INFERENCE IN BAYESIAN NETWORKS

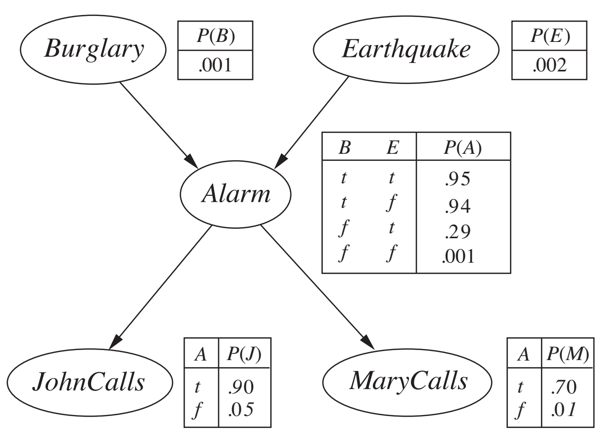
Today

- Exact inference in BN
- Approximate inference in BN

Inference in Bayesian Networks

- Probabilistic inference refers to the task of computing some desired probability given other known probabilities (evidence)
- Exact Inference
  - Enumeration
- Approximate Inference
  - Direct sampling
  - Rejection sampling
  - Likelihood weighting
  - Gibbs Sampling

Inference by Enumeration



Practice Queries:

- p(J | e)
- p(A | j, m)
- p(B | e, j)

Inference by Enumeration

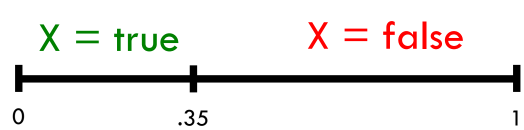
- Space complexity
  - Largest table is exponential in number of parents
- Time complexity
  - Exponential in the number of hidden variables
  - Worst case $O(d^n)$
- Variable elimination
  - Dynamic programming
  - Requires an ordering of variables for summing
  - Exponent depends on efficiency of ordering

Approximate Inference

- Analogous to uninformed/informed search algorithms that use an incremental formulation
  - Direct sampling
  - Rejection sampling
  - Likelihood weighting
- Analogous to local search algorithms that use a complete-state formulation and make local modifications
  - Gibbs sampling

Approximate Inference

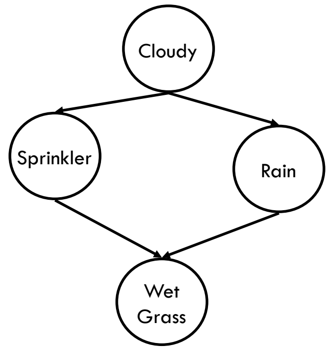
Sampling from a discrete distribution



Generate a number uniformly distributed between 0 and 1

Direct Sampling (no evidence)

Samples



P(C = true)
.50

C	P(S=true|C)
T	.10
F	.50

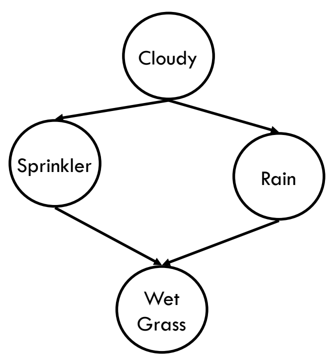
C	P(R=true|C)
T	.80
F	.20

S	R	P(W=true|S,R)
T	T	.99
T	F	.90
F	T	.90
F	F	.01

Rejection Sampling (evidence)

Query: p(R | S=true)

Samples



P(C = true)
.50

C	P(S=true|C)
T	.10
F	.50

C	P(R=true|C)
T	.80
F	.20

S	R	P(W=true|S,R)
T	T	.99
T	F	.90
F	T	.90
F	F	.01

Likelihood Weighting

- Fixes the values for the evidence so there are no wasted samples
- Sample only the non-evidence variables

Likelihood Weighting

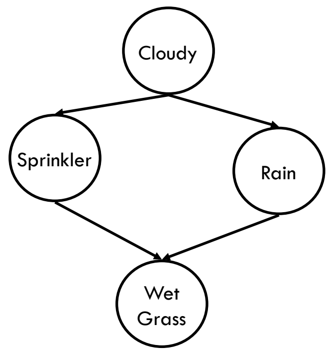
- Not every sample is created equal!
- Weight each sample by how likely the evidence is given the sampled values

weight = $p(e_1 \mid Parents(e_1)) * p(e_2 \mid Parents(e_2)) \ldots$

Likelihood Weighting

Query: p(C | S=true)

Samples



P(C = true)
.50

C	P(S=true|C)
T	.10
F	.50

C	P(R=true|C)
T	.80
F	.20

S	R	P(W=true|S,R)
T	T	.99
T	F	.90
F	T	.90
F	F	.01

Gibbs Sampling

- Analogous to a local search algorithm where we make local modifications to our current state
  - Initial state = random assignment of non-evidence variables
  - States = complete assignment of values to variables
  - Transition = sample a new value for each variable in turn

Gibbs Sampling

- Each step to a new state is recorded as a sample
- In the limit, the probability of being in a state is proportional to that state's posterior probability

Gibbs Sampling

Summary

- Bayesian Networks (graph + prob. distributions)
- Independence using graph
  - D-separation algorithm
- Inference
  - Exact
  - Approximate