P. 3361/97

ZESZYTY
NAUKOWE
POLITECHNIKI
ŚLĄSKIEJ

WOJCIECH MOCZULSKI

METODY POZYSKIWANIA WIEDZY DLA POTRZEB DIAGNOSTYKI MASZYN

MONOGRAFIA

MECHANIKA
z. 130

GLIWICE
1997

POLITECHNIKA ŚLĄSKA
ZESZYTY NAUKOWE
Nr 1382

WOJCIECH MOCZULSKI

METODY POZYSKIWANIA WIEDZY DLA POTRZEB DIAGNOSTYKI MASZYN

MONOGRAFIA

P. 3361/97

Gliwice 1997

OPINIODAWCY
Prof. dr hab. inż. Józef Korbicz
Prof. dr hab. inż. Jerzy Wróbel

KOLEGIUM REDAKCYJNE

REDAKTOR NACZELNY – Prof. dr hab. Zygmunt Kleszczewski
REDAKTOR DZIAŁU – Dr. hab. inż. Andrzej Buchacz - Profesor Politechniki Śląskiej
SEKRETARZ REDAKCJI – Mgr Elżbieta Leśko

REDAKCJA
Mgr Anna Błażkiewicz

REDAKCJA TECHNICZNA
Alicja Nowacka

Wydano za zgodą
Rektora Politechniki Śląskiej

PL ISSN 0434 - 0817

Wydawnictwo Politechniki Śląskiej
ul. Kujawska 3, 44 - 100 Gliwice

Nakł. 300 + 53 **Ark. wyd. 15** **Ark. druk. 12,875** **Papier offset. kl. III 70 x 100, 80 g**
Oddano do druku i druk ukończ. w grudniu 1997 r.

Fotokopie, druk i oprawę wykonał „ROLEK", Gliwice, ul. Kazimierza Wielkiego 4

P 5/99

Mojej Żonie, Władysławie, oraz Julii i Michałowi
chcę wyrazić szczególną wdzięczność za ich pomoc,
cierpliwość i wsparcie, bez których nie byłoby
możliwe przeprowadzenie opisanych badań
oraz przygotowanie niniejszej rozprawy

Spis treści

Przedmowa ix

Wykaz ważniejszych oznaczeń xi

1 Wstęp 1
- 1.1 Cel rozprawy . . . 5
- 1.2 Tezy . . . 5
- 1.3 Zakres rozprawy . . . 5

2 Przedmiot badań 8
- 2.1 Obiekt diagnozowania . . . 8
  - 2.1.1 Klasy obiektów . . . 10
  - 2.1.2 Cechy oraz ich wartości . . . 11
  - 2.1.3 Czas . . . 14
  - 2.1.4 Stwierdzenia, stwierdzenia dynamiczne i historia obiektu . . . 16
  - 2.1.5 Model diagnostyczny obiektu . . . 17
  - 2.1.6 Specyficzne cechy modeli diagnozowanych obiektów . . . 26
- 2.2 Zadanie diagnostyki maszyn . . . 27
- 2.3 Środki wspomagania wnioskowania diagnostycznego . . . 32
  - 2.3.1 Drzewa uszkodzeń . . . 32
  - 2.3.2 Drzewa sprawdzeń . . . 33
  - 2.3.3 Tablice decyzyjne reguł . . . 33
  - 2.3.4 Systemy doradcze . . . 34
- 2.4 Dane i wiedza w diagnostyce maszyn . . . 36
  - 2.4.1 Dane w diagnostyce maszyn . . . 37
  - 2.4.2 Wiedza diagnostyczna . . . 39
- 2.5 Środki reprezentacji danych i wiedzy w diagnostyce maszyn . . . 40
  - 2.5.1 Opis dziedziny zastosowania . . . 41
  - 2.5.2 Reprezentacja danych . . . 43
  - 2.5.3 Reprezentacja wiedzy . . . 44
  - 2.5.4 Wnioski wynikające z przyjętego sposobu reprezentacji danych i wiedzy . . . 50
- 2.6 Proces pozyskiwania wiedzy diagnostycznej . . . 50
  - 2.6.1 Źródła wiedzy diagnostycznej . . . 51
  - 2.6.2 Pozyskiwanie wiedzy diagnostycznej od specjalistów . . . 51
  - 2.6.3 Pozyskiwanie wiedzy diagnostycznej z baz danych . . . 54
- 2.7 Podsumowanie dotyczące przedmiotu badań . . . 58

3 Metody pozyskiwania wiedzy diagnostycznej 60
3.1 Pozyskiwanie relacji empirycznych od specjalistów . . . 60
3.1.1 Empiryczna relacja diagnostyczna . . . 61
3.1.2 Pozyskiwanie wiedzy od specjalisty zgodnie z modelem 1. . . . 64
3.1.3 Pozyskiwanie wiedzy od specjalisty zgodnie z modelem 2. . . . 65
3.2 Pozyskiwanie wiedzy w sposób indukcyjny na podstawie przykładów . . . 67
3.2.1 System informacyjny i tablica decyzyjna . . . 68
3.2.2 Zbiór przykładów . . . 68
3.2.3 Stosowanie pozyskanej wiedzy do klasyfikowania nowych przykładów 73
3.2.4 Selektywna indukcja reguł poprzez generowanie pokryć . . . 74
3.2.5 Indukcja reguł z zastosowaniem zbiorów przybliżonych . . . 79
3.2.6 Indukcja drzew decyzyjnych . . . 86
3.2.7 Kryteria oceny przydatności metod "uczenia maszynowego" . . . 91
3.3 Ocena pozyskanej wiedzy . . . 92
3.3.1 Sposoby oceny wiedzy pozyskanej od specjalistów . . . 93
3.3.2 Sformalizowane sposoby oceny zawartości bazy wiedzy . . . 96
3.3.3 "Krzyżowe" stosowanie sposobów oceny pozyskanej wiedzy . . . 100
3.4 Scenariusz procesu pozyskiwania wiedzy . . . 101
3.4.1 Opracowanie koncepcyjne . . . 101
3.4.2 Opracowanie prototypu . . . 104
3.4.3 Opracowanie pełnej wersji bazy wiedzy . . . 105
3.4.4 Stosowanie bazy wiedzy i nadzór autorski . . . 107
3.5 Podsumowanie dotyczące metod pozyskiwania wiedzy diagnostycznej i jej oceny . . . 109

4 Środki wspomagania procesu pozyskiwania wiedzy diagnostycznej 111
4.1 Baza danych i wiedzy *EMPREL* . . . 111
4.2 Pozyskiwanie wiedzy od specjalistów . . . 113
4.2.1 Formularz "papierowy" do pozyskiwania nowej wiedzy . . . 114
4.2.2 Formularz "papierowy" do oceny wiedzy już pozyskanej . . . 115
4.2.3 Formularz "elektroniczny" . . . 116
4.3 Pozyskiwanie wiedzy w sposób indukcyjny . . . 118
4.3.1 Przygotowanie bazy przykładów . . . 119
4.3.2 Wybór cech relewantnych . . . 119
4.3.3 Selektywna indukcja reguł . . . 120
4.3.4 Indukcja reguł z zastosowaniem zbiorów przybliżonych . . . 120
4.3.5 Indukcja drzew decyzyjnych . . . 120
4.3.6 Ocena merytorycznej poprawności bazy wiedzy . . . 120
4.4 System pozyskiwania wiedzy diagnostycznej . . . 120
4.4.1 Podsystem baz danych i wiedzy systemu *SPWD1* . . . 120
4.4.2 Oprogramowanie do wspomagania pozyskiwania wiedzy od specjalistów . . . 122
4.4.3 Oprogramowanie do "uczenia maszynowego" metodami indukcyjnymi 123
4.5 Podsumowanie dotyczące środków wspomagania procesu pozyskiwania wiedzy . . . 124

5 Badania proponowanych metod i środków pozyskiwania wiedzy 125
5.1 Badania dotyczące wyników czynnego eksperymentu diagnostycznego . . . 125
5.1.1 Charakterystyka zadania diagnostycznego . . . 126
5.1.2 Przebieg procesu pozyskiwania wiedzy . . . 126
5.1.3 Podsumowanie badań przeprowadzonych dla bazy wiedzy *RK* . . . 127
5.2 Badania przeprowadzone w ramach eksperymentu numerycznego . . . 127
5.2.1 Charakterystyka zadania diagnostycznego . . . 128
5.2.2 Przebieg procesu pozyskiwania wiedzy . . . 128
5.2.3 Podsumowanie badań przeprowadzonych dla bazy wiedzy *GD* . . . 129
5.3 Weryfikacja systemu pozyskiwania wiedzy diagnostycznej *SPWD1* . . . 130
5.4 Podsumowanie przeprowadzonych badań . . . 131

6 Podsumowanie i wnioski 133

DODATKI 139

A Ogólny opis schematu logicznego *EMPREL* 139
A.1 Opis dziedziny . . . 139
A.2 Baza danych . . . 144
A.3 Baza wiedzy . . . 146

B Sposób prowadzenia dialogu z użytkownikiem formularza "elektronicznego" 149

C Opis badań dotyczących wyników czynnego eksperymentu diagnostycznego 153
C.1 Identyfikacja zadania diagnostycznego . . . 153
C.2 Określenie projektu bazy wiedzy *RK* . . . 156
C.3 Opracowanie wersji pilotażowej bazy wiedzy oraz weryfikacja koncepcji . . . 157
C.4 Opracowanie pełnej wersji bazy wiedzy *RK* i jej weryfikacja . . . 160

D Opis badań przeprowadzonych w ramach eksperymentu numerycznego 173
D.1 Identyfikacja zadania diagnostycznego . . . 173
D.2 Określenie projektu bazy wiedzy *GD* . . . 176
D.3 Opracowanie wersji pilotażowej bazy wiedzy oraz weryfikacja koncepcji . . . 176
D.4 Opracowanie pełnej wersji bazy wiedzy *GD* i jej weryfikacja . . . 180

Spis literatury 188

Streszczenie 204

Contents

Preface ix

Table of Important Symbols xi

1 Introduction 1
- 1.1 The Goal of the Dissertation . . . 5
- 1.2 Theses . . . 5
- 1.3 The Scope of the Dissertation . . . 5

2 The Subject of Research 8
- 2.1 The Object of Diagnosing . . . 8
  - 2.1.1 Classes of Objects . . . 10
  - 2.1.2 Attributes and Their Values . . . 11
  - 2.1.3 Time . . . 14
  - 2.1.4 Statements, Dynamic Statements and History of an Object . . . 16
  - 2.1.5 Diagnostic Model of an Object . . . 17
  - 2.1.6 Specific Features of Models of Diagnosed Objects . . . 26
- 2.2 The Task of Machinery Diagnostics . . . 27
- 2.3 Means to Aid the Diagnostic Inferring . . . 32
  - 2.3.1 Fault Trees . . . 32
  - 2.3.2 Test Trees . . . 33
  - 2.3.3 Decision Tables of Rules . . . 33
  - 2.3.4 Expert Systems . . . 34
- 2.4 Data and Knowledge in Machinery Diagnostics . . . 36
  - 2.4.1 Data in Machinery Diagnostics . . . 37
  - 2.4.2 Diagnostic Knowledge . . . 39
- 2.5 Means of Representing Data and Knowledge in Machinery Diagnostics . . . 40
  - 2.5.1 Description of the Domain of Applications . . . 41
  - 2.5.2 Data Representation . . . 43
  - 2.5.3 Knowledge Representation . . . 44
  - 2.5.4 Conclusions Resulting from the Applied Way of Data and Knowledge Representation . . . 50
- 2.6 The Process of Diagnostic Knowledge Acquisition . . . 50
  - 2.6.1 Diagnostic Knowledge Sources . . . 51
  - 2.6.2 Acquisition of Diagnostic Knowledge from Experts . . . 51
  - 2.6.3 Acquisition of Diagnostic Knowledge from Databases . . . 54
- 2.7 Recapitulation Concerning the Subject of Research . . . 58

3 Methods of Acquisition of Diagnostic Knowledge 60
- 3.1 Acquisition of Empirical Relationships from Experts . . . 60
  - 3.1.1 The Empirical Diagnostic Relationship . . . 61
  - 3.1.2 Acquisition of Knowledge from an Expert in Accordance with the 1st Model . . . 64
  - 3.1.3 Acquisition of Knowledge from an Expert in Accordance with the 2nd Model . . . 65
- 3.2 Acquisition of Knowledge in an Inductive Way Based on Examples . . . 67
  - 3.2.1 Information System and Decision Table . . . 68
  - 3.2.2 The Set of Examples . . . 68
  - 3.2.3 The Application of the Acquired Knowledge to the Classification of New Examples . . . 73
  - 3.2.4 Selective Induction of Rules Through the Generation of Covers . . . 74
  - 3.2.5 Induction of Rules with the Application of Rough Sets . . . 79
  - 3.2.6 Induction of Decision Trees . . . 86
  - 3.2.7 Criteria of Estimation of the Usefulness of Machine Learning Methods 91
- 3.3 Assessment of the Acquired Knowledge . . . 92
  - 3.3.1 Ways of Assessing Knowledge Acquired from Experts . . . 93
  - 3.3.2 Formalized Ways of Assessment of the Knowledge Base Contents . . . 96
  - 3.3.3 "Cross-wise" Application of Ways of Assessing of the Acquired Knowledge . . . 100
- 3.4 The Scenario of the Knowledge Acquisition Process . . . 101
  - 3.4.1 Elaboration of the Concept . . . 101
  - 3.4.2 Elaboration of the Prototype . . . 104
  - 3.4.3 Elaboration of the Complete Version of the Knowledge Base . . . 105
  - 3.4.4 Application of the Knowledge Base and Supervision by the Authors 107
- 3.5 Recapitulation Concerning Knowledge Acquisition and Assessment Methods 109

4 Means of Aiding of the Diagnostic Knowledge Acquisition Process 111
- 4.1 The *EMPREL* Data and Knowledge Base . . . 111
- 4.2 Acquisition of Knowledge from Experts . . . 113
  - 4.2.1 The "Paper" Form for Acquisition of New Knowledge . . . 114
  - 4.2.2 The "Paper" Form for Assessment of Previously Acquired Knowledge 115
  - 4.2.3 The "Electronic" Form . . . 116
- 4.3 Acquisition of Knowledge by Using Induction . . . 118
  - 4.3.1 The Preparation of the Base of Examples . . . 119
  - 4.3.2 The Selection of Relevant Attributes . . . 119
  - 4.3.3 The Selective Induction of Rules . . . 120
  - 4.3.4 The Induction of Rules Using Rough Sets . . . 120
  - 4.3.5 The Induction of Decision Trees . . . 120
  - 4.3.6 Assessment of the Substantial Correctness of the Knowledge Base . 120
- 4.4 The System of Diagnostic Knowledge Acquisition . . . 120
  - 4.4.1 The Sybsystem of Data and Knowledge Bases of the *SPWD1* System 120
  - 4.4.2 The Software to Aid the Acquisition of Knowledge from Experts . . . 122
  - 4.4.3 The Software for "Machine Learning" by Inductive Methods . . . 123
- 4.5 Recapitulation Concerning Ways of Aiding of the Knowledge Acquisition Process . . . 124

5 Investigations of Proposed Methods and Means of Knowledge Acquisition 125
5.1 Investigations Concerning Results of the Active Diagnostic Experiment . . 125
5.1.1 Description of the Diagnostic Task 126
5.1.2 The Course of Knowledge Acquisition Process 126
5.1.3 Recapitulation of the Verification Investigations for the *RK* Knowledge Base 127
5.2 Investigations Carried-Out Within the Numerical Experiment 127
5.2.1 Description of the Diagnostic Task 128
5.2.2 The Course of Knowledge Acquisition Process 128
5.2.3 Recapitulation of the Verification Investigations for the *GD* Knowledge Base 129
5.3 Verification of the *SPWD1* System of Diagnostic Knowledge Acquisition . . 130
5.4 Recapitulation of the Carried-Out Investigations 131

6 Recapitulation and Conclusions 133

APPENDICES 139

A General Description of the *EMPREL* Logical Scheme 139
A.1 Description of the Domain 139
A.2 Data Base 144
A.3 Knowledge Base 146

B The User Interface of the "Electronic" Form 149

C Description of Investigations Concerning Results of the Active Diagnostic Experiment 153
C.1 Identification of the Diagnostic Task 153
C.2 Determination of the Design of the *RK* Knowledge Base 156
C.3 Elaboration of the Pilot Version of the Knowledge Base and Verification of the Concept 157
C.4 Elaboration of the Complete Version of the *RK* Knowledge Base and Its Verification 160

D Description of Investigations Carried-Out on the Ground of the Numerical Experiment 173
D.1 Identification of the Diagnostic Task 173
D.2 Determination of the Design of the *GD* Knowledge Base 176
D.3 Elaboration of the Pilot Version of the Knowledge Base and Verification of the Concept 176
D.4 Elaboration of the Complete Version of the *GD* Knowledge Base and Its Verification 180

References 188

Abstract 204

Przedmowa

Spośród wielu dziedzin ogólnie rozumianej eksploatacji maszyn *diagnostyka techniczna maszyn* jest jedną z tych, które podlegają burzliwemu rozwojowi. Pojawiają się nowe problemy badawcze oraz powstają metody, umożliwiające wspomaganie wnioskowania diagnostycznego. Szczególną rolę odgrywają tutaj *systemy doradcze*, wspomagające diagnostów w rozwiązywaniu wielu zadań diagnostycznych. Skuteczność działania systemów doradczych uwarunkowana jest zawartością ich *baz wiedzy*.

Przedmiotem pracy są *metody pozyskiwania wiedzy*, szczególnie przydatne dla potrzeb diagnostyki technicznej. Wybór tych metod był uwarunkowany specyfiką danych dotyczących diagnozowanych obiektów oraz wiedzy, która umożliwia uzyskanie diagnozy, dotyczącej stanu technicznego tych obiektów. Metody te były podstawą do opracowania *środków wspomagania procesu pozyskiwania wiedzy diagnostycznej*. Przeprowadzono także *badania proponowanych metod i środków pozyskiwania wiedzy*, których celem była weryfikacja niektórych z nich oraz potwierdzenie ich przydatności w procesie pozyskiwania wiedzy dotyczącej wybranych zadań diagnostycznych.

Praca posiada znaczenie teoretyczne ze względu na kompleksowe ujęcie problemów *metodologii pozyskiwania wiedzy diagnostycznej*. Ma ona także znaczenie praktyczne, ponieważ zawiera *opis scenariusza procesu pozyskiwania wiedzy* oraz obszernie udokumentowane *opisy przebiegu badań zgodnych z tym scenariuszem*, co może mieć znaczenie w wypadku planowania działań, związanych z pozyskiwaniem wiedzy diagnostycznej. Opisy te obejmują także pozyskiwanie wiedzy na podstawie przykładów, otrzymanych w wyniku przeprowadzenia czynnego eksperymentu diagnostycznego i eksperymentu numerycznego.

Przeprowadzenie znaczącej części opisanych badań było wspierane ze środków Komitetu Badań Naukowych w formie projektu badawczego nr 8 T11F 020 09 pt. *Akwizycja wiedzy dla potrzeb diagnostyki technicznej*.

Chciałbym wyrazić głęboką wdzięczność Profesorowi Wojciechowi Cholewie z Katedry Podstaw Konstrukcji Maszyn Politechniki Śląskiej — za umożliwienie mi samodzielnego podjęcia badań nad metodologią pozyskiwania wiedzy diagnostycznej i konstrukcyjnej oraz powierzenie obowiązków kierowania zespołem kolegów, współwykonujących te prace, a także za krytyczne uwagi dotyczące prowadzonych przeze mnie badań.

Na obecną postać pracy wpłynęły uwagi krytyczne, poczynione przez Profesora Józefa Korbicza i Profesora Jerzego Wróbla, którzy szczegółowo zapoznali się z maszynopisem, za co składam Im szczególne podziękowanie.

Chcę także gorąco podziękować za nieodpłatne udostępnienie oprogramowania, które było stosowane do przeprowadzonych badań:

- Prof. Jerzemu W. Grzymała-Busse, University of Kansas, KS, USA, za udostępnienie programu *LERS* do pozyskiwania wiedzy metodami zbiorów przybliżonych oraz przeprowadzenie wybranych obliczeń z wykorzystaniem dostarczonych Mu danych.

- Prof. Janowi Kicińskiemu, Instytut Maszyn Przepływowych PAN, za udostępnienie oprogramowania symulacyjnego *KINWIR-D* oraz *DYNWIR-W*, które umożliwiło przeprowadzenie eksperymentu numerycznego.

- Prof. Ryszardowi S. Michalskiemu, George Mason University, VA, USA, za udostępnienie oprogramowania do indukcji reguł z zastosowaniem pokryć (programy *AQ15c* oraz *AQ17-HCI*).

- Prof. Janowi M. Żytkowowi, University of North Carolina, NC, USA, za wspólne przeprowadzenie wstępnych badań nad możliwością odkrywania wiedzy w bazach danych diagnostycznych oraz udostępnienie w tym celu systemu programów *49er/Bacon-3*.

Przeprowadzenie badań dotyczących zastosowania opisanych metod oraz środków pozyskiwania wiedzy nie byłoby możliwe bez bardzo aktywnego współudziału Koleżanek i Kolegów z Katedry Podstaw Konstrukcji Maszyn Politechniki Śląskiej, a zwłaszcza: Krzysztofa Ciupke, Pawła Kostki, Piotra Maniaka, Anny Timofiejczuk, Ryszarda Wyczółkowskiego, Marka Wyleżoła i Bogdana Wysogląda. Pani Ewa Opoka przeprowadziła adiustację tekstu pracy, za co składam Jej serdeczne podziękowanie.

Grudzień 1997 Wojciech A. Moczulski

Wykaz ważniejszych oznaczeń

$a, a_j, atr(\,)$	cecha obiektu
a_o	zdyskretyzowana cecha obiektu
$A, A(C), A_{rel}$	zbiór cech; zbiór cech dla klasy C; podzbiór cech relewantnych
A_{ij}	częstość z próby
$B(\,), B_{ex}(\,)$	wartość stopnia przekonania o słuszności reguły
$B, \bar{B}$	podzbiór zbioru cech; relacja nierozróżnialności (przez podzbiór cech B)
$\underline{B}Y, \overline{B}Y$	B-dolna i B-górna aproksymacja zbioru Y przez podzbiór cech B
$Bn_B(Y)$	B-ograniczenie zbioru Y przez podzbiór cech B
$card(\,)$	moc (liczba kardynalna) zbioru
$class(\,)$	funkcja zwracająca identyfikator klasy danego obiektu
$c_{n,k}$	ocena stopnia pewności reguły
c_k	klasyfikator
C_i, C_k	pojęcie (koncept); klasa obiektów
d, d_i	listy identyfikatorów stanów technicznych
$D, Dom(\,)$	dziedzina cechy
D	zbiór cech decyzyjnych w systemie informacyjnym $\mathcal{S}$
$D(r), D(\Re)$	dziedzina reguły; dziedzina empirycznej relacji diagnostycznej
$Des(\,)$	opis formuły (warunkowej lub decyzyjnej)
∂_S	uogólniona decyzja w systemie informacyjnym $\mathcal{S}$
e, e_1, e_2	przykłady
$entier(\,)$	funkcja zwracająca część całkowitą swojego argumentu
E, E^u, E^t	zbiór przykładów, podzbiory przykładów uczących i testowych
$E_k, E_k^+, E_{k,j}^-$	podzbiory zbioru przykładów pozytywnych dla klasy "k", przykładów błędnie pominiętych oraz przykładów niesłusznie zaliczonych
E_s	relacja równoważności stanów obiektu
$E^{(m)}$	podzbiór przykładów, odpowiadający m-tej wartości cechy
E_{ij}	częstość oczekiwana
f	funkcja (przekształcenie)
f_{max}	maksymalna częstotliwość składowej sygnału (dla cechy ilościowej)
f_s	częstotliwość próbkowania sygnału (dla cechy ilościowej)
$head(\,)$	funkcja zwracająca pierwszy element listy (tzw. głowę listy)
H	historia eksploatacji obiektu (zapis zdarzeń i procesów związanych z użytkowaniem lub obsługiwaniem obiektu)
$id(\,)$	funkcja zwracająca identyfikator stanu technicznego
i, j, k	indeksy
I, J, K	zbiory indeksów (także nieprzeliczalne)
$M(a)$	zbiór indeksów; liczba wartości cechy a

n, n_k	prędkość obrotowa; prędkość krytyczna
$n_{err}, n_{err\|k}$	liczba przykładów błędnie sklasyfikowanych (łącznie i dla klasy k)
$nam(\,)$	nazwa (funkcja zwracająca nazwę swojego argumentu wywołania)
N, N_1, N_2	liczby całkowite
$N(\,)$	miara stopnia konieczności stwierdzenia lub reguły
$NP(\,)$	stopień prawdziwości stwierdzenia, stopień przekonania o słuszności reguły
o, o_j	obiekty
O, O_k	zbiór obiektów; podzbiory zbioru obiektów
p	stwierdzenie, formuła, przesłanka reguły
p_I	stwierdzenie dynamiczne
$p_{k,n}$	ocena prawdopodobieństwa warunkowego
p_S	znaczenie formuły p w systemie informacyjnym $\mathcal{S}$
$P(\,)$	wartość stopnia możliwości (stwierdzenia lub reguły)
$P(a, E)$	miara przydatności cechy a do rozróżniania klas w zbiorze przykładów E
q	konkluzja reguły
R	zbiór liczb rzeczywistych; system obiektu
R, R_k	hipoteza (zbiór reguł odpowiadających danemu pojęciu)
$\Re$	empiryczna relacja diagnostyczna
s_j	element stanu obiektu (abstrakcyjnego lub konkretnego)
$S(\,)$	stopień dostateczności (reguły)
S	zbiór stanów układu; przestrzeń stanów
S_1	zbiór stanów (nie zmieniających się w czasie) związanych z konstrukcją maszyny
S_2	zbiór stanów wywołanych procesem wytwarzania lub użytkowania maszyny
$\mathcal{S}$	system informacyjny
t	czas "mikro"
$tail(\,)$	funkcja zwracająca listę, z której usunięto jej pierwszy element (tzw. ogon listy)
$T(\,)$	stopień prawdziwości
$T(\Rightarrow), T(\Leftarrow)$	stopień prawdziwości implikacji "wprost" i implikacji odwrotnej
$U_1(\,), U_2(\,)$	pomocnicze funkcje schodkowe
$v, v', v_0, \ldots, v_k$	wartości cechy; liczby określające początek i koniec przedziału wartości cechy ilościowej
$val(\,)$	wartość (funkcja zwracająca wartość swojego argumentu)
$V, V(a)$	zbiór wartości wszystkich cech; zbiór wartości cechy a
$w_{ex}, w_k, w_{mit\Sigma}, w(\,)$	wagi
x_i	element wejścia obiektu (abstrakcyjnego lub konkretnego)
X	zbiór wejść obiektu
X_1	zbiór wejść diagnozowanego obiektu związanych z jego użytecznym działaniem
X_2	zbiór wejść określających warunki działania diagnozowanego obiektu
X_3	zbiór wejść interpretowanych jako szum wpływający na zmniejszenie stopnia prawdziwości relacji diagnostycznych

y_k	element wyjścia obiektu (abstrakcyjnego lub konkretnego)
Y	zbiór wyjść obiektu
Y_1	zbiór wyjść związanych z użytecznym działaniem obiektu
Y_2	zbiór wyjść związanych z występowaniem procesów resztkowych
Y_3	zbiór wyjść nie znanych lub nie obserwowanych, które są skutkiem działania maszyny
α	stała
$\alpha_B(Y)$	dokładność aproksymacji zbioru Y przez podzbiór cech B
β_1, β_2	stałe
$\gamma_B(\Sigma)$	ocena jakości klasyfikacji Σ za pomocą zbioru cech B
ΔT	odstęp próbkowania cechy ilościowej
$\Delta\theta$	długość przedziału czasu eksploatacji
$\epsilon_{cm}, \epsilon^w_{cm}$	względny i względny ważony błąd niesłusznego zaliczenia
$\epsilon_{om}, \epsilon^w_{om}$	względny i względny ważony błąd pominięcia
$\epsilon_{ov}, \epsilon^w_{ov}$	łączny i łączny ważony błąd względny
η_{ov}	łączna sprawność klasyfikatora
θ	czas "makro"; zbiór wartości uogólnionej decyzji
$\theta_0, \theta_1, \theta_2, \theta_3$	punkty czasu "makro"
Θ	przedział czasu eksploatacji; zbiór punktów czasu "makro"
π_a	zbiór par określający przekształcenie cechy ilościowej a w cechę jakościową
ρ	funkcja odpowiedzi obiektu
Σ	klasyfikacja zbioru przykładów
ϕ	funkcja zmiany stanu obiektu
χ^2	statystyka chi-kwadrat
$\propto$	oznaczenie dowolnego dwuargumentowego operatora relacji
$\approx$	operator częściowego dopasowania przesłanki reguły do przykładu
$\simeq$	operator pełnego dopasowania przesłanki reguły do przykładu
$\prec$	operator poprzedzania
$\leftarrow$	operator przypisania
$\overline{1..n}$	odcinek liczb naturalnych (podzbiór wszystkich liczb k spełniających $1 \leqslant k \leqslant n$)

Rozdział 1

Wstęp

Podstawowe zadanie diagnostyki maszyn polega na *diagnozowaniu*, a więc *rozpoznaniu lub określeniu stanu technicznego pewnego układu technicznego (np. maszyny), które następuje w warunkach ograniczonej ilości informacji* [Birger, 1978]. Istotą tego działania jest *wnioskowanie* (lub *rozumowanie*) *diagnostyczne*, które można skutecznie przeprowadzić, jeśli:

- *zostały zgromadzone odpowiednie dane, stanowiące przesłanki w prowadzonym rozumowaniu*; są to między innymi dane opisujące *wejścia obiektu* (zwłaszcza jego *warunki działania*) oraz wyjścia obiektu (wśród których należy wyróżnić wyjścia użyteczne, jak produkty lub użyteczne działanie obiektu, a także tzw. procesy resztkowe [Cempel, 1982], które są niezamierzonym skutkiem działania tego obiektu), dane dotyczące historii eksploatacji obiektu itp.,
- *znane są możliwe stany techniczne*, które są swoiste dla diagnozowanego obiektu,
- *podmiot (układ) diagnozujący dysponuje wystarczającą wiedzą*, aby na podstawie zgromadzonych dowodów wyciągnąć poprawne wnioski, odnoszące się do stanu technicznego obiektu.

Przedstawione powyżej niektóre warunki skutecznego wnioskowania diagnostycznego pozwalają na stwierdzenie, że podmiot diagnozujący winien "wiedzieć", *jakie dane i w jakiej postaci* należy gromadzić, oraz "znać" *możliwy zbiór stanów, symptomy tych stanów*, a także *reguły określające związki pomiędzy przesłankami i konkluzjami.*

Współczesne układy pomiarowe, np. stosowane do stałego nadzoru (monitorowania) tzw. maszyn "krytycznych", umożliwiają gromadzenie bardzo licznych zbiorów danych o *wyjściach* maszyny (wśród których poszukujemy symptomów stanów technicznych) oraz zwykle licznych zbiorów danych o *wejściach* tej maszyny. Skomputeryzowane układy monitorujące dostarczają takich zbiorów wyników okresowo, np. co kilka sekund [Cholewa, 1996]. Aby można było dokonywać oceny stanu monitorowanego obiektu w sposób "dynamiczny" (np. w rzeczywistej skali czasu), należy wnioskowanie diagnostyczne wspomagać komputerowo. W tym celu mogą być stosowane systemy doradcze [Cholewa, Pedrycz, 1987], których podstawowym elementem są bazy wiedzy. W bazach wiedzy gromadzona jest wiedza diagnostyczna pochodząca z wielu dostępnych źródeł, przy czym stosowane są różne metody jej reprezentacji. *Zawartość baz wiedzy w zasadniczy sposób determinuje jakość działania diagnostycznego systemu doradczego, a zwłaszcza poprawność osiąganych przez niego wniosków.*

Wymaganie znacznej szybkości działania nie jest zresztą jedyną racją uzasadniającą komputerowe wspomaganie wnioskowania diagnostycznego, a jednocześnie celowość gromadzenia wiedzy diagnostycznej pochodzącej z różnych źródeł. Diagnostyczne systemy doradcze umożliwiają stosowanie wiedzy pozyskanej od wielu specjalistów. Ponadto w pewnym stopniu "zastępują" nawet najlepszych specjalistów, którzy z wielu względów nie mogliby osobiście rozwiązać danego problemu diagnostycznego. Znane są wyniki badań, w których potwierdzono bardzo wysoką skuteczność wnioskowania systemów doradczych, porównywalną ze skutecznością działania człowieka [Harmon, King, 1985].

Praca dotyczy *pozyskiwania wiedzy dla potrzeb diagnostyki maszyn*. Termin *wiedza* wymaga tutaj wyjaśnienia. W odniesieniu do człowieka *wiedza* to ogół tego, co człowiek wie. *Wiedza specjalisty-diagnosty* (w określonej dziedzinie diagnostyki technicznej) dotyczy obiektów (np. maszyn i ich zespołów) i klas obiektów z tej dziedziny, taksonomii klas obiektów, ich własności i właściwości, związków pomiędzy obiektami i pomiędzy klasami obiektów, obejmuje także umiejętności, zrozumienie praw ogólnych, procedury postępowania diagnostycznego itp. *Wiedza zdobywana jest zwłaszcza poprzez uczenie (się) danej osoby*, np. w ramach studiów lub specjalistycznego szkolenia (odbywanych pod kierunkiem szeroko rozumianego "nauczyciela"). *Może być także gromadzona poprzez nabywanie doświadczenia i umiejętności* (najczęściej samodzielnie, co ma znamiona odkrywania) — wówczas jest wynikiem zwykle długotrwałej aktywności człowieka w pewnej dziedzinie. W dalszej części pracy termin "wiedza" będzie oznaczał zarówno wiedzę, jak i praktyczne doświadczenie specjalisty.

Należy bardzo mocno podkreślić, że pierwotnym źródłem wiedzy diagnostycznej, którego udziału nie da się pominąć w procesie tworzenia bazy wiedzy, jest człowiek-specjalista.

Część tak rozumianej wiedzy diagnostycznej (zwłaszcza elementy opisu dziedziny zastosowania) może być reprezentowana w sposób *deklaratywny*, a część w sposób *proceduralny*. W opinii autora, podstawowym źródłem wiedzy diagnostycznej reprezentowanej w sposób proceduralny jest człowiek (bezpośrednio lub pośrednio — np. poprzez przedstawienie opisu procedury postępowania diagnostycznego w jakiejś publikacji).

Zadanie innego rodzaju stanowi *pozyskiwanie wiedzy diagnostycznej reprezentowanej deklaratywnie*, która w dziedzinie diagnostyki maszyn może obejmować: klasy obiektów, relacje zachodzące pomiędzy klasami, obiekty i relacje zachodzące pomiędzy obiektami, cechy obiektów, ich rodzaje i przyjmowane przez nie wartości, sposoby przekształcania cech ilościowych w cechy jakościowe, istotne zdarzenia eksploatacyjne, stany techniczne elementarne, klasy stanów technicznych, wartości krytyczne cech stanu i symptomów diagnostycznych, reguły wnioskowania, objaśnienia dotyczące wszystkich zgromadzonych pojęć itp. Wiedza ta *może być pozyskana od specjalistów*, a także *może zostać w pewnej części wyprowadzona z odpowiednio przygotowanych baz*, w których są gromadzone dane dotyczące eksploatacji maszyn (a zwłaszcza obserwacji diagnostycznych, dotyczących tych maszyn), wyniki obliczeń symulacyjnych lub przykłady przygotowane przez zespół specjalistów.

W postępowaniu diagnostycznym zwykle mamy do czynienia z licznymi zbiorami danych. Dane, gromadzone podczas obserwacji obiektów diagnozowania, są zwykle niekompletne, sprzeczne, obciążone niepewnością, niedokładne lub mają charakter jakościowy. Wyniki wnioskowania diagnostycznego są pozornie słuszne, choć niepewne. Także wiedza diagnostyczna ma często charakter przybliżony (ang. *approximate*), czasem nawet zdroworozsądkowy.

W pracy autor skoncentrował się na zagadnieniach *pozyskiwania wiedzy reprezentowanej w sposób deklaratywny* i postanowił dla potrzeb diagnostyki maszyn opracować system pozyskiwania wiedzy, w którym możliwe będzie wykorzystanie różnych źródeł oraz odpowiadających im metod pozyskiwania wiedzy. Przyjęto, że pozyskiwana będzie wiedza dotycząca pewnej dobrze określonej dziedziny diagnostycznej (tzw. *dziedziny zastosowania*; przykładem takiej dziedziny może być diagnostyka maszyn wirnikowych). Ważne zadania w tym systemie spełnia tzw. *programista bazy wiedzy*[1], wspomagany w swoich działaniach przez różne środki, takie jak zestawy specjalnych formularzy, specjalizowane oprogramowanie komputerowe itp. Środki te działają zgodnie z opracowanym zbiorem metod pozyskiwania wiedzy. Podczas opracowywania środków wspomagania procesu pozyskiwania wiedzy postanowiono dążyć do minimalizowania udziału programisty bazy wiedzy w zapisywaniu zawartości tej bazy i jej dalszym pielęgnowaniu. Pozyskana wiedza reprezentowana jest we wspólnym, zunifikowanym formacie. Przedstawione ujęcie pozwala na połączenie wiedzy diagnostycznej pochodzącej z różnych źródeł. Podstawowa wiedza, warunkująca zapoczątkowanie działania systemu pozyskiwania wiedzy, pochodzi od specjalisty (lub specjalistów) i przede wszystkim umożliwia opis najważniejszych pojęć, definiujących dziedzinę zastosowania. Ze względu na małą efektywność pozyskiwania wiedzy od specjalistów (którą można określać co najwyżej w dziesiątkach reguł zapisywanych w ciągu jednego dnia) podjęto decyzję włączenia do systemu środków pozyskiwania wiedzy metodami tzw. "uczenia maszynowego" na podstawie przykładów, zgromadzonych w odpowiednio przygotowanych bazach danych. Dla potrzeb pozyskiwania wiedzy danymi uczącymi mogą być bądź to *wyniki pomiarów wartości cech sygnałów diagnostycznych* (jak np. w pracy [Moczulski, 1996a]) i *cech tzw. parametrów wejściowych* (procesowych) oraz ewentualne *spostrzeżenia obsługi*, *zapisy służb utrzymania ruchu* itp., bądź też *wyniki obliczeń symulacyjnych* (zob. [Moczulski, Maniak, 1997]). Postępowanie takie pozwala na wielokrotne zwiększenie efektywności procesu pozyskiwania wiedzy. Należy podkreślić, że przedstawiane metody oraz opracowany system pozyskiwania wiedzy mogą być stosowane do różnych grup zadań diagnostycznych, rozróżnianych przez C. Cempla jako zadania diagnostyki kontrolnej, eksploatacyjnej i diagnostyki procesów przemysłowych (por. [Cempel, 1989]).

Przeprowadzone badania dotyczą *metodologii pozyskiwania wiedzy dla potrzeb diagnostyki maszyn*, przy czym wiedza ta pozyskiwana jest w celu jej stosowania do komputerowego wspomagania działań z zakresu diagnostyki technicznej. Tematyka ta ma odniesienie do ogólnej problematyki *komputeryzacji prac inżynierskich*. "Szerokie wykorzystanie komputerów niemal zrewolucjonizowało (i w dalszym ciągu znacząco zmienia) wiele dziedzin twórczości inżynierskiej" [Osiński, Wróbel, 1995]. Techniki komputerowe stosowane są coraz powszechniej do wspomagania prac inżynierskich we wszystkich stadiach istnienia środka technicznego, począwszy od sformułowania założeń projektowych, a skończywszy na jego utylizacji. Rozwój tych technik może być zidentyfikowany na podstawie tematyki referatów, przedstawianych na międzynarodowych konferencjach, które obejmują problematykę komputerowego wspomagania prac inżynierskich (zob. np. [IJCAI-WK, 1995], [AAISE, 1996], [ICED, 1997]).

W stadium *projektowania i konstruowania* komputeryzacja prac inżynierskich obejmuje między innymi zadania: projektowania i konstruowania koncepcyjnego, rutynowego i innowacyjnego, zapisu konstrukcji, wyznaczenia struktury i parametrów elementów ma-

[1] Odpowiednik angielskiego terminu *knowledge engineer*; termin, zdaniem autora, jest nieprzetłumaczalny na język polski.

szyn, obliczeń z zastosowaniem metody elementów skończonych i elementów brzegowych, modelowania i identyfikacji modeli układów mechanicznych, doboru tworzywa, optymalizacji doboru cech konstrukcyjnych, kształtowania bezpieczeństwa maszyn, tworzenia dokumentacji techniczno-ruchowej, czy wreszcie zadania komputerowo zintegrowanego wytwarzania w odniesieniu do opracowania technologii oraz programów do sterowania obrabiarek. Wiele najnowszych publikacji (np. referaty na konferencji ICED'97 [ICED, 1997]) dotyczy tzw. *projektowania współbieżnego* (ang. *Concurrent Engineering*), do którego powszechnie stosuje się wspomaganie komputerowe (zwłaszcza gdy współpracują ze sobą większe zespoły ludzkie).

W stadium *wytwarzania* warto wskazać np. na rolę komputeryzacji prac inżynierskich w odniesieniu do planowania i sterowania produkcją oraz zapewnienia jakości procesów wytwórczych (np. poprzez diagnostykę procesu lub narzędzia).

W stadium *eksploatacji* wspomaganie komputerowe dotyczy między innymi badań niezawodności, gromadzenia danych dotyczących bezpieczeństwa, gromadzenia danych dotyczących eksploatacji maszyn, prowadzenia badań nieniszczących, planowania i prowadzenia obsług, napraw i remontów oraz monitorowania i diagnostyki maszyn (zwłaszcza w zakresie pomiaru sygnałów, estymacji wartości cech sygnałów, gromadzenia wyników obserwacji diagnostycznych oraz wnioskowania diagnostycznego).

Warto także wspomnieć o komputerowym wspomaganiu zapewnienia jakości we wszystkich fazach istnienia środka technicznego (np. [Moczulski, 1993]).

W opisanych wyżej stadiach istnienia środka technicznego komputerowe wspomaganie prac inżynierskich związane jest z potrzebą *działania na danych i wiedzy*. Coraz powszechniej dostrzegana jest potrzeba *zintegrowania różnych stosowanych systemów komputerowych*. Integracja taka może nastąpić np. poprzez "budowę zintegrowanych baz danych (...) z inteligentnym modułem zarządzającym" [Pokojski, Wróbel, 1995]. Wspólna, zunifikowana baza danych umożliwia *integrację łańcuchów procesowych w celu rozwiązywania problemów całościowych*, np. realizacji zleceń klienta [Banaszak *et al.*, 1997]. Nowe techniki informatyczne (bazy danych i wiedzy, "inteligentne" systemy informacyjne, sieci rozległe *InterNet*, multimedia, a zwłaszcza zastosowanie dokumentów hypertekstowych) umożliwiają praktyczną realizację koncepcji *GEN* (ang. *Global Engineering Networking*, co można przetłumaczyć jako "globalna sieć inżynierska")[2] [ICED, 1997], której celem jest zredukowanie czasu potrzebnego do opracowania nowych projektów i konstrukcji, wzrost efektywności prac inżynierskich oraz zapewnienie dostępu do wiedzy i doświadczenia inżynierskiego o poziomie światowym.

Badania autora w zakresie diagnostyki maszyn zostały zapoczątkowane pod kierunkiem Profesora Janusza Dietrycha. W zakresie metodologicznych podstaw diagnostyki maszyn zasadnicze znaczenie dla autora miały prace Profesora Czesława Cempla. Szczególny wpływ na zwrócenie przez autora uwagi na zagadnienia systemów doradczych oraz pozyskiwania wiedzy diagnostycznej miał Profesor Wojciech Cholewa. Współpraca ze środowiskiem naukowym z dziedziny tzw. "sztucznej inteligencji", a zwłaszcza dyskusje prowadzone z Profesorami Ryszardem S. Michalskim i Janem M. Żytkowem pozwoliły autorowi na zintegrowanie różnych koncepcji oraz opracowanie systemu pozyskiwania wiedzy diagnostycznej. Wiele uwag merytorycznych, związanych z weryfikacją poprawności wyników pozyskiwania wiedzy z baz danych zawierających wyniki obliczeń symulacyjnych, sformułował Profesor Jan Kiciński.

[2]Projekt *GEN* jest realizowany przez wiele ośrodków w ramach międzynarodowego programu *ESPRIT*.

Znaczącą część zadań objętych rozprawą wykonano w ramach projektu badawczego nr 8 T11F 020 09 pt. *Akwizycja wiedzy dla potrzeb diagnostyki technicznej*, finansowanego przez Komitet Badań Naukowych.

1.1 Cel rozprawy

Celem rozprawy jest *opracowanie kompleksowej metodologii pozyskiwania wiedzy dla potrzeb diagnostyki maszyn*, obejmującej: *wybrane metody reprezentacji wiedzy diagnostycznej*, *metody pozyskiwania wiedzy od specjalistów* oraz *pozyskiwania wiedzy w sposób indukcyjny na podstawie przykładów*, *metody i techniki oceny wiedzy pozyskanej*, a także *scenariusz procesu pozyskiwania wiedzy diagnostycznej*. Opracowana metodologia stanowi podstawę do utworzenia *systemu pozyskiwania wiedzy diagnostycznej*.

1.2 Tezy

1. Empiryczna relacja diagnostyczna jest środkiem reprezentacji wiedzy, ułatwiającym jej pozyskiwanie od specjalistów z zakresu diagnostyki technicznej.
2. Możliwe jest przyjęcie takiego sposobu reprezentacji wiedzy diagnostycznej, który spełnia kryteria narzucone przez metody pozyskiwania wiedzy od specjalistów oraz pozyskiwania wiedzy metodami indukcyjnymi na podstawie przykładów.

1.3 Zakres rozprawy

Podstawowa treść rozprawy jest zawarta w trzech głównych rozdziałach (rozdziały 2., 3., 4.). Badania proponowanych metod i środków pozyskiwania wiedzy, których celem była ocena przydatności tych metod w diagnostyce maszyn oraz weryfikacja metod i środków opracowanych przez autora lub według jego koncepcji, opisano w rozdziale 5.

W rozdziale 2. przedstawiono zagadnienia dotyczące przedmiotu badań, którym są problemy pozyskiwania wiedzy diagnostycznej. Przyjęto podstawowe założenie o tzw. *domknięt ości świata*, w którym rozpatrywane są zadania diagnostyczne. Założenie to jest spełnione, gdy w danej dziedzinie diagnostyki maszyn występuje skończona liczba obiektów, które należą do niewielu klas, przy czym rozpatrywane są jedynie skończone liczby własności i właściwości tych obiektów, opisywane za pomocą skończonej liczby cech, z których każda może przyjmować wartości ze skończonego (zwykle nielicznego) zbioru. Ponadto zakłada się, że wszystkie te obiekty są rozpatrywane w skończonej liczbie chwil czasu. W rozdziale zwrócono uwagę na własności i właściwości maszyn mające charakter grupowy i jednostkowy, wprowadzając w tym celu hierarchiczną strukturę klas obiektów. Przedstawiono sposób identyfikacji diagnozowanych obiektów oraz klas obiektów za pomocą cech i ich wartości. Podkreślono kluczowe znaczenie cech w procesie identyfikacji obiektów. Opisano sposób konwersji zbioru wartości cech ilościowych (o wartościach będących w ogólnym przypadku liczbami rzeczywistymi, np. uzyskanymi w wyniku pomiaru wielkości fizycznych o charakterze ciągłym) w zbiór wartości cech pochodnych jakościowych (dyskretnych). Przyjęto sposób orzekania o diagnozowanych obiektach za pomocą stwierdzeń dynamicznych. Opisano najważniejsze modele diagnostyczne obiektu, dyskutując ich cechy specyficzne. Dokonana w ten sposób szczegółowa identyfikacja obiektu badań

umożliwiła sformułowanie podstawowego zadania diagnostyki maszyn, którym jest ocena stanu technicznego na podstawie zgromadzonych symptomów diagnostycznych oraz (często niepełnych) danych o wejściu maszyny (przede wszystkim o jej warunkach działania). Wymieniono najważniejsze środki wspomagania wnioskowania diagnostycznego, zwracając uwagę na systemy doradcze. Przedstawiono specyfikę danych i wiedzy diagnostycznej, uzależniając ją od rozwiązywanego zadania diagnostycznego i zwracając uwagę na znaczenie historii wytwarzania i eksploatacji (reprezentowanej np. w postaci sekwencji zdarzeń) oraz zwykle występującą potrzebę uwzględniania upływu czasu eksploatacji. Przedstawiono krótko najważniejsze środki reprezentacji danych i wiedzy dla potrzeb diagnostyki maszyn, nieco więcej uwagi poświęcając stwierdzeniom, regułom i drzewom decyzyjnym. Wreszcie opisano główne stadia procesu pozyskiwania wiedzy diagnostycznej, najbardziej istotne źródła wiedzy diagnostycznej oraz podstawowe metody pozyskiwania wiedzy, charakterystyczne dla każdego z nich. Przedstawiony w tym rozdziale opis przedmiotu badań umożliwił przyjęcie założeń, uwzględnianych następnie w dalszych rozdziałach pracy.

W rozdziale 3. omówiono metody pozyskiwania wiedzy zastosowane w pracy dla potrzeb diagnostyki technicznej oraz sposoby oceny skuteczności procesu pozyskiwania wiedzy. Dokonano podziału metod pozyskiwania wiedzy w zależności od źródeł wiedzy, wyróżniając metody pozyskiwania wiedzy bezpośrednio od specjalistów oraz metody tzw. "uczenia maszynowego" z zastosowaniem indukcji. Do pozyskania wiedzy z zastosowaniem pierwszej grupy metod wprowadzono empiryczną relację diagnostyczną, która może być przekształcona w odpowiedni zbiór reguł wnioskowania. Wiedza specjalisty może być zapisywana w postaci relacji empirycznych. Specjalne formularze ("papierowy" lub "elektroniczny" — będący w istocie specjalnym rodzajem edytora bazy wiedzy) ułatwiają specjaliście uświadomienie sobie występowania tych relacji, przy czym rodzaj formularza określa jednocześnie zakres uczestnictwa w tym procesie programisty bazy wiedzy. W drugiej grupie opisano indukcyjne metody "uczenia maszynowego" na podstawie przykładów. Przedstawiono sposób przygotowania zbioru przykładów, opisując oryginalną metodę dekompozycji zbioru przykładów w wypadku występowania złożonej struktury zbioru stanów technicznych. Ponadto opisano metodę selektywnej indukcji reguł poprzez generowanie pokryć wraz ze sposobem wyboru podzbioru cech relewantnych, metodę indukcji reguł z zastosowaniem zbiorów przybliżonych oraz metodę indukcji drzew decyzyjnych wraz z algorytmem przycinania drzewa. Opisano także sposoby oceny wiedzy pozyskanej od specjalistów oraz sformalizowane sposoby oceny zawartości bazy wiedzy, które oparte są na miarach częstościowych. Wskazano na możliwość "krzyżowego" stosowania sposobów oceny pozyskanej wiedzy (czyli ocena merytoryczna innymi metodami niż te, które zastosowano do pozyskania wiedzy). Wreszcie opisano scenariusz procesu pozyskiwania wiedzy, który winien ułatwić właściwe zaplanowanie oraz przeprowadzenie tego procesu. Przedstawione metody stanowiły podstawę do opracowania środków wspomagania procesu pozyskiwania wiedzy, opisanych w dalszej części pracy.

W rozdziale 4. przedstawiono środki zastosowane do pozyskiwania wiedzy diagnostycznej. W opisanych badaniach podstawowe znaczenie ma baza danych i wiedzy, która jest zbudowana zgodnie ze schematem logicznym *EMPREL*. Schemat ten uwzględnia uwarunkowania związane z gromadzeniem danych i wiedzy o obiektach oraz klasach obiektów, pozbawiając bazę redundancji. We wstępnym stadium konstruowania bazy wiedzy jej zawartość pozyskiwana jest od specjalistów, do czego służy odpowiedni zestaw formularzy. Dla potrzeb systemu, aby uwzględnić indywidualne preferencje współpracujących specjalistów-diagnostów, opracowano dwie wersje zestawu formularzy: "papierową" oraz "elektroniczną". Następnie opisano stosowane oprogramowanie do "uczenia maszynowego" z użyciem metod selektywnej indukcji poprzez generowanie pokryć, indukcji drzew decyzyjnych oraz indukcji reguł z zastosowaniem zbiorów przybliżonych. Rozdział uzupełnia opis środków i sposobów integrowania poszczególnych elementów we wspólny system pozyskiwania wiedzy diagnostycznej *SPWD1*.

Wreszcie w rozdziale 5. przedstawiono przebieg badań opisanych metod pozyskiwania wiedzy diagnostycznej oraz środków zastosowanych w celu wdrożenia tych metod. Badania dotyczyły dwóch obiektów: *modelu maszyny wirnikowej Rotor Kit* (opis modelu w [Bently Nevada, 1996]), obserwowanego w ramach czynnego eksperymentu diagnostycznego, oraz *stanowiska modelowego do badań drgań wirników* (opis w [Rybczyński, 1993]), dla którego został opracowany model [Kiciński *et al.*, 1995] zastosowany przez autora w eksperymencie numerycznym.

W pracy pominięto problemy pomiaru, rejestracji i analizy sygnałów (zob. np. w [Cholewa, Moczulski, 1993]), przetwarzania cech sygnałów i wnioskowania diagnostycznego (zob. np. w [Cholewa, Kaźmierczak, 1995]), podstawowe zagadnienia systemów doradczych ([Cholewa, Pedrycz, 1987], [Cholewa, Moczulski, 1990], [Korbicz, 1994], [Korbicz, 1996] i in.).

Rozdział 2

Przedmiot badań

Celem rozdziału jest szczegółowe określenie przedmiotu badań, którym są problemy pozyskiwania wiedzy diagnostycznej. Stanowić to będzie podstawę opisu metod i środków stosowanych do pozyskiwania wiedzy diagnostycznej, zawartego w rozdziałach 3 i 4.

2.1 Obiekt diagnozowania

Przyjęto, że obiektem diagnozowania[1] jest sztuczny układ materialny (konkret) o, wytworzony według danej konstrukcji [Dietrych, 1985]. Konstrukcja wyznacza zbiór cech (konstrukcyjnych) wspólnych dla wszystkich obiektów o_j, wytworzonych zgodnie z tą konstrukcją.

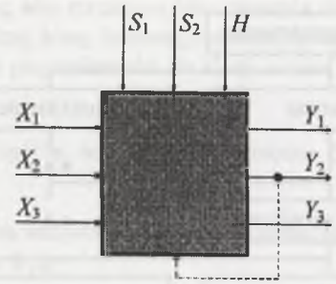
Chociaż obiektem diagnozowania może być maszyna lub urządzenie, w dalszej części pracy termin "maszyna" stosowany będzie zamiennie z terminem "obiekt".

Dla potrzeb opisu obiektu przyjęto następujące założenia metodologiczne, zgodne z ogólną teorią systemów [Mesarovic, Takahara, 1975]:

1. Obiekty są rozpatrywane jako układy względnie wyodrębnione z ich otoczenia.
2. Oddziaływania otoczenia na obiekt odbywają się poprzez wejścia.
3. Oddziaływania obiektu na otoczenie rozpatrywane są jako wyjścia tego obiektu.
4. W pracy do identyfikacji właściwości obiektu stosowany jest model "czarnej skrzynki"[2], przy czym obserwacji podlegają wejścia i wyjścia obiektu.
5. Ponieważ nie ma pewności, że obserwowane są wszystkie wejścia i wyjścia, należy ostrożnie przyjmować założenie, że mamy do czynienia z modelem przyczynowo-skutkowym.

[1] Obiekt diagnozowania zwany będzie dalej w skrócie "obiektem".

[2] Możliwe są także inne ujęcia, jak np. modelowanie wybranego zbioru właściwości obiektu, a następnie identyfikowanie odchyleń pomiędzy wartościami cech sygnałów, zmierzonych dla konkretnego obiektu, a wartościami tych cech, określonymi na podstawie modelu dla wybranych stanów obiektu (tzw. *residuów*, zob. np. [Baseville, Nikiforov, 1993], [Isermann, 1997], [Patton, Frank, Clark, 1989]). Ujęcie takie bazuje na metodach teorii sterowania i stosowane jest np. w diagnostyce procesów przemysłowych. Autor stosował je w badaniach związanych z diagnostyką kontrolną silników elektrycznych małej mocy [Barschdorff, Moczulski, Pflaging, 1986] (zob. także np. [Filbert, Dreetz, 1986]).



Rys. 2.1. Obiekt diagnozowania
Fig. 2.1. A diagnosed object

Rys. 2.1 przedstawia model obiektu diagnozowania, rozpatrywany w ustalonej chwili czasu eksploatacji $\theta \in \Theta$. Na rysunku wyróżniono podstawowe zbiory wejść i wyjść:

X_1 — zbiór wejść związanych z użytecznym działaniem maszyny (np. wejścia masowe i energetyczne);

X_2 — zbiór wejść określających warunki działania maszyny (np. obciążenie zewnętrzne, prędkość obrotowa);

X_3 — zbiór wejść, których wpływ jest trudny do oszacowania lub pomiaru i które opisuje się za pomocą wartości losowych oraz interpretuje jako szum wpływający na zmniejszenie stopnia prawdziwości relacji diagnostycznych;

S_1 — zbiór stanów związanych z konstrukcją maszyny (a więc w zasadzie nie zmieniających się w czasie);

S_2 — zbiór stanów wywołanych procesem wytwarzania i/lub użytkowania maszyny — poszukiwanie cech tych stanów jest celem procesu diagnozowania;

H — historia eksploatacji maszyny (zapis zdarzeń i procesów związanych z obsługiwaniem maszyny, który może być reprezentowany w postaci zapisu sekwencji skutków tych zdarzeń i procesów, jakimi są osiągane stany ze zbioru S_2);

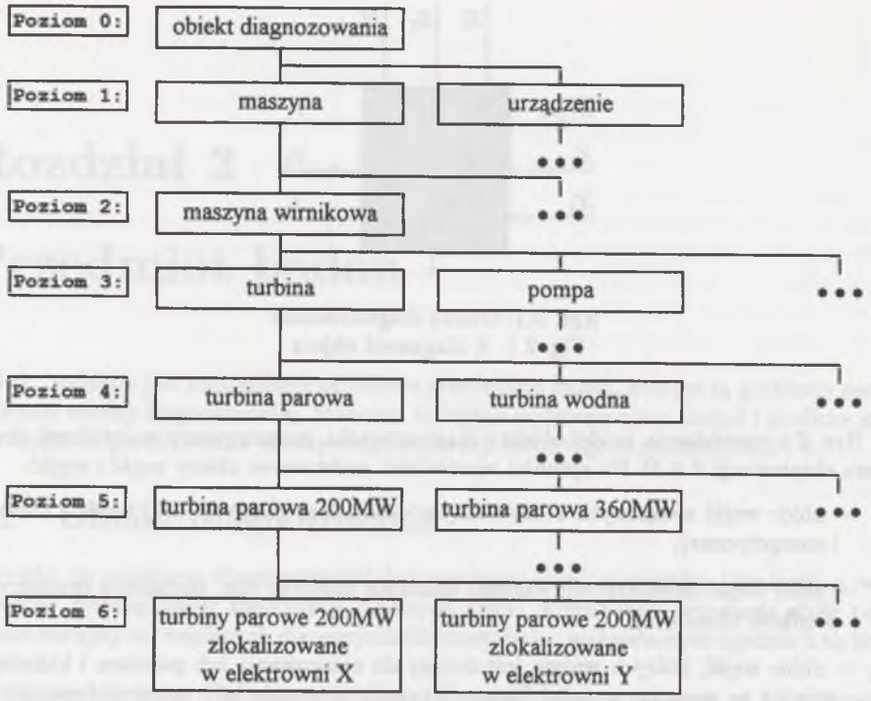
Y_1 — zbiór wyjść związanych z użytecznym działaniem maszyny (np. wytwory);

Y_2 — zbiór wyjść związanych z występowaniem procesów resztkowych [Cempel, 1982], które są niezamierzonym lub nieuniknionym skutkiem działania maszyny (np. drgania i hałas); istnieje możliwość wystąpienia destrukcyjnego sprzężenia zwrotnego (rys. 2.1);

Y_3 — zbiór wyjść nie znanych lub nie obserwowanych, które są skutkiem działania maszyny.

W dalszym ciągu pracy rozpatrywane będą zbiory wejść X obiektu, jego wyjść Y oraz stanów S w danej chwili $\theta \in \Theta$ czasu jego eksploatacji:

$$X = X_1 \cup X_2 \cup X_3 = \{x_i \mid i \in I\}, \tag{2.1}$$



Rys. 2.2. Przykładowa struktura klas
Fig. 2.2. An exemplary structure of classes

$$Y = Y_1 \cup Y_2 \cup Y_3 = \{y_k \mid k \in K\}, \tag{2.2}$$

$$S = S_1 \cup S_2 = \{s_j \mid j \in J\}, \tag{2.3}$$

gdzie I, J, K są zbiorami indeksów. Są one opisywane za pomocą cech i wartości tych cech w chwili θ. Sposób opisu elementów powyższych zbiorów przedstawiono w rozdziale 2.1.4.

2.1.1 Klasy obiektów

Obiekty mogą być elementami *klas obiektów*[3]. *Klasę* można określić poprzez wyspecyfikowanie wszystkich jej obiektów lub poprzez podanie pewnego warunku przynależności do klasy.

W pracy przyjęto, że klasy obiektów mogą tworzyć *struktury hierarchiczne* (por. rys. 2.2), w których wyróżnia się *klasy bazowe* (tzw. "przodkowie") i *klasy pochodne* (tzw. "potomkowie"). Nie wymaga się, aby struktura klas tworzyła drzewo, co oznacza możliwość występowania więcej niż jednej klasy bazowej dla danej klasy pochodnej.

[3]W takim wypadku można także mówić o zbiorze (w sensie dystrybutywnym [Marciszewski, 1988]). W pracy postanowiono stosować pojęcie "klasa" ze względu na wprowadzenie hierarchii klas oraz tzw. dziedziczenia cech, typowego dla ujęcia obiektowego (zob. np. [Stroustrup, 1991]).

Spośród wielu warunków przynależności do klasy można przykładowo wymienić (por. rys. 2.2):

- wspólne najbardziej ogólne własności dynamiczne (np. klasa "maszyna wirnikowa" — poziom 2.);
- wspólną istotę działania wszystkich obiektów należących do klasy (np. klasa "turbina parowa" — poziom 4.);
- wspólną konstrukcję (np. klasa "turbina parowa 200 MW" — poziom 5.),
- wspólną lokalizację (np. "turbiny 200 MW zlokalizowane w elektrowni X" — poziom 6.).

Własności obiektów reprezentowane są za pomocą cech tych obiektów. Wprowadzenie klas umożliwia posłużenie się pojęciem *dziedziczenia cech obiektów* oraz *przesłaniania cech*.

2.1.2 Cechy oraz ich wartości

Wejścia, *stany* i *wyjścia* diagnozowanego obiektu w danej chwili czasu eksploatacji są opisywane za pomocą *cech* oraz *wartości* tych *cech*. Cechy określane są dla klas obiektów:

$$a : (X \cup S \cup Y) \times \Theta \to D \tag{2.4}$$

$$(\forall \theta \in \Theta)(\forall z \in X \cup S \cup Y)[a : (z, \theta) \to a(z(\theta))], \tag{2.5}$$

gdzie Θ jest przedziałem czasu eksploatacji, obejmującym cały okres istnienia danego środka technicznego (od jego wytworzenia do ostatecznej likwidacji), zaś D jest dziedziną, tj. przestrzenią wartości cech.

Należy zwrócić uwagę na fakt, że takie określenie cechy różni się od jej określenia jako "elementu zbioru, który stanowi orzeczenie o przedmiocie poznania lub przedmiocie tworzenia" [Dietrych, 1985]. W pracy przyjęto, że cecha jest funkcją, której wartości opisują obiekt. Rodzajem orzeczenia o przedmiocie poznania jest *stwierdzenie* (zob. rozdział 2.1.4).

Ze względu na losowe własności procesów wytwarzania, obsługiwania i użytkowania, powodujące losowe własności obiektów, a także ze względu na losowe własności procesów obserwacji odpowiednich wejść i wyjść (wynikające np. z lokalizacji i montażu czujnika, wyboru chwili czasu rozpoczęcia rejestracji podrealizacji sygnału, szumów pomiarowych itp.), celowe jest rozpatrywanie przestrzeni zdarzeń $\Xi = \{\xi\}$, w której pojedynczym zdarzeniem jest wynik obserwacji wejścia lub wyjścia. Wejścia i wyjścia mogą być reprezentowane za pomocą (być może wielowymiarowych) procesów stochastycznych [Papoulis, 1972].

Cechy ilościowe i jakościowe

Ze względu na potrzeby rozprawy rozróżnia się *cechy ilościowe* i *cechy jakościowe*. Rozróżnienie to wprowadza się ze względu na dziedzinę danej cechy.

Wartości cechy ilościowej stanowią zwykle *bezpośredni wynik obliczenia* (np. w wypadku cech konstrukcyjnych) lub *ocenę pomiaru wielkości fizycznej*[4] (np. wartość skuteczna amplitudy prędkości drgań korpusu maszyny w określonym kierunku). Dziedziną D wielowymiarowej ($n \geqslant 1$) cechy ilościowej jest podzbiór przestrzeni:

$$R^n = \underbrace{R \times R \times \ldots \times R}_{n}. \tag{2.6}$$

Wartości cechy jakościowej mogą być określone np. *w wyniku kodowania symbolicznego* lub uzyskane *w wyniku odpowiedniej dyskretyzacji cech ilościowych* [Sobczak, Malina, 1985], [Chmielewski, Grzymała-Busse, 1992], [Dougherty, Kohavi, Sahami, 1995].

Rodzaje cech

Ze względu na różną rolę wybranych cech w procesie identyfikacji diagnozowanych obiektów oraz klas obiektów wprowadzono specjalną własność cechy, jaką jest jej *rodzaj*. Wyróżniono następujące rodzaje cech (por. rys 2.1):

- *cechy podstawowych warunków działania obiektu* (cechy wejść zawartych w podzbiorze zbioru X_2), którym to warunkom odpowiadają określone sposoby dyskretyzacji cech;
- *cechy zdarzeń eksploatacyjnych i faz eksploatacji* (które mogą być interpretowane jako cechy stanów z podzbioru $H \subset S$), umożliwiające zapis historii wytwarzania i eksploatacji danego obiektu;
- *cechy stanu technicznego* (cechy stanów zawartych głównie w podzbiorze zbioru S_2);
- *cechy wejścia* (cechy wejść z pewnego podzbioru zbiorów wejść $X_1 \cup X_2$);
- *cechy wyjścia* (cechy wyjść z pewnego podzbioru zbiorów wyjść $Y_1 \cup Y_2$).

Typy cech

Istotną własnością cechy jest jej *typ*. Wyróżniane są następujące typy cech (zastosowano nazwy typów zgodne ze stosowanymi w [Wnęk *et al.*, 1995]):

- typ `pom` (od "pomiar", odpowiedni dla cechy ilościowej o wartościach rzeczywistych),
- dla cech jakościowych:
  - typ `nom` (od "nominalna" — dla cech, których dziedziny są zbiorami nie uporządkowanymi (tj. nie wprowadzono w nich relacji porządku "<"),
  - typ `lin` (od "liniowo [uporządkowana]") — dla cech, których dziedziny są zbiorami uporządkowanymi (wprowadzono w nich relację porządku liniowego "<"),
  - typ `cyc` (od "cykliczna") - dla cech, które są okresowymi funkcjami argumentu θ (ich dziedziny są uporządkowane w sposób cykliczny — zob. poniżej).

[4] W tym wypadku można wyróżnić *cechy liczbowe* i *cechy funkcyjne*, które ze względu na zamiar pozyskiwania wiedzy z zastosowaniem komputera przedstawia się jako jednokolumnowe macierze liczb (por. [Cholewa, 1983]).

Uporządkowanie cykliczne cech jakościowych dla dziedziny $D = \{v_0, v_1, \ldots, v_{n-1}\}$ zawierającej n wartości dyskretnych, uporządkowanej relacją "<", określa się poprzez wprowadzenie relacji $next \subset D \times D$, takiej że:

$$(v_i, v_j) \in next \Leftrightarrow [(i, j < n-1 \wedge j = i+1) \vee (i = n-1 \wedge j = 0)]. \tag{2.7}$$

Relacja ta jest *funkcją*, w związku z tym zamiast $(v_i, v_j) \in next$ można ją zapisywać jako $next(v_i) = v_j$.

Wartości cech

Jak opisano powyżej, wszystkie cechy (a więc także cechy jakościowe) są określane dla klas obiektów. Również dziedzina danej cechy jest wspólna dla tej klasy.

Wartości cechy ilościowej są elementami przestrzeni $R^n, n \geqslant 1$. Dziedziną cechy ilościowej jest n-wymiarowy obszar w tej przestrzeni (który dla cechy ilościowej o wartościach będących wynikiem pomiaru jest zwykle spójny).

Wartości cechy jakościowej są dyskretne. Mogą one być np. zapisane lub zakodowane w postaci liczb (zwykle całkowitych nieujemnych) lub być termami[5], zbudowanymi ze znaków alfanumerycznych. W wypadku cech jakościowych moc dziedziny $card(D)$[6] winna być niewielka.

W celu przeciwdziałania niepotrzebnej utracie informacji o relacjach, zachodzących pomiędzy różnymi wartościami cech ilościowych, zaleca się stosowanie takich sposobów kodowania wartości cech, aby można było zachować relację porządku liniowego tam, gdzie porządek ten występuje.

Ocena zmienności wartości cech ilościowych

Dla szczególnej grupy cech, które są ocenami sygnałów, podawany jest także parametr ΔT, który ma znaczenie w wypadku operacji tzw. "sklejania" stwierdzeń dynamicznych. Istotę tej operacji opisano w rozdziale 2.1.4.

Dyskretyzacja cech ilościowych

Założenie rozpatrywania skończonej liczby stanów w przestrzeni wartości cech, a także wymóg uogólniania wiedzy diagnostycznej i jej uniezależniania od indywidualnych cech poszczególnych obiektów wywołują potrzebę *konwersji cech ilościowych* (o wartościach będących liczbami rzeczywistymi, np. uzyskanych w wyniku pomiaru wielkości fizycznych o charakterze ciągłym) *w cechy jakościowe*. W tym celu dokonywana jest *dyskretyzacja cech ilościowych*. W jej wyniku otrzymuje się *wartości cechy jakościowej*.

Niech a będzie cechą ilościową, a $Dom(a) = [v, v']$ dziedziną cechy a. Niech ponadto $Dom(a_\circ) = \{m \mid 0 \leqslant m < M(a)\}$ jest dziedziną *pochodnej cechy jakościowej* $a_\circ$, gdzie $M(a)$ oznacza liczbę różnych wartości tej cechy.

Zbiór par:

$$\pi_a = \{\langle v_i, m_i \rangle \mid i = 1, 2, \ldots, q \wedge m_i \in Dom(a_\circ)\} \tag{2.8}$$

$$v_0 = v,\ v_{i-1} < v_i,\ i = 1, 2, \ldots, q,\ v_q = v' \tag{2.9}$$

[5] *Term* jest stałą nazwową lub zmienną nazwową [Marciszewski, 1988].

[6] Dla zbioru D o skończonej liczbie elementów moc równa jest liczbie elementów tego zbioru.

określa przekształcenie

$$f : Dom(a) \rightarrow Dom(a_\circ)$$

dziedziny cechy a w dziedzinę cechy dyskretnej $a_\circ$ (która w dalszym ciągu pracy zwana będzie *cechą pochodną*), zdefiniowane następująco:

$$f(v) = \sum_{i=1}^{q} m_i \cdot U_1(v - v_{i-1}) \cdot U_2(v_i - v) + m_q \cdot U_1(v - v_q), \tag{2.10}$$

gdzie

$$U_1(v) = \begin{cases} 0, & \text{gdy } v < 0 \\ 1, & \text{gdy } v \geqslant 0 \end{cases}; \quad U_2(v) = \begin{cases} 0, & \text{gdy } v \leqslant 0 \\ 1, & \text{gdy } v > 0 \end{cases}. \tag{2.11}$$

Jeżeli jest spełniony dodatkowy warunek liniowego uporządkowania wartości cechy jakościowej $a_\circ$, to cecha ta jest cechą typu `lin`[7]. Należy nadmienić, że wartościom liczbowym cechy $a_\circ$ można przyporządkować nazwy, co umożliwia automatyczne generowanie zdań w języku naturalnym, dotyczących cech i ich wartości.

Z badań autora (zob. [Maniak, Moczulski, 1997]) wynika, że wybór liczby poziomów kwantyzacji q oraz określenie progów podziałów mają zasadniczy wpływ na informatywność uzyskanej cechy pochodnej jakościowej, co znajduje także potwierdzenie w dostępnej literaturze (zob. [Sobczak, Malina, 1985], [Chmielewski, Grzymała-Busse, 1992], [Dougherty, Kohavi, Sahami, 1995]). Oba te zagadnienia mogą być przedmiotem odrębnych badań odniesionych do potrzeb diagnostyki maszyn.

Przekształcenie dziedziny cechy a w dziedzinę cechy jakościowej $a_\circ$ może być określane dla klasy obiektów, a także indywidualnie dla pojedynczych obiektów z danej klasy (dla wyjątków). Liczba wartości progowych może być przedmiotem optymalizacji, np. ze względu na kryterium jakości klasyfikacji [Sobczak, Malina, 1985].

Dziedziczenie i przesłanianie w odniesieniu do cech

Hierarchiczna struktura klas umożliwia wprowadzenie pojęcia tzw. *dziedziczenia*, które dotyczy cech, ich wartości oraz sposobów dyskretyzacji cech ilościowych. Dzięki dziedziczeniu w klasie pochodnej określone są wszystkie te cechy, ich wartości oraz sposoby dyskretyzacji, które są określone w jednej z jej klas nadrzędnych (oraz, być może, pewne nowe). Ponadto w odniesieniu do pojedynczych obiektów możliwe jest także tzw. *przesłanianie* sposobów przekształcania cech ilościowych w cechy jakościowe, które może zachodzić w odniesieniu do obiektów. Wówczas zachodzi zjawisko "przesłaniania" przekształcenia, określonego dla klasy obiektów, przez przekształcenie, określone dla obiektu, co umożliwia *reprezentowanie wyjątków*. Dziedziczenie i przesłanianie umożliwiają uniknięcie nadmiarowości w zapisie danych o obiektach diagnozowania.

2.1.3 Czas

W diagnostyce maszyn wyróżnia się dwie zmienne czasowe: czas "mikro" i czas "makro" (zob. [Cholewa, 1983]).

[7]Warunek ten nie zawsze jest spełniony. Rozważmy np. cechę "temperatura oleju w łożysku hydrodynamicznym" o wartościach rzeczywistych z przedziału $[t_{min}, t_{max}]$ i jej cechę pochodną o tej samej nazwie i wartościach: "odpowiednia"=1 (np. z przedziału $40° \div 85°$) oraz "nieodpowiednia"=0 (spoza tego przedziału). Mamy $\pi_a = \{\langle 40, 0\rangle, \langle 85, 1\rangle, \langle t_{max}, 0\rangle\}$. Nie jest jednak możliwe określenie liniowego uporządkowania w zbiorze wartości cechy pochodnej, takiego że gdy $v < v'$, to $f(v) \leqslant f(v')$, gdzie $f(\,)$ wg (2.10).

Czas "mikro" t dotyczy procesu zachodzącego w obiekcie lub sygnału, który dla diagnosty jest nośnikiem informacji o obserwowanym obiekcie. W niniejszej rozprawie zakłada się, że wartości cech estymowane są w dziedzinie czasu t. Ponieważ (jak wspomniano) w rozprawie nie omawia się zagadnień estymacji cech obiektów na podstawie zaobserwowanych sygnałów, własności czasu "mikro" również nie będą szerzej omawiane.

Interpretacja znaczenia czasu "makro" θ zależy od grup zadań diagnostycznych, dla których wiedza jest pozyskiwana:

- W zadaniach diagnostyki kontrolnej czas "makro" ma znaczenie *identyfikatora realizacji procesu stochastycznego* [Papoulis, 1972], który opisuje własności danego egzemplarza obiektu, wykonanego zgodnie ze wspólną konstrukcją [Cholewa, 1983].

- W zadaniach diagnostyki eksploatacyjnej czas "makro" jest *czasem istnienia środka technicznego* (tzw. czasem "życia" obiektu — por. [Cempel, 1989]). W czasie "makro" następują zmiany stanu diagnozowanego procesu [Kaźmierczak, 1989]. Ponieważ jednym z podstawowych zastosowań pozyskanej wiedzy jest wnioskowanie w zakresie diagnostyki eksploatacyjnej maszyn i urządzeń, w dalszym ciągu pracy razem z określeniem "czas makro" stosowane będzie (tam, gdzie nie będzie to prowadziło do nieporozumień) także określenie "czas eksploatacji" lub po prostu "czas".

- W zadaniach diagnostyki procesów przemysłowych czas "makro" może mieć znaczenie *czasu ewolucji danego procesu.*

W diagnostyce technicznej czas może być reprezentowany za pomocą *struktur punktowych* (por. [Hajnicz, 1996]), budowanych z zastosowaniem relacji poprzedzania "$<$". Strukturę czasu punktowego $\mathcal{T}$ z porządkiem liniowym[8] stanowi para $\langle \Theta, < \rangle$, gdzie Θ jest zbiorem punktów czasu, zaś "$<$" jest dwuargumentową relacją poprzedzania, spełniającą takie aksjomaty, jak:

- przechodniość:

$$(\forall \theta_1, \theta_2, \theta_3 \in \Theta)[\theta_1 < \theta_2 \wedge \theta_2 < \theta_3 \Rightarrow \theta_1 < \theta_3]; \tag{2.12}$$

- antyzwrotność:

$$(\forall \theta_0 \in \Theta)[\neg(\theta_0 < \theta_0)]; \tag{2.13}$$

- spójność:

$$(\forall \theta_1, \theta_2 \in \Theta)[\theta_1 < \theta_2 \vee \theta_1 = \theta_2 \vee \theta_2 < \theta_1]. \tag{2.14}$$

Dla potrzeb jednolitej reprezentacji danych i wiedzy wymagających zapisu informacji o czasie ich wystąpienia postanowiono ograniczyć się do *struktur czasu z porządkiem dyskretnym*, spełniających następujący dodatkowy aksjomat dyskretności:

$$\begin{aligned}&(\forall \theta_1, \theta_2 \in \Theta)[\theta_1 < \theta_2 \Rightarrow \exists \theta_3 \in \Theta(\theta_1 < \theta_3 \wedge \neg \exists \theta_0 \in \Theta(\theta_1 < \theta_0 \wedge \theta_0 < \theta_3))],\\ &(\forall \theta_1, \theta_2 \in \Theta)[\theta_1 < \theta_2 \Rightarrow \exists \theta_3 \in \Theta(\theta_3 < \theta_2 \wedge \neg \exists \theta_0 \in \Theta(\theta_3 < \theta_0 \wedge \theta_0 < \theta_2))].\end{aligned} \tag{2.15}$$

[8]W pracy postanowiono ograniczyć się do porządku liniowego. Zagadnienia struktur czasu z porządkiem nieliniowym omawiane są np. w [Hajnicz, 1996].

W niniejszej pracy zmienna czasu eksploatacji θ stanowi zwykle *indeks szeregów czasowych* [Box, Jenkins, 1983] wartości cech, za pomocą których opisywane są własności i właściwości obiektu badań.

Spośród punktów czasu wyróżnia się *daty*, które stanowią specyfikację czasu chronologicznie ustaloną (Rescher i Urquhardt, 1971; cytowane za [Hajnicz, 1996]). Aby zapewnić możliwość badania jednoczesności występowania różnych zdarzeń, celowe jest stosowanie do zapisu dat czasu uniwersalnego GMT (co jest zgodne z innymi stosowanymi standardami — por. [MIMOSA, 1996]).

Oprócz punktów czasu w pracy rozważa się *przedziały czasu* $[\theta_1, \theta_2]$, definiowane za pomocą par punktów $\theta_1, \theta_2 \in \Theta$, określających końce tych przedziałów (zgodnie z propozycją McDermotta, 1982; cytowane za [Hajnicz, 1996]).

Ponieważ punkt czasu jest pojęciem abstrakcyjnym, w diagnostyce technicznej celowe jest rozważanie *momentu czasu*, który interpretowany jest jako pewien bardzo krótki (zwykle niepodzielny) przedział czasu. Rozważanie momentów czasu jest celowe np. w odniesieniu do operacji estymacji wartości cechy sygnału diagnostycznego.

2.1.4 Stwierdzenia, stwierdzenia dynamiczne i historia obiektu

W pracy przyjęto, że informacja o obiektach zawarta będzie w *stwierdzeniach*, *zapisach zdarzeń* oraz *historiach*.

Stwierdzenia

Stwierdzenie[9] jest zdaniem, które stanowi orzeczenie o wartości cechy danego obiektu (abstraktu lub konkretu). Stwierdzeniu może być przypisana *ocena jego stopnia prawdziwości*. W pracy przyjęto, że ze względu na właściwości obiektów oraz specyfikę zadań diagnostycznych rozpatrywane będą *stwierdzenia przybliżone* (ang. *approximate statements*; por. np. [Dubois, Prade, 1996], [Cholewa, Pedrycz, 1987]), których wartościowanie wymaga stosowania zbioru wartości logicznych. Zbiór ten zawiera więcej niż dwa elementy (tj. $\{prawda, fałsz\}$). Stwierdzenie p zapisywane jest za pomocą następującej trójki:

$$p = \langle nam(o), atr(o), val(atr(o)) \rangle, \tag{2.16}$$

gdzie: o — obiekt, $nam(\,)$ — nazwa obiektu, $atr(\,)$ — nazwa cechy tego obiektu (dotyczącej jego wejścia, stanu lub wyjścia), $val(\,)$ — wartość cechy.

Stwierdzenia dynamiczne

W większości wypadków stwierdzeniu można przypisać przedział czasu $I = [\theta_1, \theta_2]$[10], w którym stwierdzeniu p_I przysługuje stopień prawdziwości (lub stopień przekonania o prawdziwości wypowiedzi [Cholewa, 1996a]) $T(p_I)$. Stwierdzenie takie będzie nazywane *stwierdzeniem dynamicznym* [Cholewa, 1997], [Cholewa, 1997a].

W rozpatrywanym w niniejszej pracy punktowo-przedziałowym modelu temporalnym bardzo istotną własnością stwierdzeń dynamicznych jest ich podatność na łączenie (ang. *concatenability*), co umożliwia operację *sklejania stwierdzeń dynamicznych*, polegającą na połączeniu dwóch przylegających do siebie przedziałów określoności stwierdzeń dynamicznych:

$$\begin{aligned}(\forall \theta_1, \theta_2, \theta_3 \in \Theta)[\theta_1 < \theta_2 < \theta_3 \\ \wedge\, I_{1,2} = [\theta_1, \theta_2] \wedge I_{2,3} = [\theta_2, \theta_3] \wedge I_{1,3} = [\theta_1, \theta_3] \\ \wedge\, T(p_{I_{1,2}}) = T(p_{I_{2,3}}) = \alpha \Rightarrow T(p_{I_{1,3}}) = \alpha],\end{aligned} \tag{2.17}$$

gdzie α oznacza wartość stopnia prawdziwości, równą dla wszystkich rozpatrywanych stwierdzeń dynamicznych.

Zdarzenia eksploatacyjne

W każdym punkcie czasu obiekt i jego otoczenie znajdują się w pewnym *stanie*, który jest zbiorem obiektów wraz z łączącymi je funkcjami i relacjami. *Zdarzenie* jest ciągiem stanów, w którym wyróżnia się *stan początkowy* i *stan końcowy*. *Zdarzenie atomowe* jest ciągiem złożonym z dwóch kolejnych stanów. Domyślnym stanem początkowym jest aktualny stan obiektu (i jego otoczenia). Nazwa zdarzenia określana jest przez stan końcowy.

W niniejszej pracy zdarzenia stosowane są do reprezentowania tzw. *zdarzeń eksploatacyjnych* (zob. tabl. A.1).

Historia obiektu

Historią obiektu (rozpatrywanego wraz z jego otoczeniem) jest ciąg stanów dotyczących tego obiektu. Historia obiektu może także być reprezentowana przez ciąg zdarzeń dotyczących tego obiektu. *Historię parametru obiektu* stanowią szeregi czasowe wartości tego parametru.

W diagnostyce maszyn szczególne znaczenie ma zapis danych dotyczących zdarzeń zaszłych w fazie wytwarzania oraz eksploatacji (użytkowania i obsługiwania) danego środka technicznego. Dane takie, zapisane wraz z chwilą czasu eksploatacji θ, tworzą historię wytwarzania lub eksploatacji określonego środka technicznego.

Historia obiektu może być podzielona na *przekroje* (ang. *slices*), wyznaczane dla danego punktu czasowego (por. [Hajnicz, 1996]).

2.1.5 Model diagnostyczny obiektu

Dla potrzeb diagnostycznych tworzone są *modele*, opisujące związki zachodzące pomiędzy cechami wejść, wyjść i stanów obiektu [Cholewa, Kaźmierczak, 1995]. Do budowy modelu (będącego abstraktem) stosowany jest świadomie ograniczony zbiór cech obiektu, umożliwiający zbadanie wymaganych właściwości tego obiektu [Dietrych, 1985]. Modele obiektów mogą być:

- uzyskane na drodze teoretycznej (na podstawie teorii),
- identyfikowane na podstawie zgromadzonych wyników eksperymentu,
- identyfikowane na podstawie wyników obliczeń symulacyjnych,
- pozyskiwane od specjalisty (np. w postaci jakościowej).

[9] W wielu publikacjach stwierdzenie utożsamiane jest z faktem. W. Cholewa (zob. np. [Cholewa, 1997]) odróżnia stwierdzenie od faktu, zwracając uwagę na istotną właściwość stwierdzenia, które jest orzeczeniem o obiektywnie zachodzącym fakcie i jako takie może mieć charakter subiektywny.

[10] W szczególnym przypadku przedział ten może się redukować do punktu czasu $I = [\theta, \theta]$.

Ze względu na różne potrzeby diagnostyczne mogą być tworzone następujące rodzaje modeli [Cholewa, Kaźmierczak, 1995]:

- modele indywidualne, opisujące właściwości pojedynczego obiektu (zwykle konkretu),
- modele grupowe, opisujące wspólne właściwości pewnej klasy obiektów.

Dla pojedynczego obiektu można utworzyć jednoelementową klasę, zawierającą jedynie ten obiekt. Dlatego bez utraty ogólności opisu w dalszym ciągu będą rozważane *modele grupowe*, zwane od tego miejsca *modelami*.

W diagnostyce maszyn znalazły zastosowanie następujące modele:

- *model strukturalny* (np. [Cempel, 1982], [Cempel, 1989], [Morel, 1992]);
- *model matematyczny* (np. [Barschdorff, Moczulski, Pflăging, 1986], [Filbert, 1995], [Filbert, Dreetz, 1986], [Patton, Frank, Clark, 1989], [Kiciński, 1996], [Isermann, 1997]);
- *model systemowy* (por. [Cholewa, 1983], [Cholewa, Kaźmierczak, 1995]);
- *model logiczny* (por. [Mesarovic, Takahara, 1975], [Reiter, 1987]);
- *model jakościowy* (por. [Pearce, 1988], [Bratko, 1989], [Kuipers, 1994]);
- *model przybliżony* (np. model rozmyty [Cholewa, 1983]);
- *model neuronalny* oraz jego wersje rozmyte, jak np. *sieci neuronalne o logice rozmytej*, *systemy rozmyte połączone z sieciami neuronalnymi* lub *hybrydowe systemy fuzzy-neuronalne* (np. [Kościelny, Szczepaniak, 1997], [Leonhardt, Ayoubi, 1997]).

Model strukturalny

Model strukturalny (zwany również modelem fizycznym) odwzorowuje strukturę dynamiczną układu oraz wzajemne oddziaływania pomiędzy elementami tego układu. Model taki ułatwia przeprowadzenie analizy własności badanego obiektu, zwłaszcza w zakresie:

- *określenia zbioru stanów technicznych*, istotnych ze względu na działanie diagnozowanego obiektu,
- *identyfikacji wielkości fizycznych*, których zmiany winny być obserwowane za pomocą sygnałów diagnostycznych (tj. takich, których cechy są zależne od stanu technicznego),
- *wyboru cech sygnałów diagnostycznych*, które są "zorientowane uszkodzeniowo" [Cempel, 1982],
- *lokalizacji czujników pomiarowych*,
- *określenia warunków działania obiektu podczas badań*, np. w formie programu badań (w przypadku eksperymentu diagnostycznego, umożliwiającego ingerencję diagnosty w warunki działania obiektu — por. [Moczulski *et al.*, 1995]),

- *określenia związków przyczynowo-skutkowych pomiędzy cechami stanu i cechami sygnałów*, które mogą być zastosowane w procesie wnioskowania diagnostycznego; należy zwrócić uwagę na fakt, że dla potrzeb wnioskowania diagnostycznego związki te muszą być stosowane w "odwrotną" stronę, co zwykle prowadzi do możliwości wnioskowania jedynie w sposób przybliżony (ze względu na niejednoznaczność związków przyczynowo-skutkowych — zob. rozdział 2.2).

Ponieważ modelowane obiekty są złożone, podczas budowy ich modeli wymaga się licznych uproszczeń, których skutkiem jest zwykle brak zgodności szczegółowych wyników wnioskowania, dotyczących stanu obiektu, z wnioskami uzyskiwanymi na podstawie badań eksperymentalnych [Cholewa, 1983], [Cholewa, Kaźmierczak, 1995].

Model matematyczny

Model matematyczny budowany jest na podstawie zidentyfikowanego modelu fizycznego (strukturalnego). Opisuje on związki pomiędzy podstawowymi własnościami obiektu. Zwykle jest reprezentowany w postaci zależności liniowych lub nieliniowych [Patton, Frank, Clark, 1989], [Baseville, Nikiforov, 1993], [Isermann, 1997]. Przykładami zastosowań dla potrzeb diagnostycznych są: model matematyczny opisujący działanie silnika elektrycznego prądu stałego [Barschdorff, Moczulski, Pflăging, 1986], [Filbert, 1995]), układu równań różniczkowych liniowych (jak np. modele opisujące równania ruchu wirnika sztywnego [Gosiewski, Muszyńska, 1992]) lub nieliniowych (np. model opisujący przemieszczenia czopa w łożysku ślizgowym w razie wystąpienia niestabilnego działania łożyska, zwanego "biczem olejowym" [Kiciński, 1996]).

Modele matematyczne (zwane również *modelami analitycznymi*) znajdują coraz szersze zastosowanie w diagnostyce technicznej, a zwłaszcza w diagnostyce procesów przemysłowych oraz układów dynamicznych, co znajduje odbicie np. w tematyce referatów przedstawianych na międzynarodowych sympozjach IFAC SAFEPROCESS (zob. [Isermann, Ballé, 1997], [Patton, 1997]).

Ważna dziedzina zastosowań dotyczy *pozyskiwania nowej wiedzy o relacjach diagnostycznych* [Kiciński *et al.*, 1995], [Cholewa, Kiciński, 1996], zwłaszcza w odniesieniu do stanów technicznych, których prawdopodobieństwo wystąpienia jest bardzo małe lub których wystąpienie może spowodować katastrofalne uszkodzenia maszyny, stanowiące znaczne zagrożenie dla obsługi i środowiska.

Ciekawym podejściem jest budowanie modeli przyczynowo-skutkowych, odpowiadających różnym stanom technicznym maszyny (jak np. w pracy [Radkowski, 1995], gdzie modelowano wpływ deformacji zębów przekładni zębatej na cechy sygnału diagnostycznego), które mogą służyć następnie do klasyfikowania stanów na podstawie zgromadzonych wartości cech.

Modele matematyczne mogą być stosowane do estymacji wartości tzw. *residuów*. Residuum jest funkcją, której wartości są nośnikiem informacji o niesprawnościach i uszkodzeniach układu (lub procesu) [Patton, Frank, Clark, 1989]. Wartość tej funkcji określa się jako różnicę pomiędzy wartością cechy, uzyskaną w wyniku analizy sygnałów zaobserwowanych dla konkretnego obiektu, a wartością cechy obliczoną na podstawie modelu matematycznego, opisującego działanie tego obiektu [Isermann, Ballé, 1997]. Uzyskane w ten sposób residua mogą być stosowane w następujących zadaniach:

- *Detekcja uszkodzenia* (ang. *fault detection*).

  Celem postępowania jest stwierdzenie, *czy w nadzorowanym układzie występuje niesprawność/uszkodzenie* (tzw. diagnostyka dwustanowa dla klas stanów {"zdatny", "niezdatny"}). Jeśli mogą być obserwowane wejścia i wyjścia obiektu, podstawowe metody detekcji uszkodzeń bazują na *obserwatorach stanu*, *filtrze Kalmana* oraz *równaniach parzystości* [Patton, Frank, Clark, 1989], [Isermann, Ballé, 1997], [Kościelny, Szczepaniak, 1997]. Jeśli obserwacja może dotyczyć wyłącznie wyjść, stosowane są metody oparte na *modelach sygnałów* [Isermann, Ballé, 1997].

- *Diagnoza uszkodzenia* (ang. *fault diagnosis*).

  Postępowanie ma na celu określenie *rodzaju*, *wielkości*, *lokalizacji* [Isermann, Ballé, 1997], a nawet *przyczyny* [Barschdorff, Moczulski, Pfläging, 1986] uszkodzenia. Modele analityczne umożliwiają opracowanie *banków obserwatorów* oraz *residuów strukturalnych* [Kościelny, Szczepaniak, 1997], przy czym każdy obserwator (lub residuum) przypisany jest do jednego uszkodzenia [Patton, Frank, Clark, 1989]. Ważnym problemem jest rozdzielenie różnych efektów uszkodzeń [Patton, Frank, Clark, 1989]. Przykłady zastosowań liniowych modeli analitycznych do diagnostyki kontrolnej silników elektrycznych małej mocy zawierają np. prace [Filbert, Dreetz, 1986], [Barschdorff, Moczulski, Pfläging, 1986].

Bardzo ciekawe zastosowanie modeli matematycznych związane jest z *identyfikacją modeli odwrotnych*, które mogą być bezpośrednio stosowane do wnioskowania diagnostycznego [Cholewa, Kiciński, 1996], [Cholewa, Kiciński, 1997].

Model systemowy

Do określenia *modelu systemowego* danego obiektu stosuje się wybrane pojęcia ogólnej teorii systemów [Mesarovic, Takahara, 1975], w której centralną rolę spełnia pojęcie systemu R, określonego jako relacja na iloczynie kartezjańskim wejść i wyjść tego obiektu:

$$R \subset X \times Y. \tag{2.18}$$

Systemy większości obiektów technicznych są systemami zmiennymi w czasie "makro", który przyjmuje wartości z (liniowo uporządkowanego) zbioru czasu z elementem minimalnym $\theta_0 \in \Theta$. Można dla nich określić zbiory A, B stanowiące odpowiednio alfabet zbioru wejść X oraz wyjść Y, takie że:

$$X \subset A^{\Theta}, Y \subset B^{\Theta}, \tag{2.19}$$

przy czym poszczególne cechy wejść i wyjść spełniają związki:

$$\begin{aligned} &(\forall x_i \in X)\ [x_i : \Theta \to A], \\ &(\forall y_k \in Y)\ [y_k : \Theta \to B]. \end{aligned} \tag{2.20}$$

W ogólnym przypadku działanie obiektów technicznych opisuje się za pomocą systemu dynamicznego z pamięcią, realizowalnego fizycznie. Za [Mesarovic, Takahara, 1975] wprowadźmy następujące oznaczenia dla dowolnych przedziałów czasu, gdzie $\theta, \theta' \in \Theta \wedge \theta' > \theta$:

$$T_\theta = \{\theta' : \theta' \geqslant \theta\},\quad T^\theta = \{\theta' : \theta' < \theta\},\quad T_{\theta\theta'} = \{\theta^* : \theta \leqslant \theta^* < \theta'\},$$
$$\overline{T}_{\theta\theta'} = T_{\theta\theta'} \cup \{\theta'\},\quad \overline{T}^\theta = T^\theta \cup \{\theta\}$$

oraz dla dowolnego wejścia $x \in A^{\Theta}$ określając jego ograniczenia:

$$\begin{aligned} &x_\theta = x \mid T_\theta, && x^\theta = x \mid T^\theta, \\ &x_{\theta\theta'} = x \mid T_{\theta\theta'}, && \overline{x}_{\theta\theta'} = x \mid \overline{T}_{\theta\theta'}, \\ &\overline{x}^\theta = x \mid \overline{T}^\theta, && X_\theta = \{x_\theta : x_\theta = x \mid T_\theta \wedge x \in X\}, \\ &X^\theta = \{x^\theta : x^\theta = x \mid T^\theta \wedge x \in X\}, && X_{\theta\theta'} = \{x_{\theta\theta'} : x_{\theta\theta'} = x \mid T_{\theta\theta'} \wedge x \in X\}, \\ &X(\theta) = \{x(\theta) : x \in X\}. \end{aligned}$$

Ponadto wprowadza się operator sklejania "$\bullet$", który dla dowolnych wejść $u, v \in A^{\Theta}$ daje w wyniku nowe wejście $x = u^\theta \bullet v_\theta \in A^{\Theta}$, określone następująco:

$$x(\theta') = \begin{cases} u(\theta'), & \text{gdy } \theta' < \theta \\ v(\theta'), & \text{gdy } \theta' \geqslant \theta. \end{cases} \tag{2.21}$$

Niech działanie pewnego obiektu realizowalnego fizycznie będzie opisane za pomocą systemu dynamicznego R. Dla takiego obiektu można określić następujące rodziny zbiorów:

- rodzina $\overline{S}$ stanów obiektu zależnych od czasu:

$$\overline{S} = \{S_\theta \mid \theta \in \Theta\}, \tag{2.22}$$

- rodzina $\overline{\rho}$ odpowiedzi tego obiektu w danej chwili θ:

$$\overline{\rho} = \{\rho_\theta \mid \rho_\theta : S_\theta \times X_\theta \to Y_\theta \wedge \theta \in \Theta\}, \tag{2.23}$$

- rodzina $\overline{\phi}$ funkcji zmiany stanu tego obiektu, która zaszła w przedziale czasu $[\theta, \theta']$:

$$\overline{\phi} = \{\phi_{\theta\theta'} \mid \phi_{\theta\theta'} : S_\theta \times X_{\theta\theta'} \to S_{\theta'} \wedge \theta, \theta' \in \Theta \wedge \theta' > \theta\}, \tag{2.24}$$

przy czym funkcje (2.24) dodatkowo spełniają dla dowolnych $\theta, \theta' \in \Theta \wedge \theta' > \theta$:

$$\rho_\theta(s_\theta, x_\theta) \mid T_{\theta'} = \rho_{\theta'}(\phi_{\theta\theta'}(s_\theta, x_{\theta\theta'}), x_{\theta'}),\quad \text{gdzie}\quad x_\theta = x_{\theta\theta'} \bullet x_{\theta'}, \tag{2.25}$$

$$\phi_{\theta\theta'}(s_\theta, x_{\theta\theta'}) = \phi_{\theta''\theta'}(\phi_{\theta\theta''}(s_\theta, x_{\theta\theta''}), x_{\theta''\theta'}),\quad \text{gdzie}\quad x_{\theta\theta'} = x_{\theta\theta''} \bullet x_{\theta''\theta'}. \tag{2.26}$$

Rodziny funkcji (2.23), (2.24) stanowią *nośnik modelu systemowego danego obiektu*. Ich znajomość ma istotne znaczenie dla poszukiwania rozwiązania zadania diagnostycznego (por. rozdział 2.2).

Niektórzy autorzy (np. [Cholewa, Kaźmierczak, 1995]), w celu uproszczenia zadania diagnostycznego, zakładają, że rodzina funkcji odpowiedzi (2.23) oraz rodzina funkcji zmiany stanu obiektu (2.24) nie zależą (bezpośrednio) od czasu "makro" θ. Wówczas wymienione rodziny są jednoelementowe, a model reprezentowany przez te funkcje jest modelem wspólnym dla wszystkich obiektów $o \in O$. W celu zastosowania go do opisu wybranego obiektu w chwili θ należy określić wartości zmiennych stanu S_θ, które reprezentują szczególne własności rozpatrywanego obiektu (wyróżniające ten obiekt spośród innych obiektów) oraz zidentyfikować historię tego obiektu, reprezentowaną przez wartości jego wejść X_θ w przedziale czasu poprzedzającym tę chwilę (por. [Cholewa, Kaźmierczak, 1995]).

W ogólności modele obiektów diagnozowania są przyczynowo-skutkowe. Z punktu widzenia diagnostyki maszyn najbardziej istotne znaczenie ma przyczynowość czasowa, która w odniesieniu do ewolucji w funkcji czasu oznacza, że wartości wyjść obiektu w danej chwili czasu zależą wyłącznie od poprzedzających je wejść i stanów. Aby opisać wszystkie własności obiektu, należy w granicznym przypadku rozpatrzyć kompletne zbiory wejść X i wyjść Y, których liczba kardynalna jest zwykle bardzo wielka. Ze względu na ograniczone możliwości identyfikacji wejść i wyjść oraz ze względu na cel badań dokonuje się wyboru skończonego podzbioru wejść i wyjść, uwzględnianego podczas badań. Niekompletny zbiór wejść może wówczas spowodować, że uzyskany model nie będzie uznany za przyczynowo-skutkowy [Cholewa, Kaźmierczak, 1995].

W zbiorze stanów obiektu można wprowadzić *relację równoważności stanów* E_s [Mesarovic, Takahara, 1975], określoną na zbiorze wszystkich stanów:

$$\tilde{S} = \bigcup\{S_\theta \mid \theta \in \Theta\}. \tag{2.27}$$

Relacja $E_s \subset \tilde{S} \times \tilde{S}$ winna spełniać warunki:

$$(\forall[s])(\forall x_{\theta\theta'})(\exists[s'])(\forall s_\theta)(s_\theta \in [s] \Rightarrow \phi_{\theta\theta'}(s_\theta, x_{\theta\theta'}) \in [s']) \tag{2.28}$$

$$(\forall[s])(\forall s_\theta)(\forall s_\theta')(\forall x_\theta)(s_\theta \in [s] \wedge s_\theta' \in [s] \Rightarrow \rho_\theta(s_\theta, x_\theta) = \rho_\theta(s_\theta', x_\theta)). \tag{2.29}$$

Warunek (2.28) oznacza, że ewolucja obiektu w funkcji czasu zachodzi w sposób nieprzerwany, zaś warunek (2.29) oznacza, że dwa stany równoważne dają dla takiego samego wejścia identyczną odpowiedź. Przy ich spełnieniu relacja jest relacją równoważności, zaś zbiór $S = \tilde{S}/E_s$ stanowi przestrzeń stanów obiektu. Pojęcie przestrzeni stanów oraz klas stanów ma bardzo istotne znaczenie w diagnostyce maszyn (por. np. [Birger, 1978], [Cempel, 1989], [Cholewa, Kaźmierczak, 1995]).

Przykładem zastosowania modelu systemowego w diagnostyce maszyn jest *model symptomowy* ([Cempel, 1996], [Żółtowski, 1996]), który opisuje stan techniczny obiektu w funkcji obserwowanych symptomów, nie zawierających czasu dynamicznego t, lecz tylko czas θ [Żółtowski, 1996] — więc ujmuje procesy zużyciowe, zachodzące w obiekcie; znanym przykładem modelu symptomowego jest *model tribo-wibroakustyczny* (TWA — zob. [Cempel, 1989]). Szczególnym przypadkiem modelu symptomowego (o znaczeniu teoretycznym) jest *model holistyczny*, ujmujący jednocześnie dynamikę obiektu i zachodzące w nim procesy zużyciowe oraz procesy opisujące "zachowanie się obiektu w całym cyklu życia, od zamysłu koncepcji aż do likwidacji, wsparte modelowo" ([Żółtowski, 1996], s.166 oraz s.170 i n.).

Model logiczny

Inny sposób określenia modelu obiektu polega na zastosowaniu aparatu logiki matematycznej. Istotą tego postępowania jest rozważanie stwierdzeń o danym obiekcie (abstrakcie lub konkrecie) oraz wniosków, które można wyprowadzić o własnościach i zachowaniu się (działaniu) tego obiektu [Mesarovic, Takahara, 1975].

Niech F będzie językiem formalnym, zaś f zbiorem zdań w języku F, wyrażających zaobserwowane fakty lub zakładane własności układu. Niech f będzie wyczerpujący w stopniu wynikającym z naszej wiedzy, dotyczącej danego obiektu, tj. zawiera wszystkie znane lub zakładane fakty, dotyczące obiektu. Wówczas "system jest odpowiednim zbiorem dobrze zbudowanych zdań". Tak określony system jest nośnikiem *modelu logicznego* obiektu.

Wg [Mesarovic, Takahara, 1975] posiadana wiedza dotycząca danego obiektu zwykle wystarcza jedynie do utworzenia jego modelu logicznego. Takie podejście stanowi początkowy krok w odniesieniu do formalizowania zgromadzonych obserwacji.

Bardzo ważną koncepcję teoretyczną zastosowania modeli logicznych do opisu działania diagnozowanego obiektu zawiera praca [Reiter, 1987], dotycząca teorii *diagnozowania na podstawie pierwotnych zasad* (ang. *diagnosis from first principles*). Stosowany model o bardzo wysokim poziomie ogólności zapisywany jest za pomocą zdań rachunku predykatów. Służy on do opisu działania układu. Do wskazania, czy dane działanie ma miejsce w stanie zdatnym, czy niezdatnym, stosowany jest specjalny predykat $AB(\ )$ (od ang. *abnormal*). "Zachowanie się" diagnozowanego obiektu opisywane jest za pomocą wartości wejść i wyjść z uwzględnieniem stanu poszczególnych elementów. Należy podkreślić, że opis układu, wykaz jego elementów oraz obserwacje jego zachowania się są jedynymi danymi, dostępnymi dla rozwiązującego problem diagnostyczny.

Jeśli obserwacje danego obiektu nie są zgodne z opisem jego "zachowania się" (tj. obserwacje są *w konflikcie* z tym opisem), oznacza to *niesprawność danego obiektu*. Wtedy należy określić *diagnozę*, która jest podzbiorem elementów, przypuszczalnie uszkodzonych. Poszukiwanie diagnoz odbywa się na podstawie analizy zbioru konfliktów pomiędzy opisem działania obiektu a zgromadzonymi obserwacjami dotyczącymi zachowania się tego obiektu. Pomocą w zidentyfikowaniu minimalnego zbioru diagnoz jest przycięte drzewo (ang. *pruned tree*) konfliktów.

Opisywana teoria Reitera umożliwia zarówno określenie diagnoz *pojedynczych uszkodzeń*, jak i diagnoz *uszkodzeń złożonych*. Uszczegółowienie diagnoz złożonych może nastąpić wskutek zgromadzenia nowych obserwacji (zwanych *pomiarami*).

Teoria diagnozy z pierwotnych zasad, choć posiada znaczną moc heurystyczną, znajduje zastosowanie do budowy modeli jedynie bardzo prostych układów, o niewielu elementach. W pracy [Reiter, 1987] przedstawiono przykłady modeli bardzo prostych układów cyfrowych, jak np. *sumator*. Opis poprawnego działania układów mechanicznych, dokonany za pomocą rachunku predykatów, przysparza prawdopodobnie znacznych trudności, co uniemożliwia praktyczne stosowanie tej metody w diagnostyce maszyn.

Model jakościowy

Układy, których działanie może być opisane za pomocą systemu dynamicznego, mogą być także modelowane w sposób jakościowy. W dalszej części omówiono dwa odrębne podejścia, z których każde ma wiele zastosowań diagnostycznych, a mianowicie:

1. Modelowanie jakościowe z zastosowaniem *logiki wielowartościowej* (użyte między innymi w diagnostycznym systemie doradczym *KARDIO* — por. [Bratko, 1989], [Bratko, Mozetic, Lavrac, 1989]);

2. Modelowanie jakościowe działania obiektów dynamicznych z zastosowaniem *metody QSIM* (*Qualitative SIMulation* [Kuipers, 1994]).

Model jakościowy w systemie KARDIO zapisywany jest w postaci zbioru reguł i ma charakter hierarchiczny. Dla każdego poziomu abstrakcji dany model hierarchiczny może być interpretowany jako relacja:

$$model(Poziom, Stany, Symptomy), \tag{2.30}$$

gdzie *Poziom* jest liczbą naturalną oznaczającą poziom abstrakcji modelu (*Poziom* = 1 oznacza najwyższy poziom abstrakcji). Wprowadzona w ten sposób hierarchia umożliwia sterowanie przeszukiwaniem zbioru reguł. Modele odpowiadające dwóm sąsiednim poziomom abstrakcji są ze sobą zgodne w tym sensie, że model na wyższym poziomie abstrakcji jest uzyskiwany w wyniku takich operacji abstrahowania, jak:

- zastąpienie zbioru powiązanych ze sobą zmiennych (cech) poprzez jedną cechę;
- pominięcie zmiennej, która występuje w modelu bardziej szczegółowym;
- zmniejszenie liczby rozróżnianych wartości jakościowych danej zmiennej.

W wypadku wnioskowania diagnostycznego zadanie jest wstępnie rozwiązywane na wyższym poziomie abstrakcji, a znalezione tam rozwiązanie ukierunkowuje następnie poszukiwania na poziomie bardziej szczegółowym.

Model jakościowy danego układu jest identyfikowany w sposób indukcyjny na drodze tzw. "uczenia maszynowego" na podstawie przykładów i kontrprzykładów (por. rozdział 3.2). Model jakościowy jest określony poprzez:

- zbiór elementów składowych oraz połączeń pomiędzy tymi elementami, które określają strukturę modelu;
- szczegółową definicję języka opisu obiektu (dziedzin poszczególnych zmiennych).

Opisana metoda może być stosowana do modelowania jakościowego bardzo złożonych obiektów diagnozowania, jak np. serce ludzkie [Bratko, 1989]. Jej ważną zaletą jest możliwość stosowania hierarchii modeli, co pozwala na reprezentowanie tzw. *wiedzy głębokiej*.

Metoda QSIM wywodzi się z tradycyjnego podejścia do modelowania działania obiektów za pomocą układów równań różniczkowych, w których występują zmienne o charakterze ilościowym. Do modelowania obiektów o cechach zmieniających się w czasie (układów dynamicznych z pamięcią) zastosowano *jakościowe równania różniczkowe* (por. [Kuipers, 1994], [Subramanian, Mooney, 1996]). Równanie to określone jest przez zbiór zmiennych (którymi są cechy obiektu), dziedzin tych zmiennych oraz zbiór ograniczeń jakościowych, zachodzących pomiędzy poszczególnymi zmiennymi, przy czym:

- Dziedzina danej cechy reprezentowana jest w sposób jakościowy za pomocą ciągu wartości, które mogą być wyróżnione spośród pozostałych wartości dziedziny (ang. *landmark values*). Skrajne wartości stanowią jakościowe ograniczenia zmienności danej cechy. Bieżąca wartość cechy może być albo wyróżnioną wartością jakościową, albo być zawarta w przedziale o końcach określonych przez dwie wartości jakościowe;
- Ograniczenia jakościowe opisują powiązania zmiennych z innymi zmiennymi w postaci zależności arytmetycznych lub pochodnych w funkcji czasu (np. gdy wartości danej zmiennej rosną monotonicznie, wartości związanej z nią zmiennej również rosną monotonicznie). Możliwe jest także określanie związków, które zachodzą pomiędzy ograniczeniami jakościowymi więcej niż jednej zmiennej;
- Zbiór chwil czasu jest zbiorem dyskretnym punktów czasu, przyporządkowanych zajściu "interesujących zdarzeń".

Model dynamiczny obiektu reprezentowany jest w postaci:

- układu jakościowych równań różniczkowych,
- stanu początkowego modelowanego obiektu, określonego co najmniej częściowo (poprzez przypisanie cechom jakościowym obiektu ich wartości początkowych),
- drzewa, opisującego wszystkie możliwe jakościowe "zachowania" obiektu, interpretowane jako ciągi zmian stanów tych obiektów w odpowiedzi na różne możliwe historie wejść tego obiektu.

Zaletą metody *QSIM* jest łatwość budowy modelu jakościowego obiektu, dla którego znany jest tradycyjny model w postaci układu równań różniczkowych. Model *QSIM* może być także utworzony w sposób intuicyjny w odniesieniu do obiektów, których istota działania jest znana i nie jest zbyt skomplikowana. Zastosowanie tej metody do modelowania działania bardzo złożonych obiektów jest ograniczone ze względu na ogromną złożoność uzyskanego modelu. Ponadto *QSIM* generuje drzewo zachowań, którego liczba gałęzi rośnie kombinatorycznie, przy czym generowanych jest wiele gałęzi, których zasadności nie można wykluczyć ze względu na brak danych (por. [Bratko, 1989]).

Modele przybliżone

Wiele czynników, jak np.:

- losowość własności tworzyw oraz procesów wytwórczych,
- zmienność warunków fundamentowania maszyn,
- losowość przebiegu procesów użytkowania i obsługiwania tych maszyn,
- świadome ograniczenie się w procesie obserwacji wejść i wyjść obiektu do pomiaru i rejestracji skończonej liczby sygnałów,
- ograniczona dokładność estymacji cech sygnałów,
- niepewność, niezupełność oraz pewna sprzeczność wiedzy diagnostycznej,

a także inne jeszcze czynniki uniemożliwiają jednoznaczne określenie aktualnych wyjść obiektu na podstawie znajomości historii wejść oraz stanu tego obiektu. Z tych względów stosowane są modele przybliżone.

W odniesieniu do modeli systemowych stosuje się np. modele rozmyte, które są rozmytym zbiorem modeli dokładnych [Cholewa, 1983], [Cholewa, Kaźmierczak, 1995]. Modele takie umożliwiają jedynie określenie rodzin możliwych wyjść, odpowiadających aktualnemu stanowi obiektu oraz historii jego wejść.

Przybliżony model logiczny może być reprezentowany w postaci zdań logicznych, których stopień prawdziwości jest liczbą z przedziału $[0,1]$ (por. [Cholewa, 1993]) lub jest określany w logice M-wartościowej (zob. [Pacholczyk, Hunault, Pacholczyk, 1996]).

W podobny sposób można wprowadzić przybliżone modele jakościowe, zwłaszcza w razie ich reprezentacji za pomocą reguł logicznych.

2.1.6 Specyficzne cechy modeli diagnozowanych obiektów

Jak wspomniano, *obiektem diagnozowania jest maszyna obserwowana podczas procesu wytwarzania i/lub eksploatacji* (obejmującej jej *użytkowanie* i *obsługiwanie*).

Struktury i stany uzyskane w procesie wytwarzania oceniane są pod względem ich *zgodności z konstrukcją* [Dietrych, 1985], co jest przedmiotem *diagnostyki kontrolnej* [Cempel, 1989].

W poszczególnych chwilach czasu "makro" θ maszynie tej przypisane są *stany techniczne*.

Podczas eksploatacji obiektu w ramach jego użytkowania lub obsługiwania mogą wystąpić *niesprawności* i *uszkodzenia* tego obiektu, oznaczające *zmianę stanu eksploatacyjnego* oraz *zmianę stanu technicznego*. Oba te zdarzenia eksploatacyjne mogą zachodzić [Hebda, 1990]:

- *stopniowo*, np. wskutek korozji [Lawrowski, 1993] podczas przechowywania obiektu oczekującego na użytkowanie, zużycia ściernego [Chmurawa, Bińkowski, 1980], [Hebda, Wachal, 1980] lub niszczenia zmęczeniowego [Kocańda, Szala, 1991], [Kocańda, 1978],

- *nagle*, np. wskutek przekroczenia przez naprężenia wartości krytycznej wytrzymałości na złamanie lub wystąpienia tzw. przełomu doraźnego części wałka, który wcześniej uległ pęknięciu zmęczeniowemu [Kocańda, Szala, 1991],

- *w sposób przemijający*, np. w formie przemijających nieregularności działania maszyny [Isermann, Ballé, 1997], co może mieć miejsce w przypadku obrabiarki przy stopniowej zmianie wartości parametrów skrawania powodującej, że proces skrawania staje się niestateczny i w obrabiarce narastają amplitudy drgań samowzbudnych [Łączkowski, 1983].

W wypadku zmęczenia tworzywa należy zwrócić uwagę na istotne zjawisko *kumulacji uszkodzeń*, które zależy między innymi od *historii obciążeń* oraz *przebiegu zmian obciążeń* [Kocańda, Szala, 1991].

Dla pojedynczego konkretnego obiektu diagnozowania (maszyny) można budować *charakterystyki czasowe procesu eksploatacji* [Adamkiewicz *et al.*, 1983], będące zapisem historii stanów eksploatacyjnych (zdatność eksploatacyjna, obejmująca zdatność w czasie działania oraz zdatność w czasie postoju, a także niezdatność eksploatacyjna — oczekiwanie na obsługę/odnowę oraz przeprowadzenie obsługi/odnowy). Poszczególnym zmianom stanów eksploatacyjnych mogą być przypisane stany techniczne w następujący sposób:

- przejście ze stanu zdatności do stanu niezdatności jest związane z wystąpieniem uszkodzenia, które może być reprezentowane przez zmianę stanu technicznego;

- przejście ze stanu niezdatności do stanu zdatności jest związane z przeprowadzeniem obsługi lub odnowy i może być (pośrednio) reprezentowane również przez zmianę stanu technicznego.

Jeśli diagnozowaną maszynę rozpatrujemy jako obiekt złożony (w wypadku obiektów naprawialnych), zwykle istnieje możliwość wyróżnienia w niej *zespołów funkcjonalno-obsługowych* (ang. *segment* [MIMOSA, 1996a]). Zespół taki jest w istocie obiektem abstrakcyjnym, którego wejścia, wyjścia, działanie i umiejscowienie w strukturze obiektu

określane są na etapie projektowania danego środka technicznego. W maszynie-konkrecie, zrealizowanej zgodnie z projektem, abstrakcyjne zespoły zastępowane są przez *egzemplarze-konkrety* (ang. *asset* [MIMOSA, 1996a]). W ramach napraw lub remontów kompletnej maszyny może następować naprawa danego zespołu, polegająca na wymianie konkretnego egzemplarza. Każdy z egzemplarzy, stosowany jako dany zespół, posiada własną historię eksploatacji, która może być reprezentowana przez ciągi stanów technicznych, historie wejść, charakterystyki czasowe procesów eksploatacji itp. Każdemu z egzemplarzy można więc przyporządkować jego własną oś czasu "makro".

Warto także zwrócić uwagę na fakt, że stan kompletnej maszyny określony jest w danej chwili czasu eksploatacji θ przez połączenie zbioru stanów egzemplarzy zespołów, które w tej chwili czasu są zainstalowane w tej maszynie. W ramach jednej organizacji (np. przedsiębiorstwa), eksploatującej większą liczbę maszyn tej samej konstrukcji, może pomiędzy poszczególnymi megaukładami maszynowymi (zwanymi też agregatami [Łączkowski, 1983]) następować przemieszczanie egzemplarzy zespołów, związane z prowadzoną gospodarką obsługowo-remontową. Np. w rafinerii eksploatującej kilkaset pomp napędzanych silnikami elektrycznymi tej samej konstrukcji wydaje się być racjonalne posiadanie kilku silników dodatkowych, które w razie wystąpienia uszkodzenia jednej z pomp mogą umożliwić niezwłoczną wymianę, zmniejszając przestój awaryjny. Stąd w najbardziej ogólnym przypadku zapisu aktualnej historii danego megaukładu maszynowego (w jego konfiguracji odpowiadającej chwili czasu θ) wynika potrzeba reprezentowania czasu w postaci struktur nieliniowych (por. [Hajnicz, 1996]).

Podsumowując powyższe wywody, należy stwierdzić, że konkretnie istniejący megaukład maszynowy jako obiekt diagnozowania jest bardzo złożonym układem, któremu w danej chwili czasu eksploatacji θ można przyporządkować:

- *aktualną konfigurację* K_θ egzemplarzy swoich zespołów użytkowych,

- *zbiór stanów* $S_\theta = S(\theta, K_\theta)$, odpowiadających aktualnej konfiguracji egzemplarzy zespołów użytkowych, przy czym zmiany stanów mogą zachodzić w sposób stopniowy (jak w wypadku zużycia trybologicznego) lub w sposób nagły (wskutek przeciążenia układu lub np. wskutek wymiany egzemplarza danego zespołu),

- *historię wejść* $X_\theta = X(\theta, K_\theta)$, uwarunkowaną aktualną konfiguracją egzemplarzy zespołów użytkowych (a więc układ ten posiada pamięć).

Z powyższych rozważań wynika, że działanie megaukładu maszynowego (zwanego dalej maszyną) winno być reprezentowane przez system dynamiczny z pamięcią. Dla ułatwienia budowy modeli diagnostycznych obiektu zakładamy, że nie będzie rozpatrywana wymiana egzemplarzy zespołów danej maszyny, lecz wynikająca z niej zmiana własności kompletnej maszyny będzie każdorazowo uwzględniana w zbiorach wejść, wyjść i stanów.

2.2 Zadanie diagnostyki maszyn

Celem diagnostyki technicznej maszyn jest identyfikacja ich szeroko rozumianego stanu technicznego za pomocą odpowiednich metod i środków w celu osiągnięcia ich optymalnej trwałości, niezawodności i sprawności działania [Cempel, 1989]. Podstawowym zadaniem diagnostyki maszyn jest *ocena stanu technicznego* danej maszyny na podstawie zgromadzonych *symptomów diagnostycznych* (czyli wartości cech sygnałów diagnostycznych) oraz

(często niepełnych) *cech wejścia maszyny* (przede wszystkim jej *warunków działania*). Dodatkowo wymaga się, by obserwacja diagnozowanego obiektu nie zakłóciła w istotny sposób jego działania (najlepiej aby była prowadzona bez demontażu).

Sformułowane powyżej ogólne zadanie diagnostyki maszyn można podzielić na dwa podzadania o różnym stopniu szczegółowości diagnozy:

- *Monitorowanie*, polegające na zgromadzeniu wystarczającej ilości informacji o danym obiekcie, pozwalającej na ocenę ogólnego stanu tego obiektu w danej chwili, a zwłaszcza identyfikację stanu niezdatnego;
- *Diagnozowanie*, polegające na identyfikacji uszkodzonych lub niesprawnych zespołów tego obiektu (właściwa *diagnoza*) oraz określeniu prawdopodobnych przyczyn wystąpienia (*geneza*) i przebiegu lub rozwoju tych niesprawności albo uszkodzeń, zarówno do chwili badania, jak i ich dalszego rozwoju w przyszłości (*prognoza*).

Stosując oznaczenia z rozdziału 2.1, można sformułować następujące:

OGÓLNE ZADANIE DIAGNOSTYKI TECHNICZNEJ

1. Niech:

$$R \subset X \times Y \ \wedge \ X = X_1 \cup X_2 \cup X_3 \ \wedge \ Y = Y_1 \cup Y_2 \cup Y_3 \tag{2.31}$$

będzie systemem dynamicznym, opisującym działanie diagnozowanego obiektu (zob. s. 20), oraz niech model tego obiektu będzie określony przez:

(a) rodzinę $\overline{S}$ stanów tego układu zależnych od czasu (do rodziny tej należy także stan S_1, określony przez konstrukcję maszyny):

$$\overline{S} = \{S_1\} \cup \{S_\theta \mid \theta \in \Theta\}, \tag{2.32}$$

(b) rodzinę funkcji odpowiedzi tego układu:

$$\overline{\rho} = \{\rho_\theta \mid \rho_\theta : (S_\theta, X_{1,\theta}, X_{2,\theta}) \to (Y_{1,\theta}, Y_{2,\theta}) \wedge \theta \in \Theta\}, \tag{2.33}$$

(c) rodzinę funkcji zmiany stanu tego układu:

$$\overline{\phi} = \{\phi_{\theta\theta'} \mid \phi_{\theta\theta'} : (S_\theta, X_{1,\theta\theta'}, X_{2,\theta\theta'}) \to S_{\theta'} \wedge \theta, \theta' \in \Theta \wedge \theta' > \theta\}. \tag{2.34}$$

2. Dla $\theta_0, \theta, \theta' \in \Theta, \theta' > \theta \geqslant \theta_0$ (gdzie θ_0 jest elementem minimalnym zbioru czasu Θ) mogą być dane:

(a) historia $\overline{X}_{1,\theta\theta'}$ wejść związanych z użytecznym działaniem maszyny na odcinku czasu $[\theta, \theta']$,

(b) historia $\overline{X}_{2,\theta\theta'}$ warunków działania maszyny na odcinku czasu $[\theta, \theta']$,

(c) historia $\overline{Y}_{1,\theta\theta'}$ wyjść związanych z użytecznym działaniem maszyny na odcinku czasu $[\theta, \theta']$,

(d) historia $\overline{Y}_{2,\theta\theta'}$ wyjść związanych z występowaniem procesów resztkowych na odcinku czasu $[\theta, \theta']$.

3. Określić:

(a) ogólną diagnozę stanu maszyny (dla klasy stanu $[s]$):

$$\Delta_{og} : \Theta \times X_1 \times X_2 \times Y_1 \times Y_2 \to \tilde{S}/E_s \tag{2.35}$$

$$\Delta_{og} : (\overline{X}_{1,\theta\theta'}, \overline{X}_{2,\theta\theta'}, \overline{Y}_{1,\theta\theta'}, \overline{Y}_{2,\theta\theta'}) \to [s], \tag{2.36}$$

gdzie $\tilde{S}$ jest zbiorem wszystkich stanów (2.27), E_s jest relacją równoważności stanów spełniającą (2.28), (2.29), zaś $S = \tilde{S}/E_s$ jest przestrzenią klas stanów; lub

(b) uszczegółowioną diagnozę stanu maszyny (stan w chwili θ'):

$$\Delta_{det} : \Theta \times X_1 \times X_2 \times Y_1 \times Y_2 \to S \tag{2.37}$$

$$\Delta_{det} : (\overline{X}_{1,\theta\theta'}, \overline{X}_{2,\theta\theta'}, \overline{Y}_{1,\theta\theta'}, \overline{Y}_{2,\theta\theta'}) \to S_\theta, \tag{2.38}$$

lub wreszcie:

(c) przewidzieć sekwencję kolejnych stanów maszyny:

$$S_{\theta_1}, S_{\theta_2}, \ldots, S_{\theta_p} \in \tilde{S} \wedge \theta_1, \theta_2, \ldots, \theta_p \in \Theta \wedge \theta_1 < \theta_2 < \ldots < \theta_p \tag{2.39}$$

oraz prędkość ewolucji tych stanów.

Wyżej zapisane zależności zwykle zamiast dla wejść, wyjść i stanów są zapisywane dla cech wejść, cech wyjść i cech stanów (zob. [Cholewa, 1983], [Cholewa, Kaźmierczak, 1995]). Zwróćmy uwagę na fakt, że rozpatruje się także wejścia $x \in X_3$ i/lub wyjścia $y \in Y_3$, obserwowane w obecności zakłóceń.

W sformułowanym powyżej modelu należy wyróżnić dwa odmienne rodzaje związków:

- Zależności (2.33), (2.34) określają związki przyczynowo-skutkowe (dla obiektów, których działanie można opisać za pomocą systemów przyczynowo-skutkowych);
- Zależności (2.35), (2.37) stanowią podstawę wnioskowania diagnostycznego i zwane będą w dalszym ciągu pracy *relacjami diagnostycznymi*. Mają one w pewnym sensie charakter relacji odwrotnych, na co zwrócono uwagę np. w [Cholewa, White, 1993], [Cholewa, Kiciński, 1996], [Moczulski, Żytkow, 1997].

Z zagadnieniem odwracania relacji dla potrzeb wnioskowania diagnostycznego wiąże się możliwość niejednoznaczności relacji "odwrotnych", wynikająca z braku wzajemnej jednoznaczności funkcji (2.33), (2.34). Skutkiem takiej niejednoznaczności jest to, że np. dla ustalonych wejść $X_{1,\theta}$, $X_{2,\theta}$ (traktowanych jako zadane) i dla obserwowanych wyjść $Y_{1,\theta}$, $Y_{2,\theta}$ zbiory:

$$(\forall X_{1,\theta})(\forall X_{2,\theta})[\hat{S}_\theta = \{s \in S_\theta \mid \rho_\theta(s, x_\theta) = y_\theta \wedge x_\theta \in X_{1,\theta} \cup X_{2,\theta} \wedge y_\theta \in Y_{1,\theta} \cup Y_{2,\theta}\}] \tag{2.40}$$

nie są jednoelementowe. W związku z powyższym diagnoza ma zwykle charakter przybliżony (jest *funkcją w sensie uogólnionym* lub *przyporządkowaniem* — por.

[Kuratowski, 1973]). Ponadto na przybliżony charakter rozwiązania wpływają występujące szumy $X_{3,\theta}$, $Y_{3,\theta}$ odpowiednio na wejściu i wyjściu.

Zwykle obiekty diagnozowania obserwowane są w warunkach tzw. *biernego eksperymentu diagnostycznego* ([Cempel, 1982]), którego istotą jest "badanie obiektu w naturalnych warunkach jego pracy, bez znajomości cech stanu (...), a jedynie przy założeniu jego sposobu zużywania się" ([Żółtowski, 1996]). Oznacza to, że zadanie diagnostyczne należy w tym wypadku sformułować natępująco:

ZADANIE DIAGNOSTYKI TECHNICZNEJ (W EKSPERYMENCIE BIERNYM)

1. Niech:

$$R \subset X \times Y \ \wedge \ X = X_1 \cup X_2 \cup X_3 \ \wedge \ Y = Y_1 \cup Y_2 \cup Y_3 \tag{2.41}$$

będzie systemem dynamicznym, opisującym działanie diagnozowanego obiektu, oraz niech model tego obiektu jest określony przez:

   (a) rodzinę $\overline{S}' \subset \overline{S}$ stanów tego układu, gdzie $\overline{S}$ wg (2.32),

   (b) rodzinę $\overline{\rho}' \subset \overline{\rho}$ funkcji odpowiedzi tego układu, gdzie $\overline{\rho}$ wg (2.33); mogą być np. znane funkcje odpowiedzi dla właściwych podzbiorów zbiorów wejść i wyjść, odpowiednio $X_{\theta'} \subset X_\theta \wedge X_{\theta'} \neq X_\theta$ oraz $Y_{\theta'} \subset Y_\theta \wedge Y_{\theta'} \neq Y_\theta$,

   (c) rodzinę $\overline{\phi}' \subset \overline{\phi}$ funkcji zmiany stanu tego układu, gdzie $\overline{\phi}$ wg (2.34); funkcje te mogą być określone np. na podstawie ogólnej znajomości przebiegu i zaawansowania procesów zużycia, zachodzących w danym obiekcie (zob. np. [Żółtowski, 1996]).

2. Dla $\theta_0, \theta, \theta' \in \Theta, \theta' > \theta \geqslant \theta_0$ (gdzie θ_0 jest elementem minimalnym zbioru czasu Θ) mogą być dane:

   (a) historia $\overline{X'}_{1,\theta\theta'}, X_1{}' \subset X_1$ wejść związanych z użytecznym działaniem maszyny na odcinku czasu $[\theta, \theta']$,

   (b) historia $\overline{X'}_{2,\theta\theta'}, X_2{}' \subset X_2$ warunków działania maszyny na odcinku czasu $[\theta, \theta']$,

   (c) historia $\overline{Y'}_{1,\theta\theta'}, Y_1{}' \subset Y_1$ wyjść związanych z użytecznym działaniem maszyny na odcinku czasu $[\theta, \theta']$,

   (d) historia $\overline{Y'}_{2,\theta\theta'}, Y_2{}' \subset Y_2$ wyjść związanych z występowaniem procesów resztkowych na odcinku czasu $[\theta, \theta']$.

3. Określić:

   (a) ogólną (przybliżoną) diagnozę stanu maszyny (dla klasy stanu $[s]$):

$$\Delta_{og} : \Theta \times X'_1 \times X'_2 \times Y'_1 \times Y'_2 \tilde{\to} \tilde{S}/E_s \tag{2.42}$$

$$\Delta_{og} : (\overline{X'}_{1,\theta\theta'}, \overline{X'}_{2,\theta\theta'}, \overline{Y'}_{1,\theta\theta'}, \overline{Y'}_{2,\theta\theta'}) \tilde{\to} [s], \tag{2.43}$$

   lub

   (b) uszczegółowioną (przybliżoną) diagnozę stanu maszyny (stan w chwili czasu θ):

$$\Delta_{det} : \Theta \times X'_1 \times X'_2 \times Y'_1 \times Y'_2 \tilde{\to} S \tag{2.44}$$

$$\Delta_{det} : (\overline{X'}_{1,\theta\theta'}, \overline{X'}_{2,\theta\theta'}, \overline{Y'}_{1,\theta\theta'}, \overline{Y'}_{2,\theta\theta'}) \tilde{\to} S_\theta. \tag{2.45}$$

W rozpatrywanym zadaniu dokonuje się wyboru wejść i wyjść, za pomocą których obserwuje się oddziaływanie otoczenia na obiekt oraz obiektu na otoczenie. Dopuszcza się ponadto identyczność chwili początkowej θ i chwili końcowej θ', co oznacza, że historie te mogą się redukować do zbiorów $X_1(\theta'), X_2(\theta'), Y_1(\theta'), Y_2(\theta')$. Nie wszystkie z tych historii są wymagane: aby umożliwić wnioskowanie diagnostyczne, niezbędne są dane dotyczące historii niektórych wyjść związanych z występowaniem procesów resztkowych $Y_2(\theta')$.

Dostępność danych i wiedzy, odpowiadających eksperymentowi biernemu, powoduje kolejne ograniczenie dokładności rozwiązania zadania. W granicznym przypadku zadowalającym wynikiem jest zidentyfikowanie wystąpienia istotnej zmiany wartości cech obiektu, bez możliwości zinterpretowania znaczenia tej zmiany w kategoriach klas stanów technicznych diagnozowanego obiektu.

Stawianie diagnozy w diagnostyce maszyn może być interpretowane jako zadanie klasyfikacji (np. [Cholewa, Kaźmierczak, 1995]). Zakłada się, że (niezależnie od prowadzonych badań) populacja O maszyn może zostać podzielona na podzbiory O_k, $k = 1, \ldots, K$ odpowiadające wyróżnionym klasom stanu technicznego tych maszyn — por. (2.28), (2.29):

$$O = \bigcup_{k=1,\ldots,K} O_k \tag{2.46}$$

Należy podkreślić, że w ogólnym przypadku nie można zakładać, że klasy O_k, $k = 1, \ldots, K$ są rozłączne, co spowodowane jest często spotykaną jednoczesnością występowania niesprawności i uszkodzeń. Nie można więc określić relacji równoważności. Niejednokrotnie istnieje zależność przyczynowo-skutkowa pomiędzy uszkodzeniami, polegająca na tym, że jedno uszkodzenie powoduje wystąpienie następnego. W diagnostyce maszyn zależności takie mogą być reprezentowane np. za pomocą *drzew uszkodzeń* ([Giordana *et al.*, 1993], [Cholewa, Kaźmierczak, 1995]). Należy zwrócić także uwagę na możliwość wystąpienia zjawiska *synergii*, które w wypadku uszkodzeń oznacza, że łączny skutek występowania dwóch (skorelowanych) przyczyn może być istotnie większy (lub — mniejszy) niż suma skutków tych przyczyn.

W tak określanym zadaniu klasyfikacji podział populacji O może być określony za pomocą *rodziny klasyfikatorów* [Cholewa, Kaźmierczak, 1995]:

$$c_k : O \to [0, 1]; \ k = 1, \ldots, K, \tag{2.47}$$

gdzie:

$$c_k(o) = 1 \ \Leftrightarrow \ o \in O_k, \tag{2.48}$$

$$0 < c_k(o) < 1 \ \Leftrightarrow \ o \ \tilde{\in} \ O_k, \tag{2.49}$$

$$c_k(o) = 0 \ \Leftrightarrow \ o \notin O_k, \tag{2.50}$$

zaś symbol $\tilde{\in}$ oznacza, że element *częściowo* należy do danego zbioru.

W niektórych zadaniach diagnostycznych, ze względu na dynamikę zmian stanu zachodzących w obiekcie, pojawia się potrzeba udzielenia odpowiedzi w postaci konkluzji w pewnym ograniczonym czasie $(\Delta t)_{diag}$. Wymaganie to jest szczególnie ważne w wypadku tzw. *dynamicznych systemów doradczych*, działających w rzeczywistej skali czasu i realizujących zadania monitorowania w odniesieniu do obiektów "krytycznych" (zob. s. 35), gdzie wprowadza się pojęcie tzw. *czasu gwarantowanego* [Cholewa, 1996].

2.3 Środki wspomagania wnioskowania diagnostycznego

Wnioskowanie diagnostyczne jest *działaniem na informacji*. Warunkiem koniecznym skutecznego wnioskowania jest posiadanie odpowiedniej wiedzy. Efektywność procesu wnioskowania może być zwiększona dzięki zastosowaniu odpowiednich środków wspomagania, jak:

- *drzewa uszkodzeń*,
- *drzewa sprawdzeń*,
- *tablice decyzyjne reguł*,
- *systemy doradcze*.

2.3.1 Drzewa uszkodzeń

Drzewa uszkodzeń (zob. [Pau, 1981], [Schneider-Fresenius, 1985], [Giordana *et al.*, 1993]) stosowane są w badaniach, dotyczących niezawodności środków technicznych ([Adamkiewicz *et al.*, 1983], [Hebda, 1990], [Oprzędkiewicz, 1993]). Umożliwiają reprezentację zależności przyczynowo-skutkowych pomiędzy uszkodzeniami, co pozwala na odtworzenie tzw. *ścieżki uszkodzeń*.

Uszkodzenia lub niesprawności elementów układu często mogą być przyczyną wystąpienia uszkodzenia innego elementu. Np. uszkodzenie pompy olejowej w przekładni zębatej z wymuszonym obiegiem oleju może być jednocześnie przyczyną uszkodzenia zazębień i łożysk.

Drzewo uszkodzeń budowane jest dla obiektu złożonego, który jest poddawany dekompozycji na podobiekty (podzespoły) ze względu na zidentyfikowane relacje przyczynowo-skutkowe zachodzące pomiędzy stanami podobiektów i stanem obiektu nadrzędnego. Uszkodzenia lub niesprawności oraz ich przyczyny są rozpatrywane jako *zdarzenia* [Cholewa, Kaźmierczak, 1995]. Drzewo uszkodzeń może być reprezentowane za pomocą odpowiedniego grafu, w którym:

- zdarzenia są reprezentowane za pomocą *węzłów* (którym mogą być przypisane wagi, stanowiące np. prawdopodobieństwa wystąpienia danego zdarzenia pod warunkiem zajścia innych zdarzeń),
- związki przyczynowo-skutkowe pomiędzy zdarzeniami są reprezentowane za pomocą *krawędzi grafu*, którym również mogą być przypisane wagi (np. prawdopodobieństwa wystąpienia danego związku).

W praktyce do określania wag (w przypadku *ważonego drzewa uszkodzeń*) stosowane mogą być *prawdopodobieństwa subiektywne*, określane przez specjalistę.

Drzewa uszkodzeń umożliwiają wnioskowanie o *zdarzeniach podstawowych*, którymi są uszkodzenia określonych elementów, a więc możliwe *przyczyny* obserwowanego *skutku* — *zdarzenia złożonego*, będącego przedmiotem obserwacji diagnostycznej. Drzewa te budowane są więc w sposób *zstępujący* (ang. *top-down*), lecz stosowane są w sposób *wstępujący* (ang. *bottom-up*), ponieważ kierunek wnioskowania jest przeciwny do zwrotu krawędzi grafu, reprezentującego dane drzewo uszkodzeń [Pau, 1981].

2.3.2 Drzewa sprawdzeń

Drzewa sprawdzeń (por. [Pau, 1981], [Cempel, Tomaszewski, 1992]) umożliwiają reprezentowanie kolejności przeprowadzania testów i badań, co jest typowe dla *diagnostyki sekwencyjnej*. Mogą one być przedstawiane za pomocą grafów, których węzły odpowiadają poszczególnym procedurom testowym, zaś krawędzie określają sekwencje czynności związanych z postępowaniem diagnostycznym. Drzewo sprawdzeń jest zapisem sposobu działania zgodnego z *metodą zstępującą*: postępowanie jest "od ogółu", a więc od stwierdzenia niesprawności danego obiektu, "do szczegółu", tj. do zidentyfikowania *pierwotnego uszkodzenia*.

Kolejność badania diagnostycznego określana jest przez wyniki poszczególnych sprawdzeń (procedur testowych). Kolejność ta nie wynika ze związków przyczynowo-skutkowych pomiędzy uszkodzeniami i niesprawnościami pierwotnymi a niezdatnością lub niesprawnością danego obiektu, lecz może być np. ustalona na podstawie oceny kosztów lub czasu wykonywania zadań testowych.

Drzewo sprawdzeń winno umożliwić [Pau, 1981]:

- *odnalezienie sekwencji sprawdzeń*, minimalizującej oczekiwany koszt (lub oczekiwany czas) wymagany do zlokalizowania uszkodzenia,
- wybór procedur testowych do wykonania w każdym stadium badania diagnostycznego w sekwencji odpowiadającej wkładowi wyniku tego sprawdzenia w końcową diagnozę.

Przegląd metod poszukiwania optymalnych i suboptymalnych ścieżek sprawdzeń zawiera praca [Pau, 1981].

Liczne przykłady zastosowań drzew sprawdzeń, np. do wspomagania diagnozowania silników spalinowych gaźnikowych, zawarte są w [Cempel, Tomaszewski, 1992].

2.3.3 Tablice decyzyjne reguł

Tablice decyzyjne reguł[11] umożliwiają reprezentację wiedzy diagnostycznej w sposób proceduralny. Mogą być stosowane bądź do bezpośredniego wspomagania działań diagnostycznych i przyjmują wówczas specjalną postać *karty diagnostycznej* (zob. [Cempel, 1982], [Mitchell, 1981] oraz publikacje wielu firm i instytucji diagnostycznych, jak np.

[11] Nazwa wprowadzona przez autora w celu rozróżnienia dwóch znaczeń: omawianej w niniejszym rozdziale *tablicy reguł* (nazywanej przez niektórych autorów "tablicą decyzyjną" [Cholewa, 1993]) oraz *tablicy przykładów uczących* (opis w rozdziale 3.2.1), dla której będzie stosowana nazwa "tablica decyzyjna" (za [Pawlak, 1991], [Skowron, 1993], [Nowicki, Słowiński, Stefanowski, 1992] i in.).

[Solartron, 1995], [Solartron, 1995a]), bądź do wspomagania konstruowania układów wnioskujących w systemach doradczych [Cholewa, 1993], [Cholewa, Kaźmierczak, 1995].

Tablice decyzyjne reguł są specyficznym środkiem reprezentacji zbioru reguł, dotyczących zwykle jakiegoś obiektu. Reguły zapisywane są w kolumnach tablicy, której wiersze określają kolejno:

1. *warunki, pytania i testy* (ogólnie: stwierdzenia), którym przypisywane są oceny stopnia prawdziwości (co umożliwia badanie przesłanek reguł),
2. *działania* (wśród których trzeba wyróżnić specjalny ich rodzaj, polegający na uznaniu konkluzji),
3. *wyjścia* (umożliwiające organizowanie przebiegu wnioskowania, jak np. iteracja w razie udzielenia przez użytkownika sprzecznych lub błędnych odpowiedzi).

Istotne jest, że:

- kolumny interpretowane są w kolejności "od lewej do prawej" i jeśli wartość stwierdzenia (tj. stopień jego prawdziwości) spełnia warunek wyznaczony przez elementarną przesłankę reguły, następuje przejście do badania kolejnego wiersza (dla tej samej reguły),
- wiersze wykonywane są w kolejności ich zapisania (od góry do dołu), co ma znaczenie w razie występowania efektów ubocznych spowodowanych badaniem wartości warunków oraz wykonywaniem kolejnych działań.

Znaczącym rozszerzeniem zastosowań tablic decyzyjnych reguł stało się wprowadzenie *uogólnionych tablic decyzyjnych reguł* [Cholewa, 1993], które umożliwiają:

- określanie w sposób przybliżony warunków na stopnie prawdziwości stwierdzeń tworzących przesłanki,
- określenie wartości, zwracanej przez tablicę.

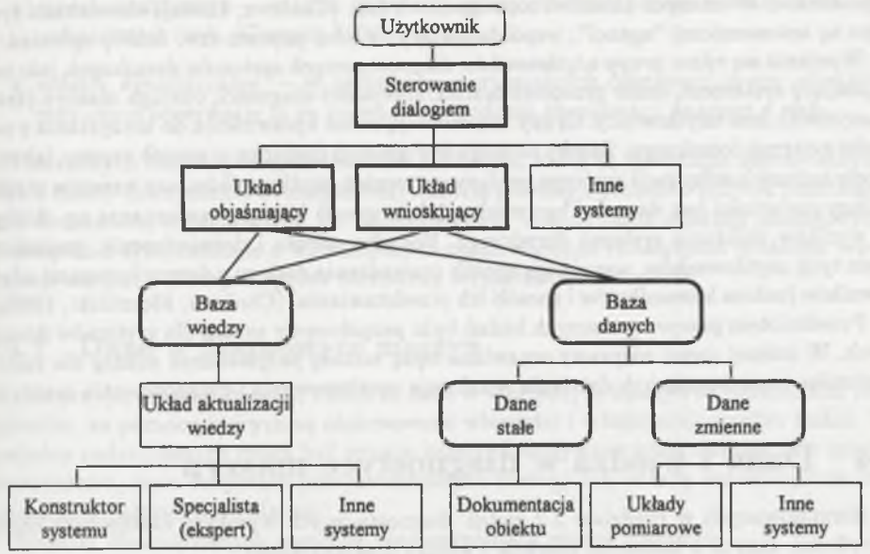
Rozszerzenie to umożliwia dokonanie dekompozycji dużych tablic decyzyjnych reguł na strukturę (np. sieć) tablic o mniejszych wymiarach, które łatwiej można modyfikować, rozbudowywać i usuwać z nich błędy (pętle reguł, pominięte reguły, powtórzone reguły) [Cholewa, 1993], [Cholewa, Moczulski, Filipowicz, 1994].

2.3.4 Systemy doradcze

System doradczy [Cholewa, Pedrycz, 1987] jest programem komputerowym (lub pakietem takich programów), który "na podstawie szczegółowej, specyficznej wiedzy z danej dziedziny, przechowywanej w systemie komputerowym, może wyciągać wnioski, działając w sposób zbliżony do sposobu rozumowania człowieka" [Cholewa, 1996].

W "klasycznym" systemie doradczym (zwanym także *systemem doradczym statycznym*, por. [Cholewa, 1996]) wyróżnia się następujące podstawowe jego elementy [Cholewa, Pedrycz, 1987], [Cholewa, 1996] (rys. 2.3):

- elementy programu, wykonywane podczas współdziałania użytkownika z systemem doradczym:



Rys. 2.3. Główne elementy statycznego systemu doradczego (wg [Cholewa, 1996])
Fig. 2.3. Main elements of a static expert system (by [Cholewa, 1996])

  - procedury sterowania dialogiem pomiędzy użytkownikiem a systemem doradczym,
  - procedury tzw. rozumowania[12],
  - procedury objaśniania,
  - procedury aktualizacji zawartości bazy wiedzy,
- elementy nie wykonywane, stanowiące odpowiednie bazy:
  - baza wiedzy,
  - baza danych stałych,
  - baza danych zmiennych.

W grupie systemów doradczych należy wyróżnić *dynamiczne systemy doradcze* (por. [Cholewa, 1996], [Cholewa, 1997], [Cholewa, 1997a]), wspomagające wnioskowanie diagnostyczne w układach monitorujących i diagnostycznych, stosowanych do nadzoru tzw. *maszyn krytycznych* (zob. np. [Cholewa, 1996a] w odniesieniu do turbozespołów energetycznych). Dynamiczny system doradczy realizuje zadania w ograniczonym, gwarantowanym czasie oraz przy ograniczonych zasobach. Dynamiczne systemy doradcze umożliwiają wspomaganie postępowania diagnostycznego w odniesieniu do obiektu z uwzględnieniem jego zmiennych warunków działania oraz ograniczonej dynamiki zmian stanu technicznego

[12]Zwykle zwane są one *procedurami wnioskowania*, lecz działanie tych procedur nie może być w typowych zastosowaniach ograniczone jedynie do wnioskowania.

tego obiektu. W znanych autorowi rozwiązaniach (np. [Cholewa, 1996a]) elementami systemu są autonomiczni "agenci", współdziałający ze sobą poprzez tzw. *tablicę ogłoszeń*.

Wyróżnia się *różne grupy użytkowników diagnostycznych systemów doradczych*, jak: zarządzający systemem, dozór przedsiębiorstwa, specjaliści-diagności, obsługa maszyn (tzw. maszyniści), inni użytkownicy. Grupy te posiadają różne uprawnienia do korzystania z zasobów systemu doradczego. Zasoby te mogą być dla nich dostępne *w sposób czynny*, jak np. w celu zmiany konfiguracji systemu, zmiany uprawnień użytkowników, czy wreszcie w celu zmiany zawartości baz danych i baz wiedzy, lub *w sposób bierny* — co oznacza np. dostęp do wyników działania systemu doradczego. Ponadto wiedza i doświadczenie, posiadane przez tych użytkowników, warunkują sposób prowadzenia dialogu z danymi grupami użytkowników (zakres komunikatów i sposób ich przedstawiania) [Cholewa, Moczulski, 1993a].

Przedmiotem przeprowadzonych badań było *pozyskiwanie wiedzy* dla systemów doradczych. W dalszej części rozprawy omawiane będą *metody pozyskiwania wiedzy dla takich systemów*, wspomagających działania w zakresie monitorowania i diagnozowania maszyn.

2.4 Dane i wiedza w diagnostyce maszyn

Ze sformułowanych w rozdziale 2.2 zadań diagnostycznych wynikają następujące cechy specyficzne danych o badanym obiekcie i wiedzy diagnostycznej:

- *duża liczba obserwowanych sygnałów*, będących nośnikami informacji o badanym obiekcie (jego wejściach i wyjściach),
- potrzeba uwzględniania zmieniających się warunków działania (zwłaszcza w razie monitorowania maszyn "krytycznych"),
- *zmienność stanów obiektu*,
- *potrzeba uwzględniania historii wytwarzania i eksploatacji* diagnozowanego obiektu,
- zgromadzone dane, będące wynikiem obserwacji diagnostycznej, mogą być *niekompletne*, *sprzeczne*, *obciążone niepewnością*, *niedokładne* lub *zapisane za pomocą wartości jakościowych* [Dubois, Prade, 1996],
- wiedza diagnostyczna często jest *niepewna*, *przybliżona* lub *niejednoznaczna*.

Jeśli ograniczyć się do danych zgromadzonych w ramach biernego eksperymentu diagnostycznego, należy rozpatrzyć następujące przypadki:

1. *Dane pochodzą z obserwacji jednego obiektu konkretnego*, przy czym obserwacje te są prowadzone w trakcie jego eksploatacji.
2. *Dane pochodzą z obserwacji $N > 1$ obiektów konkretnych*, obserwowanych w różnych chwilach ich czasu eksploatacji.

W przypadku 1. celem zadania diagnostycznego może być wyznaczenie relacji diagnostycznych dla obiektu traktowanego jako obiekt dynamiczny, ponieważ istnieje możliwość gromadzenia sekwencji zdarzeń, opisujących wejścia i wyjścia obiektu.

W przypadku 2. w zależności od planu eksperymentu biernego możliwe jest podjęcie próby identyfikacji:

- *modelu statycznego* — w razie obserwacji grupy obiektów o zróżnicowanych stanach technicznych (zob. [Cempel, 1989]),
- *modelu dynamicznego* — w przypadku powtarzalnych obserwacji grupy obiektów, przy czym obserwacje te są prowadzone podczas eksploatacji każdego z nich.

Dla różnych rozpatrywanych obiektów złożonych, które są obiektami naprawialnymi, zwykle mamy do czynienia z różną intensywnością przebiegu procesów zużycia poszczególnych elementów, występowania ich uszkodzeń oraz dokonywania operacji obsługowych. W związku z tym, ostatnie z wymienionych zadań *wymaga rozwiązania problemu reprezentacji danych i wiedzy w sposób niezależny bezpośrednio od czasu θ*[13].

2.4.1 Dane w diagnostyce maszyn

Zwykle przyjmuje się [Cholewa, 1983], że dane w diagnostyce maszyn są wartościami cech sygnałów, za pomocą których są obserwowane własności i właściwości obiektu badań. Za specjalny rodzaj danych mogą być uznane zapisy, dokonywane przez obsługę (np. opinie, spostrzeżenia, dane dotyczące przeprowadzonych obsług) lub układy automatyczne (np. komunikaty, ostrzeżenia i alarmy).

Spośród *obserwowanych sygnałów diagnostycznych* można wyróżnić:

- ze względu na dynamikę zmian wartości sygnałów:
  - sygnały statyczne i wolnozmienne,
  - sygnały szybkozmienne (dynamiczne);
- ze względu na pasmo sygnałów (związane z ilością przenoszonej informacji diagnostycznej):
  - sygnały wąskopasmowe, które mogą być opisane za pomocą pojedynczych cech liczbowych,
  - sygnały szerokopasmowe, które winny być opisane za pomocą cech funkcyjnych[14];
- ze względu na zbiór przyjmowanych wartości chwilowych:
  - sygnały dyskretne, przyjmujące co najwyżej N wartości (wraz z ważną grupą sygnałów dwustanowych, zwanych także binarnymi),
  - sygnały ciągłe, przyjmujące wartości rzeczywiste;
- ze względu na "wymiarowość":
  - sygnały jednowymiarowe,
  - sygnały wielowymiarowe [Wysogląd, 1996], które mogą być rozpatrywane jako macierze sygnałów jednowymiarowych, zwykle skorelowanych ze sobą.

[13] Przewidywane jest podjęcie badań nad tym zagadnieniem w przyszłości, np. z zastosowaniem metod syntaktycznych rozpoznawania obrazów [Fu, 1982] - zob. rozdział 6.

[14] W wypadku analizy takich sygnałów metodami cyfrowymi uzyskiwane są macierze kolumnowe wartości cech liczbowych (np. dyskretne widmo częstotliwościowe [Cholewa, Moczulski, 1993]).

Liczba obserwowanych sygnałów diagnostycznych uwarunkowana jest celem postępowania diagnostycznego i pozostaje w związku ze złożonością obiektu poddanego obserwacji diagnostycznej. W wypadku monitorowania bardzo złożonych obiektów (jak np. turbozespoły energetyczne) obserwowane są setki sygnałów statycznych i dwustanowych oraz dziesiątki sygnałów dynamicznych [Cholewa, 1996a].

Dane, gromadzone podczas obserwacji diagnostycznej obiektów, zwykle są niedoskonałe, ponieważ są [Grzymała-Busse, 1994]:

- niekompletne,
- niepewne,
- niedokładne,
- błędne.

Niekompletność danych może być spowodowana brakiem uwzględnienia ważnych cech, opisujących obiekt, lub brakiem wartości tych cech. W diagnostyce maszyn przyczyną niekompletności danych może być np. niesprawność toru pomiarowego, służącego do obserwacji sygnału. Takimi niesprawnościami mogą być np. uszkodzenie czujnika lub przerwanie kabla, łączącego ten czujnik ze wzmacniaczem pomiarowym.

Niepewność danych spowodowana jest wpływem szumów, które pojawiają się na wejściach i wyjściach obserwowanego obiektu, na wyniki obserwacji.

Niedokładność danych spowodowana jest przez własności układów pomiarowych (jak np. ich rozdzielczość) oraz technik określania wartości cech (np. błąd paralaksy przy odczycie wartości ze skali miernika). Inny rodzaj niedokładności związany jest z gromadzeniem *danych jakościowych*, związanych zwykle z opiniami lub stwierdzeniami, formułowanymi przez człowieka, lub uzyskiwanych w wyniku kwantyzacji wartości cechy ciągłej (przy czym zazwyczaj dąży się do stosowania niewielu poziomów kwantyzacji).

Dane błędne mogą zawierać błędy dotyczące jednej wartości danej cechy, lub mogą być *sprzeczne*, co może mieć miejsce np. w wypadku ich gromadzenia w układach wieloczujnikowych, które są stosowane do monitorowania względnego przesuwu osiowego walu w turbinach, w razie uszkodzenia jednego z czujników. Inną przyczyną występowania danych sprzecznych (w postaci jakościowej) mogą być rozbieżne opinie specjalistów.

Oprócz danych, zgromadzonych w wyniku obserwacji obiektu diagnozowania, ważną rolę w pozyskiwaniu wiedzy dotyczącej relacji diagnostycznych spełniają dane, pochodzące z symulacji komputerowej [Cholewa, White, 1993], [Cholewa, Kiciński, 1996], [Kiciński *et al.*, 1995]. Dane takie, będące wartościami cech takich sygnałów wibroakustycznych, jak np. wartości chwilowe przemieszczeń względnych i bezwzględnych wybranych punktów "pomiarowych", uzyskiwane są w wyniku obliczeń, przeprowadzanych za pomocą tzw. *eksperymentalnego oprogramowania komputerowego* z zastosowaniem odpowiednich *modeli symulacyjnych* [Kiciński, 1996a]. Dane, uzyskane w wyniku symulacji, mogą mieć podobne cechy jak dane zgromadzone w wyniku obserwacji (np. niekompletność lub niedokładność spowodowaną błędami zaokrągleń w wypadku obliczeń wykonywanych iteracyjnie).

Ponadto często wymagane jest gromadzenie danych umożliwiających odtworzenie historii wejść oraz historii obsług, które są reprezentowane np. w postaci sekwencji cech wejść oraz sekwencji stanów (tj. szeregów czasowych [Box, Jenkins, 1983]) w funkcji czasu θ (zob. rozdział 2.1.5).

Podsumowując, można stwierdzić, że w diagnostyce maszyn, zwłaszcza tzw. maszyn "krytycznych", które nadzorowane są za pomocą stacjonarnych układów monitorowania, mamy do czynienia z bazami danych, zawierającymi ogromne ilości danych (od kilkuset do kilku tysięcy wartości cech liczbowych, uzyskiwanych regularnie z zadanym krokiem czasowym). Stawia to szczególnie wysokie wymagania metodom wnioskowania diagnostycznego, które umożliwiają skuteczne osiągnięcie konkluzji przy ograniczonych zasobach takich systemów i w ograniczonym (gwarantowanym) czasie.

2.4.2 Wiedza diagnostyczna

Wstępny okres badań nad skutecznością zastosowania metod tzw. "sztucznej inteligencji" do rozwiązywania problemów o charakterze ogólnym (np. projekt GPS — *General Problem Solver*) doprowadził badaczy do wniosku, że warunkiem skutecznego wspomagania rozwiązywania problemów przez systemy doradcze jest odpowiednie określenie *dziedziny zastosowania*, która nie powinna być zbyt rozległa [Cholewa, Pedrycz, 1987]. Rozwiązanie danego zadania diagnostycznego poszukiwane jest wówczas w przestrzeni możliwych rozwiązań o mniejszej mocy, co istotnie wpływa na efektywność tego poszukiwania. W dalszym ciągu pracy będziemy rozpatrywali problemy pozyskiwania wiedzy diagnostycznej właśnie dla dobrze określonej i niezbyt rozległej *dziedziny zastosowania*. Przykładem takiej dziedziny może być np. diagnostyka maszyn wirnikowych lub, dziedzina jeszcze węższa, diagnostyka turbozespołów o określonej mocy [Cholewa, 1996a].

W diagnostyce maszyn szczególne znaczenie ma:

1. *wiedza podstawowa*, która dotyczy szczegółowego opisu (określenia) dziedziny zastosowania,
2. *wiedza "operacyjna"*, dotycząca relacji diagnostycznych, która stosowana jest w procesie wnioskowania lub rozumowania diagnostycznego w tej dziedzinie zastosowania.

Wiedza diagnostyczna ma zwykle charakter dynamiczny, co wynika ze sposobu opisu systemu danego obiektu, który to opis uwzględnia rodzinę funkcji zmiany stanu tego obiektu wg (2.24). Jak wspomniano na s. 21, każda z tych funkcji zależy nie tylko od długości tego przedziału

$$\Delta\theta = \theta' - \theta \ \wedge \ \theta, \theta' \in \Theta \ \wedge \ \theta \leqslant \theta',$$

ale także od samej chwili początkowej θ. W konsekwencji, związki (2.43), (2.45) pomiędzy wejściami, wyjściami i stanami także podlegają zmianom w funkcji chwili θ.

Zwykle jednak, zakładając małą dynamikę zmian stanu zachodzących w obiekcie, dopuszcza się możliwość opisu związków (2.43), (2.45) za pomocą relacji nie podlegających zmianom w funkcji czasu θ (przyjmuje się *model statyczny* w pewnym przedziale czasu $[\theta_1, \theta_2]$). W dalszym ciągu pracy zakładana będzie możliwość stosowania modeli statycznych.

Diagnostyka techniczna jest dziedziną, w której niejednokrotnie brak podstaw teoretycznych do tworzenia modeli matematycznych lub symulacyjnych (por. np. [Kiciński, 1996a]), które byłyby dostatecznie szczegółowe i mogły umożliwić praktyczne określenie relacji przyczynowo-skutkowych oraz relacji "diagnostycznych"[15]. Z tych względ-

[15] Mamy na względzie możliwość uzyskania wiarygodnych wyników obliczeń przy dostępnych zasobach i ograniczonym czasie wykonywania tych obliczeń.

dów niejednokrotnie potrzebne jest posługiwanie się wiedzą przybliżoną, niepewną lub niejednoznaczną.

Przybliżony charakter wiedzy odpowiada sposobowi rozumowania specjalisty, który czyni to także w sposób przybliżony i dlatego nie zawsze można opisać sposób jego rozumowania modelami klasycznej logiki formalnej [Cholewa, Pedrycz, 1987].

Niepewność wiedzy może być spowodowana np. rozbieżnymi opiniami specjalistów lub pozyskiwaniem tej wiedzy na podstawie zbioru przykładów, zawierających brakujące wartości, przykłady sprzeczne (zaklasyfikowane do różnych klas) lub niedokładne.

Niejednoznaczność wiedzy wynika np. z braku możliwości zidentyfikowania relacji "odwrotnych" o charakterze funkcyjnym — por. zależność (2.40) wraz z komentarzem. Innym jej przejawem jest poprawność uzyskiwanych wyników wnioskowania jedynie w większości przypadków (lecz nie zawsze [Cholewa, Pedrycz, 1987]).

Podsumowując powyższe omówienie, można stwierdzić, że dane i wiedza diagnostyczna mają swoją specyfikę, która w istotny sposób zależy od dziedziny zastosowania oraz od rozwiązywanego zadania diagnostycznego.

2.5 Środki reprezentacji danych i wiedzy w diagnostyce maszyn

Sposób reprezentacji danych i wiedzy winien umożliwić identyfikację: obiektów danej dziedziny (abstraktów i konkretów), relacji zachodzących pomiędzy tymi obiektami, cech tych obiektów, wartości tych cech oraz stopni prawdziwości relacji diagnostycznych. Jeśli mają być reprezentowane wartości cech ciągłych, zwykle zachodzi potrzeba dyskretyzacji tych wartości i wówczas wymagana jest również reprezentacja sposobów przekształcania wartości cech ciągłych w wartości dyskretne (zob. rozdział 2.5.3).

Aby możliwa była reprezentacja danych i wiedzy, wpierw należy opisać (lub zdefiniować) tzw. *dziedzinę zastosowania*. Opis ten jest zapisem *podstawowej wiedzy o dziedzinie* (zob. rozdział 2.4.2).

W diagnostycznych systemach doradczych stosowane są takie środki reprezentacji danych i wiedzy, jak:

- *stwierdzenia* i *stwierdzenia dynamiczne* (zob. rozdział 2.1.4),
- *sieci stwierdzeń* (zwłaszcza: stwierdzeń dynamicznych [Cholewa, 1996a]),
- *reguły* ("ostre"),
- *reguły rozmyte* [Korbicz, 1996], [Pieczyński, Korbicz, 1997],
- *tablice decyzyjne reguł* (zob. rozdział 2.3),
- *ramy* [Cholewa, 1993],
- *klasy* [Cholewa, Moczulski, 1990],
- *drzewa decyzyjne* [Quinlan, 1986],
- *Bayesowskie sieci przekonań* oraz *jakościowe sieci probabilistyczne* [Drużdżel, 1994], [Russell, Norvig, 1995],

- *sieci neuronalne* (tj. sztuczne sieci neuronowe) [Tadeusiewicz, 1993], [Korbicz, 1994], [Pieczyński, Korbicz, 1997],[Kosmol, 1996],
- hypertekst [Moczulski, 1994a].

Większość wymienionych środków reprezentacji danych i wiedzy została szczegółowo opisana w dostępnej literaturze (np. [Cholewa, Pedrycz, 1987], [Cholewa, Moczulski, 1990], [Cholewa, 1993], [Cholewa, 1996]). W dalszej części rozdziału zostaną omówione te środki, które stosowane były w opisywanych badaniach.

2.5.1 Opis dziedziny zastosowania

Opis ten obejmuje między innymi:

- wykaz klas obiektów (będących jednocześnie obiektami abstrakcyjnymi); wszystkie rozpatrywane klasy tworzą *rodzinę klas*:

$$\overline{C} = \{C_k \mid k \in K\}; \tag{2.51}$$

- *związki, zachodzące pomiędzy klasami* (w rodzinie klas $\overline{C}$), które umożliwiają określenie klas bazowych i klas pochodnych, będących podzbiorami klas bazowych[16];
- wykaz obiektów (przede wszystkim konkretnych), tworzących *zbiór obiektów* O:

$$O = \{o_i \mid i \in I\}; \tag{2.52}$$

- wykaz cech obiektów[17], definiowanych dla klas obiektów i tworzących łącznie *zbiór wszystkich cech* A:

$$A = \bigcup_{k \in K} A(C_k), \tag{2.53}$$

gdzie $A(C_k)$ oznacza zbiór wszystkich cech, określonych dla klasy C_k, zaś:

$$J = \sum_{k \in K} card(A(C_k)) \tag{2.54}$$

jest liczbą wszystkich cech;

- wykaz wartości (jakościowych) cech obiektów, definiowanych dla poszczególnych cech obiektów i tworzących *zbiór wszystkich wartości cech* V:

$$V = \bigcup_{k \in K} \bigcup_{a \in A(C_k)} V(a), \tag{2.55}$$

gdzie:

$$V(a) = \{v_m \mid m \in M(a)\} \tag{2.56}$$

jest zbiorem wartości (dziedziną) danej cechy a, $M(a)$ jest zbiorem wskaźników (zwykle o skończonej mocy), zaś:

$$M = \sum_{k \in K} \sum_{a \in A(C_k)} card(M(a)) \tag{2.57}$$

jest liczbą wszystkich wartości cech.

[16] Pomiędzy klasami zachodzi ważna relacja *dziedziczenia* cech i ich wartości, a także sposobów przekształcania wartości cech ciągłych na cechy pochodne o wartościach jakościowych — zob. opis na s. 14.

[17] Cechy omówiono już w rozdziale 2.1.2.

Podstawowym założeniem, czynionym przed podjęciem decyzji dotyczącej wyboru sposobu reprezentacji danych i wiedzy diagnostycznej, winno być ograniczenie mocy *przestrzeni reprezentacji danej dziedziny*. Można to osiągnąć poprzez określenie:

- skończonej liczby klas obiektów (tj. $K < \infty$),
- skończonej liczby obiektów (tj. $I < \infty$),
- skończonej liczby własności i właściwości tych obiektów oraz wejść i wyjść, które będą obserwowane za pomocą sygnałów,
- skończonej liczby cech wejść, wyjść i stanów tych obiektów (tj. $J < \infty$),
- zbioru wartości stopni prawdziwości stwierdzeń o skończonej liczności,
- skończonej liczby chwil czasu punktowego,

oraz, dodatkowo w odniesieniu do reprezentacji danych będących zapisami subiektywnych opinii i wyników obserwacji personelu, a zwłaszcza w odniesieniu do reprezentacji wiedzy (aby określić *skończoną*, a najlepiej *jak najmniej liczną przestrzeń poszukiwań*):

- skończonych (co więcej, o małej mocy, najlepiej zawierających jedynie kilka elementów) zbiorów wartości cech (ich dziedzin), przy czym zaleca się stosowanie wartości lingwistycznych (zob. [Czogała, Pedrycz, 1980], [Kacprzyk, 1986]), co implikuje także warunek na łączną liczbę wartości cech $M < \infty$.

Należy podkreślić wagę prawidłowego wyboru cech oraz zbiorów ich wartości, będą one bowiem służyły do reprezentacji danych i wiedzy diagnostycznej. Podstawowym kryterium wyboru zbioru cech winna być możliwość uchwycenia najbardziej istotnych własności obiektów danej dziedziny przy zadanym celu tego opisu. Istnieją tutaj następujące możliwości [Michalski, 1983]:

1. Wszystkie cechy *są relewantne* w odniesieniu do sformułowanego celu. W wypadku pozyskiwania wiedzy diagnostycznej istnieje więc możliwość utworzenia modelu matematycznego lub logicznego, wiążącego symptomy i wartości cech stanu obiektu. Zwykle model ten jest tworzony w sposób *indukcyjny*.
2. *Częściowa relewantność* (zbiór cech zawiera podzbiór takich cech, które są nierelewantne w odniesieniu do celu lub są wręcz nadmiarowe). Dla potrzeb pozyskiwania wiedzy diagnostycznej należy zidentyfikować cechy najbardziej relewantne i dla tych cech utworzyć odpowiedni model (np. jakościowy), stosując *selektywną indukcję*.
3. *Pośrednia relewantność* (zbiór cech nie zawiera żadnych cech relewantnych ze względu na sformułowany cel). Pozyskiwanie wiedzy polega na identyfikacji modelu diagnostycznego z zastosowaniem metod *konstruktywnej indukcji*, której istotą jest tworzenie nowych cech, będących wyrażeniami utworzonymi dla cech należących do początkowego zbioru cech [Michalski, 1997].

Z doświadczeń autora wynika, że określenie odpowiedniej przestrzeni reprezentacji dziedziny zastosowania jest kluczowym i bardzo pracochłonnym zadaniem w procesie pozyskiwania wiedzy (zob. np. [Moczulski, Maniak, 1997]). Szczególnie ważki jest właściwy wybór zbioru cech, które będą służyły do opisu własności obiektów, określenie sposobu przekształcania cech ciągłych w pochodne cechy dyskretne oraz wybór zbioru wartości cech (wraz z ewentualnymi nazwami tych wartości).

2.5.2 Reprezentacja danych

Dane gromadzone w wyniku obserwacji diagnostycznej są przechowywane w *bazach danych*. W diagnostyce maszyn stosuje się bazy danych o następujących strukturach:

- *struktura relacyjna* [Ullman, 1988],
- *struktura sieciowa* (ze szczególnym przypadkiem *struktury hierarchicznej* [Ullman, 1988]),
- *struktura obiektowa* [Figura, 1996].

W opisywanych badaniach przyjęto, że stosowane będą bazy danych o strukturze relacyjnej, określanej w taki sposób, by ograniczyć *redundancję danych*, ze złożonym układem wielu tablic, powiązanych pomiędzy sobą odpowiednimi relacjami.

Szczególne znaczenie ma stosowanie takiego środka reprezentacji danych, który umożliwia zapis wartości różnych cech obiektu i nie wymaga modyfikacji ze względu na zwiększenie stopnia szczegółowości opisu (np. poprzez wprowadzenie nowych cech). Do środków takich należą:

- stwierdzenia,
- stwierdzenia dynamiczne.

Ponadto często pojawia się potrzeba gromadzenia takich danych, jak:

- opisy słowne.

Stwierdzenia umożliwiają reprezentację wartości cech wejść, wyjść i stanów w odniesieniu do obiektu (abstraktu lub konkretu). Stwierdzenia dynamiczne umożliwiają reprezentację elementów *szeregów czasowych*. Do zapisu stwierdzeń i stwierdzeń dynamicznych można stosować zarówno *cechy ilościowe* (o wartościach ciągłych), jak i *cechy pochodne jakościowe* (o wartościach dyskretnych — por. rozdział 2.1.2).

Ciekawym przykładem definicji relacyjnej bazy danych, przeznaczonej do gromadzenia danych dla potrzeb eksploatacji maszyn, jest schemat *MIMOSA*[18] [MIMOSA, 1996], [MIMOSA, 1996a]. Schemat ten umożliwia współdziałanie wielu ośrodków diagnostycznych oraz tworzenie oprogramowania diagnostycznego w systemie otwartym. Według schematu logicznego *MIMOSA* mogą być budowane relacyjne bazy danych, w których są gromadzone dane, dotyczące m.in.:

- *struktury obiektów* (abstrakcyjnych i konkretnych);
- *gospodarki remontowej* w odniesieniu do obiektów-konkretów, przy czym wprowadza się rozróżnienie na (abstrakcyjne) *zespoły funkcjonalno-obsługowe* oraz *egzemplarze-konkrety* z ich własną historią eksploatacji (użytkowania i obsługiwania);
- *konfiguracji torów pomiarowych* do obserwacji sygnałów diagnostycznych (z podziałem tych torów na zespoły funkcjonalno-obsługowe oraz na egzemplarze-konkrety);

[18] *Machinery Information Management Open System Alliance* — projekt realizowany przez czołowe firmy z dziedziny eksploatacji maszyn, diagnostyki, wytwarzania aparatury pomiarowej, wytwarzania sprzętu komputerowego itp.

- *wyników pomiarów* sygnałów statycznych i dynamicznych;
- innych obserwacji i spostrzeżeń.

Przewiduje się [Cholewa, 1996], że zalecenia dotyczące schematu relacyjnych baz danych *MIMOSA* zostaną powszechnie zaakceptowane, co wskazuje dalszy kierunek rozwoju środków i sposobów reprezentacji danych i wiedzy dla potrzeb diagnostyki maszyn.

2.5.3 Reprezentacja wiedzy

Przyjmujemy wszystkie założenia, wymienione w rozdz. 2.5.1, s. 42 (włącznie z założeniem dotyczącym skończonych mocy dziedzin poszczególnych cech). Ponadto zakładamy, że:

- w wyniku rozumowania diagnostycznego osiągane są jedynie konkluzje, które są stwierdzeniami należącymi do wcześniej określonej przestrzeni reprezentacji,
- obliczanie stopni prawdziwości wszystkich stwierdzeń złożonych (koniunkcji, alternatywy, implikacji, agregacji itp.) jest działaniem wewnętrznym w zbiorze wartości tego stopnia.

Gdy takie założenia są spełnione, uznaje się, że poszukiwanie rozwiązania następuje *w domkniętym świecie* (ang. *closed world*).

Do reprezentacji wiedzy stosowane są zwykle *dyskretne wartości cech*. Postępowanie takie jest uzasadnione sposobem rozumowania człowieka, który zazwyczaj posługuje się pojęciami jakościowymi[19] i stosuje przybliżone metody wnioskowania. Ponadto możliwe jest zapisywanie wiedzy o charakterze ogólnym, np. dotyczącej szerokiej klasy obiektów. "Dopasowanie" ogólnej wiedzy diagnostycznej do indywidualnych właściwości poszczególnych obiektów (które w wypadku rozpatrywanych układów mechanicznych mogą być np. spowodowane znacznym rozrzutem własności dynamicznych maszyn tej samej konstrukcji, lecz działających w różnych zakładach przemysłowych i, w konsekwencji, posiadających m. in. różne własności dynamiczne układu fundamentu — zob. [Downham, Woods]) następuje wówczas przez określenie sposobów dyskretyzacji cech ciągłych *indywidualnie dla każdego obiektu*.

Reprezentacja wiedzy z zastosowaniem dyskretnych wartości cech umożliwia także ograniczenie przeszukiwania rozwiązania do przestrzeni o skończonej mocy. Dla potrzeb reprezentacji wiedzy może być wówczas użyty zapis stosowany w systemach VL_1 (ang. *variable-valued logic system 1*, tj. system logiki o zmiennej liczbie wartości) i *APC* (ang. *annotated predicate calculus*, tj. "objaśniony" rachunek predykatów) [Michalski, 1983].

Najbardziej ogólny podział sposobów reprezentacji wiedzy diagnostycznej obejmuje takie sposoby, jak:

- *proceduralny*, polegający np. na "...określeniu zbioru procedur, których działanie reprezentuje wiedzę o dziedzinie" [Cholewa, 1996],
- *deklaratywny*.

[19] Celowe jest stosowanie wartości lingwistycznych, co ułatwia użytkownikowi zrozumienie zapisów oraz upraszcza generowanie przeznaczonych dla niego komunikatów (jak: objaśnienia i konkluzje końcowe).

Relacje diagnostyczne mogą być reprezentowane w sposób proceduralny, np. w postaci modelu matematycznego, który może następnie stanowić podstawę do opracowania procedury, zapisanej w języku programowania.

W dalszym ciągu w pracy omawiana będzie przede wszystkim reprezentacja wiedzy w sposób deklaratywny. Spośród różnych środków reprezentacji, w pracy stosowane są następujące środki:

- stwierdzenia,
- stwierdzenia dynamiczne [Cholewa, 1996a],
- reguły,
- drzewa decyzyjne[20].

Stwierdzenia i stwierdzenia dynamiczne

Stwierdzenia oraz *stwierdzenia dynamiczne* omówiono w rozdziale 2.1.4. W bazach wiedzy stwierdzenia są reprezentowane w ogólnej postaci (por. [Cholewa, 1997a]):

$$\langle o, a, v, \theta, b \rangle, \tag{2.58}$$

gdzie:

o, a, v — *właściwa treść stwierdzenia* (tzn. wypowiedź: "obiektowi o przysługuje cecha a o wartości v", lub: "cecha a obiektu o ma wartość v"),

θ — *przedział czasu eksploatacji* (lub punkt tego czasu — por. przypis na s. 16),

b — *ocena stopnia prawdziwości* lub *stopnia przekonania o prawdziwości treści stwierdzenia*, która jest mu przypisana w przedziale czasu θ.

Należy zwrócić uwagę na warunek konieczny do tego, by stwierdzenie było określone:

$$p = \langle o, a, v \rangle \Rightarrow (\exists C_0 \in \{C_k \mid k \in K\})[o \in C_0 \wedge a \in A(C_0) \wedge v \in V(a)], \tag{2.59}$$

gdzie:

C_0, C_k — klasy obiektów,

$A(C_0)$ — zbiór cech, określonych dla klasy C_0,

$V(a)$ — zbiór wartości (jakościowych), określonych dla cechy a.

Warunek (2.59) oznacza, że:

1. cecha a musi być określona dla pewnej klasy obiektów, do której należy obiekt o (nie musi być ona określona dla tej klasy, do której bezpośrednio należy obiekt o, lecz może być określona dla jakiejś jej klasy bazowej — por. rozdział 2.1.1),
2. wartość cechy v ma należeć do zbioru wartości $V(a)$, zdefiniowanego dla cechy a.

[20] Drzewa decyzyjne uznawane są jednocześnie za środek reprezentacji wiedzy proceduralnej (zob. [Quinlan, 1986], [Michalski, 1997]).

Dla danego stwierdzenia p można określić funkcje (które określone są na zbiorze wszystkich stwierdzeń, a ich dziedzinami są odpowiednio zbiory obiektów, cech, wartości cech, chwil czasu "makro" i wartości stopnia prawdziwości oraz zbiór treści stwierdzeń) [Cholewa, 1996a]:

$$o(\,),a(\,),v(\,),\theta(\,),b(\,),oav(), \tag{2.60}$$

które zwracają poszczególne elementy tego stwierdzenia oraz (w wypadku ostatniej z funkcji) treść tego stwierdzenia:

$$o(p),a(p),v(p),\theta(p),b(p),oav(p). \tag{2.61}$$

Założenie "domkniętości świata" pozwala przyjąć, że mogą być utworzone następujące słowniki[21]:

- słownik (nazw) klas obiektów *DictCO*,
- słownik (nazw) obiektów *DictO*,
- słownik (nazw) cech obiektów *DictA* (określanych dla klas obiektów — por. rozdział 2.1.2),
- słownik (nazw) wartości cech obiektów *DictV*.

Elementy "treści stwierdzenia" mogą być wówczas zapisywane w bazach danych i wiedzy jako wskazania do słowników odpowiednio *DictO*, *DictA*, *DictV*.

Ponadto, w celu ograniczenia redundancji przy zapisie stwierdzeń, zwłaszcza tych, które generowane są przez automatyczne źródła danych (np. układy pomiarowe [Cholewa, 1996]), stosowany jest słownik treści stwierdzeń *DictOAV*. Zawiera on wskazania do odpowiednich słowników obiektów *DictO*, słowników cech *DictA* i słowników wartości cech *DictV*.

Stwierdzeniu p przypisywana jest *wartość stwierdzenia*, która jest miarą stopnia prawdziwości (lub stopnia przekonania o tej prawdziwości) tego stwierdzenia. Wartość stwierdzenia jest równa wartości funkcji $b(p)$.

W najprostszym przypadku wartość stwierdzenia może być jednym z elementów zbioru dwuelementowego (gdzie: 0 oznacza *fałsz* i 1 oznacza *prawdę*):

$$b(p) \in \{0,1\}. \tag{2.62}$$

Rozwiązanie takie umożliwia stosowanie *klasycznego rachunku zdań*, zwykle jednak w diagnostyce maszyn, gdzie konieczne jest reprezentowanie wiedzy niepewnej i przybliżonej, jest niewystarczające. W pracy postanowiono określać wartości stwierdzeń za pomocą miar konieczności $N(\,)$ i możliwości $P(\,)$ (zob. [Cholewa, Pedrycz, 1987], [Cholewa, 1993]), których wartości $N(p)$ i $P(p)$ mogą być interpretowane jako końce przedziału, do którego należy nieznana wartość stopnia prawdziwości $T(p)$. Zachodzą przy tym relacje:

$$(\forall p)\,[0 \leqslant N(p) \leqslant T(p) \leqslant P(p) \leqslant 1]. \tag{2.63}$$

Przy założeniu, że:

$$(\forall p)\,[N(p) = 0 \vee P(p) = 1], \tag{2.64}$$

[21] Szczegółowe opisy formatu stosowanych słowników są zawarte w [Moczulski *et al.*, 1997a].

można obie miary zastąpić jedną [Cholewa, 1996a]:

$$NP(p) \stackrel{def}{=} (N(p) + P(p))/2, \tag{2.65}$$

zwaną *stopniem prawdziwości stwierdzenia*. Na podstawie jej wartości $NP(p)$ można wyznaczyć obie wartości miar: konieczności $N(p)$ i możliwości $P(p)$ [Cholewa, 1996a]:

$$N(p) = \begin{cases} 0, & \text{gdy } NP(p) \leqslant 0,5; \\ 2 \cdot NP(p) - 1, & \text{gdy } NP(p) > 0,5, \end{cases} \tag{2.66}$$

$$P(p) = \begin{cases} 1, & \text{gdy } NP(p) \geqslant 0,5; \\ 2 \cdot NP(p), & \text{gdy } NP(p) < 0,5. \end{cases} \tag{2.67}$$

Wartości stopni prawdziwości są określane w sposób lingwistyczny, przy czym ich zbiór jest nieliczny. W odpowiednim słowniku wartości *DictB* określono wartości liczbowe miar $N(\,)$ i $P(\,)$, odpowiadające każdej z wartości lingwistycznych (por. tabl. 3.1).

Reguły

Reguły są środkiem reprezentacji wiedzy w sposób deklaratywny. Do każdej reguły przypisana jest jej *dziedzina*, określająca zbiór obiektów, dla których określona jest ta reguła[22].

Reguła r zapisywana jest w następujący sposób:

$$(\forall o \in D(r))[\text{JEŻELI } p(o) \text{ TO } q(o)], \tag{2.68}$$

gdzie:

$D(r)$ — dziedzina reguły,

$p(o)$ — przesłanka reguły,

$q(o)$ — konkluzja reguły,

W pracy rozważane są wyłącznie reguły, które spełniają kryterium *relewantności*[23].

Przesłanki reguł zawierają warunki, utworzone dla wartości cech danego obiektu, określonych dla obiektu o. Konkluzją reguły jest pojedyncze stwierdzenie $\langle o,a,v\rangle$. W wypadku bazy wiedzy diagnostycznej warunki i stwierdzenia dotyczą wartości cech (wejść, wyjść i stanów — por. rozdział 2.1.2). Oczywiste jest wymaganie, że w przesłance i konkluzji muszą występować jedynie takie cechy $A = \{a\}$, które są określone dla wszystkich obiektów $o \in D(r)$, a także takie wartości $V = \{v\}$, które są określone dla cech $a \in A$.

Przesłanki reguł mogą być *elementarne* i *złożone*. W pracy przyjęto, że *przesłankami elementarnymi* są pojedyncze warunki (elementarne). Przesłanki złożone są koniunkcjami wewnętrznych alternatyw warunków elementarnych (zob. poniżej).

Warunek elementarny dla danej reguły r określany jest w następujący sposób:

$$(\forall o \in D(r))[val(a(o)) \propto eterm(a)], \tag{2.69}$$

gdzie:

[22] Formalne wymagania, dotyczące dziedziny reguły, zostaną podane na s. 49.

[23] Kryterium to oznacza (przez analogię do implikacji relewantnej [Marciszewski, 1988]), że warunkiem koniecznym, by reguła była spełniona, jest, by istniało powiązanie znaczeniowe pomiędzy przesłanką i konkluzją (tj. by poprzednik zdania warunkowego nadawał się na przesłankę dla następnika tego zdania).

$val(a(o))$ — *wartość cechy* a dla obiektu o (*elementarny term funkcyjny*),

$\propto$ — *operator relacji* [24] wybrany spośród elementów poniższego zbioru:

$$\propto \in \{=, \neq, <, \leqslant, >, \geqslant\}, \tag{2.70}$$

$eterm(a)$ — *wartość cechy* a (term elementarny — stała).

Warunek elementarny (2.69) oznacza, że wartość funkcji, reprezentowanej przez lewą stronę wyrażenia, jest w relacji "$\propto$" z wartością prawej strony tego wyrażenia.

W celu ograniczenia redundancji zapisu reguł oraz umożliwienia zapisu reguł, pochodzących z różnych źródeł, w pracy do zapisu prawej strony warunku (2.69) stosuje się także *termy złożone*, które są *wewnętrznymi alternatywami termów elementarnych*:

$$Term(a) = eterm_1(a) \vee eterm_2(a) \vee \ldots \vee eterm_n(a). \tag{2.71}$$

Jeśli a jest cechą o dziedzinie uporządkowanej liniowo, $eterm_i(a) = v_i$ oraz $v_1 < v_2 < \ldots < v_n$ są jej kolejnymi wartościami, zamiast zapisu jak w (2.71) stosowany jest zapis:

$$Term(a) = eterm_1(a)..eterm_n(a). \tag{2.72}$$

Warunki o postaci:

$$(\forall o \in D(r))[val(a(o)) \propto Term(a)], \tag{2.73}$$

gdzie $Term(a)$ wg (2.71), zwane będą *warunkami prostymi* lub *selektorami* [Michalski, 1983]. Dla termu złożonego wg (2.73) mogą one być rozwinięte do następującej postaci:

$$(\forall o \in D(r))([val(a(o)) \propto eterm_1(a)] \vee [val(a(o)) \propto eterm_2(a)]$$

$$\vee \ldots \vee [val(a(o)) \propto eterm_n(a)]). \tag{2.74}$$

Koniunkcja warunków prostych wg (2.71):

$$(\forall o \in D(r))([val(a_1(o)) \propto_1 Term_1(a_1)] \wedge [val(a_2(o)) \propto_2 Term_2(a_2)]$$

$$\wedge \ldots \wedge [val(a_n(o)) \propto_n Term_n(a_n)]) \tag{2.75}$$

nazywana jest *warunkiem złożonym* (lub *kompleksem* [Michalski, 1983]). Rozwinięcie kompleksu do koniunkcji alternatyw wewnętrznych warunków elementarnych może spowodować wygenerowanie złożonego wyrażenia, zawierającego wiele warunków elementarnych.

Dziedzina reguły r może zawierać co najwyżej takie obiekty, które należą do klasy, dla której (lub dla jakiejś jej klasy bazowej) określone są cechy występujące w selektorach przesłanki tej reguły. Jeśli zbiór D obiektów (które są wymienione jako jego elementy lub należą do klas obiektów wymienionych jako jego podzbiory — por. opis "Dziedziny reguł" na s. 147) ma być dziedziną $D(r)$ reguły r o przesłance p:

$$p = \bigwedge_{i=1,\ldots,I} [val(a_i(o)) \propto_i Term_i(a_i)], \tag{2.76}$$

[24] Operatory inne niż "$=$" i "$\neq$" mogą być stosowane jedynie dla cech a, których dziedziny $D(a)$ są liniowo uporządkowane (por. rozdział 2.1.2).

to wymagane jest spełnienie warunku:

$$(\forall o \in D)(\forall i = 1, \ldots, I)(\exists C \in \{C_k \mid k \in K\})[o \in C \wedge a_i \in A(C)], \tag{2.77}$$

gdzie warunek $o \in C$ może być spełniony także dla jakiejś klasy bazowej (nadrzędnej) w stosunku do "własnej" klasy obiektu o. Zapis dziedziny reguły jest jedną z dodatkowych informacji (objaśnień) wymaganych przez system *APC* [Michalski, 1983].

Podobnie jak w wypadku stwierdzeń stosowane są zarówno *reguły ostre*, jak i *reguły przybliżone* [Cholewa, Pedrycz, 1987]. Regule przybliżonej można przyporządkować *stopień konieczności* $N(p \Rightarrow q)$ i *stopień dostateczności* $S(p \Rightarrow q)$ [Cholewa, Pedrycz, 1987], które dla znanej lub obliczonej wartości stopnia prawdziwości przesłanki umożliwiają określenie stopnia prawdziwości konkluzji przez układ wnioskujący systemu doradczego.

Drzewa decyzyjne

Jak wspomniano, *drzewa decyzyjne* są środkiem reprezentacji wiedzy w sposób deklaratywny oraz, w pewnym stopniu, w sposób proceduralny (gdyż określają kolejność badania warunków elementarnych, co może być uznane za pewną formę *metawiedzy*). Umożliwiają one bardzo efektywną organizację procesu wnioskowania diagnostycznego, z zastosowaniem wielu cech.

Drzewo decyzyjne jest specjalnym rodzajem klasyfikatora, czyli struktury, która jest [Quinlan, 1993]:

- bądź *liściem* wskazującym klasę (stanu),
- bądź *węzłem decyzyjnym* określającym pewien test do przeprowadzenia na wartości cechy (przypisanej do tego węzła), z którego wychodzi tyle krawędzi, ile jest możliwych wyników wspomnianego testu, a każda z tych krawędzi jest zakończona poddrzewem.

Stosowanie drzewa do klasyfikacji pewnego obiektu rozpoczyna się *od korzenia* tego drzewa i odbywa się aż *do osiągnięcia liścia*. O sposobie poruszania się wzdłuż krawędzi drzewa decydują *wyniki testów* przeprowadzanych na wartościach cech klasyfikowanego obiektu w poszczególnych węzłach decyzyjnych. Obiektowi zostaje przypisana taka klasa, która odpowiada osiągniętemu liściowi drzewa.

Drzewo decyzyjne służy do reprezentacji wiedzy diagnostycznej dotyczącej pewnej klasy C obiektów[25]. Jest ono zdefiniowane poprzez:

- *określenie jego struktury* (która jest grafem-drzewem),
- *określenie cech przypisanych do węzłów nieterminalnych* (czyli takich, które nie są węzłami końcowymi),
- *określenie uporządkowanych list wartości*, które przypisane są do poszczególnych krawędzi wychodzących z danego węzła nieterminalnego (uporządkowanie tych list jest istotne wtedy, gdy reprezentowane drzewo decyzyjne zostało poddane optymalizacji),

[25] Jeśli w bazie wiedzy ma być zapisana wiedza dotycząca więcej niż jednej klasy obiektów, wymagane jest zapisanie zbioru drzew decyzyjnych, dla którego wprowadza się nazwę "las".

- *określenie wartości cechy decyzyjnej* (np. nazwy klasy stanu), przypisanych do węzłów terminalnych (liści drzewa).

Węzłom można przyporządkować wagi, na podstawie których mogą być np. określane stopnie prawdziwości konkluzji, danej w postaci stwierdzenia "dla klasyfikowanego obiektu o cecha decyzyjna d ma wartość v".

2.5.4 Wnioski wynikające z przyjętego sposobu reprezentacji danych i wiedzy

Przedstawione opisy wybranych środków reprezentacji danych i wiedzy diagnostycznej umożliwiają sformułowanie poniższych wniosków:

1. W celu zapisania danych i wiedzy (np. w relacyjnych bazach danych) należy uprzednio zdefiniować dziedzinę zastosowania. Definicja ta może mieć formę *słowników*.
2. Definicja dziedziny zastosowania jest wspólna dla potrzeb zapisu danych i wiedzy. Z tego względu dane i wiedza winny być reprezentowane wspólnie w jednej bazie relacyjnej, w której można wyróżniać podbazy na poziomie logicznym (takie jak: baza danych stałych, baza danych zmiennych, baza wiedzy).
3. Przyjęty sposób reprezentacji danych powinien umożliwić łatwą wymianę danych z bazami zewnętrznymi o schemacie logicznym zgodnym z zaleceniami organizacji *MIMOSA*.

Uwzględniając powyższe wnioski, opracowano schemat logiczny bazy danych i wiedzy *EMPREL*, opisany skrótowo w rozdziale 4.1 wraz z dodatkiem A.

2.6 Proces pozyskiwania wiedzy diagnostycznej

Celem *procesu pozyskiwania wiedzy* jest *pozyskanie zasobu wiedzy i doświadczenia*, odpowiadających zakresowi zadań z danej dziedziny zastosowania, ze zidentyfikowanych *źródeł wiedzy*, oraz zapisanie ich w *bazie wiedzy* w sposób umożliwiający *skuteczne wspomaganie działania człowieka* podczas rozwiązywania problemów z tej dziedziny.

W procesie pozyskiwania wiedzy można wyróżnić następujące stadia:

1. określenie *dziedziny zastosowania* oraz *zakresu zadań*, których rozwiązywanie ma być wspomagane przez system doradczy, dla którego wiedza ma być pozyskana,
2. identyfikacja *źródeł wiedzy*,
3. wybór *środków reprezentacji wiedzy*,
4. *pozyskanie wiedzy* i zapis *wstępnej, prototypowej wersji bazy wiedzy* dla pewnego podzbioru typowych problemów danej dziedziny,
5. *weryfikacja wstępnej wersji bazy wiedzy* dla zbioru przypadków testowych,
6. *rozbudowa bazy wiedzy* do wersji pełnej (obejmującej pełny zakres zadań danej dziedziny zastosowania),

7. *zidentyfikowanie i usunięcie błędów* wykrytych w bazie wiedzy,
8. *walidacja bazy wiedzy* przez niezależnych specjalistów i jej *przekazanie do eksploatacji*.

Wiedza diagnostyczna jest pozyskiwana z jej źródła, a następnie zostaje przetworzona w sposób odpowiadający przyjętym sposobom reprezentacji i zapisana w bazie wiedzy systemu doradczego. Pozyskiwanie wiedzy odbywa się z zastosowaniem *odpowiednich metod oraz środków wspomagających*.

2.6.1 Źródła wiedzy diagnostycznej

Najważniejszymi źródłami wiedzy dla diagnostycznych systemów doradczych są:

- *specjaliści z zakresu danej dziedziny zastosowania*:
  - *bezpośrednio* – poprzez osobisty udział danego specjalisty w procesie pozyskiwania wiedzy,
  - *pośrednio* – specjaliści są autorami pozycji literatury fachowej, publikacji itp., których treść jest następnie interpretowana przez osoby trzecie lub przez odpowiednie oprogramowanie;
- *bazy danych*:
  - zawierające *wyniki obserwacji diagnostycznych* badanego obiektu lub zbioru obiektów (dla obiektów-konkretów),
  - zawierające *wyniki obliczeń symulacyjnych* prowadzonych z zastosowaniem odpowiedniego modelu matematycznego, realizowanego przez oprogramowanie komputerowe (*modele proste*: *przyczyna* $\Rightarrow$ *skutek* i odpowiednie *modele odwrotne* [Cholewa, White, 1993], [Cholewa, Kiciński, 1996]).

Specjaliści mogą dostarczać *opisy przypadków diagnostycznych* (ang. *case stories*) dla potrzeb pozyskiwania wiedzy diagnostycznej oraz testowania bazy wiedzy. Podobnie, bazy danych mogą być źródłem przykładów do *pozyskiwania wiedzy metodami indukcyjnymi* (z zastosowaniem tzw. *uczenia z nauczycielem*) lub umożliwiać *odkrywanie* empirycznych relacji diagnostycznych.

2.6.2 Pozyskiwanie wiedzy diagnostycznej od specjalistów

Pozyskiwanie wiedzy od specjalistów, za pośrednictwem programisty bazy wiedzy systemu doradczego lub bez tego pośrednictwa, stanowiło historycznie najwcześniej stosowaną grupę metod. Również współcześnie nie jest wskazane pominięcie tych metod przy konstruowaniu wstępnych wersji systemu doradczego. Uzasadnione jest natomiast *komputerowe wspomaganie tego procesu*, którego celem winno być zwłaszcza *znaczne ograniczenie roli programisty bazy wiedzy* (będącego dodatkowym pośrednikiem w procesie pozyskiwania wiedzy).

Jak wspomniano, wiedza od specjalistów z danej dziedziny może być pozyskiwana *w sposób bezpośredni* lub *pośredni*.

Zagadnienia pozyskiwania wiedzy bezpośrednio od specjalistów są przedmiotem wielu publikacji z zakresu socjologii i psychologii. Udział specjalisty w procesie pozyskiwania wiedzy może być czynny lub bierny.

Udział czynny wspomagany jest przez pośredników, którymi mogą być:

- *programiści bazy wiedzy*[26] (w literaturze angielskojęzycznej zwani *Knowledge Engineers*),
- *specjalne środki programowe.*

W pierwszym przypadku programista bazy wiedzy zwykle *przeprowadza wywiady ze specjalistą*, dotyczące bądź przypadków typowych dla danej dziedziny, bądź hipotetycznych problemów. Wypowiedzi specjalisty są następnie interpretowane przez programistę i zapisywane w bazie wiedzy. Taki przebieg procesu pozyskiwania wiedzy nastręcza jednak wiele problemów, związanych np. ze sposobem rozumienia używanych pojęć przez specjalistę i programistę. Przykładowy przebieg procesu pozyskiwania wiedzy, dotyczącego diagnostyki obrabiarki sterowanej numerycznie, opisano w [Moczulski, 1989].

Bierny udział specjalisty może polegać na:

- analizie *protokołów werbalnych*, tj. zarejestrowanych wypowiedzi specjalisty, "głośno myślącego" podczas rozwiązywania problemów diagnostycznych,
- *obserwacji specjalisty podczas rozwiązywania przez niego zadań diagnostycznych* z danej dziedziny zastosowania w jego środowisku działania (np. bezpośrednio przy maszynie), przy czym wskazane jest takie prowadzenie obserwacji, by w jak najmniejszym stopniu oddziaływało to na specjalistę.

Programista analizuje sposób postępowania specjalisty i zapisuje wiedzę w odpowiedniej bazie systemu doradczego. Zalecane jest przedstawienie specjaliście pozyskanej w ten sposób wiedzy do jego oceny. Również działania programisty w tej fazie procesu pozyskiwania wiedzy winny być wspomagane za pomocą odpowiednich środków programowych.

Stadia pozyskiwania wiedzy od specjalistów

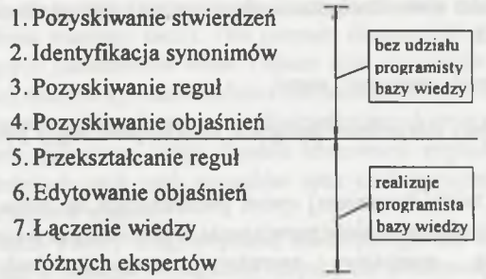
W procesie pozyskiwania wiedzy od specjalistów można wyróżnić kilka stadiów (rys. 2.4). Jednym z zadań, które należy rozwiązać, jest opracowanie takich środków wspomagania procesu pozyskiwania wiedzy, które umożliwią przeprowadzenie tego procesu bez udziału programisty bazy wiedzy. Stadia wstępne mogą być wspomagane właśnie w ten sposób.

Pozyskiwanie stwierdzeń obejmuje także pozyskiwanie pojęć (klasy obiektów, obiekty, czynności, cechy, wartości cech, relacje itp.).

Do wyrażania wartości stwierdzeń dokładnych stosowane są wartości logiczne $\{prawda, fałsz\}$, a do stwierdzeń przybliżonych — np. pary wartości $\{konieczność; możliwość\}$.

Synonimy zapisywane są w celu ułatwienia porozumienia się różnych użytkowników bazy wiedzy (specjalista / programista bazy wiedzy / użytkownik bazy wiedzy) ze sobą, a także w celu ułatwienia programiście bazy wiedzy i użytkownikowi zrozumienia znaczenia pojęć, stosowanych przez specjalistę. Z zagadnieniem synonimów wiąże się również problem identyfikacji zakresu znaczenia pojęć stosowanych przez wszystkie osoby współdziałające przy budowie bazy wiedzy, a także korzystające później z jej zawartości.

[26] W dalszym ciągu pracy zwani także "programistami".



Rys. 2.4. Stadia procesu pozyskiwania wiedzy od specjalistów
Fig. 2.4. Stages of the process of Knowledge Acquisition from experts

Stadia 1.–4. (rys. 2.4) realizowane są samodzielnie przez specjalistę, którego wspomaga i "prowadzi" w odpowiedni sposób system formularzy "papierowych" (lub formularz "elektroniczny").

Pozyskiwanie wiedzy od specjalisty wymaga wcześniejszego zgromadzenia pewnej "porcji" wiedzy, takiej jak:

- *metawiedza* (w tym wypadku — wiedza o przebiegu procesu pozyskiwania wiedzy i stosowanych metodach, która może być reprezentowana w odpowiedni sposób — np. jako *baza metawiedzy*);
- *wiedza dotychczas pozyskana*, z którą nowa "porcja" wiedzy powinna zostać zintegrowana oraz którą można "przy okazji" poddać weryfikacji przez specjalistę.

Kroki 5.-7. (rys. 2.4) wykonywane są przez programistę bazy wiedzy. Stosuje on specjalne środki programowe, jak *edytor bazy wiedzy*. Wyniki tych stadiów muszą być zweryfikowane przez autora bazy wiedzy oraz przez specjalistę(ów).

Zastosowanie formularzy do wspomagania procesu pozyskiwania wiedzy

Aby wyeliminować przyczyny problemów nomenklaturowych, dąży się do umożliwienia pozyskiwania wiedzy bez pośrednika, którym jest programista bazy wiedzy [Buchanan *et al.*, 1983]. Wówczas specjalista wspomagany jest za pomocą specjalnych środków, wśród których można wymienić różnego rodzaju formularze. Autor pracy zaproponował do tego celu specjalny *formularz papierowy* oraz specjalistyczne oprogramowanie, zwane *formularzem elektronicznym* (opis w rozdziale 4.2.3). Oba rodzaje formularzy umożliwiają tworzenie i uzupełnianie słowników pojęć, obiektów, cech tych obiektów, opisów uszkodzeń, reguł wnioskowania i reguł działania, procedur postępowania, elementów systemu objaśnień itp.

Łączenie wiedzy pozyskanej od wielu specjalistów

Potrzeba stosowania wiedzy przybliżonej pojawia się zwłaszcza w wypadku *łączenia wiedzy* pochodzącej od wielu specjalistów. Istnieje wówczas możliwość zgromadzenia rozbieżnych opinii specjalistów w zakresie:

- określania wartości stwierdzeń stosowanych np. do budowy elementarnych warunków przesłanek reguł,
- określania "stopnia pewności" reguł,
- określania wartości stwierdzeń będących konkluzjami reguł pochodzących od wielu specjalistów.

Agregacja (zwana także łączeniem) *opinii pochodzących od wielu specjalistów* stanowi istotny problem, który winien znaleźć rozwiązanie w procesie pozyskiwania wiedzy. W tym celu określane są specjalne *operatory agregacji* (zob. [Cholewa, 1985], [Cholewa, 1996a]). Współczesne techniki komputerowe umożliwiają łączenie opinii wielu specjalistów, co winno stanowić niewątpliwą zaletę i może być stosowane z pożytkiem w celu zwiększenia stopnia pewności konkluzji uzyskiwanych jako wynik wnioskowania diagnostycznego. W literaturze (np. [Cholewa, Pedrycz, 1987]) można znaleźć jednak informacje o trudnościach, które napotkano przy pozyskiwaniu wiedzy od wielu specjalistów. W związku z tym często źródłem wiedzy dla danej bazy wiedzy jest tylko jeden specjalista.

W pracy przyjęto, że łączenie wiedzy pochodzącej od wielu specjalistów może być dokonywane:

- przez programistę bazy wiedzy, który będzie określał lub aktualizował wartości stopni konieczności i stopni dostateczności poszczególnych reguł (posługując się w tym celu odpowiednią aplikacją — *edytorem bazy wiedzy*);
- bezpośrednio przez system doradczy, który będzie aktualizował wartości stopni konieczności i stopni dostateczności na drodze specyficznego uczenia się.

Zawartość bazy wiedzy może więc podlegać nieustannej modyfikacji. Wymaga to szczególnie starannego testowania zawartości bazy wiedzy za pomocą odpowiedniego zbioru danych (przypadków) testowych.

2.6.3 Pozyskiwanie wiedzy diagnostycznej z baz danych

W poprzednich rozdziałach uzasadniono potrzebę bezpośredniego udziału specjalistów w procesie pozyskiwania wiedzy, a zwłaszcza w jego początkowych stadiach. Jednak ten sposób pozyskiwania wiedzy *nie jest efektywny*, w tym sensie, że możliwe jest zapisanie niewielu reguł jako wyniku jednej sesji ze specjalistą. Utrudniona jest także *modyfikacja zawartości istniejącej bazy wiedzy* (np. dodanie nowej reguły). Dołożenie czasem tylko jednej nowej reguły do dużej bazy reguł powoduje, że system doradczy przestaje działać poprawnie. W związku z tym celowe jest zastosowanie innych, bardziej efektywnych metod pozyskiwania wiedzy. Zalicza się do nich między innymi:

- *indukcyjne metody* tzw. *uczenia maszynowego* (na podstawie przykładów wstępnie sklasyfikowanych przez nauczyciela);
- *metody odkrywania zależności jakościowych i ilościowych* (funkcyjnych) *w bazach danych* (ta klasa metod zwana jest w literaturze angielskojęzycznej *Knowledge Discovery in Databases*[27]), zob. [Żytkow, Zembowicz, 1993], [Zembowicz, Żytkow, 1992], [Michalski, 1997].

[27]Stosowana jest także nazwa *Data Mining*, która dotyczy metod umożliwiających rozwiązanie prostszych zadań pozyskiwania wiedzy z baz danych [Michalski, 1997].

Obie grupy metod mogą być stosowane do licznych zbiorów danych (zawierających wiele cech i bardzo dużo wartości cech). Dla potrzeb diagnostyki maszyn gromadzi się wiele danych dotyczących parametrów wejść i stanu oraz warunków działania badanych obiektów. Np. w bazach danych systemu nadzoru turbozespołu, z niewielkim krokiem czasowym, gromadzone są wartości cech sygnałów diagnostycznych oraz wartości parametrów roboczych turbozespołu [Cholewa, 1996]. Sposób obserwacji sygnałów diagnostycznych powoduje, że w wartościach cech tych sygnałów oraz cech parametrów "procesowych" można zidentyfikować wpływ szumu.

Zadanie pozyskiwania wiedzy diagnostycznej może polegać na zidentyfikowaniu wiedzy diagnostycznej (np. o relacjach diagnostycznych) na podstawie tych danych. Dane mogą być wstępnie sklasyfikowane poprzez przypisanie im klas stanu (przez specjalistę lub np. przez samo źródło danych diagnostycznych [Cholewa, 1996]). Nie zakłada się, że ta wstępna klasyfikacja jest zawsze wykonywana bezbłędnie.

Pozyskiwanie wiedzy indukcyjnymi metodami "uczenia maszynowego"

Do pozyskiwania wiedzy indukcyjnymi metodami uczenia maszynowego stosowany jest tzw. *model atrybutowy*, w którym dane uczące zapisywane są za pomocą jednokolumnowej macierzy *wartości cech opisujących właściwości obiektu* oraz *wartości cechy decyzyjnej* (np. nazwa lub numer klasy stanu, do której należy opisywany obiekt). Zwykle zbiór dostępnych danych uczących (*przykładów* i *kontrprzykładów* danych koncepcji — np. klas stanów) jest bardzo nieliczny w odniesieniu do liczby możliwych elementów przestrzeni reprezentacji cech. W zasadzie nie przyjmuje się żadnych założeń dotyczących rozkładu danych uczących w przestrzeni reprezentacji cech. W szczególności nie można zakładać, że rozkład ten jest równomierny.

Uzasadnienie tego faktu jest następujące:

> *Przyjmijmy, że w wyniku "uczenia maszynowego" ma być zidentyfikowana baza wiedzy, której zawartość winna umożliwić identyfikację stanu zdatnego oraz kilkunastu stanów technicznych maszyny, odpowiadających uszkodzeniom różnych podzespołów. Prawdopodobieństwo wystąpienia stanu zdatnego (poprawnej pracy) właściwie skonstruowanej, wytworzonej i eksploatowanej maszyny jest dużo większe niż prawdopodobieństwo wystąpienia każdego uszkodzenia maszyny. Jeśli dla każdej klasy stanu, która ma być rozpoznawana przez klasyfikator, dostarczono w przybliżeniu podobną liczbę przykładów uczących, to ich rozkład w przestrzeni przykładów musi być nierównomierny.*

W ramach opisanych badań, do pozyskiwania wiedzy metodami "uczenia maszynowego" stosowane są przede wszystkim:

- indukcyjna metoda pozyskiwania wiedzy *za pomocą pokryć* według algorytmu A^Q [Michalski, 1983];
- generowanie *drzew decyzyjnych* w sposób indukcyjny [Quinlan, 1986], [Quinlan, 1993];
- pozyskiwanie *reguł i klasyfikatorów przybliżonych* [Pawlak, 1982], [Mrózek, 1993], [Skowron, 1993], [Nowicki, Słowiński, Stefanowski, 1992], [Wong, Li, Ziarko, 1986], [An *et al.*, 1995], [Ziarko, Shan, 1996].

Powszechnie stosowanym źródłem danych uczących do "uczenia maszynowego" są bazy danych diagnostycznych (np. bazy danych systemów monitorowania turbozespołów).

Metody "uczenia maszynowego" mogą być także stosowane do modyfikacji i poprawiania zawartości baz wiedzy. Jedną z możliwości jest tutaj tzw. *uczenie przyrostowe* (ang. *incremental learning*) [Hong, Mozetic, Michalski, 1986], [Wnęk *et al.*, 1995], [Michalski, 1997], stosowane w przypadku metod indukcyjnego pozyskiwania reguł. Danymi wejściowymi do procesu uczenia są *reguły wstępne* (zwane *hipotezami wejściowymi*) oraz *dane uczące*. W wyniku uczenia reguły ulegają modyfikacji w taki sposób, aby jak najlepiej klasyfikować dostarczone dane uczące.

Reguły wstępne mogą być pozyskane od specjalisty lub specjalistów. Pojawia się wówczas możliwość modyfikowania zawartości bazy wiedzy za pomocą dostarczonych danych uczących (przykładów), wstępnie sklasyfikowanych przez (być może innego) specjalistę[28].

Pozyskiwanie wiedzy metodami odkryć w bazach danych

Wstępne badania, związane z zastosowaniem tej grupy metod do pozyskiwania wiedzy diagnostycznej, zostały zainicjowane przez autora w ostatnim czasie we współpracy z J. M. Żytkowem [Moczulski, Żytkow, 1997]. Badania te łączą się także z identyfikacją wiedzy o relacjach "odwrotnych".

Maszynowe (lub *automatyczne*) *odkrywanie wiedzy* różni się istotnie od *uczenia maszynowego* przede wszystkim tym, że metodami odkryć pozyskiwana jest *nowa porcja wiedzy* [Żytkow, 1996] (zob. przykład automatycznego odkrywania praw ruchu po równi pochyłej [Huang, Żytkow, 1996]), podczas gdy metodami "uczenia maszynowego" może być pozyskana jedynie wiedza już odkryta wcześniej, a celem procesu uczenia jest jej reprezentacja w bazie wiedzy danego systemu doradczego [Żytkow, 1993].

Diagnostyczne bazy danych, zawierające wartości cech, które opisują wejścia i wyjścia obserwowanych obiektów konkretnych, mogą być źródłem użytecznej wiedzy o relacjach diagnostycznych [Moczulski, Żytkow, 1997]. "Odkrywanie" wiedzy ma na celu *identyfikację regularności* istniejących w zbiorze danych, który jest zawarty w bazie. *Regularność* określana jest poprzez pewien *obraz* (ang. *pattern*) oraz *zakres*, w którym ten obraz występuje [Żytkow, Zembowicz, 1993]. Przykładami obrazów są *tablice kontyngencji*[29], *równania* i *równoważności logiczne*. Zakres występowania regularności określony jest jako podzbiór danych spełniających warunek złożony, będący koniunkcją warunków prostych (typu nierówności). Każdy warunek prosty dzieli zbiór wartości pojedynczej cechy na dwa podzbiory, w przybliżeniu równoliczne. Oprócz odkrywania regularności, w wypadku baz danych zawierających szeregi czasowe, stosowane jest podejście łączone, którego istotą jest *poszukiwanie maksimów i minimów*, powiązane z *poszukiwaniem regularności reprezentowanych w sposób ilościowy* (w postaci równań) [Żytkow, Zembowicz, 1996].

Dla potrzeb wnioskowania diagnostycznego najlepszym środkiem są *równania* [Moczulski, Żytkow, 1997]:

$$S = f(Y, X_2), \tag{2.78}$$

[28] Przykład zastosowania uczenia przyrostowego do pozyskiwania wiedzy dotyczącej doboru klas łożysk tocznych zawiera [Moczulski, 1997a]. Źródło reguł wstępnych stanowiła literatura (podręczniki i poradniki opisujące zasady konstruowania węzłów łożyskowych). Dane uczące zostały opracowane na podstawie przykładów zawartych w zaleceniach i publikacjach wydanych przez wytwórców łożysk tocznych.

[29] Ang. *contingency table*, co jest także tłumaczone jako *tablica zależności* lub *tablica wielodzielcza* [Volk, 1965].

Tabl. 2.1. Poszukiwanie różnych form wiedzy inicjowane przez odpowiednie testy [Zembowicz, Żytkow, 1996]

Sposób reprezentacji wiedzy	Przeszukiwanie inicjowane przez
Hierarchia pojęć	test relacji równoważności
Sieci podzbiorów	test podzbiorów
Funkcje	test funkcjonalności oparty na definicji funkcji
Multifunkcje	test oparty na definicji funkcjonalności i ciągłości
Drzewa decyzyjne	moc predykcji tablic kontyngencji
Sieci przyczynowe	zależności pomiędzy cechami dla wszystkich rekordów oraz w przekrojach

gdzie: Y — symptomy diagnostyczne, S — wartości cech stanu technicznego i X_2 — wartości cech warunków działania obiektu. Umożliwiają one dokładną i jednoznaczną predykcję stanów obiektu. Wyniki pomiarów i obserwacji zawierają szum, są niekompletne, przybliżone lub rozmyte, więc zbiory danych zapisanych w tych bazach są nośnikiem niekompletnej informacji o zależnościach ilościowych. Do identyfikacji regularności w bazie danych w pierwszej fazie "odkrywania" stosuje się tablice kontyngencji [Żytkow, Zembowicz, 1993]. Są one podstawowym środkiem reprezentacji dwuwymiarowych regularności, zaś inne środki reprezentacji wiedzy, jak: *równania*, *taksonomie*, *reguły* i *pojęcia*, mogą być uznane za przypadki specjalne tych tablic [Zembowicz, Żytkow, 1996].

System dokonywania odkryć w bazach danych diagnostycznych winien działać *w sposób autonomiczny*. System ten dobiera właściwy sposób reprezentacji wiedzy oraz steruje procesem przeszukiwania (zob. tablica 2.1).

Wykryte obrazy są oceniane za pomocą *miary istotności* Q, określonej jako prawdopodobieństwo losowego wygenerowania danego obrazu dla zmiennych (cech), które są niezależne. Empirycznie stwierdzono [Zembowicz, Żytkow, 1996], że należy stosować wartości $Q < 10^{-5}$, a więc jest bardzo małe prawdopodobieństwo, że wykryte obrazy mogły powstać w sposób losowy. Większe wartości Q powodują, że system odkryć identyfikuje bardzo wiele regularności, wśród których jedynie niektóre są faktycznie interesujące ze względu na małą wartość miary istotności. Powoduje to trudności dotyczące dalszego selekcjonowania regularności i wydłuża czas działania systemu odkryć.

Moc predykcji danej tablicy kontyngencji oceniana jest za pomocą miary Cramera V, określonej jako:

$$V = \sqrt{\chi^2/(N \cdot \min\{(M_{row}-1),(M_{col}-1)\}} \in [0,1], \tag{2.79}$$

gdzie wartość statystyki χ^2 (dla częstości z próby A_{ij} porównywanych z częstościami E_{ij} oczekiwanymi w wypadku hipotezy zerowej o braku jakiejkolwiek zależności pomiędzy obiema cechami) określa się jako:

$$\chi^2 = \sum_{i,j} \frac{(A_{ij} - E_{ij})^2}{E_{ij}}, \tag{2.80}$$

oraz N — liczba rekordów w tablicy zawierającej dane, M_{row}, M_{col} — odpowiednio liczba wierszy i kolumn w ocenianej tablicy kontyngencji. Im większa wartość miary V, tym bardziej unikalne predykcje można uzyskać na podstawie rozpatrywanej tablicy kontyngencji.

Istotną cechą miary V jest jej niezależność od rozmiarów tablicy oraz od liczby rekordów. Dla $V > 0,9$ zależność reprezentowana przez tablicę może być uznana za równoważność [Zembowicz, Żytkow, 1996].

Tablice kontyngencji, dla których wartość miary V jest dostatecznie duża, są przedmiotem dalszego przeszukiwania w celu znalezienia *równań* (zależności o charakterze ilościowym). Równania są poszukiwane jedynie w takim zbiorze danych, w którym zachodzi *relacja funkcjonalności*, określona w następujący sposób:

> Dany jest zbiór par
>
> $$D = \{(x_i, y_i) \mid x_i \in X \ \wedge \ i = 1, \ldots, N\}.$$
>
> y jest funkcją x wtedy i tylko wtedy, gdy dla każdego $x_0 \in X$ istnieje tylko jedno y_0, takie że $(x_0, y_0) \in D$.

Równania poszukiwane są w następujących krokach [Zembowicz, Żytkow, 1996]:

1. generowanie nowych termów funkcyjnych;
2. wybór pary termów (jako wyrażeń odpowiednio dla x i y);
3. generowanie i ocena równań dla każdej pary termów.

Do oceny stopnia dopasowania równania do danych $\{(x_i, y_i, \sigma_i)\}$ stosowany jest test χ^2, określony następująco:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{y_i - f(x_i, a_1, \ldots, a_q)}{\sigma_i} \right)^2, \tag{2.81}$$

gdzie σ_i jest odchyleniem standardowym odpowiadającej jej wartości zmiennej zależnej y_i, $y = f(x, a_1, \ldots, a_q)$ jest modelem matematycznym wzorca (jak np. $y = a_1 + a_2 x + a_3 x^2$), zaś $a_1, \ldots, a_q$ są dobrane w taki sposób, aby minimalizować wartość χ^2 wg (2.81).

Stosując oznaczenia jak w (2.78), można opisać dwa odmienne podejścia do zadania identyfikacji "odwrotnych" relacji funkcyjnych:

- wpierw identyfikowane są zależności "wprost" postaci: $Y = f(S, X_2)$, a następnie podejmowana jest próba przekształcenia takich równań do postaci "odwrotnej": $S = g(Y, X_2)$; ważnym problemem jest istnienie rozwiązania tak postawionego zadania;
- podejmowana jest próba bezpośredniej identyfikacji funkcyjnej zależności "odwrotnej" $S = g(Y, X_2)$, co następuje poprzez częściową zamianę ról pomiędzy zmiennymi niezależnymi S oraz Y.

2.7 Podsumowanie dotyczące przedmiotu badań

Reasumując, dla potrzeb prowadzonych badań przyjęto następujące podstawowe założenia dotyczące przedmiotu tych badań:

1. obiekty opisywane są za pomocą cech i ich wartości;
2. rozpatrywana jest jedynie skończona liczba klas obiektów, posiadających skończoną liczbę elementów (maszyn), które dla potrzeb wnioskowania diagnostycznego identyfikowane są jedynie za pomocą skończonej liczby cech o wartościach dyskretnych;
3. czas reprezentowany jest za pomocą struktur punktowych;
4. wartości cech obiektów reprezentowane są za pomocą stwierdzeń dynamicznych;
5. związki zachodzące pomiędzy wejściami i wyjściami diagnozowanego obiektu opisywane są za pomocą modelu systemowego "statycznego" (bez uwzględnienia sekwencji wejść, wyjść i stanów diagnozowanych obiektów);
6. celem postępowania diagnostycznego jest określenie:
   (a) co najmniej ogólnej (przybliżonej) diagnozy stanu obiektu, tj. określenie klasy stanu [s], lub
   (b) uszczegółowionej (przybliżonej) diagnozy stanu obiektu;
7. wiedza diagnostyczna, dotycząca dziedziny zastosowań, będzie reprezentowana w sposób deklaratywny;
8. wiedza będzie pozyskiwana:
   (a) od specjalistów z danej dziedziny zastosowań;
   (b) na podstawie zawartości baz danych (zawierających przykłady i kontrprzykłady danych pojęć) — z zastosowaniem metod indukcyjnych.

Szczegółowe przedstawienie przedmiotu badań w dalszym ciągu umożliwi opis wybranych metod, które stanowią logiczną podstawę do opracowania zestawu środków wspomagania procesu pozyskiwania wiedzy, zweryfikowanego w ramach badań przeprowadzonych na potrzeby niniejszej rozprawy.

Celowe jest podjęcie badań nad metodami kwantyzacji cech ilościowych, które umożliwią odpowiedni wybór liczby poziomów kwantyzacji oraz wartości progowych. Metody te winny uwzględniać wiele kryteriów, jak np. minimalnego łącznego błędu klasyfikacji, przeciwdziałając jednocześnie nadmiernemu dopasowaniu poziomów kwantyzacji do zbioru danych. Ze względu na potrzeby diagnostyki maszyn należy mieć ponadto na uwadze możliwość przypisania wartościom cechy pochodnej jakościowej wartości lingwistycznych, co winno ułatwić zrozumienie reguł przez innych specjalistów i końcowego użytkownika.

W przyszłości przewiduje się weryfikację przydatności koncepcji *autonomicznego systemu odkryć*, który zintegrowałby:

- *moduł sterujący w sposób autonomiczny przebiegiem eksperymentu*, związanego z gromadzeniem danych, które mogą być zastosowane do weryfikacji hipotez, dotyczących związków zachodzących pomiędzy cechami niezależnymi i zależnymi;
- *moduł umożliwiający uzyskanie nowego zbioru danych* przy zadanych wartościach zmiennych (cech) niezależnych;
- *moduł dokonujący odkryć* związków, zachodzących pomiędzy cechami niezależnymi i zależnymi.

Koncepcja ta może być zweryfikowana zwłaszcza na drodze eksperymentu numerycznego, realizowanego za pomocą tzw. oprogramowania symulacyjnego.

Rozdział 3

Metody pozyskiwania wiedzy diagnostycznej

W rozdziale omówiono *metody pozyskiwania wiedzy diagnostycznej*, które zostały zastosowane w niniejszej pracy. Metody te związane są z pozyskiwaniem wiedzy z podstawowych jej źródeł, jak specjaliści oraz bazy danych (zob. rozdział 2.6.1). Ponadto opisano stosowane *sposoby oceny wyników pozyskiwania wiedzy* z obu tych źródeł. Rozdział zawiera także opis opracowanego *scenariusza procesu pozyskiwania wiedzy*.

3.1 Pozyskiwanie relacji empirycznych od specjalistów

Stosowane metody dotyczą wspomagania pozyskiwania *nowej porcji wiedzy*, zgodnie z poniższymi modelami (por. [Buchanan *et al.*, 1983]):

- Model 1: *specjalista plus programista bazy wiedzy jako niezbędny pośrednik-interpretator zapisów specjalisty* (wtedy, gdy do pozyskiwania wiedzy od specjalisty stosowany jest formularz "papierowy"),
- Model 2: *specjalista plus ewentualnie programista bazy wiedzy — pomocny w edytowaniu zapisów specjalisty*, szczególnie w celu zintegrowania nowej "porcji" wiedzy z wiedzą wcześniej pozyskaną (wtedy, gdy stosowany jest formularz "elektroniczny"),

przy czym wybór modelu jest uwarunkowany przede wszystkim możliwością samodzielnego i sprawnego posługiwania się przez specjalistę sprzętem komputerowym oraz oprogramowaniem wspomagającym proces pozyskiwania wiedzy.

Pozyskiwanie wiedzy w obu tych modelach dotyczy przede wszystkim wiedzy operacyjnej w postaci *empirycznych relacji diagnostycznych*. Jednocześnie następuje pozyskiwanie wiedzy podstawowej w zakresie wymaganym do umożliwienia zapisu relacji diagnostycznych.

Bardzo istotna rola specjalistów w procesie pozyskiwania wiedzy polega na uczestnictwie w *ocenie wiedzy już pozyskanej* (zob. rozdział 3.3).

3.1.1 Empiryczna relacja diagnostyczna

Relacja diagnostyczna określa związek pomiędzy wejściami i stanem obiektu a wyjściami tego obiektu. Jeśli związek ten będzie zidentyfikowany na drodze empirycznej, stosowana będzie nazwa *empiryczna relacja diagnostyczna* [Moczulski, 1997].

Empiryczna relacja diagnostyczna $\Re$ jest podzbiorem iloczynu kartezjańskiego następujących dziedzin cech (złożonych), które są stosowane do opisu obiektu w określonej chwili czasu $\theta < \theta_{aw}$ (gdzie θ_{aw} jest chwilą czasu, w której nastąpiła awaria):

$$\Re \subset V_X \times V_S \times V_Y, \tag{3.1}$$

gdzie (zob. rozdział 2.1):

V_X — p-wymiarowa przestrzeń wartości cech $a_x \in A_X$ opisujących wejścia $X = X_1 \cup X_2$ związane z użytecznym działaniem obiektu oraz warunkami jego działania;

V_S — q-wymiarowa przestrzeń wartości cech $a_s \in A_S$ stanów (złożonych) $S = S_2$, wywołanych procesem wytwarzania lub użytkowania obiektu (wartości cech złożonych reprezentowane są przez macierze jednokolumnowe, których elementami są wartości tzw. *stanów elementarnych* — opisywanych przez cechy ilościowe lub jakościowe);

V_Y — r-wymiarowa przestrzeń wartości cech $a_y \in A_Y$ opisujących wyjścia $Y = Y_1 \cup Y_2$ obiektu, związane z jego użytecznym działaniem oraz z występowaniem procesów resztkowych.

Empiryczna relacja diagnostyczna $\Re$ jest określona jako predykat postaci [Moczulski, 1997b]:

$$\forall(o \in D(\Re))\ [P_X(o) \wedge P_S(o) \wedge P_Y(o)], \tag{3.2}$$

gdzie $D(\Re) \subset O$ jest dziedziną tej relacji (tj. podzbiorem zbioru wszystkich rozpatrywanych obiektów O).

$P_X(o), P_S(o), P_Y(o)$ oznaczają jednoargumentowe predykaty, będące koniunkcjami skończonej liczby warunków elementarnych o postaci:

$$\begin{aligned} P_X(o) &= \bigwedge_{i \in I \subset \overline{1..p}} [a_{xi}(o) \propto_i v(a_{xi})], \\ P_S(o) &= \bigwedge_{j \in J \subset \overline{1..q}} [a_{sj}(o) \propto_j v(a_{sj})], \\ P_Y(o) &= \bigwedge_{k \in K \subset \overline{1..r}} [a_{yk}(o) \propto_k v(a_{yk})], \end{aligned} \tag{3.3}$$

gdzie $\propto_i, \propto_j, \propto_k$ są dwuargumentowymi operatorami relacji spośród należących do zbioru $\{=, \neq, <, \leqslant, \geqslant, >\}$, I, J, K są zbiorami indeksów (podzbiorami odpowiednich odcinków liczb naturalnych), poszczególne cechy spełniają $a_{xi} \in A_X$, $i \in I$, $a_{sj} \in A_S$, $j \in J$, $a_{yk} \in A_Y$, $k \in K$ oraz wartości cech spełniają $v(a_{xi}) \in Dom(a_{xi})$, $v(a_{sj}) \in Dom(a_{sj})$, $v(a_{yk}) \in Dom(a_{yk})$.

Relacja diagnostyczna $\Re$ umożliwia zapisanie związku co najmniej jednoczesnego współwystępowania (w określonej chwili czasu eksploatacji θ) określonych wartości cech wejść, wyjść i stanów. Jest ona środkiem pośrednim, który pozwala specjaliście na wyrażenie opinii. Z przedstawionego określenia nie wynika wymóg istnienia związków przyczynowo-skutkowych między cechami zastosowanymi do zapisu warunków elementarnych.

Dla potrzeb wnioskowania diagnostycznego celowe jest przekształcenie danej relacji diagnostycznej $\Re$ w $card(J)$-elementowy *zbiór reguł wnioskowania* o identycznej przesłance złożonej, z których każda zawiera inną konkluzję (złożoną z pojedynczego warunku) [Moczulski, 1997b]:

$$\{r_j \mid (\bigwedge_{i\in I}[a_{x_i}(o) \propto_i v(a_{x_i})] \wedge \bigwedge_{k\in K}[a_{y_k}(o) \propto_k v(a_{y_k})]) \overset{T_j(\Rightarrow)}{\Longrightarrow} [a_{s_j}(o) \propto_j v(a_{s_j})],\ j \in J\}, \quad (3.4)$$

gdzie $T_j(\Rightarrow)$ oznacza stopień prawdziwości reguły, który może być reprezentowany np. za pomocą pary $\langle N(r_j), S(r_j)\rangle$ wartości miar konieczności $N(\ )$ i dostateczności $S(\ )$ tej reguły[1] (zob. [Cholewa, Pedrycz, 1987]).

Jeśli przyjąć, że koniunkcja warunków określających wartości cech wejścia obiektu o stanowi rodzaj kontekstu, każdą regułę r_j, $j \in J$ wg (3.4) można zapisać w postaci:

W kontekście (rozumianym jako warunek stosowania reguły)

$$C_X(o) = \bigwedge_{i\in I}[a_{x_i}(o) \propto_i v(a_{x_i})] \quad (3.5)$$

zachodzi implikacja

$$\bigwedge_{k\in K}[a_{y_k}(o) \propto_k v(a_{y_k})] \overset{T_j(\Rightarrow)}{\Longrightarrow} [a_{s_j}(o) \propto_j v(a_{s_j})] \quad (3.6)$$

Przy przekształcaniu relacji diagnostycznych w złożone reguły wnioskowania diagnostycznego należy zwracać szczególną uwagę, by zapisywane reguły wg (3.4) lub reguły wraz z kontekstem wg (3.6) były relewantne (zob. przypis 23 na s. 47). Ponadto zwykle jako $\propto_j$ stosowane są operatory: równości $=$ oraz nierówności $\neq$.

Jeśli przedmiotem rozważań byłyby reguły "dokładne", to z formalnego punktu widzenia dla $card(J) > 1$ zbiór reguł wg (3.4) zawiera $card(J)-1$ reguł sprzecznych w sensie warunku (3.92) — zob. s. 93. Ze względu na niejednoznaczność relacji diagnostycznych (zob. rozdział 2.2, s. 29) zwykle wymagane jest jednak stosowanie *reguł przybliżonych*. W wyniku zastosowania takich reguł otrzymywana jest *diagnoza przybliżona*, będąca zbiorem stwierdzeń przybliżonych (o aktualnym stanie technicznym obiektu) o określonych stopniach prawdziwości. Do wnioskowania za pomocą zbioru reguł o opisanych własnościach wymagane jest ponadto stosowanie odpowiednich procedur.

Należy podkreślić, że reguły diagnostyczne wg (3.6) są regułami złożonymi, których przesłanki są zbudowane z warunków dla wartości cech wyjścia danego obiektu $o \in D(r)$ (a więc zwykle symptomów danych stanów technicznych), zaś konkluzje dotyczą wartości cech tych stanów. Taki sposób reprezentacji reguł ma wiele zalet, gdyż np.:

- mogą one być łatwo zrozumiane przez inne osoby;
- zastosowanie reguł o złożonych przesłankach umożliwia użycie nieskomplikowanego interpretatora reguł [Cholewa, Pedrycz, 1987].

[1] Należy podkreślić, że jeśli (3.4) rozpatruje się formalnie jako zbiór implikacji, wspomniane przekształcenie powoduje, iż implikacje te mogą być spełnione dla większej liczby obiektów niż (3.2). Jeśli jednak rozważać sposób wnioskowania stosowany w systemach doradczych (gdzie nie są używane ścisłe prawa rachunku zdań), przekształcenie relacji diagnostycznej w zbiór reguł wnioskowania nie powoduje opisanego wyżej skutku.

Wspomniane powyżej zalety mają jednak wpływ na znaczną nadmiarowość (redundancję) bazy wiedzy. Przesłanka reguły wg (3.6) często zawiera bardzo liczne warunki. Co więcej, w danej bazie wiedzy mogą być reprezentowane dwie reguły r_1, r_2, których przesłanki różnią się tylko jednym warunkiem (np. badanym w ostatniej kolejności):

$$K = \{k_1, \ldots, k_{n-1}, k_n\} \subset \overline{1..r},\ v_y \neq v_{y'},$$
$$p_{r_1}(o) = [a_{y_{k_1}}(o) \propto_{k_1} v(a_{y_{k_1}})] \wedge \ldots \wedge [a_{y_{k_{n-1}}}(o) \propto_{k_{n-1}} v(a_{y_{k_{n-1}}})] \wedge [a_{y_{k_n}}(o) \propto_{k_n} v_y],$$
$$p_{r_2}(o) = [a_{y_{k_1}}(o) \propto_{k_1} v(a_{y_{k_1}})] \wedge \ldots \wedge [a_{y_{k_{n-1}}}(o) \propto_{k_{n-1}} v(a_{y_{k_{n-1}}})] \wedge [a_{y_{k_n}}(o) \propto_{k_n} v_{y'}],$$

co może być przyczyną zbyt długiego czasu oczekiwania na konkluzje. Wynika stąd potrzeba dalszego przekształcania bazy reguł, powstałej w wyniku zamiany zbioru empirycznych relacji diagnostycznych. Istotą tego przekształcenia jest wprowadzenie konkluzji pośrednich, które następnie będą stosowane jako przesłanki reguł. Tak przekształcone reguły r_k, $k \notin I \cup J$ mogą być zapisywane wówczas jako (por. rozdział 2.5.3):

Niech:

$$p_k(r) = \bigwedge_{j\in J}[val(a_j(o)) \propto_j v(a_j)]. \quad (3.7)$$

Dla każdego obiektu $o \in D(r)$ w kontekście

$$C_r(o) = \bigwedge_{i\in I}[val(a_i(o)) \propto_i v(a_i)] \quad (3.8)$$

zachodzi implikacja

$$p_k(r) \overset{T_k(\Rightarrow)}{\Longrightarrow} [val(a_k(o)) \propto_k v(a_k)], \quad (3.9)$$

gdzie k jest indeksem określającym składniki jedynego stwierdzenia, stanowiącego konkluzję reguły r_k, zaś $p_k(r)$ jest przesłanką reguły. W powyższych zależnościach zastosowano ogólne oznaczenia cech, wartości cech oraz operatorów relacji zgodne ze stosowanymi w rozdziale 2.5.3.

Jeśli reguła (3.9) ma *przesłankę elementarną* (tj. gdy przesłanka reguły jest złożona z pojedynczego warunku), można rozpatrywać *implikację dwustronną* [Cholewa, Pedrycz, 1987]:

Niech:

$$p(r) = [val(a_j(o)) \propto_j v(a_j)],\ q(r) = [val(a_k(o)) \propto_k v(a_k)],\ j \neq k. \quad (3.10)$$

Wtedy dla każdego obiektu $o \in D(r)$ w kontekście

$$C_r(o) = \bigwedge_{i\in I}[val(a_i(o)) \propto_i v(a_i)] \quad (3.11)$$

zachodzi implikacja obustronna

$$p(r) \overset{T(\Rightarrow),T(\Leftarrow)}{\Longleftrightarrow} q(r), \quad (3.12)$$

gdzie: $T(\Rightarrow), T(\Leftarrow)$ są stopniami prawdziwości odpowiednio implikacji "wprost" i implikacji odwrotnej. Przedziały zmienności tych stopni prawdziwości mogą być określane za pomocą wartości miar konieczności $N(\ )$ oraz dostateczności $S(\ )$. Implikacje dwustronne umożliwiają wnioskowanie diagnostyczne z zastosowaniem uogólnionego schematu *modus ponens* [Cholewa, 1993].

Wartości stopni prawdziwości reguł "pierwotnych" wg (3.6) mogą być wstępnie określane przez specjalistę — autora danej reguły. Mogą być także przedmiotem oceny przez innych specjalistów lub oceny dokonywanej za pomocą przykładów testowych. Gdy zgromadzono więcej niż jedną ocenę, należy przeprowadzić *agregację zbioru ocen danej reguły* [Cholewa, 1985], [Cholewa, 1997a]. Wartości stopni prawdziwości reguł przekształconych wg (3.9) określane są przez programistę bazy wiedzy m.in. na podstawie ocen reguł "pierwotnych". W razie zgromadzenia kilku ocen stopnia prawdziwości reguł przekształconych oceny tych reguł są również poddawane agregacji w celu aktualizacji wartości stopnia prawdziwości.

Kontekst $C(r)$ wg (3.8) można w szczególności interpretować jako złożony warunek wstępny, który ma być spełniony, aby przesłanka reguły mogła być badana przez interpretator reguł danego systemu doradczego.

Ponieważ każda reguła r jest w zasadniczy sposób związana z dziedziną $D(r)$, w bazie reguł reprezentowane są zarówno dziedziny, jak i reguły, przy czym w zapisie każdej reguły uwzględnia się odniesienie do odpowiedniej dziedziny.

3.1.2 Pozyskiwanie wiedzy od specjalisty zgodnie z modelem 1.

Istotę zastosowanej metody można przedstawić następująco [Moczulski, 1994]:

1. *Źródłem wiedzy diagnostycznej jest specjalista* w danej dziedzinie zastosowań.
2. Celem postępowania jest *pozyskanie wiedzy operacyjnej*, przydatnej do wnioskowania diagnostycznego.
3. *Wiedza operacyjna* może być przez specjalistę reprezentowana w postaci zbliżonej do *empirycznej relacji diagnostycznej.*
4. *Wiedza podstawowa* (o dziedzinie zastosowań) może być pozyskiwana w sposób pośredni (przy okazji zapisu empirycznych relacji diagnostycznych).
5. Przyjęta metoda winna umożliwić pozyskanie szczególnie cennej części wiedzy diagnostycznej specjalisty, obejmującej:
   (a) *relacje*, zidentyfikowane przez specjalistę na podstawie jego obserwacji dotyczących konkretnych obiektów,
   (b) *wartości dopuszczalne i krytyczne cech wyjścia* (w postaci ilościowej), umożliwiające przyporządkowanie obiektów konkretnych (na podstawie symptomów diagnostycznych, zaobserwowanych dla tych obiektów) do odpowiednich klas stanów technicznych.
6. Samodzielne działanie specjalisty w dziedzinie zastosowań *jest wspomagane prostym środkiem* (jego stosowanie nie wymaga żadnego przeszkolenia), który jednocześnie ułatwia specjaliście uświadomienie sobie relacji diagnostycznych, zachodzących dla konkretnych obiektów w danej dziedzinie.

W opisywanej metodzie postępowania znaczną rolę spełnia programista bazy wiedzy. Występuje on w procesie pozyskiwania wiedzy jako niezbędny pośrednik — interpretator zapisów specjalisty, dokonując:

1. *identyfikacji elementów opisu dziedziny*, jak: obiekty, klasy obiektów, cechy obiektów, jakościowe wartości tych cech itp., zawartych bezpośrednio w zapisach specjalisty lub wymaganych do zapisu reguł, w które zostaną przekształcone relacje empiryczne;
2. *zintegrowania zidentyfikowanych elementów opisu dziedziny z już pozyskaną wiedzą podstawową o dziedzinie zastosowań* (identyfikacja synonimów, uzupełnienie słowników itp.);
3. *identyfikacji wartości cech, stosowanych przez specjalistę do rozgraniczenia klas symptomów oraz klas stanów*, a następnie — *określenia sposobów konwersji wartości cech ilościowych w wartości jakościowe*;
4. *niezbędnej interpretacji treści zapisów*, będących nośnikami informacji o relacjach empirycznych oraz ich ewentualnego przekształcenia w reguły;
5. *dołączenia pozyskanej wiedzy operacyjnej* (w postaci reguł wnioskowania diagnostycznego) *do istniejącej bazy reguł* oraz *przypisania tym regułom wartości stopni prawdziwości lub stopni przekonania* i ewentualnej aktualizacji (lub agregacji) wartości stopni prawdziwości i przekonania w odniesieniu do reguł, które już występują w bazie.

W razie udziału w procesie pozyskiwania wiedzy więcej niż jednego specjalisty (co jest wskazane) programista bazy wiedzy ma za zadanie: określenie wspólnych słowników pojęć (oraz słowników synonimów), zinterpretowanie zapisów wszystkich specjalistów za pomocą tych pojęć oraz przypisanie otrzymanym regułom wartości stopni ich prawdziwości lub przekonania o ich słuszności i dokonanie agregacji tych stopni w razie wykrycia identycznych reguł.

Do zapisu posiadanej przez siebie wiedzy specjalista stosuje specjalny formularz (opis w rozdziale 4.2.1) [Moczulski, 1997b]. Odpowiedni układ formularza winien pomóc specjaliście w ujawnieniu znanych mu relacji diagnostycznych, zachodzących dla konkretnych obiektów w danej dziedzinie. Posługiwanie się tym formularzem nie wymaga od specjalisty żadnego uprzedniego przeszkolenia.

Oprócz elementów składowych relacji empirycznej, specjalista może umieszczać uzasadnienia oraz cytowania i referencje. Mają one znaczenie przy tworzeniu systemu objaśnień.

Ponieważ zapisy specjalisty zostają poddane interpretacji przez dodatkową osobę (programista bazy wiedzy), konieczne jest zapewnienie sprzężenia zwrotnego, którego istotą jest ocena "pierwotnych" reguł wnioskowania diagnostycznego przez specjalistę, a następnie wprowadzenie ewentualnych zmian przez programistę.

3.1.3 Pozyskiwanie wiedzy od specjalisty zgodnie z modelem 2.

Przedstawiony model w znacznej mierze odpowiada modelowi 2., opisanemu w [Buchanan *et al.*, 1983]. Zasadniczym celem przyjętego sposobu postępowania jest jak największe ograniczenie (jedynie do niezbędnych zadań) zakresu udziału pośrednika, czyli programisty bazy wiedzy, w procesie pozyskiwania wiedzy. Aby ten cel osiągnąć, specjalista uzyskuje dostęp do swoistego edytora bazy wiedzy, zwanego *formularzem elektronicznym* [Moczulski, 1996b] (opis w rozdziale 4.2.3).

Zarys koncepcji formularza "elektronicznego"

Formularz elektroniczny może wspomagać specjalistę w zakresie pozyskiwania wiedzy, reprezentowanej w sposób deklaratywny, obejmującej:

- *wiedzę podstawową o dziedzinie zastosowań*, czyli:
  - *pojęcia*, przy czym celem postępowania jest pozyskanie *opisów obiektów*, *klas obiektów*, *nazw cech* (określonych dla klas) oraz *nazw wartości* tych cech (dla cech jakościowych);
  - *taksonomie obiektów*;
  - *słowniki synonimów* ze wskazaniem podstawowych znaczeń (które będą następnie stosowane do zapisu bazy wiedzy);
  - *objaśnienia* (w odniesieniu do pojęć);
- *wiedzę operacyjną*, czyli:
  - *reguły wnioskowania*;
  - *objaśnienia* (w odniesieniu do reguł).

Scenariusz sesji pozyskiwania wiedzy od specjalisty

Na podstawie dostępnych opisów i spostrzeżeń własnych można stwierdzić, że w procedurach wnioskowania (rozumowania) systemów diagnostycznych stosuje się zwykle *kroczenie wstecz* (które jest typowe dla zadań klasyfikacji). W ramach opisywanych badań postanowiono zastosować postępowanie wstecz także w procesie pozyskiwania wiedzy od specjalistów. Oznacza to, że wpierw następuje pozyskanie końcowych konkluzji, tzn. diagnoz, a w ostatniej fazie — pozyskanie warunków wejściowych, które winny być zbadane we wstępnym stadium konsultacji.

Zdaniem autora, proponowane rozwiązanie posiada wiele zalet, między innymi:

1. Wymaga się wprowadzenia *minimalnego zbioru pojęć* (obiekty, cechy obiektów, wartości cech, relacje zachodzące pomiędzy obiektami itp.) koniecznego do określenia warunków, będących składnikami przesłanek oraz konkluzji.
2. Nie jest konieczne przewidywanie przez specjalistę, jakie pojęcia będą mu potrzebne do zdefiniowania planowanych przez niego reguł.
3. Przebieg sesji pozyskiwania wiedzy jest podobny do przebiegu konsultacji z systemem doradczym.
4. Przebieg tej sesji odpowiada także sposobowi postępowania specjalisty z dziedziny diagnostyki technicznej, gdy próbuje on postawić diagnozę dotyczącą stanu technicznego danego obiektu.

Przyjęty sposób postępowania (tj. *postępowanie wstecz*) implikuje następujący algorytm pozyskiwania reguły wnioskowania:

1. *Określenie dziedziny reguły* z użyciem elementów odpowiednich słowników nazw obiektów i nazw klas obiektów (z możliwością uzupełnienia obu tych słowników).

2. *Określenie konkluzji reguły* z użyciem elementów odpowiednich słowników nazw cech, ich wartości jakościowych oraz operatorów relacji (z możliwością uzupełnienia słowników nazw cech i słowników wartości cech).
3. *Określenie warunku prostego przesłanki reguły* (zob. rozdział 2.5.3) z użyciem elementów odpowiednich słowników nazw cech, ich wartości jakościowych oraz operatorów relacji (z możliwością uzupełnienia słowników nazw cech i ich wartości).
4. Jeśli przesłanka reguły ma zawierać dalsze warunki proste, to powrót do punktu 3.
5. *Określenie stopnia prawdziwości* pozyskanej reguły przez jej autora.

Rola programisty bazy wiedzy jest w tym wypadku znacznie mniejsza niż w wypadku realizacji metody, opisanej w rozdziale 3.1.2. Jego zadania obejmują:

1. zintegrowanie zidentyfikowanych elementów opisu dziedziny z już pozyskaną wiedzą podstawową o dziedzinie (identyfikacja synonimów, uzupełnienie słowników itp.);
2. określenie sposobów konwersji wartości cech ilościowych w wartości cech jakościowych;
3. dołączenie pozyskanej wiedzy "operacyjnej" (w postaci reguł wnioskowania diagnostycznego) do istniejącej bazy reguł oraz przypisanie tym regułom wartości stopni prawdziwości i ewentualnie aktualizacji (lub agregacji) wartości stopni prawdziwości reguł, które już występują w bazie.

3.2 Pozyskiwanie wiedzy w sposób indukcyjny na podstawie przykładów

Bezpośrednie pozyskiwanie wiedzy od specjalistów w dziedzinie zastosowań jest mało efektywne, gdyż można w ten sposób otrzymać jedynie dziesiątki reguł w ciągu jednej sesji pozyskiwania wiedzy. Również dalsza rozbudowa bazy wiedzy i dołączanie nowych "porcji" wymagają wiele czasu. W odniesieniu do różnych dziedzin praktycznych nie można w związku z tym poprzestać na pozyskiwaniu wiedzy bezpośrednio od specjalistów. Pozytywne wyniki nielicznych badań dotyczących pozyskiwania wiedzy z zakresu diagnostyki technicznej maszyn, przeprowadzonych w niektórych ośrodkach (zob. np. [Giordana *et al.*, 1993], [Nowicki, Słowiński, Stefanowski, 1992]), oraz wstępne wyniki badań własnych skłoniły autora do podjęcia wraz z zespołem intensywnych prac, związanych z zastosowaniem *indukcyjnych metod tzw. uczenia maszynowego* do pozyskiwania wiedzy diagnostycznej. Prace te koncentrują się na:

- *indukcji reguł wnioskowania*,
- *indukcji drzew decyzyjnych*.

Przy wyborze poszczególnych metod kierowano się odpowiednim zestawem kryteriów, obejmującym:

- wymagania, dotyczące zbioru danych "uczących" (możliwość występowania błędów, przykładów sprzecznych, brakujących wartości cech itp.);

- skuteczność wnioskowania diagnostycznego z zastosowaniem bazy wiedzy, pozyskanej w sposób zgodny z daną metodą;
- złożoność obliczeniową;
- stopień złożoności oraz łatwość interpretacji przez użytkownika pozyskanej "porcji" wiedzy.

Istotą zastosowanych metod jest indukcja "porcji" wiedzy diagnostycznej na podstawie przykładów zgromadzonych w odpowiednio przygotowanych bazach danych. Przykłady te odpowiadają poszczególnym koncepcjom (pojęciom), którymi w opisywanych zastosowaniach są stany techniczne (lub klasy takich stanów).

3.2.1 System informacyjny i tablica decyzyjna

Dla potrzeb opisu indukcyjnych metod pozyskiwania wiedzy wprowadza się pojęcie *systemu informacyjnego* [Pawlak, 1991] jako pary $\mathcal{S} = \langle E, A \rangle$, gdzie E jest niepustym i skończonym zbiorem pewnych obiektów e, zwanym *uniwersum*, zaś A jest niepustym i skończonym zbiorem cech tych obiektów o własności:

$$(\forall a \in A)\,[a : E \to V(a)] \tag{3.13}$$

przy czym $V(a)$ jest *dziedziną* cechy a. Każdy system informacyjny $\mathcal{S} = \langle E, A \rangle$ oraz niepusty podzbiór $B \subseteq A$ określają *funkcję B-informacyjną* [Skowron, 1993]:

$$Inf_B : E \to B \times \bigcup_{a \in B} V(a) \tag{3.14}$$

$$Inf_B(e) = \{(a, a(e)) \mid a \in B\}, \tag{3.15}$$

która każdemu obiektowi $e \in B$ przyporządkowuje zbiór wartości jego cech.

Specjalnym przypadkiem systemu informacyjnego, stosowanym w niniejszej pracy dla potrzeb pozyskiwania wiedzy metodami "uczenia maszynowego", jest *tablica decyzyjna* (zob. [Nowicki, Słowiński, Stefanowski, 1992], [Skowron, 1993]). Jest nią dowolny system informacyjny $\mathcal{S} = \langle E, A \cup D \rangle$, przy czym zbiór D spełniający warunek $D \cap A = \emptyset$ jest zbiorem specjalnie wyróżnionych cech, zwanych *cechami decyzyjnymi*. Elementy zbioru cech A nazywane są wówczas *cechami warunku*.

3.2.2 Zbiór przykładów

Aby można było pozyskać w sposób indukcyjny wiedzę na podstawie pewnego zbioru przykładów E, wszystkie te przykłady muszą dotyczyć takiego zbioru obiektów $O_E \subseteq O$, dla których można utworzyć wspólną dziedzinę D_E, rozumianą w sposób opisany np. w komentarzu do zależności (2.77). Tak więc wraz ze zbiorem przykładów określona jest dziedzina "porcji" wiedzy, która może być wyprowadzona z tych przykładów w sposób indukcyjny.

W praktycznych zastosowaniach, zgodnie z opisem zawartym w rozdziale 2.5.2, wszystkie przykłady są reprezentowane za pomocą stwierdzeń, które dotyczą wartości zarówno cech warunku, jak i cech stanu. Zapis stwierdzeń następuje w "domkniętym świecie" (zob. rozdział 2.7) z zastosowaniem skończonego zbioru cech oraz ich dyskretnych wartości (określonych w ramach tzw. wiedzy podstawowej o dziedzinie zastosowania, pozyskanej bezpośrednio od specjalistów). Wymagany opis dziedziny obejmuje (por. rozdział 2.5.1):

1. *Określenie zbioru cech warunku* (preferowane są cechy jakościowe):

$$A = \{a_j \mid j = 1, \ldots, J\} \tag{3.16}$$

   w zakresie:

   (a) *nazw cech*, stosowanych do reprezentacji przykładów,

   (b) *typów cech* (opis na s. 12);

2. Określenie *dziedziny* dla każdej cechy a_j, $j = 1, \ldots, J$:

$$V(a_j) = Dom(a_j) = \{v_{j,m} \mid m = 1, \ldots, M_j\} \tag{3.17}$$

   (jako zbioru wartości jakościowych — określonych w postaci nazw) tych cech oraz dla cech typów `lin` i **cyc** — określenie *uporządkowania zbioru wartości*;

3. Określenie *zbioru rozróżnianych stanów* (których identyfikatory będą stanowiły wartości cech decyzyjnych):

$$S = \{s_k \mid k = 1, \ldots, K\} \tag{3.18}$$

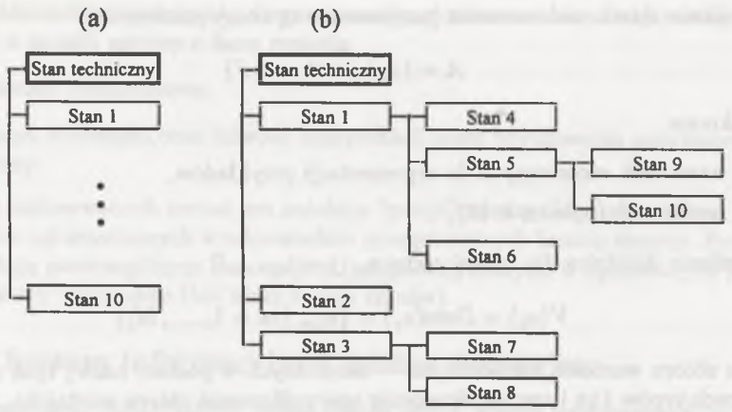
   (poprzez podanie ich nazw) oraz ewentualnie *struktury zbioru stanów technicznych* (która może być reprezentowana za pomocą *drzewa stanów* — zob. poniżej).

Przy określaniu sposobu reprezentacji przykładów należy także zdecydować, w jaki sposób będą reprezentowane brakujące wartości cech. Najczęściej przyjmuje się, że w miejsce brakującej wartości cechy a może być wstawiona jakakolwiek wartość $v \in Dom(a)$.

Struktura zbioru stanów technicznych

Obiekty diagnozowania zwykle składają się z wielu elementów (prostych i złożonych). Można więc wyróżnić wiele stanów elementarnych, charakterystycznych dla poszczególnych elementów. W pracy przyjęto (zob. rozdział 2.2), że nośnikiem informacji o stanie obiektu są sygnały, za pomocą których są obserwowane procesy resztkowe, będące skutkami stanów poszczególnych elementów obiektu złożonego oraz warunków działania obiektu. Dla diagnostyki maszyn charakterystyczne są następujące spostrzeżenia:

1. Zwykle brak *wzajemnie jednoznacznych związków* pomiędzy występowaniem pewnych stanów technicznych (określonych ilościowo lub jakościowo za pomocą cech stanu) a zaobserwowaniem określonych symptomów (wartości cech wyjść badanego obiektu).
2. W odniesieniu do skutków jednoczesnego występowania stanów zwykle nie obowiązuje *zasada superpozycji*, natomiast stwierdzono występowanie *efektu synergii*, zarówno dodatniej (tj. wzmocnienia skutków przy jednoczesnym występowaniu więcej niż jednej przyczyny), jak i ujemnej.
3. W jednym obiekcie może występować jednocześnie wiele niesprawności, mogących uniemożliwić uszczegółowioną diagnozę stanu (zob. punkt 3 na s. 30).



Rys. 3.1. Drzewa stanów dla prostej (a) i złożonej (b) struktury zbioru stanów technicznych

Fig. 3.1. State trees for simple (a) and complex (b) structure of the set of technical states

Znanych jest wiele przykładów, dotyczących spostrzeżenia 1. Np. występowanie w widmie gęstości mocy drgań względnych wirnika istotnej składowej o częstotliwości obrotów tego wirnika może być wynikiem nadmiernego jego niewyrównoważenia lub niewspółosiowości wewnętrznej (w obrębie jednego korpusu maszyny). Należy podkreślić, że opisana niejednoznaczność jest jedną z przyczyn uznawania wiedzy diagnostycznej za wiedzę przybliżoną.

Spostrzeżenie 3. dotyczy problemu braku możliwości uzyskania uszczegółowionej diagnozy w jednym kroku wnioskowania diagnostycznego. W takim wypadku diagnozowanie winno przebiegać sekwencyjnie, np. wpierw należy zgrubnie zlokalizować uszkodzenie (co może być interpretowane jako określenie zbioru możliwych stanów, które mogą być przyczyną obserwowanych objawów), a następnie dążyć do uszczegółowienia diagnozy. Warto podkreślić, że przy takim sposobie postępowania w różnych stadiach wnioskowania można stosować różne zbiory wartości cech (zob. [Giordana *et al.*, 1993]).

Do reprezentowania związków pomiędzy poszczególnymi stanami (w zbiorze stanów) autor zaproponował *drzewo stanów* (rys. 3.1). Koncepcja ta została wstępnie zweryfikowana w ramach badań prowadzonych w zespole autora [Moczulski, Kostka, 1997].

Korzeniem drzewa stanów jest "stan techniczny". Dla *prostej struktury stanów* drzewo posiada tylko krawędzie łączące korzeń z liśćmi (rys. 3.1 — drzewo "a"). W wypadku *złożonej struktury stanów* drzewo zawiera węzły nieterminalne (które odpowiadają stanom złożonym - zob. rys. 3.1 — drzewo "b"). Stany obiektu badań występują w drzewie w kolejności od stanów ogólnych (tj. złożonych) do stanów szczegółowych. Struktura drzewa stanów winna być przedmiotem optymalizacji. Jednym z kryteriów tej optymalizacji może być maksimum sprawności otrzymanej rodziny klasyfikatorów stanu technicznego.

Drzewo stanów stanowi środek reprezentacji wiedzy diagnostycznej dla danej klasy obiektów. Wiedza o związkach zachodzących pomiędzy stanami technicznymi danego obiektu może być pozyskana od specjalisty lub zidentyfikowana na podstawie przykładów.

W drugim wypadku postępowanie może polegać na grupowaniu przykładów ze względu na podobieństwo wartości cech, a następnie na przedstawieniu wyników grupowania specjaliście w celu przypisania (lub określenia) odpowiednich stanów dla poszczególnych grup przykładów.

Reprezentacja zbioru przykładów

Zbiór przykładów stosowany do pozyskiwania wiedzy w sposób indukcyjny może być zapisany w postaci:

$$E = \{e_n \mid e_n = (\underline{v}_n, d_n), n = 1, \ldots, N\}, \tag{3.19}$$

gdzie $\underline{v}_n$ jest j-elementową jednokolumnową macierzą wartości cech:

$$\underline{v}_n = [val(a_1), \ldots, val(a_J)]^T, \; n = 1, \ldots, N, \tag{3.20}$$

zaś przyporządkowania tego przykładu do wyróżnionych klas stanu są reprezentowane za pomocą listy d_n o skończonej liczbie elementów (identyfikatorów stanów technicznych):

$$d_n = (i_1, i_2, \ldots, i_k), \; i_1, \ldots, i_k \in \overline{1..K}, \; k < K, \; n = 1, \ldots, N, \tag{3.21}$$

gdzie i_k, $k = 1, \ldots, K$ są identyfikatorami wierzchołków w drzewie stanów, przy czym lista wg (3.21) nie zawiera powtórzeń i jest uporządkowana w taki sposób, że dla jej dwóch kolejnych elementów i_l, i_{l+1}, $l < k$ węzeł odpowiadający drugiemu z nich jest *potomkiem* (tj. węzłem podrzędnym) węzła odpowiadającego pierwszemu elementowi. Wprowadzenie listy identyfikatorów stanów pozwala na sekwencyjną klasyfikację (o rosnącym stopniu szczegółowości diagnoz), niejednokrotnie wymaganą w diagnostyce technicznej.

Listy identyfikatorów wierzchołków drzewa stanów umożliwiają wielokrotny podział zbioru przykładów na podzbiory, odpowiadające przykładom rozróżnianych pojęć. Niech dla dowolnego stanu $s_0 \in S$

$$E(s_0) = \{(\underline{v}_n, d_n) \in E \mid head(d_n) = id(s_0)\} \tag{3.22}$$

będzie zbiorem (o liczności $N = card(E(s_0))$) przykładów pojęcia s_0, przy czym $head(d_n)$ jest funkcją zwracającą głowę listy d_n, zaś $id(s)$ funkcją zwracającą identyfikator stanu s. Niech $s_1, s_2 \in S$ będą dwoma stanami, odpowiadającymi węzłom-potomkom węzła dla s_0:

$$s_0 \prec s_1, s_0 \prec s_2, \neg[(s_1 \prec s_2) \vee (s_2 \prec s_1)], \tag{3.23}$$

gdzie $\prec$ oznacza *operator poprzedzania* (w grafie — drzewie stanów), zaś ostatni zapis jest równoważny warunkowi, że s_1, s_2 są nieporównywalne. Jeśli ponadto jest spełniony warunek:

$$(\forall s \in S)\,[(s_0 \prec s \prec s_1 \vee s_0 \prec s \prec s_2) \Rightarrow s = s_0] \tag{3.24}$$

czyli s_1, s_2 odpowiadają węzłom, które są *bezpośrednimi potomkami* węzła przypisanego do s_0, wówczas można dokonać podziału zbioru przykładów, tworząc dwa podzbiory:

$$E(s_1) = \{(\underline{v}_n, d'_n) \mid d'_n = tail(d_n) \wedge head(d'_n) = id(s_1), n = 1, \ldots, N_1\} \tag{3.25}$$

$$E(s_2) = \{(\underline{v}_n, d'_n) \mid d'_n = tail(d_n) \wedge head(d'_n) = id(s_2), n = 1, \ldots, N_2\} \tag{3.26}$$

gdzie $N = N_1 + N_2$ oraz $tail(d_n)$ jest funkcją, która zwraca *ogon listy* d_n (tj. listę bez jej pierwszego elementu).

W procesie pozyskiwania wiedzy danymi uczącymi mogą być bądź to wyniki pomiarów wartości cech sygnałów diagnostycznych (jak np. w pracy [Moczulski, 1996a]) i cech tzw. parametrów wejściowych (procesowych), bądź też wyniki tzw. obliczeń symulacyjnych (zob. [Moczulski, Maniak, 1997]).

Przygotowanie odpowiedniego zbioru przykładów jest zadaniem bardzo ważnym, które w zasadniczy sposób warunkuje poprawność bazy wiedzy pozyskiwanej w sposób indukcyjny. Należy zwrócić uwagę na następujące problemy:

1. *Reprezentatywność zbioru przykładów* dla pozyskiwanej "porcji" wiedzy, która winna być osiągnięta za pomocą jedynie nielicznych przykładów spośród liczby możliwej ze względów kombinatorycznych.

2. *Wybór właściwego podzbioru cech* (tzw. *cechy relewantne* [Michalski, 1997]) dla każdego podzbioru przykładów odpowiadającego danym pojęciom [Moczulski, Kostka, 1997].

3. Przeprowadzenie *odpowiedniej kwantyzacji* cech ilościowych.

W zastosowaniach praktycznych, związanych z diagnostyką maszyn, zwykle dostępne są bazy danych zawierające od kilkuset do kilkudziesięciu tysięcy przykładów. Np. baza danych, opisana w pracy [Moczulski, Maniak, 1997], liczyła $N = 5076$ przykładów, co stanowi jedynie 0,75 procent wszystkich możliwych przykładów (tzn. możliwych kombinacji wartości cech). Zapewnienie reprezentatywności zbioru przykładów, postulowane w punkcie 1., wymaga odpowiedniego zaplanowania eksperymentu diagnostycznego (biernego lub czynnego) lub właściwego zaplanowania obliczeń symulacyjnych (zob. [Moczulski, Maniak, 1997]).

Wybór cech relewantnych (punkt 2.) związany jest ze zbiorem stanów, które winny być rozpoznane na podstawie danego podzbioru przykładów. Zwykle bowiem zbiór cech obserwowanych dla badanych obiektów zawiera cechy zbyt elementarne w porównaniu z cechami złożonymi, stosowanymi przez specjalistów (jak np. jednoczesne wystąpienie grupy cech o odpowiednich wartościach) [Giordana *et al.*, 1993]. Autor sformułował *hipotezę o potrzebie wyboru różnych podzbiorów cech relewantnych dla celów klasyfikacji pewnego podzbioru przykładów, odpowiadającego poddrzewu drzewa stanów*. Została ona wstępnie zweryfikowana w badaniach prowadzonych w zespole autora [Moczulski, Kostka, 1997]. Sposoby wyboru cech relewantnych opisano w dalszej części rozdziału.

Problem, sygnalizowany w punkcie 3., został już opisany w rozdziale 2.1.2. Zdaniem autora, dla podzbiorów cech relewantnych, odpowiadających poddrzewom drzewa stanów, w niektórych wypadkach może być celowe:

- bądź *grupowanie wartości jakościowych* w celu ograniczenia liczby rozróżnianych stanów (podobne postępowanie ma miejsce w odniesieniu do tablic kontyngencji — zob. opis na s. 56),

- bądź *zastosowanie innych sposobów kwantyzacji* zależnych od własności statystycznych zbioru wartości danej cechy ilościowej.

Druga z wymienionych hipotez nie została dotychczas dostatecznie zweryfikowana.

Jak wspomniano, listy d_n identyfikatorów stanów technicznych, będące następnikiem pary $e_n = (\underline{v}_n, d_n)$ stanowiącej zapis przykładu, określają kolejne przyporządkowania tego

przykładu do coraz bardziej szczegółowych koncepcji. Umożliwia to dekompozycję zadania pozyskiwania wiedzy na podzadania, w których dla mniej licznych zbiorów przykładów dokonywana jest (w identyczny sposób) indukcja mniejszej "porcji" wiedzy, dotyczącej poddziedziny rozpatrywanej dziedziny zastosowań. Z tego względu w dalszej części rozdziału, zawierającej opisy indukcyjnych metod pozyskiwania wiedzy, przyjęto, że listy d_n będą zawierały jeden element umożliwiający zidentyfikowanie pojęcia, które ilustruje dany przykład. Element ten można rozpatrywać jako wartość cechy decyzyjnej, która w omawianym przypadku jest identyfikatorem stanu technicznego.

3.2.3 Stosowanie pozyskanej wiedzy do klasyfikowania nowych przykładów

Jak wspomniano powyżej, zbiór przykładów uczących stosowany jest do indukcji pewnej "porcji" wiedzy, która jest następnie używana do wnioskowania diagnostycznego, polegającego na *klasyfikowaniu nowych danych*, opisujących własności i właściwości badanych obiektów. Dane te zwane będą *przykładami*. Są one "nowe" w takim sensie, że nie służyły do indukcji wiedzy stosowanej w procesie wnioskowania. Dalszy opis dotyczy przykładów, dla których znana jest wartość cechy decyzyjnej, określająca ich przyporządkowanie do jakiejś klasy. Rozpatrywany jest przypadek klasyfikacji "1 z N".

Ogólnie możliwe są dwa rodzaje błędów klasyfikacji:

- *Przykład nie jest klasyfikowany do właściwej klasy, którą określa wartość cechy decyzyjnej.* Błąd ten zwany jest *błędem pominięcia* (ang. *omission error*).

- *Przykład klasyfikowany jest do niewłaściwej klasy, innej niż określona przez wartość cechy decyzyjnej.* Błąd ten jest nazywany *błędem niesłusznego zaliczenia* (ang. *commission error*).

Analizując różne możliwości wystąpienia tych błędów, wyróżniono następujące przypadki [Grzymała-Busse, 1994]:

1. przykład jest klasyfikowany poprawnie (do klasy, do której faktycznie należy) i nie jest zaklasyfikowany (błędnie) do żadnej innej klasy;

2. przykład jest klasyfikowany poprawnie, lecz jednocześnie jest klasyfikowany (błędnie) do innej klasy;

3. przykład nie jest klasyfikowany poprawnie, lecz jest zaklasyfikowany (błędnie) do innej klasy;

4. przykład nie jest zaklasyfikowany do żadnej klasy.

W diagnostyce technicznej szczególne znaczenie ma klasyfikacja jednoznaczna (jak w p. 1).

Klasyfikacja jak w p. 2 może być uznana za *niejednoznaczną*. Jej przyczyną zwykle są luki w "porcji" wiedzy zidentyfikowanej w sposób indukcyjny (np. nadmierne uogólnienia, bez uwzględnienia przypadków szczególnych). Należy pamiętać, że w fazie stosowania wiedzy mamy do czynienia z przykładami (pochodzącymi np. z bieżących obserwacji obiektu), które nie zostały wstępnie sklasyfikowane przez eksperta. Wówczas, zdaniem autora, także niejednoznaczny wynik klasyfikacji jest wartościowy, gdyż jego uzyskanie może np. spowodować podjęcie decyzji o zgromadzeniu przez system klasyfikujący dodatkowych danych

(których uzyskanie może być np. bardziej kosztowne) w celu rozróżnienia klasy (spośród kilku, otrzymanych w wyniku klasyfikacji niejednoznacznej), do której obiekt należy.

Przedstawiona analiza stanowić będzie w dalszej części pracy podstawę do określenia ocen sprawności klasyfikacji oraz pewności pojedynczych stwierdzeń dotyczących przypisania danego przykładu do klasy stanu.

3.2.4 Selektywna indukcja reguł poprzez generowanie pokryć

Pozyskiwanie reguł dotyczy wyróżnionego zbioru stanów S wg (3.18). Zbiór przykładów E (zwanych *przykładami uczącymi*) dzielony jest na podzbiory:

$$E = E_1 \cup E_2 \cup \ldots \cup E_K, \quad \text{gdzie} \quad E_k = E(s_k),\ k = 1, \ldots, K, \tag{3.27}$$

przy czym dopuszcza się występowanie przykładów niejednoznacznych, tj. nie spełniających poniższego warunku:

$$(\forall i, j \in \overline{1..N})\ [e_i = (\underline{v}_i, d_i) \wedge e_j = (\underline{v}_j, d_j) \wedge \underline{v}_i = \underline{v}_j \Rightarrow d_i = d_j], \tag{3.28}$$

co może mieć miejsce zwłaszcza w wypadku gromadzenia zbioru przykładów z różnych źródeł (np. różni specjaliści) lub zgrubnej kwantyzacji wartości cech ilościowych. Warunek (3.28) jest jednym z możliwych kryteriów zgodności tablicy decyzyjnej (inaczej sformułowany warunek zgodności zawiera rozdział 3.2.5).

Niech $E = E^+ \cup E^-$ dla danego pojęcia s_{k0}, $k_0 \in \overline{1..K}$ oznacza podział zbioru E na podzbiór przykładów "pozytywnych" $E^+ = E_{k0}$ oraz jego kontrprzykładów $E^- = E \backslash E_{k0}$ (tj. pozostałych przykładów). Istotą metody (zob. np. [Michalski, 1983]) jest *generowanie pokryć* $\Pi(E^+ \mid E^-)$ zbioru E^+ przykładów danego pojęcia z wyłączeniem kontrprzykładów E^-. Pokrycie zapisywane jest w postaci alternatywy (sumy logicznej) warunków prostych wg (2.76) w zapisie zgodnym z VL_1.

Niech

$$R = \{R_k \mid k = 1, \ldots, K\} \tag{3.29}$$

oznacza zbiór hipotez wstępnych, z których każde R_k oznacza (być może pusty) zbiór "reguł", będących koniunkcjami przesłanek $p(o)$ wg (2.76) oraz (domyślnego) wyrażenia (będącego odpowiednikiem konkluzji), określającego klasę stanu:

$$R_k = \{r_i(o) \mid r_i(o) = p_i(o) \wedge q(o),\ i = 1, \ldots, n_k\}, \tag{3.30}$$

gdzie

$$p_i(o) = \bigwedge_{\substack{l=1,\ldots,m_i \\ j_1,\ldots,j_{m_i} \in \overline{1..J}}} [val(a_{j_l}(o)) \propto_l Term_l(a_{j_l})] \tag{3.31}$$

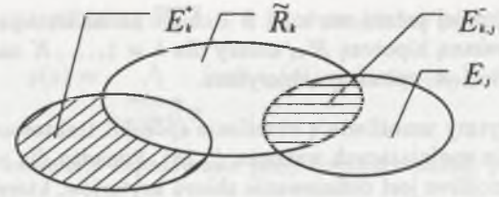
oraz

$$q(o) = [class(o) = id(s_k)] \tag{3.32}$$

gdzie $class(o)$ jest funkcją zwracającą identyfikator klasy, do której należy obiekt o (w naszym wypadku — identyfikator stanu s_k)[2].

Niepusty zbiór R hipotez wstępnych wg (3.29) umożliwia tzw. *uczenie przyrostowe*, (zob. [Hong, Mozetic, Michalski, 1986], [Wnęk *et al.*, 1995], [Moczulski, 1997a]), które

[2]Zastosowano zapis $p(o) \wedge q(o)$ zamiast $p(o) \Rightarrow q(o)$, ponieważ w praktyce można zaobserwować koniunkcje, a nie implikacje [Dietterich, 1986].



Rys. 3.2. Zależności pomiędzy zbiorami przykładów E_j, E^+_k oraz $E^-_{k,j}$ [Cholewa, Pedrycz, 1987]

Fig. 3.2. Dependences between sets of examples E_j, E^+_k and $E^-_{k,j}$ [Cholewa, Pedrycz, 1987]

może być stosowane bądź do "douczania" klasyfikatora, bądź przy rozwiązywaniu zadań klasyfikacji obiektów, których cechy wolno ewoluują w czasie "makro".

Celem procesu uczenia jest sformułowanie nowych hipotez

$$R'_k,\ k = 1, \ldots, K\,, \tag{3.33}$$

poprawnie opisujących zbiór przykładów E_k oraz wykluczających przykłady z pozostałych podzbiorów E_j, $j \neq k$.

Algorytm indukcji reguł

Ogólny algorytm A^Q indukcji reguł (R. S. Michalski, [Michalski, 1969]) można przedstawić następująco (wg [Cholewa, Pedrycz, 1987]):

1. Dla każdej hipotezy R_k są tworzone zbiory wyjątków E^+_k (przykładów błędnie pominiętych) oraz $E^-_{k,j}$ (przykładów niesłusznie zaliczonych):

$$E^+_k = E_k \setminus \tilde{R}_k \tag{3.34}$$

$$E^-_{k,j} = \tilde{R}_k \cap E_j,\ j = 1, \ldots, K \wedge j \neq k, \tag{3.35}$$

gdzie $\tilde{R}_k$ oznacza zbiór wszystkich takich przykładów $e \in E$, dla których jest spełniona hipoteza R_k (zob. rys. 3.2).

2. Dla każdego $k = 1, \ldots, K$ tworzona jest hipoteza R^-_k, opisująca wszystkie wyjątki:

$$R^-_k = \bigvee_{\substack{j=1,\ldots,K \\ j \neq k}} R^-_{k,j}, \tag{3.36}$$

gdzie

$$R^-_{k,j} = \Pi\Big(E^-_{k,j} \mid \bigcup_{i=1,\ldots,K} (\tilde{R}_i \cup E^+_i)\Big). \tag{3.37}$$

3. Budowane są nowe hipotezy R'_k, wyznaczone jako pokrycia:

$$R'_k = \Pi\Big(E_k \mid \bigcup_{\substack{i=1,\ldots,K \\ i \neq k}} (\tilde{R}_i \setminus R^-_i \cup E^+_i)\Big). \tag{3.38}$$

4. Jeśli dla co najmniej jednej wartości $k \in \overline{1..K}$ nadal istnieją przykłady $e \in E_k$ niezgodne z uzyskaną hipotezą R'_k, należy dla $k = 1, \ldots, K$ zastąpić R_k przez R'_k i powtórzyć kroki 1–4. opisanego algorytmu.

Szczegółowe algorytmy umożliwiają określenie sposobu traktowania przykładów niejednoznacznych, tj. nie spełniających warunku (3.28). Ponadto dla potrzeb oceny wygenerowanych hipotez możliwe jest definiowanie zbioru kryteriów, które są zadawane w postaci specjalnego funkcjonału leksykograficznego o nazwie *LEF* (ang. *lexicographic functional* [Wnęk *et al.*, 1995]). Spośród różnych kryteriów elementarnych można przykładowo wymienić takie, jak [Michalski, 1997]:

- maksymalizacja liczby przykładów pozytywnych, pokrytych w kolejnym kroku iteracyjnym,
- maksymalizacja ogólnej liczby pokrytych przykładów pozytywnych,
- maksymalizacja stosunku ogólnej liczby pokrytych przykładów pozytywnych do ogólnej liczby pokrytych kontrprzykładów,
- minimalizacja liczby warunków złożonych,
- maksymalizacja liczby warunków złożonych.

Możliwe jest również wpływanie na stopień ogólności reguł, w wyniku czego mogą być uzyskiwane reguły bardzo ogólne, reguły o maksymalnym stopniu szczegółowości lub reguły jak najprostsze [Wnęk *et al.*, 1995].

Klasyfikowanie nowych przykładów

Niech R oznacza zbiór reguł uzyskanych w wyniku indukcji ze zbioru przykładów uczących E^u oraz niech $R_k \subset R$ wg (3.30) oznacza podzbiór reguł odpowiadających pojęciu o identyfikatorze k. Rozpatrzmy zbiór przykładów testowych E^t. Niech $e \in E^t$ oznacza przykład, który ma zostać sklasyfikowany. Przykład ten można zapisać (uwzględniając fakt, że niektóre wartości cech mogły zostać pominięte[3]) za pomocą koniunkcji warunków dla wartości cech pewnego (domyślnego) obiektu:

$$e(o) = \bigwedge_{j=1,\ldots,J} [val(a_j(o)) \in V_j^e], \tag{3.39}$$

gdzie zazwyczaj zbiór wartości cechy a_j określony jest jako:

$$V_j^e = \begin{cases} \{v\}, & \text{gdy występuje wartość } v; \\ Dom(a_j), & \text{gdy brak wartości } v. \end{cases} \tag{3.40}$$

Celem postępowania jest przypisanie danego przykładu $e \in E^t$ do jednej z klas, określanych przez podzbiory reguł R_k, których przesłanki zapisywane są jako wyrażenia logiczne zgodnie z (3.31).

[3]Zwykle przyjmuje się, że jeśli brak jakiejś wartości cechy a, to jako wartość tej cechy może wystąpić dowolne $v \in Dom(a)$.

Niech będą dane: przykład e i reguła $r \in R_k$ o przesłance wg (2.76):

$$p(o) = \bigwedge_{\substack{l=1,\ldots,m \\ j_1,\ldots,j_m \in \overline{1..J}}} [val(a_{j_l}(o)) \propto_l Term_l(a_{j_l})]. \tag{3.41}$$

Mówimy, że istnieje *całkowite dopasowanie* (ang. *complete match*) przesłanki reguły r do przykładu e, jeśli:

$$(\forall l \in \overline{1..m})\ (\exists v \in V^e{}_{j_l})\ [val(a_{j_l}(o)) \propto_l v]. \tag{3.42}$$

Mówimy, że zachodzi *częściowe dopasowanie* (ang. *partial match*) przesłanki reguły r do przykładu e, jeśli

$$\begin{aligned} &(\exists l' \in \overline{1..m})\ [(\forall v \in V^e{}_{j_{l'}})\ \neg(val(a_{j_l}(o)) \propto_l v) \\ &\wedge\ (\forall l \in \overline{1..m},\ l \neq l')\ (\exists v \in V^e{}_{j_l})\ (val(a_{j_l}(o)) \propto_l v)]. \end{aligned} \tag{3.43}$$

Zauważmy, że dla danego przykładu występują następujące możliwości:

1. Przesłanka żadnej reguły nie jest całkowicie dopasowana do przykładu (sposób postępowania w tym wypadku opisano poniżej).
2. Przesłanka dokładnie jednej reguły jest całkowicie dopasowana do przykładu, przykład może być więc zaklasyfikowany do klasy wskazywanej przez konkluzję tej reguły.
3. Przesłanka więcej niż jednej reguły jest całkowicie dopasowana do przykładu.

W przypadku 3. należy zdecydować, do której ze wskazywanych klas przykład należy. Dla zbioru reguł $R = \cup_{k=1,\ldots,K} R_k$ sposób postępowania (z zastosowaniem metody *rform* [Wnęk *et al.*, 1995], [Grzymała-Busse, 1994]) jest następujący:

1. Dla każdej reguły $r_1, \ldots, r_m \in R_k$ obliczana jest ocena ufności tej reguły $P(r)$ jako stosunek liczby przykładów uczących E_k^u dla klasy, pokrytych przez regułę r, do ogólnej liczby przykładów uczących dla tej klasy.
2. Dla klasy o identyfikatorze k ocenę ufności $P(C_k)$ oblicza się jak prawdopodobieństwo sumy zdarzeń $P(r_1), \ldots, P(r_m)$, tj.:

$$P(C_k) = 1 - \prod_{j=1,\ldots,m} (1 - P(r_j)). \tag{3.44}$$

3. Przykład e jest klasyfikowany do klasy o takim identyfikatorze k_0, dla którego wartość oceny ufności $P(C_{k0})$ jest największa.

Gdy brak jest całkowitego dopasowania przesłanki danej reguły do przykładu, możemy mieć do czynienia z *częściowym dopasowaniem*. Dla metody *rform* przynależność do klasy określa się następująco [Wnęk *et al.*, 1995]:

1. Dla każdej reguły $r_1, \ldots, r_m \in R_k$ obliczana jest ocena ufności tej reguły $P(r)$ jako stosunek liczby selektorów (zob. rozdział 2.5.3) przesłanki reguły, dopasowanych do przykładu, do ogólnej liczby selektorów w tej regule.
2. Dla klasy o identyfikatorze k ocenę ufności $P(C_k)$ oblicza się jak w (3.44).

3. Przykład e jest klasyfikowany do klasy o takim identyfikatorze k_0, dla którego wartość oceny ufności $P(C_{k0})$ jest największa.

Aby uniemożliwić klasyfikowanie przykładów na podstawie zbyt niskich ocen ufności, autor wprowadził następujący schemat wnioskowania z zastosowaniem wartości granicznych β_1, β_2, $0 < \beta_2 < \beta_1 < 1$:

1. Gdy
$$P_{max} = \max_{k=1,\ldots,K} P(C_k) \leqslant \beta_1, \tag{3.45}$$
przykład uznaje się za *nie sklasyfikowany do żadnej klasy* (zgromadzono zbyt słabe dowody przynależności).

2. Gdy $P_{max} > \beta_1$, lecz jednocześnie
$$\max_{k=1,\ldots,K} (P_{max} - P(C_k)) \leqslant \beta_2, \tag{3.46}$$
tj. największa wartość oceny ufności zbyt mało przewyższa pozostałe, *nie da się jednoznacznie określić przynależności przykładu* (może on należeć do więcej niż jednej klasy, choć dowody wymagane do określenia przynależności do jakiejś klasy są wystarczające).

3. Gdy $P_{max} > \beta_1$ i nie jest spełnione (3.46), to przykład jest klasyfikowany do klasy, dla której ocena ufności wynosi P_{max} wg (3.45).

Sposób wyboru podzbioru cech relewantnych

Przydatność cechy a do rozróżnienia klas w zbiorze przykładów E jest określana za pomocą *miary przydatności* P. Wartości tej miary mogą być określane za pomocą:

- algorytmu zastosowanego w systemie *PROMISE* [Baim, 1982],
- algorytmu *PROMISE 2* [Kaufman, 1997], będącego modyfikacją pierwszego algorytmu.

W systemie *PROMISE* zastosowano poniższy algorytm obliczania wartości $P(a, E)$ miary przydatności cechy a do rozróżnienia klas w zbiorze przykładów E [Kaufman, 1997]:

Niech $P = 0$. Dla $m = 1, \ldots, M(a)$:

1. Niech $S = E^{(m)} = \{e \in E \mid a(e) = v_m\}$ będzie podzbiorem zbioru przykładów, dla których cecha a ma wartość v_m.
2. Dopóki S zawiera przykłady należące do więcej niż jednej klasy:
   (a) Znajdź C — najbardziej liczną spośród klas mających najmniejszą liczbę $n_C(v_m) > 0$ przykładów w zbiorze S;
   (b) Zwiększ P o liczbę $\Delta P = card(S \setminus C)/card(C)$;
   (c) Usuń wszystkie przykłady $e \in C$ ze zbioru S (tj. $S \leftarrow S \setminus C$).

Jeśli otrzymana wartość $P = P(a, E) = 0$, przydatność jest największa, podczas gdy wartość $P = 1 - (1/K)$, gdzie K jest liczbą klas, oznacza brak jakiejkolwiek przydatności cechy a do określenia przynależności przykładów $e \in E$ do klas. Do dalszego zastosowania wskazane jest przekształcenie:
$$P \leftarrow 1 + \frac{K}{1-K} \cdot P, \tag{3.47}$$
które powoduje, że wartości $P \in [0, 1]$, przy czym $P = 1$ odpowiada największej, a $P = 0$ najmniejszej wartości miary przydatności P.

Algorytm *PROMISE 2* do obliczania wartości $P = P(a, E)$ miary przydatności cechy a do rozróżnienia klas w zbiorze przykładów E określony jest następująco [Kaufman, 1997]:

Niech $P = 0$. Dla $m = 1, \ldots, M(a)$:

1. Niech $E^{(m)}$ będzie podzbiorem zbioru przykładów, dla których cecha a ma wartość v_m.
2. Znajdź C — najmniej liczną spośród klas mających największą liczbę $n_C(v_m) > 0$ przykładów w zbiorze $E^{(m)}$.
3. Zwiększ P o liczbę $\Delta P = card(E^{(m)} \cap C)/card(C)$.
   Jeśli wartość $P = P(a, E) = M(a)$, przydatność jest największa, podczas gdy wartość $P = M(a)/K$ oznacza brak jakiejkolwiek przydatności cechy a do określenia przynależności przykładów $e \in E$ do klas. Do dalszego zastosowania wskazane jest przekształcenie:
$$P \leftarrow \frac{P \cdot K - M(a)}{M(a) \cdot (K-1)} \tag{3.48}$$
   o identycznych własnościach jak (3.47).

Podzbiór cech relewantnych $A_{rel} \subseteq A$ określany jest za pomocą kryterium:
$$A_{rel} = \{a \in A \mid P(a) \geqslant \beta\}, \; 0 < \beta \leqslant 1, \tag{3.49}$$
przy czym wybór wartości stałej β winien być przedmiotem optymalizacji. Wstępnie przyjęto, że rozpatrywane są wartości $0.5 < \beta \leqslant 1$.

3.2.5 Indukcja reguł z zastosowaniem zbiorów przybliżonych

Zbiory przybliżone zostały wprowadzone przez Z. Pawlaka [Pawlak, 1982] jako nowy środek umożliwiający reprezentację wiedzy niedokładnej, niepewnej i nieprecyzyjnej. Punktem wyjścia do ich zastosowania jest spostrzeżenie, że dane nieprecyzyjne (jak np. wartości cech obiektów, obserwowane jedynie z ograniczoną dokładnością — ze względu na klasę przyrządu pomiarowego, wskutek kwantyzacji wartości cech ilościowych itp.) uniemożliwiają rozróżnienie obiektów. *Nierozróżnialność obiektów* uniemożliwia zatem ich *precyzyjną klasyfikację*.

Nierozróżnialność (ang. *indiscernibility*) może być także interpretowana jako relacja. Łatwo wykazać, że jest to relacja równoważności, w związku z czym można mówić o klasach abstrakcji w zbiorze wszystkich obiektów (którymi w naszym wypadku są poszczególne

3. Przykład e jest klasyfikowany do klasy o takim identyfikatorze k_0, dla którego wartość oceny ufności $P(C_{k0})$ jest największa.

Aby uniemożliwić klasyfikowanie przykładów na podstawie zbyt niskich ocen ufności, autor wprowadził następujący schemat wnioskowania z zastosowaniem wartości granicznych β_1, β_2, $0 < \beta_2 < \beta_1 < 1$:

1. Gdy
$$P_{max} = \max_{k=1,\ldots,K} P(C_k) \leqslant \beta_1, \tag{3.45}$$
przykład uznaje się za *nie sklasyfikowany do żadnej klasy* (zgromadzono zbyt słabe dowody przynależności).

2. Gdy $P_{max} > \beta_1$, lecz jednocześnie
$$\max_{k=1,\ldots,K} (P_{max} - P(C_k)) \leqslant \beta_2, \tag{3.46}$$
tj. największa wartość oceny ufności zbyt mało przewyższa pozostałe, *nie da się jednoznacznie określić przynależności przykładu* (może on należeć do więcej niż jednej klasy, choć dowody wymagane do określenia przynależności do jakiejś klasy są wystarczające).

3. Gdy $P_{max} > \beta_1$ i nie jest spełnione (3.46), to przykład jest klasyfikowany do klasy, dla której ocena ufności wynosi P_{max} wg (3.45).

Sposób wyboru podzbioru cech relewantnych

Przydatność cechy a do rozróżnienia klas w zbiorze przykładów E jest określana za pomocą *miary przydatności* P. Wartości tej miary mogą być określane za pomocą:

- algorytmu zastosowanego w systemie *PROMISE* [Baim, 1982],
- algorytmu *PROMISE 2* [Kaufman, 1997], będącego modyfikacją pierwszego algorytmu.

W systemie *PROMISE* zastosowano poniższy algorytm obliczania wartości $P(a, E)$ miary przydatności cechy a do rozróżnienia klas w zbiorze przykładów E [Kaufman, 1997]:

Niech $P = 0$. Dla $m = 1, \ldots, M(a)$:

1. Niech $S = E^{(m)} = \{e \in E \mid a(e) = v_m\}$ będzie podzbiorem zbioru przykładów, dla których cecha a ma wartość v_m.
2. Dopóki S zawiera przykłady należące do więcej niż jednej klasy:
   (a) Znajdź C — najbardziej liczną spośród klas mających najmniejszą liczbę $n_C(v_m) > 0$ przykładów w zbiorze S;
   (b) Zwiększ P o liczbę $\Delta P = card(S \setminus C)/card(C)$;
   (c) Usuń wszystkie przykłady $e \in C$ ze zbioru S (tj. $S \leftarrow S \setminus C$).

Jeśli otrzymana wartość $P = P(a, E) = 0$, przydatność jest największa, podczas gdy wartość $P = 1 - (1/K)$, gdzie K jest liczbą klas, oznacza brak jakiejkolwiek przydatności cechy a do określenia przynależności przykładów $e \in E$ do klas. Do dalszego zastosowania wskazane jest przekształcenie:
$$P \leftarrow 1 + \frac{K}{1-K} \cdot P, \tag{3.47}$$
które powoduje, że wartości $P \in [0, 1]$, przy czym $P = 1$ odpowiada największej, a $P = 0$ najmniejszej wartości miary przydatności P.

Algorytm *PROMISE 2* do obliczania wartości $P = P(a, E)$ miary przydatności cechy a do rozróżnienia klas w zbiorze przykładów E określony jest następująco [Kaufman, 1997]:

Niech $P = 0$. Dla $m = 1, \ldots, M(a)$:

1. Niech $E^{(m)}$ będzie podzbiorem zbioru przykładów, dla których cecha a ma wartość v_m.
2. Znajdź C — najmniej liczną spośród klas mających największą liczbę $n_C(v_m) > 0$ przykładów w zbiorze $E^{(m)}$.
3. Zwiększ P o liczbę $\Delta P = card(E^{(m)} \cap C)/card(C)$.
   Jeśli wartość $P = P(a, E) = M(a)$, przydatność jest największa, podczas gdy wartość $P = M(a)/K$ oznacza brak jakiejkolwiek przydatności cechy a do określenia przynależności przykładów $e \in E$ do klas. Do dalszego zastosowania wskazane jest przekształcenie:
$$P \leftarrow \frac{P \cdot K - M(a)}{M(a) \cdot (K-1)} \tag{3.48}$$
   o identycznych własnościach jak (3.47).

Podzbiór cech relewantnych $A_{rel} \subseteq A$ określany jest za pomocą kryterium:
$$A_{rel} = \{a \in A \mid P(a) \geqslant \beta\}, \; 0 < \beta \leqslant 1, \tag{3.49}$$
przy czym wybór wartości stałej β winien być przedmiotem optymalizacji. Wstępnie przyjęto, że rozpatrywane są wartości $0.5 < \beta \leqslant 1$.

3.2.5 Indukcja reguł z zastosowaniem zbiorów przybliżonych

Zbiory przybliżone zostały wprowadzone przez Z. Pawlaka [Pawlak, 1982] jako nowy środek umożliwiający reprezentację wiedzy niedokładnej, niepewnej i nieprecyzyjnej. Punktem wyjścia do ich zastosowania jest spostrzeżenie, że dane nieprecyzyjne (jak np. wartości cech obiektów, obserwowane jedynie z ograniczoną dokładnością — ze względu na klasę przyrządu pomiarowego, wskutek kwantyzacji wartości cech ilościowych itp.) uniemożliwiają rozróżnienie obiektów. *Nierozróżnialność obiektów* uniemożliwia zatem ich *precyzyjną klasyfikację*.

Nierozróżnialność (ang. *indiscernibility*) może być także interpretowana jako relacja. Łatwo wykazać, że jest to relacja równoważności, w związku z czym można mówić o klasach abstrakcji w zbiorze wszystkich obiektów (którymi w naszym wypadku są poszczególne

przykłady $e \in E$, przy czym własności zbioru E opisano w rozdziale 3.2.2). *Zbiór przybliżony* (ang. *rough set* [Pawlak, 1982]) określony jest jako para: dolna i górna aproksymacja danego zbioru, przy czym obie te aproksymacje zapisane są z zastosowaniem klas obiektów nierozróżnialnych.

Zastosowanie zbiorów przybliżonych umożliwia rozwiązanie następujących istotnych problemów [Nowicki, Słowiński, Stefanowski, 1992]:

1. Zredukowanie liczby przykładów nadmiarowych oraz cech nadmiarowych, tak aby można było uzyskać minimalny podzbiór cech (zwany także *reduktem*), zapewniający wysoką sprawność klasyfikacji.

2. Pozyskanie wiedzy, dotyczącej klasyfikacji przykładów, w postaci zbioru reguł.

Podstawowe pojęcia teorii zbiorów przybliżonych określono w dalszym ciągu pracy, posługując się definicją systemu informacyjnego, wprowadzonego w rozdziale 3.2.1.

Relacja nierozróżnialności

Niech $B \subset A \cup D$ będzie podzbiorem zbioru wszystkich cech (warunkowych i decyzyjnych) oraz niech $e_1, e_2 \in E$ będą dwoma dowolnymi przykładami. Przykłady te są *nierozróżnialne* przez podzbiór cech B w systemie informacyjnym S [Pawlak, 1982], co oznacza się jako $e_1 \tilde{B} e_2$, wtedy i tylko wtedy, gdy:

$$e_1 \tilde{B} e_2 \Leftrightarrow (\forall a \in B)\,[val(a(e_1)) = val(a(e_2))], \tag{3.50}$$

gdzie $val(a(e))$ jest wartością cechy a dla przykładu e.

Rodzina klas równoważności

$$E/\tilde{B} = \{F_1, \ldots, F_N\} \tag{3.51}$$

oznaczana jest jako B^*. Każda klasa równoważności F_n, $n \in \overline{1..N}$ relacji $\tilde{B}$ jest nazywana *zbiorem definiowalnym w* B [Chmielewski, Grzymała-Busse, 1992] lub *zbiorem* B-*elementarnym* w S. Zbiory A-elementarne nazywane są *atomami* w S [Nowicki, Słowiński, Stefanowski, 1992].

Aproksymacja zbiorów

Niech $Y \subseteq E$ będzie podzbiorem zbioru przykładów oraz $B \subseteq A$ podzbiorem zbioru cech. B-*dolną aproksymację* zbioru Y (oznaczoną jako $\underline{B}Y$) określa się jako [Pawlak, 1982] (zob. rys. 3.3):

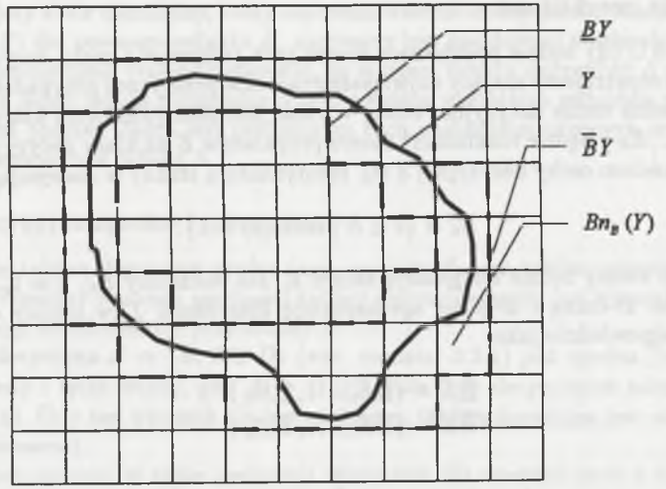
$$\underline{B}Y = \bigcup \{e \in E \mid [e]_B \subseteq Y\}, \tag{3.52}$$

zaś B-*górną aproksymację* zbioru Y (oznaczoną jako $\overline{B}Y$) określa się jako [Pawlak, 1982] (zob. rys. 3.3):

$$\overline{B}Y = \bigcup \{e \in E \mid [e]_B \cap Y \neq \emptyset\}. \tag{3.53}$$

B-*ograniczenie* (obszar wątpliwej klasyfikacji) określone jest jako:

$$Bn_B(Y) = \overline{B}Y \setminus \underline{B}Y. \tag{3.54}$$



Rys. 3.3. Ilustracja pojęć dolnej i górnej aproksymacji zbioru z uwzględnieniem relacji nierozróżnialności

Fig. 3.3. Illustration of concepts of lower and upper approximation of a file with regard to the indiscernibility relation

Dla każdego podzbioru $Y \subseteq E$ określa się *dokładność aproksymacji* zbioru Y przez podzbiór cech B, lub — w skrócie — *dokładność* Y, jako:

$$\alpha_B(Y) = \frac{card(\underline{B}Y)}{card(\overline{B}Y)}. \tag{3.55}$$

Zbiór przybliżony dla Y jest rodziną wszystkich podzbiorów zbioru przykładów E, które mają takie same dolne i górne aproksymacje dla Y [Chmielewski, Grzymała-Busse, 1992]. Na podstawie przedstawionych określeń można stwierdzić, że:

1. Zbiór $\underline{B}Y$ zawiera wszystkie takie przykłady $e \in E$, które na podstawie wartości cech z podzbioru B zbioru wszystkich cech A są klasyfikowane w sposób pewny jako elementy zbioru Y. Jest on największym definiowalnym zbiorem w B, zawartym w Y.

2. Zbiór $\overline{B}Y$ zawiera wszystkie takie przykłady $e \in E$, które na podstawie wartości cech z podzbioru B zbioru wszystkich cech A mogą być sklasyfikowane jako elementy zbioru Y. Zbiór ten jest najmniejszym definiowalnym zbiorem w B, zawierającym Y.

3. Zbiór $Bn_B(Y)$ zawiera wszystkie takie przykłady $e \in E$, które z zastosowaniem podzbioru cech B nie mogą być sklasyfikowane ani jako elementy zbioru Y, ani jako elementy zbioru $-Y$, będącego uzupełnieniem zbioru Y.

Klasyfikacja przybliżona

Niech $B \subseteq A \cup \{d\}$ będzie podzbiorem zbioru cech (włącznie z cechą decyzyjną; w dalszym ciągu rozpatrujemy mający największe znaczenie praktyczne przypadek, gdy występuje tylko jedna cecha decyzyjna, której wartość jest identyfikatorem klasy) oraz niech $\Sigma = \{E_1, \ldots, E_K\}$ będzie rozkładem[4] zbioru przykładów E na klasy decyzyjne, odpowiadające wartościom cechy decyzyjnej d (tj. identyfikatora stanu) w następujący sposób:

$$E_k = \{e \in E \mid class(e) = k\}. \tag{3.56}$$

Rozkład taki zwany będzie *klasyfikacją* zbioru E, zaś podzbiory E_k, $k = 1, \ldots, K$ zwane będą *klasami*. B-dolną i B-górną aproksymację klasyfikacji Σ w tablicy decyzyjnej $\mathcal{S}$ określa się odpowiednio jako:

$$\underline{B}\Sigma = \{\underline{B}E_1, \ldots, \underline{B}E_K\} \tag{3.57}$$

$$\overline{B}\Sigma = \{\overline{B}E_1, \ldots, \overline{B}E_K\}. \tag{3.58}$$

Liczba

$$\gamma_B(\Sigma) = \frac{\sum_{k=1}^{K} card(\underline{B}E_k)}{card(E)} \tag{3.59}$$

nazywana jest *oceną jakości klasyfikacji* Σ za pomocą zbioru cech B, lub — w skrócie — *oceną jakości klasyfikacji* Σ. Określa ona stosunek liczby przykładów B-poprawnie sklasyfikowanych do liczby wszystkich rozpatrywanych przykładów.

Dla klasyfikacji Σ i podzbioru cech $B \subseteq A \cup \{d\}$ połączenie klas decyzyjnych dolnej aproksymacji klasyfikacji Σ

$$POS_B(d) = \underline{B}E_1 \cup \ldots \cup \underline{B}E_K \tag{3.60}$$

nazywane jest *B-pozytywnym regionem* tablicy $\mathcal{S}$.

Redukcja liczby cech

Podzbiór cech $B' \subseteq A$ *zależy* od podzbioru cech $B \subseteq A$ w tablicy decyzyjnej $\mathcal{S}$, co oznacza się jako $B \to B'$, wtedy i tylko wtedy, gdy $\tilde{B} \subseteq \tilde{B'}$. Identyfikacja zależności pomiędzy podzbiorami cech umożliwia redukcję liczby cech.

Podzbiór $B \subseteq A$ jest *niezależny* w $\mathcal{S}$ wtedy i tylko wtedy, gdy

$$(\forall B' \subset B)\,[\overline{B'} \supset \tilde{B}], \tag{3.61}$$

w przeciwnym razie B jest *zależny* w $\mathcal{S}$.

Podzbiór $B \subseteq A$ jest *reduktem* zbioru A w $\mathcal{S}$ wtedy i tylko wtedy, gdy B jest zbiorem niezależnym w $\mathcal{S}$ i $B^* = A^*$. Jeśli $B \subset A$ jest reduktem zbioru cech A, to cechy ze zbioru $A \setminus B$ *są nadmiarowe*. W tablicy decyzyjnej $\mathcal{S}$ może wystąpić więcej niż jeden redukt [Nowicki, Słowiński, Stefanowski, 1992].

Identyfikacja podzbioru cech nadmiarowych może nastąpić za pomocą oceny jakości klasyfikacji $\gamma_B(\mathcal{S})$, gdyż dla reduktu $B \subseteq A$ jest $\gamma_B(\mathcal{S}) = \gamma_A(\mathcal{S})$.

[4] Wg [Kuratowski, 1973] rodzina zbiorów $\{F_n\}$ jest *rozkładem* zbioru F, jeśli zbiory te są rozłączne oraz $\bigcup F_n = F$.

Najmniejszy zbiór niezależny, który zapewnia wartość oceny jakości klasyfikacji równą wartości $\gamma_B(\mathcal{S})$ dla pewnego reduktu B, nazywany jest *podzbiorem minimalnym* w $\mathcal{S}$ (lub *reduktem minimalnym*). Należy podkreślić, że w danej tablicy decyzyjnej $\mathcal{S}$ może wystąpić więcej niż jeden redukt minimalny. Część wspólna wszystkich reduktów minimalnych nazywana jest *jądrem*. Jądro jest podzbiorem cech najbardziej istotnych ze względu na klasyfikację elementów tablicy $\mathcal{S}$.

Tablice decyzyjne zgodne i niezgodne

Rozróżnia się tablice decyzyjne *zgodne* (ang. *consistent*) oraz tablice *niezgodne* (ang. *inconsistent*). Poniższy warunek zgodności tablicy decyzyjnej może być sformułowany w kategoriach pojęć teorii zbiorów przybliżonych.

Tablica decyzyjna $\mathcal{S} = \langle E, A \cup D \rangle$ (zob. rozdział 3.2.1) jest *zgodna* (lub *zdeterminowana*) wtedy i tylko wtedy, gdy $A \to D$ (tj. zbiór cech decyzyjnych zależy od zbioru cech warunku). Gdy ten warunek nie jest spełniony, tablica decyzyjna jest *niezgodna* (lub *niezdeterminowana*).

Dla tablicy zgodnej w razie spełnienia warunków dla wartości cech $a \in A$ uzyskuje się wartości cech decyzyjnych, które *jednoznacznie określają decyzje* (np. co do zaklasyfikowania obiektu do jednej klasy). Dla tablicy niezgodnej *decyzje nie są jednoznacznie określone przez warunki*, lecz uzyskujemy więcej niż jednoelementowy zbiór wartości cech decyzyjnych odpowiadający decyzjom, które mogłyby być podjęte w razie spełnienia tych warunków [Nowicki, Słowiński, Stefanowski, 1992].

Formalne kryterium zgodności tablicy decyzyjnej może być określone za pomocą funkcji $\partial_S : E \to \overline{1..K}$, zwanej *uogólnioną decyzją w* $\mathcal{S}$ i określonej następująco [Skowron, 1993]:

$$\partial_A(e) = \{k \in \overline{1..K} \mid (\exists e' \in E)\,[e\ \tilde{A}\ e' \wedge d(e) = k]\}. \tag{3.62}$$

Tablica decyzyjna jest zgodna, gdy dla każdego $e \in E$ zachodzi $card(\partial_A(e)) = 1$. W przeciwnym razie tablica jest niezgodna. Łatwo można udowodnić, że tablica decyzyjna $\mathcal{S}$ jest zgodna wtedy i tylko wtedy, gdy $POS_A(d) = E$.

Reguły decyzyjne

Na podstawie tablicy decyzyjnej można zidentyfikować określoną "porcję" wiedzy. Wiedza ta jest reprezentowana w postaci *zbioru reguł decyzyjnych* R.

W dalszych rozważaniach ograniczymy się do przypadku, gdy w tablicy decyzyjnej $\mathcal{S} = \langle E, A \cup D \rangle$ występuje tylko jedna cecha decyzyjna d, tj. zbiór D jest jednoelementowy. Niech ponadto

$$V = V(d) \cup \bigcup_{a \in A} V(a) \tag{3.63}$$

będzie zbiorem wszystkich możliwych wartości cech (warunku i decyzyjnej).

Poniżej przedstawiono formalny sposób określenia elementarnej reguły decyzyjnej dla tablicy decyzyjnej $\mathcal{S}$ (wg [Skowron, 1993]). *Formułami atomowymi* nad $B \subseteq A \cup \{d\}$ i V są wyrażenia o postaci $a = v$, zwane *deskryptorami na* B i V, gdzie $a \in B$ oraz $v \in V(a)$. Niech $\mathcal{F}(B, V)$ oznacza zbiór wszystkich formuł nad B i V, domknięty ze względu na operatory logiczne $\vee$ i $\wedge$. Dla dowolnej formuły $p \in \mathcal{F}(B, V)$ oznaczmy przez p_S zbiór wszystkich przykładów o własności p (zwany *znaczeniem* p *w tablicy decyzyjnej* $\mathcal{S}$). Gdy

p jest formułą atomową, zbiór ten określony jest jako:

$$p_S = \{e \in E \mid a(e) = v\}, \tag{3.64}$$

zaś dla koniunkcji i alternatywy formuł $p, q \in \mathcal{F}(B, V)$ w sposób rekurencyjny, odpowiednio jako:

$$(p \wedge q)_S = p_S \cap q_S, \tag{3.65}$$

$$(p \vee q)_S = p_S \cup q_S. \tag{3.66}$$

Zbiór $\mathcal{F}(A, V)$ jest nazywany zbiorem *formuł warunkowych* tablicy decyzyjnej $\mathcal{S}$ i jest oznaczany jako $C(A, V)$ [Skowron, 1993].

Regułą decyzyjną dla tablicy decyzyjnej zgodnej $\mathcal{S}$ jest dowolne wyrażenie o postaci:

$$p \Rightarrow [d = v], \tag{3.67}$$

gdzie $p \in C(A, V)$ jest formułą warunkową i $v \in V(d)$ jest wartością cechy decyzyjnej d.

Reguła decyzyjna $p \Rightarrow [d = v]$ jest *prawdziwa* w tablicy decyzyjnej $\mathcal{S}$ wtedy i tylko wtedy, gdy zbiór znaczeń przesłanki p jest podzbiorem zbioru znaczeń formuły atomowej $[d = v]$, czyli $p_S \subseteq [d = v]_S$.

Reguła decyzyjna $p \Rightarrow [d = v]$ jest *$\mathcal{S}$-dokładna* w tablicy decyzyjnej $\mathcal{S}$ wtedy i tylko wtedy, gdy zbiór znaczeń przesłanki p jest identyczny ze zbiorem znaczeń formuły atomowej $[d = v]$, czyli $p_S = [d = v]_S$.

Regułą decyzyjną dla tablicy decyzyjnej niezgodnej $\mathcal{S}$ jest dowolne wyrażenie o postaci:

$$\alpha_\theta \Rightarrow [\partial_A = \theta], \tag{3.68}$$

gdzie $\theta \subseteq \overline{1..K}$ jest podzbiorem identyfikatorów klas, α_θ jest formułą warunkową oraz ∂_A jest uogólnioną decyzją wg (3.62).

Reguły decyzyjne mogą być generowane bezpośrednio na podstawie *zredukowanej tablicy decyzyjnej*. Niech dla rozpatrywanej tablicy decyzyjnej $\mathcal{S} = \langle E, A \cup \{d\}\rangle$ podzbiór $B \subseteq A$ będzie reduktem zbioru cech warunku A. Każda klasa równoważności F_n, $n = 1, \ldots, N$ relacji nierozróżnialności $\tilde{B}$ wg (3.51) może być opisana za pomocą jednej unikalnej formuły warunkowej wg (3.67), zwanej *opisem* [Wong, Li, Ziarko, 1986]:

$$Des(F_n) = \bigwedge_{a \in B} [a(e_n) = v], \tag{3.69}$$

gdzie $v \in V(a)$,
oraz każda klasa decyzyjna E_k, $k = 1, \ldots, K$ może być opisana za pomocą formuły decyzyjnej:

$$Des(E_k) = [d(e_k) = k], \tag{3.70}$$

gdzie $e_n \in F_n$ jest jednym z elementów należących do klasy równoważności relacji $\tilde{B}$, zaś $e_k \in E_k$ jest przykładowym elementem klasy E_k. Wówczas pierwotny zbiór reguł, które mogą być otrzymane bezpośrednio ze zredukowanej tablicy decyzyjnej $\mathcal{S} = \langle E, B \cup \{d\}\rangle$, zawiera reguły o postaci:

$$Des(F_n) \overset{c_{n,k}}{\Rightarrow} Des(E_k), \tag{3.71}$$

gdzie $c_{n,k}$ jest oceną stopnia pewności reguły, która może być np. określona jako [An *et al.*, 1995]:

$$c_{n,k} = max\{p_{k,n}, 1 - p_{k,n}\}, \tag{3.72}$$

gdzie

$$p_{k,n} = \frac{card(E_k \cap F_n)}{card(F_n)} \tag{3.73}$$

może być uznane za ocenę *prawdopodobieństwa warunkowego* wystąpienia zdarzenia "klasa E_k", pod warunkiem że klasyfikowany element należy do klasy równoważności F_n.

Dla klasyfikacji $\Sigma = \{E_1, \ldots, E_K\}$, gdzie E_k wg (3.56), zbiór reguł decyzyjnych R jest połączeniem:

- zbioru *reguł pewnych* (ang. *certain rules*) R_c, wygenerowanych na podstawie dolnej aproksymacji klasyfikacji $\underline{B}\Sigma$ wg (3.58),
- zbioru *reguł możliwych* (ang. *possible rules*) R_p, wygenerowanych na podstawie obszaru wątpliwej klasyfikacji:

$$Bn_B(\Sigma) = \{\underline{B}E_1 \setminus \overline{B}E_1, \ldots, \underline{B}E_K \setminus \overline{B}E_K\}. \tag{3.74}$$

Należy podkreślić, że opisany powyżej pierwotny zbiór reguł może zawierać warunki określone za pomocą cech, których wartości nie są potrzebne do określenia opisu danej koncepcji. Dalsze uogólnienie zbioru reguł może być osiągnięte wskutek usunięcia takich warunków zawartych w przesłance $Des(F_n)$, bez utraty zgodności. Po usunięciu maksymalnej możliwej liczby warunków uzyskuje się zbiór *reguł maksymalnie ogólnych* (ang. *maximally general rules*) [An *et al.*, 1995].

Klasyfikowanie nowych przykładów

Niech R oznacza zbiór reguł pozyskanych indukcyjnie ze zbioru przykładów uczących E^u, R_k, $k = 1, \ldots, K$ oznacza podzbiór reguł, odpowiadających pojęciu o identyfikatorze k, oraz $e \in E^t$ oznacza klasyfikowany przykład, który nie był wcześniej stosowany w procesie indukcji reguł. Decyzja dotycząca określenia klasy, do której należy przypisać ten przykład, może być podjęta na podstawie trzech czynników: *mocy, specyficzności, poparcia* (ang. odpowiednio *strength, specificity, support*) [Grzymała-Busse, 1994].

W pracy dla reguł przybliżonych, których przesłanki są całkowicie dopasowane do przykładu, określa się [Grzymała-Busse, 1994]:

- *Moc reguły* $MR(r)$ jako łączną liczbę przykładów poprawnie sklasyfikowanych przez tę regułę w fazie trenowania reguł.
- *Specyficzność reguły* $SR(r)$ jako liczbę selektorów w przesłance rozpatrywanej reguły.
- *Poparcie pojęcia* $SUPP(C_k)$ jako wyrażenie:

$$SUPP(C_k) = \sum_{\substack{r \in R_k \\ p(r) \simeq e}} MR(r) \cdot SR(r), \tag{3.75}$$

gdzie $p(r)$ oznacza przesłankę reguły r, zaś znak $\simeq$ w wyrażeniu "$p(r) \simeq e$" oznacza operator relacji pełnego dopasowania.

Moc reguły $MR(r)$ może być także określona za pomocą *przybliżonej miary sprawności* (ang. *rough performance measure*) danej reguły, obliczanej jako stosunek liczby wszystkich przykładów poprawnie sklasyfikowanych przez tę regułę do łącznej liczby przykładów, do których jest dopasowana przesłanka $p(r)$ tej reguły.

Przykład $e \in E^t$ należy do takiej klasy (odpowiadającej pewnej koncepcji C_k), dla której istnieją reguły $r \in R_k$ o przesłankach dopasowanych do przykładu e oraz dla której poparcie $SUPP(C_k)$ tego pojęcia wg (3.75) jest największe.

W razie braku pełnego dopasowania przesłanki reguły r do przykładu e rozpatruje się *częściowe dopasowanie*, które ma miejsce wtedy, gdy co najmniej jeden selektor reguły r jest dopasowany do deskryptora przykładu e. Wówczas określa się dodatkową miarę $MF(r)$ *stopnia dopasowania* przesłanki reguły do przykładu [Grzymała-Busse, 1994], równą stosunkowi liczby selektorów przesłanki reguły, dopasowanych do odpowiednich deskryptorów przykładu, do ogólnej liczby selektorów, tworzących przesłankę reguły. Przykład e jest klasyfikowany jako należący do takiej klasy (odpowiadającej pojęciu C_k), dla której następujące wyrażenie osiąga maksimum [Grzymała-Busse, 1994]:

$$SUPP(C_k) = \sum_{\substack{r \in R_k \\ p(r) \approx e}} MF(r) \cdot MR(r) \cdot SR(r), \tag{3.76}$$

gdzie $\approx$ w wyrażeniu "$p(r) \approx e$" jest operatorem częściowego dopasowania przesłanki reguły do przykładu.

Jeśli dla co najmniej dwóch pojęć C_i, C_k, $i \neq k$ zachodzi $SUPP(C_i) = SUPP(C_k)$, należy zdecydować, któremu pojęciu odpowiada przykład. W tym celu można się posłużyć heurystyką, jak np. *przypisz obiektowi e pojęcie C_k o największym identyfikatorze k* [Grzymała-Busse, 1994]. Należy podkreślić, że przedstawione rozwiązanie zależy bardzo istotnie od sposobu uporządkowania koncepcji (tj. klas stanów technicznych). Uporządkowanie to może bowiem być przypadkowe i kolejność numerowania klas stanów nie musi być nośnikiem użytecznej informacji.

3.2.6 Indukcja drzew decyzyjnych

Indukcja drzewa decyzyjnego może być prowadzona dla przykładów, zapisanych w postaci *tablicy decyzyjnej* $\mathcal{S} = \langle E, A \cup \{d\} \rangle$. Wartość cechy decyzyjnej d określa klasę stanu, do której przypisany jest dany przykład.

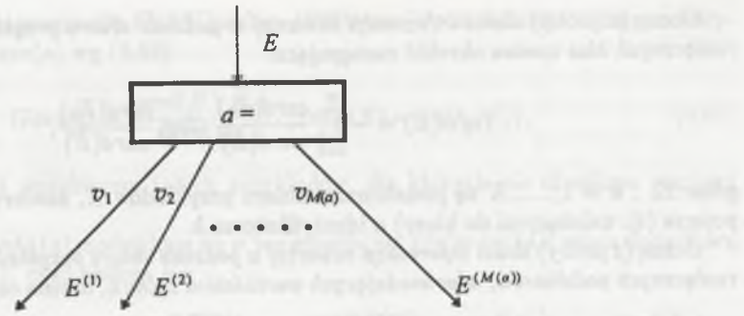
Rekurencyjny algorytm budowy drzewa decyzyjnego

Algorytm indukcji pełnego drzewa decyzyjnego ma charakter rekurencyjny. Algorytm ten dla tablicy decyzyjnej $\mathcal{S} = \langle E, A \cup \{d\} \rangle$ można przedstawić następująco:

1. Jeśli wszystkie pozostające do sklasyfikowania przykłady $e \in E$ *należą do jednej klasy* E_k, czyli spełniają warunek:

$$(\forall e \in E)\,[d(e) = k], \tag{3.77}$$

to *utwórz liść drzewa* dla tego podzbioru przykładów, przypisany do klasy o identyfikatorze k, i zakończ postępowanie.



Rys. 3.4. Węzeł drzewa decyzyjnego dla cechy a o wartościach $v_1, \ldots, v_{M(a)}$ oraz zbioru przykładów E

Fig. 3.4. A node of a decision tree for the attribute a having values $v_1, \ldots, v_{M(a)}$ and the set of examples E

2. Ze zbioru cech A, które nie zostały jeszcze zastosowane do podziału zbioru przykładów, wybierz cechę $a \in A$ oraz utwórz odpowiadający jej węzeł i tyle krawędzi zeń wychodzących, ile wartości $v \in V(a)$ ma cecha a (zob. rys. 3.4).

3. Dokonaj rozkładu zbioru E pozostających do sklasyfikowania przykładów na podzbiory:

$$E^{(m)} = [a = v_m]_{\mathcal{S}}\,, m = 1, \ldots, M(a), \tag{3.78}$$

gdzie $E = E^{(1)} \cup \ldots \cup E^{(M(a))}$ oraz $M(a) = card(V(a))$ jest mocą zbioru $V(a)$ wartości cechy a, oraz przyporządkuj je do kolejnych krawędzi, wychodzących z uprzednio utworzonego węzła.

4. Usuń cechę a ze zbioru cech A.

5. Zastosuj opisany algorytm rekurencyjnie do niepustych podzbiorów $E^{(1)}, \ldots, E^{(M(a))}$.

Sposób wyboru cechy

Podstawowym problemem jest określenie kryterium umożliwiającego wybór cechy stosowanej do rozbudowy drzewa decyzyjnego (por. punkt 2 na s. 87). Do określania zazwyczaj używanych kryteriów stosuje się *entropię*. W rozpatrywanym przypadku jest ona miarą ilości informacji zawartej w zjawisku, w którym zmienna losowa dyskretna a może przyjmować jedną ze swoich wartości (których liczba równa jest $M(a)$).

Entropia jest wyznaczana na podstawie prawdopodobieństw wystąpienia danych zdarzeń. Domyślnie zakłada się, że prawdopodobieństwo wystąpienia każdego z przykładów jest identyczne, co umożliwia estymację wartości prawdopodobieństw za pomocą odpowiednich częstości występowania przykładów z danych klas[5].

[5]Warto zwrócić uwagę na fakt, że przytoczone założenie zwykle nie jest spełnione (dla zbiorów przykładów). Np. prawdopodobieństwo stanu niezdatnego dla poprawnie skonstruowanych, wytworzonych i eksploatowanych maszyn jest znacznie mniejsze od prawdopodobieństwa stanu zdatności. Również prawdopodobieństwa przyjęcia przez cechę a swoich wartości $v \in V(a)$ zwykle nie są jednakowe.

Ocenę (z próby) *ilości informacji zawartej w podziale zbioru przykładów $E \neq \emptyset$ na K rozłącznych klas* można określić następująco:

$$Info(E) = -\sum_{k=1}^{K} \frac{card(E_k)}{card(E)} \cdot log_2 \frac{card(E_k)}{card(E)}, \tag{3.79}$$

gdzie E_k, $k = 1, \ldots, K$ są podzbiorami zbioru przykładów E, zawierającymi przykłady pojęcia (tj. należącymi do klasy) o identyfikatorze k.

Ocenę (z próby) *ilości informacji zawartej w podziale zbioru przykładów $E \neq \emptyset$ na $M(a)$ rozłącznych podzbiorów, odpowiadających wartościom cechy a*, można określić następująco:

$$SplitInfo(a) = -\sum_{\substack{m=1,\ldots,M(a) \\ E^{(m)} \neq \emptyset}} \frac{card(E^{(m)})}{card(E)} \cdot log_2 \frac{card(E^{(m)})}{card(E)}. \tag{3.80}$$

Wartość oczekiwaną ilości informacji po podziale zbioru przykładów $E \neq \emptyset$ na podzbiory $E^{(m)}$, $m = 1, \ldots, M(a)$, odpowiadające wartościom cechy a, określa się jako:

$$Info_a(E) = -\sum_{\substack{m=1,\ldots,M(a) \\ E^{(m)} \neq \emptyset}} \frac{card(E^{(m)})}{card(E)} \cdot Info(E^{(m)}). \tag{3.81}$$

Do wyboru cechy, która będzie przypisana do tworzonego węzła drzewa decyzyjnego, można stosować następujące kryteria (jako kolejną cechę stosuje się tę, dla której funkcja kryterialna ma wartość największą):

1. Kryterium *maksymalnego przyrostu ilości informacji*, spowodowanego zastosowaniem cechy a [Quinlan, 1986]:

$$Gain(a) = Info(E) - Info_a(E). \tag{3.82}$$

2. Kryterium *względnego przyrostu ilości informacji* w odniesieniu do ilości informacji zawartej w podziale zbioru przykładów odpowiednio do przyjmowanych wartości przez cechę a [Quinlan, 1986]:

$$GainRatio(a) = \frac{Gain(a)}{SplitInfo(a)}. \tag{3.83}$$

Brakujące wartości cech w przykładach uczących

Większość znanych algorytmów (np. *C4.5* [Quinlan, 1993]) umożliwia indukcję drzewa decyzyjnego na podstawie przykładów, w których brak jest co najmniej jednej wartości cechy. Wymaga to modyfikacji sposobu:

- obliczania przyrostu ilości informacji $Gain(a)$ wg (3.82),
- przydziału przykładu z brakującymi wartościami cech do podzbiorów $E^{(m)}$ wg (3.78),
- klasyfikowania nowego przykładu z brakującymi wartościami cech.

J. R. Quinlan w programie *C4.5* [Quinlan, 1993] stosuje rozwiązanie, w którym koryguje się wartość $Gain(a)$ wg (3.82):

$$Gain(a) = \frac{card(E \setminus E_*)}{card(E)} \cdot (Info(E) - Info_a(E)), \tag{3.84}$$

gdzie $E_* \subset E$ jest podzbiorem takich przykładów, dla których nie określono wartości cechy $a(e)$.

Wartość $SplitInfo(a)$ modyfikuje się w ten sposób, jak gdyby cecha a miała dodatkową wartość "nieznaną", czyli oblicza się:

$$SplitInfo(a) = -\sum_{\substack{m=1,\ldots,M(a) \\ E^{(m)} \neq \emptyset}} \frac{card(E^{(m)})}{card(E)} \cdot log_2 \frac{card(E^{(m)})}{card(E)} - \frac{card(E_*)}{card(E)} \cdot log_2 \frac{card(E_*)}{card(E)}, \tag{3.85}$$

przy czym E_* określono powyżej.

Zmianie ulega również sposób podziału zbioru przykładów E na podzbiory $E^{(m)}$, $m = 1, \ldots, M(a)$, odpowiadające kolejnym wartościom cechy a. Przykład e_* z brakującą wartością cechy a jest przypisywany do wszystkich zbiorów $E^{(m)}$ z wagą $w_m(a)$, odpowiadającą ocenie prawdopodobieństwa wystąpienia wartości v_m cechy a:

$$w_m(a) = \frac{card(E^{(m)} \setminus E_*)}{card(E \setminus E_*)}. \tag{3.86}$$

Podobny sposób zastosowano przy klasyfikacji nowego przykładu e_* z brakującymi wartościami cech [Quinlan, 1993]. Począwszy od węzła odpowiadającego cesze, dla której brakuje wartości, przykład e_* jest klasyfikowany do wszystkich klas, przypisanych do liści-potomków (bezpośrednich lub pośrednich) tego węzła. W wyniku otrzymuje się wiele ścieżek i odpowiadających im ocen klasyfikacji, ważonych przez prawdopodobieństwa wystąpienia poszczególnych wartości. Przykład e_* może być więc sklasyfikowany bądź niejednoznacznie, bądź też można przyjąć, że należy on do klasy o największym prawdopodobieństwie wystąpienia.

Przycinanie drzewa

Indukcyjny algorytm budowy drzewa decyzyjnego, opisany na s. 86, umożliwia tworzenie bardzo złożonych drzew decyzyjnych, które są *nadmiernie dopasowane* (ang. *overfit*) do dostarczonych przykładów uczących. To nadmierne dopasowanie do przykładów uczących objawia się niską sprawnością klasyfikatora (dużą liczbą błędów klasyfikacji) dla nowych przykładów testowych. Aby uniknąć nadmiernego dopasowania do danych uczących, dokonuje się *przycinania* (ang. *pruning*) drzewa uzyskanego w wyniku indukcji na podstawie zbioru przykładów uczących.

Istotą tego zabiegu jest usunięcie wybranych poddrzew (gałęzi) i zastąpienie ich węzłami terminalnymi (liśćmi). Do utworzonych w taki sposób liści są przypisywane zbiory przykładów uczących $E_\circ$, należące do więcej niż jednej klasy. Zamiast identyfikatora klasy, do której należą wszystkie przykłady, danemu liściowi przypisywany jest rozkład ocen prawdopodobieństwa p_k, $k = 1, \ldots, K$ zdarzenia "przykład $e_\circ \in E_\circ$ należy do klasy o identyfikatorze k":

$$p_k = \hat{P}\{e_\circ \in E_k\}, \; k = 1, \ldots, K. \tag{3.87}$$

Wartości niektórych z tych ocen prawdopodobieństwa mogą być równe zeru (gdy dana klasa nie występuje w zbiorze przykładów przypisanych do nowo utworzonego liścia).

Podczas gdy algorytm budowy drzewa ma charakter zstępujący (ang. *top-down*), algorytm przycinania drzewa (zob. [Quinlan, 1993]) ma charakter wstępujący (ang. *bottom-up*) z postępowaniem *w głąb* (ang. *depth-first proceeding*).

Algorytm przycinania drzewa można przedstawić następująco [Quinlan, 1993]:

1. Rozpocznij przycinanie od pierwszego liścia w drzewie.
2. Badaj najmniejsze poddrzewo, zawierające rozpatrywany liść:
   (a) Zastąp to poddrzewo liściem;
   (b) Oszacuj błąd klasyfikatora;
   (c) Jeśli błąd jest dopuszczalny dla liczby n_{err} przykładów błędnie sklasyfikowanych spośród $n = card(E_o)$ przykładów przypisanych do tego węzła, przejdź do rozpatrywania węzła nadrzędnego i jeśli nie jest on korzeniem całego drzewa[6], powtarzaj kroki 2a÷2c dla poddrzewa, którego korzeniem jest ten węzeł;
3. Kontynuuj przycinanie (począwszy od kroku 2) dla gałęzi jeszcze nie przyciętej, rozpoczynając od następnego liścia (jeśli istnieje).

Błąd obliczany w kroku 2b jest porównywany z wartością krytyczną "pesymistyczną" [Quinlan, 1993], określoną jako górna granica przedziału ufności $U_p(n_{err}, n)$ dla rozkładu dwumianowego, gdzie $p = 1 - \alpha$ jest poziomem ufności (J. R. Quinlan w programie *C4.5* domyślnie przyjmuje $p = 25\%$).

Błędy w przykładach uczących

Błędy w przykładach uczących mogą być spowodowane:

- błędnie zapisanymi wartościami cech warunku (wskutek błędnie przeprowadzonej kwantyzacji cech ilościowych lub pomyłek przy "manualnym" zapisie danych),
- błędnym przyporządkowaniem danego przykładu do klasy, którą reprezentuje.

Jeśli występowanie błędów w zbiorze przykładów uczących nie ma charakteru systematycznego, mamy do czynienia z dodatkowym *szumem*.

Niech $a \in A$ będzie cechą warunku, której wartości zawierają błędy, zaś $E(a)$ będzie podzbiorem przykładów, przypisanych do węzła, utworzonego dla cechy a. Niech ponadto

$$E^{(m)} = \{e \in E(a) \mid a(e) = v_m\}\,,\ m = 1, \ldots, M(a)$$

będą podzbiorami $E(a)$, odpowiadającymi kolejnym wartościom cechy a. Należy określić kryterium relewantności cechy a dla klasyfikacji zbioru przykładów $E(a)$. Niech:

$$\chi^2 = \sum_{\substack{m=1,\ldots,M(a) \\ E^{(m)} \neq \emptyset}} \sum_{k=1}^{K} \frac{(p_k - p'_k)^2}{p'_k}, \tag{3.88}$$

[6]Osiągnięcie korzenia całego drzewa jest równoważne zastąpieniu tego drzewa przez jeden liść, co byłoby trywialne.

gdzie $p_k = card(E_k \cap E^{(m)})$ jest liczbą przykładów, dla których cecha $a(e) = v_m$, zaś $p'_k = card(E^{(m)}))/K$ jest oczekiwaną liczbą przykładów dla losowego (ze względu na wartość cechy a) rozkładu zbioru przykładów $E(a)$. Statystyka χ^2 wg (3.88) może być zastosowana do weryfikacji hipotezy zerowej, że wartości cechy a są niezależne od wartości cechy decyzyjnej d, określającej przynależność przykładu do pewnej klasy.

Jeśli cecha a nie jest relewantna, postępuje się podobnie jak w wypadku przycinania drzewa: tworzy się liść, zawierający przykłady należące do wielu klas, a ocenę prawdopodobieństwa p_k przynależności przykładu do danej klasy określa się zgodnie z (3.87).

Budowa drzew dla cech ilościowych

Najprostszy sposób postępowania w wypadku cech ilościowych polega na bardzo zgrubnej kwantyzacji wartości tej cechy poprzez określenie wartości granicznej. Jeśli a jest cechą ilościową i $V(a) = \{v_m \mid m = 1, \ldots, M(a)\}$ jest uporządkowanym zbiorem wartości tej cechy, to wspomniana (jedyna) wartość graniczna v_{gr} może być określona jako średnia arytmetyczna jakichkolwiek dwóch sąsiadujących ze sobą wartości:

$$v_{gr} = \frac{v_m + v_{m+1}}{2}\,,\ 1 \leqslant m < M(a). \tag{3.89}$$

Wybór wartości granicznej może być przedmiotem optymalizacji, np. z uwzględnieniem kryteriów (3.82) lub (3.83) [Quinlan, 1993].

Wówczas rozkład zbioru przykładów (zob. p. 3 na s. 87) daje w wyniku dwa podzbiory:

$$E^{(1)} = [a < v_{gr}]_S\,,\ E^{(2)} = [a \geqslant v_{gr}]_S. \tag{3.90}$$

Wartości każdej cechy ilościowej można również poddać kwantyzacji. Dalsze postępowanie ze zbiorem wartości skwantowanych jest identyczne z postępowaniem opisanym dla cech o wartościach dyskretnych.

3.2.7 Kryteria oceny przydatności metod "uczenia maszynowego"

Aby umożliwić wybór metody "uczenia maszynowego", której zastosowanie byłoby optymalne ze względu na rozwiązywany problem diagnostyczny, należy utworzyć odpowiedni układ kryteriów. Na uwagę zasługują następujące grupy kryteriów (nie wyczerpujące jednak wszystkich możliwości oceny przydatności metod "uczenia maszynowego"):

1. *Poprawność działania klasyfikatora* (odpowiadającego "porcji" wiedzy pozyskanej na podstawie dostarczonych danych uczących), której miarą jest *łączna sprawność klasyfikatora* η_{ov}[7]. Jest to podstawowe kryterium oceny przydatności klasyfikatora do rozpatrywanego typu danych.
2. *Złożoność obliczeń* potrzebnych do wykonania w celu pozyskania danej "porcji" wiedzy. Kryterium to ma szczególnie istotne znaczenie w wypadku pozyskiwania wiedzy na podstawie danych o obiekcie, którego cechy ewoluują w funkcji czasu. W wypadku zagadnień "statycznych" jego znaczenie jest drugoplanowe.

[7]Miara ta jest określona w dalszej części pracy – por. zależność (3.104) na s. 97.

3. *Łatwość generowania objaśnień* w procesie wnioskowania (odpowiedzi na podstawowe pytania: *DLACZEGO* i *JAK*), bardzo istotna m.in. ze względu na problem odpowiedzialności za skutki działań podejmowanych wskutek konsultacji z systemem doradczym [Cholewa, Pedrycz, 1987].

Wielokryterialny wybór metody winien być uzależniony od specyfiki rozwiązywanego zadania pozyskiwania wiedzy.

Znaczenie problemu wyboru odpowiedniej metody "uczenia maszynowego" maleje wtedy, gdy wyniki otrzymane różnymi metodami mogą być reprezentowane w sposób umożliwiający jednoczesne ich wykorzystanie w procesie wnioskowania diagnostycznego, realizowanego przez wspólne procedury wnioskowania.

3.3 Ocena pozyskanej wiedzy

Jak wspomniano, przeprowadzone badania dotyczyły dwóch grup metod pozyskiwania wiedzy: od specjalistów oraz w sposób indukcyjny (z zastosowaniem metod "uczenia maszynowego"). Sposoby oceny pozyskanej wiedzy są przede wszystkim związane ze źródłami tej wiedzy. W niektórych wypadkach możliwe jest "krzyżowe" stosowanie metod oceny, typowych dla jednej grupy metod pozyskiwania wiedzy, do oceny wiedzy pozyskanej metodami przynależącymi do drugiej grupy.

Przedmiotem oceny mogą być:

- cała baza wiedzy — mamy wówczas do czynienia z *oceną integralną*,
- poszczególne "porcje" wiedzy bez uwzględnienia zależności zachodzących pomiędzy nimi (np. pojedyncze reguły) — jest to *ocena szczegółowa*.

Ocena wiedzy może być dokonywana *pod względem formalnym* lub *pod względem merytorycznym*.

Dla potrzeb dalszego opisu przyjęto założenie, że baza wiedzy "operacyjnej" zapisana jest w postaci zbioru reguł, przy czym wszystkie te reguły mają wspólną dziedzinę — zob. zależność (2.76) na s. 48 oraz poprzedzający ją opis.

Ocena reguły pod względem formalnym dotyczy sprawdzenia, czy możliwe jest spełnienie przesłanki tej reguły w przestrzeni poszukiwań. Ponieważ przesłanki reguły są budowane za pomocą wyrażeń atomowych, w których poszczególne termy są elementami odpowiednich słowników, formalna poprawność reguły sprowadza się do niesprzeczności jej przesłanki.

Ocena reguły pod względem merytorycznym w głównej mierze sprowadza się do oceny jej *relewantności*[8].

Integralna ocena poprawności bazy wiedzy pod względem formalnym ma na celu wykrycie [Cholewa, Pedrycz, 1987]:

- *reguł brakujących*;
- *reguł nadmiarowych*, przy czym reguły r_1, r_2 są nadmiarowe, jeśli:

$$r_1(o) = p(o) \Rightarrow q_1(o) \wedge r_2(o) = p(o) \Rightarrow q_2(o) \wedge (\forall o \in O)\,[q_1(o) = q_2(o)]; \qquad (3.91)$$

[8]Zob. przypis 23 na s. 47.

- *reguł sprzecznych*:

$$r_1(o) = p(o) \Rightarrow q_1(o) \wedge r_2(o) = p(o) \Rightarrow q_2(o) \wedge (\exists o \in O)\,[q_1(o) \neq q_2(o)]; \qquad (3.92)$$

- *reguł zapętlonych* (lub *pętli reguł*), przy czym reguły $r_i(o) = p_i(o) \Rightarrow q_i(o)$, $i = 1, \ldots, n \wedge n > 1$ są zapętlone, jeśli:

$$(\exists o \in O)\,[p_2(o) = q_1(o), \ldots, p_n(o) = q_{n-1}(o) \wedge p_1(o) = q_n(o)]. \qquad (3.93)$$

Błędy te są zwykle wykrywane za pomocą specjalnego oprogramowania, wspomagającego programistę bazy wiedzy w działaniach związanych z "pielęgnowaniem" bazy wiedzy. Postępowanie to, jako standardowe, nie wymaga dalszego opisu.

Podstawowy sposób oceny poprawności pozyskanej wiedzy (obejmujący oba względy: formalny i merytoryczny) polega na jej stosowaniu do działania systemu doradczego oraz na ocenie poprawności działania tego systemu. Ocena ta może obejmować między innymi poprawność:

1. przebiegu procesu "rozumowania",
2. realizacji procesu objaśniania,
3. uzyskiwanych konkluzji.

Dokonanie oceny bazy wiedzy z zastosowaniem wymienionych powyżej technik wymaga użycia systemu doradczego. Ponadto na uzyskaną ocenę (zwłaszcza w odniesieniu do punktów 1 i 2) znaczny wpływ ma poprawność działania procedur wnioskowania i objaśniania. Z tych względów postanowiono ograniczyć uwagę do oceny poprawności konkluzji (punkt 3) uzyskiwanych w wyniku zastosowania pozyskanej "porcji" wiedzy.

W dalszym ciągu rozdziału omówiono sposoby oceny wiedzy *pod względem merytorycznym*, opracowane lub zastosowane przez autora dla potrzeb opisanych badań.

3.3.1 Sposoby oceny wiedzy pozyskanej od specjalistów

Wg opinii W. Cholewy[9], pozyskiwanie wiedzy poprzez ocenę przez specjalistę "porcji" wiedzy już wcześniej pozyskanej jest (zwłaszcza w odniesieniu do reguł) sposobem najbardziej przystępnym dla wielu specjalistów.

Nie wykluczając możliwości integralnej oceny bazy wiedzy oraz grup reguł (np. objętych jednym kontekstem), w pracy rozpatrywano problemy *szczegółowej oceny pojedynczych reguł*.

W ramach przeprowadzonych badań do oceny reguł, pozyskanych (w sposób bezpośredni lub pośredni) od specjalistów, stosowane były następujące techniki bezpośredniej oceny przez specjalistę (lub grupę specjalistów):

- za pomocą *formularzy papierowych*,
- z zastosowaniem *formularza elektronicznego*.

[9]Komunikat prywatny.

Tabl. 3.1. Przedziały zmienności wartości stopnia przekonania o słuszności reguł

Nazwa stopnia przekonania	N(r)	P(r)
nie rozumiem treści reguły	0.00	1.00
całkowicie się zgadzam	1.00	1.00
prawie na pewno się zgadzam	0.75	1.00
raczej się zgadzam	0.55	1.00
raczej się nie zgadzam	0.00	0.45
prawie na pewno się nie zgadzam	0.00	0.25
zupełnie się nie zgadzam	0.00	0.45

Poszczególnym regułom są przypisywane oceny *stopnia przekonania o słuszności reguły* $B(\ )$. Przyjęto, że, podobnie jak to ma miejsce dla stwierdzeń, przedział zmienności stopnia przekonania $B(\ )$ można określić poprzez podanie miar *stopnia konieczności* $N(\ )$ i *stopnia możliwości* $P(\ )$, specyficznych dla każdej reguły (zob. s. 46).

Każda reguła może być oceniana przez więcej niż jednego specjalistę, co wiąże się z potrzebą łączenia opinii grupy specjalistów.

Sposób określania ocen stopnia przekonania

Wartościami stopnia przekonania $B(r)$ reguły r są liczby z przedziału $[0,1]$. Dla nieznanej wartości stopnia przekonania $B(r)$ określa się przedział zmienności jako:

$$0 \leqslant N(r) \leqslant B(r) \leqslant P(r) \leqslant 1, \tag{3.94}$$

przy czym jednocześnie można założyć, że $N(r)$ i $P(r)$ spełniają (2.64).

Aby ułatwić specjaliście wyrażanie jego przekonania o słuszności reguł (poprzez przypisanie danej regule wartości stopnia przekonania), postanowiono określić nieliczny zbiór wartości jakościowych (zob. tabl. 3.1).

Wartość stopnia przekonania o nazwie *nie rozumiem treści reguły* interpretowana jest jako brak oceny (co jest równoważne z oceną *nie mam zdania*).

Ocena za pomocą formularzy "papierowych"

Ocena dokonywana jest przez specjalistę, który otrzymuje zbiór ocenianych reguł, zapisanych w języku naturalnym z wyszczególnieniem wszystkich składników reguły, jak (por. opis na s. 47 i n.):

1. unikalny *identyfikator reguły*,
2. *dziedzina reguły* z wyszczególnieniem tworzących ją klas obiektów i/lub pojedynczych obiektów konkretnych (wraz z ich lokalizacją),
3. *przesłanka reguły* zapisana w postaci koniunkcji selektorów,
4. *koniunkcja reguły* zapisana w postaci pojedynczego warunku elementarnego.

Specjalista wyraża swoją opinię o danej regule, wybierając odpowiednią wartość stopnia przekonania i zaznaczając swój wybór na specjalnej ankiecie do rejestrowania wyników oceny. Wypełnione ankiety wraz z danymi umożliwiającymi identyfikację oceniającego specjalisty są następnie zwracane programiście bazy wiedzy, który w razie potrzeby dokonuje łączenia opinii (zob. poniżej) i aktualizuje oceny stopnia przekonania (o słuszności danych reguł) w bazie wiedzy.

Należy zwrócić uwagę na fakt, że specjalista, dysponując wydrukiem zapisów dotyczących reguł należących do pewnego zbioru, może także dokonać oceny merytorycznej reguł z tego zbioru. Oceny takie można przekazywać w formie pisemnej (jako uwagi) programiście bazy wiedzy. Programista może następnie przeprowadzić edycję reguł zapisanych w bazie, która była przedmiotem oceny specjalisty.

Ocena za pomocą formularza "elektronicznego"

Sposób postępowania jest zbliżony do opisanego uprzednio. Specjalista przegląda reguły zapisane w bazie wiedzy, wyświetlając ich pełną treść (jak w punktach 1÷4 powyżej) na ekranie. Następnie wybiera wartość stopnia przekonania o słuszności ocenianej reguły, która wpisywana jest do specjalnie dla tego celu przeznaczonej tabeli w bazie wiedzy. Oceny pochodzące od grupy specjalistów są łączone w sposób opisany poniżej i przypisywane danej regule.

Łączenie opinii specjalistów

Jeśli zgromadzono więcej niż jedną ocenę stopnia przekonania (o słuszności danej reguły), co ma miejsce w wypadku oceny reguł przez kilku specjalistów, pojawia się potrzeba *połączenia tych opinii*. W tym celu stosowany jest specjalny operator *łączenia stopni przekonania* [Cholewa, 1997a] (zwany także *operatorem agregacji*), który stosowany jest do wartości miary $NP(r)$ dla reguły r:

$$NP_{og}(r) = \frac{w(r) \cdot NP(r) + w_{ex} \cdot B_{ex}(r)}{w(r) + w_{ex}}, \tag{3.95}$$

gdzie:

$w(r)$ — *waga przypisana stopniowi przekonania reguły w bazie wiedzy* (może być uzależniona od źródła reguły, np. od renomy specjalisty–autora tej reguły),

w_{ex} — *waga przypisywana opiniom danej osoby* oceniającej reguły w bazie wiedzy,

$NP(r)$ — *dotychczasowa wartość stopnia przekonania* przypisana ocenianej regule (być może wskutek agregacji wielu opinii),

$B_{ex}(r)$ — *wartość stopnia przekonania przypisana regule przez oceniającego ją specjalistę* (zalecane są wartości wg tabl. 3.1),

przy czym miara $NP(r)$ dla reguły r określona jest poprzez wartości stopnia konieczności $N(r)$ i stopnia możliwości $P(r)$ tej reguły w taki sam sposób jak stopień prawdziwości stwierdzenia w zależności (2.65).

Istotnym zagadnieniem jest wybór wartości początkowych wag $w(r)$. Zaleca się np. przyjmować $w(r) = 1$ [Cholewa, 1997a].

3.3.2 Sformalizowane sposoby oceny zawartości bazy wiedzy

Opisane sposoby dotyczą oceny bazy wiedzy w sposób integralny pod względem merytorycznym. Przedmiotem oceny może być baza wiedzy pochodzącej od specjalistów lub pozyskanej indukcyjnie.

W ogólności, w opisywanych badaniach zastosowano następujące sformalizowane sposoby oceny [Moczulski *et al.*, 1997]:

- *bazy wiedzy* za pomocą oddzielnego zbioru przykładów testowych;
- *poprawności* przygotowanej *bazy przykładów.*

Przed opisaniem sposobów postępowania w obu tych zadaniach zostaną wprowadzone miary liczbowe ogólnie stosowane do merytorycznej oceny poprawności bazy wiedzy.

Kryteria merytorycznej poprawności bazy wiedzy

Oceny merytorycznej bazy wiedzy dokonuje się poprzez określenie wartości estymatora *sprawności klasyfikatora* w odniesieniu do odpowiedniego *zbioru przykładów testowych*, których przynależność do klas jest określona. Zakłada się przy tym, że rozkład przykładów jest losowy i reprezentatywny dla oceny klasyfikatora.

Przy założeniu że dostępny jest jedynie zbiór przykładów, pobrany losowo jako tzw. *próba prosta*[10] z populacji wszystkich możliwych przykładów, można określić następujące oceny *błędu klasyfikatora* dla liczby klas $K > 1$ [Arciszewski, Dybala, Wnek, 1992]:

1. *Łączny błąd względny* (ang. *overall error rate*):

$$\epsilon_{ov} = \frac{n_{err}}{card(E^t)}, \tag{3.96}$$

   gdzie n_{err} oznacza liczbę przykładów błędnie sklasyfikowanych;

2. *Względny błąd pominięcia* (ang. *omission error rate*):

$$\epsilon_{om} = \frac{1}{K}\sum_{k=1}^{K}\frac{card(E_k^+)}{card(E_k^t)}, \tag{3.97}$$

   gdzie E_k^+ oznacza podzbiór przykładów błędnie pominiętych (tj. przykładów $e \in E_k^t$, które nie zostają sklasyfikowane jako należące do "swojej" klasy), zaś $E_k^t \subseteq E^t$ jest podzbiorem przykładów pozytywnych dla klasy o indeksie k;

3. *Względny błąd niesłusznego zaliczenia* (ang. *commission error rate*):

$$\epsilon_{cm} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{card(E^t \setminus E_k^t)}\cdot\sum_{\substack{j=1,\ldots,K\\ j\neq k}} card(E_{kj}^-), \tag{3.98}$$

   gdzie $E_{kj}^- \subset E^t$ jest zbiorem przykładów pozytywnych klasy E_j^t (a więc jednocześnie przykładów negatywnych dla klasy E_k^t), niesłusznie zaklasyfikowanych do klasy E_k^t.

[10] Próba losowa nazywana jest próbą prostą, jeżeli przy jej pobieraniu prawdopodobieństwo pobrania każdego przykładu jest jednakowe [Chmurawa, Bińkowski, 1980].

Istotną wadą estymatorów, określonych zależnościami (3.96) do (3.98), jest nieuwzględnianie empirycznego prawdopodobieństwa wystąpienia przykładu należącego do danej klasy. Może to powodować zaakceptowanie klasyfikatora, który nie umożliwia rozpoznania przynależności obiektu do mniej licznych klas. Aby usunąć tę wadę opisanych miar, wprowadzono wagi określone następująco:

$$w_k = \frac{card(E^t)}{K\cdot card(E_k^t)}, \; k = 1,\ldots,K. \tag{3.99}$$

Niech ponadto:

$$w_\Sigma = \sum_{k=1,\ldots,K} w_k. \tag{3.100}$$

Wówczas można określić:

1. *łączny ważony błąd względny*:

$$\epsilon_{ov}^w = \frac{\sum_{k=1,\ldots,K} w_k \cdot n_{err|k}}{card(E^t)}, \tag{3.101}$$

   gdzie $n_{err|k}$ oznacza liczbę błędnie sklasyfikowanych przykładów $e \in E_k$,

2. *względny ważony błąd pominięcia*:

$$\epsilon_{om}^w = \frac{1}{w_\Sigma}\cdot\sum_{k=1,\ldots,K} w_k \cdot \frac{card(E_k^+)}{card(E_k^t)}, \tag{3.102}$$

3. *względny ważony błąd niesłusznego zaliczenia*:

$$\epsilon_{cm}^w = \frac{1}{w_\Sigma}\cdot\sum_{k=1,\ldots,K}\frac{w_k}{card(E^t \setminus E_k^t)}\cdot\sum_{\substack{j=1,\ldots,K\\ j\neq k}} card(E_{kj}^-). \tag{3.103}$$

Oceny te są bardzo wrażliwe na niejednorodny rozkład błędów w poszczególnych klasach.

Podstawowe techniki oceny sprawności klasyfikatora

Jak już wspomniano, zawartość bazy wiedzy wyznacza pewnego rodzaju *klasyfikator* stosowany do klasyfikowania nowych przypadków. Klasyfikacja polega na przypisaniu danemu przypadkowi stanu technicznego (lub klasy tego stanu) — zob. rozdział 2.2.

Ocena merytoryczna bazy wiedzy uzyskanej w sposób indukcyjny na podstawie zbioru *przykładów uczących* E^u dokonywana jest za pomocą *zbioru przykładów testowych* E^t, które nie były stosowane do identyfikacji (pozyskiwania) ocenianej "porcji" wiedzy[11].

Do oceny klasyfikatora stosuje się miarę *łącznej sprawności*[12] *klasyfikatora*, określonej jako uzupełnienie do jedynki łącznego błędu względnego wg (3.96), czyli:

$$\eta_{ov} = 1 - \epsilon_{ov} = 1 - \frac{n_{err}}{card(E^t)}, \tag{3.104}$$

[11] W literaturze angielskojęzycznej stosowane jest określenie *unseen examples* [Grzymała-Busse, 1994].

[12] Nazwa stosowana w [Cholewa, Kaźmierczak, 1995]. W literaturze angielskojęzycznej stosuje się określenie *performance* [Weiss, Kapouleas, 1989], [Michalski, 1997].

gdzie, jak poprzednio, n_{err} oznacza liczbę przykładów błędnie sklasyfikowanych.

Gdy liczba iteracji $N_I > 1$, ocena łącznej sprawności klasyfikatora uzyskiwana jest jako wartość średnia ocen $\eta_{ov,n}$ wg (3.104), określonych w poszczególnych iteracjach:

$$\overline{\eta}_{ov} = \frac{1}{N_I} \sum_{n=1}^{N_I} \eta_{ov,n}. \tag{3.105}$$

Tabl. 3.2. Porównanie technik oceny błędów względnych [Weiss, Kapouleas, 1989]

Nazwa techniki	*Hold-out*	*Random subsampling*	*Leaving-one-out*	*k-fold Cross-validation*
Przypadków uczących	$card(E^u)$	$card(E^u)$	$card(E) - 1$	$\frac{(100-k)}{100} \cdot card(E)$
Przypadków testowych	$card(E^t)$	$card(E^t)$	$card(E) - 1$	$\frac{k}{100} \cdot card(E)$
Iteracji	1	$N_I \ll card(E)$	$card(E)$	k

Podstawowe techniki oceny błędów względnych zestawiono w tablicy 3.2. Pozostawiono oryginalne nazwy tych technik wg [Weiss, Kapouleas, 1989] ze względu na ich powszechne stosowanie oraz trudności związane z ich przetłumaczeniem.

Istotę poszczególnych technik można opisać następująco [Weiss, Kapouleas, 1989]:

1. *Hold-out* polega na jednorazowym losowym podziale zbioru przykładów E na dwa rozłączne podzbiory: uczący E^u i testowy E^t. Następnie (w jednej iteracji) zbiór uczący służy do zbudowania klasyfikatora, którego sprawność jest oceniana na podstawie wyników klasyfikacji przykładów ze zbioru testowego. Wymagana jest duża liczba przykładów $N = card(E)$, lecz wg [Weiss, Kapouleas, 1989] dla $N > 1000$ niezależnych statystycznie przykładów można mieć (statystyczną) pewność, że uzyskana ocena sprawności klasyfikatora jest bardzo zbliżona do wartości określonej dla populacji generalnej przykładów.

2. *Random subsampling* polega na N_I-krotnym powtarzaniu czynności określonych przez *Hold-out*, przy czym podziały zbioru przykładów na podzbiory przykładów uczących i testowych w każdej iteracji dokonywane są *niezależnie od podziałów w innych iteracjach*. Uzyskane w każdej iteracji oceny sprawności klasyfikatora uśrednia się, stosując zależność (3.105). Uśredniona ocena ma mniejsze odchylenie standardowe niż ocena uzyskana techniką *Hold-out*.

3. *Leaving-one-out* jest techniką bardzo kosztowną obliczeniowo i wskutek tego mającą zastosowanie do mniej licznych zbiorów przykładów. Liczba iteracji jest równa liczbie przykładów. W kolejnych iteracjach ze zbioru wszystkich przykładów E usuwa się jeden przykład e, tworząc podzbiór przykładów uczących $E^u = E \setminus \{e\}$ oraz jednoelementowy zbiór przykładów testowych $E^t = \{e\}$. Uzyskana średnia ocena sprawności klasyfikatora jest zwykle "pesymistyczna", niższa od ocen uzyskiwanych innymi technikami.

4. *k-fold Cross-validation* jest uogólnieniem techniki *Leaving-one-out*. Zbiór przykładów dzielony jest na k podzbiorów $E = E_1 \cup \ldots \cup E_k$, przy czym przydział przykładów do podzbiorów jest losowy. W i-tej iteracji jako zbiór przykładów uczących

stosowany jest zbiór $E^u = E \setminus E_i$, zaś jako zbiór przykładów testowych — zbiór E_i. Jak poprzednio, uzyskane w każdej iteracji oceny sprawności klasyfikatora uśrednia się, stosując zależność (3.105).

W badaniach proponowanych metod i środków pozyskiwania wiedzy, opisanych w rozdziale 5, stosowano techniki: *10-fold Cross-validation, Random subsampling* oraz *Hold-out* dla $card(E^u)/card(E) = 70\%$.

Ocena merytoryczna za pomocą zbioru przykładów

W fazie walidacji systemu doradczego często stosuje się ocenę merytoryczną bazy wiedzy *za pomocą nowych przykładów testowych*:

1. pochodzących od specjalistów, np. w formie tzw. opisów przypadków (ang. *case stories*);

2. uzyskanych jako macierze wartości cech sygnałów (zgromadzonych podczas obserwacji diagnostycznej obiektu);

3. będących wynikami obliczeń.

Z pierwszym źródłem przykładów wiąże się problem odpowiedniej reprezentacji opisów przypadków. Zwykle opisy takie są zawarte w publikacjach i przed ich zastosowaniem do testowania z użyciem metod indukcyjnych muszą zostać przetworzone do postaci, w której zapisywane są przykłady testowe (które zwykle są reprezentowane za pomocą stwierdzeń).

Z drugim źródłem danych testowych łączy się problem wyboru obiektu konkretnego, który będzie obserwowany za pośrednictwem sygnałów. Mogą one pochodzić:

1. od obiektu, dla którego jest przygotowywana baza wiedzy będąca przedmiotem weryfikacji (zwłaszcza podczas działania tego obiektu w warunkach, w których ujawniły się pewne jego specyficzne stany techniczne);

2. ze stanowiska modelowego, przy czym zaletą takiego rozwiązania jest łatwość wywoływania stanów niesprawności obserwowanego obiektu (można prowadzić eksperymenty diagnostyczne czynne i bierne — choć jedynie dla pewnego podzbioru takich stanów).

Bardzo istotne znaczenie ma trzecie źródło przykładów testowych [Cholewa, Kiciński, 1996], [Moczulski, Maniak, 1997]. Przykłady testowe są uzyskiwane jako wyniki *eksperymentów numerycznych*, przeprowadzanych za pomocą odpowiedniego oprogramowania symulacyjnego (zwłaszcza dla stanów awaryjnych, podczas których obiekt konkretny byłby narażony na uszkodzenie lub zniszczenie, co mogłoby się wiązać z zagrożeniem życia i zdrowia pracowników obsługi oraz z zagrożeniem środowiska). Wspomniany eksperyment numeryczny może być prowadzony *w sposób systematyczny*, co umożliwia równomierne wypełnienie przestrzeni poszukiwania.

Właściwe opracowanie planu eksperymentu numerycznego jest szczególnie istotne [Mańczak, 1976]. Zdaniem autora celowe jest prowadzenie eksperymentu w taki sposób, by wartości parametrów sterujących przebiegiem obliczeń wykonywanych przez oprogramowanie symulacyjne generowane były *w sposób losowy* (przy przyjęciu pewnych rozkładów prawdopodobieństwa tych parametrów, traktowanych jako wartości zmiennych losowych).

Autor opracował odpowiednie oprogramowanie, które na podstawie opisu planu eksperymentu generuje wartości losowe parametrów, sterujące przygotowaniem przykładów testowych przez oprogramowanie symulacyjne. Przydatność takiego sposobu postępowania wymaga jeszcze potwierdzenia.

Sposób merytorycznej oceny bazy przykładów

Autor wysunął koncepcję zastosowania metod "uczenia maszynowego" do merytorycznej oceny bazy przykładów [Moczulski, 1997a]. Ocena ta dotyczy łącznie:

- *kompletności* zbioru przykładów,
- *poprawności* zbioru przykładów.

Istotą rozwiązania jest zastosowanie zweryfikowanego wcześniej zbioru reguł (zwanego dalej *zbiorem reguł zweryfikowanych*) do klasyfikowania ocenianego zbioru przykładów. Zakłada się, że zbiór reguł jest *kompletny* i *poprawny*. Zbiór ten może np. pochodzić od uznanego specjalisty w rozpatrywanej dziedzinie zastosowania.

Zbiór przykładów uznaje się za kompletny, jeśli wiedza pozyskana indukcyjnie na podstawie tych przykładów umożliwia (dla potrzeb wnioskowania diagnostycznego) określenie klasyfikatora o wystarczającej sprawności.

Poprawność zbioru przykładów rozumiana jest w ten sposób, że zgromadzone wartości cech są wystarczającą podstawą do prawidłowego określenia klasy, do której dany przykład należy.

Ocena poprawności badanych baz przykładów E_B następuje poprzez tzw. *uczenie przyrostowe* [Hong, Mozetic, Michalski, 1986], [Wnęk *et al.*, 1995] z hipotezami wstępnymi, którymi są reguły ze zbioru reguł zweryfikowanych R_W. Przy poczynionych wyżej założeniach w odniesieniu do zbioru reguł R_W oraz zbioru przykładów E_B uznaje się zbiór przykładów za *kompletny i poprawny*, gdy różnica łącznych błędów względnych wg (3.96) nie jest statystycznie istotnie różna od 0, tj. gdy

$$P\{|\,\hat{\varepsilon}_{ov}(R_W + E_B) - \hat{\varepsilon}_{ov}(R_W)\,| > \alpha\} = 0\,,\ \text{dla}\ \alpha \ll 1. \tag{3.106}$$

3.3.3 "Krzyżowe" stosowanie sposobów oceny pozyskanej wiedzy

Przez ocenę "krzyżową" rozumiana jest ocena wiedzy pozyskanej innymi metodami niż zastosowane metody oceny. Występują tutaj dwie możliwości:

1. wiedza pozyskana od specjalisty oceniana jest *za pomocą zbioru przykładów testowych*,
2. wiedza pozyskana metodami indukcyjnymi jest oceniana *bezpośrednio przez specjalistę* (lub *grupę specjalistów*).

W opisanych badaniach możliwe było zastosowanie obu sposobów, gdyż wiedza pozyskana z zastosowaniem wymienionych grup metod była reprezentowana w zunifikowany sposób. Otrzymane oceny pozwalają na modyfikację zawartości bazy wiedzy (w odniesieniu do treści reguł oraz stopni przekonania dotyczących tych reguł), co umożliwia osiągnięcie lepszych wyników stosowania bazy wiedzy do diagnozowania przypadków objętych daną dziedziną zastosowań.

3.4 Scenariusz procesu pozyskiwania wiedzy

Opisane uprzednio metody pozyskiwania wiedzy diagnostycznej oraz oceny bazy wiedzy są podstawą działań praktycznych. Ciąg takich działań, mający charakter sekwencyjno-iteracyjny, można nazwać *procesem pozyskiwania wiedzy* (zob. rozdział 2.6).

Pomiędzy procesem pozyskiwania wiedzy a procesem projektowo-konstrukcyjnym występują pewne daleko idące analogie [Cholewa, Moczulski, 1990]. Umożliwiają one zastosowanie terminologii zgodnej z używaną w nauce konstrukcji [Dietrych, 1985].

Aby ułatwić właściwe zaplanowanie oraz przeprowadzenie procesu pozyskiwania wiedzy, autor zaproponował *scenariusz procesu pozyskiwania wiedzy*. Z punktu widzenia metodologii działania twórczego może on być interpretowany jako *program postępowania*, który uznawany jest za swoisty rodzaj metody [Dietrych, 1985]. Stanowi to jednocześnie uzasadnienie umieszczenia opisu scenariusza procesu pozyskiwania wiedzy w niniejszym rozdziale.

Zapis modelu procesu pozyskiwania wiedzy diagnostycznej przedstawiono na rys. 3.5. W modelu tym można wyróżnić następujące etapy:

- *opracowanie koncepcyjne* (stadia 1÷2),
- *opracowanie prototypu* (stadia 3÷4),
- *opracowanie pełnej wersji* (stadia 5÷6).

Oprócz powyższych etapów, typowych dla procesu twórczego, w modelu występuje jeszcze:

- *przekazanie bazy wiedzy do użytkowania* i *nadzór autorski oraz okresowa kwalifikacja bazy wiedzy* (stadia 7÷8).

3.4.1 Opracowanie koncepcyjne

Etap ten obejmuje *identyfikację i opis potrzeby* oraz *określenie koncepcji*. Ma on zasadniczy wpływ na skuteczność dalszych działań, których wynikiem ma być baza wiedzy diagnostycznej.

Identyfikacja i opis potrzeby

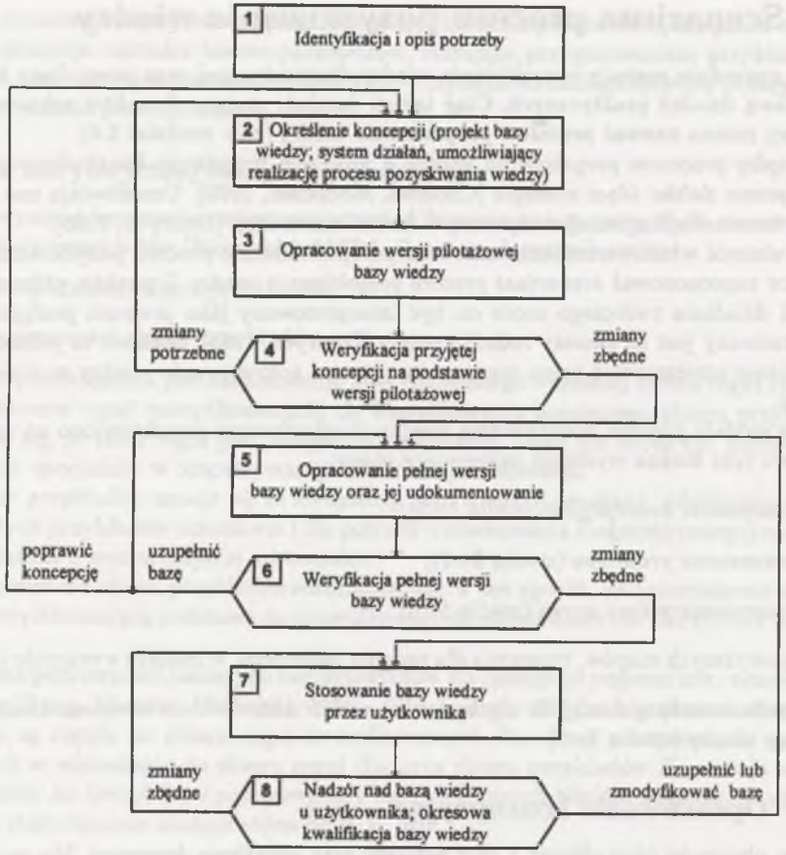
Przedmiotem identyfikacji (blok "1" na rys. 3.5) winny być w szczególności:

1. *Identyfikacja dziedziny zastosowań*

   Zaleca się, aby dziedzina zastosowań nie była zbyt szeroka. Winna ona być dobrze określona i udokumentowana. Warto się upewnić, czy stosowanie systemu doradczego do wspomagania rozwiązywania problemów diagnostycznych typowych dla tej dziedziny jest celowe. Właściwa identyfikacja dziedziny ma bowiem zasadnicze znaczenie dla dalszego przebiegu procesu pozyskiwania wiedzy.

2. *Identyfikacja wymagań w zakresie zapewnienia jakości bazy wiedzy*

   Zgodnie z zaleceniami normy dotyczącej zapewnienia jakości oprogramowania [ISO 9000-3:1994], wymagania te winny być zidentyfikowane wspólnie z przyszłym



Rys. 3.5. Model procesu pozyskiwania wiedzy diagnostycznej
Fig. 3.5. A model of the process of acquisition of diagnostic knowledge

użytkownikiem systemu doradczego. Odpowiedzialność za zidentyfikowanie wymagań spoczywa na twórcach systemu doradczego (a więc również na programistach bazy wiedzy)[13].

3. *Identyfikacja dostępnych źródeł wiedzy*

W diagnostyce technicznej identyfikacja winna dotyczyć co najmniej takich najbardziej istotnych źródeł, jak:

(a) *dokumentacja obiektu* (techniczno-ruchowa, konstrukcyjna, inna);

(b) *literatura i publikacje*, dotyczące konkretnego obiektu lub klasy obiektów (z ewentualnym uwzględnieniem analogii w wypadku obiektów podobnej konstrukcji);

[13]Bardziej szczegółowe omówienie ważnych problemów zapewnienia jakości baz wiedzy wykracza poza zakres merytoryczny niniejszej rozprawy.

(c) *specjaliści* skłonni współpracować z zespołem programistów bazy wiedzy;

(d) *bazy danych*, przy czym należy zidentyfikować:

i. jakie cechy są gromadzone i w odniesieniu do jakich obiektów (oraz inne informacje szczegółowe, związane ze sposobem reprezentacji tych danych);

ii. warunki działania obiektów, których cechy są gromadzone w bazie danych;

iii. stany techniczne, które przypisano obiektom podczas gromadzenia danych o tych obiektach w bazie;

iv. przydatność danych zgromadzonych w bazie do ich zastosowania do "uczenia maszynowego" lub do odkrywania zależności o charakterze jakościowym i ilościowym.

4. *Identyfikacja zasobów*

Dotyczy ona zasobów, które mogą być zastosowane w procesie pozyskiwania wiedzy, takich jak:

(a) *dostępne środki programowe*, które mogą być zastosowane do wspomagania procesu pozyskiwania wiedzy (np. takie, jak opisane w rozdziale 4);

(b) *dostępne środowisko sprzętowe* (sprzęt komputerowy oraz oprogramowanie systemowe) *i programowe* (oprogramowanie narzędziowe, szkieletowe systemy doradcze, oprogramowanie specjalistyczne);

(c) *programiści bazy wiedzy*, mogący zrealizować proces pozyskiwania wiedzy (w razie potrzeby winien być utworzony odpowiedni zespół).

Wyniki identyfikacji należy we właściwy sposób opisać. Opis ten winien być weryfikowany w razie potrzeby.

Określenie koncepcji

Przeprowadzona identyfikacja potrzeby umożliwia określenie koncepcji (blok "2" na rys. 3.5), obejmującej:

- *projekt bazy wiedzy*,
- *system działań, umożliwiający realizację tego projektu* (poprzez zapewnienie właściwego przebiegu procesu pozyskiwania wiedzy).

Projekt bazy wiedzy obejmuje wybór (z zastosowaniem odpowiedniego układu kryteriów) sposobu:

1. *reprezentacji wiedzy* (dotyczy opisu dziedziny zastosowań oraz wiedzy "operacyjnej" używanej do wnioskowania diagnostycznego);
2. *reprezentacji danych* (przykłady uczące i testowe);
3. *reprezentacji objaśnień*;
4. *reprezentacji metawiedzy* (dotyczącej stosowania wiedzy "operacyjnej");
5. *przechowywania zawartości bazy wiedzy i jej pielęgnowania*;

6. *dokumentowania zawartości bazy wiedzy;*

7. *oceny bazy wiedzy* (w tym — łączenia ocen pochodzących z różnych źródeł);

8. *aktualizacji zawartości bazy wiedzy.*

Określenie systemu zarządzania projektem ma szczególne znaczenie zwłaszcza w wypadku realizowania procesu pozyskiwania wiedzy przez zespół. Obejmuje to między innymi sposoby:

1. *dokumentowania prac* w zakresie realizacji danego projektu pozyskiwania wiedzy;

2. *zapewnienia wymaganej niezawodności* przechowywania zgromadzonych danych i wiedzy;

3. *oceny specjalistów* (np. ze względu na ich wiarygodność jako źródeł wiedzy) oraz uwzględniania tych ocen podczas aktualizacji bazy wiedzy;

4. *wspomagania prac projektowo-konstrukcyjnych* związanych z tworzoną bazą wiedzy.

Przyjęty system zarządzania projektem powinien być poddany wstępnej ocenie i po zaakceptowaniu przez zespół wykonawców (przy ewentualnym udziale przyszłego użytkownika) może być dopuszczony do dalszych prac.

Ponadto należy opracować wstępną wersję *planu i harmonogramu realizacji procesu pozyskiwania wiedzy.* Winny one być poddane dyskusji w zespole wykonawców, a następnie dyskusji z potencjalnym lub faktycznym użytkownikiem. Po wprowadzeniu ewentualnych modyfikacji winno nastąpić wspólne przyjęcie wstępnego planu jako podstawy do opracowania planu szczegółowego, a następnie zaakceptowanie go.

W wyniku tego etapu otrzymuje się szczegółowy opis założeń projektowych, umożliwiający podjęcie zadań *konstruowania bazy wiedzy diagnostycznej.*

3.4.2 Opracowanie prototypu

Etap ten obejmuje *opracowanie wersji pilotażowej* oraz *weryfikację koncepcji* na podstawie tej wersji. Wynikiem etapu winna być ograniczona baza wiedzy, stanowiąca model pełnej wersji tworzonej bazy wiedzy.

Opracowanie wersji pilotażowej

Pilotażową (lub wstępną, ograniczoną) *wersję bazy wiedzy* (blok "3" na rys. 3.5) opracowuje zespół programistów bazy wiedzy. Wersja ta tworzona jest dla ograniczonego podzbioru dziedziny zastosowań w celu umożliwienia praktycznej weryfikacji przyjętej koncepcji oraz jej ewentualnej modyfikacji. Prace prowadzone są zwykle przez zespół programistów bazy wiedzy (bez udziału specjalistów z zewnątrz).

Prace nad wersją pilotażową bazy wiedzy obejmują:

1. *Pozyskanie opisu dziedziny oraz wstępne pozyskanie wiedzy operacyjnej*

   Uwzględniana jest ograniczona dziedzina zastosowań. Źródłami wiedzy są zwykle:

   (a) udostępniona dokumentacja obiektów,

   (b) dostępna literatura i publikacje.

2. *Reprezentacja wiedzy i jej zapisanie w bazie wiedzy*

   Należy dokonać wyboru sposobu reprezentacji wiedzy. Zapis wiedzy powinien się odbywać w taki sposób, jaki został określony dla pełnej wersji bazy wiedzy. Do zapisania wiedzy można zastosować odpowiednie środki wspomagające.

Weryfikacja koncepcji na podstawie wersji pilotażowej

Weryfikacja koncepcji (blok "4" na rys. 3.5) obejmuje przede wszystkim [Cholewa, Pedrycz, 1987]: ocenę odpowiedniości przyjętego sposobu reprezentacji wiedzy, skuteczności wnioskowania diagnostycznego oraz przebiegu dialogu z różnymi grupami użytkowników danego systemu doradczego (zwłaszcza w odniesieniu do objaśnień). W tym celu wskazane jest przygotowanie przykładów testowych (we współpracy z wybranymi specjalistami).

Po dokonaniu weryfikacji należy opracować wnioski w celu wprowadzenia zmian dotyczących koncepcji bazy wiedzy. Mogą one obejmować: określenie dziedziny zastosowań, sposób reprezentacji wiedzy, szczegóły opisu dziedziny zastosowań, treść bazy wiedzy "operacyjnej" itp.

3.4.3 Opracowanie pełnej wersji bazy wiedzy

Etap ten obejmuje *opracowanie pełnej wersji i jej udokumentowanie* oraz *weryfikację pełnej wersji bazy wiedzy.* Wynikiem etapu winna być wersja bazy wiedzy obejmująca pełny zakres (ustalony na etapie opracowywania założeń projektowych), gotowa do przekazania użytkownikowi.

Opracowanie pełnej wersji bazy wiedzy i jej udokumentowanie

Jest to zwykle najbardziej pracochłonne stadium (blok "5" na rys. 3.5) w całym procesie pozyskiwania wiedzy. Zdaniem autora, istotny wpływ na efektywność pozyskiwania wiedzy (a więc także na czas potrzebny do opracowania pełnej bazy) ma zastosowanie indukcyjnych metod pozyskiwania wiedzy.

W etapie tym występuje:

1. *Uzupełniające pozyskanie wiedzy od specjalistów*

   Potrzeba takich działań może być związana zwłaszcza z poszerzeniem zakresu bazy wiedzy do pełnej dziedziny zastosowań. W pierwszym rzędzie uzupełnienia będą dotyczyć opisu dziedziny. Nie wyklucza się możliwości pozyskiwania wiedzy "operacyjnej" od specjalistów, choć zwykle jest to działanie mało efektywne.

2. *Przygotowanie bazy przykładów uczących dla potrzeb metod indukcyjnych pozyskiwania wiedzy*

   Źródłami przykładów mogą być:

   (a) istniejące bazy danych diagnostycznych (dane zgromadzone podczas normalnej eksploatacji obiektu, czyli w warunkach *biernego eksperymentu diagnostycznego* [Cempel, 1982]);

(b) bazy danych zgromadzonych w trakcie prowadzenia *czynnego eksperymentu diagnostycznego* (zob. np. [Kostka, 1997]);

(c) bazy danych zgromadzonych w wyniku przeprowadzenia *eksperymentu numerycznego* [Kiciński *et al.*, 1995], [Cholewa, Kiciński, 1996] (zwłaszcza w wypadku stanów przedawaryjnych lub awaryjnych, których prawdopodobieństwo wystąpienia jest bardzo małe, lecz ich wystąpienie powoduje znaczne zagrożenie dla obsługi i środowiska naturalnego).

W wypadku eksperymentu czynnego lub numerycznego duże znaczenie ma odpowiednie opracowanie planu tego eksperymentu. Ma to wpływ na rozkład punktów, reprezentujących poszczególne wyniki eksperymentów, w wielowymiarowej przestrzeni przeszukiwania (tj. przestrzeni wartości cech). Należy podkreślić, że zwykle w zbiorach wyników eksperymentów czynnych zbyt wielki udział mają przykłady, odpowiadające różnym niesprawnościom i stanowi niezdatnemu obiektów. Oznacza to, że prawdopodobieństwo wystąpienia przykładu opisującego obiekt w stanie niezdatnym jest wielokrotnie wyższe od wartości prawdopodobieństwa zawodności, które można określić na podstawie badań eksploatacyjnych tego obiektu.

Przykłady takie wymagają zwykle odpowiedniego przetworzenia (obejmującego np. zamianę wartości ilościowych na jakościowe, kodowanie, usunięcie błędów, zapis w wymaganym formacie itp.), aby mogły one być zastosowane do pozyskiwania wiedzy w sposób indukcyjny (z zastosowaniem metod tzw. "uczenia maszynowego"). Przetworzenie może także obejmować klasyfikację przykładów dokonaną przez specjalistę, który przypisuje te przykłady do określonych klas stanów technicznych (elementarnych lub złożonych).

3. *Pozyskanie wiedzy metodami uczenia maszynowego*

   Metody te można stosować, jeśli zostały zgromadzone przykłady, dla których określono przynależność do klas stanów [Cholewa, Pedrycz, 1987]). Zespół realizujący projekt winien dokonać wyboru metod "uczenia maszynowego".

4. *Pozyskanie wiedzy metodami odkryć w bazach danych*

   Metody te stosowane są głównie wtedy, gdy nie jest znana przynależność przykładu do klasy (uczenie bez nauczyciela).

5. *Połączenie wiedzy pochodzącej z różnych źródeł*

   Ze względu na wiele źródeł (np. pozyskiwanie wiedzy od różnych specjalistów oraz wieloma metodami "uczenia maszynowego") zwykle wymagane jest łączenie "porcji" wiedzy pochodzących z różnych źródeł.

Pełna wersja bazy wiedzy winna być *odpowiednio udokumentowana*. Należy rozróżnić *dokumentowanie dla potrzeb zespołu wykonawczego* i *dokumentowanie dla potrzeb użytkownika*, z których pierwsze dotyczy szczegółowych rozwiązań projektowych, wyników testowania itp., drugie zaś ma na celu ogólny opis bazy wiedzy dla potrzeb użytkownika, zwłaszcza w celu określenia zakresu bazy wiedzy (w odniesieniu do zadań diagnostycznych, których rozwiązywanie może być wspomagane przez system doradczy posługujący się tą wersją bazy wiedzy).

Weryfikacja pełnej wersji bazy wiedzy

Do przeprowadzenia weryfikacji pełnej wersji bazy wiedzy (blok "6" na rys. 3.5) potrzebny jest system doradczy. Weryfikację tę można podzielić na:

1. *weryfikację wewnętrzną* bazy wiedzy dokonaną przez jej autorów, dotyczącą:

   (a) poprawności pod względem formalnym (zob. s. 92);

   (b) poprawności pod względem merytorycznym (zob. rozdział 3.3); szczególne znaczenie ma ocena przeprowadzona za pomocą zbioru przypadków testowych;

2. *weryfikację zewnętrzną* bazy wiedzy (zwłaszcza przez osoby nie biorące udziału w dotychczasowych pracach), której celem jest przede wszystkim ocena poprawności uzyskiwanych konkluzji, przebiegu procesu rozumowania oraz generowania objaśnień.

Wszystkie działania winny być dokumentowane. Zapisy powinny zawierać identyfikator weryfikowanej "porcji" wiedzy, opis przeprowadzonego testu, otrzymane wyniki oraz ich ocenę.

Należy podkreślić szczególnie istotne znaczenie tego stadium dla powodzenia całego projektu. Wnioski wynikające z przeprowadzonej weryfikacji bazy wiedzy mogą spowodować potrzebę uzupełnienia tworzonej bazy wiedzy.

W ramach prowadzonych badań, na podstawie wyników weryfikacji bazy wiedzy, autor wielokrotnie zidentyfikował potrzebę optymalizacji:

- zbioru cech służących do opisu przykładów;
- podzbioru cech relewantnych;
- sposobów kwantowania wartości cech ilościowych.

Wyniki weryfikacji bazy wiedzy mogą również spowodować modyfikację zastosowanej koncepcji.

3.4.4 Stosowanie bazy wiedzy i nadzór autorski

Etap ten obejmuje *przekazanie bazy wiedzy do użytkowania* oraz ewentualny *nadzór autorski nad tą wersją bazy wiedzy* u użytkownika.

Przekazanie bazy wiedzy do stosowania przez użytkownika

Etap ten (blok "7" na rys. 3.5) wieńczy proces pozyskiwania wiedzy. Celem jego jest *przekazanie utworzonej bazy wiedzy użytkownikowi* oraz *końcowa ocena skuteczności* wnioskowania diagnostycznego. Skuteczność działania określana jest na podstawie wyników wnioskowania uzyskanych przez system doradczy dla przypadków zaobserwowanych podczas jego użytkowania. Ocena ta dotyczy pośrednio bazy wiedzy tego systemu utworzonej w ramach poprzednich stadiów procesu pozyskiwania wiedzy.

Ocena końcowa skuteczności systemu doradczego obejmuje:

1. *Ocenę skuteczności działania systemu doradczego*

   Ocena ta jest określana na podstawie konkluzji uzyskiwanych podczas użytkowania systemu doradczego.

2. *Ocenę sposobu działania systemu doradczego*

   Ze względu na ocenę bazy wiedzy szczególne znaczenie ma określenie poprawności przebiegu dialogu z użytkownikiem oraz sposobu udzielania odpowiedzi na pytania użytkownika (tj. generowania objaśnień).

3. *Ocenę zakresu poprawnego działania systemu doradczego*

   Ocena ta dotyczy potwierdzenia możliwości stosowania bazy wiedzy do wspomagania wnioskowania diagnostycznego w pełnej dziedzinie zastosowań. W razie stwierdzenia niezadowalającej skuteczności działania systemu doradczego dla jakiegoś podzbioru zadań z dziedziny zastosowań należy dokonać aktualizacji wiedzy dla tego zakresu i wprowadzić odpowiednie poprawki do dokumentacji bazy wiedzy.

4. *Modyfikację zawartości bazy wiedzy przed jej oddaniem do użytkowania*

   Potrzeba wprowadzenia modyfikacji może zostać stwierdzona podczas realizacji trzech poprzednich punktów. Modyfikacje te winny dotyczyć wyłącznie treści bazy wiedzy, w tym metawiedzy (np. określającej konteksty reguł lub uporządkowanie zbioru reguł). W tym stadium nie dokonuje się zmian dotyczących koncepcji lub sposobu reprezentacji wiedzy, gdyż oznaczałoby to konieczność powrotu do początkowych stadiów procesu pozyskiwania wiedzy [Cholewa, Pedrycz, 1987].

Nadzór autorski nad bazą wiedzy u użytkownika

Nadzór autorski (blok "8" na rys. 3.5) wiąże się z koniecznością zapewnienia obsługi gwarancyjnej i/lub pogwarancyjnej. Obejmuje on między innymi:

1. *Okresowe gromadzenie zapisów odpowiadających wszystkim działaniom wykonywanym przez system doradczy*

   Zgromadzone zapisy (tzw. *log*) są poddawane ocenie przez specjalistów oraz autorów bazy wiedzy. Ze względu na ocenę bazy wiedzy szczególne znaczenie ma określenie poprawności uzyskanych konkluzji, przebiegu dialogu z użytkownikiem oraz sposobu odpowiedzi na pytania użytkownika.

2. *Modyfikację zawartości bazy wiedzy w ramach nadzoru autorskiego*

   Modyfikacja ta winna dotyczyć wyłącznie treści bazy wiedzy, w tym metawiedzy (np. dotyczącej kontekstów reguł lub uporządkowania zbioru reguł). W tym stadium nie dokonuje się zmian dotyczących koncepcji lub sposobu reprezentacji wiedzy.

3. *Okresową kwalifikację bazy wiedzy*

   Okresowa kwalifikacja bazy wiedzy zalecana jest przez międzynarodowe normy zapewnienia jakości (norma ogólna [ISO 9001:1994] oraz specjalne zalecenia w odniesieniu do oprogramowania [ISO 9000-3:1994]). Celem postępowania jest ocena odpowiedniości aktualnej wersji bazy wiedzy w odniesieniu do wszystkich uzgodnionych wymagań. Wynikiem oceny mogą być wskazania dotyczące dalszej modyfikacji bazy wiedzy.

3.5 Podsumowanie dotyczące metod pozyskiwania wiedzy diagnostycznej i jej oceny

W rozdziale opisano dwie podstawowe grupy metod pozyskiwania wiedzy diagnostycznej: *od specjalistów* oraz *na podstawie indukcji ze zbioru przykładów uczących*. Pozyskana wiedza jest reprezentowana w sposób deklaratywny.

Pozyskiwanie wiedzy od specjalisty. Podstawowym środkiem reprezentacji wiedzy pozyskanej od specjalisty jest wprowadzona przez autora *relacja diagnostyczna*. Relacje są zapisywane przez specjalistę samodzielnie, za pomocą specjalnie do tego celu opracowanego *formularza papierowego*. Ponadto sformułowano koncepcję *pozyskiwania wiedzy na drodze kroczenia wstecz*, a dla jej realizacji zaproponowano specjalizowane oprogramowanie, zwane *formularzem elektronicznym*.

Zdaniem autora, dalszy rozwój metod będzie następował w nowym kierunku: *pozyskiwania wiedzy od specjalistów-diagnostów, reprezentowanej w sposób proceduralny*. Przewiduje się następujące etapy rozwoju środków wspomagających, a zwłaszcza formularza "elektronicznego":

1. *Pozyskiwanie słownych opisów procedur postępowania* wprowadzanych przez specjalistę w trybie tekstowym. Opisy te będą następnie przetwarzane przez programistę bazy wiedzy na odpowiadające im procedury, które będą mogły być wykonywane przez program sterujący systemu doradczego. Procedury te mogą być zapisane w odpowiednim języku "sztucznej inteligencji", jak np. *LISP*.

2. *Pozyskiwanie wiedzy reprezentowanej w sposób proceduralny przy użyciu schematów blokowych* zapisywanych przez specjalistę za pomocą odpowiedniego edytora graficznego. Przekształcenie w taki sposób pozyskanej wiedzy do postaci wykonywalnej przez program sterujący systemu doradczego będzie wymagało opracowania odpowiedniego kompilatora, dla którego wejściem będą dane reprezentujące schematy blokowe.

3. *Pozyskiwanie wiedzy reprezentowanej w sposób proceduralny poprzez stopniowe i systematyczne zwiększanie stopnia szczegółowości specyfikacji* (istota działania podobna do stosowanej w oprogramowaniu *CASE - Computer Aided Software Engineering*).

Nowym kierunkiem badań w diagnostyce maszyn będzie również *pozyskiwanie wiedzy diagnostycznej na podstawie sekwencji zdarzeń*. Przewiduje się jego zastosowanie do pozyskiwania wiedzy diagnostycznej dotyczącej układów dynamicznych (ewoluujących w czasie "makro").

Pozyskiwanie wiedzy indukcyjnymi metodami "uczenia maszynowego". Dokonano wyboru trzech indukcyjnych metod "uczenia maszynowego". Sformułowano wymagania dotyczące zbioru przykładów, sygnalizując trzy zasadnicze problemy, jak: *zapewnienie reprezentatywności zbioru przykładów, wybór cech relewantnych* oraz *dokonanie odpowiedniej kwantyzacji cech ilościowych*. Każdy z tych problemów wymaga podjęcia dalszych badań nad sposobami jego rozwiązania.

Zaproponowano schemat wnioskowania na podstawie ocen ufności reguł, odpowiadających poszczególnym koncepcjom (stanom). Określenie wartości stałych β_1, β_2, występujących w tym schemacie, może być przedmiotem optymalizacji. Wydaje się zasadne przyjęcie $0.7 < \beta_1 < 1$ oraz np. $0 < \beta_2 < 0.2$, co wymaga jednak zweryfikowania. Opisany schemat wnioskowania będzie stosowany nie tylko w diagnostyce; celowe jest jego zastosowanie w systemach doradczych wspomagających projektowanie i konstruowanie maszyn, zwłaszcza w odniesieniu do zadań rutynowych.

Należy wspomnieć o celowości kontynuacji rozpoczętych z inicjatywy autora [Moczulski, Żytkow, 1997] *badań nad zastosowaniem metod odkryć w bazach danych do pozyskiwania wiedzy diagnostycznej.* Celem tych badań winna być identyfikacja zależności ilościowych i jakościowych, zwłaszcza dotyczących relacji odwrotnych (*symptom* $\rightarrow$ *stan*). Kontynuacja badań wymaga rozwoju posiadanego środowiska sprzętowego oraz oprogramowania, a także współpracy z wybranymi ośrodkami zagranicznymi.

Nowym kierunkiem badań będzie opracowanie metod reprezentacji danych i wiedzy, dotyczących obiektów o własnościach zmieniających się w funkcji czasu "makro".

Metody oceny bazy wiedzy. Usystematyzowano zadania związane z oceną bazy wiedzy, ograniczając się do oceny bazy, w której wiedza jest reprezentowana za pomocą reguł wnioskowania oraz drzew decyzyjnych. Zaproponowano sposoby oceny wiedzy przez specjalistów oraz sposób łączenia opinii specjalistów, który winien być poddany dalszej weryfikacji.

Opisano kryteria *merytorycznej poprawności bazy wiedzy.* Zaproponowano wstępne wersje ocen błędów klasyfikacji, uwzględniające niezrównoważony rozkład zbioru przykładów (różne liczności podzbiorów przykładów dla poszczególnych klas). Badania w tym zakresie winny być kontynuowane w celu uzyskania ocen błędów klasyfikatora, uwzględniających prawdopodobieństwo wystąpienia przykładu należącego do danej klasy.

Przedstawiono sposób względnej oceny poprawności zbioru przykładów uczących, dokonywanej za pomocą zbioru reguł zweryfikowanych, do czego stosowane jest tzw. *uczenie przyrostowe.* Badania w tym zakresie będą kontynuowane.

Wreszcie zwrócono uwagę na możliwość "krzyżowego" stosowania sposobów oceny pozyskanej wiedzy, który może być użyty ze względu na reprezentację wiedzy w zunifikowany sposób.

Scenariusz procesu pozyskiwania wiedzy. Zaproponowany przez autora scenariusz procesu pozyskiwania wiedzy umożliwia połączenie wiedzy pochodzącej z różnych źródeł. Stwierdzono analogie pomiędzy procesem pozyskiwania wiedzy a procesem projektowo-konstrukcyjnym. Zwrócono uwagę na problemy zapewnienia jakości bazy wiedzy. Podkreślono rolę automatycznego pozyskiwania wiedzy z zastosowaniem indukcyjnych metod "uczenia maszynowego", co jest uzasadnione przez znacznie większą efektywność tych ostatnich metod.

Zaproponowany scenariusz procesu pozyskiwania wiedzy stanowił użyteczną podstawę do odpowiedniej organizacji badań proponowanych metod i środków pozyskiwania wiedzy, a także ułatwił opis tych badań na potrzeby rozprawy.

Rozdział 4

Środki wspomagania procesu pozyskiwania wiedzy diagnostycznej

W rozdziale opisano opracowane *środki wspomagania procesu pozyskiwania wiedzy diagnostycznej.* Zwrócono uwagę na integrujące znaczenie *wspólnego formatu reprezentacji danych i wiedzy.* Format ten stanowił podstawę do opracowania schematu logicznego *EMPREL.* Przedstawiono środki wspomagające *pozyskiwanie wiedzy z podstawowych jej źródeł* oraz *ocenę wiedzy już pozyskanej.* Opisano także utworzony *system pozyskiwania wiedzy diagnostycznej.*

4.1 Baza danych i wiedzy *EMPREL*

Szczegółowa analiza przedmiotu badań (zob. rozdział 2) pozwoliła autorowi stwierdzić, że:

1. *Opis obiektów badań winien następować w domkniętym świecie* (skończona liczba obiektów opisywanych za pomocą skończonej liczby cech, z których każda może przyjmować niewielką liczbę wartości).

2. Aby ograniczyć nadmiarowość baz danych i baz wiedzy, celowe jest wyróżnienie *właściwości grupowych* (charakterystycznych dla klas obiektów) oraz *indywidualnych*, traktowanych jako przypadki szczególne właściwości grupowych.

3. *Dane o obiektach (będące wartościami cech) mogą być reprezentowane w postaci stwierdzeń* (statycznych i dynamicznych). Pojawia się przy tym potrzeba jednoczesnego przechowywania *wartości cech jakościowych* i *ilościowych* (w wypadku cech, których wartości określane są przez proste układy pomiarowe, nie dokonujące kwantyzacji tych wartości).

4. *Przykłady stosowane do pozyskiwania wiedzy metodami indukcyjnymi mogą być reprezentowane za pomocą systemu informacyjnego* (który może być zapisany za pomocą zbioru stwierdzeń, określonych dla wartości cech warunku oraz cech decyzyjnych). W stwierdzeniach tych mogą występować wartości cech zarówno ilościowych, jak i jakościowych.

5. *Sposoby przekształcania wartości cech ilościowych w wartości cech jakościowych* są często uzależnione od *indywidualnych właściwości obiektów* lub *klas obiektów.*

6. *Wiedza operacyjna* (stosowana w procesie wnioskowania diagnostycznego) *może być reprezentowana za pomocą reguł ostrych i reguł przybliżonych* oraz *drzew decyzyjnych.*

7. Ponieważ warunki w przesłankach reguł oraz stwierdzenia stosowane do reprezentacji danych o obiektach są budowane z elementów składowych, pochodzących ze wspólnego opisu dziedziny zastosowania, *wysoce wskazane jest połączenie bazy danych i bazy wiedzy w jedną całość*, co ułatwi pielęgnację tej bazy podczas rozwoju i eksploatowania diagnostycznego systemu doradczego.

8. *Ponieważ istnieje możliwość częstego* (a nawet systematycznego) *uzupełniania zawartości bazy danych*, celowe jest zapewnienie *łatwego połączenia ze źródłami danych pochodzących z pomiaru i/lub symulacji.*

9. Ze względu na powyższe decyzje i uwarunkowania *wskazane jest do zapisu danych i wiedzy zastosowanie relacyjnej bazy danych*, będącej elementem systemu diagnostycznego eksploatowanego w środowisku sieci komputerowej.

Przedstawione rozumowanie (którego uzasadnienie zawarto w rozdziale 2) skłoniło autora do opracowania odpowiedniego *schematu logicznego* wspólnej bazy danych i wiedzy, zwanego *EMPREL*[1].

W bazie *EMPREL* można wyróżnić trzy główne podbazy:

1. *Opis dziedziny*, obejmujący:
   (a) słownik klas obiektów,
   (b) opis związków zachodzących pomiędzy klasami obiektów,
   (c) opis lokalizacji obiektów,
   (d) opis cech (ich nazw, rodzajów, typów i cech pochodnych oraz przekształceń cech "pierwotnych" w cechy pochodne),
   (e) słownik wartości cech jakościowych,
   (f) słownik stwierdzeń,
   (g) słownik stopni przekonania,
   (h) dane dotyczące osób (biorących udział w pracach nad bazą wiedzy) i ich adresów;

2. *Bazę danych*, zawierającą:
   (a) zapis zdarzeń pomiarowych i obserwacji,
   (b) wyniki pomiarów i/lub obliczeń symulacyjnych oraz wyniki ewentualnych obserwacji obiektów przez obsługę,
   (c) dane dotyczące konwersji wyników obserwacji w stwierdzenia (wraz ze szczegółami konwersji),

[1] W dalszym ciągu pracy baza ta będzie zwana w skrócie "bazą danych i wiedzy *EMPREL*" lub "bazą *EMPREL*".

   (d) bazę stwierdzeń dynamicznych, w której dokonywana jest specyficzna "kompresja" stwierdzeń (z zastosowaniem *sklejania stwierdzeń dynamicznych* — zob. s. 17);

3. *Bazę wiedzy* "operacyjnej" (która w obecnej wersji systemu pozyskiwania wiedzy ograniczona jest do *bazy reguł*), zawierającą:
   (a) określenie dziedzin reguł,
   (b) określenie reguł przybliżonych,
   (c) opis źródeł reguł (osoby, oprogramowanie),
   (d) oceny reguł (pochodzące przede wszystkim od niezależnych specjalistów).

Potrzeba reprezentacji reguł przybliżonych wynika z następujących przyczyn:

- w przesłankach reguł mogą występować stwierdzenia p, których stopień prawdziwości $NP(p)$ wg (2.65) spełnia $0 < NP(p) < 1$,
- mogą występować reguły, dla których stopień przekonania $NP(r)$ o słuszności danej reguły r, określony w komentarzu do zależności (3.95), spełnia $0 < NP(r) < 1$.

Część warunkowa reguły (przesłanka) określona jest przez koniunkcję selektorów, w których reprezentowane są warunki zbudowane dla wartości cech (lub zakresów wartości — w wypadku cech o zbiorze wartości uporządkowanym w sposób liniowy) z zastosowaniem operatorów relacyjnych $(=, \neq, <, \leqslant, >, \geqslant)$.

Ponadto w bazie reguł definiowane są dziedziny dla poszczególnych reguł (listy obiektów i klas tych obiektów).

Ogólny opis schematu logicznego *EMPREL* zawiera dodatek A. Opis bardziej szczegółowy zawiera raport [Moczulski *et al.*, 1997a].

4.2 Pozyskiwanie wiedzy od specjalistów

W przeprowadzonych badaniach istotną rolę odegrało pozyskiwanie wiedzy od specjalistów. Jego celem było zarówno pozyskanie nowej "porcji" wiedzy, jak i ocena wiedzy pozyskanej przez innych specjalistów (którzy nie byli źródłami tej "porcji" wiedzy).

Dla potrzeb wspomagania pozyskiwania nowej wiedzy opracowano:

- formularz "papierowy",
- formularz "elektroniczny",

natomiast dla potrzeb wspomagania oceny wiedzy pozyskanej przez inne osoby opracowano:

- formularz "papierowy" (raporty z bazy wiedzy),
- formularz "elektroniczny" (do oceny i łączenia opinii).

4.2.1 Formularz "papierowy" do pozyskiwania nowej wiedzy

Przy opracowywaniu *formularza papierowego* do pozyskiwania nowej wiedzy znalazła zastosowanie koncepcja *empirycznej relacji diagnostycznej* (zob. rozdział 3.1.1). Formularz "papierowy" do pozyskiwania nowej wiedzy przedstawiono np. na rys. C.2 (wraz z zapisaną przykładową empiryczną relacją diagnostyczną).

W omawianym formularzu występują następujące pola (por. rys. C.2):

1. nazwa obiektu;
2. lokalizacja obiektu;
3. faza eksploatacji;
4. warunki działania;
5. prawdopodobny stan techniczny;
6. podtabela dotycząca symptomów opisywanego stanu technicznego, zawierająca nazwę symptomu, jego wartość (ilościową lub jakościową) oraz wskazanie dotyczące częstości obserwowania takich symptomów;
7. uzasadnienie opisanej relacji;
8. cytowania i referencje.

Empiryczne relacje diagnostyczne, zapisane przez specjalistę za pomocą omawianego formularza "papierowego", wymagają odpowiedniej interpretacji przed ich wprowadzeniem do bazy wiedzy (zakłada się, że jest nią baza, która posiada schemat logiczny *EMPREL*). Działania takie wykonuje programista bazy wiedzy (zob. rozdział 3.1.2). Obejmują one identyfikację:

- nazw obiektów i przypisanie tych obiektów do odpowiednich klas (lub określenie nowych klas obiektów i ich powiązanie z istniejącą strukturą klas),
- cech fazy eksploatacji oraz ich wartości,
- cech opisujących warunki działania obiektów,
- cech stanów oraz wartości tych cech,
- cech sygnałów diagnostycznych,
- wartości krytycznych symptomów oraz ewentualnych przekształceń, jakim należy poddawać wartości cech ilościowych, aby otrzymać wartości jakościowe,
- uzasadnień, cytowań i referencji, które mogą być używane do objaśniania oraz formułowania odpowiedzi podczas konsultacji użytkownika z wynikowym systemem doradczym.

Podczas realizacji wszystkich powyższych działań programista bazy wiedzy powinien identyfikować *synonimy* oraz określać *nazwy główne*, które jako jedyne będą stosowane do zapisu reguł.

Wskazane jest także udostępnienie treści reguł, powstałych w wyniku interpretacji przez programistę bazy wiedzy, do oceny specjaliście, który zapisał wyjściową relację empiryczną.

4.2.2 Formularz "papierowy" do oceny wiedzy już pozyskanej

Jak wspomniano, najłatwiej jest skłonić specjalistę do oceny wiedzy już wcześniej pozyskanej (np. z innych źródeł). Aby taka ocena była skuteczna, winny być spełnione następujące kryteria:

1. Sposób reprezentacji wiedzy winien być zrozumiały dla oceniającego;
2. "Porcje" wiedzy przedstawianej do oceny winny być zapisane w języku jak najbardziej zbliżonym do języka naturalnego;
3. Zastosowana terminologia (nazwy obiektów, cech, wartości tych cech, stanów technicznych, warunków działania obiektów itp.) winna być zrozumiała dla oceniającego;
4. Nie należy stosować zbyt skomplikowanego sposobu oceniania — wskazane jest np. stosowanie ocen jakościowych o niewielu stopniach, co wpływa na zmniejszenie prawdopodobieństwa błędu oceny danej "porcji" wiedzy [Będkowski, 1991];
5. Ocenie winna również podlegać zrozumiałość danej "porcji" wiedzy.

Powyższy układ kryteriów był podstawą do opracowania rozwiązania opisanego poniżej. Wdrożenie tego rozwiązania wymagało przygotowania odpowiednich środków programowych opisanych w dalszej części rozdziału.

Istota rozwiązania

Ocenie poddaje się wiedzę reprezentowaną w postaci *reguł diagnostycznych*. Powinny to być reguły złożone, których konkluzjami są stwierdzenia dotyczące stanów technicznych. Chociaż reguły te często podlegają edycji przez programistę przed ich zapisaniem do bazy wiedzy, ich treść jest najbardziej zbliżona do treści zapisywanej przez specjalistę za pomocą empirycznych relacji diagnostycznych.

Poszczególne elementy składowe danej reguły są reprezentowane za pomocą *nazw głównych* każdego ze stosowanych pojęć (zob. s. 144). Nazwy te zostały zaproponowane przez autorów zapisów w bazie wiedzy (zwykle przez programistę bazy wiedzy) w fazie wprowadzania reguł. Należy podkreślić znaczący wpływ właściwego wyboru nazw głównych na zrozumiałość reguł przez oceniających te reguły specjalistów. Wpływ ten jest jeszcze bardziej istotny w odniesieniu do użytkowników systemu doradczego.

Każda ocena dotyczy jednej reguły. Ocena ta jest stwierdzeniem wyrażającym *subiektywny stopień przekonania o słuszności ocenianej reguły*. Stosowane są wartości stopnia przekonania określone w tablicy 3.1. Oceniający może także określić regułę jako *niezrozumiałą*, co pozwala na zidentyfikowanie zbioru takich reguł, które trzeba poddać reedycji.

Środki wspomagające ocenę wiedzy wcześniej pozyskanej

Działania, związane z ocenianiem wiedzy pozyskanej wcześniej, można podzielić na następujące fazy:

1. *Przygotowanie materiałów do oceny* (wydruk reguł, ankieta do rejestrowania ocen specjalisty/specjalistów).
2. *Właściwa ocena dokonana przez specjalistę* (działającego samodzielnie).

Nr reguły	całkowicie się zgadzam	prawie na pewno się zgadzam	raczej się zgadzam	nie mam zdania	raczej się nie zgadzam	prawie na pewno się nie zgadzam	zupełnie się nie zgadzam	nie rozumiem treści reguły	przesłanka nie pasuje do konkluzji
1									
2									
. . . . .									
N									

Rys. 4.1. Przykładowa ankieta oceny reguł
Fig. 4.1. An exemplary questionnaire for an assessment of rules

3. *Wprowadzenie ocen specjalisty do bazy wiedzy* przez programistę bazy wiedzy oraz agregacja opinii specjalisty z ocenami zapisanymi w bazie.

Do wspomagania działań wszystkich wymienionych powyżej uczestników procesu oceniania reguł opracowano następujące środki programowe, działające według koncepcji autora:

- *Generator raportów z bazy wiedzy*, umożliwiający zapis reguł dla wybranych dziedzin w sposób bardzo zbliżony do zapisu w języku naturalnym (fragment przykładowego raportu przedstawiono na rys. D.4).
- *Ankieta ocen reguł*, której kolejne wiersze zawierają identyfikatory ocenianych reguł, zaś kolumny odpowiadają wartościom oceny stopnia przekonania o słuszności reguł oraz ocenom zrozumiałości reguły i jej relewantności (przykładową ankietę przedstawiono na rys. 4.1). Określenie oceny danej reguły polega na wstawieniu w wierszu, zawierającym identyfikator ocenianej reguły, znaku $\surd$ w odpowiedniej kolumnie.
- *Specjalny edytor ocen reguł w bazie wiedzy*, ułatwiający programiście bazy wiedzy wprowadzanie ocen specjalisty/specjalistów (w postaci stopni przekonania o słuszności reguł) do bazy wiedzy oraz agregację opinii specjalistów z ocenami stopnia przekonania, już zawartymi w bazie.

4.2.3 Formularz "elektroniczny"

Formularz elektroniczny jest specjalizowaną aplikacją systemu baz danych, opracowaną według oryginalnej koncepcji autora [Moczulski, 1996]. Formularz ten winien ułatwiać specjaliście:

- *uświadomienie sobie* oraz *ujawnienie* posiadanej przez niego wiedzy dotyczącej relacji diagnostycznych,
- *identyfikację* oraz *reprezentację* niezbędnych pojęć wymaganych do zapisu relacji diagnostycznych (opisy tych pojęć utworzą opis dziedziny zastosowania),

- *reprezentację relacji diagnostycznych* w postaci deklaratywnej, z zastosowaniem przybliżonych reguł wnioskowania.

Ponadto formularz "elektroniczny" winien umożliwić:

- *ocenę merytoryczną wiedzy* wcześniej pozyskanej, która jest reprezentowana za pomocą zbioru reguł.

Kolejne wersje rozwijanej aplikacji wpierw dotyczyły oceny wiedzy wcześniej pozyskanej [Puzoń, Moczulski, 1996], [Moczulski, Wyleżoł, 1997], a następnie umożliwiały dodatkowo pozyskiwanie nowej wiedzy [Wyleżoł, 1997].

Przyjęta nazwa (*formularz elektroniczny*) nawiązuje do zastosowanego sposobu realizacji dialogu z użytkownikiem poprzez system ekranów mających postać formularzy.

W dalszej części niniejszego rozdziału zamieszczono jedynie najbardziej istotne opisy właściwości opracowanej aplikacji. Pełny opis zawarty jest w [Wyleżoł, 1997a].

Istota rozwiązania

Formularz "elektroniczny" jest aplikacją systemu baz danych, która współdziała z bazą *EMPREL* (zob. rozdział 4.1), zwaną w dalszym ciągu rozdziału "bazą roboczą (formularza)". Baza ta jest obsługiwana za pomocą techniki *ODBC* (ang. *Open Data Base Connectivity*) i jest niezbędna do działania formularza. Wyniki działania specjalisty w ramach każdej sesji są zapisywane do roboczej bazy wiedzy i mogą następnie podlegać dalszemu przetworzeniu (np. przez programistę bazy wiedzy).

Sposób posługiwania się formularzem przez specjalistę zależy od rodzaju działań, które będzie wykonywał, a mianowicie:

- wprowadzania nowej "porcji" wiedzy,
- oceny "porcji" wiedzy już pozyskanej.

Wprowadzanie nowej "porcji" wiedzy. Wprowadzanie nowej "porcji" wiedzy w postaci reguł odbywa się zgodnie z algorytmem opisanym na s. 66. Baza robocza formularza stosowanego do wprowadzania nowej "porcji" wiedzy musi zawierać jedynie podstawowe słowniki: typów i rodzajów cech, operatorów relacji porównania oraz stopni przekonania.

Reguły są reprezentowane w sposób zgodny ze schematem logicznym *EMPREL*. Każda reguła jest przypisywana do danego źródła, którym w tym wypadku jest jej autor. Definiowanie reguły obejmuje wybór *dziedziny*, określenie *przesłanki* (być może złożonej), *konkluzji* oraz *stopnia przekonania* autora reguły o jej prawdziwości. Elementy składowe reguły są budowane z nazw obiektów, cech, wartości cech, operatorów relacji porównania, wartości stopnia przekonania itp. Jeśli w słownikach bazy roboczej formularza występuje element składowy, następuje jego wybór z odpowiedniej listy (słownika). W każdej fazie definiowania reguły istnieje możliwość wprowadzenia brakującego pojęcia oraz jego wyboru jako składnika budowanej właśnie reguły.

Oprócz opisanego sposobu wprowadzania elementów słowników podczas zapisywania reguły (a więc wtedy, gdy te elementy są potrzebne), możliwe jest oddzielne wprowadzanie elementów poszczególnych słowników.

Ocena wcześniej pozyskanych reguł. Ten rodzaj działań wymaga, aby baza robocza formularza zawierała komplet zapisów umożliwiających wyświetlenie ocenianych reguł. Oprócz wspomnianych powyżej słowników, dotyczy to także co najmniej słowników: nazw klas, nazw obiektów, lokalizacji obiektów, nazw cech, nazw wartości cech.

Ocenianie reguł polega na przeglądaniu zawartości bazy wiedzy oraz przypisywaniu poszczególnym regułom ocen w postaci wartości stopnia przekonania o prawdziwości reguły. Oceny te wraz z identyfikatorem oceniającego są składowane w odpowiedniej tablicy, zawartej w roboczej bazie wiedzy. Podlegają one następnie agregacji z ocenami zawartymi już w bazie wiedzy. Wartości ocen agreguje programista bazy wiedzy (zob. s. 95).

Sposób prowadzenia dialogu z użytkownikiem

Dialog z użytkownikiem formularza elektronicznego prowadzony jest z wykorzystaniem udogodnień dostarczanych przez środowisko graficzne systemu *Microsoft Windows*. Jeśli to tylko jest możliwe, elementy składowe poszczególnych opisów dziedziny oraz reguł są wybierane z odpowiednich list lub słowników, przy czym badane jest zachowanie *więzów integralności* określonych w schemacie logicznym *EMPREL* (zob. rys. A.1 w dodatku A). Sposób prowadzenia dialogu z użytkownikiem przedstawiono na przykładach takiego dialogu, zamieszczonych w dodatku B.

Współdziałanie ze specjalistą

Formularz "elektroniczny" może być udostępniany specjaliście w formie plików na dyskietce lub np. poprzez *InterNet*. Robocze bazy wiedzy mogą być również udostępnione na dyskietce lub np. poprzez *InterNet*.

Robocze bazy wiedzy mogą być wstępnie ładowane pewnymi "porcjami" wiedzy wcześniej pozyskanej (od specjalistów lub metodami "uczenia maszynowego", co ma miejsce w wypadku oceny "krzyżowej" — por. rozdział 3.3.3). Pozwala to np. na przedstawienie specjaliście do oceny wiedzy pozyskanej wcześniej z kilku źródeł.

Zawartość roboczej bazy wiedzy, uzupełniona przez specjalistę posługującego się formularzem "elektronicznym", zostaje przesłana do ośrodka, w którym budowana jest baza wiedzy dla rozwijanego systemu doradczego. Wiedza nowo pozyskana przed jej dalszym stosowaniem zwykle będzie jeszcze wymagała edycji. Zadanie to realizuje programista bazy wiedzy, posługując się specjalnym edytorem.

4.3 Pozyskiwanie wiedzy w sposób indukcyjny

W opisywanych badaniach bardzo istotną rolę spełniały środki programowe wspomagające *pozyskiwanie wiedzy w sposób indukcyjny*. W środkach tych zastosowano metody tzw. "uczenia maszynowego", opisane w rozdziale 3.2.

Ze względu na kilkuletnią współpracę autora z ośrodkami, posiadającymi duży dorobek w zakresie rozwoju metod i oprogramowania do indukcyjnego pozyskiwania wiedzy, możliwe było zastosowanie oprogramowania udostępnionego przez te ośrodki[2]. Postępowanie takie wymagało jednak opracowania wielu programów pomocniczych. Opis tych programów nie jest przedmiotem rozprawy i jest dostępny w [Moczulski *et al.*, 1997a].

[2]Zob. rozdział "Przedmowa".

4.3.1 Przygotowanie bazy przykładów

Stosowane metody indukcyjne wymagały reprezentacji przykładów za pomocą tablic decyzyjnych (zob. rozdział 3.2.1). W opracowanym systemie pozyskiwania wiedzy diagnostycznej *SPWD1* przykłady te były reprezentowane za pomocą stwierdzeń. Przykłady gromadzono w bazie *EMPREL*.

Do weryfikacji opracowanych metod i środków pozyskiwania wiedzy metodami indukcyjnymi stosowano dwa rodzaje przykładów:

1. wyniki pomiarów prowadzonych w ramach *eksperymentu czynnego* na stanowisku modelowym *Rotor Kit*,
2. wyniki *eksperymentu numerycznego*.

Do przeprowadzenia pomiarów na stanowisku modelowym (punkt 1) zostały opracowane procedury umożliwiające pomiar sygnałów (przemieszczenia względne wału), gromadzenie zdyskretyzowanych wartości chwilowych tych sygnałów oraz estymację licznego zbioru ocen parametrów trajektorii ruchu względnego wału (opis w [Kostka, 1997], [Kostka, 1997a]).

W ramach kontynuacji tych badań (według koncepcji autora, [Moczulski, Kostka, 1997]) zmodyfikowano opisane wyżej procedury, znacznie zwiększając liczbę estymowanych cech sygnałów diagnostycznych. Dodano cechy estymowane na podstawie widma gęstości mocy (wartości skuteczne amplitud składowych o częstotliwościach: 0.25X[3], 0.33X, 0.5X, 1X, 2X, 3X) oraz widma wzajemnego (przede wszystkim przesunięcia fazowe pomiędzy wybranymi sygnałami przemieszczeń wału). Ponadto wprowadzono kryterium *ostrości dominanty widma*, umożliwiające identyfikację znaczących dominant (zob. przypis na s. 165).

Wymienione powyżej procedury pomiaru sygnałów diagnostycznych oraz estymacji cech tych sygnałów działają w środowisku *MatLab/SigLab*.

Wyniki eksperymentu numerycznego (punkt 2) uzyskano za pomocą udostępnionego oprogramowania symulacyjnego (niektóre programy systemu *MESWIR* [Kiciński *et al.*, 1995]). Do sterowania obliczeniami według planu eksperymentu numerycznego oraz do przetworzenia wyników symulacji na postać odpowiednią ze względu na klasyfikację trajektorii opracowano pakiet programów własnych [Moczulski, Maniak, 1997], [Maniak, Moczulski, 1997].

Ponadto opracowano pakiet programów i aplikacji umożliwiających kwantyzację wartości cech ilościowych oraz przygotowywanie danych uczących i testowych dla programów do pozyskiwania wiedzy metodami "uczenia maszynowego".

4.3.2 Wybór cech relewantnych

Do *wyboru cech relewantnych* z zastosowaniem algorytmu *PROMISE 2* (opis w rozdziale 3.2.4 na s. 78) zastosowano udostępniony program *AQ17-HCI*[4] [Wnęk, 1993].

[3]Zgodnie z zasadą przyjętą w publikacjach amerykańskich zapis nX oznacza częstotliwość równą n-tej krotności częstotliwości obrotów wału maszyny.

[4]Spośród wielu możliwości tego programu wykorzystywano jedynie wybór cech relewantnych.

4.3.3 Selektywna indukcja reguł

Do *selektywnej indukcji reguł* poprzez generowanie pokryć (według metody opisanej w rozdziale 3.2.4) zastosowano udostępniony program *AQ15c* [Wnęk *et al.*, 1995].

4.3.4 Indukcja reguł z zastosowaniem zbiorów przybliżonych

Do *indukcji reguł z zastosowaniem zbiorów przybliżonych* (według metody opisanej w rozdziale 3.2.5) użyto udostępnionego programu *LERS* [Grzymała-Busse, 1994]. Indukcja reguł była każdorazowo poprzedzana przez obliczanie reduktów zbioru cech warunku. Ponadto niektóre obliczenia z zastosowaniem przykładów przygotowanych według koncepcji autora rozprawy zostały wykonane przez J. W. Grzymałę-Busse, autora programu *LERS*.

4.3.5 Indukcja drzew decyzyjnych

Do *indukcji drzew decyzyjnych* (wg metody opisanej w rozdziale 3.2.6) zastosowano program *C4.5* (opis w [Quinlan, 1993]).

4.3.6 Ocena merytorycznej poprawności bazy wiedzy

Wszystkie stosowane programy do indukcji reguł umożliwiają wyznaczenie *oceny poprawności baz wiedzy* co najmniej za pomocą łącznego względnego błędu empirycznego ϵ_{ov} wg (3.96). Ponadto opracowano niezbędne programy wspomagające stosowanie podstawowych technik oceny sprawności klasyfikatora (opis w rozdziale 3.3.2). W ramach badania sprawności klasyfikatora wyznaczano ocenę wspomnianego łącznego błędu empirycznego oraz względnych empirycznych błędów: pominięcia ϵ_{om} wg (3.97) i niesłusznego zaliczenia ϵ_{cm} wg (3.98).

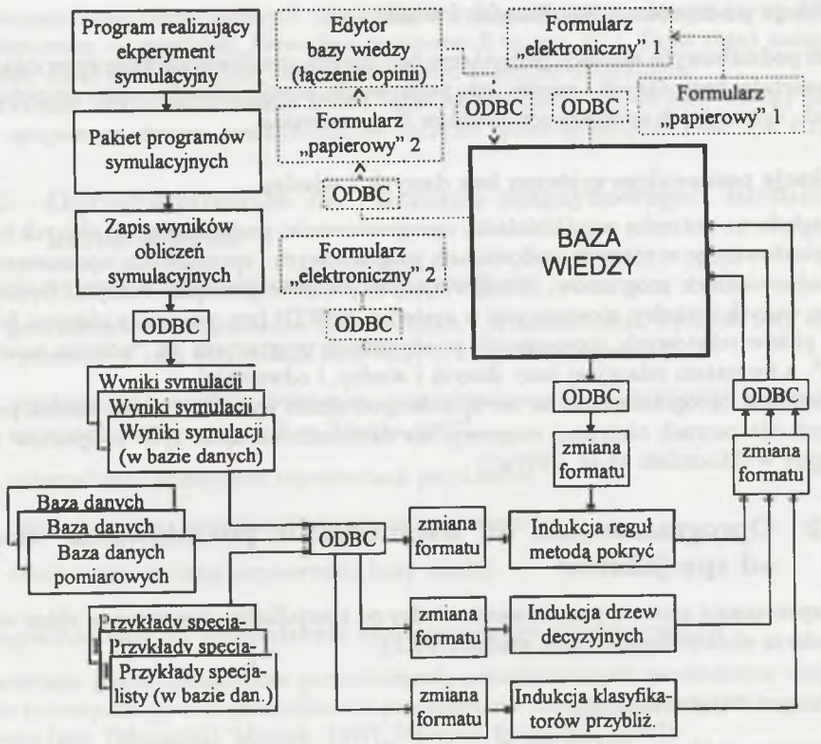
4.4 System pozyskiwania wiedzy diagnostycznej

Opracowane i wdrożone środki wspomagające przebieg procesu pozyskiwania wiedzy zostały zintegrowane w *system pozyskiwania wiedzy diagnostycznej SPWD1* (zob. rys. 4.2). Zamieszczenie pełnego opisu tego systemu, który jest jest zawarty w [Moczulski, 1997c], [Moczulski *et al.*, 1997a] nie jest wymagane w pracy. Pominięto zwłaszcza szczegółowe opisy środowiska sprzętowego i programowego systemu *SPWD1*.

4.4.1 Podsystem baz danych i wiedzy systemu *SPWD1*

Centralnym elementem systemu SPWD1 jest podsystem baz danych i wiedzy[5], na który składają się: baza danych i wiedzy oraz pakiet aplikacji do obsługi tej bazy i aplikacji pomocniczych.

[5]Baza danych i wiedzy stanowi jedną całość (zob. opis w rozdziale 4.1). Na rys. 4.2 została ona rozbita na bazę danych i bazę wiedzy dla zapewnienia większej przejrzystości tego rysunku.



Rys. 4.2. Opracowany system pozyskiwania wiedzy diagnostycznej *SPWD1*
Fig. 4.2. The developed *SPWD1* diagnostic knowledge acquisition system

Baza danych i wiedzy

Dane dotyczące obiektów diagnozowania i "porcje" wiedzy diagnostycznej zapisywane są w relacyjnej bazie danych i wiedzy *EMPREL* (por. rozdział 4.1). Możliwe jest przechowywanie danych, dotyczących licznych klas obiektów należących do wielu dziedzin zastosowań.

Dzięki wprowadzonej strukturze klas obiektów można znacznie ograniczyć nadmiarowość zapisów. Wprowadzenie stwierdzeń dynamicznych do reprezentacji obiektów o cechach ewoluujących w czasie umożliwia gromadzenie danych, które w przyszłości będą wykorzystane do pozyskiwania wiedzy na podstawie sekwencji zdarzeń.

Odpowiednio zdefiniowany schemat logiczny, sposób zapisu danych i wiedzy oraz pakiet aplikacji o charakterze pomocniczym stwarzają możliwość pełnej wymienności danych i wyników (np. pozyskanych reguł diagnostycznych) pomiędzy poszczególnymi aplikacjami podstawowymi systemu *SPWD1*.

Aplikacje podsystemu baz danych i wiedzy

Pakiet podstawowych aplikacji podsystemu baz danych umożliwia najważniejsze działania na zawartości bazy danych i wiedzy, jak: zapisywanie, edycję, odszukiwanie i uzupełnianie danych, dotyczących opisywanych obiektów diagnozowania.

Aplikacje pomocnicze systemu baz danych i wiedzy

Ze względu na potrzebę współdziałania oprogramowania, pochodzącego z różnych źródeł i eksploatowanego w różnych środowiskach programowych i sprzętowych, opracowano pakiet odpowiednich programów, umożliwiających przejście pomiędzy różnymi formatami zapisu danych i wiedzy, stosowanymi w systemie *SPWD1* (np. pomiędzy różnymi formatami plików tekstowych, stosowanymi w używanych programach do "uczenia maszynowego", a formatem relacyjnej bazy danych i wiedzy, i odwrotnie).

Ponieważ oprogramowanie to nie było bezpośrednim wynikiem badań autora prowadzonych dla potrzeb niniejszej rozprawy, nie zamieszczono opisu tych programów (opis dostępny w [Moczulski *et al.*, 1997a]).

4.4.2 Oprogramowanie do wspomagania pozyskiwania wiedzy od specjalistów

Do wspomagania procesu pozyskiwania wiedzy od specjalistów stosowane są różne wersje *formularza elektronicznego* (zob. rozdział 4.2.3).

Formularz "elektroniczny"

Opracowana aplikacja *EMPREG 2*[6] może być stosowana do pozyskiwania nowej "porcji" wiedzy (prostokąt *Formularz elektroniczny 1* na rys. 4.2) oraz do oceny wiedzy już pozyskanej (prostokąt *Formularz elektroniczny 2* na rys. 4.2).

Formularz "elektroniczny" może być także stosowany przez programistę bazy wiedzy jako środek pomocniczy do wprowadzania empirycznych relacji diagnostycznych, zapisanych przez specjalistów na *formularzach papierowych do pozyskiwania nowej wiedzy* (zob. rozdział 4.2.1; prostokąt *Formularz papierowy 1* na rys. 4.2). Empiryczne relacje diagnostyczne zapisane przez specjalistę zwykle wymagają wówczas interpretacji przez programistę bazy wiedzy.

Pobieranie zawartości tabel do roboczej bazy wiedzy formularza ("porcja" wiedzy, która ma być przedmiotem oceny, elementy opisu dziedziny itp.) odbywa się za pomocą oddzielnej aplikacji. Także wyniki pozyskiwania wiedzy pobierane są z bazy roboczej do bazy głównej za pomocą tej aplikacji.

Przygotowanie materiałów do oceny "porcji" wiedzy za pomocą formularza "papierowego"

Specjalna aplikacja generuje rodzaj *raportu z bazy wiedzy*, zawierającego reguły przeznaczone do oceny przez specjalistów za pomocą formularza "papierowego" oraz *ankietę do*

[6] Aplikację *EMPREG 2*, wg koncepcji autora rozprawy, opracował M. Wyleżoł [Wyleżoł, 1997a].

umieszczania ocen poszczególnych reguł (dla których zastosowano łączną nazwę *formularz papierowy* — prostokąt *Formularz papierowy 2* na rys. 4.2). Zapis reguł następuje w sposób możliwie najbardziej zbliżony do zapisu w języku naturalnym.

Zawartość ankiet do oceniania zbioru reguł wprowadzana jest do bazy wiedzy za pomocą *specjalnego edytora*, umożliwiającego *łączenie opinii specjalistów* (zob. rys. 4.2).

4.4.3 Oprogramowanie do "uczenia maszynowego" metodami indukcyjnymi

Ta część systemu *SPWD1* obejmuje oprogramowanie do:

- prowadzenia eksperymentów numerycznych i przekształcania wyników tych eksperymentów w celu otrzymania bazy danych z wynikami symulacji,
- przetwarzania wyników pomiarów, przeprowadzonych na stanowisku modelowym w ramach badań opisanych w [Kostka, 1997],
- optymalizacji przestrzeni reprezentacji przykładów,
- indukcji wiedzy metodami "uczenia maszynowego",
- oceny merytorycznej poprawności bazy wiedzy.

Oprogramowanie do prowadzenia eksperymentów numerycznych

Opracowano pakiet programów pomocniczych, umożliwiających prowadzenie eksperymentu numerycznego oraz gromadzenie i przetwarzanie danych, uzyskanych z tego eksperymentu (zob. [Moczulski, Maniak, 1997], [Maniak, Moczulski, 1997]).

Do właściwych obliczeń symulacyjnych zastosowano udostępniony system programów *MESWIR*, z którego do obliczeń stosowano programy (opis wg [Kiciński *et al.*, 1995]):

- *KINWIR-D*, realizujący obliczenia z zakresu kinetostatyki układu "wirnik–podpory łożyskowe" przy uwzględnieniu silnej nieliniowości charakterystyk układu. Ich wynikiem są współczynniki sztywności i tłumienia filmu smarnego, odpowiadające danej prędkości obrotowej i położeniu równowagi czopa w łożysku;
- *DYNWIR-W*, realizujący właściwe obliczenia dynamiczne w zakresie liniowym (przy założeniu że drgania wymuszone czopów w łożyskach są na tyle małe, że można przyjąć stałe wartości współczynników sztywności i tłumienia w analizowanym obszarze koła luzów). Dla zadanych wartości i położeń niewywag program oblicza parametry trajektorii ruchu bezwzględnego czopa oraz ruchu czopa względem czujnika przemieszczeń wału (trajektorie te ze względu na liniowość przyjętego modelu są elipsami).

Oprogramowanie do przetwarzania wyników pomiarów przeprowadzonych na stanowisku modelowym

Dla potrzeb badań związanych z rozprawą zmodyfikowano wg koncepcji autora (zob. opis w rozdziale 4.3.1) oprogramowanie opisane w [Kostka, 1997]. Ponadto opracowano aplikacje pomocnicze, umożliwiające optymalizację zbioru cech sygnałów, wybór optymalnego sposobu kwantyzacji wartości cech ilościowych i dokonanie tej kwantyzacji.

Oprogramowanie do indukcji wiedzy metodami "uczenia maszynowego"

Do systemu *SPWD1* zostały włączone programy do pozyskiwania wiedzy metodami indukcyjnymi na podstawie zbioru przykładów (opis programów w rozdziałach 4.3.3÷4.3.5).

Oprogramowanie do oceny merytorycznej poprawności bazy wiedzy

W celu usprawnienia weryfikacji wiedzy pozyskanej metodami "uczenia maszynowego" opracowano pakiet programów pomocniczych do oceny sprawności klasyfikatorów (zob. rozdział 4.3.6). Należy podkreślić, że wspólna reprezentacja reguł w systemie *SPWD1* umożliwia zastosowanie tego pakietu do merytorycznej oceny reguł pozyskanych bezpośrednio od specjalistów.

4.5 Podsumowanie dotyczące środków wspomagania procesu pozyskiwania wiedzy

W rozdziale opisano środki wspomagania procesu pozyskiwania wiedzy opracowane według koncepcji autora. Centralną rolę w tym procesie spełnia *wspólna baza danych i wiedzy*, która jest zorganizowana zgodnie ze schematem logicznym *EMPREL*. Dalszy rozwój bazy danych i wiedzy winien następować w kierunku jej zgodności ze schematem *MIMOSA*, co umożliwi uzyskanie dostępu do źródeł danych dotyczących eksploatacji maszyn (gromadzonych w sposób automatyczny w bazach danych systemów nadzoru diagnostycznego lub wprowadzanych do tych baz przez personel zajmujący się eksploatacją maszyn).

Rozwój środków wspomagania procesu pozyskiwania wiedzy od specjalistów winien dotyczyć przede wszystkim *pozyskiwania wiedzy reprezentowanej w sposób proceduralny*, co wymaga opracowania lub zastosowania odpowiedniego sposobu reprezentacji.

W zakresie indukcyjnego pozyskiwania wiedzy na podstawie przykładów wymagany jest *rozwój oprogramowania wspomagającego planowanie i prowadzenie eksperymentów numerycznych*. Przygotowanie przykładów do uczenia maszynowego wymaga opracowania oprogramowania wspomagającego estymację wartości licznego zbioru cech, kwantyzację cech ilościowych oraz wybór cech relewantnych. Wymagany jest także rozwój oprogramowania wspomagającego prowadzenie obliczeń dla złożonej struktury drzewa stanów technicznych.

Nowym kierunkiem badań będzie rozwój środków umożliwiających reprezentację danych, opisujących własności obiektów zmieniające się w funkcji czasu "makro". Dane takie znajdą zastosowanie do pozyskiwania wiedzy z sekwencji zdarzeń i stanów.

Rozdział 5

Badania proponowanych metod i środków pozyskiwania wiedzy

W rozdziale opisano wybrane zadania zrealizowane w związku z oceną przydatności zastosowanych metod i opracowanych środków do pozyskiwania wiedzy diagnostycznej oraz weryfikacją metod, opracowanych przez autora. Badania te przebiegały według *scenariusza procesu pozyskiwania wiedzy* przedstawionego w rozdziale 3.4. Przedstawiono wyniki badań dotyczących pozyskiwania wiedzy dla potrzeb diagnozowania stanu wirnika w odniesieniu do dwóch różnych obiektów:

1. modelu maszyny wirnikowej *Rotor Kit* [Bently Nevada, 1996], obserwowanego w ramach *czynnego eksperymentu diagnostycznego*;
2. modelu maszyny wirnikowej, zainstalowanego na stanowisku badawczym Instytutu Maszyn Przepływowych PAN w Gdańsku, który był obiektem *eksperymentu numerycznego*.

Badania te dotyczyły *skuteczności zastosowanych metod i środków pozyskiwania wiedzy diagnostycznej w odniesieniu do wybranych problemów dynamiki ruchu wirników łożyskowanych ślizgowo*. Tak określona dziedzina zastosowań obejmuje zagadnienia diagnozowania stanu wirnika, które mają kluczowe znaczenie w diagnostyce maszyn wirnikowych [Cempel, 1982], [Łączkowski, 1983], [Mitchell, 1981]. Zdaniem autora, pozytywny wynik weryfikacji opracowanych metod i środków pozyskiwania wiedzy diagnostycznej winien stanowić podstawę do uogólnienia otrzymanych wniosków na inne problemy związane z diagnozowaniem układu *wirnik – łożyska – podpory łożyskowe – fundament*. Ponadto celem badań była *weryfikacja metod pozyskiwania wiedzy, opracowanych przez autora*, oraz *środków wspomagania procesu pozyskiwania wiedzy*, opracowanych przez niego lub na podstawie jego koncepcji.

5.1 Badania dotyczące wyników czynnego eksperymentu diagnostycznego

Opisana część badań dotyczyła dziedziny zastosowań związanej z diagnozowaniem stanu wirnika modelu maszyny wirnikowej *Rotor Kit* [Bently Nevada, 1996]. Model ten zainstalowano w laboratorium Katedry Podstaw Konstrukcji Maszyn Politechniki Śląskiej.

Szczególną zaletą obiektu badań była możliwość przeprowadzenia na nim *czynnego eksperymentu diagnostycznego*.

W niniejszym rozdziale dokonano podsumowania najważniejszych wyników tych badań, którym w dalszym ciągu rozprawy przypisany będzie symbol *RK*. Obszerny opis badań zawarto w dodatku C.

5.1.1 Charakterystyka zadania diagnostycznego

Zadanie diagnostyczne polegało na *diagnozowaniu stanu wirnika modelu Rotor Kit*. Postanowiono rozpatrywać takie niesprawności, jak: *różne postaci i wielkości niewyrównoważenia wirnika*, *punktowe przycieranie wału* oraz *przeciążenie łożysk wirnika*, dopuszczając jednoczesne występowanie więcej niż jednej niesprawności. Nośnikiem informacji o stanie wirnika były wielowymiarowe sygnały przemieszczeń względnych wału.

Autor posłużył się w swoich badaniach wynikami pomiarów zgromadzonymi przez P. Kostkę [Kostka, 1997]. Złożony schemat eksperymentu, w którym obserwowano odpowiedź modelu na kombinację kilku czynników, skłonił autora do zastosowania nowego podejścia, którego istotą jest dekompozycja zbioru przykładów na podzbiory odpowiadające złożonej strukturze drzewa stanów (zob. rozdział 3.2.2).

5.1.2 Przebieg procesu pozyskiwania wiedzy

Proces pozyskiwania wiedzy obejmował zadania *pozyskiwania wiedzy od specjalistów* oraz *indukcyjnego pozyskiwania wiedzy ze zbioru przykładów* zgromadzonych w ramach obserwacji obiektu poddanego czynnemu eksperymentowi diagnostycznemu wraz z *merytoryczną oceną wiedzy pozyskanej* (178 przykładów zaobserwowanych dla 7 różnych kombinacji stanów elementarnych).

Pozyskanie wiedzy od specjalistów

Do pozyskania wiedzy od specjalistów, reprezentowanej za pomocą *empirycznych relacji diagnostycznych*, zastosowano formularz "papierowy". Pozyskane relacje zostały wprowadzone do bazy danych i wiedzy *EMPREL* za pomocą formularza "elektronicznego".

Indukcyjne pozyskiwanie wiedzy na podstawie przykładów

Stosowane metody indukcyjne wymagają reprezentacji zbioru danych w postaci *tablicy decyzyjnej* (zob. rozdz. 3.2.1), co wiąże się z potrzebą kwantyzacji cech ilościowych. Dla cech ilościowych, będących ocenami amplitud przemieszczeń względnych, dokonano *optymalnego wyboru progów kwantyzacji* ze względu na kryterium największej sprawności klasyfikatora. Przyjęte niewielkie liczby poziomów kwantyzacji (zob. opis na s. 161) *ułatwiają przypisanie wartości lingwistycznych* poszczególnym wartościom cech pochodnych jakościowych.

W wyniku wstępnej *indukcji drzew decyzyjnych* dla wybranego podzbioru cech relewantnych uzyskano odpowiadające tym drzewom klasyfikatory o niskiej sprawności (średnio 65%, zob. tabl. C.5). Podobne wyniki uzyskano podczas *indukcji reguł z zastosowaniem metody pokryć* oraz *metody zbiorów przybliżonych*.

W związku z powyższym autor dokonał *optymalizacji struktury drzewa stanów* (zob. s. 169 i n.) oraz pozyskał indukcyjnie drzewa decyzyjne zgodnie ze strukturą drzewa

stanów, przy czym dla każdego poddrzewa wybierano *podzbiór cech relewantnych*. Dla większości rozpoznawanych stanów złożonych uzyskano znaczne zwiększenie sprawności klasyfikatora (zob. tabl. C.8, C.9).

5.1.3 Podsumowanie badań przeprowadzonych dla bazy wiedzy *RK*

Badania przeprowadzone dla bazy wiedzy *RK* stanowiły podstawę do sformułowania następujących spostrzeżeń (których uzasadnienie zawiera dodatek C):

1. *Przy rowiniętym planie eksperymentu czynnego uzyskuje się niską sprawność klasyfikatorów*, jeśli nie zastosuje się specjalnego podejścia.
2. *Dekompozycja zbioru przykładów na podzbiory odpowiadające złożonej strukturze drzewa stanów umożliwia budowę klasyfikatora o wyższej sprawności* (zwłaszcza w odniesieniu do wybranych poddrzew całej struktury drzewa stanów).
3. *Wybór odpowiedniej struktury drzewa stanów* ma wpływ na wartość oceny błędu klasyfikacji.
4. *Największy błąd klasyfikacji związany jest z identyfikacją występowania przeciążenia* (zob. oceny błędu dla poddrzewa 2 oraz poddrzewa 4.2 w tabl. C.8 i tabl. C.9). Może to wskazywać na zastosowanie podczas eksperymentu zbyt małego przeciążenia, które nie powoduje objawów drganiowych typowych dla tej niesprawności.
5. *Nie stwierdzono istotnych rozbieżności pomiędzy wartościami oceny sprawności klasyfikatorów, uzyskanych z zastosowaniem różnych metod indukcyjnych* (dla badań wstępnych).

Uzyskane wyniki weryfikacji potwierdzają poprawność zastosowanych metod pozyskiwania wiedzy w odniesieniu do obiektu konkretnego, którego obserwacja jest wspomagana przez autonomiczny system gromadzenia wyników pomiarów diagnostycznych.

5.2 Badania przeprowadzone w ramach eksperymentu numerycznego

Badania dotyczyły modelu maszyny wirnikowej zainstalowanego na stanowisku badawczym Instytutu Maszyn Przepływowych PAN w Gdańsku. W celu opisu własności dynamicznych tego modelu w IMP PAN opracowano model matematyczny [Kiciński, 1996a] oraz określono podstawowe własności (jak współczynniki sztywności i tłumienia filmu olejowego, współczynniki sztywności i tłumienia utwierdzeń zewnętrznych panwi itp.) [Kiciński *et al.*, 1995], co łącznie stanowiło podstawę do przeprowadzenia obliczeń symulacyjnych. Autor miał możliwość prowadzenia *eksperymentu numerycznego* za pomocą systemu programów *MESWIR*, który pozwalał na dokonywanie obliczeń symulacyjnych drgań wymuszonych układu, spowodowanych zadanym rozkładem niewyrównoważenia wirnika wirującego z określoną prędkością obrotową (obliczenia prowadzono dla trzech różnych prędkości).

W niniejszym rozdziale dokonano podsumowania najważniejszych wyników tych badań, którym w dalszym ciągu rozprawy przypisany będzie symbol *GD*. Obszerny opis badań zawarto w dodatku D.

5.2.1 Charakterystyka zadania diagnostycznego

Zadanie diagnostyczne polegało na *identyfikacji stanu wyrównoważenia wirnika*, podpartego w dwóch łożyskach hydrodynamicznych i wirującego z jedną z prędkości obrotowych: podrezonansową, rezonansową lub ponadrezonansową. Rozpatrywano *pięć różnych stanów*: zgrubne wyrównoważenie dynamiczne oraz cztery podstawowe stany niewyrównoważenia [Łączkowski, 1979]: statyczne, quasi-statyczne, momentowe i dynamiczne. Wymienione stany należało rozpoznać na podstawie symptomów diagnostycznych, którymi były *wartości cech trajektorii drgań bezwzględnych i drgań względnych*, wyznaczonych dla obu czopów wału. Dla potrzeb opisanych badań cechy te wyznaczane były przez odpowiednie oprogramowanie, umożliwiające prowadzenie obliczeń symulacyjnych.

5.2.2 Przebieg procesu pozyskiwania wiedzy

Postanowiono pozyskać wiedzę zarówno od specjalistów (pośrednio — na podstawie opisów dostępnych w literaturze), jak i metodami "uczenia maszynowego" na podstawie zbioru przykładów. W ramach prowadzonego *eksperymentu numerycznego* przygotowano liczny zbiór przykładów (5076 przykładów o zbliżonej liczności podzbiorów tych przykładów dla każdego z rozpatrywanych stanów). Należy nadmienić, że w stosowanym zbiorze przykładów *nie występują stany złożone*, gdyż rozpatrywane stany wzajemnie się wykluczają.

Pozyskanie wiedzy na podstawie dostępnej literatury

Za pomocą formularza "elektronicznego" zapisano *relacje diagnostyczne*, które zidentyfikowano w dostępnej literaturze diagnostycznej (np. [Cempel, 1989], [Kiciński, 1996a], [Łączkowski, 1979], [Łączkowski, 1983], [Mitchell, 1981], [Morel, 1992], [Solartron, 1995], [Solartron, 1995a]). Relacje te miały postać *reguł wnioskowania* i dotyczyły *ogólnej wiedzy o sposobach identyfikacji stanu niewyrównoważenia wirnika*. Reguły te zostały poddane ocenie przez innych specjalistów, dokonanej za pomocą formularza "papierowego" do oceny reguł. Oceny reguł wprowadzono do bazy wiedzy.

Indukcyjne pozyskiwanie wiedzy na podstawie przykładów

Ta część badań wymagała *przygotowania zbioru przykładów uczących* oraz zastosowania *wybranych indukcyjnych metod uczenia maszynowego* do pozyskania wiedzy. Ponadto dokonano *weryfikacji możliwości łączenia wiedzy pozyskanej z różnych źródeł* (od specjalistów oraz w wyniku "uczenia maszynowego").

Przygotowanie zbioru przykładów. Wiele uwagi poświęcono odpowiedniemu przygotowaniu zbioru przykładów uczących. Dla potrzeb "uczenia maszynowego" postanowiono zastosować zbiór przykładów przygotowany za pomocą udostępnionego oprogramowania symulacyjnego. Rozwiązano pomyślnie zadania *wyboru zbioru symptomów diagnostycznych*, którymi są wartości cech trajektorii drgań względnych i bezwzględnych czopów

wirnika. Szczególne znaczenie miał *odpowiedni dobór sposobów kwantyzacji wartości cech jakościowych*, określających wymiary trajektorii, a także *wybór wartości progów kwantyzacji*, przy którym uwzględniano wpływ warunków działania obiektu (reprezentowanych przez prędkość obrotową — zob. opis na s. 184 i n.). W obu przypadkach dokonano wyboru spośród kilku rozwiązań, stosując kryterium *największej sprawności klasyfikatora*. W wyniku tego wyboru przyjęto *logarytmiczną skalę długości półosi elips z pięcioma poziomami kwantyzacji* (zob. tabl. D.5) oraz *liniową skalę wartości kątowych*. Mała liczba wartości dyskretnych ułatwia przypisanie im wartości lingwistycznych. W ten sposób osiągnięto jednocześnie dwie korzyści: *łatwość przedstawienia w języku naturalnym reguł pozyskanych w sposób automatyczny* (aby były zrozumiałe dla użytkownika systemu doradczego) oraz *pełną wymienność reguł pozyskanych od specjalistów z regułami pozyskanymi metodami uczenia maszynowego* (zarówno w zakresie jej stosowania, jak i oceniania).

Pozyskanie wiedzy na podstawie przykładów. Badania przeprowadzono w trzech seriach: serii BW, w której uwzględniano wszystkie cechy (zarówno trajektorii drgań bezwzględnych, jak i względnych), serii B (tylko cechy drgań bezwzględnych) oraz serii W (tylko cechy drgań względnych). Zarówno *indukcja drzew decyzyjnych*, jak i *indukcja reguł* (metodą pokryć i metodą zbiorów przybliżonych) pozwoliły na wyznaczenie klasyfikatorów o wysokiej sprawności (lepszej niż 90%). Najlepsze wyniki uzyskano w wypadku uwzględnienia pełnego zbioru cech (seria BW). Bardziej szczegółowe zestawienie i omówienie uzyskanych wyników zawiera dodatek D.

Połączenie wiedzy pozyskanej z różnych źródeł. Zastosowano *metodę uczenia przyrostowego* z regułami pozyskanymi od specjalisty jako *hipotezami wstępnymi*. Reguły te *modyfikowano za pomocą zbioru przykładów uczących*. Testowanie pozyskanych w ten sposób reguł odbyło się metodą *Hold-out* oraz za pomocą *specjalnie przygotowanego oddzielnego zbioru przykładów testowych*. Przeprowadzone badanie (opis na s. 188) potwierdziło możliwość zastosowania takiego sposobu postępowania w celu *modyfikacji* (uogólnienia i/lub uszczegółowienia) *reguł pozyskanych od specjalistów*. Metoda ta umożliwia *znaczne zwiększenie efektywności procesu pozyskiwania wiedzy diagnostycznej*.

5.2.3 Podsumowanie badań przeprowadzonych dla bazy wiedzy *GD*

Badania przeprowadzone dla bazy wiedzy *GD* dotyczyły stanowiska badawczego (modelu maszyny wirnikowej), którego wybrane cechy były modelowane za pomocą systemu programów symulacyjnych. Otrzymane wyniki badań pozwalają na sformułowanie następujących wniosków (których pełne uzasadnienie zawiera dodatek D):

1. *Eksperyment numeryczny został zaplanowany we właściwy sposób*, co umożliwiło wygenerowanie bazy przykładów, która stanowiła podstawę do pozyskania wiedzy umożliwiającej budowę klasyfikatorów o wysokiej sprawności η_{ov}.

2. Porównanie wyników testowania sprawności klasyfikatorów (uzyskanych poprzez indukcję reguł metodą generowania pokryć) metodą *Hold-out* oraz za pomocą odrębnego zbioru przykładów testowych (zob. tabl. D.6) wskazuje na *nadmierne dopasowanie klasyfikatora do danych*. Wydaje się celowe generowanie reguł bardziej

ogólnych, dla których nie powinny występować znaczące różnice sprawności klasyfikatora w razie badania go za pomocą nowego zbioru przykładów testowych.

3. *Sprawność klasyfikatora zależy od zastosowanego podzbioru cech warunku* i jest najwyższa wówczas, gdy są uwzględniane cechy opisujące zarówno trajektorie drgań bezwzględnych, jak i trajektorie drgań względnych.

4. Na sprawność klasyfikatora znaczący wpływ ma *sposób kwantyzacji tych cech o wartościach ilościowych, które opisują rozmiary trajektorii drgań.* W przeprowadzonych badaniach znacznie wyższe sprawności klasyfikatorów otrzymano w wypadku kwantyzacji, której dokonano z zastosowaniem progów określonych w skali logarytmicznej.

5. Przyjęte liczby poziomów kwantyzacji *ułatwiają przypisanie poszczególnym wartościom jakościowym takich nazw, które mogą być zaakceptowane przez specjalistę.* Umożliwia to połączenie wiedzy pochodzącej od specjalistów z wiedzą pozyskaną metodami indukcyjnymi.

6. Wstępnie potwierdzono *przydatność uczenia przyrostowego* (zob. opis na ss. 188 i n.) do łączenia wiedzy pochodzącej od specjalistów z wiedzą pozyskaną indukcyjnie.

Otrzymane wyniki weryfikacji w pełni potwierdzają możliwość pozyskiwania wiedzy na podstawie zbioru przykładów, zgromadzonych w sposób systematyczny w wyniku odpowiednio zaplanowanego *eksperymentu numerycznego.*

5.3 Weryfikacja systemu pozyskiwania wiedzy diagnostycznej *SPWD1*

Celem badań weryfikacyjnych w odniesieniu do systemu pozyskiwania wiedzy diagnostycznej *SPWD1* była ocena prawidłowości współdziałania poszczególnych elementów tego systemu. Badania te zostały zaplanowane kompleksowo, obejmując weryfikację:

- poprawności działania bazy danych i wiedzy *EMPREL*;
- możliwości odpowiedniego wybierania zestawów danych uczących oraz ich eksportowania w formacie odpowiednim dla stosowanego oprogramowania do indukcyjnego "uczenia maszynowego";
- możliwości importowania reguł zapisanych w formacie stosowanym w poszczególnych programach "uczenia maszynowego";
- możliwości wyboru reguł ze względu na określoną dziedzinę oraz eksportowania reguł zapisanych w formacie odpowiednim dla stosowanych programów "uczenia maszynowego";
- poprawności działania oprogramowania sterującego przebiegiem eksperymentu numerycznego;
- możliwości "krzyżowej" oceny reguł pozyskanych od specjalistów za pomocą zbioru przykładów testowych oraz "uczenia przyrostowego";

- możliwości "krzyżowej" oceny reguł, pozyskanych metodami indukcyjnymi ze zbioru przykładów uczących, przez specjalistów.

W rozprawie nie zamieszczono szczegółowego opisu przeprowadzonych badań, który jest zawarty w raporcie [Moczulski *et al.*, 1997a]. W wyniku przeprowadzonych badań potwierdzono poprawność przyjętego rozwiązania, w którym szczególną rolę spełnia baza danych i wiedzy *EMPREL.* Potwierdzono pełną wymienność stosowanego sposobu reprezentacji wiedzy ze sposobami, używanymi w udostępnionym oprogramowaniu, oraz możliwość łączenia wiedzy pozyskanej z różnych źródeł.

5.4 Podsumowanie przeprowadzonych badań

Przeprowadzone badania opracowanych metod oraz środków, umożliwiających działania zgodne z tymi metodami, pozwoliły na potwierdzenie poprawności ogólnych koncepcji.

W odniesieniu do poszczególnych grup metod i środków wspomagania procesu pozyskiwania wiedzy stwierdzono, co następuje:

1. *Metody pozyskiwania nowej wiedzy bezpośrednio od specjalistów*

   Potwierdzono ogólną poprawność opracowanych metod oraz środków wspomagania procesu pozyskiwania wiedzy z zastosowaniem tych metod. Dalsze prace nad oprogramowaniem do pozyskiwania wiedzy od specjalistów (bez udziału programisty bazy wiedzy) winny mieć na celu ułatwienie posługiwania się tym oprogramowaniem przez osoby o niewielkim doświadczeniu w stosowaniu technik komputerowych.

2. *Metody merytorycznej oceny wiedzy uprzednio pozyskanej*

   Oba opracowane środki (formularz "papierowy" oraz formularz "elektroniczny" do oceny reguł) umożliwiają skuteczną ocenę wcześniej pozyskanej wiedzy. Wyróżnienie niewielu wartości stopnia przekonania o słuszności reguł ułatwia dokonanie oceny. Należy doskonalić aplikację umożliwiającą łączenie opinii różnych osób, oceniających reguły zapisane w bazie wiedzy. Należy uzupełnić system oceny o możliwość oceniania związków zachodzących pomiędzy regułami oraz oceniania elementów metawiedzy (np. uporządkowania w zbiorze reguł, przypisanych do jednej dziedziny lub do jednego kontekstu).

3. *Metody indukcyjnego pozyskiwania wiedzy*

   W pełni potwierdzono przydatność zastosowanych metod indukcyjnych do pozyskiwania wiedzy diagnostycznej oraz wysoką efektywność tych metod. Przeprowadzone badania potwierdziły szczególne znaczenie prawidłowego opracowania planu eksperymentu (co dotyczy zarówno eksperymentu czynnego, jak i eksperymentu numerycznego). Zasadnicze znaczenie ma optymalizacja przestrzeni reprezentacji cech (w odniesieniu do wyboru cech relewantnych oraz odpowiedniego określenia dziedziny wartości jakościowych każdej cechy). Istotne jest także uzupełnienie schematów wnioskowania o przypadki nieokreślonych konkluzji (gdy wartość miary dopasowania klasyfikowanego przykładu do przesłanek reguł jest zbyt niska) oraz konkluzji wieloznacznych (gdy wartości wspomnianej miary dla kilku klas są zbliżone do siebie).

4. *Metoda dekompozycji zbioru przykładów dla przypadku złożonej struktury stanów*

   Potwierdzono przydatność opisanej metody postępowania zwłaszcza w przypadku złożonej struktury stanów. Stwierdzono także, że opracowana metoda winna być w sposób skuteczny wspomagana za pomocą odpowiednich środków programowych (które zostały opracowane przez P. Kostkę i zweryfikowane podczas realizacji opisanych badań [Moczulski, Kostka, 1997]).

5. *Metoda łączenia wiedzy pochodzącej od specjalistów z wiedzą pozyskaną indukcyjnie*

   Przeprowadzone badania wstępnie potwierdziły możliwość łączenia wiedzy pochodzącej z wymienionych powyżej źródeł. Do łączenia wiedzy pochodzącej z różnych źródeł może być stosowane "uczenie przyrostowe" z regułami wstępnymi pozyskanymi od specjalisty. Oprócz istotnego zwiększenia efektywności procesu pozyskiwania wiedzy, "uczenie przyrostowe" umożliwia doskonalenie istniejącej zawartości bazy wiedzy.

6. *System pozyskiwania wiedzy diagnostycznej SPWD1*

   Potwierdzono istotną rolę wspólnej reprezentacji danych i wiedzy diagnostycznej, która może być podstawą do opracowania schematu logicznego bazy danych (schemat *EMPREL*), umożliwiającej gromadzenie oraz udostępnianie danych i wiedzy diagnostycznej. Baza danych i wiedzy spełnia centralną rolę w opracowanym systemie, umożliwiając "krzyżowe" stosowanie metod i środków pozyskiwania nowej wiedzy oraz oceny wiedzy już pozyskanej.

Ponadto badania potwierdziły, że skuteczne pozyskiwanie wiedzy diagnostycznej wymaga, aby osoby realizujące dane zadanie pozyskiwania wiedzy posiadały głęboką wiedzę dotyczącą dziedziny zastosowań. Dotyczy to związków przyczynowo-skutkowych, związków zachodzących pomiędzy cechami obiektów, sygnałami diagnostycznymi itp. W kategoriach pojęć stosowanych w niniejszej rozprawie oznacza to potrzebę reprezentowania bardzo bogatej wiedzy o dziedzinie zastosowań, której posiadania specjalista sobie zwykle nie uświadamia.

Zdaniem autora, jest to zasadniczy argument przemawiający za niecelowością wyeliminowania człowieka-specjalisty z wszystkich stadiów procesu pozyskiwania wiedzy.

Dalsze badania winny zmierzać w kierunku potwierdzenia możliwości stosowania opracowanych metod i środków pozyskiwania wiedzy w odniesieniu do:

1. *bardziej złożonych obiektów*, w których ujawniają się *nieliniowe związki pomiędzy cechami stanu i symptomami diagnostycznymi*, co zwykle jest związane z wielkością tych obiektów;

2. zadań diagnostycznych, w których winny być rozpatrywane złożone stany techniczne (tj. *współwystępowanie kilku stanów technicznych*) z możliwością występowania synergii (dodatniej i ujemnej) w odniesieniu do symptomów diagnostycznych tych stanów (co ma miejsce w wypadku *złożonych układów maszynowych*, których przykładami są również turbozespoły energetyczne).

Należy nadmienić, że zespół pod kierunkiem autora rozpoczął już wstępne badania dotyczące pozyskiwania wiedzy o turbozespołach energetycznych (w ramach projektu KBN nr PBZ-038-06).

Rozdział 6

Podsumowanie i wnioski

Praca jest wynikiem badań autora nad metodologią pozyskiwania wiedzy dla potrzeb wspomagania działań w zakresie diagnostyki maszyn. Stanowi ona próbę syntezy obejmującej całokształt zagadnień związanych z pozyskiwaniem wiedzy potrzebnej do diagnozowania wybranych stanów technicznych pewnej wyróżnionej klasy maszyn, począwszy od identyfikacji przedmiotu badań, a skończywszy na określeniu programu postępowania, jakim jest scenariusz pozyskiwania wiedzy diagnostycznej.

Sformułowany we wstępie cel rozprawy został w zakresie *opracowania kompleksowej metodologii pozyskiwania wiedzy* osiągnięty poprzez:

- *szczegółową identyfikację przedmiotu badań* (jakim są problemy pozyskiwania wiedzy diagnostycznej, zob. rozdział 2), wymagającą analizy własności obiektu diagnozowania, określenia zadania diagnostyki maszyn, uwzględnienia specyficznych cech danych i wiedzy w diagnostyce maszyn, analizy wybranych metod reprezentacji danych i wiedzy diagnostycznej oraz identyfikacji przebiegu procesu pozyskiwania wiedzy diagnostycznej (od specjalistów, a także z baz danych zawierających wyniki obserwacji lub wyniki obliczeń symulacyjnych);

- opis *wybranych metod pozyskiwania wiedzy diagnostycznej*, dotyczący metod pozyskiwania wiedzy od specjalistów w danej dziedzinie zastosowań (rozdział 3.1) oraz metod pozyskiwania wiedzy w sposób indukcyjny na podstawie wstępnie sklasyfikowanego zbioru przykładów (rozdział 3.2);

- określenie *metod i technik oceny wiedzy pozyskanej*, włącznie z określeniem "krzyżowego" zastosowania opracowanych sposobów oceny pozyskanej wiedzy (rozdział 3.3);

- opracowanie *scenariusza procesu pozyskiwania wiedzy diagnostycznej*, stanowiącego rodzaj metody (zapisanej w postaci programu postępowania), a także modelu procesu pozyskiwania wiedzy diagnostycznej (rozdział 3.4).

Mając na uwadze *praktyczne zastosowania opracowanej metodologii*, utworzono *zintegrowany system pozyskiwania wiedzy diagnostycznej*. Centralnym elementem tego systemu jest *baza danych i wiedzy*, która opracowana jest zgodnie ze schematem logicznym *EMPREL*. Baza umożliwia deklaratywny sposób reprezentacji danych i wiedzy, który ma uniwersalne zastosowanie, niezależnie od źródła pochodzenia danych i wiedzy. Mając na względzie zasadnicze znaczenie specjalistów jako źródeł wiedzy, których udział w procesie pozyskiwania wiedzy jest bardzo pożądany (zwłaszcza w początkowym stadium, gdy

identyfikowany jest opis dziedziny zastosowania oraz ograniczony zasób wiedzy "operacyjnej" warunkującej możliwość wstępnego wnioskowania diagnostycznego), opracowano środki wspomagające pozyskiwanie wiedzy mające charakter formularzy ("papierowych" lub "elektronicznych"). Umożliwiają one zmniejszenie roli programisty bazy wiedzy jako dodatkowego uczestnika procesu pozyskiwania wiedzy. Ze względu na małą efektywność procesu pozyskiwania wiedzy od specjalistów do opisywanego systemu włączono elementy umożliwiające pozyskiwanie wiedzy drogą indukcji ze zbioru odpowiednio przygotowanych przykładów. Przykłady te mogą być gromadzone w trakcie *obserwacji diagnostycznej obiektów* (w ramach czynnych i biernych eksperymentów diagnostycznych), a także mogą być uzyskane jako wyniki *eksperymentów symulacyjnych*. Wspomniany uniwersalny sposób reprezentacji danych i wiedzy umożliwia także "krzyżowe" pozyskiwanie nowej wiedzy (z zastosowaniem "uczenia przyrostowego") oraz "krzyżowe" sposoby oceny wiedzy pozyskanej, gdzie specjalista ocenia reguły pozyskane w sposób indukcyjny, lub reguły zapisane przez specjalistę są poddawane ocenie za pomocą odpowiedniego zbioru przykładów testowych.

Dla oceny przydatności opisanych w rozprawie metod i środków pozyskiwania wiedzy diagnostycznej oraz weryfikacji metod opracowanych przez autora przeprowadzono oddzielne badania:

- dotyczące wyników *czynnego eksperymentu diagnostycznego* obejmującego diagnozowanie stanu wirnika modelu maszyny wirnikowej (rozdział 5.1);
- w ramach *eksperymentu symulacyjnego*, prowadzonego za pomocą systemu programów pozwalającego na dokonywanie obliczeń symulacyjnych drgań wymuszonych układu, spowodowanych zadanym rozkładem niewyrównoważenia wirnika oraz jego wirowaniem z różnymi prędkościami obrotowymi (rozdział 5.2).

Potwierdziły one poprawność przyjętych koncepcji dotyczących systemu pozyskiwania wiedzy oraz skuteczność opracowanych środków wspomagania procesu pozyskiwania wiedzy w odniesieniu do typowych zadań diagnostycznych.

Badania, zrealizowane w opisanym powyżej zakresie, umożliwiają sformułowanie następujących stwierdzeń dotyczących tez pracy:

1. Słuszność tezy 1. wykazano, *opracowując odpowiednie środki* do pozyskiwania wiedzy od specjalistów, jak formularz "papierowy" (rozdział 4.2.1) oraz formularz "elektroniczny" (rozdział 4.2.3), umożliwiające pozyskanie nowej "porcji" wiedzy reprezentowanej za pomocą empirycznych relacji diagnostycznych, a następnie stosując te środki do pozyskania wiedzy od specjalistów w ramach przeprowadzonych badań weryfikacyjnych.
2. Słuszność tezy 2. wykazano, *opracowując schemat logiczny bazy danych i wiedzy EMPREL* (rozdział 4.1), a następnie opracowując na jego podstawie relacyjną bazę danych i wiedzy, w której mogą być jednocześnie zapisywane "porcje" wiedzy diagnostycznej pozyskanej od specjalistów, przykłady zgromadzone dla potrzeb pozyskania wiedzy metodami indukcyjnymi oraz "porcje" wiedzy pozyskanej indukcyjnie na podstawie tych przykładów.

   Tezę tę dodatkowo potwierdza możliwość stosowania wiedzy (zapisanej zgodnie z przyjętym sposobem reprezentacji) w sposób niezależny od pierwotnego źródła

tej wiedzy. W zakresie działań opisanych w rozprawie ma to miejsce podczas oceniania przez specjalistów wiedzy wcześniej pozyskanej (np. za pomocą formularza "papierowego" do oceny reguł — zob. rozdział 4.2.2, lub formularza "elektronicznego" wtedy, gdy stosowany jest do oceny reguł — zob. rys. B.1) lub podczas tzw. "uczenia przyrostowego" (rozdział D.4).

Oceniając wykonaną pracę, autor uważa, że jej najważniejszymi, oryginalnymi (jego zdaniem) elementami są:

1. pojęcie *empirycznej relacji diagnostycznej* jako podstawowego środka stosowanego do pozyskiwania wiedzy od specjalistów (rozdział 3.1.1);
2. *metoda dekompozycji zbioru przykładów w wypadku złożonej struktury drzewa stanów* (określenie w rozdziale 3.2.2 i weryfikacja w rozdziale C.4);
3. *scenariusz procesu pozyskiwania wiedzy diagnostycznej* (rozdział 3.4);
4. *schemat logiczny bazy danych i wiedzy EMPREL* (rozdział 4.1), a zwłaszcza spostrzeżenie, że nie należy oddzielać od siebie opisu dziedziny, bazy danych i bazy wiedzy (punkt 7 na s. 112);
5. koncepcja *formularza elektronicznego* (rozdział 4.2.3), a zwłaszcza koncepcja pozyskiwania elementów reguł na drodze "kroczenia wstecz" (rozdział 3.1.3);
6. *system pozyskiwania wiedzy diagnostycznej* integrujący wszystkie zastosowane środki pozyskiwania wiedzy (rozdział 4.4).

Ponadto istotna część pracy (rozdział 5 wraz z dodatkami C i D) dotyczy oceny przydatności wprowadzonych metod i zastosowanych środków pozyskiwania wiedzy diagnostycznej oraz weryfikacji metod zaproponowanych przez autora i środków opracowanych w kierowanym przez niego zespole.

Przeprowadzone badania, które opisano w rozprawie, pozwalają na określenie *kierunków dalszych badań*, których realizacja wiąże się z *wdrożeniami opracowanych metod i środków pozyskiwania wiedzy diagnostycznej*, a także z potrzebą *podjęcia nowych kierunków badań*. Uzasadnia to celowość *dalszego rozwoju metodologii pozyskiwania wiedzy diagnostycznej*. Autor zamierza kontynuować te badania wraz z grupą współdziałających z nim pracowników Katedry Podstaw Konstrukcji Maszyn Politechniki Śląskiej, zwłaszcza poprzez *realizację odpowiednich projektów badawczych*, *podejmowanie przez współpracowników tematów prac doktorskich* oraz *współpracę międzynarodową* (w tym w ramach odpowiednich programów badawczych).

Kierunki dalszych badań oraz nowe kierunki badań zostały już wstępnie zasygnalizowane w treści podsumowań kończących poszczególne rozdziały rozprawy. Należy ponownie zaakcentować potrzebę kontynuacji lub podjęcia badań, które dotyczyć będą między innymi takich problemów, jak:

1. *W zakresie rozwoju sposobów reprezentacji danych i wiedzy:*

   (a) *Rozwój metod kwantyzacji cech ilościowych dla potrzeb diagnostyki maszyn.*

   Problem dotyczy określenia odpowiedniej liczby poziomów kwantyzacji oraz zbioru wartości progowych. Należy rozważyć wiele kryteriów, jak np. minimalny

łączny błąd klasyfikacji, możliwość łączenia wiedzy pochodzącej z różnych źródeł. Jednocześnie należy przeciwdziałać nadmiernemu dopasowaniu sposobu kwantyzacji do zbioru danych.

(b) *Rozwój sposobów reprezentacji danych i wiedzy o obiektach, których cechy zmieniają się w funkcji czasu.*

Rozwiązanie takich problemów wymagane jest ze względu na potrzeby diagnostycznych systemów monitorowania maszyn. Przewiduje się reprezentację danych za pomocą *sekwencji zdarzeń*, a reprezentację zmian wartości cech obiektów za pomocą opisów stosowanych w *metodach syntaktycznego rozpoznawania obrazów*. Do reprezentacji wiedzy może być zastosowany np. tzw. *deterministyczny system działań*, wyznaczony przez zbiór reguł oraz pola sterowania, określające działania w wypadku spełnienia przesłanek danej reguły i w wypadku gdy te przesłanki nie są spełnione.

(c) *Rozwój metod proceduralnej reprezentacji wiedzy diagnostycznej.*

Podczas rozwiązywania problemów diagnostycznych specjalista często postępuje zgodnie z wypracowaną przez siebie procedurą, która nie zawsze ma charakter zdeterminowany. Celowe jest reprezentowanie w sposób proceduralny wiedzy o sposobach postępowania specjalisty. Do reprezentacji procedur zdeterminowanych mogą być np. stosowane schematy blokowe lub zapis w odpowiednim języku programowania (np. *LISP*). Należy opracować metody reprezentacji procedur, w których wybór odpowiednich kroków postępowania nie jest zdeterminowany.

2. *W zakresie rozwoju metod pozyskiwania wiedzy i oceny wiedzy pozyskanej:*

(a) *Opracowanie metod pozyskiwania i weryfikacji wiedzy diagnostycznej, reprezentowanej w sposób proceduralny.*

Wiedza reprezentowana proceduralnie może być pozyskiwana np. poprzez stopniowe zwiększanie stopnia szczegółowości specyfikacji. Możliwe jest także zastosowanie metod syntezy *deterministycznych systemów działań*, dokonywanej na podstawie odpowiednio przygotowanych tablic decyzyjnych.

(b) *Zastosowanie metod odkryć w bazach danych do pozyskiwania wiedzy diagnostycznej.*

Nastąpi kontynuacja rozpoczętych już badań nad zastosowaniami tych metod do odkrywania zależności o charakterze jakościowym i ilościowym, opisujących relacje "wprost" i relacje odwrotne.

(c) *Rozwój metod i technik oceny wiedzy pozyskanej.*

Metody te winny uwzględniać prawdopodobieństwa wystąpienia danego stanu oraz inne kryteria, np. ryzyka spowodowanego wystąpieniem danego uszkodzenia lub awarii.

3. *W zakresie rozwoju środków pozyskiwania wiedzy:*

(a) *Rozwój środków pozyskiwania wiedzy reprezentowanej w sposób proceduralny.*

Przewiduje się dalszy rozwój koncepcji *formularza elektronicznego* m.in. poprzez dołączenie możliwości edycji prostych schematów blokowych. Ponadto

nastąpi rozwój oprogramowania umożliwiającego syntezę deterministycznych systemów działań.

(b) *Rozwój środków pozyskiwania wiedzy na podstawie sekwencji zdarzeń i stanów.*

Przewiduje się opracowanie oprogramowania, wspomagającego pozyskiwanie wiedzy w sposób interaktywny, na podstawie odpowiednio reprezentowanych opisów przypadków (ang. *case stories*).

(c) *Weryfikacja przydatności koncepcji autonomicznego systemu odkryć.*

Moduł sterujący takiego systemu winien *w sposób dynamiczny modyfikować plan prowadzonego eksperymentu* (czynnego lub numerycznego), związanego z gromadzeniem danych, które mogą być zastosowane do weryfikacji hipotez dotyczących związków zachodzących pomiędzy cechami. Moduł ten powinien sterować działaniem modułu przygotowującego przykłady uczące (jak np. systemem programów symulacyjnych lub programem wspomagającym przeprowadzenie eksperymentu czynnego oraz dokonującym pomiaru sygnałów diagnostycznych i estymacji wartości cech tych sygnałów). Przykłady te stanowią dane dla modułu dokonującego odkryć związków zachodzących pomiędzy cechami niezależnymi i zależnymi.

(d) *Dalszy rozwój systemu pozyskiwania wiedzy diagnostycznej.*

Oprócz zintegrowania z systemem *SPWD1* nowych narzędzi programowych, przewiduje się rozwój bazy *EMPREL* w celu umożliwienia pozyskiwania przykładów uczących z systemu baz danych *MIMOSA*.

4. *W zakresie dalszej weryfikacji metod i środków pozyskiwania wiedzy:*

(a) *Weryfikacja opracowanych metod i środków w odniesieniu do bardziej złożonych obiektów.*

W przypadku bardziej złożonych obiektów ujawniają się *nieliniowe związki pomiędzy cechami stanu i symptomami diagnostycznymi*, co zwykle jest związane z wielkością tych obiektów.

(b) *Dalsza weryfikacja metod pozyskiwania wiedzy w przypadku złożonej struktury drzewa stanów.*

Podejmowane będą zadania diagnostyczne, w których winny być rozpatrywane złożone stany techniczne (tj. *współwystępowanie kilku elementarnych stanów technicznych*) z możliwością występowania synergii (dodatniej i ujemnej) w odniesieniu do symptomów diagnostycznych tych stanów (co ma miejsce w wypadku *złożonych megaukładów maszynowych*, których przykładami są turbozespoły energetyczne).

Zdaniem autora, realizacja zadań wynikających z zarysowanych kierunków dalszych badań oraz zgromadzona wiedza i doświadczenie dotyczące metodologii pozyskiwania wiedzy dla potrzeb diagnostyki maszyn uzasadniają w przyszłości potrzebę opracowania monografii poświęconej tej metodologii.

Dodatek A

Ogólny opis schematu logicznego *EMPREL*

W dodatku umieszczono ogólny opis schematu bazy danych i wiedzy *EMPREL* (zob. rys. A.1). Jak wspomniano w rozdziale 4.1, w bazie *EMPREL* można wyróżnić trzy główne podbazy: *opis dziedziny*, *bazę danych* i *bazę wiedzy operacyjnej*.

A.1 Opis dziedziny

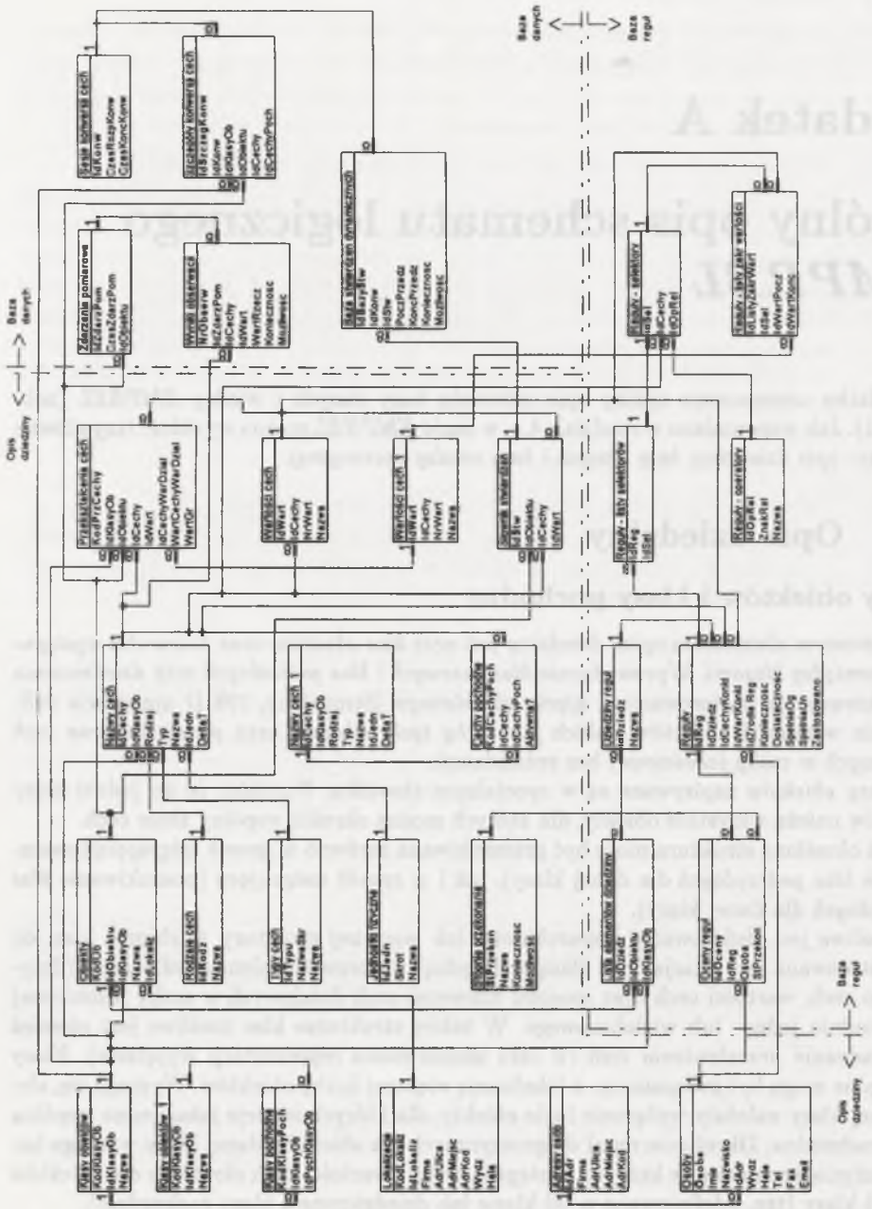
Klasy obiektów i klasy pochodne

Podstawowym elementem opisu dziedziny jest opis *klas obiektów* oraz *hierarchii występującej pomiędzy klasami*. Wprowadzenie *klas bazowych* i *klas pochodnych* oraz *dziedziczenia wielobazowego* (z zastosowaniem *ujęcia obiektowego* [Stroustrup, 1991]) umożliwia definiowanie własności obiektów (takich jak *cechy tych obiektów* oraz *przekształcenia cech ilościowych w cechy jakościowe*) bez redundancji.

Klasy obiektów zapisywane są w specjalnym słowniku. Przyjęto, że do jednej klasy obiektów należą wszystkie obiekty, dla których można określić wspólny zbiór cech.

Tak określona struktura może być przeszukiwana zarówno *w sposób zstępujący* (poszukiwanie *klas podrzędnych* dla danej klasy), jak i *w sposób wstępujący* (poszukiwanie *klas nadrzędnych* dla danej klasy).

Możliwe jest definiowanie *hierarchicznej* lub *sieciowej* struktury w zbiorze klas, do czego stosowana jest relacja "być klasą podrzędną" z wprowadzeniem *dziedziczenia* (dotyczącego cech, wartości cech oraz sposobu konwersji cech ilościowych w cechy jakościowe) odpowiednio jedno- lub wielobazowego. W takiej strukturze klas możliwe jest również wprowadzanie *przesłaniania cech* (w celu umożliwienia reprezentacji wyjątków). Klasy podrzędne mogą być związane np. z lokalizacją większej liczby obiektów. Wymaga się, aby do danej klasy należały wyłącznie takie obiekty, dla których istnieje jakaś jedna wspólna klasa nadrzędna. Określanie reguł diagnostycznych dla obiektów danej klasy wymaga bowiem użycia warunków, w których występują cechy i wartości cech określone dla obiektów z danej klasy (tzn. zdefiniowane w tej klasie lub dziedziczone z klasy nadrzędnej).



Rys. A.1. Uproszczony schemat logiczny bazy danych i wiedzy *EMPREL*
Fig. A.1. Simplified logical scheme of the *EMPREL* data and knowledge base

Tabl. A.1. Rodzaje cech

Rodzaj	Opis
A	cechy podstawowych warunków działania maszyny
E	cechy do reprezentowania zdarzeń eksploatacyjnych (obsługa, naprawa lub wymiana elementu itp.)
F	cechy do reprezentowania faz eksploatacji
S	elementarne cechy stanu technicznego obiektu
X	elementarne cechy wejścia, które przede wszystkim określają chwilowe warunki działania obiektu
Y	elementarne cechy wyjścia, dotyczące procesów użytecznych (wytwórczych) oraz tzw. procesów "resztkowych"

Obiekty i ich lokalizacja

Możliwe jest gromadzenie danych, dotyczących *obiektów* (takich jak maszyna, podzespół maszyny itp.). W obecnej wersji bazy *EMPREL* nie wprowadzono jeszcze stosowanego w bazie danych *MIMOSA* rozróżnienia na zespoły funkcjonalno-obsługowe i egzemplarze-konkrety[1]. Nie są też jeszcze reprezentowane relacje "być podzespołem", które umożliwiłyby określenie struktury funkcjonalnej diagnozowanych obiektów.

Dla poszczególnych obiektów-konkretów podaje się ich *lokalizację*, przez którą rozumie się zwykle ściśle określony wydział oraz halę zakładu lub przedsiębiorstwa.

Cechy

Najbardziej rozbudowana część opisu dziedziny dotyczy określenia *cech*. Cechy przysługują klasom obiektów. Dla cech określono ich rodzaje związane z kategorią własności odnoszących się do opisywanego obiektu. Rozróżnienie to jest istotne ze względu na możliwość organizowania list wyboru dla potrzeb wprowadzania danych.

W obecnej wersji bazy *EMPREL* mogą być stosowane zarówno *cechy ilościowe*, jak i *cechy jakościowe*. W tym celu wprowadzono *typy cech*.

Baza *EMPREL* nie jest przeznaczona do gromadzenia szczegółowych danych, dotyczących torów pomiarowych, wartości mierzonych, kierunków pomiaru itp. W kolejnej wersji bazy wiedzy przewiduje się w tym celu możliwość jej zintegrowania z bazami danych w standardzie *MIMOSA*.

Rodzaje cech. Istotną właściwością cechy jest jej *rodzaj*. Stosowane rodzaje cech zestawiono w tablicy A.1.

Wprowadzenie rodzaju cechy ułatwia opracowywanie oprogramowania wspomagającego pozyskiwanie wiedzy od specjalistów. Oprogramowanie to zostało opisane w dalszej części niniejszego rozdziału.

[1]Zob. omówienie na s. 26.

Tabl. A.2. Typy cech

Typ	Opis
`cyc`	cecha jakościowa o cyklicznie uporządkowanym zbiorze wartości
`lin`	cecha jakościowa o liniowo uporządkowanym zbiorze wartości
`nom`	cecha jakościowa nominalna, o nie uporządkowanym zbiorze wartości
`pom`	cecha ilościowa "pomiarowa" (o wartościach rzeczywistych)

Typy cech. *Typy cech* umożliwiają rozróżnienie *cech ilościowych* od *cech jakościowych* oraz określenie podstawowych właściwości zbiorów wartości tych cech. Stosowane typy cech przedstawiono w tablicy A.2.

Wartości cech ilościowych określane są zwykle w wyniku pomiaru wielkości ciągłych (np. temperatura, ciśnienie, wartość skuteczna prędkości drgań itp.). Mogą one być także uzyskane np. w wyniku obliczeń symulacyjnych. Założono, że cechy ilościowe mają wartości rzeczywiste. Dla cech ilościowych określa się sposoby ich przekształcania w pochodne cechy jakościowe.

Pozostałe typy dotyczą cech jakościowych. Cechy jakościowe stosowane są m. in. do zapisu stwierdzeń oraz warunków w przesłankach reguł i konkluzji tych reguł. Wprowadzenie sposobu opisu własności obiektów za pomocą cech jakościowych jest jednym z warunków ograniczenia przestrzeni reprezentacji cech obiektów do tzw. "domkniętego świata".

Jednostki fizyczne. Słownik *jednostek fizycznych* wprowadzono w celu umożliwienia generowania poprawnych raportów, zawierających wartości cech ilościowych (takich jak wyniki pomiarów i/lub obliczeń).

Nazwy cech. Podstawowe znaczenie ma słownik *nazw cech*. Cechy są określane dla klas obiektów, posiadają rodzaj i typ, a ich wartościom jest przypisana jednostka fizyczna.

Dla cech ilościowych określana jest wartość parametru ΔT, umożliwiająca ocenę tempa zmian danej cechy (a pośrednio — jej pasma częstotliwościowego). Wartość ta jest stosowana przy operacji *sklejania stwierdzeń dynamicznych* (zob. rozdział 2.1.4).

Wartości cech jakościowych

Wartości cech jakościowych określone są również za pomocą słownika. Dla cech o typie `lin` i `cyc` wymaga się określenia uporządkowania w zbiorze wartości cech.

Ponieważ wartości cech są przypisane danej cesze jakościowej, są one jednocześnie przypisane klasie obiektów, dla której jest określona dana cecha. Ze względu na występującą strukturę klas cechy wraz z ich wartościami mogą być dziedziczone przez klasy podrzędne, a więc także przez obiekty należące do tych klas.

Przekształcenia cech ilościowych

Przekształcenia cech ilościowych umożliwiają kwantyzację ich wartości w celu uzyskania zbioru wartości cech jakościowych (zwykle o typie `lin`). Utworzony w ten sposób zbiór wartości przypisuje się cesze jakościowej, zwanej *cechą pochodną*.

Przekształcenia cech są w zasadzie określone dla klas obiektów. Dopuszcza się określanie przekształceń dla pojedynczych obiektów (np. w wypadku obiektów o nadmiernej aktywności wibracyjnej można przyjmować wyższe wartości progowe, rozdzielające przedział wartości dopuszczalnych od niedopuszczalnych).

Cechy pochodne. Określenie *cech pochodnych* następuje w podobny sposób jak określenie klas pochodnych: dla cechy ilościowej "pierwotnej" można określić co najmniej jedną cechę pochodną.

Przekształcenia cech ilościowych w cechy jakościowe. *Przekształcenia cech* mogą być określone dla:

- klas obiektów,
- pojedynczego obiektu.

Drugi z przypadków umożliwia określenie przekształcenia cech dla obiektów o specjalnych właściwościach, wyróżniających te obiekty spośród innych obiektów danej klasy (np. dla obiektów nadmiernie wibroaktywnych lub eksploatowanych w specjalnych warunkach).

Przekształcenia cech są zdefiniowane dla ściśle określonych warunków działania obiektu, identyfikowanych za pomocą wartości cech rodzaju "A". Przekształcenia cech zwykle są określane co najmniej dla przypadku "*warunki działania*"="*nominalne*".

W bazie *EMPREL* stosowane są bardzo proste przekształcenia określone zależnością (2.10). Jako definicję takiego przekształcenia (dla określonych warunków działania obiektu) można uznać zbiór wartości progowych

$$\{v_1, v_2, \ldots, v_{M(a)}\} \tag{A.1}$$

wg zależności (2.9), przy czym w bazie zapisywane są tylko granice podprzedziałów (przyjmuje się ograniczenie, że kolejnym podprzedziałom odpowiadają rosnące wartości cechy jakościowej). Nie jest wymagane stosowanie podprzedziałów o równej długości. Odpowiedni wybór progów v_m , $m = 1, \ldots, M(a)$ może być np. przedmiotem optymalizacji [Chmielewski, Grzymała-Busse, 1992]. Wartością cechy jakościowej, otrzymaną w wyniku zastosowania przekształcenia wartości v cechy ilościowej, jest indeks m_0 spełniający:

$$v_{m0} = \inf_{m=1,\ldots,M(a)} \{v_m \mid v < v_m\}. \tag{A.2}$$

Jeśli przekształcenie nie jest określone bezpośrednio dla danego obiektu, przyjmuje się, że stosowane będą przekształcenia określone dla klasy obiektów, do której ten obiekt należy. Poszukiwanie odpowiedniego przekształcenia rozpoczyna się od określonego dla "najwęższej" klasy i w razie potrzeby jest kontynuowane dla klas nadrzędnych, zgodnie ze strukturą klas.

Stopnie przekonania

Stopnie przekonania (o słuszności reguł) reprezentowane są za pomocą wartości miar konieczności i możliwości. Miara ta jest stosowana do oceny reguł przez niezależnych specjalistów oraz do reprezentowania stopni słuszności reguł.

Wartości stopni przekonania stosowane w obecnej wersji bazy *EMPREL* przedstawiono w tablicy 3.1.

Słownik stwierdzeń

Słownik stwierdzeń stosowany jest do zapisu bazy stwierdzeń dynamicznych, określanych wyłącznie dla wartości cech jakościowych (zob. rozdział A.2). Umożliwia on znaczne ograniczenie redundancji tej bazy.

Osoby i ich adresy

Słownik osób odnosi się do autorów reguł oraz autorów ocen reguł. Adresy osób ułatwiają kontakt z tymi osobami.

Synonimy nazw słownikowych i objaśnienia

Opis dziedziny obejmuje także *synonimy nazw słownikowych*. Potrzeba ich stosowania wynika z występowania wielu synonimów nazw: obiektów i ich podzespołów, cech tych obiektów oraz wartości cech. Lista synonimów przypisywana jest jednej *nazwie głównej*, która będzie służyła do reprezentowania danego pojęcia. Synonimy tej nazwy są dostępne jako elementy składowe systemu objaśnień.

Każdy ze słowników zawiera także pola umożliwiające gromadzenie opisów i elementów objaśnień (tekstowych i graficznych). Ułatwia to generowanie objaśnień, co ma miejsce podczas konsultacji z diagnostycznym systemem doradczym (na życzenie użytkownika).

A.2 Baza danych

Baza danych jest podbazą służącą do gromadzenia danych dotyczących konkretnych obiektów. Gromadzi się:

- *Wyniki obserwacji* dokonanych przez obsługę danej maszyny. Dotyczą one spostrzeżeń odnoszących się do eksploatacji oraz obsług i napraw. Wyniki obserwacji są reprezentowane za pomocą stwierdzeń dynamicznych.
- *Wyniki pomiarów* sygnałów reprezentowane za pomocą wartości cech:
  - *jakościowych*, gdy źródłem wartości cech jest *autonomiczny agent pomiarowy*, dokonujący oceny zmierzonych wartości i ich wstępnej klasyfikacji,
  - *ilościowych*, gdy wartości cech podlegają jedynie przetwarzaniu analogowo-cyfrowemu.
- *Wyniki obliczeń symulacyjnych*, które (podobnie jak wyniki pomiarów) mogą być reprezentowane za pomocą wartości cech jakościowych lub ilościowych.

Zdarzenia pomiarowe i obserwacje oraz ich wyniki

Baza *EMPREL* umożliwia gromadzenie *wyników obserwacji*, *pomiarów* oraz *eksperymentów numerycznych*. Aby ułatwić posługiwanie się tymi wynikami, wprowadzono pojęcie *zdarzenia*, które jest związane z obiektem oraz punktem czasu, w którym dokonano obserwacji lub pomiaru albo który jest przypisany wynikom obliczeń symulacyjnych w celu ich identyfikacji.

Zdarzenia pomiarowe i obserwacje. Wyniki gromadzone w bazie danych są przypisywane do odpowiednich *zdarzeń pomiarowych* lub *obserwacji*. Przyjmuje się, że wyniki odpowiadające jednemu zdarzeniu pomiarowemu zostały określone *jednocześnie*. Za specjalny rodzaj zdarzenia pomiarowego uznawany jest jeden zestaw wyników *eksperymentu numerycznego*.

Każde zdarzenie pomiarowe, obserwacja lub eksperyment numeryczny jest przypisane jednemu obiektowi. Ponadto zapisywany jest czas zaobserwowania lub pomiaru (jako czas uniwersalny GMT).

Wyniki obserwacji. *Wyniki obserwacji* są zapisywane za pomocą stwierdzeń, reprezentowanych jako trójka <*obiekt*, *cecha*, *wartość*>, gdzie obiekt określony jest pośrednio poprzez zdarzenie pomiarowe/obserwację, zaś cecha jest określona poprzez identyfikator nazwy cechy.

Wartości cechy określane są:

- dla cech jakościowych (w obecnej wersji bazy *EMPREL* typy `nom`, `lin`, `cyc`) — poprzez podanie identyfikatora wartości cechy;
- dla cech ilościowych (typ `pom`; dla takich cech nie ma określonego zbioru wartości) — poprzez podanie wartości cechy jako liczby rzeczywistej.

Wartości cech ilościowych są następnie przekształcane (poddawane konwersji) w wartości cech jakościowych.

Dla stwierdzeń dotyczących wyników pomiarów/obserwacji mogą być określane stopnie prawdziwości poprzez podanie pary wartości miar konieczności i możliwości. Domyślnie przyjmuje się, że wynik pomiaru/obserwacji jest pewny.

Aby umożliwić jednolitą reprezentację wszystkich wyników pomiarów, przyjęto wspólny sposób gromadzenia danych zarówno dotyczących cech liczbowych (jednowymiarowych), jak i cech funkcyjnych, zwykle zapisywanych za pomocą macierzy liczb. Pozwala to na ujednolicenie operacji, związanych z zapisem, aktualizacją oraz wyszukiwaniem wartości cech.

Baza stwierdzeń dynamicznych i konwersje wyników obserwacji

Przyjęto, że wszystkie zgromadzone wyniki pomiarów, obserwacji oraz obliczeń symulacyjnych będą reprezentowane za pomocą *stwierdzeń dynamicznych*. Stwierdzeniu dynamicznemu przypisuje się *przedział czasu*, w którym to stwierdzenie ma określoną *wartość* (reprezentowaną za pomocą jego *stopnia prawdziwości*) [Cholewa, 1996].

Przekształcenie wyników obserwacji, pomiarów lub eksperymentów numerycznych w stwierdzenia dynamiczne pociąga za sobą potrzebę:

- przekształcenia wartości cech "pierwotnych" ilościowych w wartości pochodnych cech jakościowych;
- *sklejenia* sąsiadujących ze sobą przedziałów czasu (aby powiększyć przedział, w którym dane stwierdzenie ma określoną wartość).

Wiedza dotycząca przekształcania cech ilościowych w cechy jakościowe została określona w opisie dziedziny. Obejmuje ona:

- nazwy cech pochodnych o charakterze jakościowym,
- nazwy wartości cech pochodnych,
- związki pomiędzy cechami pierwotnymi i pochodnymi,
- wartości graniczne.

Sesje konwersji cech. Przekształcanie cech ilościowych w cechy jakościowe odbywa się w ramach *sesji konwersji cech*. Konwersje wartości cech nie są wykonywane w sposób automatyczny (tj. niezwłocznie po zgromadzeniu danego wyniku pomiaru lub obserwacji), lecz są inicjowane przez użytkownika. Sesja dotyczy wartości wszystkich cech zgromadzonych w ściśle określonym przedziale czasu. Obejmuje ona:

- *przekształcenie wartości cech ilościowych w wartości wszystkich ich cech pochodnych jakościowych*, które są aktywne w chwili dokonywania konwersji (wartości cech jakościowych, będące bezpośrednimi wynikami obserwacji lub pomiaru — np. w przypadku sygnałów dwustanowych — nie wymagają przekształcenia);
- *zapis w tabeli odpowiedniego stwierdzenia wraz z przedziałem czasu i oceną stopnia jego prawdziwości*; ta ostatnia operacja może się wiązać z tzw. "sklejaniem" przedziałów czasu, w których określone są stwierdzenia o identycznej wartości.

Baza stwierdzeń dynamicznych. W *bazie stwierdzeń dynamicznych* są zapisywane rekordy, zawierające identyfikatory stwierdzeń (ze słownika stwierdzeń) wraz z przedziałem czasu i oceną stopnia prawdziwości tych stwierdzeń. Jak wspomniano, ograniczenie liczby rekordów w tej bazie może zostać osiągnięte w wyniku tzw. "sklejania" przedziałów czasu, w których określone są stwierdzenia dla podanej wartości stopnia prawdziwości. Algorytm tej operacji jest następujący:

1. Sprawdzić, czy w bazie stwierdzeń dynamicznych występuje rekord zawierający takie samo stwierdzenie i o tej samej wartości stopnia prawdziwości, dla którego końcowa chwila przedziału czasu jego "ważności" oddalona jest o mniej niż ΔT (wartość ta jest określona dla każdej cechy — zob. s. 13);
2. Jeśli znaleziono taki rekord, wydłużyć przedział czasu, w którym stwierdzenie jest "ważne", aby kończył się chwilą czasu, w której zaobserwowano przetwarzany wynik;
3. W przeciwnym razie (gdy nastąpiła utrata ciągłości obserwacji cechy ilościowej) utworzyć nowy rekord dla generowanego stwierdzenia dynamicznego, określając jego punktowy przedział "ważności" odpowiadający chwili czasu, w której zaobserwowano przetwarzany wynik.

A.3 Baza wiedzy

Baza wiedzy zawiera reguły przybliżone, reprezentowane w sposób opisany na s. 47 i n., zbliżony do stosowanego przez R. S. Michalskiego (zapis w logice VL_1 [Michalski, 1983]),

lecz z uwzględnieniem pewnych ograniczeń. Rozwiązanie takie zostało przyjęte, by umożliwić reprezentację reguł pochodzących z różnych źródeł. Aby spełnić założenia reprezentacji obiektów i ich cech w "domkniętym świecie", reguły określane są dla cech jakościowych.

Ponadto w bazie wiedzy definiowane są *dziedziny reguł* reprezentowane w postaci list obiektów i list klas. Baza zawiera także dane dotyczące *źródeł reguł* (osoby, oprogramowanie) i *ocen reguł* przez niezależnych specjalistów.

Operatory relacji stosowanych do zapisu warunków w przesłankach reguł

Warunki reguł mają postać zdań zapisanych w logice VL_1. Do ich zapisu stosowane są *operatory porównania*, takie jak $=, \neq, <, \leqslant, \geqslant, >$. Operatory te są reprezentowane w postaci słownika.

Dziedziny reguł

Dziedzina reguły określana jest poprzez podanie listy obiektów i/lub klas obiektów. W szczególnym przypadku dziedzina reguły może zawierać jeden obiekt.

Przyjęto, że:

- Klasy obiektów są tworzone ze względu na jakieś kryterium. Podstawowym kryterium jest wspólna konstrukcja obiektów należących do jednej klasy. Dopuszcza się możliwość tworzenia klas ze względu na wspólną lokalizację obiektów.
- Lista, określająca dziedzinę reguły, może zawierać także pojedyncze obiekty określone w opisie dziedziny (zob. rozdział A.1).

Dla wszystkich obiektów tworzących jedną dziedzinę reguł musi istnieć wspólna klasa[2], do której te obiekty należą i dla której określono cechy zastosowane później do budowy selektorów w przesłankach oraz konkluzji reguł.

Reguły

W bazie reguł reprezentowane są reguły przybliżone o postaci:

JEŻELI <*przesłanka*> TO <*konkluzja*> ZE <*stopniem przekonania*>,

gdzie (zob. rozdział 2.5.3):

- przesłanka (część warunkowa) reguły jest pojedynczym selektorem wg (2.71) lub kompleksem wg (2.75);
- konkluzja jest stwierdzeniem.

Każdej regule przypisane są dwie miary: konieczności i dostateczności. Umożliwiają one określenie stopnia prawdziwości konkluzji dla stwierdzeń dopasowanych do przesłanki reguły i posiadających swoje stopnie prawdziwości. Sposób obliczania wartości stopni prawdziwości konkluzji opisano w [Cholewa, Pedrycz, 1987].

[2]Może ona być wspólną klasą nadrzędną dla wszystkich obiektów lub nadzbiorem wszystkich klas tworzących dziedzinę reguły.

Przykładami selektorów są (zapis zgodny z VL_1; poszczególne elementy składowe dysjunkcji wewnętrznej rozdziela się znakiem przecinka):

[*obciążenie* >= *średnie*]

[*prędkość_obrotowa* = *mała..duża*]

[*przesunięcie_fazowe* = 0..30, 80..100] lub [*warunki_działania* = *rozruch*, *rozbieg*].

Przykładowy zapis złożonej przesłanki (kompleksu):

[*obciążenie* >= *średnie*] [*prędkość_obrotowa* = *mała..duża*] .

Reguła określona jest przez:

1. wskazanie jej *dziedziny*;
2. podanie *identyfikatora reguły* i *źródła reguły*;
3. określenie *przesłanki reguły* (jako kompleksu, poprzez podanie listy selektorów dla danej reguły);
4. określenie poszczególnych *selektorów* przez podanie dla każdego z nich:
   (a) identyfikatora *cechy*, dla której jest określany ten selektor, oraz identyfikatora *operatora relacji porównania*,
   (b) *listy zakresów wartości* (dla cechy typu `lin`) lub *listy wartości* (dla cech pozostałych typów);
5. określenie *konkluzji reguły* (będącej pojedynczym stwierdzeniem) oraz oceny *stopnia przekonania* o słuszności reguły (za pomocą wartości miar konieczności i dostateczności tej reguły).

Elementy składowe reguły zapisywane są w oddzielnych tabelach w taki sposób, by ograniczyć redundancję.

Źródłem reguły może być specjalista, programista bazy wiedzy (wówczas gdy dokonuje interpretacji i edycji "porcji" wiedzy pozyskanej od specjalisty w postaci *empirycznych relacji diagnostycznych*) lub oprogramowanie (np. do "uczenia maszynowego" lub do odkrywania zależności o charakterze jakościowym i funkcyjnym).

Przyjęty sposób reprezentacji umożliwia zapis w bazie wiedzy reguł, będących wynikami działania różnych aplikacji, działających w kilku najbardziej popularnych środowiskach sprzętowych i programowych.

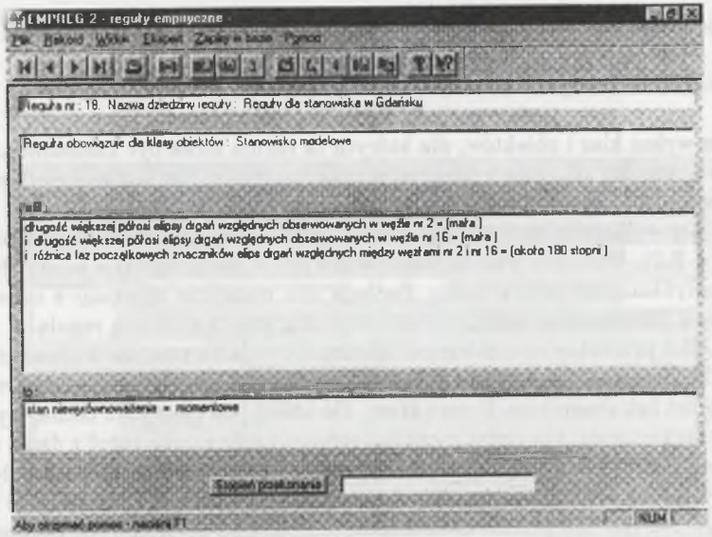
Oceny reguł

W bazie wiedzy gromadzone są także *oceny reguł* sformułowane przez niezależnych specjalistów. Oceny są formułowane za pomocą wartości stopnia przekonania o słuszności danej reguły. Jednej regule może odpowiadać wiele ocen, każda pochodząca od innego specjalisty. Oceny te są agregowane z ocenami przechowywanymi wraz z regułami, przy czym uwzględnia się wartości wag przypisanych poszczególnym oceniającym specjalistom. Oceny reguł w bazie wiedzy są aktualizowane po otrzymaniu nowych ocen od specjalistów.

Dodatek B

Sposób prowadzenia dialogu z użytkownikiem formularza "elektronicznego"

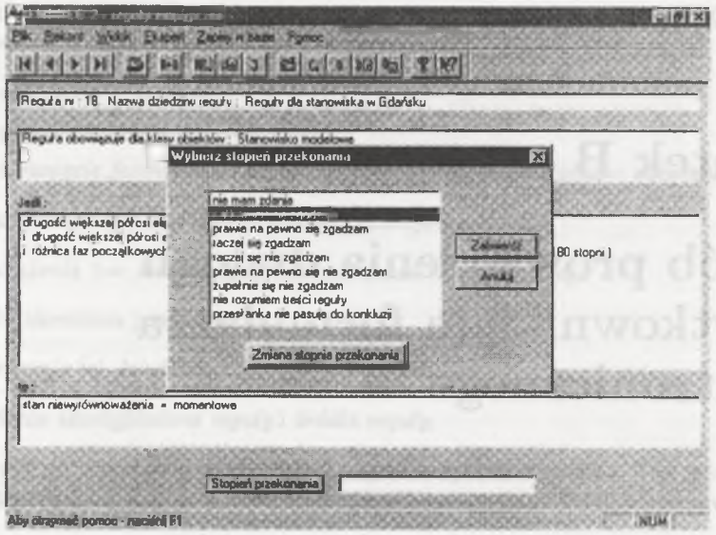
Sposób prowadzenia dialogu z użytkownikiem przedstawiono na kolejnych rysunkach[1].



Rys. B.1. Zawartość ekranu formularza "elektronicznego" w fazie oceniania reguły
Fig. B.1. Content of the screen of the "electronic" form in the phase of rule evaluation

Na rys. B.1 przedstawiono ekran wyświetlany przez program podczas działań w fazie *oceny reguł* (zwykle pozyskanych wcześniej, przede wszystkim z innego źródła). Wyświetlane są podstawowe części składowe reguły, takie jak: jej numer, nazwa dziedziny

[1]Przedstawiane kopie ekranów zostały przygotowane przez M. Wyleżoła. Uzyskano je w wyniku zastosowania aplikacji *EMPREG 2* [Wyleżoł, 1997a].



Rys. B.2. Wybór stopnia przekonania w celu oceny reguły
Fig. B.2. Selection of a belief factor for rule evaluation

reguły oraz wykaz klas i obiektów, dla których ta reguła może być zastosowana, złożona przesłanka w postaci koniunkcji selektorów oraz konkluzja. Nie jest wyświetlana wartość stopnia przekonania reguły (aby nie sugerować oceniającego).

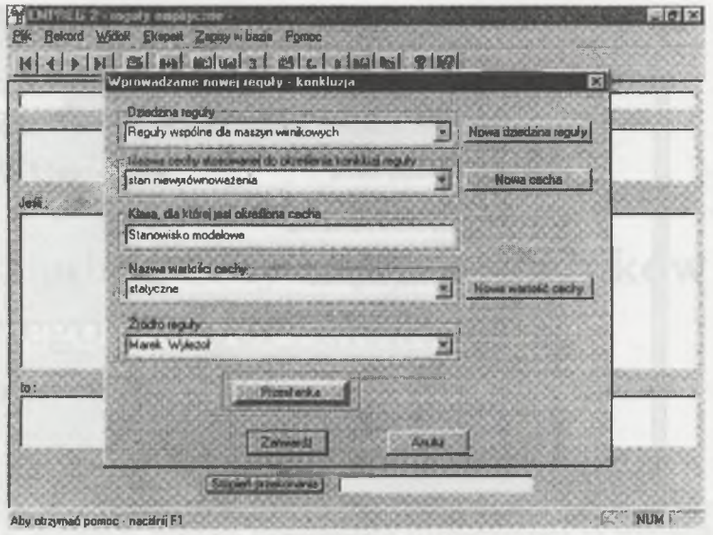
Oceniający wybiera jedną z wartości stopnia przekonania, określonych w roboczej bazie wiedzy (rys. B.2). Wybrana wartość zapisywana jest w roboczej bazie wiedzy formularza wraz z identyfikatorem autora oceny. Podlega ona następnie agregacji z oceną stopnia przekonania o słuszności tej reguły, przechowywaną wraz z ocenianą regułą.

Na rys. B.3 przedstawiono zawartość ekranu formularza podczas wprowadzania konkluzji reguły. Czynność ta polega na wyborze poszczególnych elementów składowych z odpowiednich list lub słowników. Nazwa klasy, dla której jest określona cecha, wyświetlana jest w celu sprawdzenia, czy cecha może być stosowana do zapisu reguł z danej dziedziny.

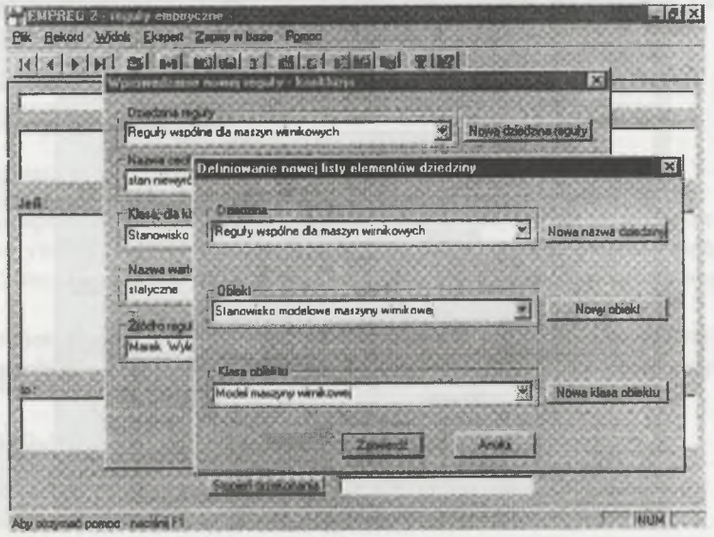
Zgodnie z przyjętym algorytmem postępowania, podczas definiowania reguły możliwe jest wprowadzanie nowych elementów, które właśnie są potrzebne specjaliście do zapisu danej reguły. Na rys. B.4 przedstawiono sposób określania dodatkowych elementów dziedziny reguły, którymi mogą być klasy obiektów lub obiekty-konkrety. W razie potrzeby specjalista może wprowadzić dane dotyczące nowych obiektów lub klas obiektów, które chce dołączyć do listy elementów rozpatrywanej dziedziny reguły.

Na rys. B.5 przedstawiono sposób wprowadzania danych określających nową cechę. Rodzaj i typ cechy są określone przez stałą zawartość bazy *EMPREL*, nie można więc tych słowników uzupełniać. W tym stadium mogą być natomiast uzupełnione pozostałe elementy opisu cechy.

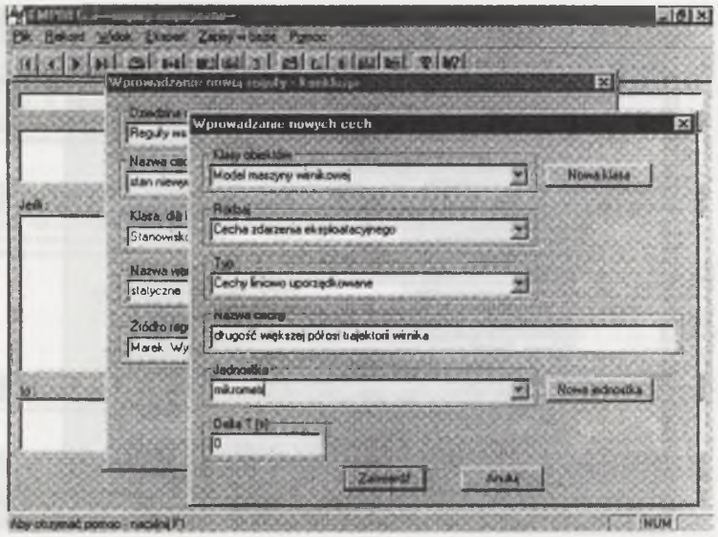
Wreszcie na rys. B.6 przedstawiono zawartość ekranu formularza podczas wprowadzania przesłanki reguły. Do określenia przesłanki wymagane jest określenie listy selektorów (warunków prostych) — por. opis na s. 47 i n.



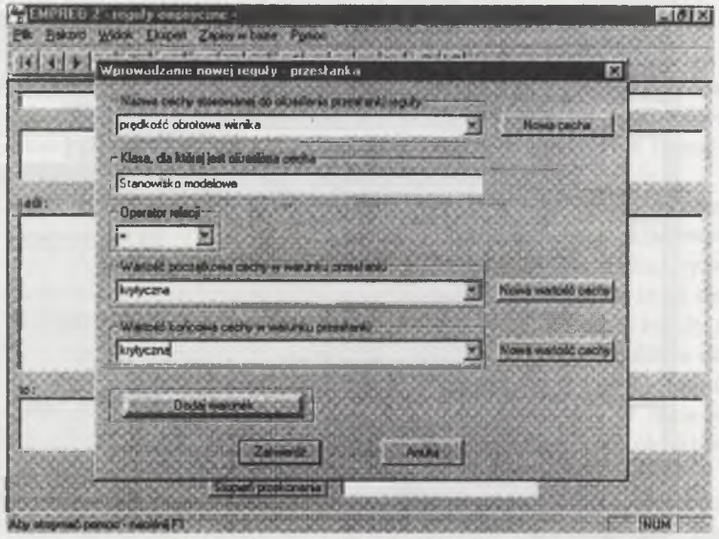
Rys. B.3. Zawartość ekranu formularza "elektronicznego" podczas określania dziedziny oraz konkluzji nowej reguły
Fig. B.3. Content of the screen of the "electronic" form in the phase of defining of a domain and conclusion of a new rule



Rys. B.4. Zawartość ekranu formularza "elektronicznego" podczas wprowadzania nowej listy elementów dziedziny reguły
Fig. B.4. Content of the screen of the "electronic" form when inputting a new list of elements of a domain of a rule



Rys. B.5. Sposób wprowadzania danych określających nową cechę
Fig. B.5. A way of inputting data that define a new attribute



Rys. B.6. Zawartość ekranu formularza "elektronicznego" podczas wprowadzania przesłanki nowej reguły
Fig. B.6. Content of the screen of the "electronic" form in the phase of inputting of a premise of a new rule

Dodatek C

Opis badań dotyczących wyników czynnego eksperymentu diagnostycznego

W dodatku zamieszczono opis badań przeprowadzonych z wykorzystaniem danych zgromadzonych w warunkach czynnego eksperymentu diagnostycznego na stanowisku *Rotor Kit*.

C.1 Identyfikacja zadania diagnostycznego

Zadanie diagnostyczne polega na diagnozowaniu stanu wirnika modelu *Rotor Kit*. Postanowiono rozpatrywać następujące niesprawności:

1. *różne postaci i wielkości niewyrównoważenia* wirnika;
2. *punktowe przycieranie wału*;
3. *przeciążenie łożysk wirnika*.

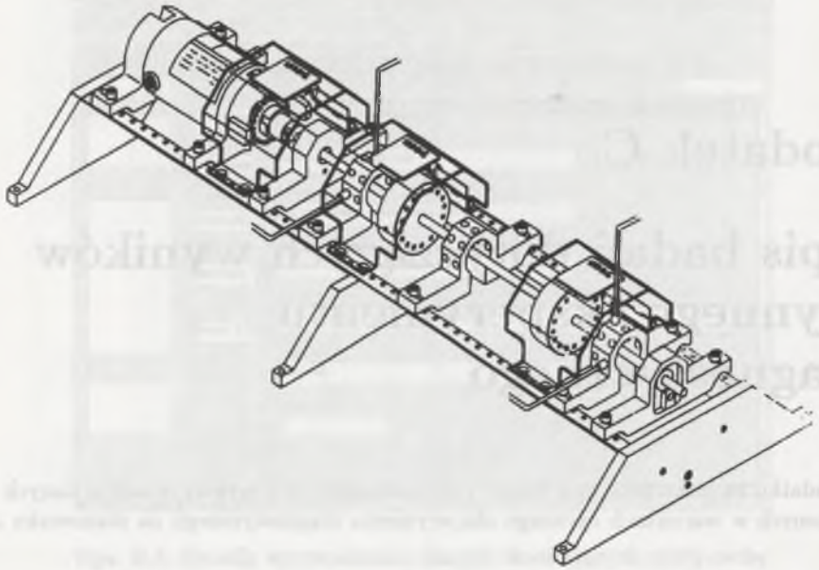
Określenie niesprawności winno nastąpić na podstawie znajomości symptomów diagnostycznych, kórymi są wartości cechy wielowymiarowych sygnałów przemieszczeń względnych wału. Zaleca się zastosowanie w tym celu cech opisujących trajektorie względnego ruchu wału.

Wskazane jest, aby ocena stanu wirnika była wspomagana za pomocą systemu doradczego, co wywołuje potrzebę pozyskania odpowiedniej "porcji" wiedzy diagnostycznej.

Stanowisko do prowadzenia czynnych eksperymentów diagnostycznych

W skład stanowiska do prowadzenia czynnych eksperymentów diagnostycznych wchodzą następujące elementy:

- model maszyny wirnikowej *Rotor Kit*;
- układ pomiarowy umożliwiający pomiar sygnałów oraz ich dyskretyzację i zapis do pomocniczej *bazy realizacji sygnałów diagnostycznych*.



Rys. C.1. Model maszyny wirnikowej *Rotor Kit* (wirnik posadowiony w łożyskach ślizgowych) [Bently Nevada, 1996]
Fig. C.1. The *Rotor Kit* model (the rotor supported in slide bearings) [Bently Nevada, 1996]

Model maszyny wirnikowej *Rotor Kit*. Ogólną postać modelu maszyny wirnikowej *Rotor Kit* wraz z układem czujników przedstawiono na rys. C.1. Do istotnych własności dynamicznych stanowiska należą wartości jego prędkości krytycznych oraz wrażliwość na zakłócenia działania wirnika, takie jak przycieranie o elementy nieruchome i podatność na przeciążenia. Własności te w zasadniczy sposób zależą od konfiguracji stanowiska, którą określają:

- *rozmieszczenie tarcz*, na których umieszczane są specjalne ciężarki powodujące niewyrównoważenie wirnika (dla stanowiska modelowego *Rotor Kit* czynnik ten ma największy wpływ na wartości prędkości krytycznych wirnika);
- *rodzaj łożysk* (w opisywanych badaniach wirnik był podparty w łożyskach ślizgowych smarowanych smarem stałym);
- *rozkład niewywag* na tarczach;
- *położenie wspornika* z otworem gwintowanym, w który wkręcana jest śruba pozwalająca na uzyskanie efektu przycierania, a także *tworzywo* tej śruby i *stan powierzchni* jej części dociskanej do wału;
- *położenie ramy* z dodatkowym łożyskiem zawieszonym za pomocą czterech sprężyn (niewidocznych na rys. C.1), za pomocą których można wywołać efekt przeciążenia.

Podczas przygotowań do eksperymentu czynnego przeprowadzono identyfikację prędkości krytycznych wirnika (związanych z precesją współbieżną [Łączkowski, 1983]), określając prędkości krytyczne dla ustalonej i nie zmienianej podczas eksperymentu konfiguracji stanowiska jako równe $n_{k,1} = 1720$ [min^{-1}] oraz $n_{k,2} = 5680$ [min^{-1}].

Konfiguracja stanowiska umożliwiała celowe wprowadzanie między innymi tych niesprawności, które opisano na s. 153, przy czym:

1. *niewyrównoważenie wirnika* powodowane jest przez odpowiednie rozmieszczenie dodatkowych mas na tarczach;
2. efekt *punktowego przycierania wału* osiąga się za pomocą śruby mosiężnej dociskanej do powierzchni wału, przy czym istnieje możliwość różnorodnej lokalizacji miejsca przycierania;
3. *przeciążanie łożysk wirnika* powodowane jest przez różne stany napięcia układu sprężyn podtrzymującego dodatkowe łożysko toczne osadzone na wale.

Układ pomiarowy. Układ pomiarowy umożliwia obserwację dwóch trajektorii ruchu wału względem ruchomego układu odniesienia, określonego przez wzajemnie do siebie prostopadłe osie par czujników, zlokalizowanych w płaszczyznach promieniowych w pobliżu łożysk ślizgowych (zob. rys. C.1). Rejestrowane mogą być dwa sygnały przemieszczeń względnych w każdej płaszczyźnie pomiarowej. Ponadto można obserwować sygnał znacznika położenia kątowego wału (ang. *KeyPhasor*), który umożliwia pomiar chwilowej prędkości obrotowej wirnika. Po demodulacji sygnały napięciowe proporcjonalne do przemieszczeń względnych są podawane na wejścia układu pomiarowego *SigLab* [SigLab, 1996]. Po filtracji dolnoprzepustowej sygnały są poddawane dyskretyzacji, co obejmuje próbkowanie chwil czasu (z odpowiednio dobraną częstotliwością próbkowania f_s) oraz kwantyzację amplitudy sygnału dyskretnego za pomocą 20-bitowego przetwornika analogowo-cyfrowego. W wyniku opisanego sposobu postępowania, którego przebiegiem steruje mikrokomputer interpretujący program zapisany w języku *MatLab* [Kostka, 1997], uzyskuje się realizacje czterowymiarowych sygnałów cyfrowych.

Dostępne źródła wiedzy diagnostycznej

W ramach opisanych badań możliwe było pozyskanie wiedzy z następujących źródeł:

- *literatura diagnostyczna*, jak np. [Cempel, 1989], [Gosiewski, Muszyńska, 1992], [Kiciński, 1996a], [Łączkowski, 1983], [Mitchell, 1981], [Morel, 1992], [Solartron, 1995], [Solartron, 1995a], [Żółtowski, 1996];
- *specjaliści* (zwłaszcza zespół dokonujący pomiarów na stanowisku *Rotor Kit* w laboratorium Katedry);
- *czynny eksperyment diagnostyczny* (jako źródło danych, które mogą być zastosowane do pozyskania wiedzy w sposób indukcyjny — z nauczycielem, określającym przyporządkowanie poszczególnych wyników pomiarów do klas stanu technicznego [Michalski, 1983]).

Zasoby dostępne w procesie pozyskiwania wiedzy

Proces pozyskiwania wiedzy prowadzony był w dwóch etapach, w których dostępne były różne środki programowe. I tak:

- we wstępnych badaniach prowadzonych przez P. Kostkę był to system doradczy wspomagający stosowanie wielowymiarowej analizy sygnałów, działający w środowisku *MatLab* [Kostka, 1997], [Kostka, 1997a], przy czym na szczególną uwagę, ze względu na cel badań prowadzonych dla potrzeb niniejszej rozprawy, zasługują takie jego elementy, jak:
  - oprogramowanie do pomiaru wielowymiarowych sygnałów diagnostycznych, obserwowanych na stanowisku modelowym,
  - oprogramowanie do estymacji wartości cech sygnałów (wstępny zbiór cech);
- w trakcie badań przeprowadzonych przez autora — system pozyskiwania wiedzy diagnostycznej *SPWD1* (opis w rozdziale 4.4), obejmujący między innymi:
  - bazę danych i wiedzy *EMPREL*,
  - oprogramowanie do pozyskiwania wiedzy od specjalistów,
  - pakiety programów do pozyskiwania wiedzy metodami indukcyjnymi (na podstawie bazy przykładów uczących).

Ponadto autor miał dostęp do danych pomiarowych (obejmujących *bazę realizacji sygnałów diagnostycznych* oraz wstępną wersję *bazy wartości cech sygnałów*), które zostały zgromadzone przez P. Kostkę w trakcie badań opisanych w [Kostka, 1997]. Badania te przeprowadzono w warunkach eksperymentu czynnego (schemat opisany w rozdziale C.3).

C.2 Określenie projektu bazy wiedzy *RK*

Ze względu na zakres zadania oraz bardzo mały zespół wykonawców projekt określał przede wszystkim sposoby:

1. reprezentacji danych,
2. reprezentacji wiedzy,
3. weryfikacji pozyskanej "porcji" wiedzy.

Do reprezentacji danych i wiedzy w sposób deklaratywny zastosowano bazę danych i wiedzy, która posiada schemat logiczny *EMPREL* (opis w rozdziale 4.1).

Reprezentacja wiedzy. Postanowiono zastosować deklaratywny sposób reprezentacji wiedzy stanowiącej *opis dziedziny zastosowań*, przede wszystkim określając nazwy cech oraz odpowiadające im zbiory wartości jakościowych.

Wiedza o sposobach uzyskania wartości cech zbioru realizacji sygnałów obserwowanych, a także wiedza dotycząca konwersji wartości ilościowych na jakościowe, została zawarta w pakiecie procedur opracowanych w środowisku *MatLab* oraz *Microsoft Excel*.

Wiedzę operacyjną stosowaną do wnioskowania diagnostycznego postanowiono reprezentować w sposób deklaratywny w postaci *reguł wnioskowania* oraz *drzew decyzyjnych*.

Reprezentacja danych. Przyjęto, że dane źródłowe (tj. realizacje czterowymiarowego sygnału cyfrowego) będą reprezentowane w formacie stosowanym w środowisku *MatLab*, w oddzielnej *bazie realizacji sygnałów diagnostycznych*. Uzasadnieniem takiego rozwiązania jest przewidywana potrzeba dalszych badań w zakresie optymalizacji przestrzeni reprezentacji cech sygnałów.

Wartości cech sygnałów diagnostycznych reprezentowane są za pomocą stwierdzeń. Stwierdzenia te zapisano w bazie danych i wiedzy *EMPREL*.

Sposoby weryfikacji pozyskanej "porcji" wiedzy. Postanowiono zastosować opracowane sposoby weryfikacji bazy wiedzy opisane w rozdziale 3.3. Ponieważ przewidywano pozyskiwanie wiedzy zarówno od specjalistów, jak i metodami indukcyjnymi na podstawie przykładów, możliwa była jednoczesna weryfikacja większości metod oceny bazy wiedzy, zwłaszcza metod oceny pod względem merytorycznym.

C.3 Opracowanie wersji pilotażowej bazy wiedzy oraz weryfikacja koncepcji

Zgromadzenie danych pomiarowych *RK*

Dane pomiarowe zostały zgromadzone przez P. Kostkę w trakcie badań opisanych w [Kostka, 1997]. Pomiary przeprowadzono w warunkach *eksperymentu czynnego*. Poniżej zamieszczono najważniejsze informacje mające znaczenie dla badań przeprowadzonych przez autora rozprawy (szczegółowy opis zawiera cytowana powyżej praca).

Schemat eksperymentu czynnego. W planie eksperymentu czynnego uwzględniono takie czynniki, jak:

1. *różne postaci i wielkości niewyrównoważenia wirnika*;
2. *punktowe przycieranie wału*;
3. *przeciążenie łożysk wirnika*.

Pełny schemat eksperymentu czynnego zawarty jest w pracy [Kostka, 1997]. Skrócone zestawienie zgromadzonych wyników eksperymentu zawiera tabl. C.1 (skrót "Przyc+przec" oznacza łączne występowanie przycierania i przeciążenia). Dla poszczególnych kombinacji stosowanych czynników podano liczbę dokonanych obserwacji (pomiarów). Pomiary te prowadzone były przy różnych prędkościach obrotowych n, odpowiadających zakresom prędkości obrotowych dla *strefy przedrezonansowej*, *strefy okołorezonansowej* i *strefy ponadrezonansowej*.

Analizując liczności obserwacji przeprowadzonych dla różnych kombinacji poszczególnych czynników (tabl. C.1), autor stwierdził, że w pracy [Kostka, 1997] przyjęto bardzo złożony plan eksperymentu, którego właściwości były następujące:

1. Plan ten obejmuje czynniki, które mogą być opisane przez trzy elementarne cechy stanu, w związku z czym uzyskuje się wiele kombinacji wartości tych cech stanu;

Tabl. C.1. Liczby obserwacji dokonanych dla różnych kombinacji czynników, uwzględnianych w eksperymencie czynnym na stanowisku *Rotor Kit*

Stan wyrównoważenia	Dodatkowe czynniki				Razem
	Brak	Przycieranie	Przeciążenie	Przyc+przec	
zgrubne wyrównoważenie dynamiczne	13	–	12	–	25
niewyrównoważenie momentowe	7	–	–	–	7
niewyrównoważenie quasi-statyczne	16	98	23	9	146
Razem	36	98	35	9	178

2. Klasy stanów nie są rozłączne w tym sensie, że symptomy poszczególnych uszkodzeń i niesprawności przesłaniają się;

3. W ramach jednej klasy zgrupowane są wyniki pomiarów zgromadzone podczas działania z różnymi prędkościami obrotowymi (ze strefy przedrezonansowej, strefy okołorezonansowej i strefy ponadrezonansowej — zob. tabl. C.2), co oznacza np. różne postaci linii ugięcia wału.

Tabl. C.2. Liczby obserwacji dokonanych dla różnych zakresów prędkości obrotowej wirnika stanowiska *Rotor Kit*

Zakres prędkości obrotowych [min^{-1}]	Liczba rejestracji
$n < n_{k,1}$	21
$n \simeq n_{k,1}$	17
$n_{k,1} < n < n_{k,2}$	104
$n \simeq n_{k,2}$	9
$n > n_{k,2}$	27
Razem	178

Podsumowując, można stwierdzić, że ze względu na znaczną złożoność eksperymentu należy oczekiwać niskich wartości oceny sprawności klasyfikatora. Ponieważ jednak w praktyce diagnostycznej często mamy do czynienia ze znacznymi ograniczeniami dotyczącymi możliwości przeprowadzenia eksperymentu czynnego, autor zdecydował się na podjęcie próby weryfikacji stosowanych metod z wykorzystaniem właśnie tych "trudnych" danych.

Wstępny zbiór cech sygnałów wielowymiarowych. W opisywanych badaniach obserwowano realizacje czterowymiarowego sygnału analogowego, który następnie, po poddaniu go filtracji dolnoprzepustowej (w celu wyeliminowania zjawiska *maskowania częstotliwości* [Cholewa, Moczulski, 1993]) próbkowano z częstotliwością $f_s = 512$ [Hz], 1280 [Hz] lub 2560 [Hz] uzależnioną od prędkości obrotowej wirnika n (która była stała

w czasie prowadzenia danego pomiaru). Wartości sygnału dyskretnego były następnie kwantowane za pomocą przetwornika analogowo-cyfrowego, w wyniku czego uzyskano realizacje czterowymiarowych sygnałów cyfrowych, zapisane następnie w oddzielnej *bazie realizacji sygnałów diagnostycznych.*

Do opisu postaci i niektórych ocen ilościowych trajektorii stosowano cechy opisane w [Wysogląd, 1996], które dla potrzeb badań wstępnych zostały zmodyfikowane w sposób opisany w [Kostka, 1997]. Opisy pełnego zbioru cech oraz określenia poszczególnych estymatorów zawarte są w cytowanych powyżej pracach.

Wstępna wersja bazy przykładów *RK*. Wstępna wersja *bazy przykładów RK* jest tablicą decyzyjną (zob. rozdział 3.2.1) o $N = 178$ wierszach. Liczba cech warunku wynosi $J = 32$, zaś cech decyzyjnych $L = 1$. Jedyna cecha decyzyjna określa numer klasy stanu (według podziału wstępnego na $K = 6$ klas — zob. tabl. C.3).

Tabl. C.3. Wyróżnione klasy stanu technicznego, stosowane we wstępnej wersji bazy przykładów *RK*

Nr klasy	Nazwa stanu technicznego
1	zgrubne wyrównoważenie dynamiczne
2	niewyrównoważenie momentowe
3	niewyrównoważenie quasi-statyczne
4	przycieranie
5	przeciążenie
6	przycieranie i przeciążenie

Pilotażowa wersja bazy wiedzy *RK*

Pilotażowa wersja bazy wiedzy została opracowana przez P. Kostkę w ramach badań opisanych w [Kostka, 1997]. Reguły zostały pozyskane metodą indukcyjną z zastosowaniem zbiorów przybliżonych (zob. rozdział 3.2.5) za pomocą udostępnionego programu *LERS* [Grzymała-Busse, 1994]. Otrzymano $M = 45$ reguł (pewnych i przybliżonych) umożliwiających klasyfikację obiektów do jednej z $K = 6$ klas stanu technicznego na podstawie wartości rozpatrywanych cech warunku.

Pilotażową wersję bazy wiedzy testowano za pomocą zbioru przykładów testowych (test *hold-out*, zob. rozdział 3.3.2), uzyskując bardzo niską wartość oceny $\bar{\eta}_{ov} = 54\%$ *łącznej sprawności klasyfikatora* wg (3.104).

Wnioski z weryfikacji pilotażowej wersji bazy wiedzy *RK*

Na podstawie analizy przebiegu badań opisanych w [Kostka, 1997] oraz występowania składowych dla częstotliwości charakterystycznych dla poszczególnych uszkodzeń autor sformułował następujące hipotezy dotyczące przyczyn małej sprawności klasyfikatora:

1. Schemat eksperymentu czynnego jest zbyt złożony (zob. tabl. C.1 oraz s. 157);
2. Nie we wszystkich przypadkach, które odpowiadają wynikom pomiarów przypisywanym do *zjawiska przycierania*, udało się uzyskać trajektorie ruchu wału typowe dla tego zjawiska (a więc takie, w których występowałyby składowe podharmoniczne częstotliwości obrotów wirnika, a zwłaszcza składowa 0.5X). Jednocześnie obserwowane są typowe symptomy *przeciążenia łożysk* (składowe harmoniczne, a zwłaszcza 2X), co oznacza, że zamiast przycierania wywołano *przeciążenie*[1];
3. Zbiór cech jest niepełny: brak uwzględnienia zależności fazowych pomiędzy trajektoriami zataczanymi przez przeciwległe końce wału.

W związku z powyższym autor postanowił:

- W odniesieniu do przyczyny 1. zastosować *dekompozycję zbioru przykładów* odpowiednio do prawdopodobnej struktury stanów technicznych (zob. rozdział 3.2.2).
- W odniesieniu do przyczyny 2. *nie modyfikować przypisania poszczególnych przykładów do klas stanu* (co byłoby możliwe w wyniku ponownej obserwacji postaci trajektorii ruchu wału, odtworzonych na podstawie zarejestrowanych realizacji sygnałów cyfrowych). Podjęcie takiej decyzji oznacza wprawdzie zmniejszenie pracochłonności badań przy opracowywaniu pełnej wersji bazy wiedzy, lecz jednocześnie nieuniknione jest uzyskanie niskich wartości ocen sprawności klasyfikatora.
- Uzupełnić zbiór cech, a następnie przygotować nową wersję zbioru przykładów *RK* zawierających również wartości nowych cech, aby usunąć przyczynę 3.

Opisane działania podjęto w ramach badań, prowadzonych pod kierunkiem autora.

C.4 Opracowanie pełnej wersji bazy wiedzy *RK* i jej weryfikacja

Do opracowania pełnej wersji bazy wiedzy zastosowano obie grupy metod:

- pozyskiwania wiedzy od specjalistów (w celu uzupełnienia "porcji" wiedzy pozyskanej w trakcie badań wstępnych),
- pozyskiwania wiedzy w sposób indukcyjny (metodami "uczenia maszynowego"),

oraz zastosowano środki działające zgodnie z tymi metodami.

Uzupełniające pozyskanie wiedzy od specjalistów

W tym stadium nastąpiło: pozyskanie nowej "porcji" wiedzy w postaci *empirycznych relacji diagnostycznych*, do czego zastosowano formularz "papierowy" (nie stosowano formularza "elektronicznego"), jej wprowadzenie do bazy danych i wiedzy za pomocą formularza "elektronicznego" oraz przedstawienie tej wiedzy do oceny innym specjalistom i wprowadzenie ich ocen do bazy wiedzy.

[1]Opisany problem może być interpretowany jako występowanie błędnych wartości cechy decyzyjnej [Grzymała-Busse, 1994]. Jego usunięcie wymaga dokonania *wstępnego przetwarzania bazy przykładów.*

Ponadto wprowadzono nowe cechy sygnałów przemieszczeń względnych walu, bazujące na gęstościach widmowych tych sygnałów (własnej i wzajemnej). Umożliwiło to wykorzystanie znanych relacji diagnostycznych, wiążących występowanie określonych dominant tych gęstości widmowych z typowymi niesprawnościami maszyn wirnikowych (zob. [Mitchell, 1981], [Cempel, 1982], [Solartron, 1995], [Solartron, 1995a]).

Zastosowanie formularza "papierowego" do pozyskania nowej wiedzy. Posługując się formularzem "papierowym" (opis w rozdziale 4.2.1), autor zapisał podstawowe empiryczne relacje diagnostyczne, które zaobserwował podczas działania stanowiska *Rotor Kit* w warunkach występowania niewyrównoważenia, przeciążenia i przycierania (przykładowe relacje przedstawiają rys. C.2÷C.4).

Zastosowanie formularza "elektronicznego" do wprowadzenia pozyskanej wiedzy. Relacje empiryczne zapisane za pomocą formularza "papierowego" (zob. powyżej) zostały zinterpretowane przez programistę bazy wiedzy, w wyniku czego do bazy wiedzy wprowadzono zbiór reguł diagnostycznych. Interpretacja dotyczyła przede wszystkim nazw cech oraz ich wartości jakościowych. Następnie elementy opisu dziedziny (nazwy i wartości cech) oraz reguły wnioskowania diagnostycznego zostały wprowadzone do bazy wiedzy *RK* za pomocą formularza "elektronicznego" (opis w rozdziale 4.2.3).

Ocena pozyskanej wiedzy przez niezależnych specjalistów. Treść reguł poddano ocenie za pomocą formularza "papierowego" wraz z ankietą (opis w rozdziale 4.2.2). Uzyskane oceny wprowadzono do bazy wiedzy, stosując ponownie formularz "elektroniczny".

Uzupełnienie opisu dziedziny zastosowań. Po przeanalizowaniu zbioru cech stosowanych do opracowania pilotażowej wersji bazy wiedzy, a także programu eksperymentu czynnego i występujących w nim czynników autor zaproponował rozszerzenie zbioru cech o cechy będące nośnikiem informacji o występowaniu:

- *przeciążeń* (podwyższone wartości amplitud składowych 2X, 3X);
- *przycierania* (podwyższone wartości amplitud składowych podharmonicznych 0.5X, 0.33X, 0.25X);
- *określonych postaci niewyrównoważenia* (jak np. niewyrównoważenie momentowe charakteryzujące się przeciwnymi fazami drgań naprzeciwległych końców wału).

Ponadto autor zaproponował sposób kwantyzacji wartości cech ilościowych uwzględniający możliwość stosowania ogólnie akceptowalnych nazw wartości jakościowych:

- dla wartości cech nie będących miarami kąta (np. wartości skutecznych amplitudy składowej częstotliwościowej widma drgań) w wyniku kwantyzacji winien być otrzymany zbiór zawierający co najwyżej 5 wartości, przy czym:
  - wartości bezwzględne przed zakodowaniem podlegają odpowiedniemu przekształceniu logarytmicznemu,
  - wartości względne przekształcane są w sposób liniowy;

RELACJA EMPIRYCZNA

UWAGA: proszę nie wypełniać pól w kolorze szarym.

		Identyfikator
Dla następujących obiektów (lub pojedynczego obiektu):	1 *Model maszyny wirnikowej Rotor Kit*	2
zlokalizowanych(nego) w:	3 *laboratorium Katedry PKM Politechniki Śląskiej (s. 14)*	4
będących w fazie eksploatacji:	5 *normalne użytkowanie do celów badawczych*	6
i działających (cego) w nast. warunkach[1]:	7 *prędkość obrotowa wirnika poniżej pierwszej prędkości krytycznej*	8
z następującym stanem technicznym:	9 *duże niewyrównoważenie o charakterze momentowym*	10

łączę takie symptomy tego stanu, jak:

	11 Symptom (podaj jego nazwę)	12 ma wartość (wpisz wartość liczbową lub określ jakościowo)	można obserwować[2]: 15 zawsze	16 często	17 rzadko	13 Identyfikator symptomu	14 Identyfikator wartości
a	*kształt trajektorii*	*eliptyczny*	X				
b	*wartość skuteczna amplitudy przemieszczeń względnych wału*	*znacznie podwyższona*	X				
c	*składowa 1X w widmie przemieszczeń względnych wału*	*dominująca*	X				
d	*znacznik położenia kątowego wału*	*1 na 1 obrót wału; nieruchomy*	X				
e	*przesunięcie fazowe m. sygnałami przem. względych. w obu łożyskach*	*około 180°*	X				

18 Próba krótkiego uzasadnienia związku między stanem technicznym a wypisanymi powyżej symptomami:
Dla prędk. obrotowej w zakresie podrezonansowym duże niewyrównoważenie o charakterze momentowym powoduje powstanie pary sił odśrodkowych, wywołujących ruch drgający wału o eliptycznych trajektoriach, znaczny wzrost wartości skutecznej amplitudy przemieszczeń drgań, przesunięcie fazowe między sygnałami przemieszczeń=180°; w widmie przemieszczeń dominuje składowa 1X.....

19 Cytowania, referencje: *Cz. Cempel, Diagnostyka wibroakustyczna maszyn, PWN, Warszawa 1989*
J. S. Mitchell, Machinery Analysis and Monitoring, PennWell, Tulsa, OK, 1981
B. Wysogląd, Metody reprezentacji drgań wałów maszyn wirnikowych w diagnostycznych bazach danych, pr. doktorska, KPKM Pol. Śl., Gliwice, 1996.

Data	Autor	Adres/tel/e-mail	Identyfikator:		Data otrz.:
			Wprowadził:	Nazwa pliku:	Data wpr.:

[1] Np. obciążenie, prędkość obrotowa, rodzaj produktu lub realizowanego działania itp.
[2] Proszę wstawić znak „X" w odpowiednim polu.

Rys. C.2. Przykładowa relacja empiryczna dotycząca niewyrównoważenia wirnika modelu *Rotor Kit*

Fig. C.2. An exemplary empirical relationship concerning imbalance of the *Rotor Kit* model

RELACJA EMPIRYCZNA

UWAGA: proszę nie wypełniać pól w kolorze szarym.

		Identyfikator
Dla następujących obiektów (lub pojedynczego obiektu):	1 *Model maszyny wirnikowej Rotor Kit*	2
zlokalizowanych(nego) w:	3 *laboratorium Katedry PKM Politechniki Śląskiej (s. 14)*	4
będących w fazie eksploatacji:	5 *normalne użytkowanie do celów badawczych*	6
i działających (cego) w nast. warunkach[1]:	7 *prędkość obrotowa wirnika poniżej pierwszej prędkości krytycznej*	8
z następującym stanem technicznym:	9 *znaczne przeciążenie łożysk*	10

łączę takie symptomy tego stanu, jak:

	11 Symptom (podaj jego nazwę)	12 ma wartość (wpisz wartość liczbową lub określ jakościowo)	można obserwować[2]: 15 zawsze	16 często	17 rzadko	13 Identyfikator symptomu	14 Identyfikator wartości
a	*kształt trajektorii*	*typowa ósemka*	X				
b	*wartość skuteczna amplitudy przemieszczeń wału*	*nieznacznie większa od stanu bez przec.*	X				
c	*znacznik położenia kątowego wału*	*widać 1 (na 2 obr. wału); nieruchomy*	X				
d	*amplituda składowej 2X w widmie przemieszczeń względnych wału*	*znacząca (równa ampl.składowej 1X)*	X				
e							

18 Próba krótkiego uzasadnienia związku między stanem technicznym a wypisanymi powyżej symptomami:
Dla prędk. obrotowej w zakresie podrezonansowym występują typowe symptomy dużego przeciążenia: trajektoria ruchu wału o kształcie ósemki, nieznaczny wzrost wartości skutecznej amplitudy przemieszczeń względnych wału oraz składowe wyższych rzędów (2X) w widmie przemieszczeń względnych.

19 Cytowania, referencje: *Cz. Cempel, Diagnostyka wibroakustyczna maszyn, PWN, Warszawa 1989*
J. S. Mitchell, Machinery Analysis and Monitoring, PennWell, Tulsa, OK, 1981
B. Wysogląd, Metody reprezentacji drgań wałów maszyn wirnikowych w diagnostycznych bazach danych, pr. doktorska, KPKM Pol. Śl., Gliwice, 1996.

Data	Autor	Adres/tel/e-mail	Identyfikator:		Data otrz.:
			Wprowadził:	Nazwa pliku:	Data wpr.:

[1] Np. obciążenie, prędkość obrotowa, rodzaj produktu lub realizowanego działania itp.
[2] Proszę wstawić znak „X" w odpowiednim polu.

Rys. C.3. Przykładowa relacja empiryczna dotycząca przeciążenia łożysk ślizgowych modelu *Rotor Kit*

Fig. C.3. An exemplary empirical relationship concerning excessive load of the slide bearings of the *Rotor Kit* model

RELACJA EMPIRYCZNA

UWAGA: proszę nie wypełniać pól w kolorze szarym.

		Identyfikator
Dla następujących obiektów (lub pojedynczego obiektu):	1 *Model maszyny wirnikowej Rotor Kit*	2
zlokalizowanych(nego) w:	3 *laboratorium Katedry PKM Politechniki Śląskiej (s. 14)*	4
będących w fazie eksploatacji:	5 *normalne użytkowanie do celów badawczych*	6
i działających (cego) w nast. warunkach[1]:	7 *prędkość obrotowa wirnika znacznie powyżej pierwszej prędkości krytycznej (bliska drugiej prędkości krytycznej)*	8
z następującym stanem technicznym:	9 *przycieranie wirnika o element nieruchomy (punktowe)*	10

łączę takie symptomy tego stanu, jak:

	11 Symptom (podaj jego nazwę)	12 ma wartość (wpisz wartość liczbową lub określ jakościowo)	można obserwować[2]: 15 zawsze	16 często	17 rzadko	13 Identyfikator symptomu	14 Identyfikator wartości
a	*kształt trajektorii*	*z jedną pętlą zewnętrzną*	X				
b	*wartość skuteczna amplitudy przemieszczeń wału*	*większa o 50% od stanu bez przycier.*	X				
c	*znaczniki położenia kątowego wału*	*2; nieruchome (nie przemieszczają się)*	X				
d	*amplituda składowej 1/2X w widmie przemieszczeń względnych wału*	*znacząca (ok. 50% ampl.składowej 1X)*	X				
e							

18 Próba krótkiego uzasadnienia związku między stanem technicznym a wypisanymi powyżej symptomami:

Dla prędkości obrotowej w zakresie ponadrezonansowym występują typowe symptomy przycierania: trajektoria ruchu wału z zewnętrzną pętlą, wzrost wartości skutecznej amplitudy przemieszczeń wału oraz znaczące składowe połówkowe (1/2X) w widmie przemieszczeń względnych wału.

19 Cytowania, referencje: *Cz. Cempel, Diagnostyka wibroakustyczna maszyn, PWN, Warszawa 1989*

J. S. Mitchell, Machinery Analysis and Monitoring, PennWell, Tulsa, OK, 1981

B. Wysogląd, Metody reprezentacji drgań wałów maszyn wirnikowych w diagnostycznych bazach danych, pr. doktorska, KPKM Pol. Śl., Gliwice, 1996.

Data	Autor	Adres/tel/e-mail			
			Identyfikator:		Data otrz.:
			Wprowadził:	Nazwa pliku:	Data wpr.:

[1] Np. obciążenie, prędkość obrotowa, rodzaj produktu lub realizowanego działania itp.
[2] Proszę wstawić znak „X" w odpowiednim polu.

Rys. C.4. Przykładowa relacja empiryczna dotycząca przycierania wirnika modelu *Rotor Kit*

Fig. C.4. An exemplary empirical relationship concerning rubbing of the *Rotor Kit* model

wartościom tym mogą być przypisywane takie nazwy jak np. "bardzo małe", "małe", "średnie", "duże", "bardzo duże";

- dla wartości cech, będących miarami kąta, w wyniku kwantyzacji otrzymuje się zbiór zawierający co najwyżej 12 elementów, którym przypisuje się nazwy odpowiadające położeniu godzin na tarczy zegara;
- dla wartości cech, będących różnicami kątów, w wyniku kwantyzacji otrzymuje się również zbiór zawierający co najwyżej 12 elementów, lecz nazwy wartości określa się jako "około 30 stopni", "około 60 stopni" itp.

Zaproponowane zmiany w odniesieniu do zbioru cech oraz sposobów kwantyzacji wartości ilościowych wymagały opracowania dodatkowych procedur umożliwiających uzyskanie zmodyfikowanej bazy przykładów *RK*.

Przygotowanie bazy przykładów uczących *RK*

Do reprezentacji przykładów zastosowano cechy opisujące postać i niektóre oceny ilościowe trajektorii. Dokonano optymalizacji zbioru cech ze względu na maksymalizację wartości oceny sprawności klasyfikatora, rozpoczynając te działania od zbioru cech trajektorii stosowanych w [Kostka, 1997]. Zbiór tych cech został następnie zmodyfikowany i uzupełniony według zaleceń autora o cechy bazujące na gęstościach widmowych (własnych i wzajemnych) [Moczulski, Kostka, 1997], [Kostka, 1997a].

Opisy pełnego zbioru cech oraz określenia estymatorów zawarte są w cytowanych powyżej pracach. Pełny wykaz nazw cech stosowanych do opisu poszczególnych przypadków dla potrzeb "uczenia maszynowego" zawarty jest w tabl. C.4. W rozprawie nie zamieszczono szczegółowego opisu wszystkich stosowanych estymatorów umożliwiających uzyskanie wartości stosowanych cech. Odpowiednie określenia zawiera [Moczulski, Kostka, 1997] (gdzie uwzględniono estymatory opisane w [Kostka, 1997] i [Kostka, 1997a] oraz dodano nowe estymatory, zaproponowane przez autora).

Baza przykładów RK jest tablicą decyzyjną (zob. rozdział 3.2.1) o $N = 178$ wierszach. Liczba cech warunku wynosi $J = 56$. Liczba cech decyzyjnych $L = 3$ (ze względu na złożoną strukturę zbioru stanów — zob. s. 69). W omawianej tablicy decyzyjnej brakuje wielu wartości cech, zwłaszcza odpowiadających niektórym podharmonicznym i harmonicznym wyższych rzędów, dla których nie da się z dostatecznie wysoką pewnością określić poziomów tych składowych[2]. Brakujące wartości cech mają wpływ na dalsze istotne obniżenie oceny sprawności klasyfikacji.

Pozyskanie wiedzy metodami indukcyjnymi

Pozyskiwanie wiedzy indukcyjnymi metodami "uczenia maszynowego" na podstawie przykładów stanowi najważniejszą część opisywanych badań przeprowadzonych dla danych *RK*.

Wstępne badanie dotyczyło indukcji reguł bez uwzględnienia struktury zbioru stanów, z zastosowaniem przykładów opisywanych za pomocą wstępnego zbioru cech warunku (wg

[2] W pracy [Moczulski, Kostka, 1997] zastosowano zaproponowane przez autora kryterium ostrości dominanty, które sprowadza się do wymagania, by przewyższała ona w statystycznie istotnym stopniu otaczające ją składowe gęstości widmowej.

Tabl. C.4. Pełny wykaz nazw cech stosowanych do zapisu bazy przykładów zgromadzonych w eksperymencie czynnym na stanowisku *Rotor Kit*

Lp.	Oznaczenie	Nazwa cechy
1	n	prędkość obrotowa, [min^{-1}]
2	dim1x	rozmiar trajektorii 1. w układzie obserwacji — kierunek x
3	dim1y	rozmiar trajektorii 1. w układzie obserwacji — kierunek y
4	xy1	stosunek rozmiarów trajektorii 1. w układzie obserwacji: x/y
5	dim1x-r	rozmiar trajektorii 1. w układzie obróconym — kierunek x
6	dim1y-r	rozmiar trajektorii 1. w układzie obróconym — kierunek y
7	xy1-r	stosunek rozmiarów trajektorii 1. w układzie obróconym: x/y
8	ang1	kąt obrotu układu współrzędnych trajektorii 1.; $[0, 2\pi)$
9	f-tr1	częstotliwość trajektorii 1., [Hz]
10	dir1-l	kierunek wirowania trajektorii 1.: lewy
11	dir1-r	kierunek wirowania trajektorii 1.: prawy
12	shape1	ocena kształtu trajektorii 1.
13	noise1	zaszumienie trajektorii 1., [V]
14	noise1-n	zaszumienie trajektorii 1. odniesione do średniej wymiarów trajektorii (w obróconym układzie współrzędnych)
15	dim2x	rozmiar trajektorii 2. w układzie obserwacji — kierunek x
16	dim2y	rozmiar trajektorii 2. w układzie obserwacji — kierunek y
17	xy2	stosunek rozmiarów trajektorii 2. w układzie obserwacji: x/y
18	dim2x-r	rozmiar trajektorii 2. w układzie obróconym — kierunek x
19	dim2y-r	rozmiar trajektorii 2. w układzie obróconym — kierunek y
20	xy2-r	stosunek rozmiarów trajektorii 2. w układzie obróconym: x/y
21	ang2	kąt obrotu układu współrzędnych trajektorii 2., [rad]
22	f-tr2	częstotliwość trajektorii 2., [Hz]
23	dir2-l	kierunek wirowania trajektorii 2.: lewy
24	dir2-r	kierunek wirowania trajektorii 2.: prawy
25	shape2	ocena kształtu trajektorii 2.
26	noise2	zaszumienie trajektorii 2., [V]
27	noise2-n	zaszumienie trajektorii 2. odniesione do średniej wymiarów trajektorii (w obróconym układzie współrzędnych)
28	05-1	składowa 0.5X widma sygnału 1.
29	1-1	składowa 1X widma sygnału 1.
30	2-1	składowa 2X widma sygnału 1.
31	3-1	składowa 3X widma sygnału 1.
32	025-2	składowa 0.25X widma sygnału 2.
33	033-2	składowa 0.33X widma sygnału 2.
34	05-2	składowa 0.5X widma sygnału 2.
35	1-2	składowa 1X widma sygnału 2.
36	2-2	składowa 2X widma sygnału 2.
37	3-2	składowa 3X widma sygnału 2.
38	025-3	składowa 0.25X widma sygnału 3.
39	033-3	składowa 0.33X widma sygnału 3.
40	05-3	składowa 0.5X widma sygnału 3.

Tablica C.4. (c.d.)

Lp.	Oznaczenie	Nazwa cechy
41	1-3	składowa 1X widma sygnału 3.
42	2-3	składowa 2X widma sygnału 3.
43	3-3	składowa 3X widma sygnału 3.
44	05-4	składowa 0.5X widma sygnału 4.
45	1-4	składowa 1X widma sygnału 4.
46	2-4	składowa 2X widma sygnału 4.
47	3-4	składowa 3X widma sygnału 4.
48	05-12ph	faza składowej 0.5X widma wzajemnego sygnałów 1. i 2.; $[0, 2\pi)$
49	1-12ph	faza składowej 1X widma wzajemnego sygnałów 1. i 2.; $[0, 2\pi)$
50	2-12ph	faza składowej 2X widma wzajemnego sygnałów 1. i 2.; $[0, 2\pi)$
51	05-34ph	faza składowej 0.5X widma wzajemnego sygnałów 3. i 4.; $[0, 2\pi)$
52	1-34ph	faza składowej 1X widma wzajemnego sygnałów 3. i 4.; $[0, 2\pi)$
53	2-34ph	faza składowej 2X widma wzajemnego sygnałów 3. i 4.; $[0, 2\pi)$
54	05-13ph	faza składowej 0.5X widma wzajemnego sygnałów 1. i 3.; $[0, 2\pi)$
55	1-13ph	faza składowej 1X widma wzajemnego sygnałów 1. i 3.; $[0, 2\pi)$
56	2-13ph	faza składowej 2X widma wzajemnego sygnałów 1. i 3.; $[0, 2\pi)$

[Kostka, 1997]), z którego wybrano podzbiór cech relewantnych. Biorąc pod uwagę niską sprawność uzyskanych klasyfikatorów, postanowiono następnie podjąć próbę dekompozycji zbioru przykładów z uwzględnieniem struktury zbioru stanów. Ponadto uzupełniono zbiór cech warunku o dalsze cechy, które stosowane są do zapisu reguł przez specjalistów z zakresu dynamiki maszyn wirnikowych oraz są spotykane w dostępnej literaturze.

Wstępna indukcja drzew decyzyjnych. Przed przeprowadzeniem indukcji drzew decyzyjnych dokonano wyboru podzbioru cech relewantnych, posługując się programem *AQ17-HCI*, w którym zastosowano algorytm *PROMISE 2* (zob. opis na s. 79). Otrzymano podzbiór $J_{rel} = 27$ cech warunku (pierwszych 27 cech w tabl. C.4) oraz jedną cechę decyzyjną ($L = 1$).

Do indukcji drzew decyzyjnych zastosowano program *C4.5* [Quinlan, 1993]. Ocenę sprawności klasyfikatora uzyskano w wyniku zastosowania trzykrotnie powtórzonego badania drzew decyzyjnych techniką *Random subsampling* ($N_I = 3$; zob. opis na s. 98), za każdym razem dzieląc losowo zbiór E zawierający $N = 178$ przykładów na podzbiór uczący E^u, zawierający 70% przykładów, oraz podzbiór testowy E^t, zawierający 30% przykładów. Zestawienie oceny *łącznego błędu względnego* wg (3.96) dla każdego losowania (dla uzyskanych drzew decyzyjnych oraz po ich przycięciu) zawiera tablica C.5.

Zbiorcze zestawienie wyników oceny sprawności klasyfikacji za pomocą drzew decyzyjnych (po ich przycięciu) zawiera tablica C.6.

Tablica C.7 zawiera dane dotyczące wyników klasyfikowania przykładów testowych dla wyróżnionych klas stanu. Na jej podstawie można stwierdzić, że niektóre klasy (np. o numerach 2, 6) nie są w ogóle rozpoznawane, podczas gdy klasa nr 4 jest rozpoznawana z niewielkim procentem błędów.

Ponadto obliczono wartości następujących błędów:

Tabl. C.5. Oceny łącznego błędu względnego dla poszczególnych losowań (indukcja drzew decyzyjnych dla wstępnej wersji bazy przykładów *RK*)

Nr losowania	Ocena łącznego błędu względnego [%]	
	przed przycinaniem	po przycinaniu
1	43.4	43.4
2	29.6	29.6
3	34.0	32.1

Tabl. C.6. Zbiorcze zestawienie wyników oceny sprawności klasyfikacji (indukcja drzew decyzyjnych dla wstępnej wersji bazy przykładów *RK*)

Nr klasy	Nazwa stanu technicznego	Łączna liczba przykładów				Udział przykładów poprawnie sklas. [%]
		testowych	poprawnie sklas.	pominiętych	niesłusznie zaliczonych	
1	zgrubne wyrównoważenie dynamiczne	10	2	8	5	20.0
2	niewyrównoważenie momentowe	8	0	8	4	0.0
3	niewyrównoważenie quasi-statyczne	13	7	6	7	53.8
4	przycieranie	88	77	11	28	87.5
5	przeciążenie	31	18	13	12	58.1
6	przycieranie i przeciążenie	10	0	10	0	0.0
Razem		160	104	56	56	65.0

- łącznego błędu względnego wg (3.96): $\widehat{\varepsilon}_{ov} = 35.0\%$,
- względnego błędu pominięcia wg (3.97): $\widehat{\varepsilon}_{om} = 63.4\%$,
- względnego błędu niesłusznego zaliczenia wg (3.97): $\widehat{\varepsilon}_{cm} = 9.8\%$.

Przedstawione wyniki potwierdzają wstępną ocenę, że ze względu na złożony plan eksperymentu nie jest możliwe uzyskanie klasyfikatora o wysokiej sprawności.

Wstępna indukcja reguł metodą pokryć. Dla wyżej opisanego zbioru $N = 178$ przykładów, zapisanego za pomocą tablicy decyzyjnej o $J = 56$ cechach warunku i $L = 1$ cesze decyzyjnej, dokonano wstępnej indukcji reguł metodą pokryć. Ocenę sprawności klasyfikatora uzyskano w wyniku trzykrotnie powtórzonego badania techniką *Random subsampling* (podział zbioru przykładów uczących w stosunku 70:30), uzyskując średnią

Tabl. C.7. Wyniki klasyfikowania przykładów testowych (indukcja drzew decyzyjnych dla wstępnej wersji bazy przykładów *RK*)

Nr klasy	Nazwa stanu technicznego	Liczba przykładów sklasyfikowanych jako należących do klasy o numerze					
		1	2	3	4	5	6
1	zgrubne wyrównoważenie dynamiczne	2	2	1	2	3	–
2	niewyrównoważenie momentowe	–	–	5	3	–	–
3	niewyrównoważenie quasi-statyczne	–	2	7	4	–	–
4	przycieranie	3	–	–	77	8	–
5	przeciążenie	2	–	1	10	18	–
6	przycieranie i przeciążenie	–	–	–	9	1	–

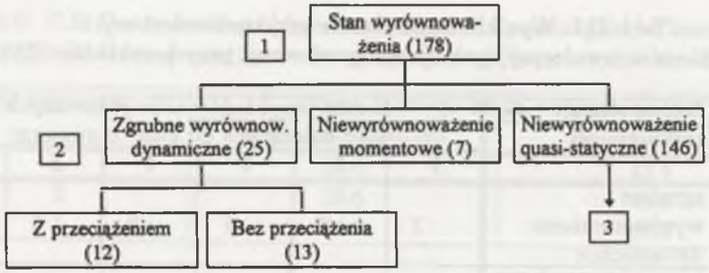
wartość *łącznego błędu względnego* $\widehat{\varepsilon}_{ov} = 41.3\%$. Wartość ta jest znaczna, co potwierdza wstępną ocenę autora, że nie jest możliwe uzyskanie klasyfikatora stanu o wysokiej sprawności.

Wstępna indukcja reguł metodą zbiorów przybliżonych. Do indukcji reguł zastosowano zmodyfikowany algorytm *LEM2*[3], który umożliwia pozyskanie reguł na podstawie przykładów uczących, biorąc pod uwagę jedynie te cechy warunku, dla których we wszystkich przykładach określone są ich wartości [Grzymała-Busse, Wang, 1997]. Wartości oceny łącznego błędu względnego uzyskano techniką *10-fold cross validation*, otrzymując wartość $\widehat{\varepsilon}_{ov} = 34.3\%$. Przy klasyfikowaniu przykładów testowych na podstawie reguł wzięto pod uwagę wszystkie cechy (także te, których wartości nie były określone dla jakiegoś podzbioru przykładów). Uzyskana wartość oceny łącznego błędu względnego jest bardzo zbliżona do ocen obliczonych dla klasyfikatorów otrzymanych za pomocą innych metod indukcyjnych.

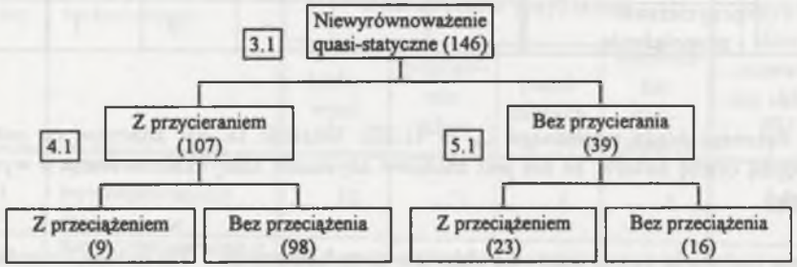
Struktura drzewa stanów. Autor postawił hipotezę, że przyczyną niskiej sprawności klasyfikatorów zidentyfikowanych w ramach badań wstępnych była zbyt złożona struktura zbioru stanów technicznych, wywoływanych w ramach eksperymentu czynnego prowadzonego za pomocą modelu *Rotor Kit*. Postanowiono wprowadzić różne stopnie szczegółowości diagnoz oraz dążyć do zwiększenia stopnia szczegółowości diagnozy, postępując "od ogółu do szczegółu".

Zastosowany algorytm optymalizacji struktury zbioru stanów przypomina algorytm budowy optymalnego drzewa decyzyjnego (zob. rozdział 3.2.6). Wprowadzono trzy pomocnicze cechy decyzyjne (będące elementarnymi cechami stanu):

[3]Obliczenia zostały wykonane przez J. W. Grzymałę-Busse, autora oprogramowania, w którym zastosowano algorytm *LEM2*.



Rys. C.5. Drzewo stanów technicznych dla danych *RK*
Fig. C.5. A tree of technical states of the *RK* data

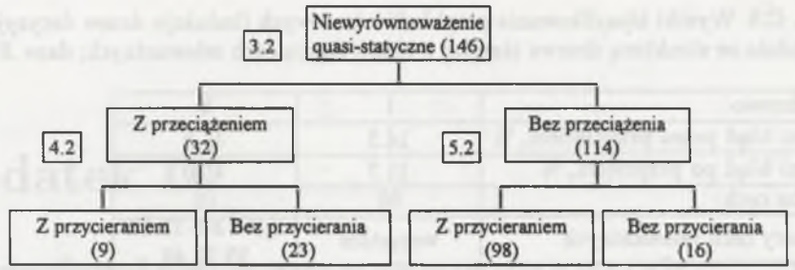


Rys. C.6. Poddrzewo stanów technicznych dla danych *RK*
(cecha *przycieranie*)
Fig. C.6. A subtree of technical states of the *RK* data (attribute *rubbing*)

- cecha *stan wyrównoważenia* o wartościach:
  {"zgrubnie wyrównoważony dynamicznie", "niewyrównoważenie momentowe", "niewyrównoważenie quasi-statyczne" } ;
- cecha *przeciążenie łożysk* o wartościach:
  {"występuje", "nie występuje" };
- cecha *przycieranie* o wartościach:
  {"występuje", "nie występuje" }.

Uzyskane drzewo stanów przedstawiono na rys. C.5. Poddrzewo dla cechy "przycieranie" przedstawiono na rys. C.6, zaś poddrzewo dla cechy "przeciążenie" — na rys. C.7.

Powtórna indukcja drzew decyzyjnych. Indukcję drzew decyzyjnych przeprowadzono zgodnie ze strukturą drzewa stanów, jednocześnie dokonując optymalizacji struktury tego drzewa. Dla poszczególnych poddrzew wykonywano następujące operacje:



Rys. C.7. Poddrzewo stanów technicznych dla danych *RK*
(cecha *przeciążenie*)
Fig. C.7. A subtree of technical states of the *RK* data (attribute *excessive load*)

Tabl. C.8. Wyniki klasyfikowania przykładów testowych (indukcja drzew decyzyjnych zgodnie ze strukturą drzewa stanów; bez identyfikacji cech relewantnych; dane *RK*)

Poddrzewo	1	2	
Średni błąd przed przycięciem, %	15.0	45.8	
Średni błąd po przycięciu, %	12.5	45.8	
Poddrzewo	3.1	4.1	5.1
Średni błąd przed przycięciem, %	17.4	13.3	17.4
Średni błąd po przycięciu, %	17.4	7.2	17.4
Poddrzewo	3.2	4.2	5.2
Średni błąd przed przycięciem, %	18.2	36.7	4.8
Średni błąd po przycięciu, %	19.9	43.3	5.7

1. wybór podzbioru przykładów odpowiadających klasie stanu przypisanej do korzenia tego poddrzewa,
2. identyfikacja podzbioru cech relewantnych (z użyciem algorytmu *PROMISE 2*),
3. identyfikacja dwóch poddrzew decyzyjnych: poddrzewa zbudowanego z zastosowaniem wszystkich 56 cech warunku oraz poddrzewa zbudowanego z zastosowaniem podzbioru cech relewantnych (dla tego węzła), a następnie ocena względnego błędu klasyfikacji uzyskanego po zastosowaniu każdego z analizowanych poddrzew.

Podzbiory cech relewantnych odpowiadających poszczególnym węzłom drzewa stanów są zwykle o mniejszej liczności niż pełny zbiór 56 cech.

Wartości błędu względnego klasyfikacji dla poszczególnych poddrzew uzyskanych bez identyfikacji podzbioru cech relewantnych zestawiono w tabl. C.8.

Wartości błędu względnego klasyfikacji dla poszczególnych poddrzew uzyskanych przy wymogu każdorazowej identyfikacji podzbioru cech relewantnych zestawiono w tabl. C.9. Należy nadmienić, że obie te tablice uzyskano dla wyników klasyfikacji takich samych zbiorów przykładów testowych.

Tabl. C.9. Wyniki klasyfikowania przykładów testowych (indukcja drzew decyzyjnych zgodnie ze strukturą drzewa stanów; z identyfikacją cech relewantnych; dane *RK*)

Poddrzewo	1	2	
Sredni błąd przed przycięciem, %	14.2	45.8	
Średni błąd po przycięciu, %	11.7	45.0	
Liczba cech	56	10	
Numery cech relewantnych	wszystkie	2 3 4 8 9 12 13 25 27 49	
Poddrzewo	3.1	4.1	5.1
Średni błąd przed przycięciem, %	14.4	15.5	11.1
Średni błąd po przycięciu, %	18.2	9.3	11.1
Liczba cech	15	56	11
Numery cech relewantnych	1 2 3 4 8 12 13 14 17 21 25 27 49 52 55	wszystkie	wszystkie
Poddrzewo	3.2	4.2	5.2
Średni błąd przed przycięciem, %	17.4	43.3	7.8
Średni błąd po przycięciu, %	16.7	40.0	8.8
Liczba cech	56	56	14
Numery cech relewantnych	wszystkie	1 2 3 8 9 15 25 27 49 52 55	1 2 4 8 9 12 13 14 21 25 27 29 49 55

Połączenie wiedzy pozyskanej z różnych źródeł

Połączenie wiedzy pozyskanej z różnych źródeł możliwe było ze względu na przyjęty wspólny sposób reprezentacji wiedzy diagnostycznej w systemie *SPWD1*. Nie stwierdzono występowania błędów formalnych w bazie reguł (zob. rozdział 3.3, s. 92). Nie stwierdzono także potrzeby łączenia ocen reguł, gdyż nie wystąpiły reguły nadmiarowe o różnych wartościach stopnia przekonania.

Uwagi dotyczące stosowania bazy wiedzy

Pozyskana wiedza diagnostyczna zawarta w bazie wiedzy *RK* może być próbnie stosowana na stanowisku modelowym *Rotor Kit*. Przed jej zastosowaniem wymagana jest aktualizacja zawartości tablicy decyzyjnej reguł używanej w tym programie. Zidentyfikowana struktura drzewa stanów może być wykorzystana do odpowiedniej dekompozycji tablicy decyzyjnej reguł na podtablice o mniejszych rozmiarach.

Dodatek D

Opis badań przeprowadzonych w ramach eksperymentu numerycznego

W dodatku zamieszczono opis badań przeprowadzonych z wykorzystaniem danych uzyskanych w ramach eksperymentu numerycznego. W eksperymencie wykorzystano model matematyczny oraz system programów *MESWIR* opisany w [Kiciński, 1996a].

D.1 Identyfikacja zadania diagnostycznego

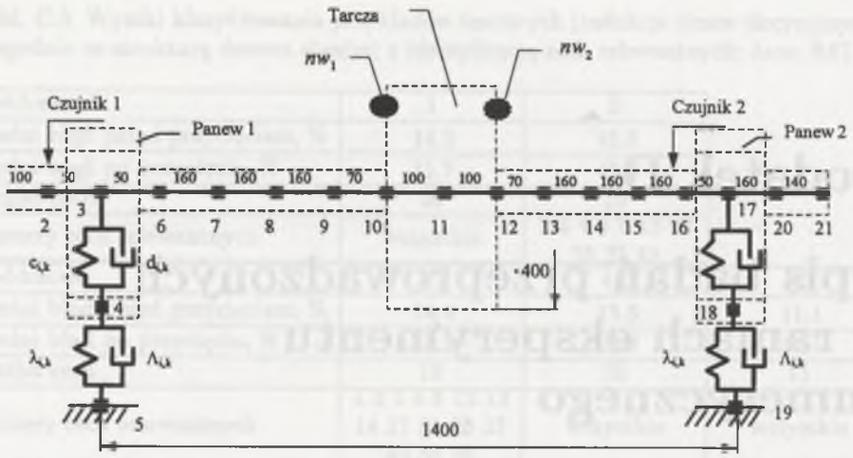
Zadanie diagnostyczne polega na *identyfikacji stanu wyrównoważenia wirnika* podpartego w dwóch łożyskach hydrodynamicznych i wirującego z jedną z prędkości obrotowych: podrezonansową, rezonansową lub ponadrezonansową. Należy rozpoznać następujące stany:

1. *wirnik zgrubnie wyrównoważony dynamicznie*,
2. *niewyrównoważenie statyczne*,
3. *niewyrównoważenie quasi-statyczne*,
4. *niewyrównoważenie momentowe*,
5. *niewyrównoważenie dynamiczne*.

Określenie stanu wyrównoważenia winno nastąpić na podstawie znajomości symptomów diagnostycznych, którymi są wartości cech trajektorii ruchu (bezwzględnego i względnego) wybranych węzłów wału. Dla potrzeb opisanych badań cechy te wyznaczane były przez odpowiednie oprogramowanie umożliwiające prowadzenie obliczeń symulacyjnych.

Opis modelu obiektu stosowanego w eksperymencie numerycznym

Modelowany obiekt (maszyna wirnikowa) umożliwia prowadzenie badań własności dynamicznych układu: *wirnik – podpory łożyskowe*. Zespół wirnika tego stanowiska jest złożony z wału oraz osadzonej na nim tarczy o znacznej masie. Zespół ten łożyskowany jest w dwóch łożyskach hydrodynamicznych. Dla potrzeb badań symulacyjnych utworzono



Rys. D.1. Teoretyczny model wirnika (na podstawie [Kiciński *et al.*, 1995], z modyfikacjami)
Fig. D.1. A theoretic model of the rotor (by [Kiciński *et al.*, 1995], modified)

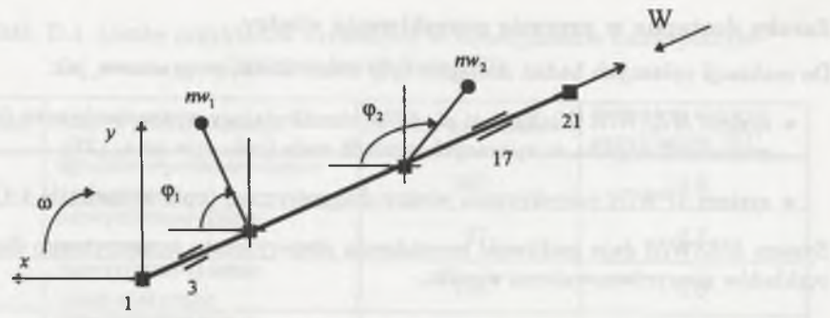
model stanowiska badawczego przedstawiony na rys. D.1. Opis stanowiska oraz oryginalnej wersji modelu zawiera m. in. praca [Kiciński *et al.*, 1995]. Modyfikacja modelu dla potrzeb opisanych badań polegała na innym podziale wału za pomocą węzłów.

Ponieważ model opisuje właściwości układu konkretnego, możliwe jest bezpośrednie porównanie wyników obliczeń symulacyjnych z wynikami pomiarów przeprowadzonych na stanowisku. Umożliwia to weryfikację wielu danych uczących.

Przedmiotem badań symulacyjnych są drgania wymuszone spowodowane niewyrównoważeniem przyłożonym w różnych płaszczyznach wirnika. Stosowane są dwie *niewywagi* nw_1, nw_2 przykładane w płaszczyznach promieniowych przecinających geometryczną oś wału w wybranych węzłach. *Kąty fazowe niewywag* ϕ_1, ϕ_2 wraz z *numerami węzłów* loc_1, loc_2 w pełni identyfikują położenie niewywag. Wymienionych sześć wielkości jednoznacznie określa *rozkład niewyrównoważenia* (rys. D.2).

Celem eksperymentu numerycznego było wyznaczenie (na drodze obliczeń komputerowych z zastosowaniem programów systemu *MESWIR* [Kiciński *et al.*, 1995]) ocen drgań wymuszonych wirnika w wybranych węzłach wału odpowiadających lokalizacji czujników drgań na modelowanym obiekcie konkretnym. Podobnie jak w [Kiciński *et al.*, 1995] przyjęto, że przedmiotem analizy będą zjawiska zachodzące w układzie działającym w warunkach nie powodujących utraty przez ten układ stabilności hydrodynamicznej (objawiającej się wystąpieniem wiru, a nawet bicza olejowego [Gosiewski, Muszyńska, 1992]). Także wielkości niewywag dobierano w taki sposób, by można było zjawiska zachodzące w układzie *wirnik – podpory łożyskowe* uznać za liniowe.

Ponieważ w każdej z płaszczyzn pomiarowych rzeczywistego stanowiska zainstalowane są po dwa czujniki o osiach największej czułości wzajemnie do siebie prostopadłych, możliwa była estymacja wartości cech trajektorii ruchu środka czopa (zarówno bezwzględnego, jak i względem podpory łożyska). Model liniowy przyjęty przy opracowaniu zastosowa-



Rys. D.2. Przykładowe rozmieszczenie niewywag [Kiciński *et al.*, 1995]
Fig. D.2. An exemplary location of imbalances [Kiciński *et al.*, 1995]

nego oprogramowania (zob. s. 123) oraz wymienione powyżej ograniczenia powodują, że trajektorie te są elipsami. Opisany model teoretyczny charakteryzuje siedem parametrów:

- wartości niewywag nw_1, nw_2;
- kąty fazowe niewywag ϕ_1, ϕ_2;
- lokalizacje płaszczyzn loc_1, loc_2, na których leżą wektory niewywag;
- prędkość obrotowa n wirnika.

Pozostałe dane warunkujące przeprowadzenie obliczeń symulacyjnych dotyczą: cech geometrycznych wirnika i łożysk, sztywności podpór łożyskowych oraz temperatury oleju. Przyjęto je za [Kiciński *et al.*, 1995] (warto wspomnieć, że odpowiadają one wartościom cech określonym podczas eksperymentów prowadzonych na konkretnym obiekcie).

Dostępne źródła wiedzy diagnostycznej

W ramach opisanych badań możliwe było pozyskanie wiedzy z następujących źródeł:

- *literatura diagnostyczna* (zob. s. 128),
- *specjaliści* (zwłaszcza zespół dokonujący pomiarów na stanowisku badawczym zlokalizowanym w Instytucie Maszyn Przepływowych PAN w Gdańsku),
- *eksperyment numeryczny* (jako źródło danych, które mogą być zastosowane do pozyskania wiedzy w sposób indukcyjny).

Możliwość przeprowadzenia eksperymentu numerycznego stanowiła szczególny czynnik wyboru opisanego modelu maszyny wirnikowej jako obiektu badań.

Zasoby dostępne w procesie pozyskiwania wiedzy

Do realizacji opisanych badań dostępne były różne środki programowe, jak:

- system *MESWIR* [Kiciński *et al.*, 1995] umożliwiający wyznaczenie ocen drgań wymuszonych wirnika w wybranych węzłach wału (zob. opis na s. 123);
- system *SPWD1* pozyskiwania wiedzy diagnostycznej (opis w rozdziale 4.4).

System *MESWIR* daje możliwość prowadzenia eksperymentu numerycznego dla różnych rozkładów niewyrównoważenia wirnika.

D.2 Określenie projektu bazy wiedzy *GD*

Ze względu na prostotę zadania projekt obejmował jedynie sposoby reprezentacji danych i wiedzy oraz weryfikacji bazy wiedzy.

Przyjęto, że dane i wiedza będą reprezentowane w sposób deklaratywny, w postaci stwierdzeń, reguł i drzew decyzyjnych. Dane i wiedza winny być gromadzone w odpowiedniej bazie *EMPREL*.

D.3 Opracowanie wersji pilotażowej bazy wiedzy oraz weryfikacja koncepcji

Pilotażowa wersja bazy wiedzy została uzyskana w sposób indukcyjny na podstawie wstępnej wersji bazy danych uczących *GD*, której swoistą cechą były znaczne różnice w licznościach przykładów przypisanych do poszczególnych klas (szczegółowy opis wstępnej wersji bazy przykładów *GD* oraz wartości poszczególnych parametrów rozkładu niewyrównoważenia zawiera raport [Moczulski, Maniak, 1997]). Poniżej przedstawiono jedynie najważniejsze wyniki wpływające na przebieg dalszych prac związanych z badaniami dotyczącymi pozyskiwania wiedzy ze zbioru przykładów uzyskanych w wyniku eksperymentu numerycznego.

Wstępna wersja bazy *GD* wyników eksperymentu numerycznego

Wstępną wersję bazy wyników eksperymentu numerycznego przygotował P. Maniak według schematu zaproponowanego przez autora. Za pomocą udostępnionego systemu programów *MESWIR* obliczono parametry elips drgań w wyróżnionych węzłach wału. Dokonano także optymalizacji sposobu reprezentacji cech elips drgań ze względu na kryterium sprawności klasyfikatora. Tak przygotowaną wstępną wersję bazy *GD* zastosowano do pozyskiwania wiedzy metodą indukcji. Wnioski opracowane na podstawie tej fazy badań umożliwiły modyfikację planu eksperymentu oraz wprowadzenie nowych cech elips drgań.

Wstępny schemat eksperymentu numerycznego. Według wstępnego schematu eksperymentu numerycznego wygenerowano zbiór wyników obliczeń, reprezentujących parametry elips drgań dla poszczególnych klas, przy czym liczności wyników obliczeń, odpowiadające poszczególnym klasom, znacznie różnią się pomiędzy sobą (zob. tabl. D.1).

Tabl. D.1. Liczby przykładów uzyskanych w eksperymencie numerycznym (wstępna wersja bazy *GD*)

Nr klasy	Stan wyrównoważenia	Liczba przykładów	Udział w zbiorze przykładów [%]
1	zgrubne wyrównoważenie dynamiczne	840	23.6
2	niewyrównoważenie statyczne	27	0.7
3	niewyrównoważenie quasi-statyczne	108	3.0
4	niewyrównoważenie momentowe	324	9.1
5	niewyrównoważenie dynamiczne	2268	63.6
Razem		3567	100.0

Optymalizacja sposobu reprezentacji cech elips drgań. Program *DYNWIR-W* dla zadanych parametrów rozkładu niewywag oraz prędkości obrotowej (przy ustalonych wartościach pozostałych danych warunkujących przeprowadzenie obliczeń symulacyjnych) oblicza dla każdej trajektorii wartości czterech współczynników A, B, C, D w pełni opisujących te trajektorie, przy czym:

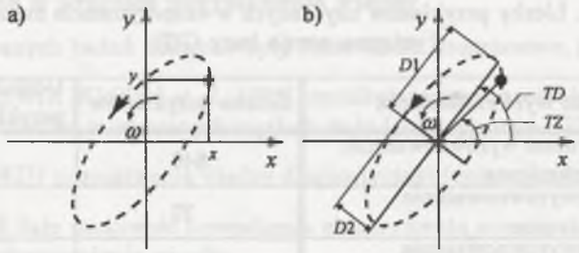
$$x = A\sin\tau + B\cos\tau,$$
$$y = C\sin\tau + D\sin\tau, \tag{D.1}$$

gdzie $\tau = \omega t$ jest bezwymiarowym czasem [rad], zaś x, y są odpowiednio składowymi: poziomą i pionową elipsy drgań (zob. rys. D.3.a).

Pomimo że zależności (D.1) w pełni opisują trajektorie ruchu środka walu względem panewki łożyska, opis ten dla potrzeb diagnostycznych jest nieprzydatny. Dla tych potrzeb (co wynika między innymi ze sposobu oceny trajektorii przez specjalistów) istotnymi cechami trajektorii są:

- *długości półosi elipsy*,
- *stosunek długości dłuższej półosi do długości krótszej półosi* (będący oceną wydłużenia trajektorii),
- *kąt pochylenia dłuższej półosi elipsy*,
- *kierunek obiegu elipsy* (określający rodzaj precesji: współbieżna lub przeciwbieżna),
- *przesunięcie fazowe pomiędzy promieniami trajektorii* zataczanych przez dwa przeciwległe końce wału.

Parametrów tych nie da się ocenić bezpośrednio na podstawie znajomości parametrów A, B, C, D (bez dodatkowych obliczeń). Z tych względów autor wprowadził do opisu trajektorii zbiór cech, które przedstawiono na rys. D.3.b. Ponadto wprowadzono cechy określające różnice kątów pochylenia dłuższych półosi elips i różnice faz początkowych określane dla trajektorii drgań, zarówno względnych, jak i bezwzględnych, obserwowanych na obu przeciwległych czopach wału wirnika.



Rys. D.3. Parametry opisu trajektorii: a) stosowane w programie *DYNWIR-W* [Kiciński *et al.*, 1995], b) stosowane dla potrzeb "uczenia maszynowego"
Fig. D.3. Parameters describing orbits: a) used in the *DYNWIR-W* program [Kiciński *et al.*, 1995], b) applied for "machine learning" purposes

Pełny wykaz nazw cech stosowanych do opisu trajektorii drgań bezwzględnych i względnych, których parametry obliczane były za pomocą systemu programów *MESWIR* w ramach prowadzonego eksperymentu numerycznego, zamieszczono w tabl. D.2.

Optymalizacja sposobów kwantowania wartości cech ilościowych. W celu kwantowania wartości cech ilościowych, na podstawie analizy postaci histogramów wartości tych cech dokonano wyboru sposobów określania ograniczeń podziałów π_a wg (2.9) zbiorów wartości poszczególnych cech $Dom(a)$, $a \in A$:

- w odniesieniu do rozmiarów liniowych trajektorii postanowiono wartości progowe podziałów określać w sposób logarytmiczny, wstępnie przyjmując 10 wartości jakościowych;
- w odniesieniu do miar kąta (kąty pochylenia głównej osi elipsy oraz różnice kątów) postanowiono zastosować podział liniowy, wstępnie przyjmując 10 wartości.

Indukcja reguł na podstawie wstępnej wersji bazy *GD*. Pilotażowa wersja bazy wiedzy *GD* zawiera opis dziedziny (dotyczący określenia nazw cech oraz zbioru ich wartości jakościowych, w tej fazie badań reprezentowanych za pomocą identyfikatorów liczbowych) oraz zbiór reguł pozyskanych indukcyjnie na podstawie wstępnej wersji bazy przykładów *GD*. W tablicy D.3 zamieszczono wyniki cząstkowe i łączną ocenę sprawności obu klasyfikatorów, uzyskanych poprzez indukcję z dwóch różniących się zbiorów przykładów, z których pierwszy zawiera wartości jakościowe cech uzyskane wskutek kwantowania z zastosowaniem progów określonych w skali liniowej, zaś drugi — z zastosowaniem progów określonych w skali logarytmicznej.

Po przeprowadzeniu szczegółowej analizy uzyskanych wyników badań wstępnych stwierdzono, co następuje:

1. Po wprowadzeniu logarytmicznego sposobu określania progów kwantyzacji istotnie wzrosła sprawność klasyfikatora (wzrost o około 20%), przy czym zaobserwowano

Tabl. D.2. Pełny wykaz nazw cech stosowanych do zapisu bazy przykładów zgromadzonych w eksperymencie numerycznym

Lp.	Oznaczenie	Nazwa cechy
1	RtSp	prędkość obrotowa, [min^{-1}]
2	A04D1	długość większej półosi elipsy drgań bezwzględnych obserwowanych w węźle nr 4
3	A04D2	długość mniejszej półosi elipsy drgań bezwzględnych obserwowanych w węźle nr 4
4	A04TD	kąt pochylenia większej półosi elipsy drgań bezwzględnych obserwowanych w węźle nr 4
5	A04TZ	faza początkowa znacznika elipsy drgań bezwzględnych obserwowanych w węźle nr 4
6	A18D1	długość większej półosi elipsy drgań bezwzględnych obserwowanych w węźle nr 18
7	A18D2	długość mniejszej półosi elipsy drgań bezwzględnych obserwowanych w węźle nr 18
8	AdifTD	różnica pomiędzy kątami pochylenia większej półosi elipsy drgań bezwzględnych w węźle nr 4 i nr 18 (A04TD-A18TD)
9	AdifTZ	różnica pomiędzy fazami początkowymi znacznika elipsy drgań bezwzględnych w węźle nr 4 i nr 18 (A04TZ-A18TZ)
10	R02D1	długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 2
11	R02D2	długość mniejszej półosi elipsy drgań względnych obserwowanych w węźle nr 2
12	R02TD	kąt pochylenia większej półosi elipsy drgań względnych obserwowanych w węźle nr 2
13	R02TZ	faza początkowa znacznika elipsy drgań względnych obserwowanych w węźle nr 2
14	R16D1	długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 16
15	R16D2	długość mniejszej półosi elipsy drgań względnych obserwowanych w węźle nr 16
16	RdifTD	różnica pomiędzy kątami pochylenia większej półosi elipsy drgań względnych w węźle nr 2 i nr 16 (R02TD-R16TD)
17	RdifTZ	różnica pomiędzy fazami początkowymi znacznika elipsy drgań względnych w węźle nr 2 i nr 16 (R02TZ-R16TZ)

Tabl. D.3. Oceny sprawności klasyfikatorów uzyskanych w eksperymencie numerycznym (wstępna wersja bazy *GD*)

Nr klasy	Stan wyrównoważenia	Udział przykładów poprawnie sklasyfikowanych [%] dla wartości progowych określonych w skali	
		liniowej	logarytmicznej
1	zgrubne wyrównoważenie dynamiczne	95.2	99.0
2	niewyrównoważenie statyczne	22.2	22.2
3	niewyrównoważenie quasi-statyczne	64.8	69.4
4	niewyrównoważenie momentowe	96.6	90.4
5	niewyrównoważenie dynamiczne	64.9	95.2
Razem		74.5	94.4

niewielki wzrost błędu w odniesieniu do klasy nr 4 przy jednoczesnym znacznym obniżeniu wartości błędu dla klasy nr 5 (która jest najbardziej liczna);

2. Wskazane jest uwzględnienie różnic kątów pochylenia elips drgań oraz różnic faz początkowych znaczników elipsy drgań (w odniesieniu do elips, zataczanych przez geometryczne środki czopów po przeciwległych końcach wału);

3. Celowe jest wprowadzenie zmian do planu eksperymentu numerycznego, aby można było uzyskać zbliżone liczności zbiorów przykładów dla poszczególnych klas.

D.4 Opracowanie pełnej wersji bazy wiedzy *GD* i jej weryfikacja

Dalszy etap badań obejmuje opracowanie pełnej wersji bazy wiedzy. W tym celu postanowiono przygotować nową bazę przykładów uczących *GD*, która była następnie stosowana do indukcji reguł. Ponadto postanowiono uzupełnić wiedzę "operacyjną" o dodatkowe reguły pozyskane od specjalisty, dotyczące rozpoznawania różnych stanów wyrównoważenia. Oprócz techniki *Random subsampling*, do weryfikacji pozyskanej wiedzy postanowiono zastosować *specjalnie opracowany zbiór przykładów*. Wreszcie postanowiono zweryfikować możliwość zastosowania hipotez wstępnych, doskonalonych następnie za pomocą zbioru przykładów uczących (tzw. *uczenie przyrostowe*).

Pozyskanie wiedzy od specjalistów

W pełnej wersji bazy wiedzy postanowiono zapisać za pomocą reguł wiedzę pozyskaną od specjalistów oraz metodami "uczenia maszynowego". Ponieważ specjaliści posługują się

zwykle niewieloma wartościami cech o charakterze jakościowym, postanowiono zmniejszyć liczbę poziomów kwantyzacji cech ilościowych. Po zmodyfikowaniu opisu dziedziny zapisano reguły reprezentujące ogólną wiedzę dotyczącą identyfikacji rodzaju niewyrównoważenia wirnika 2-podporowego. Uzyskane reguły poddano ocenie przez innych specjalistów, a następnie wyniki oceny wprowadzono do bazy wiedzy.

Określenie liczb poziomów kwantyzacji. W odniesieniu do opisu dziedziny zastosowań autor zaproponował zmianę liczby poziomów kwantyzacji wartości cech. W wyniku optymalizacji przyjęto następujące sposoby kwantyzacji cech:

1. Dla wszystkich długości półosi elips przyjęto *logarytmiczny sposób kwantyzacji* za pomocą przekształcenia:

$$v_\circ = \max\{0, entier(\alpha(n) \lg v + \beta(n))\}, \tag{D.2}$$

gdzie $\alpha(n), \beta(n)$ były stałymi dobranymi tak, aby móc przekształcić zakres $[v_{min}, v_{max}]$ (określony globalnie dla wszystkich długości półosi, niezależnie od węzła, dla którego wyznaczono elipsę drgań) w odcinek liczb $\overline{0..4}$, co daje 5 różnych wartości. Wartości tych stałych są uzależnione od prędkości obrotowej n. Małą liczbę wartości przyjęto ze względu na łatwość przypisania każdej z nich jednej z wartości lingwistycznych, np. *bardzo mała*, *mała*, *średnia*, *duża*, *bardzo duża*.

2. Wartości kątowe *kwantowano w sposób liniowy*, dokonując podziału kąta pełnego na 12 kątów po $\pi/6$ [rad], przy czym dwusieczne kolejnych kątów tworzą z osią Ox kąt $k(\pi/6)$. Przyjęty sposób kwantyzacji umożliwia przyporządkowanie zakresowi kątów $[(k-0.5)\pi/6, (k+0.5)\pi/6)$:

   (a) w wypadku *kątów pochylenia głównej osi elipsy* oraz *faz początkowych znacznika elipsy* — odpowiedniej godziny k (zgodnie z położeniem godzin na tarczy zegara, przy czym godzina 12:00 interpretowana jest jako 00:00), która może być interpretowana jako wartość lingwistyczna *około godziny k-tej*;

   (b) w wypadku *różnic kątów* — opisywanie otrzymanych wartości za pomocą wartości lingwistycznych *około* $30 \cdot k$ *stopni*.

Pozyskanie dodatkowej "porcji" wiedzy od specjalisty. Na podstawie dostępnej literatury (zob. s. 155) opracowano zbiór reguł, umożliwiający identyfikację rozpatrywanych rodzajów niewyrównoważenia wirnika 2-podporowego (dla zakresu użytecznych prędkości obrotowych poniżej 1. prędkości krytycznej układu "wirnik-podpory łożyskowe"). Reguły te zapisano w bazie wiedzy za pomocą formularza "elektronicznego". Fragment raportu z bazy wiedzy zawierającej te reguły przedstawiono na rys. D.4.

Zapisany zbiór reguł poddano ocenie za pomocą formularza "elektronicznego". Osoby oceniające nie były zaangażowane w pozyskanie ocenianych reguł. Oceny w postaci wartości stopnia przekonania o słuszności reguł zostały wprowadzone do bazy wiedzy przez programistę, który dokonał agregacji tych ocen z ocenami uzyskanymi uprzednio.

Przygotowanie bazy przykładów *GD*

Rozpatrując przypadek wirnika sztywnego, autor określił jakościowe opisy analizowanych stanów wyrównoważenia, a następnie, uwzględniając dodatkowo wnioski otrzymane

Reguły - EMPREL2
KPKM Pol. Śl.

Reguła: **1**
Dziedzina:
Klasy obiektów:
Stanowisko modelowe

Jeżeli długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 16 = średnia
i długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 2 = średnia
i różnica faz początkowych znaczników elips drgań względnych między węzłami nr 2 i nr 16 = około 180 stopni,
to stan wyrównoważenia = niewyrównoważenie momentowe.

Reguła: **2**
Dziedzina:
Klasy obiektów:
Stanowisko modelowe

Jeżeli długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 16 = średnia
i długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 2 = średnia
i różnica faz początkowych znaczników elips drgań względnych między węzłami nr 2 i nr 16 = około 0 stopni,
to stan wyrównoważenia = niewyrównoważenie statyczne.

Reguła: **3**
Klasy obiektów:
Stanowisko modelowe

Jeżeli długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 16 = duża
i długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 2 = średnia
i prędkość obrotowa wirnika = podkrytyczna
i różnica pomiędzy kątami pochylenia większej półosi elipsy drgań względnych w węźle nr 2 i nr 16 = około 0 stopni,
to stan wyrównoważenia = niewyrównoważenie quasi-statyczne.

Reguła: **4**
Dziedzina:
Klasy obiektów:
Stanowisko modelowe

Jeżeli długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 16 = mała
i długość większej półosi elipsy drgań względnych obserwowanych w węźle nr 2 = mała
i prędkość obrotowa wirnika = podkrytyczna,
to stan wyrównoważenia = zgrubne wyrównoważenie dynamiczne.

Rys. D.4. Fragment raportu zawierającego reguły wnioskowania o stanie wyrównoważenia
Fig. D.4. A part of a report containing inferring rules on the state of balancing

w trakcie badań wstępnych, zmodyfikował plan eksperymentu numerycznego. W celu uzyskania pełnej wersji bazy wiedzy *GD* opracowano bazę przykładów uczących oraz dodatkowy zbiór przykładów testowych.

Jakościowe opisy stanów wyrównoważenia. *Stan niewyrównoważenia wirnika* jest jednoznacznie określony bądź poprzez wektor główny niewyrównoważenia i moment główny niewyrównoważenia, bądź przez dwa wektory niewyrównoważenia, w ogólnym wypadku różne co do wartości i kierunku, które leżą w dwóch dowolnie przyjętych płaszczyznach prostopadłych do osi wirnika [Łączkowski, 1979].

W opisywanym eksperymencie numerycznym różne stany niewyrównoważenia uzyskiwane były poprzez zadanie odpowiednich rozkładów niewywag. Niewywagi te mogły być zlokalizowane jedynie na płaszczyznach promieniowych przecinających oś wału w wyróżnionych węzłach (rys. D.1). Uwzględniając te ograniczenia oraz rozpatrując wyłącznie przypadek wirnika sztywnego [Łączkowski, 1979], można przyjąć następujące jakościowe charakterystyki rozpatrywanych stanów wyrównoważenia:

1. *Wirnik zgrubnie wyrównoważony dynamicznie*:

   Możliwe są różne rozkłady niewywag, lecz wartości tych niewywag powinny być mniejsze niż dopuszczalne.

2. *Niewyrównoważenie statyczne*:

   Dowolna lokalizacja kątowa niewywagi znajdującej się w płaszczyźnie równo oddalonej od obu podpór łożyskowych (płaszczyzna promieniowa w węźle nr 11).

3. *Niewyrównoważenie quasi-statyczne*:

   Niewywaga położona w jednej płaszczyźnie, która jest zlokalizowana w węźle nie leżącym w połowie odległości między łożyskami (węzeł o numerze innym jak nr 11).

4. *Niewyrównoważenie momentowe*:

   Dwie bardzo zbliżone co do wielkości niewywagi położone na płaszczyznach promieniowych umieszczonych w węzłach równo oddalonych od łożysk, przy czym różnica faz niewywag jest bliska kątowi półpełnemu[1].

5. *Niewyrównoważenie dynamiczne*:

   Dowolny rozkład niewywag, których wartości są większe od dopuszczalnych, nie spełniający kryteriów 2÷4.

Schemat eksperymentu numerycznego dla uzyskania zbioru przykładów uczących. Autor określił plan eksperymentu numerycznego, w którym dokonano obliczeń wybranych parametrów drgań wymuszonych w czterech węzłach (zlokalizowanych jak na rys. D.2):

- *trajektorii ruchu środka czopa wirnika względem panwi* (węzły nr 2 i nr 16),
- *trajektorii ruchu bezwzględnego panwi* (węzły nr 4 i nr 18).

[1]Niewyrównoważenie posiadające takie cechy można nazwać *zbliżonym do momentowego*.

Tabl. D.4. Liczby przykładów uzyskanych dla różnych kombinacji czynników, uwzględnianych w eksperymencie numerycznym (końcowa wersja bazy *GD*)

Nr klasy	Stan wyrównoważenia	Prędk. obr.	Niewywaga				Razem
			bardzo mała	mała	średnia	duża	
1	zgrubne wyrównoważenie dynamiczne	p	360	–	–	–	360
		k	360	–	–	–	360
		n	360	–	–	–	360
2	niewyrównoważenie statyczne	p	–	108	108	108	324
		k	–	108	108	108	324
		n	–	108	108	108	324
3	niewyrównoważenie quasi-statyczne	p	–	108	108	108	324
		k	–	108	108	108	324
		n	–	108	108	108	324
4	niewyrównoważenie momentowe	p	–	108	108	108	324
		k	–	108	108	108	324
		n	–	108	108	108	324
5	niewyrównoważenie dynamiczne	p	–	120	120	120	360
		k	–	120	120	120	360
		n	–	120	120	120	360
Razem			1080	1332	1332	1332	5076

Prędkość obrotowa: p = podkrytyczna k = krytyczna n = nadkrytyczna

Wymuszenie odpowiadało pięciu różnym stanom wyrównoważenia (zob. tabl. D.4), z uwzględnieniem opisów jakościowych rozkładów niewywag (zob. s. 183). Wielkościom niewywag przypisano określenia jakościowe na podstawie [Kiciński *et al.*, 1995] oraz opinii J. Kicińskiego[2]. W eksperymencie numerycznym zastosowano prędkości obrotowe identyczne z opisanymi w [Kiciński *et al.*, 1995].

Stosowany sposób kwantyzacji wartości cech jakościowych. Aby można było zastosować sposób kwantyzacji długości półosi elips, opisany na s. 181 (punkt 1), należało dobrać wartości stałych α, β występujących w (D.2). Ponieważ kodowaniu miały być poddane długości półosi elips [μm], oznaczało to potrzebę określenia wartości progowych odniesionych do *wartości szczytowych przemieszczenia drgań wału*, które (zob. np. [Cempel, 1982], [Łączkowski, 1983]) uzależnione są od prędkości obrotowej wału maszyny.

Przy określaniu wartości progowych przyjęto następujące rozwiązanie:

1. Wartości progowe określane są oddzielnie dla długości półosi elips drgań względnych i drgań bezwzględnych, przy czym są one wspólne dla obu podpór łożyskowych;

2. Wybór wartości progowej dla pierwszego przedziału wartości cechy ilościowej winien umożliwić rozróżnienie (z jak najmniejszą liczbą błędów, co stanowiło kryterium przy optymalnym wyborze wartości progowych) przykładów, odpowiadających stanowi zgrubnego wyrównoważenia dynamicznego, od pozostałych przykładów[3];

3. Pomiędzy dwiema sąsiednimi wartościami progowymi v_m, v_{m+1} winien zachodzić związek $v_{m+1}/v_m = 2.5$ (wg normy ISO 2372 taka różnica wartości skutecznych drgań oznacza zmianę klasy stanu).

[2]Komunikat prywatny.

Tabl. D.5. Wartości progowe stosowane do kwantyzacji długości półosi elips drgań [μm]

Nazwa wartości	Elipsa drgań bezwzględnych Prędkość obr. [min^{-1}]			Elipsa drgań względnych Prędkość obr. [min^{-1}]		
	2050	2972	3509	2050	2972	3509
bardzo mała	0.50	0.34	0.29	1.20	0.83	0.70
mała	1.25	0.86	0.73	3.00	2.07	1.75
średnia	3.12	2.16	1.83	7.50	5.27	4.38
duża	7.81	5.39	4.56	18.75	12.93	10.95

Wartości progowe ustalone w wyniku opisanej powyżej optymalizacji przedstawiono w tabl. D.5. Odpowiednio dobrane stałe α, β w (D.2) umożliwiają przekształcenie zakresu zmienności każdej cechy ilościowej w odcinek liczb $\overline{0..4}$ (co najwyżej 5 różnych wartości).

Uzyskana baza przykładów *GD*. *Baza przykładów GD* jest tablicą decyzyjną (zob. rozdział 3.2.1) o $N = 5076$ wierszach. Liczba cech warunku wynosi $J = 17$. Liczba cech decyzyjnych $L = 1$. W przeciwieństwie do bazy danych *RK*, w omawianej tablicy decyzyjnej nie ma brakujących wartości cech. Cechą charakterystyczną bazy przykładów *GD* jest zrównoważona liczba przykładów dla każdej wyróżnianej klasy stanu (tj. klasy, odpowiadającej *stanowi wyrównoważenia*). Należy jednak podkreślić, że prawdopodobieństwo zdarzenia, że dany przykład $e \in C_k, k = 1, \ldots, 5$ (równe w przybliżeniu 20 %), nie odpowiada prawdopodobieństwu wystąpienia danego stanu wyrównoważenia w praktyce eksploatacyjnej (gdzie np. wystąpienie niewyrównoważenia dynamicznego jest znacznie bardziej prawdopodobne niż wystąpienie niewyrównoważenia momentowego).

Dodatkowy zbiór przykładów testowych. Uwzględniając opisy jakościowe rozkładów niewywag (zob. s. 183) określono dodatkowy zbiór przykładów testowych. Przykłady te generowane były według nieco innego planu niż przykłady uczące. Zbiór ten postanowiono zastosować do weryfikacji pozyskanej "porcji" wiedzy oraz do weryfikacji wyników *uczenia przyrostowego.*

Pozyskanie wiedzy metodami indukcyjnymi

Do pozyskania wiedzy zastosowano wszystkie trzy metody indukcyjne, przy czym podstawowe badania przeprowadzono metodą indukcji reguł z zastosowaniem pokryć.

[3]Do oceny poprawności doboru wartości progowych autor zastosował histogramy wartości cech ilościowych, określane oddzielnie dla każdej z rozpatrywanych klas stanu.

Tabl. D.6. Zestawienie wyników badań — indukcja reguł metodą pokryć (końcowa wersja bazy *GD*)

Symbol badania	Test *Hold-out*			Odrębny zbiór przykładów testowych		
	Wartości ocen błędów względnych [%]			Wartości ocen błędów względnych [%]		
	ϵ_{ov}	ϵ_{om}	ϵ_{cm}	ϵ_{ov}	ϵ_{om}	ϵ_{cm}
BW	10.4	10.1	2.6	20.0	20.0	5.0
B	11.2	11.2	2.8	21.7	21.7	5.4
W	15.8	15.8	4.0	26.7	26.7	6.7

Indukcja reguł metodą pokryć. Indukcję reguł przeprowadzono łącznie z identyfikacją podzbioru cech relewantnych, przy czym porównywane podzbiory cech określono na podstawie dodatkowej wiedzy diagnostycznej dotyczącej sposobów obserwacji maszyn wirnikowych. Wykonano dwie serie badań, różniące się od siebie sposobem określenia ocen błędów klasyfikatorów (zob. rozdział 3.3.2):

1. badanie sprawności klasyfikatora z zastosowaniem do oceny błędów klasyfikatora techniki *Hold-out* (podział zbioru przykładów *GD* na zbiór uczący i zbiór testowy w proporcji 70:30),

2. badanie sprawności klasyfikatora za pomocą *odrębnie przygotowanego zbioru przykładów testowych* (zob. powyżej).

W każdej serii wykonywano badania, w których stosowano różniące się podzbiory cech (zob. tabl. D.2):

Badanie BW: cechy określone dla elips drgań bezwzględnych (obserwacja w węzłach nr 4 i 18) oraz dla drgań względnych (obserwacja w węzłach nr 2 i 16) — obejmowało wszystkie cechy;

Badanie B: cechy określone dla elips drgań bezwzględnych (obserwacja w węzłach nr 4 i 18) — obejmowało cechy o identyfikatorach $1 \div 9$;

Badanie W: cechy określone dla elips drgań względnych (obserwacja w węzłach nr 2 i 16) — obejmowało cechy o identyfikatorach: 1 oraz $10 \div 17$.

Uzyskane wyniki indukcji reguł metodą pokryć zestawiono w tabl. D.6.

Indukcja reguł metodą drzew decyzyjnych. Pozyskiwanie wiedzy reprezentowanej za pomocą drzew decyzyjnych przeprowadzono[4] z zastosowaniem opisanej metody indukcyjnej. Zestawienie ocen błędów względnych przedstawia tabl. D.7.

Uzyskane wyniki oceny sprawności klasyfikatora oraz wyniki klasyfikowania przykładów testowych przedstawiono odpowiednio:

[4]Obliczenia wg koncepcji autora wykonał P. Kostka.

Tabl. D.7. Zestawienie wyników badań — indukcja drzew decyzyjnych (końcowa wersja bazy *GD*)

Symbol badania	Wartości ocen błędów względnych [%]		
	ϵ_{ov}	ϵ_{om}	ϵ_{cm}
BW	5.7	5.6	5.6
B	16.5	16.5	4.1
W	11.9	11.7	3.0

- dla badania BW: w tablicach D.8 i D.9,
- dla badania B: w tablicach D.10 i D.11,
- dla badania W: w tablicach D.12 i D.13.

Tabl. D.8. Zbiorcze zestawienie wyników oceny sprawności klasyfikacji (indukcja drzew decyzyjnych dla końcowej wersji bazy przykładów *GD*; badanie BW)

Nr klasy	Nazwa stanu technicznego	Łączna liczba przykładów				Udział przykładów poprawnie sklas. [%]
		testowych	poprawnie sklas.	pominiętych	niesłusznie zaliczonych	
1	zgrubne wyrównoważenie dynamiczne	972	905	67	7	93.1
2	niewyrównoważenie statyczne	876	866	10	73	98.9
3	niewyrównoważenie quasi-statyczne	876	810	66	71	92.5
4	niewyrównoważenie momentowe	875	868	7	66	99.2
5	niewyrównoważenie dynamiczne	970	858	112	45	88.5
Razem		4569	4307	262	262	94.3

Indukcja reguł metodą zbiorów przybliżonych. Biorąc pod uwagę *pełny zestaw cech* opisujących elipsy drgań względnych i bezwzględnych (jak w badaniu BW, zob. s. 186), dokonano ponadto indukcji reguł przybliżonych[5]. Do oceny sprawności klasyfikatora zastosowano metodę *10-fold cross validation*, w wyniku której 182 przykłady testowe zostały błędnie sklasyfikowane, co daje ocenę błędu względnego $\widehat{\epsilon}_{ov} = 3.59\%$.

[5]Obliczenia wykonał Jerzy W. Grzymała-Busse.

Tabl. D.9. Wyniki klasyfikowania przykładów testowych
(indukcja drzew decyzyjnych dla końcowej wersji bazy przykładów *GD*; badanie BW)

Nr klasy	Nazwa stanu technicznego	Liczba przykładów sklasyfikowanych jako należących do klasy o numerze				
		1	2	3	4	5
1	zgrubne wyrównoważenie dynamiczne	905	–	1	66	–
2	niewyrównoważenie statyczne	–	866	6	–	4
3	niewyrównoważenie quasi-statyczne	–	25	810	–	41
4	niewyrównoważenie momentowe	7	–	–	868	–
5	niewyrównoważenie dynamiczne	–	48	64	–	858

Połączenie wiedzy pozyskanej z różnych źródeł

W ramach prowadzonych badań poddano ocenie możliwość łączenia wiedzy pozyskanej z różnych źródeł. Wpierw pozyskano wiedzę od specjalisty, przy czym została ona zapisana w bazie wiedzy *EMPREL* (zob. s. 180). Wiedza ta była reperezentowna za pomocą reguł. Do połączenia wiedzy pozyskanej od specjalisty oraz wiedzy pozyskanej indukcyjnie zastosowano *uczenie przyrostowe* [Hong, Mozetic, Michalski, 1986]. Reguły pozyskane od specjalisty wprowadzono jako *hipotezy wstępne* do procesu indukcyjnego pozyskiwania wiedzy (z zastosowaniem pokryć), w którym zbiór przykładów uczących stanowiła zawartość bazy *GD*. Zastosowanie tych przykładów umożliwiło specjalizację (uszczegółowienie) reguł wstępnych.

Podobnie jak w wypadku indukcji reguł metodą pokryć dla końcowej wersji bazy *GD* (zob. s. 186), przeprowadzono dwie serie badań:

- techniką *Hold-out* związaną z podziałem zbioru przykładów na zbiór przykładów uczących i zbiór przykładów testowych w stosunku 70:30,
- z zastosowaniem dodatkowego zbioru przykładów testowych (opis na s. 185), przy czym zbiór danych uczących zawierał wszystkie przykłady.

W każdej serii prowadzono badanie BW, B, W. Dla badania BW stosowano trzy różne podzbiory hipotez wstępnych: BW (warunki zarówno dla wartości cech drgań bezwzględnych, jak i drgań względnych), B (warunki dla wartości cech drgań bezwzględnych), W (warunki dla wartości cech drgań względnych). Uzyskane wyniki badania w postaci ocen błędów względnych przedstawiono w tabl. D.14.

Uwagi dotyczące stosowania bazy wiedzy

Baza wiedzy *GD*, uzyskana w sposób opisany powyżej, może być przekazana do próbnego użytkowania. Aby umożliwić stosowanie tej bazy wiedzy dla potrzeb wnioskowania diagnostycznego, wymagane jest zastosowanie odpowiedniego systemu doradczego.

Tabl. D.10. Zbiorcze zestawienie wyników oceny sprawności klasyfikacji
(indukcja drzew decyzyjnych dla końcowej wersji bazy przykładów *GD*; badanie B)

Nr klasy	Nazwa stanu technicznego	Łączna liczba przykładów				Udział przykładów poprawnie sklas. [%]
		testowych	poprawnie sklas.	pominiętych	niesłusznie zaliczonych	
1	zgrubne wyrównoważenie dynamiczne	972	895	77	15	92.1
2	niewyrównoważenie statyczne	876	753	123	285	86.0
3	niewyrównoważenie quasi-statyczne	876	582	293	154	66.5
4	niewyrównoważenie momentowe	875	859	15	77	98.3
5	niewyrównoważenie dynamiczne	972	726	246	223	74.7
Razem		4569	3815	754	754	83.5

Tabl. D.11. Wyniki klasyfikowania przykładów testowych
(indukcja drzew decyzyjnych dla końcowej wersji bazy przykładów *GD*; badanie B)

Nr klasy	Nazwa stanu technicznego	Liczba przykładów sklasyfikowanych jako należących do klasy o numerze				
		1	2	3	4	5
1	zgrubne wyrównoważenie dynamiczne	895	–	–	77	–
2	niewyrównoważenie statyczne	–	753	52	–	71
3	niewyrównoważenie quasi-statyczne	–	141	582	–	152
4	niewyrównoważenie momentowe	15	–	–	859	–
5	niewyrównoważenie dynamiczne	–	144	102	–	726

Tabl. D.12. Zbiorcze zestawienie wyników oceny sprawności klasyfikacji (indukcja drzew decyzyjnych dla końcowej wersji bazy przykładów *GD*; badanie W)

Nr klasy	Nazwa stanu technicznego	Łączna liczba przykładów				Udział przykładów poprawnie sklas. [%]
		testowych	poprawnie sklas.	pominiętych	niesłusznie zaliczonych	
1	zgrubne wyrównoważenie dynamiczne	971	897	74	84	92.4
2	niewyrównoważenie statyczne	876	809	67	163	92.4
3	niewyrównoważenie quasi-statyczne	876	732	144	86	83.6
4	niewyrównoważenie momentowe	874	847	27	109	96.9
5	niewyrównoważenie dynamiczne	972	740	232	102	76.1
Razem		4569	4025	544	544	88.1

Tabl. D.13. Wyniki klasyfikowania przykładów testowych (indukcja drzew decyzyjnych dla końcowej wersji bazy przykładów *GD*; badanie W)

Nr klasy	Nazwa stanu technicznego	Liczba przykładów sklasyfikowanych jako należących do klasy o numerze				
		1	2	3	4	5
1	zgrubne wyrównoważenie dynamiczne	897	–	–	73	1
2	niewyrównoważenie statyczne	34	809	1	–	32
3	niewyrównoważenie quasi-statyczne	–	52	732	23	69
4	niewyrównoważenie momentowe	19	–	8	847	–
5	niewyrównoważenie dynamiczne	31	111	77	13	740

Tabl. D.14. Zestawienie wyników badań — indukcja reguł metodą pokryć z hipotezami wstępnymi (końcowa wersja bazy *GD*)

Zbiór cech	Zbiór reguł	Test *Hold-out*			Odrębny zbiór przykładów testowych		
		Wartości ocen błędów względnych [%]			Wartości ocen błędów względnych [%]		
		ϵ_{ov}	ϵ_{om}	ϵ_{cm}	ϵ_{ov}	ϵ_{om}	ϵ_{cm}
BW	BW	28.3	28.8	7.1	26.7	26.7	6.7
	B	15.9	15.7	4.0	20.0	20.0	5.0
	W	23.7	23.8	5.9	31.7	31.7	7.9
B	B	17.0	17.4	4.3	25.0	25.0	6.3
W	W	28.7	28.6	7.2	36.6	36.6	9.2

Spis literatury

[AAISE, 1996] *Applications of Artificial Intelligence in Structural Engineering*, Proc. of the 3rd Workshop of the European Group for Structural Engineering Applications of Artificial Intelligence (EG-SEA-AI), University of Strathclyde, UK, August 12-13, 1996, Abstracts, Internet `http://www.strath.ac.uk/Departments/Civeng/egseaai`.

[Adamkiewicz *et al.*, 1983] W. Adamkiewicz, L. Hempel, A. Podsiadło, R. Śliwiński, *Badania i ocena niezawodności maszyny w systemie transportowym*, WKiŁ, Warszawa.

[An *et al.*, 1995] A. An, N. Shan, Ch. Chan, N. Cercone, W. Ziarko, *Discovering Rules from Data for Water Demand Prediction*, [W:] B. Julien, S. J. Fenves, T. Arciszewski (Eds.), *Machine Learning in Engineering*, Workshop Notes of the IJCAI-95, Montreal.

[Arciszewski, Dybala, Wnek, 1992] T. Arciszewski, T. Dybala, J. Wnek, *A Method for Evaluation of Learning Systems*, [W:] *HIEURISTICS, The Journal of Knowledge Engineering*, Special Issue on Knowledge Acquisition and Machine Learning, Vol. 5, No. 4, Winter 1992.

[Baim, 1982] P. W. Baim, *The PROMISE Method for Selecting Most Relevant Attributes for Inductive Learning Systems*, Report No. ISG 82-1, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL.

[Banaszak *et al.*, 1997] Z. Banaszak, D. Gattner, K. Mazur-Łukomska, W. Muszyński, *Zarządzanie operacjami*, Wydawnictwo Pol. Zielonogórskiej, Zielona Góra.

[Barschdorff, Moczulski, Pfläging, 1986] D. Barschdorff, W. Moczulski, G. Pfläging, *Motordiagnose unter Anlaufbedingungen mit Hilfe von Mustererkennungsverfahren*, [W:] Mustererkennung 1986, 8. DAGM-Symposium, Informatik-Fachberichte 125, ss. 149-153, Springer-Verlag, Berlin Heidelberg.

[Baseville, Nikiforov, 1993] M. Baseville, I. V. Nikiforov, *Detection of Abrupt Changes — Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ.

[Bently Nevada, 1996] Bently Nevada, *Rotor Kit. Instruction Manual No. TW8029319*, Minden NV.

[Będkowski, 1991] L. Będkowski, *Elementy diagnostyki technicznej*, WAT, Warszawa.

[Birger, 1978] I. A. Birger, *Techničeskaja diagnostika*, Mašinostroenie, Moskva.

[Box, Jenkins, 1983] G. E. P. Box, G. M. Jenkins, *Analiza szeregów czasowych: prognozowanie i sterowanie*, PWN, Warszawa.

[Bratko, 1989] I. Bratko, *Qualitative Modelling and Learning in KARDIO*, 5th Australian Conference on Applications of Expert Systems, Sydney 1989, [W:] B. G. Buchanan, D. C. Wilkins (Eds.), *Readings in Knowledge Acquisition and Learning*, pp. 616-626, Morgan Kaufmann, San Mateo, CA.

[Bratko, Mozetic, Lavrac, 1989] I. Bratko, I. Mozetic, N. Lavrac, *Kardio: a Study in Deep and Qualitative Knowledge for Expert Systems*, MIT Press, Cambridge MA.

[Buchanan *et al.*, 1983] B. G. Buchanan *et al.*, *Constructing an Expert System*, [W:] F. Hayes-Roth, D. A. Waterman (Eds.), *Building Expert Systems*, pp. 127-168, Addison-Wesley, Reading MA.

[Cempel, 1982] Cz. Cempel, *Podstawy wibroakustycznej diagnostyki maszyn*, WNT, Warszawa.

[Cempel, 1989] Cz. Cempel, *Diagnostyka wibroakustyczna maszyn*, PWN, Warszawa.

[Cempel, 1996] Cz. Cempel, *Ewolucyjne modele symptomowe w diagnostyce maszyn*, [W:] Kongres Diagnostyki Technicznej, t. I, ss. 31-38, KPKM Pol. Śląskiej i IMP PAN, Gdańsk.

[Cempel, Tomaszewski, 1992] Cz. Cempel, F. Tomaszewski (Red.), *Diagnostyka maszyn. Zasady ogólne. Przykłady zastosowań*, MCNEMT, Radom.

[Chmielewski, Grzymała-Busse, 1992] M. R. Chmielewski, J. W. Grzymała-Busse, *Global Discretization of Continuous Attributes as Preprocessing for Inductive Learning*, TR-92-7, Department of Computer Science, University of Kansas, Lawrence, KS.

[Chmurawa, Bińkowski, 1980] M. Chmurawa, Wł. Bińkowski, *Podstawy niezawodności i eksploatacji maszyn roboczych*, Skrypt uczeln. Pol. Śląskiej nr 936, Gliwice.

[Cholewa, 1983] W. Cholewa, *Metoda diagnozowania maszyn z zastosowaniem zbiorów rozmytych*, ZN Pol. Śląskiej nr 764, Seria: Mechanika z. 79, Gliwice.

[Cholewa, 1985] W. Cholewa, *Aggregation of Fuzzy Opinions — an Axiomatic Approach*, Fuzzy Sets and Systems, 17 (1985), pp. 249-258.

[Cholewa, 1993] W. Cholewa, *Frames in Diagnostic Reasoning*, Journal of Applied Mathematics and Computer Science, vol. 3 (1993), No. 3, pp. 595-612.

[Cholewa, 1996] W. Cholewa, *Dynamiczne systemy doradcze w diagnostyce technicznej*, [W:] Kongres Diagnostyki Technicznej, t. I, ss. 57-84, KPKM Pol. Śląskiej i IMP PAN, Gdańsk.

[Cholewa, 1996a] W. Cholewa, *Diagnostyczny system doradczy DT3D100. Założenia ogólne*, raport częściowy z realizacji projektu PBZ-038-06, KPKM Pol. Śląskiej, Gliwice.

[Cholewa, 1997] W. Cholewa, *Real-Time Expert Systems for Technical Diagnostics*, [W:] 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, Proceedings Vol.6, pp. 715-20, Berlin, 24-30.08.1997.

[Cholewa, 1997a] W. Cholewa, *Diagnostyczny system doradczy DT3D100. Organizacja procesu wnioskowania*, Raport częściowy nr DT6D131 z realizacji projektu PBZ-038-06, KPKM Pol. Śląskiej, Gliwice.

[Cholewa, Kaźmierczak, 1995] W. Cholewa, J. Kaźmierczak, *Data Processing and Reasoning in Technical Diagnostics*, WNT, Warszawa.

[Cholewa, Kiciński, 1996] W. Cholewa, J. Kiciński, *Metody identyfikacji reguł dla diagnostycznych systemów doradczych*, [W:] XXXV Sympozjon "Modelowanie w Mechanice", ZN Katedry Mechaniki Technicznej, Zeszyt 1, ss. 57-66, Pol. Śląska, Gliwice.

[Cholewa, Kiciński, 1997] W. Cholewa, J. Kiciński (Red.), *Diagnostyka maszyn. Odwrotne modele diagnostyczne*, Monografie — Budowa i Eksploatacja Maszyn, Wyd. Pol. Śląskiej, Gliwice.

[Cholewa, Moczulski, 1990] W. Cholewa, W. Moczulski, *Systemy doradcze w diagnostyce maszyn. Cz. 1: Istota działania, Cz. 2: Zasady konstruowania*, Zagadnienia Eksploatacji Maszyn, Z. 2-3 (1990) ss. 331-342 i Z. 4 (1990) ss. 519-527.

[Cholewa, Moczulski, 1993] W. Cholewa, W. Moczulski, *Diagnostyka techniczna maszyn. Pomiary i analiza sygnałów*, Skrypt uczelniany Pol. Śląskiej nr 1758, Gliwice.

[Cholewa, Moczulski, 1993a] W. Cholewa, W. Moczulski, *Systemy wnioskowania diagnostycznego*, referat zamawiany, Materiały X Szkoły Diagnostyki "Diagnostyka'92", ss. 51-64, Polskie Towarzystwo Diagnostyki Technicznej i Zespół Diagnostyki Sekcji Eksploatacji Maszyn Komitetu Budowy Maszyn PAN, Poznań-Zajączkowo.

[Cholewa, Moczulski, Filipowicz, 1994] W. Cholewa, W. Moczulski, M. Filipowicz, *Knowledge-Based Diagnosing of Technical State of Machinery*, [W:] [Dąbrowski, Michalewicz, Raś, 1994], pp. 148-161.

[Cholewa, Pedrycz, 1987] W. Cholewa, W. Pedrycz, *Systemy doradcze*, Skrypt uczelniany Pol. Śląskiej nr 1447, Gliwice.

[Cholewa, White, 1993] W. Cholewa, M. F. White, *Inverse Modelling in Rotordynamics for Identification of Unbalance Distribution*, Machine Vibration, 2 (1993), pp. 157-167.

[Ciupke, 1997] K. Ciupke, *Przykład zastosowania teorii zbiorów przybliżonych do akwizycji wiedzy diagnostycznej*, [W:] II Krajowa Konferencja Naukowo-Techniczna "Diagnostyka procesów przemysłowych", ss. 99-102, Łagów 08-11.09.1997.

[Czogała, Pedrycz, 1980] E. Czogała, W. Pedrycz, *Elementy i metody teorii zbiorów rozmytych*, Skrypt uczelniany Pol. Śląskiej nr 989, Gliwice.

[Dąbrowski, Michalewicz, Raś, 1993] M. Dąbrowski, M. Michalewicz, Z. Raś (Eds.), *Intelligent Information Systems*, Proceedings of the Conference "Practical Aspects of Artificial Intelligence II", Augustów, Inst. Podstaw Informatyki PAN, Warszawa.

[Dąbrowski, Michalewicz, Raś, 1994] M. Dąbrowski, M. Michalewicz, Z. Raś (Eds.), *Intelligent Information Systems*, Proceedings of the Conference "Practical Aspects of Artificial Intelligence III", Wigry, Inst. Podstaw Informatyki PAN, Warszawa.

[Dietrych, 1985] J. Dietrych, *System i konstrukcja*, WNT, Warszawa.

[Dietterich, 1986] T. G. Dietterich, *Learning at the Knowledge Level*, Machine Learning, 1, pp. 287-316, Kluwer, Boston.

[Dougherty, Kohavi, Sahami, 1995] J. Dougherty, R. Kohavi, M. Sahami, *Supervised and Unsupervised Discretization of Continuous Features*, [W:] A. Prieditis, S. Russell (Eds.), *Machine Learning: Proceedings of the 12th International Conference*, Morgan Kaufman, San Francisco, CA.

[Downham, Woods] E. Downham, R. Woods, *The Rationale of Monitoring Vibration on Rotating Machinery in Continuously Operating Process Plant*, ASME Paper No. 71 – Vibr – 96.

[Drużdżel, 1994] M. J. Drużdżel, *Some Useful Properties of Probabilistic Knowledge Representations From the Point of View of Intelligent Systems*, [W:] [Dąbrowski, Michalewicz, Raś, 1994], pp. 278-292.

[Dubois, Prade, 1996] D. Dubois, H. Prade, *Approximate and Commonsense Reasoning: From Theory to Practice*, [W:] [Raś, Michalewicz, 1996], pp. 19-33.

[Figura, 1996] D. Figura, *Obiektowe bazy danych*, Akademicka Oficyna Wydawnicza PLJ, Warszawa.

[Filbert, 1995] D. Filbert, *A Non-Linear Model Approach to the Diagnosis of Mechanical Faults from the Current of Electric Motors*, [W:] Proc. of 2nd Int. Symposium "Acoustical and Vibratory Surveillance Methods and Diagnostic Techniques", Vol. 1, pp. 33-38, Paris-Senlis.

[Filbert, Dreetz, 1986] D. Filbert, E. Dreetz, *Ein rechnergestütztes Prüfsystem für elektrische Kleinmotoren nach dem Prinzip einer neuartigen, modell-gestützten Messtechnik*, Ber. Inst. Allg. Elektrotechn., TU Berlin.

[Fu, 1982] K. S. Fu, *Syntactic Pattern Recognition and Applications*, Prentice-Hall, Englewood Cliffs, NJ.

[Giordana *et al.*, 1993] A. Giordana, L. Saitta, F. Bergadano, F. Brancadori, D. De Marchi, *ENIGMA: A System that Learns Diagnostic Knowledge*, IEEE Trans. on Knowledge and Data Engineering, Vol. 5, No.1, Feb. 1993, pp. 15-28.

[Gosiewski, Muszyńska, 1992] Z. Gosiewski, A. Muszyńska, *Dynamika maszyn wirnikowych*, Wydawnictwo Uczelniane WSI, Koszalin.

[Grzymała-Busse, 1994] J. W. Grzymała-Busse, *Managing Uncertainty in Machine Learning from Examples*, [W:] [Dąbrowski, Michalewicz, Raś, 1994], pp. 70-84.

[Grzymała-Busse, Wang, 1997] J. W. Grzymala-Busse, A. Wang, *Modified Algorithms LEM1 and LEM2 for Rule Induction from Data with Missing Attribute Values*, [W:] Proc. of the 5th International Workshop on Rough Sets and Soft Computing (RSSC'97) at the 3rd Joint Conference on Information Sciences (JCIS'97), Research Triangle Park, NC.

[Hajnicz, 1996] E. Hajnicz, *Reprezentacja logiczna wiedzy zmieniającej się w czasie*, Akademicka Oficyna Wydawnicza PLJ, Warszawa.

[Harmon, King, 1985] P. Harmon, D. King, *Expert Systems. Artificial Intelligence in Business*, J. Wiley & Sons, New York.

[Hebda, 1990] M. Hebda, *Elementy teorii eksploatacji systemów technicznych*, MCNEMT, Radom.

[Hebda, Wachal, 1980] M. Hebda, A. Wachal, *Trybologia*, WNT, Warszawa.

[Hong, Mozetic, Michalski, 1986] J. Hong, I. Mozetic, R. S. Michalski, *AQ15: Incremental Learning of Attribute-Based Descriptions from Examples. The Method and User's Guide*, ISG Report 86-5, Department of Computer Science, University of Illinois at Urbana-Champaign, IL.

[Huang, Żytkow, 1996] K.-M. Huang, J. M. Żytkow, *Robotic Discovery: the Dilemmas of Empirical Equations*, [W:] Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery, Univ. of Tokyo, Nov. 1996, pp. 217-224.

[ICED, 1997] ICED 97, *Proceedings of the 11th International Conference on Engineering Design*, A. Riitahuhta (Ed.), Tampere University of Technology, Tampere.

[IJCAI-WK, 1995] B. Julien, S. J. Fenves, T. Arciszewski (Eds.), Notes of the Workshop *Machine Learning in Engineering*, International Joint Conference on Artificial Intelligence, Montreal.

[Isermann, 1997] R. Isermann, *Supervision, Fault-Detection and Fault-Diagnosis Methods — An Introduction*, Control Engineering Practice, Vol. 5, No. 5, pp. 639-652.

[Isermann, Ballé, 1997] R. Isermann, P. Ballé, *Trends in the Application of Model-Based Fault-Detection and Diagnosis of Technical Processes*, Control Engineering Practice, Vol. 5, No. 5, pp. 709-719.

[ISO 9000-3:1994] ISO 9000-3:1994: *Quality Management and Quality Assurance Standards — Part 3: Guidelines for the Application of ISO 9001 to the Development, Supply and Maintenance of Software.*

[ISO 9001:1994] ISO 9001:1994: *Quality Systems - Model for Quality Assurance in Design/Development, Production, Installation and Servicing.*

[Kacprzyk, 1986] J. Kacprzyk, *Zbiory rozmyte w analizie systemowej*, PWN, Warszawa.

[Kaufman, 1997] K. A. Kaufman, *INLEN: A Methodology and Integrated System for Knowledge Discovery in Databases*, dissertation, George Mason University, Fairfax, VA.

[Kaźmierczak, 1989] J. Kaźmierczak, *Zastosowanie liniowych modeli procesów losowych do prognozowania w diagnostyce maszyn*, ZN Pol. Śląskiej nr 1037, Seria: Mechanika, z.95, Gliwice.

[Kiciński, 1996] J. Kiciński, *Hydrodynamiczne poprzeczne łożyska ślizgowe*, Wydawnictwo IMP PAN, Gdańsk.

[Kiciński, 1996a] J. Kiciński, *Analiza komputerowa w diagnostyce technicznej maszyn*, Kongres Diagnostyki Technicznej, t. I, ss. 93-122, KPKM Pol. Śląskiej i IMP PAN, Gdańsk.

[Kiciński *et al.*, 1995] J. Kiciński, P. Materny, R. Drozdowski, A. Markiewicz, *Pozyskiwanie relacji diagnostycznych metodą badań symulacyjnych, Część I*, Raport nr 311/95, IMP PAN, Gdańsk.

[Kocańda, 1978] St. Kocańda, *Zmęczeniowe niszczenie metali*, WNT, Warszawa.

[Kocańda, Szala, 1991] St. Kocańda, J. Szala, *Podstawy obliczeń zmęczeniowych*, PWN, Warszawa.

[Korbicz, 1994] J. Korbicz, *Sztuczne sieci neuronowe*, Akademicka Oficyna Wydawnicza, Warszawa.

[Korbicz, 1996] J. Korbicz, *Metody sztucznej inteligencji w diagnostyce technicznej. Przegląd struktur i technik*, [W:] Kongres Diagnostyki Technicznej, Materiały, T. II, ss. 373-380, KPKM Pol. Śląskiej i IMP PAN, Gdańsk.

[Kosmol, 1996] J. Kosmol (Red.), *Monitorowanie ostrza skrawającego: Metody konwencjonalne i sieci neuronowe*, WNT, Warszawa.

[Kostka, 1997] P. Kostka, *System doradczy wspomagający stosowanie wielowymiarowej analizy sygnałów do badania drgań wirników*, praca magisterska (promotor: W. Cholewa), KPKM, Pol. Śląska, Gliwice.

[Kostka, 1997a] P. Kostka, *Model systemu diagnostycznego działającego w środowisku MatLab*, [W:] I Krajowa Konferencja "Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim", materiały ss. 151-158, Akademia Górniczo-Hutnicza, Kraków.

[Kościelny, Szczepaniak, 1997] J. M. Kościelny, P. Szczepaniak, *Terminologia oraz klasyfikacja metod detekcji i diagnostyki procesów przemysłowych*, II Krajowa Konferencja Naukowo-Techniczna "Diagnostyka procesów przemysłowych", ss. 57-68, Łagów 08-11.09.1997.

[Kuipers, 1994] B. Kuipers, *Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge*, MIT Press, Cambridge, MA.

[Kuratowski, 1973] K. Kuratowski, *Wstęp do teorii mnogości i topologii*, wyd. 6., PWN, Warszawa.

[Lawrowski, 1993] Z. Lawrowski, *Tribologia. Tarcie, zużywanie i smarowanie*, PWN, Warszawa.

[Leonhardt, Ayoubi, 1997] S. Leonhardt, M. Ayoubi, *Methods of Fault Diagnosis*, Control Engineering Practice, Vol. 5, No. 5, pp. 683-692.

[Łączkowski, 1979] R. Łączkowski, *Wyważanie elementów wirujących*, WNT, Warszawa.

[Łączkowski, 1983] R. Łączkowski, *Wibroakustyka maszyn i urządzeń*, WNT, Warszawa.

[Maniak, Moczulski, 1997] P. Maniak, W. Moczulski, *Przykład indukcji reguł dotyczących postaci niewyrównoważenia wirnika*, [W:] I Krajowa Konferencja "Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim", materiały s. 159-164, Akademia Górniczo-Hutnicza, Kraków.

[Mańczak, 1976] K. Mańczak, *Technika planowania eksperymentu*, WNT, Warszawa.

[Marciszewski, 1988] W. Marciszewski (Red.), *Mała encyklopedia logiki*, Ossolineum, Wrocław.

[Mesarovic, Takahara, 1975] M. D. Mesarovic, Y. Takahara, *General Systems Theory: Mathematical Foundations*, Academic Press, New York.

[Michalski, 1969] R. S. Michalski, *Recognition of Total or Partial Symmetry in a Completely or Incompletely Specified Switching Function*, [W:] Proceedings of the IV IFAC Congress, Vol. 27, pp. 109-129, Warszawa.

[Michalski, 1983] R. S. Michalski, *A Theory and Methodology of Inductive Learning*, Artificial Intelligence 20 (1983), pp. 111-161.

[Michalski, 1997] R. S. Michalski, *Machine Learning, Data Mining and Knowledge Discovery. Principles and Applications*, Konferencja "Intelligent Information Systems" IIS'97, Zakopane, Inst. Podstaw Informatyki PAN, Warszawa.

[MIMOSA, 1996] *MIMOSA Common Relational Information Schema (CRIS)*, Internet `http://www.hsb.com/mimosa`.

[MIMOSA, 1996a] *MIMOSA for Machine Condition Monitoring Databases. CORE Support Requirements*, Internet `http://www.hsb.com/mimosa`.

[Mitchell, 1981] J. S. Mitchell, *Machinery Analysis and Monitoring*, PennWell, Tulsa, OK.

[Moczulski, 1989] W. Moczulski, *Problemy identyfikacji i gromadzenia danych informacyjnych*, ZN Pol. Śląskiej nr 1027, Seria: Mechanika nr 92, ss. 193-198, Gliwice.

[Moczulski, 1993] W. Moczulski, *Komputerowe wspomaganie sterowania jakością (ISO serii 9000)*, "Maszyny górnicze", 1993 nr 41, wkładka (ss.1-14).

[Moczulski, 1994] W. Moczulski, *Problems of Knowledge Acquisition for Diagnostic Expert Systems*, [W:] Proc. IMEKO XIII World Congress, vol. 2, pp. 1224-27, Torino.

[Moczulski, 1994a] W. Moczulski, *Metody tworzenia systemów objaśniających w procesach komputerowego wspomagania prac inżynierskich*, Raport RMT6215 z realizacji pracy BK-30/RMT-6/94, KPKM, Pol. Śląska, Gliwice.

[Moczulski, 1996] W. Moczulski, *Metody akwizycji wiedzy diagnostycznej*, [W:] XXIII Sympozjum "Diagnostyka Maszyn", Referaty, ss. 44-49, Instytut Transportu, Pol. Śląska, Katowice-Węgierska Górka.

[Moczulski, 1996a] W. Moczulski, *Diagnostic Knowledge Identification by Machine Learning Methods*, [W:] Proc. of Int. Conference on Computer Aided Manufacturing, vol. I, pp. 243-50, Zakopane 1996, Wydział Mech. Technologiczny, Pol. Śląska, Gliwice.

[Moczulski, 1996b] W. Moczulski, *Metody akwizycji wiedzy dla diagnostycznych systemów doradczych*, [W:] Kongres Diagnostyki Technicznej, Materiały, T. III, ss. 79-84, KPKM Pol. Śląskiej i IMP PAN, Gdańsk.

[Moczulski, 1997] W. Moczulski, *Koncepcja empirycznej relacji diagnostycznej*, [W:] XXIV Sympozjum "Diagnostyka Maszyn", Referaty, ss. 157-162, Instytut Transportu, Pol. Śląska, Zeszyt 1/97, Katowice-Węgierska Górka.

[Moczulski, 1997a] W. Moczulski, *Inductive Learning in Design: A Method and Case Study Concerning Design of Antifriction Bearing Systems*, [W:] R. S. Michalski, I. Bratko, M. Kubat (Eds.), *Methods and Applications of Machine Learning, Data Mining and Knowledge Discovery*, John Wiley & Sons.

[Moczulski, 1997b] W. Moczulski, *Diagnostic Knowledge Acquisition Using Empirical Relationships*, [W:] Proceedings of the IMP'97 Conference "Modelling and Design in Fluid-Flow Machinery", pp. 257-262, Wydawnictwo IMP PAN, Gdańsk.

[Moczulski, 1997c] W. Moczulski, *Wspomaganie procesu pozyskiwania wiedzy w systemie DT200 diagnostyki turbozespołów energetycznych*, [W:] III Konferencja "Problemy Badawcze Energetyki Cieplnej", Prace naukowe — Konferencje z. 15, tom II, ss.97-104, Oficyna Wydawnicza Pol. Warszawskiej, Warszawa.

[Moczulski *et al.*, 1995] W. Moczulski, B. Wysogląd, J. Maraszewski, R. Wyczółkowski, *Przykład oceny nadmiernych drgań turbozespołu*, Przegląd Mechaniczny, 1-2'95, ss. 38-41.

[Moczulski *et al.*, 1997] W. Moczulski, K. Ciupke, P. Maniak, M. Wyleżoł, *Pozyskiwanie wiedzy projektowej i konstrukcyjnej metodami indukcyjnymi*, [W:] XI Konferencja "Metody i środki projektowania wspomaganego komputerowo", ss. 255-62, Inst. Podstaw Budowy Maszyn, Pol. Warszawska, Warszawa.

[Moczulski *et al.*, 1997a] W. Moczulski i inni, *Akwizycja wiedzy dla potrzeb diagnostyki technicznej*, Sprawozdanie z realizacji projektu badawczego nr 8 T11F 020 09 (PBU-74/RMT-6/95), KPKM, Pol. Śląska, Gliwice.

[Moczulski, Kostka, 1997] W. Moczulski, P. Kostka, *Pozyskiwanie reguł metodą indukcji drzew decyzyjnych dla przypadku złożonej struktury zbioru stanów*, Raport RMT6374, KPKM, Pol. Śląska, Gliwice.

[Moczulski, Maniak, 1997] W. Moczulski, P. Maniak, *Pozyskiwanie reguł metodą selektywnej indukcji z wyników obliczeń symulacyjnych dotyczących różnych postaci niewyrównoważenia*, Raport RMT6357, KPKM, Pol. Śląska, Gliwice.

[Moczulski, Wyleżoł, 1997] W. Moczulski, M. Wyleżoł, *Pozyskiwanie relacji empirycznych dla potrzeb diagnostyki maszyn*, [W:] III Krajowa Konferencja "Inżynieria wiedzy i systemy ekspertowe", t. II, ss. 160-164, Oficyna Wydawnicza Pol. Wrocławskiej, Wrocław.

[Moczulski, Żytkow, 1997] W. Moczulski, J. M. Żytkow, *Automated Search for Knowledge on Machinery Diagnostics*, [W:] Proceedings of the Workshop "Intelligent Information Systems" IIS'97, pp. 194-203, Inst. Podstaw Informatyki PAN, Warszawa-Zakopane.

[Morel, 1992] J. Morel, *Drgania maszyn i diagnostyka ich stanu technicznego*, PTDT, Warszawa.

[Mrózek, 1993] A. Mrózek, *Some Aspects of Rough Set Applications*, [W:] [Dąbrowski, Michalewicz, Raś, 1993], ss. 47-60.

[Nowicki, Słowiński, Stefanowski, 1992] R. Nowicki, R. Słowiński, J. Stefanowski, *Rough Sets Analysis of Diagnostic Capacity of Vibroacoustic Symptoms*, Computers Math. Applic. 24 (1992), No. 7, pp. 109-123.

[Oprzędkiewicz, 1993] J. Oprzędkiewicz, *Wspomaganie komputerowe w niezawodności maszyn*, WNT, Warszawa.

[Osiński, Wróbel, 1995] Z. Osiński, J. Wróbel, *Teoria konstrukcji*, PWN, Warszawa.

[Pacholczyk, Hunault, Pacholczyk, 1996] D. Pacholczyk, G. Hunault, J.-M. Pacholczyk, *A Symbolic Probability Theory and Its Application to Qualitative Reasoning Under the Uncertain Statements of the Natural Language*, [W:] 9th Int. Symposium ISMIS'96 "Methodologies for Intelligent Systems", Proc. of Poster Session, pp. 55-66, Zakopane.

[Papoulis, 1972] A. Papoulis, *Prawdopodobieństwo, zmienne losowe i procesy stochastyczne*, WNT, Warszawa.

[Patton, 1997] R. J. Patton (Ed.), *IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes: SAFEPROCESS'97*, Aug. 26 - 28, 1997, The Univ. of Hull, Kingston upon Hull.

[Patton, Chen, 1997] R. Patton, J. Chen, *Observer-Based Fault Detection and Isolation: Robustness and Applications*, Control Engineering Practice, Vol. 5, No. 5, pp. 671-682.

[Patton, Frank, Clark, 1989] R. Patton, P. Frank, R. Clark (Eds.), *Fault Diagnosis in Dynamic Systems — Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ.

[Pau, 1981] L.-F. Pau, *Failure Diagnosis and Performance Monitoring*, Marcel Dekker, New York.

[Pawlak, 1982] Z. Pawlak, *Rough Sets*, International Journal of Information and Computer Sciences, 11 (1982), No. 5, pp. 341-356.

[Pawlak, 1991] Z. Pawlak, *Rough Sets. Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers.

[Pearce, 1988] D. Pearce, *The Induction of Fault Diagnosis Systems from Qualitative Models*, [W:] Proc. AAAI-88 Conf., Saint Paul, MN.

[Pieczyński, Korbicz, 1997] A. Pieczyński, J. Korbicz, *Zastosowanie agregacji operacji na rozmytych sygnałach obiektowych w detekcji uszkodzeń*, [W:] II Krajowa Konferencja Naukowo-Techniczna "Diagnostyka procesów przemysłowych", ss. 93-98, Łagów 08-11.09.1997.

[Pokojski, Wróbel, 1995] J. Pokojski, J. Wróbel, *Inteligentne bazy danych dla systemów komputerowo wspomaganego projektowania maszyn*, [W:] XVII Sympozjon Podstaw Konstrukcji Maszyn, Referaty Naukowe cz. II, ss. 731-736, Wydział Mechaniczny, Politechnika Lubelska, Lublin.

[Puzoń, Moczulski, 1996] J. Puzoń, W. Moczulski, *Akwizycja wiedzy diagnostycznej od ekspertów dziedzinowych z zastosowaniem formularzy*, [W:] Kongres Diagnostyki Technicznej, T. III, ss.189-194, KPKM Pol. Śląskiej i IMP PAN, Gdańsk.

[Quinlan, 1986] J. R. Quinlan, *Induction of Decision Trees*, Machine Learning, 1 (1986), pp. 81-106.

[Quinlan, 1993] J. R. Quinlan, *C4.5 Programs for Machine Learning*, Morgan Kaufmann, San Mateo, CA.

[Radkowski, 1995] St. Radkowski, *Low-energy Components of Vibroacoustic Signal as the Basis for Diagnosis of Defect Formation*, Machine Dynamics Problems, Vol. 12, Inst. Podstaw Budowy Maszyn, Pol. Warszawska, Warszawa.

[Raś, Michalewicz, 1996] Z. W. Raś, M. Michalewicz (Eds.), *Foundations of Intelligent Systems*, Proc. of 9th Int. Symposium ISMIS'96, Zakopane 1996. Lecture Notes in AI No. 1079, Springer, Berlin-Heidelberg.

[Reiter, 1987] R. Reiter, *A Theory of Diagnosis from First Principles*, Artificial Intelligence (32) pp. 57-95.

[Russell, Norvig, 1995] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, Englewood Cliffs, NJ.

[Rybczyński, 1993] J. Rybczyński, *Stanowisko badawcze dynamiki wirników i łożysk w Laboratorium Układów Modelowych Maszyn Wirnikowych*, [W:] Konferencja Naukowa "Problemy badawcze energetyki cieplnej", Referaty, Warszawa.

[Schneider-Fresenius, 1985] W. Schneider-Fresenius (Ed.), *Technische Fehlerfrühdiagnose — Einrichtungen*, R. Oldenbourg Verlag, München.

[SigLab, 1996] *SigLab Version 2.13. User's Guide*, DSP Technology, Inc., Fremont, CA.

[Skowron, 1993] A. Skowron, *A Synthesis of Decision Rules: Applications of Discernibility Matrix Properties*, [W:] [Dąbrowski, Michalewicz, Raś, 1993], ss. 30-46.

[Sobczak, Malina, 1985] W. Sobczak, W. Malina, *Metody selekcji i redukcji informacji*, WNT, Warszawa.

[Solartron, 1995] Solartron Instruments, *Common Vibration Symptoms for Turbomachinery in Distress*, materiały firmowe.

[Solartron, 1995a] Solartron Instruments, *Turbomachinery Problem/Symptom Matrix*, materiały firmowe.

[Stroustrup, 1991] B. Stroustrup, *The C++ Programming Language*, 2nd Edition, Addison-Wesley, Reading, MA.

[Subramanian, Mooney, 1996] S. Subramanian, R. J. Mooney, *Qualitative Multiple-Fault Diagnosis of Continuous Dynamic Systems Using Behavioral Models*, [W:] Proc. of 13th National Conference on Artificial Intelligence AAAI-96.

[Tadeusiewicz, 1993] R. Tadeusiewicz, *Sieci neuronowe*, Akademicka Oficyna Wydawnicza, Warszawa.

[Ullman, 1988] J. Ullman, *Systemy baz danych*, WNT, Warszawa.

[Volk, 1965] W. Volk, *Statystyka stosowana dla inżynierów*, WNT, Warszawa.

[Weiss, Kapouleas, 1989] S. M. Weiss, I. Kapouleas, *An Empirical Comparison of Pattern Recognition, Neural Nets, and Machine Learning Classification Methods*, [W:] Proceedings of 11th International Joint Conference on Artificial Intelligence.

[Wnęk, 1993] J. Wnęk, *Hypothesis-Driven Constructive Induction*, dissertation, George Mason University, Fairfax, VA.

[Wnęk *et al.*, 1995] J. Wnęk, K. Kaufman, E. Bloedorn, R. S. Michalski, *Selective Induction Learning System AQ15c: The Method and User's Guide*, Center for Machine Learning and Inference, George Mason University, Fairfax, VA.

[Wong, Li, Ziarko, 1986] S. K. M. Wong, Ye Li, W. Ziarko, *Comparison of Rough Sets and Statistical Methods in Inductive Learning*, Int. Journal of Man-Machine Studies, Vol. 24, pp. 53-72.

[Wyleżoł, 1997] M. Wyleżoł, *Elektroniczny formularz–edytor do pozyskiwania reguł empirycznych dla celów diagnostycznych*, [W:] I Krajowa Konferencja "Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim", materiały ss. 179-184, Akademia Górniczo-Hutnicza, Kraków.

[Wyleżoł, 1997a] M. Wyleżoł, *Formularz - edytor elektroniczny EMPREG 2. Instrukcja użytkownika*, Raport RMT6410 z realizacji projektu PBU-74/RMT-6/95, KPKM, Pol. Śląska, Gliwice.

[Wysogląd, 1996] B. Wysogląd, *Metody reprezentacji drgań wałów maszyn wirnikowych w diagnostycznych bazach danych*, praca doktorska, ZN Pol. Śląskiej nr 1347, Mechanika z.126, KPKM, Gliwice.

[Zembowicz, Żytkow, 1992] R. Zembowicz, J. M. Żytkow, *Discovery of Equations: Experimental Evaluation of Convergence*, [W:] Proc. of the AAAI-92, pp. 70-75, AAAI Press, Menlo Park, CA.

[Zembowicz, Żytkow, 1996] R. Zembowicz, J. M. Żytkow, *From Contingency Tables to Various Forms of Knowledge in Databases*, [W:] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (Eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press.

[Ziarko, Shan, 1996] W. Ziarko, N. Shan, *Database Mining Using Rough Sets*, Computer Science Department, University of Regina, Canada, Saskatchewan.

[Żółtowski, 1996] B. Żółtowski, *Podstawy diagnostyki maszyn*, Wydawnictwo Uczelniane ATR, Bydgoszcz.

[Żytkow, 1993] J. M. Żytkow, *Introduction: Cognitive Autonomy in Machine Discovery*, Machine Learning, 12 (1993), pp. 7-16.

[Żytkow, 1996] J. M. Żytkow, *Automated Discovery of Empirical Laws*, Fundamenta Informaticae, 27, 1996 pp. 299-318.

[Żytkow, Zembowicz, 1993] J. M. Żytkow, R. Zembowicz, *Database Exploration in Search of Regularities*, Journal of Intelligent Information Systems, 2 (1993), pp. 39-81.

[Żytkow, Zembowicz, 1996] J. M. Żytkow, R. Zembowicz, *Mining Patterns at Each Scale in Massive Data*, [W:] [Raś, Michalewicz, 1996], pp. 139-148.

METODY POZYSKIWANIA WIEDZY DLA POTRZEB DIAGNOSTYKI MASZYN

Streszczenie

Praca stanowi próbę syntezy obejmującej całokształt zagadnień związanych z pozyskiwaniem wiedzy dla potrzeb wspomagania diagnozowania wybranych stanów technicznych pewnej wyróżnionej klasy maszyn. Zakłada się, że wiedza ta jest pozyskiwana w celu jej zapisania w bazie wiedzy systemu doradczego. Przykładem zastosowania są systemy doradcze, używane do wspomagania diagnozowania lub nadzoru tzw. maszyn krytycznych.

Omówienie zagadnień, związanych z identyfikacją przedmiotu badań, rozpoczęto od opisu sposobu identyfikacji obiektu diagnozowania. Umożliwiło to sformułowanie zadania diagnostyki maszyn, którym jest ocena stanu technicznego danej maszyny na podstawie zgromadzonych symptomów diagnostycznych oraz wartości cech wejścia maszyny (głównie jej warunków działania). Opisano wybrane środki wspomagania wnioskowania diagnostycznego. Omówiono środki reprezentacji danych i wiedzy diagnostycznej, jak: stwierdzenia, reguły i drzewa decyzyjne. Wreszcie opisano przebieg procesu pozyskiwania wiedzy diagnostycznej, podkreślając znaczenie dwóch podstawowych źródeł tej wiedzy: specjalistów i baz danych diagnostycznych. Szczegółowa identyfikacja przedmiotu badań umożliwiła wybór metod pozyskiwania wiedzy oraz metod oceny zawartości bazy wiedzy.

W pracy opisano metody pozyskiwania wiedzy diagnostycznej z obu jej źródeł. Do pozyskiwania wiedzy od specjalistów wprowadzono empiryczną relację diagnostyczną oraz dwie metody: pozyskiwania wiedzy z udziałem programisty bazy wiedzy i bez tego udziału. Omówiono także sposób przygotowania przykładów uczących i opisano trzy podstawowe metody "uczenia maszynowego": selektywną indukcję reguł poprzez generowanie pokryć, indukcję drzew decyzyjnych i indukcję reguł z zastosowaniem zbiorów przybliżonych. Przedstawiono sposoby oceny wiedzy pozyskanej od specjalistów oraz na drodze indukcji, opisując formalne techniki określania sprawności klasyfikatorów. Ponadto opisano "krzyżowe" stosowanie sposobów oceny pozyskanej wiedzy. Dla uzupełnienia opracowanej metodologii zaproponowano scenariusz procesu pozyskiwania wiedzy diagnostycznej.

Wspomniane metody były podstawą do opracowania środków wspomagających realizację procesu pozyskiwania wiedzy. Zaproponowano schemat logiczny wspólnej bazy danych i wiedzy *EMPREL*. Do pozyskiwania wiedzy od specjalistów zastosowano formularze "papierowe" i "elektroniczne". Do pozyskiwania wiedzy w sposób indukcyjny zastosowano udostępnione pakiety programów "uczenia maszynowego" na podstawie przykładów. Wszystkie te środki zostały połączone we wspólny system pozyskiwania wiedzy *SPWD1*, w którym zasadniczą rolę spełnia baza danych i wiedzy *EMPREL*.

Weryfikacja opracowanych metod i środków wspomagających przebieg procesu pozyskiwania wiedzy nastąpiła w ramach dwóch eksperymentów: czynnego eksperymentu diagnostycznego, prowadzonego na stanowisku modelowym, oraz eksperymentu numerycznego, realizowanego za pomocą oprogramowania. Wyniki weryfikacji w pełni potwierdzają poprawność stosowanych metod i środków pozyskiwania wiedzy diagnostycznej.

Słowa i zwroty kluczowe: diagnostyka maszyn, system doradczy, baza wiedzy, reguły, stwierdzenia, drzewa decyzyjne, pozyskiwanie wiedzy od specjalistów, formularz elektroniczny, uczenie maszynowe, scenariusz procesu pozyskiwania wiedzy

Dziedziny: diagnostyka techniczna maszyn, eksploatacja maszyn

METHODS OF KNOWLEDGE ACQUISITION FOR THE NEEDS OF MACHINERY DIAGNOSTICS

Abstract

This dissertation is an attempt to synthesize the entirety of issues connected with knowledge acquisition for the needs of aiding the diagnosing of selected technical states of a given class of machinery. It is assumed that this knowledge is acquired in order to put it down into a knowledge base of an expert system. An example of an application is an expert system used for aiding of the diagnosing or supervising of critical machines.

The discussion of problems concerning identification of the subject of the research has begun with a description of how to identify a diagnosed object. Then the task of the machinery diagnostics was formulated i.e. estimation of the technical state of the given machine based upon collected diagnostic symptoms and input values (mainly operating conditions of the machine). Selected tools of aiding the diagnostic inferring were described. Then several means of data and diagnostic knowledge representation were discussed, as statements, rules and decision trees. Finally, a course of the process of diagnostic knowledge acquisition was described with the accents placed on two basic sources of knowledge: experts and diagnostic databases. The subject identification enabled the selection of methods of both knowledge acquisition and estimation of correctness of the knowledge base contents.

In the dissertation methods of knowledge acquisition from the both sources are dealt with. To acquire knowledge from experts, an empirical diagnostic relationship was proposed and two methods: acquisition of knowledge with and without participation of a knowledge engineer. The way of preparing training examples was described. Three basic "Machine Learning" methods were presented: selective induction of rules by covers generation, induction of decision trees and induction of rules using rough sets. Ways of estimation of correctness of knowledge acquired from experts and inductively and formal techniques of estimation of performance of the rulesets were presented. Moreover, cross-estimation of classifiers' performance was described. To complete the methodology, a scenario of diagnostic knowledge acquisition process was proposed.

The methods mentioned were the basis for development of tools which aided the accomplishment of the knowledge acquisition process. A logical scheme *EMPREL* of the common data and knowledge base was proposed. The "paper" and "electronic" forms were applied for knowledge acquisition from domain experts. The inductive Machine Learning software packages accessible to the author were applied for acquisition of knowledge from examples. All these means have been combined into a *SPWD1* knowledge acquisition system where the *EMPREL* data and knowledge base played the central role.

The developed methods and tools which aid the course of knowledge acquisition process were verified within two experiments: an active diagnostic experiment on a model kit and a simulations experiment conducted by means of software. The obtained results fully confirmed that the used methods and tools are correct.

Key words and phrases: machinery diagnostics, expert system, knowledge base, rules, statements, decision trees, knowledge acquisition from domain experts, electronic form, machine learning, scenario of knowledge acquisition process

Domains: technical diagnostics of machinery, exploitation of machinery

BIBLIOTEKA GŁÓWNA
Politechniki Śląskiej
P.3361/97/130
Druk: Drukarnia Gliwice, ul. Zwycięstwa 27, tel. 230-49-50