

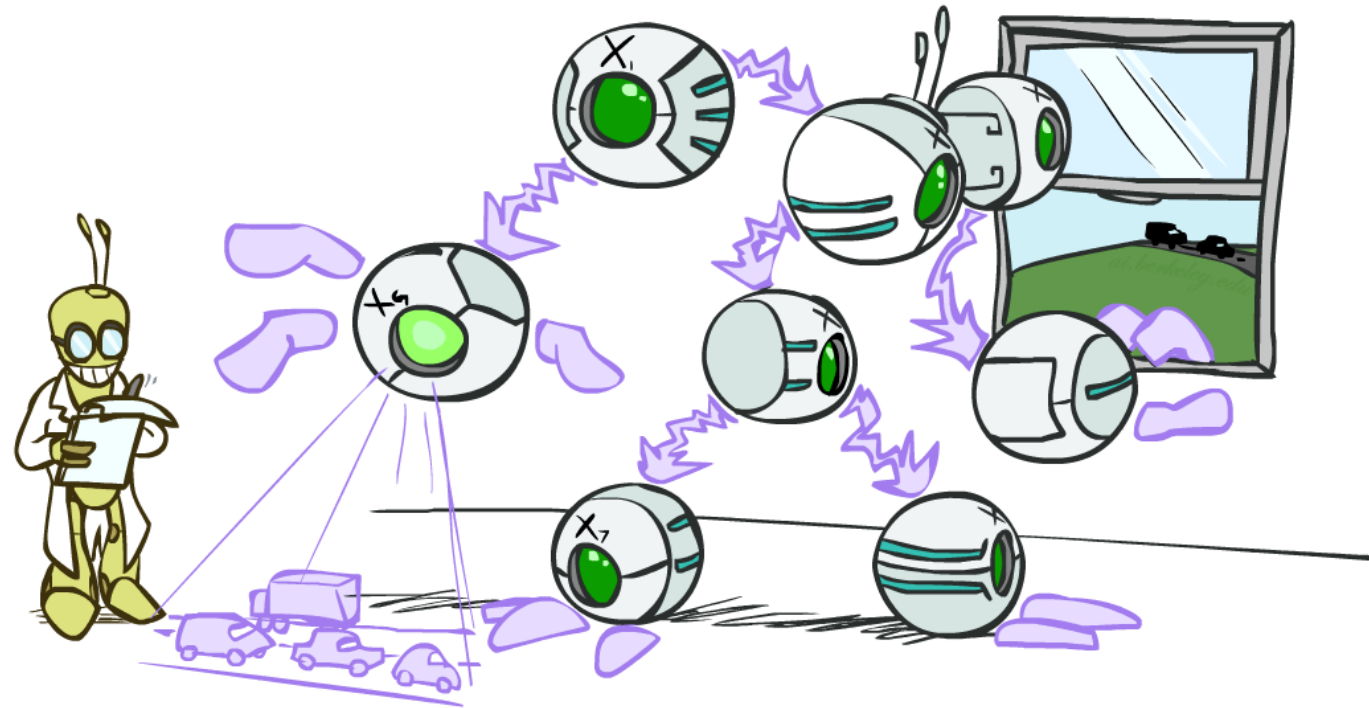
شبکه‌های بیز: استنتاج

سید ناصر رضوی n.razavi@tabrizu.ac.ir

۱۳۹۵

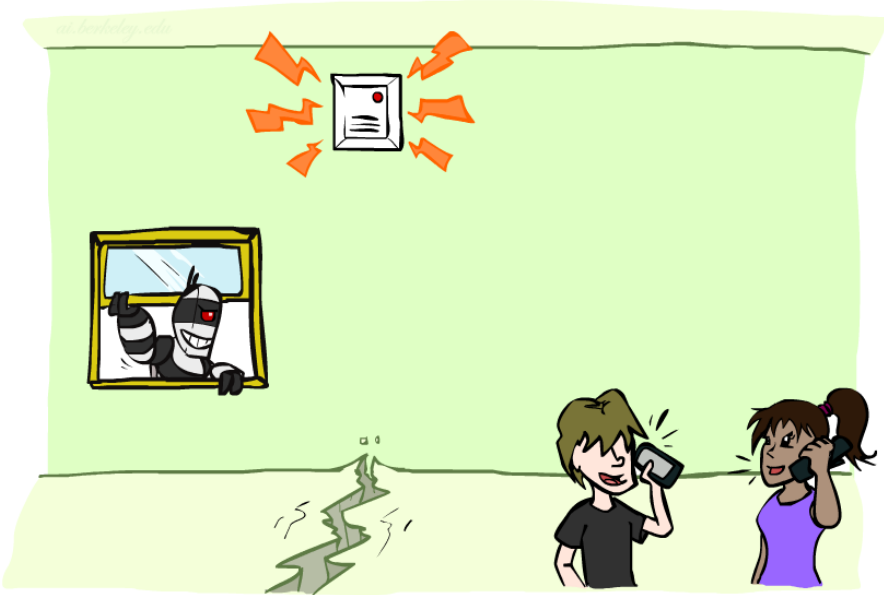
شبکه‌های بی‌سیم: استنتاج

۲

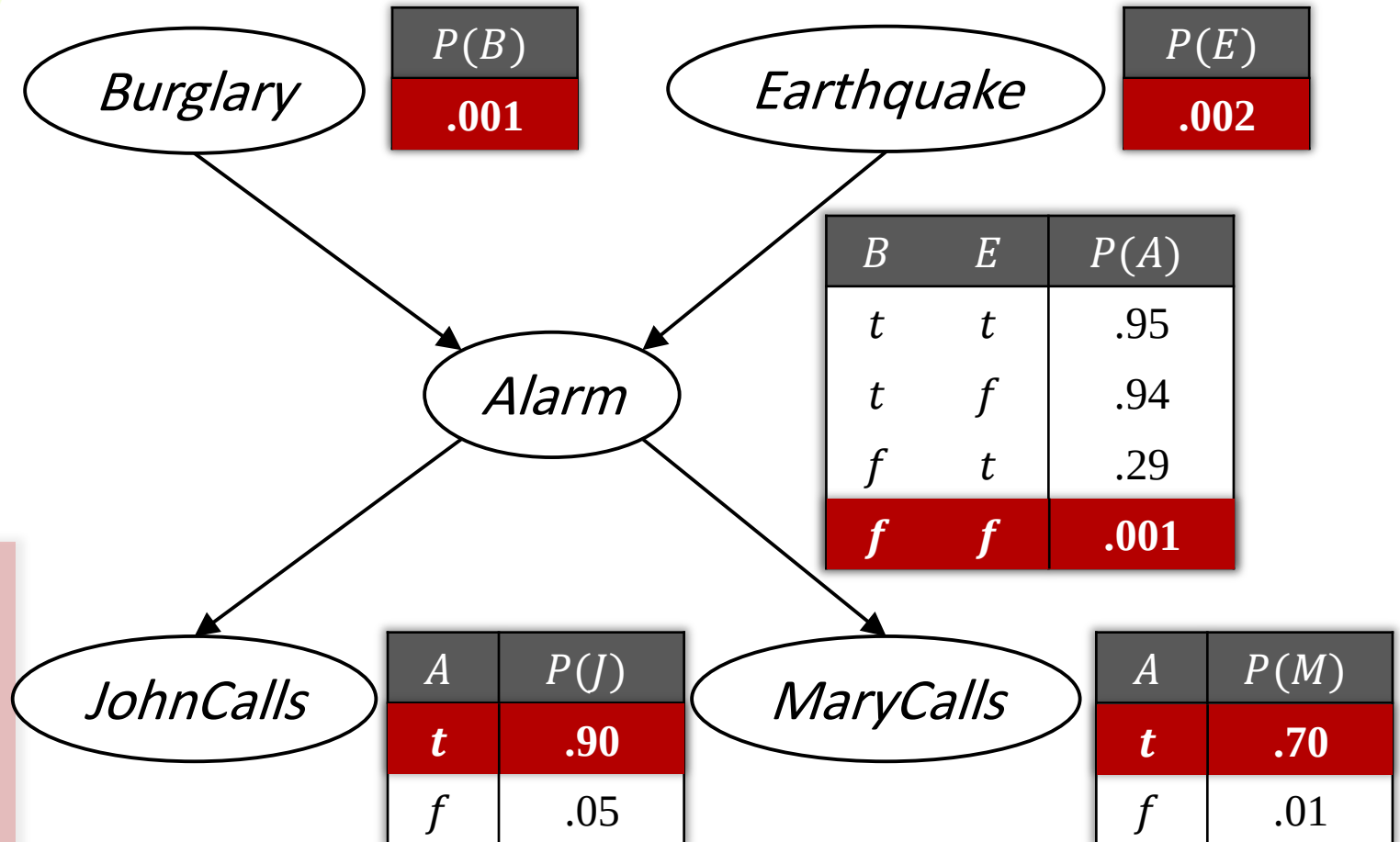


مثال: زنگ هشدار

۴



$$\begin{aligned}
 &P(j, m, a, \neg b, \neg e) \\
 &= P(j|a)P(m|a)P(a|\neg b, \neg e)P(\neg b)P(\neg e) \\
 &= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.988 \\
 &= 0.000628
 \end{aligned}$$



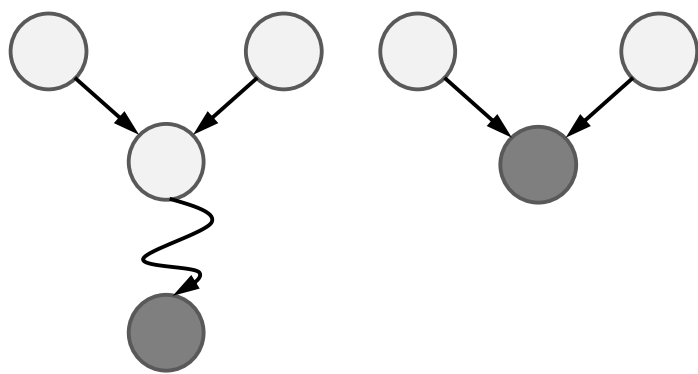
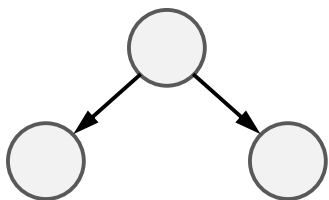
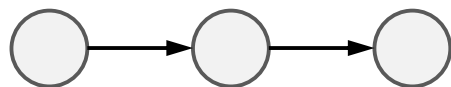
استقلال در شبکه‌های بیز (یادآوری)

۵

□ قاعده‌ی زنجیری. به ازای هر توزیع توأم داریم:

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | x_{i-1}, \dots, x_1)$$

□ فرضیات در نظر گرفته شده با تعریف یک شبکه‌ی بیز.



سه تایی‌های فعال

$$P(x_i | x_{i-1}, \dots, x_1) = P(x_i | \text{parents}(x_i))$$

□ D -جدایی. استخراج استقلال‌های شرطی از گراف.

□ دو متغیر d -جدا هستند اگر و فقط اگر هیچ مسیر فعالی بین آنها وجود نداشته باشد.

□ مسیر فعال، مسیری است که تمام سه تایی‌های موجود در آن فعال باشند.

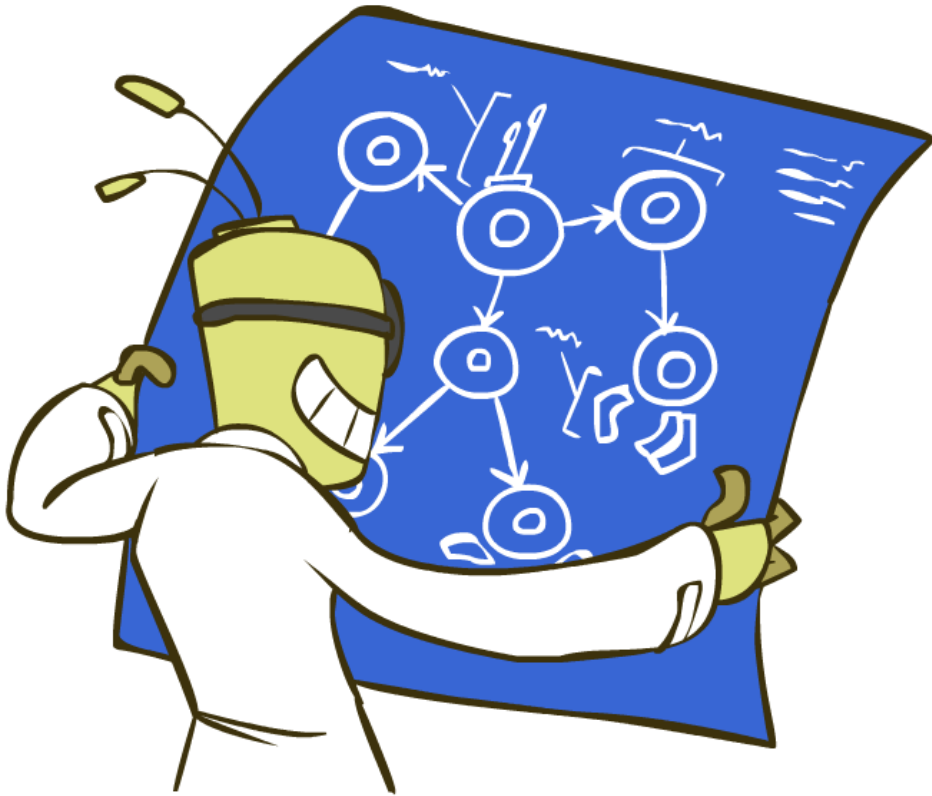
تعیین استقلال‌های شرطی (یادآوری)

۶

□ با داشتن یک شبکه‌ی بیز، می‌توان یک لیست کامل از تمامی استقلال‌های شرطی به صورت زیر ایجاد نمود:

$$X_i \perp X_j \mid \{X_{k_1}, X_{k_2}, \dots, X_{k_n}\}$$

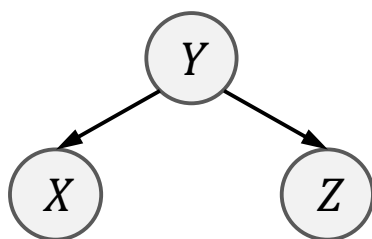
□ این لیست بیانگر مجموعه‌ی تمام توزیع‌های احتمالاتی است که به وسیله‌ی شبکه‌ی بیز قابل بازنمایی هستند.



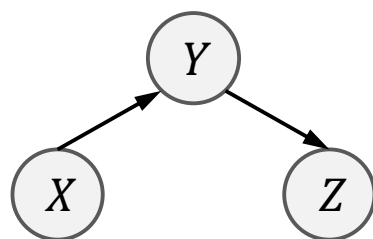
محاسبه‌ی استقلال‌های شرطی (یادآوری)

۷

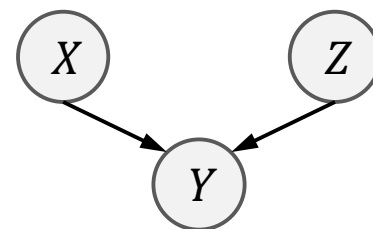
□ محاسبه‌ی تمام استقلال‌های شرطی.



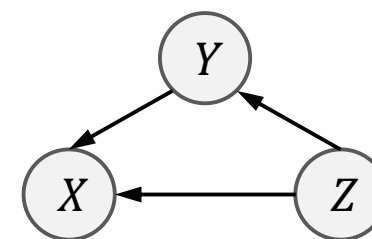
$\{X \perp Z \mid Y\}$



$\{X \perp Z \mid Y\}$



$\{X \perp Z\}$



$\{\}$

□ **توجه.** افزودن یالها باعث می‌شود تعداد توزیع‌های قابل بازنمایی به وسیله‌ی شبکه‌ی بیز بیشتر شود، اما هزینه‌ی استنتاج نیز افزایش می‌یابد.

استنتاج در شبکه‌های بیز

۸

□ استنتاج احتمالاتی.

□ استنتاج یک مسئله‌ی **ان پی-کامل** است.

□ روش‌های استنتاج دقیق.

□ استنتاج به وسیله‌ی شمارش

■ پیچیدگی نمایی!

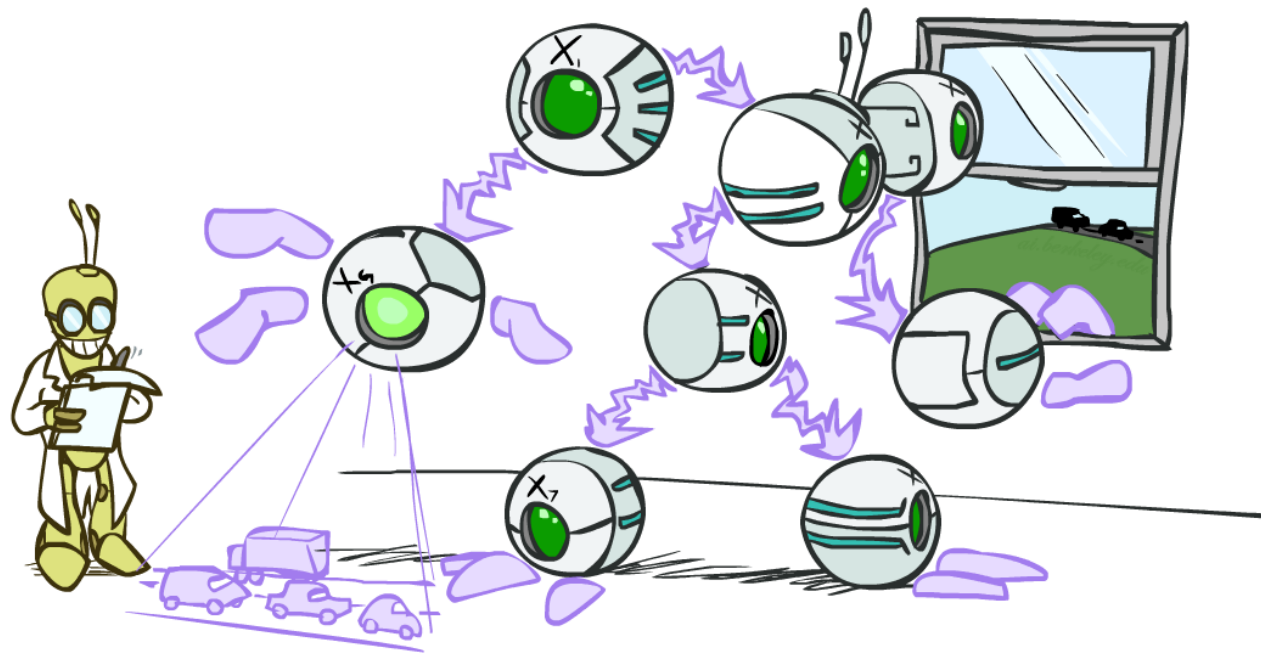
□ حذف متغیر

■ بدترین حالت نمایی، اما اغلب بسیار بهتر از نمایی!

□ روش‌های استنتاج تقریبی.

□ نمونه‌برداری!

□ یادگیری شبکه‌ی بیز از روی داده‌ها.

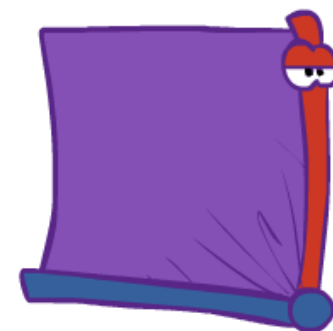
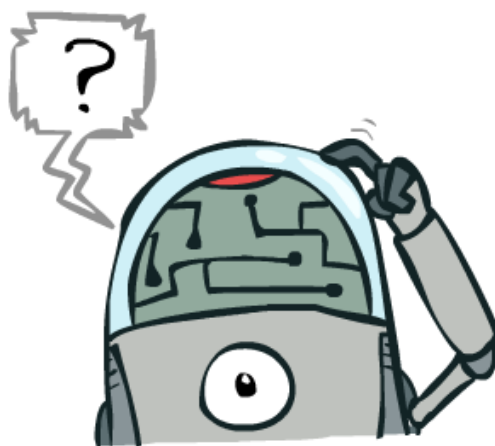
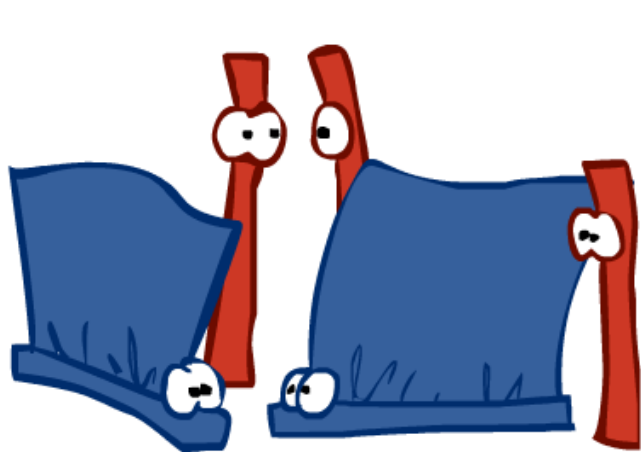


- استنتاج. محاسبه‌ی یک کمیت مفید از روی
- احتمال پسین

$$P(Q|E_1 = e_1, \dots, E_k = e_k)$$

- محتمل‌ترین توضیح

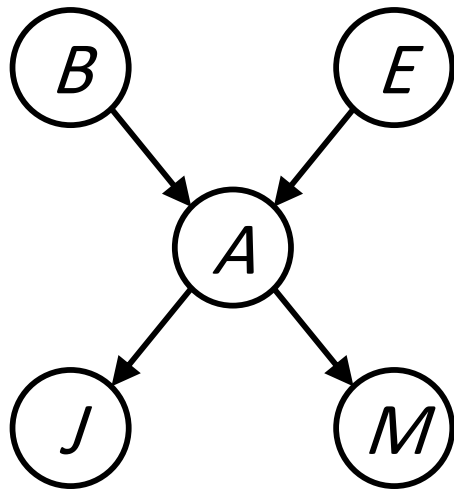
$$\arg \max_q P(Q = q|E_1 = e_1, \dots, E_k = e_k)$$



استنتاج به وسیله‌ی شمارش

۱۰

□ با فرض داشتن زمان نامحدود، استنتاج در شبکه‌های بیز بسیار ساده است.



□ دستورالعمل.

□ تعیین توزیع‌های حاشیه‌ای مورد نیاز

□ مشخص کردن تمام احتمال‌های اتمیک مورد نیاز

□ محاسبه و ترکیب

□ مثال. محاسبه‌ی $P(+b|+j, +m)$

$$P(+b|+j, +m) = \frac{P(+b, +j, +m)}{P(+j, +m)}$$

$$P(+b, +j, +m) = \sum_A \sum_E P(+b, +j, +m, A, E)$$

استنتاج به وسیله شمارش

۱۱

$$P(+b|+j, +m) = \frac{P(+b, +j, +m)}{P(+j, +m)}$$

$$P(+b, +j, +m) = \sum_A \sum_E P(+b, +j, +m, A, E)$$

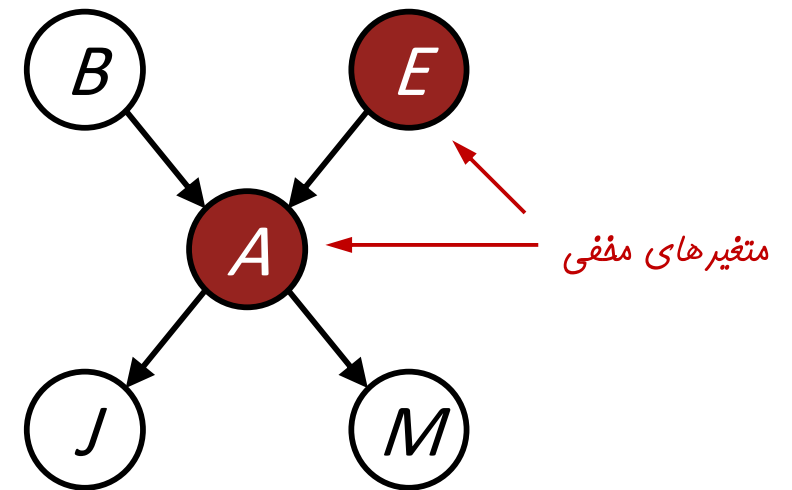
$$= P(+b, +j, +m, +a, +e) \longrightarrow P(+b)P(+e)P(+a|+b, +e)P(+j|+a)P(+m|+a)$$

$$+ P(+b, +j, +m, -a, +e) \longrightarrow P(+b)P(+e)P(-a|+b, +e)P(+j|-a)P(+m|-a)$$

$$+ P(+b, +j, +m, +a, -e) \longrightarrow P(+b)P(-e)P(+a|+b, -e)P(+j|+a)P(+m|+a)$$

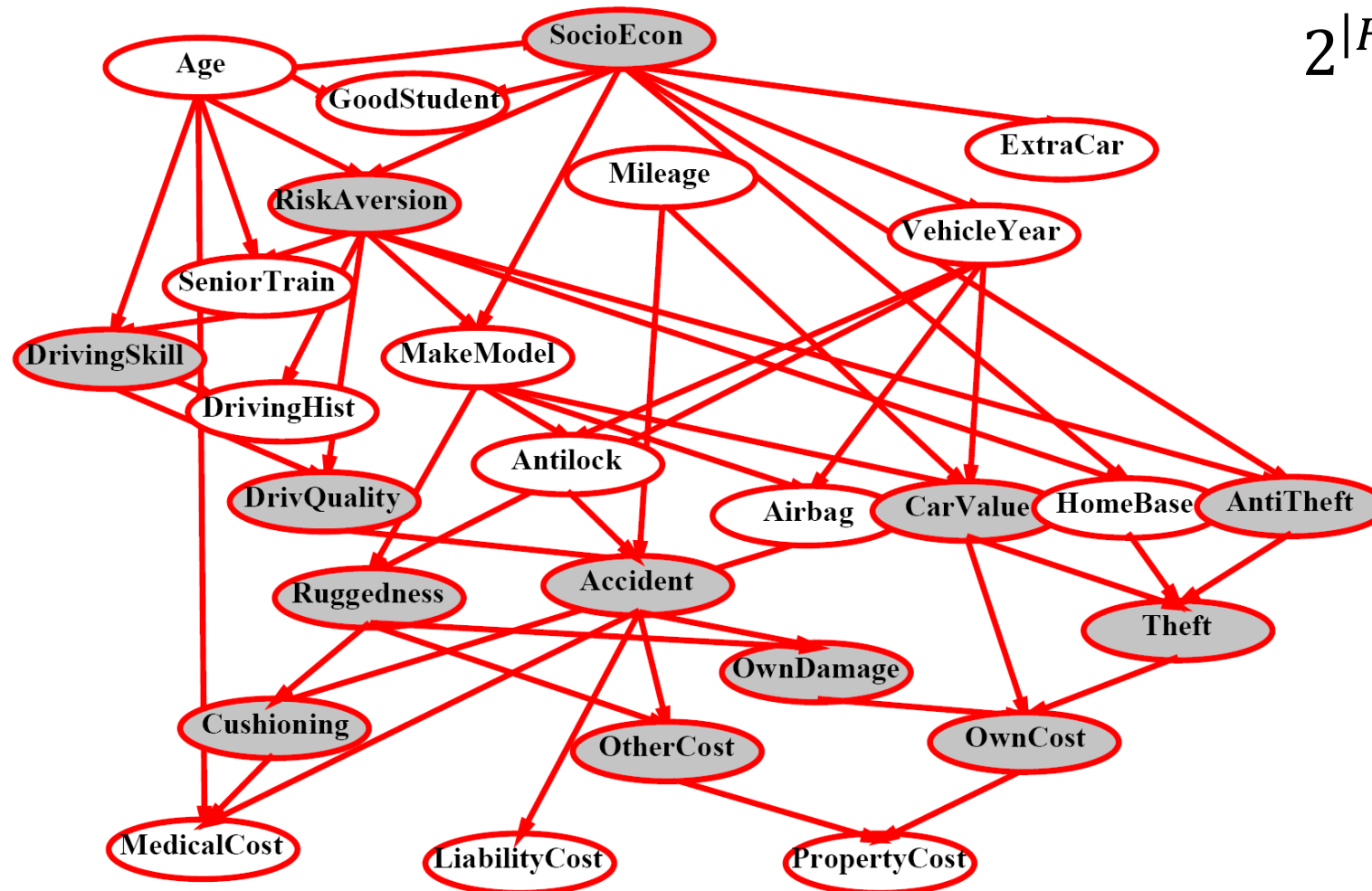
$$+ P(+b, +j, +m, -a, -e) \longrightarrow P(+b)P(-e)P(-a|+b, -e)P(+j|-a)P(+m|-a)$$

□ مثال. محاسبه‌ی $P(+b|+j, +m)$



استنتاج به وسیلهی شمارش: تحلیل

۱۲



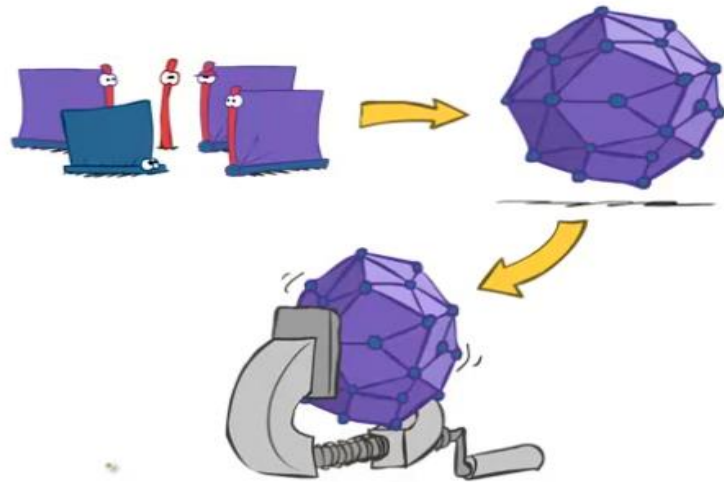
□ تعداد جملات جمع شونده = $2^{|H|}$

تعداد متغیرهای مفق (دورویی)

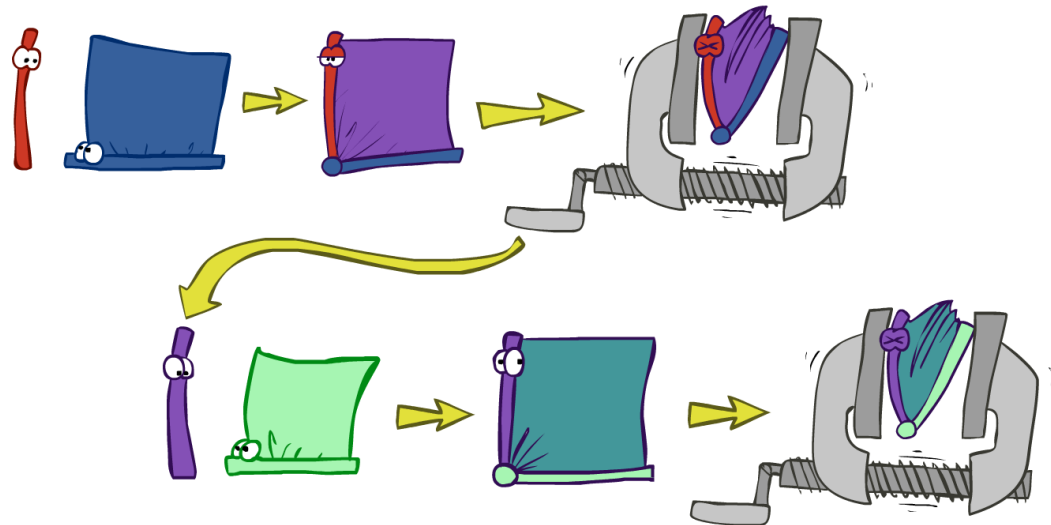
استنتاج به وسیله شمارش: تحلیل

۱۳

- چرا استنتاج به وسیله شمارش این قدر کند است؟
- به دلیل ایجاد کل توزیع توأم قبل از حذف متغیرهای مخفی.



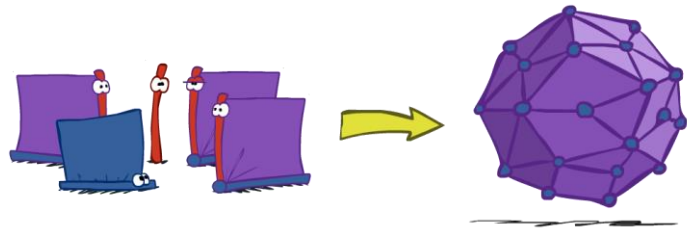
- یک ایده‌ی بهتر. الگوریتم حذف متغیر!
- انجام عمل الحاق و حاشیه‌گیری به صورت یک در میان.
- اگرچه زمان اجرا هنوز نمایی است، اما معمولاً در عمل بسیار سریع‌تر از استنتاج به وسیله شمارش است!



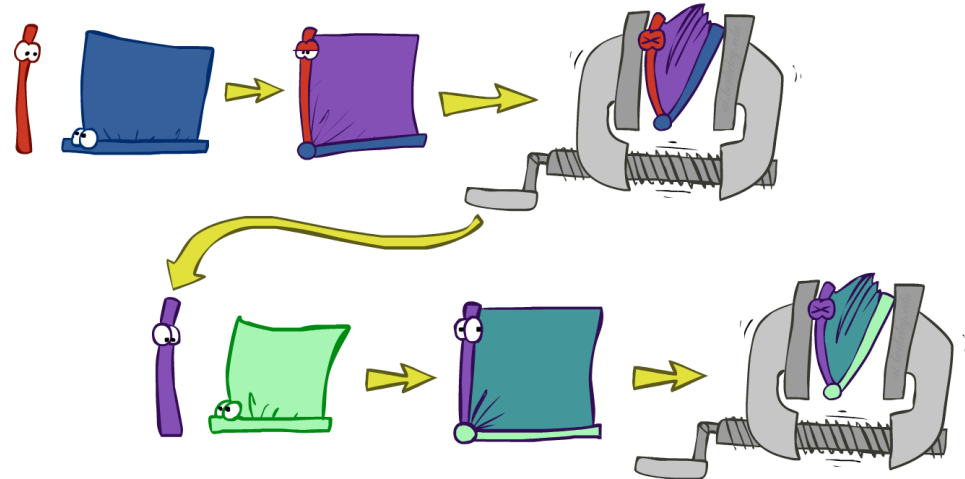
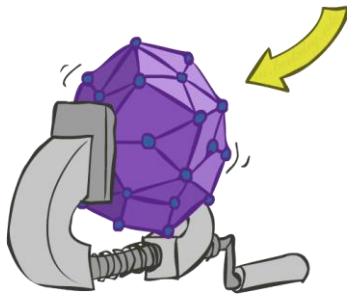
- ابتدا نیاز به تعریف یک نماد جدید داریم: فاکتورها!

استنتاج به وسیله شمارش: تحلیل

۱۴



- چرا استنتاج به وسیله شمارش این قدر کند است؟
- به دلیل ایجاد کل توزیع توأم قبل از حذف متغیرهای مخفی.



- یک ایده‌ی بهتر. الگوریتم حذف متغیر!
- انجام عمل الحاق و حاشیه‌گیری به صورت یک در میان.
- اگرچه زمان اجرا هنوز نمایی است، اما معمولاً در عمل بسیار سریع‌تر از استنتاج به وسیله شمارش است!

- ابتدا نیاز به تعریف یک نماد جدید داریم: فاکتورها!

فاکتورها (۱)

۱۵

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P(cold, W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

□ توزیع توأم: $P(X, Y)$

□ یک خانه به ازای هر مقدار ممکن برای x و y

□ یک ماتریس با ابعاد $|X| \times |Y|$

□ مجموع همه‌ی مقادیر برابر با یک

□ یک برش از توزیع توأم: $P(x, Y)$

□ مقدار x ثابت است.

□ یک خانه به ازای هر مقدار ممکن برای y

□ مجموع همه‌ی مقادیر برابر با $P(x)$

□ ابعاد جدول = تعداد حروف بزرگ!

فاکتورها (۲)

۱۶

$$P(W|cold)$$

T	W	P
cold	sun	0.4
cold	rain	0.6

$$P(W|T)$$

T	W	P
hot	sun	0.8
hot	rain	0.2
cold	sun	0.4
cold	rain	0.6

$$P(W|hot)$$

$$P(W|cold)$$

□ توزیع شرطی: $P(Y | x)$

□ مقدار x ثابت است.

□ یک خانه به ازای هر مقدار ممکن برای y

□ مجموع همه‌ی مقادیر برابر با یک

□ یک خانواده از توزیع‌های شرطی: $P(Y | X)$

□ چند توزیع شرطی

□ یک خانه به ازای هر مقدار ممکن برای x و y

□ مجموع همه‌ی مقادیر برابر با $|Y|$

فاکتورها (۳)

۱۷

$P(rain|T)$

T	W	P
hot	rain	0.2
cold	rain	0.6

$P(rain|hot)$

$P(rain|cold)$

□ یک خانواده مشخص: $P(y | X)$

□ مقدار y ثابت است.

□ یک خانه به ازای هر مقدار ممکن برای x

□ مجموع همه‌ی مقادیر برابر با ... [بی معنی]

□ خلاصه‌ی فاکتورها. $P(Y_1 \dots Y_N | X_1 \dots X_M)$

□ هر فاکتور یک آرایه‌ی چند بعدی است.

□ مقادیر فاکتور $P(y_1 \dots y_N | x_1 \dots x_M)$

□ با انتساب مقدار به هر متغیر، یک بعد از ابعاد جدول کاسته می‌شود.

مثال: دامنهی ترافیک

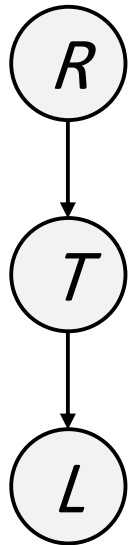
۱۸

متغیرها. □

□ R: هوا بارانی است.

□ T: ترافیک وجود دارد.

□ L: کلاس دیر شده است.



$P(R)$	
$+r$	0.1
$-r$	0.9

$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

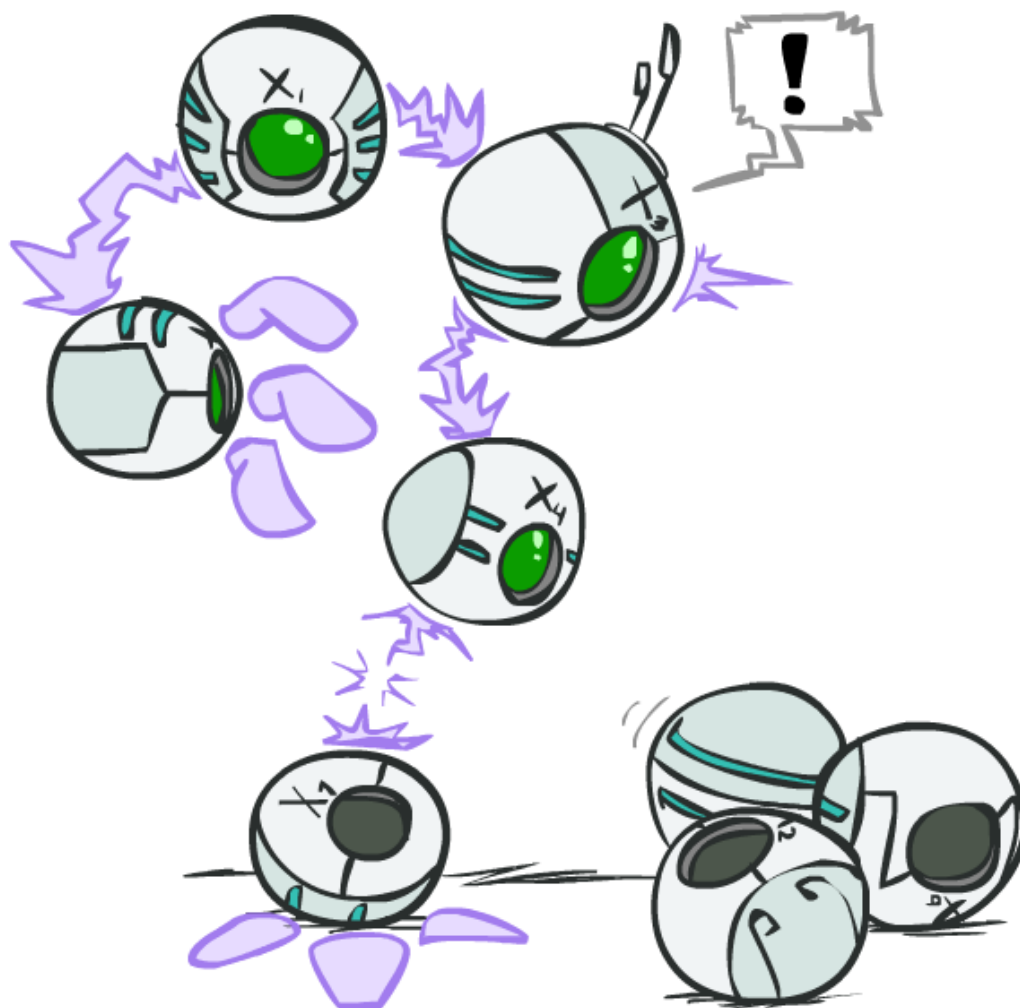
$$P(L) = ?$$

$$= \sum_{r,t} P(r, t, L)$$

$$= \sum_{r,t} P(r)P(t|r)P(L|t)$$

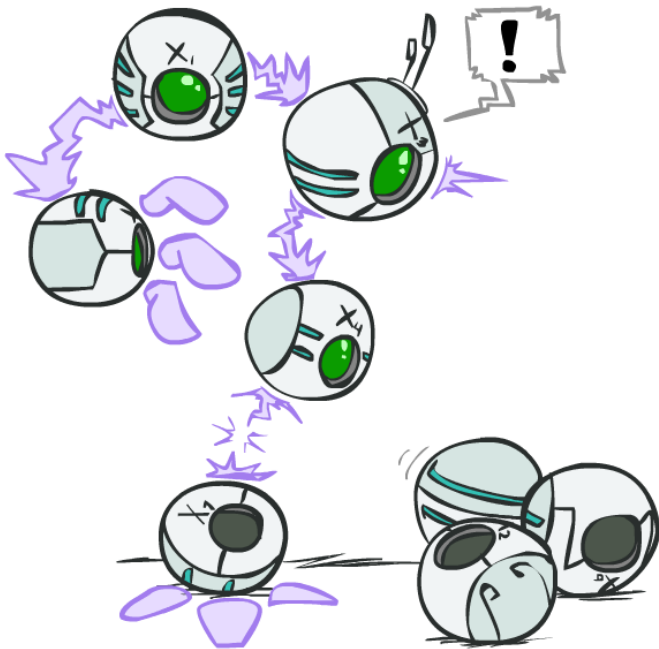
مذف متغير

١٩



استنتاج به وسیله‌ی شمارش: طرح کلی

۲۰



□ اشیای پایه‌ای. فاکتورها!

□ فاکتورهای اولیه. جداول توزیع شرطی محلی [یک جدول به ازای هر گره]

$P(R)$	
$+r$	0.1
$-r$	0.9

$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

□ مقادیر شناخته شده را انتخاب می‌کنیم.

□ مثلاً اگر بدانیم $L = +l$ ، در این صورت فاکتورهای اولیه برابرند با:

$P(R)$	
$+r$	0.1
$-r$	0.9

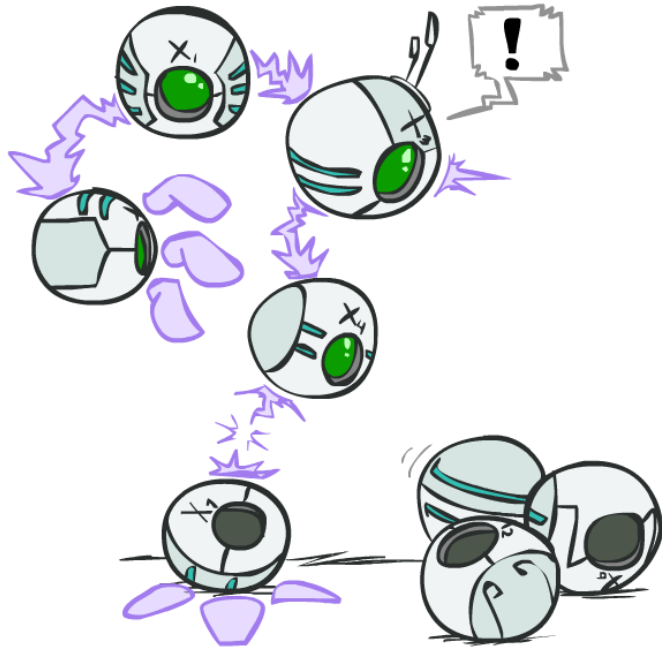
$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$-t$	$+l$	0.1

استنتاج به وسیله‌ی شمارش:
الحاق تمام فاکتورها و سپس حذف
متغیرهای مخفی!

مذف متغیر: طرح کلی

۲۱



□ اشیای پایه‌ای. فاکتورها!

□ فاکتورهای اولیه. جداول توزیع شرطی محلی [یک جدول به ازای هر گره]

$P(R)$	
$+r$	0.1
$-r$	0.9

$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

□ مقادیر شناخته شده را انتخاب می‌کنیم.

□ مثلاً اگر بدانیم $L = +l$ ، در این صورت فاکتورهای اولیه برابرند با:

$P(R)$	
$+r$	0.1
$-r$	0.9

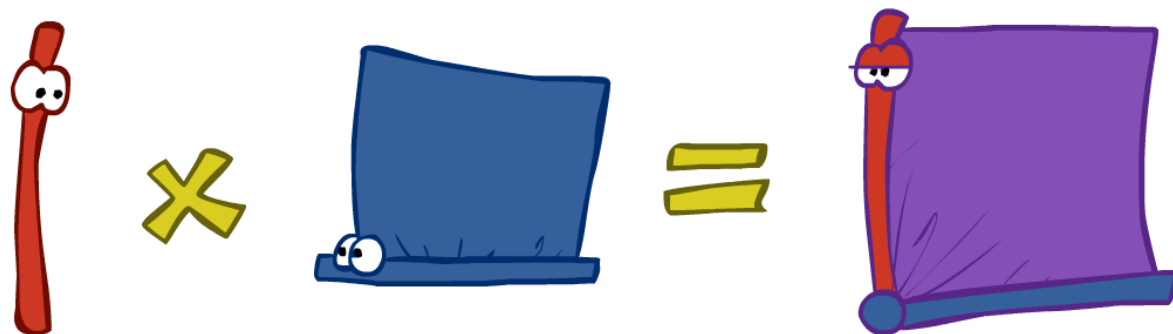
$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$-t$	$+l$	0.1

الگوریتم حذف متغیر:
الحاق فاکتورها و حذف متغیرها به
صورت یک در میان!

عمل اول: الحاق فاکتورها

۲۲

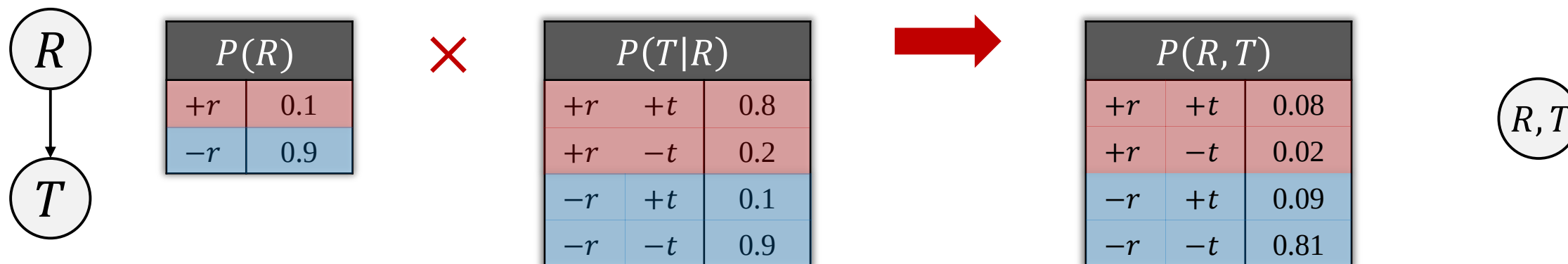


□ الحاق فاکتورها. [برای حذف متغیر Z]

□ انتخاب تمام فاکتورهای شامل متغیر Z .

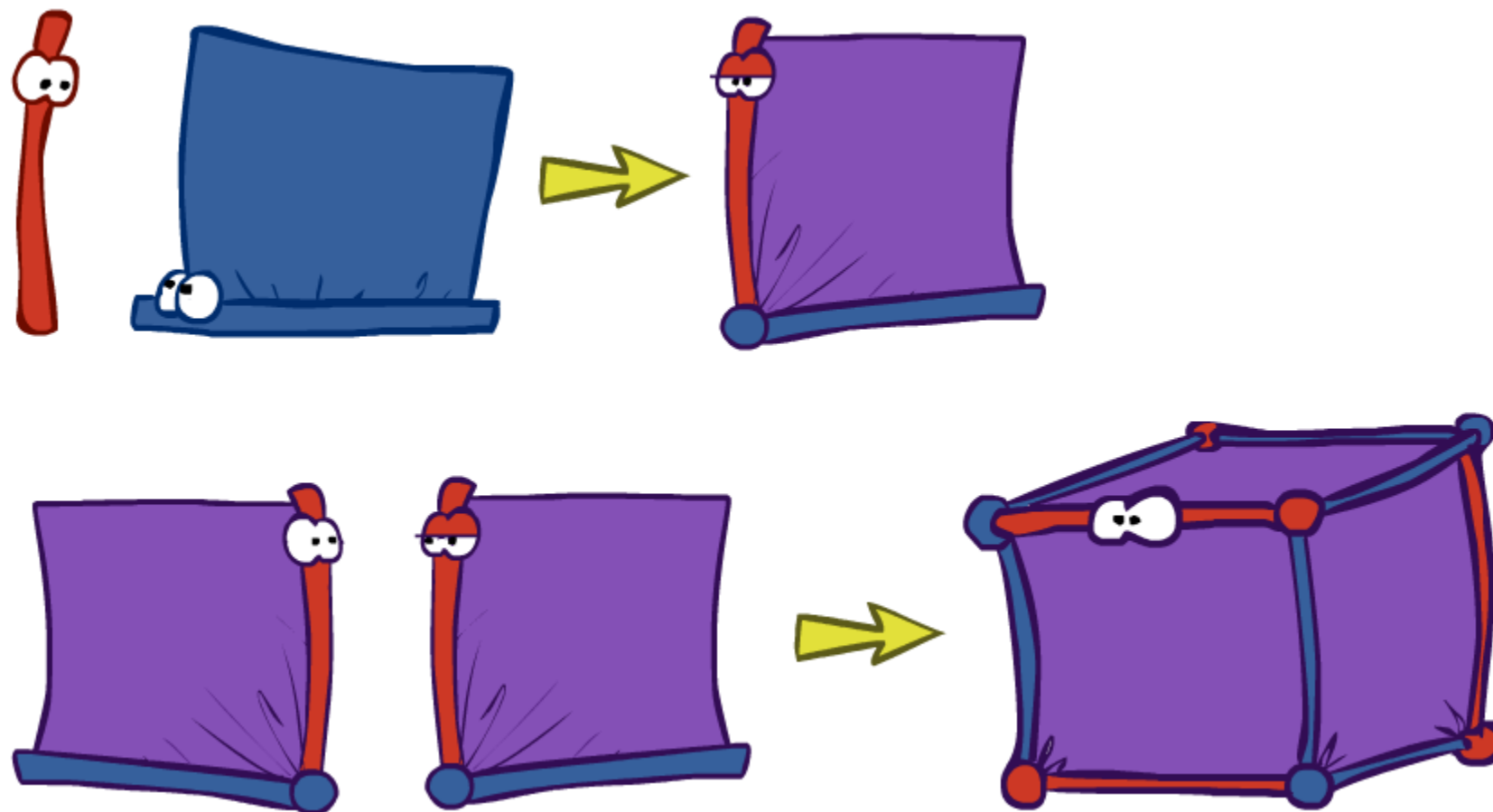
□ ایجاد یک فاکتور جدید بر روی اجتماع متغیرها.

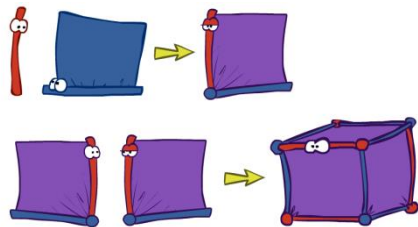
□ مثال. الحاق روی R



انجام چند الحاق

۲۳





مثال: انجام چند الحاق

۲۴

$P(R)$	
$+r$	0.1
$-r$	0.9

$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

Join R



R

T

L

$P(R,T)$		
$+r$	$+t$	0.08
$+r$	$-t$	0.02
$-r$	$+t$	0.09
$-r$	$-t$	0.81

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

Join T



R, T

L

$P(R,T,L)$			
$+r$	$+t$	$+l$	0.024
$+r$	$+t$	$-l$	0.056
$+r$	$-t$	$+l$	0.002
$+r$	$-t$	$-l$	0.018
$-r$	$+t$	$+l$	0.027
$-r$	$+t$	$-l$	0.063
$-r$	$-t$	$+l$	0.081
$-r$	$-t$	$-l$	0.729

R, T, L

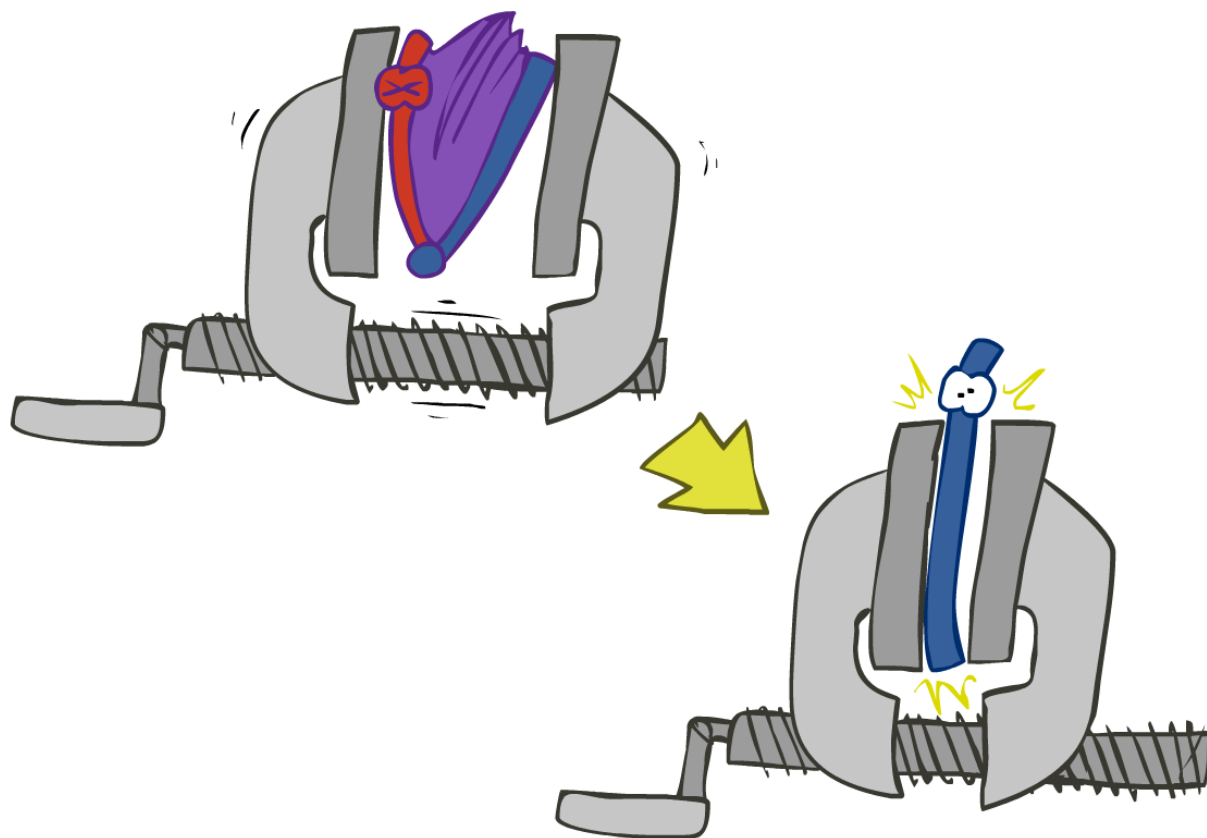
عمل دوم: حذف با ماشیه‌گیری

۲۵

□ حاشیه‌گیری.

□ انتخاب یک فاکتور و جمع زدن بر روی یک متغیر.

□ کوچک‌تر کردن اندازه‌ی یک فاکتور. [کاهش یک بعد]



□ مثال.

$P(R, T)$		
$+r$	$+t$	0.08
$+r$	$-t$	0.02
$-r$	$+t$	0.09
$-r$	$-t$	0.81

Sum R
→

$P(T)$	
$+t$	0.17
$-t$	0.83

(R, T, L)

$P(R, T, L)$			
$+r$	$+t$	$+l$	0.024
$+r$	$+t$	$-l$	0.056
$+r$	$-t$	$+l$	0.002
$+r$	$-t$	$-l$	0.018
$-r$	$+t$	$+l$	0.027
$-r$	$+t$	$-l$	0.063
$-r$	$-t$	$+l$	0.081
$-r$	$-t$	$-l$	0.729

Sum R

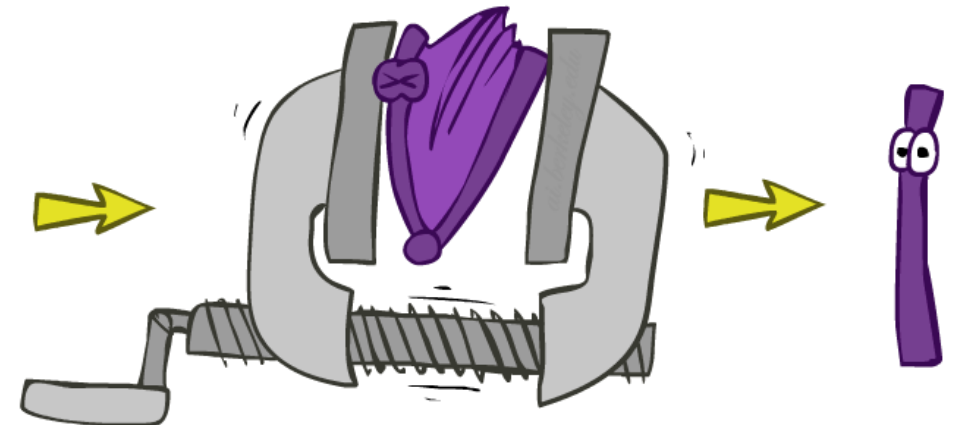
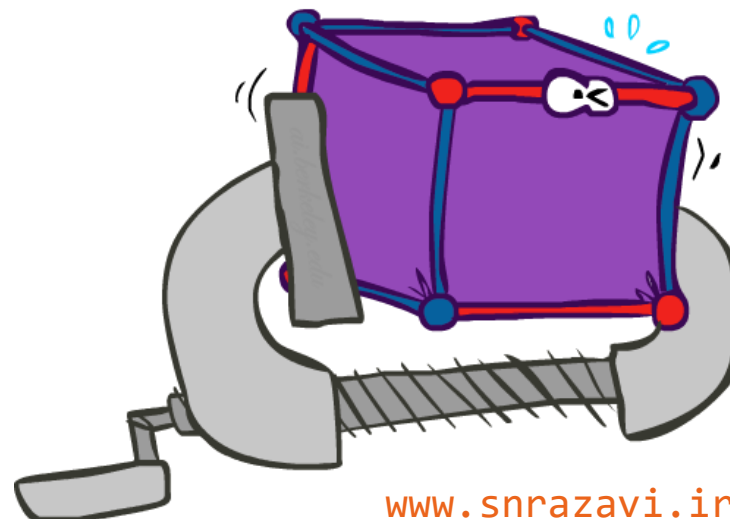
(T, L)

$P(T, L)$		
$+t$	$+l$	0.051
$+t$	$-l$	0.119
$-t$	$+l$	0.083
$-t$	$-l$	0.747

Sum T

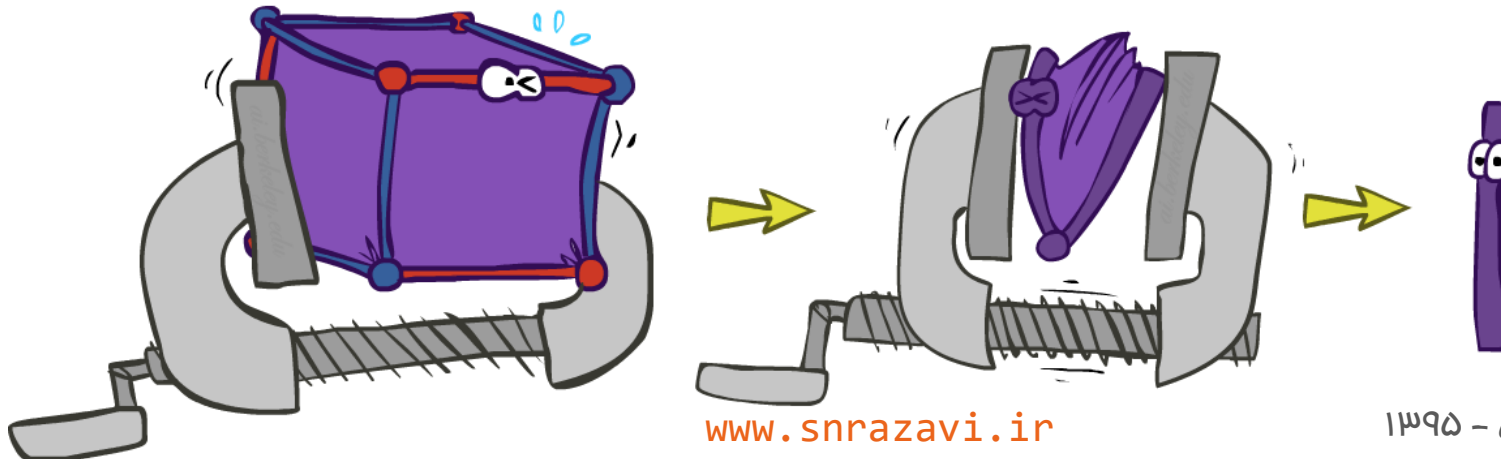
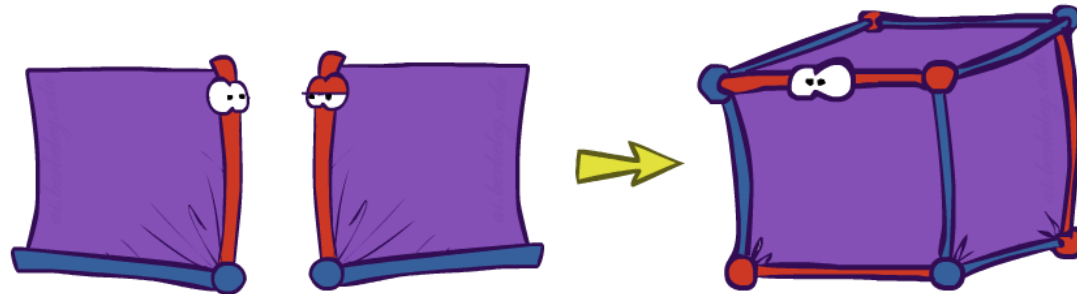
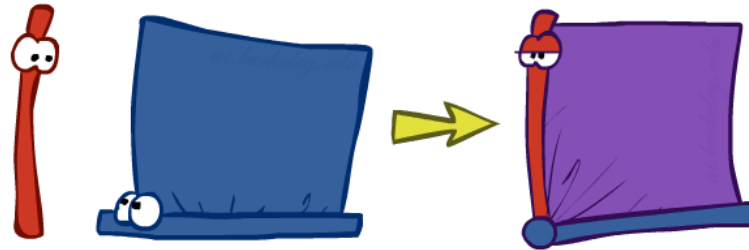
(L)

$P(L)$	
$+l$	0.134
$-l$	0.866



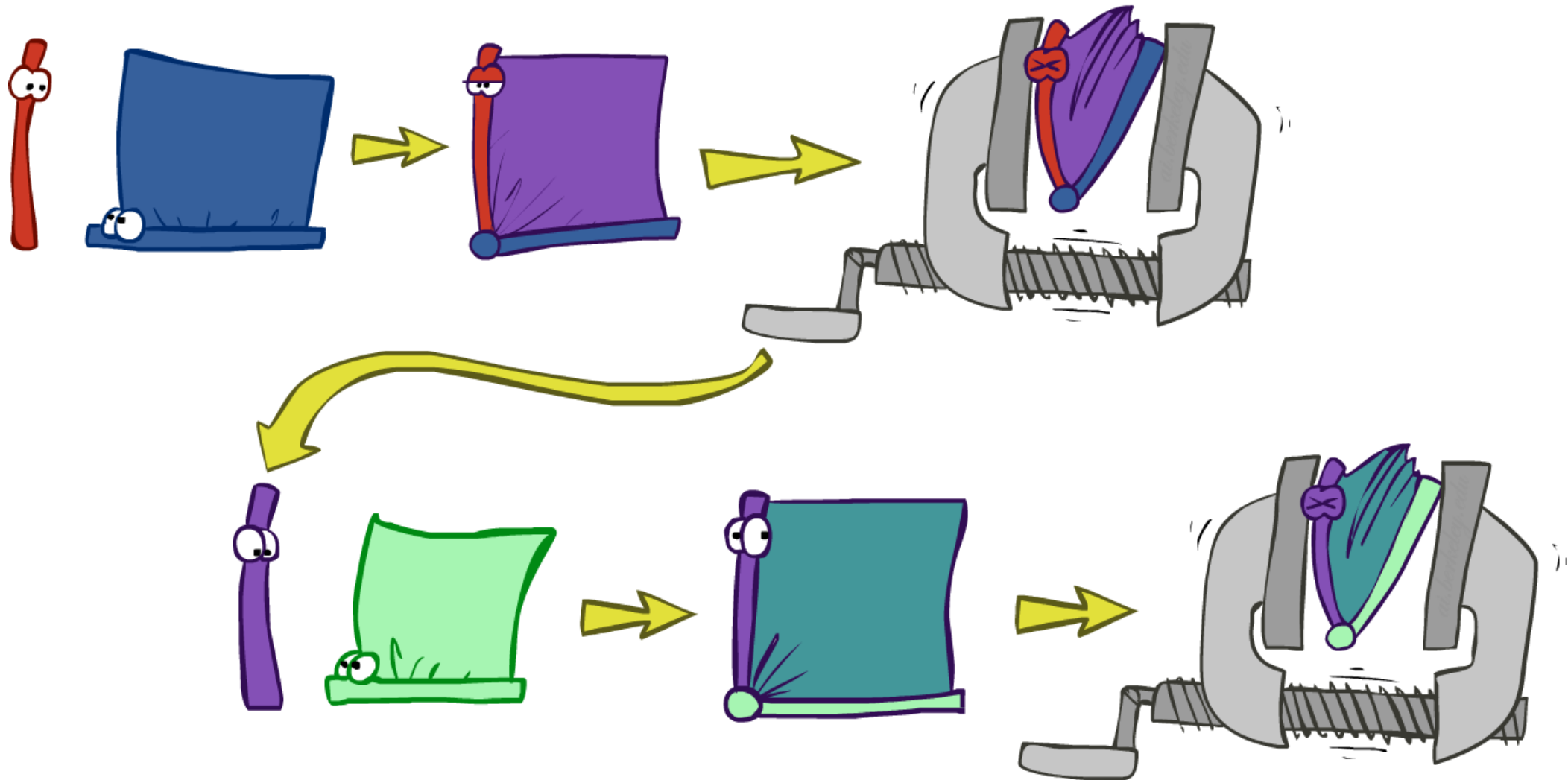
استنتاج به وسیلهی شمارش (= چند الحاق و چند حذف)

۲۷



استنتاج به روش حذف متغیر (= ماشینگیری زود هنگام)

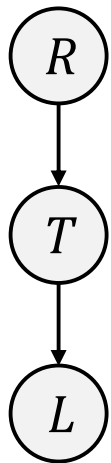
۲۸



مثال: دامنه‌ی ترافیک

۲۹

$$P(L) = ?$$



□ استنتاج به روش حذف متغیر.

$$= \sum_t P(L|t) \underbrace{\sum_r \underbrace{P(r)P(t|r)}_{\text{الاق روی } R}}_{\text{مذف } R} \underbrace{}_{\text{الاق روی } T} \underbrace{}_{\text{مذف } T}$$

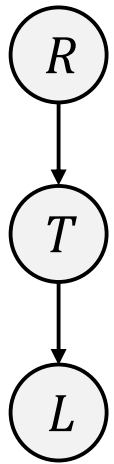
□ استنتاج به وسیله‌ی شمارش.

$$= \sum_t \sum_r \underbrace{P(L|t) \underbrace{P(r)P(t|r)}_{\text{الاق روی } R}}_{\text{الاق روی } T} \underbrace{}_{\text{مذف } R} \underbrace{}_{\text{مذف } T}$$

مثال: دامنه‌ی ترافیک

۳۰

$P(L) = ?$



□ استنتاج به روش حذف متغیر.

$$\begin{aligned} &= \sum_t P(L|t) \sum_r P(r)P(t|r) && \text{الاق روی } R \\ &= \sum_t P(L|t) \sum_r P(r, t) && \text{حذف } R \\ &= \sum_t P(L|t)P(t) && \text{الاق روی } T \\ &= \sum_t P(t, L) && \text{حذف } T \\ &= P(L) \end{aligned}$$

□ استنتاج به وسیله‌ی شمارش.

$$\begin{aligned} &= \sum_t \sum_r P(L|t)P(r)P(t|r) && \text{الاق روی } R \\ &= \sum_t \sum_r P(L|t)P(r, t) && \text{الاق روی } T \\ &= \sum_t \sum_r P(r, t, L) && \text{حذف } R \\ &= \sum_t P(t, L) && \text{حذف } T \\ &= P(L) \end{aligned}$$

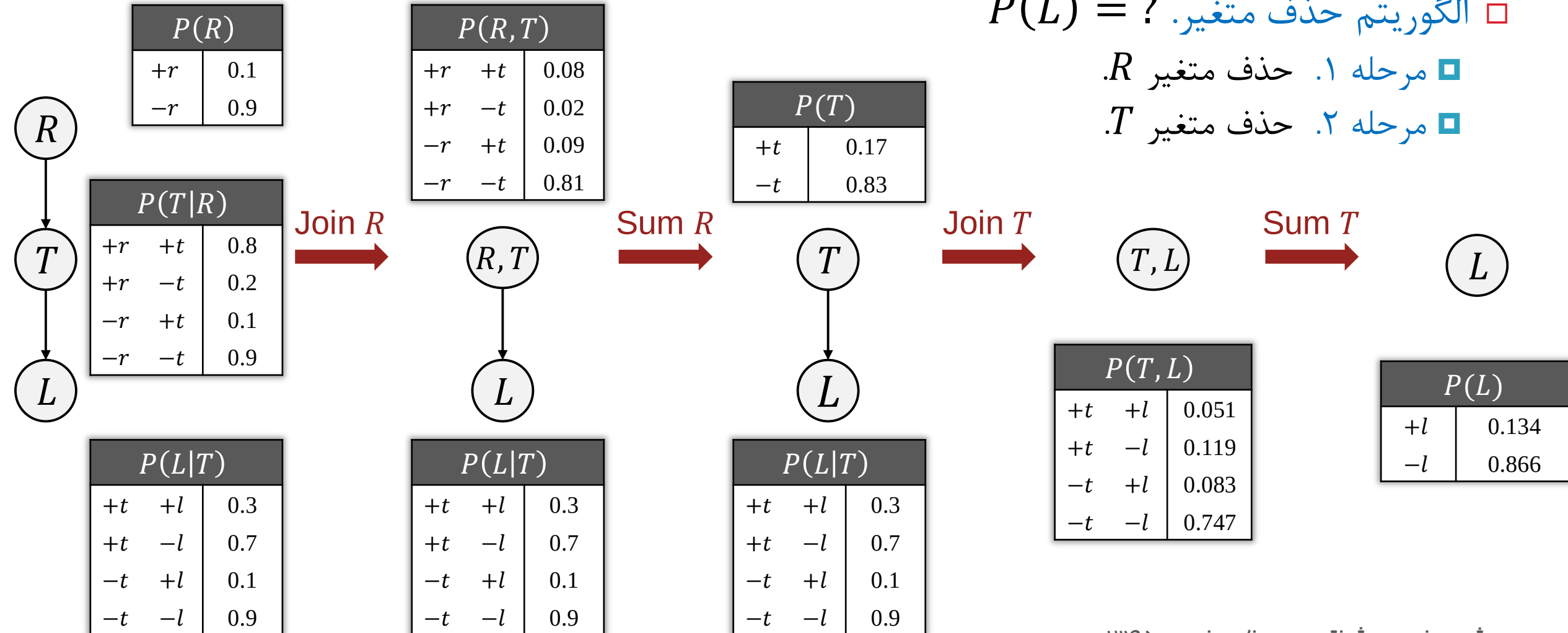
مثال: حذف متغیر

۳۱

□ الگوریتم حذف متغیر. $P(L) = ?$

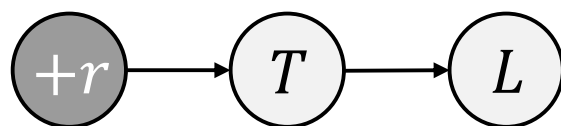
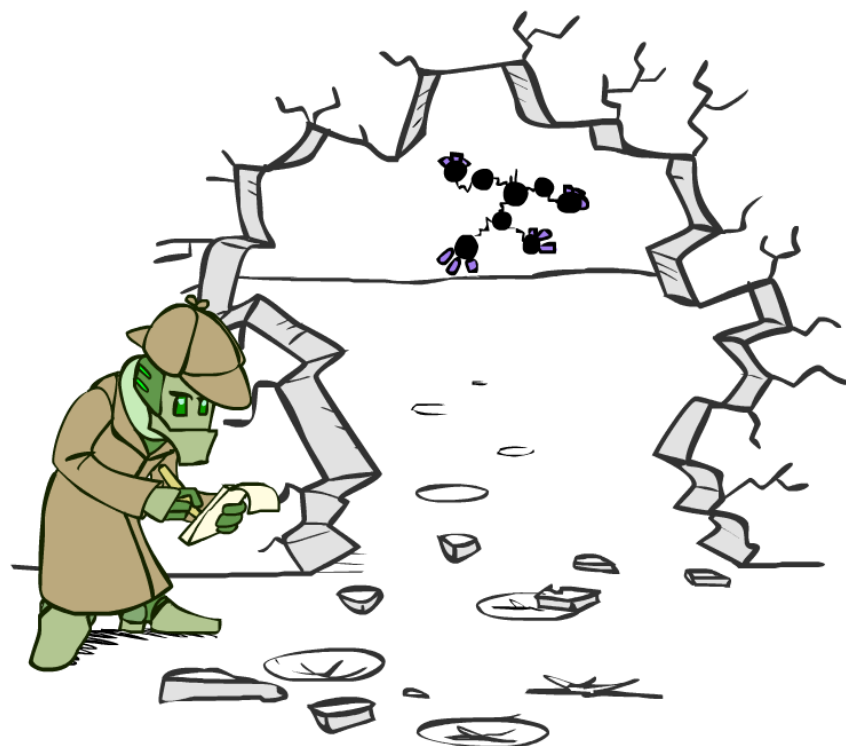
□ مرحله ۱. حذف متغیر R .

□ مرحله ۲. حذف متغیر T .



استفاده از شواهد

۳۲



مزف متغیرهای مففی با استفاده از الگوریتم مزف متغیر

□ شواهد. $R = +r$

□ با فاکتورهایی که آن شواهد را انتخاب کرده‌اند، شروع می‌کنیم.

$P(R)$	
$+r$	0.1
$-r$	0.9

$P(T R)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2
$-r$	$+t$	0.1
$-r$	$-t$	0.9

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

□ فاکتورهای اولیه برای محاسبه $P(L|+r)$

$P(+r)$	
$+r$	0.1

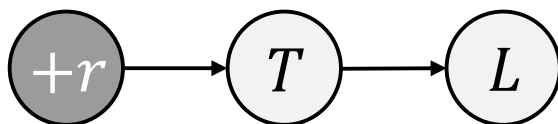
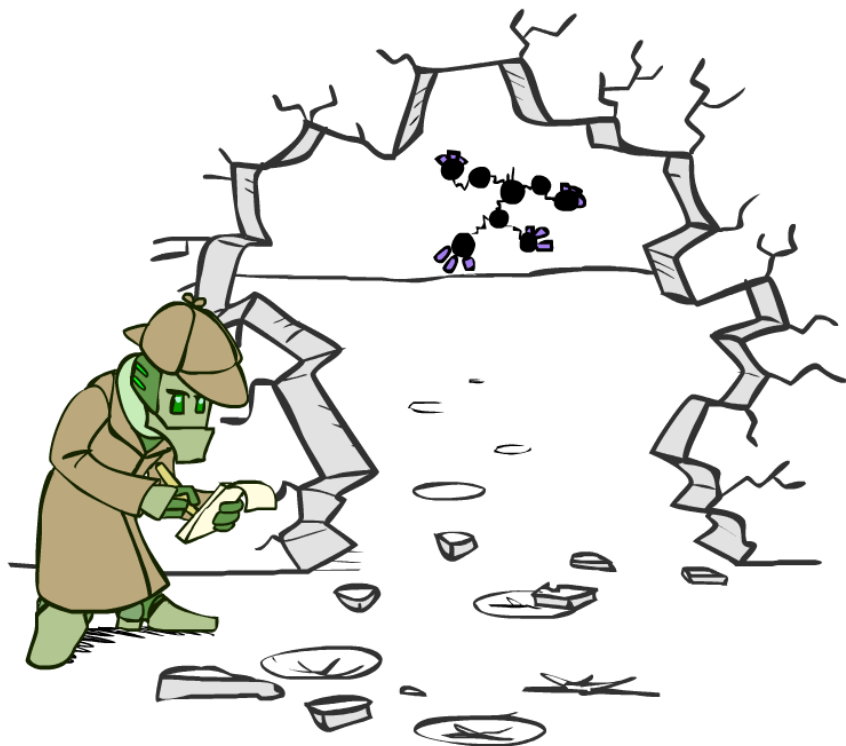
$P(T +r)$		
$+r$	$+t$	0.8
$+r$	$-t$	0.2

$P(L T)$		
$+t$	$+l$	0.3
$+t$	$-l$	0.7
$-t$	$+l$	0.1
$-t$	$-l$	0.9

استفاده از شواهد

۳۳

□ نتیجه. یک برش از متغیرهای شواهد و پرس و جو.
□ مثلاً برای $P(L|+r)$ نتیجه به صورت زیر خواهد بود.



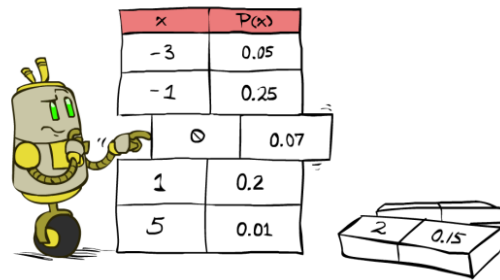
$P(+r, L)$		
$+r$	$+t$	0.026
$+r$	$-t$	0.074



$P(L +r)$		
$+r$	$+t$	0.26
$+r$	$-t$	0.74

الگوریتم حذف متغیر

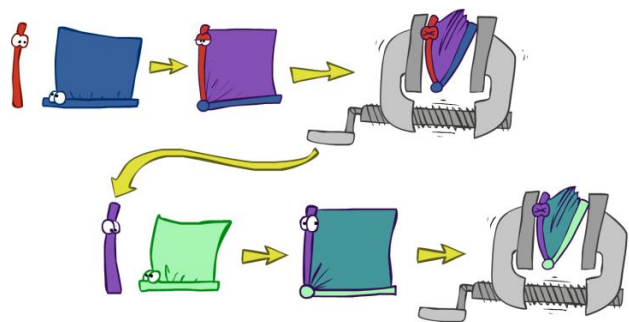
۳۴



□ پرس و جو: $P(Q|E_1 = e_1, \dots, E_k = e_k)$

□ با فاکتورهای اولیه شروع کن.

□ جداول شرطی محلی [نمونه‌دهی شده به وسیله‌ی شواهد]



□ تا وقتی متغیر مخفی وجود دارد:

□ یک متغیر مخفی مانند H را انتخاب کن،

□ همه‌ی فاکتورهای شامل H را الحاق کن،

□ متغیر H را حذف کن. [با جمع زدن روی H]

$$\text{Variable} \times \text{Factor} = \text{New Factor} \times \frac{1}{Z}$$

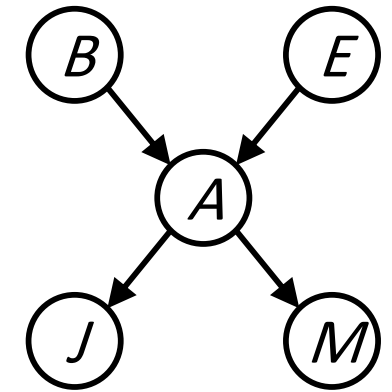
□ تمام فاکتورهای باقیمانده را الحاق و سپس نرمال‌سازی کن.

مثال: حذف متغیر

۳۵

$$P(B|j, m) \propto P(B, j, m)$$

$P(B)$	$P(E)$	$P(A B, E)$	$P(j A)$	$P(m A)$
--------	--------	-------------	----------	----------

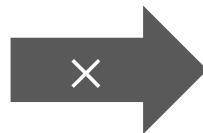


حذف A □

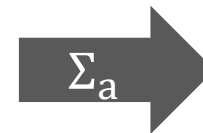
$$P(A|B, E)$$

$$P(j|A)$$

$$P(m|A)$$



$$P(j, m, A|B, E)$$



$$P(j, m|B, E)$$

$P(B)$	$P(E)$	$P(j, m B, E)$
--------	--------	----------------

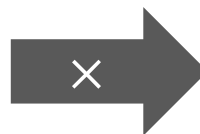
مثال: حذف متغیر

۳۶

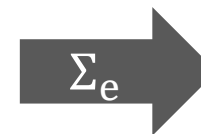
$$P(B|j, m) \propto P(B, j, m)$$

$P(B)$	$P(E)$	$P(j, m B, E)$
--------	--------	----------------

$$P(E)$$

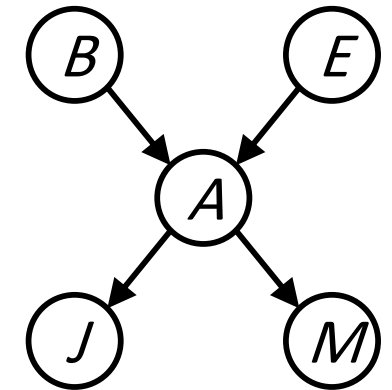


$$P(j, m, E|B)$$



$$P(j, m|B)$$

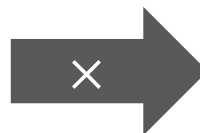
$$P(j, m|B, E)$$



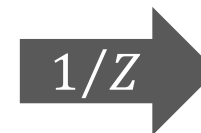
حذف E □

$P(B)$	$P(j, m B)$
--------	-------------

$$P(B)$$



$$P(j, m, B)$$



$$P(B|j, m)$$

$$P(j, m|B)$$

همان مثال به صورت ریاضی

۳۷

$$P(B|j, m) \propto P(B, j, m)$$

$P(B)$

$P(E)$

$P(A|B, E)$

$P(j|A)$

$P(m|A)$

$$P(B|j, m) \propto P(B, j, m)$$

$$= \sum_e \sum_a P(B, j, m, e, a)$$

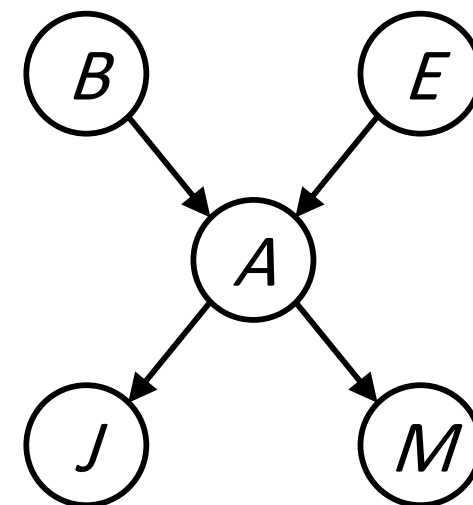
$$= \sum_e \sum_a P(B)P(e)P(a|B, e)P(j|a)P(m|a)$$

$$= \sum_e P(B)P(e) \sum_a P(a|B, e)P(j|a)P(m|a)$$

$$= \sum_e P(B)P(e)\phi_1(B, e, j, m)$$

$$= P(B) \sum_e P(e)\phi_1(B, e, j, m)$$

$$= P(B)\phi_2(B, j, m)$$



الگوریتم حذف متغیر: یک مثال دیگر

۳۸

$$P(X_3 | Y_1 = y_1, Y_2 = y_2, Y_3 = y_3)$$

$$P(Z)P(X_1|Z)P(X_2|Z)P(X_3|Z)P(y_1|X_1)P(y_2|X_2)P(y_3|X_3)$$

$$\phi_1(Z, y_1) = \sum_{x_1} P(X_1|Z)P(y_1|X_1)$$

حذف متغیر X_1

$$P(Z)\phi_1(Z, y_1)P(X_2|Z)P(X_3|Z)P(y_2|X_2)P(y_3|X_3)$$

$$\phi_2(Z, y_2) = \sum_{x_2} P(X_2|Z)P(y_2|X_2)$$

حذف متغیر X_2

$$P(Z)\phi_1(Z, y_1)\phi_2(Z, y_2)P(X_3|Z)P(y_3|X_3)$$

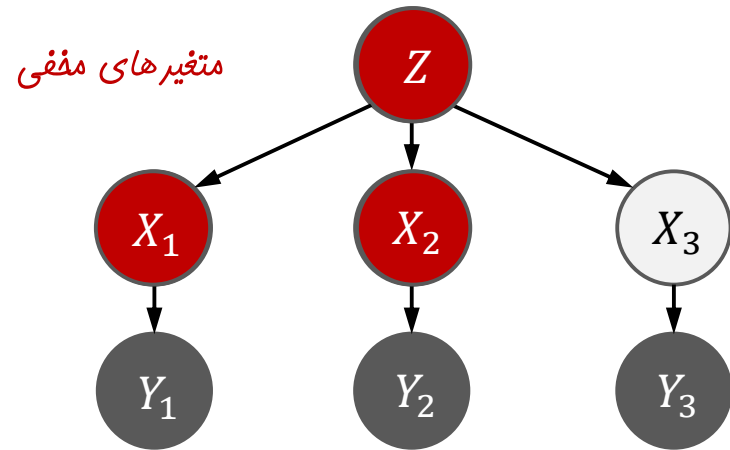
$$\phi_3(y_1, y_2, X_3) = \sum_Z P(Z)\phi_1(Z, y_1)\phi_2(Z, y_2)P(X_3|Z)$$

حذف متغیر Z

$$\phi_3(y_1, y_2, X_3)P(y_3|X_3)$$

$$\phi_4(y_1, y_2, y_3, X_3) = \phi_3(y_1, y_2, X_3)P(y_3|X_3) \xrightarrow{1/Z} P(X_3 | y_1, y_2, y_3)$$

نرمال سازی



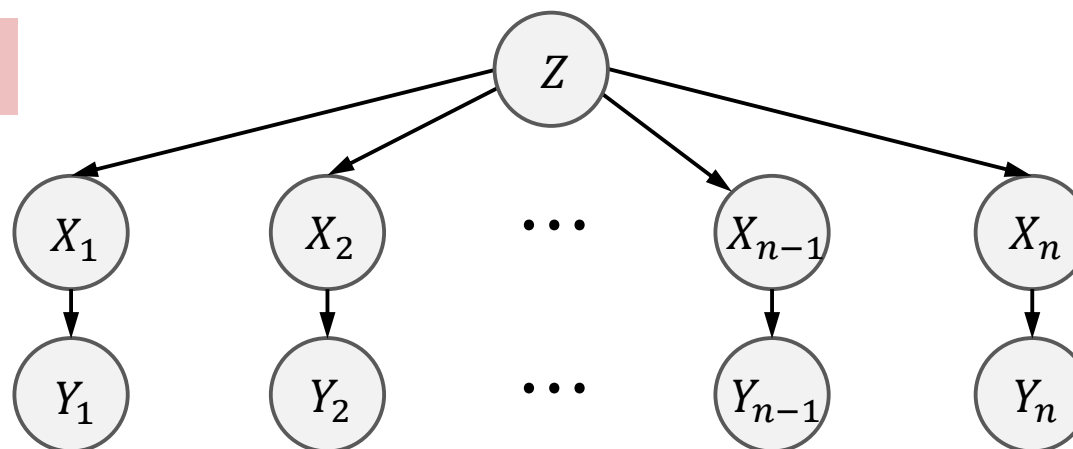
پیچیدگی زمانی الگوریتم حذف متغیر به اندازه‌ی فاکتورهای تولید شده در طول اجرای الگوریتم بستگی دارد. اندازه‌ی یک فاکتور برابر است با تعداد مقادیر موجود در جدول مربوط به آن فاکتور. در این مثال (با فرض متغیرهای دودویی) اندازه‌ی فاکتورهای تولید شده برابر با ۲ است، زیرا هر یک از این فاکتورها تنها شامل یک متغیر است (به ترتیب Z ، X_3 و Z).

ترتیب حذف متغیرها

۳۹

□ **ترتیب حذف متغیرها.** فرض کنید برای پرس و جوی زیر الگوریتم حذف متغیر را یک بار به ترتیب $X_1, \dots, X_{n-1}, Z$ و یک بار دیگر به ترتیب $Z, X_1, \dots, X_{n-1}$ انجام دهیم. در این صورت اندازه‌ی بزرگ‌ترین فاکتور تولید شده در طول اجرای الگوریتم برای هر یک از این دو ترتیب کدام است؟

$$P(X_n | y_1, y_2, \dots, y_n)$$

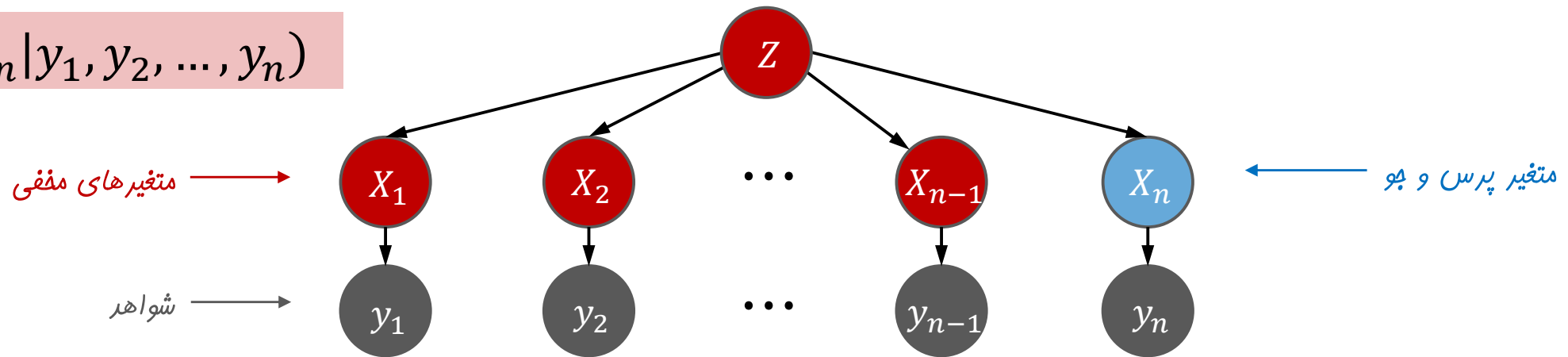


ترتیب حذف متغیرها

۴۰

□ **ترتیب حذف متغیرها.** فرض کنید برای پرس و جوی زیر الگوریتم حذف متغیر را یک بار به ترتیب $X_1, \dots, X_{n-1}, Z$ و یک بار دیگر به ترتیب $Z, X_1, \dots, X_{n-1}$ انجام دهیم. در این صورت اندازه‌ی بزرگ‌ترین فاکتور تولید شده در طول اجرای الگوریتم برای هر یک از این دو ترتیب کدام است؟

$$P(X_n | y_1, y_2, \dots, y_n)$$



2^2

□ **پاسخ.** فاکتورهای تولید شده اگر متغیرهای مخفی به ترتیب $X_1, \dots, X_{n-1}, Z$ حذف شوند.

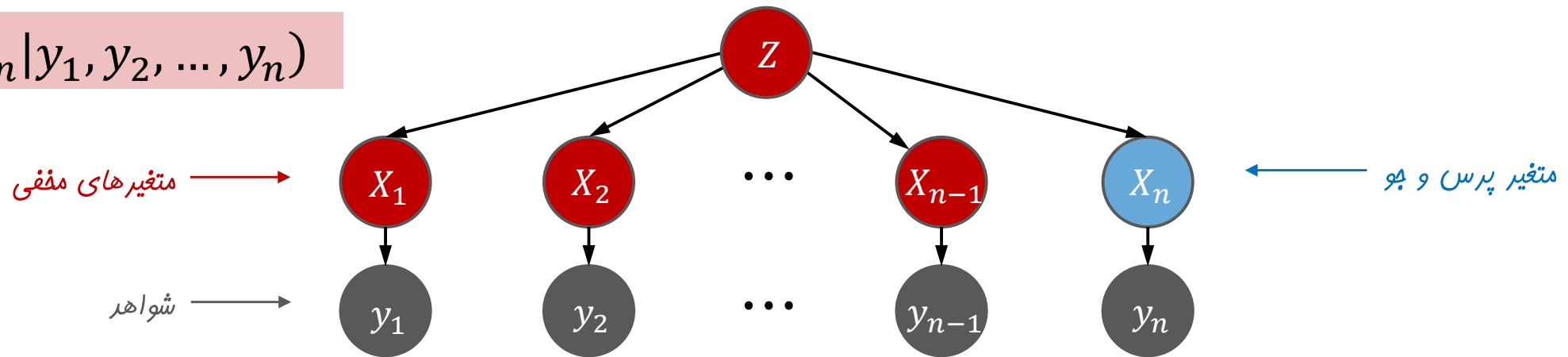
$$\phi_1(Z, y_1) = \sum_{x_1} \underbrace{P(X_1|Z)P(y_1|X_1)}_{\phi'_1(Z, X_1, y_1)} \quad \phi_2(Z, y_2) \quad \dots \quad \phi_{n-1}(Z, y_{n-1}) \quad \phi_n(y_1, y_2, \dots, y_n, X_n)$$

ترتیب حذف متغیرها

۴۱

□ **ترتیب حذف متغیرها.** فرض کنید برای پرس و جوی زیر الگوریتم حذف متغیر را یک بار به ترتیب $X_1, \dots, X_{n-1}, Z$ و یک بار دیگر به ترتیب $Z, X_1, \dots, X_{n-1}$ انجام دهیم. در این صورت اندازه‌ی بزرگ‌ترین فاکتور تولید شده در طول اجرای الگوریتم برای هر یک از این دو ترتیب کدام است؟

$$P(X_n | y_1, y_2, \dots, y_n)$$



2^{n+1}

□ **پاسخ.** فاکتورهای تولید شده اگر متغیرهای مخفی به ترتیب $Z, X_1, \dots, X_{n-1}$ حذف شوند.

$$\phi_1(X_1, X_2, \dots, X_n) = \sum_z P(Z)P(X_1|Z)P(X_2|Z) \cdots P(X_n|Z)$$

حذف متغیر Z

$$\phi_2(y_1, X_2, \dots, X_n) = \sum_{x_1} \phi_1(X_1, X_2, \dots, X_n)P(y_1|X_1)$$

حذف متغیر X_1

الگوریتم حذف متغیر: تحلیل

۴۲

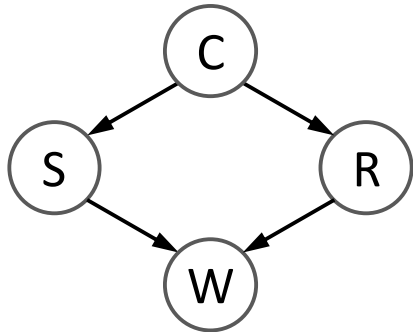
- پیچیدگی زمان و حافظه.
- به اندازه‌ی بزرگ‌ترین فاکتور تولید شده در طول اجرای الگوریتم بستگی دارد.
- ترتیب حذف متغیرها.
- می‌تواند به شدت بر روی اندازه‌ی بزرگ‌ترین فاکتور تولید شده تأثیر بگذارد.
- آیا همواره ترتیبی وجود دارد که باعث تولید فاکتورهایی با اندازه‌ی کوچک شود؟ **خیر!**
- یافتن بهترین ترتیب برای حذف متغیرها خود یک **مسئله‌ی ان‌پی-سخت** است.
- تحلیل بدترین حالت. [بر حسب تعداد متغیرهای مخفی]
- نمایی!

چند درخت‌ها

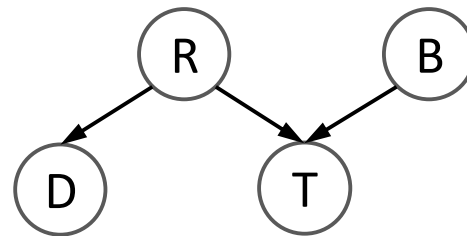
۴۳

□ چند درخت. یک گراف جهت‌دار بدون دور. [بدون در نظر گرفتن جهت یالها]

□ همانند درخت، با این تفاوت که هر گره می‌تواند بیش از یک والد داشته باشد!



این گراف بدون دور در نظر گرفتن جهت یالها دارای دور است.

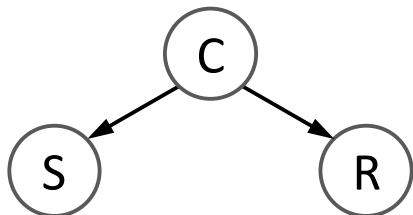


یک مثال از یک چند درخت.
گره T دارای دو گره والد است

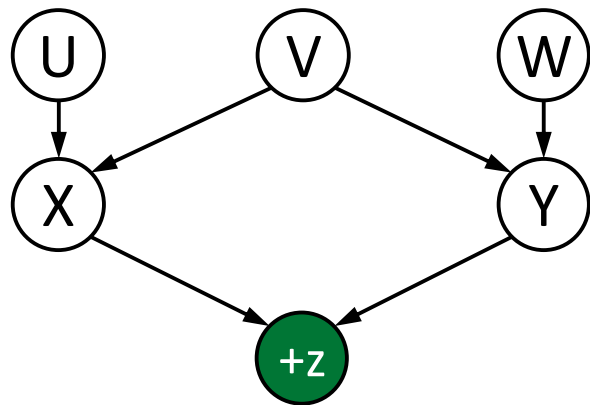
□ توجه. برای چند درخت‌ها همواره یک ترتیب کارا به منظور حذف متغیرهای مخفی وجود دارد.

□ مجموعه برشی برای استنتاج در شبکه‌های بیز.

□ یک زیرمجموعه از متغیرها که پس از حذف آنها تنها یک «چند درخت» باقی می‌ماند.



یک چند درخت



□ هدف. محاسبه $P(U|+z)$