Building Bayesian Influence Ontologies
Literature Review

Tamlin Love
1438243

March 14, 2019

1 Introduction

In Section 2, we discuss a variety of classical approaches used to measure similarity. In Section 3, we discuss the concept of a Bayesian Network. In Section 4, we discuss methods used to learn Bayesian network structures and evaluate them. In Section 5, we discuss the application of Bayesian Networks to measuring similarity.

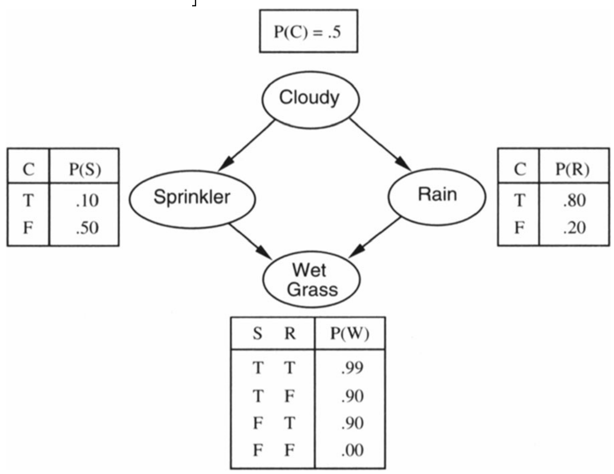
2 Similarity Metrics

3 Bayesian Networks

When considering a joint probability distribution across n random variables, classical probability states that the number of parameters needed to represent the distribution grows exponentially in n [Koller and Friedman 2009]. Even in the simple case of binary variables, we would still need $2^n - 1$ parameters to describe the distribution. This is clearly unfeasible for practical applications, in which the number of random variables can grow very large.

Bayesian networks, originally developed by Pearl [1988], present a way of reducing the number of parameters needed to represent a joint distribution. A Bayesian network is a directed acyclic graph (DAG) whose nodes represent random variables and whose edges represent influence of one variable on another. This structure can also be thought of as a representation of the conditional independencies between the random variables [Koller and Friedman 2009]. Indeed, it is through the exploitation of these independency assumptions that a Bayesian network can more compactly represent a joint distribution.

Figure 1: A famous example of a Bayesian network, showing how a complete representation of any random variable X requires considering only those variables who are parents of X in the graphical representation [Norvig and Russell 1994].



An important notion in Bayesian networks is that of d-separation, first presented by Pearl [1986], which is used to determine whether or not two sets of random variables are conditionally independent given another set. The set $\mathcal{I}(\mathcal{G})$ of conditional independencies in the graph $\mathcal{G}$ is used as the basis for an equivalence relation, I-equivalence, for which any two I-equivalent graphs represent the same independency assumptions [Verma and Pearl 1991].

4 Score-Based Structure Learning

5 Bayesian Similarity

References

[Koller and Friedman 2009] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models - Principles and Techniques.* MIT Press, 2009.

[Norvig and Russell 1994] Peter Norvig and Stuart J. Russell. *Artificial Intelligence: A Modern Approach.* Prentice Hall, December 1994.

[Pearl 1986] Judea Pearl. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, 29(3), September 1986.

[Pearl 1988] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, 1988.

[Verma and Pearl 1991] Thomas Verma and Judea Pearl. Equivalence and synthesis of causal models. In *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence*, UAI '90, pages 255–270, New York, NY, USA, 1991. Elsevier Science Inc.