

XIII MIĘDZYNARODOWE KOŁOKWIUM

"MODELE W PROJEKTOWANIU I KONSTRUOWANIU MASZYN"

13th INTERNATIONAL CONFERENCE ON

"MODELS IN DESIGNING AND CONSTRUCTIONS OF MACHINES"

25-28.04. ZAKOPANE

Wojciech MOCZULSKI

Instytut Mechaniki i Podstaw Konstrukcji Maszyn
Politechnika ŚląskaPROBLEMY IDENTYFIKACJI I GROMADZENIA DANYCH INFORMACYJNYCH
DLA POTRZEB DIAGNOSTYCZNEGO SYSTEMU DORADCZEGO

Streszczenie. Omówiono problemy metodologii osiągania i gromadzenia danych informacyjnych dla potrzeb diagnostycznych systemów doradczych. Opisano potencjalne źródła danych diagnostycznych oraz stadia procesu pozyskiwania danych. Przedstawiono również podstawowe modele procesu.

1. Wprowadzenie

W referacie omawiane są problemy metodologii osiągania i gromadzenia danych informacyjnych dla potrzeb diagnostycznych systemów doradczych. System doradczy to pewien specjalny system informacyjny [6], wspomagający użytkownika w rozwiązywaniu złożonych problemów, które "wymagają udziału specjalisty-doradcy wtedy, gdy rozwiązywane są przez człowieka" [8]. Istotną jego cechą jest wykorzystywanie danych informacyjnych, zapisanych w jego bazach danych, w celu osiągania pożądanej informacji diagnostycznej. W celu właściwego działania stosowane są odpowiednie środki techniczne, jak np. mikrokomputery. Można uznać, że w ten sposób zostaje utworzony model sposobu postępowania doświadczonych ekspertów - doradców, których właściwością jest przedstawianie danych informacyjnych w określonej (wąskiej) dziedzinie.

Systemy doradcze określane są w tekstach angielskich terminem "expert systems". Ponieważ w IMiPKM prowadzi się badania nad zastosowaniem systemów doradczych do rozwiązywania problemów diagnostyki technicznej oraz wspomagania prac inżynierskich, postanowiono zastosować nazwę "systemy doradcze" [3] [4].

Osiąganie, identyfikacja i gromadzenie danych informacyjnych (w literaturze określane także wspólnym terminem "pozyskiwanie") - to działania w sferze abstrakcji, mające na celu zapis w bazie danych systemu doradczego: danych informacyjnych oraz sposobów (metod i technik) rozwiązywania problemów należących do pewnej klasy (np. problemów diagnozowania maszyn wirnikowych). Z tymi działaniami łączy się także modyfikacją zawartości bazy ze względu na przyjęte układy kryteriów, (np. ze względu na zmniejszenie niepewności działania użytkownika). Potencjalnymi źródłami danych są eksperci, literatura fachowa, inne bazy danych, wyniki obserwacji

lub pomiarów, dane statystyczne i własne doświadczenie osoby, zapisujące te dane w pamięci komputera.

Wstępna wersja referatu przedstawiona została na Seminarium Nauki Konstrukcji, prowadzonym przez Profesora Janusza Dietrycha.

2. Bazy danych informacyjnych diagnostycznego systemu doradczego

W stosowanych obecnie diagnostycznych systemach doradczych można zidentyfikować następujące kategorie danych informacyjnych:

- Dane o obiekcie diagnozowania i np. o innych obiektach tej samej konstrukcji (struktura, stan, objawy niesprawności, ważniejsze zdarzenia zasze podczas eksploatacji, jego niezawodność, ...). Dane te są zapisywane w bazie danych stałych.
- Opisy sposobów (metod i technik) postępowania diagnostycznego oraz relacji diagnostycznych. Są to zapisane dane informacyjne, których pierwotnym źródłem są wiedza i umiejętności ekspertów-specjalistów w przedmiotowej dziedzinie. Są one zapisywane w tzw. "bazie wiedzy" (ang. Knowledge Base).

Wymienione bazy w przypadku diagnostycznych systemów doradczych nie muszą być rozłączne. Możliwość ich identyfikacji uzależniona jest od stosowanego sposobu zapisu danych informacyjnych obu kategorii. Dane zgromadzone w bazach są wynikiem identyfikacji wiedzy i umiejętności ekspertów, a więc wynikiem działalności myślowej człowieka lub wynikiem procesu poznania. Dlatego mogą być niepewne, ważne, niedokładne lub rozmyte. Metody, którymi posługuje się ekspert, mogą być metodami heurystycznymi [6]. Także relacje diagnostyczne mogą być słuszne jedynie częściowo. Nie da się w praktyce utworzyć odpowiednich modeli, które opisywałyby działanie badanych obiektów i określały relacje między stanem technicznym i cechami obserwowanych sygnałów. Z tych względów bazuje się na opiniach specjalistów, często niejednoznacznych i (z natury) subiektywnych. Istotną właściwością danych zapisywanych w bazach jest więc możliwość wystąpienia ich "niepewności" lub "niedokładności".

3. Ogólne metody pozyskiwania danych informacyjnych

Pozyskiwanie danych informacyjnych jest jednym z najtrudniejszych stadiów tworzenia systemu doradczego. Proces ten obejmuje:

- sformułowanie problemu, do którego rozwiązania stosowany będzie dany system doradczy,
- reprezentację zgromadzonych faktów, relacji i procedur,
- weryfikację i doskonalenie baz danych.

Opisywane procesy mają charakter sekwencyjno-iteracyjny. Wyróżnia się w nich stadia [9]: identyfikacji, koncyptowania, formalizacji, implementacji oraz weryfikacji.

Najczęściej (np. [9]) wyróżnia się następujące metody pozyskiwania danych informacyjnych dla potrzeb systemów doradczych:

- 1) pozyskiwanie danych informacyjnych od eksperta/ekspertów przy pomocy specjalisty-metodologa z zakresu "inżynierii wiedzy",
- 2) pozyskiwanie danych informacyjnych od eksperta przy pomocy odpowiednio opracowanego oprogramowania,
- 3) pozyskiwanie danych informacyjnych w drodze tzw. "uczenia maszynowego",
- 4) pozyskiwanie danych informacyjnych w wyniku automatycznej analizy tekstów.

Największe znaczenie ma pozyskiwanie danych informacyjnych od ekspertów. Wiedza fachowa i umiejętności eksperta obejmują definicje pojęć, fakty (stwierdzenia), algorytmy, strategie i heurysty-

ki. Oprócz stwierdzeń-faktów (których prawdziwość jest stwierdzona lub akceptowana) ważną rolę spełniają także stwierdzenia-przekonania (wyobrażenia/hipotezy dotyczące pewnych nieobserwowalnych sytuacji) oraz reguły heurystyczne (będące proceduralnymi wskazówkami lub niepełnymi sposobami rozwiązania zadania). Postępowanie eksperta rzadko odpowiada pewnemu rygorystycznemu, ściśle algorytmicznemu procesowi. Należy podkreślić, że oprócz wiedzy ekspert posiada także umiejętności. Między wiedzą a umiejętnością (ang. skill) zachodzi istotna różnica, bowiem "umiejętność to posiadanie odpowiedniej wiedzy i jej efektywne stosowanie" [9].

3.1. MODEL 1 : Ekspert-->Specjalista-metodolog-->System doradczy

Model ten należy obecnie do najbardziej popularnych. Występuje w nim dwie strony: ekspert, tj. specjalista posiadający wiedzę i doświadczenie w pewnej dziedzinie, oraz specjalista-metodolog z zakresu osiągania i gromadzenia danych informacyjnych dla potrzeb systemów doradczych (w literaturze angielskojęzycznej określane jako Knowledge Engineer - "inżynier wiedzy"), który spełnia rolę pośrednika między ekspertem a systemem doradczym.

Zadaniem specjalisty-metodologa jest pomoc ekspertowi w identyfikacji posiadanej przez niego wiedzy i umiejętności, ich strukturalizacji i formalizacji. Pomoc ta może polegać np. na [1]:

- interpretacji i uogólnianiu odpowiedzi eksperta na stawiane mu pytania,
- szkicowaniu analogii, które pomogą ekspertowi w systematyzowaniu dziedziny wiedzy lub uzmysłowią mu ważne fakty,
- szukaniu kontrprzykładów i nowych problemów do rozwiązania.

Specjalista-metodolog musi uważnie przysłuchiwać się wypowiedziom eksperta, starać się jak najlepiej pojąć jego sposób rozumowania, zidentyfikować używane przez niego algorytmy. Po dokonaniu (samodzielnie) formalizacji i strukturalizacji zidentyfikowanych danych informacyjnych, przedstawia zapis wyników swej pracy ekspertowi. Według [10], aby tworzony system doradczy mógł skutecznie działać, wiedza eksperta musi zostać usystematyzowana i odpowiednio zinterpretowana. Ekspertowi należy także ogólnie przedstawić istotę działania opracowywanego systemu doradczego.

Najpopularniejszymi metodami współdziałania eksperta ze specjalistą-metodologiem (mającymi wiele cech wspólnych) są wywiady i analiza protokołów. Wywiad prowadzi specjalista-metodolog, stawiając ekspertowi pytania i problemy do rozwiązania. Wypowiedzi eksperta mają umożliwić identyfikację faktów, reguł i procedur heurystycznych. Analiza protokołów z kolei ułatwia identyfikację postępowania eksperta podczas rozwiązywania problemów z jego specjalności. W tym celu prosi się go o "głośne myślenie" i werbalne przedstawianie swoich procesów myślowych. Protokół, będący zapisem wypowiedzi eksperta oraz zawierający spostrzeżenia specjalisty-metodologa, jest później szczegółowo analizowany.

Metody te mają wiele wad [5]. Nie pozwalają one na skuteczne ujawnienie "niemej" (niewyartykułowanej) wiedzy eksperta, co może spowodować luki w bazach danych systemu doradczego. Wybrany sposób reprezentacji danych w pewnym stopniu steruje przebiegiem procesu pozyskiwania danych informacyjnych, dając w wyniku bazę, której zawartość bardziej jest uzależniona od tego sposobu, niż od wiedzy, która ma być w tej bazie zapisana. Ponadto ekspert w czasie wywiadu może używać innej terminologii, niż podczas rozmowy z fachowcem ze swej specjalności (gdyż specjalista-metodolog jest często laikiem lub nowicjuszem w dziedzinie eksperta). Inne wady to:

- wywoływanie introspekcji i powodowanie werbalnego artykułowa-

- nia wiedzy przez eksperta,
- subiektywna interpretacja przez specjalistę-metodologa przekazywanych przez eksperta informacji,
- wzajemne psychologiczne oddziaływanie na siebie obu uczestników procesu pozyskiwania danych informacyjnych.
- trudności połączone z opisaniem przez eksperta procesów myślowych, powiązanych z aktualnie rozwiązywanym problemem (szczególną trudność sprawia opisanie heurystyk).

3.2. MODEL 2 : Ekspert-->"Inteligentny" edytor-->Baza wiedzy

Opisane uprzednio metody pozyskiwania wiedzy z udziałem specjalisty-metodologa powodują, że sam proces staje się czasochłonny i nużący. Stoi to w jawnej sprzeczności z powszechnym wyobrażeniem o ekspercie jako o bardzo zajęтым człowieku. Stąd dążenie (będące zarazem jednym z najważniejszych celów badawczych) do tego, by z procesu pozyskiwania wiedzy wyeliminować pośrednika, jakim jest właśnie specjalista-metodolog. Rozwiązaniem byłoby wyposażenie eksperta w odpowiednie środki techniczne (sprzęt elektroniczny oraz oprogramowanie), aby był w stanie samodzielnie zapisać oraz wprowadzić do bazy wiedzy systemu doradczego dane informacyjne, odpowiadające posiadanej przez niego wiedzy i doświadczeniu. Formułowane są przy tym następujące istotne kryteria:

- wymagane od eksperta przygotowanie informatyczne winno być jak najmniejsze (najlepiej, gdy nie jest wymagane jakiegokolwiek),
- na zapis danych informacyjnych należy poświęcić tak mało czasu, jak to tylko jest możliwe,
- dostarczone ekspertowi środki techniczne powinny umożliwić utworzenie systemu doradczego, który będzie jak najwierniej naśladował postępowanie eksperta podczas rozwiązywania problemów z danej dziedziny wiedzy.

W tym modelu można wyodrębnić [9] trzy główne zadania: wprowadzanie danych informacyjnych do pamięci komputera, identyfikacja (i ew. poprawienie) błędów oraz doskonalenie i uzupełnianie bazy wiedzy w celu uzyskania wymaganej skuteczności działania systemu doradczego. Realizację każdego z tych zadań wspierają dostępne programy, jak np. edytory bazy wiedzy.

Zgodny z omawianym modelem proces pozyskiwania wiedzy musi być więc poprzedzony wyborem wspomagających go środków. Największe znaczenie ma wybór właściwego edytora bazy wiedzy, który jest równoznaczny z wyborem sposobu reprezentacji danych informacyjnych i należy do eksperta (któremu może pomóc specjalista-metodolog).

Istotą najbardziej zaawansowanych metod jest pozyskiwanie danych od eksperta w sposób automatyczny. Może być w tym celu zastosowany taki układ o "inteligentnych" cechach, który "prowadzi" wywiad z ekspertem, na bieżąco aktualizując bazę wiedzy. Rozważane są różne sposoby współdziałania eksperta z takim układem [7]:

- ekspert wprowadza opisy przypadków, a "prowadzący" z nim wywiad program rejestruje te przypadki w swoich zasobach pamięciowych i poprzez analogie identyfikuje relacje między nimi,
- program "dostarcza" ekspertom przypadki do oceny, zaś oni komentują relacje między przypadkami, umożliwiając programowi na drodze analogii identyfikację stosowanych przez siebie reguł,
- eksperci i program wzajemnie na siebie oddziałują, formułując reguły oraz identyfikując i gromadząc wiedzę.

Przewiduje się, że gdy zostaną skutecznie rozwiązane problemy odpowiednio niezawodnej komunikacji z człowiekiem z komputerem (rozumienie mowy, przetwarzanie języka naturalnego), pozyskiwanie wiedzy będzie polegało na instruowaniu komputera przez eksperta oraz

konwersacji w języku naturalnym, wspartej przedstawianymi przez eksperta danymi w postaci graficznej (forma graficzna ma ogromne znaczenie w procesie przekazywania danych).

3.3. MODEL 3 : Dane-->System doradczy

Model ten opisuje formę tzw. uczenia maszynowego (ang. Machine Learning), połączonego z przyswojeniem sobie przez uczący się podmiot sposobów zastosowania nabytej wiedzy. Szersze omówienie metod uczenia maszynowego jako formy pozyskiwania danych informacyjnych zawiera praca [4], w której wyróżniono następujące metody ogólne (dwie ostatnie mają charakter indukcyjny):

- uczenie się poprzez przekazanie informacji,
- uczenie się poprzez wykorzystanie analogii (np. [7]),
- uczenie się na podstawie przykładów (i ew. kontrprzykładów),
- uczenie się na podstawie samodzielnych obserwacji (tj. bez udziału "nauczyciela" - ang. Unsupervised Learning).

W [4] są także opisane pewne szczegółowe metody pozyskiwania wiedzy dla baz, zapisanych za pomocą reguł. Jedną z nich jest generowanie reguł za pomocą drzew decyzyjnych - metoda łatwa do realizacji i o dużej efektywności, stosowana wtedy, gdy reguły wykorzystywane są do klasyfikacji obiektów, identyfikowanych za pomocą par <Atrybut; Wartość>. Zalecane jest stosowanie tej metody do modyfikacji struktury reguł sformułowanych przez eksperta.

Inną metodą jest metoda generacji pokryć [4], wymagająca dostępności pewnego początkowego zbioru reguł oraz zbioru przykładów. Generowane są reguły o przesłankach zapisanych z wykorzystaniem logiki o zmiennej liczbie wartości logicznych. Postępowanie ma charakter indukcyjny, a istotą jest tworzenie "pokrycia" zbioru przykładów, eliminującego jednocześnie kontrprzykłady. Pokrycie takie generowane jest dla każdego nowo dostarczonego przykładu.

Jedną z metod indukcyjnych jest pozyskiwanie wiedzy poprzez grupowanie zdarzeń [2]. Dokonywane są systematyczne obserwacje, przy czym wykrywane są różne współzależności między wynikami pomiarów. Są one istotne: w diagnostyce niesprawności maszyny na podstawie obserwacji symptomów tej niesprawności w jednej maszynie nie można w pełni zidentyfikować znaczenia pewnych cech diagnostycznych, jeśli nie jest dostępna wiedza o relacjach między symptomami a daną niesprawnością. Takie współzależności między zjawiskami mogą być uzyskane poprzez indukcję na podstawie wielokrotnych obserwacji. W tym celu wykorzystywane są algorytmy grupowania.

4. Podsumowanie

W referacie omówiono problemy osiągnięcia i gromadzenia danych informacyjnych dla potrzeb diagnostycznego systemu doradczego. Stwierdzono, że jest to obecnie jedno z najtrudniejszych zadań, dla którego nie jest jeszcze znane ogólne rozwiązanie. Zadanie to ma jednak rozwiązania szczególne, czego dowodzi skuteczne działanie wielu systemów doradczych. Problem ten wymaga dalszych badań oraz rozwoju programów wspomagających proces pozyskiwania wiedzy.

LITERATURA

- [1] BUCHANAN B.G.: New Research on Expert Systems. [W:] HAYES J.E., MICHIE D., PAO Y-H (Eds.), Machine Intelligence 10, s.269-299. Ellis Horwood, Chichester 1984.
- [2] CHIU D.K.Y., WONG A.K.C.: Synthesizing Knowledge: A Cluster Analysis Approach Using Event Covering. IEEE Trans., Vol. SMC-16, No.2, March/Apr. 1986, s.251-9.
- [3] CHOLEWA W.: Automacyjne układy i metaukłady doradcze. Materiały I Krajowej Konf.Naukowo-Techn. "Diagnostyka Techniczna Urządzeń i Systemów", Kielce-Sielc 1986, t.2, ss.47-56.
- [4] CHOLEWA W., PEDRYCZ W.: Systemy doradcze. Skrypt Pol. Sl. Nr 1447, Gliwice 1987.
- [5] COOKE N.M., McDONALD J.E.: A Formal Methodology for Acquiring and Representing Expert Knowledge. Proc. IEEE, 74, No.10, Oct. 1986, s.1422-30.
- [6] DIETRYCH J.: System i konstrukcja. WNT, Warszawa 1985.
- [7] ELIOT L.B.: Analogical Problem-Solving and Expert Systems. IEEE Expert, 1(1986), No.2, s.17-28.
- [8] FEIGENBAUM E.A.: Expert Systems in the 1980. [W:] "State of the Art Report on Machine Intelligence", BOND A. (Ed.). Pergamon-Infotech.
- [9] HAYES-ROTH F., WATERMAN D.A., LENAT D.B.: Building Expert Systems. Addison-Wesley, Reading, Mass. 1983.
- [10] NOELKE U.: Das Wesen des Knowledge Engineering. [W:] SAVORY S.H. (Ed.), Künstliche Intelligenz und Expertensysteme, s.109-123. R. Oldenbourg Verlag, München 1985.

**ПРОБЛЕМЫ ИДЕНТИФИКАЦИИ И СОБИРАНИЯ ИНФОРМАЦИОННЫХ ДАННЫХ
ДЛЯ ПОТРЕБНОСТЕЙ ДИАГНОСТИЧЕСКОЙ ЭКСПЕРТНОЙ СИСТЕМЫ**

Р е з ю м е

Обсуждено проблемы методологии достижения и собирания информационных данных для потребностей диагностических экспертных систем. Описано потенциальные источники диагностических данных, а также этапы процесса приобретения данных. Представлено также основные модели этого процесса.

PROBLEMS OF KNOWLEDGE ACQUISITION FOR DIAGNOSTIC EXPERT SYSTEM

S u m m a r y

The paper deals with some problems of the methodology of knowledge acquisition for the needs of diagnostic expert systems. Potential sources of diagnostic knowledge are described and stages of the knowledge acquisition process are discussed. The most important models of the process are presented as well.

Recenzent: doc. dt hab. A. Lipowczan

Wpłynęło do Redakcji 5.I.1989 r.