SYSTEM DIAGNOSTYCZNY OPARTY NA LOGICE DOMNIEMAŃ

Ewa Madalińska

na podstawie prac:

[1] R. Reiter. A Logic for Default Reasoning. *Artificial Intelligence Journal*, **13**, 81-132, 1980.

[2] W. Łukaszewicz. (1988) "Considerations on Default Logic: An Alternative Approach". *Computational Intelligence*, 44[1], 1-16.

[3] R. Reiter. (1987) "A theory of Diagnosis from First Principles". *Artificial Intelligence*, 32, 57-95.

Problem diagnozy

Problemem diagnozy nazywamy następujące zadanie:
dane jest urządzenie U, składające się z elementów $e_1, \ldots, e_n$, oraz zestaw obserwacji $o_1, \ldots, o_k$ dotyczących jego działania. Należy sprawdzić, czy obserwacje $o_1, \ldots, o_k$ są zgodne z zakładanym działaniem urządzenia i, jeśli nie, określić te elementy spośród $e_1, \ldots, e_n$, których wadliwe działanie może mieć na to wpływ.

Podejścia do rozwiązania problemu

Dwa podejścia:

- Diagnoza formalna
  - opis systemu – urządzenia fizycznego czy dziedziny świata rzeczywistego
  - obserwacja zachowania się tego systemu.

  Niezgodność oznacza, że pewne jego elementy działają wadliwie. Powstaje wtedy problem diagnostyczny polegający na wyznaczeniu elementów, których wadliwe działanie tłumaczyłoby rozbieżnośc między obserwowanym a przewidywanym zachowaniem systemu.
- Diagnoza empiryczna
  - informacja statystyczna o poprzednich zachowaniach systemu, np. "Jeżeli obserwujesz takie to a takie wadliwe działanie, to w 90% zepsuty jest element X".

Zakłada się, że zarówno opis działania badanego urządzenia U jak i obserwacje $o_1, \ldots, o_n$ zapisane są w języku logiki. W celu określenia czy obserwacje zgodne są z zakładanym działaniem urządzenia, jak również znalezienia ewentualnych wadliwych elementów, wykorzystuje się niemonotoniczną logikę domniemań.

Język reprezentacji systemów – język logiki pierwszego rzędu.

Opis systemu – specyfikacja jego zachowania przy założeniu, że wszystkie elementy działają poprawnie.

Problem diagnostyczny – obserwowane działanie jest niezgodne z przewidywanym.

Diagnoza nie musi być unikalna; może być wiele alternatywnych wyjaśnień dla tego samego wadliwego systemu.

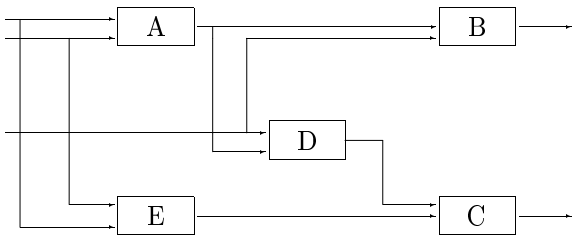
Problem obliczeniowy – wyznaczenie wszystkich możliwych diagnoz dla pewnego wadliwego systemu.

Formalizacja problemu

Definicja 1
System jest to para (OS,ELEM), gdzie:
(1) OS – opis systemu – jest zbiorem zdań logiki pierwszego rzędu.
(2) ELEM – elementy systemu – skończony zbiór stałych indywiduowych.

W opisie systemu wyróżnia się predykat unarny AB(.), interpretowany jako "wadliwy" (ang. abnormal).

Przykład 1
Rozważmy układ zbudowany z bramek logicznych:



A i B są bramkami XOR (sumy wyłącznej), D i E są bramkami AND, C jest bramką OR.

Układ ten może być reprezentowany przez system z 5 elementami A,B,C,D,E oraz następujący opis:

a) opis funkcjonowania poszczególnych elementów

$\mathrm{AND}(x) \wedge \neg ab(x) \Rightarrow out(x) \equiv in1(x) \wedge in2(x)$,
$\mathrm{XOR}(x) \wedge \neg ab(x) \Rightarrow out(x) \equiv \neg in1(x) \wedge in2(x) \vee in1(x) \wedge \neg in2(x)$,
$\mathrm{OR}(x) \wedge \neg ab(x) \Rightarrow out(x) \equiv in1(x) \vee in2(x)$

b) opis rodzajów poszczególnych elementów

AND(D),
AND(E),
XOR(A),
XOR(B),
OR(C)

c) opis powiązań poszczególnych elementów

$out(A) \equiv in2(D)$,
$out(A) \equiv in1(B)$,
$out(D) \equiv in1(C)$,
$out(E) \equiv in2(C)$,
$in1(D) \equiv in2(B)$,
$in1(A) \equiv in1(E)$,
$in2(A) \equiv in2(E)$

Obserwacje systemu

Definicja 2
Obserwacją systemu nazywamy skończony zbiór zdań logiki pierwszego rzędu.
System z obserwacjami oznaczać będziemy jako trójkę (OS,ELEM,OBS).∎

Przykład 2
Załóżmy, że do systemu z przykładu 1. podano na wejściu 1,0,1 i otrzymano na wyjściu 1,0. Obserwacja ta może być zapisana jako,

$$in1(A), \neg in2(A), in1(D), out(B), \neg out(C)$$

Obserwacja ta wskazuje, że układ jest wadliwy; oba wyniki na wyjściu są niepoprawne.∎

Wadliwość systemu (OS,$\{c_1, ..., c_n\}$,OBS) – obserwacje OBS są rozbieżne z wynikającym z opisu zachowaniem systemu.

Diagnozy

$\{\neg ab(c_1), ..., \neg ab(c_n)\}$ – przypuszczenie, że wszystkie elementy systemu działają właściwie.

Fakt, że obserwacje OBS są niezgodne z działaniem systemu przy założeniu, że wszystkie jego elementy działają poprawnie, może być sformalizowany przez:

$$\mathrm{OS} \cup \{\neg ab(c_1), ..., \neg ab(c_n)\} \cup \mathrm{OBS} \qquad (1)$$

jest sprzeczny.

Intuicyjnie, diagnoza – założenie, że pewne elementy są wadliwe, a pozostałe działają dobrze.

Zadanie – wyspecyfikować wadliwe elementy tj. wyjaśnić źródło sprzeczności (1).

Naturalną drogą jest usunięcie dostatecznej ilości przypuszczeń $\neg ab(c_i)$, gdzie $1 \leq i \leq n$, aż do otrzymania niesprzeczności.

Racjonalnie będzie zastosować zasadę:

Diagnoza jest założeniem, że pewien minimalny zbiór elementów jest wadliwy.

Definicja 3
Diagnozą dla (OS, ELEM, OBS) jest minimalny zbiór $\Delta \subseteq$ ELEM taki, że

$$\text{OS} \cup \text{OBS} \cup \{ab(c) \mid c \in \Delta\} \cup \{\neg ab(c) \mid c \in \text{ELEM} - \Delta\} \quad (2)$$

jest niesprzeczny.▪

Diagnoza jest wyznaczona przez minimalny zbiór elementów o następującej własności:

założenie, że każdy z tych elementów jest wadliwy (abnormal) wraz z założeniem, że wszystkie pozostałe są dobre (not abnormal) jest zgodne (niesprzeczne) z opisem systemu i obserwacją.

Przykład 3
Z układem z przykładu 2. związane są 3 diagnozy: {A}, {B,C}, {B,D}.▪

Własności diagnoz

Twierdzenie 1
Dla systemu (OS,ELEM,OBS) istnieje diagnoza wtt OS $\cup$ OBS jest niesprzeczny.▪

Twierdzenie 2
{ } jest diagnozą (i jest to jedyna diagnoza) dla (OS,ELEM,OBS) wtt

$$\text{OS} \cup \text{OBS} \cup \{\neg ab(c) \mid c \in \text{ELEM}\}$$

jest niesprzeczny.▪

Znaczy to, że obserwacje są zgodne z opisem systemu i wszystkie jego elementy działają poprawnie.

Twierdzenie 3
Jeżeli Δ jest diagnozą dla (OS,ELEM,OBS), to dla każdego $c_i \in \Delta$

$$\text{OS} \cup \text{OBS} \cup \{\neg ab(c) \mid c \in \text{ELEM} - \Delta\} \models ab(c_i)$$

▪

Twierdzenie 4
$\Delta \subseteq$ ELEM jest diagnozą dla (OS,ELEM,OBS) wtt Δ jest minimalnym zbiorem takim, że

$$\text{OS} \cup \text{OBS} \cup \{\neg ab(c) \mid c \in \text{ELEM} - \Delta\}$$

jest niesprzeczny. ▪

Wpływ dodatkowych obserwacji na diagnozy

Załóżmy, że (OS,ELEM,OBS) ma więcej niż jedną diagnozę. Wybór jednej z nich wymaga dalszych informacji o systemie tj. dokonania pomiarów np.: przeprowadzenia testu laboratoryjnego na pacjencie, wstawienia miernika w obwód. Wynik pomiaru będzie skończonym ciągiem zdań pierwszego rzędu POM.

Jaki jest związek pomiędzy diagnozami dla (OS,ELEM,OBS) a (OS,ELEM,OBS $\cup$ POM) ?

Definicja 4
Diagnoza Δ dla (OS,ELEM,OBS) pociąga Π (zdanie pierwszego rzędu) wtt

$$\text{OS} \cup \text{OBS} \cup \{ab(c) \mid c \in \Delta\} \cup \{\neg ab(c) \mid c \in \text{ELEM} - \Delta\} \models \Pi$$▪

Twierdzenie 5
Załóżmy, że każda diagnoza dla (OS,ELEM,OBS) pociąga jedną z Π i $\neg\Pi$. Wtedy:

1. Każda diagnoza dla (OS,ELEM,OBS), która pociąga Π jest diagnozą dla (OS,ELEM,OBS $\cup\ \Pi$).
2. Żadna diagnoza dla (OS,ELEM,OBS), która pociąga $\neg\Pi$ nie jest diagnozą dla (OS,ELEM,OBS $\cup\ \Pi$).
3. Diagnoza dla (OS,ELEM,OBS $\cup\ \Pi$), która nie jest diagnozą dla (OS,ELEM,OBS), jest właściwym nadzbiorem pewnej diagnozy dla (OS, ELEM,OBS), która pociąga $\neg\Pi$.▪

Z twierdzenia wynika, że:

- pomiar, który nie jest sprzeczny z żadną diagnozą, nie wnosi nowej informacji.
- dodatkowa informacja Π pozwala rozróżniać konkurencyjne diagnozy w ten sposób, że:
  - zachowują się te diagnozy, które pociągają Π,
  - odrzuca się te, które pociągają $\neg\Pi$,
  - mogą powstać nowe diagnozy.

Obliczanie diagnoz nie jest możliwe w najbardziej ogólnym przypadku, ponieważ definicja diagnozy odwołuje się do testu niesprzeczności dla zbioru formuł pierwszego rzędu. Problem ten jest **nierozstrzygalny**.

Niemniej jednak jest wiele praktycznych zastosowań opierających się na takich podzbiorach języka logiki pierwszego rzędu, dla których niesprzeczność jest rozstrzygalna i diagnozy są obliczalne. Należy do nich między innymi przypadek sieci logicznych. W problemach, które mają modele prowadzące do teorii nierozstrzygalnych należy zastosować podejście heurystyczne.

Zastosowanie logiki domniemań do obliczania diagnoz

Wnioskowanie diagnostyczne – niemonotoniczne: dodatkowe informacje mogą obalić pewne diagnozy systemu.

Istnieje ścisły związek między wnioskowaniem diagnostycznym opartym na logice pierwszego rzędu a logiką domniemań.

Z systemem (OS,ELEM) wraz z obserwacjami OBS wiążemy teorię domniemań $DT = \langle A, D\rangle$, gdzie:

$A = \text{OS} \cup \text{OBS}$

$D = \{: \neg ab(c)/\neg ab(c) \mid c \in \text{ELEM}\}$.

Wzajemnie jednoznaczną odpowiedniość pomiędzy diagnozami dla (OS,LELEM,OBS) a rozszerzeniami powyższej teorii domniemań wyznacza następujące twierdzenie:

Twierdzenie 6
E jest rozszerzeniem dla DT wtt dla pewnej diagnozy Δ dla (OS,ELEM,OBS)

$$E = \{\Pi \mid \Delta \text{ pociąga } \Pi\}.$$

▪

Posługując się pojęciem domniemań generujących rozszerzenie, twierdzenie to można by sformułować następująco:

E jest rozszerzeniem dla DT wtt dla pewnej diagnozy Δ zachodzi:

$$\Delta = \text{ELEM} - \{c \in \text{ELEM} \mid (: \neg ab(c)/\neg ab(c)) \in GD(E)\}.$$