The Reasoner

Volume 13, Number 11
November 2019

thereasoner.org
ISSN 1757-0522

Contents

Editorial

Dear Reasoners,

It is with great pleasure that I introduce you to this issue of The Reasoner, featuring an interview with Jeffrey Helzner. A decision theorist by training, Jeffrey started an academic career, but then switched to industry – improving greatly his happiness. Currently he is Decision Scientist at Blackboard Insurance in New York. Jeffrey's professional choice went against the default according to which (bright) PhDs should aim at getting tenure and insist until they manage to do so, perhaps after some globetrotting. But Academia does not make everyone happy, a bit of information that can only be acquired through personal and sometimes tough experience. This should then defeat the default. But this is often more easily said than done, for a decision to this effect may arrive at a time when one is well beyond their 20s and the market for interesting job is populated with young competitors. Moreover, non academic jobs seem to be more readily available to PhDs in quantitative subjects. What should the others do then? If you have any thoughts on this, we would love to hear from you at features@thereasoner.org.

Many thanks to Jeff for his time and for sharing his view with the readers of The Reasoner on a subject which, I think, should be discussed more openly by academics and non academics alike.

Hykel Hosni
University of Milan

Features

Interview with Jeffrey Helzner

Hykel Hosni: Can you tell us a bit about your background?

Jeffrey Helzner: I grew up in Philadelphia. My parents are medical doctors, as are two of my uncles, but I've never had any interest in medicine – too messy and too much memorization for my tastes. For as long as I can remember, I've spent a good amount of my time thinking about thinking. I did and still do enjoy programming. The programmable toy vehicle – Big Trak – was a game changer for me, as was my father's copy of *Gödel, Escher and Bach*. I really hated school as a kid (and really didn't think much of it until grad school). Most of my friends in high school were musicians or artists or trouble makers or freaks. I never hung around with the math and science kids in high school.

HH: But you did eventually get to graduate school . . .

JH: After drifting around for a while, not really knowing what I wanted to do, I eventually decided to study categorical logic with the great F.W. Lawvere. When I finished my masters in mathematics at Buffalo – recognizing that mathematics was a means rather than an end for me – I moved to the *Logic, Computation, and Methodology* program at Carnegie Mellon. There I continued my studies in categorical logic with Steve Awodey and wrote a master's thesis on that topic. Awodey is obviously an outstanding category theorist and analytical philosopher, but somewhere around that time my interests started to shift from analytic philosophy to pragmatism and from logic to decision theory.

HH: Some people see a great deal of overlap between logic and decision theory, but you seem to think otherwise. What is in your view the key difference between logic and decision theory?

JH:

I do think there's overlap, but it might not qualify as what you mean by a "great deal of overlap". I have an unpublished paper "Admissibility in a Logical Framework" where I attempted a logic-based characterization of some rational choice norms. I guess I'm not entirely clear on the boundaries of logic and decision theory.

I view logic as providing a norm/standard for full belief, much as the probability axioms, or some generalizations of them, provide a standard for partial beliefs. You could view (normative) decision theory as providing a standard on consistency for admissibility (e.g., the standard of expected utility theory requires that the admissible alternatives maximize expected utility with respect to the set of available alternatives). To me this would leave out much of what I find interesting about decision making. One difference that is important to me is that I find it more natural to connect decisions to observables than trying to go more directly from reasoning to observables. Without accounting for preferences/tastes/values, it seems more difficult to distinguish between an illogical agent and one who just doesn't care about the task at hand. Not saying that this isn't an issue at some level for decision theory – e.g., Kahneman and Tversky make experimental assumptions in their demonstrations of irrationality – but decision theory, unlike logic, has structure to account for this in its attention to desires and beliefs.

HH: I see, thanks. You were talking about how your interestes shifted from pure mathematical logic to decision theory . . .

JH: Yes, and that is why I decided to work under Teddy Seidenfeld for my dissertation – I've been focused on decision making ever since. I feel very fortunate to have studied decision theory with Teddy. There aren't many people on this planet who have thought as deeply about decision theory as he has. Before moving on, I also want to mention the important influence that Jeremy Avigad and Horacio Arlo Costa had on me. Jeremy served on the committee for my master's thesis and really gave me my first research experience that eventually led to a joint publication. Horacio served on my dissertation committee and introduced me to many interesting topics in philosophical logic. We eventually collaborated on some projects, became good friends and remained so until his tragic death in 2011. I still miss him.

HH: I can imagine how tough that might have been. But going back to your graduate work, what was your dissertation about?

JH: It was mainly about relaxing the ordering assumption in empirical tests of a decision theory's descriptive accuracy. For example, in many tests of the descriptive adequacy of expected utility theory it is simply assumed – either explicitly (as in when the subject is asked to rank the alternatives) or implicitly (as when all of the items are based on pairwise choice) – that the subject has a complete preference ordering with respect to the given alternatives.

HH: Which of course is a very demanding assumption in many realistic cases. Were you happy with your experience in graduate school?

JH: I have fond memories of my time at Carnegie Mellon. The faculty really gave me the impression that philosophy (at least what is emphasized in the *Logic, Computation, and Methodology* program at CMU) is a living and active discipline with an important role to play in our modern world. I'd like to think that my current work embodies some of the spirit that the CMU faculty intended to convey.

HH: Indeed, let us move on to your current work. You started an academic career in Philosophy at Columbia University, but then you switched to industry. Can you tell us what motivated you in this decision?

JH: I started as an assistant professor at Columbia and left as associate professor (without tenure). I was miserable at Columbia. I really didn't fit in at all. I'll never be able to understand the appeal of a Judith Butler. Too fancy for me.

HH: I see what you mean . . .

JH: In retrospect, I stayed at Columbia and in academic philosophy much longer than I should have for my development, but there was other stuff going on in my life (e.g., breakup of my first marriage, Hoarcio's tragic death) that made the idea of a career change seem overwhelming at the time. So when I (thankfully) failed to get tenure at Columbia, it was natural for me to move to industry.

HH: How did it all begin?

JH: My first position in industry was leading a small behavioral science team as part of AIG. We were a small part of the larger Science team that had received a lot of support while Peter Hancock was the CEO. Roughly speaking, our focus was on developing interventions to improve human decision making. There's a lot of judgment and decision making under uncertainty in commercial insurance. This is true on the underwriting side as well as the claims side (e.g., as when an adjuster tries to estimate a case reserve).

HH: That sounds indeed a bit more more hands-on that dealing with forgetful functors! Are you still in that position?

JH: No, my current position is as a decision scientist on the R&D team at Blackboard Insurance, which is AIG's insurtech startup.

HH: Can you list advantages and disadvantages of coming to that job from an academic path?

JH: Well, I can list those with respect to my academic path and my position in industry, but I'm not sure how much they generalize. The education I received at Carnegie Mellon – foundations of cognitive science, decision theory, probability,

utility, measurement theory, and computation – serves me almost every day in my current role. The value comes less from any particular results that have been obtained in those subjects and more from the robust framework that they provide for structuring real world problems. Disadvantages? Not sure, really. Probably learning to balance a concern for conceptual elegance and practical value.

HH: Your experience suggests that a well designed (post)graduate programmes don't have to be aimed explicitly at industry placement to be relevant for the needs of industry. Based on your experience, do you think industrial employers are aware of this potential, or even show particular interest in recruiting PhDs in theoretical disciplines like Philosophy?

JH: Not that I've noticed. That said, the R&D team at Blackboard includes PhDs in various fields, from theoretical physics to statistics to computer science to civil engineering.

HH: Quantitative PhDs seem more obviously attractive to industry. Moving to a more general topic, in many areas (both academic and geographic) people who take up non academic jobs after their PhD (or maybe postdoc) are perceived like "opting out" academia – as if the default option of anyone who starts a PhD is to become a professor, unless of course they become too old to keep trying. I think this view vastly underestimates the value of doing a PhD, but I'm afraid it matches the way many PhD programmes are structured. Do you have any thoughts on this and possibly on how to fix it?

JH: Yeah, this is unfortunate since some people might even find that they prefer industry research. I certainly do.

No, I'm not sure how to fix it. Probably best to look at why the attitude has evolved in certain parts of academia. Who or what is served by the attitude in question, right?

HH: I think that is the problem and I guess this can only be tackled by opening a frank conversation on the issue. There are certainly many PhD students among the readers of The Reasoner who may be interested in pursuing non academic careers, and perhaps some may be motivated by your story. What would you advise them to do to have a picture of their options?

JH: Spend some time thinking about how your interests and skill set could be applied/developed in industry.

HH: Thanks very much Jeff.

The Reasoner Speculates

Legg-Hutter universal intelligence implies classical music is better than pop music for intellectual training

In their thought-provoking paper, Legg and Hutter (2007: Universal intelligence: A definition of machine intelligence, https://arxiv.org/pdf/0712.3329.pdf, Minds and machines 17(4), 391–444) consider a certain abstraction of an intelligent agent, and define a *universal intelligence* measure, which assigns every such agent a numerical intelligence rating. We will briefly summarize Legg and Hutter's paper, and then give a tongue-in-cheek argument that if one's goal is to become more intelligent by cultivating music appreciation, then it is better to use classical music (such as Bach, Mozart, and Beethoven) than to use more recent pop music. The same argument could be adapted to other media: books beat films, card games beat first-person shooters, parables beat dissertations, etc.

Legg and Hutter paint an ingenious portrait of the arbitrary intelligent agent. It is an open problem to define what intelligent agents *are*, but that difficulty is swept under the rug by instead considering what intelligent agents *do*. Legg and Hutter formalize both the agents, and the environments in which they operate. The idea is that an agent should be flexible enough that it can be placed in any computable environment, where it will make an initial *observation* and receive an initial *reward*, and then it will respond by taking an *action*. The environment will respond to the action with a new observation and reward. The agent will then respond by taking a second action. This process continues forever: observation, reward, action, observation, reward, action, ... The sum of the rewards from an environment measures how well the agent performs in that environment. Legg and Hutter formalize this in detail (made more difficult because they allow non-determinism, something we will ignore for sake of brevity).

Legg and Hutter's goal is to assign each agent a numerical *universal intelligence* rating, in such a way that each agent's intelligence rating captures, in some way, how well the agent performs (i.e., how much reward it extracts) across the whole universe of all computable environments. The problem is, there are infinitely many computable environments. How can we distill performances across an infinitude of environments into one single number? Legg and Hutter's answer is to use a weighted infinite sum. Multiply the reward from each environment by a weighing-factor, giving exponentially less weight to more complex environments (environments with higher Kolmogorov complexity). Under certain technical assumptions about the environments, these weighing-factors can be chosen in such a way that every agent's sum converges to a finite real number– which Legg and Hutter call the agent's *universal intelligence*. This is like the Occam's Razor of intelligence measurement: if we are judging a robot's general-purpose abilities, we probably care more about how well the robot performs in routine household environments than how well the robot performs in complicated and contrived environments like "dodge oncoming traffic while juggling a hundred chainsaws".

To summarize so far: an agent's universal intelligence rating is defined as a weighted infinite sum of the agent's raw performance numbers across the whole space of computable environments, giving exponentially smaller weights to more complex environments, in such a way that the infinite sum always converges for every agent.

Assuming Legg and Hutter have given us an accurate glimpse at intelligence, I offer an unexpected argument. I will argue that if you want to become more intelligent by cultivating music appreciation, then you will get better results from classical music than from pop.

The key difference between classical and pop is that a classical piece is entirely determined by its musical score, whereas a pop song is determined by its performance. Casually copy Bach's score by hand and you'll get the exact same piece, even though you have different handwriting than Johann Sebastian Bach. On the other hand, if you cover a Beatles song, it will be universally considered a distinct new piece, regardless of whether or not you manage to play the right notes. To truly capture a Beatles piece in the same way as a Bach piece, the "score" would have to specify the exact contours of each singer's lungs and diaphragm and every tiny movement of the guitars.

From the above observations, our argument writes itself. The

environment "Listen to Beethoven's *Moonlight Sonata* and derive pleasure from it" is much less complex than the environment "Listen to Pink Floyd's *Welcome to the Machine* and derive pleasure from it", because the former can be expressed with little more than its musical score (a few hundred kilobytes), and the latter basically requires a full recording (megabytes). If we were shopping for a general-purpose robot, the latter environment would be much more contrived, and therefore we ought to give more weight to the robot's performance at the former.

We hope this playful argument will inspire some original thoughts about the nature of intelligence and its measurement.

SAMUEL ALEXANDER

NEWS

Calls for Papers

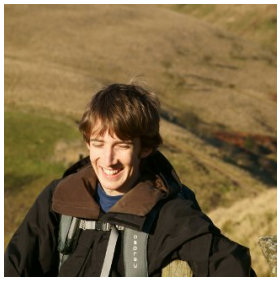
NANCY CARTWRIGHT'S PHILOSOPHY OF SCIENCE: special issue of *Theoria*, deadline 1 November.

IDEALIZATION, REPRESENTATION, EXPLANATION ACROSS THE SCIENCES: special issue of *Studies in History and Philosophy of Science Part A*, deadline 15 January.

TRUTH AND FALSITY: special issue of *Kairos*, deadline 28 February.

WHAT'S HOT IN . . .

Uncertain Reasoning



A "benchmark" was originally a mark that surveyors would make in some structure at a known height. This then allowed them to make measurements consistently from a known starting point. The term "benchmark" is now used across a number of fields to mean a standardised test or battery of tests used to consistently assess, for example performance at some task.

I've been thinking recently about what benchmarks we might use to assess performance of a formal model of uncertain reasoning. One approach that might be helpful here is the "Challenge Problems" put forward about fifteen years ago by a number of engineers and statisticians. (Oberkampf et al (2004). "Challenge problems: uncertainty in system response given uncertain parameters." Reliability Engineering and System Safety, 85:11–19.)

The challenge problems are a set of statistical inference problems that vary in terms of the information you are given, and the way it must be used. For example, you might be given an estimate of some values, and then you are required to calculate some function of those values (taking into account the uncertainties involved). Or, you are given a collection of estimates of upper and lower bounds on some values, those values are parameter values for a simple physical model, and you must predict something about the behaviour of the physical system. The above mentioned paper is in a special issue of the journal, and most of the other papers in that issue are various attempts to solve the challenge problems using a variety of statistical approaches (bayesian theory, lower previsions, random sets, belief functions and many more). A summary of the approaches is provided in Ferson et al (2004). ("Summary from the epistemic uncertainty workshop: consensus amid diversity." Reliability Engineering and System Safety, 85:355–369.)

Something about this idea appeals to me. I guess in psychology, models of uncertain inference are measured againts the benchmark of the empirical data about how people actually reason.

In a logical approach, the typical touchstones of a theory's performance are its expressive power, its metatheoretical properties (soundness and completeness) and also, possibly, properties related to algorithmic complexity.

In philosophy I think the situation is a little less clear. Different mathematical theories of uncertain belief are assessed on a pretty ad hoc basis, and perhaps it would help to have a standard set of issues that we would like a theory to deal with. So for example, John Norton has criticised Bayesian theory for failing to properly accommodate ignorance (Norton, John D. "Ignorance and indifference." Philosophy of Science 75.1 (2008): 45-68.) Imprecise Probabilities – which arguably do better as regards ignorance – are criticised for, among other things, not being able to properly accommodate learning (Vallinder, Aron. "Imprecise bayesianism and global belief inertia." The British Journal for the Philosophy of Science 69.4 (2017): 1205-1230.) Maybe we want a theory of rational belief to have sensible solutions to, say, Newcomb's problem or Sleeping Beauty.

I don't want to suggest that I have anything close to a complete list of desiderata for a formal theory of uncertain reasoning, but I think it's an interesting exercise to explore what things one might want from such a theory. I.J. Good claimed that there were 46,656 varieties of Bayesianism (Good, I. J. "Good thinking: the foundations of probability and its applications. 1983."), and I suspect that we'd have a similar diversity of formal theories tailored to different people's priorities as regards theoretical desiderata. And that's the point. Until we agree on the criteria for a good theory of uncertain reasoning, we can't know whether our arguments about what is a good theory or not are genuine disagreements, or whether we are just talking past one another. That's the question I'll leave you with: what are the benchmarks against which we should be assessing the performance of theories of uncertain reasoning?

Science Policy

Society looks up to scientists for answers to questions about health, technology, social policies, etc. They are experts, expected to bring progress to society. However, scientists themselves are only humans and are prone to mistakes. Some of these mistakes are unconscious, such as cognitive biases, but some also represent normative misconducts. Though, due to the human nature, normative misconduct cannot be completely eliminated, there are measures that we can take to reduce it.

Not every scientific misconduct is equally strong. However, all of them have negative consequences for the scientific community. Consider for instance a case of a publication that did not reach all the standards for making a solid conclusion, but it

is also not reporting anything purposefully false. Rather, it is a publication based on unstable and hypothetical results. Such a publication *prima facie* might not seem overly problematic, at least from the perspective of the author. However, from the perspective of the scientific community as whole it adds to the noise in the literature, making it harder to parse, and might even mislead other researchers. Thus, academic honesty apart from being a character virtue, also has positive epistemic consequences for the scientific community.

Contemporary scientists report strong existential pressure. Early-career researchers often do not have stable jobs or cannot easily chose a country in which they will find one. They perceive their future as rather insecure. Because of these pressures, scientists feel the need to publish as much as possible, even when their results are not completely solid. More publications guarantee them better career prospects. Thus, one obvious (and often discussed) solution for scientific misconduct is offering more stable jobs. Once the existential pressure is reduced, the need for publishing becomes less desperate.

Apart from negative motivation, positive measures to stimulate the scientific productivity are also taken. One of them is giving financial incentives for publishing in high impact journals. Award based system provides for a healthier atmosphere in scientific funding, since academic merits are favorized. Yet, even positive measures should be emplaced with caution. They only give results when basic core funding is guaranteed. Using formal models Le Maux, Necker & Rocaboy (forthcoming in Research Policy, 2019) showed that only awarding publications in highly-cited journals leads to the loosening of scientific norms and in the last stance scientific misconduct. To overcome this problem, they proposed giving financial incentive for all publications including the ones in journals of lower rank.

Additional pressure leading to misconduct about scientific results comes from the fact that positive results are overrepresented in scientific journals. In synergy with publication dependent funding, the pressure on researcher to test only safe hypotheses increased. As specific measure to avoid favouring positive research outcomes, some journals allow for publishing registered reports. A registered report is reviewed on the basis of the originality and quality of the proposed methods and hypotheses, while the journal guarantees the publication of the results no matter whether the outcome will be positive or not. The system of registered reports does not work equally well for all research questions, e.g., it is not best for exploratory studies, but it represents a stepping stone in the right direction.

The ideal funding system is one based on merits, where merits are defined as efforts that follow norms of good conduct. Since scientific discoveries are unpredictable, these efforts might not always be fruitful. While basic and stable funding can be provided for a larger number of researchers, thus guaranteeing existential security, special merits should still be respected and rewarded. Finally, the selection of scientists should be based on their research and motivation, but also on their academic honesty, which is both an epistemic and character virtue.

Vlasta Sikimić
University of Belgrade

Events

November

RaE: Reasoning About Evidence, University of Ghent, 4–6 November.

PT&O: Post-Truth and the Objectivity of Epistemic Norms, University of Sussex, 7–8 November.

KoK: Kinds of Knowledge, University of Connecticut, 15–16 November.

WSS: What Sample Size do I Need?, Cass Business School, London, 25 November.

December

EaUM: Explanation and Understanding within Mathematics, Vrije Universiteit Brussel, 5–6 December.

D-SiL: Decision-Support in Litigation, University of Edinburgh, 6 December.

CML: Causal Machine Learning, Vancouver, 13–14 December.

PoS&PoM: Philosophy of Science and Philosophy of Mind, University of Edinburgh, 16 December.

January

MetaExp: Metaphysical Explanation III, Lund University, 8–9 January.

PWoDD: Practical Workshop and Data Dive, Belfast, 21–22 January.

Courses and Programmes

Courses

SSA: Summer School on Argumentation: Computational and Linguistic Perspectives on Argumentation, Warsaw, Poland, 6–10 September.

Programmes

APhil: MA/PhD in Analytic Philosophy, University of Barcelona.
Master Programme: MA in Pure and Applied Logic, University of Barcelona.
Doctoral Programme in Philosophy: Language, Mind and Practice, Department of Philosophy, University of Zurich, Switzerland.
Doctoral Programme in Philosophy: Department of Philosophy, University of Milan, Italy.
LogiCS: Joint doctoral program on Logical Methods in Computer Science, TU Wien, TU Graz, and JKU Linz, Austria.
HPSM: MA in the History and Philosophy of Science and Medicine, Durham University.
Master Programme: in Statistics, University College Dublin.
LoPhiSC: Master in Logic, Philosophy of Science and Epistemology, Pantheon-Sorbonne University (Paris 1) and Paris-Sorbonne University (Paris 4).
Master Programme: in Artificial Intelligence, Radboud University Nijmegen, the Netherlands.
Master Programme: Philosophy and Economics, Institute of Philosophy, University of Bayreuth.
MA in Cognitive Science: School of Politics, International Studies and Philosophy, Queen's University Belfast.
MA in Logic and the Philosophy of Mathematics: Department of Philosophy, University of Bristol.
MA Programmes: in Philosophy of Science, University of Leeds.
MA in Logic and Philosophy of Science: Faculty of Philosophy, Philosophy of Science and Study of Religion, LMU Munich.
MA in Logic and Theory of Science: Department of Logic of the Eotvos Lorand University, Budapest, Hungary.
MA in Metaphysics, Language, and Mind: Department of Philosophy, University of Liverpool.
MA in Mind, Brain and Learning: Westminster Institute of Education, Oxford Brookes University.
MA in Philosophy: by research, Tilburg University.
MA in Philosophy, Science and Society: TiLPS, Tilburg University.
MA in Philosophy of Biological and Cognitive Sciences: Department of Philosophy, University of Bristol.
MA in Rhetoric: School of Journalism, Media and Communication, University of Central Lancashire.
MA programmes: in Philosophy of Language and Linguistics, and Philosophy of Mind and Psychology, University of Birmingham.
MRes in Methods and Practices of Philosophical Research: Northern Institute of Philosophy, University of Aberdeen.
MSc in Applied Statistics: Department of Economics, Mathematics and Statistics, Birkbeck, University of London.
MSc in Applied Statistics and Datamining: School of Mathematics and Statistics, University of St Andrews.
MSc in Artificial Intelligence: Faculty of Engineering, University of Leeds.
MSc in Cognitive & Decision Sciences: Psychology, University College London.
MSc in Cognitive Systems: Language, Learning, and Reasoning, University of Potsdam.
MSc in Cognitive Science: University of Osnabrück, Germany.
MSc in Cognitive Psychology/Neuropsychology: School of Psychology, University of Kent.
MSc in Logic: Institute for Logic, Language and Computation, University of Amsterdam.
MSc in Mind, Language & Embodied Cognition: School of Philosophy, Psychology and Language Sciences, University of Edinburgh.
MSc in Philosophy of Science, Technology and Society: University of Twente, The Netherlands.
MRes in Cognitive Science and Humanities: Language, Communication and Organization: Institute for Logic, Cognition, Language, and Information, University of the Basque Country (Donostia San Sebastián).
Open Mind: International School of Advanced Studies in Cognitive Sciences, University of Bucharest.
Research Master in Philosophy and Economics: Erasmus University Rotterdam, The Netherlands.

Jobs and Studentships

Jobs

Post-doc: in Theoretical Philosophy, University of Konstanz, Germany, deadline open until filled.
Post-doc: in Statistical Inference for Mechanistic Models, University of Nottingham, deadline 5 November.
Lecturer: in Philosophy of Artificial Intelligence/Cognitive Science, University of Kent, deadline 10 November.
Tenure track position: in Applied Bayesian Statistics, University of Oulu, Finland, deadline 15 November.