OSNOVE UMETNE INTELIGENCE
2018/19

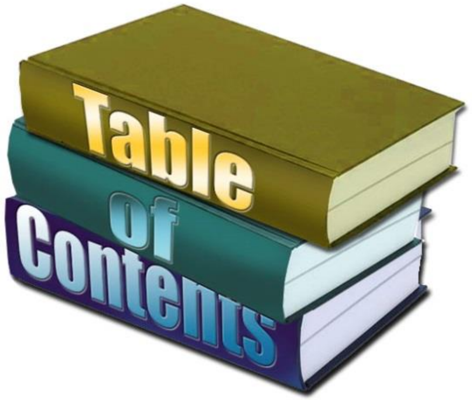
razporejanje opravil

verjetnostno sklepanje z bayesovskimi mrežami

© Zoran Bosnić

del gradiva povzet po: Bratko: Prolog programming for AI, Pearson (2011) in Russell, Norvig: AI: A Modern Approach, Pearson (2010)

Pregled

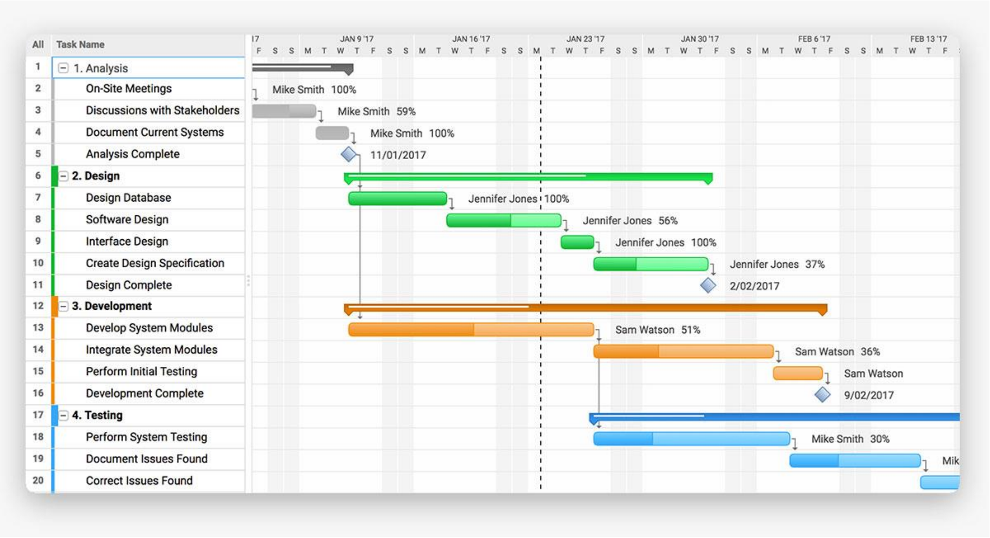


- planiranje
  - planiranje s "klasičnim" preiskovanjem prostora stanj
  - planiranje s sredstvi in cilji
  - planiranje z regresiranjem ciljev
  - razporejanje opravil

- bayesovske mreže
  - definicija
  - odvisnosti v bayesovski mreži
  - neodvisnosti v bayesovski mreži
  - verjetnostno sklepanje
  - ekvivalenca bayesovskih mrež

Planiranje in razporejanje opravil

- do sedaj (klasično planiranje): **kaj narediti** in v kakšnem **vrstnem redu**
- pristopi:
  - planiranje kot preiskovanje prostora stanj
  - planiranje s sredstvi in cilji
  - planiranje z regresiranjem ciljev skozi akcije
- v realnosti imamo številne **dodatne omejitve**:
  - časovne omejitve (začetki aktivnosti, trajanja aktivnosti, roki zaključkov)
  - resursi (omejeno število procesorjev, kadra, bencina, denarja, surovin, …)



Razporejanje opravil

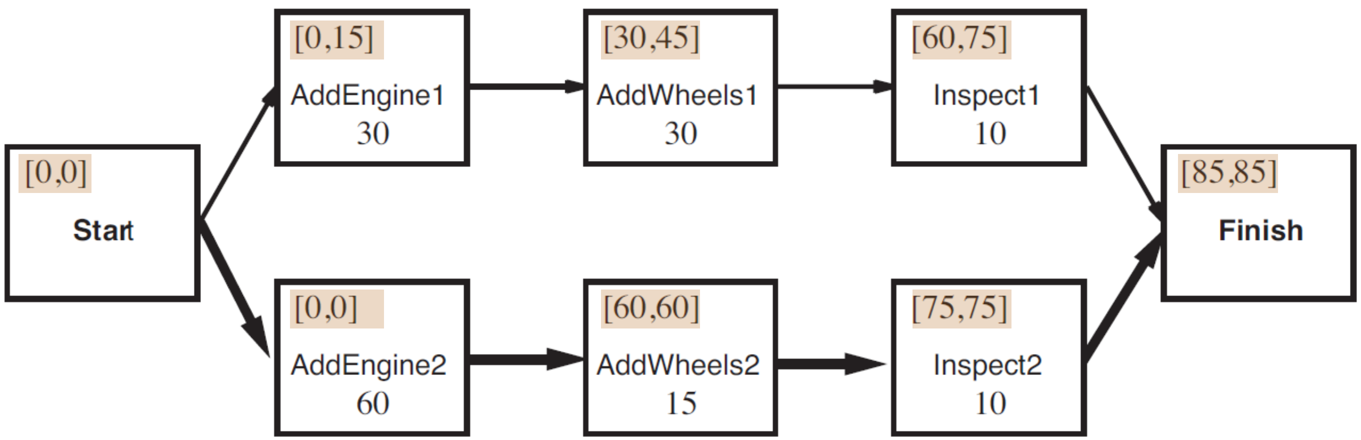
- delno urejen plan: vrstni red podmnožice aktivnosti je lahko urejen
- razširimo lahko notacijo (PDDL):
  - **`Akcija1 ≺ Akcija2`**: pomeni, da se mora Akcija1 zgoditi pred Akcijo2
  - **Resources** podaja števila razpoložljivih resursov
  - **DURATION** opredeljuje trajanje posamezne akcije
  - **CONSUME** opredeljuje (trajno) porabo določene količine resursov
  - **USE** opredeljuje (začasno) zasedenost količine resursov med izvajanjem akcije

```
Jobs (AddEngine1 ≺ AddWheels1 ≺ Inspect1,
      AddEngine2 ≺ AddWheels2 ≺ Inspect2 )
Resources (EngineHoists(1), WheelStations(1), Inspectors(2), LugNuts(500))

Action (AddEngine1 , DURATION:30,
        USE:EngineHoists(1))
Action (AddEngine2 , DURATION:60,
        USE:EngineHoists(1))
Action (AddWheels1 , DURATION:30,
        CONSUME:LugNuts(20), USE:WheelStations(1))
Action (AddWheels2 , DURATION:15,
        CONSUME:LugNuts(20), USE:WheelStations(1))
Action (Inspect i, DURATION:10,
        USE:Inspectors (1))
```

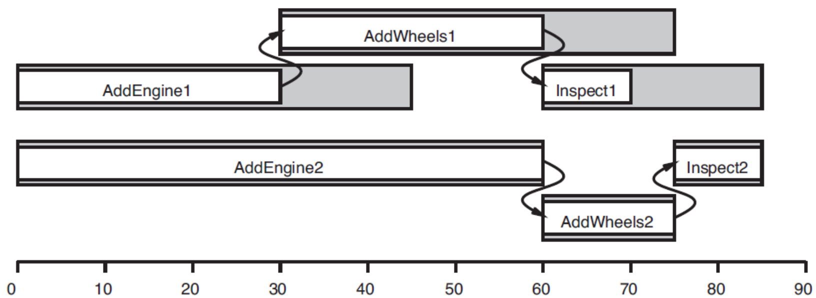
Razporejanje opravil

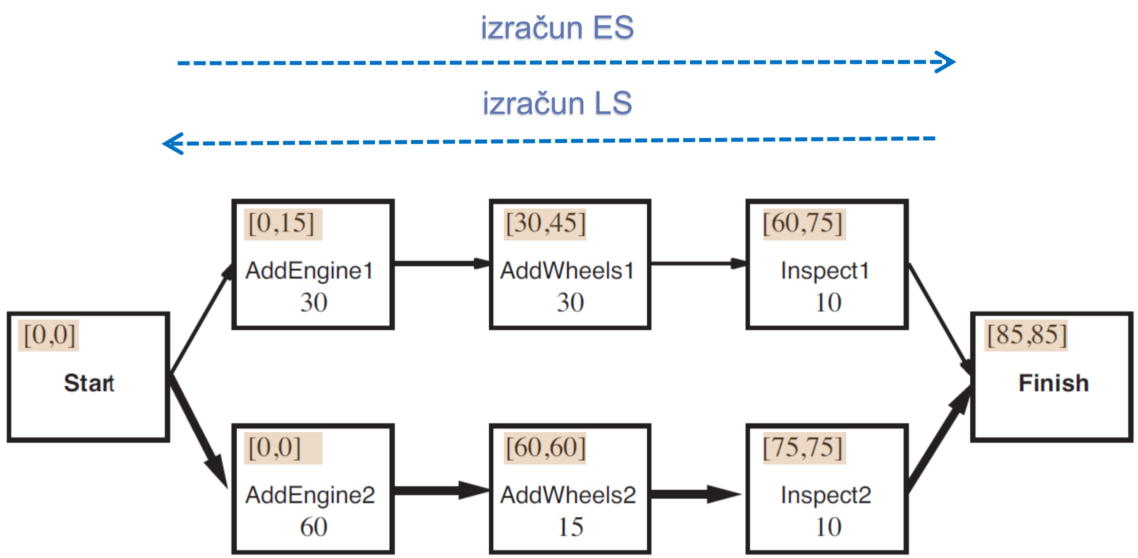
- za začetek: samo časovne omejitve
- **metoda kritične poti**
  - kritična pot: pot, ki je najdaljša in določa dolžino trajanja celotnega plana (krajšanje vzporednih poti ne vpliva na trajanje plana)
  - vsaki akciji priredimo par **[ES, LS]**:
    - **ES** – najbolj zgodnji možen začetek (angl. *Earliest Start*)
    - **LS** – najbolj pozen možen začetek (angl. *Latest Start*)



Razporejanje opravil

$$
\begin{aligned}
&ES(Start) = 0\\
&ES(B) = \max_{A \prec B} ES(A) + Duration(A)\\
&LS(Finish) = ES(Finish)\\
&LS(A) = \min_{A \prec B} LS(B) - Duration(A)\\
&rezerva\ (slack) = LS - ES
\end{aligned}
$$

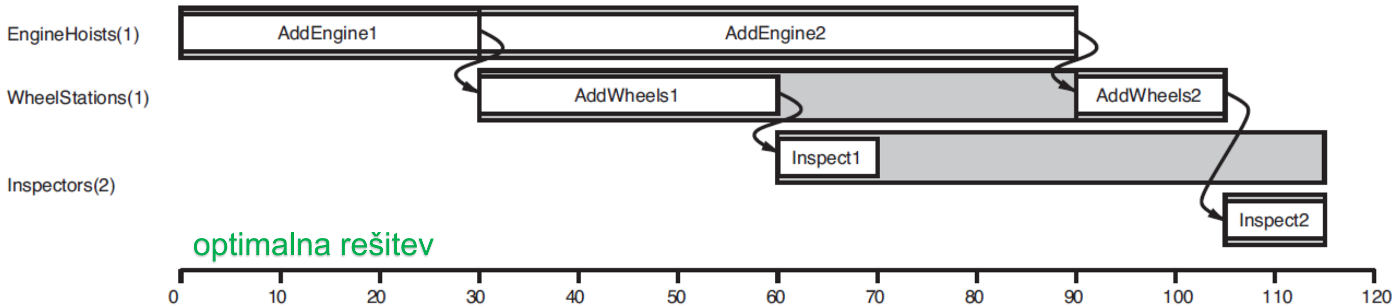




- časovna zahtevnost algoritma: $O(Nb)$, N – število akcij, b – faktor vejanja

Razporejanje opravil

- dodatno: upoštevanje tudi resursov
- uvede **omejitev**, da se aktivnosti, ki potrebujeta iste resurse, ne smeta prekrivati



- sprememba časovne zahtevnosti: $O(Nb)$ → NP-težek problem (!)
- primer izziv iz leta 1963 nerešen 23 let:
  - resursi: 10 strojev, 10 nalog, 100 akcij
  - preizkušene metode: simulirano ohlajanje, tabu search, razveji in omeji, …
- primerna hevristika: algoritem **najmanjše časovne rezerve** (angl. *minimum slack algorithm*)
  - na vsaki iteraciji dodeli **najbolj zgodnji možen začetek** akciji, ki ima **izpolnjene vse predhodnike** in ima **najmanj časovne rezerve**,
  - nato posodobi [ES in LS] za celotni graf in ponovi.

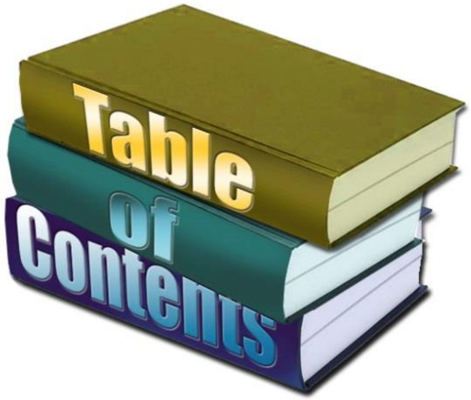
Razporejanje opravil

Diskusija:

- Kakšen je rezultat simulacije algoritma najmanjše časovne rezerve na obravnavanem problemu?
- Ali je rešitev enaka optimalni? Zakaj?

- Kako upoštevati omejitve v zaporedju akcij pri pristopih za planiranje?
- Kako upoštevati omejitve v omejenem številu resursov?

Pregled

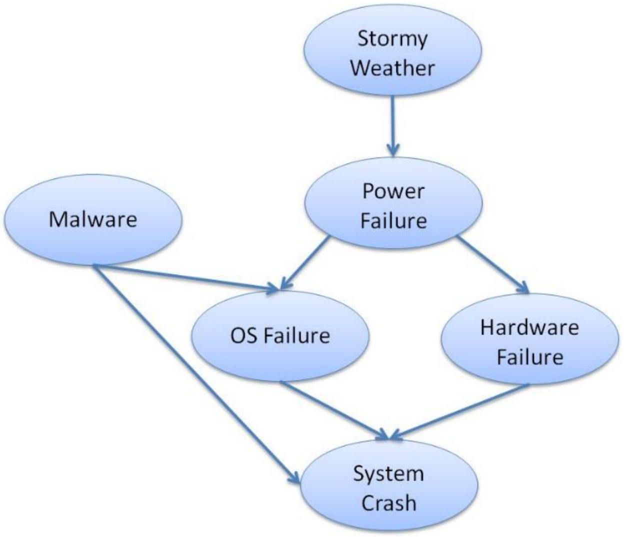


- planiranje
  - planiranje s "klasičnim" preiskovanjem prostora stanj
  - planiranje s sredstvi in cilji
  - planiranje z regresiranjem ciljev
  - razporejanje opravil

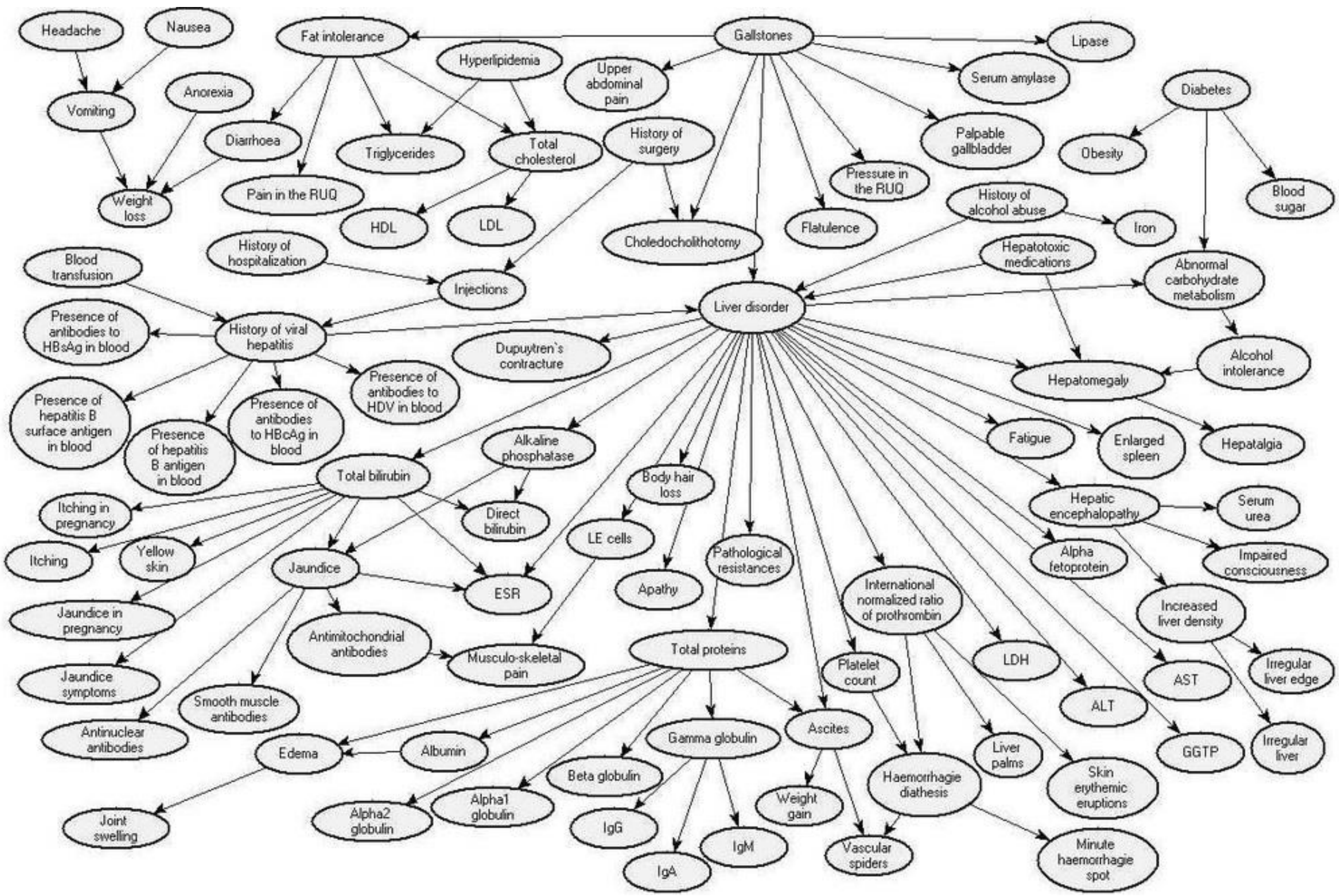
- bayesovske mreže
  - definicija
  - odvisnosti v bayesovski mreži
  - neodvisnosti v bayesovski mreži
  - verjetnostno sklepanje
  - ekvivalenca bayesovskih mrež

Bayesovske mreže

- so verjetnostni model, s katerim predstavimo **odvisnosti med slučajnimi spremenljivkami**
- pristop za **obravnavo negotovosti** v bazah znanja, ki je matematično dobro utemeljen v verjetnosti
- model je predstavljen z **usmerjenim acikličnim grafom**:
  - vozlišča: slučajne spremenljivke (dejstva, hipoteze),
  - povezave: odvisnosti med spremenljivkami (vpliv starša na naslednika)
- primeri uporabe:
  - splošno: za predstavitev verjetnostnega znanja in verjetnostno sklepanje
  - medicina: povezave med boleznijo in simptomi (diagnostika), napovedovanje izida operacije
  - ekspertni sistemi: ocenjevanje kvalitete vode, …
  - sklepanje: kako verjetno so določene trditve, če vemo, da so druge trditve resnične?



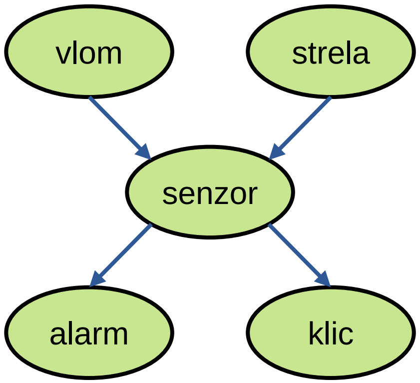
Primer iz medicine



A. Onisko, M. J. Druzdzel, and H. Wasyluk, “A bayesian network model for diagnosis of liver disorders,” in Proceedings of the Eleventh Conference on Biocybernetics and Biomedical Engineering, vol. 2. Citeseer, 1999, pp. 842–846.

Bayesovske mreže

- stanje sveta povzamemo z **vektorjem (logičnih) spremenljivk**
- agent sklepa na **verjetnost** resničnosti določene spremenljivke
- upoštevamo lahko, da so določene spremenljivke med seboj neodvisne, kar predstavimo z bayesovsko mrežo, ki odraža te neodvisnosti (nepovezana vozlišča niso odvisna)
- primer:
  - senzor se sproži ob vlomu v hišo
  - včasih lahko tudi udar strele nehoteno sproži senzor
  - senzor ima nalogo, da sproži alarm in izvede opozorilni telefonski klic
- odvisnosti, ki izhajajo iz mreže:
  - senzor je odvisen od vloma in strele
  - alarm je odvisen od senzorja
  - klic je odvisen od senzorja



Bayesovske mreže

- z zapisom $P(X)$ okrajšamo $P(X = true)$, z zapisom $P(XY)$ pa konjunkcijo
- za opis stanja sveta, ki ima n spremenljivk, bi morali poznati **popolno verjetnostno porazdelitev** ($2^n - 1$ podatkov – možnih stanj vseh logičnih spremenljivk)
  - spremenljivke: V, St, Se, A, K
  - popolna verjetnostna porazdelitev:

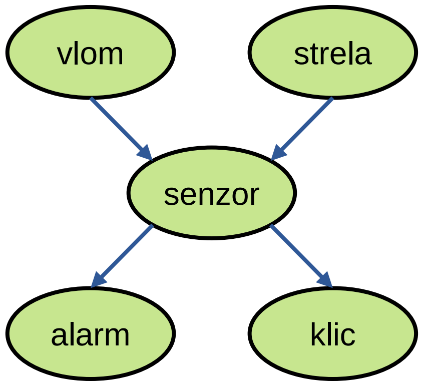
    $P(V\ St\ Se\ A\ K) =$ ...

    $P({\sim}V\ St\ Se\ A\ K) =$ ...

    $P(V{\sim}St\ Se\ A\ K) =$ ...

    $P({\sim}V{\sim}St\ Se\ A\ K) =$ ...

    ...
  - potrebujemo $2^5 - 1 = 31$ verjetnosti
  - nepraktično ali nemogoče za veliko število spremenljivk

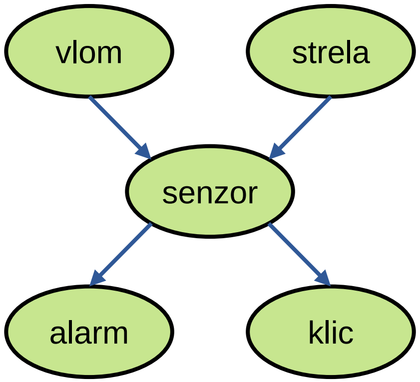


- verjetnost pojubnega dogodka (npr. $P(VK)$) izračunamo z vsoto vseh kombinacij vrednosti spremenljivk St, Se, A (pozitivna ali negirana) pri vrednostih $V = true$ in $K = true$.

Pogojne verjetnosti

- ker bayesovska mreža opredeljuje odvisnosti spremenljivk, lahko opredelimo problem samo s pogojnimi verjetnostmi:

$$P(vlom) = 0{,}001$$
$$P(strela) = 0{,}02$$
$$P(senzor \mid vlom \ \wedge strela) = 0{,}9$$
$$P(senzor \mid vlom \ \wedge \sim strela) = 0{,}9$$
$$P(senzor \mid \sim vlom \ \wedge strela) = 0{,}1$$
$$P(senzor \mid \sim vlom \ \wedge \sim strela) = 0{,}001$$
$$P(alarm \mid senzor) = 0{,}95$$
$$P(alarm \mid \sim senzor) = 0{,}001$$
$$P(klic \mid senzor) = 0{,}95$$
$$P(klic \mid \sim senzor) = 0$$



- podamo torej 10 podatkov namesto $2^5 - 1 = 31$
- za spremenljivke, ki niso med seboj odvisne, ne potrebujemo vseh kombinacij verjetnosti:
  - če sta X in Y **odvisna**, v splošnem velja $P(XY) = P(X) \cdot P(Y|X)$ (potrebujemo $P(Y|X)$)
  - če sta X in Y **neodvisna**, velja: $P(XY) = P(X) \cdot P(Y)$ ($P(Y|X)$ ne potrebujemo, ker zaradi neodvisnosti velja $P(Y|X) = P(Y)$)

Pogojne verjetnosti

- pogojne verjetnosti lahko predstavimo tudi s tabelami pogojnih verjetnosti

$P(vlom)$
0,001

$P(strela)$
0,02

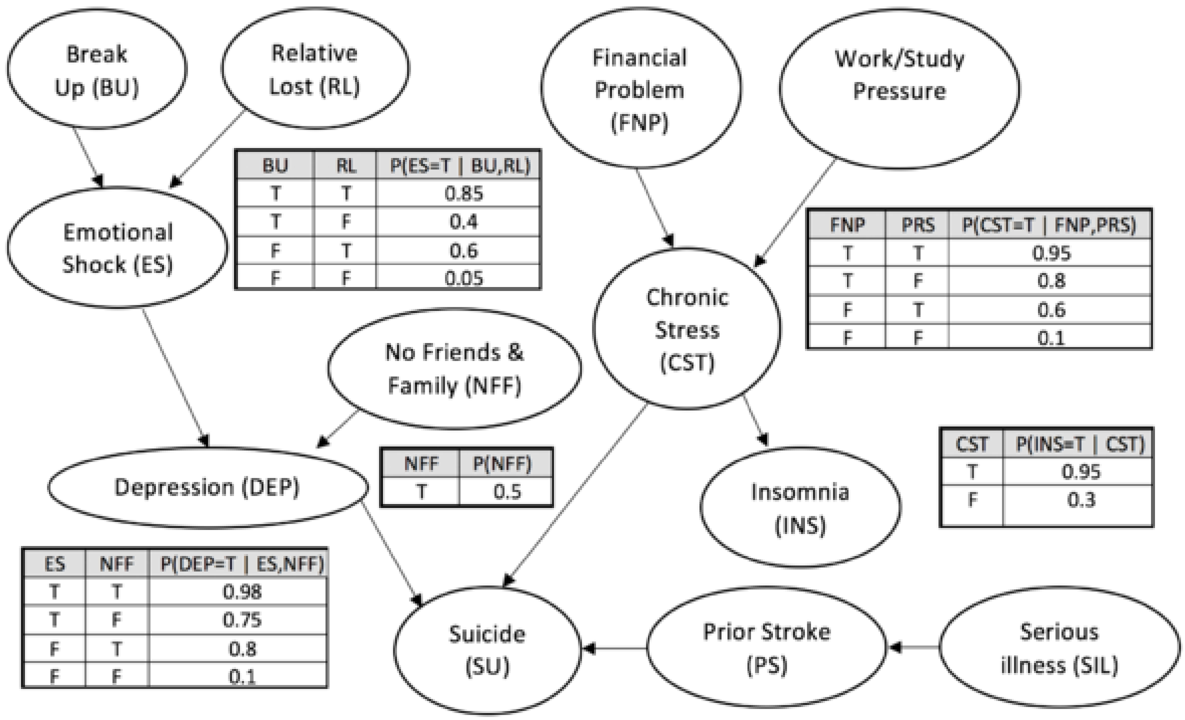
$vlom$	$strela$	$P(senzor)$
true	true	0,9
true	false	0,9
false	true	0,1
false	false	0,001

$senzor$	$P(alarm)$
true	0,95
false	0,001

$senzor$	$P(klic)$
true	0,95
false	0

- verjetnostni značaj modeliranja z bayesovskimi mrežami: vsota pogojnih verjetnosti $P(X|Y)$ in $P(X|{\sim}Y)$ ni enaka 1, kar nakazuje, da obstajajo za X tudi drugi, neopredeljeni razlogi, ki niso zajeti v predstavitvi problema s podano mrežo

Primer s podanimi verjetnostmi



BU	P(BU)
T	0.2

RL	P(RL)
T	0.1

FNP	P(FNP)
T	0.3

PRS	P(PRS)
T	0.3

BU	RL	P(ES=T | BU,RL)
T	T	0.85
T	F	0.4
F	T	0.6
F	F	0.05

FNP	PRS	P(CST=T | FNP,PRS)
T	T	0.95
T	F	0.8
F	T	0.6
F	F	0.1

NFF	P(NFF)
T	0.5

CST	P(INS=T | CST)
T	0.95
F	0.3

ES	NFF	P(DEP=T | ES,NFF)
T	T	0.98
T	F	0.75
F	T	0.8
F	F	0.1

DEP	CST	PS	P(SU=T | DEP,CST,PS)
T	T	T	0.8
T	T	F	0.95
T	F	T	0.65
T	F	F	0.75
F	T	T	0.3
F	T	F	0.4
F	F	T	0.02
F	F	F	0.05

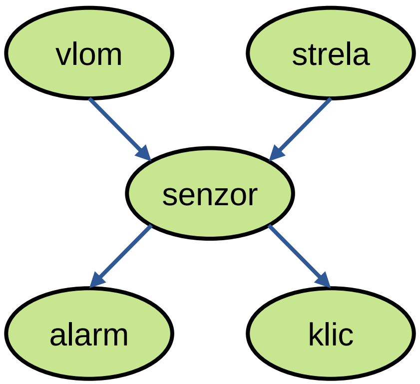
SIL	P(PS=T | SIL)
T	0.95
F	0.2

SIL	P(SIL)
T	0.2

The University of British Columbia: CPSC522/Bayesian Networks, Student’s Death Diagnosis Bayesian Network Model

Izračun verjetnosti dogodka

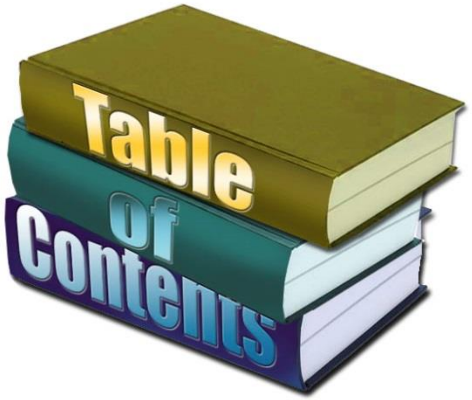
- s pogojnimi verjetnostmi **lahko izračunamo verjetnost dogodka** iz popolne verjetnostne porazdelitve
- primer: kakšna je verjetnost $P(V \sim St\ Se\ A\ K)$?
  $P(V \sim St\ Se\ A\ K) = P(V) \cdot P(\sim St\ Se\ A\ K|V) =$
  $= P(V) \cdot P(\sim St\ |V) \cdot P(Se|V{\sim}St) \cdot P(A|V{\sim}St\ Se) \cdot P(K|V{\sim}St\ Se\ A)$
- zaradi neodvisnosti, podanih v mreži, velja:
  $P(\sim St|V) = P(\sim St)$
  $P(A|V \sim St\ Se) = P(A|Se)$
  $P(K|V \sim St\ Se\ A) = P(K|Se)$



- torej:
  $P(V \sim St\ Se\ A\ K) = P(V) \cdot P(\sim St\ Se\ A\ K|V) =$
  $= P(V) \cdot P(\sim St) \cdot P(Se|V{\sim}St) \cdot P(A|Se) \cdot P(K|Se)$
  $= 0{,}001 \cdot 0{,}98 \cdot 0{,}9 \cdot 0{,}95 \cdot 0{,}9 = 0{,}00075$
- v splošnem velja:

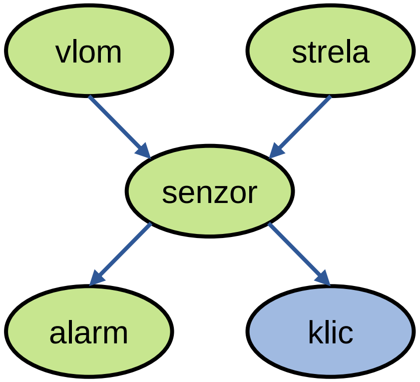
$$P(X_1 X_2 \ldots X_n) = \prod_{i=1}^{n} P(X_i|starši(X_i))$$

Pregled

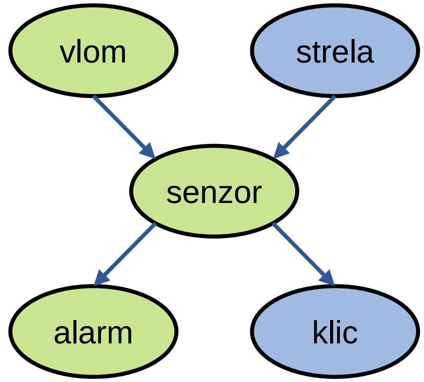


- bayesovske mreže
  - definicija
  - odvisnosti v bayesovski mreži
  - neodvisnosti v bayesovski mreži
  - verjetnostno sklepanje
  - ekvivalenca bayesovskih mrež

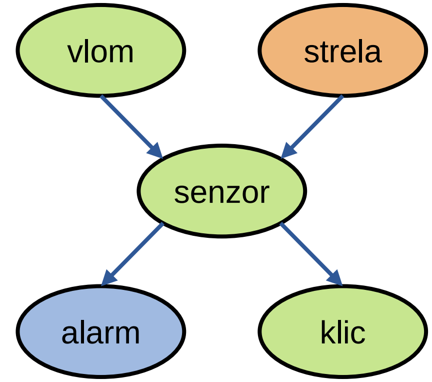
Verjetnostno sklepanje



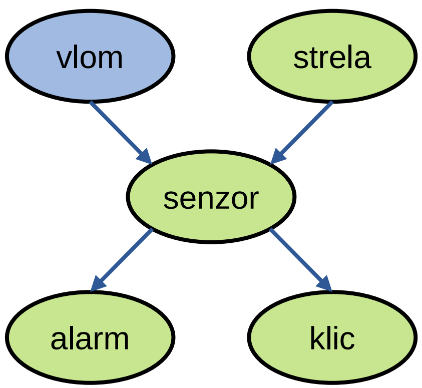
$P(V \mid K) = ?$



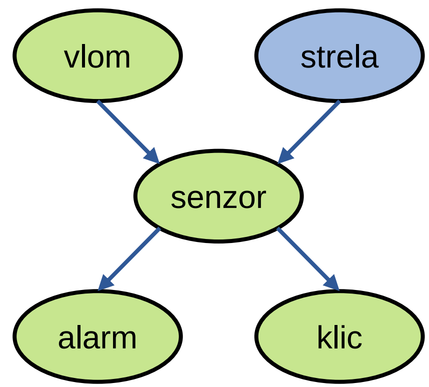
$P(V \mid K\ St) = ?$



$P(V \mid A\ {\sim}St) = ?$



$P(A\ {\sim}K \mid V) = ?$

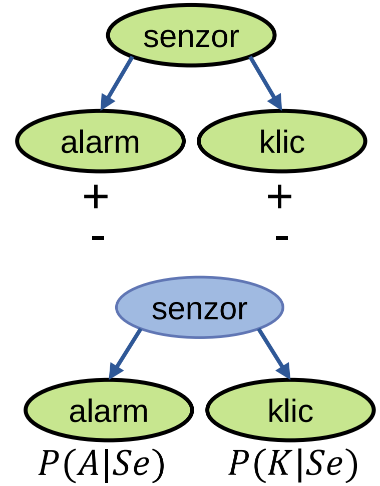


$P(V \mid St) = ?$

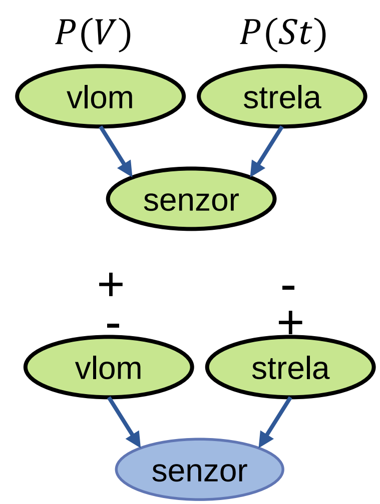
- možni sta dve smeri sklepanja:
  - **vzročno** (od vzrokom k posledicam): npr. $P(A \mid V\ St) = ?$
  - **diagnostično** (od posledic k vzrokom): npr. $P(V \mid A) = ?$

Odvisnosti v mreži

- **skupni prednik**:
  - *alarm* in *klic* sta **odvisna;** če vemo, da je eden od njiju resničen, vpliva to tudi na naše verjetje o resničnosti drugega (če se je sprožil alarm, se je verjetno izvedel tudi klic); $P(A|K) \neq P(A)$, $P(K|A) \neq P(K)$
  - vendar: poznavanje resničnosti prednika *senzor* omogoči, da *alarm* in *klic* obravnavamo kot **neodvisna** (vemo, da se je sprožil *senzor*; torej se je z določeno verjetnost tudi sprožil *alarm* in z določeno (neodvisno) verjetnostjo izvedel tudi *klic;* $P(A|Se\ K) = P(A|Se)$, $P(K|SeA) = P(K|Se)$

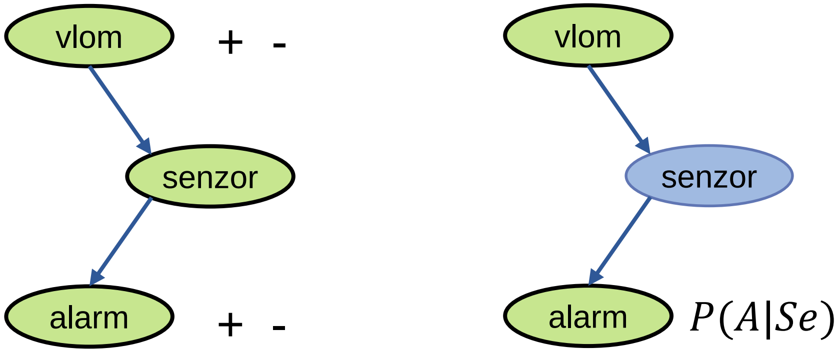


- **skupni naslednik**:
  - *vlom* in *strela* sta medseboj **neodvisna** (vedenje, da se je zgodil vlom, ne vpliva na verjetje o dogodku strele) $P(V|St) = P(V), P(St|V) = P(St)$
  - vendar: poznavanje resničnosti tega, da se je sprožil *senzor* povzroči, da dogodka *vlom* in *strela* postaneta **odvisna**; ker sta oba vzroka za sproženje *senzorja*, velja, da resničnost enega zmanjšuje verjetnost drugega in obratno

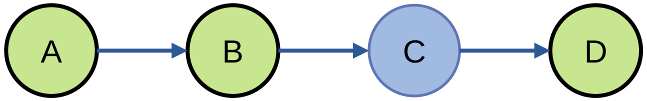


Odvisnosti v mreži

- **veriga**
  - *vlom* in *alarm* sta **odvisna**; poznavanje resničnosti enega od njiju vpliva na naše verjetje o resničnosti drugega
  - vendar: če vemo, da je resničen tudi *senzor*, postaneta *vlom* in *alarm* **neodvisna**: poznavanje resničnosti spremenljivke *alarm* ni pogojena s poznavanjem *vloma* in obratno
  - pravimo, da vozlišče senzor *blokira* vpliv vozlišča *vlom* na vozlišče *alarm*

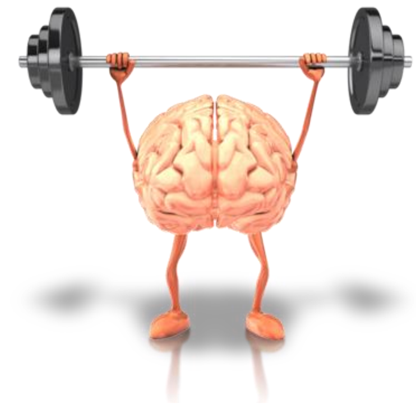


- pravilo lahko posplošimo na daljše verige:

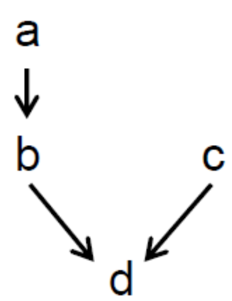


$$P(D|ABC) = P(D|C)$$

Vaja



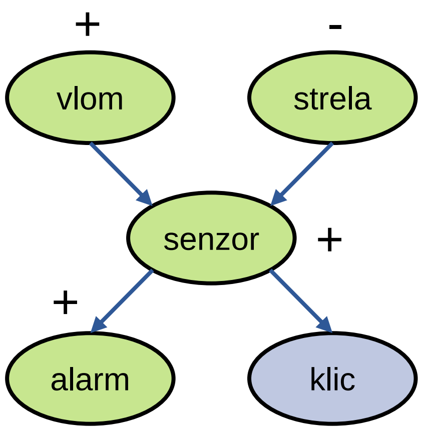
- podana je naslednja bayesovska mreža:



- a in c naj bosta redka dogodka in vse povezave v mreži naj predstavljajo vzročnost med dogodki (torej npr. $P(b|a) \gg P(b|\sim a)$). Brez računanja oceni relacije (<, > ali =) med naslednjimi verjetnostmi:
  a) $P(a)$: $P(a|c)$
  b) $P(a)$: $P(a|d)$
  c) $P(a|d)$: $P(a|cd)$
  d) $P(d|bc)$: $P(d|abc)$

Primer

- podobno sklepanje lahko uporabimo na našem primeru:
  - vlom je sam po sebi malo verjeten dogodek
  - denimo, da prejmemo opozorilni klic
  - zaradi prejetega klica se verjetnost proženja senzorja poveča
    (in ravno tako verjetnost alarma)
  - ker vlom sproža senzor, se poveča tudi verjetnost vlomu
  - ali: izvemo, da je doma bila nevihta s strelami; ker je strela možen vzrok za proženje senzorja, se verjetnost vloma zmanjša



Bayesovske mreže: odvisnosti, sklepanje