هوش مصنوعی

فصل ۱۳

استدلال احتمالاتی

Probabilistic Reasoning

کاظم فولادی قلعه
دانشکده مهندسی، پردیس فارابی
دانشگاه تهران

http://courses.fouladi.ir/ai

هوش مصنوعی

استدلال احتمالاتی

۱

بازنمایی دانایی در یک دامنه‌ی نامطمئن

شبکه‌های بیزی

BAYESIAN NETWORKS

شبکه‌ی بیزی

Bayesian Network

یک نمادگذاری ساده و گرافیکی برای بیان استقلال شرطی

(و در نتیجه برای مشخص‌سازی متراکم توزیع‌های توأم کامل)

شبکه‌های بیزی

نحو

BAYESIAN NETWORKS

شبکه‌ی بیزی
Bayesian Network

یک گراف جهت‌دار بدون دور

پیوندها *Links*	**گره‌ها** *Nodes*
نشان‌دهنده‌ی **رابطه‌ی تأثیر مستقیم**	نشان‌دهنده‌ی **متغیرهای تصادفی**
(link ≈ "directly influences")	**توزیع شرطی** *Conditional Distribution*
	برای هر گره، توزیع شرطی آن گره به‌شرط والدهای آن را داریم:
	$\mathbf{P}(X_i \mid Parents(X_i))$
	در قالب جدول احتمال شرطی (CPT)

شبکه‌های بیزی

جدول احتمال شرطی

CONDITIONAL PROBABILITY TABLE

توزیع شرطی به طور ساده در قالب جدول احتمال شرطی بازنمایی می‌شود.

جدول احتمال شرطی
Conditional Probability Table

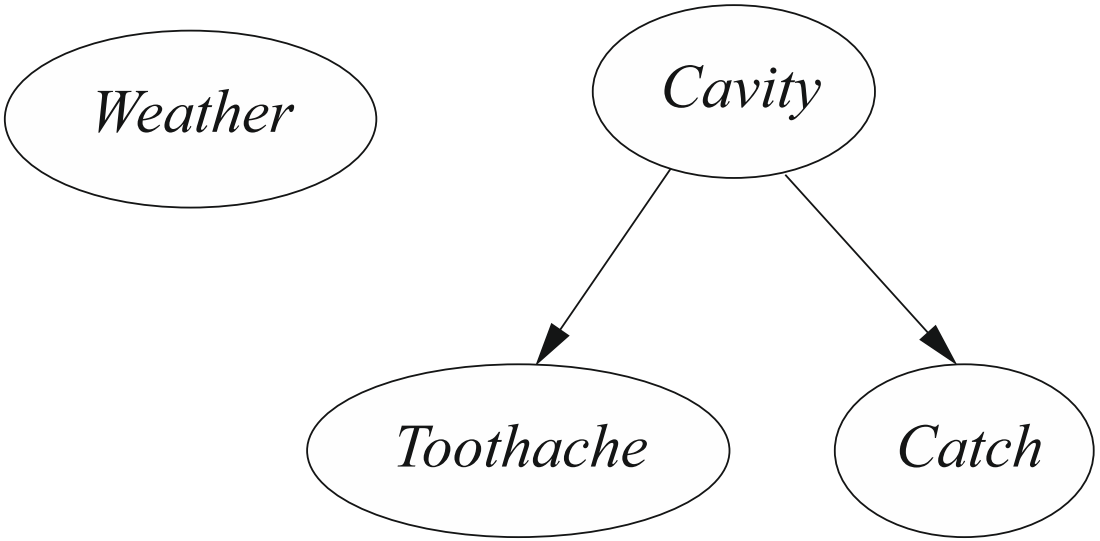
توزیع برای متغیر X_i برای هر ترکیب از مقادیر والدهای آن

CPT

شبکه‌های بیزی

مثال

BAYESIAN NETWORKS



$Weather$ is independent of the other variables

$Toothache$ and $Catch$ are conditionally independent given $Cavity$

توپولوژی شبکه‌ی بیزی بیان استقلال شرطی را کدگذاری می‌کند.

شبکه‌های بیزی

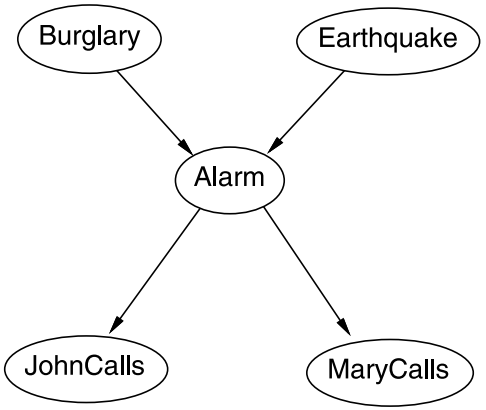
مثال: سیستم هشدار سرقت منزل

BAYESIAN NETWORKS

شما دو همسایه به نام‌های مری و جان دارید.
آن‌ها قول داده‌اند که در صورت شنیدن صدای زنگ هشدار سرقت با شما در محل کارتان تماس بگیرند.
جان همیشه وقتی تماس می‌گیرد که صدای زنگ هشدار را بشنود،
اما گاهی صدای زنگ هشدار را با صدای زنگ تلفن اشتباه می‌گیرد و با شما تماس می‌گیرد.
مری موسیقی با صدای بلند گوش می‌دهد و گاهی صدای زنگ هشدار را نمی‌شنود.
البته گاهی زمین‌لرزه‌ی خفیف هم باعث به صدا در آمدن زنگ هشدار می‌شود.
می‌خواهیم با دانستن فرد تماس‌گیرنده، احتمال سرقت را تخمین بزنیم.

Variables: *Burglar*, *Earthquake*, *Alarm*, *JohnCalls*, *MaryCalls*

سرقت زمین‌لرزه هشدار تماس جان تماس مری



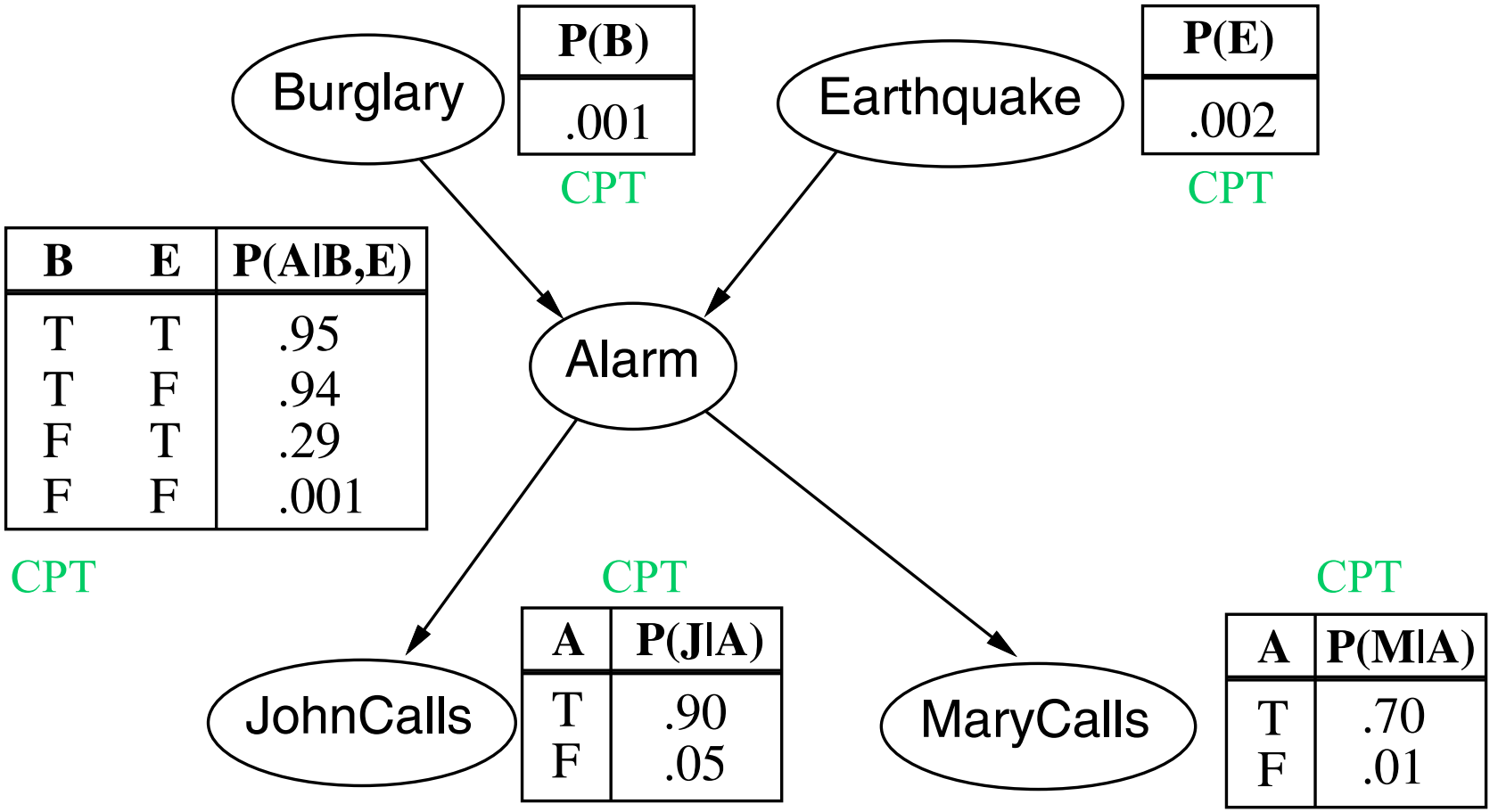
توپولوژی شبکه‌ی بیزی، دانایی «علّی» را منعکس می‌کند:

- وقوع یک سرقت می‌تواند زنگ هشدار را به صدا درآورد.
- وقوع یک زمین‌لرزه می‌تواند زنگ هشدار را به صدا درآورد.
- زنگ هشدار می‌تواند باعث شود مری تماس بگیرد.
- زنگ هشدار می‌تواند باعث شود جان تماس بگیرد.

شبکه‌های بیزی

مثال: سیستم هشدار سرقت منزل

BAYESIAN NETWORKS



Variables: *Burglar*, *Earthquake*, *Alarm*, *JohnCalls*, *MaryCalls*

B E A J M

شبکه‌های بیزی

مفهوم «تراکم»

COMPACTNESS

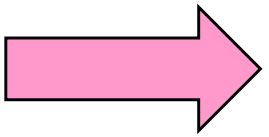
یک CPT برای متغیر تصادفی بولی X با k والد بولی
دارای 2^k سطر برای ترکیب‌های مختلف مقادیر والدهاست.

هر سطر یک عدد p برای $X = true$ نیاز دارد.
(عدد $1-p$ برای $X = false$ مشخص است)

اگر هریک از n متغیر حداکثر k والد داشته باشد، کل شبکه به $O(n \cdot 2^k)$ عدد برای CPTها نیاز دارد.

$O(2^n)$

برای توزیع‌های توأم کامل
(نمایی بر حسب n)



$O(n \cdot 2^k)$

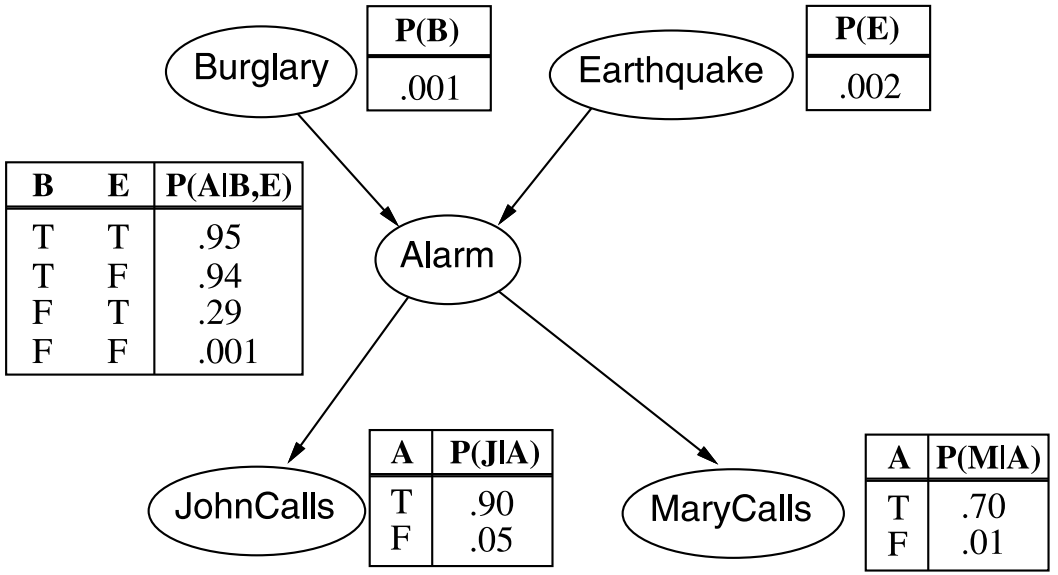
برای توزیع‌های شرطی
(خطی بر حسب n)

شبکه‌های بیزی

مفهوم «تراکم»: مثال

COMPACTNESS

برای مثال سیستم هشدار سرقت



$O(2^n)$

برای توزیع‌های توأم کامل
(نمایی بر حسب n)

$2^5 = 32$

$O(n \cdot 2^k)$

برای توزیع‌های شرطی
(خطی بر حسب n)

$1 + 1 + 4 + 2 + 2 = 10$

هوش مصنوعی

استدلال احتمالاتی

۲

معناشناسی شبکه‌های بیزی

معناشناسی شبکه‌های بیزی

THE SEMANTICS OF BAYESIAN NETWORKS

معناشناسی شبکه‌های بیزی	
معناشناسی سراسری *Global Semantics*	**معناشناسی محلی** *Local Semantics*

معناشناسی شبکه‌های بیزی

معناشناسی سراسری

THE SEMANTICS OF BAYESIAN NETWORKS

معناشناسی شبکه‌های بیزی	
معناشناسی محلی *Local Semantics*	**معناشناسی سراسری** *Global Semantics*

توزیع توأم کامل = حاصل‌ضرب توزیع‌های شرطی محلی

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

معناشناسی شبکه‌های بیزی

معناشناسی سراسری: مثال

THE SEMANTICS OF BAYESIAN NETWORKS

معناشناسی شبکه‌های بیزی

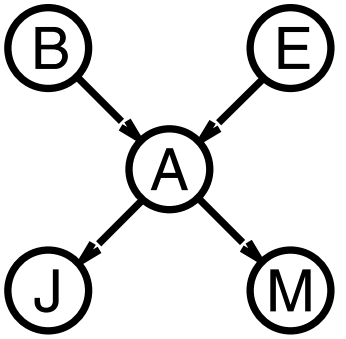
معناشناسی محلی *Local Semantics*	**معناشناسی سراسری** *Global Semantics*

توزیع توأم کامل = حاصل‌ضرب توزیع‌های شرطی محلی

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

e.g., $P(j \wedge m \wedge a \wedge \neg b \wedge \neg e)$

$$\begin{aligned} &= P(j|a)P(m|a)P(a|\neg b, \neg e)P(\neg b)P(\neg e) \\ &= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 \\ &\approx 0.00063 \end{aligned}$$



معناشناسی شبکه‌های بیزی

معناشناسی محلی

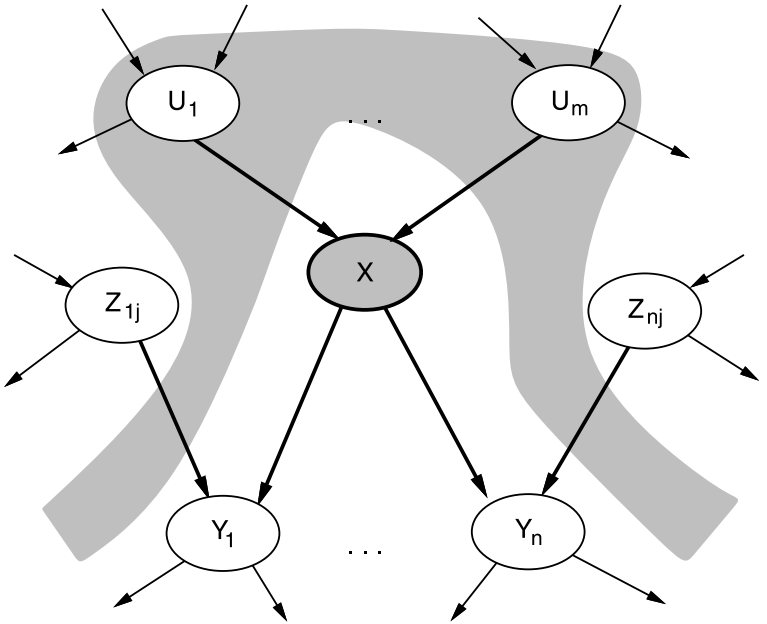
THE SEMANTICS OF BAYESIAN NETWORKS

معناشناسی شبکه‌های بیزی

معناشناسی محلی
Local Semantics

معناشناسی سراسری
Global Semantics

هر گره با داشتن والدهایش از گره‌های غیرنواده‌اش مستقل شرطی است.



معناشناسی شبکه‌های بیزی

قضیه‌ی هم‌ارزی معناشناسی سراسری با معناشناسی محلی

THE SEMANTICS OF BAYESIAN NETWORKS

معناشناسی شبکه‌های بیزی	
معناشناسی سراسری *Global Semantics*	**معناشناسی محلی** *Local Semantics*

Theorem: Local semantics $\Leftrightarrow$ global semantics

معناشناسی شبکه‌های بیزی

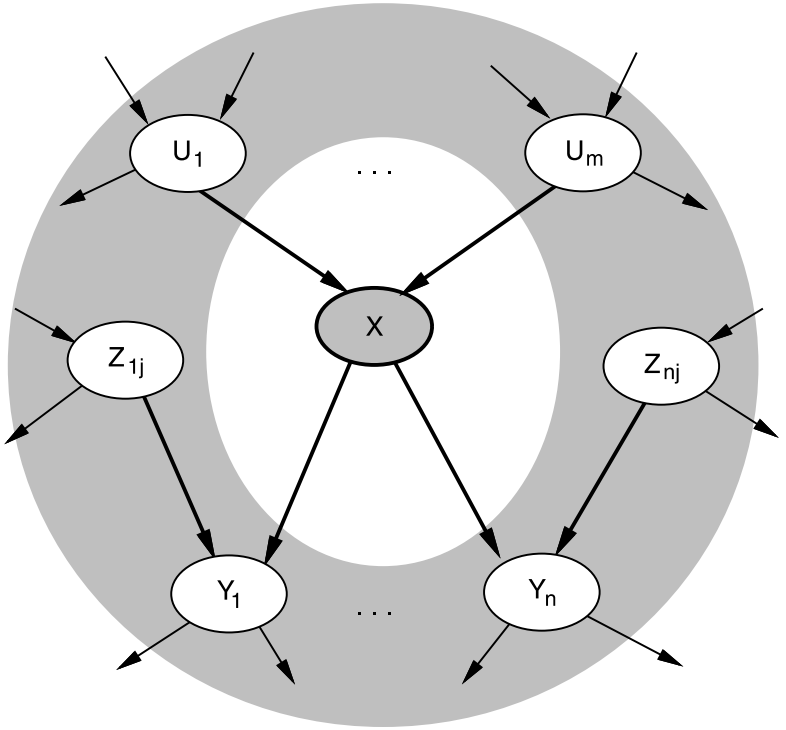
پتوی مارکوف

MARKOV BLANKET

هر گره از سایر گره‌ها مستقل شرطی است به‌شرط داشتن پتوی مارکوف آن گره

پتوی مارکوف
Markov Blanket

پتوی مارکوف برای هر گره = والدهای آن + فرزندان آن + والدهای فرزندان آن



ساختن شبکه‌های بیزی

CONSTRUCTING BAYESIAN NETWORKS

به روشی نیاز داریم که
با بررسی یک سری از بیان‌های آزمون‌پذیر به طور محلی در مورد استقلال شرطی
بتواند معناشناسی سراسری مورد نیاز را تضمین کند.

1. Choose an ordering of variables $X_1, \ldots, X_n$
2. For $i = 1$ to n
   add X_i to the network
   select parents from $X_1, \ldots, X_{i-1}$ such that
   $$\mathbf{P}(X_i|Parents(X_i)) = \mathbf{P}(X_i|X_1, \ldots, X_{i-1})$$

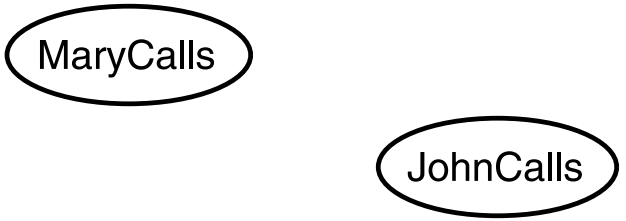
انتخاب والدها، معناشناسی سراسری را تضمین می‌کند:

$$\begin{aligned}\mathbf{P}(X_1, \ldots, X_n) &= \prod_{i=1}^{n} \mathbf{P}(X_i|X_1, \ldots, X_{i-1}) \quad \text{(chain rule)} \\ &= \prod_{i=1}^{n} \mathbf{P}(X_i|Parents(X_i)) \quad \text{(by construction)}\end{aligned}$$

ساختن شبکه‌های بیزی

مثال (۱ از ۶)

Suppose we choose the ordering M, J, A, B, E

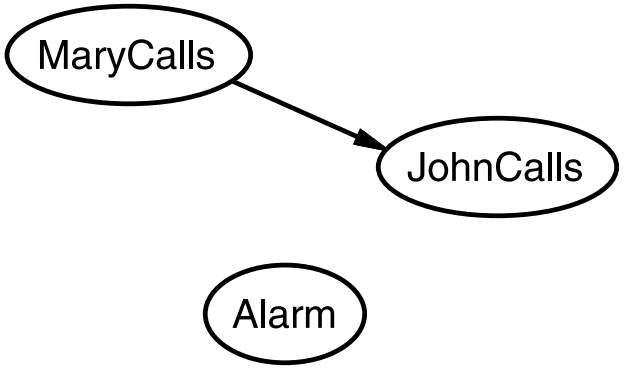


$P(J|M) = P(J)$?

ساختن شبکه‌های بیزی

مثال (۲ از ۶)

Suppose we choose the ordering M, J, A, B, E

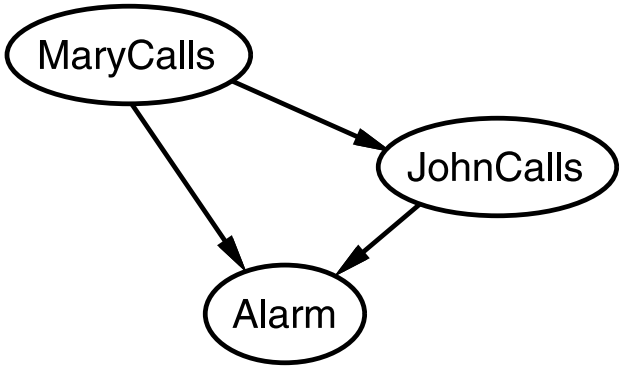


$P(J|M) = P(J)$? No
$P(A|J, M) = P(A|J)$? $P(A|J, M) = P(A)$?

ساختن شبکه‌های بیزی

مثال (۳ از ۶)

Suppose we choose the ordering M, J, A, B, E

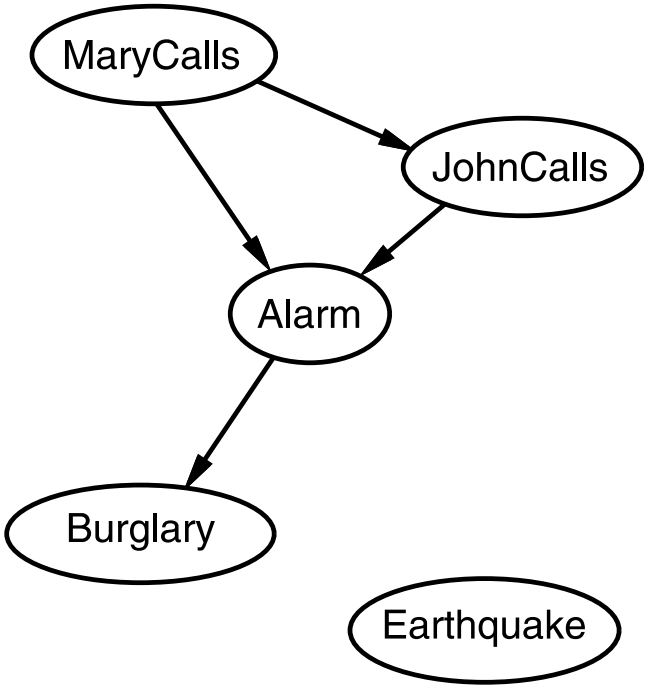


$P(J|M) = P(J)$? No
$P(A|J, M) = P(A|J)$? $P(A|J, M) = P(A)$? No
$P(B|A, J, M) = P(B|A)$?
$P(B|A, J, M) = P(B)$?

ساختن شبکه‌های بیزی

مثال (۴ از ۶)

Suppose we choose the ordering M, J, A, B, E



$P(J|M) = P(J)$? No

$P(A|J, M) = P(A|J)$? $P(A|J, M) = P(A)$? No

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

$P(B|A, J, M) = P(B)$? No

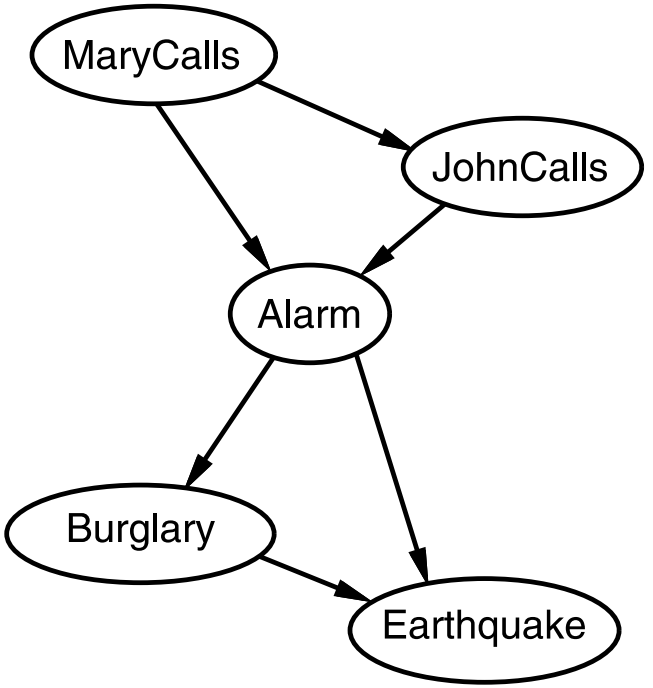
$P(E|B, A, J, M) = P(E|A)$?

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

ساختن شبکه‌های بیزی

مثال (۵ از ۶)

Suppose we choose the ordering M, J, A, B, E



$P(J|M) = P(J)$? No

$P(A|J, M) = P(A|J)$? $P(A|J, M) = P(A)$? No

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

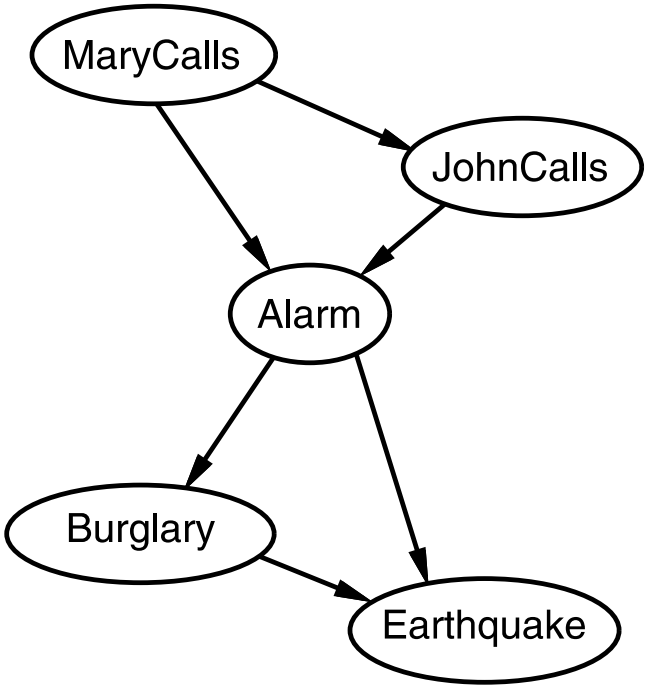
$P(B|A, J, M) = P(B)$? No

$P(E|B, A, J, M) = P(E|A)$? No

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

ساختن شبکه‌های بیزی

مثال (۶ از ۶)



تصمیم‌گیری در مورد استقلال شرطی در جهت‌های غیرعلّی دشوار است.

(به نظر می‌رسد مدل‌های علّی و استقلال شرطی برای انسان‌ها سیم‌بندی سخت شده است)

سنجش احتمالات شرطی در جهت‌های غیرعلّی دشوار است.

در این مثال تراکم شبکه کم است: فقط $1 + 2 + 4 + 2 + 4 = 13$ عدد لازم داریم.

ساخت شبکه‌های بیزی

مثال: تشخیص عیب خودرو

EXAMPLE: CAR DIAGNOSIS

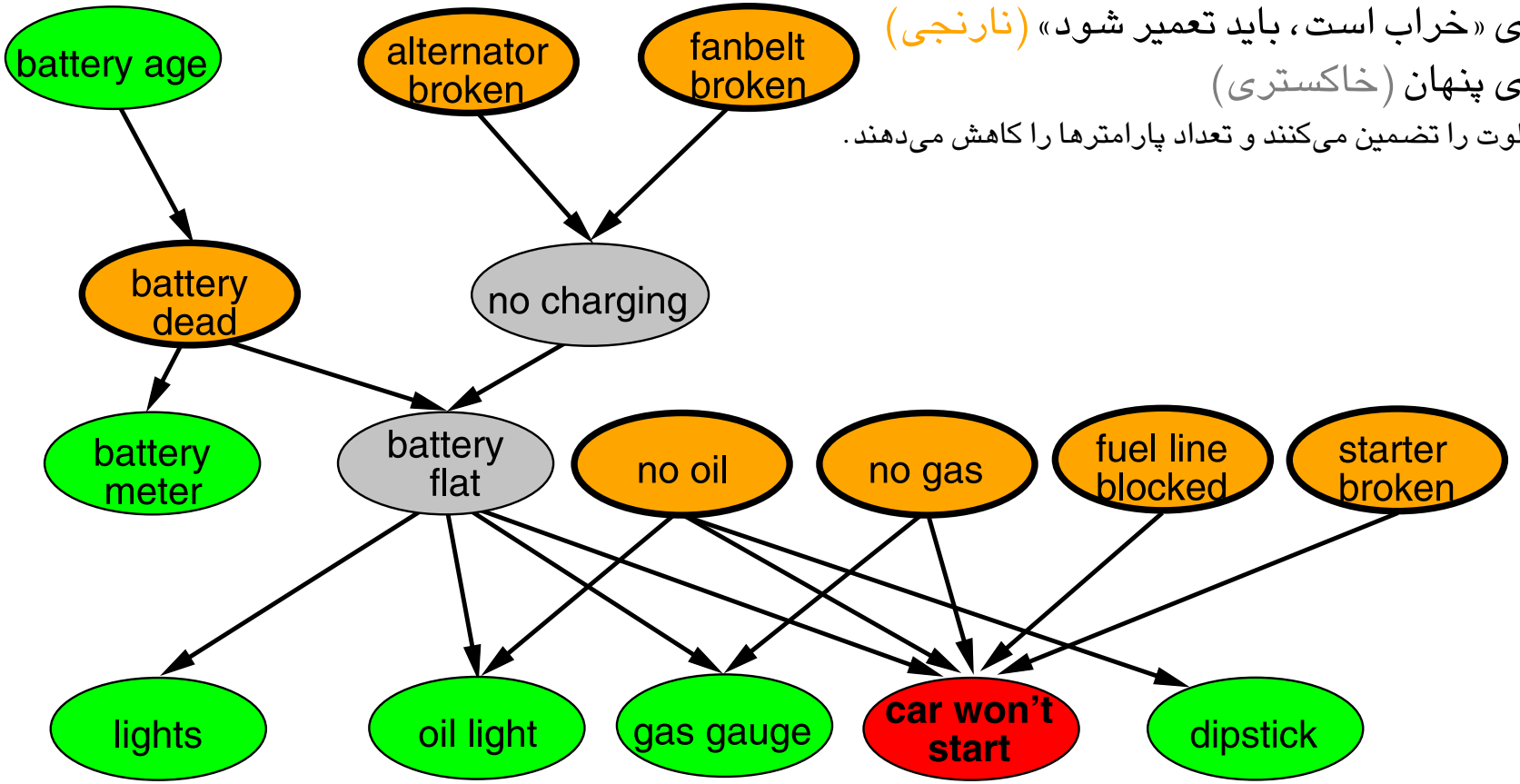
شاهد آغازین: **خودرو روش نمی‌شود**

متغیرهای آزمون‌پذیر (سبز)

متغیرهای «خراب است، باید تعمیر شود» (نارنجی)

متغیرهای پنهان (خاکستری)

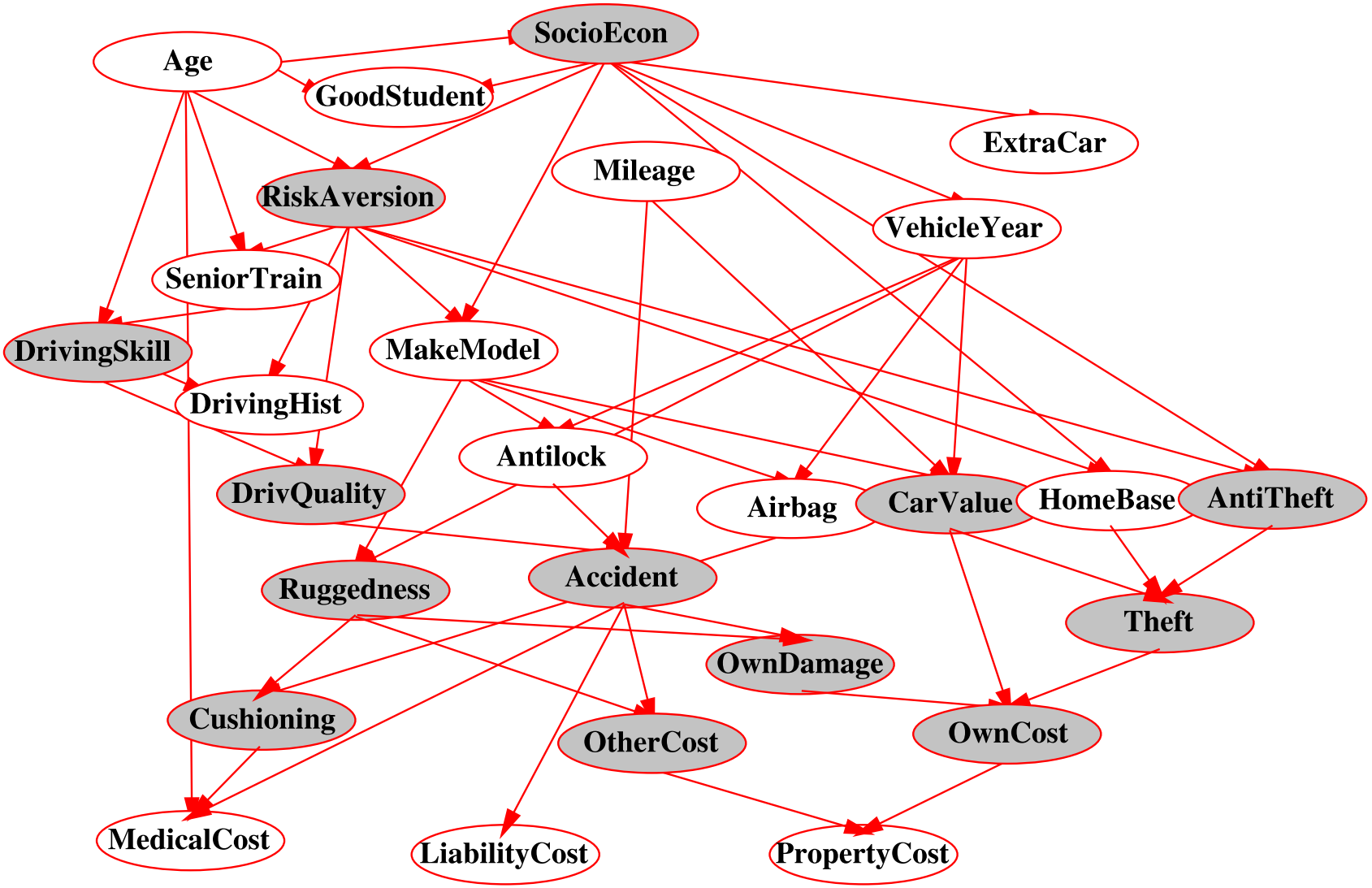
ساختار خلوت را تضمین می‌کنند و تعداد پارامترها را کاهش می‌دهند.



ساخت شبکه‌های بیزی

مثال: بیمه‌ی خودرو

EXAMPLE: CAR INSURANCE



هوش مصنوعی

استدلال احتمالاتی

۳

بازنمایی کارآمد توزیع‌های شرطی

توزیع‌های شرطی متراکم

COMPACT CONDITIONAL DISTRIBUTIONS

مشکل

اندازه‌ی CPT با تعداد والدها به صورت نمایی رشد می‌کند.

حل مشکل

برای گره‌های قطعی **(deterministic nodes)**

استفاده از روابط قطعی بین گره‌ها

برای روابط نامطمئن **(uncertain relationships)**

استفاده از روابط منطقی «نویزی» مانند مدل **noisy-OR**

اندازه‌ی CPT با والدها یا فرزندان پیوسته–مقدار بی‌نهایت می‌شود.

حل مشکل

متغیرهای پیوسته، والدهای گسسته + پیوسته

استفاده از یک تابع چگالی شرطی

متغیرهای گسسته، والدهای پیوسته

استفاده از یک تابع چگالی شرطی

توزیع‌های شرطی متراکم

برای گره‌های قطعی

DETERMINISTIC NODES

مقدار گره‌ی قطعی از روی مقدار والدهای آن به‌صورت قطعی مشخص می‌شود.

$$X = f(Parents(X)) \text{ for some function } f$$

برای مثال: توابع بولی:

$$NorthAmerican \Leftrightarrow Canadian \vee US \vee Mexican$$

برای مثال: روابط عددی بین متغیرهای پیوسته

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

توزیع‌های شرطی متراکم

برای رابطه‌های نامطمئن

UNCERTAIN RELATIONSHIPS

رابطه‌های غیرقطعی را می‌توان
با استفاده از روابط منطقی «نویزی» مانند مدل **noisy-OR** مشخص کرد.

$$Child \Leftarrow Parent_1 \vee Parent_2 \vee \cdots \vee Parent_k$$

در اینکه والدها می‌توانند موجب درست شدن فرزند شوند، عدم اطمینان مجاز شمرده می‌شود.

توزیع‌های **noisy-OR** علت‌های غیرمتعامل چندگانه را مدل می‌کنند؛ با دو شرط

۱ والدها شامل همه‌ی علت‌ها باشند: $U_1, U_2, \ldots, U_k$

همیشه می‌توان سایر علت‌ها را با اضافه کردن یک **گره‌ی نشتی (leak node)** پوشش داد.

۲ هر علت به‌تنهایی دارای احتمال شکست مستقل q_i باشد.

q_i : **احتمال ممانعت فردی (individual inhibition probabilities)**

در این صورت داریم:

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

توزیع‌های شرطی متراکم

برای رابطه‌های نامطمئن : مثال

UNCERTAIN RELATIONSHIPS

در منطق گزاره‌ای می‌توان گفت $Fever \Leftrightarrow Cold \vee Flu \vee Malaria$

$Cold$	Flu	$Malaria$	$P(Fever)$	$P(\neg Fever)$
F	F	F	**0.0**	1.0
F	F	T	0.9	**0.1** داده شده
F	T	F	0.8	**0.2** داده شده
F	T	T	0.98	$0.02 = 0.2 \times 0.1$
T	F	F	0.4	**0.6** داده شده
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	$0.012 = 0.6 \times 0.2 \times 0.1$

بر اساس احتمالات ممانعت فردی،
می‌توان کل CPT را کامل کرد.

$$q_{\text{cold}} = P(\neg fever \,|\, cold, \neg flu, \neg malaria) = 0.6 ,$$
$$q_{\text{flu}} = P(\neg fever \,|\, \neg cold, flu, \neg malaria) = 0.2 ,$$
$$q_{\text{malaria}} = P(\neg fever \,|\, \neg cold, \neg flu, malaria) = 0.1 .$$

$$P(x_i \,|\, parents(X_i)) = 1 - \prod_{\{j:X_j = true\}} q_j ,$$

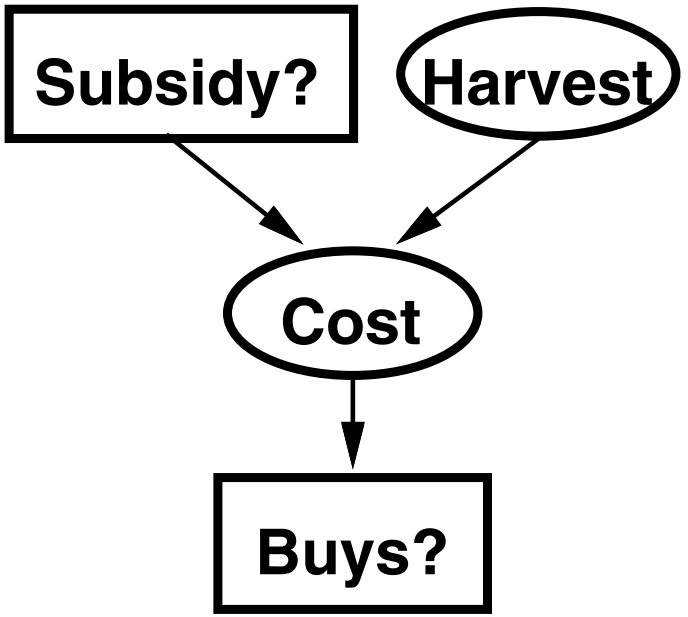
تعداد پارامترهای مورد نیاز بر حسب تعداد والدها، **خطی** است.

توزیع‌های شرطی متراکم

برای شبکه‌های بیزی هیبرید (گسسته + پیوسته)

HYBRID (DISCRETE+CONTINUOUS) NETWORKS

Discrete (*Subsidy?* and *Buys?*); continuous (*Harvest* and *Cost*)



چگونگی برخورد با متغیرهای پیوسته

گزینه ۱) استفاده از گسسته‌سازی: ❋ مشکل خطای احتمالی بالا ❋ مشکل CPTهای بزرگ

گزینه ۲) استفاده از خانواده‌های کانونیک توزیع‌های پارامتری متناهی

توزیع‌های شرطی متراکم

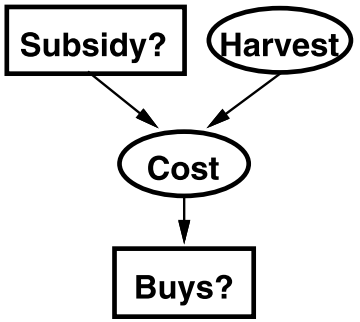
برای شبکه‌های بیزی هیبرید (گسسته + پیوسته): برای متغیرهای پیوسته فرزند

HYBRID (DISCRETE+CONTINUOUS) NETWORKS

نیاز داریم به یک تابع چگالی شرطی
برای متغیر فرزند پیوسته با داشتن والدهای پیوسته به‌ازای هر انتساب ممکن برای والدهای گسسته

معمولاً مدل **گاوسی خطی** برای تابع چگالی شرطی متداول‌ترین گزینه است، مثلاً:

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



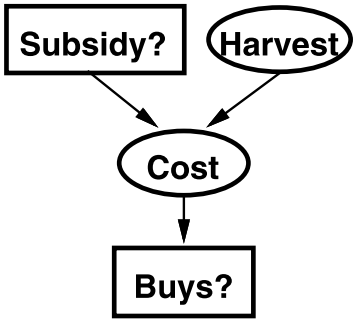
متوسط فرزند ($Cost$) به صورت خطی با والد ($Harvest$) تغییر می‌کند؛ واریانس ثابت است.

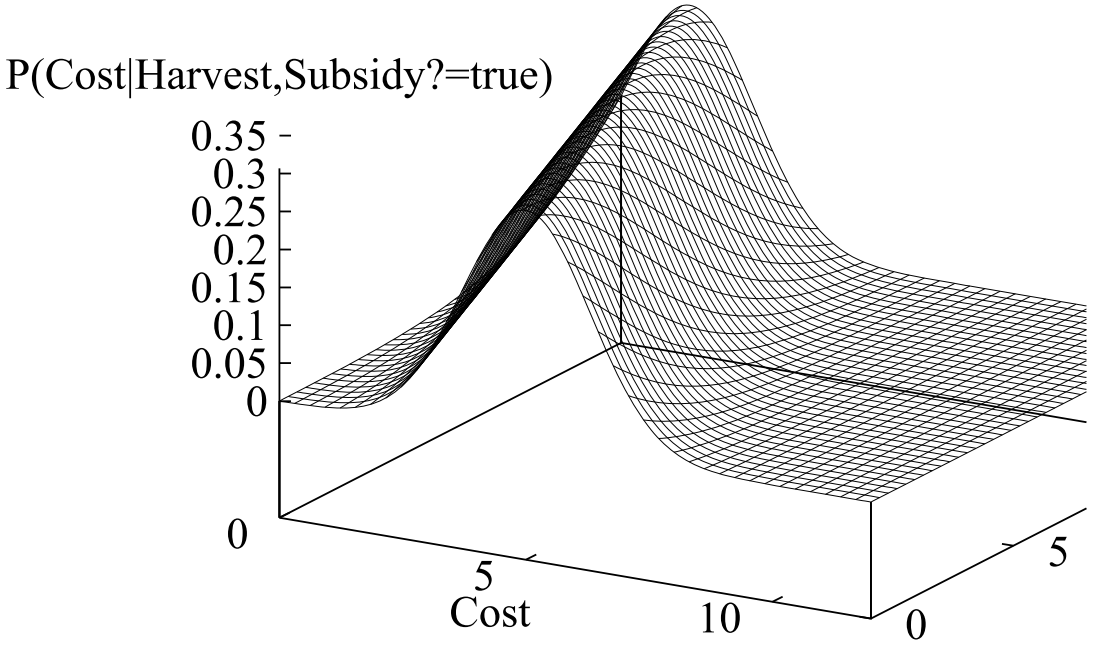
تغییر خطی بر روی یک بازه‌ی بزرگ غیرمنطقی است،
اما به‌خوبی کار می‌کند اگر بازه‌ی احتمالی $Harvest$ باریک باشد.

توزیع‌های شرطی متراکم

برای شبکه‌های بیزی هیبرید (گسسته + پیوسته): برای متغیرهای پیوسته فرزند: توزیع گاوسی خطی

HYBRID (DISCRETE+CONTINUOUS) NETWORKS





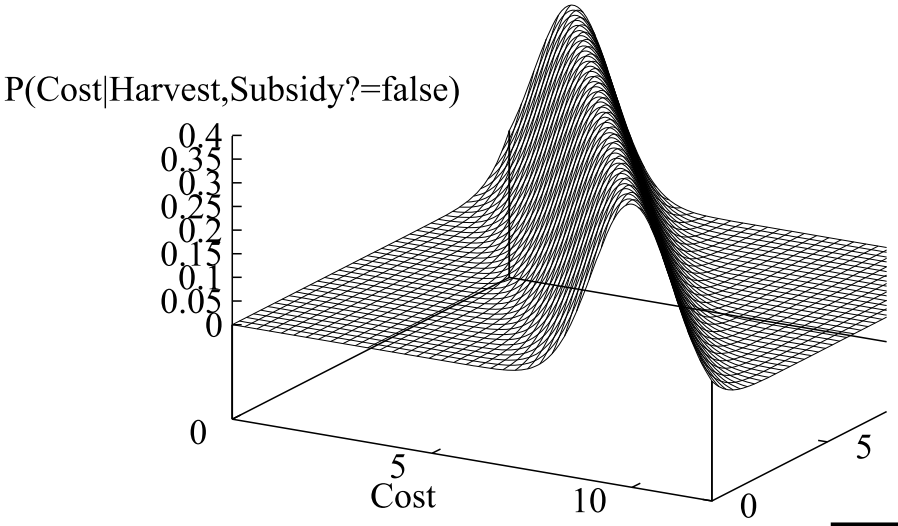
یک شبکه که تنها از متغیرهای پیوسته با توزیع گاوسی خطی تشکیل شده است
دارای یک توزیع توأم کامل با گاوسی چندمتغیره است.

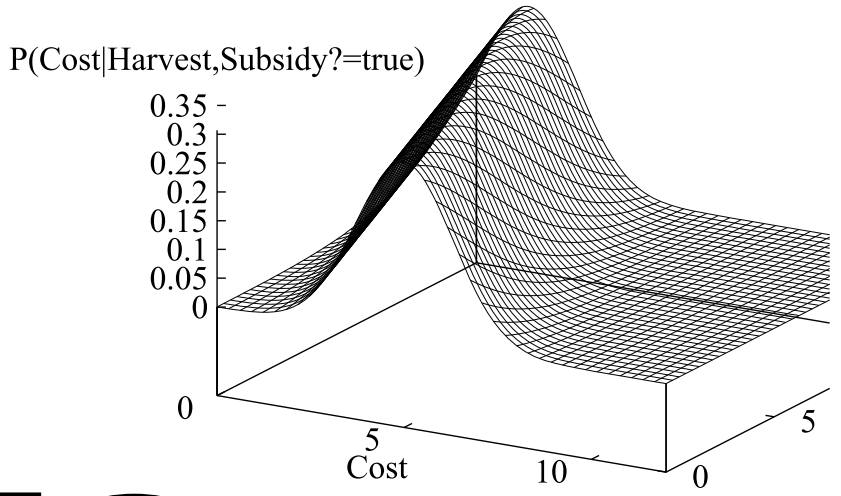
شبکه‌ی گاوسی خطی گسسته + پیوسته، یک شبکه‌ی گاوسی شرطی است.
یعنی: یک گاوسی چندمتغیره بر روی همه‌ی متغیرهای پیوسته
برای هر ترکیب از مقادیر متغیرهای گسسته وجود دارد.

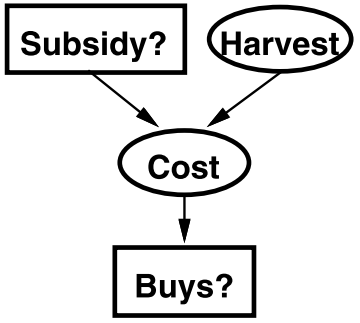
توزیع‌های شرطی متراکم

برای شبکه‌های بیزی هیبرید (گسسته + پیوسته): برای متغیرهای پیوسته فرزند: توزیع گاوسی خطی

HYBRID (DISCRETE+CONTINUOUS) NETWORKS







شبکه‌ی گاوسی خطی گسسته + پیوسته، یک شبکه‌ی گاوسی شرطی است.

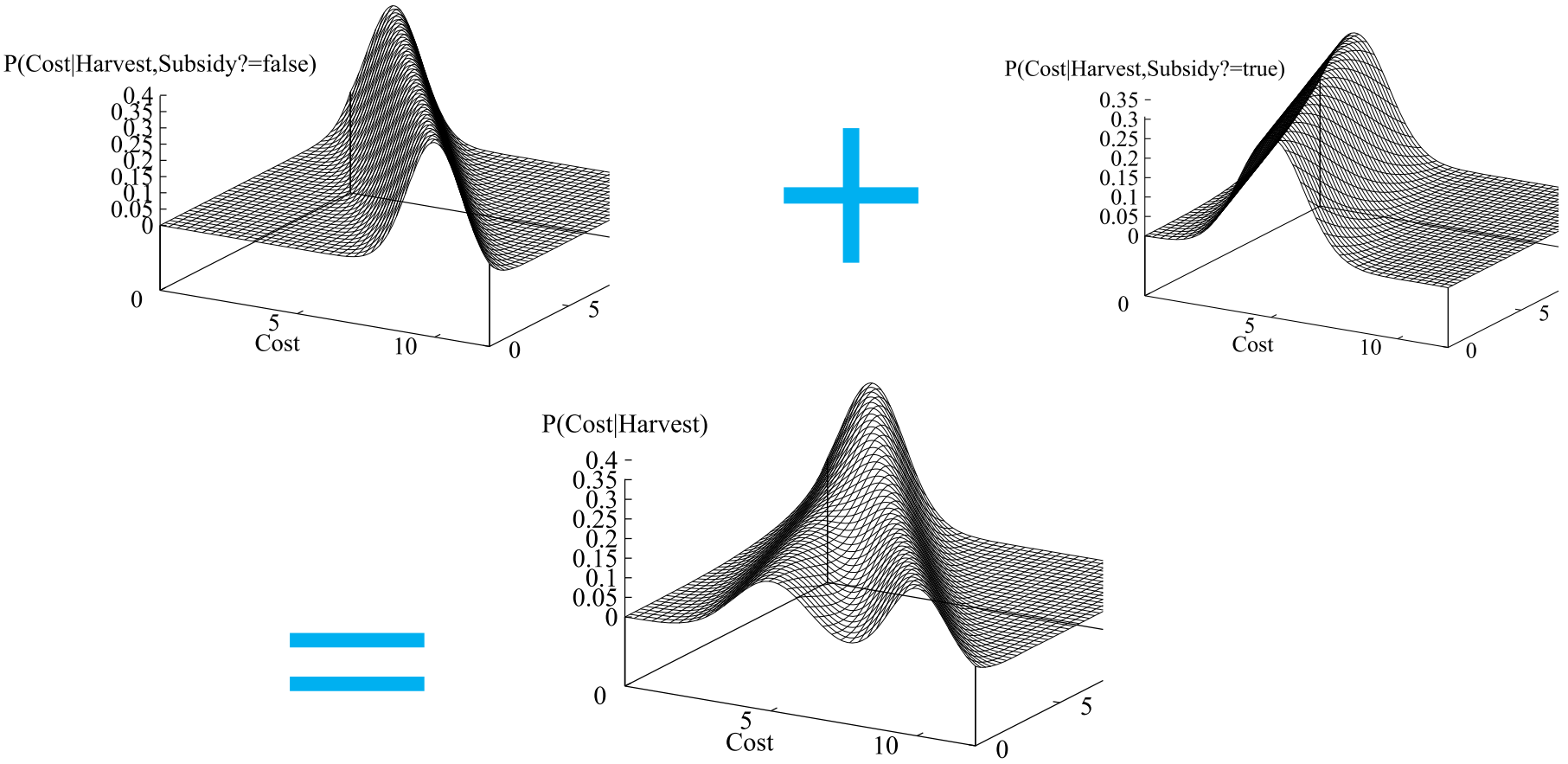
یعنی: یک گاوسی چندمتغیره بر روی همه‌ی متغیرهای پیوسته

برای هر ترکیب از مقادیر متغیرهای گسسته وجود دارد.

توزیع‌های شرطی متراکم

برای شبکه‌های بیزی هیبرید (گسسته + پیوسته): برای متغیرهای پیوسته فرزند: توزیع گاوسی خطی

HYBRID (DISCRETE+CONTINUOUS) NETWORKS



مجموع دو توزیع گاوسی، یک گاوسی است.

توزیع‌های شرطی متراکم

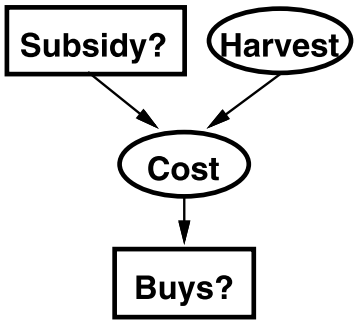
برای شبکه‌های بیزی هیبرید (گسسته + پیوسته): برای متغیرهای گسسته فرزند و والدهای پیوسته

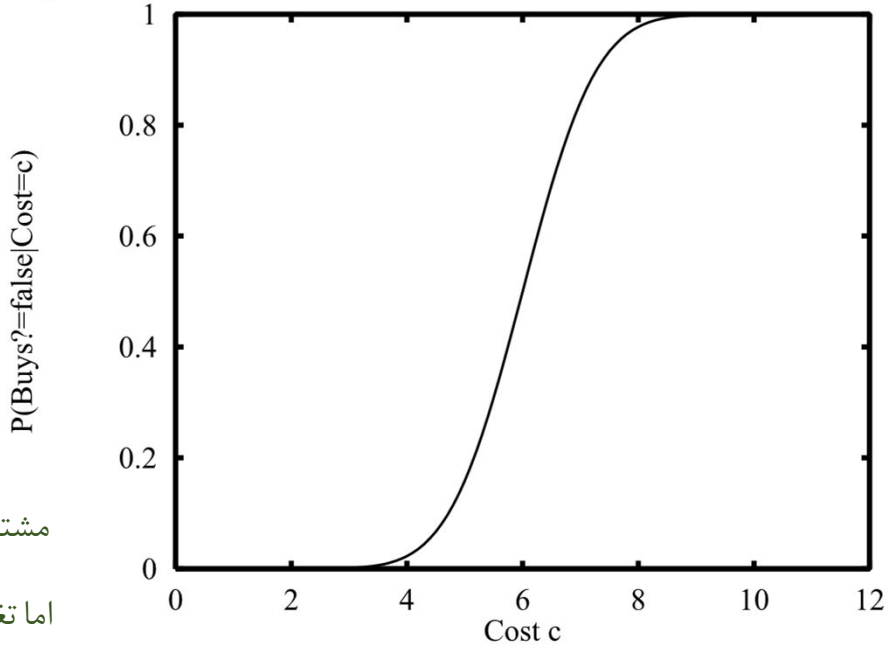
HYBRID (DISCRETE+CONTINUOUS) NETWORKS

احتمال فرزند گسسته به شرط والد پیوسته باید یک مقدار آستانه‌ی نرم باشد (مثل توزیع پروبیت).

(نباید تغییر ناگهانی داشته باشد)

Probability of *Buys?* given *Cost* should be a "soft" threshold:





مشتری در صورتی خرید می‌کند که قیمت پایین باشد و
در صورت بالا بودن قیمت خرید نخواهد کرد؛
اما تغییرات احتمال خرید در ناحیه‌ی میانی ملایم است.

توزیع **پروبیت (Probit)** از انتگرال گاوسی استفاده می‌کند:

$$\Phi(x) = \int_{-\infty}^{x} N(0,1)(x)dx$$
$$P(Buys? = true \mid Cost = c) = \Phi((-c + \mu)/\sigma)$$

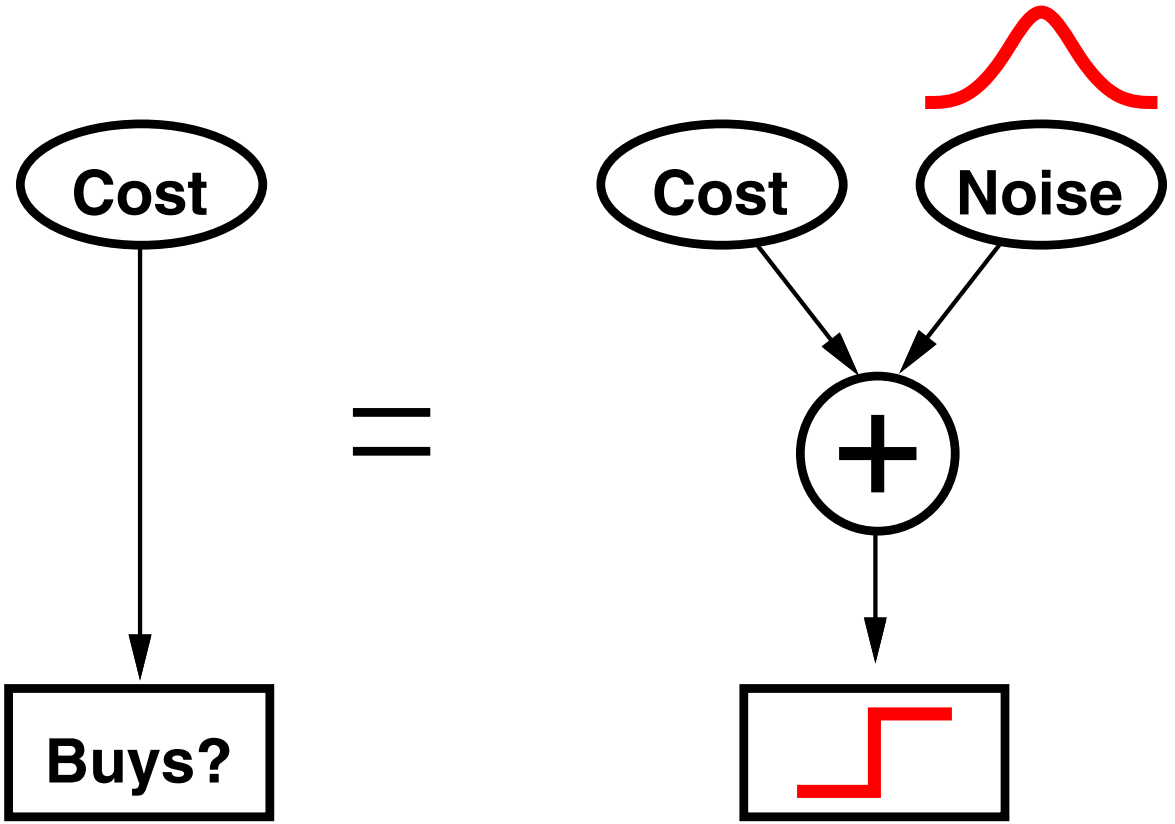
توزیع‌های شرطی متراکم

برای شبکه‌های بیزی هیبرید (گسسته + پیوسته) : برای متغیرهای گسسته فرزند و والدهای پیوسته : توزیع پروبیت

HYBRID (DISCRETE+CONTINUOUS) NETWORKS

ویژگی توزیع **پروبیت** (Probit)

فرآیند تصمیم‌گیری دارای مقدار آستانه‌ی سخت است، اما مکان دقیق آستانه در معرض نویز گاوسی تصادفی قرار دارد.



توزیع‌های شرطی متراکم

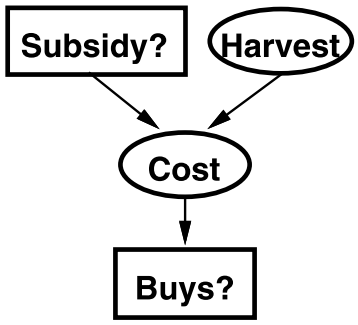
برای شبکه‌های بیزی هیبرید (گسسته + پیوسته): برای متغیرهای گسسته فرزند و والدهای پیوسته: توزیع لوگیت

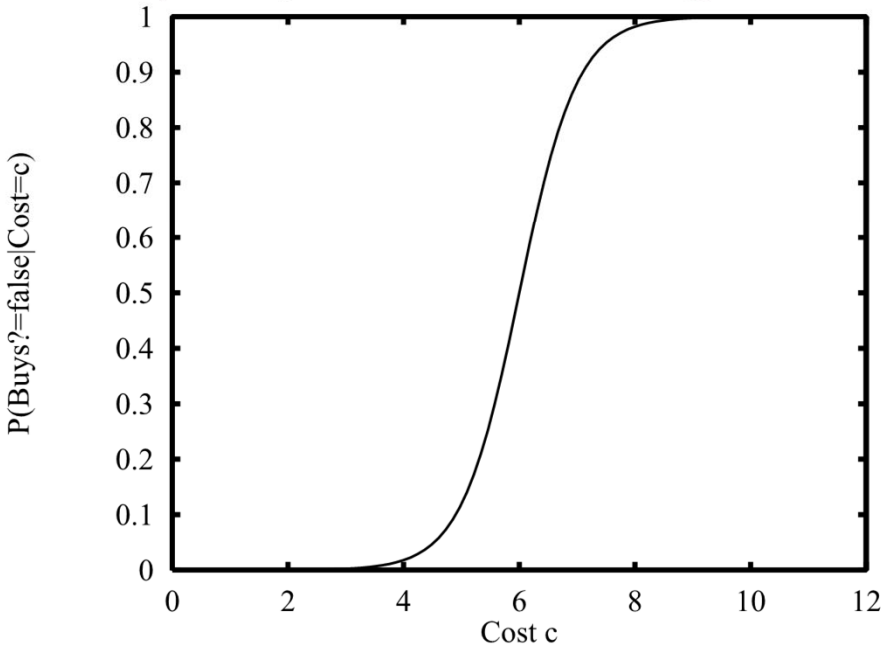
HYBRID (DISCRETE+CONTINUOUS) NETWORKS

ویژگی توزیع لوگیت (Logit) (یا سیگموئید (Sigmoid))

مشابه پروبیت اما دم‌های آن طولانی‌تر است.

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





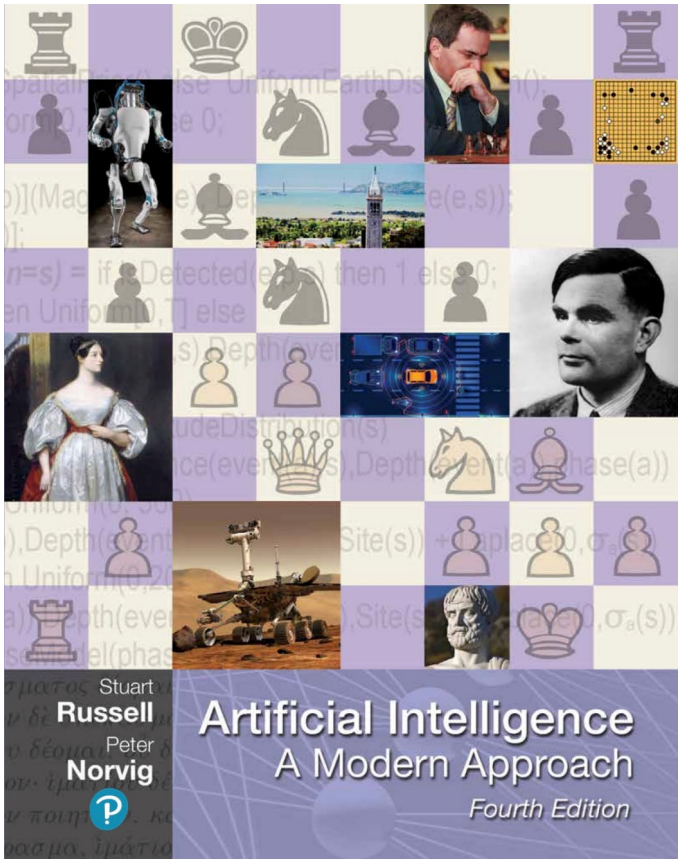
هوش مصنوعی

استدلال احتمالاتی

۴

منابع، مطالعه، تکلیف

منبع اصلی



Stuart Russell and Peter Norvig,
Artificial Intelligence: A Modern Approach,
4th Edition, Prentice Hall, 2020.

Chapter 13