Sampling Based Inference

Il-Chul Moon
Dept. of Industrial and Systems Engineering
KAIST
icmoon@kaist.ac.kr

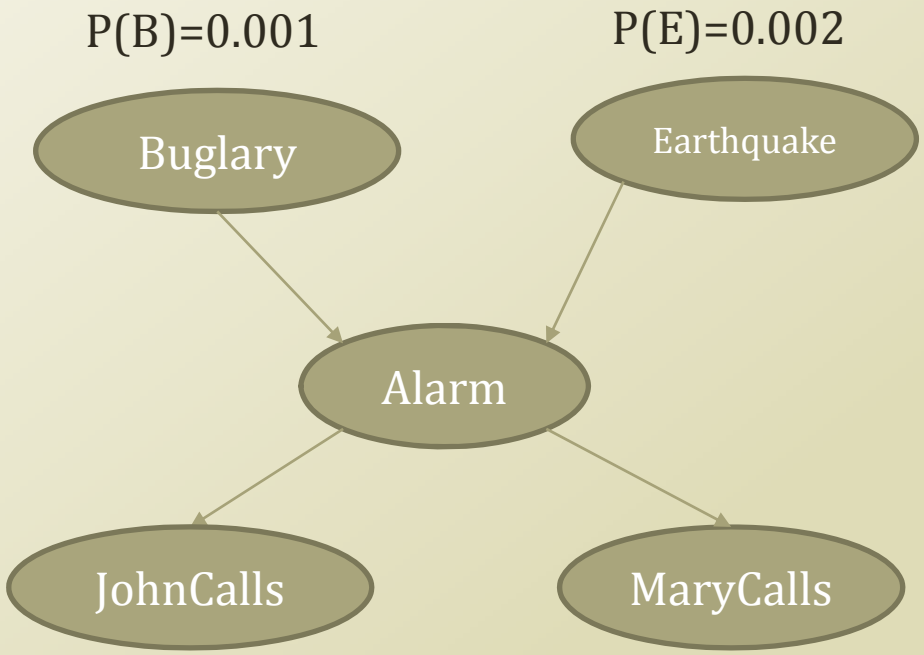
Weekly Objectives

- Learn basic sampling methods
  - Understand the concept of Markov chain Monte Carlo
  - Able to apply MCMC to the parameter inference of Bayesian networks
  - Know the mechanism of rejection sampling
  - Know the mechanism of importance sampling
- Learn sampling based inference
  - Understand the concept of Metropolis-Hastings algorithm
  - Know the mechanism of Gibbs sampling
- Know a case study of sampling based inference
  - Understand the latent Dirichlet allocation model
  - Know the collapsed Gibbs sampling
  - Know how to derive Gibbs sampling formula for LDA

BASIC SAMPLING METHODS

Forward Sampling

- Generate a sample from the Bayesian network
  - Follow topological order
    - Buglary → false
    - Earthquake → false
    - Alarm|B=F,E=F→true
    - JC|A=T→true
    - MC|A=T→false
  - Create such sample many, many, many times
- Then, count the samples match the case
  - P(E=T|MC=T)=?
    - Count the cases of E=T and MC=T
    - Count the cases of MC=T
- Any problem?



B	E	P(A|B,E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

A	P(J|A)
T	0.90
F	0.05

A	P(M|A)
T	0.70
F	0.01

Forward Sampling in GMM

$$P(x) = \sum_{k=1}^{K} P(z_k)P(x|z)$$
$$= \sum_{k=1}^{K} \pi_k N(x|\mu_k, \Sigma_k)$$

- Forward sampling of GMM
  - Sample z from π
    - z is the indicator of the mixture distribution
  - With selected z, sample x from $\mathrm{N}(\mu_z, \Sigma_z)$
- After many, many sampling, you can draw the histogram of the mixture distribution
- You have an empirical PDF, so you can ask a query like $P(0 \leq x \leq 5|\pi, \mu, \Sigma)$