# Notes on Judea Pearl’s Probabilistic Reasoning in Intelligent Systems

D. Gueorguiev 9/20/23

## Notes on Chapter 1 Uncertainty in AI Systems: An Overview

// the whole paragraph is answering the question Why reasoning with exceptions is a minefield

Reasoning about any realistic domain requires that some simplifications be made. The very act of preparing knowledge to support reasoning requires that we leave many facts unknown, unsaid, or crudely summarized. For example, if we choose to encode knowledge and behavior in rules such as “Birds fly” or “Smoke suggests fire” the rules will have many exceptions which we cannot afford to enumerate, and the conditions under which the rules apply (e.g., seeing a bird or smelling smoke) are usually ambiguously defined or difficult to satisfy precisely in real life. Reasoning with exceptions is like navigating minefield -most steps are safe, but some can be devastating. If we know their location, we can avoid or defuse each mine, but suppose we start our journey with a map the size of a postcard, with no room to mark down the exact location of every mine or the way they are wired together. An alternative to the extremes of ignoring or enumerating exceptions is to *summarize* them, i.e., provide some warning signs to indicate which areas of the minefield are more dangerous than others. Summarization is essential if we wish to find a reasonable compromise between safety and speed of movement. We will a language in which summaries of exceptions in the minefield of judgement can be represented and processed. //The last sentence is a nice summary of the purpose of Pearl’s book

1.1.2 Why Is It a Problem?

One way to summarize exceptions is to assign to each proposition a numerical measure of uncertainty and then combine these measures according to uniform syntactic principles, the way truth values are combined in logic.

## Notes on Chapter 3 Markov and Bayesian Networks

From Numerical to Graphical Representation

Widely believed idea is that in order to construct an adequate representation of probabilistic knowledge, we must define a joint distribution function on all propositions and their combinations, where this function serves as a primary basis for all inferred judgements. While useful from mathematical standpoint in facilitation of rigorous mathematical analysis this view on probability theory is totally inadequate for representing and modeling human reasoning.

Consider for example the problem of encoding an arbitrary join distribution for propositional variables. To store explicitly would require a table with entries. Even if we found an economical way of storing – or rules of generating it – there would remain the problem if computing from it the probabilities which are relevant for humans in specific context. For example, computing the marginal probability would require summing over all combinations of the remaining variables. Similarly, computing the conditional probability as:

Would entail dividing two marginal probabilities, each a result of summation over an exponentially large number of variable combinations. Human performance shows the opposite pattern of complexity: probabilistic judgements on a small number of propositions (especially two-component conditional statements such as the likelihood that a patient suffering from a given disease will develop a certain type of complication) are issued swiftly and reliably, while judging the likelihood of a conjunction of propositions entails much difficulty and hesitancy. This suggests that the elementary building blocks of human knowledge are not entries of a joint-distribution table. Rather, they are low-order marginal and conditional probabilities defined over small clusters of propositions.

Another problem with purely numerical representations of probabilistic information is their lack of *psychological meaningfulness*. The numerical representation can produce coherent probability measures for all propositional sentences, but often leads to computations that the human reasoner would not use. // Comment to myself: That makes sense although the term *psychological meaningfulness* is murky. As a result, the process leading from the premises to the conclusions cannot be followed, tested, or justified by the users, or even the designers, of the reasoning system. Even simple tasks such as computing the impact of a piece of evidence on a hypothesis via

require large number of meaningless arithmetic operations, unsupported by familiar mental processes.