Chapter 9

Представяне на несигурни знания и вероятностни разсъждения

9.1 Лекции

9.1.1 Необходимост от разсъждения с несигурни знания

- Човек обикновено разсъждава в термините на несигурни предположения и твърдения (не 100% истина).
- Класическата логика изисква 100% истинна.
- Директното прилагане на теорията на вероятностите изисква наличието на прекалено много данни (условни вероятности).
- Правилата с фактори за сигурност са прекалено опростен модел.

9.1.2 Представяне на несигурни знания с вероятности

- *Случайна променлива.* Терм в езика за представяне на знания, който може да има няколко (вкл. безкрайно 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* (число между 0 и 1) = мярка за увереност в дадено твърдение: $P(A) = 0 \rightarrow 100\%$ увереност, че твърдението A е лъжа. $P(A) = 1 \rightarrow 100\%$ увереност, че твърдението A е истина.
- *Вероятност = мярка за незнание*: Ако твърдението има вероятност между 0 и 1, това не значи, че то е истина "в известна степен", а по-скоро, че ние не знаем със сигурност неговата истинностна стойност.
- *Вероятностното разпределение* задава вероятността на всяка възможна стойност на променливата. Например ако $dom(x) = \{v_1, ..., v_n\}$, 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: $P(\textit{температура} = \textit{повишена} \mid \textit{болест} = \textit{грип})$.
- *Основни формули* (A, B – твърдения):
  - $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
  - A, B – *независими* (знанието на едното не променя вероятността на другото), $P(A \wedge B) = P(A)P(B)$
  - A, B – *несъвместими* (никога не се случват заедно), $P(A \wedge B) = 0$.
  - Дефиниция на условна вероятност: $P(A|B) = \frac{P(A \wedge B)}{P(B)}$
  - Бейс: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

- Вероятностен модел на проблемната област:
  - *Атомарно събитие* – $(x_1 = v_1) \wedge (x_2 = v_2) \wedge ... \wedge (x_n = v_n)$, където $x_1, ..., x_n$ са случайни променливи. Описва конктретно състояние на проблемната област.
  - *Съвместно разпределение* – n-мерна таблица с m_i клетки $(i = 1, ..., n)$ 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* (взаимно изключващи се) и таблицата съдържа *всички* атомарни събития, сумата от стойностите на всички клетки е 1. Например:

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

  - При дадено съвместно разпределение могат да се изчисляват вероятностите на произволни твърдения. Например (за краткост стойностите на променливите са пропуснати):
    $P(\textit{кариес}) = 0.04 + 0.06 = 0.1$ *(сума на реда)*
    $P(\textit{кариес} \vee \textit{зъбобол}) = 0.04 + 0.01 + 0.06 = 0.11$
    $P(\textit{кариес} \mid \textit{зъбобол}) = P(\textit{кариес} \wedge \textit{зъбобол}) \ / \ P(\textit{зъбобол}) = 0.04/(0.04 + 0.01) = 0.8$

9.1.3 Механизми за извод

- Използуване на съвместно разпределение.
- Използуване на формулата на Бейс.
  - Дадени са e – множество от симптоми и $d_1, ..., d_n$ – диагнози. Известни са $P(d_i)$ и $P(e|d_i)$ за $i = 1, ..., n$.
  - Изчисляваме $P(d_i|e)$ за всяко i и намираме най-вероятната диагноза при дадените симтоми e.

$$P(d_i|e) = \frac{P(d_i)P(e|d_i)}{P(e)}$$

  - $P(e)$ може да се изчисли (ако е необходимо) от формулата:

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

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

  - Изчисляването на $P(e|d_i)$ изисква изброяване на всички възможни комбинации от стойности на елементарни симтоми (участвуващи в множеството e), които причиняват d_i (*експоненциална сложност*).
- Разсъждения при предположение за независимост на симтомите ("naive" Bayes).
  - Ако приемем, че елементарните симтоми, съставящи e, $e = \{e_1, ..., e_k\}$, са *независими*, то

$$P(e|d_i) = \prod_{j=1}^{k} P(e_j|d_i)$$

  - e_j $(j = 1, ..., k)$ могат да се определят лесно от експерт или да се извлекат от експериментални данни.
  - Пример:

Вероятност	*здрав*	*грип*	*алергия*
$P(d)$	0.9	0.05	0.05
$P(\textit{кихане}|d)$	0.1	0.9	0.9
$P(\textit{кашлица}|d)$	0.1	0.8	0.7
$P(\textit{температура}|d)$	0.01	0.7	0.4

    Ако симтомите e са кихане и кашлица без повишена температура, то:
    $P(\textit{здрав}|e) = (0.9)(0.1)(0.1)(0.99)/P(e) = 0.0089/P(e)$
    $P(\textit{грип}|e) = (0.05)(0.9)(0.8)(0.3)/P(e) = 0.01/P(e)$
    $P(\textbf{алергия}|e) = (0.05)(0.9)(0.7)(0.6)/P(e) = 0.019/P(e)$
    $P(e) = 0.0089 + 0.01 + 0.019 = 0.0379$
    $P(\textit{здрав}|e) = 0.23$
    $P(\textit{грип}|e) = 0.26$
    $P(\textbf{алергия}|e) = 0.50$

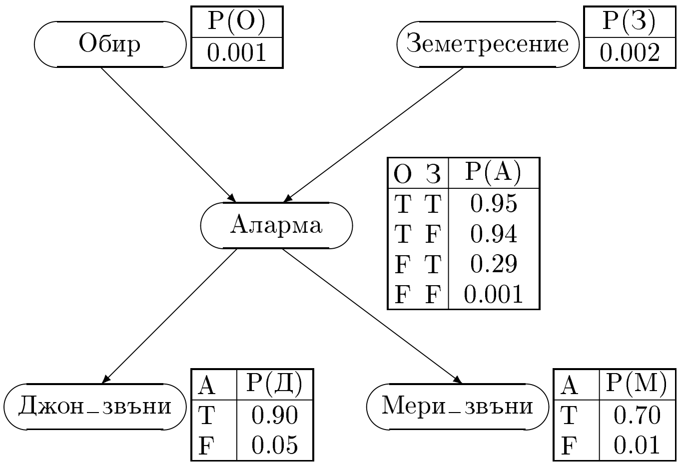

Figure 9.1: Бейсова мрежа

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

- Бейсови мрежи – използуване на *насочен ацикличен граф* за представяне на зависимостите между променливите с цел сбито описание на съвместното им разпределение.

9.1.4 Откъде се появават вероятностите?

- Честотен подход – експерименти.

- Обективен подход – вероятностите са проява на съществуващи свойства на реалния свят.

- Субективен подход – описание на степента на увереност на агента (знанието му за дадено обективно съществуващо явление).

- Избор на клас за сравнение (reference class).

- Пример: каква е вероятността слънцето за изгрее утре? (неопределена, 1, $1-\epsilon$, $\frac{d+1}{d+2}$).

9.1.5 Бейсови мрежи (БМ)

- За всяка случайна променлива – възел.

- Връзките задават директни *причинно-следствени връзки.*

- За всеки възел е дефинирана *таблица с условни вероятности*, която задава вероятността на всяка стойност на променливата във възела в зависимост от всяка една комбинация от стойности на променливите в родителските възли (експоненциален брой редове на таблиците). Пример: фигура 9.1.

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

$$P(v_1, ..., v_n) = \prod_{i=1}^{n} P(v_i | Parents(x_i)),$$

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

  $P(\textit{Джон\_звъни, Мери\_звъни, Аларма}, \neg \textit{Обир}, \neg \textit{Земетресение}) = (0.9)(0.7)(0.001)(0.999)(0.998) = 0.000628$

- Условна независимост (използува се при извода). Ако при премахване на подмножество от възлите на мрежата S възлите x и y станат несвързани, то казваме, че x и y са *независими относно* S, т.е. $P(x|y,S) = P(x,S)$.

- Типове връзки:
  - $x \leftarrow y \rightarrow z$: x и z са независими относно y.
  - $x \leftarrow y \leftarrow z$ или $x \rightarrow y \rightarrow z$: x и z са независими относно y.
  - $x \rightarrow y \leftarrow z$: x и z са независими относно y без наличието на информация за стойността на y. При дедена стойност на y (или негови наследници) те стават зависими. Пример: *Обир* $\rightarrow$ *Аларма* $\leftarrow$ *Земетресение*. Ако вече има аларма и се е случило земетресение, вероятността за обир е по-малка.
- Видове извод в БМ. При дадена стойност на подмножество от променливите (наблюдаеми, evidence variables) да се определи вероятността на стойностите на друго подмножество променливи (търсени, query variables).
  - Диагностика – от следствието към причината: $P(\mathit{Обир}|\mathit{Джон\_звъни})$?
  - Предсказване – от причината към следствието: $P(\mathit{Джон\_звъни}|\mathit{Обир})$?
  - Междупричинен извод – между причините на дадено следствие: $P(\mathit{Обир}|\mathit{Земетресение})$?
  - Смесен – комбинация на горните три: $P(\mathit{Аларма}|\mathit{Джон\_звъни} \wedge \neg \mathit{Земетресение})$?
- Алгоритми за извод в БМ.
  - Изчисляване на пълното съвместно разпределение.
  - Разпространение на увереността – изчисляване на вероятността на всички променливи в съответствие с наблюденията. Итеративна модификиция на вероятностните вектори за всяка променлива чрез предаване на съобщения между възлите: "наследници→родители" и "родители→наследници", започвайки от наблюденията.
  - Алгоритми, водени от целевите (query) променливи – не изчисляват всички променливи.
  - В общия случай задачата е *експоненциална* по броя на променливите.
  - За едносвързани мрежи (без ненасочени цикли) съществуват *полиномиални* алгоритми.

9.2 Семинарни упражнения

- Изчисляване на априорни и апостериорни вероятности чрез честотен анализ на данни от вида: `example(Id,Diagnosis,[S1=V1,S2=V2,...,Sn=Vn])` (файл tumor.pl).
- Реализация на алгоритъм за класификация (диагностика) "naive Bayes" за работа с данни от горния вид.
- Извод в БМ чрез изчисляване на пълното съвместно разпределение (БМ от фигура 9.1).
- Създаване на БМ за диагностика на автомобили чрез използуване на базата от правила с фактори за сигурност (cars.pl).

9.3 Лабораторни упражнения

- Програмиране на предикати за изчисляване на априорни и апостериорни вероятности чрез честотен анализ на данни (файл tumor.pl).
- Програмиране на алгоритъм за класификация (диагностика) "naive Bayes" за работа с данни от горния вид.
- Експерименти с примерната БМ (bnet_dat.pl) интерпретатора на БМ (bnet.pl).
- Програмиране на БМ от фигура 9.1 и изпълнение с интерпретатора на БМ (bnet.pl).
- Програмиране на БМ за диагностика на автомобили и изпълнене с интерпретатора на БМ (bnet.pl).