بسمه تعالی

هوش مصنوعی

شبکه های بیزی

نیمسال اول ۱۴۰۱-۱۴۰۰

دکتر مازیار پالهنگ

آزمایشگاه هوش مصنوعی

دانشکدهٔ مهندسی برق و کامپیوتر

دانشگاه صنعتی اصفهان

مقدمه

- شبکه های بیزی روشی برای نمایش گرافیکی استقلال شرطی و مطلق و به دنبال آن مشخص کردن فشرده احتمال توزیع توأم کلی
- دستور
  - مجموعه ای از رئوس: یکی برای هر متغیر تصادفی
  - یک گراف جهتدار غیردوری (یال به معنای نفوذ مستقیم)
  - یک توزیع احتمال شرطی برای هر رأس به شرط داشتن والدین خود

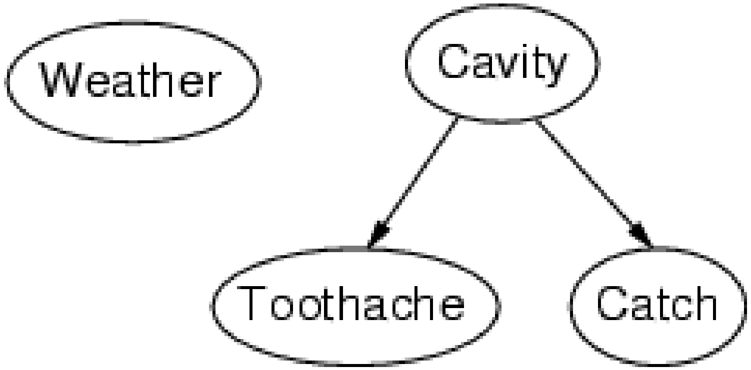
$$P\,(X_i \mid \text{Parents}\,(X_i))$$

در ویرایش جدید: $\theta(X_i|Parents(X_i))$

- در حالت ساده توزیع شرطی بصورت یک جدول احتمال شرطی (conditional probablity table – CPT) نمایش داده می شود که توزیع روی X_i برای هر ترکیب مقادیر والدینش را نمایش نشان می دهد.

مثال

- همبندی شبکه استقلالها را نشان می دهد.

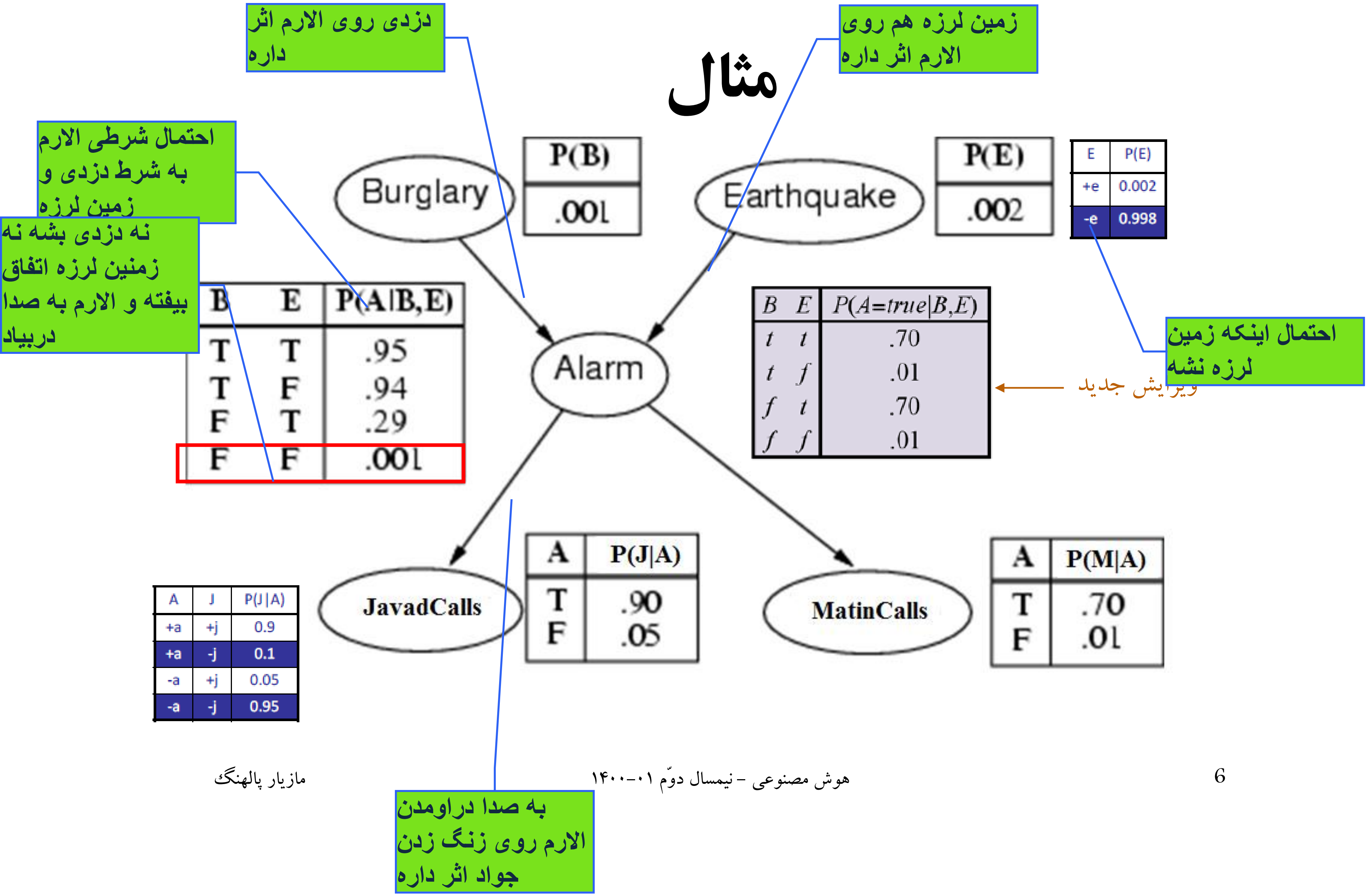


- هوا از دیگر متغیرها مستقل است.
- دندان درد و کشیدن با داشتن کرم خوردگی مستقل هستند.

مثال

- ناصر سرکار است، همسایه اش جواد تلفن زده و می گوید که زنگ دزدگیر او به صدا در آمده است، اما همسایهٔ دیگرش متین تلفن نزده است. زنگ دزدگیر گاهی با زمین لرزه به صدا در می آید. آیا یک دزدی انجام شده است؟
- متغیرها: دزدی، زلزله، زنگ، تلفن جواد، تلفن متین

مثال



P(B)
.001

P(E)
.002

E	P(E)
+e	0.002
-e	0.998

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

B	E	P(A=true|B,E)
t	t	.70
t	f	.01
f	t	.70
f	f	.01

ویرایش جدید

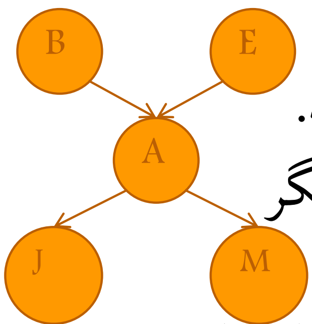
A	P(J|A)
T	.90
F	.05

A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	P(M|A)
T	.70
F	.01

جدول توزیع شرطی

فشردگی



- یک CPT برای متغیر بولی X_i با k ولی بولی دارای 2^k سطر برای ترکیب مقادیر والدین می باشد.
- هر سطر دارای مقدار p برای هنگامی X_i درست می باشد.
- هنگامی که X_i نادرست است مقدار آن $1-p$ است که دیگر نمایش داده نمی شود.
- اگر هر متغیر بیش از k ولی نداشته باشد به $O(n.2^k)$ مقدار نیاز داریم.
- رشد خطی، در صورتی که برای توزیع توأم کلی $O(2^n)$
- برای مثال گفته شده: $1+1+4+2+2=10$ در مقابل $2^5-1=31$

معنا

- توزیع توأم کلی

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

- مثال:

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

$$\begin{aligned} P(j,m,a,\neg b,\neg e) &= P(j|a)P(m|a)P(a|\neg b \wedge \neg e)P(\neg b)P(\neg e) \\ &= 0.90 \times 0.70 \times 0.01 \times 0.999 \times 0.998 = 0.00628. \end{aligned}$$

$\longleftarrow$ ویرایش جدید

ساخت شبکه های بیزی

- قانون ضرب:

$$P(x_1,\ldots,x_n) = P(x_n \mid x_{n-1},\ldots,x_1)P(x_{n-1},\ldots,x_1)$$

- قانون زنجیری:

$$\begin{aligned} P(x_1,\ldots,x_n) &= P(x_n \mid x_{n-1},\ldots,x_1)P(x_{n-1} \mid x_{n-2},\ldots,x_1) \cdots P(x_2 \mid x_1)P(x_1) \\ &= \prod_{i=1}^{n} P(x_i \mid x_{i-1},\ldots,x_1)\,. \end{aligned}$$

- مقایسه:

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

- معادل با

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

- اگر

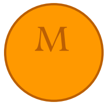
$$Parents(X_i) \subseteq \{X_{i-1},\ldots,X_1\}$$

- مشخص کردن متغیرهای تصادفی لازم
- مرتب کردن $X_1 \ldots X_n$
- هر ترتیبی کار می کند ولی بهتر است ابتدا سببها بعد آثار
- برای $i=1$ تا n
  - اضافه کردن رأس X_i
  - انتخاب ولیها از میان X_1 تا X_{i-1} بطوریکه

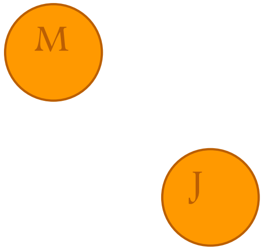
$$\mathbf{P}(X_i \mid X_{i-1}, \ldots, X_1) = \mathbf{P}(X_i \mid Parents(X_i))$$

  - رسم یالها
  - نوشتن توزیع

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچید گی را افزایش دهد.
- فرض: ترتیب متغیرها M، J، A، B، E

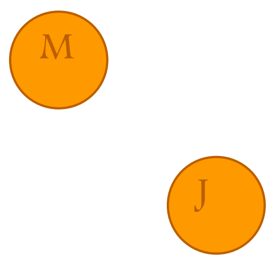


- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها E، B، A، J، M



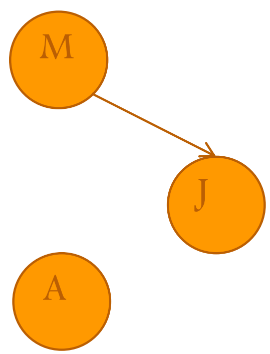
- $p(J|M)=p(J)$؟

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها M، J، A، B، E



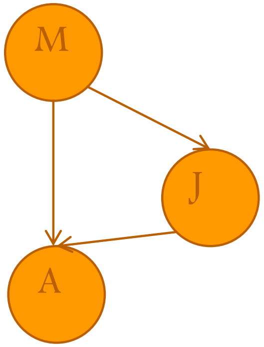
- $p(J|M)=p(J)$؟ نه

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها M، J، A، B، E



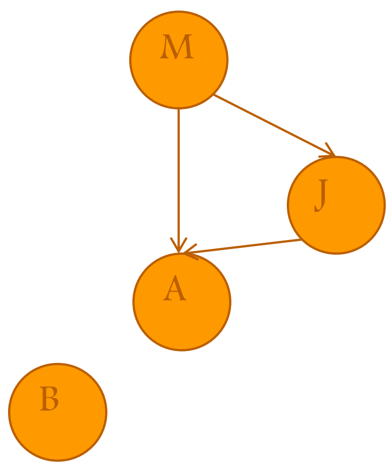
- $p(J \mid M) = p(J)$؟ نه
- $P(A \mid J, M) = P(A \mid J)$؟ $P(A \mid J, M) = P(A)$؟

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها M، J، A، B، E



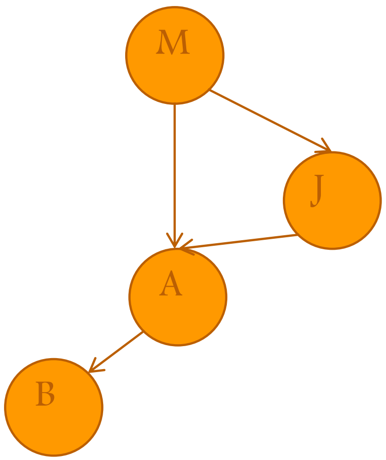
- $p(J|M)=p(J)$؟ نه
- $P(A \mid J, M) = P(A \mid J)$? $P(A \mid J, M) = P(A)$? نه

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها M، J، A، B، E



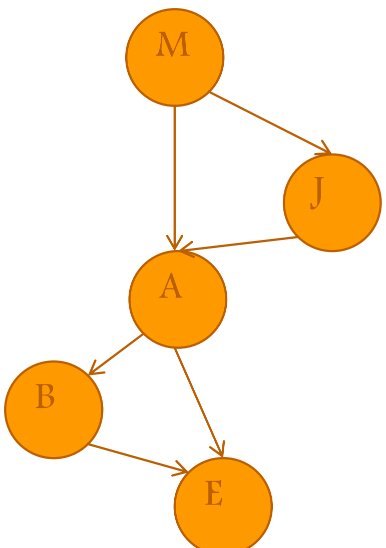
- $p(J|M)=p(J)$؟ نه
- $P(A \mid J, M) = P(A \mid J)$? $P(A \mid J, M) = P(A)$? نه
- $P(B \mid A, J, M) = P(B \mid A)$?
- $P(B \mid A, J, M) = P(B)$?

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها E، B، A، J، M



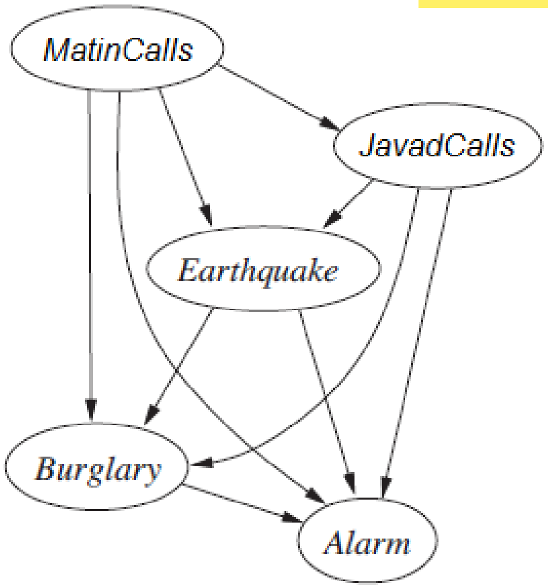
- $p(J|M)=p(J)$؟ نه
- $P(A \mid J, M) = P(A \mid J)$؟ $P(A \mid J, M) = P(A)$؟ نه
- $P(B \mid A, J, M) = P(B \mid A)$؟ بله
- $P(B \mid A, J, M) = P(B)$؟ نه

- اگر ترتیب خوب انتخاب نشده باشد می تواند پیچیدگی را افزایش دهد.
- فرض: ترتیب متغیرها M، J، A، B، E



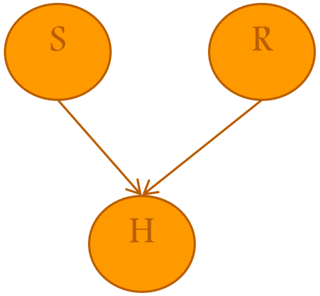
- $p(J \mid M)=p(J)$؟ نه
- $P(A \mid J, M) = P(A \mid J)$؟ $P(A \mid J, M) = P(A)$؟ نه
- $P(B \mid A, J, M) = P(B \mid A)$؟ بله
- $P(B \mid A, J, M) = P(B)$؟ نه
- $P(E \mid B, A, J, M) = P(E \mid A)$؟ نه
- $P(E \mid B, A, J, M) = P(E \mid A, B)$؟ بله

- اگر سعی کنیم مدل تشخیصی بجای مدل سببی طراحی کنیم مجبور به مشخص کردن وابستگیهای اضافی خواهیم شد.
- یک ترتیب بد متغیرها
- احتیاج به مشخص کردن ۳۱ احتمال
- به اندازه توزیع توأم کلی

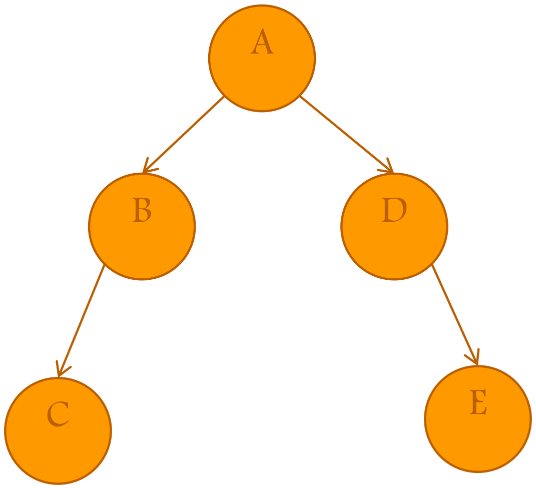


توضیح دادن explaining away

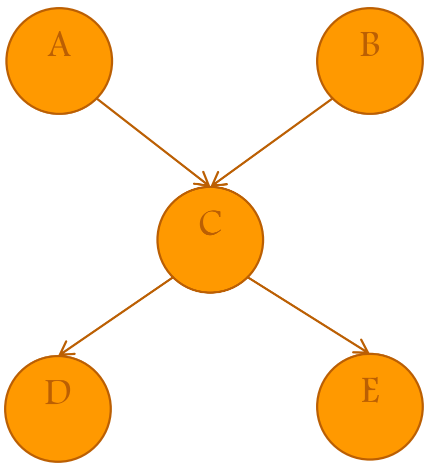
- اثر مشترک



جدائی d (d-separation)



- A ، C
- C، A|B
- C،D
- C،D|A
- E،C
- E،C|A
- **در این حالت**: دو متغیر از هم مستقل هستند اگر بوسیله متغیرهای ناشناس به هم مرتبط نباشند.



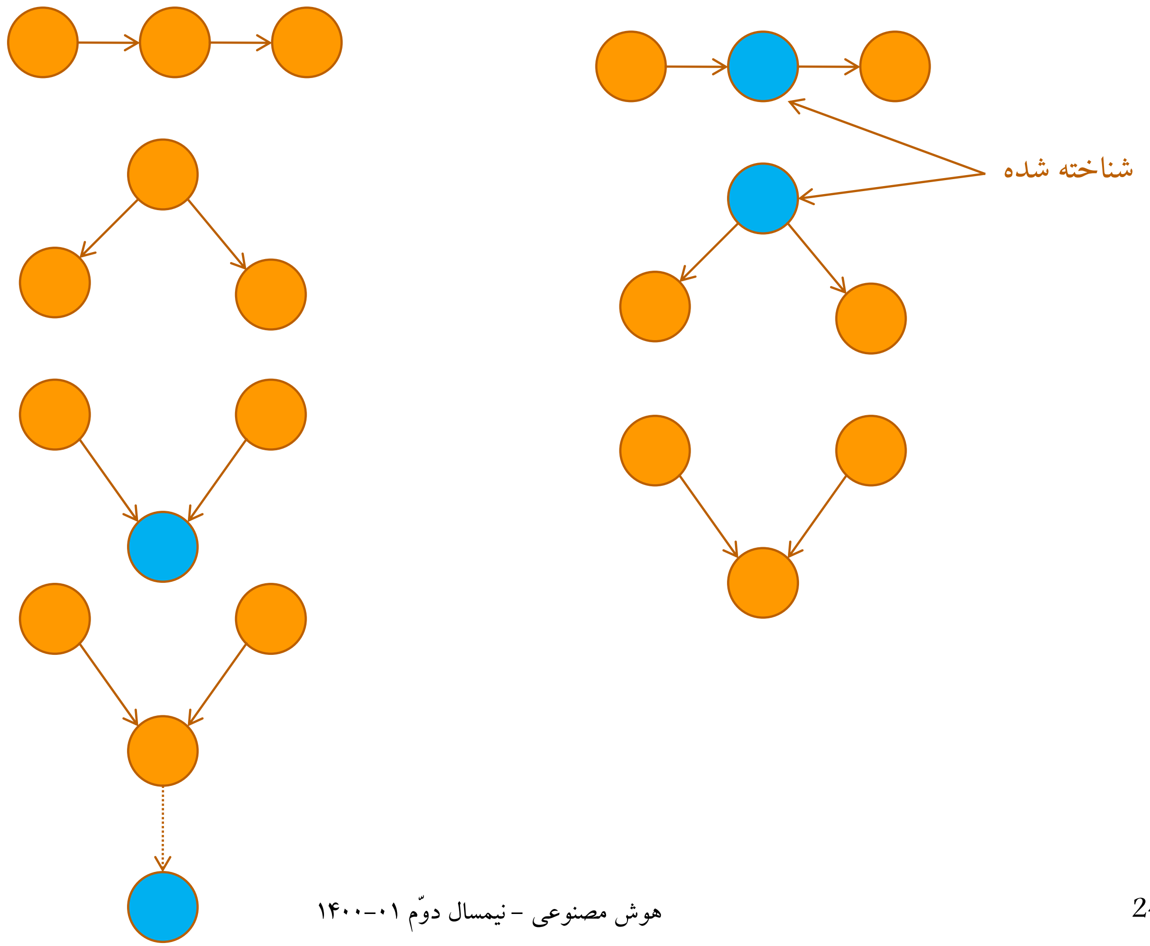
- A،C
- A،E
- A،E | B
- A،E | C
- A،B
- A،B | C

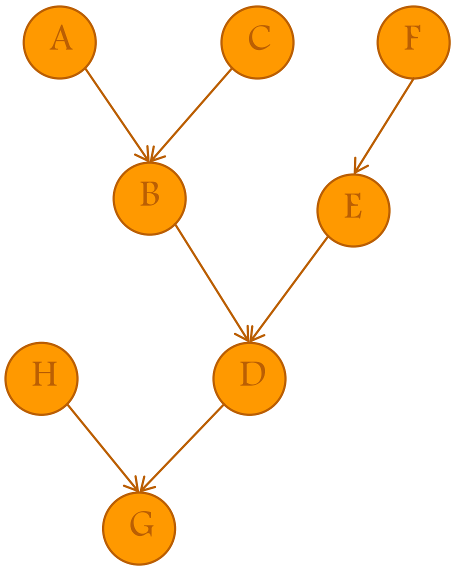
جدائی D (D-separation)

- بررسی استقلال متغیرها از روی گراف
- بررسی سه تائیهای در طول مسیر بین متغیرها
- سه تائیهای فعال = وابسته
- سه تائیهای غیرفعال = غیروابسته

فعال (وابسته)

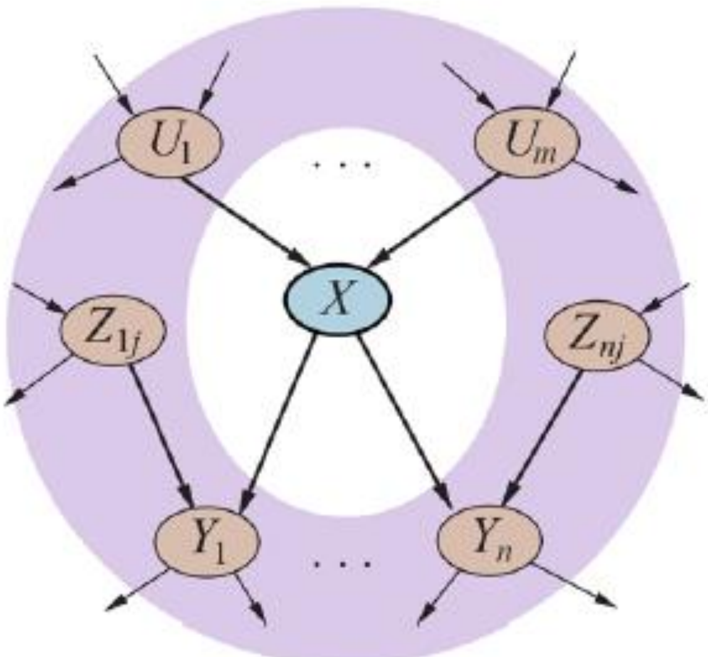
غیر فعال (غیر وابسته)





- F،A
- F،A | D
- F،A | G
- F،A | H

25



- هر رأس بصورت شرطی از همهٔ دیگر رئوس مستقل است به شرط داشتن پتوی مارکف (Markov blanket) خود
- پتوی مارکف: والدین+ فرزندان + والدین فرزندان

استنتاج با فهرست کردن

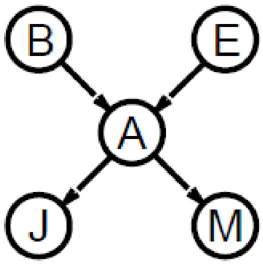
- ممکن است مایل باشیم توزیع احتمال یک متغیر (متغیر سؤال) را به شرط داشتن برخی از متغیرهای دیگر (متغیرهای دلیل) بدست آوریم.

$$\mathbf{P}(X \mid \mathbf{e})$$

$$\mathbf{P}(X \mid \mathbf{e}) = \alpha\, \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y})\,.$$

استنتاج با فهرست کردن

- در مثالی که داشتیم:



$$\begin{aligned}&\mathbf{P}(B|j,m)\\&= \mathbf{P}(B,j,m)/P(j,m)\\&= \alpha\mathbf{P}(B,j,m)\\&= \alpha\ \textstyle\sum_e \sum_a \mathbf{P}(B,e,a,j,m)\end{aligned}$$

- با استفاده از گراف:

$$\begin{aligned}&\mathbf{P}(B|j,m)\\&= \alpha\ \textstyle\sum_e \sum_a \mathbf{P}(B)P(e)\mathbf{P}(a|B,e)P(j|a)P(m|a)\\&= \alpha\mathbf{P}(B)\ \textstyle\sum_e P(e)\ \sum_a \mathbf{P}(a|B,e)P(j|a)P(m|a)\end{aligned}$$

$$\mathbf{P}(B\mid j,m) = \alpha\,\langle 0.00059224, 0.0014919\rangle \approx \langle 0.284, 0.716\rangle$$

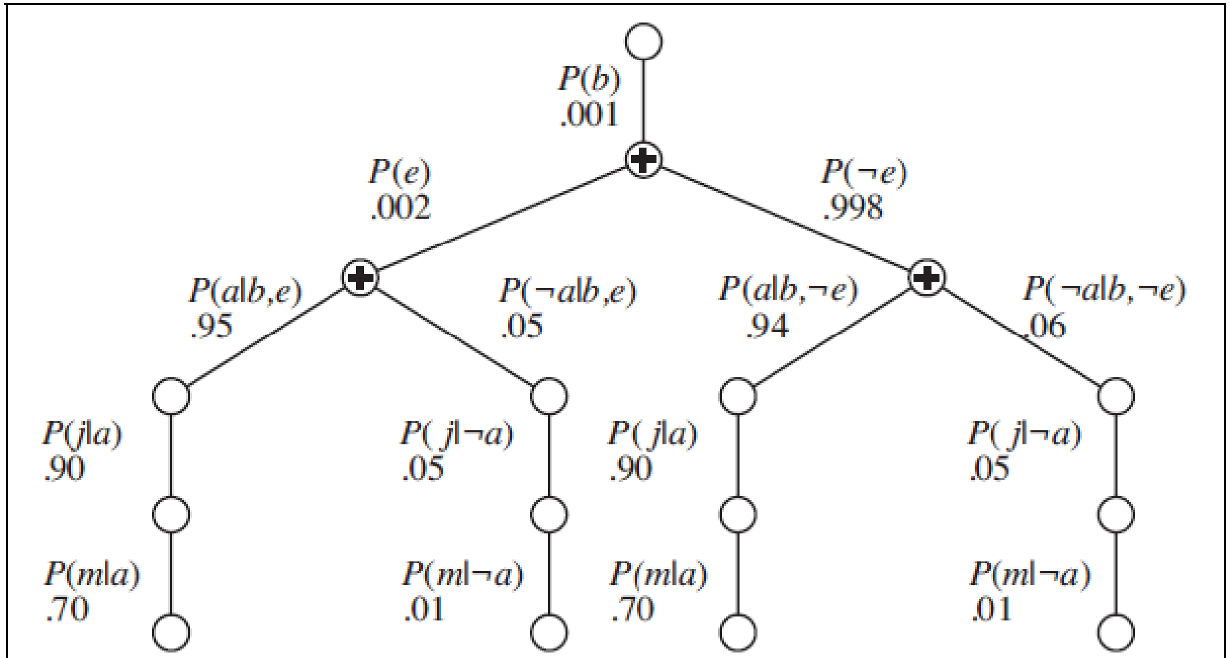


Figure 14.8 The structure of the expression shown in Equation (14.4). The evaluation proceeds top down, multiplying values along each path and summing at the "+" nodes. Notice the repetition of the paths for j and m.

- .

در ویرایش جدید کتاب فراموش کرده که برخی از مقادیر را بروز کند!

اصفهان – بوستان شهرستان

- دقت نمائید که پاورپوینت ابزاری جهت کمک به یک ارائهٔ شفاهی می باشد و **به هیچ وجه یک جزوهٔ درسی نیست** و شما را از خواندن مراجع درس بی نیاز نمی کند.
- لذا حتماً مراجع اصلی درس را مطالعه نمائید.
- در تهیه این اسلایدها، از اسلایدهای سایت کتاب و برخی منابع از اینترنت استفاده شده است.