Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Doing more with less: meta-reasoning and meta-learning in humans and machines☆

Thomas L Griffiths[1], Frederick Callaway[1], Michael B Chang[2], Erin Grant[2], Paul M Krueger[1] and Falk Lieder[3]

Artificial intelligence systems use an increasing amount of computation and data to solve very specific problems. By contrast, human minds solve a wide range of problems using a fixed amount of computation and limited experience. We identify two abilities that we see as crucial to this kind of general intelligence: meta-reasoning (deciding how to allocate computational resources) and meta-learning (modeling the learning environment to make better use of limited data). We summarize the relevant AI literature and relate the resulting ideas to recent work in psychology.

Addresses
[1] Princeton University, USA
[2] University of California at Berkeley, USA
[3] Max Planck Institute for Intelligent Systems, Tübingen, Germany

Corresponding author: Griffiths, Thomas L (tomg@princeton.edu)

Current Opinion in Behavioral Sciences 2019, **29**:24-30

This review comes from a themed issue on **Artificial intelligence**

Edited by **Matt Botvinick** and **Sam Gershman**

https://doi.org/10.1016/j.cobeha.2019.01.005

2352-1546/Published by Elsevier Ltd.

The current trend in artificial intelligence research is to focus on solving more ambitious problems using more computational power and more training data. One recent analysis, presented in a blog post by OpenAI [1], charted breakthroughs in AI as a function of time and computational resources, finding that the path from the first deep networks producing high performance on image classification to recent successes in playing games like Go involved an exponential increase in computational resources. No such trend applies to human cognition: humans have finite computational resources — those that can be carried around inside our heads — and are limited to the data that can be obtained in the course of a lifetime.

In this paper, we will argue that the constraints that characterize human cognition are also intrinsic to developing key components of what we recognize as intelligence [2]. Since the human mind has only a limited amount of computational resources, it has to allocate them in an adaptive manner to solve complex problems efficiently. Having finite time in which to learn means that humans have to be able to make the most of each piece of data available to them, building and then drawing upon a rich model of the world in which learning takes place. As a consequence of developing these abilities, humans become efficient, general-purpose learners — a strong contrast to current AI systems that require huge amounts of training data and are highly specialized to particular tasks.

These two components of human intelligence — efficient use of computational resources [3,4] and efficient use of data — relate to two problems that have been studied in the AI literature: meta-reasoning and meta-learning. In the next two sections of this paper we review recent progress in these two areas, highlighting how it relates to human cognition. In the final section we consider how this approach can lead to more human-like AI systems: systems that do more with less.

Meta-reasoning

The term 'meta-reasoning' contains within it the solution to the problem of how to efficiently deploy computational resources: meta-reasoning is reasoning about reasoning, which means making intelligent decisions about how to think [5]. In research on artificial intelligence, meta-reasoning plays a central role in the definition and design of rational agents that can operate on performance-limited hardware and interact with their environment in real-time [2,6–8,4]. Considering this problem led to the notion of 'metalevel rationality' [7], under which the rationality of an agent is not assessed by the expected utility of the actions they take, but by how well the algorithm they follow in order to select those actions trades off expected utility with the costs of taking more time and expending more computation before acting. From this perspective, rationality is not just about making good decisions and drawing good inferences, but also about employing efficient cognitive strategies. Subsequent work refined the

☆ This research was supported by contract number FA9550-18-1-0077 from the Air Force Office of Scientific Research and by contract number FA8650-18-2-7832 from the Defence Advanced Research Projects Agency.

notion of metalevel rationality into the concept of 'bounded optimality' that also takes into account the computational constraints on metareasoning itself [9,7]. Bounded optimality defines an optimal program for an agent with performance-limited hardware that has to interact with its environment in real time as the solution to a constrained optimization problem over the space of programs that the hardware can execute. Here, the program's performance is measured by the utility of the sequence of world states that would result from letting the program run on the agent's hardware as it interacts with its environment.

Having defined this criterion for assessing the rationality of resource-bounded agents, the question arises of how agents can achieve it. Rational meta-reasoning [6] provides one solution: formulated in these terms, the problem of deciding how to think can itself be expressed using the language of statistical decision theory, and many of the familiar tools of AI can be applied to it [10,11]. Formally, rational meta-reasoning reduces to the problem of computing the value of computation (VOC) for each computation c that could be executed. The VOC is the difference between the increase in expected utility that would be gained by executing computation c and the cost that would be incurred by doing so. In the simplest case, when at most one step of computation can be performed, this can be expressed as

$$\mathrm{VOC}(c,b) = \mathbb{E}_{p(b'|b,c)}[\max_{a'}\mathbb{E}[U(a')|b'] - \max_{a}\mathbb{E}[U(a)|b]] \\ -\mathrm{cost}(c)$$

where b is the agent's current belief, b' is the refined belief resulting from executing computation c, and $\mathbb{E}[U(a)|b]$ is the expected utility of taking action a over the distribution of outcomes corresponding to belief b.

The rational agent should pursue the computation with the highest VOC, or, if no computation has positive VOC, not pursue any computation at all. The challenge here is that computing the VOC itself would be extremely costly if it required executing each computation to determine the resulting action. Consequently, research has focused on how to efficiently approximate the VOC or identify special cases where it can be computed more easily [2,9–15,6,16,17••,18,19].

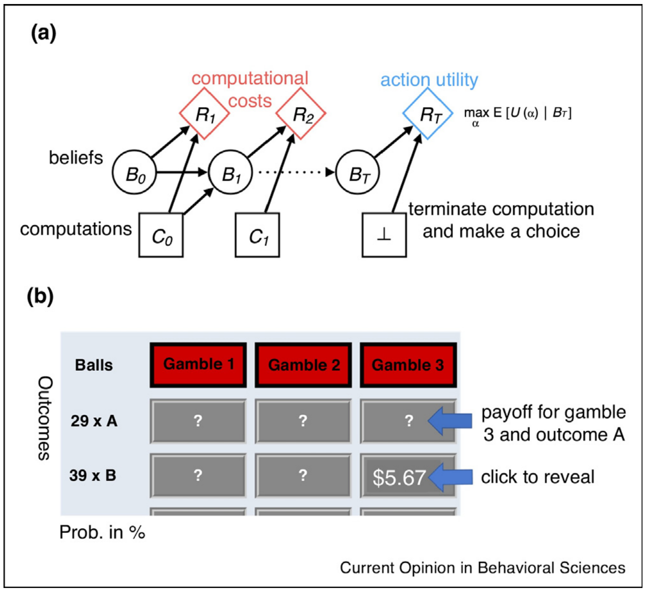
We see meta-reasoning as a key component of human intelligence, with the potential to explain a wide range of aspects of human cognition and to shed light on the factors that differentiate human minds from current AI systems. Previous work on human metacognition (e.g. [20,50••]) and active learning (e.g. [21,22]) have explored aspects of human cognition relevant to meta-reasoning — namely awareness of one's own internal states such as the accuracy of a memory or confidence in a judgment, and reasoning intelligently about how to gather information. However, meta-reasoning is a far more general phenomenon, characterizing the process of selecting or discovering the cognitive processes that will be used to tackle a task.

Part of what makes people smart is their capacity to make decisions about how they make decisions [23,24]. This makes decision-making a natural starting point for studying human meta-reasoning. An extensive literature has suggested that people make decisions by following heuristics [25]. A heuristic in itself could constitute a bounded optimal strategy, finding a good trade-off between accuracy and the time required to make a decision. However, since human decisions take place in a variety of environments and the relative cost of errors and time vary across those environments, it makes more sense to have a corresponding variety of heuristics and decide intelligently which strategy to follow in which situation [26••]. Previous research has explored human strategy selection, identifying how the strategies that people use vary based on the task and suggesting relatively simple schemes for selecting strategies based on ideas from reinforcement learning [27,28]. Formulated as a meta-reasoning problem, strategy selection becomes a matter of choosing the strategy that has highest VOC. Lieder and Griffiths [26••] showed that taking this approach, assuming that people develop simple internal models of the accuracy and time cost associated with applying different strategies to different problems, better captured human performance than previous psychological models of strategy selection.

Being able to select between existing strategies is only one aspect of what makes people able to adapt to different problems. More generally, people need to be able to efficiently deploy their cognitive resources in situations that they have never encountered before — to discover new strategies. This is a much more challenging meta-reasoning problem. However, in many situations it is possible to use a simplified version of the problem to make progress in strategy discovery. The key observation is that the construction of a strategy often reduces to a sequence of decisions about which computation to perform next [15]. Each computation reveals some piece of information or reduces uncertainty to some extent, providing data that can be used in selecting the next computation. In this case, strategy discovery reduces to a sequential decision problem, and can be expressed as a meta-level Markov decision process (MDP; see Figure 1) [10]. The MDP is a standard way of formulating reinforcement learning problems, and as a consequence we can leverage powerful tools from that literature and apply them to meta-reasoning [13,15].

One application of this approach is deriving optimal strategies for decision-making in environments where just gathering information about the options is costly. For example, a significant amount of research on

Figure 1



Meta-reasoning. **(a)** Illustration of a meta-level Markov decision process. The initial meta-level rewards capture the cost of computation and the final meta-level reward captures the benefits of computation by the expected object-level reward for choosing an action based on the final belief state. **(b)** Illustration of the Mouselab paradigm. The Mouselab paradigm externalizes computations by clicks, belief states by revealed information, and the cost of each computation by the fee charged for the corresponding click.

decision-making has used the Mouselab task illustrated in Figure 1(b). In this paradigm people are presented with a set of gambles to choose between but can only reveal the properties of those gambles by moving their computer mouse to particular locations on the screen. This task is supposed to externalize the process of attending to different pieces of information when making decisions, providing a way for psychologists to identify the strategies that people are following. Previous work using this approach has typically focused on strategies identified by the researchers, examining whether people change their strategies as the parameters of the task are varied. However, this task can be formulated as a strategy discovery problem and (approximately) solved using methods for solving MDPs. This solution provides a rational explanation for previously proposed heuristics such as Take-The-Best and Satisficing; but it also predicts novel combinations and extensions of these hand-generated heuristics, some of which people were found to actually use [29,15]. The result is a far richer picture of the kinds of strategies that can and should be used, and a more accurate characterization of the strategies that people actually follow.

The meta-reasoning approach to strategy discovery has potential applications that range far beyond decision-making. For example, developing an adaptive planning strategy — deciding which paths to explore in a multi-step decision or problem-solving task — can be expressed in these terms, allowing us to identify optimal planning algorithms that can be compared against existing cognitive models and human behavior [30•]. The computation of the VOC at the heart of rational meta-reasoning also has close parallels with ideas that have been proposed in the neuroscience literature on cognitive control [31•,32]. More generally, the rational meta-reasoning framework provides us with a set of tools for exploring how people identify the strategies or algorithms that guide their behavior in any cognitive task, from the choices that they make about how to manage their memory to choices about where to deploy their attention.

Meta-learning

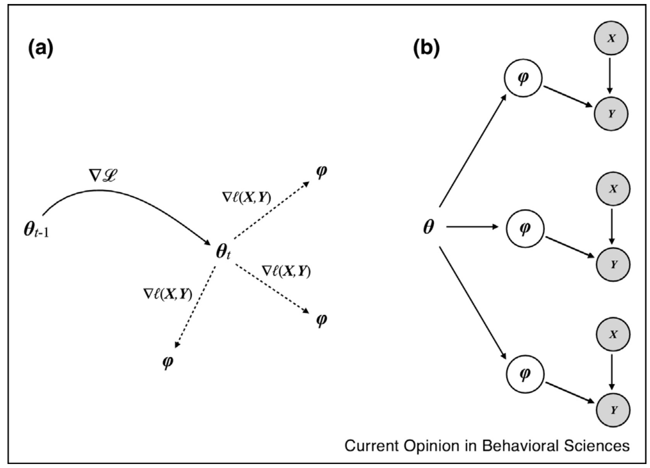
While meta-reasoning focuses on the efficient use of cognitive resources, a related challenge is the efficient use of data. One of the most striking aspects of human learning is the ability to learn new words or concepts from limited numbers of examples [33]. This ability contrasts strongly with traditional machine learning methods, which typically require many labeled examples in order to even begin to classify stimuli accurately. Recent work in machine learning has aimed to narrow this gap, exploring problems of 'few-shot' or 'one-shot' learning [34].

One of the frameworks that has been most effective in developing machine learning systems capable of solving these problems is 'meta-learning' [35,36]. In meta-learning, the learner is not presented with a single task — such as learning a particular concept — but with many tasks that all have a similar character. The system aims to leverage commonalities across these tasks in order not only to become better at solving each individual task but also to solve future tasks better and more quickly, effectively 'learning to learn.'

Recent approaches to meta-learning operate by estimating a single set of hyperparameters that parameterize a task-general component, such as a metric space [34], a memory-augmented neural network [39], or a recurrent neural network [40,41], that is shared across all tasks. In one particular approach, gradient-based meta-learning, learners are adapted for better performance on each specific task using an optimization algorithm such as gradient descent [42,43]. Meta-learning is implemented by also learning the parameterization of the learners that adapt to each task. The goal is to find a set of parameters that work well across all of the different tasks so that the learners start with a bias that allows them to perform well despite receiving only a small amount of task-specific data.

Gradient-based meta-learning is an approach with a long history [35,44,45] that has flourished due to the recent success of deep learning systems, which make use of large artificial neural networks with parameters trained by gradient descent. In cognitive science, models of cognition based on artificial neural networks are often presented as an alternative to Bayesian models, which explain human judgments as the result of rational statistical inference. It may thus come as a surprise that gradient-based meta-learning has a close relationship with one of the key components of Bayesian models of cognition, hierarchical Bayesian inference.

Figure 2



Gradient-based meta-learning. **(a)** The model-agnostic meta-learning (MAML) [37] algorithm optimizes the parameters θ of a set of models so that when one or a few gradient descent steps are taken from the initialization at θ using a small sample of task data (X, Y) to compute the negative log-likelihood $\ell(X, Y)$, each model obtains new parameters ϕ that result in good generalization performance on another sample of data from the same task as measured by the marginal negative log-likelihood $\mathcal{L}$. **(b)** The probabilistic graphical model for which MAML provides a parameter estimation procedure [38••]. Each task-specific parameter ϕ is distinct from but influences the estimation of the others through the parameters of a prior θ shared across all ϕ.

Hierarchical Bayesian inference captures the idea that learning should take place at multiple levels of abstraction. This idea has been widely used in Bayesian models of cognition. For example, hierarchical Bayesian inference can be used to learn about the properties of objects that words tend to label (such as shape) at the same time as learning the meaning of individual words [46], and to learn about the kinds of causal relationships that exist at the same time as learning those relationships [47]. While Bayesian inference generically indicates how a learner should combine data with a prior distribution over hypotheses, a hierarchical Bayesian model learns that prior distribution through experience.

There is clearly a loose analogy between hierarchical Bayesian inference and any approach to meta-learning: a meta-learning algorithm tries to establish an inductive bias for new tasks from old tasks, equivalent to 'learning a prior.' Grant *et al.* [38••] showed that this analogy is, in fact, exact for a prominent form of gradient-based meta-learning, model-agnostic meta-learning (MAML; see Figure 2) [37•]. The key idea is that the few steps of gradient descent performed by the task-specific learners results in an approximation to the Bayesian estimate of the parameter values for that task with a prior that depends on the initial parameterization, so learning the initial parameterization is equivalent to learning a prior.

Since much of the previous literature on learning-to-learn in cognitive science has focused on hierarchical Bayesian models, this connection provides a way to translate insights from modeling human learning into contemporary machine learning systems. It also opens the door to defining models using the rich and expressive language of probabilistic generative models that are able to take advantage of the large-scale learning algorithms used in deep learning. As a consequence, it may be possible to develop models that are able to capture the process of learning to learn at a resolution and scale that comes closer to that of human cognition.

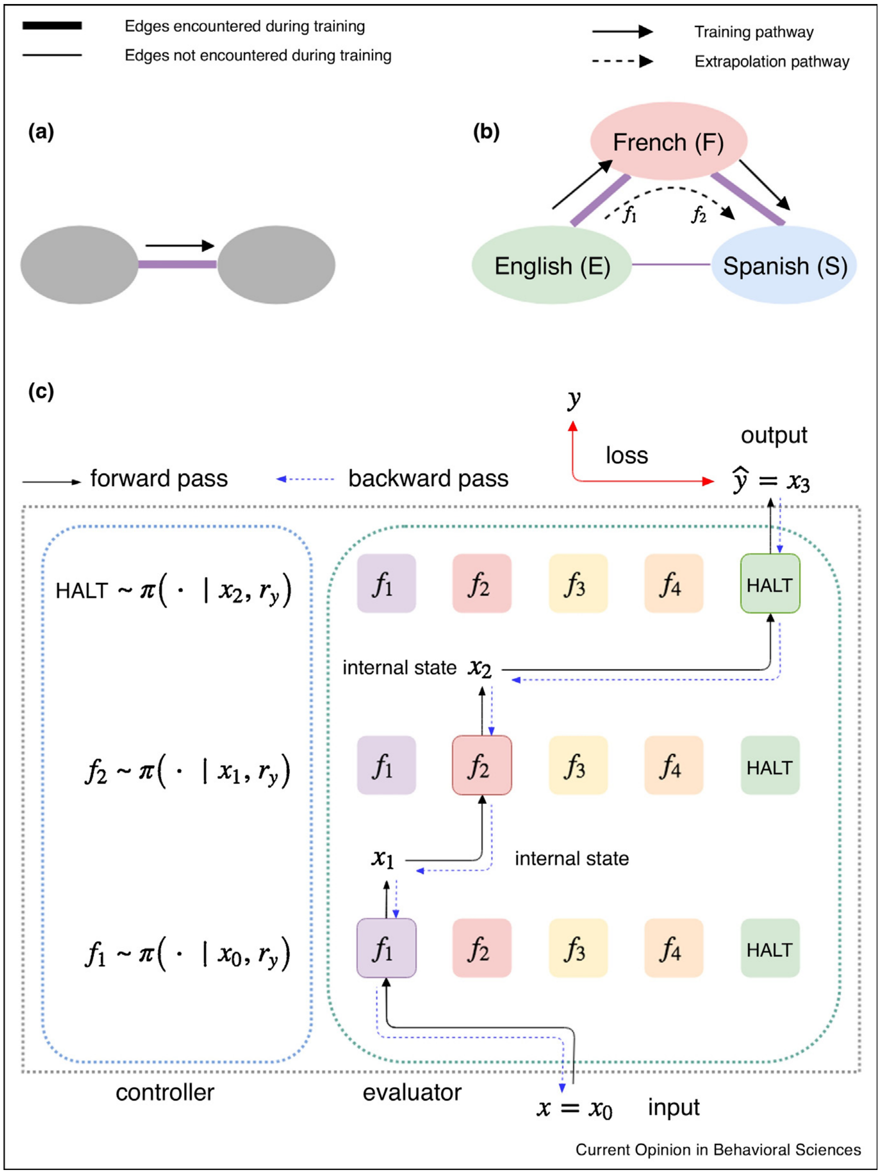
Towards general intelligence

Meta-reasoning and meta-learning individually capture key components of human intelligence, but it may be by combining them that we come closest to identifying what is necessary to achieve the kind of general intelligence that currently eludes AI systems. While meta-learning has begun to explore problems that involve learning to perform many different tasks in parallel, those tasks remain relatively similar to one another. By contrast, what characterizes human environments is the need to perform many different tasks that are genuinely different from one another: the same system drives cars, plays chess, loads the dishwasher, looks after small children, creates art, and writes scientific papers. In order to make AI systems that demonstrate the same kind of general intelligence, we need to think about how to train a single system on sets of tasks that display the same diversity.

Part of what allows humans to solve this broad range of problems is the capacity to intelligently reuse elements of their cognitive and motor skills when they encounter a new problem [48]. To translate this into the language of deep learning, we are making intelligent decisions about how to deploy learned neural modules dynamically, fluidly constructing new network architectures out of old components. This approach has elements of both meta-reasoning and meta-learning: we learn to perform well across different tasks by efficiently allocating cognitive resources. Recent work exploring this approach has shown that it can be used to automatically construct neural network architectures for novel problems in a way that supports far broader generalization than traditional neural network learning algorithms [49•].

Figure 3 shows how reasoning about what modules to deploy to solve a problem supports a distinctive form of

Figure 3



Compositional recursive learner. **(a)** In the standard supervised learning problem, the learner receives supervision for mapping an instance x from one distribution r_x to its corresponding instance y in another distribution r_y. **(b)** In the extrapolation problem, the learner receives only supervision for certain mappings but not for others. For example, suppose that during training the learner receives supervision for learning to translate English to French and to translate French to Spanish. A compositional recursive learner attempts to distill out primitive modules that transform English to French (f_1) and French to Spanish (f_2) and compose them to extrapolate to English-to-Spanish problems. **(c)** The compositional recursive learner consists of a controller π, an evaluator, and a set of reusable computational modules f_k. Its goal is to transform its input x_0 into a target representation r_y by composing together learned modules. At step j, the controller observes the internal state x_j and the target representation r_y and selects a module f_k. The evaluator applies f_k to transform x_j into a new internal state x_{j+1}. When π selects `HALT`, a loss is computed by comparing the current internal state with the desired output. The loss is backpropagated through the modules, and the controller is trained with policy optimization algorithms.

generalization. The key idea is that such a system can solve problems that are *composed* of previously solved subproblems, in the formal sense of function composition. For example, a system that has learned to translate English to French and French to Spanish can automatically translate English to Spanish by first translating to French. This supports long-range generalization beyond the tasks that the system has been trained upon — much like what we see in human behavior.

We anticipate that learning how to do more with less will become an increasingly important aspect of AI as researchers begin to hit the limits of their computational and data resources, and that hitting those limits will paradoxically result in systems that more closely resemble human cognition. This convergence could be accelerated by including computational constraints into the definition of the benchmark problems that drive the development of machine learning and artificial intelligence.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Amodei D, Hernandez D: *AI and Compute*. 2018 http://blog.openai.com/ai-and-compute.

2. Russell SJ, Wefald E: *Do the Right Thing: Studies in Limited Rationality*. Cambridge, MA: MIT Press; 1991.

3. Griffiths TL, Lieder F, Goodman ND: **Rational use of cognitive resources: levels of analysis between the computational and the algorithmic**. *Top Cognit Sci* 2015, **7**:217-229.

4. Gershman SJ, Horvitz EJ, Tenenbaum JB: **Computational rationality: a converging paradigm for intelligence in brains, minds, and machines**. *Science* 2015, **349**:273-278.

5. Cox MT, Raja A: *Metareasoning: Thinking about Thinking*. Cambridge, MA: MIT Press; 2011.

6. Russell SJ, Wefald E: **Principles of metareasoning**. *Artif Intell* 1991, **49**:361-395.

7. Russell SJ: **Rationality and intelligence**. *Artif Intell* 1997, **94**:57-77.

8. Horvitz EJ, Zilberstein S: **Computational tradeoffs under bounded resources**. *Artif Intell* 2001, **126**:1-4 ISSN 00043702.

9. Russell SJ, Subramanian D: **Provably bounded-optimal agents**. *J Artif Intell Res* 1994, **2**:575-609.

10. Hay N, Russell SJ, Tolpin D, Shimony S: **Selecting computations: theory and applications**. In In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence.* Edited by de Freitas N, Murphy K. *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence* Corvallis, OR: AUAI Press; 2012.

11. Lieder F, Plunkett D, Hamrick JB, Russell SJ, Hay N, Griffiths TL: **Algorithm selection by rational metareasoning as a model of human strategy selection**. *Advances in Neural Information Processing Systems*. 2014:2870-2878.

12. Milli S, Lieder F, Griffiths TL: **When does bounded-optimal metareasoning favor few cognitive systems?** In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, vol. 31.* Edited by Singh SP, Markovitch S. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, vol. 31* Palo Alto, CA: AAAI Press; 2017.

13. Callaway F, Gul S, Kruegera P, Griffithsa TL, Lieder F: **Learning to select computations**. *Uncertainty in Artificial Intelligence: Proceedings of the Thirty-Fourth Conference* 2018.

14. Lin CH, Kolobov A, Kamar E, Horvitz E: **Metareasoning for Planning Under Uncertainty**. In *Proceedings of the 24th International Conference on Artificial Intelligence*. AAAI Press; 2015:1601-1609:. ISBN 978-1-57735-738-4.

15. Lieder F, Krueger PM, Griffiths TL: **An automatic method for discovering rational heuristics for risky choice**. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society (CogSci)* 2017.

16. Vul E, Goodman N, Griffiths TL, Tenenbaum JB: **One and done? Optimal decisions from very few samples**. *Cognit Sci* 2014, **38**:599-637.

17. •• Russell SJ: **Rationality and intelligence: a brief update**. In *Fundamental Issues of Artificial Intelligence.* Edited by Müller VC. Springer; 2016:7-28.

A survey of recent progress in artificial intelligence research on meta-reasoning and bounded optimality that clarifies the differences between perfect rationality, calculative rationality, meta-level rationality, and bounded optimality.

18. Kocsis L, Szepesvári C: **Bandit based Monte-Carlo planning**. *European Conference on Machine Learning*. Springer; 2006:282-293.

19. Lieder F, Griffiths T, Goodman N: **Burn-in, bias, and the rationality of anchoring**. *Advances in Neural Information Processing Systems*. 2012:2690-2798:. vol 25.

20. Dunlosky J, Metcalfe J: *Metacognition*. Thousand Oaks, CA: Sage Publications; 2008.

21. Oaksford M, Chater N: **A rational analysis of the selection task as optimal data selection**. *Psychol Rev* 1994, **101**:608-631.

22. Gureckis TM, Markant DB: **Self-directed learning: a cognitive and computational perspective**. *Perspect Psychol Sci* 2012, **7**:464-481.

23. Payne JW, Bettman JR, Johnson EJ: *The Adaptive Decision Maker*. Cambridge University Press; 1993.

24. Boureau Y-L, Sokol-Hessner P, Daw ND: **Deciding how to decide: self-control and meta-decision making**. *Trends Cognit Sci* 2015, **19**:700-710.

25. Gigerenzer G, Gaissmaier W: **Heuristic decision making**. *Annu Rev Psychol* 2011, **62**:451-482.

26. •• Lieder F, Griffiths TL: **Strategy selection as rational metareasoning**. *Psychol Rev* 2017, **124**:762-794.

An application of rational reasoning to modeling people's capacity for adaptive strategy selection and how it is acquired through learning.

27. Rieskamp J, Otto PE: **SSL: a theory of how people learn to select strategies**. *J Exp Psychol: Gen* 2006, **135**:207.

28. Erev I, Barron G: **On adaptation, maximization, and reinforcement learning among cognitive strategies**. *Psychol Rev* 2005, **112**:912.

29. Gul S, Krueger PM, Callaway F, Griffiths TL, Lieder F: **Discovering rational heuristics for risky choice**. *The 14th Biannual Conference of the German Society for Cognitive Science, GK* 2018.

30. • Callaway F, Lieder F, Das P, Gul S, Krueger PM, Griffiths TL: **A resource-rational analysis of human planning**. *Proceedings of the 40th Annual Conference of the Cognitive Science Society, Cognitive Science Society, Austin, TX* 2018.

Rational metareasoning is applied to discover optimal planning strategies and the resulting predictions are tested experimentally.

31. • Lieder F, Shenhav A, Musslick S, Griffiths TL: **Rational metareasoning and the plasticity of cognitive control**. *PLoS Comput Biol* 2018, **14**:e1006043.

This paper links the idea of 'value of computation' from the AI literature to an independently proposed idea of 'value of control' in cognitive neuroscience.

32. Shenhav A, Musslick S, Lieder F, Kool W, Griffiths TL, Cohen J, Botvinick M: **Toward a rational and mechanistic account of mental effort**. *Annu Rev Neurosci* 2017, **40**:99-124.

33. Lake BM, Salakhutdinov R, Tenenbaum JB: **Human-level concept learning through probabilistic program induction**. *Science* 2015, **350**:1332-1338.

34. Vinyals O, Blundell C, Lillicrap T, Wierstra D *et al.*: **Matching networks for one shot learning**. *Advances in Neural Information Processing Systems (NIPS)*. 2016:3630-3638.

35. Schmidhuber J: *Evolutionary Principles in Self-Referential Learning (Ph.D. thesis)*. Institut für Informatik, Technische Universität München; 1987.

36. Thrun S, Pratt L: *Learning to Learn*. Springer Science & Business Media; 2012.

37. Finn C, Abbeel P, Levine S: **Model-agnostic meta-learning for fast adaptation of deep networks**. *Proceedings of the 34th International Conference on Machine Learning (ICML)* 2017.
The model-agnostic meta-learning (MAML) algorithm learns a parameter initialization for gradient-based optimization of function approximators such as artificial neural networks; the result is a model that is easy to fine-tune on a variety of tasks.

38. •• Grant E, Finn C, Levine S, Darrell T, Griffiths TL: **Recasting gradient-based meta-learning as hierarchical Bayes**. *Proceedings of the 6th International Conference on Learning Representations (ICLR)* 2018.
A particular gradient-based approach to meta-learning can be reformulated as a classic method for parameter estimation in a hierarchical Bayesian model (HBM). The connection provides a means for research domains that traditionally make use of HBMs, such as cognitive science, to scale to high-dimensional models such as neural networks and therefore naturalistic stimuli such as images.

39. Santoro A, Bartunov S, Botvinick M, Wierstra D, Lillicrap T: **Meta-learning with memory-augmented neural networks**. *Proceedings of the 33rd International Conference on Machine Learning (ICML)* 2016:1842-1850.

40. Duan Y, Schulman J, Chen X, Bartlett PL, Sutskever I, Abbeel P: *RL^2: Fast Reinforcement Learning via Slow Reinforcement Learning*. 2016. arXiv preprint arXiv:1611.02779.

41. Wang JX, Kurth-Nelson Z, Tirumala D, Soyer H, Leibo JZ, Munos R, Blundell C, Kumaran D, Botvinick M: *Learning to Reinforcement Learn*. 2017. arXiv preprint arXiv:1611.02779.

42. Andrychowicz M, Denil M, Gomez S, Hoffman MW, Pfau D, Schaul T, de Freitas N: **Learning to learn by gradient descent by gradient descent**. *Advances in Neural Information Processing Systems (NIPS), vol. 29* 2016:3981-3989.

43. Ravi S, Larochelle H: **Optimization as a model for few-shot learning**. *Proceedings of the 4th International Conference on Learning Representations (ICLR)* 2017.

44. Bengio S, Bengio Y, Cloutier J, Gecsei J: **On the optimization of a synaptic learning rule**. *Proceedings of the Conference on Optimality in Artificial and Biological Neural Networks* 1992:6-8.

45. Hochreiter S, Younger A, Conwell P: **Learning to learn using gradient descent**. *Proceedings of the International Conference on Artificial Neural Networks (ICANN)* 2001:87-94.

46. Kemp C, Perfors A, Tenenbaum JB: **Learning overhypotheses with hierarchical Bayesian models**. *Dev Sci* 2007, **10**:307-321.

47. Mansinghka VK, Kemp C, Tenenbaum JB, Griffiths TL: **Structured priors for structure learning**. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence, UAI'06*; *AUAI Press, Arlington, Virginia, United States: 2006:324-331. ISBN 0-9749039-2-2*.

48. Taatgen NA: **The nature and transfer of cognitive skills**. *Psychol Rev* 2013, **120**:439.

49. • Chang MB, Gupta A, Levine S, Griffiths TL: **This paper reframes the extrapolation problem as a problem of learning algorithmic procedures over learned transformations between representations.**. *Automatically Composing Representation Transformations as a Means for Generalization*. 2018. arXiv preprint arXiv:1807.04640.

50. •• Ackerman R, Thompson VA: **Meta-reasoning: monitoring and control of thinking and reasoning**. *Trends Cognit Sci* 2017, **21**:607-617.
Provides an overview of recent results on human metacognition.