In: Bouchon, B., Yager, R.R. (eds) Uncertainty in Knowledge-Based Systems. IPMU 1986. Lecture Notes in Computer Science, vol 286. Springer, Berlin, Heidelberg, 1987. https://doi.org/10.1007/3-540-18579-8_19

TECHNICAL REPORT
R-50
1986

PROBABILISTIC REASONING USING GRAPHS

Judea Pearl
Cognitive Systems Laboratory
UCLA Computer Science Department, Los Angeles, California 90024

EXTENDED ABSTRACT

Probability theory is shunned by most researchers in Artificial Intelligence. New calculi, claimed to better represent human reasoning under uncertainty, are being invented and reinvented at an ever-increasing rate. A major reason for the emergence of this curious episode has been the objective of making reasoning systems *transparent*, i.e., capable of producing *psychologically meaningful* explanations for the intermediate steps used in deriving the conclusions.

Traditional probability theory, admittedly, has erected cultural barriers against meeting this requirement. For example, scholarly textbooks on probability theory create the impression that to construct an adequate representation of probabilistic knowledge we must start, literally, by defining a *joint distribution* function $P(x_1, x_2, \cdots x_n)$ on all propositions and their combinations, and that this function should serve as the sole basis for all inferred judgments. As a result, even simple tasks such as computing the impact of evidence e on a hypothesis h via $P(h \mid e) = P(h, e) / P(e)$ appear to require a horrendous number of meaningless arithmetic operations, unsupported by familiar mental processes.

Another example is the striking disparity between numerical definitions of independence, e.g., $P(h, e) = P(h)P(e)$, and the ease and conviction with which people distinguish dependencies from independencies, being so unwilling to provide precise numerical estimates of probabilities.

Contrary to this tradition, we argue that a more natural representation of probabilistic knowledge is provided by *dependency graphs*. The nodes in these graphs represent propositions (or variables), and the arcs represent local dependencies among conceptually-related propositions. Graphs permit us to specify dependencies explicitly and qualitatively; and preserve these dependencies despite numerical imprecision. We further argue that the basic steps invoked while people query and update their knowledge, correspond to mental tracings of preestablished links in such graphs, and it is the degree to which an explanation mirrors these tracings that determines whether it is considered "psychologically meaningful."

The first part of the talk will examine what properties of probabilistic models can be captured by graphical representations and will compare the properties of two such representations: Markov Networks and Bayes Networks. A Markov network is an undirected graph where the links represent symmetrical probabilistic dependencies, while a Bayes network is a directed acyclic graph (DAG), where the arrows represent causal influences or object-property relationships. The analysis rests heavily on the theory of GRAPHOIDS, uncovers the logical basis of information dependencies and ties it to vertex-separation conditions in graphs. Given an initial set of such dependencies, the axioms established permit us to infer new dependencies by non-numeric, logical manipulations. Graphs provide an economical language for representing these dependencies and an efficient inference calculus for distinguishing the relevant from the irrelevant.

The second part will introduce a calculus for performing inferences in Bayes Networks. The impact of each new evidence is viewed as a perturbation that propagates through the network via local communication among neighboring concepts. We show that in singly-connected networks such auto-

nomous propagation mechanism can support both predictive and diagnostic inferences, that it is guaranteed to converge in time proportional to the network's diameter and that every proposition is eventually accorded a measure of belief consistent with the axioms of probability theory. This mechanism resolves some long-standing philosophical problems associated with Jeffrey's rule of updating and Polya's "logic of Plausible Inference" and also provides a reasonable model of neural nets performing low level cognition. In multiply-connected networks, clustering and conditioning techniques are available which conduct uncertainty propagation in abstract tree-structured topologies.

In addition to belief updating, the network model also facilitates distributed revision of composite beliefs, i.e., the categorical instantiation of a subset of hypotheses which constitute the most satisfactory explanation of the evidence at hand. We show that, in singly-connected networks, the most satisfactory explanation can be found in linear time by a message-passing algorithm similar to the one used in belief updating. In multiply- connected networks, the problem may be exponentially hard but, if the network is sparse, topological considerations can be used to render the interpretation task tractable. In general, finding the most probable combination of hypotheses is no more complex than computing the degree of belief for any individual hypothesis.

In conclusion, we will show that the current trend of abandoning probability theory as the standard formalism for managing uncertainty is grossly premature -- taking graph propagation as the basis for probabilistic reasoning nullifies most objections against the use of probabilities in reasoning systems. In particular, the graph representation allows us to:

Construct consistent knowledge bases naturally, modularly and incrementally

Distinguish ignorance from uncertainty, and "probable" from "possible" or "plausible"

Distinguish conflicting evidence from uncertain or insufficient evidence

Admit judgmental evidence at any level of abstraction

Trace back the sources of beliefs and produce sound explanations

Optimize the acquisition of data

Answer queries and introduce new constraints

Commit beliefs to the most plausible multi-hypotheses explanation

Learn structures and parameters from empirical data

REFERENCES

1. Pearl, J., ''Fusion, Propagation and Structuring in Belief Networks,'' *Artificial Intelligence,* Vol. 29, No. 3, September 1986, pp. 241-288.

2. Pearl, J., ''How to Do with Probabilities What People Say You Can't,'' *Proceedings,* 2nd Conference on AI Applications, Miami, FL., December 1985, pp. 6-12.

3. Pearl, J. & Paz, A., ''GRAPHOIDS: a Graph-Based Logic for Reasoning about Relevance Relations,'' *Proceedings,* European Conference on Artificial Intelligence-86, Brighton, U.K., June 1986.

4. Pearl, J., ''Distributed Revision of Composite Beliefs,'' *Proceedings,* 2nd AAAI Workshop on

Uncertainty in Artificial Intelligence, Philadelphia, PA., August 1986, pp. 201-209; also to appear in *Artificial Intelligence,* North-Holland, 1987.

5. Pearl, J., ''Bayes and Markov Networks: a Comparison of Two Graphical Representations of Probabilistic Knowledge,'' UCLA Computer Science Department *Technical Report 860024 (R-46),* October 1986.

6. Geffner, H. & Pearl, J., ''A Distributed Diagnosis of Systems with Multiple Faults,'' *Proceedings,* 3rd IEEE Conference on Artificial Intelligence Applications, Orlando, Florida, February 1987, pp. 156-162.

7. Pearl, J., & Verma,T., "The Logic of Representing Dependencies by Directed Graphs," UCLA Cognitive Systems Laboratory *Technical Report 870004 (R-79),* February 1987, *Proceedings,* AAAI Conference, Seattle, WA. July, 1987.