Formal Epistemology

NUS Presidential Conference

May 25-27, 2023

Republic of Singapore Yacht Club (RSYC)
52 W Coast Ferry Rd, 126887

About

What principles of rationality govern credences, beliefs, values, and so on? To what formal rules might those principles conform? And what, more generally, is the connection between formal and non-formal methods in philosophy? These questions, and others like them, have received considerable attention from philosophers in recent decades. The purpose of this conference—which is part of a larger series of conferences in the philosophy department at NUS—is to bring researchers together, to facilitate discussion of these issues in formal epistemology.

Speakers

Jennifer Carr (UCSD)
Lisa Cassell (UMBC)
Kevin Dorst (MIT)
Zach Barnett (NUS)
Kenny Easwaran (Texas A&M)
Melissa Fusco (Columbia)
Al Hájek (ANU)
Julia Staffel (Boulder)
Snow Zhang (Berkeley)

Official Discussants

Boris Babic (Toronto)
Ben Blumson (NUS)
Joseph Cobon (Oxford)
J. Dmitri Gallow (ACU)
Jie Gao (Zhejiang)
Mikayla Kelley (Stanford)
Gabrielle Kerbel (Michigan)
Boris Kment (Princeton)
Michael Nielsen (Sydney)
Abelard Podgorski (NUS)
Hsueh Ming Qu (NUS)
Neil Sinhababu (NUS)
Laura Soter (Duke)
Weng Hong Tang (NUS)

Organized by Ashley Chay (NUS) and Isaac Wilhelm (NUS).

Sponsored by an NUS Presidential Young Professorship grant.

Schedule

May 25

9:30–9:55am	Coffee and refreshments
9:55–10am	Opening remarks
10–11:30am	*Do Gambler's Commit Fallacies?* Kevin Dorst (MIT)
11:30–11:45am	Break
11:45–1:15pm	*Unfinished Business, or: When Being Akratic is a Good Idea* Julia Staffel (Boulder)
1:15–3pm	Lunch: RSYC
3–4:30pm	*Title TBA* Jennifer Carr (UCSD)
4:30–4:45pm	Break
4:45–6:15pm	*Against Risk-aversion* Zach Barnett (NUS)
7pm	Dinner: Whole Earth (76 Peck Seah St., 079331)

May 26

9:30–10am	Coffee and refreshments
10–11:30am	*Learning by Maximizing Expected Accuracy in an Infinite, Non-Partitional Context* Kenny Easwaran (Texas A&M), joint work with Michael Nielsen (Sydney)
11:30–11:45am	Break
11:45–1:15pm	*The Retrospective Account of Bayesian Updating* Lisa Cassell (UMBC)
1:15–3pm	Lunch: RSYC
3–4:30pm	*Why Image?* Snow Zhang (Berkeley)
4:30–6:30pm	Activity: Haw Par Villa
7pm	Dinner: Publico (1 Nanson Rd., 238909)

May 27

9:30–10am	Coffee and refreshments
10–11:30am	*Absolution of a Causal Decision Theorist* Melissa Fusco (Columbia)
11:30–11:45am	Break
11:45–1:15pm	*Consequentialism, Cluelessness, Clumsiness, and Counterfactuals* Al Hájek (ANU)
1:15–1:20pm	Closing remarks
2–6pm	Activity: MacRitchie hike
7pm	Dinner: The Dragon Chamber (2 Circular Rd., 049358)

Abstracts

May 25

Do Gambler's Commit Fallacies?

Kevin Dorst (MIT)

The "gambler's fallacy" is the widely-observed tendency for people to expect random processes to "switch"—for example, to think that after a string of tails, a heads is more likely. Is it irrational? Understood narrowly, it is—but we have little evidence that people exhibit it. Understood broadly, I'll argue that it follows from reasonable uncertainty combined with rational management of a limited memory.

Unfinished Business, or: When Being Akratic is a Good Idea

Julia Staffel (Boulder)

This talk has two main objectives. The first is to introduce the core ideas of my current book project, *Unfinished Business*. The book investigates the nature of the attitudes we form during ongoing deliberations, and the norms that apply to them. The second objective is

to take a closer look at the relationship between first-order and higher-order attitudes while our reasoning is in progress, and to ask how they should relate to one another. I will argue that cases of in-progress reasoning suggest a surprising claim about epistemic akrasia. While epistemic akrasia is almost always irrational when it affects the conclusions of our reasoning, it is typically rational when it obtains while our deliberations are still ongoing. I call this version of rational akrasia *transient akrasia*.

Now, What Have We Learned?

Jennifer Carr (UCSD)

NA

Against Risk-aversion

Zach Barnett (NUS)

Risk-aversion can be understood as a tendency to favor known outcomes over gambles—even when the gambles would maximize expected utility. Although orthodox decision theory forbids this sort of risk-aversion, it is widely thought that some degree of risk-aversion can be rationally permissible. And to respect the rationality of risk-aversion, some have proposed alternative decision theories, which make room for a range of different ways of valuing risky options. But an examination of these views reveals problems. Specifically, it can be argued that permitting risk-aversion commits us to choices which are expected to harm some and help none. After seeing how this difficulty arises, we'll see how this argument against risk-aversion differs from existing criticisms, which traditionally appeal to the fact that risk-averse agents are liable to lose out in the long run. The argument developed here does not rely upon this observation and can be presented in the context of a single decision.

May 26

Learning by Maximizing Expected Accuracy in an Infinite, Non-Partitional Context

Kenny Easwaran & Michael Nielsen (Texas A&M, Sydney)

Greaves and Wallace (2006) justify Bayesian conditionalization as the optimal update strategy for an agent considering finitely many possibilities, who is about to undergo a learning event where the potential propositions that she might learn form a partition — she is already certain that exactly one of the propositions she might learn is true, and that she will learn the one that is in fact true. In recent years, several philosophers have objected that learning often takes place in less idealized circumstances. If the agent has only imperfect access to her evidence, then there will generally be multiple possible evidence propositions that are true in each world, and she may be unsure which of these will be learned in any given world. Various proposals have been made for how best to respond to these learning scenarios, though all look something like Bayesian conditionalization on some proposition that is closely related to the one that she learns. (Schoenfield, 2017; Gallow, 2021; Carr, 2021; Isaacs and Russell, 2022; Schultheis, 2023) Much of this is motivated by an externalist epistemological view.

We have earlier extended the results of Greaves and Wallace to learning situations in which the agent has an infinite set of worlds, but the potential propositions learned still form a partition. (Easwaran, 2013; Nielsen, 2022b) Here we seek to do the same for these externalist update rules in which the propositions do not form a partition.

In moving from the finite to the infinite case, one important step involves moving from a partition to a σ-algebra. Some apparent paradoxes (Dubra and Echenique, 2004) can be resolved if we are careful about how the move is made (Tóbiás, 2021). We show how to make the parallel move for non-partitional learning situations.

The Retrospective Account of Bayesian Updating

Lisa Cassell (UMBC)

This paper argues for a new account of Bayesian updating by taking a retrospective approach to diachronic coherence. This approach says that an agent is diachronically coherent whenever the information she has revised her beliefs on satisfies whatever constraint we would want our evidence to satisfy. This approach contrasts with a common way of thinking about the Bayesian framework, according to which it treats evidence as a black box. The aim of this

paper is to provide a different interpretation of Bayesianism's main updating constraint by replacing this black box with a Bayesian account of evidence.

Why Image?

Snow Zhang (Berkeley)

According to the orthodox, there are two kinds of supposition, indicative and subjunctive, and they call for different hypothetical revisions of one's degrees of belief (credences). Rational credal revision under indicative supposition goes by Bayesian conditionalization, whereas rational credal revision under subjunctive supposition goes by imaging. But why is imaging the rational way to revise one's credences under subjunctive supposition (assuming that it is)? This talk proposes an answer to this question: I prove a Dutch book theorem and an accuracy-dominance theorem for imaging, parallel to de Finetti's Dutch-book and accuracy-dominance theorems for conditionalization. I'll also discuss how these results shed light on the status of rational norms for credence in conditionals that have been proposed in the literature, such as the Stalnaker's Thesis and the principle of Preservation.

May 27

Absolution of a Causal Decision Theorist

Melissa Fusco (Columbia)

I respond to a dilemma for Causal Decision Theory (CDT) under determinism, posed in Adam Elga's paper "Confessions of a Causal Decision Theorist". The treatment I present highlights (i) the status of laws as predictors, and (ii) the consequences of decision dependence, which arises natively out of Jeffrey Conditioning and CDT's characteristic equation.

My argument leverages decision dependence to work around a key assumption of Elga's proof: to wit, that in the two problems he presents, the CDTer must employ subjunctive-suppositional (rather than evidential) transformations of a shared prior.

Consequentialism, Cluelessness, Clumsiness, and Counterfactuals

Al Hájek (ANU)

According to objective consequentialism, a morally right action is one that *has the best consequences*. (These are not just the immediate consequences of the actions, but the *long-term* consequences, perhaps until the end of history.) I will argue that on one understanding this makes no sense, and on another understanding, it has a startling metaphysical presupposition concerning counterfactuals. Objective consequentialism has faced various objections, including the problem of "cluelessness": we have no idea what most of the consequences of our actions will be. I think that on these understandings, objective consequentialism has a far worse problem: its very foundations are highly dubious. Even granting these foundations, a worse problem than cluelessness remains, which I call "clumsiness". Moreover, I think that these problems quickly generalise to a number of other moral theories. But the point is most easily made for objective consequentialism, so I will focus largely on it.

I will consider three ways that objective consequentialism might be improved:
1) Appeal instead to short-term consequences of actions;
2) Understand consequences with objective probabilities;
3) Understand consequences with subjective/evidential probabilities.
But even here, there be dragons.