logisch-matematische Untersuchungen über den Begriff der Zahl*. Breslau.
- GARDIN J. C. (1965), *SYNTOL*, Rutgers series on systems for the intellectual organization of information. vol.II. New Brunswick: The Rutgers University Press.
- GARDIN J. C., de GROLLIER E., LEVERY F. (1964), *L'organisation de la documentation scientifique*, Paris: Gauthier — Villars.
- GOFFMAN W. (1970), *A General Theory of Communication*, [W: SARACEVIC T. (ed.), 726-747.
- GOFFMAN W. (1964), „On relevance as a measure”, *Information Storage and Retrieval* vol.2: No 3, 201-223.
- GOFFMAN W., NEVILL V. A., *Communication and Epidemic Process*, Proceedings of the Royal Society, A 298, No 1454, 316-334.
- GOFFMAN W., NEVILL V. A. (1966), „Methodology for Test and Evaluation of Information Retrieval Systems”, *Information Storage and Retrieval* vol. 3: No 1, 19-25.
- GOLDSTEIN Ch. M., FORD W. (1978), „The User-cordial Interface”, *On-line Review* vol. 2: No 3, 270-278.
- GOODMAN N. (1951), *The Structure of Appearance*, Harvard.
- GRABOWSKA M. (1986), „Amerykański zautomatyzowany system informacji prawniczej oraz informacji prasowej — LEXIS/NEXIS”, *Zagadnienia Informacji Naukowej* nr 2(49), 89-102.
- GRABOWSKA M. (1987), „Zautomatyzowane katalogi centralne w bibliotekach USA”, *Zagadnienia Informacji Naukowej* nr 2(51), 103-120.
- GREGORY N. (1979), „The US Congress on-line Users as Policy Makers”, *On-line Review* vol. 3: No 4.
- GRICE H. P. (1957), „Meaning”, *Philosophical Review* 66, 377-388.
- GRICE H. P. (1969), „Utterer's Meaning and Intentions”, *Philosophical Review* 78, 147-177.
- GRICE H. P. (1975), *Logic and Conversation*, [W: COLE P., MORGAN J. (eds.), (1975)]
- GUIDE to DIALOG Searching (1979), Palo Alto, California: Lockheed Information Systems. vol. I-IV.
- HAMMERL R., SAMBOR J. (1990), *Statystyka dla językoznawców*, Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- HARMAN P. (1976), „Funkci analiza faktograficeskich tabulek”, *Ceskoslovenska Informatika* nr 9, 239-244.
- HEMPLIŃSKI M., 1974, *Brytyjska filozofia analityczna*, Warszawa.
- HENRY W. M., LEIGH J. A., TEDD L. A., WILLIAMS P. W. (1980), *Online Searching*, London: Butterworths.
- HEWITT C., BISHOP P., STEIGER R. (1973), *An Universal Modular a Formalism for Artificial Intelligence*, Stanford, California.
- HILLMAN D. J. (1977), „Model for the On-line Management of Knowledge Transfer”, *On-line Review* vol.1: No 1, 23-30.
- HILLMAN D. J. (1964), „The Notion of Relevance”, *American Documentation* vol.15: No 1, 26-34.
- HINTIKKA J. (1966), *Studies in the Logic of Existence and Necessity*, *Monist*, 55-76.
- HJEMSLEV L. (1979), *Prolegomena do teorii języka*. [W: KURKOWSKA H., WEINSBERG A., (red.), (1979)].
- HOLÓWKA T. (1986), *Myślenie potoczne. Heterogeniczność zdrowego rozsądku*, Warszawa: Państwowy Instytut Wydawniczy.
- HOOVER R. E. (1980), *The Library and Information Manager's Guide to Online Services*, New York: Knowledge Industry Publication Inc.
- HOYLE W. G. (1965), „On the Number of Categories for Classification”, *Information Storage and Retrieval*, vol. 5: No 1, 1-6.
- HUSSERL E. (1967), *Idee czystej fenomenologii i fenomenologicznej filozofii*, tłum. D. Gierulanka, Warszawa.
- HUTCHINS W. J. (1978), *Languages of Indexing and Classification. Linguistic Study and Functions*, Stevenage: Peter Peregrinus Ltd.

- ISO 2788-1974 (Documentation — Guidelines for the Establishment and Development of Monolingual Thesauri).
- IVANKIN V. J. (1973), „O metodach wyboru klucowych słów przy koordynatnym indeksowaniu. Postanowka problemu”, *Naučno-techničeskaja informacija* Ser.2 nr 11, 14-9.
- JACKENDOFF R. (1983), *Semantics and Cognition*, Cambridge: The Mit Press, 283.
- JADACKA H. (1979), „O formie, znaczeniu i desygnacji terminu”, *Zagadnienia Naukoznawstwa* nr 4 (56).
- JAMES W. (1957), *Pragmatyzm*, przeł. M. Kozłowski, Warszawa.
- KANT I. (1781), *Kritik der reinen Vernunft*.
- KATZ J. J., FODOR J. A. „The Structure of a Semantic Theory”, *Language* 52, 1-17.
- KELLY M. J., CHAPANIS N. (1977), „A Limited Vocabulary Natural Language Dialogue”, *International Journal of Man-Machine Studies* No 9, 479-501.
- KEMP D. A. (1974), „Relevance. Pertinence and Information System Development”, *Information Storage and Retrieval* No 10(2), 37-47.
- KERVEGANT D. (1962), *Introduction a la Documentation Agronomique. La Classification Bibliographique*, Paris.
- KEYNES J. M. (1921), *A Treatise on Probability*, London.
- KOCHEN M. (1974), *Principles of Information Retrieval*, Los Angeles: Mellville Publishing Company.
- KOLODNER J. (1984), *Organisation and Retrieval in a Conceptual Memory for Events*. [W: Schank R. C., 1984].
- KOPALIŃSKI W. (1978), *Słownik wyrazów obcych i zwrotów obcojęzycznych*, Warszawa: Wiedza Powszechna.
- KOZAKCHOV L. S. (1969), „Relevancija w informatyce i naukovedenii”, *Naučno-techničeskaja informacija* Ser. 2 nr 8, 3-11.
- KORNER S. (1974), *Categorical Frameworks*, Oxford.
- KRIPKE S. A. (1959), „A Completeness Theorem in Modal Logic”, *Journal of Symbolic Logic* nr 24, 1-15.
- KRISTALNYJ B. V., RASKINA A. A., SIDOROV I. S. (1976), „O ponjatii „faktograficeskaja informacija”, *Voprosy informacionnoj teorii i praktiki* Ser. 1. Moskwa, 7-16.
- KURKOWSKA H., WEINSBERG A. (red.), (1979), *Językoznawstwo strukturalne*, Warszawa: Państwowe Wydawnictwo Naukowe.
- KWAPISZ A. (1981), *Informacja patentowa*, Warszawa: CIINTE nr 19.
- KYBURG H. (1970), *Probability and Inductive Logic*, London.
- LAKOFF G., JOHNSON M. (1988), *Metafory w naszym życiu*, Warszawa: Państwowy Instytut Wydawniczy.
- LANCASTER F. W. (1968), *Information Storage and Retrieval System. Characteristics, Testing, Evaluation*, New York.
- LANGACKER R. W. (1987), *Foundations of Cognitive Grammar. Part 1.*, Stanford.
- LANGACKER R. W. (1995), *Wykłady z gramatyki kognitywnej*, Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- LEECH G. N. (1983), *Principles of Pragmatics*, London.
- LEWIS D. A. (1980), „Today's Challenge — Tomorrow's Choice Change or be Changed or the Doomsday Scenario MK 2”, *Journal of Information Science* nr 2, 59-74.
- LINDSAY P. H., NORMAN D. A. (1984), *Procesy przetwarzania informacji przez człowieka. Wprowadzenie do psychologii.*, tłum. A. Kowaliszyn, H. Szafraniec, dodatek A, Rudnicki A. J. dodatek B, Warszawa: Państwowe Wydawnictwo Naukowe.
- LOTKA A. J. (1926), „The Frequency Distribution of Scientific Productivity”, *Journal of the Washington Academy of Sciences* No 16(12), 283-350.
- LUCE R. D., BUSH R. R., GALANTER E. (eds.), (1963), *Handbook of Mathematical Psychology*, vol. 1 New York: Wiley.
- MADALIŃSKA-BUGAJ E. (1983), *Metody dedukcji w systemach pytań i odpowiedzi*, [W: BOLC L. (red.), (1983)], 85-102.

- MAŁA encyklopedia logiki* (1970), Marciszewski W. (red.), Wrocław: Zakład Narodowy im. Ossolińskich.
- MARCISZEWSKI W. (1973), „Relacje tematyczno-hierarchiczne w językach deskryptorowych”, *Zagadnienia Informatyki Naukowej* nr 2(23), 35-52.
- MARCISZEWSKI W. (1977), *Metody analizy tekstu naukowego*, Warszawa: Państwowe Wydawnictwo Naukowe.
- MARCISZEWSKI W. (red.), (1987), *Logika formalna. Zarys encyklopedyczny*. Warszawa: Państwowe Wydawnictwo Naukowe.
- MARON M. E., KUHN J. L. (1960), „On Relevance. Probabilistic Indexing and Information Retrieval”, *Journal of the Association for Computing Machinery* No 7(3), 216-244.
- MARULLI L., KOENIG M. E. D. (1979), „Bradford Distribution of Data Elements”, *Journal of American Society for Information Science* vol. 30: No 2, 107.
- MAYENOWA M. R. (red.), (1980), *Tekst. Język. Poetyka*, Warszawa: Zakład Narodowy im. Ossolińskich.
- MCCULLOCH W. S., PITTS W. (1943) „A logical calculus of the ideas immanent in nervous activity”, *Bulletin of Mathematical Biophysics* No 5, 115-133.
- Mc DERMOTT D. (1978), *The Last Survey of Representation of Knowledge*, Proceedings of the AISBG/GI Conference on AI, Hamburg, 206-221.
- MEADOW Ch. T. (1972), *Analiza systemów informacyjnych. Wyszukiwanie, organizacja i przetwarzanie informacji*, Warszawa: Wydawnictwa Naukowo-Techniczne.
- MELČUK I., ZHOLKOVSKY A. (1984), *Explanatory Combinatorial Dictionary of Modern Russian. Semantico-Syntactic Studies of Russian Vocabulary*, Wiener Slavistischer Almanach Sonderband 14, Vienna.
- MANHEIM K. (1952), *The Problem of a Sociology of Knowledge*, [W: *Essays on the Sociology of Knowledge*]. London.
- MEYER R. K., ROUTLEY R. (1972), „Algebraic Analysis of Entailment”, *Logique et Analyse* vol. XV, 407-428.
- MIĘDZYNARODOWA Klasyfikacja Patentowa (1980-1981) IV edycja. 9 tomów. Warszawa: Urząd Patentowy PRL.
- MINSKY M. (ed.), (1968), *Semantic Information Processing*, Cambridge: MIT Press.
- MINSKY M., PAPER T. (1969), *Perceptrons*, Cambridge: MIT Press.
- MINSKY M. (1975), *A Framework for Representing Knowledge*, [W: WINSON P. (ed.), (1975)].
- MOOERS C. N. (1951), „Zotocoding Applied to Mechanical Organization of Knowledge”, *American Documentation* No 1, 20-32.
- MOORE E. G. (1920), *External and Internal Relations*, Proc. Aristotelian Society, 42-62.
- MORYK A. (1981), „Wprowadzenie do III edycji Międzynarodowej Klasyfikacji Patentowej”, *Informator Patentowy OPT* nr 2.
- MYKOWIECKA A. (1987), *Przegląd systemów automatycznej generacji tekstu w języku naturalnym*, Prace Instytutu Podstaw Informatyki PAN 614, listopad.
- MYKOWIECKA A. (1992), *Podstawy przetwarzania języka naturalnego. Metody generowania tekstów*, Warszawa: Akademicka Oficyna Wydawnicza.
- NAGEL E. (1970), *Struktura nauki*, przeł. E. Gedymin, B. Rassalski, H. Eilstein, Warszawa.
- NEUMANN von J. (1958), *The Computer and the Brain*, New Haven: Yale University Press.
- NEWELL A. (1973), *Production Systems: Models of Control Structure*, [W: Chase W., (1973)], 463-526.
- NEW Kind of „Weight-relators” in Indexing” (1975), *Studi ci Cerc. Docum.* nr 3-4, 294-296.
- NIKITINA S. E. (1978), *Tezaurus po teoreticeskoj i prikladnoj lingvistike (Avtomatičeskaja obrabotka teksta)*, Moskwa: Izdat. Nauka.
- NILSSON N. (1983), *Artificial Intelligence Research and Applications*, Menlo Park: SRI International Inc.
- NILLSON N. (1971), *Problem-solving Methods in Artificial Intelligence*, New York: McGraw-Hill.

- NONIEWICZ U. (1988), *Definiując niezdefiniowane. Próba prezentacji znaczenia terminu „system informacyjno-wyszukiwawczy”*. Praca magisterska wykonana pod kierunkiem doc. dr hab. B. Bojar., Warszawa: Instytut Bibliotekoznawstwa i Informatyki UW.
- NORMAN D. A. (ed.), (1970), *Models of Human Memory*, New York — London: Academic Press.
- O'CONNOR J. (1967), „Relevance Disagreements and Unclear Request Forms”, *American Documentation* No 18(3), 165-177.
- O'CONNOR J. (1968), „Some Question Concerning „Information Need”, *American Documentation* vol. 19: No 2, 200-203.
- OGÓRKIEWICZ W. (1985), „Wskaźniki roli w deskryptorowym języku informacyjno-wyszukiwawczym”, *Zagadnienia Informatyki* nr 2(47), 71-94.
- OXFORD *Advanced Learners Dictionary of Current English* (1981), London.
- PARRY W. T. (1933), „Ein Axiomensystem für eine neue Art von Implikation (Analytische Implikation)”, *Ergebnisse eines mathematischen Kolloquiums* vol. 4, 5-6.
- PEIRCE Ch. S. (1960), *Collected Papers*, Cambridge, Massachusetts.
- PERREAULT I. (1965), „Categories and Relators, a New Schema”, *Review Intern. Docum.* vol. 32: No 4, 136-144.
- PERRY J. W., KENT A. (1956), *A. Machine Literature Searching*, New York: Interscience Publishers.
- PERRY W. J., KENT A. (1958), *Tools for Machine Literature Searching. Semantic Code Dictionary. Equipment. Procedures*, New York.
- POLANYI M. (1958), *Personal Knowledge*, Chicago.
- POPOWSKA H. (1985), „Bazy danych. Zagadnienia klasyfikacji i serwisy on-line”, *Aktualne Problemy Informatyki i Dokumentacji* nr 5, 22-30.
- POPOWSKA H. (1988), „Bazy danych — klasyfikacja i typologie”. Część 1, *Aktualne Problemy Informatyki i Dokumentacji* nr 5, 14-23.
- POPOWSKA H. (1989), „Bazy danych — klasyfikacja i typologie”. Część 2, *Aktualne Problemy Informatyki i Dokumentacji* nr 3, 16-22.
- PRICE D. J. de S. (1965), „Networks of Scientific Papers”, *Science* No 149 (3683), 510-515.
- PSZCZOŁOWSKI T. (1978), *Mała encyklopedia prakseologii i teorii organizacji*, Warszawa: Zakład Narodowy im. Ossolińskich.
- QUILLIAN M. (1968), *Semantic Memory*, [W: MINSKY M. (ed.), (1968)], 227-270.
- RANGANATHAN R. S. (1967), *Prolegomena to Library Classification*, 3rd ed. London: Asis Publishing Houses.
- RASIOWA H. (1975), *Wstęp do matematyki współczesnej*, Warszawa: Państwowe Wydawnictwo Naukowe.
- REES A. M., SARACEVIC T. (1963), „Conceptual Analysis of Questions in Information Retrieval Systems”, *Proceedings of the American Documentation Institution. Part 2*, Washington, 175-177.
- REES A. M., SARACEVIC T. (1966), „The Measurability of Relevance”, *Proceedings of the American Documentation Institute*, Washington No 3, 225-234.
- REFERENCE GUIDE to SOCRATES (1985), The Online Library Catalog of Stanford University.
- RIESBECK Ch. K. (1978), *An Expectation Driven Production System for Natural Language Understanding*, New York: Academic Press Inc.
- ROBOWSKI J. (1974), *Języki deskryptorowe. Analiza porównawcza języków deskryptorowych i innych języków informacyjno-wyszukiwawczych*, Warszawa: CIINTE.
- ROSCHE E. (1977), *Human Categorization*, [W: WARREN N. (ed.), (1977)].
- ROSENBLATT F. (1958), „The perceptron: A probabilistic model for information storage and organization in the brain”, *Psychological Review* No 65, 386-408.
- RUSTIN R. (1975), *Natural Language Processing*, New York: Algorithmic Press.
- RYLE G. (1974), *Świat nauki i świat życia codziennego*, [W: Hempliński M. (1974)].
- SADOWSKA J. (1982), *Kategorie semantyczne określników w języku haseł przedmiotowych*, Rozprawa doktorska wykonana pod kierunkiem doc. dr hab. B. Bojar, Uniwersytet Warszawski, Wydział Neofilologii.
- SADOWSKA J. (1983), „Przedmiot i temat w teorii i praktyce katalogowania przedmiotowego”, *Zagadnienia Informatyki* nr 2(43), 53-68.

- SALTON G., SMART. (1975), *Automatyczny system wyszukiwania informacji*, Warszawa: Wydawnictwa Naukowo-Techniczne.
- SAMBOR J. (1972), *Słowa i liczby. Zagadnienia językoznawstwa statystycznego*, Wrocław: Zakład Narodowy im. Ossolińskich.
- SARACEVIC T. (1970), *Introduction to Information Science*, New York: Bowker, 726-747.
- SARACEVIC T. (1975), „Relevance: a Review of and a Framework for the Thinking on the Notion in Information Science”, *Journal of American Society for Information Science* No 6., 321-343.
- SARACEVIC T., REES A. M. (1963), „Conceptual Analysis of Questions in Information Retrieval Systems”, *Proceedings of the American Documentation Institute*, Washington No 2, 175-177.
- SATOSI WANATABE (1960), „Theorem of Ugly Duckling. A Mathematical Explication of Inductive Inference”, *Proceedings of International Colloquium*, Tihany, Hungary.
- SAUSSURE de F. (1961), *Kurs językoznawstwa współczesnego*, Warszawa: Państwowe Wydawnictwo Naukowe.
- SCHANK R. C. (1975), *Conceptual Information Processing. Fundamental Studies in Computer Science*. vol. 3., North Holland.
- SCHANK R. C. (1984), *Dynamic Memory. A Theory of Reading and Learning in Computers and People*, Cambridge — London — New York: Cambridge University Press.
- SCHANK R. C. (1980), „Language and Memory”, *Cognitive Science* vol. 4, 243-284.
- SCHANK R. C., ABELSON R. P. (1975), *Scripts, Plan and Understanding*, New York: Lawrence Erlbaum.
- SCHANK R. C., COLBY K. M. (1973), *Computer Models of Thought and Language*, San Francisco: Freeman.
- SCHELER M. (1973), *Phenomenology and Theory of Knowledge*, [W: *Selected Philosophical Essays*].
- SCHELER M. (1973), *Selected Philosophical Essays*, Evanston.
- SCHMIDT C., SRIDHARAN N. (1977), „The Plan Recognition Problem: A Hypothesize and Revise Paradigm”, *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Boston, Massachusetts.
- SCHUTZ A. (1964), *Collected Papers*. vol. 2, The Hague, 73.
- SCHUTZ A. (1970), *Reflections on the Problem of Relevance* New Haven: Yale University Press.
- SEXTUS EMPIRICUS (1929), *Outline of Pyrrhonism and Against the Logicians*, Eng. transl. by Rev. R. G. Bury., Cambridge: Harvard University Press.
- SHORTLIFFE E. (1976), *Computer-based Medical Consultations: MYCIN*, New York: Elsevier.
- SKOROCHOD'KO E. F., PSENICNAJA L. E., KARJALAJNE I. N. (1968), *Informacionno-poiskowaja sistema „BIT”*, Kiev: Naukova Dumka.
- SŁOWNIK ENCYKLOPEDYCZNY TERMINOLOGII JĘZYKÓW I SYSTEMÓW INFORMACYJNYCH, B. Bojar (red.), (1993), Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- SŁOWNIK JĘZYKA POLSKIEGO, 1981, M.Szymczak (red.), t.3,4, Warszawa: Państwowe Wydawnictwo Naukowe.
- SŁOWNIK terminologiczny informacji naukowej, 1979, Wrocław: Zakład Narodowy im. Ossolińskich.
- SOSIŃSKA B. (1981), „Struktura wyrażen języków informacyjnych a konotacyjna i denotacyjna interpretacja znaczenia”, *Zagadnienia Informacji Naukowej* nr 2(39), 41-60.
- SOSIŃSKA B. (1983), „O tzw. relacjach fazowych w teorii języków informacyjnych”, *Zagadnienia Informacji Naukowej* nr 1(42), 111-114.
- SOSIŃSKA B. (1983), „Międzynarodowe Centrum Koordynacyjne Sieci Systemów Informacji Bibliotecznej — nowa funkcja OCLC”, *Zagadnienia Informacji Naukowej* nr 2(43), 99-114.
- SOSIŃSKA B. (1985), „Reprezentacja wiedzy w systemach informacji dokumentacyjnej”, *Zagadnienia Informacji Naukowej* nr 1(46), 19-36.
- SOSIŃSKA B. (1986), „Denotacja wyrażen języka informacyjno-wyszukiwawczego”, *Zagadnienia Informacji Naukowej* nr 1(48), 29-40.
- SOSIŃSKA B. (1986), „Konotacja wyrażen języków informacyjno-wyszukiwawczych”, *Zagadnienia Informacji Naukowej* nr 2(49), 15-38.

- SOSIŃSKA B. (1987), *Relacje między planem treści i planem wyrażenia w języku informacyjno-wyszukiwawczym*. Praca doktorska wykonana pod kierunkiem doc. dr hab. B. Bojar, Warszawa: Uniwersytet Warszawski, Wydział Neofilologii.
- SPANG-HANSEN H. (1976), *Roles and Links Compared with Grammatical Relations in Natural Languages*, Lyngby: Dansk Dansk Teknisk Litteraturselskab.
- SPERBER D., WILSON D. (1986), *Relevance: Communication and Cognition*, Basil Blackwell Ltd.
- STRAWSON P. (1964), „Intention and Cognition in Speech Acts”, *Philosophical Review* 73, 439-460.
- SWETS J. A. (1963), „Information Retrieval Systems”, *Science* No 141 (3577), 245-250.
- SWIFT D. F. (1977), „A Multi-Modal Approach to Indexing and Classification”, *International Classification* vol. 4: No 2, 90-94.
- SZREJDER J. A. (1975), *Równość, podobieństwo, porządek*, Warszawa: Wydawnictwa Naukowo-Techniczne.
- ŚCIBOR E. (1982), *Typologia strukturalna języków informacyjnych*, Warszawa: IINTE.
- TABAKOWSKA E. (1990), „Językoznawstwo kognitywne a poetyka przekładu”, W: *Teksty Drugie*.
- TADEUSIEWICZ R. (1993), *Sieci neuronowe*, Warszawa: Akademicka Oficyna Wydawnicza RM.
- TAUBE M. A. (1965), „Note on the Pseudomathematics of Relevance”, *American Documentation* vol. 6, No 2, 69-72.
- TAUBE M., GULL C. D., WATCHEL I. S. (1952), „Unit Terms in Coordinate Indexing”, *American Documentation* No 4, 213-218.
- TAYLOR W. K. (1960), *The computer and the nervous system. Models and Analogues in Biology*, Cambridge University Press.
- The LIBRARY and Information Manager's Guide to Online Services* (1980), Hoover R. E. (ed.), New York: Knowledge Industry Publications Inc.
- The „MIXED Term-links” — a New Category of Relators in Indexing (1975), *Studi ci Cerc. Doc.* No 3-4, 288-293.
- THESAURUS of Engineering Terms. A list of Engineering Terms and their Relationships for Use in Vocabulary Control in Indexing and Retrieving Engineering Information* (1964), New York: Engineers Joint Council.
- UNGURIAN O. (1976), *Elementy teorii języków informacyjnych*, Warszawa: Ośrodek Informacji Naukowej PAN.
- UNGURIAN O. (1982), „Ogólna struktura języka informacyjnego. Propozycje „szkieletu organizacyjnego” słownictwa”, *Zagadnienia Informacji Naukowej* nr 2(41), 21-40.
- UNGURIAN O. (1975), *Teoria i praktyka klasyfikacji fasetowej S. R. Ranganathana*, Warszawa: IINTE.
- UNGURIAN O. (1978), *Wprowadzenie do Uniwersalnej Klasyfikacji Dziesiętnej*, Warszawa: Wydawnictwa Komunikacji i Łączności.
- URQUART D. J. (1959), „Use of scientific Periodicals”, *Proceedings of the International Conference on Scientific Information*, Washington: National Academy of Sciences, 277-290.
- VENEZIANO V. (1980), „Library Automation: Data Processing and Processing for Data”, *Annual Review of Information Science and Technology* vol. 15, 110-145.
- WALLACE M. (1984), *Communicating with Database in Natural Language*, Ellis Horwood Series in Artificial Intelligence, New York.
- WARREN N. (ed.), (1977), *Advances in Cross-Cultural Psychology*, New York: Academic Press.
- WERESZCZYŃSKA-CISŁO B. (1984), *Przydatność specyfikacji relacji kojarzeniowych w procesie wyszukiwania informacji z zakresu wybranej gałęzi technologii żywności*. Praca doktorska wykonana pod kierunkiem prof. dr hab. W. Marciszewskiego. Warszawa: Uniwersytet Warszawski, Wydział Neofilologii.
- WERESZCZYŃSKA-CISŁO B. (1985), „Relacje kojarzeniowe w językach informacyjnych”, *Zagadnienia Informacji Naukowej* nr 1(46), 69-80.
- WERESZCZYŃSKA-CISŁO B., OGÓRKIEWICZ W. (1986), „Pytania informacyjne w ujęciu teoretycznym”, *Zagadnienia Informacji Naukowej* nr 1(48), 41-60.
- WIERZBICKA A. (1972), *Semantic Primitives*, Athenäum Verlag.

- WILKS Y. (1973), *An Artificial Intelligence Approach to Machine Translation*, [W: SCHANK R.C., COLBYK.M. (1973)].
- WILSON P. (1973), „Situational Relevance”, *Information Storage and Retrieval* vol. 9, No 8, 457-471.
- WINOGRAD T. (1980), „What does it mean to understand Language?”, *Cognitive Science* vol. 4, 209-241.
- WINSON P. (ed.), (1975), *The Psychology of Computer Vision*, New York: McGraw-Hill.
- WITTGENSTEIN L. (1972), *Dociekania filozoficzne*, Warszawa: Państwowe Wydawnictwa Naukowe.
- WOJTASIEWICZ O. A., BOJAR B. (1979), *Kod semantyczny SINTO*, Sprawozdanie z prac za lata 1977 i 1978 (maszynopis).
- WOJTASIEWICZ O. A., SOSIŃSKA B. (1982), „O koncepcji tworzenia makrotezaurusa nauk społecznych”, *Zagadnienia Informatyki Naukowej* nr 1(40), 3-20.
- WOODS W. A. (1975), *An Experimental Parsing System for Transition Network Grammars*, [W: RUSTIN R., (1975)], 111-154.
- ZINOWJEW A. (1976), *Logika nauki*, Warszawa: Państwowe Wydawnictwo Naukowe.

SKOROWIDZ RZECZOWY

- aktualizacja bazy wiedzy, 234
- aktualny kontekst przetwarzania, 156
- alternatywność, 33
- analiza fasetowo-fazowa, 127
- analiza frazowa, 217
- analiza morfologiczna, 219
- automatyczne indeksowanie, 219

- baza danych
 - konceptualna, 143
- bazy danych
 - nie-bibliograficzne, 199
 - zakres tematyczny, 193
- Bidirectional Associative Memory (BAM), zob. dwukierunkowa pamięć skojarzeniowa

- cechy sytuacyjne obiektów
 - UKD, 125
- charakterystyka wyszukiwawcza dokumentu, 20
- charakterystyka wyszukiwawcza obiektu, 20

- definicja korelacyjna, 34
- dekodowanie informacji, 52
- denotacja, 20
 - bezpośrednia, 89
 - pośrednia, 90
- desygnat bezpośredni, 89
- desygnat pośredni, 89
- dopowiedzenie, 98
- dwukierunkowa pamięć skojarzeniowa, 170
- dstrybutor baz danych
 - funkcje, 187
- dstrybutor informacji, 186

- efekty kontekstowe, 60; 63
- efektywność przetwarzania informacji, 66
- eksperyment cranfieldski, 76
- eksplicacja, 35
- eksplicacja semantyczna
 - zasady, 35
- eksplicacja znaczenia
 - JIW, 100

- ekstensja wyrażen, 89
- ekwiwalencja, 31
- elementarne cechy semantyczne, 151
 - słowniki, 151
- elementarne działanie, 154
- entailment, 32; 37; 38; 40; 56; 79; 88; 222
- epidemiczna teoria komunikacji, 75
- etnometodologia, 26
- etykiety pól wyszukiwawczych, 195
- explicandum, 35
- explicatum, 35

- fakt, 145
 - definicja, 145
- faseta
 - wykładniki, 99
- filozofia fenomenologiczna, 23
- formalne procedury dowodzenia, 147
- formuła zdaniowa
 - systemy online, 203
- funkcja confirmacji, 32; 56
- funkcja metatekstowa, 176
- funkcja użyteczności informacji, 81
- funkcje przypadku (Kasus functions), 218

- generowanie odpowiedzi w języku naturalnym, 220
- glossematyka, 88
- gramatyka częściowo pozycyjna, 102
- gramatyka języków deskryptorowych
 - wiedza o rzeczywistości, 134
- gramatyka jhp
 - wiedza o rzeczywistości, 129
- gramatyka JIW
 - systemy online, 203
- gramatyka Klasyfikacji Dwukropkowej
 - wiedza o rzeczywistości, 127
- gramatyka kodów semantycznych
 - wiedza o rzeczywistości, 137
- gramatyka Międzynarodowej Klasyfikacji Patentowej
 - wiedza o rzeczywistości, 131
- gramatyka niepozycyjna, 102

- gramatyka pozycyjna, 102
 gramatyka UKD
 wiedza o rzeczywistości, 124
 gramatyka zerowa, 102
 gramatyki generatywno-transformacyjne, 217
- hiperonimia**, 92
hipoteza wspólnej wiedzy, 53
- idealne modele konceptualne**, 140
implikacja, 32; 37; 56; 63
 analityczna, 43; 59
 dla rachunku zdań, 45
 materialna, 38
 mocna, 40
 nietrywialna, 58
 ściśła, 39; 40; 45
 relewantna, 50
 syntetyczna, 59
implikacje kontekstowe, 60
impresywna funkcja
 JIW, 110
- indeks**
 wyrażeń relewantnych, 189; 196
indeksowanie ankietowe
 reprezentacja wiedzy, 136
indeksowanie relacyjne Farradane'a, 137
indeksy
 funkcja, 100
 pomocnicze online
 funkcje, 202
 rdzeni terminów, 203
 rdzeni wyrazowych, 203
 unitermów, 203
- informacja dokumentacyjna**, 89
informacja encyklopedyczna, 55
informacja faktograficzna, 89
informacja irrelewantna, 30
informacja konceptualna, 57
informacja leksykalna, 57
informacja logiczna, 57
informacja metodologiczna, 183
informacja relewantna, 23
informacja relewantna negatywnie, 33
informacja relewantna pozytywnie, 33
informacja teoretyczno-koncepcyjna, 183
instrukcja słowna, 178
instrukcja wyszukiwawcza, 175; 177; 181
 formułowanie, 195
 formułowanie online, 203
 modyfikacja, 191
instrumenty informacyjne SIW, 185
integrowanie zasobów leksykalnych JIW
 bazy danych, 202
inteligentne systemy faktograficzne, 222
intencja informacyjna, 54
intencja komunikacyjna, 54
intensja wyrażeń, 89
interakcyjny program wyszukiwawczy, 191
interpretacja wypowiedzenia, 52
interpretacja zdarzenia
 system inteligentny, 229
interpretacja znaczenia, 89
 denotacyjna, 89
 inferencyjna, 53
 konotacyjna, 89
- język deskryptorowy**, 198
 system SIMTOJ, 122
język haseł przedmiotowych, 197
język haseł przedmiotowych (jhp), 107
 wiedza o rzeczywistości, 108
język kwerend, 189
język manipulowania danymi, 177
język opisu danych, 177
język predykatów pierwszego rzędu, 223
język RX-kodów, 110; 117
język słów kluczowych, 96; 197
język SYNTOL, 110
języki deskryptorowe, 120
 wiedza o rzeczywistości, 120
języki informacyjno-wyszukiwawcze
 specjalistyczne, 103; 117; 129
 systemy online, 206
 ortografia, 208
 uniwersalne, 103; 129
językoznawstwo strukturalne, 140
jhp
 określniki lokalizujące, 109
 określniki ogólne, 109
 określniki szczegółowe, 109
 tematy chronologiczne, 108
 tematy formalne, 108
 tematy geograficzne, 108
 tematy językowo-etniczne, 108
 tematy klasowe, 108
 tematy osobowe, 108
 tematy przedmiotowe, 107
 tematy rzeczowe, 108
- kartoteki wzorcowe**
 systemy online, 201; 215
katologi przedmiotowe
 wiedza o rzeczywistości, 130
kategorie ontologiczne Arystotelesa, 139
kategorie ontologiczne Ranganathana, 106

kategorie semantyczne, 94
 kategorie syntaktyczne, 94
 kategoryzacja obiektów, 199
 Międzynarodowa Klasyfikacja Patentowa, 118
 system SIMTOJ, 122
 kategoryzacja określników, 109
 kategoryzacja semantyczna, 99; 102; 152
 język hasel przedmiotowych, 107
 język RX-kodów, 116
 języki deskryptorowe, 120
 Klasyfikacja Dwukropkowa Ranganathana, 106
 kod do opisu dyscyplin sportowych, 111
 kod SINTO, 110; 111
 kod W.Perry`ego i A.Kenta, 114
 Międzynarodowa Klasyfikacja Patentowa, 118
 poddziały kategorii treściowych UKD, 105
 poddziały pomocnicze UKD, 104
 poddziały wspólne analityczne UKD, 105
 polskie czasowniki ruchu, 112
 UKD, 104
 kategoryzacja semantyczna słownictwa, 94
 kategoryzacja syntaktyczna, 99
 Międzynarodowa Klasyfikacja Patentowa, 119
 kategoryzacja tematów, 108
 kategoryzacja wyrażeń
 język RX-kodów, 117
 Klasyfikacja Dwukropkowa Ranganathana
 wiedza o języku, 106
 wiedza o rzeczywistości, 107
 klasyfikacja patentowa
 międzynarodowa, 198
 klasyfikacje biblioteczne
 systemy online, 213
 klasyfikacje patentowe
 narodowe, 198
 klatka, 159
 klucz wyszukiwawczy, 175
 kod semantyczny ASM-WRU, 114
 kod semantyczny SINTO, 110; 111; 122
 kod semantyczny W.Perry`ego i A.Kenta, 111; 114
 kodowanie informacji, 52
 kody indeksowe, 119
 kody semantyczne, 110
 funkcja informacyjna, 110
 wiedza o rzeczywistości, 110
 komplementarność, 92
 komunikacja
 ostensywna, 52; 54
 werbalna, 52
 komunikacja użytkownika z systemem komendy, 193
 komunikacja z systemem
 schematy przebiegu dialogu, 220
 schematy retoryczne, 220; 221
 komunikowanie się z użytkownikiem
 system inteligentny, 215
 konceptualne struktury informacji, 140
 konfirmacja
 ilościowa, 36
 klasyfikacyjna, 36
 porównawcza, 36
 konotacja, 14
 wyrażeń JIW, 104
 kontekst przetwarzania założeń, 62
 kontekstualizacja, 60
 kontekstualizacja informacji, 61
 konwersywność, 92
 koordynacja, 18
 prosta, 102
 korelacja epistemiczna, zob. definicja korelacyjna
 kryteria relewancji
 pozajęzykowe, 94
 wykładniki językowe, 91
 kryteria relewancji informacji, 20
 kwalifikatory, 99
 kwerenda, 177

 logiczna interpretacja relewancji, 74
 logiczna relewancja informacji, 72
 logika indukcji, 31
 logiki modalne, 39; 40
 logiki relewancji, 38; 41

 łańcuch klasyfikacyjny, 99

 maksymy relewancji, 54
 maskowanie, 208; 209
 mechanizm transformacji informacji, 58
 mechanizmy wnioskowania, 230
 systemy inteligentne, 222
 meta-metainformacja, 192
 metainformacja, 15; 192
 metainformacyjna funkcja
 JIW, 110; 184
 metawiedza, 145
 metoda pojęć elementarnych, 111
 metody kognitywne, 139
 metody selekjonowania informacji
 techniczne, 175
 metody statystyczne

- wskaźniki wagi, 136
- Międzynarodowa Klasyfikacja Patentowa, 117
 - formuła zdaniowa, 132
 - funkcje, 117
 - kod indeksowy, 132
 - wiedza nerelewantna, 131; 146
 - wiedza relewantna, 131
- minimalny warunek relewancji systemowej, 189
- mnożniki semantyczne, 114
- możliwy świat, 40
- model danych, 17
 - hierarchiczny, 17
 - relacyjny, 17
 - sieciowy, 17
- model komórki nerwowej, 168
- model komunikacji
 - inferencyjny, 52
 - pragmatyczny, 53
 - semiotyczny, 52
- model pamięci asocjacyjnej, 148
- model pamięci dynamicznej Schanka, 165
- model pamięci konceptualnej, 218
- model struktury neuronu, 169
 - warstwy, 169
- modele języków informacyjno-wyszukiwawczych
 - uniwersalistyczne, 184
- modele nauczania
 - arbitralne, 184
 - zindywidualizowane, 184
- modele pamięci, 170
- modele przetwarzania języka naturalnego, 152
- modele systemu nerwowego, 168
- modem, 190
- moduł komunikacji w języku naturalnym, 215
- moduł komunikacji z bazami danych, 216
- moduły komunikacji
 - systemy online, 215
- modyfikatory, 102
- możliwy świat, 40
- myślenie potoczne, zob. wiedza potoczna

- nauka o prawdopodobieństwie, 27
- neurokomputering, 167
- neutralizacja znaku, 208
- nośnik informacji w systemie, 16

- obcinanie (truncation), 208
- odpowiedniość wyszukiwawcza, 92
- odsyłacz
 - całkowity, 98
 - orientacyjny, 98
 - porównawczy, 98
- systematyczny, 98
- uzupełniający, 98
- odwzorowanie rzeczywistości
 - mentalne, 53
- odwzorowanie wiedzy
 - nauki przyrodnicze, 183
 - nauki społeczne, 183
 - podjęcie przedmiotowe, 183
- określanie zakresu tematycznego zbiorów, 184
- określniki jhp
 - wiedza o rzeczywistości, 130
- opcje wyszukiwawcze
 - system online, 209
- operacje telekomunikacyjne, 189
- operacyjna interpretacja relewancji, 75
- operator systemu, 186
- operatory algebry Boole'a, 206
- operatory logiczne, 192
- operatory tekstowe, 192; 210
 - kontekstowe, 212
 - odległości, 212
 - wykładniki, 211
- organizacja planu wyrażania JIW, 95
- organizacja zbiorów informacji, 16
 - inwersyjna, 16
 - prosta, 16
- organizacyjna funkcja JIW, 185
- ostensja, 55
- ostensywne zachowanie, zob. komunikacja ostensywna

- paradygmatyka języków informacyjno-wyszukiwawczych, 103
- parser, 215
- pełny tekst, 210
- percepcja, 53
- perceptron, 168
- pertynencja, 77
 - definicja, 78
- plany językowe, 88
 - notacyjny, 88
 - treści dokumentów, 88
 - werbalny, 88
- podjęcie deklaratywne, 21
- podobieństwo treściowe dokumentów, 70
- podobieństwo znaczeniowe dokumentów, 72
- pojęcie, 57
- poła pełnotekstowe, 199
- poła wyszukiwawcze, 192; 194
 - bazy danych dokumentacyjne, 193
 - bazy danych faktograficzne, 193
- pole semantyczne

- Międzynarodowa Klasyfikacja Patentowa, 119
- struktura, 96
- struktura płaska, 96
- struktura polirelacyjna, 96
- pole semantyczne JIW
 - organizacja, 106
- poprawnie zbudowana formuła, 147
- porządek pionowy JIW, 91
- postkoordynacja, 18
- potrzeba informacyjna, 72; 78; 83
- potrzeby informacyjne, 183
 - podejście arbitralne, 184
 - podejście zindywidualizowane, 184
- powiązanie znaczeniowe, 48
- poziomy znaków JIW, 89
- prawda, 28
- prawdopodobieństwo, 28; 31
 - interpretacje logiczne, 31
 - interpretacje statystyczne, 31
 - logiczne, 34
- prawdopodobieństwo relewancji, 74
- prawdopodobne orzekanie, 31
- prawdziwość postrzeżeń, 28
- prawdziwość wiedzy, 145
- preferencje użytkownika, 80
- prekoordynacja, 18
- prezentacja znaczenia, 95; 96
- probabilistyczna interpretacja relewancji, 74
- problem komiwojażera, 169
- procedury decyzyjne, zob. formalne procedury dowodzenia
- procedury dedukcji, 223
 - generator funkcji, 224
 - podejście bezpośrednie, 223
 - systemy inteligentne, 223
 - zasada rezolucji, 223
- procedury interrogacyjne, 76; 191
- procesy inferencyjne człowieka, 79
- producent informacji, 186
- protokół konwersacji, 189
- prototyp obiektu, 219
- prymitywy semantyczne, zob. elementarne cechy semantyczne
- przekład automatyczny, 151
- przetwarzanie informacji
 - model kognitywny, 138
- przypadki głębokie Fillmore'a, 137
- psychologia kognitywna, 139
- pytanie, 178
 - rozstrzygnięcia, 178
 - uzupełnienia, 178
- pytanie informacyjne, zob. zapytania informacyjne
- rachunek entailment, 41; 45
- rachunek predykatów, 148
- rachunek predykatów pierwszego rzędu, 147
- rachunek zdań, 147
- rama
 - strukturalizacja wiedzy, 159
- rama modalna, 218
- ramy, 150; 163; 181
 - hierarchia dziedziczenia, 159
- ramy egzemplarza prototypu, 159
- rama prototypowa, 159
- ramy przypadków (case frames), 218
- reguły dedukcji naturalnej, 147
- reguły eliminacji, 58
- reguły heurystyki, 230
- reguły poprawności konceptualnej, 140
- reguły produkcji, 154
 - zastosowanie, 157
- reguły transformacji informacji, 143
- reguły wnioskowania
 - modus ponens, 223
 - skrypty, 162
- reguły wprowadzania, 58; 59
- rekursywne gramatyki sieciowe, 217
- relacja, 80
- relacja częściowego podobieństwa, 71
- relacja ekwiwalencji, 75
- relacja nieokreślona
 - Klasyfikacja Dwukropkowa, 128
- relacja podobieństwa, 29; 70; 71
- relacja podtrzymywania, 80
- relacja prototyp - egzemplarz prototypu, 159
- relacja równorzędności, 97
- relacja ukierunkowania
 - Klasyfikacja Dwukropkowa, 128
- relacja współrzędności, 97
- relacje
 - asocjacyjne, 91
 - hierarchiczne, 91
 - generyczne, 92
 - przeciwstawności znaczenia, 92
- relacje fazowe, 134
 - Klasyfikacja Dwukropkowa, 128
- relacje fazowe A.C.Fosketta, 138
- relacje paradygmatyczne
 - asocjacyjne, 93
 - hierarchiczne
 - hierarchii tematycznej, 92
 - mereologiczne, 92
 - kod W.Perry'ego i A.Kenta, 117

- krzyżowania, 91
- nadrzędności, 91
- podrzędności, 91
- przeciwstawności, 92
- równoważności, 91
- wykluczania, 91
- relacje syntagmatyczne
 - łączące, 102
 - kod W.Perry'ego i A.Kenta, 115
 - specyfikujące, 102
 - typologia, 101
- relacje sytuacyjne
 - instrumentalizacja, 153
 - kauzatywne, 153
 - Klasyfikacja Dwukropkowa, 128
 - Międzynarodowa Klasyfikacja Patentowa, 133
 - systemy relatorów UKD, 127
 - UKD, 141
- relatemy, 116
- relatory
 - Klasyfikacja Dwukropkowa, 128
- relewancja, 22; 26; 33; 38; 68
 - całkowicie pozytywna, 33
 - definicja, 62; 66; 78; 85
 - definicja Bradforda, 69
 - definicja Saracevica, 76
 - dla użytkownika, 75
 - ekstremalnie pozytywna, 33
 - ewidencyjna, 79
 - formalna, 69; 89
 - inżynierska, 69; 85
 - informacji, 77
 - interpretacja binarna, 69
 - interpretacyjna, 29; 79
 - logiczna, 38; 72; 80
 - motywacyjna, 29; 30
 - niewymuszona, 29
 - pomiar probabilistyczny, 75
 - praktyczna, 81
 - przedmiotowa, 29; 79; 104
 - kod W.Perry'ego i A.Kenta, 116
 - UKD, 106
 - psychologiczna, 80
 - semantyczna, 45
 - systemowa, 75
 - sytuacyjna, 79
 - bezpośrednia, 80
 - pośrednia, 80
 - techniczna, 69; 85
 - typologia interpretacji, 67
 - ujęcie socjologiczne, 26
 - w filozofii, 23
 - w informacji naukowej, 67
 - w logice, 31
 - warunkowo prawdopodobna, 76
 - wielowartościowa, 76
 - wymuszona, 29
- relewancja informacji
 - zasady, 82
- relewancja sytuacyjna
 - kod W.Perry'ego i A.Kenta, 114
- relewantny, zob. relewancja
- reprezentacja wiedzy
 - definicja, 87
- reprezentacja ideologii
 - skrypty, 224
- reprezentacja wiedzy, 87
 - arbitralna, 104
 - aspekt filozoficzny, 139
 - aspekt psychologiczny, 139
 - deklaratywna, 146
 - elementarne cechy semantyczne, 146
 - gramatyka JIW, 124
 - gramatyka UKD, 124
 - modularna, 146
 - o rzeczywistości, 220
 - ograniczenia, 146
 - proceduralna, 146; 154; 155; 156
 - reguły produkcji, 146
 - schematy, 145
 - sieci neuronowe, 185
 - sieci semantyczne, 146
 - systemy faktograficzne, 138
 - systemy inteligentne, 145
 - systemy logiczne, 146
 - UKD, 104
- reprezentacja wiedzy o rzeczywistości
 - kategoryzacja
 - systemy online, 214
- reprezentacja wiedzy sytuacyjnej
 - ramy, 158
 - skrypty, 158
- reprezentacja wypowiedzenia
 - mentalna, 52
 - semantyczna, 52
- role semantyczne
 - język RX-kodów, 137
 - kod SINTO, 111; 137
 - skrypty, 162
 - SYNTOL, 137
- rozpoznawanie sygnałów
 - sieci neuronowe, 170
- rozszerzone modele gramatyk sieciowych, 217
- rozumienie tekstów
 - system inteligentny, 228

rozumienie wypowiedzi, 63
 rzeczywistość dokumentacyjna, 90
 rzeczywistość pozadokumentacyjna, 90
 rzeczywistość pozajęzykowa, 20

słowniki JIW
 udostępnianie online, 200

scenariusze, 150
 sceny, 164
 schemat założenia, 55
 schematy kognitywne, 25
 scope note, 88
 selekcjonowanie informacji, 173; 191; 196
 chemicznej, 200
 selekcjonowanie informacji online
 cechy formalne, 207
 cechy treściowe, 204
 semantic primitives, zob. elementarne cechy semantyczne
 semantyka referencyjna, 151
 semantyka relacyjna, 42
 sfery relewancji, 26
 sfery znaczenia, 26
 siła założeń, 61
 wzmacnianie, 61
 sieci neuronowe, 167
 sieć Hopfielda, 172
 struktura, 169
 zastosowanie, 170
 sieci semantyczne, 148; 218
 baza notacyjna, 148; 163
 organizacja, 149
 rama sytuacyjna, 149
 relacje sytuacyjne, 149
 sieci teleinformatyczne, 214
 SIW
 struktura pamięci, 163
 skrypty, 160; 163; 224
 funkcjonowanie, 230
 role uczestników, 230
 socjologia życia codziennego, 26
 socjologia fenomenologiczna, 26
 spójnik entailment 41
 status wyszukiwawczy dokumentu, 74
 stereotypowe działanie, 219
 stereotypowa sytuacja, 219
 stop lista, 200; 211
 strategia wyszukiwawcza, 175
 definicja, 175; 176; 177
 formułowanie, 175; 191
 realizacja online, 207
 system dialogowy, 189
 streszczanie, 174
 struktura języka informacyjno-wyszukiwawczego, 88
 struktury monorelacyjna, 94
 struktura polirelacyjna, 94
 struktury pamięci
 dynamiczne, 163
 kieszenie, 164
 szczegółowe, 164
 uniwersalne, 164
 synonimia, 91
 synonimia wyszukiwawcza, 92
 system bibliograficzny, 17
 system DAI, 44
 system dialogowy, 17
 kryteria relewancji, 187
 system dokumentacyjny, 17; 20; 73
 system doradczy, 17
 system E, zob. rachunek entailment
 system ekspercki, 17; 215
 system ekspertowy, zob. system ekspercki
 system faktograficzny, 17; 20; 73
 system informacji, zob. system informacyjno-wyszukiwawczy
 dialogowy, 174
 kryteria relewancji, 192
 dokumentacyjnej, 14
 faktograficznej, 14
 tradycyjny, 174; 181
 system informacyjno-wyszukiwawczy, 16
 definicja, 13; 14
 system informacyjny, zob. system informacyjno-wyszukiwawczy
 system inteligentny, 215
 system interakcyjny, 17
 system jednodziedziny, 16
 system konwersacyjny, 17
 system małej mechanizacji, 16
 system manualny, 16
 system offline, zob. system wsadowy
 system online, zob. system dialogowy
 system PAI, 43
 system R, zob. logiki relewancji
 system S, 51
 system skomputeryzowany, 16
 system tradycyjny, 18
 system transdyscyplinowy, 16
 system wielodziedziny, 16
 system wsadowy, 17
 system wyszukiwania informacji, zob. system informacyjno-wyszukiwawczy
 system z bazą wiedzy, 215
 system z wiedzą, zob. system ekspercki
 systemy faktograficzne

- inteligentne, 165
- systemy informacji
 - dedukcyjne, 173
 - dialogowe, 186
 - JIW, 198
 - leksyka, 189
 - nie-dedukcyjne, 173
- systemy inteligentne
 - faktograficzne, 174
 - reguły wnioskowania, 157
- systemy online
 - ocena, 214
- systemy pełnotekstowe, 17
- systemy przetwarzające język naturalny, 217
- systemy relatorów, 102
- systemy relewancji, 30
- systemy rozumiejące język naturalny, 217
 - generacja druga, 217
 - generacja pierwsza, 217
- systemy teleinformatyczne, 192

- środowisko poznawcze człowieka, 53

- teoria elementarnych cech semantycznych
 - Wilksa, 219
- teoria indukcji i prawdopodobieństwa R. Carnapa, 31
- teoria konotacji składniowej, 219
- teoria mocnego wynikania, 51
- teoria pamięci
 - dynamicznej, 139
- teoria prototypów, 57
- teoria przypadków głębokich Fillmore'a, 218
- teoria pytań
 - Ajdukiewicza, 174; 178
 - Marciszewskiego, 174; 197
- teoria ram i skryptów, 146; 158
- teoria relewancji D. Sperbera i D. Wilson, 52
- teoria zależności konceptualnej R. Schanka, 152; 219
- terminologiczne banki danych, 201
- terminy, 116
- tezaurus
 - technologii preparatów enzymatycznych, 128
- transformacja
 - prawdziwościowa, 173
 - pytań badawczych, 180
 - semantyczna, 173
 - zapytań informacyjnych, 180
- transformacja informacji, 14
 - reguły heurystyki, 234
 - system inteligentny, 215; 224
 - wewnątrzsystemowa, 216
 - transformacja semantyczna, 21
 - tryb wyszukiwania informacji
 - interakcyjny, 175
 - nie-interakcyjny, 175
- użyteczność informacji, 72; 77
- użytkownik informacji, 77
- uczenie się sieci neuronowych, 170
- udostępnianie informacji, 173
 - bez transformacji, 173
 - z transformacją semantyczną, 173
- udostępnianie wiedzy
 - system inteligentny, 216
- UKD
 - wiedza o rzeczywistości językowej, 106
 - wiedza o rzeczywistości pozajęzykowej, 106
- Uniwersalna Klasyfikacja Dziesiętna, 104
- uporządkowanie formalno-semantyczne, 100
- uporządkowanie pomocnicze
 - wykładniki, 97

- wagi, zob. wskaźniki wagi
 - synaptyczne, 169
- wartość confirmacji, 60; 61
- wiarygodność postrzeżeń, 27
- wiedza
 - definicja, 138
 - system faktograficzny, 143
- wiedza deklaratywna, 222
 - transformacja, 162
- wiedza encyklopedyczna, 63; 143
- wiedza heurystyczna, 156
- wiedza jednostki, 78; 85
- wiedza językowa, 143
- wiedza ogólna, 85
- wiedza potoczna, 23; 24; 25; 26
 - struktura, 25
- wiedza proceduralna, 222
- wiedza socjokulturowa, 26
- wiedza społecznie akceptowana, 78; 85
- wiedza użytkownika, 80; 85
- wiedza zdroworozsądkowa, zob. wiedza potoczna
- wielowymiarowość pola semantycznego, 96
- wnioskowanie, 52
 - analityczne, 61
 - dedukcyjne, 222
 - domyślne, 155
 - formalne, 222
 - indukcyjne, 31
 - probabilistyczne, 79
 - przez analogię, 222

- przez domyślanie się (default reasoning), 222
- przez rozszerzoną aktywizację, 149
- przez uogólnienie, 222
- spontaniczne, 55; 59
 - dedukcyjne, 57
- sterowane celami, 227
- sterowane zadaniami (goal-directed), 224
- syntetyczne, 61
- wskaźnik pertynencji, 75
- wskaźnik relewancji, 75
- wskaźniki roli, 102
 - funkcja, 135
 - funkcja językowa, 136
 - funkcja pozajęzykowa, 136
 - kod Perry'ego i Kenta, 137
- wskaźniki wagi
 - funkcja, 136
- wskaźniki więzi, 135
 - funkcja, 135
- współczynniki wagowe, zob. wskaźniki wagi
- wyjaśnianie, 174
- wykładniki kategorii, 99
- wykładniki relacji syntagmatycznych
 - izolowane, 102
 - nieizolowane, 102
- wykładniki relewancji
 - kategorie semantyczne, 91
 - relacje paradygmatyczne, 91
- wykluczanie, 33
- wyrażenia
 - asemantyczne, 200
 - autosyntaktyczne, 94; 104
 - synsyntaktyczne, 94; 104
- wyszukiwanie informacji, 14
 - bez transformacji semantycznej, 185
 - dokumentacyjnej, 73; 79; 85
 - faktograficznej, 73; 85
 - model Swetsa, 75
 - wyszukiwanie kontrolowane
 - ocena, 213
 - wyszukiwanie pełnotekstowe, 17; 210
 - ocena, 213
 - warunki relewancji, 211
 - wyszukiwawcza funkcja, 185
- założenia pytań, 179
- założenie faktualne, 55
- założenie o rzeczywistości, 53
- zainteresowania użytkownika, 80
- zakres tematyczny zbiorów, 16
- zapytanie informacyjne, 80; 177
 - przekład, 174
 - transformacja, 178
- zasada akcja-reakcja, 174
- zasada dziedziczenia cech, 149
- zasada maksymalizacji przetwarzania informacji, 60
- zasada produktywności, 86
- zasady kooperacji konwersacyjnej, 54
- zasady użytecznego następstwa Ranganathana, 97
- zasięg systemu, 16
- zasób leksykalny JIW
 - udostępnianie online, 200
- zbiór wiedzy, 78
- zbiory informacyjne
 - meta-metainformacja, 192
- zdanie analityczne, 57
- zdanie syntetyczne, 57
- zdarzenia, 198
- związki znaczeniowe
 - wykładniki, 97

INDEKS NAZWISK

- Abelson R. P., 158
Ackerman W., 40
Addis T. R., 217
Ajdukiewicz K., 89; 174; 178
Anderson A. R., 38
Apresjan J. D., 151
Artowicz E., 120
Arystoteles, 31; 37; 52; 139
- Bar-Hillel Y., 70
Belnap N. D., 38
Bielicka L., 120; 126
Bobrov D. G., 162
Bojar B., 94; 110; 113; 117
Bookstein A., 75
Bradford S. C., 65; 67
Buchanan B., 103
- Carnap R., 22; 31; 56; 70; 71; 83; 85; 87; 90; 98
Chmielewska-Gorczyca E., 18
Chomsky N. A., 217
Choroś K., 136
Chryzopos, 27
Chryzyp z Soloi, 45
Cooper W. S., 72; 76; 79; 83; 92
CRG (Classification Research Group), 94
Cuadra C. A., 72; 76
Czerny A. I., 176
- Dao Thi Qui, 120; 121; 124; 136
Diodor Chronos, 45
Dunn J. M., 44
- Epstein R. L., 45
- Feyes T., 39
Fillmore Ch., 218
Fodor J. A., 57; 217; 218; 219
Foskett D. J., 78; 94; 97; 103; 120
Frege G., 34
- Gardin J. C., 110; 117
Goffman W., 75
Grice P., 53
Grollier de E., 94
- Hillman J. D., 69; 85; 103
Hintikka J., 34
Hjemslev L., 88
Hołówka T., 23
Husserl E., 26; 28
Hutchins W. J., 94; 120
- Jackendoff R., 140
James W., 26
- Kant I., 39; 43
Karneades z Cyreny, 27; 166
Katz J. J., 217; 218; 219
Kemp D. A., 78
Kent A., 110
Kervegant D., 127
Klejtomach, 27
Kochen M., 81
Kozakchov L. S., 69
Kripke S., 40; 42
Kuhn J. L., 74
Kyburg H., 34
- Lancaster F. W., 176
Langacker R., 143
Lewis D., 40
Lotka A. J., 69
- Manheim K., 25
Marciszewski W., 174; 178
Maron M. E., 74
McCulloch W. S., 167
Mc Dermott D., 87
Meadow Ch. T., 175
Meyer R. K., 42
Minsky M., 158
Moore G. E., 23; 38
Mykowiecka A., 219
- Nevill A., 75
- Ogórkiewicz W., 135; 178
- Parry W., 43
Peano G., 34

Perrault I., 127
Perry W., 110
Price D. J. de S., 69

Quillian R., 148; 218
Quine W. V., 40

Ranganathan S. R., 88; 94; 96; 97; 103; 106
Rosch E., 139
Routley R., 42
Russel B., 37; 40

Sadowska J., 129
Salton G., 77; 176
Saracevic T., 67; 73;
Saussure F. de, 52; 88
Schank R. C., 139; 152; 158; 164; 219; 224
Scheler M., 25
Schutz A., 22; 24; 26; 28; 30; 38; 81; 166
Sextus Empiryk, 27
Shannon C., 52
Skorochocko E. F., 110
Sosińska B., 88; 94; 95; 106; 110; 127
Sperber D., 52; 143; 166; 222
Swets J. A., 74

Ścibor E., 95; 120; 126

Tadeusiewicz R., 167
Tarski A., 37
Taube M. A., 77

Ungurian O., 97; 98; 106; 117; 120; 176
Urquart D. J., 69

Vickery B. C., 70; 94

Wereszczyńska-Cisło B., 178
Whorf B., 24
Wierzbicka A., 151
Wilks Y., 151
Wilson D., 52; 143; 166; 222
Wilson P., 79; 80
Winograd T., 163; 219
Wittgenstein L., 139
Wojtasiewicz O. A., 94; 110; 117
Wright M. von, 39

Zinowjew A., 50
Zipf G. K., 69

WYDAWNICTWO

SBP



Halo! Tu FO-KA

Drodzy Czytelnicy!

W ramach serii „Propozycje i Materiały” rozpoczęliśmy cykl wydawniczy, który będzie rozpoznawalny po symbolu graficznym FO-KA. To sympatyczne zwierzę będzie patronowało książkom poświęconym **FORMATOM, i KARTOTEKOM HASEŁ WZORCOWYCH**. Dzięki inicjatywie mgr Marii Burchard i decyzji dyrektora Biblioteki Uniwersytetu Warszawskiego dr. Henryka Hollendra staliśmy się wydawcami prac, które powstały w **CENTRUM FORMATÓW i KARTOTEK HASEŁ WZORCOWYCH** zlokalizowanym w Bibliotece Uniwersyteckiej.

Celem powołania **Centrum** jest utrzymanie spójnej bazy haseł wzorcowych oraz tworzenie zasad umożliwiających pełną wymianę danych bibliograficznych między bibliotekami wykorzystującymi zintegrowane systemy biblioteczne.

Kartoteki wzorcowe stanowią elementarny budulec do tworzenia pełnego rekordu bibliograficznego. Kooperatywne tworzenie kartotek wzorcowych w bazie centralnej zapewnia ich szybki przyrost, eliminuje dublowanie prac, gwarantuje spójność danych, a w konsekwencji jest podstawą do pełnej wymiany danych między poszczególnymi bibliotekami.

Już możecie Państwo kupić następujące książki:

A. Paluszkiewicz — „Struktura danych bibliograficznych w zintegrowanych systemach bibliotecznych”.

M. Lenartowicz, A. Paluszkiewicz — „Katalogowanie książek i wydawnictw ciągłych w formacie USMARC. Poradnik”.

M. Lenartowicz — „Format USMARC rekordu bibliograficznego dla książki”.

M. Burchard — „Format USMARC rekordu bibliograficznego dla druku muzycznego”.

W przygotowaniu: **„Mała KABA”**

Wyrażam nadzieję, że ta inicjatywa wydawnicza spotka się z życzliwym przyjęciem naszych Czytelników a szczególnie tych, którzy myślą o komputeryzacji bibliotek lub są w trakcie tego trudnego procesu.

Janusz Nowicki
Dyrektor Wydawnictwa SBP

Zamówienia prosimy kierować:

Dział Promocji i Kolportażu
ul. Hankiewicza 1, 02-103 Warszawa
lub

Wydawnictwo SBP ul. Konopczyńskiego 5/7, 00-953 Warszawa

WYDAWNICTWO
SBP



WYDAWNICTWO
SBP



oferuje

nową, bardzo ciekawą książkę prof.
Pauliny Buchwald-Pelcowej

CENZURA W DAWNEJ POLSCE **Między prasą drukarską a stosem.**

Książka ta jest owocem wieloletniej pracy badawczej Autorki. Obejmuje okres od panowania Zygmunta Starego do końca istnienia I Rzeczypospolitej. Oto tytuły rozdziałów:

- I. Edykty i ingerencje królewskie.
- II. Sprawy książek na Sejmach i Sejmikach.
- III. Działania urzędów i urzędników w dziedzinie cenzury. Jurysdykcja marszałkowska.
- IV. Orzeczenia Trybunału w sprawach książek i ich autorów.
- V. Cenzorskie poczynania władz miejskich.
- VI. Akademia Krakowska jako instytucja sprawująca cenzurę.
- VII. Kościół Katolicki wobec problemów cenzury książek.
- VIII. Nowe Kościoły reformowane a problemy cenzury.
- IX. Księgi nieprawomyślne i zabronione przez Kościół Prawosławny.
- X. Cenzura książek polskich za granicą i ingerencje zewnętrzne w Rzeczypospolitej.
- XI. Próba podsumowania.
Źródła i opracowania.

Książka zawiera indeksy a także jest bogato ilustrowana.

**JEST TO POZYCJA N I E Z B Ę D N A D L A P O Ł O N I S T Ó W , H I S T O R Y K Ó W ,
B I B L I O T E K A R Z Y O R A Z S T U D I U J Ą C E J M Ł O D Z I E Ż Y**

Stron 290, cena 20 zł.

Do nabycia:

Dział Promocji i Kolportażu SBP
02-103 Warszawa - Ochota
ul. S. K. Hankiewicza 1 pok 104
tel 822-43-45

Przyjmujemy zamówienia telefoniczne

Sprzedaż odręczna:

Warszawa, Hankiewicza 1 pok 104
Warszawa, Konopczyńskiego 5/7

WYDAWNICTWO

SBP



WYDAWNICTWO

SBP



SERIA

PROPOZYCJE I MATERIAŁY

dostarcza Państwu systematycznie materiały z najciekawszych konferencji poświęconych problemom nurtującym środowisko bibliotekarskie.
Czternasty tom, pt.

ZAWÓD BIBLIOTEKARZA DZIŚ I JUTRO

zawiera materiały z ogólnopolskiej konferencji zorganizowanej w październiku 1996 r. przez Sekcję Bibliotek Naukowych w Łodzi.

Książka zawiera referaty wybitnych przedstawicieli środowiska biblioteka naukowych. Ich rozważania poparte są wieloletnimi badaniami i obserwacjami naszego środowiska.

Autorzy:

- Jan JANIĄK, „*Bibliotekarze i biblioteki '96 w świetle ankiety*”
- Henryk HOLLENDER, „*Pragmatyka służbowa w bibliotekach naukowych w świetle ankiety*”
- Radosław CYBULSKI, Hanna KOLENDO, „*Bibliotekarze dyplomowani – frustracje, oczekiwania potrzeby*”
- Maria ŚLIWIŃSKA, „*Bibliotekarska elita i pariasi – przygotowanie do zadań i realizowane funkcje*”
- Jan WOŁOSZ, „*Bibliotekarz – zawód z jaką przyszłością?*”
- Bolesław HOWORKA, „*Status prawny bibliotekarzy w szkole wyższej*”
- Adam RUSEK, „*Zawód bibliotekarza okiem socjologa. Czy badania środowisk bibliotekarskich mogą przydać się praktykom – reformatorom?*”
- Małgorzata KŁOSSOWSKA, „*Na marginesie klasyfikacji zawodów*”

stron 120, cena 12 zł

Zamówienia prosimy kierować:

Dział Promocji i Kolportażu
ul. Hankiewicza 1, 02-103 Warszawa
lub

Wydawnictwo SBP ul. Konopczyńskiego 5/7, 00-953 Warszawa

ISBN 83-87629-01-4

**SERIA WYDAWANA
Z INICJATYWY
INSTYTUTU INFORMACJI NAUKOWEJ
I STUDIÓW BIBLIOLOGICZNYCH
UNIwersytetu warszawskiego
ORAZ WYDAWNICTWA SBP**