A note on probabilistic logics and probabilistic networks

Jon Williamson

Draft of April 22, 2008

In a classical logic we are typically faced with the following kind of question: do some given premisses $\varphi_1, \ldots, \varphi_n$ entail a given conclusion ψ? This question can be written

$$\varphi_1, \ldots, \varphi_n \approx\!\!\!\mid\; \psi?$$

Here $\varphi_1, \ldots, \varphi_n, \psi$ are premisses of some formal language, such as a propositional language or a predicate language. $\approx\!\!\!\mid$ is an *entailment* relation: the entailment holds if all models of the premisses also satisfy the conclusion, where the logic provides some suitable notion of 'model' and 'satisfy'. Proof theory is normally invoked to answer a question of this form: one tries to prove the conclusion from the premisses in a finite sequence of steps, where at each step one invokes an axiom or applies a rule of inference.

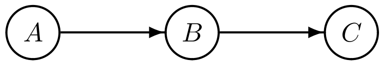
Probabilistic logics come in various guises but we shall look at logics where probabilities attach to sentences. Thus $\forall x(Ux \rightarrow Vx)^{0.8}$ says that the probability that all Us are Vs is 0.8. Probabilistic logics have great potential in any application in which logical or structural constraints operate, but where they only operate in a certain proportion of cases or where the constraints are uncertain—e.g., in inferring meaning in natural language, predicting protein folding in biology or modelling scientific theory change.

In a probabilistic logic, the fundamental question takes a different form to that of classical logic. While we might have premisses of the form $\varphi_1^{X_1}, \ldots, \varphi_n^{X_n}$ where $X_1, \ldots, X_n$ are probabilities or sets of probabilities, it is rare that we are presented with a conclusion of the form ψ^Y and asked whether the conclusion follows from the premisses. More typically, there is a conclusion proposition ψ of interest and we want to know what probability or set of probabilities Y to attach to ψ. Thus the fundamental question of probabilistic logic can be written

$$\varphi_1^{X_1}, \ldots, \varphi_n^{X_n} \approx\!\!\!\mid\; \psi^?$$

Since this question differs from that of classical logic, one might anticipate that the means to solve the question differ too. In fact, determining Y is essentially a question about probability, so methods of probabilistic inference are more appropriate than the standard notion of proof. In (2008: *Probabilistic logics and probabilistic networks*, available here) Rolf Haenni, Jan-Willem Romeijn, Gregory Wheeler and I explore the use of *probabilistic networks* to answer this question.

A *Bayesian network*—the simplest kind of probabilistic network—consists of a directed acyclic graph on a finite set of variables, together with the probability distribution of each variable conditional on its parents in the graph. For example, given propositional variables A, B and C, the following constitutes a Bayesian network:



$P(A) = 0.7$	$P(B\|A) = 0.2$	$P(C\|B) = 0.9$
	$P(B\|\neg A) = 0.1$	$P(C\|\neg B) = 0.4$

By assuming what is called the *Markov Condition*, which says that each variable is probabilistically independent of its non-descendants in the graph conditional on its parents, a Bayesian network suffices to determine the joint probability distribution over the set of variables. In our example the Markov condition says that C is independent of A conditional on B. Probabilities over the whole set of variables are multiples of corresponding conditional probabilities: $P(A \wedge \neg B \wedge C) = P(C|\neg B)P(\neg B|A)P(A)$.

A *credal network* is just like a Bayesian network except that the conditional probabilities are only identified to within closed intervals. Thus the above graph together with the following constraints determines a credal network:

$P(A) \in [0.7, 0.8]$	$P(B\|A) = 0.2$
	$P(B\|\neg A) \in [0.1, 1]$

$P(C\|B) \in [0.9, 1]$
$P(C\|\neg B) \in [0.4, 0.45]$

While a Bayesian network represents a single probability function, a credal network represents a convex set of probability functions. A wide variety of algorithms have been developed for constructing probabilistic networks and for calculating probabilities from them. The use of probabilistic networks can greatly reduce the computational burden of probabilistic inference: broadly speaking, the sparser the graph, the quicker it is to draw inferences.

In the context of probabilistic logic it turns out that for a range of natural semantics the $X_1, \ldots, X_n$ are normally probabilities or intervals of probabilities, and consequently the premisses determine a convex set of probability functions. A probabilistic network can be used to represent that set of functions and to infer an appropriate set Y of probabilities to attach to the conclusion sentence. While the probabilistic network itself depends on the chosen semantics, the machinery for calculating Y does not—see our (2008: §8.2).

A surprisingly broad range of approaches to probabilistic inference can be invoked to provide semantics for probabilistic logics. Under the *standard probabilistic semantics* a probability function P satisfies φ^X iff $P(\varphi) \in X$; premisses entail a conclusion iff all probability functions that satisfy the premisses also satisfy the conclusion. According to *probabilistic argumentation*, Y is the probability of worlds for which the premisses *force* the conclusion sentence ψ to be true. With *evidential probability*, the φ_i include statistical statements, ψ is inferred by certain rules for manipulating reference classes, and Y quantifies the level of risk associated with this inference. *Classical statistics* can also be used to provide a semantics since probabilistic argumentation or evidential probability can capture fiducial probability. According to *Bayesian statistical inference* the premisses contain information about prior probabilities and likelihoods, and the entailment holds if the conclusion follows by Bayes' theorem. With *objective Bayesian epistemology*, the entailment holds if any agent with evidence characterised by the premisses should believe ψ to degree within Y.

In sum, probabilistic logics admit a range of natural semantics and probabilistic networks can be used to answer the queries that such a logic faces. The interested reader is urged to consult our (2008) for more details.