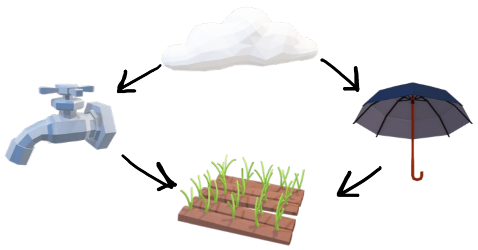


Inferencing on Bayesian Networks: Sampling Methods

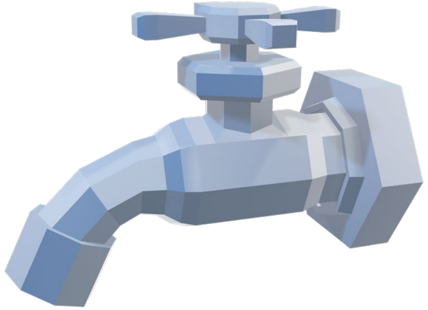
CSC384 – Introduction to Artificial Intelligence
University of Toronto

Chandra Gummaluru

Based on content from Fahiem Bacchus, Sonya Allin, and Steve Engels.

A Bayesian Problem

A farmer buys a new "smart" sprinkler which is supposed to water his crops when it does not rain.



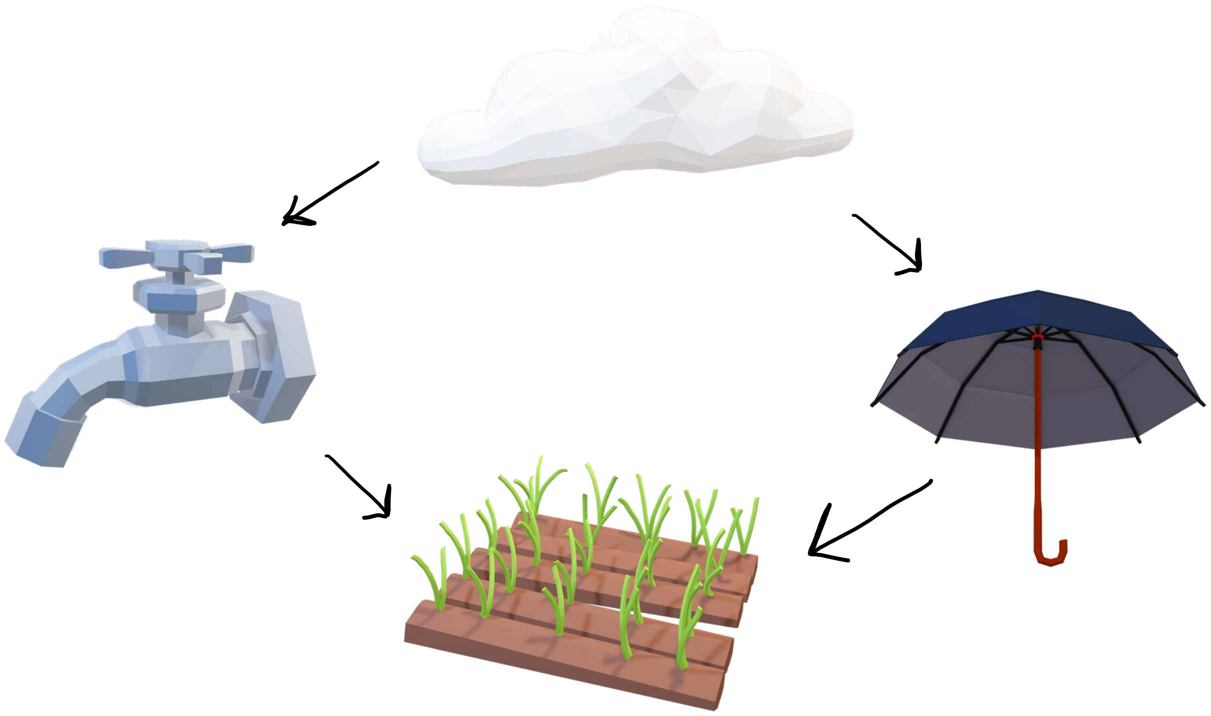
C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

The sprinkler comes equipped with a daylight sensor that lets it know if it is cloudy outside.

It is programmed to water the crops if it is cloudy. However, the sensor is not perfect and the farmer does not want his crops to be over-watered.

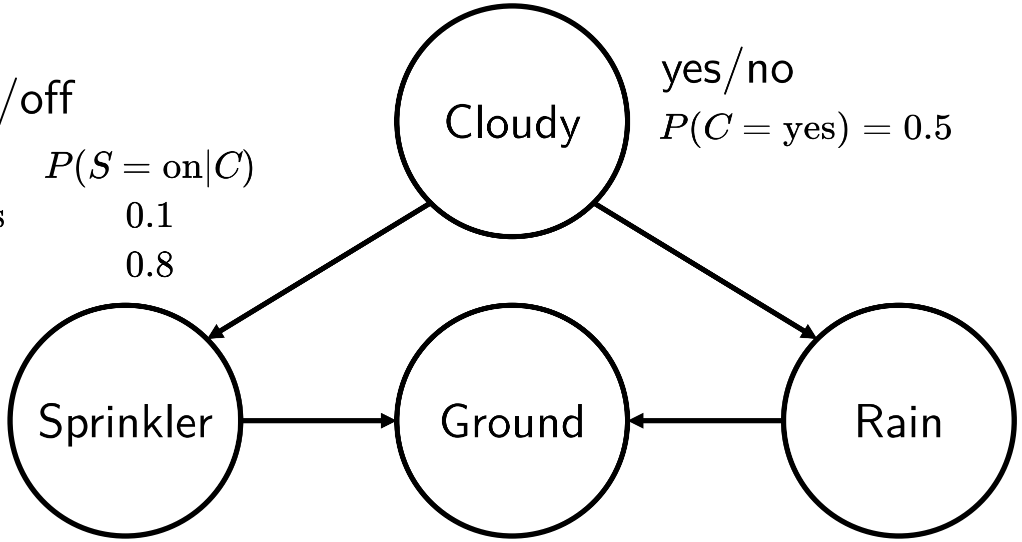
Creating a Bayesian Network

We can model the situation using a Bayesian network by identifying the causal relationships.



Creating a Bayesian Network

We can model the situation using a Bayesian network by identifying the causal relationships.



yes/no

$P(C = \text{yes}) = 0.5$

on/off

C	$P(S = \text{on}\mid C)$
yes	0.1
no	0.8

yes/no

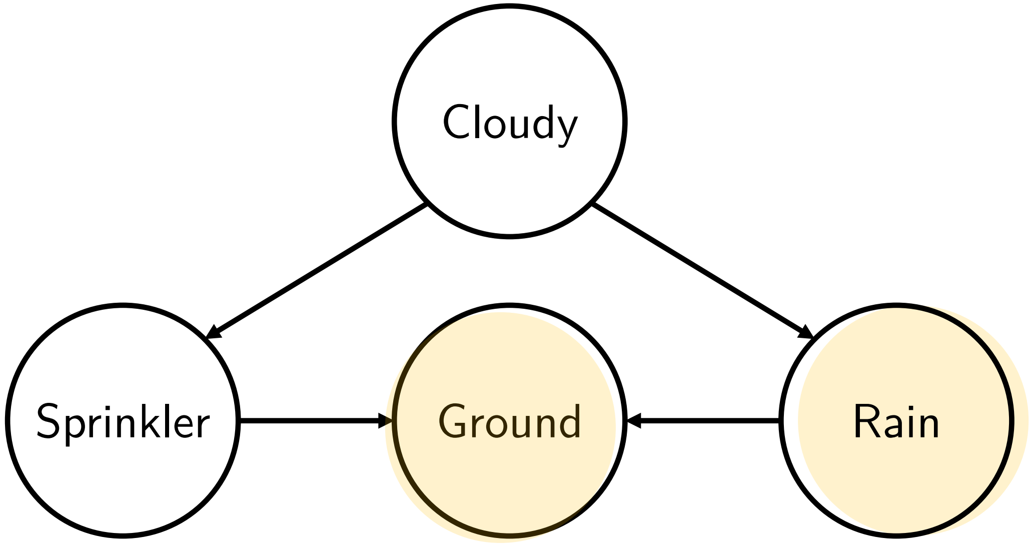
C	$P(R = \text{yes}\mid C)$
yes	0.8
no	0.1

dry/damp/wet

S	R	$P(G = \text{dry}\mid S, R)$	$P(G = \text{damp}\mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

Inference on a Bayesian Network

The farmer wants to know how likely it is for his crops to be over-watered if it rains, i.e., $P(G = \text{wet} | R = \text{yes})$



We can use variable elimination.

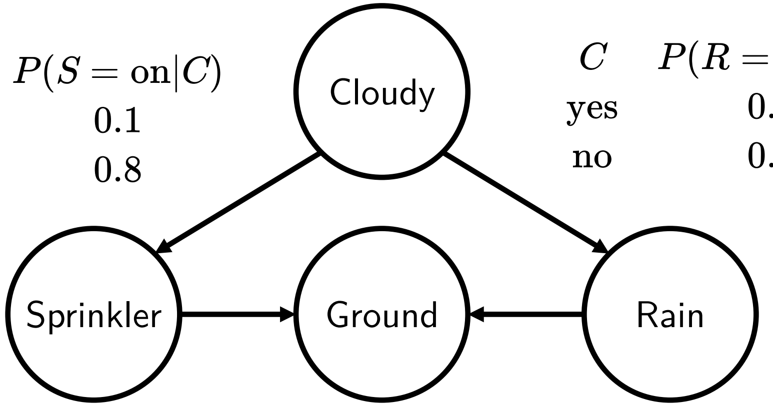
Inference via Variable Elimination

Find $P(G = \text{wet} | R = \text{yes})$

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

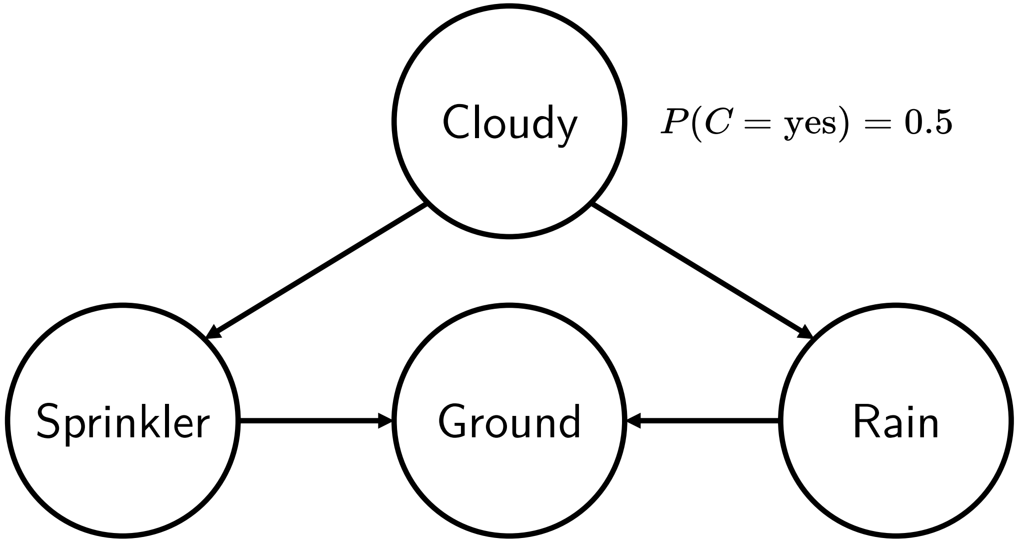


S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

Using variable elimination, we determine $P(G = \text{wet} | R = \text{yes})$ to approximately, 19.68%.

Inference via Sampling

Let's now consider finding $P(G = \text{wet}|R = \text{yes})$ using sampling methods.

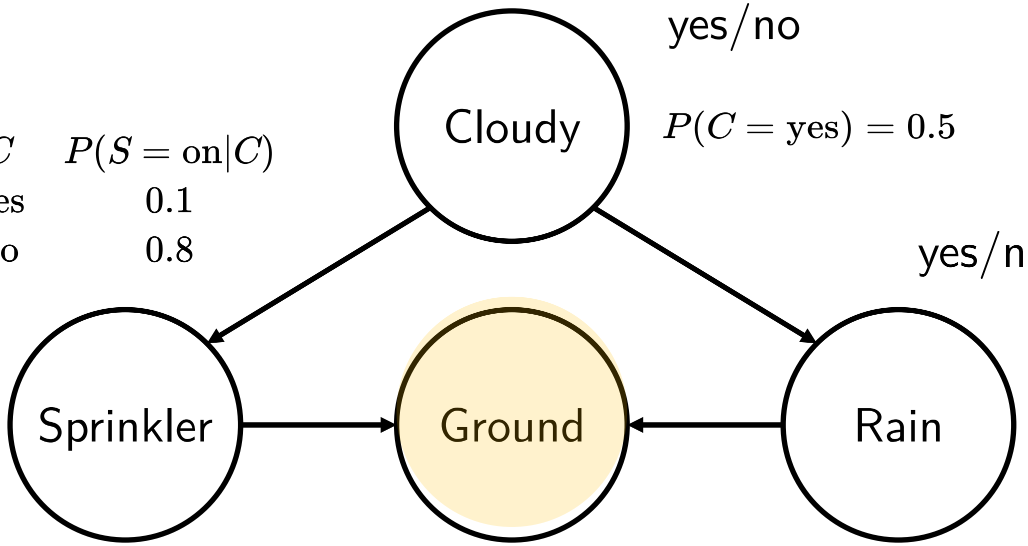


C	$P(R = \text{yes}|C)$
yes	0.8
no	0.1

We can sample the state of a node using its probability table... but, we must know the states of its parents.

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



yes/no

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

yes/no

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

on/off

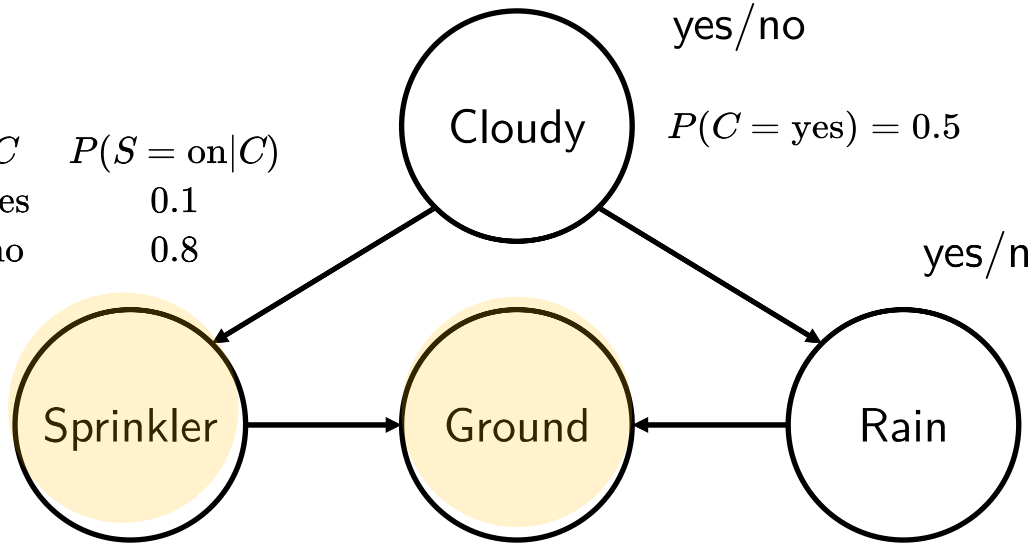
S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

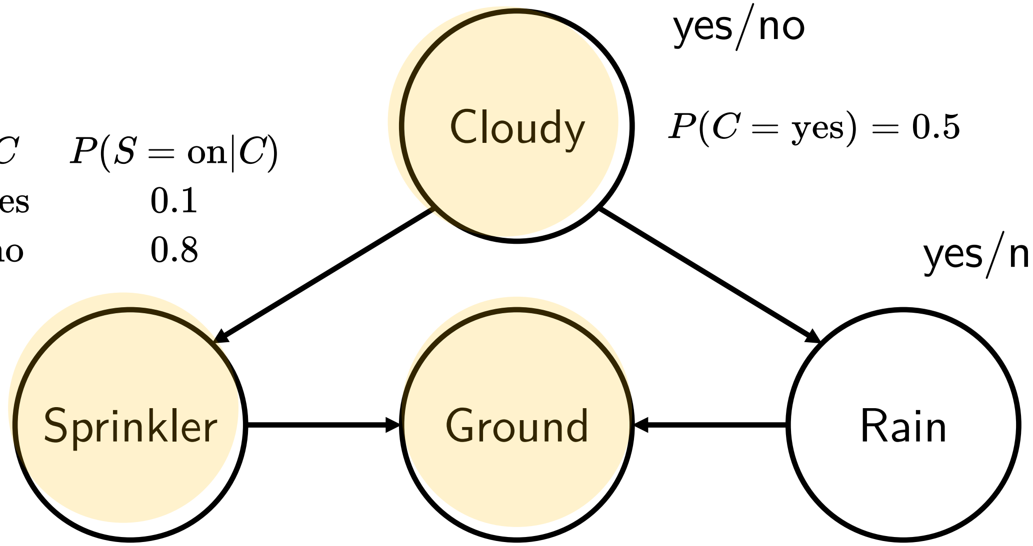
C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

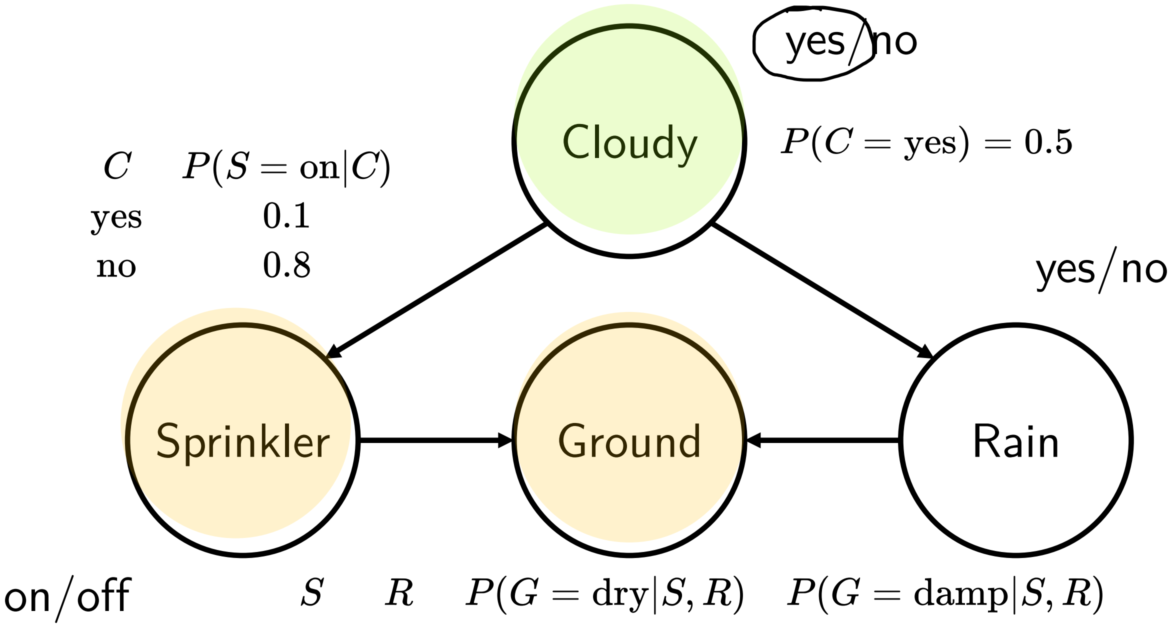
C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



C	$P(S = \text{on}\vert C)$
yes	0.1
no	0.8

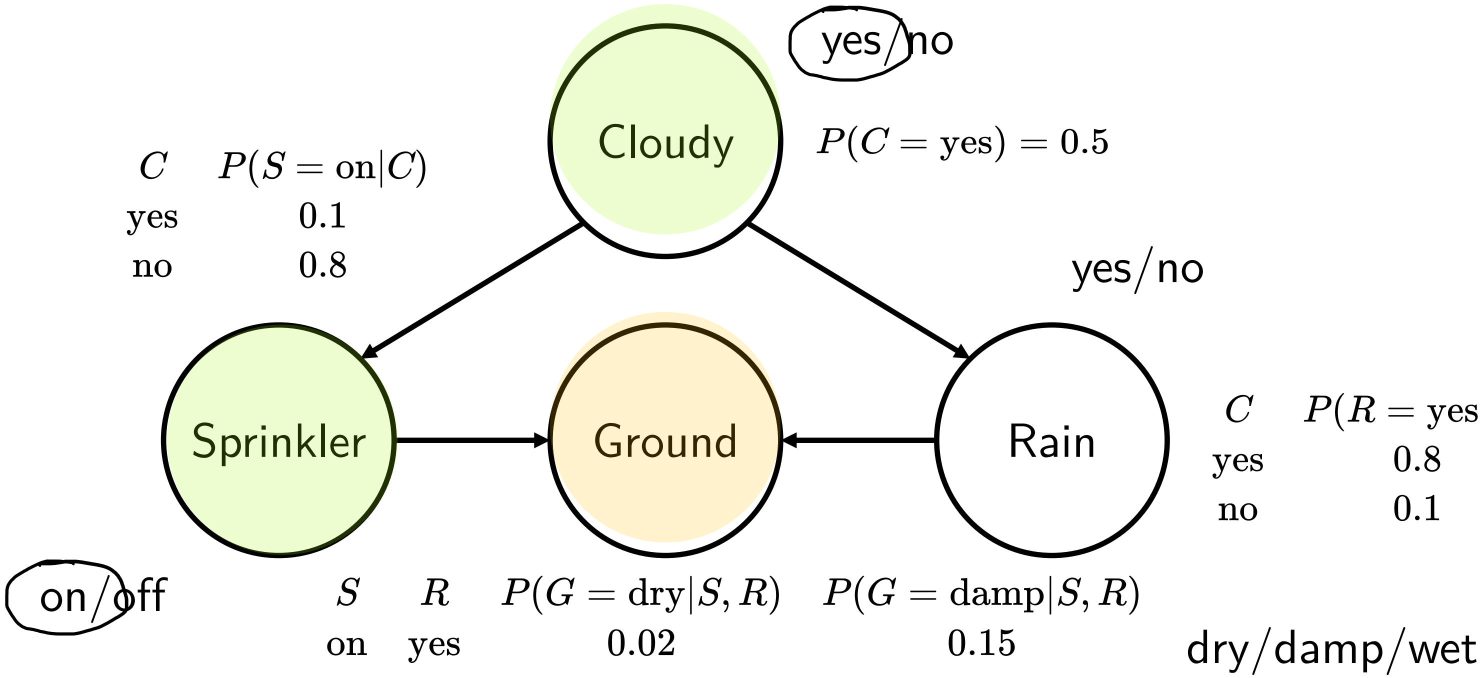
C	$P(R = \text{yes}\vert C)$
yes	0.8
no	0.1

S	R	$P(G = \text{dry}\vert S, R)$	$P(G = \text{damp}\vert S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

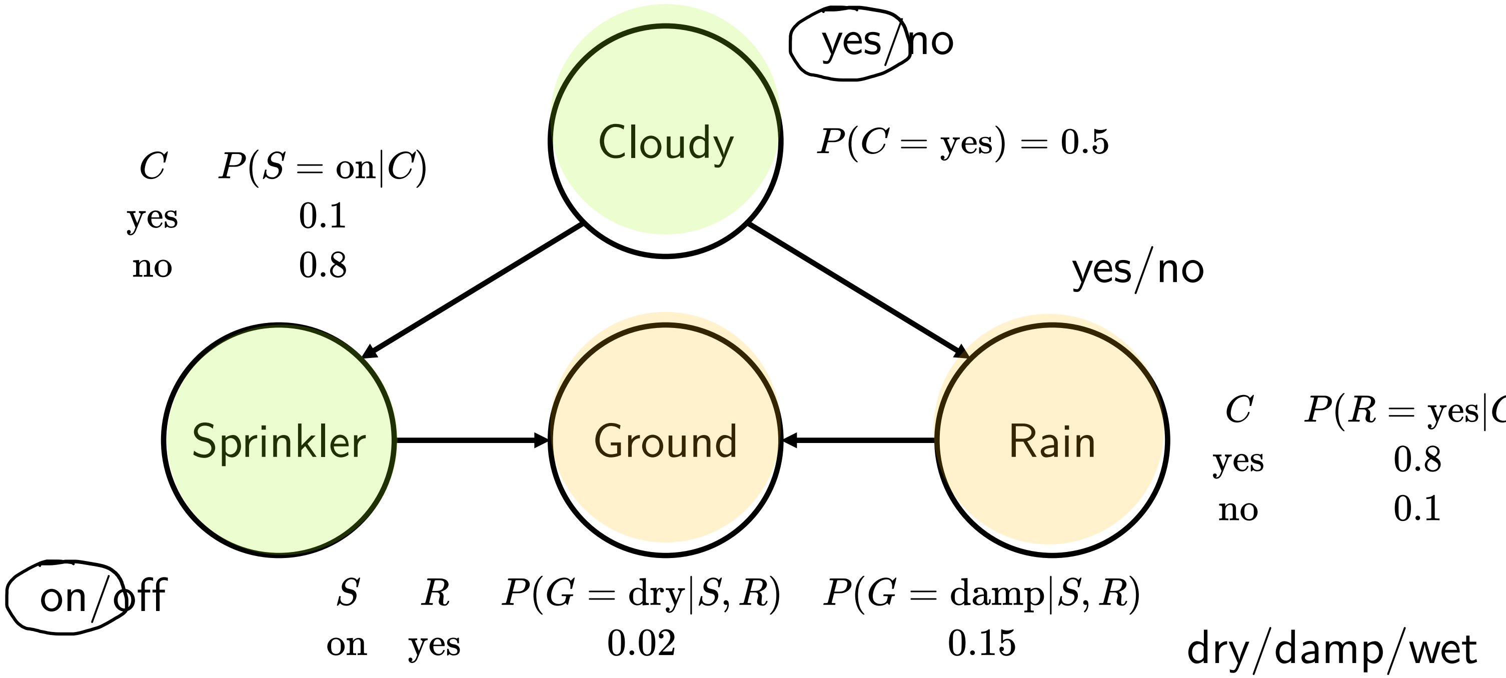
C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

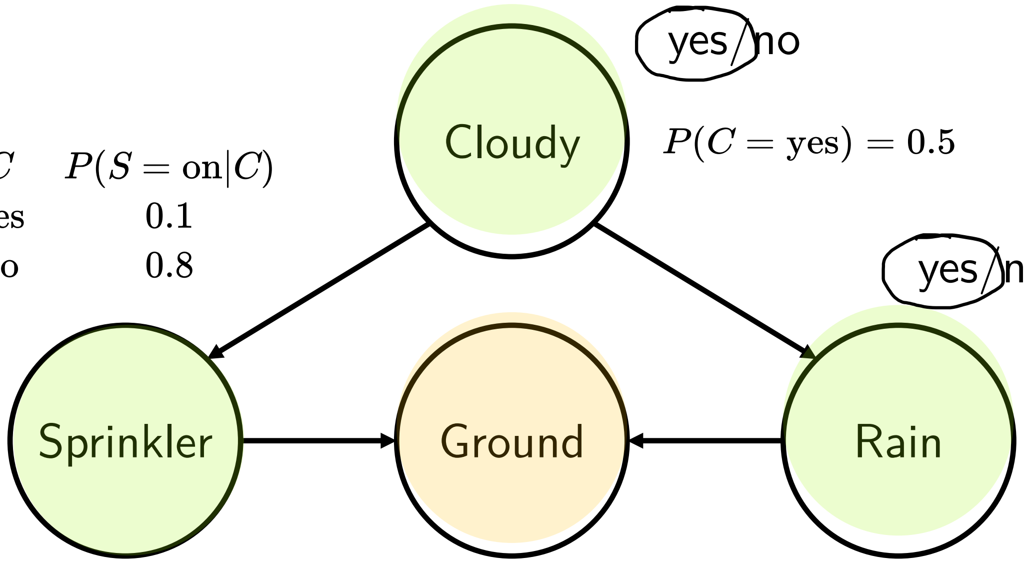
C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.



yes/no

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

yes/no

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

on/off

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

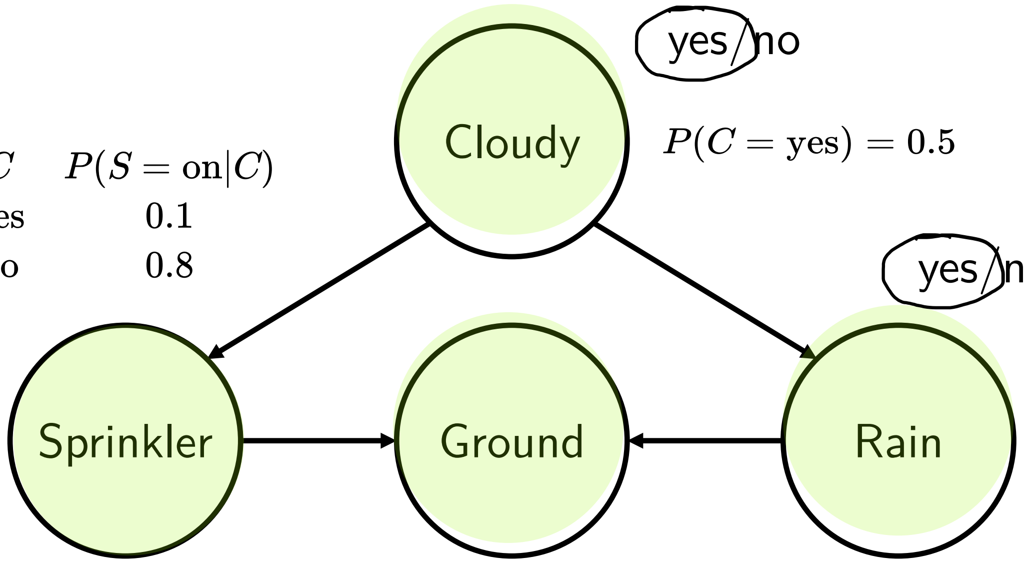
Inference via Sampling: Algorithm

Starting with the goal node, we recursively sample the parents' states and use the results to sample their child, i.e., the current node.

yes/no

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on}\|C)$
yes	0.1
no	0.8



yes/no

C	$P(R = \text{yes}\|C)$
yes	0.8
no	0.1

on/off

S	R	$P(G = \text{dry}\|S, R)$	$P(G = \text{damp}\|S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

We have created one sample, $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

To inference, we will need to have a LOT of samples.

Suppose we sample 10 times and obtain the following results.

To estimate $P(G = \text{wet} | R = \text{yes})$ we first consider the samples that are consistent with the evidence.

$(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

$(G = \text{damp}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{yes})$

$(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

To inference, we will need to have a LOT of samples.

Suppose we sample 10 times and obtain the following results.

To estimate $P(G = \text{wet} | R = \text{yes})$ we first consider the samples that are consistent with the evidence.

We then look at those samples which also match the query, $G = \text{wet}$

$(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

$(G = \text{damp}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{yes})$

$(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

To inference, we will need to have a LOT of samples.

Suppose we sample 10 times and obtain the following results.

To estimate $P(G = \text{wet} | R = \text{yes})$ we first consider the samples that are consistent with the evidence.

We then look at those samples which also match the query, $G = \text{wet}$

$(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

$(G = \text{damp}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{yes})$

$(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

Inference via Sampling: Algorithm

To inference, we will need to have a LOT of samples.

Suppose we sample 10 times and obtain the following results.

To estimate $P(G = \text{wet} | R = \text{yes})$ we first consider the samples that are consistent with the evidence

We then look at those samples which also match the query, $G = \text{wet}$

We estimate $P(G = \text{wet} | R = \text{yes})$ to be about 3 out of 5, or 60%.

$(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

$(G = \text{damp}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{damp}, S = \text{on}, R = \text{no}, C = \text{yes})$

$(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{on}, R = \text{no}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{dry}, S = \text{off}, R = \text{no}, C = \text{yes})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

$(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

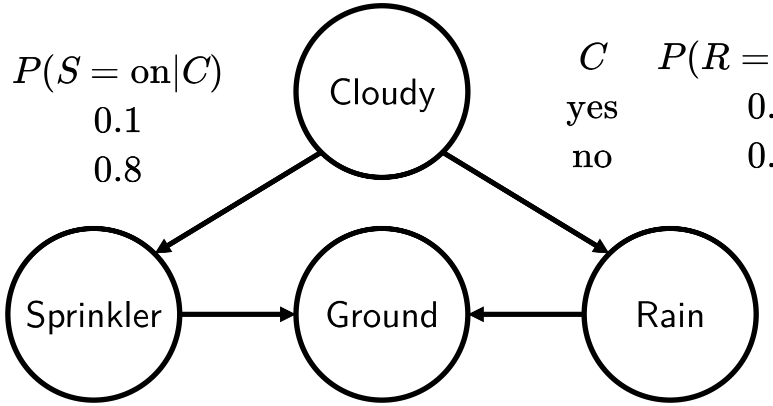
Inference via Sampling: Code

Find $P(G = \text{wet} | R = \text{yes})$

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

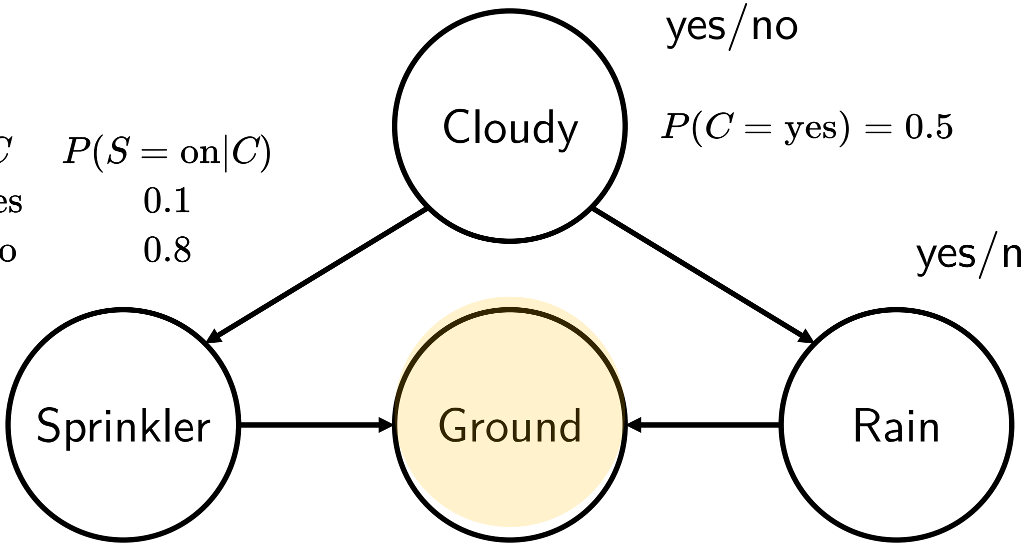


S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

Using sampling, we estimate $P(G = \text{wet} | R = \text{yes})$ to be approximately 19.68%.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

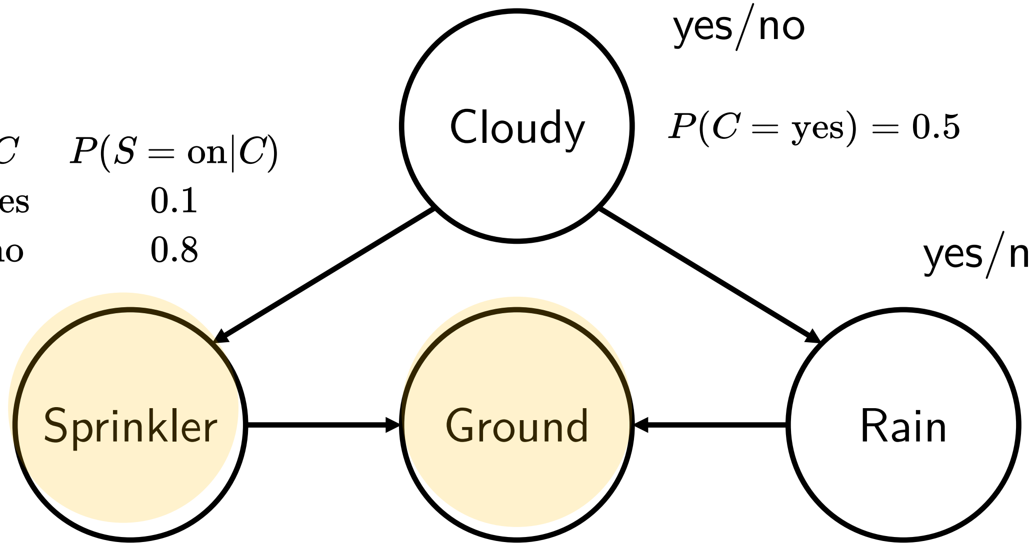
S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

$$W = 1$$

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.



Cloudy: yes/no

$P(C = \text{yes}) = 0.5$

Sprinkler: on/off

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

Rain: yes/no

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

Ground: dry/damp/wet

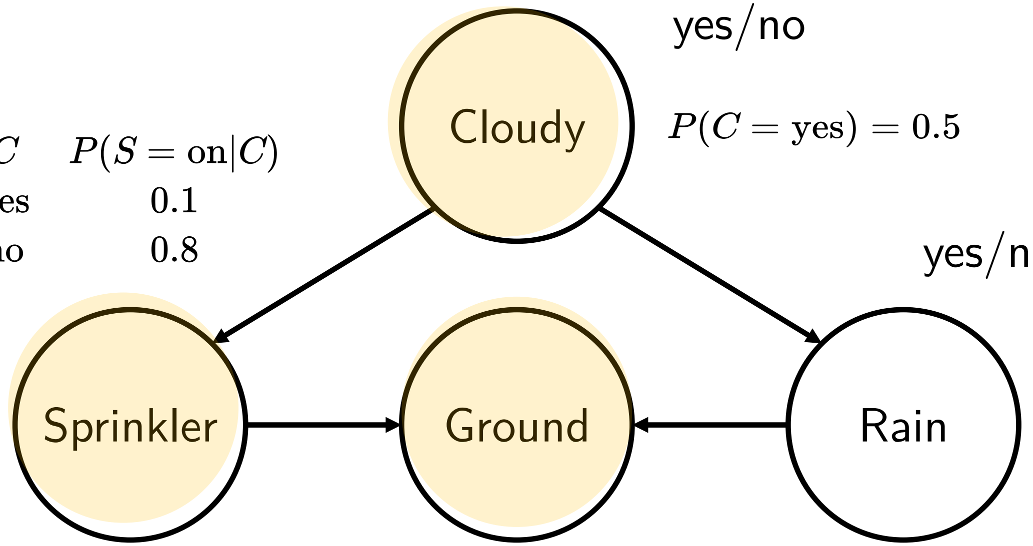
S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

$$W = 1$$

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.



yes/no

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

$$W = 1$$

yes/no

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

on/off

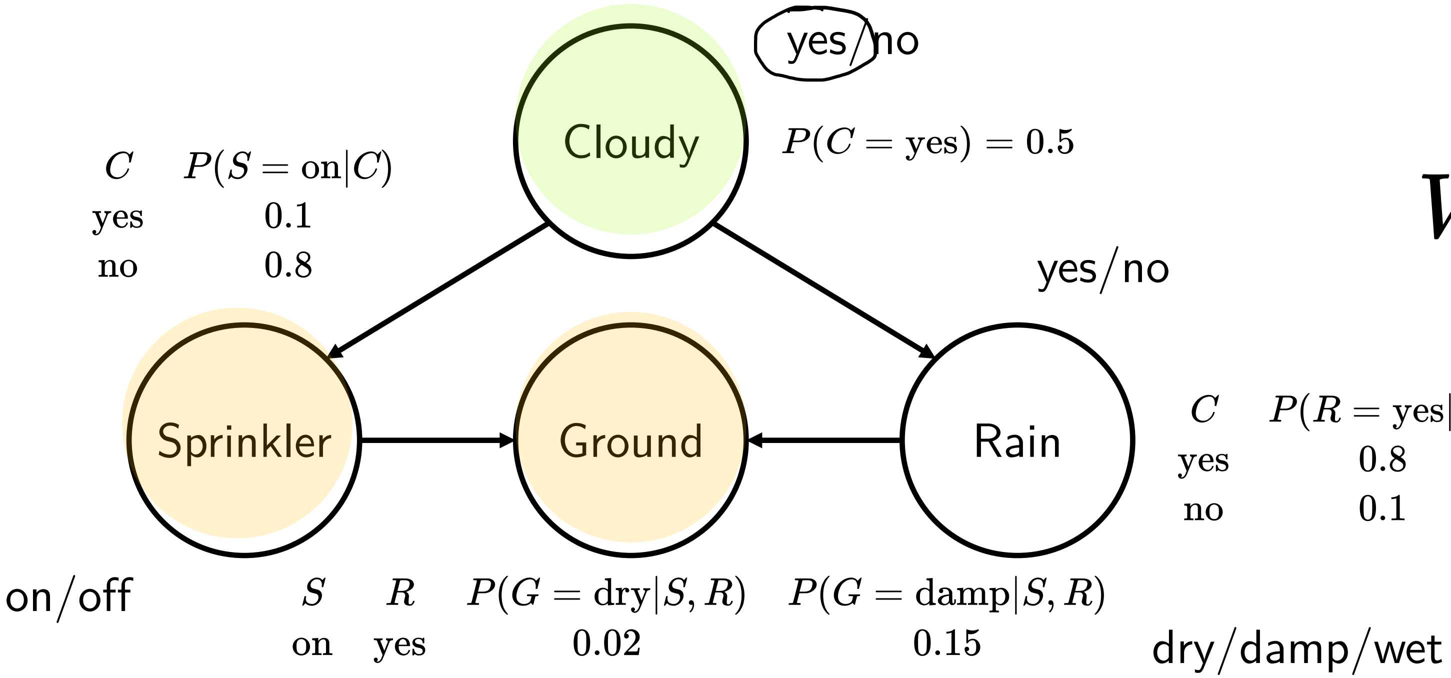
S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on}\|C)$
yes	0.1
no	0.8

C	$P(R = \text{yes}\|C)$
yes	0.8
no	0.1

S	R	$P(G = \text{dry}\|S, R)$	$P(G = \text{damp}\|S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

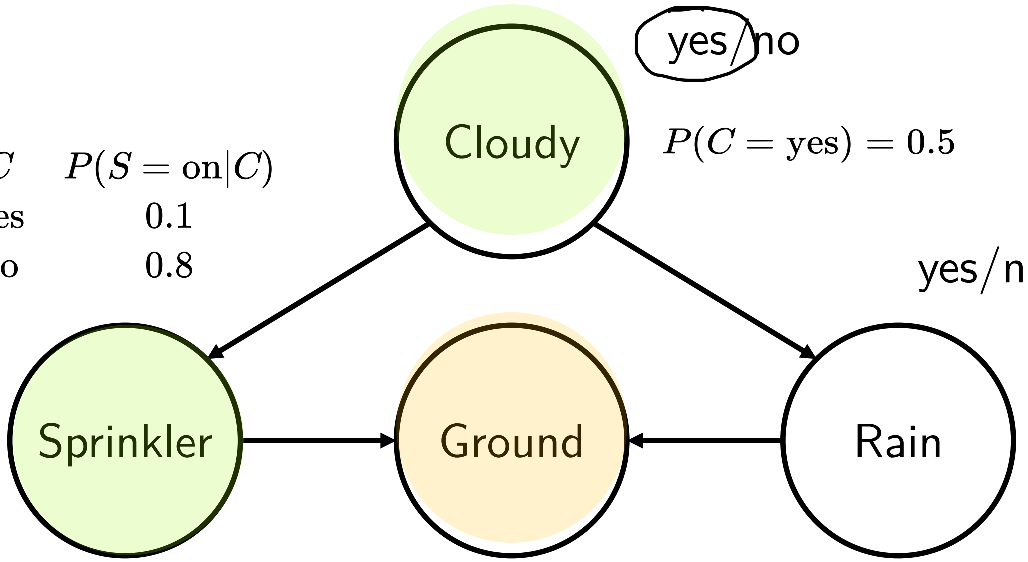
$$W = 1$$

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.

yes/no



$P(C = \text{yes}) = 0.5$

$$W = 1$$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

yes/no

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

on/off

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

Solution: Associate the sample with a weight, W, which fits the evidence.

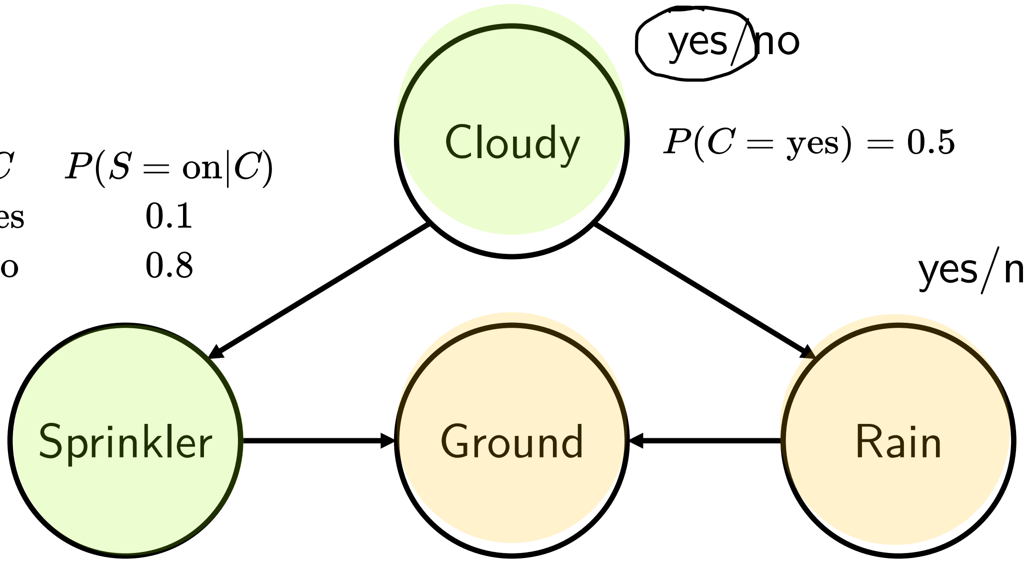
Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.

yes/no

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on}\vert C)$
yes	0.1
no	0.8



$$W = 1$$

yes/no

C	$P(R = \text{yes}\vert C)$
yes	0.8
no	0.1

on/off

S	R	$P(G = \text{dry}\vert S, R)$	$P(G = \text{damp}\vert S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.

yes/no

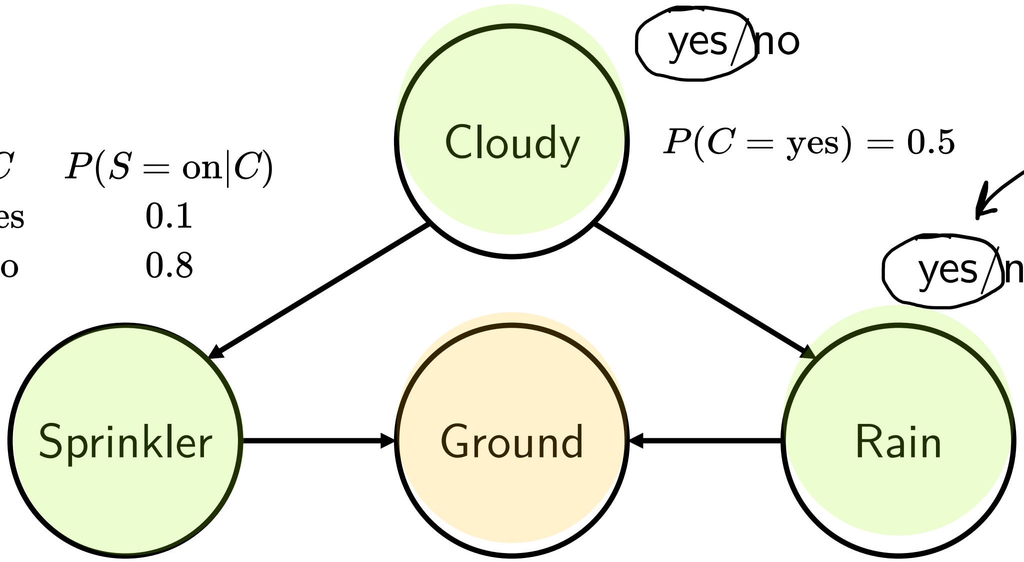
$P(C = \text{yes}) = 0.5$

force this

yes/no

$$W = 0.8$$

C	$P(S = \text{on}\mid C)$
yes	0.1
no	0.8



C	$P(R = \text{yes}\mid C)$
yes	0.8
no	0.1

on/off

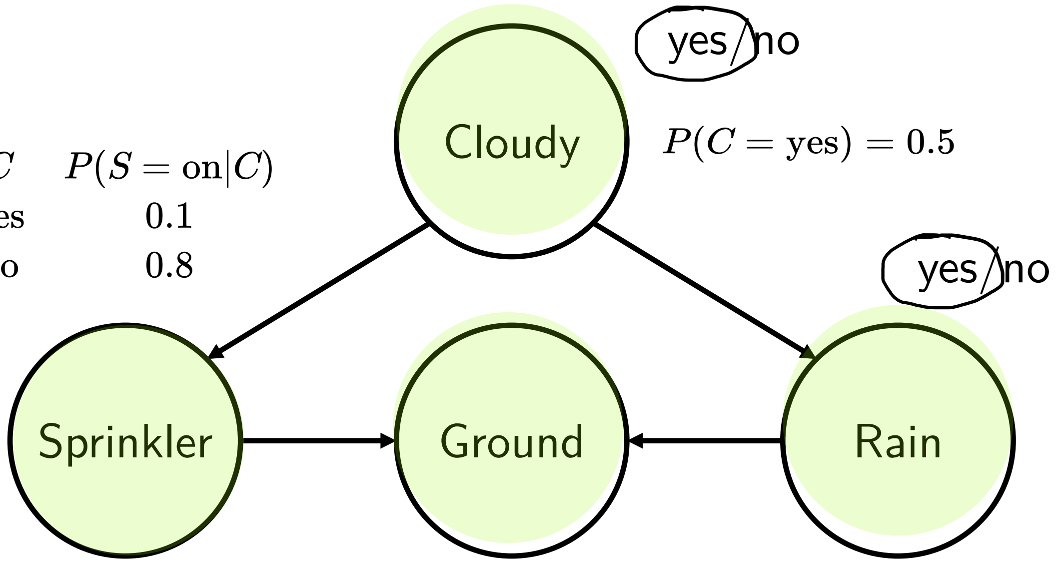
S	R	$P(G = \text{dry}\mid S, R)$	$P(G = \text{damp}\mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (w/ LW): Algorithm

If only a few samples that satisfy the evidence, then our estimate might be poor and if we force the evidence, our sampling will be biased.



$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \mid C)$
yes	0.1
no	0.8

C	$P(R = \text{yes} \mid C)$
yes	0.8
no	0.1

$$W = 0.8$$

on/off

S	R	$P(G = \text{dry} \mid S, R)$	$P(G = \text{damp} \mid S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

dry/damp/wet

Solution: Associate the sample with a weight, W, which fits the evidence.

Inference via Sampling (with likelihood weighting)

To inference, we will need to have a LOT of samples.

Suppose we sample 10 times and obtain the following results.

To estimate $P(G = \text{wet}|R = \text{yes})$ we simply take total weight of the samples which match the query and normalize by the total weight of all samples.

0.8 $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

0.2 $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

0.2 $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

0.2 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

0.8 $(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

0.2 $(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

0.8 $(G = \text{dry}, S = \text{off}, R = \text{yes}, C = \text{no})$

0.2 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

0.8 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

0.8 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

Inference via Sampling (with likelihood weighting)

To inference, we will need to have a LOT of samples.

Suppose we sample 10 times and obtain the following results.

To estimate $P(G = \text{wet}|R = \text{yes})$ we simply take total weight of the samples which match the query and normalize by the total weight of all samples.

We estimate $P(G = \text{wet}|R = \text{yes})$ to be about 3.2 over 5, or 64%.

0.8 $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

0.2 $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

0.2 $(G = \text{wet}, S = \text{on}, R = \text{yes}, C = \text{yes})$

0.2 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

0.8 $(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

0.2 $(G = \text{dry}, S = \text{on}, R = \text{yes}, C = \text{no})$

0.8 $(G = \text{dry}, S = \text{off}, R = \text{yes}, C = \text{no})$

0.2 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

0.8 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{no})$

0.8 $(G = \text{wet}, S = \text{off}, R = \text{yes}, C = \text{yes})$

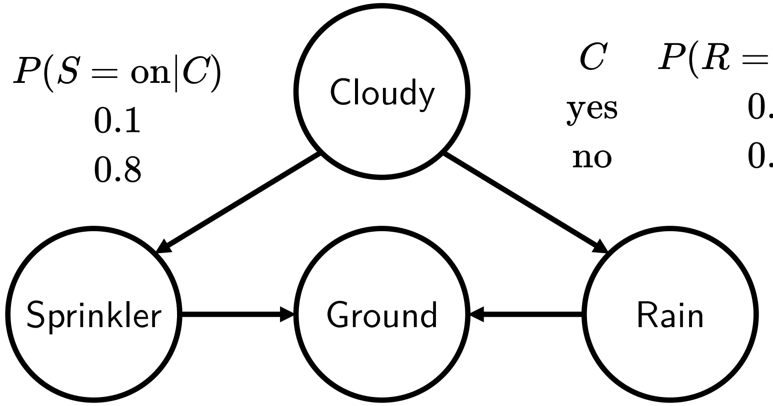
Inference via Sampling (w/ LW): Code

Find $P(G = \text{wet} | R = \text{yes})$

$P(C = \text{yes}) = 0.5$

C	$P(S = \text{on} \vert C)$
yes	0.1
no	0.8

C	$P(R = \text{yes} \vert C)$
yes	0.8
no	0.1



S	R	$P(G = \text{dry} \vert S, R)$	$P(G = \text{damp} \vert S, R)$
on	yes	0.02	0.15
on	no	0.05	0.85
off	yes	0.04	0.90
off	no	0.95	0.04

(try coding this yourself)

The source code is available here.

(the implementation is optimized for readability)

A copy of these slides is available here.