# 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.

This approach often yields unpredictable and counterintuitive results. A problem: uncertainty measures stand for something totally different than truth values. Whereas truth values in logic characterize the formulas under discussion, uncertainty measures characterize invisible facts, i.e., exceptions not covered in the formulas. Accordingly, while the syntax of the formula is a perfect guide for combining the visibles, it is nearly useless when it comes to combining the invisibles. For example, the machinery of Boolean algebra gives us no clue as to how the exceptions to interact with those of to yield the exceptions to . These exceptions may interact in an intricate and clandestine ways, robbing us of the modularity and monotonicity that make classical logic computationally attractive.

The interactions in logic expressions are visible and we can calculate the impact of each new fact in stages by a process of derivation that resembles the propagation of wave // this could also be a reference to the backprop algorithm We compute the impact of a new fact on a set of syntactically related sentences , store the results, then propagate the impact from to another set of sentences , and so on, without having to return to . The problem is that this divide-and-conquer approach so basic to logical deduction cannot be justified under uncertainty unless one makes some restrictive assumptions on *independence*.

Another feature we lose in going from logic to uncertainty is *incrementality* (that is, the availability of *recursive algorithms*). When we have several items of evidence, we would like to account for the impact of each of them individually: compute the effect of the first item, then absorb the added impact of the next item, and so on.

1.1.3 Approaches to Uncertainty

Three formal schools: *logicist*, *neo-calculist*, *neo-probabilist*.

This taxonomy is superficial. The taxonomy which captures the semantic variations – *extensional* vs *intensional* approaches. Extensional approach other names – rule-based systems, procedure-based systems. It treats uncertainty as a generalized truth value attached to formulas and computes the uncertainty of any formula as a function of the uncertainties of its subformulas. Intensional approach a.k.a declarative/model-based uncertainty exists in the context of “states of affairs” or subsets of “possible worlds”. Extensional systems are computationally efficient but semantically sloppy, while intensional systems are semantically clear but computationally clumsy. // This computational clumsiness is obsolete – my bets are on the intensional (model-based) systems in the modern days.

1.1.4 Extensional vs Intentional Approaches

Extensional approach: Certainty factors calculus // does anybody use it today, I don’t think so

Example: the certainty of the conjunction is given by some function (e.g., the minimum or the product) of the certainty measures assigned to A and B individually.

Intensional approach: Probability theory

Certainty measures are assigned to sets of worlds and the connectives combine sets of worlds by set theory operations. For example: the probability is given by the weight assigned to the intersection of two sets of worlds – those in which is true and those in which is true – but cannot be determined from the individual probabilities and .

Rules in extensional systems provide licenses for certain symbolic activities.

Example: could mean “If you see A, then you are given the license to update the certainty of B by certain amount which is a function of the rule strength ”. The rules are interpreted as a summary of past performance of the problem solver, describing the way an agent normally reacts to problem situations or to items of evidence.

Rules in intensional systems denote elastic constraints about the world.

Example: In [Dempster-Shafer formalism](https://en.wikipedia.org/wiki/Dempster%E2%80%93Shafer_theory) the rule does not describe how the agent reacts on finding A, but asserts that the set of worlds in which and simultaneously hold has low likelihood and hence should be excluded with probability .

## 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.