A Natural Generalization of Jeffrey Conditioning

The Story

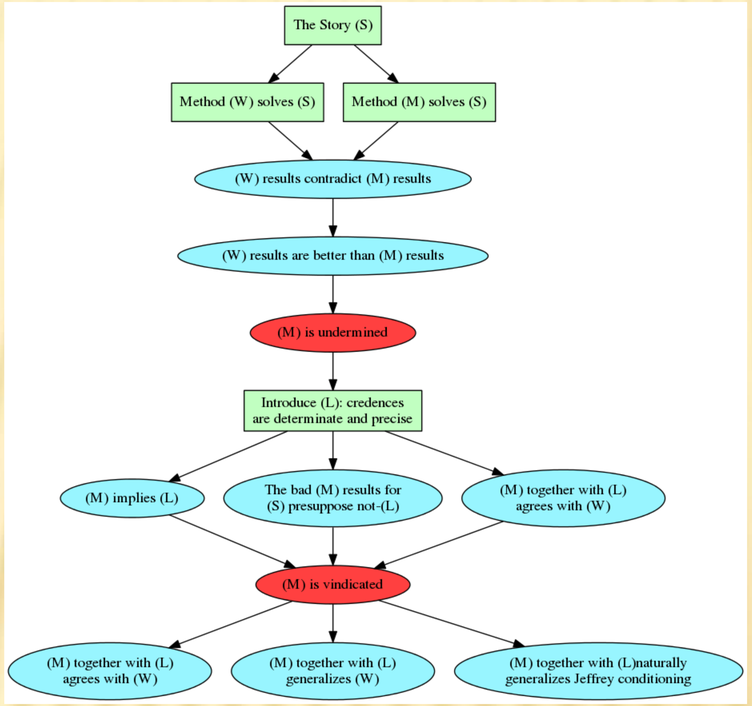
You encounter the native of a certain foreign country and wonder whether he is a Catholic northerner, a Catholic southerner, a Protestant northerner, or a Protestant southerner. Your prior probability over these possibilities (based, say, on population statistics and the judgment that it is reasonable to regard this individual as a random representative of his country) is (0.2,0.3,0.4,0.1), respectively. The individual now utters a phrase in his native tongue which, due to the aural similarity of the phrases in question, might be a traditional Catholic piety, an epithet uncomplimentary to Protestants, an innocuous southern regionalism, or a slang expression used throughout the country in question. After reflecting on the matter you assign subjective probabilities (0.4,0.3,0.2,0.1), respectively, to these alternatives. In the light of this new evidence how should you revise your prior probabilties for the identity of the native?

The Laplacean Principle

Bayesians used to subscribe to the claim that sharp credences can be elicited from rational agents with respect to well-defined events. This is no longer true. There are Bayesians who advocate that the credences of a rational agent may be imprecise and/or indeterminate, considering that evidence may be incomplete or ambiguous. Let **(L)** be the Laplacean principle, which adds to the usual Bayesian convictions that the credences of a rational agent are always precise and determinate (or sharp).

Carl Wagner's Solution

Carl Wagner develops a generalization of Jeffrey Conditioning which can handle our story. The generalization is based on a plausible intuition: leave the ratio of probabilities alone unless you have a reason to change them in the updating process, given your new evidence. Call this generalization **(W)**.



The Problem: MAXENT undermined

Another plausible intuition suggests that we should gain as little information as possible in the updating process. This intuition leads to an updating method called *Infomin*. It is a relative of MAXENT, the principle of maximum entropy. We will call this method **(M)**. Wagner shows that **(W)** contradicts **(M)** with its results for our story and that **(W)**'s results are much more plausible. This is a problem for advocates of **(M)** as a method that is supposed to work in general.

Remark about the Relationship between (M) and (L)

(L) Is currently highly controversial in the literature. My assumption is that although many Bayesians reject **(L)**, no advocates of **(M)** do. The reason is that advocates of **(M)** tend to be more "strict" in their interpretation of Bayesianism and fall more naturally on the side of sharp credences. I think it would be difficult to articulate **(M)** for mushy credences, since a measure of information would have to be defined on sets of permissible probability distributions. Even if this were to succeed, my paper could be used to show that, without intending to, Wagner has conclusively demonstrated that **(M)** and non-**(L)** are incompatible.

The Conclusion

My paper shows that formally Wagner's undermining of **(M)** is only effective if non-**(L)** is assumed. On any reasonable interpretation of **(M)**, **(M)** implies **(L)**, so undermining **(M)** by assuming non-**(L)** is not going to be persuasive. **(M)** is vindicated and offers a much more integrated natural generalization of Jeffrey conditioning which agrees with all of Wagner's intuitions about keeping ratios stable as much as the evidence allows.