COMMENTS ON ROMEIJN *et al.*'S "PROBABILISTIC LOGICS AND PROBABILISTIC NETWORKS"

Jonathan Weisberg
University of Toronto

Three Core Claims

I. Probabilistic logics are underused; because they are disparate, complicated, and computationally unfriendly.

II. But several probabilistic logics can be brought under the unifying umbrella of a single framework.

III. Within that framework, Bayesian and credal networks offer a way to make these systems computationally tractable.

A Critical Perspective

Is the project outlined in 1-3 well-founded? I think it tries to force things into an unhelpful and infelicitous framework, one that does not fit, and obscures the natural shape of things:

A. The systems JW wants to unify in a single framework don't really fit that framework; a traditional taxonomy makes more sense.

B. Probabilistic logics don't seem to be 'underused' in the way JW suggests.

C. Nets make no contribution to probabilistic *logic* as such; discussing them in this context misleadingly suggests a special role for logic in applications.

The Very Notion of Probabilistic Logic

Two ways to put the probability into probabilistic logic:

1. *Probabilize the Language.* Take an existing logic, like propositional logic, and augment the language to include probabilistic locutions.

   ▷ Analogy: creating modal logic by adding modal locutions.

2. *Probabilize the Consequence Relation.* Allow $\vDash$ to admit of degrees, so that logical entailment comes in various strengths, depending on the probabilistic relationships. Or keep it qualitative, but still dependent on probabilistic relationships.

   ▷ Analogy: creating default logic by re-conceptualizing entailment.

ILLUSTRATIONS

1. *Probabilizing a Language*:
   - ▷ Add to the language of propositional logic the function symbol $p(\ )$, $=$, and names for the reals in $[0, 1]$.
   - ▷ Augment the semantics, adding probability functions to the models.
   - ▷ Now we can say and evaluate things like,

$$p(A) = .3 \ \models \ p(\neg A) = .7$$

   Written more compactly,

$$A^{.3} \ \models \ \neg A^{.7}$$

2. *Probabilizing Consequence*:

- ▷ Replace the standard consequence relation of propositional logic with a quantitative version, $\models^x$.
  - ○ Define $\models^x$ using a privileged probability function p:

$$\Gamma \models^x \psi \quad \textit{iff} \quad p(\psi|\Gamma) = x$$

- ▷ Or, use p to define a qualitative yet probabilistic $\models$:

$$\Gamma \models \psi \quad \textit{iff} \quad p(\psi|\Gamma) \geq .95$$

EQUATION (1) AS A UNIFYING FRAMEWORK

JW is taking the first approach, probabilizing the language.

▷ His unifying framework is his Equation (1):

$$\varphi_1^{X_1}, \varphi_2^{X_2}, \ldots, \varphi_n^{X_n} \models \psi^Y$$

Here the language is modified to allow the expression of probability assignments. Note two things about this framework:

1. Formal: the central question of progic in the Equation (1) framework is, "given the probability assignments on the φ's, what assignment follows for ψ?"
2. Semantic: $\models$ retains a standard definition, where entailment holds iff all models of the LHS are models of the RHS.

Are these features of the framework really universal for progics?

EVIDENTIAL PROBABILITY AND EQUATION (1)

- ▷ Evidential Probability uses a set of frequency data to assign probabilities to single-case propositions.
  - ○ Given that 1% of people know the capital of Algeria, the EP that Albert knows is .01: $\Gamma \models^{.01} \psi$
  - ○ Assume further that Albert is a geography PhD, and 99% of geography PhDs know. Then the EP is .99: $\Gamma^{+} \models^{.99} \psi$
- ▷ We have two mismatches with the Equation (1) framework:
  1. Formal: wrong kind of question (wrong relata and relation).
  2. Semantic: wong kind of entailment (non-monotonic).
- ▷ Talk of second-order probabilities just changes the subject.

Bayesian Inference and Equation (1)

- Given priors and evidence, Bayesian inference infers a posterior:

$$\varphi_1^{X_1}, \ldots, \varphi_n^{X_n}, \chi_1, \ldots, \chi_m \models \psi^Y$$

This looks more Equation (1)-ish, but …

  - Semantic mismatch again: this $\models$ is non-monotonic.

- A better representation is actually:

$$\varphi_1^{X_1}, \ldots, \varphi_n^{X_n}, \chi_1, \ldots, \chi_m \models^Y \psi$$

but that just makes the mismatch more patent.

- ▷ JW prefers to represent Bayesian inference along the lines

$$\varphi_1^{X_1}, \ldots, \varphi_n^{X_n} \models (\psi | \chi_1, \ldots, \chi_m)^Y$$

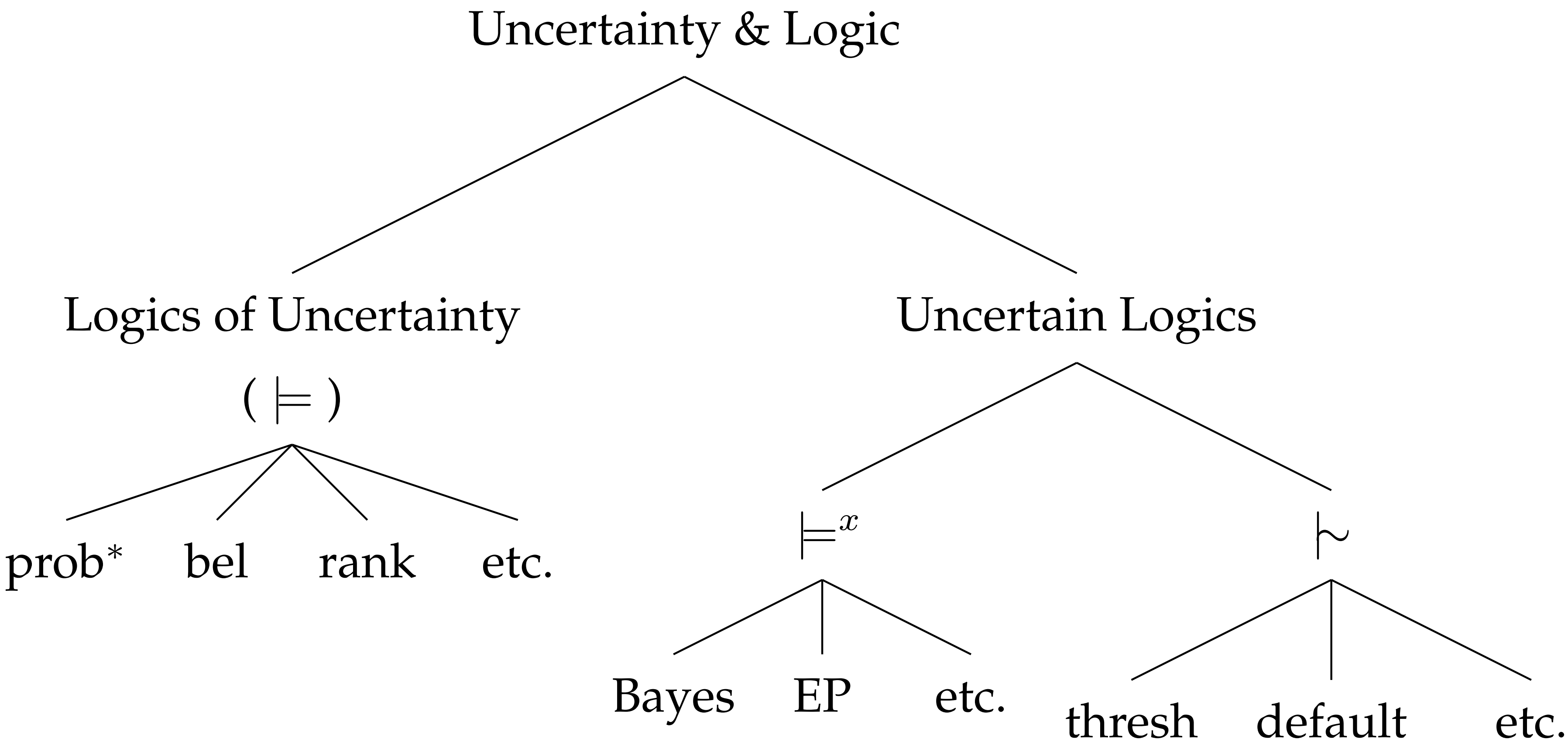
But that's just the standard probability logic in disguise:

$$\varphi_1^{X_1}, \ldots, \varphi_n^{X_n} \models p(\psi, \chi_1, \ldots, \chi_m)/p(\chi_1, \ldots, \chi_m) = Y$$

- ▷ Recall the notorious pitfall we emphatically flag for students:
  1. Conditional probability is a *locution*.
  2. Conditionalization is a *principle of inference*.
- ▷ Fitting Bayesian inference into the Equation (1) framework takes all the juice out of the theory.
  - ◦ No longer a theory of Bayesian *inference*; just probability theory with conditional probabilities.

What Moral to Draw?

- These systems don't fit the framework because they are systems of uncertain inference, not systems for reasoning about uncertainty.
- Come to think of it, the framework isn't even essentially probabilistic; it's a framework for reasoning about uncertainty, probabilistic or otherwise.
  - Eqn. (1) fits the logic of Dempster-Shafer theory just as well.
- Better to use a traditional taxonomy then:



*progics, in the sense of Equation (1) with probabilistic models.

PROGICS UNDERUSED?

- ▷ Why make a logic of probability? Why bother formalizing probability theory?
- ▷ Standard reasons for formalizing a mathematical theory:
  1. Explore its logical properties (axiomatizations, decidability, etc.).
  2. Program computers to use the theory, *especially syntactically*.
- ▷ Worthy pursuits, but no contributions here (*cf*. Halpern *et al.*).
- ▷ Also, not the applications JW is after — how are his applications helped by logical formalization?
  - ○ Nets may be useful, but that's a separate matter.
- ▷ If the crux is just that nets can help us solve constraints problems, what are we doing talking about logic?