Bayesian Reasoning

Andrew Prendergast

October 2014

1 Example Bayesian Networks

Example Bayesian networks will be demonstrated at the start of the session, with links to the Netica (http://www.norsys.com/) and Excel posted online at http://blog.andrewprendergast.com/ afterwards.

- Animal identification example from Netica
- Pathfinder medical diagnosis example from Netica
- Course/grade Intelligence plate model example from PGMs book

2 Practical Example: Bayesian Network

We will use Netica to model the following scenario as a bayesian network:

James has been sneezing today. Construct a Bayesian network to model all of the possible causes of his sneezing. Some of these causes may not be directly observable, so the model must take into consideration additional information that might be observable within James' environment. Once built, use the model to calculate the posterior probability of what might have caused James' sneezing given that we observe different environmental states.

2.1 Building our model

In class we will apply a group-think approach to bayesian analysis:

1. Identify entities that might make useful random variables
2. Arrange their structure according to the Causal Markov Assumption
3. Define states for each random variable
4. Assess conditional probability distributions (CPDs) for each random variable

2.2 Querying our model

Once our Bayesian network is built, we will use it to answer three types of query:

- Diagnosis: effects $\rightarrow$ possible causes (evidential reasoning)
- Prediction: causes $\rightarrow$ effects (causal reasoning)
- Explaining away & intercausal reasoning

For more info see Russell & Norvig Section 14.2. For a definition of the Causal Markov Assumption see Neapolitan & Jiang 2013, Section 6.3.2.2.

3 Practical example: Hidden Markov Model (HMM)

As a second example, lets look at a hidden Markov model implemented in both Excel & Netica, compare the answers and talk through how it would be implemented as a dynamic Bayesian network:

When James isn't overcome by sneezing, he works as a burglar. James has been watching a house every evening and wants to know when the occupants are away on 'vacation'. The only data James has is his observation that the lights are on or off. If James makes the nightly lighting observations of TFFFFFTFTTTTTTTT, use filtering & smoothing to calculate on which evenings it would have been safe enough for James to enter.

For more info see Russell & Norvig Section 15.5 and Figure 15.16.

4 Notation & Useful Axioms

Sets of random variables:

$$P(X_1, \ldots, X_n) \;=\; P(\overline{X})$$

The probability distribution for X (use uppercase for random variables):

$$P(X) \text{ where } X \in \{x_1, \ldots, x_n\}$$

The probability that X is in state x_1 (use lowercase for variable states):

$$P(X = x_1)$$

Bayes rule:

$$P(X \big| Y) \;=\; \frac{P(X, Y)}{P(Y)} \;=\; \frac{P(Y \big| X) P(X)}{P(Y)}$$

The probability chain rule:

$$P(X_1, \ldots, X_n) = P(X_n | X_{n-1}, \ldots, X_1) \cdot P(X_{n-1}, \ldots, X_1)$$

The bayes-net chain rule:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P\left(X_i \,\middle|\, \text{Parents}(X_i)\right)$$

Conditioning:

$$P(X, Y \big| Z) = \frac{P(X, Y, Z)}{P(Z)}$$

Law of total probability:

$$\sum_X P(X) = 1$$

$$\sum_X P(X \big| Y) = 1$$

$$\sum_{X,Y} P(X, Y \big| Z) = 1$$

Marginalisation:

$$P(Y) = \sum_X P(X, Y)$$

$$P(Y \big| Z) = \sum_X P(Y, X \big| Z)$$

$$P(Y) = \sum_X P(Y \big| X) P(X)$$

$$P(Y \big| Z) = \sum_X P(Y \big| X, Z) P(X \big| Z)$$

5 Further Information

If you want to find out more about Probabilistic AI using Bayesian Networks, these books are easy to digest:

- Artificial Intelligence: A Modern Approach (3rd Edition) - Russell & Norvig, 2010. Chapters 13 & 14 (make sure you have the latest edition).
- Probabilistic Methods for Financial and Marketing Informatics - Neapolitan & Jiang, 2007.
- Bayesian Networks in R with Applications in Systems Biology - Nagarajan, Scutari & Lebre, 2013.
- Contemporary Artificial Intelligence - Neapolitan & Jiang, 2013. Work through the examples in Chapter 5 (NB. the same are also in Neapolitan's 2007 book).

6 Top Researchers

The main researchers to follow in this field are:

- Ann Nicholson & Kevin Korb - Based at Monash Uni, wrote the well known Bayesian AI book which includes a clear description of the belief propagation / message passing inference algorithm.
- Judea Pearl - Wrote the Causality book which started the Bayesian revolution.
- Daphne Koller - Teaches AI at Stanford, winner of the Turing Award.
- Richard Neapolitan - Wrote the primary textbooks on the subject.
- Sebastian Thrun - Leads the Robotics Group at Stanford. Originally from CMU.
- Andrew Ng - Leads the Machine Learning Group at Stanford.
- Ron Howard - Directs teaching in decision analysis at Stanford.