

Heuristics, hacks, and habits: Boundedly optimal approaches to learning, reasoning and decision making

Ishita Dasgupta¹, Eric Schulz¹, Jessica B. Hamrick² & Joshua B. Tenenbaum³

¹Department of Psychology, Harvard University

²DeepMind, London, UK

³Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology

Humans regularly perform tasks that require combining information across several sources of information to learn, reason, and make decisions. Bayesian models provide a computational framework, and a normative account, for how humans carry out these tasks. However, exact inference is intractable in most real-world situations, and extensive empirical work shows that human behavior often deviates significantly from the Bayesian optimum. A promising possibility is that people instead approximate rational solutions using bounded available resources. In this workshop, we bring together leading researchers from cognitive science, neuroscience and machine learning to build a better understanding of boundedly optimality in how humans learn, reason and make decisions.

Keywords: Heuristics; Resource rationality; Reasoning; Decision making; Reinforcement learning; Machine learning

Introduction

This workshop will cover work that casts human and machine learning, decision making and reasoning as *boundedly optimal*. In particular, we will focus on meta-reasoning, reinforcement learning, active information acquisition, and probabilistic reasoning.

The notion that the mind approximates rational (Bayesian) inference has had a strong influence on thinking in psychology since the 1950s. However, people deviate from Bayesian ideals in several well-documented instances (Gilovich, Griffin, & Kahneman, 2002), giving rise to the idea that they rely on heuristic rules instead (Gigerenzer & Brighton, 2009). Nonetheless, people can behave in ways that approximate Bayesian inference in complex domains such as (active) learning (Bramley, Dayan, Griffiths, & Lagnado, 2017), reasoning (Battaglia, Hamrick, & Tenenbaum, 2013) and decision making (Wu, Schulz, Speekenbrink, Nelson, & Meder, 2018). How can these apparently contradictory findings be explained?

One idea is that people approximate rational solutions using limited available resources, a proposal sometimes discussed under the terms of resource or computational rationality (Gershman, Horvitz, & Tenenbaum, 2015; Griffiths, Lieder, & Goodman, 2015). In light of limited resources, boundedly optimal solutions to complex problems can take the form of sampling-based approximations (Dasgupta, Schulz, & Gershman, 2017), simplified decision rules (Şimşek, 2013), pruning of low-value options (Huys, Eshel, O’Nions, Sheridan, Dayan, & Roiser, 2012), or through an adaptation of information acquisition to the structure of the

task (Ruggeri & Lombrozo, 2015). However, how exactly the different approaches should be combined to produce a fully-developed theory of bounded optimality that transfer across domains and tasks is still an open question, with some researchers proposing that intelligent agents can meta-reason about which strategies to apply (Lieder, Plunkett, Hamrick, Russell, Hay, & Griffiths, 2014), and others stressing the connections between heuristic and Bayesian inference (Parpart, Jones, & Love, 2018) and the role of inductive biases (Hamrick, Allen, Bapst, Zhu, McKee, Tenenbaum, & Battaglia, 2018).

Goal and scope

The aim of this workshop is to bring together scientists who have a joint interest in how resource-constrained agents solve realistic problems, such as making decisions, finding rewards, acquiring information or reasoning and learning about the world. We have invited leading researchers from cognitive science and machine learning interested in the computational foundations of bounded optimality. In particular, our goal is to facilitate discussion and help build a more unified notion of rationality that takes resource and computational limitations into consideration. Key questions of discussion will include:

- How can we formalize theories of bounded optimality?
- What is a good framework and what are good domains in which to benchmark progress in developing such theories?
- What can we learn from past debates on and formalizations of rationality?
- Do agents learn different context-specific boundedly optimal strategies? How might they recognize when to apply which strategy?
- What does a bounded agent optimize, if at all? How can bounded optimality cope with the curse of dimensionality?

Target audience

This workshop fits well with this year’s focus on “Creativity + Cognition + Computation”. These key elements of cognition are precisely those that drive modern accounts of bounded optimality and are features of human intelligence that modern theories of rationality seek to explain. Our target audience is interdisciplinary and almost as broad as the conference as a whole — we expect this workshop to be of interest to cognitive psychologists, linguists, developmental psychologists, neuroscientists, philosophers and machine learning researchers alike. The workshop’s webpage can be found at: <https://hacksandhabits.github.io>

Organizers and presenters

Ishita Dasgupta (Organizer) is a PhD-student at Harvard University working in Samuel Gershman’s Computational Cognitive Science lab. Ishita’s work explores how people and machines make resource rational approximations to difficult problems, in particular in the domains of probability estimation, hypothesis generation, and intuitive physics.

Eric Schulz (Organizer) is a Data Science Postdoctoral Fellow at Harvard University. Eric studies generalization as function learning with a particular focus on compositionality and reinforcement learning.

Jessica B. Hamrick (Organizer) is a Research Scientist at DeepMind. Her research focuses on cognitive science-inspired theories of machine learning. In particular, she focuses on the role of mental simulation and resource rational approximations.

Joshua B. Tenenbaum (Organizer) is Professor of cognitive science at MIT. Josh’s lab sits at the intersection of cognitive science and machine learning, with a focus on hallmarks of human intelligence; in particular, the ability to learn efficiently and flexibly from limited data.

Paula Parpart is a postdoc at the University of Warwick working with Prof. Neil Stewart. Her research has focused on reconciling heuristic and Bayesian views of rationality in decision making.

Falk Lieder leads the Rationality Enhancement Group at the Max Planck Institute for Intelligent Systems in Tübingen. His mission is to build a scientific foundation and practical tools for helping people become more effective by supporting cognitive growth, goal setting, and goal achievement.

Tom Griffiths is a Professor of Psychology and Computer Science at Princeton University. Tom develops mathematical models of higher level cognition to understand the formal principles that underlie people’s ability to solve everyday computational problems.

Özgür Şimşek is a Senior Lecturer in Machine Learning at the University of Bath. Her research is on algorithms that can learn from limited experience in complex, real-world environments, with a focus on reinforcement learning.

Neil Bramley is a Lecturer of Cognitive Psychology at the University of Edinburgh. His work focuses on how people actively construct and use causal models to guide their interactions with the natural world.

Azzurra Ruggeri is a Max Planck Research Group Leader at the MPI for Human Development in Berlin. Her research focuses on how children and adults actively search for information when making decisions, drawing causal inferences and solving categorization tasks.

Kelsey Allen is a graduate student advised by Josh Tenenbaum at MIT. She uses computational models and behavioral experiments to study the development of intuitive theories, in particular intuitive physics in planning and reinforcement learning contexts.

Peter Dayan is a director at the Max Planck Institute for Biological Cybernetics in Tübingen. His research focuses on the

computational neuroscience of learning and decision making, with a focus on neuromodulation, meta-control and computational psychiatry.

Workshop structure

We propose a full-day workshop consisting of three parts. The first two parts will be a series of 20 minute talks. The final part will be a panel discussion about the limits and future of bounded optimality in cognitive science.

The morning session will consist of the following talks:

Presenter	Topic
Eric Schulz	Optimizing with confidence
Paula Parpart	Heuristics as Bayesian inference
Falk Lieder	Learning how to decide
Ishita Dasgupta	Learning to infer
Josh Tenenbaum	Computational rationality

The afternoon session will consist of the following talks:

Presenter	Topic
Jessica Hamrick	Resource-rational mental simulation
Tom Griffiths	Bridging Marr’s levels
Özgür Şimşek	Exploiting the statistical properties of decision environments
Neil Bramley	Neurath’s ship: Incremental active theory-building
Azzurra Ruggeri	Ecological active learning
Kelsey Allen	Hacks in intuitive theories
Peter Dayan	Slothful serial; perilous parallel processing

The final 45 minutes will be a *panel discussion*.

References

- Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, (p. 201306572).
- Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing neurath’s ship: Approximate algorithms for online causal learning. *Psychological review*, 124(3), 301.
- Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive psychology*, 96, 1–25.
- Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristics: Why biased minds make better inferences. *Topics in cognitive science*, 1(1), 107–143.
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.
- 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(2), 217–229.
- Hamrick, J. B., Allen, K. R., Bapst, V., Zhu, T., McKee, K. R., Tenenbaum, J. B., & Battaglia, P. W. (2018). Relational inductive bias for physical construction in humans and machines. *arXiv preprint arXiv:1806.01203*.
- Huys, Q. J., Eshel, N., O’Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS computational biology*, 8(3), e1002410.

- Lieder, F., Plunkett, D., Hamrick, J. B., Russell, S. J., Hay, N., & Griffiths, T. (2014). Algorithm selection by rational metareasoning as a model of human strategy selection. In *Advances in neural information processing systems*, (pp. 2870–2878).
- Parpart, P., Jones, M., & Love, B. C. (2018). Heuristics as bayesian inference under extreme priors. *Cognitive psychology*, *102*, 127–144.
- Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203–216.
- Şimşek, Ö. (2013). Linear decision rule as aspiration for simple decision heuristics. In *Advances in Neural Information Processing Systems*, (pp. 2904–2912).
- Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*(12), 915.