NARSIS Workshop

Training on Probabilistic Safety Assessment for Nuclear Facilities

September 2-5, 2019, Warsaw, Poland

Bayesian Networks – Practical Session

Varenya Kumar D Mohan,

Delft University of Technology

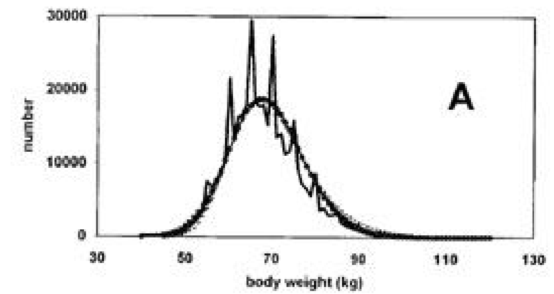
Session Objectives

- Refresh Bayesian network (BN) building steps with an example
- Practical
  - Represent a problem with an uncertain outcome, in a BN
  - Use BN software to calculate the probabilities of different outcomes of interest
  - Understand how new evidence changes our current probability estimates

Probability Theory - Basics

- **Useful definitions:**

- *Random Variables*: A variable whose possible values are outcomes of a random phenomenon.
  - ⇒ Discrete RV: takes distinct, separate values – e.g. outcomes of the roll of dice; X = {1, 2, 3, 4, 5, 6}
  - ⇒ Continuous RV: can take any value in a given range – e.g. weight of a randomly selected person in Warsaw (not real data)



- *Probability Distribution*: Table, equation or graph that links each outcome with its probability of occurrence
  - ⇒ e.g. X is a RV denoting outcome of a coin toss

Value of X	Probability of Occurrence
Heads	0.5
Tails	0.5

- *Marginal Probability*: P(A), is the probability of a variable without reference to the values of the other variables in the problem domain.

- *Conditional Probability*: P(A | B), is the probability that A occurs given that B has occurred. It follows that:

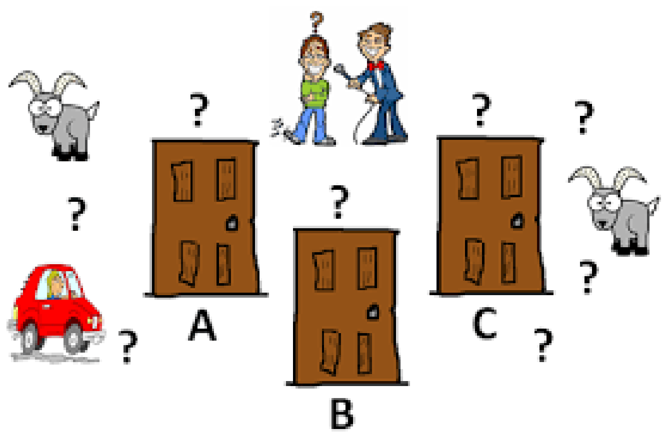
$$P(A \cap B) \text{ or } P(AB) = P(BA) = P(A).P(B \mid A) = P(B).P(A \mid B)$$

Bayesian Networks

- **A Bayesian network is a 'directed' (the dependence between variables is given by arrows) and acyclic (the arrows don't cycle back) probabilistic graphical model**
  - **Random variables represented by nodes**
  - **Dependencies given by arcs**

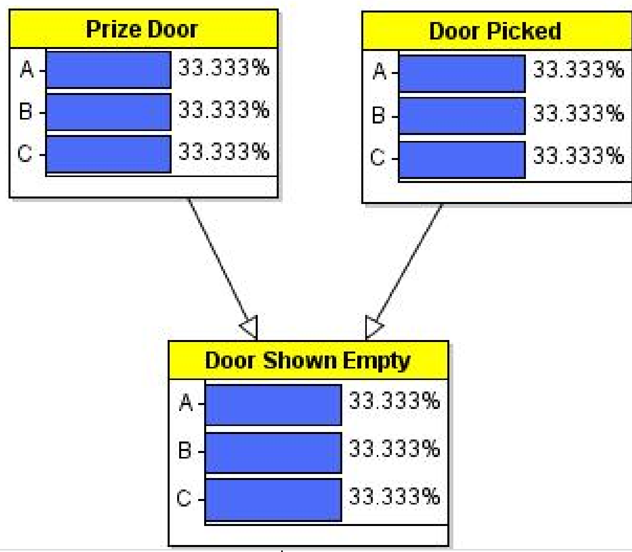
BN Example – The Monty Hall Problem

- **Example of BN – The Monty-Hall Problem**
  - You're on a game show. You're given a choice between three closed doors.
  - There is a car behind one door, and nothing behind the other two.
  - You want the car (usually), and choose one door, say Door A
  - Monty, the host, knows what's behind each door. So he opens another door, say Door B which of course, is empty.
  - Now you are asked to either stick with Door A or switch to Door C. What will you do?



BN Example – The Monty Hall Problem

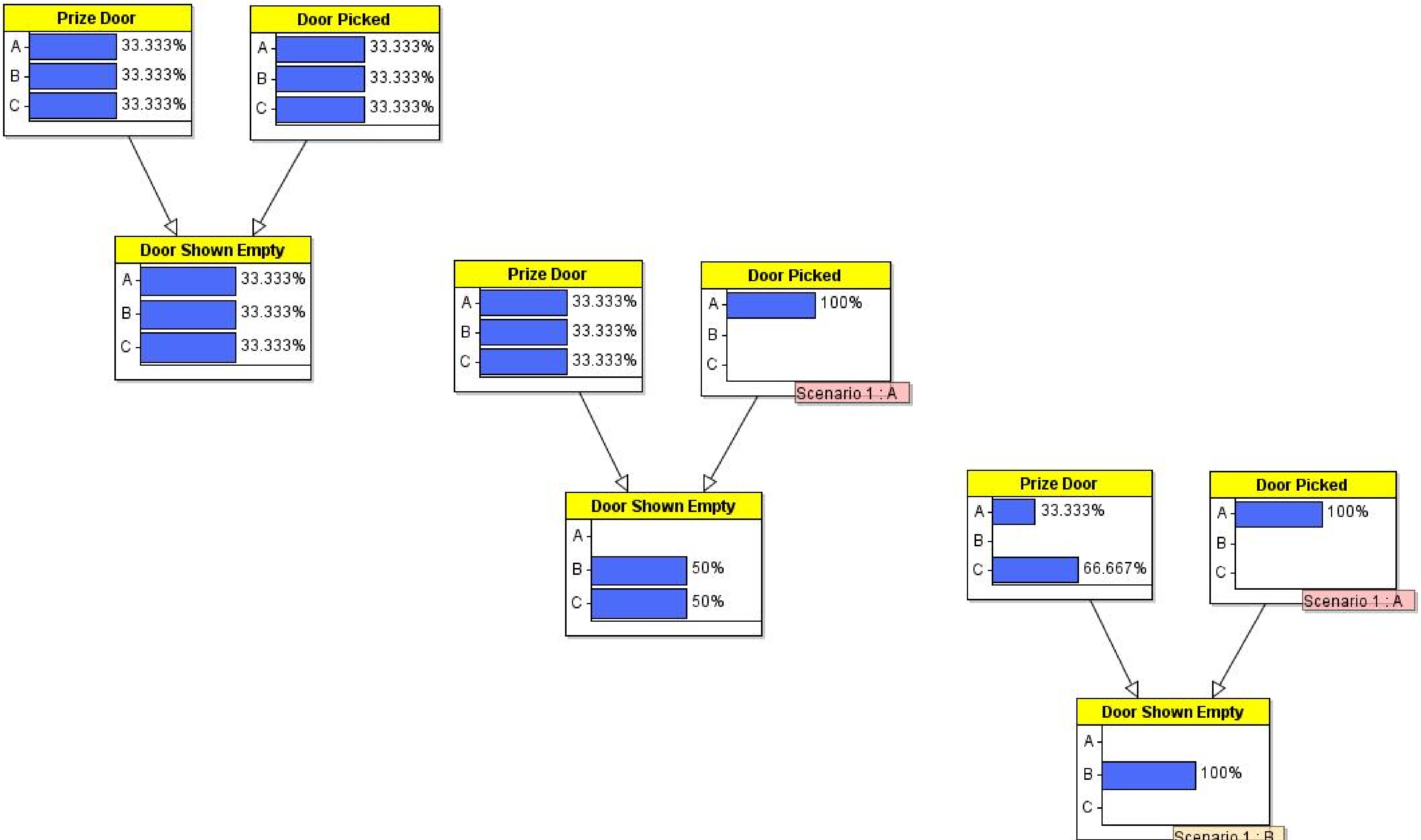
PRIOR DISTRIBUTIONS



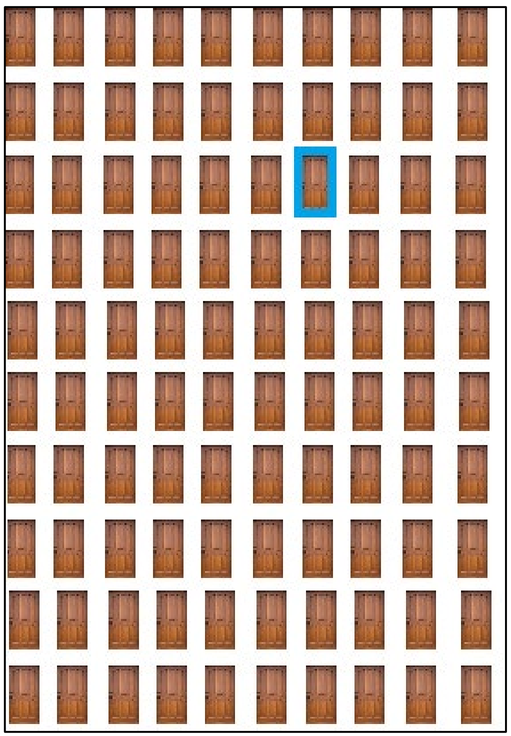
Conditional Probability Table/Distribution (CPD) for "Door Shown Empty"

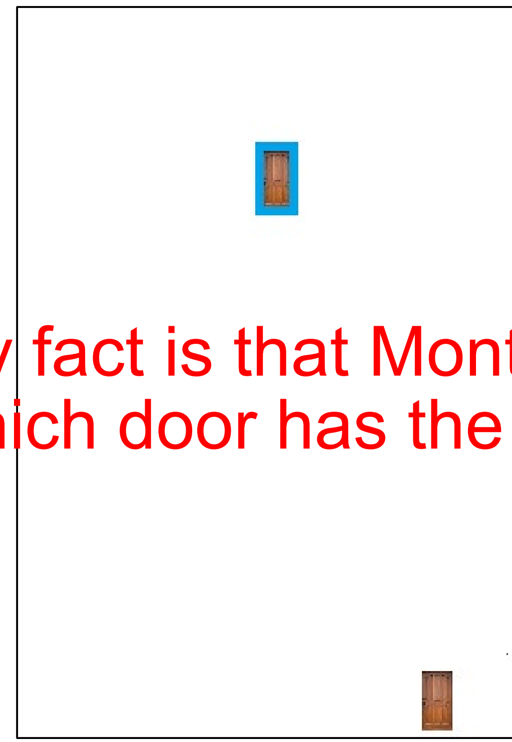
Prize Door	**Door A**			**Door B**			**Door C**		
Door Picked	*Door A*	*Door B*	*Door C*	*Door A*	*Door B*	*Door C*	*Door A*	*Door B*	*Door C*
Door A	0.0	0	0	0	0.5	1	0	1	0.5
Door B	0.5	0	1	0	0	0	1	0	0.5
Door C	0.5	1	0	1	0.5	0	0	0	0

BN Example – The Monty Hall Problem



BN Example – The Monty Hall Problem





The key fact is that Monty knows which door has the car

Practical - Problem

- EDG operating condition, Markov model
- Damage States D0 (new) to D6 (failure)
- For every passing year:
  - No change in damage state – 40% of the time
  - Deterioration by one damage state – 30%
  - Deterioration by two damage states – 20%
  - Deterioration by three damage states – 10%

- Model the damage progression of the EDG using a Bayesian network starting from when the EDG is new (Year0) to the end of a 5-year period (Year5)
- Q1. What is the accumulated probability of failure after 5 years?
- Q2. At the end of three years, a deterministic inspection is performed that definitively determines the condition of the EDG to be in D2. Given this information, the accumulated probability of failure 5 years?

Practical - Hints

- EDG operating condition, Markov model
- Damage States D0 (new) to D6 (failure)
- For every passing year:
  - No change in damage state – 40% of the time
  - Deterioration by one damage state – 30%
  - Deterioration by two damage states – 20%
  - Deterioration by three damage states – 10%
- Model the damage progression of the EDG using a Bayesian network starting from when the EDG is new (Year0) to the end of a 5-year period (Year5)
- Q1. What is the accumulated probability of failure after 5 years?
- Q2. At the end of three years, a deterministic inspection is performed that definitively determines the condition of the EDG to be in D2. Given this information, the accumulated probability of failure 5 years?

Practical - Hints

Hints:

- Formulate random variables using ‘Years’ and ‘Damage States’, and the given probability values
- The discrete states of the variables should be able to directly answer Q1 and Q2.
- The BN will have 6 discrete random variables.

Practical – Random Variables

Which if these would be a good RV to use in the BN?

(Opt.1) DS - damage state of the EDG within the five year period

(Opt. 2) Yn – damage state of the EDG at the end of the nth year.

(Opt. 3) F5 – failure probability at the end of year 4

Practical – Random Variables

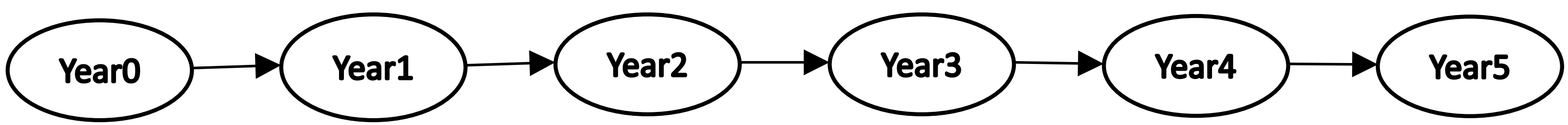
- **Year0 is the damage state of the EDG at time 0 (EDG is new)**
- **Year1 is the damage state at the end of the first year**
- **Hence, RVs in the BN are Year0, Year1…Year5**

Random Variable	States
Year0	D0
Year1	D0, D1, D2, D3
Year2	D0, D1, D2, D3, D4, D5, D6
Year3	D0, D1, D2, D3, D4, D5, D6
Year4	D0, D1, D2, D3, D4, D5, D6
Year5	D0, D1, D2, D3, D4, D5, D6

Practical – BN Structure

- What is the dependence between the random variables Year0, Year1…Year5?

- Markov chain



- If there is a circular dependence in the problem, BNs cannot be used (acyclic graphical model)

NETICA files

Practical - Problem

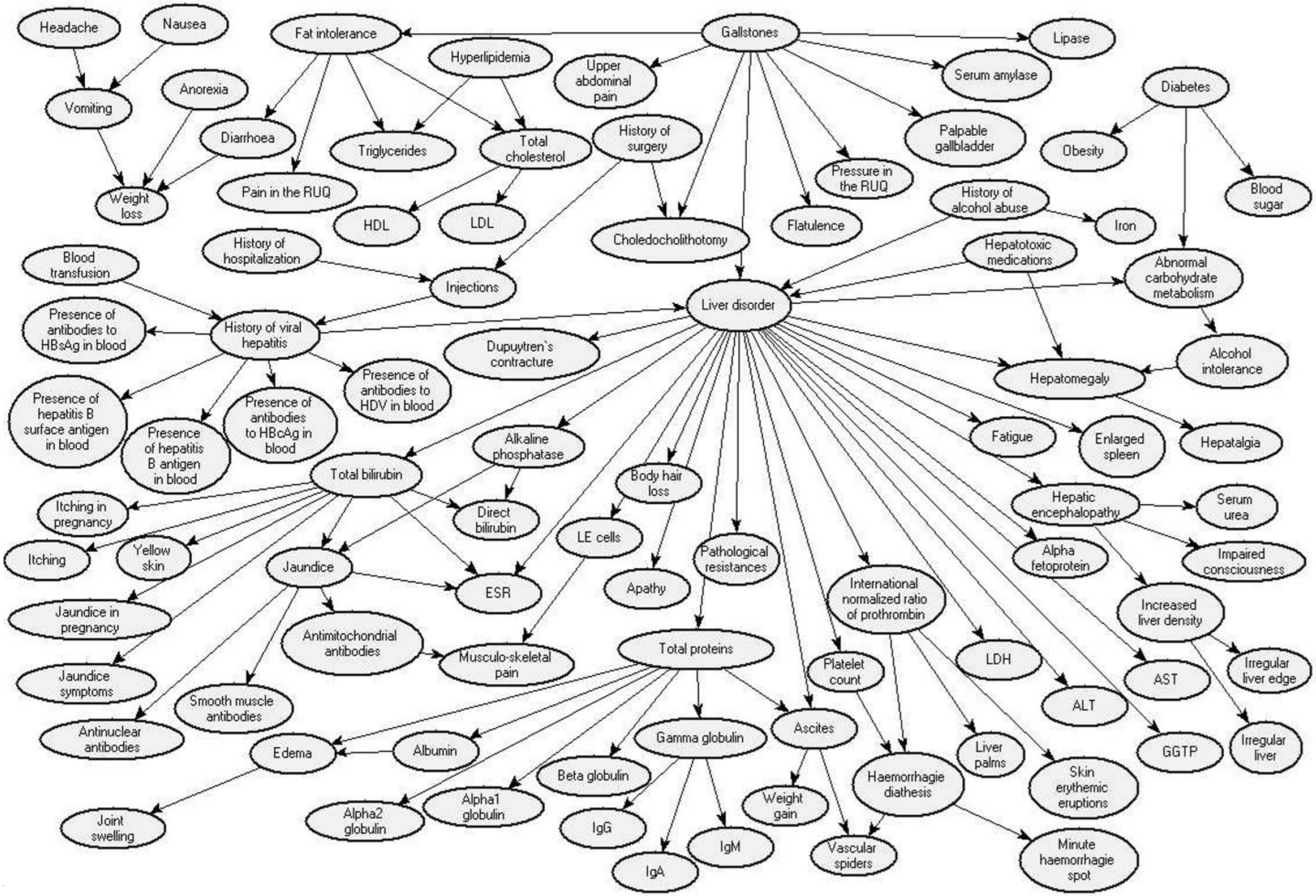
- Bonus Q3. Instead of a deterministic inspection, now the inspection after two years has a probability of 0.4, 0.8, 0.9, 0.95, 0.98 and 1 of identifying states 1 to 6. This inspection does not detect anything. Now, what is the accumulated probability of failure 5 years?

Bayesian Networks

Applications include

- Finance
- Medical Diagnosis
- Speech Recognition
- Image Processing
- Spam Filtering
- Engineering Risk Assessment

Bayesian Networks



Summary

- BNs are effective to represent problems with uncertain outcomes, e.g. risk assessments involving multiple dependent variables
- Probabilistic estimates of outcome can be calculated based on prior ‘beliefs’/probabilities obtained from data or expert judgement
- Estimates can be updated based on new evidence

Questions