

ФАКУЛТЕТ ЗА ИНФОРМАТИЧКИ НАУКИ И КОМПЈУТЕРСКО ИНЖЕНЕРСТВО

Веројатносно расудување

Од каде потекнува несигурноста?

- Логичките системи со кои се претставуваат проблемите од реалниот живот не се комплетни.
- Но, агентите мора да дејствуваат и во вакви услови.
- Расудувањето во услови на некомплетност, неконзистентност и немонотоност се вика расудување во услови на несигурност.

Причини за несигурноста според Charles R. Dyer

- Несигурност која се должи на случајноста (на пр. фрлање паричка или коцка): вистинска несигурност
- Отсуство на знаење кое се должи на некомплетноста на системот (на пр. медицинска дијагноза): теоретско незнаење
- Несигурност која се должи на преобемноста на релевантни фактори, или на неможноста да се опфатат сите меѓурекултати: мрзливост
- Несигурност која се должи на недоволно прибрани информации во врска со доменот: практично незнаење

Трите принципа на несигурноста

1. Принцип на максимизација на несигурноста (Principle of Uncertainty Maximization)
2. Принцип на минимизација на несигурноста (Principle of Uncertainty Minimization)
3. Принцип на непроменливост на несигурноста (Principle of Uncertainty Invariance)

Максимизација на несигурноста (1)

- Основна идеја: при индуктивното расудување да се искористат сите расположливи информации, но не повеќе од тоа, така што ќе се обезбеди да не се појави ниту една дополнителна информација што е несакајќи додадена.
- Применлив е кај системите во кои е вклучено амплијативното расудување.
- Амплијативно е тоа расудување кај кое претпоставките не ги повлекуваат со себе заклучоците.

Максимизација на несигурноста (2)

- Според овој принцип, заклучоците при амплијативното заклучување ја максимизираат релевантната несигурност во условите што ги прикажуваат претпоставките.
- Принципот гарантира дека заклучоците максимално ќе одбегнуваат директен одговор во однос на информациите што не се содржани во претпоставките.
- Кај стохастичките системи, принципот на максимизација на несигурноста се поистоветува со принципот за максимална ентропија.

Минимизација на несигурноста

- Основна идеја: при дедуктивното расудување, да се изгубат што е можно помалку информации.
- Упростувањето се врши со примена на три стратегии:
 - ☐ елиминирање на некои ентитети од системот
 - ☐ обединување на некои ентитети од системот
 - ☐ разбивање на целиот систем во соодветни потсистеми
- Применлива е и во областа на разрешување конфликтни проблеми.

Непроменливост на несигурноста

- Кога еден систем се трансформира, тогаш количеството информации во системот што се добива треба да бидат што е можно поблиски до оригиналот.
- При преминот:
 - количеството на несигурност сврзано со ситуацијата да се запази при преминот од теоријата $T1$ во $T2$
 - степенот на доверба во $T1$ да се распредели на составните делови од $T2$ со соодветна скала
- Принципот на непроменливост на несигурноста се вика и принцип за зачувување на информацијата.

Практични решенија за несигурноста

- Чисто логичко расудување со игнорирање на несигурноста
- Веројатносно расудување
- Расудување според дадена вредност (Default reasoning) кое се базира врз верувањето во нешто, додека не се појави нешто во кое може повеќе да се верува
- Правила со фактор на несигурност
- Ракување со незнаењето (D-S теорија)
- Неодреденост (непрецизирана логика, матна логика – fuzzy logic)

Врската меѓу логиката и веројатноста

- Веројатноста и логиката користат ист онтолошки посветен (commitment) пристап
- Секој исказ се третира како настан на кој му е доделен степен на верувањето (degree of belief)
- Од друга страна, на настаните им се доделува вредност на вистинитоста

Како се дефинира веројатноста на исказите

- Нека е даден еден исказ A
- Нему му се доделува веројатноста $P(A)$
- Притоа, и тука важи условот дека
$$0 \leq P(A) \leq 1$$
- Ако исказот A е вистинит, тогаш $P(A)=1$
- Ако исказот е лажен, тогаш $P(A)=0$
- Исказот мора секогаш да биде или вистинит или лажен, но $P(A)$ го претставува нашиот степен на верување дека е вистинит или лажен.

Пример:

- Ако кажете дека утре времето ќе биде променливо облачно, во тој случај во метеоролошката прогноза ќе стои и процентуалната можност за дожд (на пр. 60%)
- Тоа значи дека веруваме дека можноста да заврне има веројатност 0,6, па затоа:
 $P(\text{време} = \text{дожд}) = 0,6$
- Ако вредностите во доменот се:
сонце, дожд, снег, облаци,
тогаш важат следниве аксиоми на веројатноста:

Основните аксиоми на веројатноста за примерот:

- $0 \leq P(\text{сонце}) \leq 1$
- $0 \leq P(\text{дожд}) \leq 1$
- $0 \leq P(\text{снег}) \leq 1$
- $0 \leq P(\text{облаци}) \leq 1$

$$P(\text{сонце}) + P(\text{дожд}) + P(\text{снег}) + P(\text{облаци}) \\ = 1$$

Ова равенство произлегува од условот дека вредностите во доменот се исцрпни и заемно исклучиви.

Мало потсетување

- Кога се случуваат настаните:
 - $A \cup B$
 - $A \cap B$
- Кој настан се означува со \bar{A} ?
- А што значат B/A и A/B ?
- Настаните при расудувањето се логички искази, па тогаш операторите над настаните преминуваат во логички оператори.

Премин кон расудувањето

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cap B) = P(A) P(B/A)$$

$$P(A \cap B) = P(B) P(A/B)$$

$$P(\bar{A}) = 1 - P(A)$$

т.е.

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

$$P(A \wedge B) = P(A) P(B/A)$$

$$P(A \wedge B) = P(B) P(A/B)$$

$$P(\neg A) = 1 - P(A)$$

Донесување одлука во услови на несигурност

Пример од Расел и Норвиг:

- Треба да стигнете на аеродром.

Претпоставете дека верувате во следново:

$$P(A_{25} \text{ ќе ве донесе на време} \mid \dots) = 0.04$$

$$P(A_{90} \text{ ќе ве донесе на време} \mid \dots) = 0.70$$

$$P(A_{120} \text{ ќе ве донесе на време} \mid \dots) = 0.95$$

$$P(A_{1440} \text{ ќе ве донесе на време} \mid \dots) = 0.9999$$

при што A_t значи дека патникот ќе тргне кон аеродромот t минути пред полетувањето.

- Која активност ќе ја преземете, односно кое време на тргнување ќе се одбере?

Несигурност - Uncertainty

Нека е акцијата A_t = да се тргне кон аеродромот t минути пред летот

Дали A_t ќе ме доведе на аеродромот на време?

Проблеми:

1. делумна можност за набљудување (состојба на патот, плановите на другите возачи итн)
2. шум во сензорите (сообраќајни извештаи)
3. несигурност во резултатите од акциите (дупната гума итн)
4. комплексност својствена за моделирањето и предвидувањето на сообраќајот

оттука чистиот логички приод (a purely logical approach)
или

1. ризикува неточност (falsehood): “ A_{25} ќе ме доведе таму на време”,

или

1. води до заклучоци кои се премногу слаби за донесување на одлуки (decision making):

“ A_{25} ќе ме доведе таму на време ако нема сообраќајка на мостот и ако не врне и ако не се дупне некоја гума ИТН ИТН”

(A_{1440} изгледа разумно дека ќе ме доведе таму на време, ама ќе треба да преспијам на аеродромот ...)

Методи за справување со несигурностите

- Предодредена (Default) или немонотона (nonmonotonic) логика:
 - Претпоставка дека колата нема дупната гума
 - Претпоставка A_{25} врши работа освен ако не е противречно со новите сознанија
- Прашања: Кои претпоставки се разумни? Како да ги наброиме сите претпоставки (што сè може да влијае врз исходот)? Како да се справиме со противречностите?
- Веројатност (Probability)
 - Се моделира степенот на верување (belief) кај агентот
 - Според достапните сознанија (сведоштва, докази) (evidence),
 - A_{25} ќе ме доведе на време на аеродромот, со веројатност 0,04

Веројатност (Probability)

Веројатносните тврдења ги сумираат ефектите на

- мрзливост: неуспех во набројувањето на исклучоците, квалификациите итн.
- незнаење (ignorance): недостиг на релевантни факти, почетни услови итн.

Субјективна веројатност:

- Веројатностите ги поврзуваат исказите за сопствената состојба на знаењето на агентот

$$\text{пр., } P(A_{25} \mid \text{без извештаи за сообраќајки}) = 0,06$$

- Веројатностите на исказите се менуваат со нови сознанија (сведоштва, докази):

$$\text{пр., } P(A_{25} \mid \text{без извештаи за сообраќајки, 5 наутро}) = 0,15$$

Одговорот на прашањето „Кога да тргнете кон аеродромот?“

- Зависи ...

од она што се претпочита: да се пропушти авионот или да се чека предолго.

- Теоријата на корисност (Utility theory) се користи за да го претстави и заклучи она што се претпочита (represent and infer preferences)
- Теорија на одлука (Decision theory) = теорија на веројатност + теорија на корисност

Заедничка распределба на веројатноста

- Ако за еден домен сме нашле доволно случајни променливи со кои доменот добро се опишува, тогаш се дефинира таканаречената целосна заедничка распределба на веројатноста (full joint probability distribution FJPD).
- Со избраните случајни променливи се специфицираат целосно сите информации, па затоа FJPD доделува вредности на сите можни комбинации на случајните променливи.
- $P(V_1 = v_1, V_2 = v_2, \dots, V_n = v_n)$

Атомни настани

- Еден настан е атомен (atomic event) ако дава целосна спецификација на светот за кој агентот е несигурен.
- Атомните настани се заемно исклучителни и исцрпно го опишуваат светот.
- На пример, две булови променливи дефинираат четири заеднички атомни настани.
- Тие се поврзани со конјункција.

Пример за FJPD

- Светот на Коко може да се претстави преку случајните променливи: птица, лета и мала.

птица	лета	мала	веројатност
T	T	T	0,19
T	T	F	0,25
T	F	T	0,01
T	F	F	0,02
F	T	T	0,01
F	T	F	0,01
F	F	T	0,32
F	F	F	0,19

Што дознаваме од FJPD?

- $P(\text{птица} = T) = 0,19 + 0,25 + 0,01 + 0,02 = 0,47$
- $P(\text{птица} = T, \text{лета} = T) =$
 $P(\text{птица}, \text{лета}, \text{мала}) +$
 $P(\text{птица}, \text{лета}, \neg \text{мала}) = 0,19 + 0,25 = 0,44$
- Според ова, одговорот на прашањето дали Коко лета би било: Веројатноста дека Коко лета е 0,44.
- Со тоа, импликацијата станува матна (fuzzy) импликација, бидејќи не е сосема сигурна.
- Колкава е веројатноста дека Коко е мал?
- А колкава е веројатноста дека не е птица?

Условната веројатност

- Условната веројатност е клучна при расудувањето, затоа што го формализира прибирањето докази и обновувањето на веројатностите.
- Пример: Колкава е веројатноста да врне во Скопје?
 - статистички е покажано дека $P(\text{дожд}) = 0,21$
 - оваа веројатност се вика “а приори” веројатност или безусловна веројатност
 - ако времето е веќе облачно, тогаш веројатноста за дожд се зголемува и се оценува со $P(\text{дожд} / \text{облаци})$
 - Оваа веројатност се вика “а постериори” веројатност или веројатност која се базира на некој нов доказ (evidence).

Условната веројатност во претходниот пример

$$\begin{aligned}
 1. \quad & P(\neg \text{птица} / \text{лета}) = \\
 & P(\neg \text{птица}, \text{лета}) / P(\text{лета}) = \\
 & (P(\neg \text{птица}, \text{лета}, \text{мала}) + P(\neg \text{птица}, \text{лета}, \\
 & \neg \text{мала})) / P(\text{лета}) = \\
 & (0,01 + 0,01) / 0,46 = 0,0435
 \end{aligned}$$

$$P(A \wedge B) = P(A) P(B/A)$$

=>

$$P(B/A) = P(A \wedge B) / P(A)$$

пти ца	лет а	мал а	веројатност
T	T	T	0,19
T	T	F	0,25
T	F	T	0,01
T	F	F	0,02
F	T	T	0,01
F	T	F	0,01
F	F	T	0,32
F	F	F	0,19

Условната веројатност во претходниот пример

1. $P(\neg \text{птица} / \text{лета}) =$
 $P(\neg \text{птица}, \text{лета}) / P(\text{лета}) =$
 $(P(\neg \text{птица}, \text{лета}, \text{мала}) + P(\neg \text{птица}, \text{лета}, \neg \text{мала})) / P(\text{лета}) =$
 $(0,01 + 0,01) / 0,46 = 0,0435$
2. Како е пресметана маргиналната веројатност $P(\text{лета})$?
3. Дали важи $P(\text{птица} / \text{лета}) = 1 - P(\neg \text{птица} / \text{лета})$?

Условна независност

- Една променлива V е условно независна од множество променливи V_i за зададено множество променливи V_j ако важи дека

$$P(V/V_i, V_j) = P(V/V_j)$$

- Оваа состојба се означува со $I(V, V_i/V_j)$
- Една променлива V_i е независна од друга променлива V_j во множеството V ако важи дека

$$P(V_i, V_j / V) = P(V_j / V) * P(V_i / V)$$

- Променливите V_1, \dots, V_k се заемно условно независни ако

$$P(V_1, V_2, \dots, V_k / V) = \prod P(V_i / V_{i-1}, \dots, V_1, V)$$

- Ако V е празно, тогаш важи дека:

$$P(V_1, V_2, \dots, V_k) = P(V_1) * P(V_2) * \dots * P(V_k)$$

Комбинирање повеќе докази користејќи JPD

- За одредување на веројатноста за истовремено исполнување на n услови се користи правилото на верига (Chain Rule):

$$P(A, B, C, D) = P(A/B, C, D) * P(B/C, D) * P(C/D) * P(D)$$

- Редоследот на променливите не влијае врз резултатот.

Пример за условна веројатност која се базира на повеќе докази

$$P(\text{птица} / \text{лета, мала}) = ?$$

$$P(\text{птица} / \text{лета, мала}) =$$

$$P(\text{птица, лета, мала}) / (P(\text{лета} / \text{мала}) * P(\text{мала})) =$$

$$P(\text{птица, лета, мала}) / (P(\text{лета, мала})) =$$

$$P(\text{птица, лета, мала}) / (P(\text{птица, лета, мала}) + P(\neg \text{птица, лета, мала})) =$$

$$0,19/(0,19+0,01)=0,95$$

Резиме

- Веројатноста е формализам кој одлично ги решава проблемите на несигурното знаење
- Заедничката дистрибуција на веројатностите ја одредува веројатноста на атомните настани.
- Одговорите на прашањата се сведуваат на сумирање на веројатностите на атомните настани.
- Независноста на настаните го намалува напорот за одредување на заедничката веројатност.
- Но, овој услов ретко кога може да биде исполнет.

Како да се олесни пресметувањето?

- JPD е тешко да се определи, посебно кога бројот на настаните расте.
- Во голем број случаи, не може да се определи условната веројатност.
- Алтернативно решение се мрежите на Bayes.
- Тие претставуваат податочна структура преку која се одредува зависноста меѓу променливите и дава спецификација на нивното врзување.
- Мрежите на Bayes се викаат и мрежи на верувањето (Belief Networks)

Причинско-последичен (каузален) поглед



- Непозната хипотеза H за која сакаме да направиме порценка за неизвесноста и да донесеме некоја одлука.
- Почнуваме со некое независно уверување за H (на пр., еден од секои 1000 луѓе ја има болеста).
- Ова наше верување / знаење се менува во зависност од доказниот материјал (се зголемува или намалува во зависност од веројатноста сврзана за тој доказ).

Мотивација*

- Користење на мамографијата како прогноза за дијагнозата
- Независната веројатност значи дека една од 100 тестирани жени има рак
 $P(h) = 0,01$
- Ратата на потврдна дијагноза и лажна позитивна дијагноза се:
 $P(e|h) = 0,8$ и
 $P(e/\text{not } h) = 0,1$

$$\begin{aligned} P(h|e) &= \frac{P(e|h)P(h)}{P(e)} \\ &= \frac{P(e|h)P(h)}{P(e|h)P(h) + P(e|\neg h)P(\neg h)} \\ &= \frac{0.8 \times 0.01}{0.8 \times 0.01 + 0.1 \times 0.99} \\ &= \frac{0.008}{0.008 + 0.099} \\ &= \frac{0.008}{0.107} \\ &\approx 0.075 \end{aligned}$$

Примена на правилото на Bayes*

- Се применува често во ситуации во кои ги знаеме распределбите на зависните веројатности и сакаме да пресметаме некоја независна
- Пример: поставување медицинска дијагноза
- Лекарот знае дека менингитисот (m) предизвикува вкочанетост на вратот (b) во 50% случаи.
- Веројатноста пациентот да има менингитис е $1/50000$, а веројатноста да се има вкочанет врат е $1/20$.

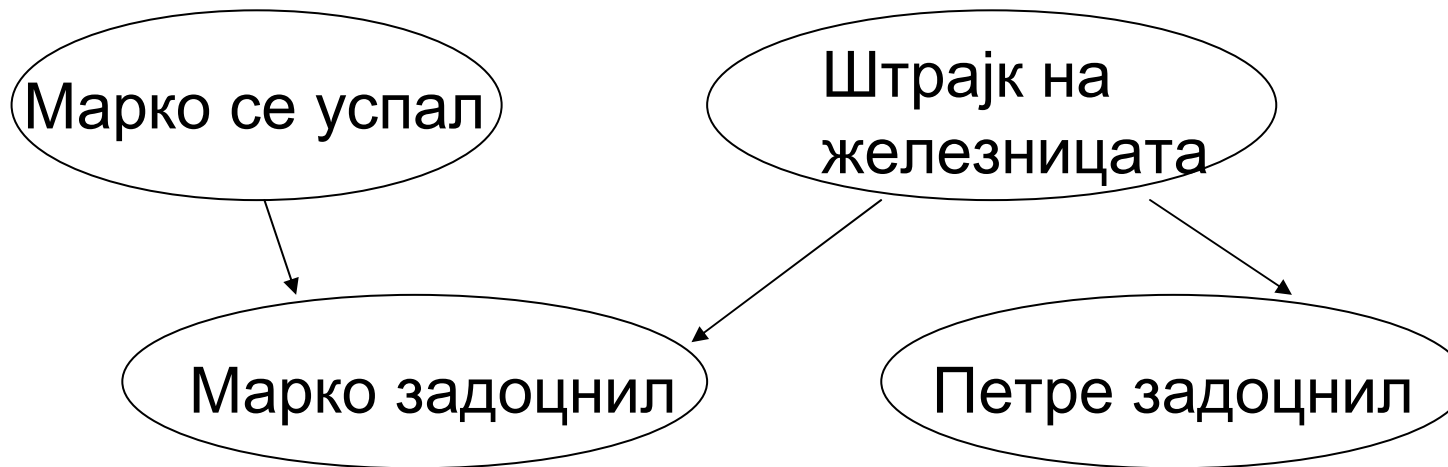
$$P(b/m) = 0,5 \quad P(m) = 1/50000 \quad P(b) = 1/20$$

$$P(m/b) = P(bm) / P(b) = P(b/m) * P(m) / P(b) = 0,0002$$

Здружување на сведоштвото

- Ажурирањето на правилото на Bayes дозволува доказите да се вклучуваат еден по еден (нови информации, нови сведоштва).
- Со секој нов доказ, веројатноста на непознатата променлива се зголемува пропорционално со фактор зависен од новата информација.
- Системите со поголема робусност правилно ги одразуваат промените во околината.
- Независноста меѓу различните променливи ги поедноставува пресметките при користењето на правилото на Bayes (што резултира со помал број зависимости при пресметката).

Здружување на сведоштвото



- Ако постои доказ дека Марко задоцнил, тоа значи дека уверувањето на агентот дека двете можни причини се случиле се зголемила. На пример, веројатноста на штрајк се зголемила од 0,1 на 0,13, а за успивање од 0,4 на 0,41.
- Но, ако имаме сведоштво (доказ) дека и Петре задоцнил, според правилото на Bayes веројатноста за штрајк се зголемува на 0,54, а дека Марко се успал на 0,44.

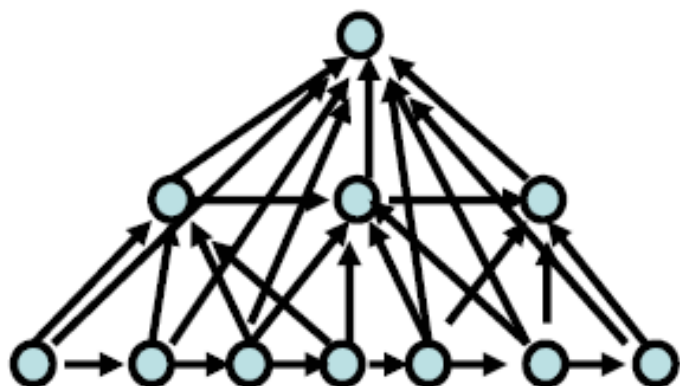
Мрежите на Bayes

- Една мрежа на Bayes е насочен ацикличен граф чии јазли се означени со случајни променливи.
- Таа условува (stipulate) секој јазел V_i во графот да биде независен од останатото подмножество јазли.
- Оваа претпоставка е предуслов за условна независност (conditional independence)
- Ако $A(V_i)$ е било кое множество јазли кои не се наследници на V_i , тогаш нека $P(V_i)$ се сите следни наследници во графот.
- Во тој случај важи дека:

$$P(V_1, V_2, \dots, V_k) = \prod P(V_i / P(V_i))$$

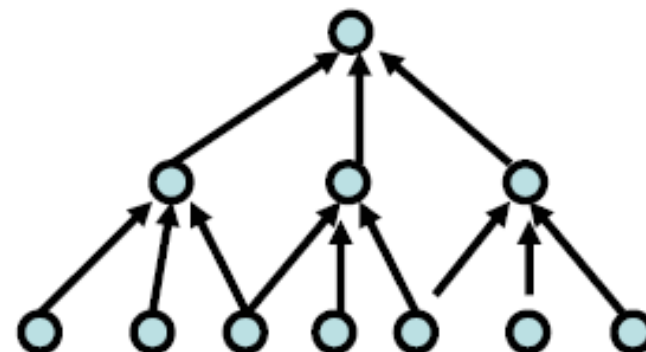
Важноста на независноста*

n binary nodes,
fully connected



$2^n - 1$ independent numbers

n binary nodes
each node max. 3 parents



less than $2^3 \cdot n$
independent numbers

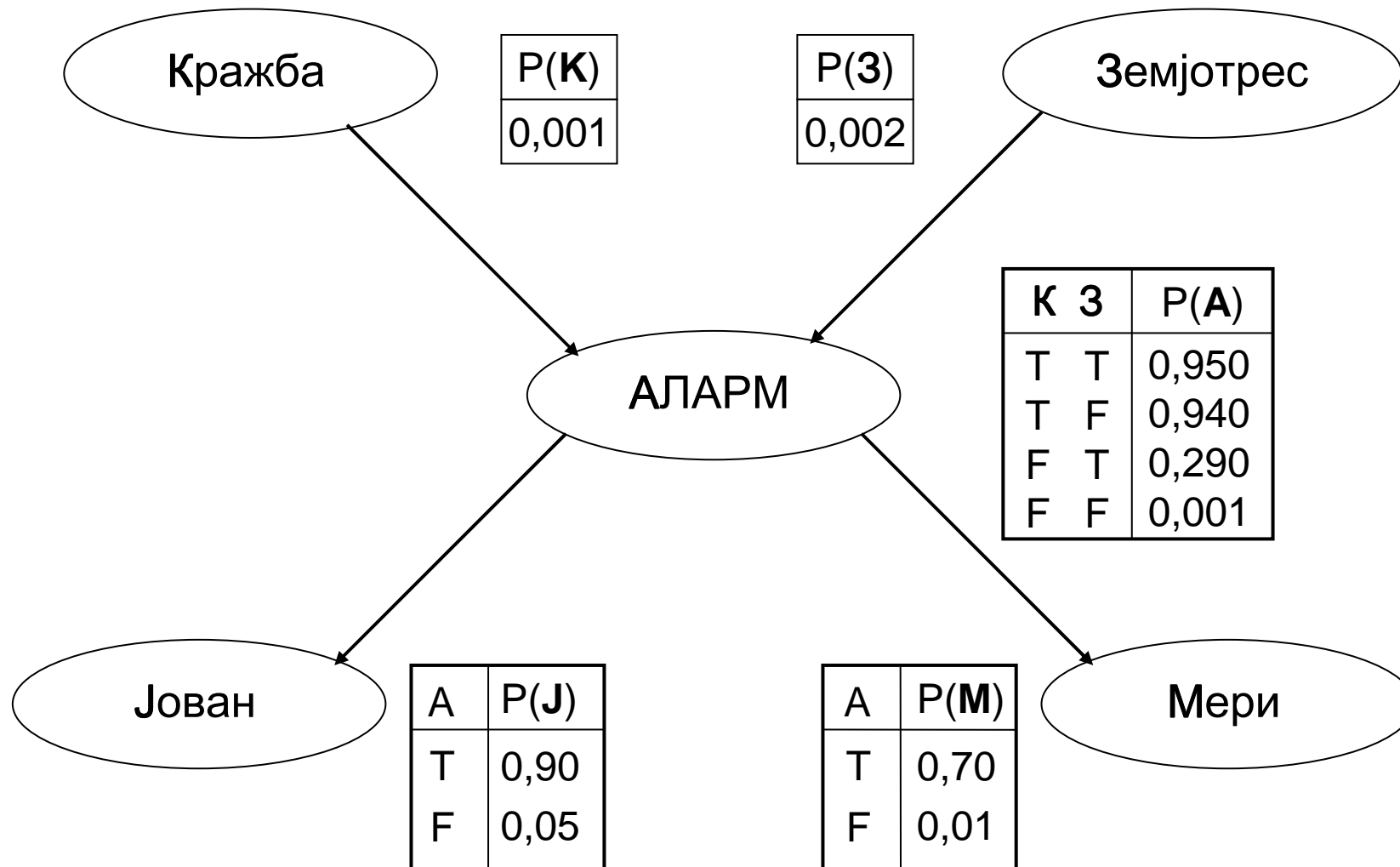
* Слайдот е преземен од <http://www.blutner.de/uncert/index.htm>

Пример 1: Кражба или земјотрес?

- Сте набавиле нов аларм против кражба
- Тој се активира секогаш кога се врши обид за кражба, но, реагира и на мали земјотреси.
- Со соседите Јован и Мери имате договор, секогаш кога ќе го слушнат алармот, да ви се јават.
- Јован се јавува секогаш кога ќе го слушне алармот, но често го меша со телефонското звонење (се јавува и тогаш).
- Мери слуша силна музика, па понекогаш не го дослушува алармот.
- Базирајќи се врз фактот кој се јавил, а кој не, да се процени веројатноста за кражба.

- Примерот е од Judea Pearl, а преземен е од Расел и Норвиг.

Јован се јавува, а Мери не. Дали дошол крадецот?



Веројатноста е ...

- Алтернативниот запис на мрежите на Bayes е:

$$P(x_1, \dots, x_n) = \prod P(x_i / \text{Roditeli}(x_i))$$

- Притоа, од топологијата на мрежата зависи кои јазли се родителски.
- Ако не се знае дали Јован навистина го слушнал алармот тогаш:

$$P(J, \neg M, A, K, \neg 3) + P(J, \neg M, \neg A, K, \neg 3) = 0,00025625646$$

Детално:

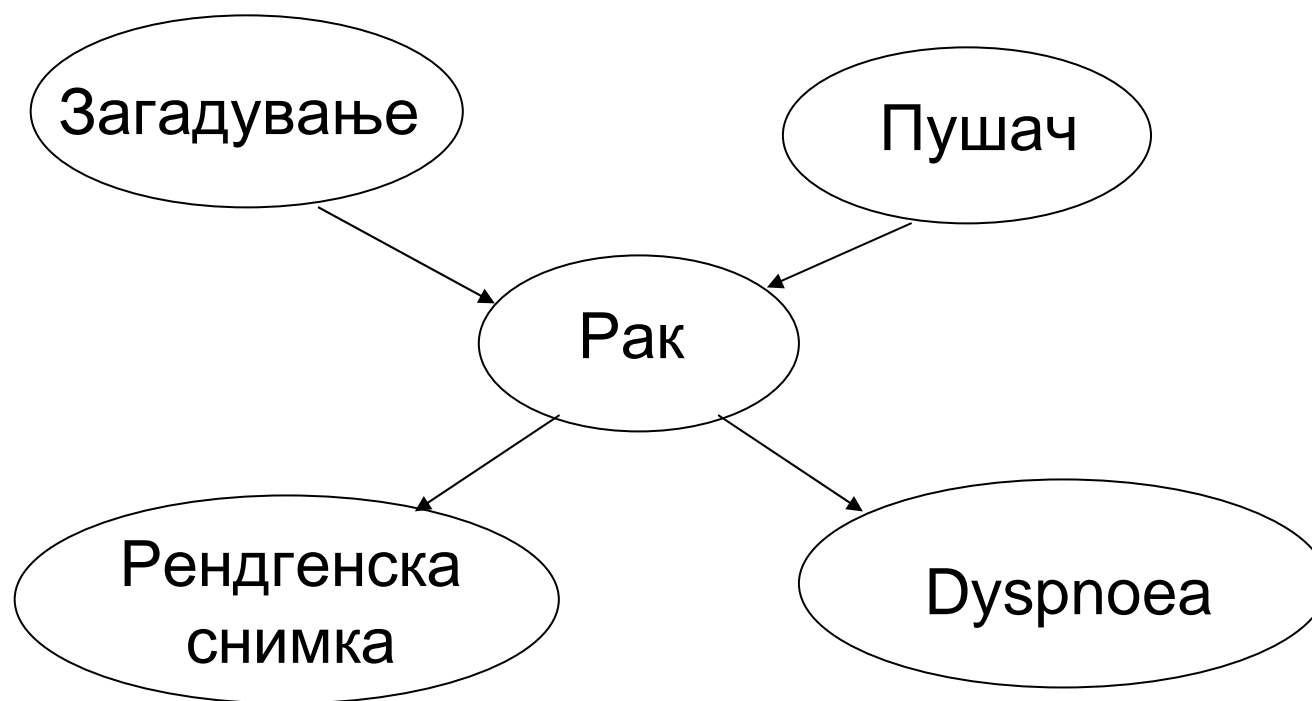
$$\begin{aligned} \blacksquare P(J, \neg M, A, K, \neg 3) &= \\ P(J/A) * P(\neg M/A) * P(A/K, \neg 3) * P(K) * P(\neg 3) &= \\ 0,90 * 0,30 * 0,940 * 0,001 * 0,998 &= \\ 0,0002532924 \end{aligned}$$

$$\begin{aligned} \blacksquare P(J, \neg M, \neg A, K, \neg 3) &= \\ P(J/\neg A) * P(\neg M/\neg A) * P(\neg A/K, \neg 3) * P(K) * & \\ P(\neg 3) &= \\ 0,05 * 0,99 * 0,060 * 0,001 * 0,998 &= \\ 0,00000296406 \end{aligned}$$

Пример 2 за Бејесова мрежа

- Пациентот чувствува тешкотии во движењето (Dyspnoea) и исплашен дека можеби има рак на белите дробови оди на лекар.
- Лекарот знае дека овој симптом може да биде последица и на други заболувања, како на пример, туберкулоза (ТВ) или воспаление на белите дробови.
- Тој исто така знае дека постојат други релевантни информации кои треба да се земат предвид, како што е податокот дали пациентот е пушач и какво е загадувањето на средината во која живее.
- Рендгенската снимка на белите дробови може да индицира постоење на ТВ или на рак на белите дробови.

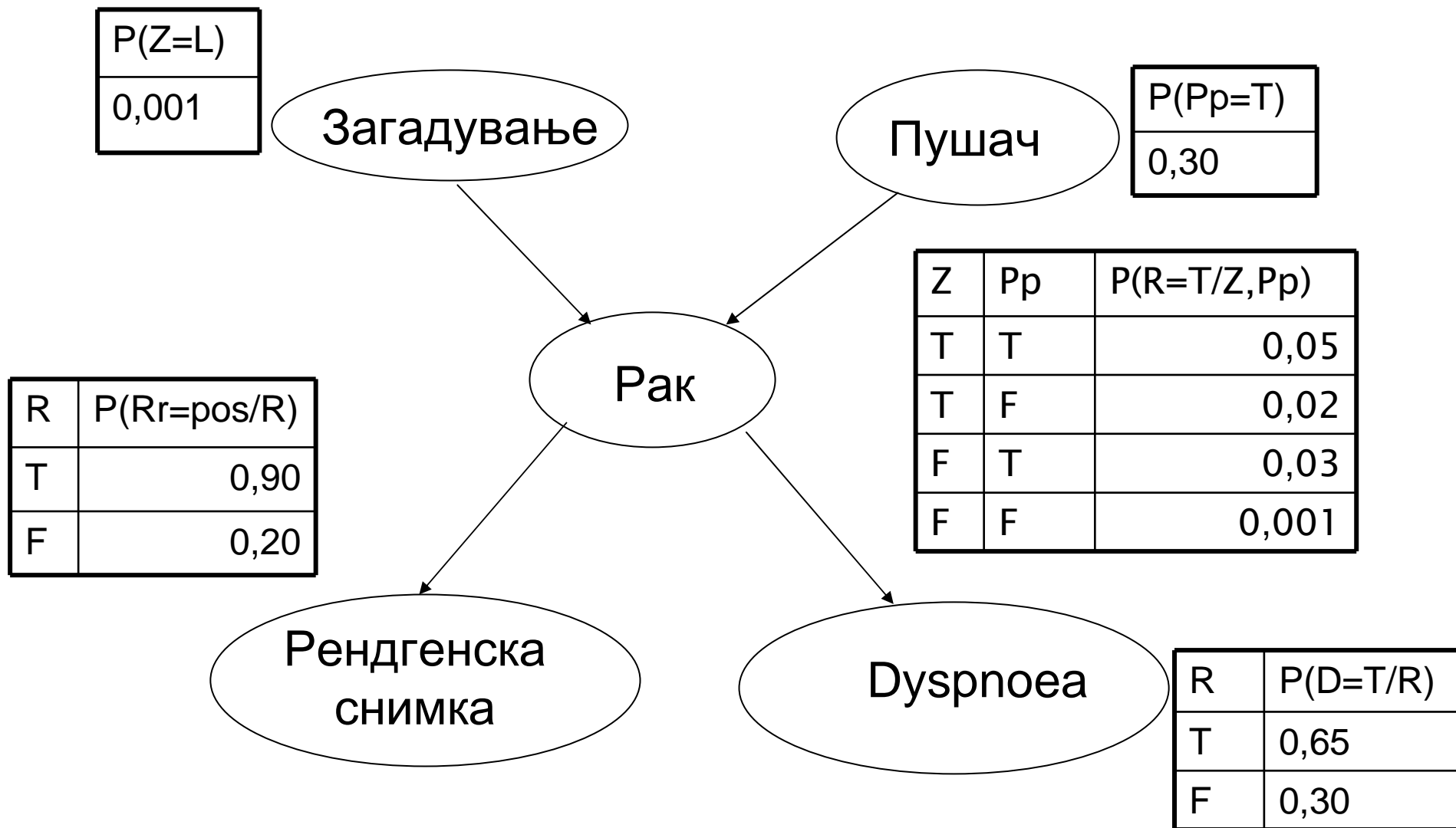
Илустрација на примерот



Фактори вклучени во Бејесовата мрежа

Node name	Type	Values
<i>Pollution</i>	Binary	$\{low, high\}$
<i>Smoker</i>	Boolean	$\{T, F\}$
<i>Cancer</i>	Boolean	$\{T, F\}$
<i>Dyspnoea</i>	Boolean	$\{T, F\}$
<i>XRay</i>	Binary	$\{pos, neg\}$

Продолжение на примерот



Како се градат мрежите на Bayes

- Се одбира множество случајни променливи кои го опишуваат проблемот
- Случајните променливи се означуваат како X_1, \dots, X_n
- Од $i=1$ до n
 - Отвори нов јазел за X_i во мрежата
 - Множеството родители на X_i е најмалото можно множество јазли кои веќе се во мрежата, а се условно независни со X_i . Тие стануваат $Roditeli(X_i)$
 - Воспостави лак од секој $Roditeli(X_i)$ кон X_i
 - Ако X_i има барем еден родител, тогаш дефинирај ја табелата на условната веројатност. Во спротивно, дефинирај ја веројатноста на јазелот.

Практичен начин за конструкција на мрежата

- Најдобро е да се започне со јазелот / јазлите кои се причина, а потоа се додаваат јазлите врз кои тие делуваат итн., сè додека не се дојде до лисјата во графот (каузална / причинско-последична структура)
- Каузалниот редослед овозможува директна репрезентација на условните веројатности, бидејќи секој друг редослед ќе бара внесување нови врски во мрежата за воспоставување на независност.
- Ако се почне со други јазли или се користи друг редослед се добиваат посложени структури кои бараат повеќе пресметки / веројатности (колку што има заедничката распределба).

Сведоштво - практично

- *Конкретен доказ* (пример, дека пациентот е пушач)
- *Негативен доказ* (пример, дадена променлива не е во дадена состојба, што значи дека е во некоја од преостанатите)
- *Несигурен или можен доказ* кога изворот на информации не е сигурен за тоа (пример, радиологот е 80% сигурен дека рендгенската снимка е позитивна)
- Расудувањето со помош на Бејесовите мрежи е флексибилно: се додава ново сведоштво за било кој јазел (информации за околината), се ажурираат уверувањата за сите други.

Расудувањето во мрежите на Bayes

- Да се најде условната (posterior) веројатност за одредени променливи - хипотези врз основа на вредностите кои ги имаме како доказ (evidence variables)
- Агентот ги добива веројатностите за доказните променливи врз основа на своите перцепции или слични расудувања и поставува прашање за можните веројатности на другите променливи, за да може да одлучи која акција да ја одбере.

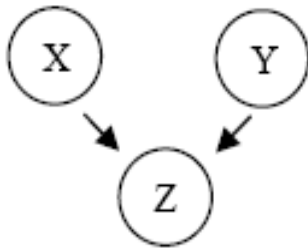
Резиме за мрежите на Bayes

- Се создаваат интуитивно, но притоа, ако од јазелот X води стрелка кон Y , во тој случај, X има директно влијание врз Y .
- Се викаат и причинско-последични мрежи (causal networks) во кои лаците ги претставуваат директните причини.
- Пресметките се вршат со помош на табели на условната веројатност (conditional probability tables CPTs).
- Услови:
 - ☐ настаните се независни
 - ☐ треба да се запази редоследот причина – последица

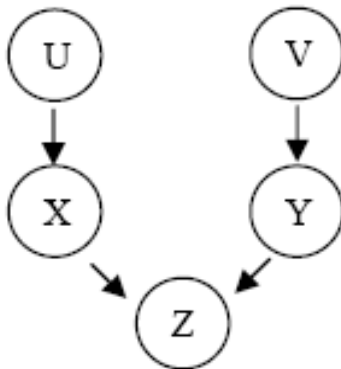
D-одделување (D-Separation)*

- Мрежите на Bayes може да се градат само ако секој јазел е условно независен од сите јазли што не се негови наследници
- Ова е услов на локалната семантика.
- D-одделувањето е критериум кој е построг од локалната семантика.

Илустрација на двата критериума*



Гледано локално, X и Y се независни од претходните јазли.

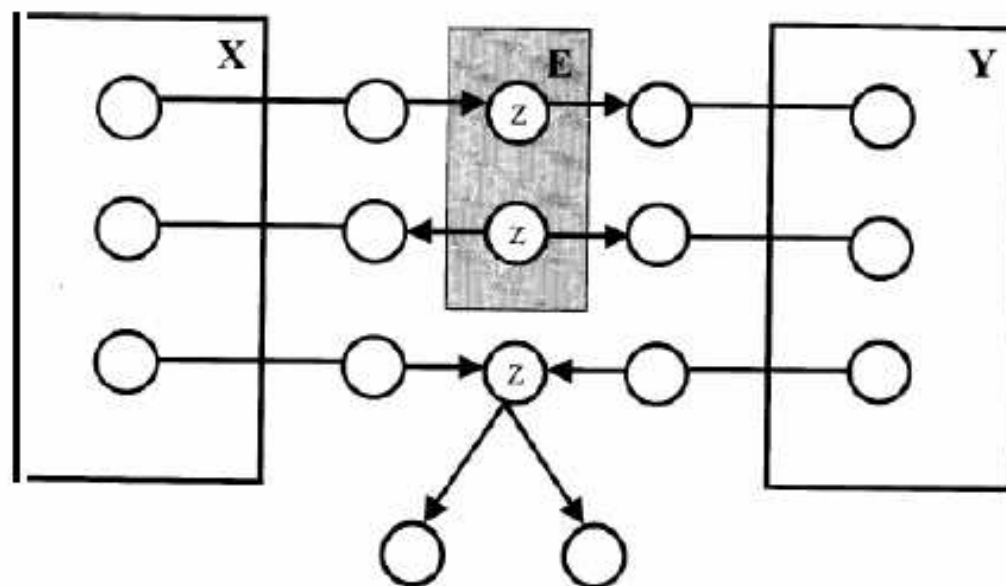


X и Y веќе не се независни од претходните јазли, но сепак се меѓусебно независни.

Како да се одреди дали едно множество јазли X е независно од друго множество Y , во услов на множество докази E ?*

- Ако секој ненасочен пат од јазел во X кон јазел во Y е одделен со E , тогаш X и Y се условно независни преку E .
- Множеството јазли E ги одделува множествата јазли од X и Y ако секој ненасочен пат од јазелот од X кон јазел од Y е блокиран од E .
- Патот е блокиран со множеството јазли од E ако постои јазел Z на патот, за кој важи:
 1. Z е во E и Z има стрелка која води кон неа и стрелка која води надвор од неа (ланец)
 2. Z е во E и двете стрелки на Z водат надвор од неа (заедничка причина)
 3. Ниту Z , ниту било кој наследник на Z не е во E и двете стрелки водат кон Z (заеднички ефект)

Илустрација на критериумите за блокирање на патот*



E d-separates X and Y

* Сладот е преземен од <http://www.blutner.de/uncert/index.htm>

Теоријата на Демпстер и Шефер (Dempster-Shafer Theory)

- Според Шефер, во теоријата на доказите се дефинираат две мерки:
 - мерка за верување (belief measure): Bel
 - мерка за веродостојност (plausibility measure): Pl
- Според Демпстер, овие функции се функции на верување и тие се наменети за пресликувања со повеќе вредности.

Теоријата на Демпстер и Шефер (Dempster-Shafer Theory)

- Теоријата се базира врз две идеи:
 - Да се одреди мерка на верување (degree of belief) за дадено прашање врз основа на субјективна веројатност за сродно прашање
 - Комбинирање на мерките на верување кога тие се засновани на независни докази
- Мерката за верување може, но не мора да ги има својствата на веројатноста
- Мерката за верување се вика и функција на верување $Bel(x)$

Пример 1

- Фрлање на паричка кое ви го предлага сомнителен тип кој ви вели дека можете да заработите пари (не знаете дали не си поиграл со паричките)
 - Бидејќи немате доказ, степенот на вашето верување е исто и за двата одговора петка и глава: $Bel(\text{Глава}) = 0,5$; $Bel(\neg\text{Глава}) = 0,5$
 - Ако дојде експерт за парички, можно е да каже со 90% сигурност дека паричките не се менувани: $Bel(\text{Глава}) = 0,9 \times 0,5 = 0,45$
- Теоријата не дава одговор на секое прашање

Пример 2

- Имате субјективен впечаток за тоа дали Бети ја зборува вистината или не
 - 0,9 дека ја зборува вистината и 0,1 дека не
 - Ако дојде и ви каже дека некој ја удрил вашата кола, тоа што и верувате само 90% дека ја кажува вистината ќе ви значи дека веројатноста некој навистина да ве удрил е 0,9. Но верувањето дека никој не ви ја удрил колата е 0 (не е 0,1, бидејќи немаме доказ). Тоа нема исто значење како веројатноста 0, дека никој не ја удрил колата.
 - Вредностите на функцијата на верување се 0,9 и 0 соодветно.

Пример 2, проширување

- Имате субјективен впечаток и за тоа дали Марија ја кажува вистината
 - 0,9 дека ја зборува вистината и 0,1 дека не
 - Ако дојде Марија и таа исто така ви каже дека некој ви ја удрил колата, веројатноста за верување на двете е $0,9 \cdot 0,9 = 0,81$; веројатноста дека не може да им се верува е $0,1 \cdot 0,1 = 0,01$, а веројатноста дека и веруваме барем на едната е $1 - 0,01 = 0,99$ (вашето верување дека некој ви ја удрил колата, бидејќи и двете во кажење е 0,99).
- Имаме уверување за едно прашање (дали некој ви ја удрил колата?) врз основа на субјективната веројатност за друго (колку им веруваме?)

Интерпретација

- Интерпретација 1: се дефинира праг на веројатности
 - Пред да дојде експертот: $Bel(\text{Глава}) = [0,1]$ по доаѓањето: $Bel(\text{Глава}) = [0.45, 0.55]$.
 - Опсегот на интервалот посочува дали ви се потребни повеќе докази
- Долниот праг е мерка за верување, а горниот се вика мерка за веродостојност (plausability)

Непрецизирана (fuzzy) претстава (логика)

- Прецизно дефинирано множество
(crisp)
 $A = \{a, b, c, d, e\}$
- Се дефинира функција на припадност
(membership function)
 $M = \{1, 1, 1, 1, 1\}$

Замислете кога функцијата на припадност би можела да
прима вредности во интервалот $[0, 1]$

Fuzzy = матно, замаглено

- Fuzzy \neq фази, оти фази е множина од фаза

Значи:

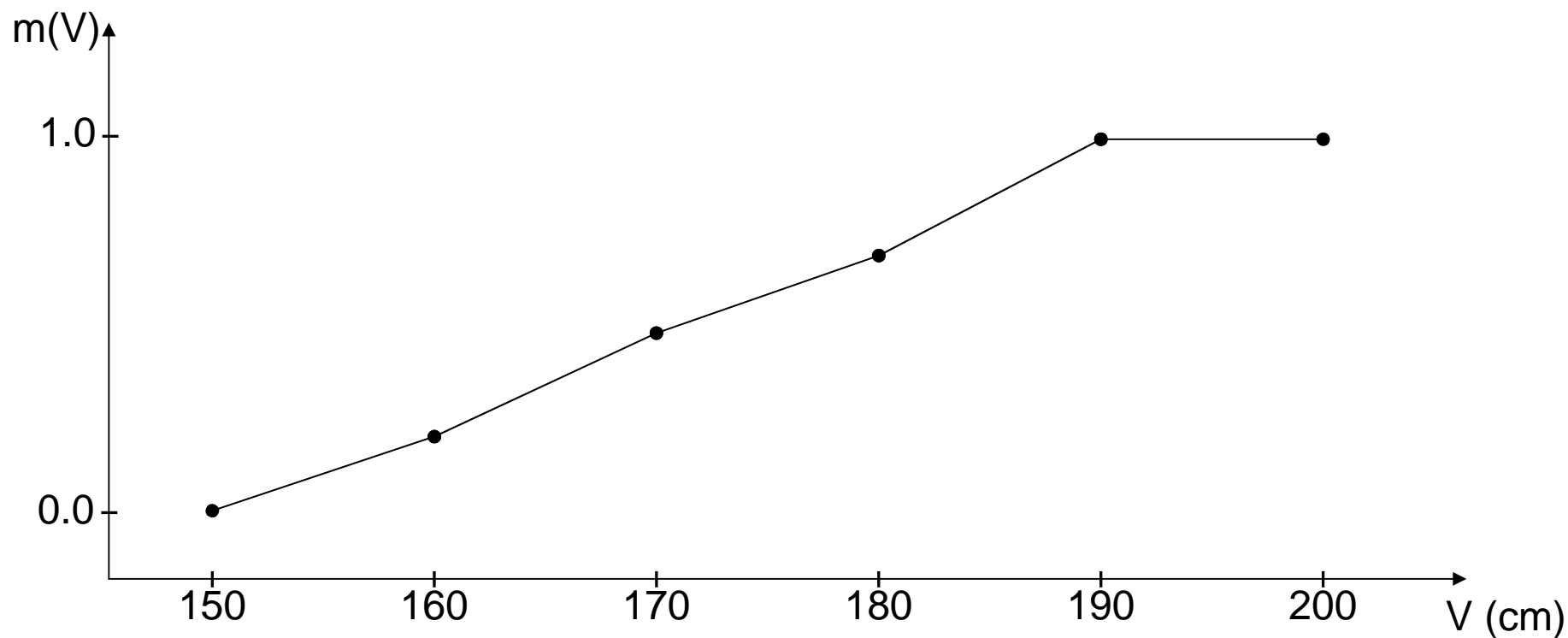
- Fuzzy Logic = матна логика,
- Fuzzy concepts = матни, заматени или замаглени концепти,
- Fuzzification = заматување, замаглување
- Defuzzification = одмаглување

Непрецизирани множества

- $M = \{0.7, 0.6, 0.2, 1.0\}$ на пример да биде множество на припадност на дадени елементи кон некое множество
- Па вака може да се дефинираат и некои поими (концепти)
 - Пример: Висок. Што е тоа висок човек?
 - $V = \{150, 160, 170, 180, 190, 200\}$
 $m(V) = \{0.0, 0.2, 0.5, 0.7, 1.0, 1.0\}$

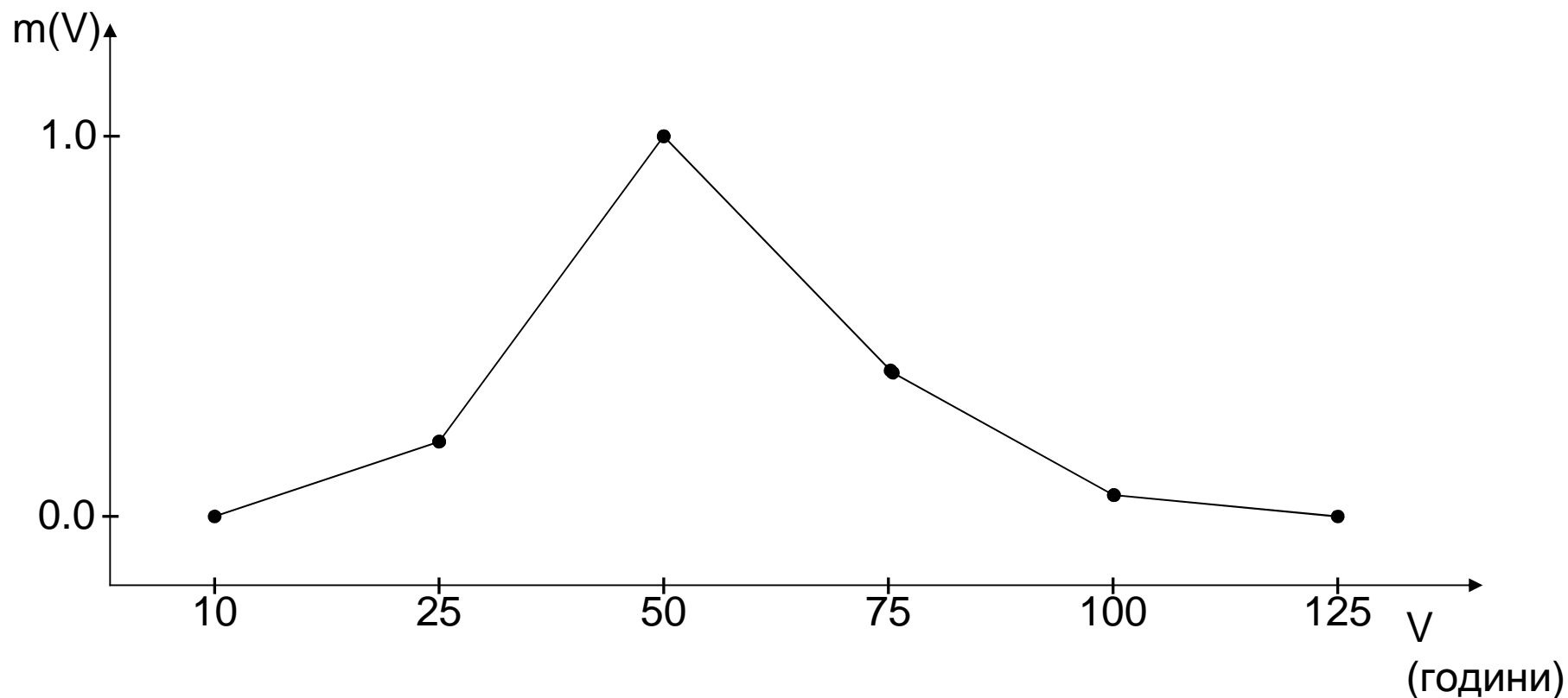
ПОИМОТ ВИСОК ЧОВЕК

- $V = \{150, 160, 170, 180, 190, 200\}$
 $m(V) = \{0.0, 0.2, 0.5, 0.7, 1.0, 1.0\}$



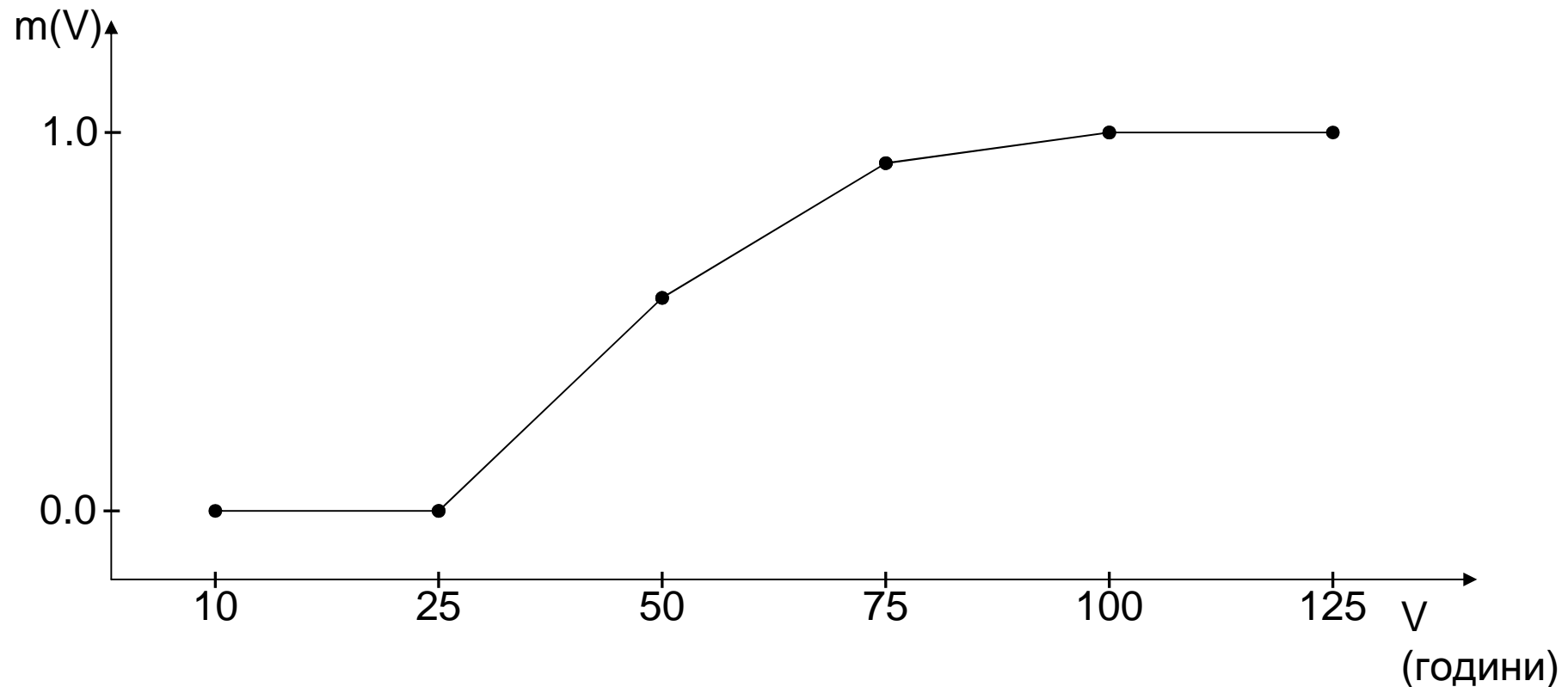
Поимот средовечен човек

- $V = \{10, 25, 50, 75, 100, 125\}$
 $m(V) = \{0.0, 0.2, 1.0, 0.3, 0.1, 0.0\}$



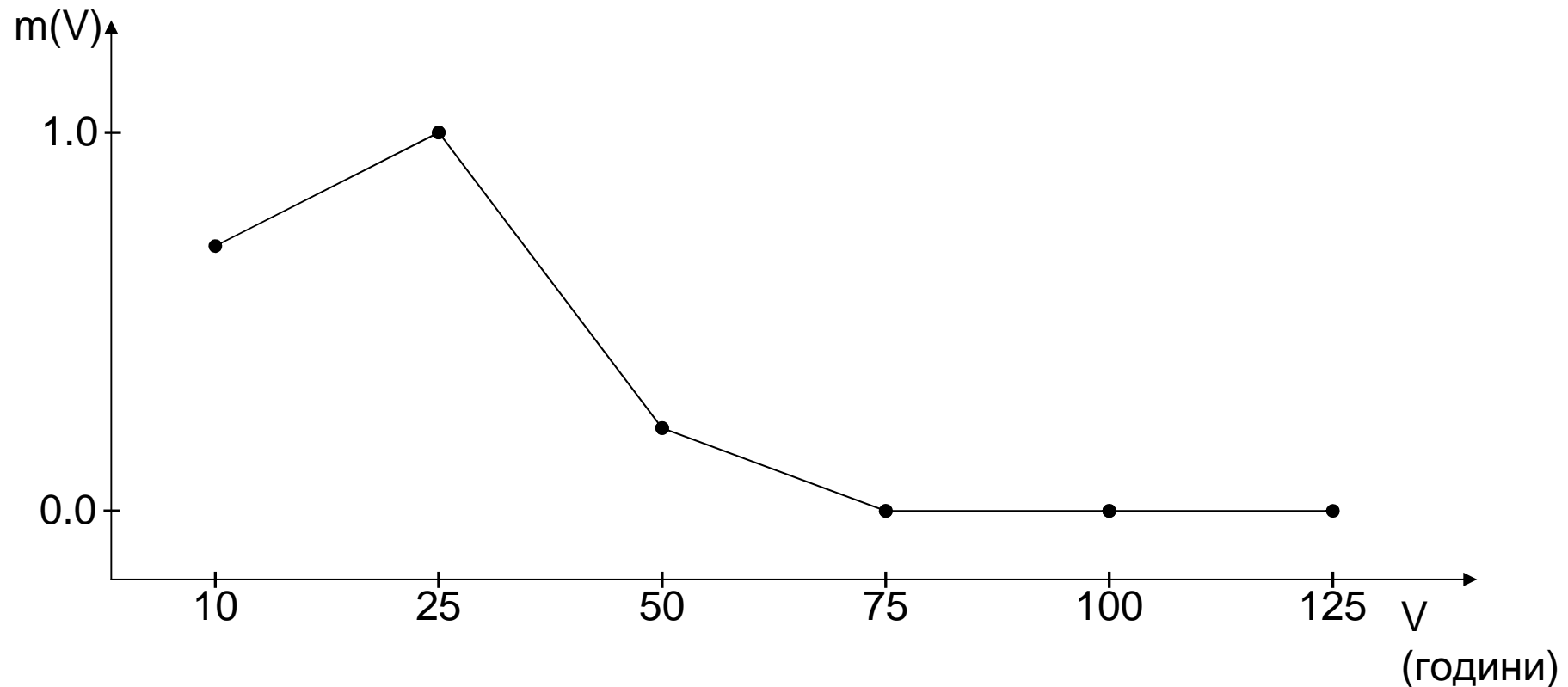
Поимот стар човек

- $V = \{10, 25, 50, 75, 100, 125\}$
 $m(V) = \{0.0, 0.0, 0.5, 0.9, 1.0, 1.0\}$



Поимот млад човек

- $V = \{10, 25, 50, 75, 100, 125\}$
 $m(V) = \{0.7, 1.0, 0.2, 0.0, 0.0, 0.0\}$



Користена литература

- Artificial Intelligence, A Modern Approach
2nd edition, Russel and Norvig
- Artificial Intelligence, A New Synthesis, Nils J.
Nilsson



Прашања?