

REPREZENTACJE WIEDZY I WNIOSKOWANIE W SYSTEMACH EKSPERCKICH DRUGIEJ GENERACJI

Wiesław Traczyk

Instytut Automatyki i Informatyki Stosowanej
Politechnika Warszawska

1 Wstęp

W niezbyt długiej historii systemów eksperckich (SE) wyróżnia się dwa etapy rozwoju – *generacje*. Pierwsza generacja to dosyć proste systemy, zazwyczaj wykorzystujące *reguły produkcji* oraz wnioskowanie w przód i wstecz. Cechy charakterystyczne drugiej generacji to:

- mieszane reprezentacje wiedzy, łączące wiedzę ilościową z jakościową, reguły z ramami (lub podobnymi opisami sieciowymi) i modelami, a niekiedy także z sieciami neuronowymi,
- nowe metody wnioskowania – na podstawie modelu, na podstawie znanych przypadków, na zasadzie podobieństwa,
- tendencje do automatycznego pozyskiwania wiedzy, wobec trudności z jej uzyskiwaniem od ludzi-ekspertów.

Możliwość użycia tych rozwiązań sprawiła, że liczba wariantów tworzenia SE dla różnych zastosowań zwiększyła się znacznie i wybór odpowiedniego SE przestał być prosty. Kłopoty z syntezą takich systemów powiększa dodatkowo brak jednolitych i pełnych podstaw teoretycznych. Wprawdzie niektóre dziedziny związane z projektowaniem SE (np zbiory rozmyte i przybliżone, wnioskowanie przybliżone, modelowanie jakościowe, uczenie maszyn) mają obszerną i bardzo zaawansowaną literaturę, nie rozwiązuje ona jednak na wiele problemów praktycznych i teoretycznych powstających w czasie tworzenia systemu dla konkretnych potrzeb.

Prezentowany projekt ¹ dotyczył kilku zagadnień typowych dla drugiej generacji SE i mających istotne znaczenie praktyczne. Trzy główne tematy projektu to:

modele i wnioskowanie jakościowe – charakterystyczna reprezentacja wiedzy i związany z nią sposób wnioskowania, z wykorzystaniem transformacji opisów jakościowych w ilościowe,

kryteria wyboru wariantów w syntezie systemów eksperckich, sformułowane na podstawie utworzonego opisu formalnego zasad i strategii wnioskowania,

generowanie reguł z przykładów, wykorzystujące rachunek podziałów dla tworzenia reguł o zadanych cechach.

Metody i wyniki tych prac są – w zarysie – przedstawione niżej.

2 Modele i wnioskowanie jakościowe

Ważnymi cechami wielu SE jest uwzględnianie niedoskonałości wiedzy (niedokładności, niepewności itp), oraz tendencja do redukcji wiedzy do rozmiarów rzeczywiście potrzebnych w rozważanej dziedzinie. Bardzo pomocne mogą w tym być *metody jakościowe*, silnie rozwijane w ostatnich latach na użytek różnych dziedzin. Główne zalety metod jakościowych to:

¹grant KBN nr 3 0606 91 01

- maskowanie szczegółów, gdy np zamiast wartości temperatury 36,5 lub 36,7 wystarczy jakościowe określenie NORMALNA,
- ukrywanie niedokładności, gdy np zamiast ryzykownej prognozy, że akcje Wedla podrożeją o 0,5% można przewidywać, że podrożeją NIEZNACZNIE,
- czytelność wartości i mniejsza ich liczba, co ułatwia posługiwanie się danymi i ich przetwarzanie.

Oczywista wada metod jakościowych – brak precyzji – w wielu przypadkach nie ma dużego znaczenia, zwłaszcza gdy istotna jest ocena relacji (WIĘKSZY - MNIEJSZY), tendencji (ROŚNIE - MALEJE) itp.

Pierwotne wartości zmiennych występujących w SE mają niekiedy postać numeryczną, która dla celów analizy jakościowej jest przetwarzana na postać lingwistyczną (często z zastosowaniem teorii zbiorów rozmytych), istotne są więc związki między zbiorami:

A – zmienne: x, y

U – wartości numeryczne: 1995, 3,14

V – wartości lingwistyczne: BARDZO-MAŁE, PODWYŻSZONA

czyli

$$\text{War: } A \rightarrow U$$

$$\text{War}^*: A \rightarrow V$$

$$\text{Ozn: } U \rightarrow V$$

Przyjęcie określonych oznaczeń (Ozn) umożliwia zastąpienie pierwotnego wartościowania War przez jakościowe wartościowanie wtórne War^* , ale nie rozwiązuje problemu operacji na jakościowych wartościach lingwistycznych. Proste i dobrze znane działania w zbiorze U nie mogą być łatwo przeniesione na zbiór V, i fakt ten stanowi główną przeszkodę w rozwoju metod jakościowych. Każda ze znanych metod definiowania operacji (algebry znaków, przedziałów, rzędu wielkości itd) ma wady, ograniczające zakres zastosowań do pewnych przypadków szczególnych, poszukuje się więc nowych metod, rozszerzających obszar zastosowań.

Podejście proponowane w pracach objętych projektem ma swe źródło w obserwacji, że w wielu przypadkach wartości lingwistyczne (*termy*) konkretnej zmiennej tworzą zbiór uporządkowany; dla atrybutu (zmiennej) a będzie więc:

$$V(a) = \langle v_{-m}, \dots, v_{-1}, v_0, v_1, \dots, v_n \rangle$$

np

$$V(a) = \langle \text{MAŁE}, \text{PRZECIĘTNE}, \text{DUŻE} \rangle.$$

Numery porządkowe (indeksy) tych wartości, odpowiednio dobrane w zależności od treści przenoszonych wartościami, mogą być zastępczymi argumentami działań na zmiennych lingwistycznych, likwidując niektóre problemy związane z tymi działaniami.

Indeksacja

$$\text{Ind} : V \rightarrow N \quad i(v_i) = i$$

i poziom zmiennej $\psi(a)$:

$$(\psi(a) = i) \Leftrightarrow (a = v_i)$$

sprowadzają operacje na zmiennych lingwistycznych do działań na zbiorze liczb całkowitych N (w innej wersji – liczb rzeczywistych).

Użyteczność takiego podejścia będzie zademonstrowana na przykładzie reguł produkcji i równań modelu.

2.1 Reguły

Klasyczna wersja zbioru reguł produkcji z lingwistycznymi wartościami atrybutów, opisujących np. tablicę decyzyjną, może być przedstawiona w postaci;

$$(a' = v'_1) \wedge (a'' = v''_1) \wedge \dots \Rightarrow (a = v_1)$$

$$(a' = v'_2) \wedge (a'' = v''_2) \wedge \dots \Rightarrow (a = v_2)$$

...

$$(a' = v'_z) \wedge (a'' = v''_z) \wedge \dots \Rightarrow (a = v_z)$$

Gdyby udało się wartość atrybutu decyzyjnego a wyrazić w postaci funkcji wartości atrybutów warunkowych $a', a'', \dots$, cały zbiór reguł można byłoby zastąpić jedną regułą o postaci:

$$(a' = x') \wedge (a'' = x'') \wedge \dots \Rightarrow (a = \eta(x', x'', \dots))$$

z zastosowaniem funkcji jakościowej:

$$[\eta(x', x'', \dots)]_a = v_j \quad j = \eta^*(\iota(x'), \iota(x''), \dots)$$

Okazuje się, że funkcję tę można wyrazić za pomocą następujących operacji: podobieństwo

$$[v'_i]_a = v_i$$

przeciwieństwo

$$[\ominus v'_i]_a = \begin{cases} v_{-i} & \text{jeśli } i \neq 0 \\ v_0 & \text{jeśli } i = 0 \end{cases}$$

suma

$$[v'_i \oplus v'_j]_a = v_{i+j} \quad [K \oplus v_i]_a = v_{K+i}$$

różnica

$$[v'_i \ominus v'_j]_a = v_{i-j} \quad [K \ominus v_i]_a = v_{K-i}$$

iloczyn

$$[K \otimes v_i]_a = v_{i \times K}$$

Działania te mogą wyprowadzać poza dopuszczalny zakres wyniku i dlatego wprowadza się dodatkowe ograniczenia, np.

nasycenie

$$[v_i]_a^S = \begin{cases} v_i & \text{jeśli } -m \leq i \leq m \\ v_{-n} & i < -m \\ v_n & i > m \end{cases}$$

ekstrema

$$[v_i]_a^k = \begin{cases} v_i & \text{jeśli } -k \leq i \leq k \\ v_{-2k-i} & i < -k \\ v_{2k-i} & i > k \end{cases}$$

Wprowadzone operacje umożliwiają formułowanie funkcji $\eta(x', x'', \dots)$, a przez to – zapisywanie zbioru konwencjonalnych reguł produkcji w postaci jednej reguły funkcyjnej. Na przykład, nawet przy licznych zbiorach wartości atrybutów, przydatność pracownika może być określona jedną regułą:

JEŚLI PRACOWITOŚĆ JEST x I ZDOLNOŚCI SĄ y
TO PRZYDATNOŚĆ JEST $[x \oplus y]^S$

Założono tu, że na przydatność pracownika jednakowo wpływają jego pracowitość i zdolności; wagi tych cech można łatwo zmienić za pomocą stałego mnożnika K .

2.2 Równania

Przy opisywaniu formalnych modeli zjawisk i obiektów powszechnie wykorzystuje się równania i nierówności algebraiczne. Jeśli zależności te mają dotyczyć parametrów jakościowych, pojawiają się kłopoty z interpretacją relacji między zmiennymi lingwistycznymi, dlatego zamiast konwencjonalnej równości ilościowej $a = f(a', a'', \dots)$ proponuje się *równość jakościową*, dotyczącą poziomów zmiennych:

$$a \stackrel{Q}{=} F(a', a'', \dots) \Leftrightarrow \psi(a) = \psi(F(a', a'', \dots))$$

gdzie

$$(\psi(a) = i) \Leftrightarrow (a = v_i).$$

Operacje na symbolach zmiennych mają podobną postać do przedstawionych wyżej działań na wartościach zmiennych, np

$$\psi(a' \oplus a'') = \psi(a') + \psi(a''),$$

i umożliwiają opisywanie zależności jakościowych między zmiennymi o różnym charakterze. Na przykład dla tranzystora:

$$I_c \stackrel{Q}{=} [I_b \oplus V_c]^S$$

W podobny sposób można wprowadzać *nierówności jakościowe*, rozszerzające możliwości modelu.

3 Kryteria wyboru wariantów w syntezie systemów eksperckich

Zastosowania SE obejmują coraz szersze obszary ludzkiej działalności, same systemy muszą więc być dostosowywane do różnorodnych i specyficznych potrzeb konkretnej dziedziny. Sprzyja temu duża różnorodność znanych wariantów tych systemów, ale zadanie wyboru utrudnia brak jednolitego opisu i analizy możliwości licznych rozwiązań.

W ramach projektu zaproponowano model procesu wnioskowania na tyle ogólny, że umożliwia objęcie bardzo zróżnicowanych i licznych rozwiązań, i na tyle szczegółowy, że może być podstawą do generowania praktycznych algorytmów.

3.1 Model procesu wnioskowania

Zjawiska zachodzące podczas wnioskowania w SE można przedstawić za pomocą zależności:

$$\{C, F, W\} \vdash_{\mathcal{R}} \mathcal{E}$$

oznaczającej, że wynik ekspertyzy $\mathcal{E}$ powstaje jako logiczny wywód (za pomocą reguł wnioskowania ze zbioru $\mathcal{R}$) z posiadanej wiedzy W i faktów F , przy określonym celu postępowania C .

Powiązane ze sobą cele i wyniki ekspertyzy tworzą kilka charakterystycznych typów, które za pomocą predykatów P i współczynników oceny α można, dla różnych rodzajów zadań, przedstawić w sposób następujący:

-zadania analizy

- | | | |
|-----------------|---|--|
| a- $C = P$, | <i>Czy to prawda że P?</i> | $\mathcal{E} = \{P\}$ lub $\mathcal{E} \in \{Y, N\}$ |
| b- $C = P(x)$, | <i>Dla jakich wartości x jest $P(x)$?</i> | $\mathcal{E} = \{P(a)\}$ |
| c- $C = xP$, | <i>W jakim stopniu zachodzi P?</i> | $\mathcal{E} = \{\alpha P\}$ |
| d- $C = X$, | <i>Co wynika z dostarczonych faktów?</i> | $\mathcal{E} = \{P_a\}$ |

-zadania syntezy

- e- $C = \langle X_a, X_b, \dots, X_z \rangle$, *Jakie akcje składają się na $\mathcal{E}$?*
 $\mathcal{E} = \{ \langle P_1, P_2, P_3 \rangle \}.$

Współczynniki oceny mogą dotyczyć dokładności, pewności, niezawodności itp.

Do wnioskowania potrzebne są wszystkie składniki modelu, ale jeden ze zbiorów C, F, W jest zwykle podstawą postępowania, a sposób uzyskiwania wyników określony jest typem reguły, wybranej ze zbioru $\mathcal{R}$.

3.2 Zasady wnioskowania

Zbiory $C, F, W, \mathcal{R}, \mathcal{E}$ są zazwyczaj redukowane (do $C, F, W, \mathcal{E}$) w procesie dostosowywania do konkretnego przypadku. Jeśli podstawą dla dalszych rozważań jest zbiór C – mamy *wnioskowanie na podstawie celu*, jeśli zbiór F – *wnioskowanie na podstawie faktów*, jeśli zbiór W – *wnioskowanie na podstawie modelu lub na podstawie znanych przypadków*. Te trzy podstawowe grupy można dalej rozszerzać, uzyskując metody wnioskowania na podstawie:

- celu
- danych, rozszerzające
- danych, modyfikujące
- danych, zężające i precyzujące
- modelu przyczynowego
- modelu funkcyjnego
- znanych przypadków

Dla każdego z tych przypadków tworzony jest opis, określający sposób wyznaczania wyników. Na przykład dla prostego przypadku wnioskowania rozszerzającego na podstawie danych będzie to wielofazowy proces iteracyjny, opisany zależnościami:

$$\begin{aligned} D_0 &= F_p; & \{C, D_i, W\} &\vdash_{\mathcal{R}} D_i^+; & i &= 0, 1, \dots, k \\ D_{i+1} &= D_i \cup D_i^+; & E &\in D_k^+. \end{aligned}$$

Inicjującymi danymi są fakty początkowe F_p , z których wywodzi się (np. przez dedukcję) nowe dane. Są one dodawane do poprzednich i proces powtarza się tak długo, aż w zbiorze uzyskanych danych znajdzie się poszukiwany wynik ekspertyzy (lub skończy się możliwość wyvodu).

3.3 Reguły wnioskowania

Wywód dokonywany w każdej fazie opisanego wyżej procesu może odbywać się z użyciem różnych reguł wnioskowania (ze zbioru $\mathcal{R}$). Rozważono następujące metody:

- dedukcja
- indukcja
- rezolucja
- dziedziczenie
- przyczynowość
- eliminacja

Dla każdej z nich tworzy się *schematy*, wprowadzające – obok wersji podstawowej – wersje z uzgodnieniami, podobieństwem i współczynnikami oceny. Na przykład dla prostej dedukcji schematy te mają postać następującą:

1. Schemat podstawowy

$$\frac{G, G \Rightarrow H}{H}$$

2. Uzgodnienia

$$\frac{G', G \Rightarrow H, G' \Theta = G \Theta}{H \Theta}$$

3. Podobieństwo

$$\frac{G', G \Rightarrow H, G' \sim G}{\gamma H}$$

4. Współczynniki oceny

$$\frac{\alpha_1 G_1, \alpha_2 G_2, \beta_1 (G_1 \Rightarrow H), \beta_2 (G_2 \Rightarrow H)}{\gamma H}$$

Trójwymiarowa przestrzeń, powstała przez uwzględnienie zasady wnioskowania, typu reguł i rodzaju schematu, zawiera wszystkie znane metody. Jednolita postać opisu i analizy ułatwia sformułowanie zasad wyboru wariantu, najbardziej przydatnego w konkretnym zastosowaniu.

4 Generowanie reguł z przykładów

Najsłabszym ogniwem w procesie tworzenia użytecznych SE jest zwykle etap pozyskiwania odpowiednio wiarygodnej, dokładnej i aktualnej wiedzy, będącej podstawą wnioskowania. Człowiek-ekspert często nie potrafi lub nie chce przekazać swych umiejętności, istotnego znaczenia nabierają więc metody automatycznego zdobywania wiedzy. Bogatym jej źródłem mogą być zwłaszcza komputerowe bazy danych.

Relacyjna baza danych z zależnościami funkcyjnymi między wartościami atrybutów tworzy *tablicę decyzyjną*, np. o postaci:

$$\begin{matrix} & c_1 & c_2 & \dots & c_n & d \\ \begin{matrix} 1 \\ 2 \\ \vdots \\ N \end{matrix} & \left[\begin{matrix} & & & & \\ & & & & \\ & & \vdots & & \\ & \dots & v_{ij} & & w_i \\ & & & & \end{matrix} \right] \end{matrix}$$

Wiersze opisują przypadki (obiekty, zdarzenia itp) w których wartości atrybutów warunkowych c determinują wartość atrybutu decyzyjnego d . Zbiór wierszy znanej tablicy decyzyjnej może być traktowany jako *zbiór treningowy*:

$$TS = \bigcup_{i=1}^K \langle V_i, w_i \rangle$$

służący do wyznaczania ogólnych zależności opisanych tablicą.

Każdy element z TS, np

$$\langle v_1, v_2, \dots, v_n, w \rangle,$$

może być określony regułą decyzyjną:

$$\text{JEŚLI } [c_1 = v_1] \wedge [c_2 = v_2] \wedge \dots \wedge [c_n = v_n] \text{ TO } [d = w].$$

Bezpośrednia transformacja tablic w reguły produkcji generuje bardzo dużo bardzo złożonych i specjalizowanych reguł, dlatego poszukuje się metod uzyskiwania zbioru reguł, optymalnego wg przyjętych kryteriów, a także właściwie klasyfikującego możliwie dużo przypadków spoza zbioru treningowego.

Istnieje wiele takich metod; argumentem za celowością poszukiwania dalszych są trudności w uzyskaniu reguł o zadanych właściwościach, co jest bardzo ważne z praktycznego punktu widzenia.

Proponowane tu podejście wykorzystuje *rachunek podziałów* do generowania reguł w sposób sterowany przez użytkownika. Umożliwia to wprowadzenie następujących kryteriów wyboru reguł:

- liczność zbioru reguł uzyskiwanych z TS,
- liczba selektorów w regułach,
- liczba aktywnych atrybutów,
- waga aktywnych atrybutów.

Przyjęta metodologia ułatwia także podejmowanie dodatkowych decyzji, dotyczących:

- formy selektorów,
- precyzji wartości atrybutów.

Oznacza to, że – oprócz typowych selektorów równościowych (o postaci $a = v$) – możliwe staje się uzyskiwanie selektorów nierównościowych, zbiorowych ($a \in V$) itp, a dyskretyzacja ciągłych wartości atrybutów sugeruje potrzebą precyzję ich wyznaczania.

Dodatkową zaletą proponowanej metody jest możliwość jej zastosowania w pewnych zadaniach *odkrywania wiedzy* z baz danych bez wydzielonych atrybutów decyzyjnych.

Literatura

- [1] W.Traczyk: "Learning best rules from examples", Intern.Conf. on Systems Science, Wrocław 1992.
- [2] R.Dzięciolowski, W.Traczyk: "Heuristic knowledge in formal reasoning on animal habitat quality", *Folia Forestalia Polonica*, n.35, 1993.
- [3] W.Traczyk: "Kryteria wyboru wariantów w syntezie systemów ekspertowych", II Krajowa Konf. *Inżynieria Wiedzy i Systemy Ekspertowe*, Wrocław 1993.
- [4] W.Traczyk: "Partitions in knowledge learning and discovery", *Foundations of Computing and Decision Sciences*, Vol.19, Nr 1-2, 1994.
- [5] W.Traczyk: "Modele jakościowe dla wnioskowania głębokiego", Krajowa Konf. *Analiza Decyzyjna, Systemy Ekspertowe, Zastosowania Systemów Komputerowych*, Warszawa 1994.
- [6] W.Traczyk: "Metody wnioskowania w systemach ekspertowych", rozdział w książce *Problemy Sztucznej Inteligencji* (red.W.Traczyk), Wiedza i Życie, 1995.
- [7] W.Traczyk: "Linguistic variables in qualitative modeling", Int.Conf. *System, Modeling, Control*, Zakopane, 1995.
- [8] W.Traczyk: "Qualitative models with quantitative operations", Int.Conf. *Intelligent Information Systems*, Augustów, 1995.
- [9] W.Traczyk: "Qualitative representation of deep knowledge", Int.Conf. *Knowledge Acquisition, Representation, and Processing*, Auburn, USA, 1995.