bnet.builder

About Bayesian Belief Networks

charles river analytics

www.cra.com

Copyright

Charles River Analytics, Inc.
625 Mount Auburn Street
Cambridge, MA 02138

Examples and Data

Trademarks and Patents

Document Information

About Bayesian Belief Networks, for BNet Version 1.0
Last updated April 22, 2008

bnet.builder

ABOUT BAYESIAN BELIEF NETWORKS

This document provides a general introduction to Bayesian belief networks. It is intended as background information for users of the BNet suite. It offers a comprehensive introduction to the structure and terminology of Bayesian belief networks, and briefly discuss the underlying mathematics and methods for working with belief networks.

More detailed information on BNet.Builder™ and numerous examples of Bayesian belief networks are provided in the *BNet.Builder User's Guide* and the *BNet.EngineKit Developer's Guide*. Please see these guides if you are interested in more technical information or for additional examples of constructing and using Bayesian belief networks.

This document includes the following key topics:

- Overview
- Bayes' Rule
- Structure of a Bayesian belief network
- States
- Conditional probability tables
- Beliefs and evidence
- References

Overview

Bayesian belief networks are powerful tools for modeling causes and effects in a wide variety of domains. They are compact networks of probabilities that capture the probabilistic relationship between variables, as well as historical information about their relationships.

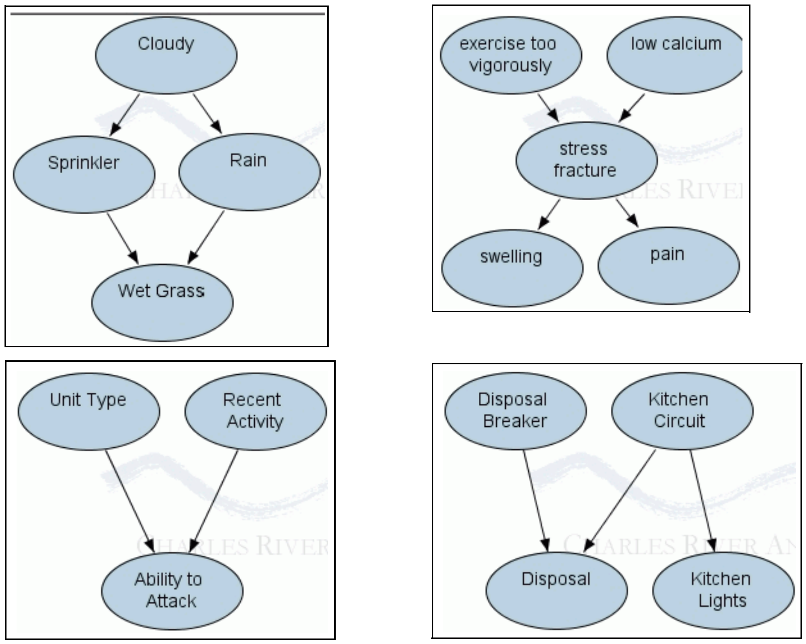
Bayesian belief networks are very effective for modeling situations where some information is already known and incoming data is uncertain or partially unavailable (unlike rule-based or “expert” systems, where uncertain or unavailable data results in ineffective or inaccurate reasoning). These networks also offer consistent semantics for representing causes and effects (and likelihoods) via an intuitive graphical representation. Because of all of these capabilities, Bayesian belief networks are being increasingly used in a wide variety of domains where automated reasoning is needed.

An important fact to realize about Bayesian belief networks is that they are not dependent on knowing exact historical information or current evidence. That is, Bayesian belief networks often produce very convincing results when the historical information in the conditional probability tables or the evidence known is inexact. Given that humans are excellent at vague linguistic representations of knowledge (for example, “it will probably rain tomorrow”), and less adept and providing specific estimates, the ability to be effective despite vagaries in the input information is particularly advantageous. This robustness in the face of imperfect knowledge is one of the many reasons why Bayesian belief nets are increasingly used as an alternative to other AI representational formalisms.

In simpler terms, a Bayesian belief network is a model. It can be a model of anything: the weather, a disease and its symptoms, a military battalion, even a garbage disposal. Belief networks are especially useful when the information about the past and/or the current situation is vague, incomplete, conflicting, and uncertain.

Uncertainty arises in many situations. For example, experts may be uncertain about their own knowledge, there may be uncertainty inherent in the situation being modeled, or uncertainty about the accuracy and availability of information. Because Bayesian belief networks offer consistent semantics for representing uncertainty and an intuitive graphical representation of the interactions between various causes and effects, they are a very effective method of modeling uncertain situations that depend on cause and effect.

Figure 1 Some Bayesian belief networks

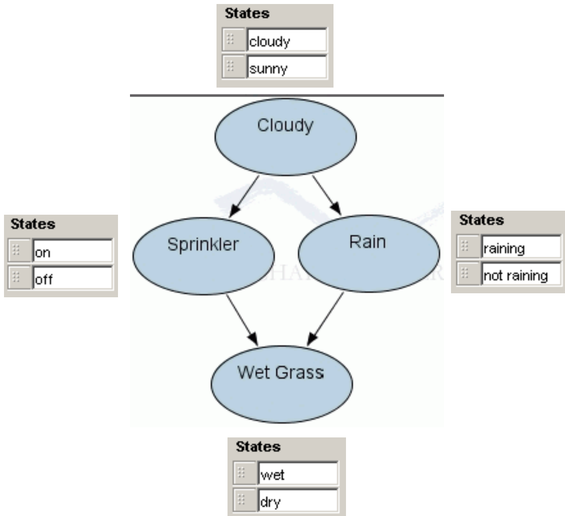


Each of the variables in the Bayesian belief network are represented by *nodes*. A variable in a belief network could be whether a light switch is on, the proximity of an enemy battalion, or the RPM of an engine.

Each node has *states*, or a set of probable values for each variable. For example, the weather could be cloudy or sunny, an enemy battalion could be near or far, symptoms present or not present, and the garbage disposal working or not working.

Nodes are connected to show causality with an arrow indicating the direction of influence. These arrows are called *edges*.

Figure 2 Nodes, edges, and states in one Bayesian belief network



In the simple model shown in Figure 2, the sky is either sunny or cloudy. Whether it is raining or not depends on cloudiness. The grass can be wet or dry, and the sprinkler can be on or off. There is also some causality: If the weather is rainy, it will make the grass wet directly. However, sunny weather can also make the grass wet indirectly, by causing a homeowner to turn on the sprinkler.

A Bayesian belief network is a model that represents the possible states of a given domain. A Bayesian belief network also contains probabilistic relationships among some of the states of the domain. For example, when probabilities are entered into this Bayesian belief network that represent real world weather and sprinkler usage, this belief network can be used to answer questions like the following:

- If the lawn is wet, was it more likely to be caused by rain or by the sprinkler?
- How likely is it that I will have to water my lawn on a cloudy day?

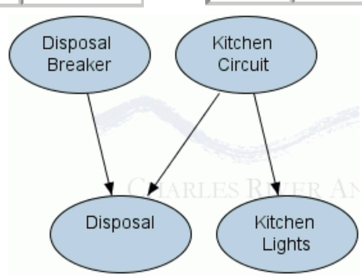
The probability of any node in the Bayesian belief network being in one state or another without current evidence is described using a *conditional probability table*. Probabilities on some nodes are affected by the state of other nodes, depending on causality. Prior information about the relationships among nodes may indicate that the likelihood that a node is in one state is dependent on another node's state.

For example, prior information may show that if it is cloudy, the likelihood of rain is higher; if you have a stress fracture, the likelihood that you will have pain and swelling in your legs is higher; or if the garbage disposal stops working, the likelihood that the cause is a thrown circuit breaker is higher. An example of the conditional probabilities for a Bayesian belief network is shown in Figure 3.

Figure 3 Conditional probability tables in another Bayesian belief network

Disposal Breaker	
Tripped	Not Tripped
0.100	0.900

Kitchen Circuit Breaker	
Tripped	Not Tripped
0.110	0.890



Disposal Breaker	Kitchen Circuit Breaker	Disposal	
		Works	Won't Work
Tripped	Tripped	0.000	1.000
Tripped	Not Tripped	0.000	1.000
Not Tripped	Tripped	0.000	1.000
Not Tripped	Not Tripped	1.000	0.000

Kitchen Circuit Breaker	Kitchen Lights	
	Can turn them on	Cannot turn them on
Tripped	0.000	1.000
Not Tripped	1.000	0.000

With the historical information stored in the conditional probability tables, Bayesian belief networks can be used to help make decisions, or as a way of automating a decision-making process. You can use Bayesian belief networks to perform inductive reasoning (diagnosing a cause given an effect) and deductive reasoning (predicting an effect given a cause).

For more information on using Bayesian belief networks, see the *BNet.Builder User's Guide* and the *BNet.EngineKit Developer's Guide*.

Bayes' Rule

Bayesian belief networks are based on the work of the mathematician and theologian Rev. Thomas Bayes, who worked with conditional probability theory in the late 1700s to discover a basic law of probability, which was then called Bayes' rule.

Most simply, Bayes' rule can be expressed as:

$$P(b|a) = \frac{P(a|b) \times P(b)}{P(a)}$$

where *P(a)* is the probability of *a*, and *P(a|b)* is the probability of *a* given that *b* has occurred.

For example, let's suppose that we know that meningitis can cause a stiff neck about 50% of the time. Suppose we also know from population studies that one in 50,000 people have meningitis, and one in 20 have a stiff neck. We want to know how likely it is that a patient complaining of a stiff neck has meningitis. That is, how probable is meningitis, given a stiff neck?

To figure this out, we compute:

$$P(\text{meningitis}|\text{stiff neck}) = \frac{P(\text{stiff neck}|\text{meningitis}) \times P(\text{meningitis})}{P(\text{stiff neck})}$$

$$= \frac{0.5 \times 1/50{,}000}{1/20} = 0.0002$$

That is, if a patient complains of a stiff neck, the likelihood that it is caused by meningitis is only 0.02%.

A more complex way of expressing Bayes' rule includes a hypothesis, past experience, and evidence:

$$P(H|E, c) = \frac{P(H|c) \times P(E|H,c)}{P(E|c)}$$

where we can update our belief in hypothesis *H* given the additional evidence *E*, and the background context (past experience), *c*.

The left-hand term, *P(H|E,c)* is called the *posterior probability*, or the probability of hypothesis *H* after considering the effect of the evidence *E* on past experience *c*.

The term *P(H|c)* is called the *a-priori probability* of *H* given *c* alone.

The term *P(E|H,c)* is called the *likelihood* and gives the probability of the evidence assuming the hypothesis *H* and the background information *c* is true.

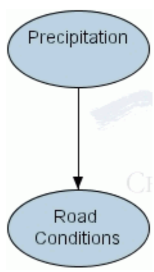
Finally, the last term $P(E|c)$ is independent of H and can be regarded as a normalizing or scaling factor. (Niedermayer, 2003)

Bayesian belief networks capture Bayes' rule in a graphical model. The BNet suite expands the rule to dynamically compute probabilities for large numbers of multi-valued variables as probabilities and evidence are changed. These algorithms quickly become very long and complex to compute by hand, but the BNet suite solves them easily and efficiently.

Structure of a Bayesian belief network

Graphically, Bayesian belief networks are models in which each variable is represented by a *node*, and causal relationships are denoted by an arrow, called an *edge*.

Figure 4 Two nodes and an edge in a very simple Bayesian belief network



Nodes

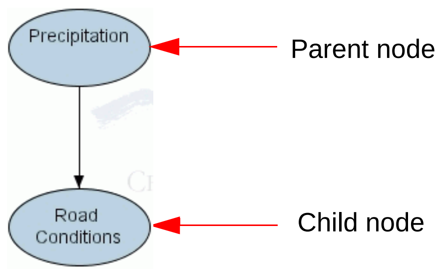
A node represents a variable in the situation being modeled. A node is often represented graphically by a labeled oval. The simple example in Figure 4 shows two nodes, Precipitation and Road Conditions.

Edges

An edge represents a causal relationship between two nodes. It is represented graphically by an arrow between nodes; the direction of the arrow indicates the direction of causality. The intuitive meaning of an edge drawn from node X to node Y is that node X has a direct influence on node Y. For example, in Figure 4, the edge shows that the level of precipitation directly influences road conditions. *How* one node influences another is defined by the conditional probability tables for the nodes.

Edges also determine some qualifying terms for nodes. When two nodes are joined by an edge, the causal node is called the *parent* of the other node. In this example, Precipitation is a parent of Road Conditions, and Road Conditions is the child of Precipitation. Child nodes are conditionally dependent upon their parent nodes.

Figure 5 An edge indicates causality and conditional dependence

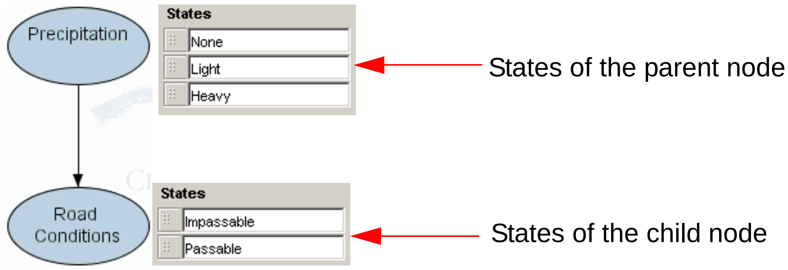


Other terminology you might encounter includes the term *root node* for any node without parents and *leaf node* for any node without children. These terms are more commonly seen in larger belief networks.

States

The values taken on by a variable (represented by a node) are referred to as *states*. For example, the important states of the Precipitation variable are None, Light, and Heavy. We know that precipitation causes a road to be passable or impassible. Those become the states of the Road Conditions node, as shown in Figure 6.

Figure 6 States are values that can be taken on by a node



Conditional probability tables

Every node also has a conditional probability table, or *CPT*, associated with it. Conditional probabilities represent likelihoods based on prior information or past experience.

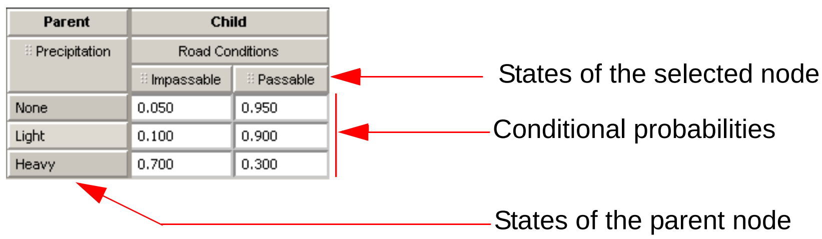
A conditional probability is stated mathematically as $P(x|p1, p2, ..., pn)$, i.e. the probability of variable X in state *x* given parent P1 in state *p1*, parent P2 in state *p2*, ..., and parent Pn in state *pn*.

That is, for each parent and each possible state of that parent, there is a row in the CPT that describes the likelihood that the child node will be in some state. For example, the first cell of the CPT shown in Figure 7 for the Road Conditions node can be read as: “If parent Precipitation is in state None, then the probability that Road Conditions will be in the state Impassable is 5%.” Each cell of the CPT can be read in this way.

One graphical representation of a CPT is in table form as shown in Figure 7. The upper left column header is always labeled Parent, and directly below that are the names of all nodes having causal influence over the node in question. In this case, the node only has one parent, so there is only one column on that side of the table, but there may be more. On the right hand side of the table, the upper right column header always gives the name of the node the CPT is associated with; directly below this the state names for that node are shown. The rest of the table holds the probabilities themselves.

The CPT shown in Figure 7 contains answers to questions about this road such as “Given that precipitation is heavy, what is the probability that the road will be impassable?” That question corresponds to the last row in the table, and the answer is 0.700 or 70%.

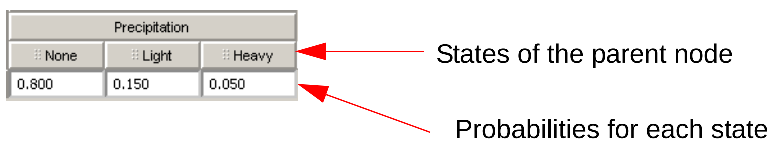
Figure 7 Conditional probability table for a child node



Parent	Child	
Precipitation	Road Conditions	
	Impassable	Passable
None	0.050	0.950
Light	0.100	0.900
Heavy	0.700	0.300

Nodes with no parents also have CPTs, but they are simpler and consist only of the probabilities for each state of the node under consideration.

Figure 8 Conditional probability table for a parent node



Precipitation		
None	Light	Heavy
0.800	0.150	0.050

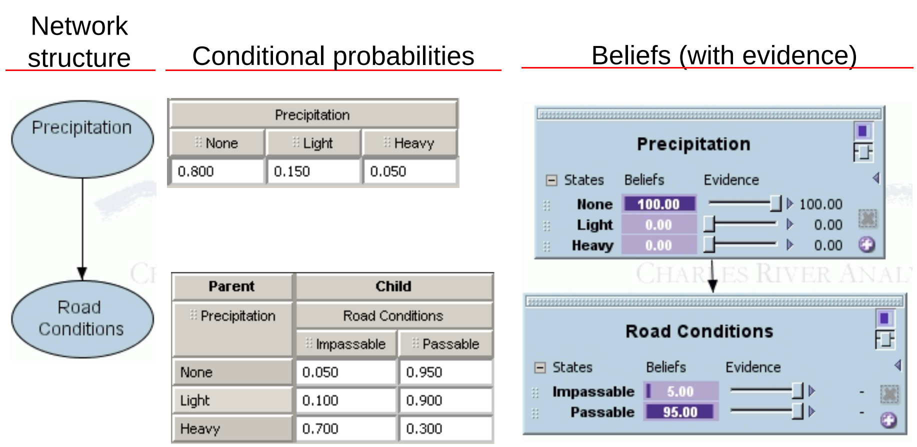
Beliefs and evidence

Beliefs are the probability that a variable will be in a certain state based on the addition of evidence in a current situation. *A-priori beliefs* are a special case of beliefs that are based only on prior information. A-priori beliefs are determined only by the information stored in the belief network’s CPTs.

Evidence is information about a current situation. For example, in the simple Bayesian belief network shown in Figure 9, we have evidence that there is currently no precipitation. The effect of this evidence on current beliefs is reflected in the Beliefs column of the Precipitation node: we are 100% sure that there is no precipitation. Bayesian belief networks support vague, conflicting, and incomplete evidence by allowing you to enter a probability for evidence of a variable being in each state.

The belief network also shows how the evidence changes the current beliefs about the condition of the road: the current beliefs shown in the Beliefs column of the Road Conditions node shows that the likelihood that the road is impassable was reduced from 70% to 5% when we had evidence that there was no precipitation.

Figure 9 Belief network showing the effect of precipitation on road conditions



There are two kinds of evidence:

- *Hard evidence*
  Evidence that a node is 100% in one state, and 0% in all other states. (Hard evidence was posted to the belief network shown in Figure 9.)
- *Soft evidence*
  Any evidence that is not hard evidence. In other words, evidence that a node is less than 100% in one state, and/or greater than 0% in other states. Soft evidence is often used for information about which there is some uncertainty, such as from conflicting reports or an unreliable source.

You can post evidence to a Bayesian belief network and analyze current beliefs to predict a result or to diagnose a cause. For more information on using Bayesian belief networks, see the *BNet.Builder User's Guide* and the *BNet.EngineKit Developer's Guide*.

References

Artificial Intelligence. Stuart J. Russell and Peter Norvig. Pearson Education, Upper Saddle River, NJ. 2003.

An Introduction to Bayesian Networks and Their Contemporary Applications. Daryle Niedermayer. 2003. `http://www.niedermayer.ca/papers/bayesian/bayes.html`.

Web Site

To learn more about our Bayesian Belief Network modeling tool, BNet.Builder, visit www.cra.com/bnet.