Extending Rationality

Emmanuel M. Pothos[1] (Emmanuel.pothos.1@city.ac.uk), Jerome R. Busemeyer[2] (jbusemey@indiana.edu), Tim Pleskac[3] (pleskac@ku.edu), James M. Yearsley[1] (James.Yearsley@city.ac.uk), Joshua B. Tenenbaum[4] (jbt@mit.edu), Noah D. Goodman[5] (ngoodman@stanford.edu), Michael Henry Tessler[4] (tessler@mit.edu), Thomas L. Griffiths[6] (tomg@princeton.edu) , Falk Lieder[7] (falk.lieder@tuebingen.mpg.de), Ralph Hertwig[8] (hertwig@mpib-berlin.mpg.de), Thorsten Pachur[8] (pachur@mpib-berlin.mpg.de), Christina Leuker[8] (leuker@mpib-berlin.mpg.de) & Richard M. Shiffrin[2] (shiffrin@indiana.edu)

[1]Department of Psychology, City, University of London, Northampton Square London, EC1V 0HB, UK
[2]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405, USA
[3]Tim Pleskac, Department of Psychology, University of Kansas, Lawrence, KS 66045, USA
[4]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[5]Noah D. Goodman, Department of Psychology, Stanford University, Stanford, CA 94305, USA
[6] Department of Psychology, Princeton University, Princeton, NJ 08540, USA
[7]Max Planck Institute for Intelligent Systems, Tübingen 72076, Germany
[8]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin 14195, Germany

Keywords: rationality, bounded rationality, fallacies, heuristics, resource-rational, probabilistic programming language, classical and quantum probability theory

Fallacies?

Since antiquity, we have wondered about the foundations of our (apparent) intellectual superiority. A way to approach this issue is to seek rational standards in decision making and examine convergence between such standards and behavior. However, establishing a rational framework is not straightforward. One of the most unique contributions of cognitive science is the varied perspectives it has provided for rationality. With many recent advances in decision theory (including novel techniques in probabilistic inference and sophisticated frameworks for heuristics-driven reasoning), it is particularly timely to reevaluate rational standards and our assumptions regarding rational behavior. This is the purpose of this interdisciplinary symposium, bringing together expertise in psychology, computer science, mathematics, physics, and philosophy of mind.

Cognitive science research has already instigated major shifts in our perception of rationality and optimality. For most of our history, it has been considered that classical logic is the source of human rationality and the appropriate normative standard against which to assess decisions. Wason sought a general test of whether natural reasoning is consistent with classical logic, by asking participants to select which evidence was best suited to test a given rule. Logic prescribes selections which can definitely falsify the rule (a falsificationist mentality which has had a pervasive influence in scientific reasoning, including in frequentist statistics), but instead participants selected evidence with potential to confirm the rule. Oaksford and Chater (1994) proposed that participants prefer the cards which minimize the information-theoretic uncertainty regarding the validity of the rule, employing Anderson's (1990) idea of optimal adaptation. Classical probability theory (CPT) thus revealed an alternative perspective for the 'correct' selections in Wason's task.

CPT is currently recognized as the right starting point for understanding rational decision making, benefiting from powerful formal justifications and excellent descriptive coverage. Equally, it has been increasingly appreciated that a *baseline* CPT framework is unlikely to provide either a complete descriptive framework for cognition or indeed an appropriate normative framework, without suitable extensions (e.g., Tenenbaum et al., 2011). One influential source of indication that this is the case concerns reports of persistent apparent violations of CPT principles, usually called fallacies. Tversky, Kahneman and their colleagues have produced some of the most evocative examples, for example, the conjunction fallacy, according to which naïve observers are quite happy to accept that *Prob*(A&B)>*Prob*(A) (Tversky & Kahneman, 1983). The most telling instantiation of this result involves the probability of a Scandinavian person having blue eyes and blond hair vs. just having blond eyes (Tentori et al., 2004). Imagining a line-up of Scandinavian individuals makes it immediately obvious why the conjunction fallacy is, well, a fallacy, and yet the conjunctive statement still feels natural – it is this persistence that makes fallacies so puzzling. There are several similar results. For example, a famous Gallup poll study showed a *Prob*(Clinton is honest) of 50% when this question was first but 57% after a similar question for Gore (Moore, 2002); in another famous study, a mixture of weak and strong evidence had less impact than just the strong evidence (the dilution effect; Nisbett et al., 1981).

Such findings appear to challenge our expectation of rationality. But do they have to? Over the last decades, new, sophisticated techniques and ideas have emerged, which require drastic revision to our perception of applicability of *baseline* CPT frameworks in thought. In this symposium we explore four approaches, some of which directly extend baseline CPT ideas while others are motivated from baseline CPT ideas to develop in more alternative directions, with sometimes surprising implications for empirical coverage and normative evaluation.

Resource-rational analysis: Griffiths, Lieder

Baseline CPT inference is expensive, and practical models often involve some kind of sampling-based approximation to posterior probabilities. In the tradition of bounded

rationality, the resource-rational analysis is about finding the optimal balance between the accuracy of probabilistic approximations and resource allocation, with the latter formulated in terms of computational cost (Griffiths et al., 2015). This approach can recover previously-identified heuristics and discover new ones, as well as shed light in the way resource limitations can lead to apparent deviations from CPT prescription.

Quantum: Busemeyer, Pleskac, Yearsley, Pothos

Another way in which CPT probabilistic inference can be made more tractable is by limiting the size of the probabilistic space. The logical structure of CPT is a Boolean algebra, but for Quantum probability theory (QPT) it is a partial Boolean algebra, which means a collection of smaller (simpler) parts, which are classical individually, but inconsistencies/ contextuality/ apparent fallacies arise when reasoning between parts. We think that QPT representations are more likely when e.g. participants are unfamiliar with a problem or unwilling to engage thoughtfully. We show how QPT can reveal rational perspectives to established fallacies (Pothos et al., 2017) and further consider whether QPT can shed light on rational status of behavior in strategic games, in situation when decisions appear inconsistent with the Nash equilibrium or sub game perfect equilibrium.

Heuristics: Hertwig, Pachur, Leuker

Rather than simplify or approximate CPT inference through e.g. more efficient sampling procedures, an alternative, influential possibility is that the mind adopts heuristics. Heuristics can be as accurate and sometimes even more accurate than strategies that employ the greatest possible amount of information and computation. Can such advantages generalize to situations involving interactions with other intelligent, competitive actors? We will explore the effectiveness of heuristics in stationary games against nature and in strategic games and show that heuristics are particularly competitive when the level of epistemic uncertainty is high. We will also consider in general the ecological structures that heuristics can harness, and how theories of heuristics can be integrated with other frameworks of human choice.

Probabilistic language of thought: Tenenbaum, Goodman, Tessler

An important extension to baseline CPT frameworks concerns incorporating language-like properties (such as compositionality), representations, and pragmatic reasoning in probabilistic inference. The probabilistic programming language (PPL) / probabilistic language of thought (PLoT) can more naturally apply to richer forms of reasoning, including everyday reasoning under uncertainty (e.g., Goodman et al., 2015). Furthermore, enriching these models with an understanding of natural language pragmatics can explain apparent fallacies in classical reasoning tasks (e.g., Tessler & Goodman, 2014). Assuming a communicative context to a task involving language allows a reasoner in a PPL/PLoT model to incorporate the goals of a speaker (e.g., assuming the speaker intends to be informative), so providing a rational perspective on reasoning fallacies. We will also consider the way resource limitations guide practical models in PPL.

Discussion: Shiffrin

The discussion part of the symposium will address these varying perspectives on rationality and bring together the themes raised in the presentations. The overarching questions concern what is rationality, and whether 'bounded rational' approaches capture enough of what humans mean by this concept. The discussion will be open to all presenters and the audience.

Acknowledgments

EMP was supported by ONRG grant N62909-19-1-2000.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Goodman, N. D., Tenenbaum, J. B., & Gerstenberg, T. (2015). Concepts in probabilistic language of thought. In Eds. E .Margolis & S. Laurence, *New Directions in the Study of Concepts*, pp.623-653. MIT Press: Cambridge.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7, 217-229.

Moore, D. W. (2002). Measuring new types of question-order effects. *Public Opinion Quarterly*, 66(1), 80-91.

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, 13, 248-277.

Oaksford, M. & Chater, N. (1994). A Rational Analysis of the Selection Task as Optimal Data Selection. *Psychological Review*, 101, 608-631.

Pothos, E. M., Busemeyer, J. R., Shiffrin, R. M., & Yearsley, J. M. (2017). The rational status of quantum cognition. *Journal of Experimental Psychology: General*, 146, 968-987.

Tenenbaum, J. B, Kemp, C., Griffiths, T. L., & Goodman, N. (2011). How to grow a mind: statistics, structure, and abstraction. *Science*, 331, 1279-1285.

Tentori, K., Bonini, N., & Osherson, D. (2004). The conjunction fallacy: a misunderstanding about conjunction? *Cognitive Science*, 28, 467-477.

Tessler, M.H., & Goodman, N. (2014). Some arguments are probably valid: Syllogistic reasoning as communication. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjuctive fallacy in probability judgment. *Psychological Review*, 90, 293-315.