

Metody wnioskowania w systemach ekspertowych

Lidia Szyper

1. Wstęp

System ekspertowy (SE) jest programem komputerowym sugerującym, w ściśle określonej dziedzinie, rozwiązania pewnych problemów, przy tym mogącym wyjaśniać zasady uzyskiwania tej ekspertyzy oraz korzystać z informacji niedokładnych, niepełnych lub niepewnych, zarówno numerycznych jak i symbolicznych.

Głównymi składnikami **SE** są:

- a) *baza wiedzy* – wiedza o dziedzinie, której ma dotyczyć ekspertyza (np. związek między symptomami i nazwami jednostek chorobowych);
- b) *baza faktów* – szczegółowe dane o rozpatrywanym przypadku (np. symptomy choroby u konkretnej osoby);
- c) *maszyna wnioskująca* – wykorzystuje informacje z bazy faktów i bazy wiedzy do logicznego wnioskowania, prowadzącego do konkluzji czyli wyniku ekspertyzy; przy niejednoznacznym wyniku i braku jego pewności system wyposaża sugestie w pewne oceny wagi lub wiarygodności (np. zamiast konkluzji: „Jest to choroba A”; możemy uzyskać wynik: „Prawie pewne, że jest to choroba A”; „Możliwe, że jest to choroba A”);
- d) *blok pośredniczący (interfejs)* – umożliwia użytkownikowi korzystanie z systemu przez system menu, okien, itp.;
- e) *blok wyjaśniający* – wyjaśnia (na żądanie) zasadność zadawanych pytań i sugerowanych wniosków.

Wszystkimi blokami steruje centralny układ sterujący, zazwyczaj zlokalizowany w maszynie wnioskującej, często bowiem korzysta z tych samych lub podobnych mechanizmów wnioskowania.

Informacje opisujące dziedzinę zawarte są w bazach wiedzy i faktów, dzięki czemu maszyna wnioskująca oraz bloki pośredniczący i wyjaśniający

mogą być w pewnym stopniu od niej niezależne. Umożliwia to budowanie, tzw. *systemów szkieletowych* o pustych bazach wiedzy i faktów, które w miarę potrzeby zapełniane są przez użytkownika informacjami z interesującej go wiedzy.

Głównym procesem wykonywanym w **SE** jest wnioskowanie, z zakresu którego podstawowe metody zostaną przedstawione.

2. Model procesu wnioskowania.

Wnioskowanie (rozumowanie) w **SE** polega na stosowaniu pewnych ogólnych reguł do zgromadzonej wiedzy o dziedzinie i dostarczonych lub wywnioskowanych danych o konkretnym przypadku, dla uzyskania pożytecznych wniosków, będących materiałem do dalszego wnioskowania lub końcowym wynikiem ekspertyzy. Charakter poszukiwanego wyniku i sposób wnioskowania zależą od celu postawionego przed systemem. Proces uzyskiwania wyniku (sesję) można traktować jako dowodzenie odpowiedniej wiedzy i właściwych reguł wnioskowania. Ogólna postać:

$$\{C, F, W\} \vdash_R E$$

(jest to wyprowadzenie wyników ekspertyzy **E** z faktów (danych) **F** i wiedzy **W**, przy określonym celu **C** i przy zastosowaniu reguł wnioskowanie ze zbioru **R**).

Reguły wnioskowania – ogólne zasady, pokazujące w jaki sposób można wykorzystać pewien element wiedzy i określone dane do uzyskania nowych informacji. Postać ogólna

$$\frac{\text{przesłanka 1, przesłanka 2, \dots}}{\text{wniosek}}$$

(gdy prawdziwe są wszystkie przesłanki, to prawdziwy jest również wniosek (konkluzja)).

Wyróżnia się:

wnioskowanie płytke – związek między przyczyną i skutkiem nie musi być oczywisty czy uzasadniony i często ma charakter heurystyczny;

wnioskowanie głębokie – każdy krok rozumowania ma swe uzasadnienie tworzące zwykle logiczny łańcuch przyczynowo–skutkowy.

Pierwszym etapem pracy złożonych **SE** jest dostosowanie zasobów systemu do postawionego zadania, czyli dokonywany jest wybór odpowiednich podzbiorów *C, F, W, R, E* zbiorów **C, F, W, R, E**. Następnym etapem jest właściwe wnioskowanie, składające się z sekwencji *cykli* o powtarzających się czynnościach.

Wyróżnia się trzy metody wnioskowania w zależności od składnika (cel, fakty, wiedza, reguły) będącego podstawą działań. Są to: wnioskowanie na podstawie faktów (danych), wnioskowanie na podstawie celu i wnioskowanie na podstawie wiedzy.

3. Wnioskowanie na podstawie danych

Wnioskowanie na podstawie danych (w przód, wstępujące) polega na stosowaniu posiadanej wiedzy do znanych faktów, przez użycie reguł wnioskowania. Wnioski, będące wynikiem zastosowania reguł, stanowią podstawę powtarzającego się procesu wnioskowania (do momentu wygenerowania wyniku spełniającego cel ekspertyzy lub napotkania innych ograniczeń), wpływają w różny sposób na zbiór danych w systemie i dlatego istnieje kilka wariantów rozwiązań:

- a) *rozszerzanie danych*: na podstawie faktów początkowych F_p wyselekcjonowanej wiedzy w dziedzinie W i etapowego celu C dedukuje się dane D_1 , służące do dalszego wnioskowania o nowych danych D_i . Wszystkie te dane gromadzi się, szukając w nich wyniku E , określonego przez cel.

$$D_0 = F_p; \quad \{C, D_i, W\} \vdash_R D_i^+; \quad i = 0, 1, \dots, k-1$$

$$D_{i+1} = D_i \cup D_i^+; \quad E \in D_k^+.$$

W przypadku rozległej wiedzy i licznych danych, cykl wyprowadzania danych D_i^+ rozbijany jest na dwie fazy:

- a) $\{C, D_i, W\} \vdash_R W_i$ – selekcja danych;
 b) $\{C, D_i, W\} \vdash_R D_i^+$ – generacja nowych danych, wykorzystuje się w tej fazie dedukcję i rezolucję.

- b) *modyfikowanie danych*: zmniejszanie zasobów danych,

$$D_0 = F_p; \quad \{C, D_i, W\} \vdash_R \langle D_i^+, D_i^- \rangle$$

$$D_{i+1} = D_i \setminus D_i^- \cup D_i^+; \quad E \in D_k^+ \text{ lub } E = D_k.$$

Regułą najczęściej stosowaną jest dedukcja.

- c) *zawężanie danych*: stopniowa redukcja przestrzeni możliwych rozwiązań, bliższa zasadom wnioskowania na podstawie modelu, przedstawionych niżej.

*Reguły wnioskowania***Dedukcja**

Podstawą dedukcji jest reguła *modus ponens*, opisywana schematem

$$\frac{P, P \Rightarrow Q}{Q}$$

(jeśli prawdziwy jest predykat P i zachodzi implikacja $P \Rightarrow Q$, to prawdziwy jest też predykat Q).

Systemy regułowe, wykorzystujące implikację lub jej odpowiednik będący częścią reguły modus ponens, zwane są często *regułami produkcji*. W regułach tych występują zwykle *formuły predykatowe* – predykaty połączone są w nich spójnikami **i**, **lub**, **nie**.

Postać ogólna dedukcji, gdzie G i H są formułami:

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

Schemat z uzgodnieniami – uzgodnienie (unifikacja) zmiennych w regułach produkcji ze stałymi wartościami w faktach i wnioskach:

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

gdzie $G \Theta$ oznacza zastosowanie unifikatora Θ do formuły G .

Schemat z podobieństwem –

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

gdzie s jest stopniem podobieństwa dwóch formuł lub predykatów,
 γ jest stopniem oceny pewności konkluzji (jest funkcją s).

Schemat ze współczynnikami oceny:

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

gdzie $\gamma = f(\alpha_1, \alpha_2, \beta_1, \beta_2)$ jest stopniem pewności konkluzji składowych. Przyjmuje się pewien próg współczynnika oceny, przekroczenie którego jest niezbędne, by oznaczony współczynnikiem oceny element – reguły (fakt, formuła, implikacja) mógł uczestniczyć w dalszym wnioskowaniu.

Wyżej wymienione schematy mogą być stosowane łącznie.

Przedstawione zasady dedukcji dotyczą podstawowego przypadku wnioskowania z rozszerzeniem danych. Przy modyfikacji danych implikacja

(reguła produkcji), w swym następniku, musi zawierać dwie możliwości: dołączenie nowych danych do poprzedniego zbioru i usunięcie niektórych danych ze zbioru (przez wprowadzenie akcji, czyli składnika następnika implikacji, który poleca wykonanie pewnej czynności, np. typu dołącz i usuń).

Rezolucja

Rezolucja polega na zastąpieniu implikacji w regułach wnioskowania równoważną postacią dyzjunkcyjną (alternatywą literałów, czyli predykatów lub ich negacji), zwaną często *klauzulą*. Stosuje się ją, aby kierunek wnioskowania wymuszony przez występowanie implikacji w regułach wnioskowania, nie ograniczał zakresu zastosowania wiedzy, opisanej przez te implikacje.

Schemat podstawowy

$$\frac{G_1^c \vee P, G_2^c \vee \neg P}{G_1^c \vee G_2^c},$$

gdzie G^c – klauzula odpowiadająca formułom predykatywnym G ,
 P – predykat.

Schemat z uzgodnieniem:

$$\frac{G_1^c \vee P', G_2^c \vee \neg P, P' \Theta = P \Theta}{G_1^c \Theta \vee G_2^c \Theta}$$

Schemat ze współczynnikami oceny:

$$\frac{\beta_1(G_1^c \vee P), \beta_2(G_2^c \vee \neg P)}{\gamma(G_1^c \vee G_2^c)}.$$

Rezolucji na podstawie podobieństwa nie stosuje się.

4. Wnioskowanie na podstawie celu

Wnioskowanie na podstawie celu (wstecz, zstępujące) polega na stosowaniu posiadanej wiedzy poczynając od predykatu opisującego cel wnioskowania. Cel i dostępna wiedza o dziedzinie umożliwiają określenie *podcelów* – predykatów bezpośrednio warunkujących spełnienie celu. Każdy podcel może być uznany za spełniony lub jest traktowany jako cel chwilowy i dalej rozwijany w zbiór swoich własnych podcelów.

Podcel jest *spełniony* jeśli:

- istnieje zgodny z nim fakt (identyczny, podobny, lub dający się uzgodnić); lub

- dostępna jest procedura wyliczająca predykat zgodny z podcelem; lub
- istnieje możliwość zapytania użytkownika o prawdziwość (i ewentualnie wartość zmiennych) predykatu opisującego podcel i odpowiedź jest zgodna z tym podcelem; lub
- nastąpi wykonanie akcji, jeśli podcel jest akcją.

Cel pierwotny jest prawdziwy, gdy wszystkie podcele na różnych poziomach okażą się prawdziwe dla jakiegoś podstawienia zmiennych i brak jest dalszych podcelów i pytań. W przeciwnym przypadku, system melduje brak możliwości udowodnienia zadanego celu.

Warianty rozwiązań.

$$C_0 = C, D_0 = F_p \quad \{C_i, D_i, W\} \vdash_R \langle P_{t_i}, C_{i+1} \rangle$$

$$D_{i+1} = D_i \cup Od_i \quad \{C_k, D_k, W\} \vdash_R \emptyset$$

$$E = C \Theta \text{ lub } E = \emptyset$$

gdzie P_t – pytania, Od – odpowiedzi, a k -ty krok oznacza zakończenie wnioskowania.

Fazy cyklu wnioskowania:

a) $\{C_i, D_i, W\} \vdash_R \langle C'_i, C''_i, P_{t_i} \rangle,$

jest to *faza generowania podcelów*, gdzie podcele będące akcjami wykonywane są natychmiast; C' – podcele dające się uzgodnić z dostępnymi danymi lub wyliczyć dostępnymi procedurami; C'' – cele chwilowe; P_t – podcele, o które można (lub trzeba) pytać użytkownika. Jeśli (P_t) nie jest pusta – to system generuje odpowiednie pytania;

b) $\{C'_i, C''_i, Od_i, D_i, W_i\} \vdash_R C_{i+1},$

jest to *faza selekcji celów* – analiza ewentualnych odpowiedzi użytkownika i wartościowanie użytecznych celów. Wybór z grupy celów chwilowych, celów udowodnionych w kolejnym cyklu.

W obu fazach stosuje się przeszukiwanie:

- *w głąb* – spośród wariantów zastąpienia podcelami celu głównego lub chwilowego, wybiera się jeden i dla niego kontynuowane jest wnioskowanie;
- *w szerz* – równoczesne analizowanie wszystkich wariantów.

*Reguły wnioskowania***Indukcja**

Indukcja, w przeciwieństwie do dedukcji, sugeruje, by dla wskazania prawdziwości H , gdy zachodzi $G \Rightarrow H$, udowodnić prawdziwość G .

Schemat podstawowy

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

gdzie $H?$ i $G?$ oznaczają hipotezy formuł H i G .

Schemat z uzgodnieniami

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

Schemat z podobieństwem

$$\frac{H?, G \Rightarrow H', H' \approx H?}{\gamma G?},$$

gdzie γ – współczynnik oceny roli predykatu $G?$ w dowodzeniu celu $H?$.

Stosuje się dwuetapowe postępowanie w przypadku, gdy dostępne fakty, odpowiedzi na pytania lub reguły produkcji nie są pewne:

- 1) dowodzi się prawdziwości celu przy użyciu podanych wyżej zależności dla wnioskowania indukcyjnego, wyznaczając potrzebne dane i reguły;
- 2) dla wyznaczonych dróg dowodzenia (od faktów do celu) wylicza się kolejno współczynniki oceny (analogicznie do metody dedukcyjnej).

Jeśli dróg dowodzenia jest kilka, system może porządkować rozwiązania wg wartości globalnych współczynnika oceny lub zaproponować tylko wyniki najlepsze.

Rezolucja indywidualna

Implikacje z koniunkcją pozytywnych predykatów w poprzedniku i pojedynczym pozytywnym predykatem w następniku można zastąpić klauzłą Horna klauzłą o jednym tylko pozytywnym predykatem. Jeśli dla pozytywnego predykatu jednej klauzuli odnajdzie się taki sam lecz zanegowany predykat w innej klauzuli, to wynikiem będzie jedna, prostsza klauzula. Mechanizm dowodzenia polega na określeniu, za pomocą klauzuli, podcelów celu głównego i analizie każdego z nich. Jeśli brak jest danych zgodnych z tym podcelem oraz procedury mogącej go określić, to poszukuje się klauzuli pozwalającej zastąpić (z pomocą rezolucji) cel chwilowy przez jego własne

podcele. Gdy nie istnieje odpowiednia klauzula, system pyta użytkownika o prawdziwość (i ewentualne wartości zmiennych) predykatu opisującego cel chwilowy i uwzględnia odpowiedź.

Uzgadnianie wartości zmiennych oraz wprowadzanie współczynników oceny odbywa się tu na takich samych zasadach jak w powyższej rezolucji.

Rezolucja globalna

W sposób globalny bada się przydatność zbioru klauzul do udowodnienia znanego celu, przy pomocy sprzeczności i falsyfikacji (*reductio ad absurdum*).

Jeśli ze zbioru formuł lub klauzul wyciągnąć można przeciwstawne wnioski, to zbiór taki uznaje się za sprzeczny. Do wyciągania wniosków służy rezolucja, wówczas dowodem sprzeczności jest możliwość wprowadzenia klauzuli pustej \square

$$\frac{P, \neg P}{\square}$$

Jeśli zbiór klauzul uznanych za prawdziwe nie jest sprzeczny, a staje się sprzeczny po powiększeniu o pewną klauzulę, to oznacza, że dołączona klauzula nie jest prawdziwa.

Reguła falsyfikacji

$$\frac{\Psi^c, \{\Psi^c, \neg G^c\} \vdash \square}{G^c},$$

jeśli Ψ^c jest niesprzecznym zbiorem klauzul, a $\Psi^c \cup \neg G^c$ jest zbiorem sprzecznym, to formuła $\neg G^c$ nie jest prawdziwa, czyli prawdziwa jest klauzula G^c .

Uzgodnienia i współczynniki oceny wprowadza się analogicznie, jak przy wnioskowaniu na podstawie danych.

W złożonych **SE** często stosuje się *dwukierunkowe wnioskowanie* (w przód i wstecz) wykorzystujące do wnioskowania zarówno dane jak i cel (lub cele chwilowe).

5. Wnioskowanie na podstawie modelu

Podstawą wnioskowania, obok danych i celu, może być wiedza zawarta w bazie wiedzy. Jeśli jest to *wiedza strukturalizowana* (zawiera *zgrupowania* informacji i *relacje* między nimi) opisująca z odpowiednią dokładnością budowę, działanie lub strukturę dziedziny zainteresowań, to uznaje się ją za *model* dziedziny. Poniższe rozważania dotyczyć będą modelu w postaci drzewa, w którym wiedza związana z węzłami zawiera informacje, które użytkownik może dostarczyć lub których oczekuje od systemu. Są one tym bardziej precyzyjne, im wyższy poziom drzewa.

Celem wnioskowania jest zazwyczaj osiągnięcie takiego węzła (lub węzłów) drzewa, w którym informacje są na tyle precyzyjne, że satysfakcjonują użytkownika, a zakres informacji, które system chce uzyskać nie przekracza możliwości użytkownika (często celem tym są liście drzewa).

$$W_0^* = W_0, D_0 = F_p \quad \{C, D_i, W_i^*, W_i\} \vdash_R \langle P_{t_i}, W_{i+1}^* \rangle$$

$$D_{i+1} = D_i \cup Od_i \quad E \in W_k^*,$$

gdzie W_0 – wiedza związana z korzeniem,

W_i – wiedza związana z poziomami drzewa,

W_i^* – wiedza adekwatna, wiedza W_i zawężona w wyniku wnioskowania,

D_i – posiadane dane,

P_t – pytania systemu mające na celu zdobycie nowych danych.

Dwie fazy cyklu:

$$a) \{D_i, W, W_i^*\} \vdash_R \langle P_{t_i}, W_{i+1}^+, W_{i+1}^- \rangle,$$

W fazie tej następuje: *generacja hipotez* – wszystkich węzłów wyższego poziomu związanych z rozwiązaniem; wstępna *weryfikacja hipotez* – badanie zgodności informacji w węzłach z dostępnymi danymi; *generacja pytań* – dla dokładniejszej weryfikacji węzłów.

$$b) \{W_{i+1}^+, Od_i, C\} \vdash_R W_{i+1}^*,$$

w fazie tej ma miejsce dalsza *weryfikacja hipotez* na podstawie uzyskanych odpowiedzi.

Reguły wnioskowania

Podstawowe pojęcia:

a) *rama* – zgrupowane informacje w węzłach

$$\mathbf{RM} = \langle \mathbf{RM}_P, \mathbf{RM}_K, \mathbf{RM}_W, \mathbf{RM}_D \rangle,$$

gdzie \mathbf{RM}_P – przesłanki ramy,

\mathbf{RM}_K – konkluzje ramy,

\mathbf{RM}_W – wskaźniki ramy,

\mathbf{RM}_D – dodatkowe informacje.

b) *rama potwierdzona* – jeśli po dotarciu przez system do ramy \mathbf{RM} prawdziwe są jej przesłanki, to prawdziwe są również jej konkluzje.

c) *dziedziczenie* informacji z niższych poziomów.

Generowanie hipotez

Generowanie hipotez polega na przechodzeniu od ram poziomu niższego, bardziej ogólnych już potwierdzonych, do związanych z nimi (wskaźnikami) ram poziomu wyższego.

Schemat podstawowy

$$\frac{RM_p, PM' \in \Phi(RM)}{RM_k?},$$

gdzie $RM_k?$ hipoteza prawdziwości zbioru predykatów RM_k ;

$\Phi(RM)$ zbiór wszystkich związanych z RM ram wyższego poziomu.

Schemat z podobieństwem

$$\frac{RM_p, RM_p \stackrel{S}{\sim} RM'_p, RM'' \in \Phi(RM')}{\gamma RM_k?}$$

Schemat ze współczynnikami oceny

$$\frac{\alpha RM_p, RM' \in \Phi(RM)}{\gamma RM_k?},$$

gdzie γ zależy od wartości α .

Weryfikacja hipotez

Weryfikacja hipotez polega na uznaniu hipotezy za prawdziwą, po stwierdzeniu prawdziwości wszystkich predykatów – przesłanek, czyli przynależności tych predykatów do zbioru stwierdzonych faktów (danych) D .

Schemat podstawowy

$$\frac{RM_k?, RM_p \subseteq D}{RM_k}$$

Schemat z podobieństwem

$$\frac{RM_k?, RM_p \stackrel{S}{\sim} D}{\gamma RM_k}$$

Schemat ze współczynnikami oceny

$$\frac{\alpha RM_k?, \beta(RM_p \subseteq D)}{\gamma RM_k},$$

gdzie β może określać zakres w jakim predykaty zbioru RM_p znajdują pokrycie w danych D .

Zapewnienie zgodności przesłanek z danymi bywa rozwiązane przez podział przesłanek na dwie grupy:

- 1) zbiór warunków koniecznych: zgodnych całkowicie z danymi;
- 2) zbiór warunków pomocniczych: podwyższają wiarygodność hipotezy jeśli są zgodne z danymi.

Weryfikacja hipotez, oparta na badaniu zgodności przesłanek z danymi, może być uzupełniona o *eliminację hipotez*, polegającą na wykluczaniu z rozważań ram, których przesłanki są sprzeczne z posiadanymi danymi:

$$\frac{RM_k?, ((RM_p, D) \vdash \square)}{\neg RM_k}$$

Wnioskowanie z hierarchicznym generowaniem i weryfikacją hipotez rozpoczyna się od zaproponowania hipotezy dostatecznie ogólnej by mieć pewność, że jest prawdziwa; wygenerowane z niej hipotezy wyższego poziomu będą podlegały weryfikacji. Ramy zweryfikowane (potwierdzone) służą do dalszej generacji itd.

Uwagi końcowe

Przedstawione zasady wnioskowania są podstawowymi metodami wykorzystywanymi w górnych poziomach decyzji podejmowanych przez konstruktora *SE* lub użytkownika wybierającego system dla konkretnego zastosowania. Niewystarczające w nich okazuje się traktowanie niedokładności i niepewności informacji wyłącznie w aspekcie numerycznym, w związku z tym przydatne okazało się wykorzystanie zbiorów przybliżonych i rozmytych pozwalające wnioskować z niepewnych i niedokładnych przesłanek.

Bibliografia

1. Bubnicki Z., *Wstęp do systemów ekspertowych*, PWN, 1990.
2. Szałas A., *Zarys dedukcyjnych metod automatycznego wnioskowania*, Akademicka Oficyna wydawnicza, 1992.
3. Traczyk W., *Kryteria wyboru wariantów w syntezie systemów ekspertowych*, Mat. IIKKN Inżynieria Wiedzy i Systemy Ekspertowe, Wrocław, 1993.
4. Wójcik M., *Zasada rezolucji. Metoda automatycznego wnioskowania*, PWN Warszawa, 1991.