Algoritmos y Programación III

Andrés A. Aristizábal P. aaaristizabal@icesi.edu.co
Angela Villota apvillota@icesi.edu.co

Departamento de Computación y Sistemas Inteligentes
UNIVERSIDAD ICESI

2024-2

1. **Reasoning with uncertainty**
   - Bayesian networks
   - Exercises

Bayesian networks- Intro

Bayesian networks?

Bayesian networks- Intro

Bayesian networks?

They are key a computer technology for :

- Modeling Uncertainty: They effectively handle uncertainty by using probabilities.
- Inference: They allow us to compute the likelihood of certain outcomes given partial evidence.
- Causal Reasoning: Bayesian networks can model causal relationships, not just correlations, which is useful in decision-making.

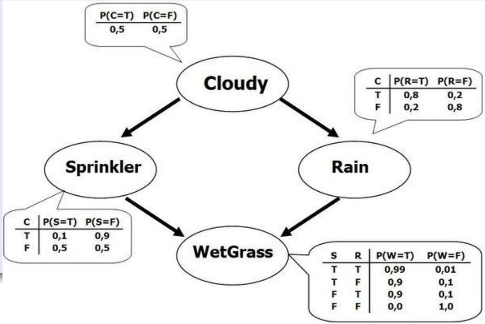
Examples

- Diagnosis and Decision Support: Aid in medical and system fault diagnosis.
- Prediction and Forecasting: Predict outcomes like weather, stock market trends.
- Risk Analysis: Assess risks in business, insurance, and operations.

Bayesian networks

A **Bayesian network** is a probabilistic, graphical model which represents a set of variables and their conditional dependencies using a directed acyclic graph.

It is also called a Bayes network, belief network, decision network, or Bayesian model.



When do we need a Bayesian network?

Bayesian networks

When do we need a Bayesian network?

- Real world applications are probabilistic in nature, and to represent the relationship between multiple events, we need a Bayesian network.
- It can also be used in various tasks including prediction, anomaly detection, diagnostics, automated insight, reasoning, time series prediction, and decision making under uncertainty.
- Bayesian Network can be used for building models from data and experts opinions.

Which are the two parts in which a Bayesian network is composed of?

Bayesian networks

Which are the two parts in which a Bayesian network is composed of?

- Directed Acyclic Graph
- Conditional Probabilities Tables (CPT).

Directed Acyclic Graph

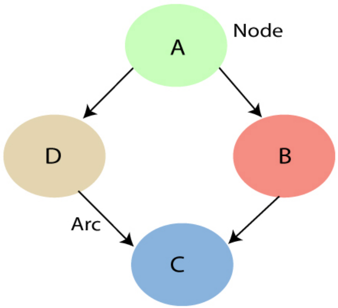
Bayesian networks

Directed Acyclic Graph

Each node corresponds to the random variables

Arcs or directed arrows represent the causal relationship or conditional probabilities between random variables.

These links represent that one node directly influence the other node, and if there is no directed link that means that nodes are independent with each other.



Bayesian networks

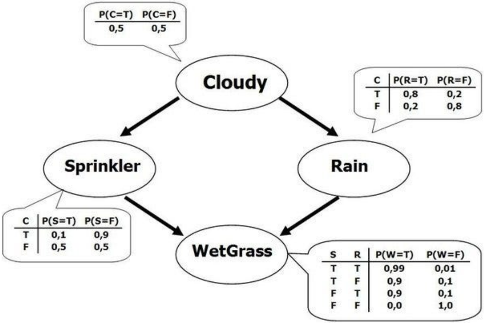
And the tables??

Bayesian networks

And the tables??

Each node in the Bayesian network has condition probability distribution $P(X_i|Parent(X_i))$, which determines the effect of the parent on that node.

This probability distribution is presented in the CPT associated to each node



Let us look at an example

Let us look at an example

Example

I have a burglar alarm that is sometimes set off by minor earthquakes. My two neighbors, John and Mary, promised to call me at work if they hear the alarm.

- *Inference task: suppose Mary calls and John doesn't call. What is the probability of a burglary.*

Bayesian networks

What are the random variables?

Bayesian networks

What are the random variables?

Example

- *Burglary (B)*
- *Earthquake(E)*
- *Alarm(A)*
- *John Calls(J)*
- *Mary calls(S)*

Bayesian networks

What are the direct influence relationships?

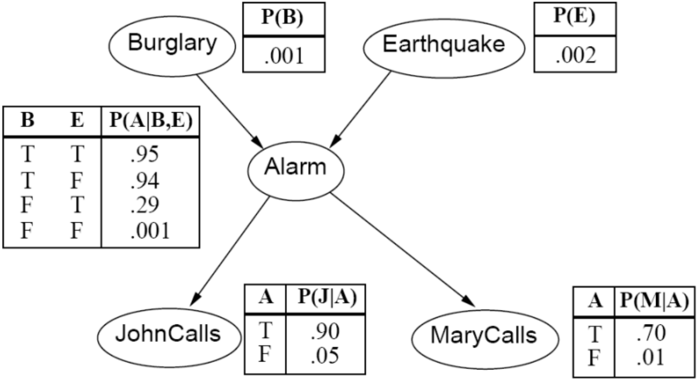
Bayesian networks

What are the direct influence relationships?

Example

- *A burglar can set the alarm on*
- *An earthquake can set the alarm on*
- *The alarm can cause Mary to call*
- *The alarm can cause John to call*

Example: Burglar Alarm



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

Bayesian networks

What to take into account?

Bayesian networks

What to take into account?

- Key property: **each node is conditionally independent** of its non-descendants given its parents
- Suppose the nodes $X_1, ..., X_n$ are sorted in topological order
- To get the joint distribution $P(X_1, ..., X_n)$, use the chain rule:

$$\begin{aligned} P(X_1,\ldots,X_n) &= \prod_{i=1}^{n} P\left(X_i \mid X_1,\ldots,X_{i-1}\right) \\ &= \prod_{i=1}^{n} P\left(X_i \mid Parents(X_i)\right) \end{aligned}$$

Conditional independence?

Bayesian networks

Conditional independence?

- The conditional independence describes a situation where two events are independent of each other given the knowledge of a third event.
- Knowing whether B occurs does not provide further information about A once C is known.
- $P(A, B|C) = P(A|C) \cdot P(B|C)$

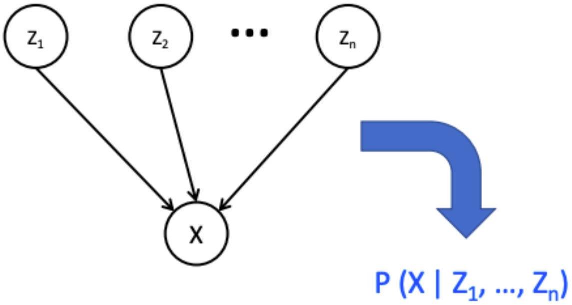
Bayesian networks

Then what?

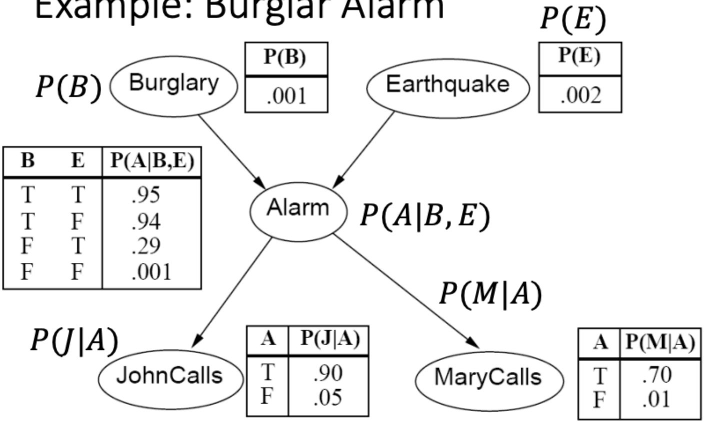
Then what?

Solution

- *We need to specify the full joint distribution.*
- *Therefore, we need to specify a conditional distribution for each node given its parents:* $P(X|Parents(X))$



Example: Burglar Alarm



What do we have here?

Bayesian networks

What do we have here?

Solution

- *That a model is a complete specification of the dependencies.*
- *The conditional probability tables are the model parameters.*

Bayesian networks

How do we estimate posteriori probabilities?

How do we estimate posteriori probabilities?

Solution

Inference Task: *suppose Mary calls and John doesn't call. What is the probability of a burglary.*

- *We find the joint probability of B (and $\neg$B) and M (and $\neg$M) and any other variables that are necessary in order to link these two together.*
- $P(B, E, A, M) = P(B) \cdot P(E) \cdot P(A|B, E) \cdot P(M|A)$

Bayesian networks

How do we do that?

How do we do that?

Solution

$P(BEAM)$	$\neg M, \neg A$	$\neg M, A$	$M, \neg A$	M, A
$\neg B, \neg E$	0.986045	2.99×10^{-4}	9.96×10^{-3}	6.98×10^{-4}
$\neg B, E$	1.4×10^{-3}	1.7×10^{-4}	1.4×10^{-5}	4.06×10^{-4}
$B, \neg E$	5.93×10^{-5}	2.81×10^{-4}	5.99×10^{-7}	6.57×10^{-4}
B, E	9.9×10^{-8}	5.7×10^{-7}	10^{-9}	1.33×10^{-6}

What is the following step?

What is the following step?

Solution

- *We marginalize (add) to get rid of the variables we don't care about.*
- $P(B, M) = \sum_{E,\neg E} \sum_{A,\neg A} P(B, E, A, M)$

Bayesian networks

And we get

And we get

Solution

$P(B,M)$	$\neg M$	M
$\neg B$	0.987922	0.011078
B	0.000341	0.000659

What next?

What next?

Solution

- *We ignore the column that didn't happen.*

And we get

And we get

Solution

$P(B,M)$	M
$\neg B$	0.011078
B	0.000659

And last?

And last?

Solution

- *We use the definition of conditional probability.*
- $P(B|M) = \dfrac{P(B, M)}{P(B, M) + P(B, \neg M)}$

And we get

And we get

Solution

$P(B\|M)$	M
$\neg B$	0.943883
B	0.056117

In order to answer the question we have that ...

In order to answer the question we have that ...

Solution

- *A burglary is so unlikely that, if only Mary calls or only John calls, the probability of a burglary is still only about 5 %.*

Exercises

Exercise

Implement a bayesian network in a Jupyter notebook for the first example and upload it to intu.