# Bayesove mreže

## Poglavlje 14.1–3

Prema slajdovima Stuarta Russella (hvala)!

# Sadržaj

- ♢ Sintaksa
- ♢ Semantika
- ♢ Parameterizirane distribucije

# Bayesove mreže

Jednostavna, grafička notacija za tvrdnje o uvjetnoj nezavisnosti i za sažetu specifikaciju zajedničke (združene) distribucije vjerojatnosti.

Sintaksa Bayesove mreže:

– skup čvorova, jedan po svakoj varijabli (varijabla = ime čvora),
– usmjereni, aciklički graf (engl. "directed acyclic graph", DAG)
  smjer veze ≈ "direktni utjecaj" roditelja na dijete,
– uvjetna distribucija za svaki čvor, za dane njegove roditelje:
  $\mathbf{P}(X_i \,|\, Parents(X_i))$ — kvantificira utjecaj roditelja na dijete.

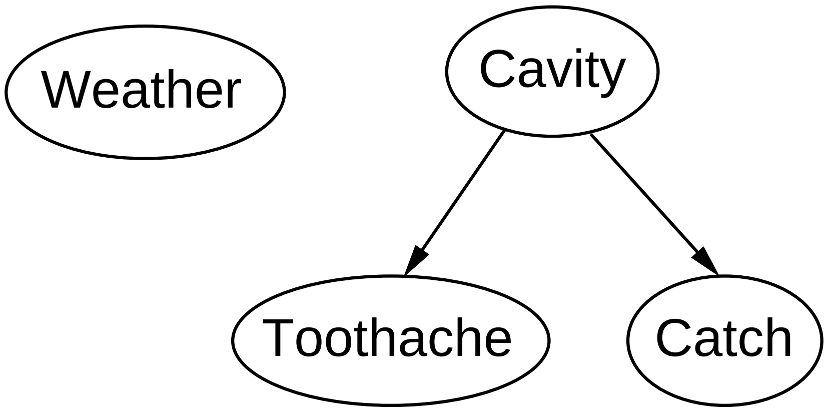
U diskretnom slučaju, uvjetna distribucija je reprezentirana

tablicom uvjetne vjerojatnosti (engl. "conditional prob. table", CPT)

koja daje distribuciju preko svih vrijednosti $X_i$, za svaku kombinaciju vrijednosti roditelja.

# Jednostavni primjer

Topologija mreže (grafa) — čvorovi i veze među njima,
zadaje odnose uvjetne nezavisnosti u domeni:




*Weather* je nezavisna o svim preostalim varijablama — nema veza.

Veze označavaju da je *Cavity* direktni uzrok za *Toothache* i *Catch*.

*Toothache* i *Catch* su uvjetno nezavisne za dan (poznat) *Cavity*
nema direktne veze između njih.

# Primjer

Priča: Imam novi protuprovalni alarm kod kuće. On je dosta pouzdan u otkrivanju provale, ali katkad reagira i na manje potrese (Kalifornija).

Imam dvoje susjeda, Johna i Mary, koji su obečali da će me nazvati na posao, ako čuju alarm.

John gotovo uvijek nazove kad čuje alarm, ali ga ponekad zamijeni sa zvonjavom telefona, pa me i tad nazove.

S druge strane, Mary voli glasnu muziku i često uopće ne čuje alarm.

Problem: Ako znam tko je ili nije zvao (dokazi), želim znati
kolika je vjerojatnost stvarne provale.

Na primjer, John je zvao, a Mary nije.

# Primjer — nastavak

Varijable: $Burglary$, $Earthquake$, $Alarm$, $JohnCalls$, $MaryCalls$

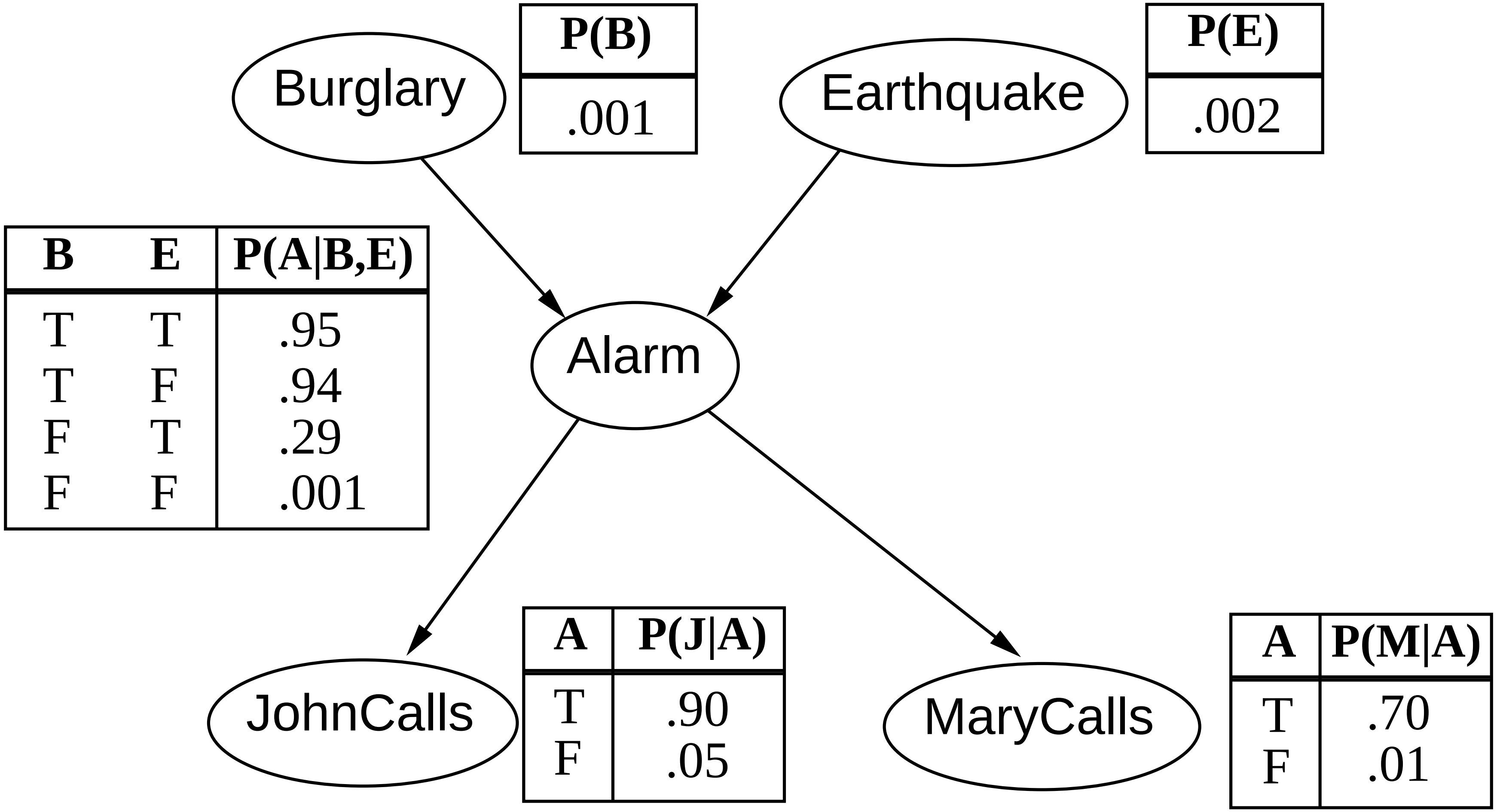
Katkad koristimo skraćene oznake prvim slovom: $B$, $E$, $A$, $J$, $M$

Topologija mreže reflektira "kauzalno" = uzročno znanje:
- Provalnik može aktivirati alarm
- Potres može aktivirati alarm
- Alarm može prouzročiti Johnov poziv
- Alarm može prouzročiti Maryin poziv

Uočiti:
- Poziv ovisi samo o alarmu, a ne o stvarnoj provali ili potresu
- Ako čuju alarm, pozivi su nezavisni (ne dogovaraju se o tome).

Evo i mreže, zajedno s tablicama uvjetnih vjerojatnosti u čvorovima.

# Primjer — Bayesova mreža




| 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 |

# Zapis vjerojatnosti preko CPT

Tablice uvjetnih vjerojatnosti (CPT) u čvorovima sadrže:

– po jedan redak za svaku moguću kombinaciju vrijednosti svih roditeljskih čvorova (tzv. "uvjetni slučaj"),

– po jedan stupac za svakog roditelja (prednji dio tablice),

– po jedan stupac za svaku moguću vrijednost varijable u tom čvoru (stražnji dio tablice = popis vjerojatnosti).

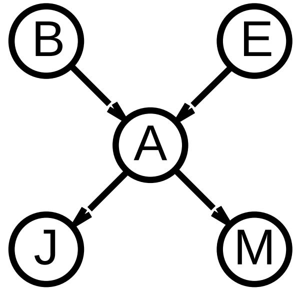
U čvoru bez roditelja, CPT ima samo jedan redak — to su početne (a priori) vjerojatnosti za svaku moguću vrijednost varijable.

Uočiti: zbroj svakog retka tablice (u stražnjem dijelu) mora biti 1, zato stupac za "zadnju" vrijednost varijable smijemo ispustiti!

Tablični prikaz uvjetnih vjerojatnosti je prikladan za diskretne varijable.

Postoje i drugi prikazi — neki su prikladni i za neprekidne varijable.

# Kompaktnost — sažetost zapisa




CPT za Booleovu varijablu $X_i$ s $k$ Booleovih roditelja ima $2^k$ redaka — za sve kombinacije vrijednosti roditelja.

Svaki redak zahtijeva samo jedan broj $p$ za $X_i = true$ (broj za $X_i = false$ je upravo $1 - p$).

Sažetost zapisa potpune združene distribucije vjerojatnosti preko CPT:
ako imamo $n$ Booleovih varijabli (radi jednostavnosti)
i svaka varijabla/čvor ima najviše $k$ roditelja,
cjelokupna mreža zahtijeva $O(n \cdot 2^k)$ brojeva.

Ovo raste linearno s $n$, nasuprot $O(2^n)$ za potpunu distribuciju.

Za provalničku mrežu, imamo samo $1 + 1 + 4 + 2 + 2 = 10$ brojeva (preostalih $10$ su $1 - p$), nasuprot punih $2^5 - 1 = 31$.

Za $n = 30$ i $k = 5$, dovoljno je $960$ brojeva (prema $2^{30}$).

# Semantika Bayesovih mreža

Do sada — opisali što je mreža (kako izgleda), ali ne i njezino značenje.

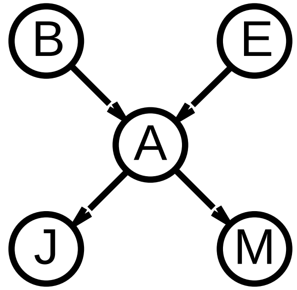
Postoje dva pogleda na semantiku — naravno, ekvivalentna.

Prvi = globalni: mreža je prikaz združene distribucije vjerojatnosti
pomaže u tome kako treba konstruirati mreže.

Drugi = lokalni: mreža je prikaz skupa činjenica o uvjetnoj nezavisnosti
pomaže u projektiranju postupaka zaključivanja.

Na razini sintakse, Bayesova mreža je
usmjereni aciklički graf (DAG),
s nekim numeričkim parametrima vezanim za svaki čvor.

# Globalna semantika




Globalna semantika definira potpunu združenu distribuciju kao produkt odgovarajućih numeričkih parametara iz zadanih tablica u čvorovima.

Svaka pojedina vrijednost u združenoj distribuciji je vjerojatnost konjunkcije konkretnih vrijednosti svake varijable, u oznaci $P(X_1 = x_1, \ldots, X_n = x_n) = P(x_1, \ldots, x_n)$.

Globalna semantika definira

$$P(x_1, \ldots, x_n) = \text{def.} = \prod_{i=1}^{n} \theta(x_i \,|\, parents(X_i)),$$

gdje je $\theta(x_i \,|\, parents(X_i))$ odgovarajući parametar iz tablice u čvoru $X_i$, za dane vrijednosti čvora ($x_i$) i njegovih roditelja ($parents(X_i)$).

# Globalna semantika — nastavak

Iz te definicije slijedi da ovi parametri moraju biti upravo odgovarajuće uvjetne vjerojatnosti $P(x_i \,|\, parents(X_i))$, pa vrijedi

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \,|\, parents(X_i)).$$

Dakle, tablica u svakom čvoru je lokalna uvjetna distribucija

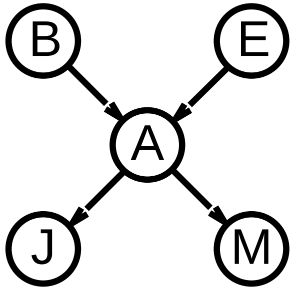
$$\mathbf{P}(X_i \,|\, Parents(X_i)).$$

Drugim riječima, početna interpretacija je korektna.

Ovo je "numerička" semantika — faktorizacija združene distribucije.

Može se koristiti i za zaključivanje — zbrajanjem (kao prije). Međutim, ima i boljih načina.

# Globalna semantika — ilustracija




Globalna semantika definira potpunu združenu distribuciju kao produkt lokalnih uvjetnih distribucija:

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \,|\, parents(X_i)).$$

Nađimo vjerojatnost da se alarm aktivirao, s tim da nije bilo ni provale ni potresa, a zvali su i John i Mary.

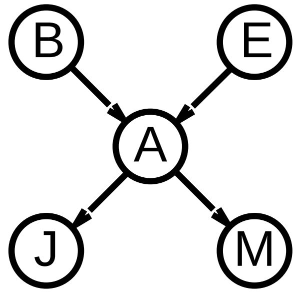
U skraćenim oznakama, tražimo $P(j, m, a, \neg b, \neg e)$.

Zapisom preko lokalnih uvjetnih vjerojatnosti i množenjem vrijednosti iz tablica, dobivamo

$$\begin{aligned} P(j, m, a, \neg b, \neg e) &= P(j|a)\, P(m|a)\, P(a \,|\, \neg b, \neg e)\, P(\neg b)\, P(\neg e) \\ &= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 \\ &\approx 0.000628 \end{aligned}$$

# Lokalna semantika

Lokalna ili "topološka" semantika — kako je opisana u strukturi grafa: svaki čvor je uvjetno nezavisan od svojih nesljedbenika (nepotomaka) ako su dani njegovi roditelji (kao uvjet/dokaz).




Na primjer, u provalničkoj Bayesovoj mreži, $JohnCalls$ ne ovisi o $Burglary$, $Earthquake$ i $MaryCalls$, uz uvjet da znamo vrijednost za $Alarm$.

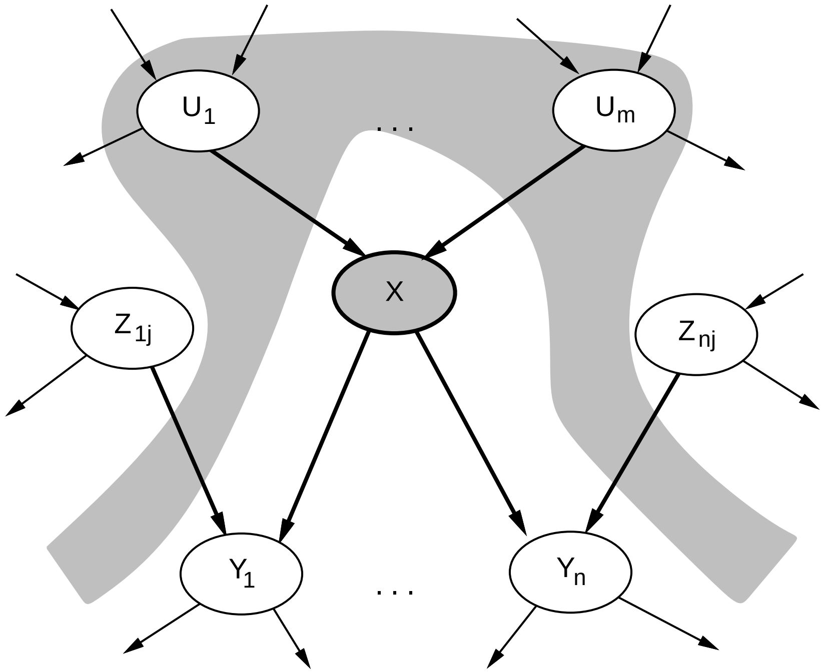
U lokalnoj semantici, osim tih činjenica o uvjetnoj nezavisnosti, tablica u svakom čvoru je, po definiciji, lokalna uvjetna distribucija vjerojatnosti $\mathbf{P}(X_i \,|\, Parents(X_i))$.

Teorem: Lokalna semantika $\iff$ globalna semantika, tj. potpunu združenu distribuciju (po točkama) dobivamo kao malo prije

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \,|\, parents(X_i)).$$

# Lokalna semantika — nastavak

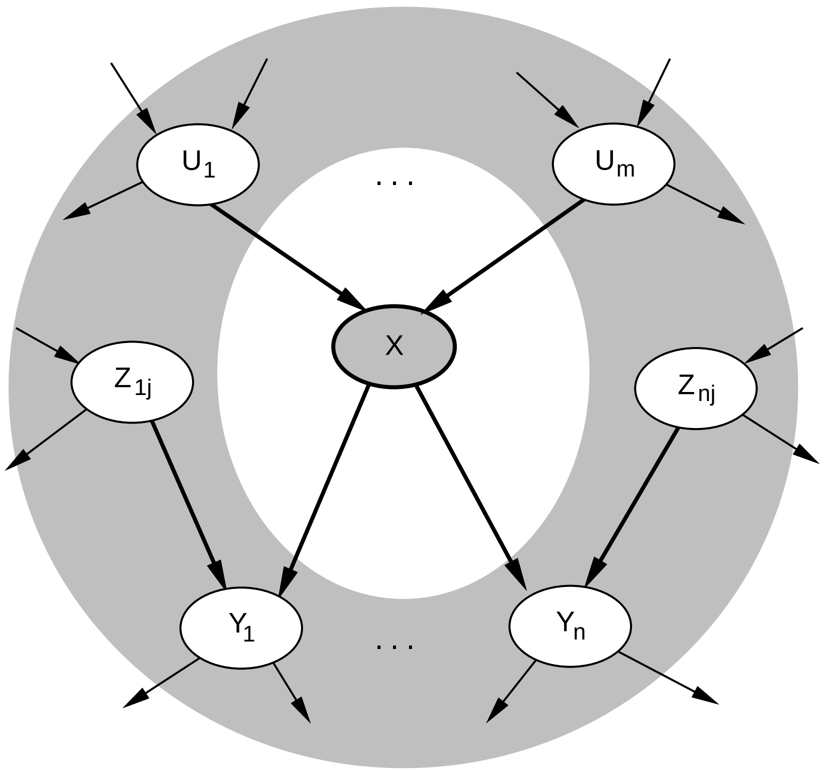
Lokalna ili "topološka" semantika — kako je opisana u strukturi grafa:
svaki čvor je uvjetno nezavisan od svojih nesljedbenika (nepotomaka) ako su dani njegovi roditelji (kao uvjet/dokaz).




Na primjer, $X$ je uvjetno nezavisan od nepotomaka $Z_{ij}$, uz uvjet svojih roditelja $U_i$ — prikazanih u sivoj zoni.

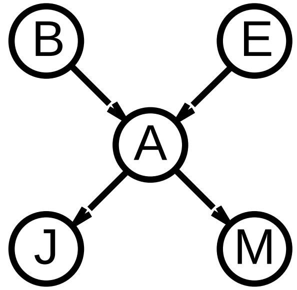
# Topološka semantika i Markovljev pokrivač

Topološka semantika implicira još jedno važno svojstvo nezavisnosti:
svaki čvor je uvjetno nezavisan od svih ostalih čvorova u mreži, ako su dani (kao uvjet) njegovi roditelji + djeca + roditelji djece, tj. njegov Markovljev pokrivač — za $X$, označen sivo na slici



# Primjer i grafovska interpretacija

Na primjer, ako znamo $Alarm$ i $Earthquake$, onda je $Burglary$ uvjetno nezavisan od $JohnCalls$ i $MaryCalls$.




Intuitivno "objašnjenje" nezavisnosti i uvjetne nezavisnosti na "grafovskom" nivou — preko puteva u DAG:

nezavisnost dva čvora = ne postoji put između njih,

uvjetna nezavisnost $X$ i $Y$, uz poznati dokaz $Z$
= dokaz $Z$ prekida ili siječe put između $X$ i $Y$, i to neovisno o smjeru puteva/strelica — kao da su neusmjereni putevi.

Smjer strelica označava samo kauzalnost veze, a pojam nezavisnosti je simetričan!

# Nezavisnost i uvjetna nezavisnost — grafovski

Proširenje ove interpretacije na skupove varijabli (čvorova) $\mathbf{X}$, $\mathbf{Y}$ i $\mathbf{Z}$ je očito.

Postoji opći "topološki" kriterij (algoritam) zvan **d–separacija**
za otkrivanje je su li dva skupa čvorova $\mathbf{X}$ i $\mathbf{Y}$ uvjetno nezavisna obzirom na dani treći skup $\mathbf{Z}$.

Za kratki opis, pogledati
MO-Web za UI, Predavanje 7 = Bayesove mreže, str. 27–28.

# Konstrukcija Bayesovih mreža

Traži se metoda za konstrukciju Bayesove mreže tako da

– dobivena združena distribucija dobro prikazuje domenu (korektna globalna semantika),

– topologija mreže korektno odražava relacije uvjetne nezavisnosti među varijablama (korektna lokalna semantika).

Podloga je lančano pravilo za distribucije, pisano unatrag po varijablama

$$\mathbf{P}(X_1, \ldots, X_n) = \prod_{i=1}^{n} \mathbf{P}(X_i \,|\, X_{i-1}, \ldots, X_1).$$

Usporedba s relacijom iz globalne semantike

$$\mathbf{P}(X_1, \ldots, X_n) = \prod_{i=1}^{n} \mathbf{P}(X_i \,|\, Parents(X_i))$$

daje da za svaku varijablu $X_i$ u mreži treba vrijediti

$$\mathbf{P}(X_i \,|\, X_{i-1}, \ldots, X_1) = \mathbf{P}(X_i \,|\, Parents(X_i)).$$

Ova jednakost osigurava = garantira globalnu semantiku.

# Konstrukcija Bayesovih mreža — nastavak

Da bismo zadovoljili zadnju jednakost

$$\mathbf{P}(X_i \,|\, X_{i-1}, \ldots, X_1) = \mathbf{P}(X_i \,|\, Parents(X_i)),$$

za početak, treba biti

$$Parents(X_i) \subseteq \{X_{i-1}, \ldots, X_1\}.$$

Dakle, čvorove (varijable) $X_i$ treba numerirati u poretku koji je konzistentan s parcijalnim uređajem implicitno zadanim u strukturi grafa — odražava kauzalne veze.

Medjutim, to nije dovoljno za jednakost — fali još nezavisnost.

Bayesova mreža je korektna reprezentacija domene, samo ako je svaki čvor uvjetno nezavisan od svojih ostalih prethodnika (u poretku čvorova), uz uvjet da znamo njegove roditelje (kao dokaze).

# Konstrukcija Bayesovih mreža — skica algoritma

Prethodni uvjet možemo zadovoljiti na sljedeći način:

1. Čvorovi: izaberi neki redoslijed (uređaj) varijabli $X_1, \ldots, X_n$
   Bilo koji poredak je načelno dobar, ali mreža je kompaktnija, ako poštujemo kauzalni poredak — uzroci prethode efektima.

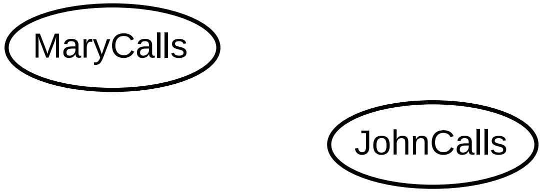
2. Veze: Za $i = 1$ do $n$
   dodaj $X_i$ u mrežu;
   odaberi minimalan skup roditelja od $X_i$ između $X_1, \ldots, X_{i-1}$
   takav da vrijedi $\mathbf{P}(X_i \,|\, Parents(X_i)) = \mathbf{P}(X_i \,|\, X_1, \ldots, X_{i-1})$
   dodaj u mrežu vezu od svakog roditelja na $X_i$;
   za čvor $X_i$, spremi CPT koja sadrži $\mathbf{P}(X_i \,|\, Parents(X_i))$.

Ova metoda garantira da nema ciklusa — dobivamo DAG.

Intuitivno, roditelji čvora $X_i$ su oni čvorovi iz $X_1, \ldots, X_{i-1}$ koji izravno utječu na $X_i$ (izravna kauzalnost) — ali nije nužno!

# Primjer — neprirodan poredak

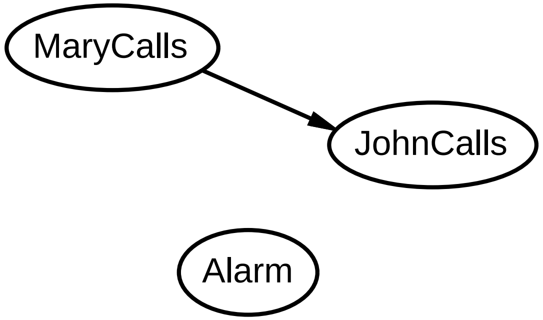
Pretpostavimo da smo uzeli poredak $M$, $J$, $A$, $B$, $E$




$P(J|M) = P(J)$?

# Primjer — neprirodan poredak

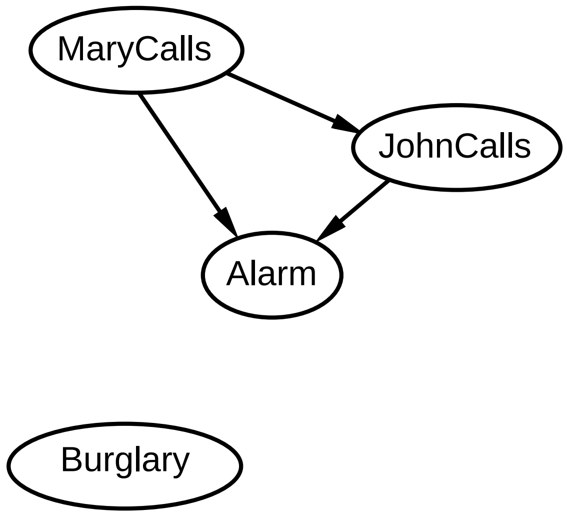
Pretpostavimo da smo uzeli poredak $M$, $J$, $A$, $B$, $E$




$P(J|M) = P(J)$? Ne
$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$?

# Primjer — neprirodan poredak

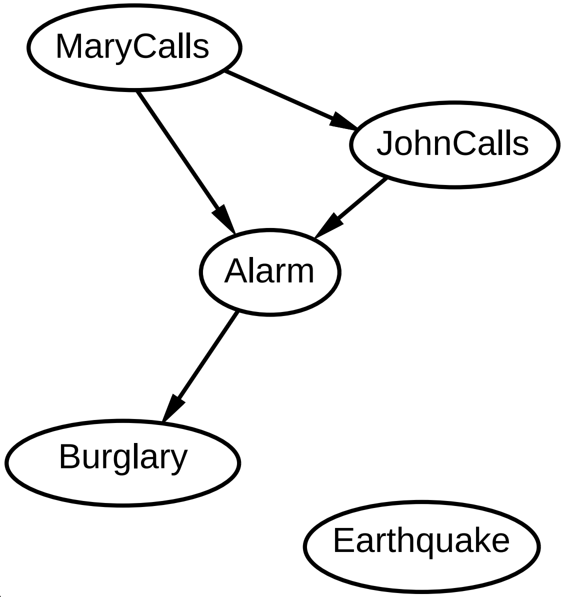
Pretpostavimo da smo uzeli poredak $M$, $J$, $A$, $B$, $E$




$P(J|M) = P(J)$? Ne
$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$? Ne
$P(B|A,J,M) = P(B|A)$?
$P(B|A,J,M) = P(B)$?

# Primjer — neprirodan poredak

Pretpostavimo da smo uzeli poredak $M$, $J$, $A$, $B$, $E$




$P(J|M) = P(J)$? Ne

$P(A|J, M) = P(A|J)$? $P(A|J, M) = P(A)$? Ne

$P(B|A, J, M) = P(B|A)$? Da

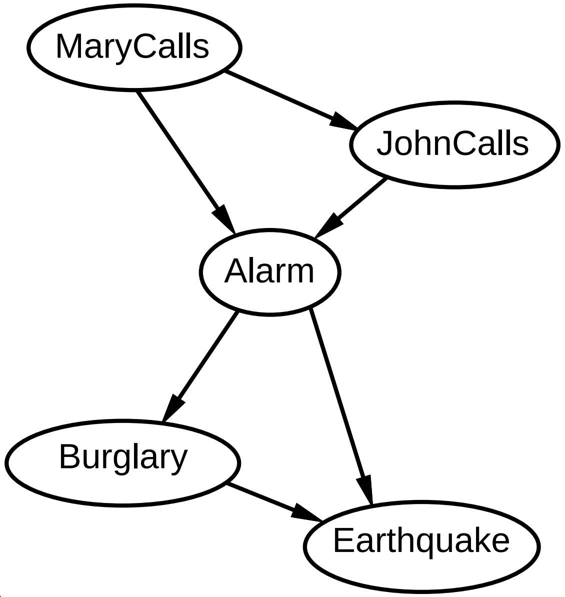
$P(B|A, J, M) = P(B)$? Ne

$P(E|B, A, J, M) = P(E|A)$?

$P(E|B, A, J, M) = P(E|A, B)$?

# Primjer — neprirodan poredak

Pretpostavimo da smo uzeli poredak $M$, $J$, $A$, $B$, $E$




$P(J|M) = P(J)$? Ne

$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$? Ne

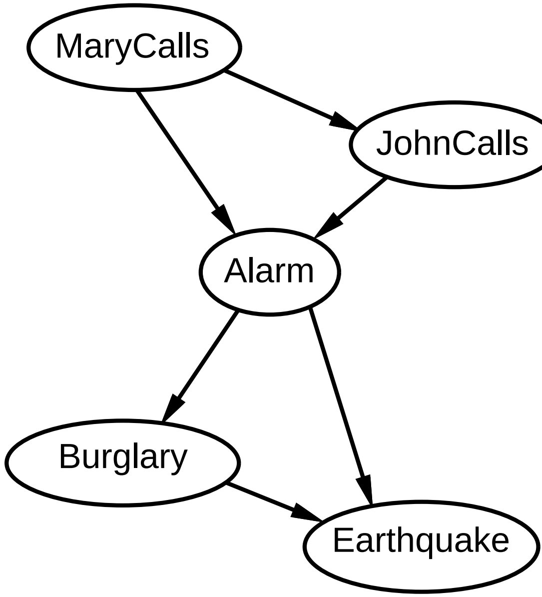
$P(B|A,J,M) = P(B|A)$? Da

$P(B|A,J,M) = P(B)$? Ne

$P(E|B,A,J,M) = P(E|A)$? Ne

$P(E|B,A,J,M) = P(E|A,B)$? Da

# Primjer — neprirodan poredak, komentar




Odlučivanje o uvjetnoj nezavisnosti je teško u nekauzalnim smjerovima.

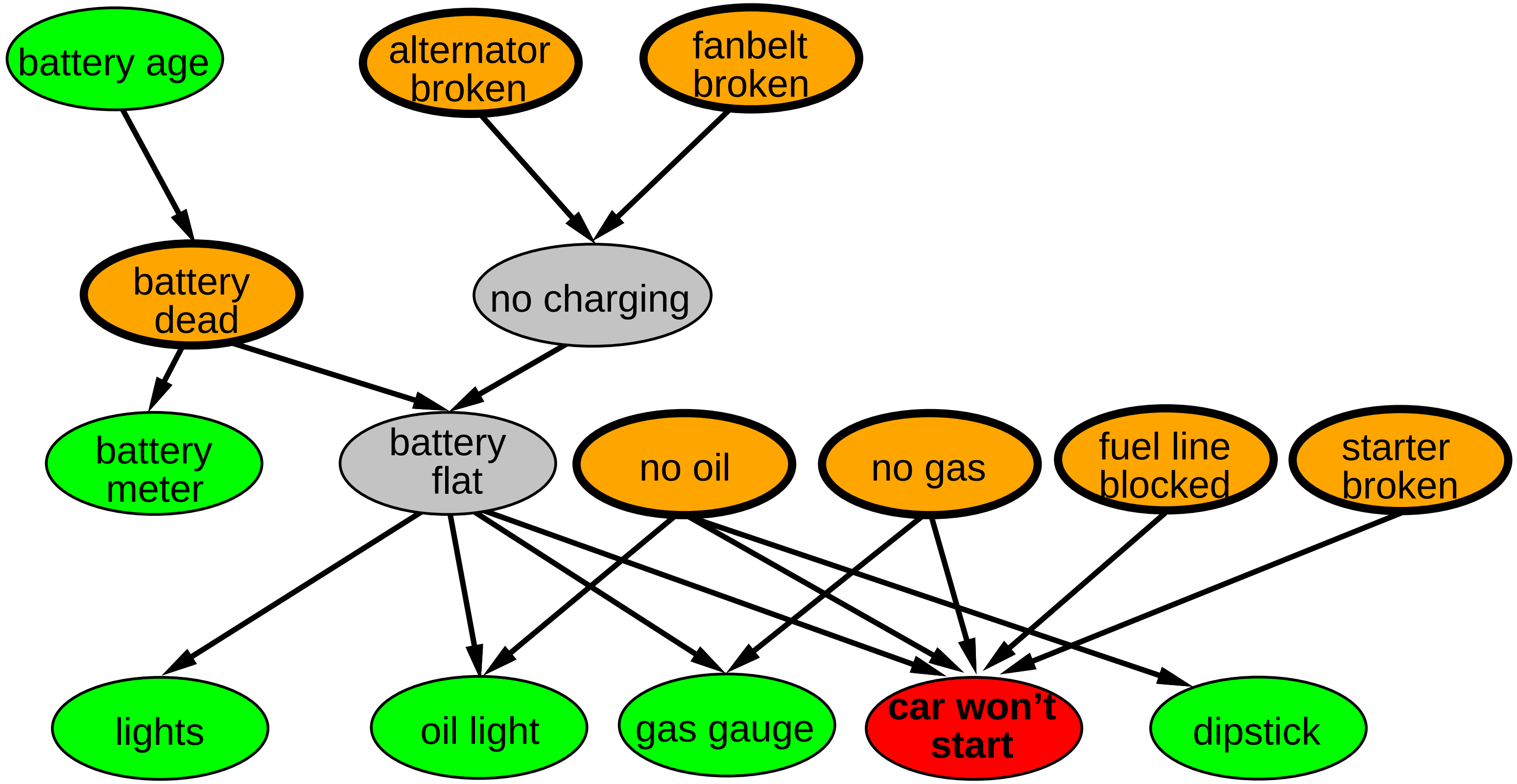
(Za ljude su kauzalni modeli i uvjetna nezavisnost čvrsto vezani!)

Procjena uvjetne vjerojatnosti je teška u nekauzalnim smjerovima.

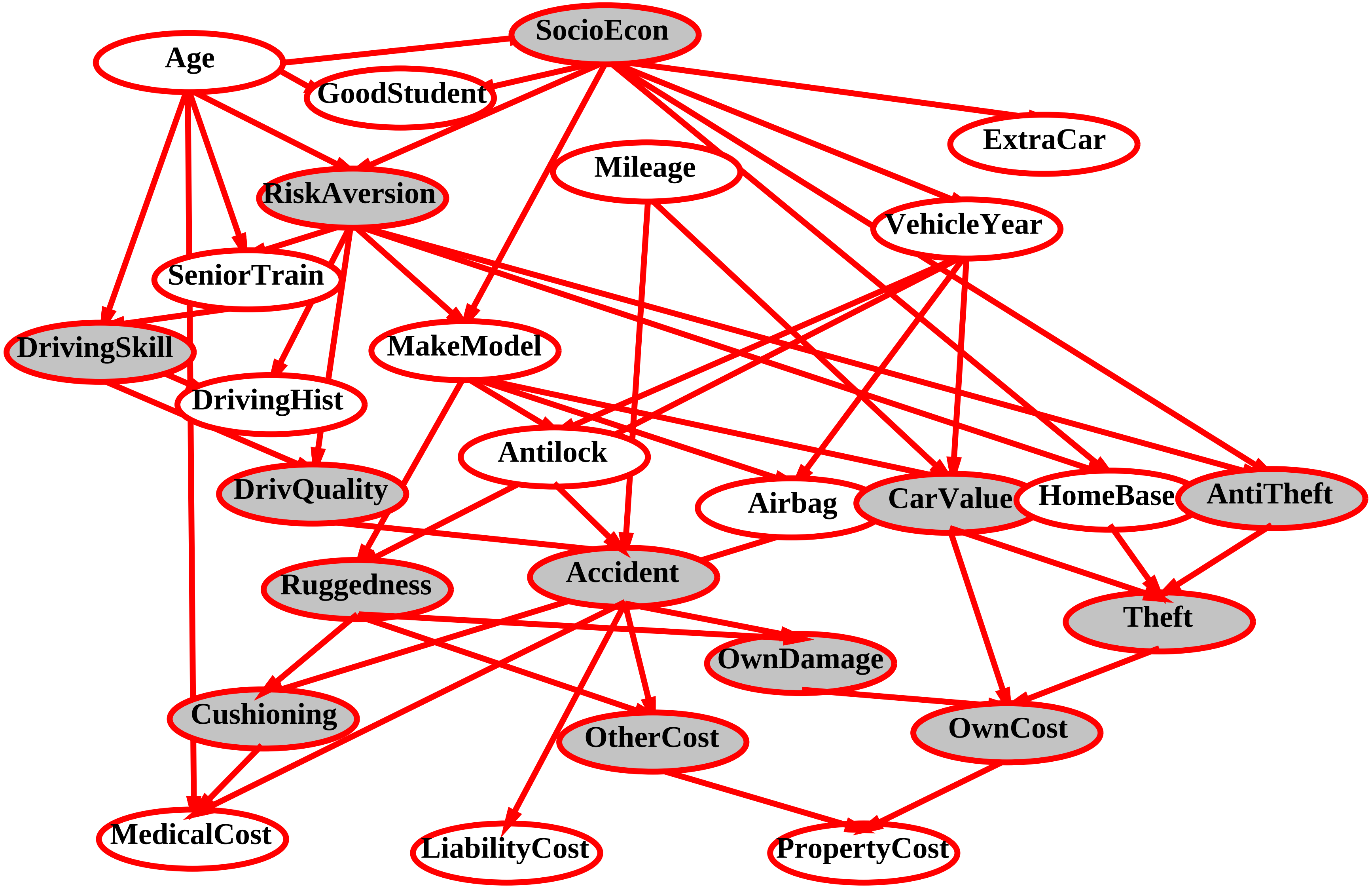
Mreža je manje kompaktna: $1+2+4+2+4=13$ brojeva je potrebno.

# Primjer: dijagnoza kvara automobila

Početni "dokaz": automobil ne želi startati.
Test–varijable (zeleno), "pokvareno, popravi" varijable (narančasto),
Skrivene varijable (sivo) daju rijetku strukturu, reduciraju parametre.



# Primjer: osiguranje automobila



# Sažeti prikaz uvjetnih distribucija

Problem: Veličina CPT raste eksponencijalno ($2^k$) s brojem roditelja $k$.
Ako su roditelj ili dijete neprekidne, CPT postaje beskonačna.

Rješenje: kanonske distribucije koje su sažeto (kompaktno) definirane,
prema nekom standardnom "predlošku",
eventualno uz zadavanje nekoliko parametara.

Deterministički čvorovi su najjednostavniji slučaj:
$X = f(Parents(X))$ za neku funkciju $f$.

Na primjer, Booleove funkcije
$NorthAmerican \;\Leftrightarrow\; Canadian \vee US \vee Mexican$

Na primjer, numeričke relacije među neprekidnim varijablama

$$\frac{\partial Level}{\partial t} = inflow + precipitation - outflow - evaporation$$

# Sažeti prikaz uvjetnih distribucija — nastavak

Nesigurne (nedeterminističke) veze često se mogu opisati tzv. logičkim vezama "sa šumom" ili "bukom" (engl. "noisy").

Standardni primjer je tzv. "ILI–sa šumom" (engl. "noisy–OR") veza, kao generalizacija logičkog ILI.

ILI–sa šumom distribucija modelira višestruke nezavisne uzroke (nema međudjelovanja). Pretpostavke modela:

1. Roditelji $U_1, \ldots, U_k$ uključuju sve moguće uzroke. Ako neki uzroci fale, dodamo pomoćni čvor za njih (tzv. "leak node").
2. Izostanak jednog uzroka je nezavisan od izostanka ostalih uzroka.

Trebamo samo vjerojatnosti neuspjeha $q_i$ za svaki uzrok zasebno.

Za sve kombinacije uzroka onda vrijedi

$$P(X \mid U_1, \ldots, U_j, \neg U_{j+1}, \ldots, \neg U_k) = 1 - \prod_{i=1}^{j} q_i$$

# Sažeti prikaz — primjer za ILI–sa šumom

Uzmimo da samo prehlada, gripa i malarija izazivaju groznicu.

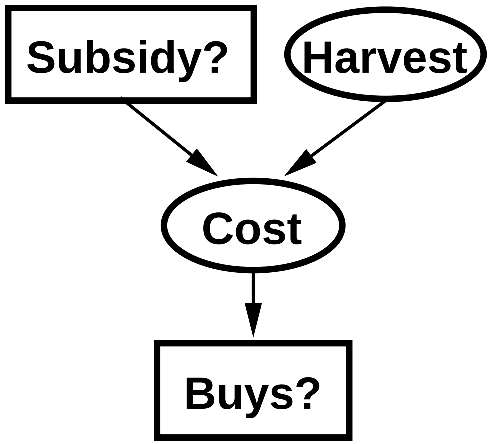
Trebamo znati $q_i$ = nemam groznicu, a imam točno jedan od tri uzroka.

| $Cold$ | $Flu$ | $Malaria$ | $P(Fever)$ | $P(\neg Fever)$ |
|---|---|---|---|---|
| F | F | F | $\mathbf{0.0}$ | $1.0$ |
| F | F | T | $0.9$ | $\mathbf{0.1}$ |
| F | T | F | $0.8$ | $\mathbf{0.2}$ |
| F | T | T | $0.98$ | $0.02 = 0.2 \times 0.1$ |
| T | F | F | $0.4$ | $\mathbf{0.6}$ |
| T | F | T | $0.94$ | $0.06 = 0.6 \times 0.1$ |
| T | T | F | $0.88$ | $0.12 = 0.6 \times 0.2$ |
| T | T | T | $0.988$ | $0.012 = 0.6 \times 0.2 \times 0.1$ |

Broj parametara je **linearan** u broju roditelja.

# Hibridne (diskretne + neprekidne) mreže

Varijable: diskretne ($Subsidy?$ i $Buys?$), neprekidne ($Harvest$ i $Cost$).




Pristupi:

– Diskretizacija neprekidnih varijabli po intervalima — često izaziva velike greške, vrlo velike CPT.

– Kanonske familije neprekidnih f-a gustoće s konačno parametara.

Ovisno o tipu djece, trebamo dvije nove vrste uvjetnih distribucija:

1. Neprekidne varijable, diskretni+neprekidni roditelji (na pr. $Cost$)
2. Diskretne varijable, neprekidni roditelji (na pr. $Buys?$)

# Neprekidne varijable djece

Za $Cost$: trebamo po jednu uvjetnu funkciju gustoće za varijablu-dijete neprekidnih roditelja, za svaku vrijednost diskretnih roditelja.

Najčešće se koristi linearni Gaussov (LG) model, na pr.:

$$
\begin{aligned}
P(Cost = c \mid Harvest = h, Subsidy? = true) \\
&= N(a_t h + b_t, \sigma_t^2)(c) \\
&= \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left( -\frac{1}{2} \left( \frac{c - (a_t h + b_t)}{\sigma_t} \right)^2 \right)
\end{aligned}
$$

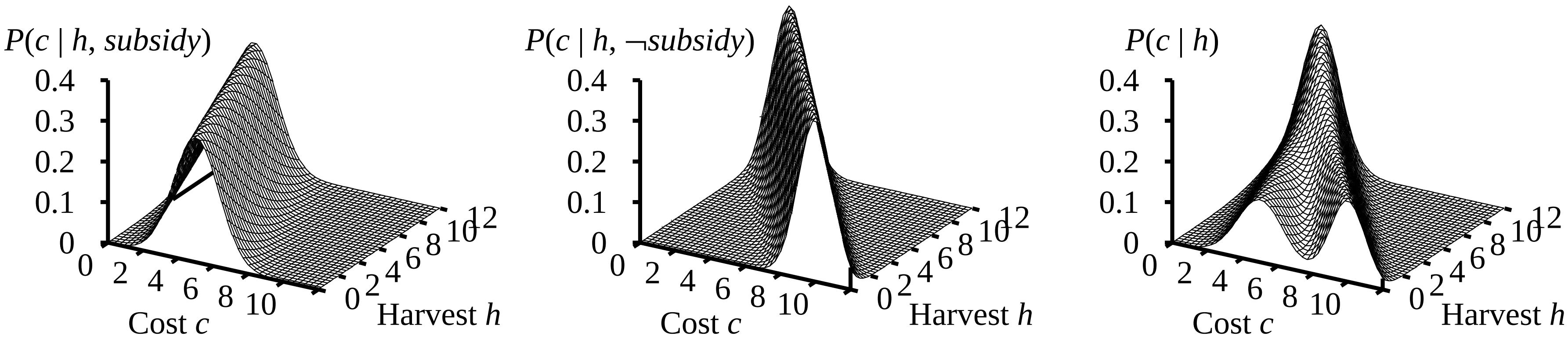
Analogno za $Subsidy? = false$, s parametrima $a_f$, $b_f$, $\sigma_f$.

Očekivanje za $Cost$ varira linearno s $Harvest$, varijanca je fiksna.

Linearna varijacija očekivanja je nerazumna za cijeli raspon,
ali radi dobro ako je očekivani raspon varijable $Harvest$ uzak.

# Neprekidne varijable djece — nastavak

Tri uvjetne distribucije $P(Cost \mid Harvest)$ — za $subsidy$, $\neg subsidy$ i u prosjeku (a priori pola-pola, za $subsidy$ i $\neg subsidy$):

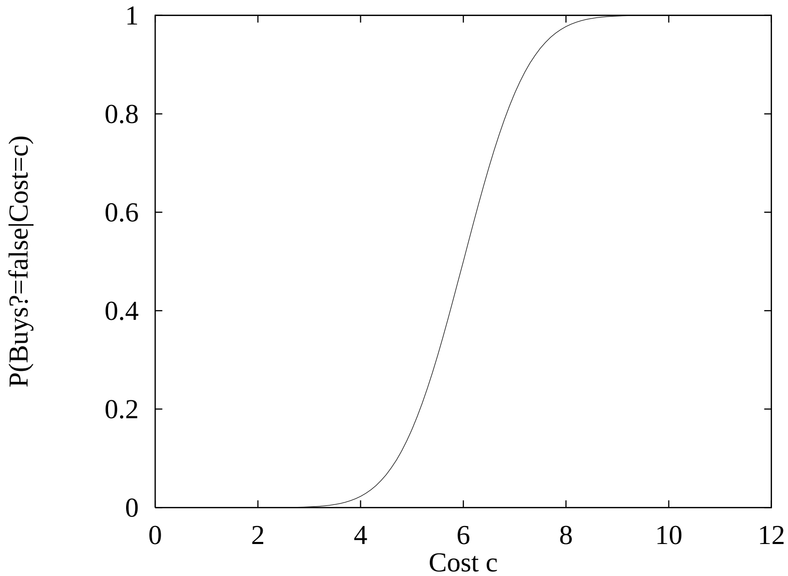



Potpuno neprekidna mreža s LG distribucijama
$\Longrightarrow$ potpuna zajednička distribucija je multivarijantna Gaussova.

Diskretna+neprekidna LG mreža je uvjetno Gaussova mreža, tj. multivarijantna Gaussova po svim neprekidnim varijablama, za svaku kombinaciju vrijednosti diskretnih varijabli.

# Diskretne varijable s neprekidnim roditeljima

Vjerojatnost od $Buys?$ za dani $Cost$ trebala bi imati "meku" granicu:




Probit distribucija koristi integral Gaussove gustoće = f-a distribucije:

$$\Phi(x) = \int_{-\infty}^{x} N(0,1)(x)dx$$

$$P(Buys? = true \mid Cost = c) = \Phi((-c + \mu)/\sigma).$$

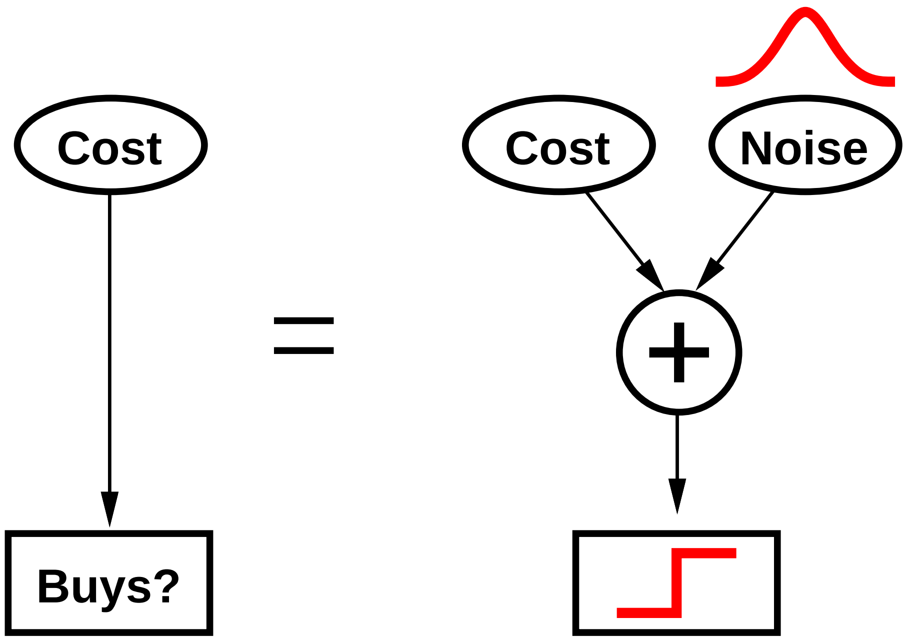
Na slici je $P(Buys? = false) = 1 -$ ovo iznad $= \Phi((c - \mu)/\sigma)$.

# Zašto probit?

Probit = kratica za "probability unit".

1. Ima dobar oblik — za neprekidnu vezu između vrijednosti 0 i 1.

2. Pogled = "čvrsta" granica, čija lokacija ovisi o Gaussovom "šumu":



# Odakle probit?

Definicija funkcije probit: Ako je $\Phi$ funkcija distribucije za $N(0,1)$, funkcija probit je njezin inverz, tj. probit $= \Phi^{-1}$ — nije distribucija!

Razlog za primjenu na diskretne varijable s neprekidnim roditeljima:

Stvarno, imamo samo dvije vrijednosti za vjerojatnost: $0$ i $1$, tj. imamo "čvrstu" ili "oštru" granicu $=$ stepenasta ili Heavisideova funkcija.

Traži se "omekšanje" ove skokovite granice za vrijednost $Y$ $=$ neprekidna veza ("fitting") između dvije diskretne granice.

Ideja: linearizirati zavisnu varijablu $Y$ nekom funkcijom $F : [0,1] \to \mathbf{R}$, a zatim naći pravac koji je odgovarajuća aproksimacija (fitting)

$$F(Y) = \text{pravac u } X,$$

i onda "vratiti" aproksimaciju u $Y = F^{-1}(\text{pravac})$.

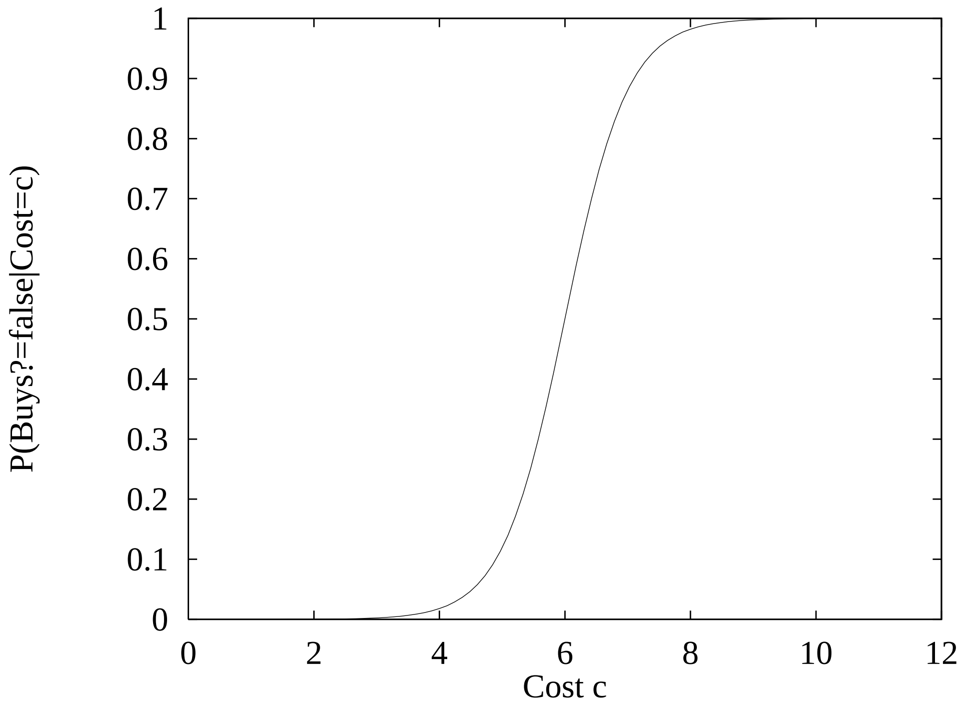
Vrlo prirodan izbor je baš $F = \Phi^{-1}$, pa je $Y = \Phi(\text{pravac})$.

# Diskretne varijable s nepr. roditeljima — nastavak

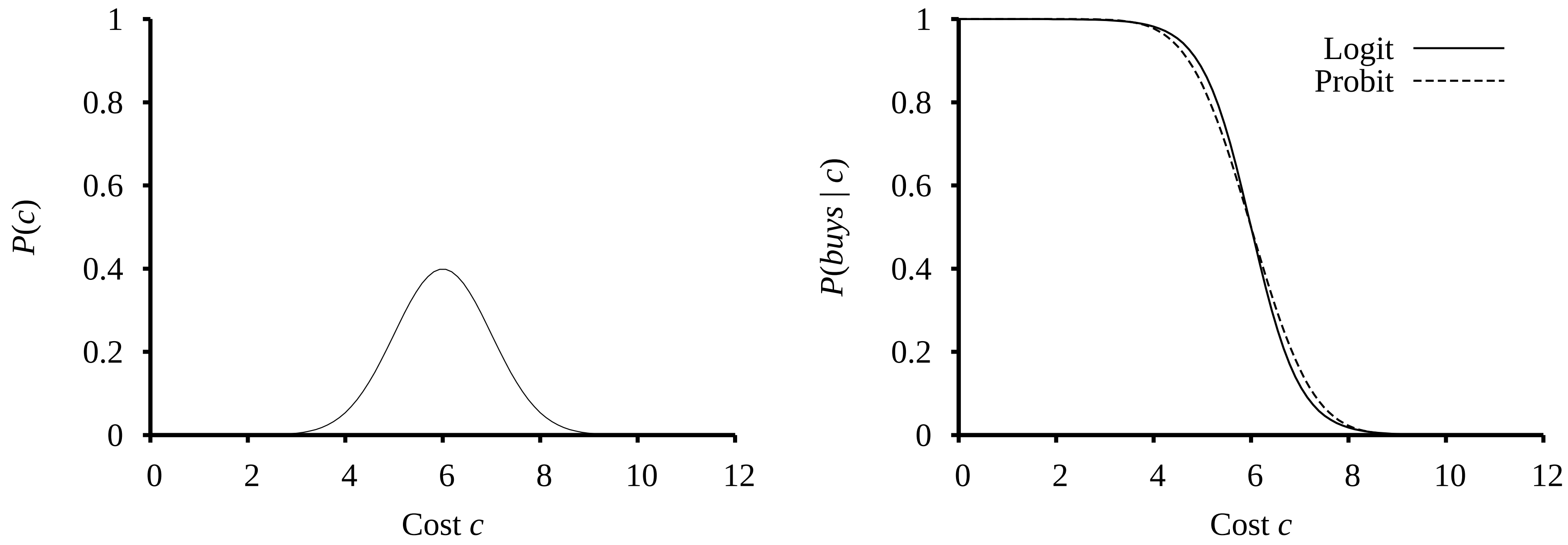
Sigmoidna ili logit distribucija, $F^{-1} = 1/(1+e^{-x})$, često se koristi u neuralnim mrežama:

$$P(Buys? = true \mid Cost = c) = \frac{1}{1+\exp(-2\frac{-c+\mu}{\sigma})}$$

Logit distribucija ima sličan oblik kao probit, ali mnogo dulje repove:



# Diskretne varijable s nepr. roditeljima — nastavak




Normalna (Gaussova) distribucija za granicu cijene $Cost = c$,
s očekivanjem $\mu = 6.0$ i standardnom devijacijom $\sigma = 1.0$,

Logit i probit distribucije za vjerojatnost $P(Buys? = true \mid Cost = c)$,
s parametrima $\mu = 6.0$ i $\sigma = 1.0$.

## Sažetak

Bayesove mreže daju prirodnu reprezentaciju uzročno induciranih uvjetnih nezavisnosti

Topologija $+$ CPT $=$ sažeta reprezentacija zajedničke distribucije

Općenito jednostavno i (ne)ekspertima za konstrukciju

Kanonske (standardne) distribucije (na pr., ILI–sa šumom) $=$ sažeta reprezentacija CPT

Neprekidne varijable $\Rightarrow$ parameterizirane distribucije (na pr., linearna Gaussova)