# Fusion, Propagation, and Structuring in Belief Networks

**Judea Pearl**

Cognitive Systems Laboratory, Computer Science Department,

University of California. Los Angeles, CA 90024, U.S.A

Recommended by Patrick Hayes

### ABSTRACT

Belief networks are directed acyclic graphs in which the nodes represent propositions (or variables), the arcs signify direct dependences between the linked propositions, and the strengths of these dependencies are quantified by conditional probabilities. A network of this sort can be used to represent the generic knowledge of a domain expert, and it turns into a computational architecture if the links are used not merely for storing factual knowledge but also for directing and activating the data flow in the computations which manipulate this knowledge.

The first part of the paper deals with the task of fusing and propagating the impacts of new information through the networks in such a way that, when equilibrium is reached, each proposition will be assigned a measure of belief consistent with the axioms of probability theory. It is shown that if the network is singly connected (e.g. tree-structured), then probabilities can be updated by local propagation in an isomorphic network of parallel and autonomous processors and that the impact of new information can be imparted to all propositions in time proportional to be longest path in the network.

The second part of the paper deals with the problem of finding a tree-structured representation for a collection of probabilistically coupled proposition using auxiliary (dummy) variables, colloquially called “hidden causes.” It is shown that if such a tree-structured representation exists, then it is possible to uniquely uncover the topology of the tree by observing pairwise dependencies among the available propositions (i.e. the leaves of the tree). The entire tree structure, including the strengths of all internal relationships, can be reconstructed in time proportional to n log n, where n is the number of leaves.

### 1.Introduction

This study was motivated by attempts to devise a computational model for humans’ inferential reasoning, namely, the mechanism by which people integrate data from multiple sources and generate a coherent interpretation of that data. Since the knowledge from which inferences are drawn is mostly judgmental-subjective, uncertain and incomplete-a natural place to start would be to cast the reasoning process in the framework of probability theory. However, the mathematician who approaches this task from the vantage point of probability theory may dismiss it as a rather prosaic exercise. For, if one assumes that human knowledge is represented by a joint probability distribution, P(x1,…,xn), on a set of propositional variables , x1,…,xn, the task of drawing inferences form observations amounts to simply computing the probabilities of a small subset, H1,…,Hk, of variables called hypotheses, conditioned upon a group of instantiated variables, e1,…,em, called evidence. Indeed, computing P(H1,…,Hk|e1,…,em) from a given joint distribution on all propositions is merely arithmetic tedium, void of theoretical or conceptual interest.

It is not hard to see that this textbook view of probability theory presents a rather distorted picture of human reasoning and misses its most interesting aspects. Consider, for example, the problem of encoding an arbitrary joint distribution, P(x1,…,xn), on a computer. If we need to deal with n propositions, then to store P(x1,…,xn) explicitly would require a table with 2n entries-an unthinkably large number, by any standard. Moreover, even if we found some economical way of storing P(x1,…,xn) (or rules for generating it), there would still remain the problem of manipulating it to compute the probabilities of propositions which people consider interesting. For example, computing the marginal probability P(xi) would require summing P(x1,…,xn) over all 2n-1 combinations of the remaining