

The tool we will use to work with probabilistic reasoning is called a **Bayesian Network**. This consists of two major components:

1. Directed Acyclic Graph shows causality
2. Numbers/Nodes represent probability

Notation:

- Variable — X
- Value — x
- Variable Probability Distribution — $\Pr(x) = \Pr(X=x)$
- Set of Variables — X
- Instantiation — x
- World Probability Distribution — $\Pr(X) = \text{truth table}$

Remember that if $\Pr(\alpha|\beta, \gamma) = \Pr(\alpha|\gamma)$, then α & β are independent given γ

We can apply the same idea to sets of variables:

X and Y are independent given Z if $\Pr(x|y, z) = \Pr(x|z)$

We write this as $I(X, Z, Y)$

Consider the following

!DirectedAcyclicGraph(G)](https://paper-attachments.dropbox.com/s13F2D47E7D3DCA10BE34C69E62E331D2224F9C5F91BEA8639D242F2792B8651715901156979drawing.jpg)

Logic

The alarm is triggered by burglary or earthquake. A radio report is caused by an earthquake. A call from the neighbor may come post-alarm.

Independence

R & C are independent given A — $I(R, A, C)$

E & B are independent — $I(E, \phi, B)$

This can be much more efficient than a table.

The independence and parameters uniquely identify the graph.

We introduce a few terms:

- $\text{Parents}(V)$ — the nodes pointing through V via an edge
 $\text{Parents}(A) = \{E, B\}$ & $\text{Parents}(R) = \{E\}$
- $\text{Descendants}(V)$ — the nodes reachable by V
 $\text{Descendants}(E) = \{R, A, C\}$ & $\text{Descendants}(B) = \{A, C\}$ & $\text{Descendants}(X) = \{\}$
- $\text{Non-Descendants}(V)$ — the nodes unreachable from V, excluding the self & parents
 $\text{Non-descendants}(A) = \{R\}$ & $\text{Non-descendants}(E) = \{B\}$

We call these independence statements the **Marcovion Assumptions**

\equiv all the defined independence statements of a Bayesian Network

These are of the form $I(V, \text{Parents}(V), \text{Non-Descendants}(V))$

$\text{Marcov}(G) = \{ I(C, A, BE), I(R, E, ABC), I(A, BE, R), I(B, \phi, ER), I(E, \phi, B) \}$

We can parametrize the structure by assigning weight to the connections.

4

Note that $Pr(\neg d, b, \neg c) = \theta_{\{\neg d|b, \neg c\}}$

Consider an arbitrary row abcde

$$Pr(abcde) = \theta_a \theta_{b|a} \theta_{c|a} \theta_{d|bc} \theta_{e|c} = 0.7$$

$$Pr(ab\neg cd\neg e) = \theta_a \theta_{b|a} \theta_{\neg c|a} \theta_{d|b\neg c} \theta_{\neg e|c} = 0.216$$

$$Pr(\neg(abcde)) = \theta_{\neg a} \theta_{\neg b|\neg a} \theta_{\neg c|\neg a} \theta_{\neg d|\neg b\neg c} \theta_{\neg e|\neg c} = 0.09$$

We can evaluate the independence of any two nodes are independent given a third by

Deseperation

This uses an expansion on Markov's much more efficient than probability theory.

If setting Z blocks paths $x \rightarrow y$, x & y are d-separated given Z

There are three ways for a node to be connected in a path:

- sequential $\equiv \rightarrow w \rightarrow$
- divergent $\equiv \leftarrow w \rightarrow$
- convergent $\equiv \rightarrow w \leftarrow$

These are each treated differently:

¬[](*https* : *//paper* –
attachments.dropbox.com/s_13F2D47E7D3DCA10BE34C69E62E331D2224F9C5F91BEA8639D242F2792B86517_15901184517
drawing + 2.jpg)

We can clarify by giving an example:

8