Bayesian Networks in Epistemology and Philosophy of Science

Lecture 3: Applications in Epistemology

Stephan Hartmann

Center for Logic and Philosophy of Science
Tilburg University, The Netherlands

Formal Epistemology Course
Northern Institute of Philosophy
Aberdeen, June 2010

Plan

Bayesian Epistemology is a well-established sub-branch of Formal Epistemology. It is also taken seriously by mainstream epistemologist as the inclusion of survey articles in standard encyclopedias shows:

1. Hájek, A. and S. Hartmann: Bayesian Epistemology. In: J. Dancy, E. Sosa, and M. Steup (eds.), *Blackwell Companion to Epistemology*. Oxford: Blackwell ^{2}2010, 93-106.
2. Hartmann, S. and J. Sprenger: Bayesian Epistemology, to appear in: S. Bernecker and D. Pritchard (eds.), *Routledge Companion to Epistemology*. London: Routledge 2010.

While much work in Bayesian Epistemology can be done without Bayesian Networks, they are of much help for many problems.

My main meta-philosophical claim in this third lecture will be that formal epistemology can gain considerably from "going empirical" and from connecting to the psychology of reasoning.

Overview

Lecture 1: Bayesian Networks

1. Probability Theory
2. Bayesian Networks
3. Partially Reliable Sources

Lecture 2: Applications in Philosophy of Science

1. A Survey
2. Intertheoretic Reduction
3. Open Problems

Lecture 3: Applications in Epistemology

1. A Survey
2. Bayesianism Meets the Psychology of Reasoning
3. Open Problems

Applications of Bayesian Networks in Epistemology

In lecture 2 I mentioned that Bayesian Networks are also applied in philosophy of science without subscribing to Bayesianism. The debate about causal discovery is a case in point.

In epistemology, all applications so far are – to the best of my knowledge – in a Bayesian setting.

The three most important areas of application are:

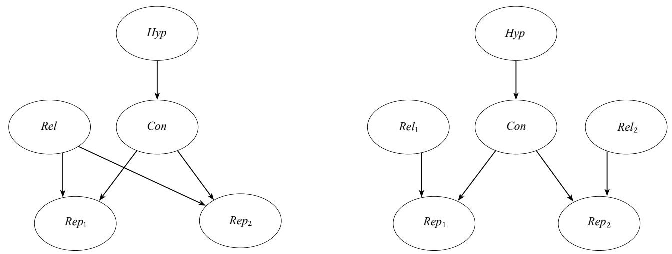
1. Confirmation
2. Testimony
3. Coherence

1. Confirmation

- What does it mean that a piece of evidence E confirms a hypothesis H?
- Qualitative accounts face various problems (the raven paradox, tacking), which motivates a quantitative account.
- A quantitative account is also psychologically motivated as we are able to judge (at least) the relative strength of evidence.
- Example: Sir Henry was killed in his castle; H: Butler James is the murderer. E_1: The bloody knife was found in James' room. E_2: Henry had an affair with James' wife.
  Clearly, $E_1 \wedge E_2$ confirms H more than E_1 alone.
- As philosophers, we should make sense out of this practice.

Extension 1

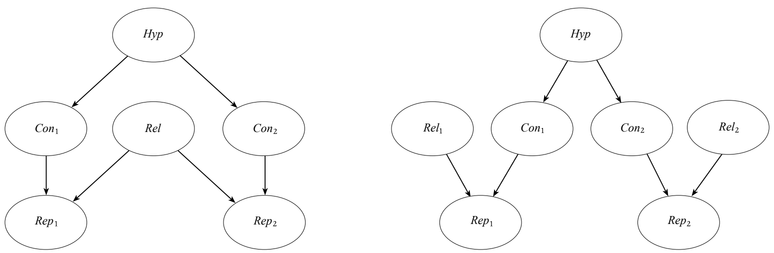
With the help of Bayesian Networks, the following situations, involving partially reliable measuring instruments, can be analyzed:
1. The consequence of a hypothesis is tested, once with dependent measurement instruments, once with independent measurement instruments.



Which testing scenario is epistemically advantageous?
Answer: It depends. . .

Extension 2

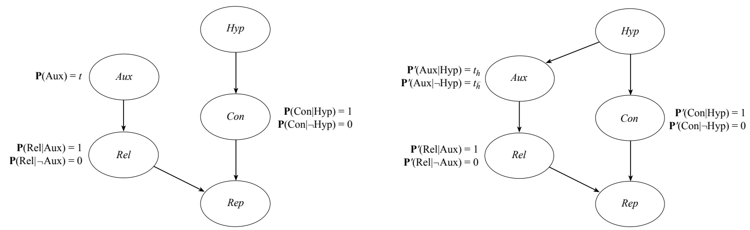
2. A hypothesis has several testable consequences. These are tested, once with dependent measurement instruments, once with independent measurement instruments.



Which testing scenario is epistemically advantageous?
Answer: It depends. . .

Extension 3

3. A hypothesis is tested, but to do so, an auxiliary assumption has to be made. This assumption may or may not depend on the hypothesis under test.



Which testing scenario is epistemically advantageous?
Answer: It depends. . .

2. Testimony

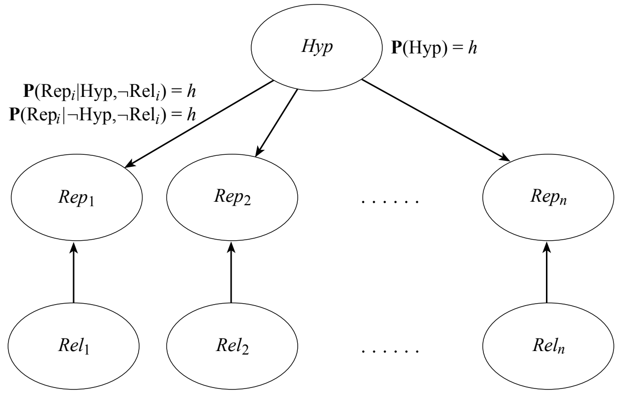
Most of what we know comes from testimony (our parents and teachers, books, etc.). Are we justified to do so?

This problem is especially important as many sources are at best partially reliable.

Specific questions:

1. Which role do confirming testimonies (by more or less dependent sources/witnesses) play?
2. Which role does coherence play?

Low Priors, High Posteriors



3. Coherence

- According to the Coherence Theory of Justification, a set of propositions is justified if it coheres well.
- But what does "coherence" mean? Answers using phrases such as "hanging together well", "dovetail" etc are not really helpful, and so the theory is (at best) not clearly stated.
- It's also hard to see how, in an informal way, the coherence-truth link can be studied. While it is clear that coherence is in general not truth-conducive (example: fairy tales), it may well be truth-conducive under certain conditions. But what are these conditions?
- There are several Bayesian attempts to formulate a coherence measure and to explore the philosophical consequences of these measures. Coming soon...

Challenges for Bayesian Epistemology

- Bayesian Epistemology has been criticized on several grounds. Some of these challenges are not fully addressed yet.
- However, in the light of its enormous explanatory success, I suggest to *not* take these challenges as refutations of Bayesianism.
- Instead, I recommend treating Bayesian Epistemology like a scientific theory, with its strengths, limitations and time-dependent domain of applicability.
- But this does not mean that there aren't any problems that deserve attention. I focus on two of them:
  1. The Descriptive-Adequacy Problem
  2. The Plurality-of-Measures Problem

The Descriptive-Adequacy Problem

- Arguably no scientist or ordinary person proceeds in the way Bayesians recommend it.
- We are also not able to do so. After all, who can handle a probability distribution over a large set of variables?
- Standard responses: (i) Bayesianism is a normative theory, (ii) it must (at least) be possible to provide a Bayesian *reconstruction* of how people are reasoning.
- These replies are not satisfactory: Due to its close link to Folk Psychology, Bayesanism must relate to what people are doing.
- Hence, I suggest to explore the link between Bayesianism and empirical psychology in some detail.
- To do so, we examine three examples. The first two will also demonstrate the Plurality-of-Measures Problem.

1. Measures of Confirmation

All Bayesians agree that E confirms H iff $P(\mathrm{H}|\mathrm{E}) > P(\mathrm{H})$. But how can we measure the evidential strength? Here are some proposals:

- Distance measure: $d(\mathrm{H},\mathrm{E}) := P(\mathrm{H}|\mathrm{E}) - P(\mathrm{H})$
- Ratio measure: $r(\mathrm{H},\mathrm{E}) := \log\left[P(\mathrm{H}|\mathrm{E})/P(\mathrm{H})\right]$
- Log-likelihood measure: $l(\mathrm{H},\mathrm{E}) := \log\left[P(\mathrm{E}|\mathrm{H})/P(\mathrm{E}|\neg\mathrm{H})\right]$
- Joyce-Christensen measure: $s(\mathrm{H},\mathrm{E}) := P(\mathrm{H}|\mathrm{E}) - P(\mathrm{H}|\neg\mathrm{E})$
- *Z*-measure: $Z(\mathrm{H},\mathrm{E}) := d(\mathrm{H},\mathrm{E})/P(\neg\mathrm{H})$ if $d(\mathrm{H},\mathrm{E}) > 0$ and $Z(\mathrm{H},\mathrm{E}) := d(\mathrm{H},\mathrm{E})/P(\mathrm{H})$ otherwise

Problems:

1. These measures are not ordinally equivalent (Fitelson 2002).
2. Presumably only one measure can be right, but which?

Way Out 1: Philosophical Arguments

- Argue for conditions that any suitable measure should satisfy.
- Example: The evidential strength that a hypothesis receives from two independent pieces of evidence should be the sum of the evidential strengths the hypothesis receives from each piece of evidence. (Fitelson)
  Result: Only the *l*-measure fulfills this condition.
- Problem: There are other plausible requirements that favor other measures, and so this strategy does not seem to be conclusive.

Way Out 2: Pluralism

- Argue that different measures reflect different aspects of the confirmation relation (Huber, Joyce).
- Problem: It is not clear how this proposal relates to the practice of ordinary reasoning and of scientific reasoning.

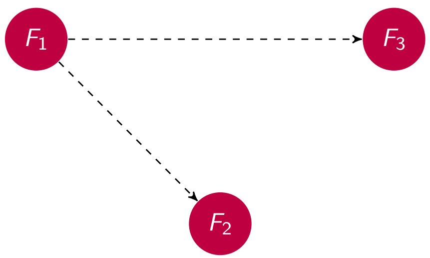
Way Out 3: Experiments

- Go empirical and study experimentally which (if any) of these measures is preferred empirically.
- Crupi, Gonzales and Tentori (2007) showed in a series of experiments that the Z-measure does best. The authors also present a number of normative reasons in favor of this measure and argue that the Z-measure is singled out on descriptive and normative grounds.
- I do not want to endorse this strong and arguably preliminary conclusion. There is still much work to be done to support it.
- All I want to do here is to point to this line of research and argue that a combination of formal and experimental work has the potential of being fruitful in epistemology.

2. Measures of Coherence

- A similar claim can be made about coherence measures. We are clearly able to judge the relative coherence of various information sets, but how can this be formalized? And what is the relation between coherence and truth?
- Unfortunately this important debate is stuck. There are several proposals for a coherence measure, but purely philosophical criteria (combined with our intuitions) do not suffice to single-out one of them.
- The proposals can be divided into two classes:
  - (i) The Non-Witness Approach, and
  - (ii) The Witness Approach.

The Non-Witness Approach



The Non-Witness Approach

One set of measures identified coherence with positive relevance between the propositions in an information set.

The Shogenji Measure ($n = 2$)

$$c_S(\mathrm{A}_1, \mathrm{A}_2) = \frac{P(\mathrm{A}_1|\mathrm{A}_2)}{P(\mathrm{A}_1)} = \frac{P(\mathrm{A}_1, \mathrm{A}_2)}{P(\mathrm{A}_1)\cdot P(\mathrm{A}_2)}$$

Other measures identify coherence with positive overlap in probability space.

The Olsson Measure ($n = 2$)

$$c_O(\mathrm{A}_1, \mathrm{A}_2) = \frac{P(\mathrm{A}_1, \mathrm{A}_2)}{P(\mathrm{A}_1 \vee \mathrm{A}_2)}$$

Note: Given such a measure, information sets can always be ordered according to their coherence.

The Shogenji Measure Generalized

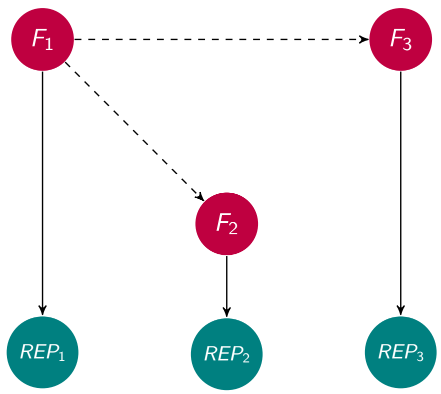
The Shogenji Measure

$$c_S(\mathrm{A}_1, \ldots, \mathrm{A}_n) = \frac{P(\mathrm{A}_1,\ldots,\mathrm{A}_n)}{P(\mathrm{A}_1)\cdots P(\mathrm{A}_n)}$$

Fitelson (2003) presents the following two criticisms:

1. If the E_i are logically equivalent (hence $P(E_i) = p$), then $c_S(S) = p/p^n = p^{1-n}$. This is *unintuitive* as we would expect the coherence to be maximal and independent of the prior in this case.
2. Shogenji's measure is based on the n-wise independence of the set. It is possible, for example, that two sets differ on all $(n-1)$-wise independencies, but have the same degree of n-wise independence and hence assign the same the (Shogenji) degree of coherence. This is *unintuitive*.

The Witness Approach



The Witness Approach

- An example of a measure that follows the witness approach is the Bovens-Hartmann measure and its generalizations by Douven and Meijs.
- This measure is defined as follows:

$$c_{BH} = \frac{\frac{P(\mathrm{F}_1,\ldots,\mathrm{F}_n|\mathrm{Rep}_1,\ldots,\mathrm{Rep}_n)}{P(\mathrm{F}_1,\ldots,\mathrm{F}_n)}}{\left(\frac{P(\mathrm{F}_1,\ldots,\mathrm{F}_n|\mathrm{Rep}_1,\ldots,\mathrm{Rep}_n)}{P(\mathrm{F}_1,\ldots,\mathrm{F}_n)}\right)_{max}}$$

- Note: This "measure" depends on the reliability r of the sources. Argue: A set S is more coherent than a set S' if, for all values of r, $c_{BH}(S) > C_{BH}(S')$. This leads to an Impossibility Theorem.
- For more on this, see Bovens & Hartmann 2003: chs. 1 and 2.

Experiments

- Harris and Hahn (2009) have conducted some preliminary experiments and explored which measure does best empirically.
- Harris and Hahn show that the Bovens-Hartmann measure does best, but there is more work to be done and the last word is certainly not yet spoken on this issue.

The Linda Problem

Tversky and Kahneman (1983) presented the following problem to the participants in an experiment:

Linda is in her early thirties. She is single, outspoken, and very bright. As a student she majored in philosophy and was deeply concerned with issues of discrimination and social justice.

Which of the following propositions is more probable?

(B) *Linda is a bank teller.*
(B & F) *Linda is a bank teller and a feminist.*

Results and (Provisional) Conclusions

- In the original experiment, 85% of the participants judged B & F to be more likely than B. This contradicts the probability calculus, or so it seems ("conjunction fallacy").
- What shall we conclude? That 85% of us are irrational?
- Tversky and Kahneman's studies triggered a tremendous amount of work in Cognitive Psychology, but also in Philosophy as the issue of rationality is at stake.
- The motivating assumption behind the corresponding research program is that we are doing quite well in our ordinary reasoning, and so examples such as the Linda case suggest that we should reconsider our theory of rationality and perhaps come up with an alternative that includes non-empirical and empirical considerations.

Ways Out

So how can the experimental findings be explained? Here are four proposals:

1. People implicitly add "and not a feminist" to proposition B.
2. People have problems with the notion of probability. If one uses *frequencies* instead, the effect will disappear. In fact, the number of people who commit the conjunction fallacy goes down but the effect does not disappear (Gigerenzer *et al.*).
3. People do not read *and* in B & F as the logical operator $\wedge$.
4. People ask which of the two propositions B and B & F is better *confirmed* by the background story (Fitelson *et al.*).

We now discuss a fifth proposal (and a follow-up) in more detail.

The Witness Model

- We assume that the participants in the experiments address the following question: *Which of the two options (i.e.* B *and* B & F) *is more probable given that a partially reliable source (i.e. the experimenter) informs you about them?*
- More formally, we assume that the participants compare the conditional probabilities $P(\mathrm{B}, \mathrm{F}|\mathrm{Rep_B}, \mathrm{Rep_F})$ and $P(\mathrm{B}|\mathrm{Rep_F})$. Note that both condition on different background information, and so $P(\mathrm{B}, \mathrm{F}|\mathrm{Rep_B}, \mathrm{Rep_F})$ can be larger than $P(\mathrm{B}|\mathrm{Rep_F})$.
- To study this proposal in more detail, we construct a concrete model: We assume that the variables B and F are probabilistically independent and that each proposition is uttered by a partially reliable witness.

The Witness Model (cont'd)

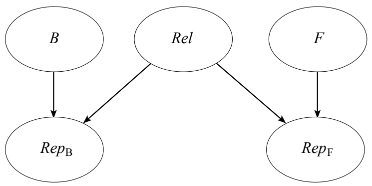


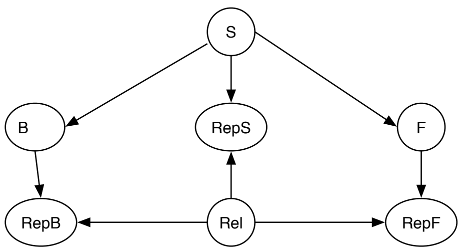
Figure: Bayesian Network representation of the Linda Problem

One can then show that $P(\mathrm{B}, \mathrm{F}|\mathrm{Rep_B}, \mathrm{Rep_F}) > P(\mathrm{B}|\mathrm{Rep_B})$ if $P(\mathrm{F}) > P(\mathrm{B})$ and another (arguably plausible) condition holds (Bovens and Hartmann 2003: ch. 3).

The Sophisticated Witness Model

- Observation: It does not seem to be correct that B and F are independent.
- In fact, both variables are negatively relevant given the background story S about Linda (i.e. that she is in her early thirties,...).
- The Sophisticated Witness Model therefore takes the background story S into account. S is also uttered by the experimenter, and the participants consider S in their probabilistic judgment.
- Claim: This approach makes sense psychologically as, e.g., the participants in a psychological experiment expect to be fooled. So why should they take as certain what the experimenter says?

The Sophisticated Witness Model (cont'd)



The Sophisticated Witness Model (cont'd)

Note that, by assumption, $P(\mathrm{B}|\mathrm{S}) < P(\mathrm{B})$ and $P(\mathrm{F}|\mathrm{S}) > P(\mathrm{F})$. We can now state our main result (Hartmann and Meijs 2009).

Theorem

The following claims are equivalent:
(i) $P(\mathrm{B}, \mathrm{F}, \mathrm{S}|\mathrm{Rep_B}, \mathrm{Rep_F}, \mathrm{Rep_S}) > P(\mathrm{B}, \mathrm{S}|\mathrm{Rep_B}, \mathrm{Rep_S})$
(ii) $P(\mathrm{F}|\mathrm{S}, \mathrm{B}) > P(\mathrm{F}) \cdot (1 + \delta)$.
(iii) $c_S(B, F, S) > c_S(B, S) \cdot (1 + \delta)$, with the Shogenji coherence measure c_S. δ is a (typically small) error measure.

Recall that c_S is defined as follows:

Shogenji Measure

$$c_S(B, S) = \frac{P(\mathrm{B}, \mathrm{S})}{P(\mathrm{B}) \cdot P(\mathrm{S})} \quad ; \quad c_S(B, F, S) = \frac{P(\mathrm{B}, \mathrm{F}, \mathrm{S})}{P(\mathrm{B}) \cdot P(\mathrm{F}) \cdot P(\mathrm{S})}$$

Bayesianism Reconsidered

- Let's return to our two problems:
  1. The Descriptive-Adequacy Problem
  2. The Plurality-of-Measures Problem
- The second problem can presumably be addressed by combining conceptual analysis, modeling, and experiments.
- Hence, the following picture emerges: Bayesianism consists of a "hard core" (beliefs come in degrees, degrees of belief are probabilities, a set of appropriate updating rules,...) plus various measures etc. which are fixed by empirical data and, perhaps, additional principles that are tentatively held.
- The data and the principles may have to be balanced out against each other until a relfective equilibrium is reached.
- This still leaves us with the first problem ...

The Descriptive-Adequacy Problem Reconsidered

- There is gap between actual human reasoning and the Bayesian representation of it. How bridging it?
- **Main idea:** Introduce a medium level of concepts.
- These concepts should (i) be grounded in human reasoning and (ii) allow for a Bayesian explication.
- What are the concepts of this medium level? Evidential strength, coherence and simplicity come to mind.
- As our third example ("Linda") shows, people can (arguably) make the right probabilistic judgments if they follow coherence considerations.
- All this results in a (hopefully progressive) philosophical research program which we call Naturalizing Bayesianism.

Open Problems

1. Solve the reflective-equilibrium problem.
2. Study the coherence-truth link for other information-gathering scenarios.
3. Extend Bayesianism to Social Epistemology.
4. Explore how to systematically combine logical and probabilistic information.
5. Explore the limits of the Bayesian approach.

Final Remarks

1. Bayesian Networks are powerful tools to address problems from epistemology and the philosophy of science.
2. In my three lectures, I have (i) introduced the theory of Bayesian Networks, and discussed various applications in (ii) philosophy of science and (iii) epistemology.
3. I have argued that the formal machinery needs empirical input:
   - Generalizations from case studies (lecture 2)
   - Experimental data from cognitive psychology (lecture 3)
4. Bayesian Networks are especially suited to integrate these findings.
5. I have also shown that there are many open questions. So please join in if you like.

Thanks. . .

for your attention!

Some Useful References

- Bovens, L. and S. Hartmann (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.
- Crupi, V., and S. Hartmann (2009). Formal and Empirical Methods in Philosophy of Science, to appear in F. Stadler et al. (eds.), *The Present Situation in the Philosophy of Science*. Berlin: Springer.
- Hajek, A. and S. Hartmann (2009). Bayesian Epistemology, to appear in: M. Steup (ed.), *Blackwell Companion to Epistemology*. Oxford: Blackwell.
- Hartmann, S. and W. Meijss (2009). Walter the Banker: The Conjunction Fallacy Reconsidered, to appear in *Synthese*.
- Hartmann, S. and J. Sprenger (2009). *Bayesian Epistemology*, to appear in: S. Bernecker and D. Pritchard (eds.), *Routledge Companion to Epistemology*. London: Routledge.

Further References

- Crupi, V., B. Fitelson and K. Tentori (2008). Probability, confirmation and the conjunction fallacy. *Thinking and Reasoning* 14, pp. 182-199.
- Crupi, V., K. Tentori and M. Gonzalez (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science* 74, 229-252.
- Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science* 66, S362-S378.
- Harries, A. and U. Hahn (2009): Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1366-1373.

Further References (cont'd)

- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences* 10(10), 464-470.
- Oaksford, M. and N. Chater (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning.* Oxford: Oxford University Press.
- Schupbach J. (2009). Is the conjunction fallacy tied to probabilistic confirmation? To appear in *Synthese*.
- Tentori, K., V. Crupi and D. Osherson (2007). Determinants of confirmation. *Psychonomic Bulleting and Review* 14, 877-883.
- Tversky, A. and D. Kahneman (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review* 90.