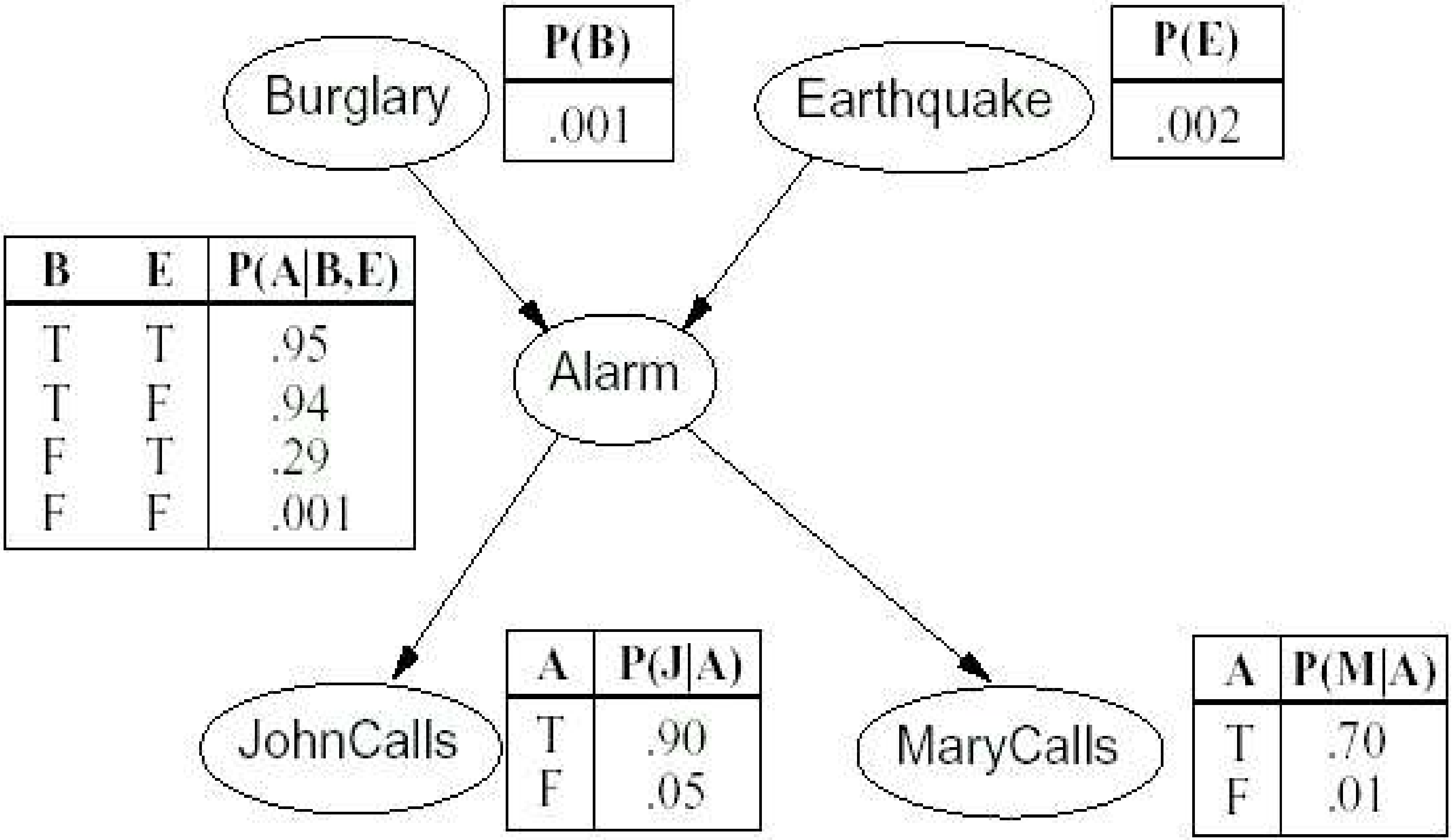
CSC421 Intro to Artificial Intelligence

UNIT 23:
Probabilistic Reasoning

Midterm Review

- Edit distance

Example



P(B)
.001

P(E)
.002

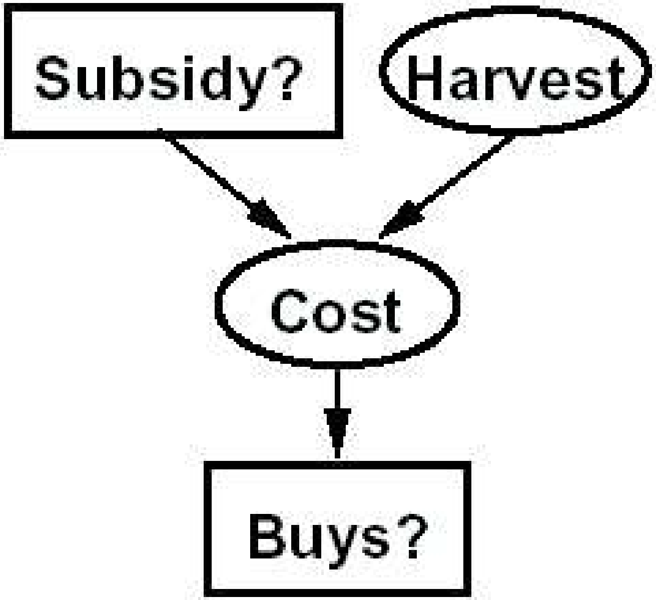
B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01

Hybrid (discrete + continuous) networks

Discrete (Subsidy?, Buys?)
Continuous (Harvest, Cost)



Option 1: Discretization - possibly large erros, large CPTs
Option 2: Finitely parametrized canonical families

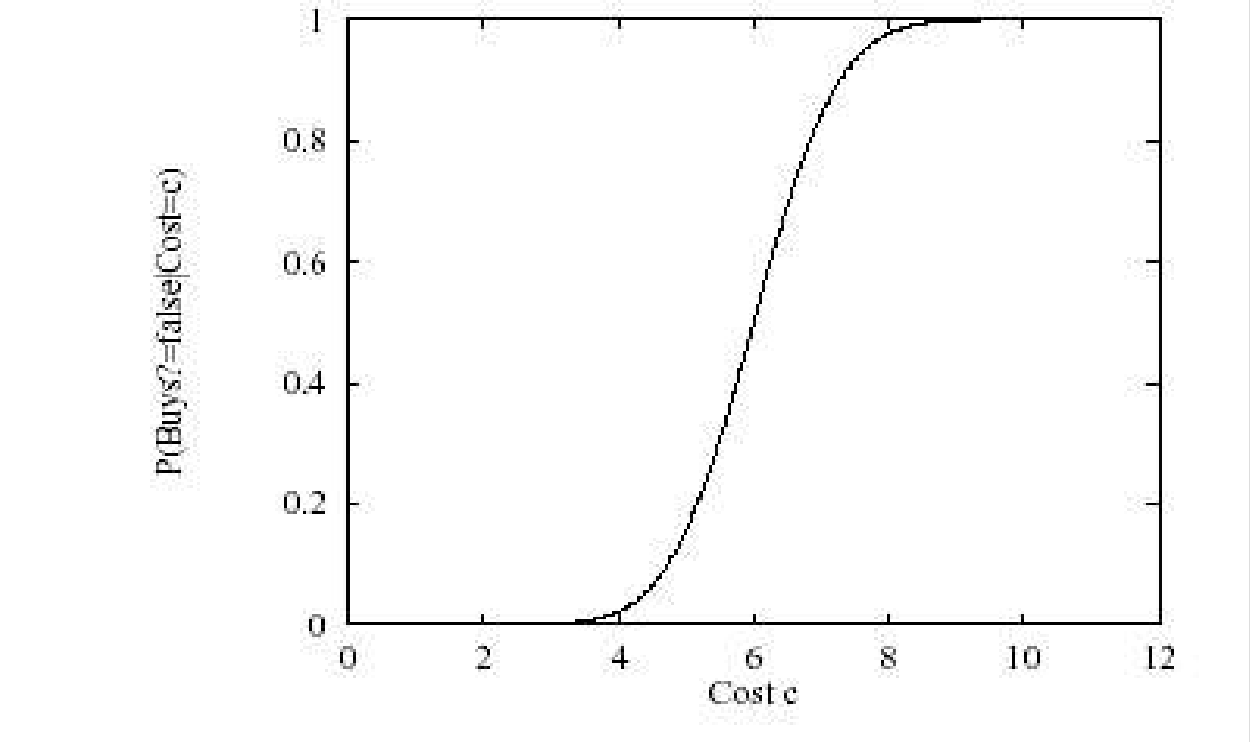
1) Continuous variable (discrete & continuous parents) (e.g. Cost)
2) Discrete variable (continuous parents) (e.g Buys?)

Continuous Child Variables

- Need one conditional density function for child variable given continuous parents for each possible assignment to discrete parents
- Most common linear Gaussian model:
  - P(Cost = c | Harvest =h, Subsidy = true) = $N(a_t h + b_t, \sigma_t)(c)$
- All continuous network with LGs then joint distribution is multivariate Gaussian
- Discrete + continuous = conditional Gaussian (multivariate Gaussian for every setting of all discrete variables)

Discrete Variable with continuous parents

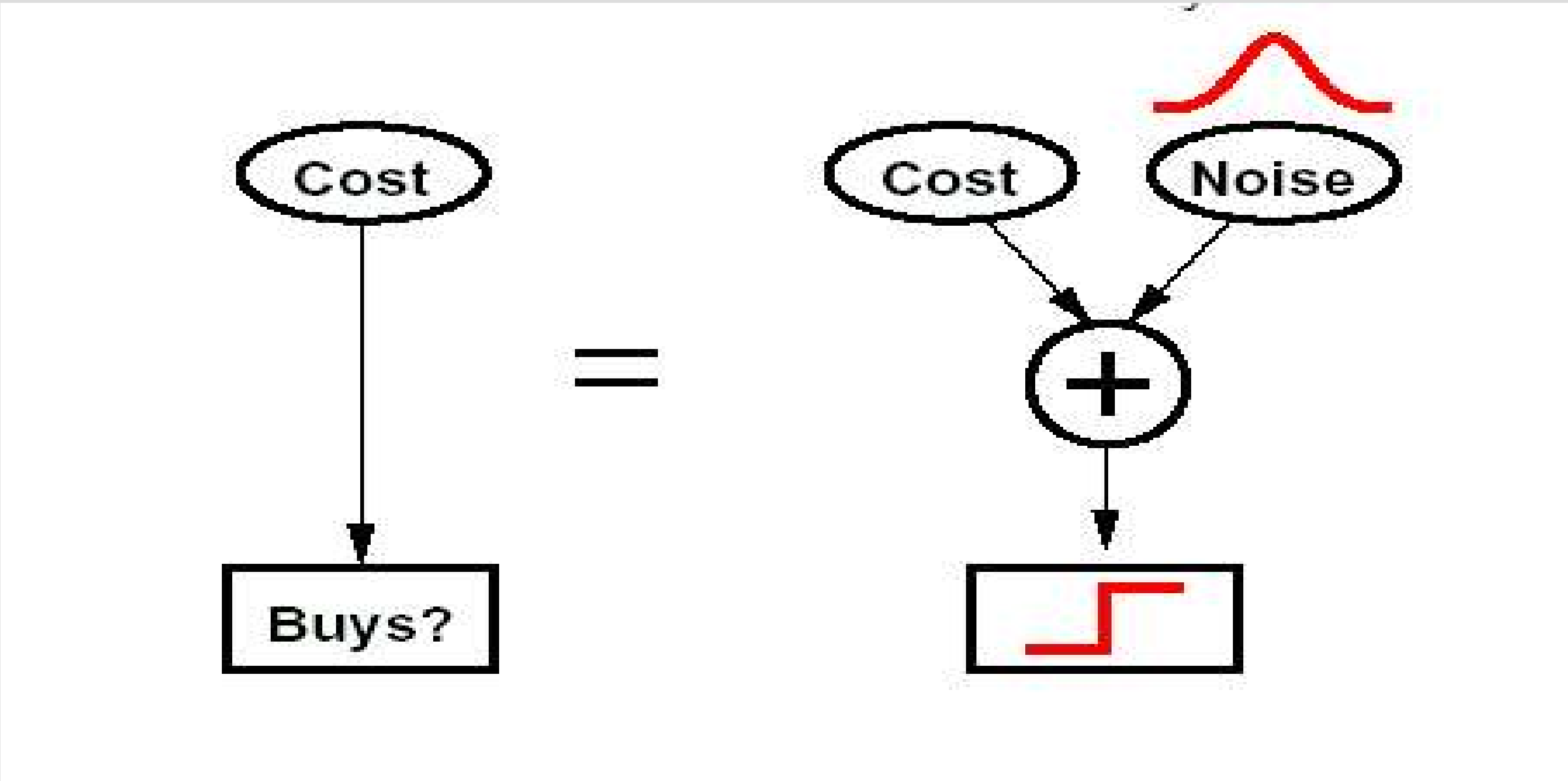
- Probability of Buys ? given Cost should be a soft threshold



Probit distribution = integral of Gaussian
Sigmoid (or Logit) distribution = also used in Neural Networks

Why probit ?

- It's sort of the right shape
- Can be viewed as a hard threshold whose location is subject to noise



Summary

- Bayesian networks provide a natural representation for (causally induced) conditional independence
- Topology + CPT = compact reprsentation of joint distribution
- Generally easy for (non) experts to construct
- Continuous variables => parametrized distributions (e.g linear Gaussian)
- Extra:
  - Canonical distributions (noisy-OR)

Inference in BN

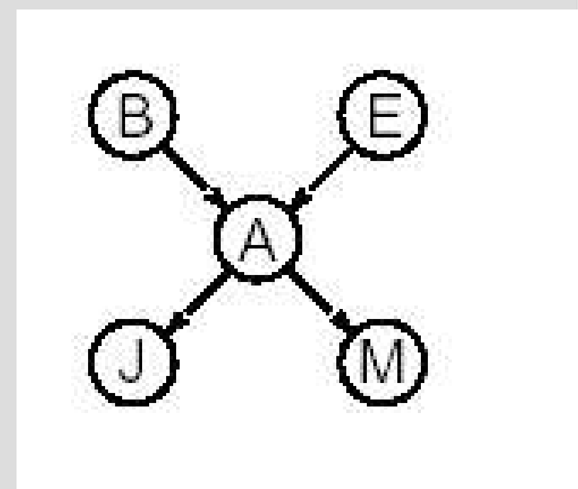
- Exact Inference by enumeration
- Exact Inference by variable elimination
- Approximate inference by stochastic simulation
- Approximate infrence by Markov Chain Monte Carlo

Inference Tasks

- Simple queries
  - P(NoGas | Gauge = empty, Lights = on, Starts =false)
- Conjuctive queries
  - $P(X_i, X_j \mid E = e) = P(X_i \mid E = e)\ P(X_j \mid X_i, E = e)$
- Query variables
- Event (assignment of values to set of evidence variables)
- Non-evidence variables (hidden variables)

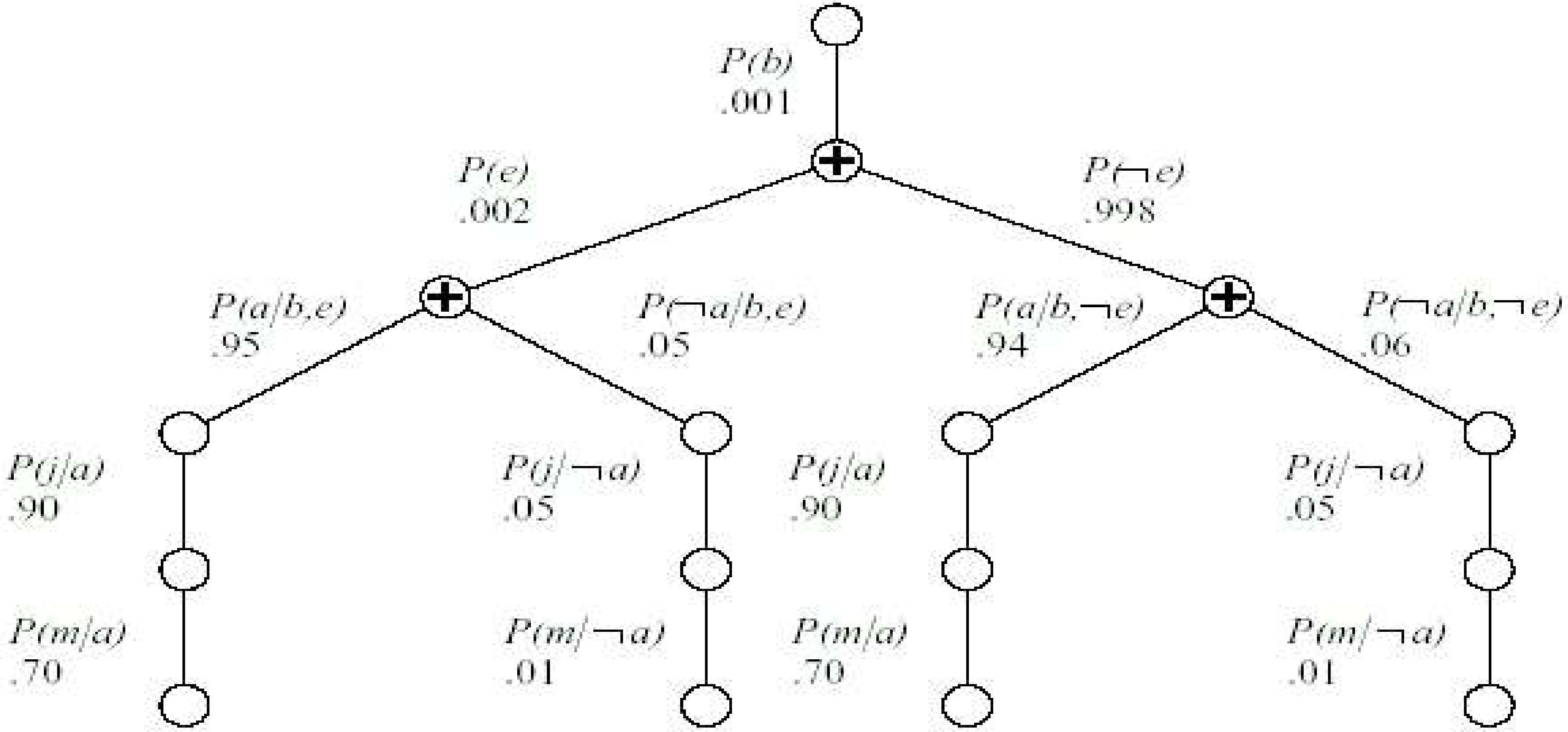
Inference by enumeration

- Sum out variables from the joint without actually constructing it's explicit representation



- Simple query:
  - $P(B \mid j, m) = P(B, j, m) \,/\, P(j,m) =$
    $\alpha\, P(B, j, m) = \alpha\, \Sigma_e \Sigma_a P(B, e, a, j, m)$
- Rewrite full joint entries using product of CPT entries
  - $P(B|j,m) = \alpha\, \Sigma_e \Sigma_a P(B)\, P(e)\, P(a|B, e)\, P(j|a)\, P(m|a) =$
    $\alpha\, P(B)\, \Sigma_e P(e)\, \Sigma_a P(a|B, e)\, P(j|a)\, P(m|a)$
- Recursive depth-first search enumeration

Evaluation Tree



Enumeration is inefficient: repeated computation
e.g. Computer P(j|a) P(m|a) twice for each value of e

Inference by variable elimination

- Variable elimination: carry out summations right-to-left, storing intermediate results (factors) to avoid recomputation