The Reasoner

Volume 12, Number 9
September 2018

thereasoner.org
ISSN 1757-0522

Contents

Editorial

Dear Readers,

as you know well, our goal is to give you a monthly selection of reasoning bites, delivered to your inbox in a latex-compiled, unretweetable and unlikeable pdf. We received a number of suggestions to the effect that The Reasoner should go social. We thought a lot about those kind suggestions, but we felt we should keep it this way – after all reasoning benefits from taking it slow.

One slight change that we think will improve The Reasoner however, is to do with theming some issues. We would like to start with a number whose Features will be devoted to *Reasoning about Algorithms*. This is a very broad topic which can be tackled from a very wide array of perspectives, and of course we are interested in all of them! The Reasoner gathers a strong multidisciplinary community and few topics need all voices to be heard more than this. From the foundational questions on accountability, auditing and fairness, to the applications in cryptocurrencies to those in medicine (as discussed brilliantly below by our very own Daniel Auker-Howlett), we have an urgent need to understand what do we reason about when we reason about algorithms. If you would like to guest-edit the *Reasoning about Algorithms* issue, please get in touch with us by September 30th 2018 at features@thereasoner.org

If you'd rather send us a contribution for the themed issue, please follow the standard editorial guidelines of The Reasoner, but make sure you specify "Reasoning about Algorithms" in your submission which should be sent to us also by September 30th 2018.

Hykel Hosni
Università di Milano

Features

Theistic Ethics: A Conceptual Cartography

In his book, *Morality* (1972), Bernard Williams concludes the chapter on the relationship between morality and religion with the statement that the 'trouble with religious morality comes not from morality's being inescapably pure, but from religion's being incurably unintelligible.' (p. 86) In his later work, *Ethics* *and the Limits of Philosophy* (1985), he renews his criticism of religious ethics by remarking that 'the development of the ethical consciousness means the collapse of religion' not because a religious ethics 'is logically debarred from being ethical' but rather for the dialectical reason that 'if the self-understanding of religion is not to be left behind by the ethical consciousness, it has to move in a direction that will destroy religion.' (p. 33) However, in view of the recent rise of intense interest in the relationship between morality and religion, this paper is devoted to delineating some major logical and conceptual hurdles on the frontiers and terrain of theistic ethics. Without a reasonably respectable resolution of these logico-conceptual difficulties, the practitioners and products of theistic ethics cannot avoid becoming, in Williams' word, 'unintelligible'.

One may start by first recognizing that the concept of theistic ethics is multivocal in the sense of being subject to two broad types of scrutiny. First, the idea of theistic ethics may be examined through two different approaches:

(A) Theistic ethics as a set of moral values whereby a type of normative ethics based on God and religion is being offered

and

(B) Theistic ethics as a source of moral values whereby a type of meta-ethics founded on Divine Command Theory is being offered.

Second, one may examine the idea of theistic ethics from two perspectives:

(a) Theistic ethics from an internal (*ab intra*) or "within the community of believers" perspective

and

(b) Theistic ethics from an external (*ab extra*) or "without the community of believers" perspective thereby involving the community of "all". There are thus four possible combinations of examining the content and character of theistic ethics.

In the first stage of sketching this four-fold schema, one may look at theistic ethics by applying the internal perspective to the two foregoing approaches: that is, probing the problems and issues arising from taking theistic ethics as a normative theory and then as a meta-ethical theory. From an internal perspective, the normative interpretation of theistic ethics faces two central concerns:

(i) The phenomenon of moral difference and disagreement among believers in terms of what ethical values are sanctioned by the scriptural sources. Patently the fact of moral difference and divergence – possibly as a consequence of the variety of scriptural interpretation due to, for example, context sensitivity of understanding – is not in itself a critical cause of concern until one appreciates the *absence* of a relevant *decision procedure* to resolve such conflicts and clashes. The problem is heightened if, as the result of the development of our ethical consciousness, we come to subscribe to the doctrine of ethical conflict-regulation whereby it is stipulated that moral requirements must be capable of authoritatively regulating ethical conflicts.

(ii) A related, though separate, second problem in the same category is what may be called *Abraham's Sorites* where the prophet presented God with a sorites series in his intercession for Sodom when God was intent on destroying the *whole* cities and Abraham posing the question: 'Will you indeed sweep away the righteous with the wicked?' (*Genesis* 18: 20-33). Abraham thus sets up the sorites by asking God: 'Suppose there are fifty righteous within the city; will you then sweep away the place and not forgive it for the fifty righteous who are in it?' God's response to Abraham is: 'If I find at Sodom fifty righteous in the city, I will forgive the whole place for their sake.' Immediately Abraham unleashes the sorites by asking God: 'Suppose five of the fifty righteous are lacking? Will you destroy the whole city for lack of five?', to which God says, 'I will not destroy it if I find forty-five there.' Abraham seizes upon this and pushes the number of righteous down to forty and thereby engages God in a sorites of highest ethical proportion! In the extant text, Abraham goes as far as ten righteous people with the same response from God and then suddenly the conversation is brought to an end without a clear indication of what if, for example, there is only one righteous person in the city. Wittingly or otherwise, what is very significant about this sorites by Abraham is to highlight the complexity and intricacy of the *epistemology* of making moral judgments. How does one make a moral judgment?

This concern conveniently connects to the third issue (iii) arising from the application of the internal perspective to theistic ethics not as a normative theory but as a metaethical theory when Abraham in his pleading with God says that it 'is impossible' for God to 'kill the innocent with the guilty'. In other words, Abraham is setting a *constitutive constraint* on the nature of divinity by requiring that the 'judge of all the earth *has* to act justly' (emphasis added), thereby implying, if not declaring outright, the independence of a significant *source* of morality from God. In fact, it is not surprising that historically one comes across sects of, for example, various Abrahamic or Semitic religions that explicitly impose the condition of justice on divinity such as Karaites in the Jewish tradition and Motazalites among Muslims.

In the second stage of using the four-fold division, one may look at theistic ethics by applying the external perspective to it as a normative theory and then as a metaethical theory. From an external perspective, the normative interpretation of theistic ethics faces the following five interrelated issues: (iv) Problem of Universality: how universal are the moral values emanating from a theistic normative theory? (v) Problem of Compatibility: how compatible are the moral values of a theistic normative system with non-theistic moral values? (vi) Problem of Partiality: how does a theistic normative theory account for the *partiality* promised and presumed in such frameworks towards certain chosen or favored people? (vii) Emergence of New Moral Values: how does a theistic normative account handle and regulate the advent of new moral values? (viii) Problem of Incompleteness: how does a theistic normative theory explain one of the lessons of the process of what Williams calls 'the development of the ethical consciousness' that our moral outlook is ultimately incomplete?

Finally, in covering the last step of the four-fold permutations, the external perspective as applied to theistic ethics in the form of a metaethical theory draws our attention to the following six subjects: (ix) Euthyphro Dilemma: this is too classical an issue to need any articulation! (x) Abraham's Sorites: in this incarnation, the sorites can be used to sow the seeds of moral skepticism and consequently to cast doubt on the viability of any theistic ethics. (xi) Problem of Subjectivity: can God know what it feels like to be a non-divine moral agent and thereby questioning the fairness or justness of God's sitting in judgment on such moral agents? This problem has an interesting connection with a variant of the paradox of omnipotence, viz., the paradox of sin: can God commit sin? (xii) God's Command of Moral Values and Omniscience: does God really know what moral commands to make in light of the occurrence of *divine regret*? Having observed 'how wicked everyone on earth was and how evil their thoughts were all the time', God laments that 'he was *sorry* that he had ever made them and put them on the earth. He was so filled with *regret* that he said, "I will wipe out these people I have created, and also the animals and the birds, because I am *sorry* that I made any of them." (*Genesis* 6: 5-7, emphasis added) Generally, there is a tremendous tension between omniscience and regret. (xiii) God's Existence and Moral Motivation: can the existence of God provide motivation for acting morally? This question actually manifests itself in three different forms: (1) the Socratic version in the form of Euthyphro Dilemma, (2) the Kantian version in the form of Categorical Imperative, and (3) the Humean version in the form of "is/ought" or naturalistic fallacy. (xiv) God's Creation of Moral Agents: can God create moral agents that freely always choose the good? Obviously, the question has an important overlap with the traditional problem of evil. And, finally, (xv) Problem of Moral Luck: if, as part of the development of the ethical consciousness, we have come to realize the significance and impact of moral luck on our actions and inactions, how does a theistic ethics deal with *this* pervasive trait of our lives?

Having identified a number of key logico-conceptual challenges in both normative and metaethical components of theistic ethics from an internal as well as external perspective, one may cautiously conclude: without some satisfactory grip on these problems, the theoretical and practical tenets of theistic ethics seem unable to account for the profoundly human phenomenon of ethical consciousness.

Majid Amini
Philosophy, Virginia State University

The Reasoner Speculates

Mathematical shortcomings in a simulated universe

I present an argument that for any computer-simulated civilization we design, the mathematical knowledge recorded by that civilization has one of two limitations. It includes falsehoods, or it is weaker than our own mathematical knowledge. This is paradoxical because it seems that nothing prevents us from building in all sorts of advantages for the inhabitants of said simulation.

Imagine we were to program a simulation of a human-like civilization. Imagine the simulation were sophisticated enough that its inhabitants began stating mathematical axioms. Imagine we were to contrive that a monolith descends to earth inside the simulation. A monolith equipped with a keyboard and with instructions. The instructions implore them to use the keyboard to record mathematical axioms into the monolith. They are instructed to always continue recording the strongest mathematical axioms they can think of (in, say, the language of set theory). And they are warned never to record any mathematical falsehood into the monolith.

Having designed this simulation, we would know its source-code. From that source-code, we could infer a code for the list L of all axioms which are ever recorded into the monolith.

Could we trust the statements in L to be true? A priori, the answer depends on the simulation. At one end of the spectrum, the inhabitants might be afraid of the monolith and never record anything: their assertions would vacuously be true. On the other extreme, they might immediately record "1 = 0", just to see what happens.

If we could trust the truth of the inhabitants' statements, then we would necessarily be smarter than those inhabitants. We would be smarter than them because we would know a mathematical truth Φ which dominates all the axioms they ever record: namely,

$$\Phi = \text{“all statements in } L \text{ are true”}$$

(there are technical problems with this, which I address later).

Here's a paradox. Suppose we engineer the simulation in such a way that its inhabitants have vast motivation and resources to do mathematics. Let your imagination run wild (limited only by Turing computability).

- We could program the inhabitants to have huge brains.
- We could endow their world with vast energy reserves.
- We could program the world to constantly encourage the recording of strong mathematical axioms. For example, we could embed encouraging messages into the simulation, like in the paper by Steven Hsu and Anthony Zee (2006: Message in the Sky, https://arxiv.org/pdf/physics/0510102.pdf, Modern Physics Letters A **21** (19), 1495–1500).
- We could engineer the simulation to discourage the recording of falsehoods. For example, maybe the monolith is preceded by fake monoliths which punish people for asserting pre-selected falsehoods (we would have to pre-select some falsehoods because the full set of falsehoods is non-computable).

In spite of all such contrivances, the inhabitants' recorded mathematical knowledge would fall short in at least one of two ways: it is untrustworthy, or it is weaker than our own. This seems paradoxical because we ourselves do not have such huge brains, such vast energy reserves, or such great motivation.

Now I will address three technical problems.

- (Impracticality) Running the simulation might be prohibitively expensive. This is irrelevant: we do not need to actually run the simulation to reason about what would happen if we did run it. We only need to know its source-code. (Since it is a computer simulation, it is deterministic.)
- (Undefinability of truth) Even if we know Φ, we might not be able to *state* it in the language of set theory (which we would need to do in order to meaningfully contrast Φ

against the statements in L). We can avoid this problem by using codes for computable ordinals. Computable ordinals have the property that there is a computable function f such that, given a code n for a computably enumerable list of codes of computable ordinals $\alpha_1, \alpha_2, \ldots$, $f(n)$ is a code for a computable ordinal α greater than all the α_i. Having the source-code for the simulation, we could infer a code n for the list of all numbers k such that the statement "k is a code of a computable ordinal" is implied by the inhabitants' axioms. If we know the inhabitants do not assert false axioms, we can infer $f(n)$ is a code of a computable ordinal. Thus, we know a computable ordinal bigger than all the computable ordinals the inhabitants know.

- (Self-reference) Perhaps, by diagonalization and self-reference, L already contains a statement equivalent to Φ. This objection can be defeated by the same computable ordinal trick used above.

(The above computable ordinal trick is reminiscent of how ordinals were used by I.J. Good (1969: Gödel's Theorem is a Red Herring, https://doi.org/10.1093/bjps/19.4.357, The British Journal for the Philosophy of Science **19** (4), 357–358) to shed light on the Lucas-Penrose argument.)

The more incentives and resources we give them to record strong axioms, the more likely the inhabitants will accidentally record a falsehood. The more we discourage them from recording falsehoods, the more afraid they will be to record strong axioms. Our argument suggests this dilemma is irreconcilable. We are incapable of designing computable incentives and resources strong enough to simulate a civilization smarter than us while balancing those with computable disincentives strong enough to ensure its trustworthiness.

The above reasoning makes us wary of consensus in our own world. We acknowledge that false axioms are occasionally proposed, and this is okay because over time they are corrected. But suppose we decide that once per millenium, we shall single out one new axiom we're really really sure of, never to be revised. Knowing the risk, we take many precautions to avoid choosing an error. Whatever incentives we impose on ourselves to avoid error, without crippling the strength of the chosen axioms, such safeguards apparently must fail: if they could succeed for us, then why not for the people we simulate?

Samuel Alexander

News

The 7th biennial conference of the Society for Philosophy of Science in Practice, 30 June–02 July

The 7th biennial conference of the Society for Philosophy of Science in Practice (SPSP 2018) was held from June 30 until July 02 at Ghent University, Belgium, and was preceded by a pre-conference workshop on June 29. SPSP 2018 was sponsored by the Research Foundation Flanders (FWO) and the research fund of the Faculty of Arts and Philosophy of Ghent University. SPSP 2018 was organized by the Centre for Logic and Philosophy of Science, which is part of the Department of Philosophy and Moral Science.

SPSP advocates a philosophy of scientific practice, which takes scientific theory, scientific practice, and the world into consideration simultaneously, rather than only theory-practice relations or theory-world relations. SPSP aims to offer a detailed and systematic study of scientific practices. This study includes, inter alia i) the usage of scientific knowledge and how this knowledge is shaped by its intended usage, ii) how artifacts, such as laboratory instruments and models, mediate be-tween theory and the world and how they shape scientific practice, iii) in and across all scientific disciplines, iv) and assessed from both a history and philosophy of science perspective, in interaction with practicing scientists, engineers, and policy makers.

A major venue for the discussion and dissemination of results on the philosophical study of scientific practices is the SPSP conference series. The conference brought together 220 participants from all over the world who presented talks and posters on a great variety of topics related to the philosophical study of scientific practices.

The pre-conference workshop focused on the use of case studies and examples in philosophy of science. The thematically organized presentations of the main conference covered a great variety of issues, many more than can be mentioned here. These issues included the diverse uses of models and other representational tools in a variety of scientific domains. Another set of presentations focused on how scientific practice bears on metaphysical issues, such as the individuation of genes. Other issues that were discussed included the role of values in scientific investigation, issues concerning scientific experimentation, causality, types of scientific explanation (e.g., mechanistic explanation, causal explanation), policies in science, the use of big data in science, and the study of scientific practice from an integrated history and philosophy of science perspective. These issues were discussed in the context of many scientific domains, ranging from astrophysics to archeology, cognitive science, medicine, engineering, and biology.

The keynotes were delivered by William Bechtel (University of California at San Diego), Sabina Leonelli (University of Exeter), Maarten van Dyck (Ghent University), and Alison Wylie (University of British Colombia).

In his keynote, *Visualization practices*, William Bechtel gave on overview of visual representations used by scientists to formulate hypotheses, represent experimental instruments and protocols, present data, and communicate conclusions, and subsequently assessed two specific scientific visualization practices in more detail: the growing practice of publishing visual abstracts and the development of platforms to represent and analyze networks.

In her keynote, *Philosophy of Science in the age of big and open data: what does studying practice involve?*, Sabina Leonelli examined three (intertwined) modes of analysis characteristic of philosophy of science in practice: i) the study of science aimed at identification of concerns and features of research in such a way that these can be examined empirically, ii) critical reflection on the epistemological, social, ethical and/or ontological significance of the practices and output of scientific research, and iii) the engagement of society in these research endeavors. She illustrated these modes through philosophy of big and open data.

In his keynote, *Did Galileo's experiments confirm the law of fall?*, Maarten van Dyck elaborated the historical debate on the question whether Galileo actually executed experiments with bodies descending from an inclined plane and, in case he did, whether their results could confirm his law of fall. In this exposition, Maarten van Dyck sketched the development of Galileo?s thinking and the roles played by experimentation therein and in the confirmation of the law of fall. Based on this analysis, some general reflections were offered on the notions of 'confirmation' and 'scientific revolution'.

In her keynote, *Philosophy in the field: witnessing and translating*, Alison Wylie elaborated a research cluster, Indigenous/Science, based at the University of British Columbia, that builds collaborative partnerships between archeological science and Indigenous Peoples that are designed to bring the tools of the former to bear on indigenous-led research questions in a way that is respectful and reconciliatory. Alison Wylie traced the background of Indigenous/Science in terms of several key junctures in her own practice-focused investigations, and explored the question what philosophers can contribute to these endeavors.

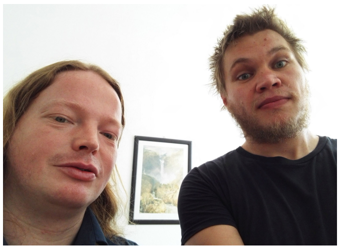
Dingmar van Eck
Erik Weber
Ghent University

Calls for Papers

Formalization of Arguments: special issue of *Dialectica*, deadline 31 July.
Reliability: special issue of *Synthese*, deadline 11 November.
Instrumentalism about Epistemic Rationality: For and Against: special issue of *Synthese*, deadline 30 October.

What's Hot in ...

Mathematical Philosophy



An imprecise probability model $\underline{P}$ is called two-monotone if it satisfies the inequality $\underline{P}(A) + \underline{P}(B) \leq \underline{P}(A \cap B) + \underline{P}(A \cup B)$ for arbitrary events A and B. A much quoted phrase from Peter Walley, *the* luminous figure in imprecise probabilities, is that he doesn't *"...know any rationality argument for two-monotonicity, beyond its computational convenience"*. And obviously: Using a model $\underline{P}$ that is two-monotone is computationally very convenient, since for instance lower and upper expectations of arbitrary (bounded) gambles can be computed by closed formulas. So, more interestingly: What about the second part of the statement, the "beyond" part? While many of the classical and well-motivated imprecise probabilistic models, such as belief functions induced by set-valued mappings or ε-contamination models, turn out to be two-monotone anyway and, accordingly, avoid the necessity of an explicit such assumption, the statement still raises the following question: Are there reasons for assuming two-monotonicity when it is not a direct implication of some well-motivated model? In other words: Are there reasons for explicitly assuming two-monotonicity as a model primitive? Indeed, there are. For instance in financial risk measurement the property of two-monotonicity is neatly related to the (in a certain sense essentially equivalent) property of comonotone additivity: Informally speaking, two random variables X and Y are comonotone if they can be interpreted as being increasing functions of a common risk driver and the "main idea behind comonotonic risk measures is that merging two comonotone positions X and Y will not lead to diversification and, hence, the amount of risk capital required for the aggregated position $X + Y$ should correspond to the sum of the individual capital requirements" [Koch-Medina, Munari and Svindland (2017). Comonotonic risk measures in a world without risk-free assets. *Technical report, University of Zurich*].
Two-monotonicity often plays an important technical role when it comes to generalizing classical concepts to imprecise probabilities and as both of us were scientifically socialized at a department of statistics, we would like to briefly scratch on the implications of two-monotonicity in the context of generalizing first order stochastic dominance to imprecise probabilities. Therefore, first recall that for two random variables X and Y mapping from some probability space $(\Omega, \mathcal{A}, P)$ to some linearly ordered set $(A, \geq)$, variable X is said to (weakly) stochastically dominate Y if every expected utility maximizer whose utility respects the ordering $\geq$ weakly prefers X to Y, i.e. if for every $\geq$-isotone utility function $u : A \to \mathbb{R}$ the expected utility $\mathbb{E}_P(u \circ X)$ of X is larger than or equal to the expected utility $\mathbb{E}_P(u \circ Y)$ of Y. Generalizing stochastic dominance to imprecise probabilities can be done in (at least) three different ways: A first approach is to replace the probability measure P by some non-additive measure $C : \mathcal{A} \to [0, 1]$ (sometimes also called a *capacity*). Then, in order to generalize stochastic dominance we require a concept of integration with respect to capacities, where the most common choice (outside the Walley universe) seems to be the (upper) Choquet integral [Choquet (1953). Theory of capacities. *Annales de l'Institut Fourier*, 5, 131–295]. A straightforward generalization, developed and analyzed in [Dyckerhoff and Mosler (1993). Stochastic dominance with non-additive probabilities. *Zeitschrift für Operations Research*, 37(3), 231-256], then can be obtained by using the same definition as above, where, however, the integral $\mathbb{E}_P(u \circ X)$ is replaced by the Choquet integral $\mathbb{E}_C(u \circ X)$. Instead of working with non-additive measures, one can alternatively rely on replacing P by the *credal set* $\mathcal{M}$ of all probability measures that are in accordance with the information about the underlying uncertainty. A second approach then would be to generalize stochastic dominance by replacing the expectation $\mathbb{E}_P(u \circ X)$ through the associated *upper envelope* $\sup_{P \in \mathcal{M}} \mathbb{E}_P(u \circ X)$, i.e. the supremum of all expectations corresponding to compatible probabilities (this construction is closely related to the so-called *natural extension*, which is the common "integral" within the Walley universe). A third approach is obtained by demanding that $\mathbb{E}_P(u \circ X) \geq \mathbb{E}_P(u \circ Y)$ for every $P \in \mathcal{M}$ and every $\geq$-isotone utility function u.

But how to decide for one of the approaches in a specific situation? First, it is useful to observe that if the underlying credal set generates upper envelopes that are two-monotone, then approaches one and two coincide. However, as Walley indicates in the quote, assuming two-monotonicity/comonotone additivity is not always self-evident and in cases where the two approaches do not coincide there are certain arguments of rationality that frame the use of upper envelopes as the only reasonable approach and, in particular, as the only approach with a clear behavioral interpretation. Accordingly, a good way could be to start with considering envelopes of credal sets and use the mathematical toolbox for Choquet integration only if these envelopes happen to be two-monotone. Note that, of course, Schmeidler's axiomatic approach [Schmeidler (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3), 571-587] also gives behavioral meaning to the Choquet integral, but it uses comonotonic independence as a model primitive and thus only gives justification for using the Choquet integral for situations where the approaches coincide anyway. On the other hand, as is shown for example in [Papamarcou and Fine (1996). A note on undominated lower probabilities. *The Annals of Probability*, 2(14), 710-723], there are situations in which there exists no single probability measure that is compatible with a given capacity. Clearly, in such situations the generalization of stochastic dominance via the Choquet integral is still well-defined whereas approaches two and three are not.

Whatever a decision for or against a particular approach may turn out, within each approach one actually has to decide about further subtleties: For example within the Choquet approach one has to decide, if one wants to use the upper or the lower Choquet integral, the same for the credal set approach. And for the third approach, what is the exact rationale for demanding $\mathbb{E}_P(u \circ X) \geq \mathbb{E}_P(u \circ Y)$ for all $P \in \mathcal{M}$ and not only for one P? All these issues led to a plethora of proposed generalizations in the literature, see e.g., [Langewisch, Choobineh (1996). Stochastic dominance tests for ranking alternatives under ambiguity. *European Journal of Operational Research*, 95(1), 139-154] or [Denoeux (2009): Extending stochastic ordering to belief functions on the real line. *Information Sciences*, 179(9), 1362-1376], to name just a few further sources.

And as if that conceptual reservations weren't enough, the computational elegance of using Choquet integration looses its bite if it comes to looking at the situation, where the space $(A, \geq)$ is only *partially* ordered: In this case, one seemingly cannot derive easy to evaluate characterizations of stochastic dominance.

Therefore, without having a good answer to it, we would like to reformulate Walley's quote as: *"Are there any rigorous reasons for demanding two-monotonicity or comonotone additivity beyond diversification arguments?"*

Christoph Jansen and Georg Schollmeyer
Munich Centre for Mathematical Philosophy

Evidence-Based Medicine

A recent study found that an artificial intelligence (AI) system using novel deep learning architecture was able to intepret and refer eye scans with comparable accuracy to experts (5.5% error rates, compared to 6.7% and 6.8% for the two best performing experts in the study). This system was produced by Google's DeepMind AI company in collaboration with Moorfield Eye Hospital in London.

The eye scans being assessed were produced by Optical Coherence Tomography (OCT), a technique that measures reflection of near infrared light at a resolution of ~5μm. Opthalmologists can use this technique to form high-resolution 3D images of cross sections of the tissue at the centre of the retina, the macula. It is a technique analogous to 3D ultrasound but using light instead of sound, and is one of the most common imaging procedures currently in use: 5.3 million OCT scans were performed on the US medicare population in 2014, and it is widely and routinely used on the NHS as part of an initial assessment and triage of patients with acute and chronic sight loss. The AI system was fed 877 OCT scans to 'train' it to convert raw scans into a 'detailed tissue segmentation map'. 14,884 OCT scans were then segmented by the first stage of the system, with the resultant maps used to enable the system to 'learn' how to classify the segmentation map to provide diagnosis and referral outputs. Novel OCTs were then inputted to the system, and outputs were then compared against expert judgements. The Novel OCTs were obtained from case reports, and the assessments in the study were matched up against the case report outcomes (which were designated the 'gold standard' reference). Error rates were thus calculated against the actual outcomes of the case reports. This is a brief detailing of the system, and readers interested in the workings of the AI system are referred to the study.

I would like to focus instead on some notable medical implications, beyond the findings of comparable accuracy. The first medically important thing to note is that while the title of the study refers to diagnosis, the AI system here did not diagnose eye conditions. Rather it interpreted and triaged them as urgent, semi-urgent, routine and observation. It was this classification that matched the performance of experts. This does not mean the results of the paper are not important, but it does thankfully (for doctors at least) mean we are not replacing doctors just yet. If the results of this study hold then the implementation of the technology promises to speed up an existing process that places excessive demands on already strained services. Eye-related diseases are increasing due to an aging population and it has been suggested that the UK alone will "experience a growth in...numbers of about 44% for glaucoma, 50% for operable cataract, and 59% for AMD [Age-Related Macular Degeneration]". On top of this, there is a distinct lack of experts available to interpret and triage the increasing number of OCT scans performed, resulting in some alarming false referral statistics, e.g 40%-false positive referalls for suspected glaucoma. An AI system that can perform this triage task would speed up the process and ease the pressure on eye-care health professionals, leaving them more time to perform the diagnosis and treatment they do-out perform machines at.

Another interesting positive is very relevant to the aims of EBM. This is that the judgement process performed by the AI system is "readily viewable by a clinical expert and integrates into clinical workflows". Typically, deep learning systems face a black-box problem: the process by which results are arrived at is not available to those having to work with the outputs. The DeepMind system has a structure that closely follows the clinical-decision making process, "separating judgements about the scan itself from the subsequent referral decision." Clinicians are thus not merely presented with the diagnosis and referral suggestion, but also with the segmentation maps and subsequent classifications. This makes the whole procedure explicit, and opens up the 'reasoning' process of the AI to evaluation. Clinical judgements (by people) are always recommended to be made in a clear and transparent manner, citing evidence and reasons. Not only can this AI system potentially alleviate pressure on a strained healthcare system it could do so in a way that meets the foundational principles of EBM. This has the added consequence that the system could be used in an educational and training setting, as the maps can be used to show the relevant structures to identify for different pathologies.

Surely these results means we should be ramping up production and implementation? The authors are now seeking evidence for efficacy of the framework in an RCT. This may seem like an extra uneccessary step considering the high degree of accuracy here and the fact that this is not a medical intervention. However, an RCT would offer evidence on whether the accuracy of the system results in clinically important outcomes. So far, accuracy is proven only against the 'gold standard' reference, and we have no data on whether patients would benefit from being referred by the system. A final remaining issue is that this study was conducted in part by DeepMind employees, who obviously have a vested interest in the project's success. Publication bias is a risk here, and so independently performed trials may mitigate this risk. If an independent RCT does show that this system benefits patients, then because of it's ability to improve efficiency in a strained health-care system, this is a promising avenue for AI to have a positive impact on health more broadly.

D.J. Auker-Howlett
Philosophy, Kent

Events

September

PLP: The 5th Workshop on Probabilistic Logic Programming, Ferrara, Italy, 1 September.
ILP: Inductive Logic Programming, Ferrara, Italy, 2–4 September.
KoR: Kinds of Reasoning, Vercelli and Novara, Italy, 3–4 September.
PhilSci: Workshop and Conference on Philosophy-Science, University of Aix-Marseille, 4–6 September.
GrammInfer: International Conference on Grammatical Inference, Poland, 5–7 September.
BLTfCDMBayesian Learning Theory for Complex Data Modelling, Grenoble, France: 6–7 September,

.

DegBel: What Are Degrees of Belief?, University of Leeds, 11–12 September.
BayBtS: Bayes By the Sea: Formal Epistemology, Statistics, and Probability, Anconca, Italy, 13–14 September.
WAW: Warsaw Argumentation Week, Warsaw, Poland, 6–16 September.
EPINON: Epistemology in Ontologies, Cape Town, South Africa, 17–18 September.
Logi&QuanPhy: Logic and Quantum Physics, Universitá della Svizzera Italiana, 21–22 September.

October

FAaAHR: Bridging the Gap Between Formal Argumentation and Actual Human Reasoning, Ruhr University Bochum, 4–5 October.
PITEaAiA: Philosophical Imagination, Thought Experiments and Arguments in Antiquity, University of Maribor, Slovenia, 9 October.
MLaEiCS: Machine learning and explanation in cognitive science, Institute of Philosophy, Czech Academy of Sciences, 15–16 October.
BayComp: Afternoon Meeting on Bayesian Computation, University of Reading, 17 October.
FormCaus: Formal Causation, Rostock, Germany, 22–23 October.
ARiS&M: Analogical Reasoning in Science and Mathematics, Munich, 26–28 October.
HRaL: Hybrid Reasoning and Learning, Tempe, Arizona, USA, 28 October.

Courses and Programmes

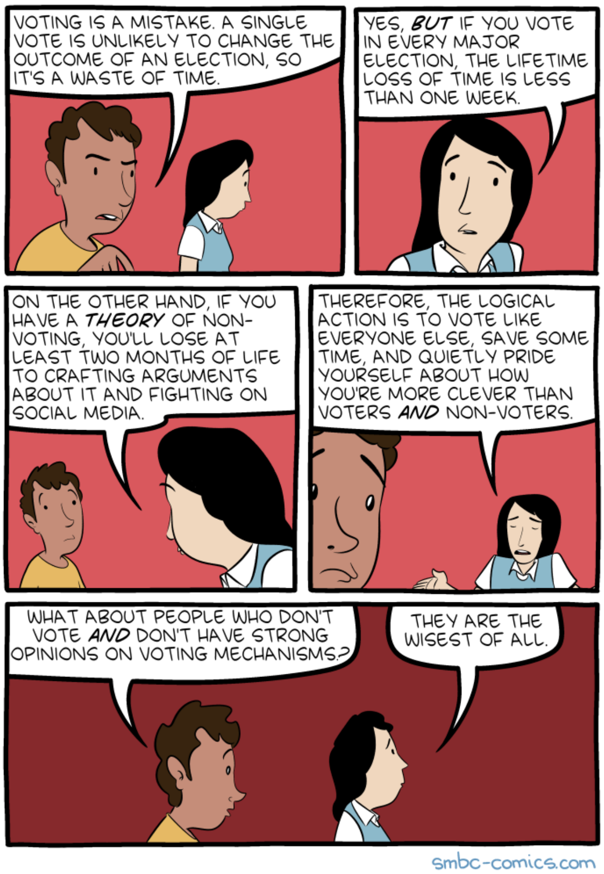
Courses

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

Lecturer: in Logic, Ludwig-Maximilians-University, Munich, deadline 1 September.
Assistant Professor: in Statistics, TU Delft, deadline 30 September.
Professor: in Theoretical Philosophy, University of Vienna, deadline 30 September.
PhD: in Computational Probability, TU Delft, deadline 31 October.