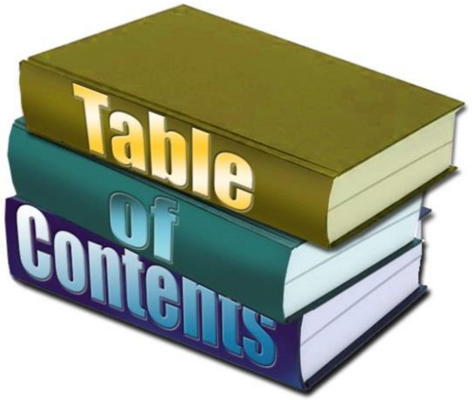
OSNOVE UMETNE INTELIGENCE

2018/19

bayesovske mreže:
- odvisnosti
- sklepanje

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

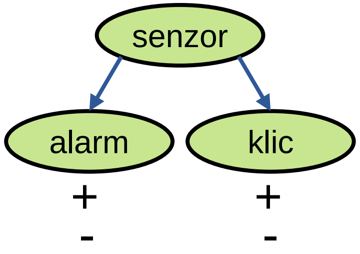
Pregled

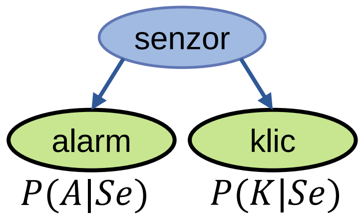


- bayesovske mreže
  - definicija
  - odvisnosti v bayesovski mreži
  - neodvisnosti v bayesovski mreži
  - verjetnostno sklepanje
  - ekvivalenca bayesovskih mrež

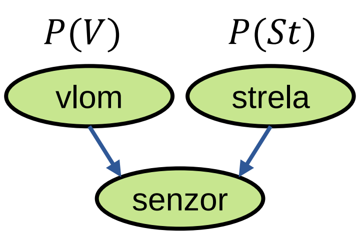
Odvisnosti v mreži

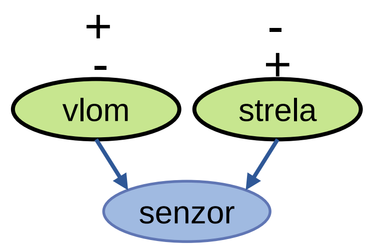
- **skupni prednik**:
  - *alarm* in *klic* sta **odvisna;** če vemo, da je eden od njiju resničen, vpliva to tudi na naše verjetje o resničnosti drugega (če se je sprožil alarm, se je verjetno izvedel tudi klic); $P(A|K) \neq P(A)$, $P(K|A) \neq P(K)$
  - vendar: poznavanje resničnosti prednika *senzor*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





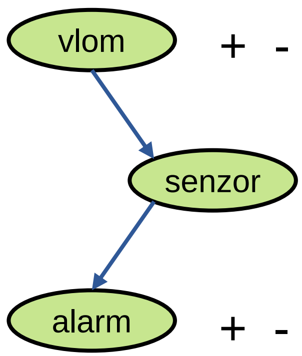
- **skupni naslednik**:
  - *vlom* in *strela* sta medseboj **neodvisna** (vedenje, da se je zgodil vlom, ne vpliva na verjetje o dogodku strele) $P(V|St) = P(V), P(St|V) = P(St)$
  - vendar: poznavanje resničnosti tega, da se je sprožil *senzor* povzroči, da dogodka *vlom* in *strela* postaneta **odvisna**;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

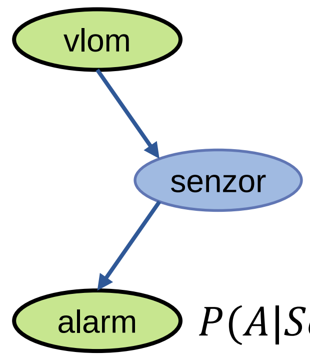




Odvisnosti v mreži

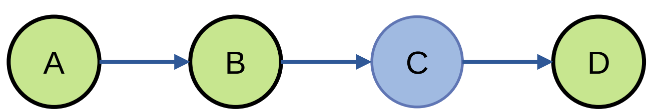
- **veriga**
  - *vlom* in *alarm* sta **odvisna**; poznavanje resničnosti enega od njiju vpliva na naše verjetje o resničnosti drugega
  - vendar: če vemo, da je resničen tudi *senzor*, postaneta *vlom* in *alarm* **neodvisna**: poznavanje resničnosti spremenljivke *alarm* ni pogojena s poznavanjem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





$P(A|Se)$

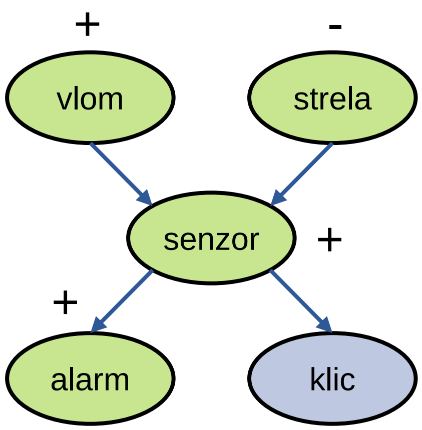
- pravilo lahko posplošimo na daljše verige:



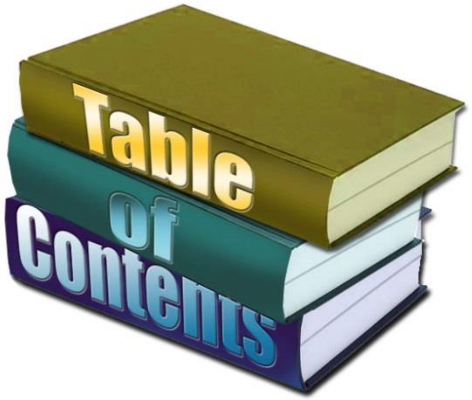
$$P(D|ABC) = P(D|C)$$

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



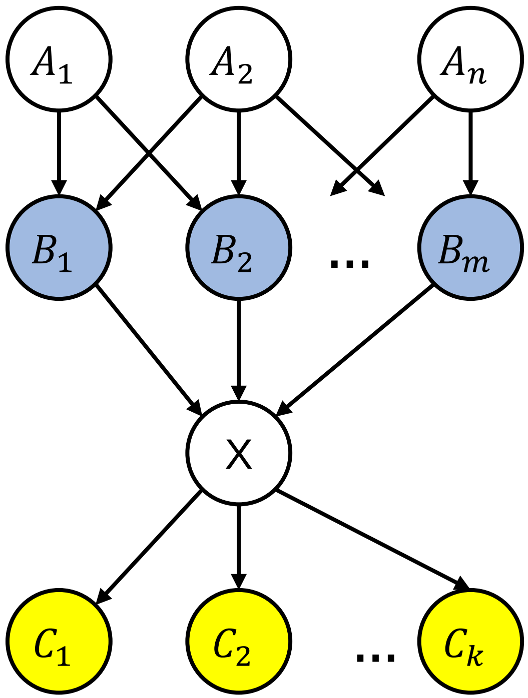
Pregled



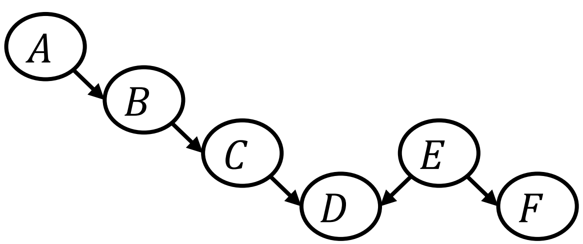
- bayesovske mreže
  - definicija
  - odvisnosti v bayesovski mreži
  - neodvisnosti v bayesovski mreži
  - verjetnostno sklepanje
  - ekvivalenca bayesovskih mrež

Neodvisnost v mreži

- dosedanje pravilo o verigah lahko posplošimo:
  - če so **podani starši** vozlišča X,
    je X **neodvisen samo od svojih nenaslednikov** (predhodnikov staršev)
    $P(X|A_1 \dots A_n B_1 \dots B_m C_1 \dots C_k) = P(X|B_1 \dots B_m C_1 \dots C_k)$



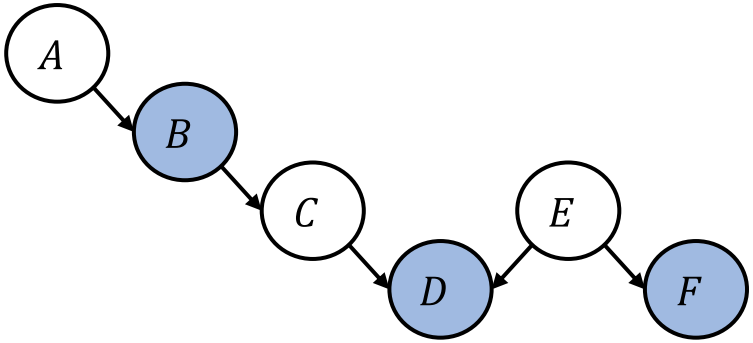
- pozor:
  - zgornje velja le, če so podani **vsi** starši $B_1, \dots, B_m$
  - zgornje velja le, če so podani **samo** starši $B_1, \dots, B_m$.
  - če so podana tudi druga vozlišča, je potrebno upoštevati tudi njihove neposredne ali posredne vplive
  - primer:



velja $P(C|AB) = P(C|B)$,
ne pa tudi $P(C|BDF) = P(C|BD)$!!! Zakaj?

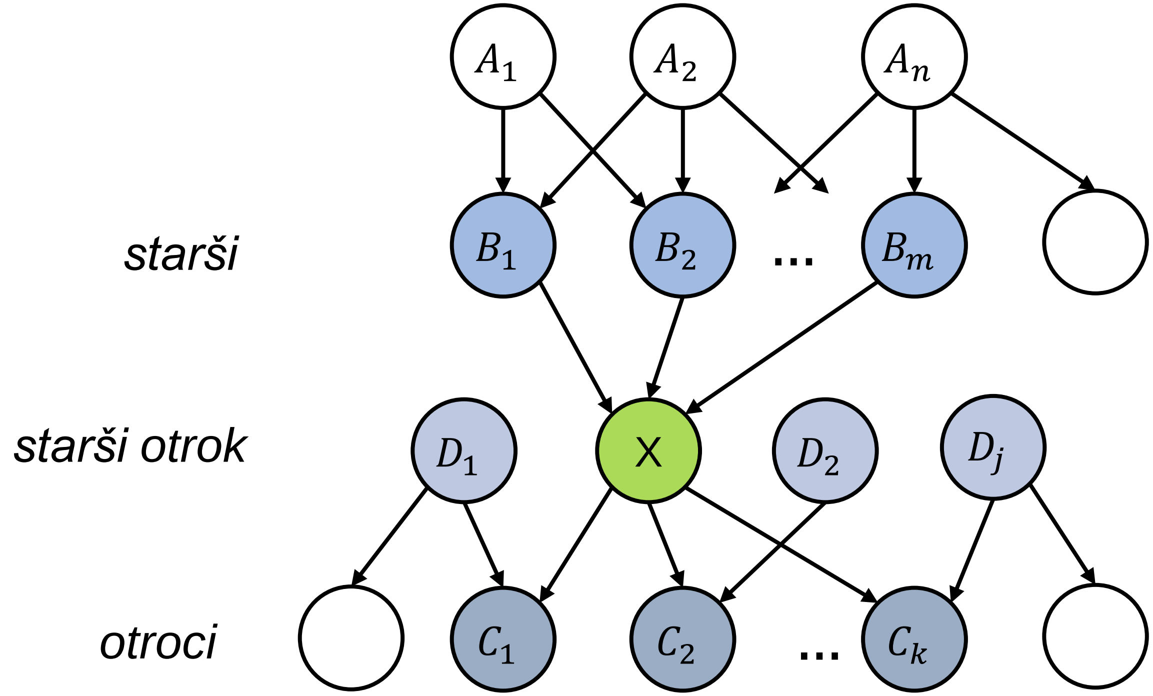
Neodvisnost v mreži

- $P(C|BDF) \neq P(C|BD)$
- čeprav je F nenaslednik od C, vpliva na to, da je E bolj verjeten in zato C manj verjeten.
- pravilo o neodvisnosti od nenaslednikov velja torej samo, če so podani izključno starši vozlišča C



Ovojnica Markova

- prejšnje pravilo lahko še bolj posplošimo
- **ovojnica Markova** (angl. *Markov blanket*) definira, od katerih vozlišč je odvisno opazovano vozlišče
- če so podani **starši, otroci in starši otrok**, je vozlišče X **neodvisno** od vseh ostalih vozlišč



d-ločevanje

- angl. *d-separation (direction-dependent separation*)
- še večja **posplošitev določanja neodvisnih vozlišč**
- pravilo: če sta A in B dve vozlišči (spremenljivki) v mreži, sta ti vozlišči **neodvisni**, če obstaja množica vozlišč E, ki d-ločuje A in B
- to pomeni, da velja:

$$P(AB|E) = P(A|E) \cdot P(B|E)$$

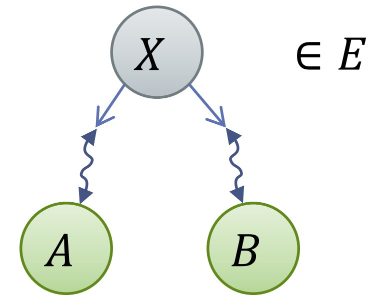
in

$$P(A|EB) = P(A|E)$$

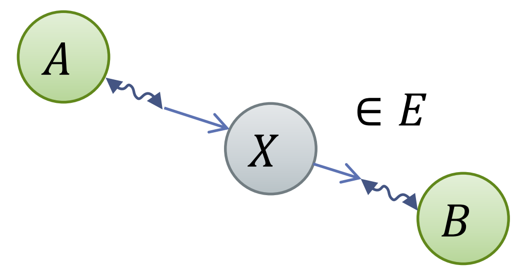
- kako najdemo množico E, ki d-ločuje vozlišči A in B ?

d-ločevanje

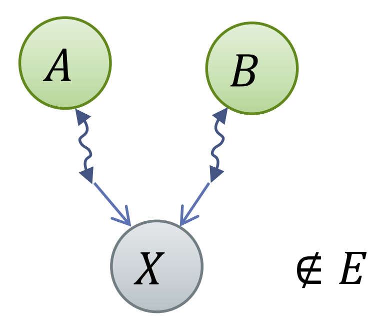
- kako najdemo množico E, ki d-ločuje vozlišči A in B ?
- množica E d-ločuje vozlišči A in B, če obstaja takšno podano vozlišče X na vsaki (neusmerjeni) poti med A in B, da blokira to pot. Vozlišče X blokira pot na enega od naslednjih načinov:



1. X je **divergentno vozlišče** (skupni vzrok, angl. *fork*) – iz njega kažeta povezavi v A in B. Tedaj velja $X \in E$.



2. X je **zaporedno vozlišče** (bolj neposreden vzrok za B kot za A, angl. *serial, chain*). Tedaj velja $X \in E$.



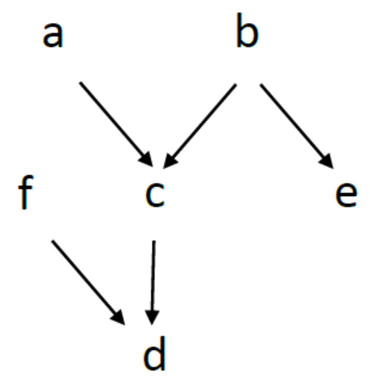
3. X je **konvergentno vozlišče** (skupna posledica, angl. *converging, collider*) – vanj kažeta povezavi iz A in B. Tedaj velja za X in za vse njegove naslednike, da $\notin E$.

Izpitna naloga

- 1. izpitni rok, 30. 1. 2018

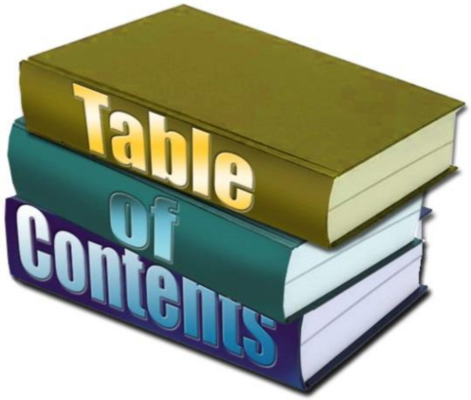
<u>4. NALOGA:</u>

Podana je Bayesovska mreža, ki je prikazana na sliki. Predpostavimo, da vse povezave v mreži predstavljajo pozitivno vzročnost med dogodki.
Odgovori na naslednja vprašanja:



a) Kolikšno število verjetnosti je potrebno podati, da je prikazana mreža dobro definirana? Kakšen je prihranek v številu verjetnosti glede na število podatkov v popolni verjetnostni porazdelitvi?
b) Katera vozlišča so v ovojnici Markova vozlišča c? Zapiši, katera neodvisnost izhaja iz pravila o ovojnici Markova za vozlišče c in kateri pogoji morajo biti za to izpolnjeni?
c) Če je možno, čim bolj poenostavi pogojni del v izrazu: P(f|abe). Odgovor utemelji.
d) Če je možno, čim bolj poenostavi pogojni del v izrazu: P(f|abde). Odgovor utemelji.
e) Izrazi verjetnost P(d|c) z verjetnostmi, ki so podane v mreži.
f) Zapiši vse množice vozlišč, ki d-ločujejo vozlišči a in e.

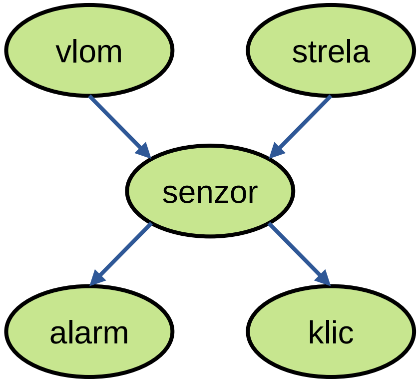
Pregled



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

Primeri možnih vprašanj

- $P(V \mid K) = ?$
- $P(V \mid K\ St) = ?$
- $P(V \mid St) = ?$
- $P(V \mid A \sim St) = ?$
- $P(A \sim K \mid V) = ?$



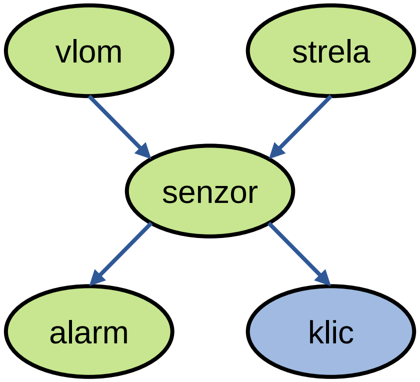
- možni sta dve smeri sklepanja:
  - **vzročno** (od vzrokom k posledicam):
    $P(A \mid V\ St) = ?$
  - **diagnostično** (od posledic – naslednikov k vzrokom – predhodnikom); uporabimo Bayesovo formulo
    $P(V \mid A) = ?$
    $P(V|A) = P(V)\frac{P(A|V)}{P(A)}$

Pravila verjetnostnega sklepanja

1. Verjetnost **konjunkcije**:
   $P(X_1 X_2|C) = P(X_1|C) \cdot P(X_2|X_1 C)$
2. Verjetnost **gotovega** dogodka:
   $P(X|\ldots X \ldots) = 1$
3. Verjetnost **nemogočega** dogodka:
   $P(X|\ldots \sim X \ldots) = 0$
4. Verjetnost **negacije**:
   $P(\sim X|C) = 1 - P(X|C)$
5. Če pogoj **vključuje naslednika** Y (vzvratno sklepanje), uporabi posplošeno Bayesovo formulo:
   $P(X|YC) = P(X|C) \cdot \frac{P(Y|XC)}{P(Y|C)}$
6. Če pogoj C **ne vključuje naslednika** od X, potem:
   a) če X **nima** staršev: $P(X|C) = P(X)$
   b) če **ima** X starše P: $P(X|C) = \sum_{s \in stanja\ staršev(X)} P(X|S) \cdot P(S|C)$

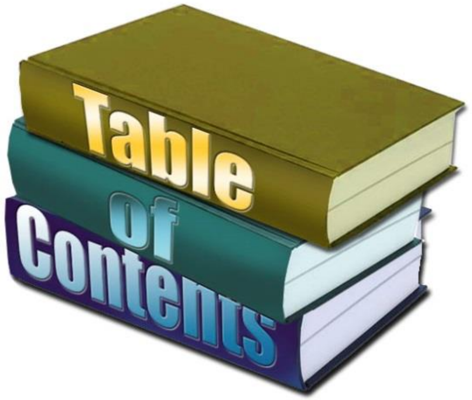
Primer

Kolikšna je verjetnost, da se je zgodil vlom, če smo prejeli alarmni klic?
$P(V|K) = ?$



- uporabimo Bayesovo formulo:
  $P(V|K) = P(V) \cdot \frac{P(K|V)}{P(K)}$
- $P(V)$ je podan
- $P(K|V) = ?$
  pogoj ne vključuje naslednika od K, uporabimo pravilo 6b:
  $P(K|V) = P(K|Se) \cdot P(Se|V) + P(K|{\sim}Se) \cdot P({\sim}Se|V)$
- $P(Se|V) = ?$
- $P(K) = ?$
  ravno tako uporabimo pravilo 6b:
  $P(K) = P(K|Se) \cdot P(Se) + P(K|{\sim}Se) \cdot P({\sim}Se)$

Pregled

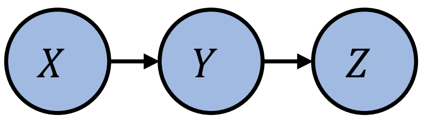


- bayesovske mreže
  - definicija
  - odvisnosti v bayesovski mreži
  - neodvisnosti v bayesovski mreži
  - verjetnostno sklepanje
  - ekvivalenca bayesovskih mrež

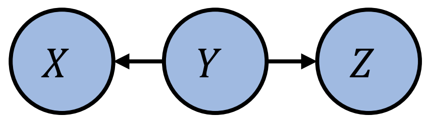
Ekvivalenca mrež

Za dve mreži pravimo, da sta ekvivalentni, če je z verjetnostmi ene mreže možno izraziti vse verjetnosti druge mreže, tako da mreži še vedno izražata iste odvisnosti

- ideja: vzročno ali diagnostično smer sklepanja lahko obrnemo z Bayesovo formulo
- primeri ekvivalentnih mrež:



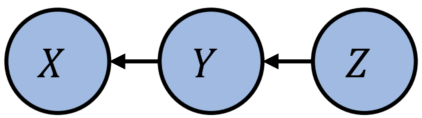
$\boldsymbol{P(X), P(Y|X), P(Z|Y)}$



$\boldsymbol{P(Y), P(X|Y), P(Z|Y)}$

$P(X) = P(X|Y) \cdot P(Y) + P(X|{\sim}Y) \cdot P({\sim}Y)$

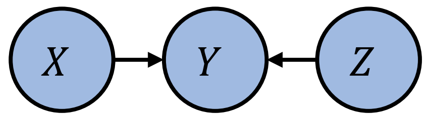
$P(Y|X) = P(Y) \cdot P(X|Y) / P(X)$



$\boldsymbol{P(Z), P(X|Y), P(Z|Y)}$

izrazi za vajo!

- primer neekvivalentne mreže zgornjim:

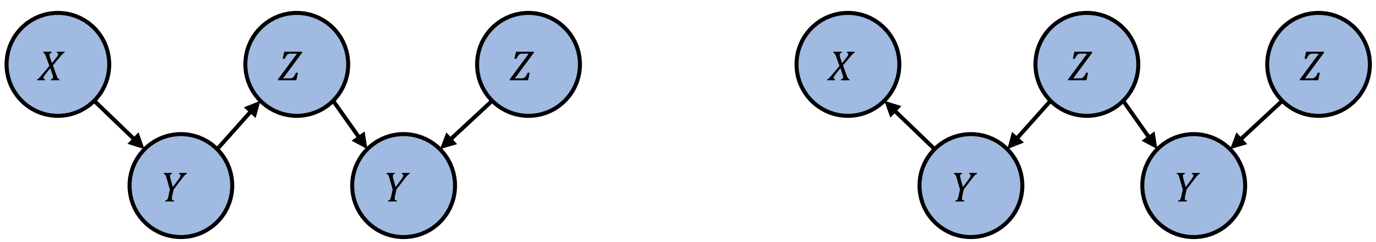


$\boldsymbol{P(X), P(Z), P(Y|XZ)}$

Ekvivalenca mrež

formalno: mreži sta I-ekvivalentni (I-equivalence, independence-equivalence), če:

- imata enako strukturo (ob ignoriranju usmerjenosti povezav),
- imata ista konvergentna vozlišča.



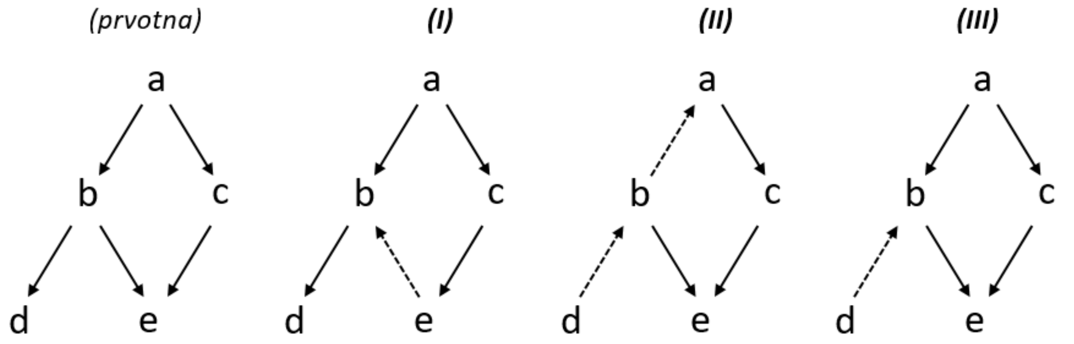
- opomba: obstajajo tudi ekvivalentne bayesovske mreže, ki ne ustrezajo zgornjim zahtevam (npr. polni grafi – drugačna konvergentna vozlišča, ni neodvisnosti)

Izpitna naloga

- 2. izpitni rok, 15. 2. 2018

3. NALOGA (25t):

Na spodnji sliki je podana prvotna bayesovska mreža (skrajno levo) in tri mreže, ki so izpeljane iz prvotne (označene z rimskimi številkami). Vsaka spremenjena mreža ima drugače usmerjeno eni ali dve povezavi (spremembe so označene s črtkano puščico). Predpostavimo, da so s prvotno mrežo podane vse verjetnosti, ki so potrebne za definiranje te mreže. Za dve mreži pravimo, da sta *ekvivalentni*, če je z verjetnostmi ene mreže možno izraziti vse verjetnosti druge mreže, tako da mreži še vedno izražata iste odvisnosti.



a) **(4t) Od treh mrež na desni strani je natanko ena ekvivalentna prvotni mreži – katera, zakaj?**
b) **(7t) Izrazi verjetnosti v ekvivalentni spremenjeni mreži (odgovor iz prejšnje točke) z verjetnostmi iz prvotne mreže.**
c) **(7t) Izrazi verjetnost P(abcd~e) z verjetnostmi, ki so podane s prvotno mrežo.**
d) **(7t) Katere množice vozlišč d-ločujejo vozlišči c in d v prvotni mreži?**

Predstavitev znanja