دانشگاه صنعتی شریف
دانشکده‌ی مهندسی کامپیوتر

درس هوش مصنوعی

استاد محمدحسین رهبان

تمرین چهارم | شبکه‌ی بیز | مصطفی قدیمی

سؤال ۱. تعداد بیشینه‌ی یال‌ها

می‌دانیم که شبکه‌ی بیز یک گراف جهت‌دار بدون دور است. بنابراین باید تعداد بیشینه‌ی یال‌هایی را که بدون تشکیل دور در آن می‌توان داشت محاسبه کنیم. تعداد آن برابر با $n(n-1)/2$ است.

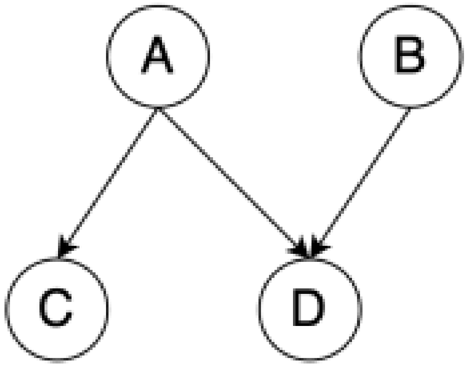
اثبات:

یک شبکه‌ی بیز روی $X_1, \ ..., \ X_n$ در نظر بگیرید. بین هر $X_i, \ X_j$ در صورتی که $i < j$ باشد، یک یال داریم. تعداد کل یال‌های این گراف برابر است با:

$$1 + 2 + ... + (n-2) + (n-1) = n(n-1)/2$$

برای این‌که نشان دهیم چنین شبکه‌ای وجود دارد، باید اثبات کنیم دور جهت‌داری در این گراف وجود ندارد. با فرض خلف، تصور می‌کنیم که یک دور به فرم $X_{i1}, \ X_{i2}, \ ..., \ X_{im}, \ X_{i1}$ وجود دارد. اما به کمک فرضیات هنگام ساخت گراف، می‌دانیم که $i_1 < i_2 ... < i_m < i_1$ که بیانگر $i_1 < i_1$ است و تناقض است. بنابراین ثابت می‌شود که دوری (جهت‌دار) وجود ندارد.
این مقدار بیشینه نیز هست؛ زیرا هر گراف جهت‌دار با داشتن بیش‌از $n(n-1)/2$ یال باید حداقل از یکی از رئوس بیش‌از یک یال داشته باشد. این به این معناست که حداقل یک یال در هر دو جهت دارد که منجر به تشکیل یک دور می‌شود.
برای متغیر A این تعداد برابر با چهار است.

سؤال ۲. CPT



شکل ۱: شبکه‌ی بیز

$$P(A, B, C, D) = P(A)\ P(B)\ P(C|A)\ P(D|A,\ B)$$

- آ) کم‌ترین تعداد مقادیر احتمالاتی برای نمایش جدول‌های این شبکه، ۱۱ است.

برای متغیر B، ۲ مقدار کافی است. زیرا هم پدری ندارد و هم طبق قوانین احتمالات باید جمع احتمال‌های مربوط به یک متغیر برابر با ۱ شود؛ بنابراین نیازی به حالت سوم نیست و از متمم‌گیری به دست می‌آید.

برای متغیر A هم مشابه توضیحاتی که برای متغیر B دادیم صدق می‌کند و حداقل به ۱ مقدار برای آن نیاز داریم.

برای احتمال $P(C|A)$ حداقل ۲ مقدار نیاز است. زیرا هر کدام از Posterior و Priorاش، ۲ حالت دارند. برای هر حالت ثابتی از Posteriorها باید جمع احتمالات Prior برابر با یک شود.

برای احتمال $P(D|A,\ B)$ حداقل ۶ مقدار نیاز است. زیرا A، ۲ مقدار و B، ۳ مقدار دارد؛ بنابراین طبق اصل ضرب $2 \times 3 = 6$ حالت دارند که چون متغیر D ۲ حالت دارد، برای هر کدام از آن‌ها ۱ حالت کافی است و دیگری با توجه به قوانین احتمال و متمم‌گیری به دست می‌آید.

- ب)

$$P(C|d_1) = \frac{P(C, d_1)}{P(d_1)} = \frac{\Sigma_{a\in A}\Sigma_{b\in B}P(A, B, C, d_1)}{P(d_1)} = \frac{\Sigma_{a\in A}\Sigma_{b\in B}P(A)P(B)P(C|A)P(d_1|A, B)}{P(d_1)}$$

- ج)

$$C \perp\!\!\!\perp D|A,\ C \perp\!\!\!\perp B|A$$

سؤال ۳. استقلال و شبکه‌ی بیز

برای متغیر X باید مقدار آن را ثابت در نظر گرفته و روی مقادیر دیگر جمع بزنیم.

X	P
0	0.6
1	0.4

همانند آن‌چه برای X داشتیم، برای Y هم داریم.

Y	P
0	0.7
1	0.3

همانند آن‌چه برای X داشتیم، برای Z هم داریم.

Z	P
0	0.332
1	0.668

حال برای به دست آوردن احتمال X, Y باید روی مقادیر Z جمع بزنیم.

X	Y	P
0	0	0.42
0	1	0.18
1	0	0.28
1	1	0.12

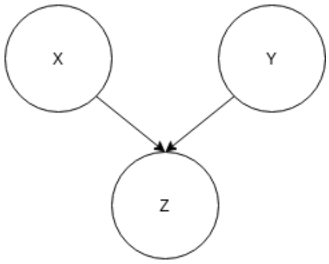
با توجه به $P(X)P(Y) = P(X,Y)$ است به این نتیجه می‌رسیم که X و Y مستقل از یک‌دیگرند.
حال برای به دست آوردن احتمال X, Z باید روی مقادیر Y جمع بزنیم.

X	Z	P
0	0	0.096

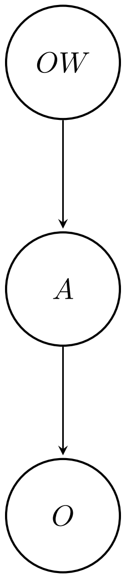
با توجه به $P(X)P(Z) \neq P(X,Z)$ به این نتیجه می‌رسیم که دو متغیر X و Z از یک‌دیگر مستقل نیستند.
ال برای به دست آوردن احتمال Y, Z باید روی مقادیر Z جمع بزنیم. بنابراین نیازی نیست بقیه سطرهای جدول را محاسبه کنیم.

Y	Z	P
0	0	0.182

با توجه به $P(Y)P(Z) \neq P(Y,Z)$ به این نتیجه می‌رسیم که دو متغیر Y و Z از یک‌دیگر مستقل نیستند.به همین خاطر نیاز نیست بقیه سطرهای جدول را محاسبه کنیم.
فقط یک شبکه بیز وجود دارد.



سؤال ۴. احتمال



OW: وزن اضافی، A: صدای آژیر و O: خاموش شدن آسانسور

$$P(OW|\sim a, o) = P(o|\sim a) \times (P(a|ow)P(ow) + P(a|\sim ow)P(\sim ow))$$

$$P(OW|\sim a, o) = 0.02\,(0.2 \times 0.1 + 0.9 \times 0.9) = 0.166$$

سؤال ۵.

- آ) c. زیرا بیانگر این است که هر سه ژن فرزند، پدر و مادر از یک‌دیگر مستقل هستند؛ بنابراین نباید بین هیچ‌کدام از آن‌ها یالی باشد.
- ب) a. با وجود این‌که a و b هر دو درست هستند اما در b یال‌های اضافی بین متغیرهای پنهان که نشان‌دهنده‌ی رابطه‌ی علّی مستقیم نیستند، وجود دارد.
- ج)

G_{mother}	G_{father}	$P(G_{child} = l \| G_{mother}, G_{father})$	$P(G_{child} = r \| G_{mother}, G_{father})$
l	l	$1 - m$	m
l	r	0.5	0.5
r	l	0.5	0.5
r	r	m	$1 - m$

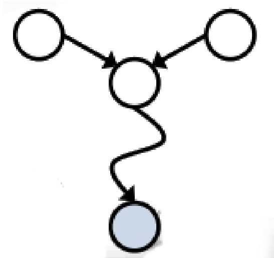
- د)

$$P(G_{child} = l) = \Sigma_{g_m, g_f} P(G_{child} = l | g_m, g_f) P(g_m, g_f) = \Sigma_{g_m, g_f} P(G_{child} = l | g_m, g_f) P(g_m) P(g_f)$$

$$= (1 - m)q^2 + 0.5q(1 - q) + 0.5(1 - q)q + m(1 - q)^2 = m + q - 2mq$$

سؤال ۶. استقلال

- خیر مستقل نمی‌باشند؛ زیرا فقط یک مسیر دارند که آن هم از مسیرهای فعال است.



- خیر، مستقل نمی‌باشند؛ زیرا بین F، G و E رابطه‌ی Common Cause وجود دارد.
- متغیر C

سؤال ۷.

با توجه به قاعده‌ی موجود در شبکه‌های بیز داریم:

$$P(A,B,C,D,E) = P(A)P(B|A)P(E|A,B)P(D|E,C)P(C)$$

برای به دست آوردن $P(+D)$ داریم:

$$P(+D) = \Sigma_{a\in A}\Sigma_{b\in B}\Sigma_{c\in C}\Sigma_{e\in E}P(A)P(B|A)P(E|A,B)P(+D|E,C)P(C)$$

با توجه الگوریتم Variable Elimination باید کاری کنیم که سیگماهایی که به یک‌دیگر وابسته نیستند را از هم جدا کنیم. با توجه به CPTهای داده شده داریم:

۱. B را حذف می‌کنیم:

$P(+E,+A|+B) = 0.3,\ P(+E,+A|-B) = 0.1,\ P(+E,-A|+B) = 0.2,\ P(+E,-A|-B) = 0.5$
$\rightarrow P(+E|+A) = 0.14,\ P(+E|-A) = 0.38$

۲. A را حذف می‌کنیم:

$P(+E,+A) = P(+E|+A)P(+A) = 0.098,\ P(+E,-A) = P(+E|-A)P(-A) = 0.114$
$\rightarrow P(+E) = 0.212$

۳. E را حذف می‌کنیم:

$P(+D,+E|+C) = 0.0212,\ P(+D,+E|-C) = 0.106,\ P(+D,-E|+C) = 0.106,\ P(+D,-E|-C) = 0.3152$
$\rightarrow P(+D|+C) = 0.3364, P(+D|-C) = 0.4212$

۴. C را حذف می‌کنیم:

$P(+D,+C) = 0.1682,\ P(+D,-C) = 0.2106$
$\rightarrow P(+D) = 0.3788$

برای به دست آوردن مقدار $P(+E|-C)$ با توجه به این که از یک‌دیگر مستقل هستند، بنابراین داریم:

$P(+E|-C) = P(+E) = 0.212$