Program of Work

Epistemology is the philosophical discipline concerned with knowledge. Formal epistemology is the subdiscipline of epistemology that uses formal or mathematical methods to investigate epistemological questions. More narrowly, however, one could say that formal epistemology, in contrast to traditional epistemology, thinks that partial beliefs (which come in degrees) are particularly suited to scientific inquiry and that their quantitative nature lends itself to the use of mathematical methods.

My postdoctoral project, *Formal Theory of Partial Belief Updating*, seeks to contribute to the establishment of axiomatic foundations for partial belief epistemology. The dominant articulation of formal epistemology is the Bayesian approach. There is another approach to reasoning with partial beliefs that is often considered to be ancillary to Bayesian methods: information theory. My research is shaped by the idea that information theory is more foundational than has been appreciated for understanding the norms of reasoning with partial beliefs. Information theory can serve to justify many of the formal methods of Bayesian epistemology and extend them to cases that are highly controversial even among Bayesians.

Bayesian epistemology is characterized by its commitment to prior probabilities and standard conditioning. My project is concerned not with prior probabilities but with the norms that govern the updating of an agent's probability. An agent learns something new, possibly by observation or experiment. The question is whether there are rationality constraints on how the agent should change her partial beliefs in light of the new evidence. Bayesians think that there are such constraints. They have a core commitment to standard conditioning, a mathematical formula that keeps ratios of partial beliefs unaffected on the condition that the new evidence does not specifically affect them. In this way, Bayesians have developed a logic of partial beliefs, not dissimilar to the more traditional logic of full beliefs with which we are familiar. Conditioning is a norm that aims at diachronic consistency (compare this to the synchronic consistency that we customarily require of full beliefs in logic, such that an agent should not believe "*p* and *q*" and "not-*q*" at the same time), ensuring that an agent's revised degrees of belief take evidence into account in an appropriate manner.

There are conceptual and formal problems with diachronic consistency requirements. These include: how to represent new evidence formally, how to justify consistency requirements and updating rules, how to evaluate the Bayesian position, and, most importantly for my project, how to extend the Bayesian position to cases where standard conditioning does not apply.

Information theory unifies the formal relationships at the heart of many of these questions. Information theory was developed independently of formal epistemology. It is a discipline about the encoding, transmission, and decoding of messages. Its mathematical apparatus is distinct from the mathematical apparatus of probability theory, but connects to it in substantial and sometimes surprising ways.

Information theory proposes to govern the updating of probabilities by a powerful and simple intuition: the fundamental diachronic consistency requirement is that an agent ought not to inform herself more than necessary and less than required by the intervening evidence. This idea is mathematically appealing in the sense that the use of Lagrange Multipliers in information theory mirrors the Bayesian instinct to keep ratios constant. Even though the methods of probability and information theory are different, information theory supports Bayesian norms such as standard conditioning. Information theory also confirms some proposed extensions of Bayesian consistency requirements (for example Richard Jeffrey's and Carl Wagner's) -- and disconfirms others (for example Hannes Leitgeb's and Richard Pettigrew's).

Bayesians have developed several strategies to justify the norms of their theory. The traditional approach is to highlight the pragmatic virtues of these norms. If an agent makes decisions based on Bayesian norms, the agent maximizes his or her expected outcome. Other approaches have come to supplement the traditional approach. Richard Cox pioneered an approach which looks at the formal virtues of a theory. This approach asks questions about which formal approach yields unique solutions depending on a number of axioms. Very recently, a new approach to justification has garnered a lot of attention among formal epistemologists: Jim Joyce's "accuracy first" epistemology seeks justification for updating norms by investigating how well these norms fare in terms of expected accuracy, where accuracy is some notion of proximity to the truth. Given these three approaches highlighting pragmatic, formal, and epistemic virtues, information theory has the potential to illuminate strengths and weaknesses for all of them and add substantial elements to them.

My doctoral dissertation, *Information Theory and Partial Belief Reasoning*, addresses concerns about the relevance of information theory to epistemology, especially the sweeping applicability of controversial principles such as the Principle of Maximum Entropy or the Principle of Minimum Cross-Entropy. My postdoctoral research will be focused on the development and justification of a positive formal theory aimed at the articulation and justification of norms for updating probabilistic beliefs on the basis of information theory. The projects are continuous but different. At the centre of my dissertation is a discussion of a crucial set of counterexamples and conceptual difficulties with the use of information theory in

Bayesian epistemology. A majority of Bayesians reject information theory as anything more than a tool kit for calculations. Bayesians often reject information-theoretic tools when they superficially conflict with intuitions. For this reason, my dissertation undertook the negative task of disarming the major objections to information theory.

My positive project, to be undertaken as my postdoctoral research, is to demonstrate convincingly that information theory is a key conceptual and formal component of the justificatory structure supporting the Bayesian view. To come to these conclusions, I need to present a unified formal theory of probability updating based on information theory that accords with our most basic epistemic intuitions. There are several mathematical and philosophical challenges along the way. I will name just two of them, related (respectively) to the pragmatic and epistemic approaches to justification. The first is that in betting scenarios (which are representative of a much wider range of decision problems) credal states that are inconsistent with information theory may do much better than their counterparts. The second is that information theory insists on heavily non-trivial asymmetry patterns in its similarity measures or scoring rules for probability distributions. I cannot present these problems here in any detail, but my hope is that they will lead to an information-theoretically based logic of partial beliefs that is richer and formally more integrated than existing accounts.

To give a flavour of the kinds of questions that my postdoctoral research will pursue I am providing a list of the following promising projects:

- ◆ What is the nature of information and what is its proper connection to Bayesian probability theory? I am primarily interested in the axiomatic approach where, both in information theory and probability theory, norms are tested by axioms which justify them. It is a substantial question which axioms to include and which to exclude, and it is often surprising how specific the recommendations of relatively weak axioms are.

- ◆ What is the topology induced by similarity measures of probability distributions based on information theory? Currently, a Euclidean topology is favoured for the epistemic justification of both synchronic and diachronic partial belief norms. A Euclidean topology is inconsistent with information theory. More strongly, information theory is also inconsistent with any non-Euclidean metric. The similarity measure recommended by information theory is asymmetric. This asymmetry is non-trivial and intransitive. Prima facie, these are troubling facts about the use of information theory in partial belief logic. I am hoping to contribute an account of justification in favour of information theory (thereby challenging the simple metrics that are used today for epistemic justification). It is an open question how the current lively debate about scoring rules (a scoring rule measures how accurately a partial belief function reflects the truth compared to other

partial belief functions) relates to the questions of topology that are of great interest to me.

◆ In decision theory, credal states (states which quantitatively reflect what the agent believes) can be tested with respect to their success in betting scenarios. It appears that credal states inconsistent with information theory do better than credal states that conform. We need a more fully formed formal account of how to use information theory in decision theory to address this problem. One application for this formal account is the stock market. There is a potential that information theory can provide rational constraints on buying and selling based on available information. It is important to note that information theory clarifies the epistemic nature of partial beliefs. Partial beliefs, sometimes expressed in subjective probabilities, are epistemic states of an agent and have an intricate conceptual relationship with objective chances. The norms associated with partial beliefs are based on what is accessible to the agent: it is not useful to require an agent to have partial beliefs that match objective chances independent of available information, just as it is not useful to require an agent to have true full beliefs independent of justification.

◆ There has been good progress on justifying Bayesian commitments (especially probabilism and standard conditioning) epistemically rather than pragmatically. These epistemic justifications rest on assumptions that are inconsistent with information theory. It is an open question whether the axioms that are currently inconsistent with information theory can be revised to provide equally strong justifications in favour of these Bayesian commitments. I am optimistic that it can be done and hope to do so.

The result of my postdoctoral research is intended to be a set of published papers in journals such as *Synthese* , the *British Journal for the Philosophy of Science*, *Philosophy of Science*, *Philosophical Studies*, *Philosophical Perspectives*, *Philosopher's Imprint*, *Journal of Philosophy*, *Erkenntnis*, and the *International Journal of Approximate Reasoning*. These papers, together with additional material that I expect to draw from my dissertation and to develop during the tenure of my postdoctoral fellowship, may form the core of a book, tentatively titled *Information Theory and the Logic of Quantitative Belief.* I have submitted a book proposal to Oxford University Press, which has requested sample chapters from me. I am currently working on these sample chapters and hope that SSHRC will be able to fund further research to make this project possible.