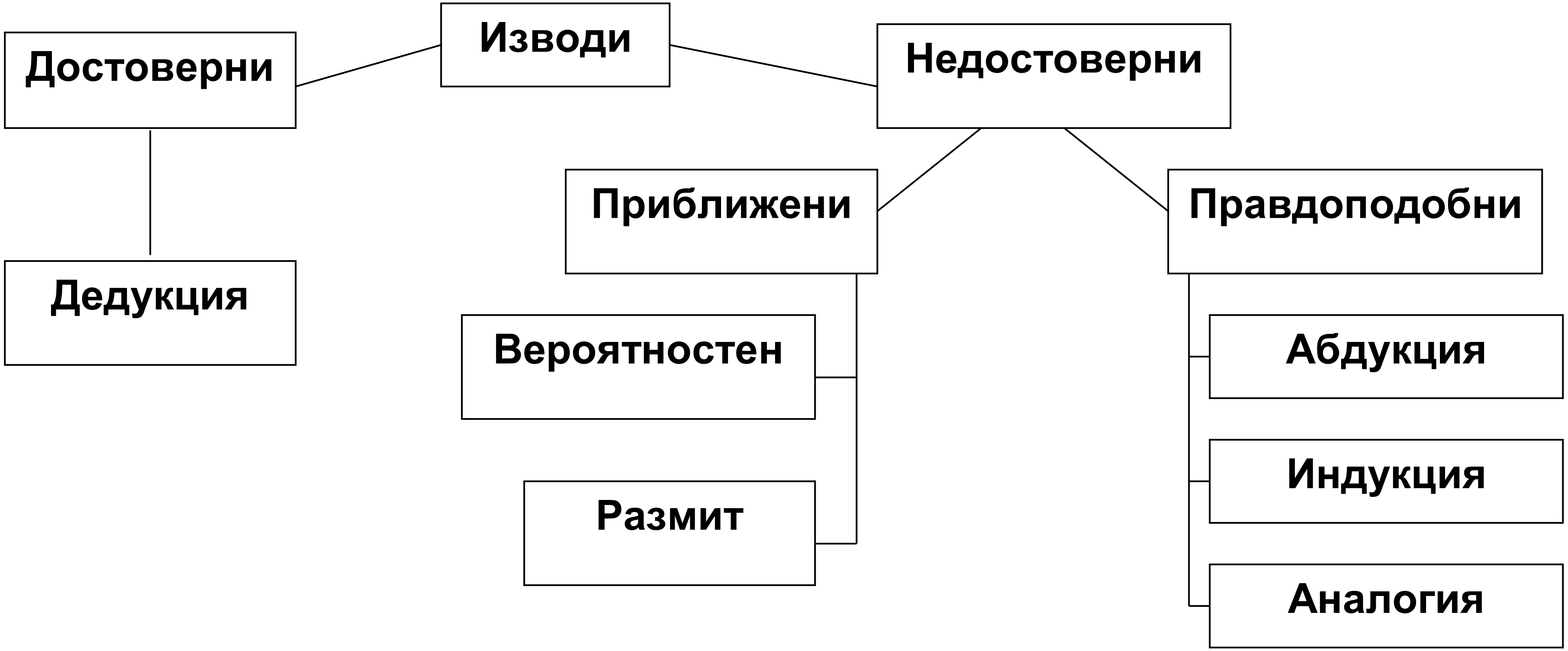
Изкуствен интелект - летен семестър, 2023/2024 учебна година

Тема 9:
Работа с несигурни знания в СОЗ

Видове изводи в СОЗ



Дедукция. Теоретична основа на дедукцията е правилото за извод Modus Ponens (MP). Същност на MP:

(ако A, то B)	условно вярно твърдение
A	изпълнено условие
B	изведено заключение

По-точно, дедуктивният извод има формата:

$$\frac{\begin{array}{l}(\forall x)\ (\text{ако } A(x), \text{ то } B(x))\\ A(a)\end{array}}{B(a)}$$

При това като модел на твърденията от вида (ако A, то B) обикновено се използва традиционната импликация A -> B. В действителност тя не означава непременно причинно-следствена връзка между A и B, но в повечето случаи върши работа (същевременно импликацията е добре изучена, а причинността е много сложно понятие).

Интерпретаторът на правилата в системите, основани на правила, по същество извършва дедуктивен извод (прав или обратен).

Абдукция. Абдукцията е генериране на правдоподобни обяснения за това, което наблюдаваме около нас. Тя може да се разглежда в следната форма:

$$\frac{(\text{ако } A, \text{ то } B) \quad B}{A}$$

По-точно, абдукцията би трябвало да се разглежда като правило за извод от вида

$$\frac{(\text{причина } ?x\ ?y) \quad ?y}{?x}$$

Пример. Когато хората са пияни, те не могат да пазят равновесие. Ако Джак не може да пази равновесие, бихме могли да предположим, че той е пиян. Естествено, това е само едно предположение, което може да се окаже и невярно (причината за неспособността му да пази равновесие може да бъде съвсем друга).

Индукция. Индуктивният извод е опит за обобщение на базата на общи признаци, наблюдавани у голям брой конкретни обекти. При натрупване на допълнителни знания за средата достоверността на обобщението може съществено да се повиши. От тази гледна точка може да се твърди, че способността за индуктивен извод е съпоставима със способността на човека за обучение и самообучение.

Аналогия. При извода по аналогия на базата на знания за сходство между два обекта по някои признаци се генерира хипотезата, че тези обекти са сходни и по други признаци, които са установени в единия обект, но все още не са установени в другия. В този смисъл при извода по аналогия се извършва пренасяне (трансформиране) на информация от единия обект към другия.

Представяне на несигурни знания и вероятностни разсъждения

Представяне на несигурни знания с вероятности

- *Случайна променлива.* Величина в езика за представяне на знания, която може да има няколко (вкл. безброй много) възможни стойности.
- *Област на променлива*: *dom(x)* = множеството от възможни стойности на *x*.
- *Твърдение*: булев израз от присвоявания на променливи ($x_i = v_j$). Например: *(време = дъждовно)* $\vee$ *(болест = грип)* $\vee$ $\neg$*(температура = повишена)*.

- *Вероятност* = мярка за увереност в дадено твърдение (реално число между 0 и 1). $P(A)=0 \rightarrow$ 100% увереност, че твърдението *A* е лъжа; $P(A)=1 \rightarrow$ 100% увереност, че твърдението *A* е истина.
- *Вероятностното разпределение* задава вероятността на всяка възможна стойност на променливата. Ако $dom(x) = \{v_1, v_2, \dots, v_n\}$, то $\sum_{i=1}^{n} P(x = v_i) = 1$.
- *Априорна вероятност* – вероятност при отсъствие на каквато и да е информация.
- *Условна вероятност* – вероятност при наличието на информация за стойностите на други случайни променливи. Например: *P(температура = повишена | болест = грип)*.

- *Основни зависимости* (*A*, *B* – твърдения):
  - $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
  - *A* и *B* са *независими* (т.е. знанието на едното не променя вероятността на другото), когато $P(A \wedge B) = P(A)P(B)$
  - *A* и *B* са *несъвместими* (т.е. никога не могат да се случат заедно), когато
    $P(A \wedge B) = 0$
  - Дефиниция на условна вероятност: $P(A|B) = \frac{P(A \wedge B)}{P(B)}$

    Следователно, $P(A \wedge B) = P(A|B)\, P(B)$.

- *Условна независимост* на *A* и *B* при дадено *C*: ако $P(A|B \wedge C) = P(A|C)$ и $P(B|A \wedge C) = P(B|C)$
- Формула (теорема) на Бейс: $P(A|B) = \dfrac{P(B\,|\,A)P(A)}{P(B)}$

- Вероятностен модел на предметната област:
  - *Атомарно събитие*: $(x_1 = v_1) \wedge (x_2 = v_2) \wedge \ldots \wedge (x_n = v_n)$, където x_i са случайни променливи. Описва конкретно състояние на предметната област.
  - *Съвместно разпределение*: n-мерна таблица с m_i ($i = 1,2, \ldots , n$) клетки по всяка размерност (ако x_i има m_i възможни стойности). Във всяка клетка се записва вероятността на съответното атомарно събитие. Тъй като атомарните събития са несъвместими (т.е. взаимно изключващи се) и таблицата съдържа всички атомарни събития, то сумата от стойностите на всички клетки е 1.

Механизми за извод

- Използване на съвместното разпределение
  Дадено е съвместното разпределение на няколко случайни променливи, например

	зъбобол = да	*зъбобол = не*
кариес = да	0.04	0.06
кариес = не	0.01	0.89

Тогава могат да се изчисляват вероятностите на произволни твърдения. Например (за краткост са пропуснати стойностите на променливите):

$P(кариес) = 0.04 + 0.06 = 0.1$ (сумата на реда)

$P(кариес \vee зъбобол) = 0.04 + 0.06 + 0.01 = 0.11$

$P(кариес | зъбобол) = P(кариес \wedge зъбобол)/P(зъбобол) = 0.04/(0.04+0.01) = 0.8$

- Използване на формулата на Бейс
  - Дадени са: e – множество от симптоми ($e = e_1 \wedge e_2 \wedge \ldots \wedge e_k$) и $d_1, d_2, \ldots , d_n$ – изчерпващо множество от диагнози. Предполага се, че елементарните симптоми $\{e_i\}$ са независими. Известни са $P(d_i)$ и $P(e|d_i)$ за $i = 1, \ldots , n$ (по-точно, $P(e_j|d_i)$ за $j = 1, \ldots , k$ и $i = 1, \ldots , n$).
  - Задачата е да се пресметнат $P(d_i|e)$, $i = 1, \ldots , n$ и да се намери най-вероятната диагноза при дадените симптоми e.
  - Според формулата на Бейс

    $P(d_i \mid e) = \frac{P(d_i)P(e|d_i)}{P(e)}$ за всяко $i = 1, \ldots , n$

- Предполага се, че елементарните симптоми $\{e_i\}$ са независими, следователно

$$P(e \mid d_i) = \prod_{j=1}^{k} P(e_j \mid d_i) \quad \text{за всяко } i = 1, \ldots, n$$

- *P(e)* може да се намери по следния начин:

$$\sum_{i=1}^{n} P(d_i \mid e) = \sum_{i=1}^{n} \frac{P(d_i)P(e|d_i)}{P(e)} = 1, \text{ следователно}$$

$$P(e) = \sum_{i=1}^{n} P(d_i)P(e \mid d_i)$$

- Пример:

вероятност	*здрав*	*грип*	*алергия*
P(d)	0.9	0.05	0.05
P(кихане|d)	0.1	0.9	0.9
P(кашлица|d)	0.1	0.8	0.7
P(температура|d)	0.01	0.7	0.4

Нека симптомите *e* са кихане и кашлица без повишена температура.

Тогава

$e = e_1 \wedge e_2 \wedge e_3$,

e_1 = *кихане*, e_2 = *кашлица*, $e_3 = \neg$(*повишена температура*)

d_1 = *здрав*, d_2 = *грип*, d_3 = *алергия*

$$P(здрав|e) = \frac{P(здрав)P(e\mid здрав)}{P(e)} = \frac{(0.9)P(e\mid здрав)}{P(e)};$$

$$P(e|здрав) = \prod_{j=1}^{3} P(e_j \mid здрав) = (0.1)(0.1)(1\text{-}0.01)$$

Следователно,

$$P(\textit{здрав}|e) = \frac{(0.9)(0.1)(0.1)(0.99)}{P(e)} = \frac{0.0089}{P(e)}$$

$$P(\textit{грип}|e) = \frac{(0.05)(0.9)(0.8)(0.3)}{P(e)} = \frac{0.01}{P(e)}$$

$$P(\textit{алергия}|e) = \frac{(0.05)(0.9)(0.7)(0.6)}{P(e)} = \frac{0.019}{P(e)}$$

$$P(e) = \sum_{i=1}^{3} P(d_i)P(e \mid d_i) = 0.0089 + 0.01 + 0.019 = 0.0379$$

Следователно,

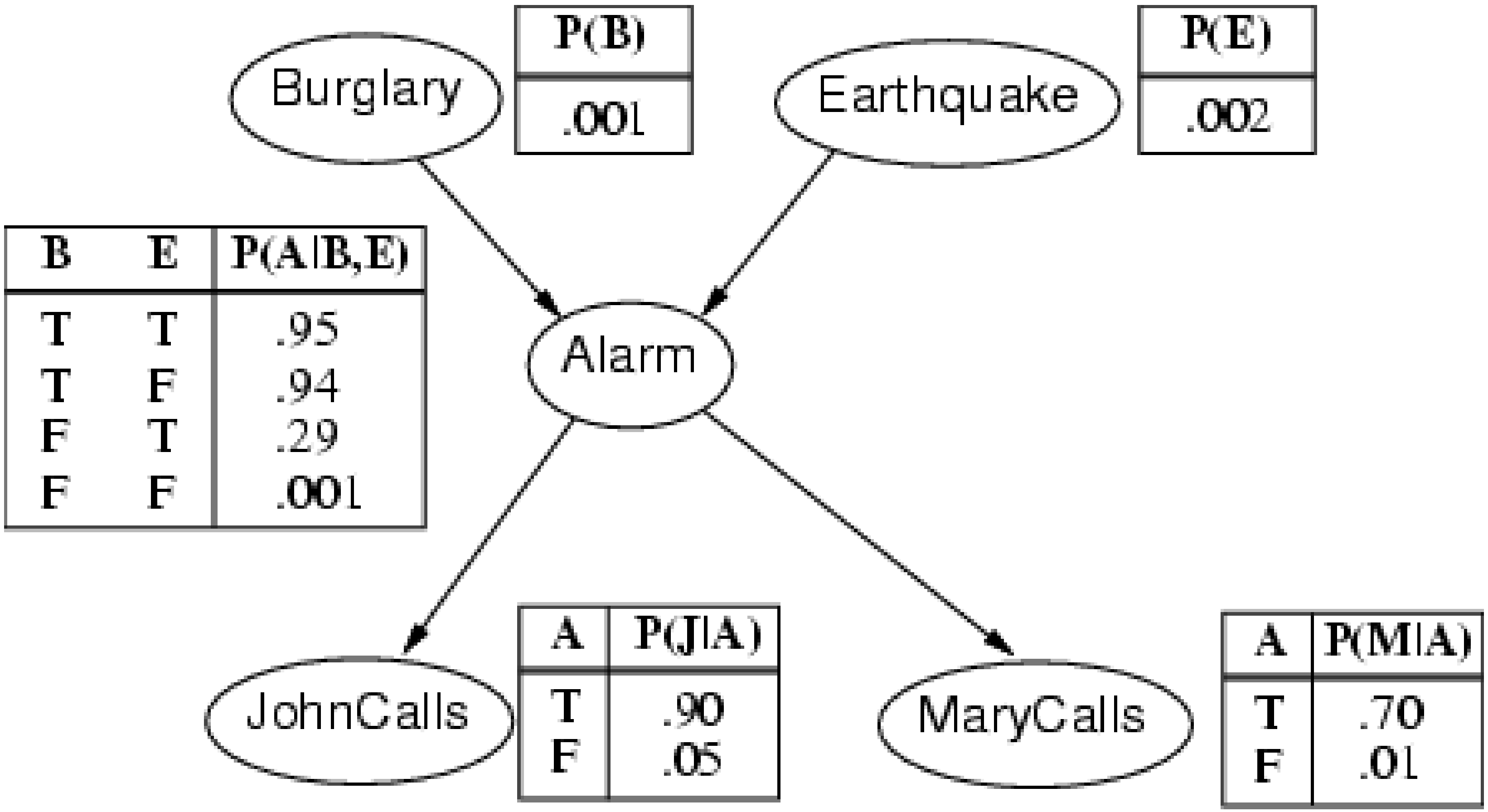
$P(\textit{здрав}|e) = 0.23$; $P(\textit{грип}|e) = 0.26$; $P(\textit{алергия}|e) = 0.50$

- Проблем: предположението за независимост на елементарните симптоми е прекалено силно и нереалистично.

Бейсови мрежи (БМ; Belief Networks)

- Използване на ацикличен ориентиран граф за представяне на зависимостите между променливите с цел сбито (компактно) описание на съвместното им разпределение.
- На всяка случайна променлива съответства отделен възел от мрежата. Дъгите от мрежата задават *причинно-следствени връзки*. Интуитивното значение на дъгата от възела X към възела Y е, че X оказва *директно влияние* върху Y.
- За всеки възел е дефинирана таблица с условни вероятности, която задава вероятността на всяка стойност на променливата във възела в зависимост от всяка възможна комбинация от стойности на променливите в родителските възли.

Пример:



P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Примерна предметна област. В жилището си имате монтирана нова сигнална инсталация (аларма). Тя е чувствителна и реагира на опит за проникване в жилището ви (в частност, при опит за обир), но също и на (дори слаби) земетресения. Имате също двама съседи, Джон и Мери, които са обещали да ви се обаждат по телефона в службата винаги когато чуят, че алармата във вашето жилище се е включила. Джон винаги ви се обажда, когато чуе алармата, но понякога я обърква със звъна на телефона и тогава също ви се обажда. Мери пък обича да слуша силна музика и понякога е възможно да не чуе алармата у вас.

Ако е известно кой от двамата ви се е обадил или не се е обадил, може да се установи например вероятността в жилището ви да е извършен обир.

- БМ задават неявно съвместното разпределение на променливите си. Нека $x_1, x_2, \ldots , x_n$ са случайни променливи и $P(v_1,v_2, \ldots ,v_n)$ е съвместната вероятност те да получат съответно стойности $v_1, v_2, \ldots , v_n$. Тогава

$$P(v_1, v_2, \ldots, v_n) = \prod_{i=1}^{n} P(v_i \mid Parents(x_i)),$$

където $P(v_i \mid Parents(x_i))$ е условната вероятност за $x_i = v_i$ при условие, че са дадени стойностите на родителските променливи $Parents(x_i)$ на x_i.

Например:

$P(JohnCalls, MaryCalls, Alarm, \neg Burglary, \neg Earthqake) =$

$P(J|A).P(M|A).P(A|\neg B \wedge \neg E).P(\neg B).P(\neg E) =$

$(0.9)(0.7)(0.001)(0.999)(0.998) = 0.000628$

- Видове извод в БМ. При дадени стойности на подмножество от променливите (наблюдаеми, evidence variables) да се определи вероятността на стойностите на друго подмножество от променливите (търсени, query променливи).
  - Диагностика – от следствието към причината: $P(Burglary \mid JohnCalls) = ?$
  - Предсказване – от причината към следствието: $P(JohnCalls \mid Burglary) = ?$
  - Междупричинен извод – между причините за дадено следствие: $P(Burglary \mid Earthquake) = ?$
  - Смесен извод – комбинация на горните три: $P(Alarm \mid JohnCalls \wedge \neg Earthquake) = ?$