Unanticipated possibilities

Marcello/Rafal

1 The problem

1.1 Basic idea

How does the Bayesian framework deal with the problem of unanticipated possibilities?

- There seems to be two sides to this problem:

  a. The space of possibilities could get larger, say, first one considers only two suspects, Michael and Nero, and then a third suspect, Tito, becomes plausible. The priors of M and N were initially 50:50 (or 75:25, whatever), and T got nothing. Actually, T was not in the algebra to begin with. But, now T is in the algebra and gets a non-zero probability. How can this be handled in the Bayesian framework?

  b. The space of possibilities become more fine-grained. Consider a DNA evidence case. At first, one of the hypothesis is that the defendant is the source of the crime stain. But as the investigation proceeds, it becomes clear that it makes a difference whether the defendant is the source of the crime stains on the pillow or on the sofa. Perhaps, the stain on the pillow were left innocently, but the stains on sofa were left by the perpetrator. So the space of possibilities becomes more fine-grained.

- Not sure if the two items above are actually different, but they seem to be. In the first, the possibilities expand. In the second, the possibilities become more fine-grained. The first can be represented by adding a possible world, while the second, by making the equivalence classes among possible world more fine grained.

- We need to be careful about what we are trying to represent here. Say we are trying to represent what is going on in the jurors' or judges' minds as they hear evidence in the trial and they process the evidence. We are trying to give a plausible picture of the processing of the evidence and hypotheses at stake.

- Are we trying to give an empirically adequate picture or just a philosophically or formal plausible picture?

1.2 Relevant literature

1.2.1 Ronald Allen

Quotations from Allen's The Nature of Juridical Proof: Probability as a Tool in Plausible Reasoning:

> The decision-makers in all legal systems must learn about the issue being litigated, whether this is through serial hearings as in some continental systems or concentrated trials in the American tradition ... in no system does the fact-finder have substantial knowledge about the case until evidence is presented. Obviously. That means, though, that it is literally impossible to form a probability space of mutually exclusive hypotheses, to assign them initial probabilities. (p. 138)

Some of you may think that I overstate the case and that the mutually exclusive hypotheses are obvious: guilt or innocence or liability or no liability ... Guilt, innocence, liability and no liability are not factual hypotheses; they are legal conclusions that are applied to characterise the implications of the facts that are found. It should be obvious that one cannot condition 'guilt' on some evidence without going through a factual proposition that is the actual object of proof. (p. 138)

Complicating matters further is the **emergence of new theories** ... Like society itself, the litigation process is dynamic, not static, and frequently new theories will emerge. When they do, the probability space must be reconfigured. It is thus pointless to configure the probability space even if it could be done at some point prior to receiving all the evidence and hearing all the arguments. (p. 139)

And finally, the **spectre of Old Evidence** raises its head. After the theories of the parties are finally identified and the logical space defined, initial assignments of probability must be made in order for Bayes' Theorem to be employed. But those initial assignments of probability will obviously include consideration of the trial evidence, as well as pre-existing knowledge of the fact-finder. Once the probability space is determined and initial probabilities assigned, all the trial evidence is old evidence that has already been accommodated. There is no work for Bayes' Theorem to do. (p. 139)

What, then, are the **tasks for the future**? ... For the legal analysts, there are two issues that should preoccupy us going forward. The first is the perennial problem of constructing methods to **decrease the probability that unreliable evidence is being provided** by purported experts and scientists ... that means engaging with the underlying forensic science to ensure that appropriate methodologies are employed and that the evidence offered at trial does not overleap what those methodologies can deliver. The second is the more fundamental problem noted above of **integrating that evidence with the rest of the evidence** in the case ... What that means in turn, of course, is to flesh out the nature of rational thought ... The discrete question to be addressed is the implications of rational thought writ large for the legal system. (p. 141)

1.2.2 Chihara

2 Approaches / solutions

2.1 Possible worlds / equivalence classes

- We start with all possible worlds S. This is the entirety of the logical space. Each world is a complete description of how things could have been.
    - Language approach. Each world is a maximally consistent set of sentences. The language here should be assumed the most informative language possible, or at least, a very expressive language, say English. Anything that cannot be expressed in that language is left out and will never be under consideration.
    - Model theoretic approach. Note sure.
- At some point during trial, a common ground $\mathscr{G} = (G, \sim)$ is formed—it becomes clear what is at issue, what questions are being asked, what is not at issue, what is disputed, what is agreed upon, etc. This common ground, formally, has two dimensions:
    a. some possible worlds are removed, i.e. all possible worlds that make true propositions that are demonstrable false, assumed to be false or simply not considered at

issue, so we should only focus on a subset of *S*, that is, $G \subseteq S$

    b. an equivalence class $\sim$ is induced over the remaining set *G* of possible worlds. This equivalence class fixes how fine grained the discussion is going to be (at the level of source/not-sourse or pillow/sofa, etc.). Worlds that are in the same equivalence classes are indistinguishable and thus treated as equivalent. Notation: let $[G]_\sim$ be the set of all equivalence classes induced by $\sim$ in *G*.

- The probability measure P() applies within the common ground as defined in points (a) and (b) above. So if $[G]_\sim$ is the set of all equivalence classes induced by $\sim$ in *G*, any probability measure P() is a function from all subsets of $[G]_\sim$ into $[0,1]$, that is, $\mathsf{P}() : \wp([G]_\sim) \to [0,1]$.

- Two side notes:

  – The Bayesian updating, in the traditional sense, occurs within the common ground $\mathscr{G} = (G, \sim)$. Presumably, since the common ground is fairly limited in terms of how much is at under dispute—the litigant at trial cannot disc ussu everything—the application os probability to these limited domain is not unfeasible. To be discussed further. This might address Allen's complexity concerns about Bayesianism.

  – Suppose what is being disputed is just M (=Michael did it) versus N (=Nero did it). Then, M and N form an exhaustive and exclusive pair. Which means that $\mathsf{P}(E|M)/\mathsf{P}(E\ vertN)$ is the same same thing as $\mathsf{P}(E|M)/\mathsf{P}(E\ vert\neg M)$. This is true, however, relative to a fixed common ground. It is not true in general since *M* and *N* dod nt cover teh entire logical space *S*.

- Unanticipated possibilities may arise (a) either because possible worlds that were taught to be false, assumed to be false or simply excluded become relevant—for example, when Tito becomes a suspect besides Nero and Michael—(b) or because the equivalence class become more fine-grained—for example, when the focus become pillow/sofa and not simply source/non-source.

- There seems to be two dynamic updates going on here:

      a. dynamic update within a fixed common ground, this the classical Bayesian update

      b. meta-level dynamic update about what the common ground

- Both updates are plausible. For judges and jurors may be asked (a) to assess evidence relative to a fixed common ground or (b) assess the evidence relative to a new common ground (say, new sets of hypothesis).

- The formal framework should be able to go back and forth between the two level (simply Bayesian updates within the common ground and meta-level update about the common ground itself). There should bridge rule or constraints that allow to go from level to another without loss of information.

  – What we want to avoid is that, once a new common ground is formed, the evidence assessment is redone completely afresh from scratch. The epistemic work done previously in asessing evidence relative to the old common ground shouod be retained if possible.

  – Note that we speak as though the evidence is kept the same in going from one common ground to the next. This makes sense intuitively. How can we make plausible sense of this sameness of evidence across common grounds? Strictly speaking the evidence is not the same. Since the evidence is a propositions of some kind (=set of possible world), any change in common ground will trigger a change in the evidence as a proposition. So we need to specify conditions by which the evidence stays the same.

- One question here is whether we should do this syntactically or model-theoretically. The

problem of doing it model-theoretically are things like logical omniscience.

2.2 Bayesian network updating and comparison

- The Bayesian network need not be incompatible with the possible world approach. The two might be complementary.
- Rafal identified two dynamic processes:

    a. Once a Bayesian network is constructed, and all probabilities are assigned, Bayesian updating takes place within the network.

    b. A meta-level updating that may consist in (b1) refining the Bayesian network itself by adding arrows and nodes or (b2) comparing a Bayesian network to another Bayesian network.

- How does the meta-level network updating (refining of a network or comparison across networks) relate to the meta-level updating of the possible world approach outlined earlier?