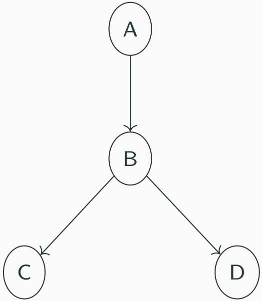
Lecture 25: Approximate Inference in Bayesian Networks

Last Lecture Takeaway

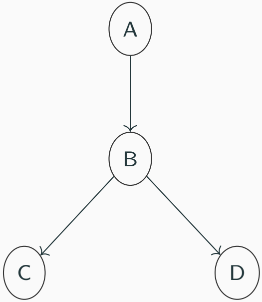
Inference-By-Enumeration/Variable Elimination algorithms allow for General Inference on Bayesian Networks

Complexity of Exact Inference



$$
\begin{aligned}
P(A|d) &= \alpha \sum_b \sum_c P(A, b, c, d) \\
&= \alpha P(A) \sum_b \sum_c P(b|A) P(c|b) P(d|b) \\
&= \alpha P(A) \sum_b P(b|A) P(d|b) \sum_c P(c|b) \\
&= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)} \underbrace{P(d|b)}_{f_3(B)} \underbrace{\sum_c P(c|b)}_{=1} \\
&= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B) \\
&= \alpha f_1(A) \times f_4(A) \\
&= \dots
\end{aligned}
$$

Complexity of Exact Inference



$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

$$= \alpha P(A) \sum_b \sum_c P(b|A) P(c|b) P(d|b)$$

$$= \alpha P(A) \sum_b P(b|A) P(d|b) \sum_c P(c|b)$$

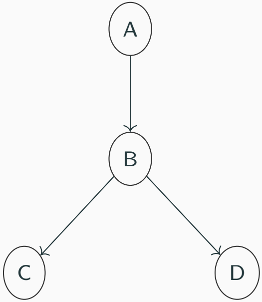
$$= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)} \underbrace{P(d|b)}_{f_3(B)} \underbrace{\sum_c P(c|b)}_{=1}$$

$$= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B)$$

$$= \alpha f_1(A) \times f_4(A)$$

$$= \dots$$

Complexity of Exact Inference



$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

$$= \alpha P(A) \sum_b \sum_c P(b|A)P(c|b)P(d|b)$$

$$= \alpha P(A) \sum_b P(b|A)P(d|b) \sum_c P(c|b)$$

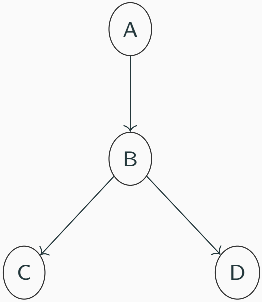
$$= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)} \underbrace{P(d|b)}_{f_3(B)} \underbrace{\sum_c P(c|b)}_{=1}$$

$$= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B)$$

$$= \alpha f_1(A) \times f_4(A)$$

$$= \dots$$

Complexity of Exact Inference



$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

$$= \alpha P(A) \sum_b \sum_c P(b|A)P(c|b)P(d|b)$$

$$= \alpha P(A) \sum_b P(b|A)P(d|b) \sum_c P(c|b)$$

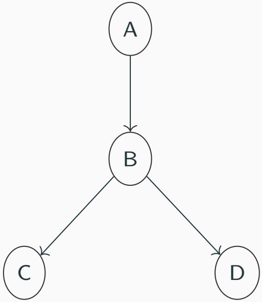
$$= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)}\underbrace{P(d|b)}_{f_3(B)}\underbrace{\sum_c P(c|b)}_{=1}$$

$$= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B)$$

$$= \alpha f_1(A) \times f_4(A)$$

$$= \ldots$$

Complexity of Exact Inference



$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

$$= \alpha P(A) \sum_b \sum_c P(b|A)P(c|b)P(d|b)$$

$$= \alpha P(A) \sum_b P(b|A)P(d|b) \sum_c P(c|b)$$

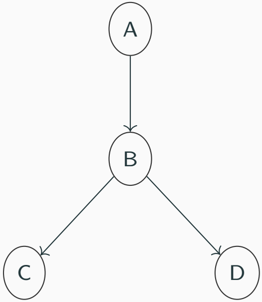
$$= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)} \underbrace{P(d|b)}_{f_3(B)} \underbrace{\sum_c P(c|b)}_{=1}$$

$$= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B)$$

$$= \alpha f_1(A) \times f_4(A)$$

$$= \dots$$

Complexity of Exact Inference



$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

$$= \alpha P(A) \sum_b \sum_c P(b|A)P(c|b)P(d|b)$$

$$= \alpha P(A) \sum_b P(b|A)P(d|b) \sum_c P(c|b)$$

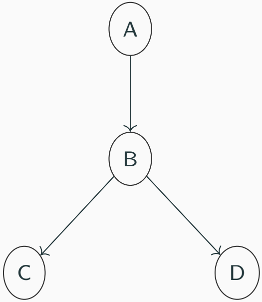
$$= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)} \underbrace{P(d|b)}_{f_3(B)} \underbrace{\sum_c P(c|b)}_{=1}$$

$$= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B)$$

$$= \alpha f_1(A) \times f_4(A)$$

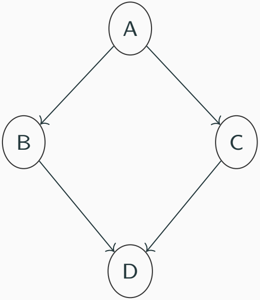
$$= \ldots$$

Complexity of Exact Inference



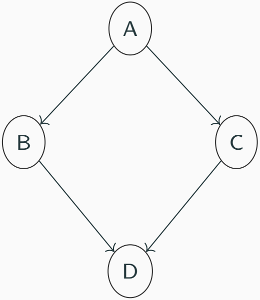
$$
\begin{aligned}
P(A|d) &= \alpha \sum_b \sum_c P(A, b, c, d) \\
&= \alpha P(A) \sum_b \sum_c P(b|A) P(c|b) P(d|b) \\
&= \alpha P(A) \sum_b P(b|A) P(d|b) \sum_c P(c|b) \\
&= \alpha \underbrace{P(A)}_{f_1(A)} \sum_b \underbrace{P(b|A)}_{f_2(B,A)} \underbrace{P(d|b)}_{f_3(B)} \underbrace{\sum_c P(c|b)}_{=1} \\
&= \alpha f_1(A) \sum_b f_2(B, A) \times f_3(B) \\
&= \alpha f_1(A) \times f_4(A) \\
&= \dots
\end{aligned}
$$

Complexity of Exact Inference



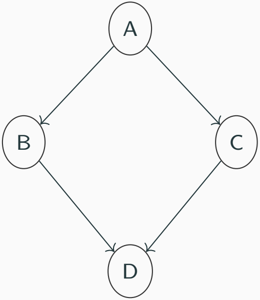
$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

Complexity of Exact Inference



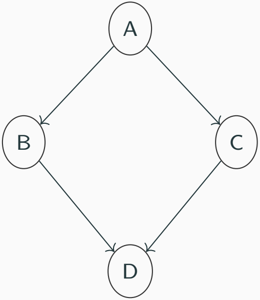
$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$
$$= \alpha P(A) \sum_b P(b|A) \sum_c P(c|A) P(d|b, c)$$

Complexity of Exact Inference



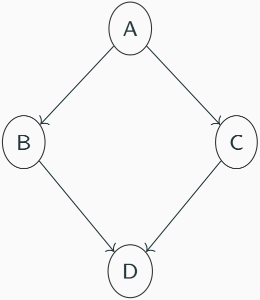
$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$
$$= \alpha P(A) \sum_b P(b|A) \sum_c P(c|A) P(d|b, c)$$
$$= \alpha f_1(A) \sum_b f_2(B, A) \sum_c f_3(C, A) f_4(B, C)$$

Complexity of Exact Inference



$$
\begin{aligned}
P(A|d) &= \alpha \sum_b \sum_c P(A, b, c, d) \\
&= \alpha P(A) \sum_b P(b|A) \sum_c P(c|A) P(d|b, c) \\
&= \alpha f_1(A) \sum_b \color{red}{f_2(B, A)} \color{black} \sum_c f_3(C, A) \color{red}{f_4(B, C)}
\end{aligned}
$$

Complexity of Exact Inference

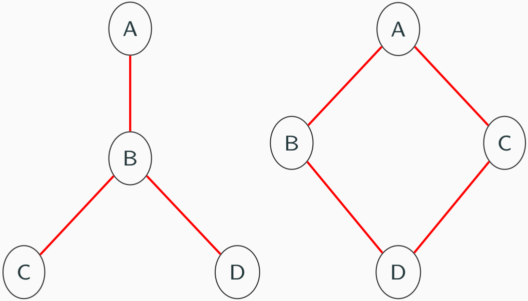


$$P(A|d) = \alpha \sum_b \sum_c P(A, b, c, d)$$

$$= \alpha P(A) \sum_b P(b|A) \sum_c P(c|A) P(d|b, c)$$

$$= \alpha f_1(A) \sum_b f_2(B, A) \sum_c f_3(C, A) f_4(B, C)$$

Complexity of Exact Inference

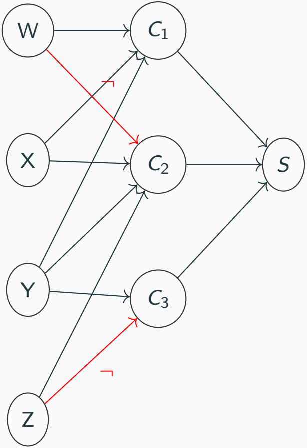


Complexity of Exact Inference

$(W \vee X \vee Y) \wedge (\neg W \vee Y \vee Z) \wedge (X \vee Y \vee \neg Z)$

Complexity of Exact Inference

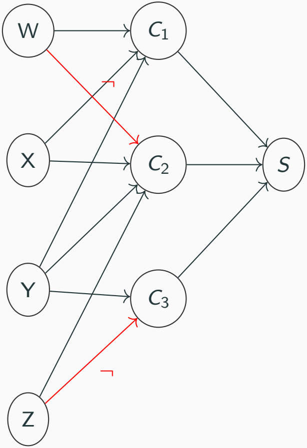
$(W \vee X \vee Y) \wedge (\neg W \vee Y \vee Z) \wedge (X \vee Y \vee \neg Z)$

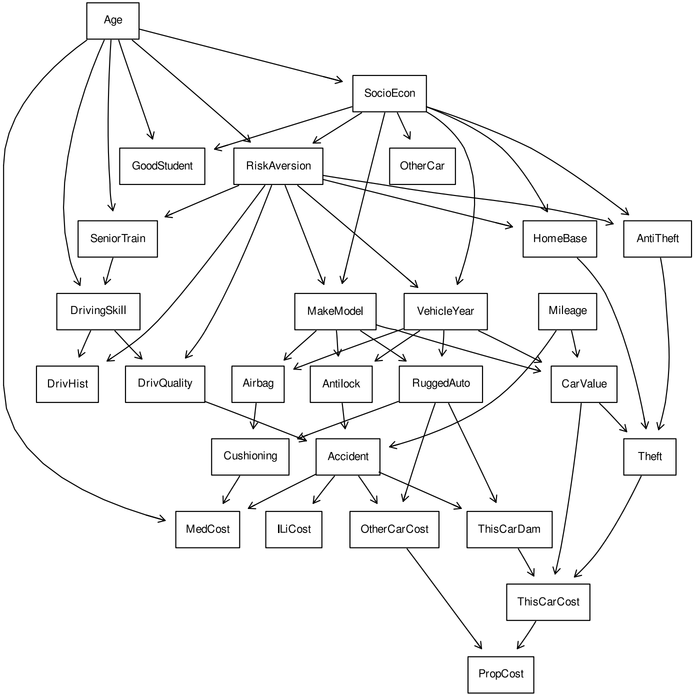


Complexity of Exact Inference

$(W \vee X \vee Y) \wedge (\neg W \vee Y \vee Z) \wedge (X \vee Y \vee \neg Z)$

$P(S = true)$?




Age
SocioEcon
GoodStudent
RiskAversion
OtherCar
SeniorTrain
HomeBase
AntiTheft
DrivingSkill
MakeModel
VehicleYear
Mileage
DrivHist
DrivQuality
Airbag
Antilock
RuggedAuto
CarValue
Cushioning
Accident
Theft
MedCost
ILiCost
OtherCarCost
ThisCarDam
ThisCarCost
PropCost

Approximate Inference

- Heads

$$P(Coin) = \langle 0.5, 0.5 \rangle$$

Approximate Inference

- Heads
- Tails

$$P(Coin) = \langle 0.5, 0.5 \rangle$$

Approximate Inference

- Heads
- Tails
- Tails

$$P(Coin) = \langle 0.5, 0.5 \rangle$$

Approximate Inference

$$P(Coin) = \langle 0.5, 0.5 \rangle$$

- Heads
- Tails
- Tails
- Heads

Approximate Inference

$$P(Coin) = \langle 0.5, 0.5 \rangle$$

- Heads
- Tails
- Tails
- Heads
- Tails

Approximate Inference

$P(Coin) = \langle 0.5, 0.5 \rangle$

- Heads
- Tails
- Tails
- Heads
- Tails
- Heads

Approximate Inference

$P(Coin) = \langle 0.5, 0.5 \rangle$

- Heads
- Tails
- Tails
- Heads
- Tails
- Heads
- Tails

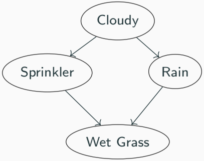
Approximate Inference

$P(Coin) = \langle 0.5, 0.5 \rangle$

- Heads
- Tails
- Tails
- Heads
- Tails
- Heads
- Tails
- Tails

Direct Sampling without Evidence

P(C)
0.5



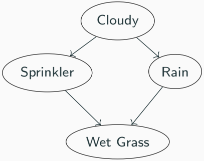
C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

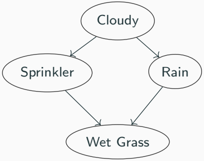
C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$$P(Cloudy) = \langle 0.5, 0.5 \rangle$$

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

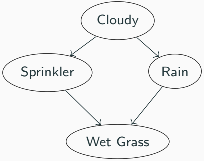
C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

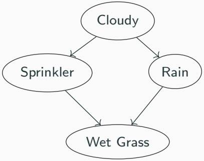
$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$P(Sprinkler | Cloudy = true) = \langle 0.1, 0.9 \rangle$

Direct Sampling without Evidence

P(C)
0.5

C	P(S)
t	0.10
f	0.50



C	P(R)
t	0.8
f	0.2

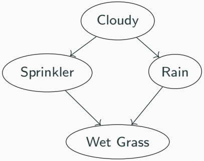
S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$P(Sprinkler | Cloudy = true) = \langle 0.1, 0.9 \rangle$ *false*

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

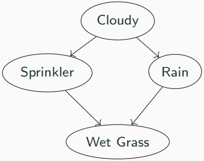
$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$P(Sprinkler | Cloudy = true) = \langle 0.1, 0.9 \rangle$ *false*

$P(Rain | Cloudy = true) = \langle 0.8, 0.2 \rangle$

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

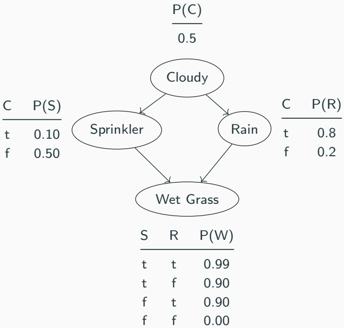
S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$P(Sprinkler|Cloudy = true) = \langle 0.1, 0.9 \rangle$ *false*

$P(Rain|Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

Direct Sampling without Evidence



$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

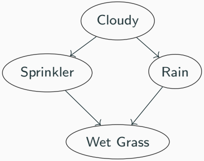
$P(Sprinkler | Cloudy = true) = \langle 0.1, 0.9 \rangle$ *false*

$P(Rain | Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

$P(WetGrass | Sprinkler = false, Rain = true) = \langle 0.9, 0.1 \rangle$

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

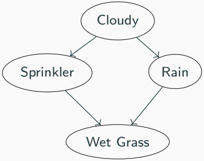
$P(Sprinkler|Cloudy = true) = \langle 0.1, 0.9 \rangle$ *false*

$P(Rain|Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

$P(WetGrass|Sprinkler = false, Rain = true) = \langle 0.9, 0.1 \rangle$ *true*

Direct Sampling without Evidence

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$P(Sprinkler | Cloudy = true) = \langle 0.1, 0.9 \rangle$ *false*

$P(Rain | Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

$P(WetGrass | Sprinkler = false, Rain = true) = \langle 0.9, 0.1 \rangle$ *true*

Event Returned: $[c, \neg s, r, w]$

Direct Sampling without Evidence

Probability of Generating a given sample: $[x_1, \ldots x_n]$

$$S(x_1, \ldots x_n) =$$

Direct Sampling without Evidence

Probability of Generating a given sample: $[x_1, \ldots x_n]$

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

Direct Sampling without Evidence

Probability of Generating a given sample: $[x_1, \dots x_n]$

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

$$\frac{N(x_1, \dots, x_n)}{N}$$

Direct Sampling without Evidence

Probability of Generating a given sample: $[x_1, \ldots x_n]$

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

$$\lim_{n\to\infty} \frac{N(x_1, \ldots, x_n)}{N} = S(x_1, \ldots, x_n) = P(x_1, \ldots, x_n)$$

Direct Sampling without Evidence

Probability of Generating a given sample: $[x_1, \dots x_n]$

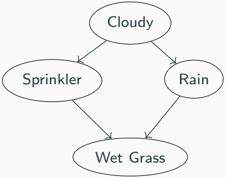
$$S(x_1, \dots x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i)) = P(x_1, \dots, x_n)$$

$$\lim_{n \to \infty} \frac{N(x_1, \dots, x_n)}{N} = S(x_1, \dots, x_n) = P(x_1, \dots, x_n)$$

$$P(x_1, \dots, x_n) \approx \frac{N(x_1, \dots, x_n)}{N}$$

Rejection Sampling

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$$P(Rain|s)$$

$[\neg c, s, r, w]$

$[c, s, r, \neg w]$

$[c, \neg s, r, \neg w]$

$[\neg c, \neg s, r, w]$

$[c, s, r, w]$

$[c, \neg s, \neg r, w]$

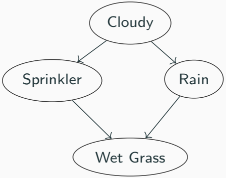
$[\neg c, s, \neg r, \neg w]$

$[c, s, r, w]$

$\dots$

Rejection Sampling

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$$P(Rain|s)$$

$[\neg c, s, r, w]$

$[c, s, r, \neg w]$

~~$[c, \neg s, r, \neg w]$~~

~~$[\neg c, \neg s, r, w]$~~

$[c, s, r, w]$

~~$[c, \neg s, \neg r, w]$~~

$[\neg c, s, \neg r, \neg w]$

$[c, s, r, w]$

...

Rejection Sampling

$$P(X|e)$$

Rejection Sampling

$$P(X|e)$$

$$\frac{N(X, e)}{N(e)} \approx \frac{P(X, e)}{P(e)}$$

Rejection Sampling

$$P(X|e)$$

$$\frac{N(X, e)}{N(e)} \approx \frac{P(X, e)}{P(e)} = P(X|e)$$

Rejection Sampling: $P(Rain|Sprinkler = true)$

100 Samples:

- In 73, $Sprinkler = false$; In 27, $Sprinkler = true$
- Of the 73 where $Sprinkler = false$, 42 have $Rain = true$; 31 have $Rain = false$
- Of the 27 where $Sprinkler = true$, 8 have $Rain = true$; 19 have $Rain = false$

$$P(Rain|Sprinkler = true) \approx$$

Rejection Sampling: $P(Rain|Sprinkler = true)$

100 Samples:

- In 73, $Sprinkler = false$; In 27, $Sprinkler = true$
- Of the 73 where $Sprinkler = false$, 42 have $Rain = true$; 31 have $Rain = false$
- Of the 27 where $Sprinkler = true$, 8 have $Rain = true$; 19 have $Rain = false$

$$P(Rain|Sprinkler = true) \approx \alpha\langle 8, 19\rangle = \langle 0.296, 0.704\rangle$$

Rejection Sampling: $P(Rain|Sprinkler = true)$

100 Samples:

- In 73, $Sprinkler = false$; In 27, $Sprinkler = true$
- Of the 73 where $Sprinkler = false$, 42 have $Rain = true$; 31 have $Rain = false$
- Of the 27 where $Sprinkler = true$, 8 have $Rain = true$; 19 have $Rain = false$

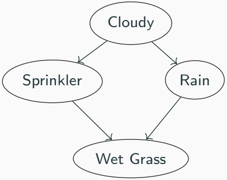
$$P(Rain|Sprinkler = true) \approx \alpha\langle 8, 19\rangle = \langle 0.296, 0.704\rangle$$

(True Answer: $\langle 0.3, 0.7\rangle$)

Rejection Sampling: Problems

P(C)
0.5

C	P(S)
t	0.001
f	0.005



C	P(R)
t	0.8
f	0.2

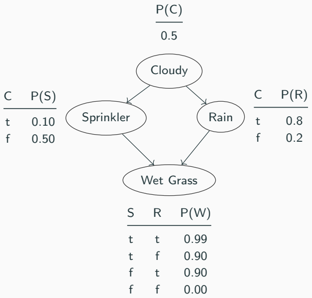
S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$$P(Rain | Sprinkler = true)$$

Likelihood Weighting (Importance Sampling)

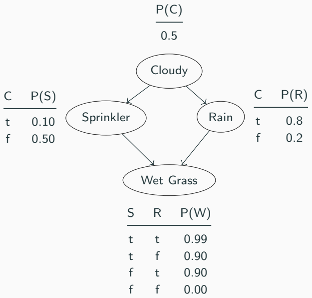
Likelihood Weighting Idea: Let's only generate samples consistent with the evidence

Likelihood Weighting



$$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$$

Likelihood Weighting



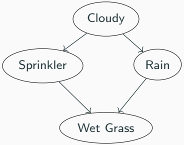
$$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$$

$$w = 1$$

Likelihood Weighting

P(C)
0.5

C	P(S)
t	0.10
f	0.50



C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

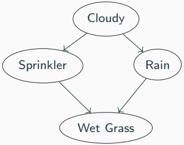
$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$

$w = 1$

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

Likelihood Weighting

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

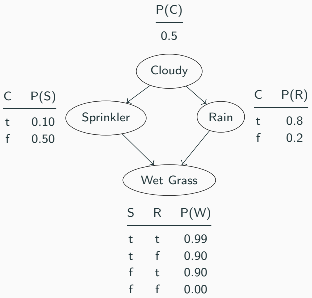
$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$

$w = 1$

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$w \leftarrow w \times P(Sprinkler = true | Cloudy = true) = 0.1$

Likelihood Weighting



$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$

$w = 1$

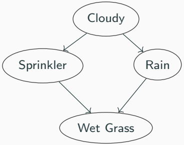
$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$w \leftarrow w \times P(Sprinkler = true|Cloudy = true) = 0.1$

$P(Rain|Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

Likelihood Weighting

P(C)
0.5



C	P(S)
t	0.10
f	0.50

C	P(R)
t	0.8
f	0.2

S	R	P(W)
t	t	0.99
t	f	0.90
f	t	0.90
f	f	0.00

$$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$$

$w = 1$

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

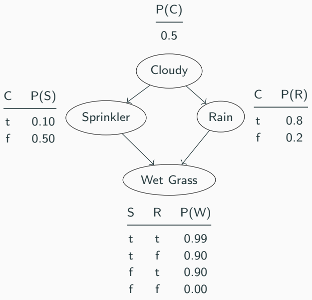
$w \leftarrow w \times P(Sprinkler = true|Cloudy = true) = 0.1$

$P(Rain|Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

$w \leftarrow w \times P(WetGrass = true|Sprinkler = true, Rain = true)$

$= 0.099$

Likelihood Weighting



$\mathbf{P}(\mathbf{Rain}|\mathbf{Sprinkler} = \mathbf{true}, \mathbf{WetGrass} = \mathbf{true})$

$w = 1$

$P(Cloudy) = \langle 0.5, 0.5 \rangle$ *true*

$w \leftarrow w \times P(Sprinkler = true | Cloudy = true) = 0.1$

$P(Rain | Cloudy = true) = \langle 0.8, 0.2 \rangle$ *true*

$w \leftarrow w \times P(WetGrass = true | Sprinkler = true, Rain = true)$

$= 0.099$

Sample $[true, true, true, true]$ with weight 0.099 tallied under $Rain = true$

Likelihood Weighting: Why it Works

Probability of Generating a sample: $[z_1, \ldots, z_l, e_1, \ldots, e_m]$

$$S(z, e) =$$

Likelihood Weighting: Why it Works

Probability of Generating a sample: $[z_1, \ldots, z_l, e_1, \ldots, e_m]$

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

Likelihood Weighting: Why it Works

Probability of Generating a sample: $[z_1, \ldots, z_l, e_1, \ldots, e_m]$

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

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

Likelihood Weighting: Why it Works

Probability of Generating a sample: $[z_1, \ldots, z_l, e_1, \ldots, e_m]$

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

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

$$\underbrace{\prod_{i}^{l} P(z_i | parents(Z_i))}_{S(z,e)} \underbrace{\prod_{i=1}^{m} P(e_i | parents(E_i))}_{w(z,e)}$$

Likelihood Weighting: Why it Works

Probability of Generating a sample: $[z_1, \ldots, z_l, e_1, \ldots, e_m]$

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

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

$$\underbrace{\prod_{i}^{l} P(z_i | parents(Z_i))}_{S(z,e)} \underbrace{\prod_{i=1}^{m} P(e_i | parents(E_i))}_{w(z,e)}$$

Likelihood Weighting: Why it Works

Probability of Generating a sample: $[z_1, \ldots, z_l, e_1, \ldots, e_m]$

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

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

$$\underbrace{\prod_{i}^{l} P(z_i | parents(Z_i))}_{S(z,e)} \underbrace{\prod_{i=1}^{m} P(e_i | parents(E_i))}_{w(z,e)} = P(z, e)$$

Markov Chain Monte Carlo

MCMC Idea: "Random Walk" – Generate an event from the previous event, rather than generate all events from scratch.

Markov Chain Monte Carlo (Gibbs Sampling)

1. Start with initial state
2. Repeat:
   2.1 Choose a random non-evidence variable Z_i
   2.2 Set a new value for Z_i by sampling $P(Z_i|mb(Z_i))$
   2.3 Add current state to the count

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize $Cloudy$ and $Rain$ randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
    - Sample $Cloudy$ from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
    - New state $[false, true, false, true]$
    - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
    - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize $Cloudy$ and $Rain$ randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample $Cloudy$ from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
    - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
    - New state $[false, true, false, true]$
    - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
    - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

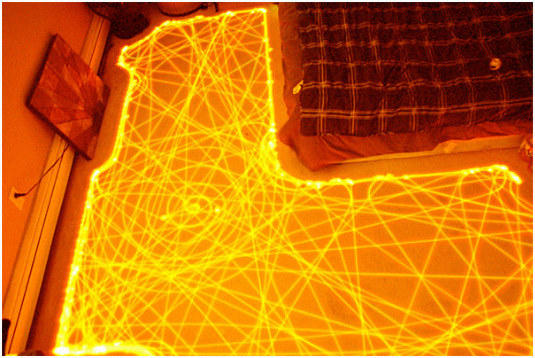
- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

MCMC: $P(Rain|Sprinkler = true, WetGrass = true)$

- Fix $Sprinkler = true$ and $WetGrass = true$
- Initialize *Cloudy* and *Rain* randomly
- **Initial State:** $[true, true, false, true]$
- Repeat:
  - Sample *Cloudy* from $P(Cloudy|Sprinkler = true, Rain = false)$ (false)
  - New state $[false, true, false, true]$
  - Sample Rain from $P(Rain|Sprinkler = true, Cloudy = false, WetGrass = true)$ (true)
  - New state $[false, true, true, true]$

$$\frac{N(r)}{N} \approx \alpha P(r|s, w)$$

(Bonus) Markov Chain Monte Carlo: Why?



(Bonus) Markov Chain Monte Carlo: Why?

The sampling process settles into a "dynamic equilibrium" in which the long-term fraction of time spent in each state is exactly proportional to its posterior probability

(Bonus) Markov Chain Monte Carlo: Why?

The sampling process settles into a "dynamic equilibrium" in which the long-term fraction of time spent in each state is exactly proportional to its posterior probability

$$q(x \to x')$$

$$\pi_{t+1}(x') = \sum_{x} \pi_t(x) q(x \to x')$$

(Bonus) Markov Chain Monte Carlo: Why?

The sampling process settles into a "dynamic equilibrium" in which the long-term fraction of time spent in each state is exactly proportional to its posterior probability

$$
\begin{aligned}
&q(x \to x') \\
&\pi_{t+1}(x') = \sum_x \pi_t(x) q(x \to x')
\end{aligned}
$$

$$\pi(x') = \sum_x \pi(x) q(x \to x') \quad \text{for all } x'$$

Summary

- Complexity of Exact Inference
- Approximate Inference
    - Rejection Sampling
    - Likelihood Weighting
    - Markov Chain Monte Carlo