

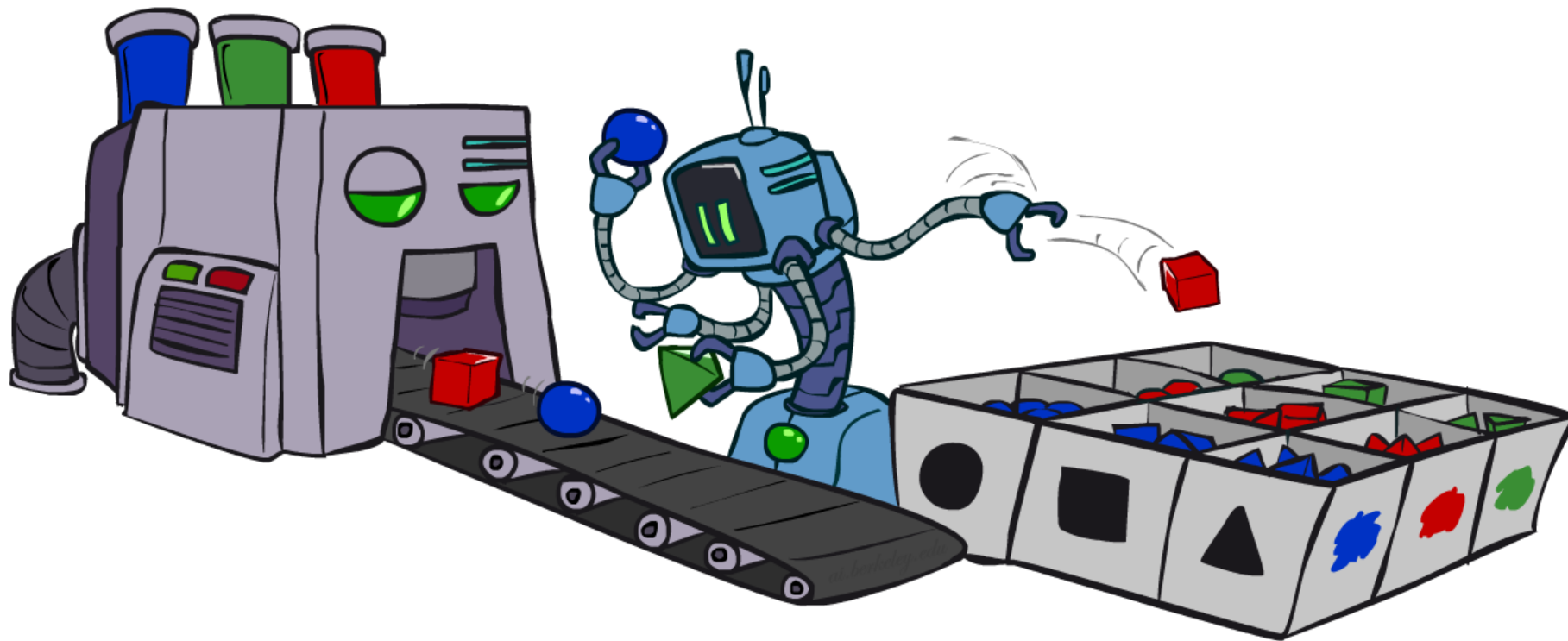
شبکه‌های بیز: نمونه‌برداری

سید ناصر رضوی www.snrazavi.ir

۱۳۹۸

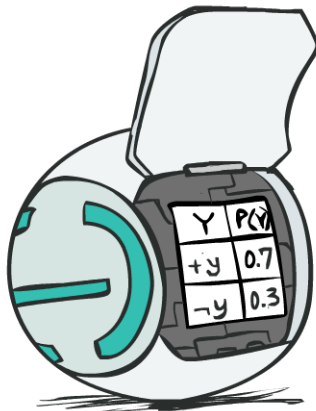
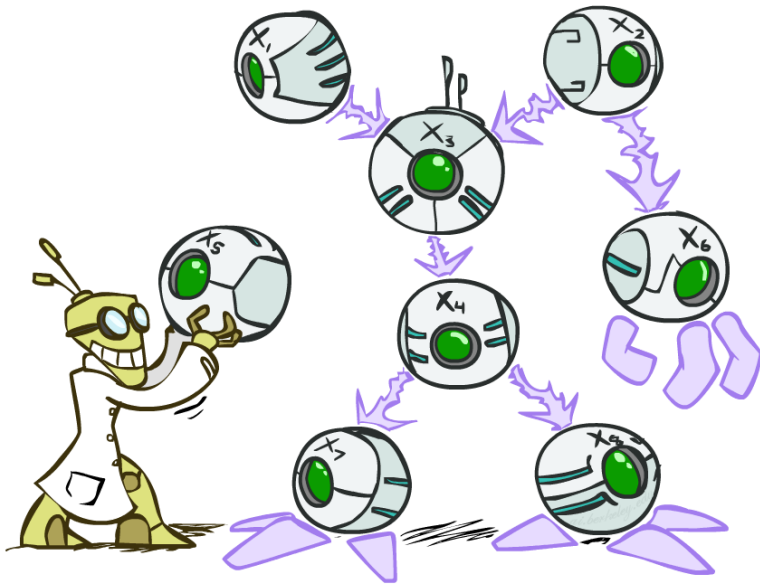
نمونه برداری

۲



شبکه‌های بیز: بازنمایی (یادآوری)

۳



□ شبکه بیز. یک گراف جهت‌دار و بدون دور.

□ یک رأس به ازای هر متغیر تصادفی

□ یک جدول احتمال شرطی به ازای هر رأس

$$P(X_i | \text{parents}(X_i))$$

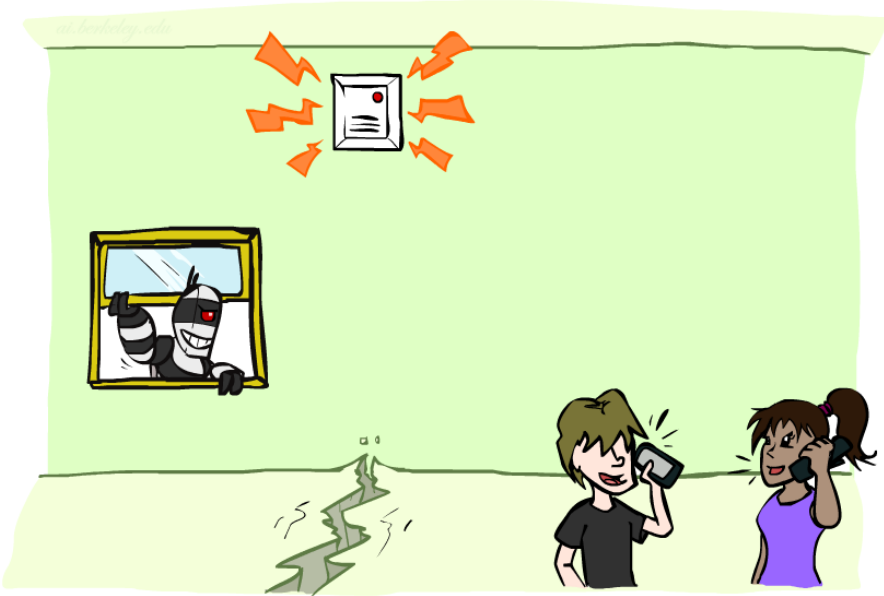
□ هر شبکه بیز به طور ضمنی بیانگر یک توزیع توأم است.

□ به صورت حاصل ضرب توزیع‌های شرطی محلی

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

مثال: زنگ هشدار (یادآوری)

۴

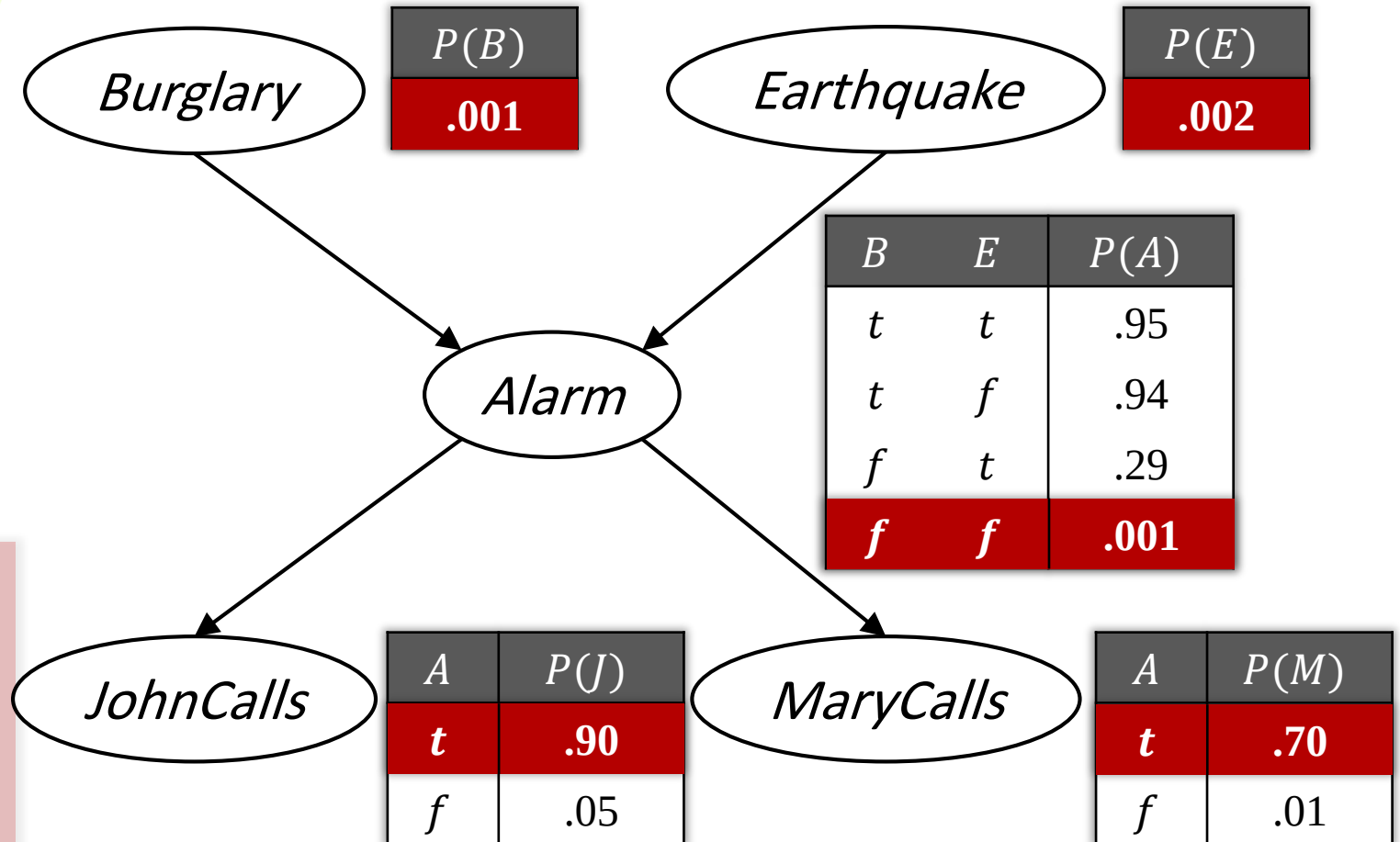


$$P(j, m, a, \neg b, \neg e)$$

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

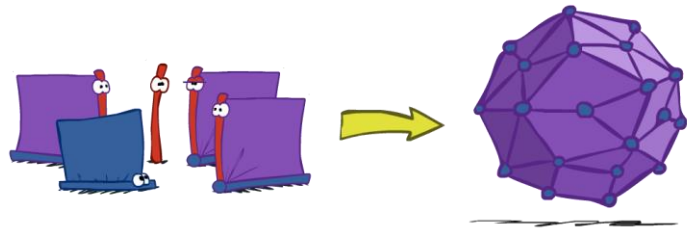
$$= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.988$$

$$= 0.000628$$

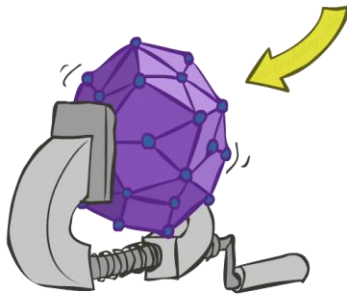


استنتاج به وسیله شمارش: تحلیل

۵



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

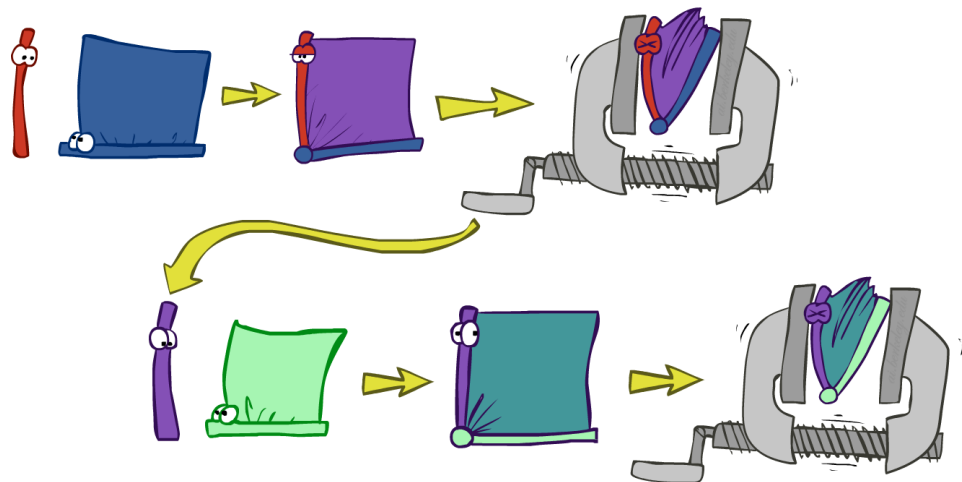


- یک ایده بهتر. الگوریتم حذف متغیر!

□ 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!**

الگوریتم حذف متغیر (یادآوری)

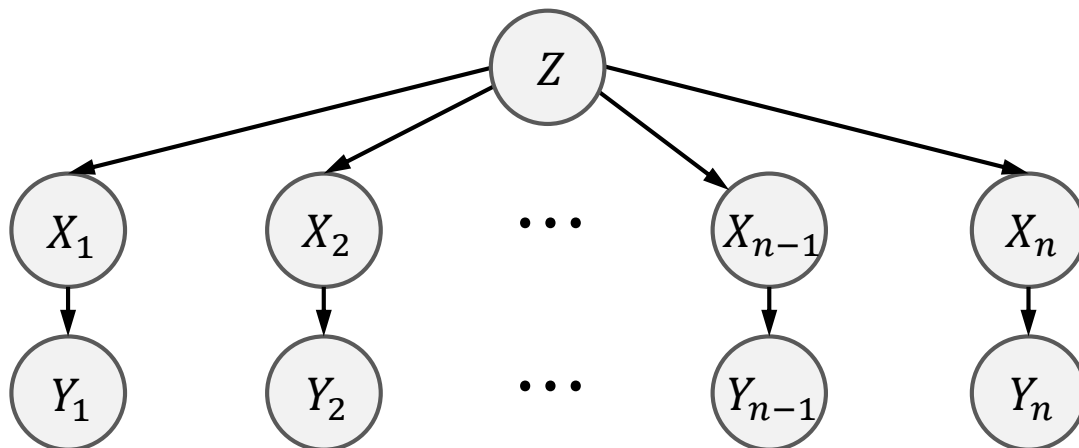
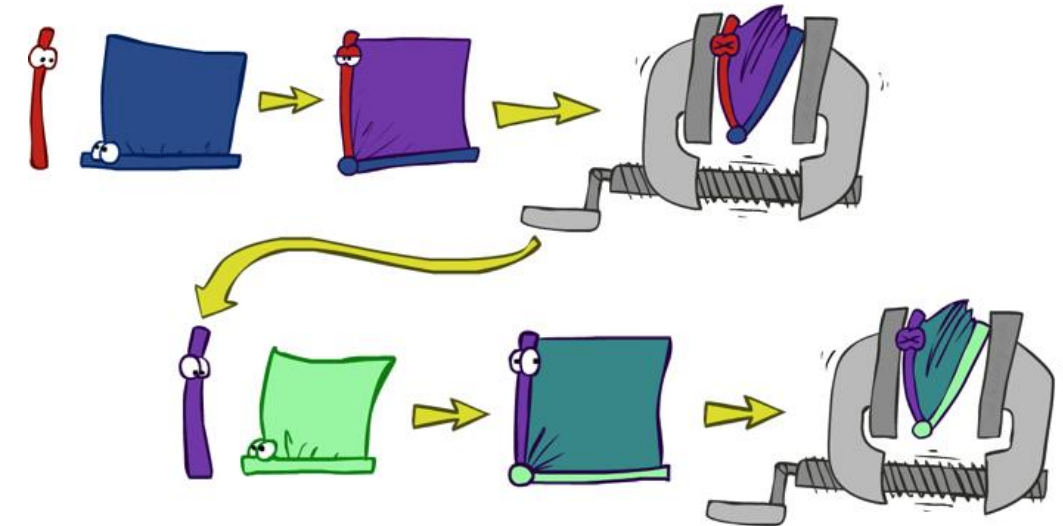
۶

□ ایده. انجام عملیات الحاق (ضرب) و حاشیه‌گیری (جمع) به صورت یک در میان.

□ اندازه فاکتور. تعداد مقادیر محاسبه شده برای یک فاکتور بر روی k متغیر با اندازه دامنه d برابر است با d^k .

□ ترتیب حذف متغیرهای مخفی. ترتیب حذف متغیرها می‌تواند به شدت بر روی اندازه فاکتورهای تولید شده تأثیر بگذارد.

□ زمان اجرا در بدترین حالت. نمایی!



چند درخت‌ها (یادآوری)

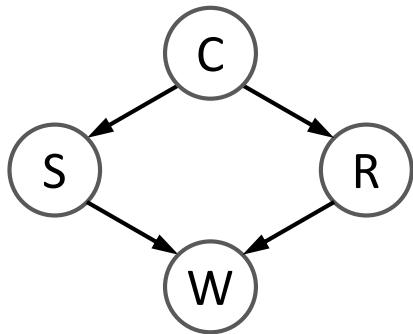
□ چند درخت. یک گراف جهت‌دار بدون دور. [بدون در نظر گرفتن جهت یالها]

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

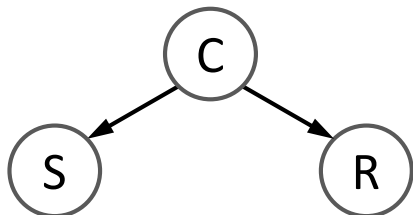
□ توجه. برای چنددرخت‌ها همواره یک ترتیب کارا به منظور حذف متغیرهای مخفی وجود دارد.

□ مجموعه برشی برای استنتاج در شبکه‌های بیز.

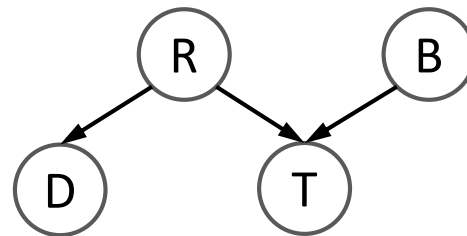
□ یک زیرمجموعه از متغیرها که پس از حذف آنها تنها یک «چند درخت» باقی می‌ماند.



این گراف بدون دور در نظر گرفته می‌شود
چون یالها دارای دور است.



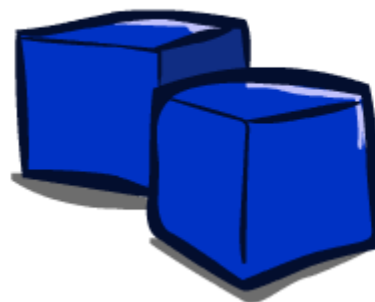
یک چند درخت



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

استنتاج تقریبی: نمونه‌برداری

۸



نمونه برداری

□ نمونه برداری شباهت زیادی با شبیه سازی دارد. □ چرا نمونه برداری؟

□ ایده اصلی.

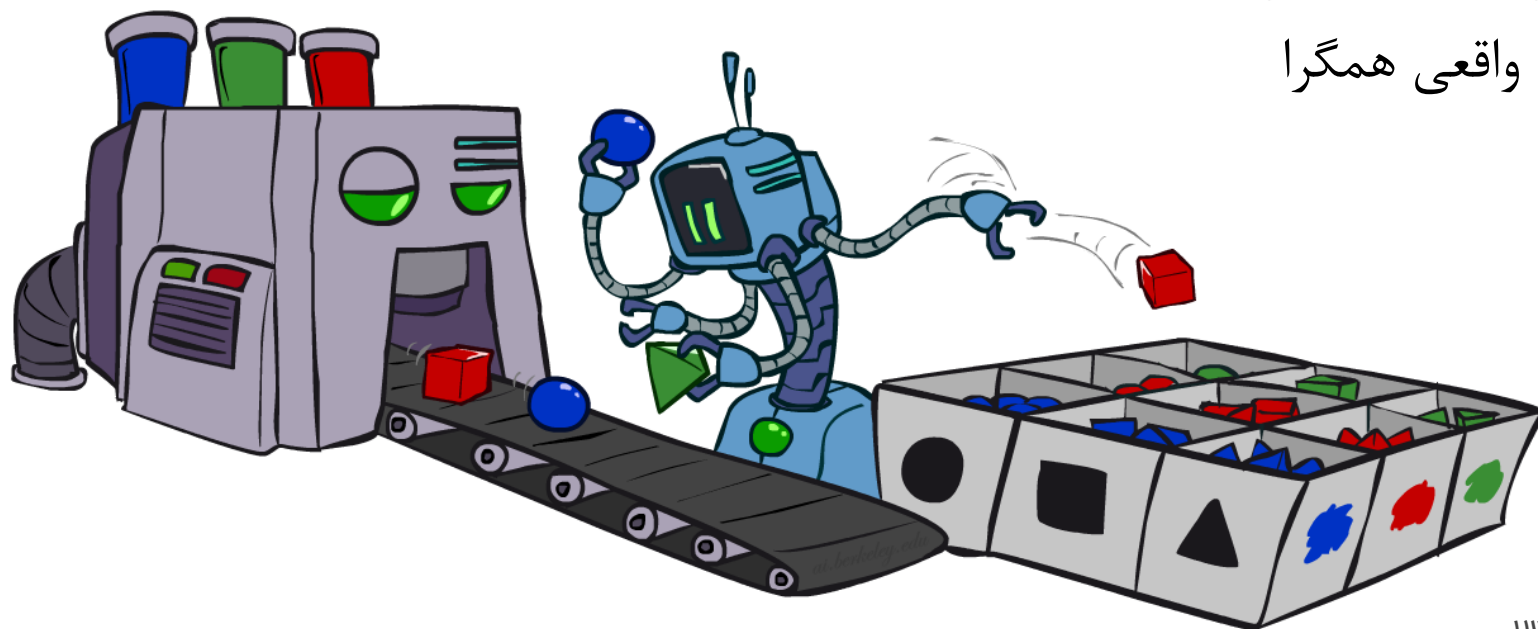
□ یادگیری: نمونه برداری از یک توزیع ناشناخته.

□ استنتاج: نمونه برداری نسبت به محاسبه پاسخ دقیق، سریع تر است.

□ از توزیع S به تعداد N بار نمونه برداری کن.

□ از روی نمونه ها، احتمال پسین را تقریب بزن.

□ نشان بده این مقدار به احتمال واقعی همگرا می گردد. [سازگاری]



نمونه برداری

۱۰

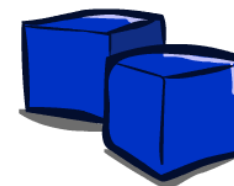
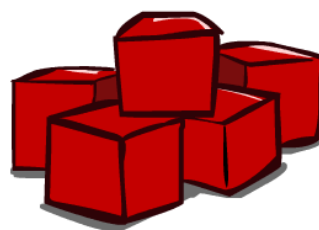
□ نمونه برداری از یک توزیع داده شده.

<i>Color</i>	<i>P(Color)</i>
red	0.6
green	0.1
blue	0.3

□ گام ۱: برداشتن نمونه u از یک توزیع یکنواخت بر روی بازه $[0,1)$.

■ مثلاً با استفاده از تابع `random()` در جاوا یا پایتون.

□ گام ۲: نمونه برداری از توزیع داده شده با استفاده از نمونه u .



$$0.0 \leq u < 0.6 \rightarrow C = red$$

$$0.6 \leq u < 0.7 \rightarrow C = green$$

$$0.7 \leq u < 1.0 \rightarrow C = blue$$

$$\hat{P}(r) = \frac{5}{8},$$

$$\hat{P}(b) = \frac{2}{8},$$

$$\hat{P}(g) = \frac{1}{8}$$

نمونه برداری

۱۱



ویلارد گیبس
(۱۸۳۹ - ۱۹۰۳)

□ نمونه برداری در شبکه های بیز.

□ نمونه برداری پیشین

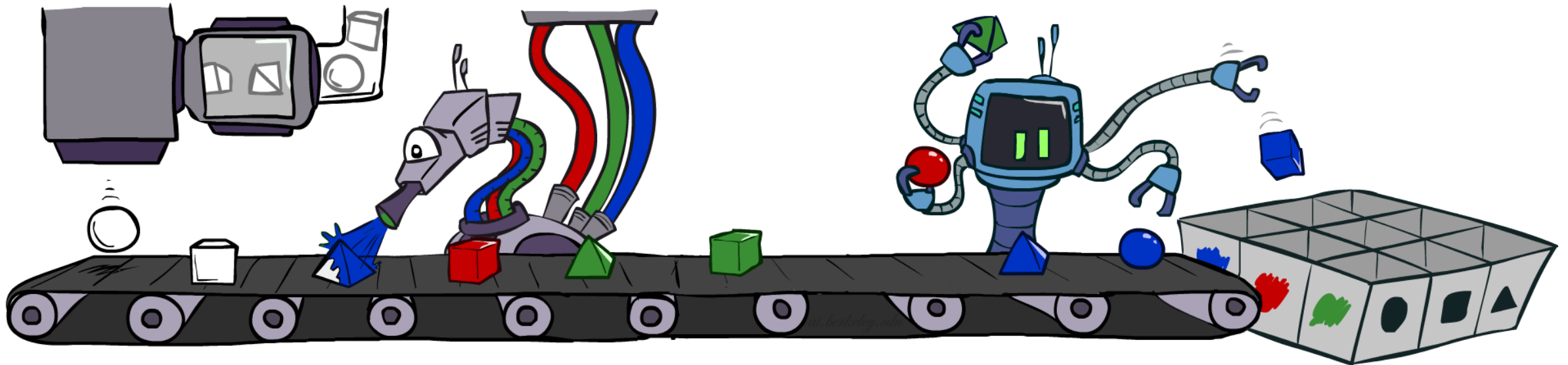
□ نمونه برداری با امکان رد کردن

□ وزن دهی نمونه ها

□ نمونه برداری «گیبس»

نمونه‌برداری پیشین

۱۲



نمونه برداری پیشین

۱۳

□ مرتب سازی توپولوژیکی رئوس.

$$C \rightarrow S \rightarrow R \rightarrow W$$

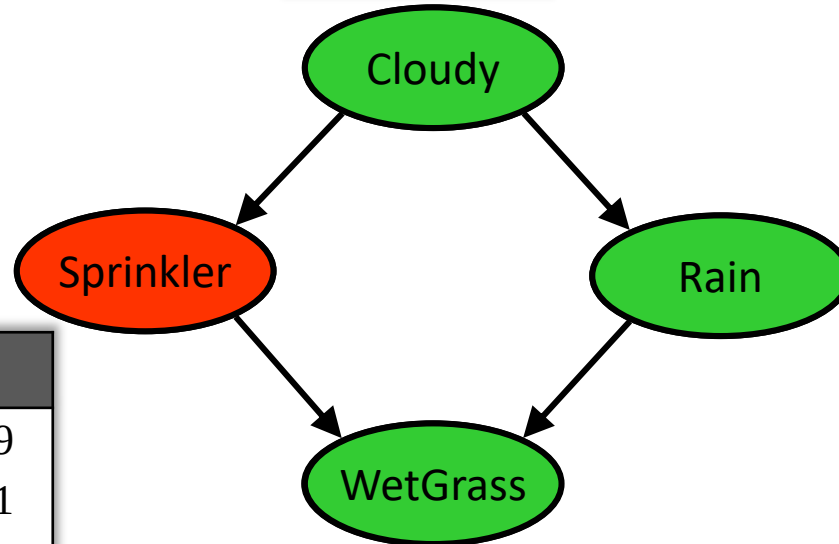
□ نمونه ها.

$+c \quad -s \quad +r \quad +w$

$P(S C)$		
$+c$	$+s$	0.1
$+c$	$-s$	0.9
$-c$	$+s$	0.5
$-c$	$-s$	0.5

$P(W S, R)$			
$+s$	$+r$	$+w$	0.99
$+s$	$+r$	$-w$	0.01
$+s$	$-r$	$+w$	0.90
$+s$	$-r$	$-w$	0.10
$-s$	$+r$	$+w$	0.90
$-s$	$+r$	$-w$	0.10
$-s$	$-r$	$+w$	0.01
$-s$	$-r$	$-w$	0.99

C	$P(C)$
$+c$	0.5
$-c$	0.5



$P(R C)$		
$+c$	$+r$	0.8
$+c$	$-r$	0.2
$-c$	$+r$	0.2
$-c$	$-r$	0.8

نمونه برداری پیشین

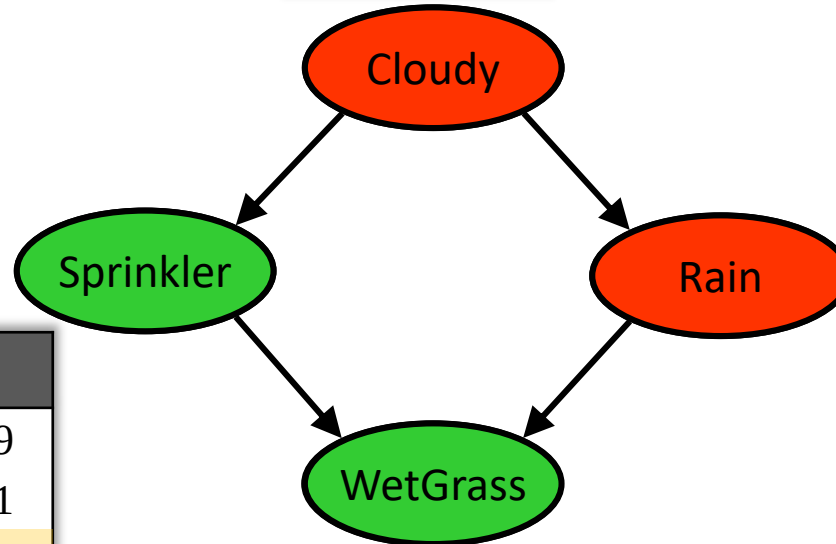
۱۴

□ مرتب سازی توپولوژیکی رئوس.

$$C \rightarrow S \rightarrow R \rightarrow W$$

$P(S C)$		
$+c$	$+s$	0.1
$+c$	$-s$	0.9
$-c$	$+s$	0.5
$-c$	$-s$	0.5

C	$P(C)$
$+c$	0.5
$-c$	0.5



$P(R C)$		
$+c$	$+r$	0.8
$+c$	$-r$	0.2
$-c$	$+r$	0.2
$-c$	$-r$	0.8

□ نمونه ها.

$+c \quad -s \quad +r \quad +w$
 $-c \quad +s \quad -r \quad +w$

$P(W S, R)$			
$+s$	$+r$	$+w$	0.99
$+s$	$+r$	$-w$	0.01
$+s$	$-r$	$+w$	0.90
$+s$	$-r$	$-w$	0.10
$-s$	$+r$	$+w$	0.90
$-s$	$+r$	$-w$	0.10
$-s$	$-r$	$+w$	0.01
$-s$	$-r$	$-w$	0.99

□ احتمال تولید نمونه $\langle -c, +s, -r, +w \rangle$.

$$P(-c)P(+s|-c)P(-r|-c)P(+w|+s, -r) = P(-c, +s, -r, +w)$$

نمونه‌برداری پیشین

۱۵

□ نمونه‌برداری پیشین. هر نمونه با یک احتمال تولید می‌شود:

$$S_{PS}(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{parents}(X_i)) = P(x_1, x_2, \dots, x_n)$$

فرمول بازسازی شبکه بیز

$$N_{PS}(x_1, x_2, \dots, x_n)$$

□ تعداد نمونه‌های یک رویداد:

$$\begin{aligned} \lim_{N \rightarrow \infty} \hat{P}(x_1, x_2, \dots, x_n) &= \lim_{N \rightarrow \infty} N_{PS}(x_1, x_2, \dots, x_n) / N \\ &= S_{PS}(x_1, x_2, \dots, x_n) \\ &= P(x_1, x_2, \dots, x_n) \end{aligned}$$

□ بنابراین:

□ نتیجه. روش نمونه‌برداری پیشین یک روش نمونه‌برداری سازگار است!

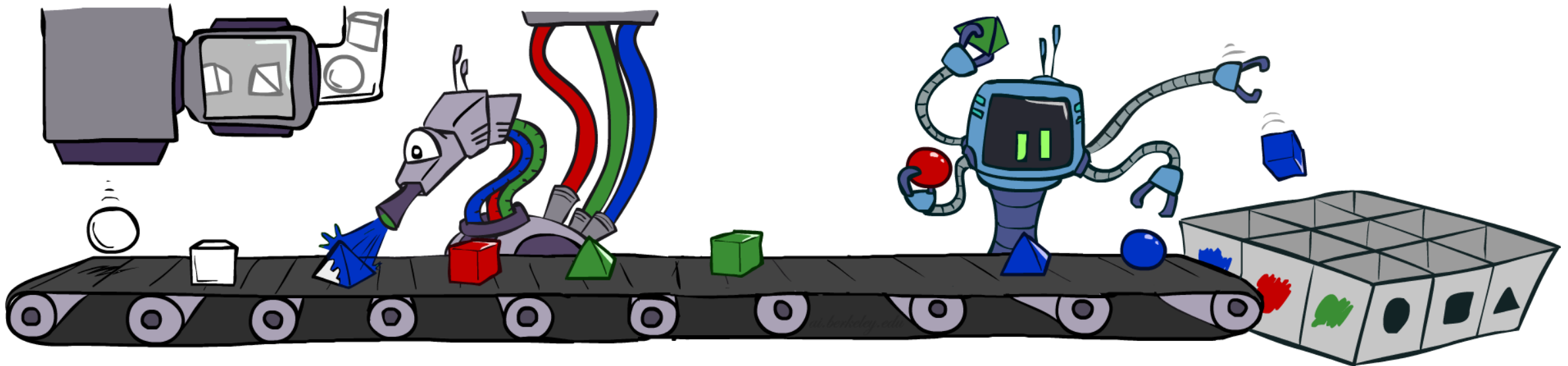
نمونه برداری پیشین

۱۶

```
for i = 1, 2, ..., n
```

```
    sample  $x_i$  from  $P(X_i | \text{parents}(X_i))$ 
```

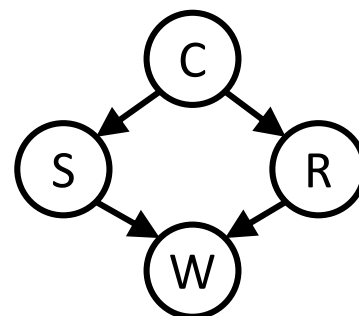
```
return  $(x_1, x_2, \dots, x_n)$ 
```



مثال (۱)

۱۷

□ نمونه برداری. ابتدا از شبکه بیز چند مرتبه نمونه برداری می کنیم:



$+c$	$-s$	$+r$	$+w$
$+c$	$+s$	$+r$	$+w$
$-c$	$+s$	$+r$	$-w$
$+c$	$-s$	$+r$	$+w$
$-c$	$-s$	$-r$	$+w$

□ مثال. محاسبه احتمال $P(W)$.

$$\langle +w: 4.0, -w: 1.0 \rangle$$

□ محاسبه فراوانی:

$$P(W) = \langle +w: 0.8, -w: 0.2 \rangle$$

□ نرمال سازی:

□ با افزایش تعداد نمونه ها، توزیع تخمین زده شده به توزیع واقعی نزدیک تر می شود.

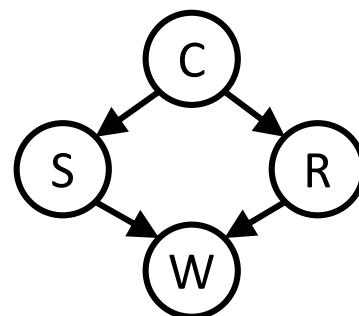
□ به همین روش می توان مقادیر دیگر را نیز تخمین زد.

مثال (۲)

۱۸

□ نمونه برداری. ابتدا از شبکه بیز چند مرتبه نمونه برداری می کنیم:

$+c$	$-s$	$+r$	$+w$
$+c$	$+s$	$+r$	$+w$
$-c$	$+s$	$+r$	$-w$
$+c$	$-s$	$+r$	$+w$
$-c$	$-s$	$-r$	$+w$



□ مثال. محاسبه احتمال $P(+s|+r)$

□ محاسبه فراوانی: $\langle +r: 4.0, +s, +r: 2.0 \rangle$

□ نرمال سازی: $P(+s|+r) = \langle 0.5 \rangle$

□ با افزایش تعداد نمونه ها، توزیع تخمین زده شده به توزیع واقعی نزدیک تر می شود.

□ به همین روش می توان مقادیر دیگر را نیز تخمین زد.

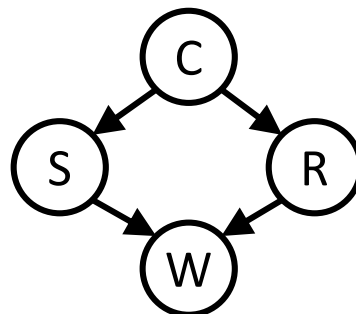
$+c$	$-s$	$+r$	$+w$
$+c$	$+s$	$+r$	$+w$
$-c$	$+s$	$+r$	$+w$
$+c$	$-s$	$+r$	$+w$

مثال (۳)

۱۹

□ نمونه برداری. ابتدا از شبکه بیز چند مرتبه نمونه برداری می کنیم:

$+c$	$-s$	$+r$	$+w$
$+c$	$+s$	$+r$	$+w$
$-c$	$+s$	$+r$	$-w$
$+c$	$-s$	$+r$	$+w$
$-c$	$-s$	$-r$	$+w$



□ مثال. محاسبه احتمال $P(C|-r, -w)$

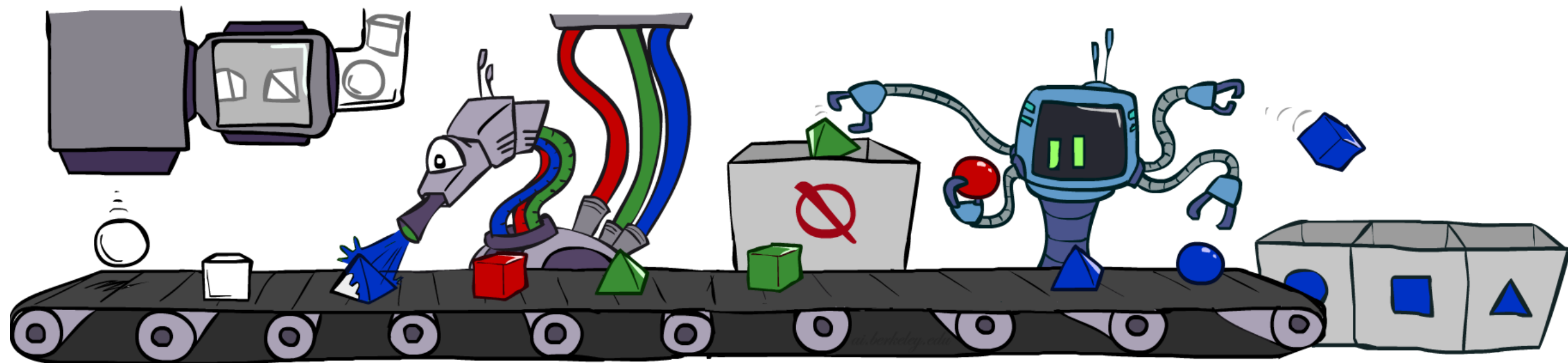
□ محاسبه فراوانی: **صفر**

□ بنابراین با توجه به نمونه های موجود، احتمال $P(C|-r, -w)$ قابل تخمین نیست!

□ سرعت: در صورت کمبود زمان می توان از نمونه های کمتری استفاده نمود. (ایراد؟)

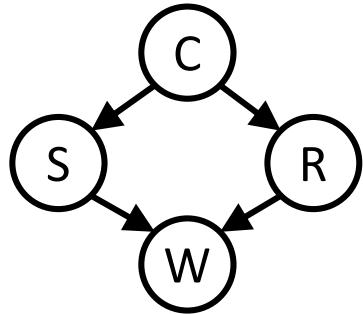
نمونه‌برداری با امکان رد کردن نمونه‌ها

۲۰



نمونه‌برداری با امکان رد کردن نمونه‌ها

۲۱



□ تخمین $P(C)$

- در نمونه‌برداری، نیازی به ذخیره کردن همه نمونه‌ها با هم نداریم.
- می‌توانیم همزمان با تولید نمونه‌ها، عمل شمارش را نیز انجام دهیم.

□ تخمین $P(C|+s)$

- همانند قبل، اما این بار نمونه‌هایی را که با شواهد $(+s)$ منطبق نیستند رد کن.
- به این روش، **نمونه‌برداری با امکان رد کردن** می‌گویند.
- این روش نمونه‌برداری نیز یک روش **سازگار** است.

نمونه‌های ۱، ۴ و ۵ رد می‌شوند.

(1)	$+c$	$-s$	$+r$	$+w$
(2)	$+c$	$+s$	$+r$	$+w$
(3)	$-c$	$+s$	$+r$	$-w$
(4)	$+c$	$-s$	$+r$	$+w$
(5)	$-c$	$-s$	$-r$	$+w$



$P(C +s)$		
$+s$	$+c$	0.5
$+s$	$-c$	0.5

■ با افزایش تعداد نمونه‌ها، احتمال تخمینی به احتمال واقعی همگرا می‌شود.

نمونه‌برداری با امکان رد کردن نمونه‌ها

۲۲

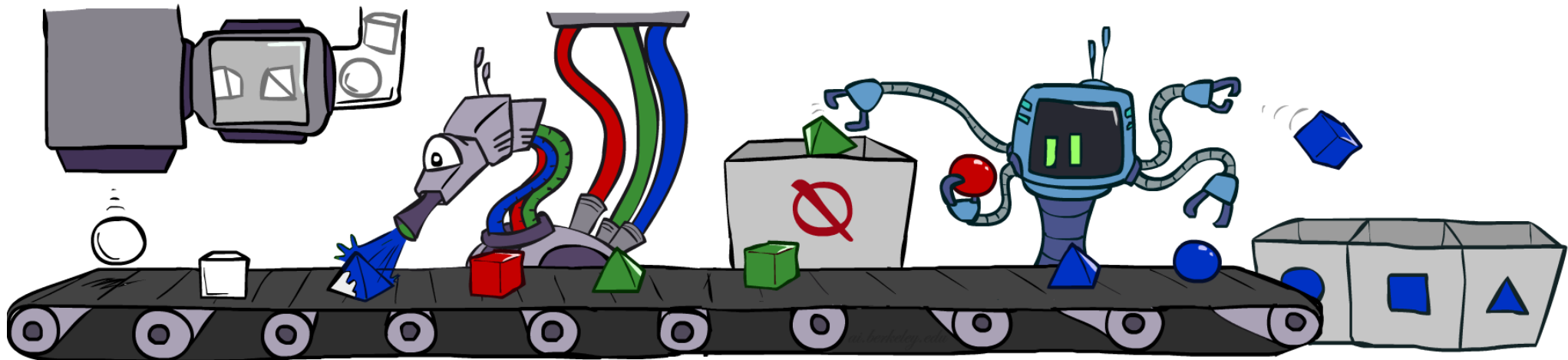
INPUT: evidence instantiation

for $i = 1, 2, \dots, n$

sample x_i from $P(X_i | \text{parents}(X_i))$

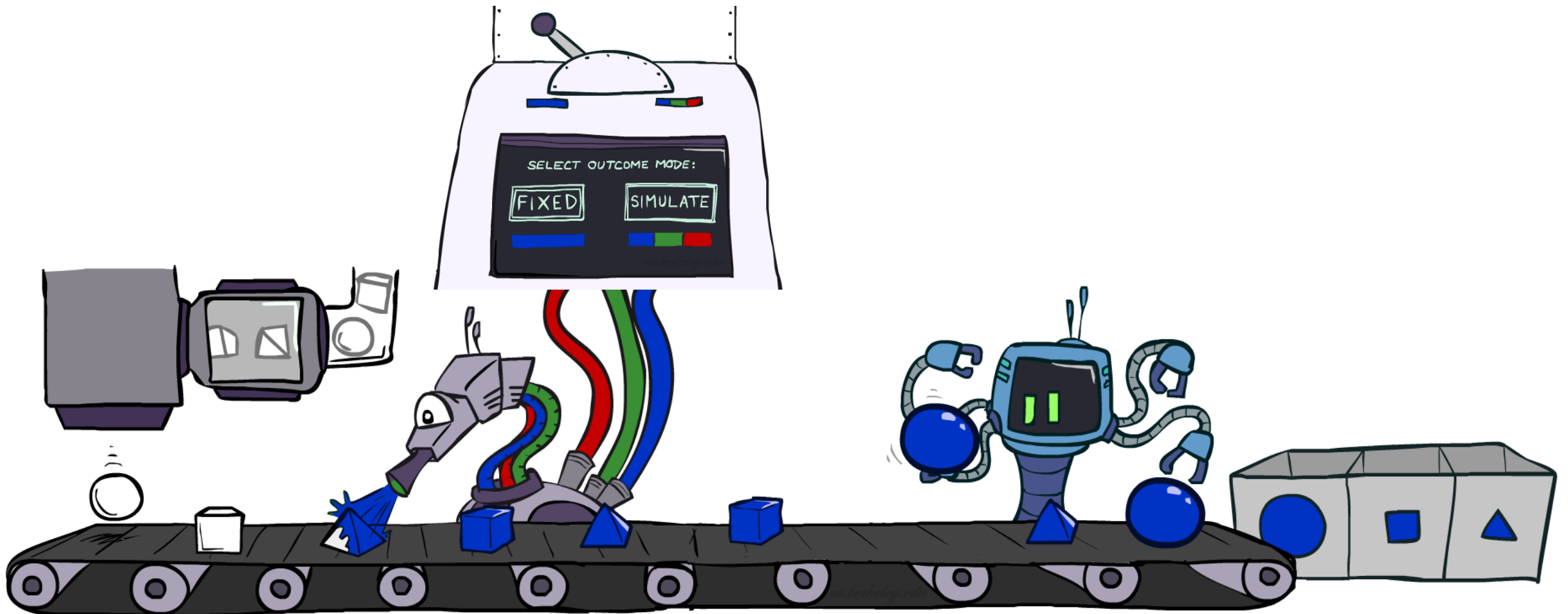
if x_i is not consistent with evidence return null

return $(x_1, x_2, \dots, x_n)$



وزن‌دهی نمونه‌ها

۲۳



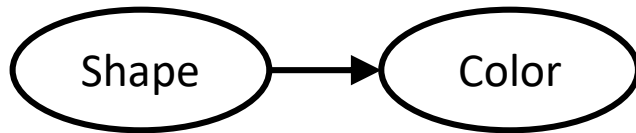
وزن دهی نمونه‌ها

۲۴

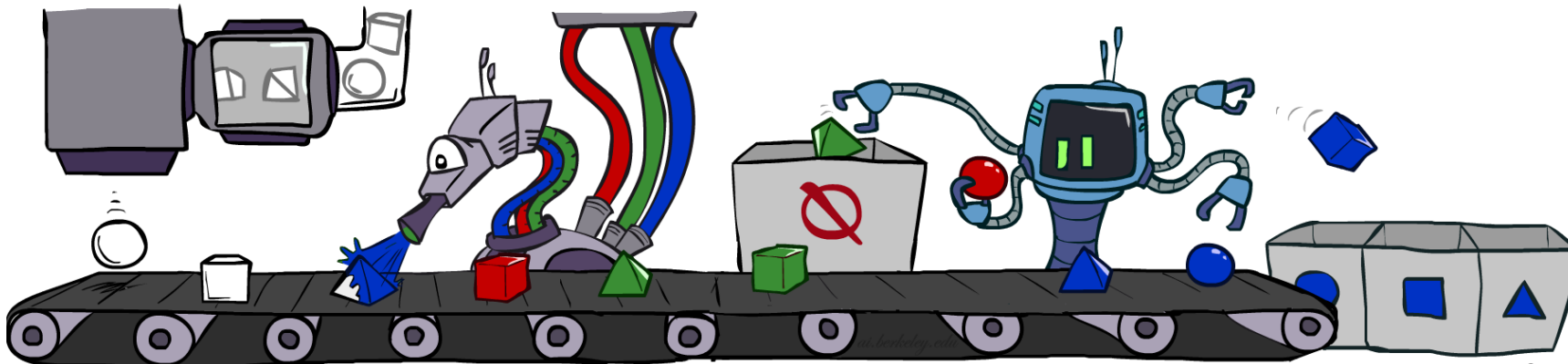
❑ مشکلات نمونه‌برداری با امکان رد کردن.

❑ رد کردن بسیاری از نمونه‌ها در صورت پایین بودن احتمال شواهد.

❑ عدم بهره‌برداری از شواهد در حین نمونه‌برداری.



❑ فرض کنید می‌خواهیم $P(\text{Shape} | \text{blue})$ را تخمین بزنیم.



نمونه‌ها

هرم،	سبز
هرم،	قرمز
کره،	آبی
مکعب،	قرمز
کره،	سبز

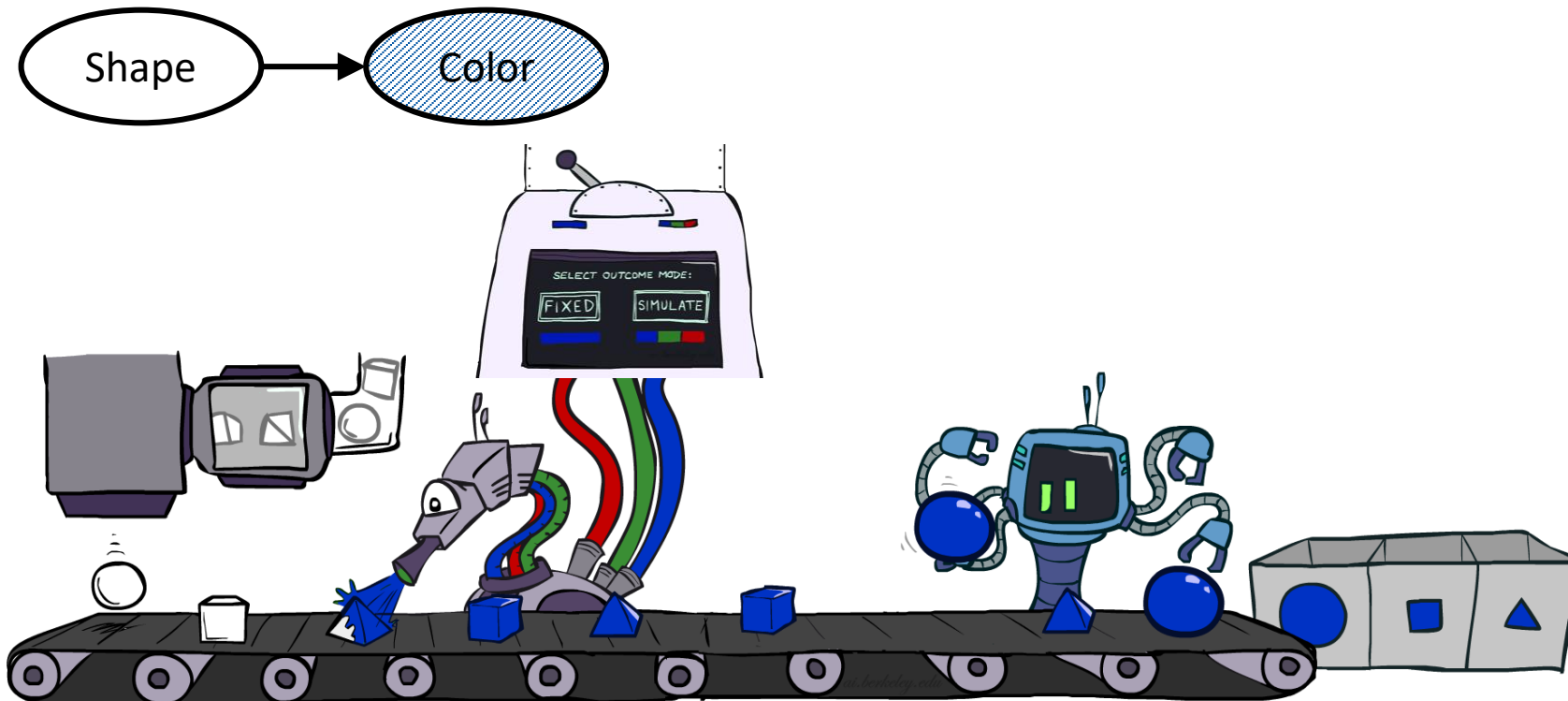
وزن‌دهی نمونه‌ها

۲۵

□ ایده. مقدار متغیرهای مشاهده شده را ثابت در نظر بگیر و از بقیه متغیرها نمونه‌برداری کن.

□ مشکل: توزیع نمونه‌ها سازگار نیست.

□ راه‌حل: ضرب کردن در احتمال شرطی شواهد با داشتن والدین.



نمونه‌ها

هرم،	آبی
هرم،	آبی
کره،	آبی
مکعب،	آبی
کره،	آبی

وزن دهی نمونه‌ها: مثال

۲۶

□ مرتب‌سازی توپولوژیکی رئوس.

$$C \rightarrow S \rightarrow R \rightarrow W$$

□ نمونه‌ها.

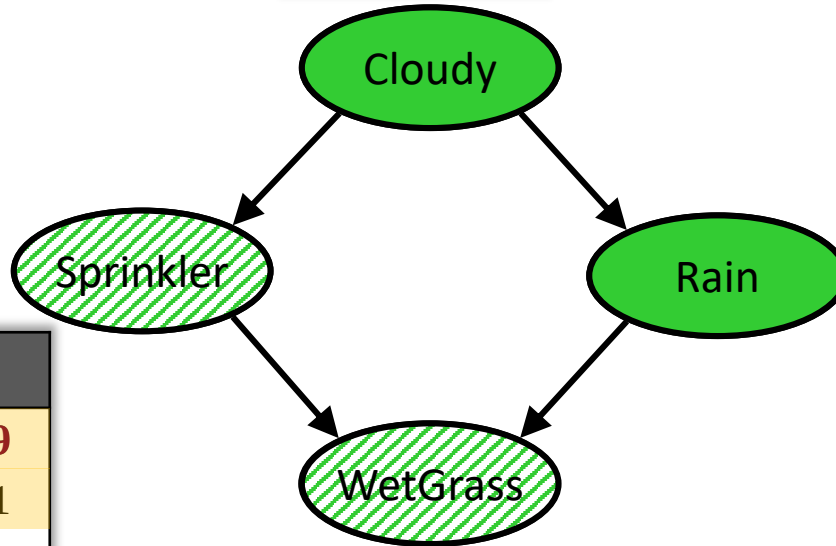
$$+c \quad +s \quad +r \quad +w$$

C	$P(C)$
$+c$	0.5
$-c$	0.5

$P(R C)$		
$+c$	$+r$	0.8
$+c$	$-r$	0.2
$-c$	$+r$	0.2
$-c$	$-r$	0.8

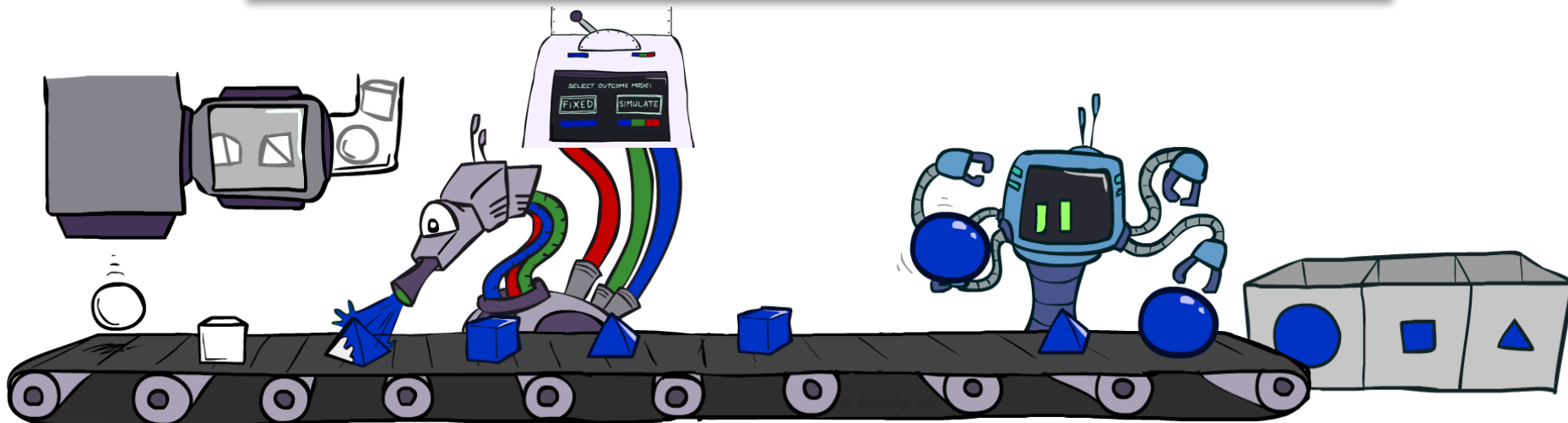
$P(S C)$		
$+c$	$+s$	0.1
$+c$	$-s$	0.9
$-c$	$+s$	0.5
$-c$	$-s$	0.5

$P(W S, R)$			
$+s$	$+r$	$+w$	0.99
$+s$	$+r$	$-w$	0.01
$+s$	$-r$	$+w$	0.90
$+s$	$-r$	$-w$	0.10
$-s$	$+r$	$+w$	0.90
$-s$	$+r$	$-w$	0.10
$-s$	$-r$	$+w$	0.01
$-s$	$-r$	$-w$	0.99



$$W = 1.0 \times 0.1 \times 0.99$$

```
INPUT: evidence instantiation  
 $w = 1.0$   
for  $i = 1, 2, \dots, n$   
  if  $X_i$  is an evidence variable  
     $X_i = \text{observation } x_i \text{ for } X_i$   
     $w = w \times P(x_i | \text{parents}(X_i))$   
  else  
    sample  $x_i$  from  $P(X_i | \text{parents}(X_i))$   
return  $(x_1, x_2, \dots, x_n), w$ 
```



وزن‌دهی نمونه‌ها

۲۸

□ توزیع نمونه‌ها، اگر Z نمونه‌برداری شده باشد و e ثابت باشد.

$$S_{WS}(z, e) = \prod_{i=1}^l P(z_i | \text{parents}(Z_i))$$

□ در نظر گرفتن وزن مربوط به نمونه‌ها.

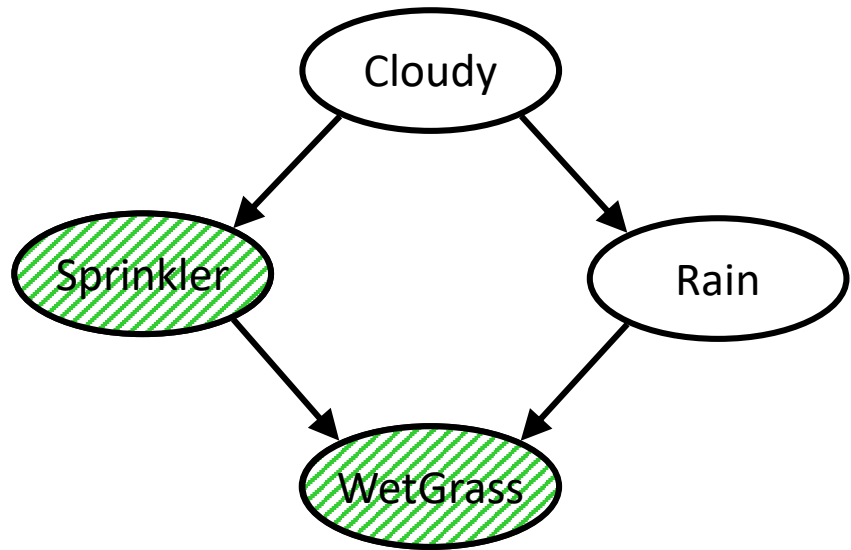
$$w(z, e) = \prod_{i=1}^m P(e_i | \text{parents}(E_i))$$

□ در نتیجه، نمونه‌برداری با وزن‌دهی نمونه‌ها یک روش سازگار است.

$$\begin{aligned} S_{WS}(z, e) \times w(z, e) &= \prod_{i=1}^l P(z_i | \text{parents}(Z_i)) \prod_{i=1}^m P(e_i | \text{parents}(E_i)) \\ &= P(z, e) \end{aligned}$$

وزن دهی نمونه‌ها: مثال

۲۹



نمونه برداری

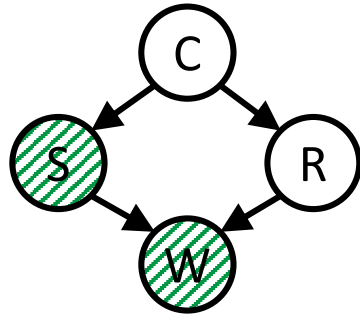
$+c$	$+s$	$+r$	$+w$	0.2
$+c$	$+s$	$-r$	$+w$	0.3
$-c$	$+s$	$+r$	$+w$	0.1

گام ۲. محاسبه $P(+r|+s, +w)$

$$P(+r|+s, +w) = \frac{0.2 + 0.1}{0.2 + 0.3 + 0.1} = 0.5$$

وزن‌دهی نمونه‌ها

۳۰



□ وزن‌دهی نمونه‌ها خوب است.

□ در حین تولید نمونه‌ها، شواهد در نظر گرفته می‌شوند.

□ با توجه به شواهد دریافتی، نمونه‌های بیشتری بیانگر حالت محیط هستند.

□ اما وزن‌دهی نمونه‌ها همه مشکلات ما را حل نمی‌کند.

□ در حین نمونه‌برداری، شواهد تنها بر روی متغیرهای بعد از خود (در ترتیب توپولوژیکی) تأثیر می‌گذارند و بر روی متغیرهای قبل از خود هیچ گونه تأثیری ندارند.

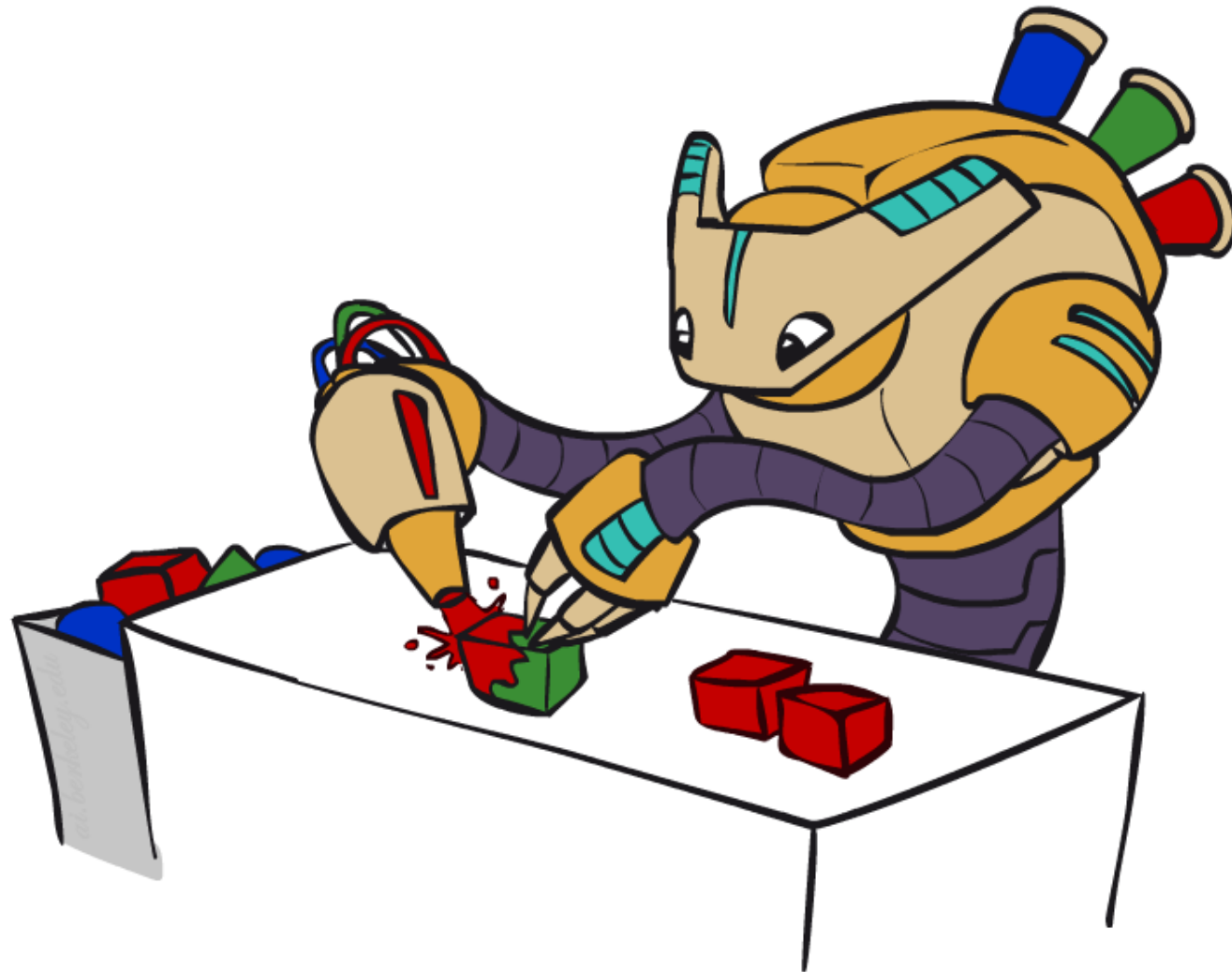
□ به عنوان مثال، مشاهده مقادیر S و W هیچ تأثیری در نمونه‌برداری از متغیر C ندارد.

□ ایده. در نظر گرفتن شواهد هنگام نمونه‌برداری از تمام متغیرها.

□ نمونه‌برداری گیبس!

نمونه برداری گیس

۳۱



نمونه برداری گیس

۳۲

□ نمونه برداری گیس.

□ با یک مقداردهی دلخواه به متغیرها که با شواهد سازگار است شروع کن.

□ هر بار با شرطی سازی بر روی مقدار سایر متغیرها از یک متغیر نمونه برداری کن، اما مقدار متغیرهای مشاهده شده را ثابت نگهدار.

□ این عمل را برای یک مدت زمان طولانی تکرار کن.

□ ویژگی. اگر رویه بالا به تعداد نامحدودی تکرار شود، مانند این است که نمونه به دست آمده از توزیع واقعی نمونه برداری شده باشد.

□ مزیت. هم مقدار متغیرهای بعدی و هم مقدار متغیرهای قبلی از شواهد تأثیر می پذیرند.

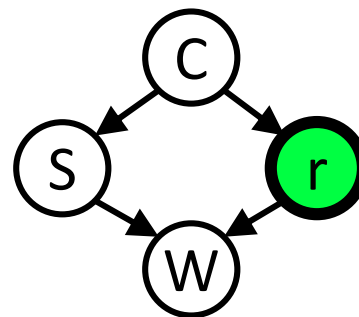
□ مقایسه. در وزن دهی نمونه ها، شواهد تنها بر روی مقدار متغیرهای بعد از خود تأثیر می گذارند. این امر باعث می شود وزن های به دست آمده گاهی اوقات بسیار کوچک باشند. مجموع وزن همه نمونه ها بیانگر تعداد نمونه های «مؤثر» به دست آمده از نمونه برداری است و در نتیجه داشتن وزنها بزرگ برای نمونه ها مطلوب تر است.

نمونه‌برداری گیبس: محاسبه $P(S|+r)$

۳۳

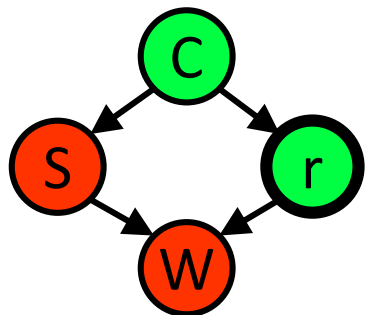
گام ۱. ثابت کردن مقدار شواهد. □

$$R = +r \quad \blacksquare$$



گام ۲. مقداردهی به متغیرهای دیگر. □

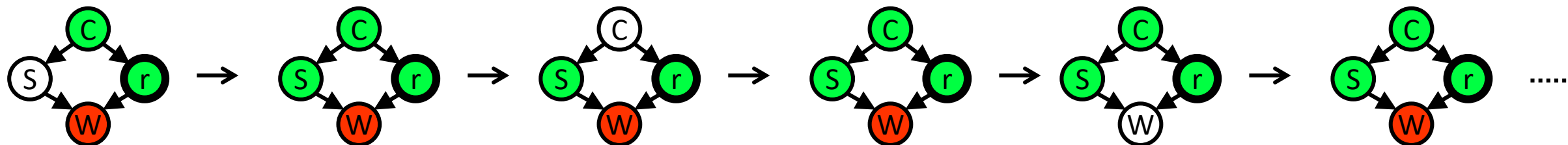
به صورت تصادفی! □



گام ۳. تکرار مراحل زیر: □

انتخاب یک متغیر مشاهده نشده مانند X به صورت تصادفی □

نمونه‌برداری مجدد X از توزیع شرطی $P(X|all other variables)$ □



$$P(S|+c, -w, +r)$$

$$P(C|+s, -w, +r)$$

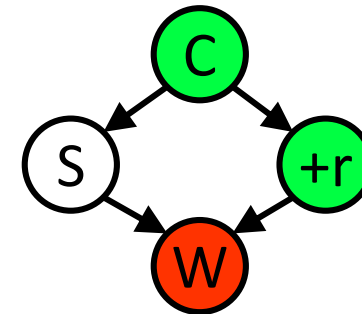
$$P(W|+s, +c, +r)$$

نمونه‌برداری مجدد یک متغیر به صورت کارا

۳۴

□ نمونه‌برداری از $P(S|+c, +r, -w)$

$$\begin{aligned} P(S|+c, +r, -w) &= \frac{P(S, +c, +r, -w)}{P(+c, +r, -w)} \\ &= \frac{P(S, +c, +r, -w)}{\sum_s P(s, +c, +r, -w)} \\ &= \frac{P(+c)P(S|+c)P(+r|+c)P(-w|S, +r)}{\sum_s P(+c)P(s|+c)P(+r|+c)P(-w|s, +r)} \\ &= \frac{P(+c)P(S|+c)P(+r|+c)P(-w|S, +r)}{P(+c)P(+r|+c) \sum_s P(s|+c)P(-w|s, +r)} \\ &= \frac{P(S|+c)P(-w|S, +r)}{\sum_s P(s|+c)P(-w|s, +r)} \end{aligned}$$



□ توجه. تنها جداولی که شامل متغیر S هستند باید در نظر گرفته شده و با یکدیگر الحاق شوند.

مطالعه بیشتر*

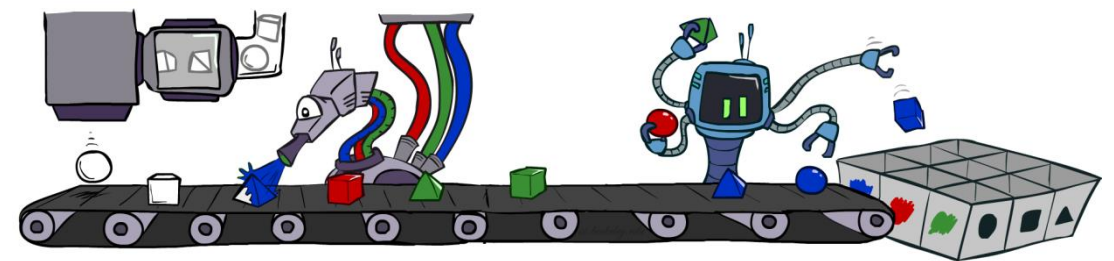
۳۵

- نمونه‌برداری گیبس. به صورت حدی با انجام تعداد نامحدودی از نمونه‌برداری‌های مجدد، نمونه‌هایی از پرس و جوی $P(Q|e)$ تولید می‌کند.
- ایده. به جای این که هر بار یک نمونه را از ابتدا تولید کنی، نمونه‌هایی تولید کن که هر کدام شبیه به آخرین نمونه تولید شده هستند.
- زنجیره مارکوف مونت کارلو. نمونه‌برداری گیبس یک حالت خاص از یک خانواده از روش‌های عمومی‌تر نمونه‌برداری است که روش‌های زنجیره مارکوف مونت کارلو نام دارند.
 - الگوریتم متروپلیس-هاستینگز یکی از معروف‌ترین الگوریتم‌های متعلق به این خانواده است.
 - در واقع نمونه‌برداری گیبس یک حالت خاص از الگوریتم متروپلیس-هاستینگز است.
- روش‌های مونت کارلو. برای مطالعه‌ی بیشتر در این زمینه می‌توانید روش‌های مونت کارلو را مطالعه کنید.
 - روش‌های مونت کارلو روش‌هایی برای نمونه‌برداری هستند.

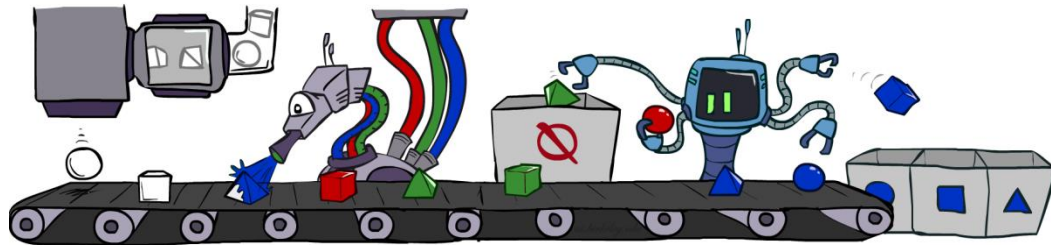
نمونه‌برداری از شبکه ییز: خلاصه

۳۶

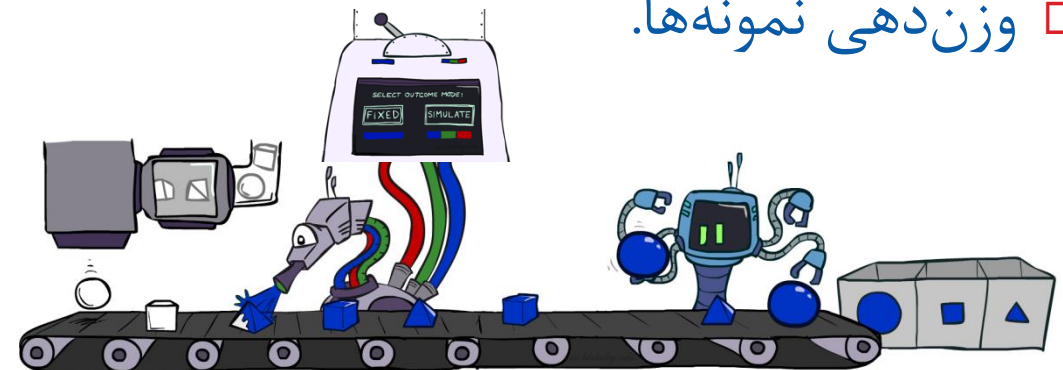
□ نمونه‌برداری پیشین.



□ نمونه‌برداری با امکان رد کردن.



□ وزن‌دهی نمونه‌ها.



□ نمونه‌برداری گیبس.

