Bayesian belief network - BBN

PK8200 – Risk Influence Modelling

Jørn Vatn

Bayes theorem

- $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$ and $\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}$
- Combining gives Bayes theorem:
  - $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(\, B|A)\Pr(A)}{\Pr(B)}$
- Law of total probability $\Pr(B) = \Pr \sum (B|A_i) \Pr(A_i)$, thus we also have
  - $\Pr(A|B) = \frac{\Pr(\, B|A)\Pr(A)}{\sum \Pr(B|A_i) \Pr(A_i)}$

What is a BBN

- A Bayesian belief network is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed ***acyclic*** graph
  - For example, a Bayesian network could represent the probabilistic relationships between diseases and symptoms
  - Given symptoms, the network can be used to compute the probabilities of the presence of various diseases

Working with BBN

- The theory behind BBN is hard to grasp
- In order to efficient perform calculation in a BBN we also need theory which is not easy to approach
- However, there exist computerized tools that can do the necessary manipulations
  - Hugin
  - Netica
  - More…

Motivation

- BBN has been used in Norwegian research projects on risk influence modelling
- Both a full BBN approach, and in the hybrid BBN approach to be covered in this course

Definitions and concepts

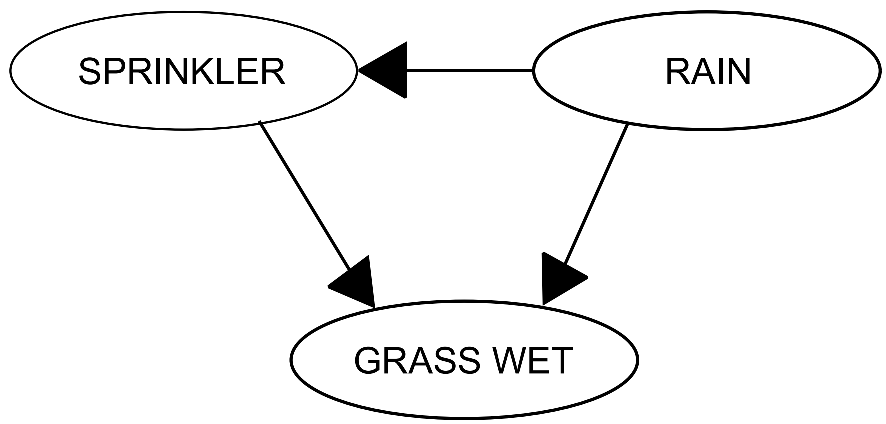
- Bayesian networks are directed acyclic graphs (DAG) whose nodes represent variables, and whose arcs (edges) encode conditional independencies between the variables
- If there is an arc from node A to another node B, A is called a *parent* of B, and B is a *child* of A

Definitions and concepts

- A DAG is a BBN relative to a set of variables if the ***joint distribution*** of the node values can be written as the ***product*** of the local distributions of each node and its parents
- $\Pr(X_1, X_2, \ldots, X_n) = \prod_i \Pr(X_i \,|\text{parents}(Xi))$
  - If X_i has no parents its prob. distribution is *unconditional*
  - If the value of X_i is observed, it is an *evident* node

Wikipedia example – Why is the grass wet?

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

CPT – Conditional Probability Table

- A CPT is a table specifying the *conditional probabilities* for a variable (node) given the value of it's parents
- E.g., "Grass wet" given Sprinkler/Rain

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

CPT

- The CPTs are established by experts and/or data analysis prior to manipulation of the BBN
- Even for a reasonable size of a BBN this is often a very demanding task
- When the CPT are established (prior distribution), we may update the CPT by means of standard Bayesian techniques

Calculation formula

- We have:
- $\Pr(G, S, R) = \Pr(G \mid S, R)\Pr(S \mid R)\Pr(R)$
  - where *G* = *Grass wet*, *S* = *Sprinkler*, and *R* = *Rain*
  - The right hand side are given by the CPTs

Calculation example

$$\Pr(G,S,R) = \Pr(G \mid S,R)\Pr(S \mid R)\Pr(R)$$

$$\mathrm{P}(R=T \mid G=T) = \frac{\mathrm{P}(G=T,R=T)}{\mathrm{P}(G=T)} = \frac{\sum_{S\in\{T,F\}}\mathrm{P}(G=T,S,R=T)}{\sum_{S,R\in\{T,F\}}\mathrm{P}(G=T,S,R)}$$

$$= \frac{(0.99 \times 0.01 \times 0.2 = 0.00198_{TTT}) + (0.8 \times 0.99 \times 0.2 = 0.1584_{TFT})}{0.00198_{TTT} + 0.288_{TTF} + 0.1584_{TFT} + 0_{TFF}} \approx 35.77\%.$$

P(*G* | *S*,*R*)

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

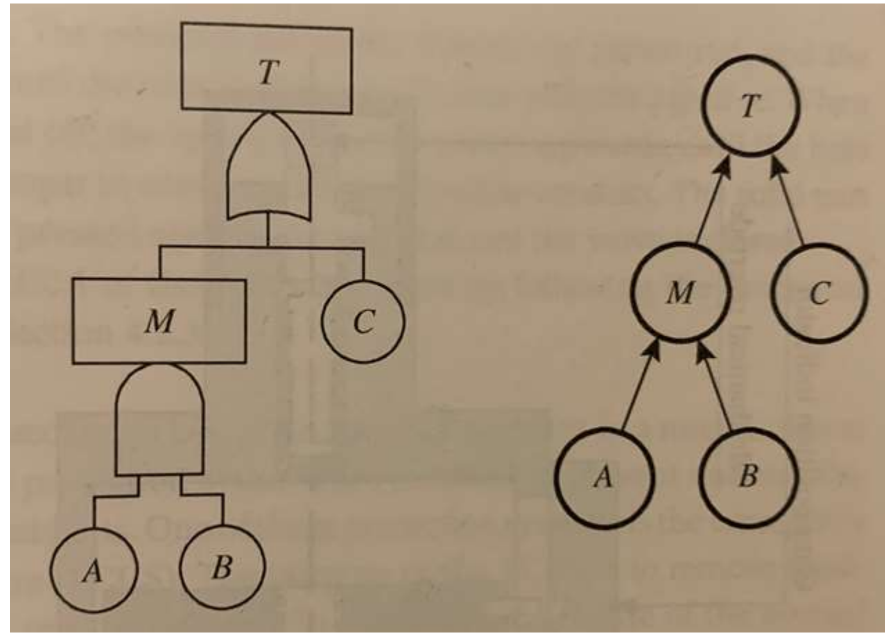
P(*S* | *R*)

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99

P(*R*)

RAIN	
T	F
0.2	0.8

FTA and BBN



- A fault tree can be represented as a BBN
- The basic events are always ***parent*** nodes
- The gates are child notes, but could be parent nodes of other child nodes

Conditional probability tables

- For the “gates”, the CPTs are “binary”, and also referred to as truth tables (1=“occurs”)

A	B	M | A, B
0	0	0
0	1	0
1	0	0
1	1	1

C	M	T | C, M
0	0	0
0	1	1
1	0	1
1	1	1

Calculation

- Since *A* and *B* have no parents, their *joint probability* is found by the product rule:
- $\Pr(M = 1) = \Pr(A = 1 \cap B = 1) = \Pr(A = 1)\Pr(B = 1) = q_A q_B$
- Similarly, since *C* has no parents, it is independent of *M*, hence
- $Q_0 = \Pr(T = 1) = 1 - \Pr(\mathrm{T} = 0) = 1 - \Pr(M = 0 \cap C = 0) = 1 - (1 - q_A q_B)(1 - q_c)$

Pros and Cons

- BBN can do whatever the FTA can do
- To specify the truth tables might be more tedious than using AND and OR gates in the FTA
- If the truth tables and the CPTs are specified, BBN algorithm can in principle solve the quantification, but it might be slow (explosion in the sample space)
- BBN can explicit define dependencies between components in the CPTs
- If two basic events are affected by the same risk factor, we can develop the risk factor structure as part of the BBN

Additional comments

- BBN is a strong tool to model dependencies, and structure of dependencies
- Efficient alogorithms have been developped, but some problems need Monte Carlo simulation
- Often we use BBN as a calculation too to find probabilities (Q_0 and $\Pr(R=T \mid G=T)$), but we can also use BBN to efficiently update the CPTs given new evidence (data)
- In particular if we have evidence in a hierachical structure, BBN is a strong tool for statistical inference