Lecture 2

Bayesian inference in risk analysis

Jan-Erik Holmberg, Alessandro Mancuso
Systeemianalyysin laboratorio
Matematiikan ja systeemianalyysin laitos
Aalto-yliopiston perustieteiden korkeakoulu
PL 11100, 00076 Aalto
jan-erik.holmberg@aalto.fi

Contents

- Bayes’ theorem
- Bayesian estimation of reliability parameters
- Bayesian networks for reliability and risk analysis

Probability measure

Definition: Probability P is a function that maps all events $A \subset S$ onto real numbers and satisfies the following three axioms:

1. If S is the set of all possible outcomes, then $P(S) = 1$
2. $0 \leq P(A) \leq 1$
3. If A and B are mutually exclusive ($A \cap B = \emptyset$) then
$$P(A \cup B) = P(A) + P(B)$$

Probability measure

From the three axioms it follows that:

I. $P(\emptyset) = 0$

II. If $A \subset B$, then $P(A) \leq P(B)$

III. $P(\bar{A}) = 1 - P(A)$

IV. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

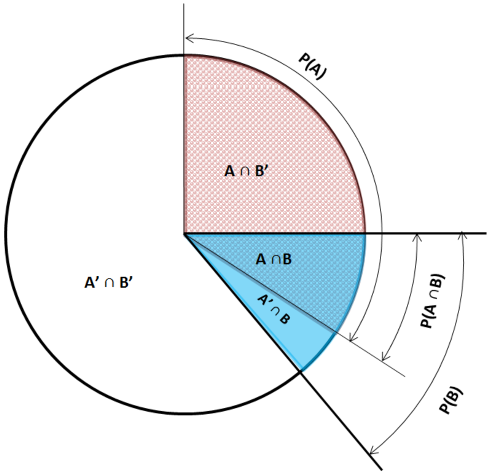
Conditional probability

- Definition of (statistical) independence: Two events A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

- Conditional probability $P(A|B)$ of A given that B has occurred is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



Conditional probability

- If A and B are independent, the probability of A (B) does not depend on whether B (A) has occurred or not:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\,P(B)}{P(B)} = P(A)$$

Law of total probability

- If $B_1,\ldots, B_n$ are mutually exclusive and collectively exhaustive events, then

$$P(A) = P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)$$

- Most frequent use of the law of total probability:
  - Events B and $\bar{B}$ are mutually exclusive and collectively exhaustive
  - Probabilities $P(A|B)$, $P(A|\bar{B})$, and $P(B)$ are known
  - These can be used to compute

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

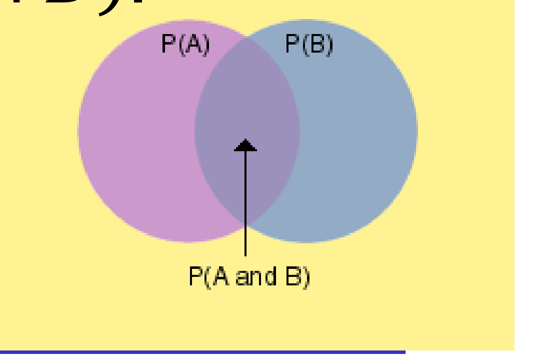
Bayes' rule

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

It follows from the definition of conditional probability:

$P(A|B) = \frac{P(A \cap B)}{P(B)}$; $P(B|A) = \frac{P(B \cap A)}{P(A)}$

Commutative laws: $P(B \cap A) = P(A \cap B)$.



Bayes’ rule example

- The probability of a fire in a certain building is 1/10000 any given day.
- An alarm goes off whenever there is an actual fire, but also once in every 200 days for no reason (false alarm).
- Suppose the alarm goes off. What is the probability that there is a fire?

Bayes' rule example, solution

$F = Fire, \bar{F} = No\ fire, A = Alarm, \bar{A} = No\ alarm$

$P(F) = 0.0001, P(\bar{F}) = 0.9999, P(A|F) = 1,$
$P(A|\bar{F}) = 0.005$

Bayes: $P(F|A) = \frac{P(A|F)P(F)}{P(A)} = \frac{1 \cdot 0.0001}{0.0051} \approx 2\%$

Law of total probability:

$$P(A) = P(A|F)P(F) + P(A|\bar{F})\,P(\bar{F}) = 0.0051$$

Parameter estimation

- Random variable X
- Probability distribution function $F(x;\theta_1,\ldots,\theta_r)$
- Parameters $(\theta_1,\ldots,\theta_r)$ unknown
- Observations (sample, evidence, data) $(x_1,\ldots,x_n)$

- Which parameter values best fit the observations for the given probability model?

Classical parameter estimation

- Maximum likelihood method (MLE)
- Method of moments

- Example: failure probability per demand, p
  - Data: k failures in n trials
  - Classical (MLE/method of moments) estimate

$$\hat{p} = \frac{k}{n}$$

- What if $k = 0$?
  - Should we conclude that the failure probability is zero?

Bayesian estimation of parameters

- We define:
  - the vector $\theta = \{\theta_1, \dots, \theta_r\}$ of parameters, which are treated as random variables
  - the set $E = \{x_1, \dots, x_n\}$ of observations (evidence)
- Then:
  - Likelihood function $L[E|\theta]$ is the conditional probability to observe E given certain parameters values
  - $p(\theta)$ = prior probability distribution for parameters
- Application of Bayes' rule:

$$P(\theta|E) = \frac{L(E|\theta)p(\theta)}{\int_{\theta\in\Theta} L(E|\theta)p(\theta)d\theta}$$

  which is the posterior probability distribution for parameters
- Expected values of $P(\theta|E)$ are often applied as point estimates for the parameters

Bayesian estimation of parameters Binomial distribution (failure probability per demand)

- Hypothesis: $X \sim Binomial(p, n)$
  - n is the number of trials
  - p is the probability of failure per trial
  - X is the number of failures in n trials
- Data: k failures in n trials
- Likelihood function of probability p is

$$L[k|p,n] = P[X = k|p,n] = \binom{n}{k} p^k (1-p)^{n-k}$$

Aalto-yliopiston perustieteiden korkeakoulu
Matematiikan ja systeemianalyysin laitos

Bayesian estimation of parameters Binomial distribution

- Typical choice for a prior distribution is Beta because it is a conjugate distribution for binomial sampling → Posterior distribution is also Beta

$$f(p;\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}$$

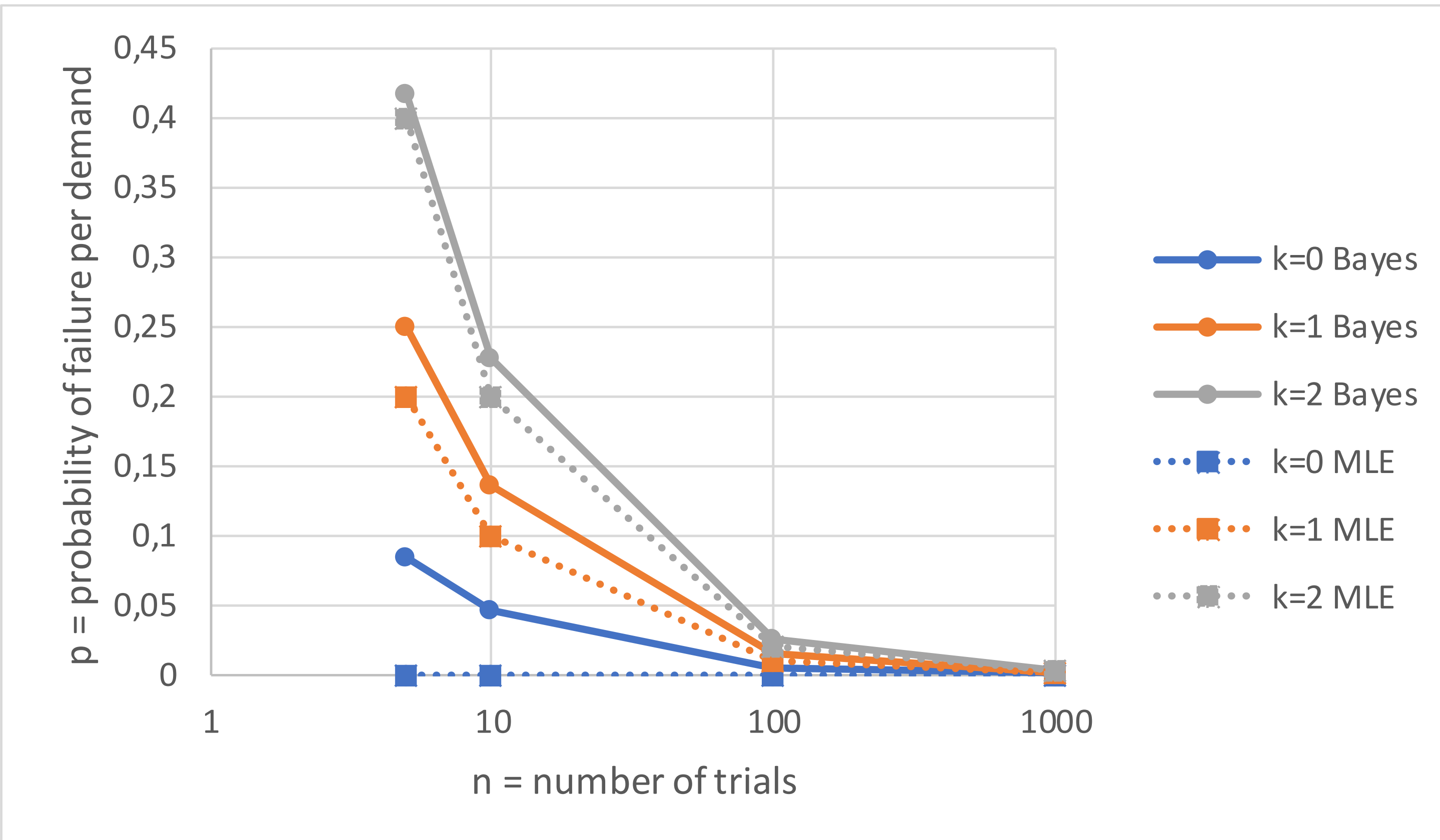
$$p\sim Beta(\alpha+k,\beta+n-k)$$

- Thus, the posterior mean is

$$\mathbb{E}[p|k,n] \ =\frac{k+1/2}{n+1}$$

based on non-informative prior $Beta(0.5,0.5)$ which is the most common in risk analysis.

Bayesian estimation of parameters Binomial distribution, Example



Bayesian estimation of parameters Poisson distribution (exponential failure rate)

- Hypothesis: $X \sim Poisson(\lambda T)$
  - λ is the failure rate
  - T is the time interval
  - X is the number of failures during $[0, T]$
- Data: k failures during $[0, T]$
- Likelihood function of failure rate λ

$$L[k|\lambda, T] = P[X = k|\lambda, T] = \frac{(\lambda T)^k}{k!} \exp(-\lambda T)$$

Aalto-yliopiston perustieteiden korkeakoulu
Matematiikan ja systeemianalyysin laitos

Bayesian estimation of parameters Poisson distribution

- Typical choice for a prior distribution is Gamma because it is a conjugate distribution for Poisson sampling → Posterior distribution is also Gamma

$$f(\lambda; \alpha, \beta) = \frac{\beta^{\alpha-1}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$\lambda \sim Gamma(\alpha + k, \beta + T)$$

- Thus, the posterior mean is

$$\mathbb{E}[\lambda | k, T] = \frac{k + 1/2}{T}$$

based on non-informative prior $\mathrm{Gamma}(0.5{,}0)$ which is the most common in risk analysis.

POLITECNICO DI MILANO

lasar
laboratory of signal and risk analysis

S2D2
Security, Safety, Defense, Disaster Management and Recovery Cluster del Politecnico di Milano

MASTER COURSE

Bayesian Networks for Reliability and Risk Analysis

POLITECNICO MILANO 1863

M.Sc. Alessandro Mancuso

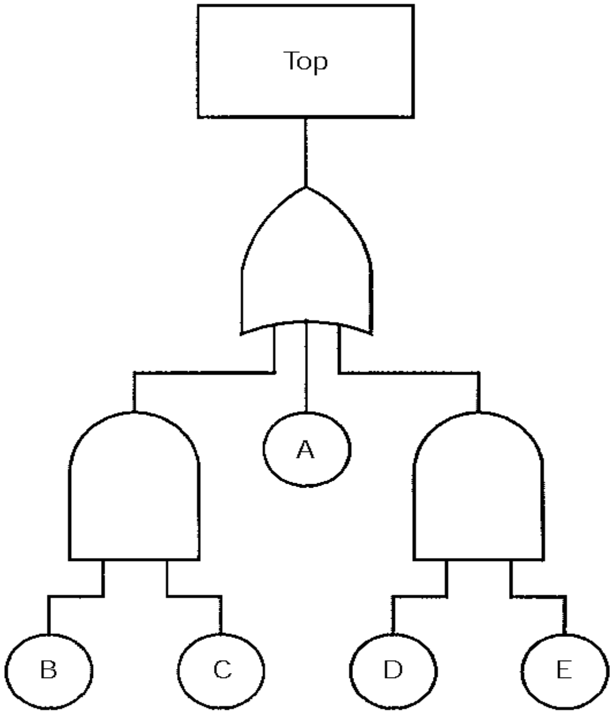
A. Y. 2018 – 2019

Primary techniques in Probabilistic Risk Assessment (PRA)

Scenario modeling and quantification are pursued through:

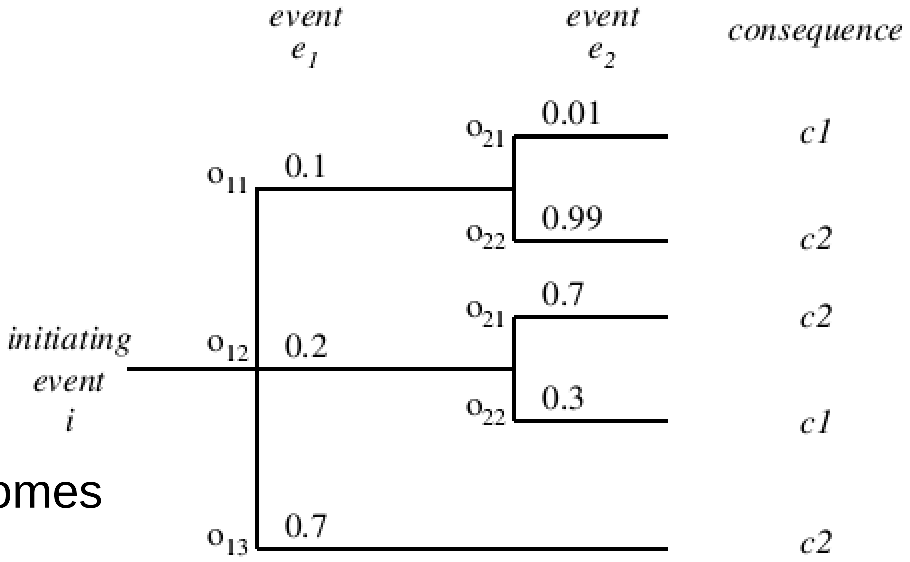
FAULT TREE ANALYSIS (FTA)

1. Events are **binary events** (operating/not-operating);
2. Events are **statistically independent**;
3. Relationships between events and causes are represented by logical AND and OR **gates**;
4. The undesirable event, called Top Event, is postulated and the possible ways for the occurrence of this event are systematically deduced.



Scenario modeling and quantification are pursued through:

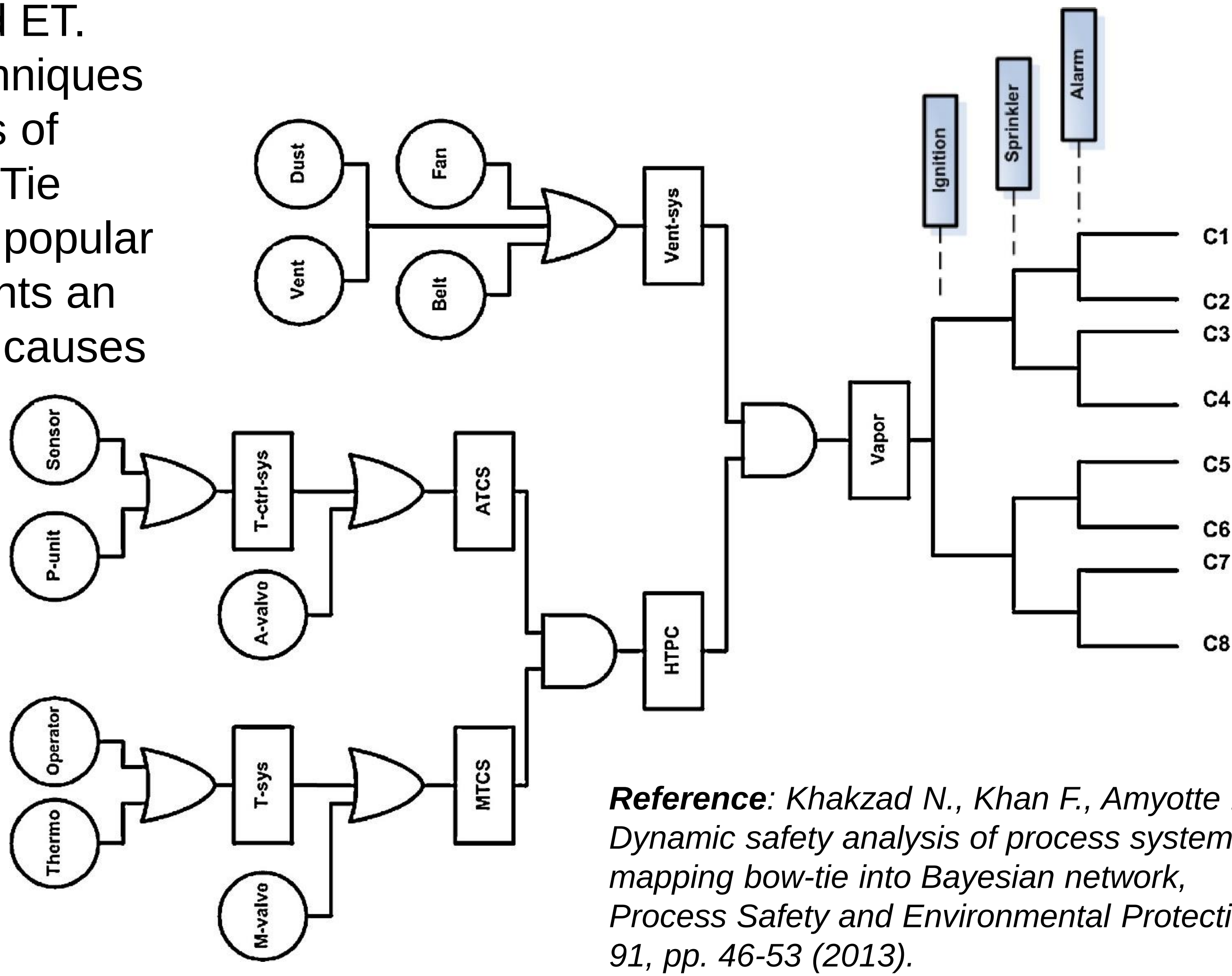
EVENT TREE ANALYSIS (ETA)



1. System evolution following the hazardous occurrence is divided into **discrete events**;
2. System evolution starts from an **initiating event**;
3. Each event has a finite set of outcomes (commonly there are two outcomes: **occurring event or not occurring**) associated with the occurrence probabilities;
4. The leafs of the event tree represent the **consequence scenarios** to be analyzed.

Bow-Tie analysis

Bow-Tie (BT) combines the scenario modeling and quantification of FT and ET. Among the various techniques used for safety analysis of process systems, Bow-Tie analysis is becoming a popular technique as it represents an accident scenario from causes to effects.



***Reference**: Khakzad N., Khan F., Amyotte P., Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network, Process Safety and Environmental Protection 91, pp. 46-53 (2013).*

Limitations of Bow-Tie analysis

The application of BT in reliability and risk analysis is limited due to:

1. The **static nature** of its components, Fault Tree and Event Tree.

   ⮡ Restrictions in describing system dynamical behavior.

2. Inability to represent **conditional dependence**.

   ⮡ Event dependence is common among primary events and safety barriers.

3. Difficulties in handling **imprecise information** (probabilities).

   ⮡ Limited availability of data is frequent in process safety analysis.

To overcome these limitations, the Bow-Tie can be mapped into a **Bayesian Network** (BN). Formally, a BN is a directed acyclic graph consisting of:

- **Nodes** (circles) represent the BT random events whose combination can lead to system failure. In particular, when the BT is converted into BN, some BT events can be merged to the same node;
- **Directed arcs** indicate conditional dependencies among nodes. Specifically, the arc (i, j) which connects node j to node i shows that the event at node j is conditionally dependent to the event at node i.

Bayesian Networks have been **successfully applied in many fields**, such as ecology, medicine, reliability, computer science…

Why Bayesian Networks?

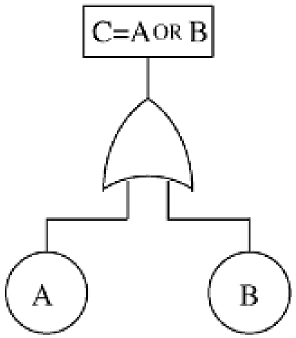
Bayesian networks are **probabilistic graphical models**, which offers a convenient and efficient way of generating joint distribution of all its events.

- **Convenient**: causal relationships between events are easy to model.
- **Efficient**: no redundancies in terms of graphical modelling and probability computations.
- **Flexible**: capable of handling imprecise information by capturing quantitative and qualitative data.

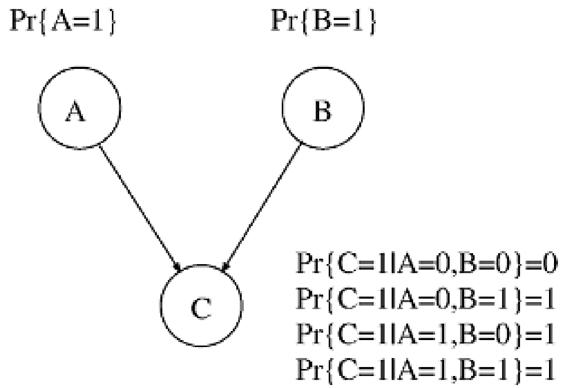
Wide applicability, for instance:

- Diagnosis: Microsoft trouble shooting wizard, Medical expert system
- Classification: Spam email classification
- Voice recognition

Mapping FT into Bayesian Network

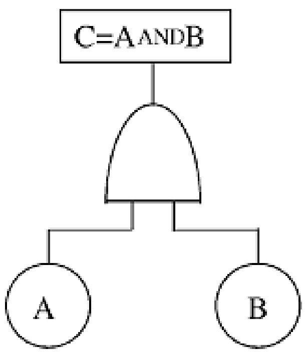


FAULT - TREE: OR Gate

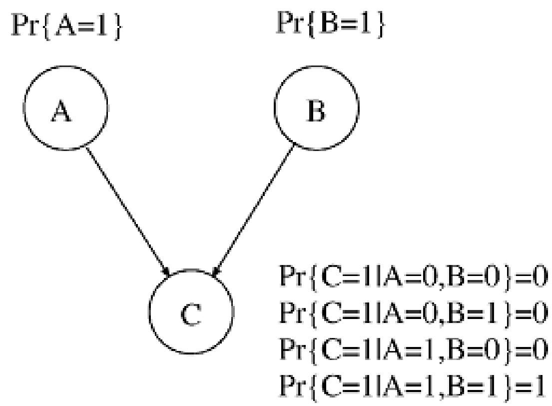


BAYESIAN NETWORK: OR Node

Switching from binary logic to probabilistic logic!



FAULT - TREE: AND Gate



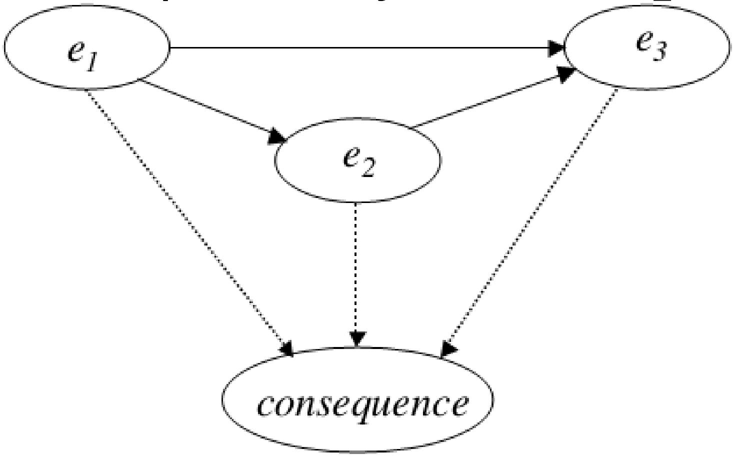
BAYESIAN NETWORK: AND Node

***Reference**:* Bobbio A., Portinale L., Minichino M., Ciancamerla E., "*Improving the analysis of dependable systems by mapping fault trees into Bayesian networks",* Reliability Engineering and System Safety 71, pp. 249–260 (2001) .

Any event tree with three events e_1, e_2, and e_3 can be represented by the BN shown below.

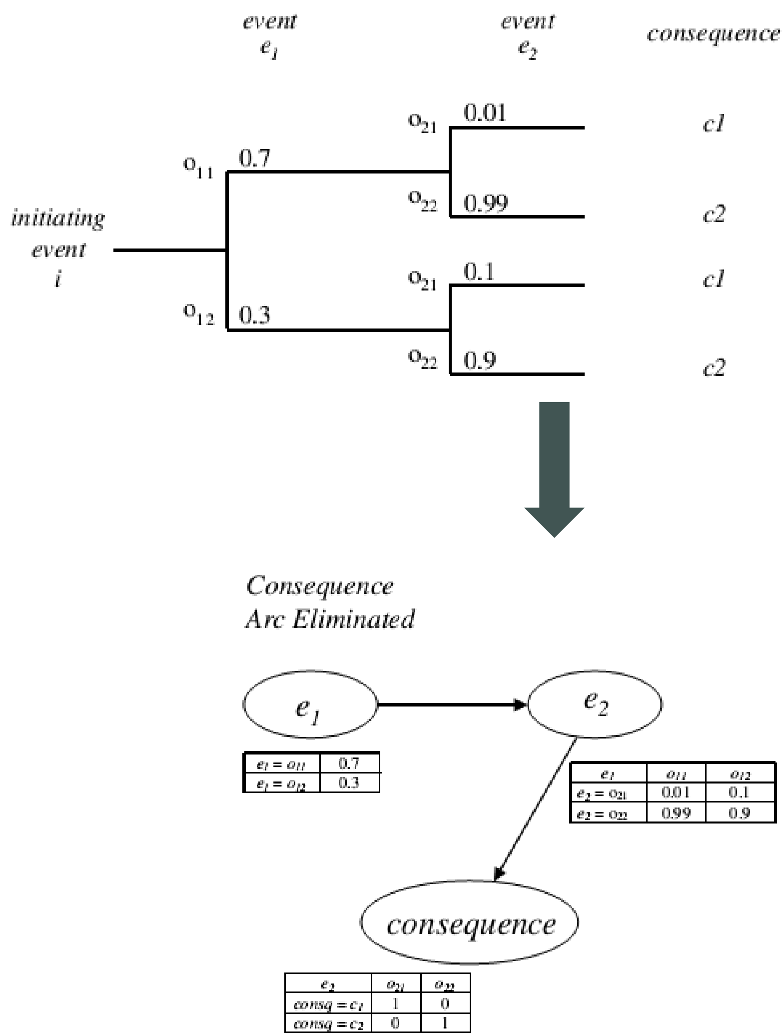
Two types of directed arc complete the network:

- Consequence arcs (shown as dotted lines) connect each event node to the consequence node. This relationship is deterministic: the probability table for the consequence node encodes the logical relationship between the events and the consequences.
- Causal arcs (shown as solid lines) connect each event node to all events later in time. For instance, event e_1 is a causal factor for event e_2, thus it influences the probability of event e_2.



Reference: Bearfield G., Marsh W., *"Generalizing Event Trees Using Bayesian Networks with a Case Study of Train Derailment"*, Computer Safety, Reliability, and Security (2005).

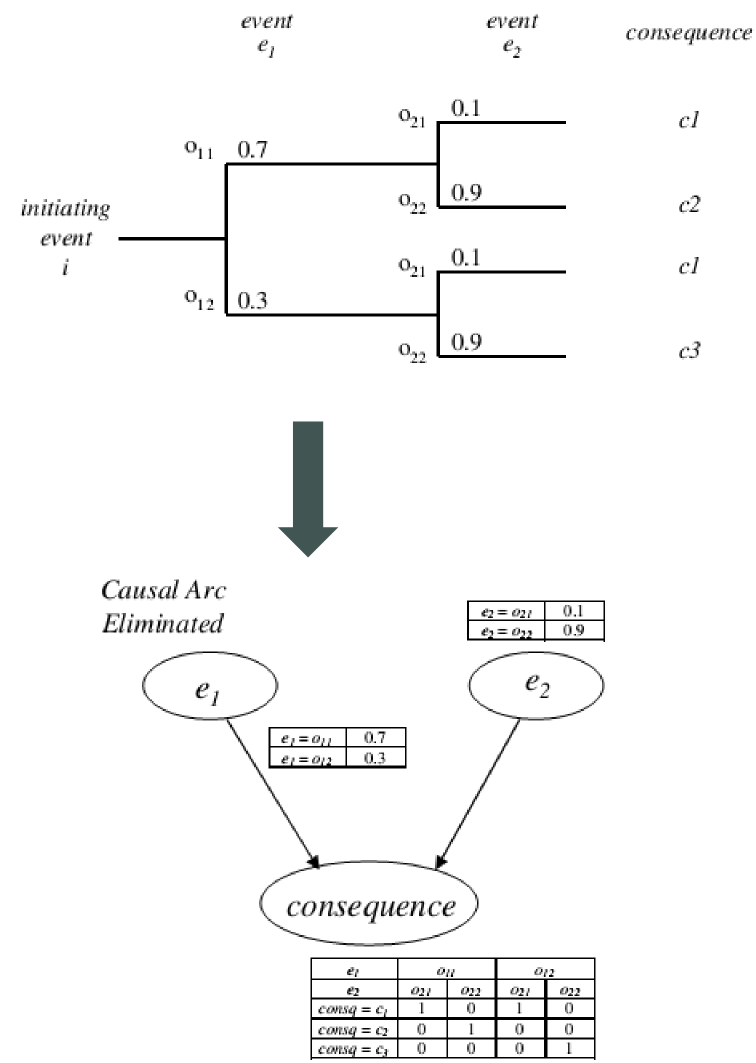
Mapping ET into Bayesian Network (alternatives)



$e_1 = o_{11}$	0.7
$e_1 = o_{12}$	0.3

e_1	o_{11}	o_{12}
$e_2 = o_{21}$	0.01	0.1
$e_2 = o_{22}$	0.99	0.9

e_2	o_{21}	o_{22}
$consq = c_1$	1	0
$consq = c_2$	0	1



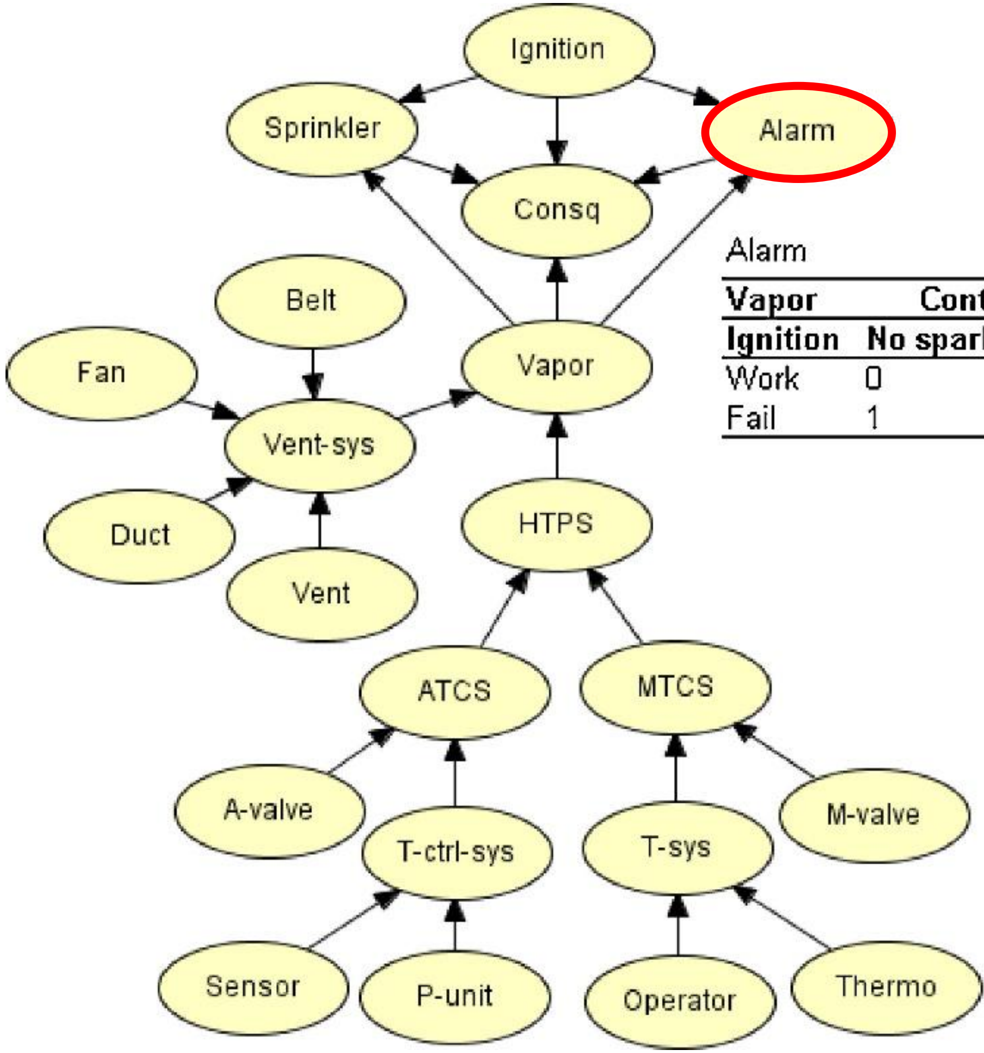
$e_2 = o_{21}$	0.1
$e_2 = o_{22}$	0.9

$e_1 = o_{11}$	0.7
$e_1 = o_{12}$	0.3

e_1	o_{11}		o_{12}	
e_2	o_{21}	o_{22}	o_{21}	o_{22}
$consq = c_1$	1	0	1	0
$consq = c_2$	0	1	0	0
$consq = c_3$	0	0	0	1

Advantages of Bayesian Networks

Mapping of Bow Tie (slide 4) into **Bayesian Network**:



Alarm

Vapor	Controlled		Overflow	
Ignition	**No spark**	**Spark**	**No spark**	**Spark**
Work	0	0	0.775	0.9987
Fail	1	1	0.225	0.0013

<u>Advantages</u>

- Multi-state modeling.
- Extension of concepts of AND/OR gates through probability distributions.
- Combining expert judgement and quantitative knowledge to estimate the risk.

Failure probabilities

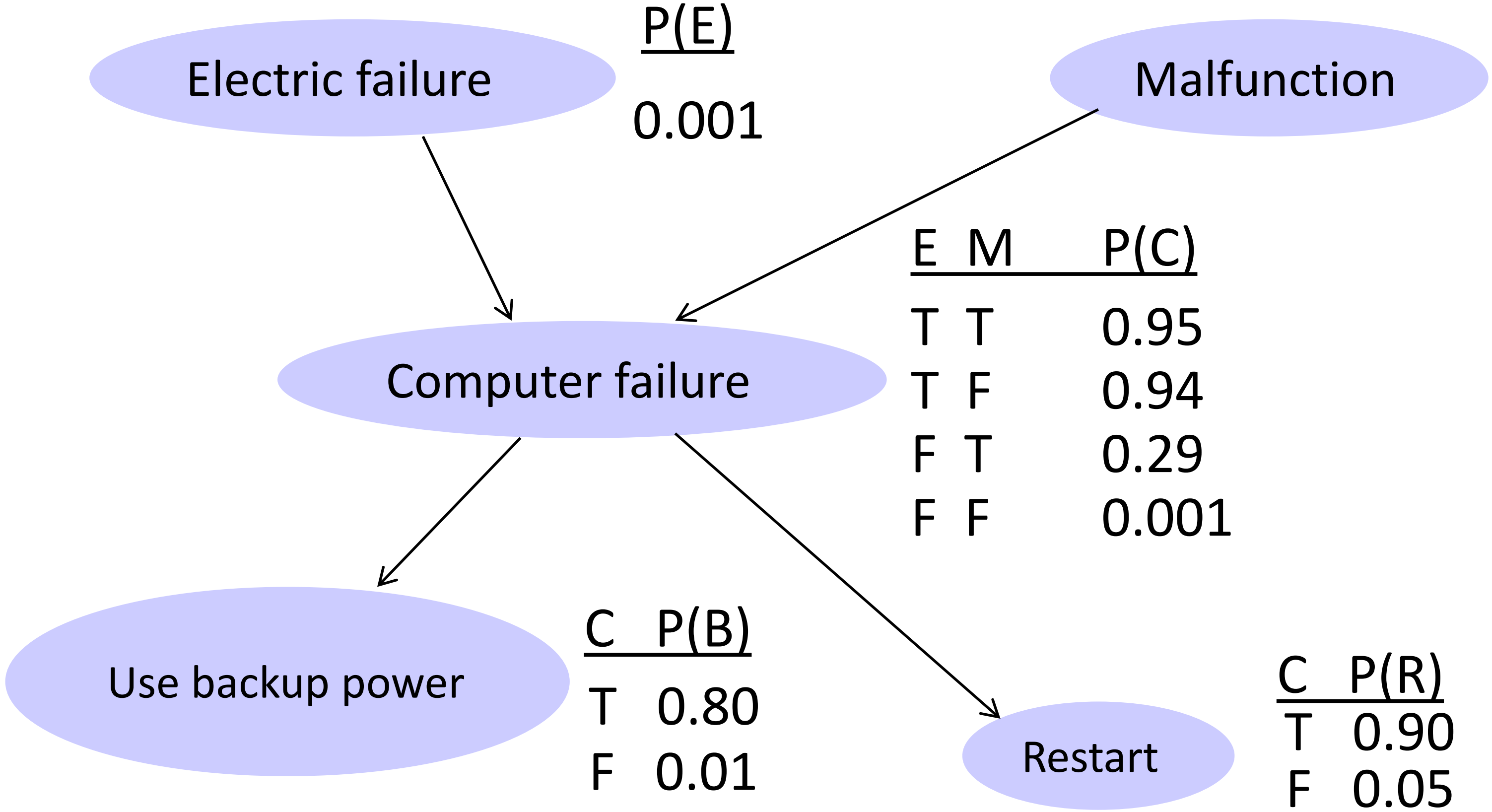
Information sources

- Information provided by AND/OR gates in BT.
- Statistical analyses / Simulations.
- Expert elicitation.

The probabilities of events are defined as follows:

- Initiating events → failure probabilities of system components.
- Intermediate and top events → conditional probability tables (CPT).

Bayesian Network representation



P(E)
0.001

P(M)
0.02

E	M	P(C)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

C	P(B)
T	0.80
F	0.01

C	P(R)
T	0.90
F	0.05

Query: Assuming to collect some observations (**evidence**) from the system, how would this evidence impact the probabilities of the events?

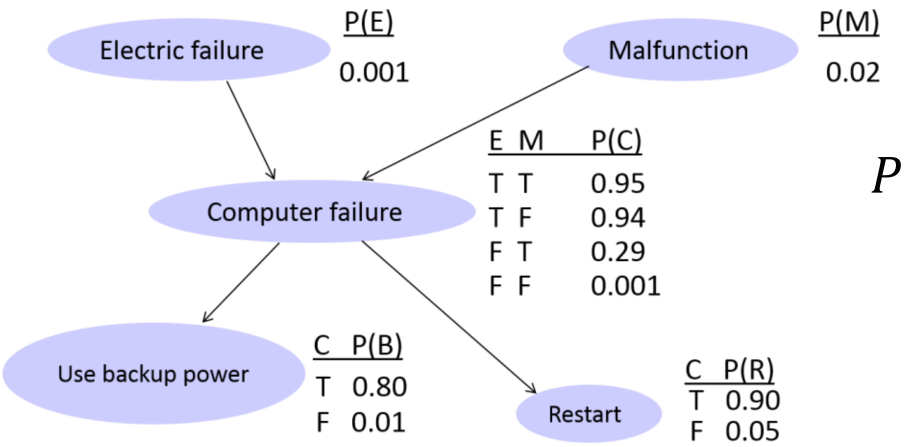
The conditional probability of a random event given the evidence is known as **a posteriori belief**, useful in case of:

- **Prediction**: computing the probability of an outcome event given the starting condition → Target is a descendent of the evidence!
- **Diagnosis**: computing the probability of disease/fault given symptoms → Target is an ancestor of the evidence!

Note: the direction between variables does not restrict the directions of the queries → Probabilistic inference can combine evidence from all parts of the network!

Examples of inference
Prediction

For the computer failure example, what it is the probability that the backup power is working given an electrical failure?



$$P(B|E) = P(B|C)P(C|E) + P(B|\bar{C})P(\bar{C}|E)$$

$$P(C|E) = P(C|\bar{M},E)P(\bar{M}) + P(C|M,E)P(M)$$

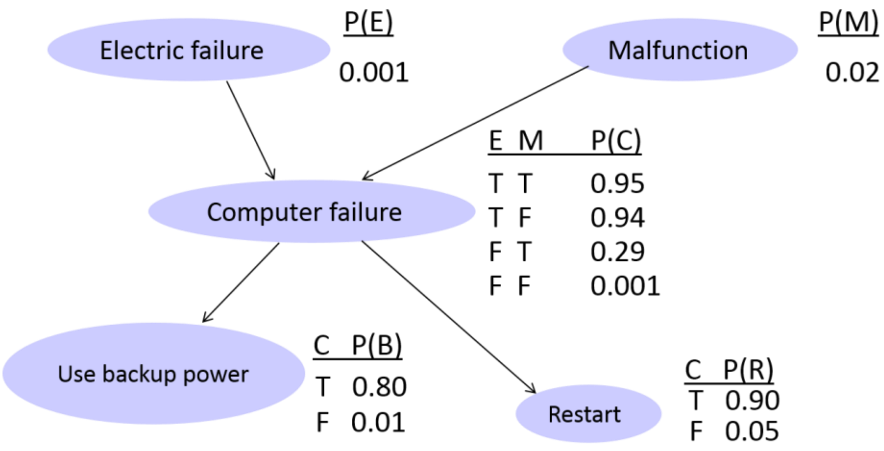
$$P(C|E) = 0.98 \cdot 0.94 + 0.02 \cdot 0.95 = 0.94$$

$$P(B|E) = P(C|E) \cdot 0.8 + P(\bar{C}|E) \cdot 0.01$$

$$P(B|E) = 0.94 \cdot 0.8 + 0.06 \cdot 0.01 = 0.75$$

Examples of inference
Diagnosis

For the computer failure example, what it is the probability that the electricity is working given a backup power failure?



$$P(E|B) = \frac{P(B|E)P(E)}{P(B)}$$

$$P(E|B) = \frac{0.75 \cdot 0.001}{0.0161} = 0.0468$$

$$P(C) = P(C|E,M)P(E)P(M) + P(C|E,\bar{M})P(E)P(\bar{M}) + P(C|\bar{E},M)P(\bar{E})P(M) + P(C|\bar{E},\bar{M})P(\bar{E})P(\bar{M}) = 0.0077$$

$$P(B) = P(B|C)P(C) + P(B|\bar{C})P(\bar{C})$$

$$P(B) = 0.8 \cdot 0.0077 + 0.01 \cdot 0.9923 = 0.0161$$

Computational issues

Theorem: Computing event probabilities in a Bayesian network is **NP-hard**.

Hardness does not mean it is impossible to perform inference, but

- There is no general procedure that works efficiently for all networks.
- For particular families of networks, there are proved efficient procedures.
- Different algorithms are developed for inferences in Bayesian networks.

There are available softwares that efficiently perform Bayesian Network inference through a library of functions for several popular algorithms, among those:

- **GeNIe Modeler**: https://www.bayesfusion.com/genie-modeler
- **HUGIN Expert**: https://www.hugin.com/