



Tasks for aligning human and machine planning

Bas van Opheusden and Wei Ji Ma

Research on artificial intelligence and research on human intelligence rely on similar conceptual foundations and have long inspired each other [1,2^{*}]. However, achieving concrete synergy has been difficult, with one obstacle being a lack of alignment of the tasks used in both fields. Artificial intelligence research has traditionally focused on tasks that are challenging to solve, often using human performance as a benchmark to surpass [3–7]. By contrast, cognitive science and psychology have moved towards tasks that are simple enough to allow for detailed computational modeling of people’s choices. These divergent objectives have led to a divide in the complexity of tasks studied, both in perception and cognition. The purpose of this paper is to explore the middle ground: are there tasks that are reasonably attractive to both fields and could provide fertile ground for synergy?

Address

Center for Neural Science and Dept of Psychology, New York University,
4 Washington Place, New York, NY 10003, United States

Current Opinion in Behavioral Sciences 2019, **29**:127–133

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

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

<https://doi.org/10.1016/j.cobeha.2019.07.002>

2352-1546/© 2019 Published by Elsevier Ltd.

The challenge of aligning tasks between fields

Research on artificial intelligence and research on human intelligence rely on similar conceptual foundations and have long inspired each other [1,2^{*}]. However, achieving concrete synergy has been difficult, with one obstacle being a lack of alignment of the tasks used in both fields. Artificial intelligence research has traditionally focused on tasks that are challenging to solve, often using human performance as a benchmark to surpass [3–7]. By contrast, cognitive science and psychology have moved towards tasks that are simple enough to allow for detailed computational modeling of people’s choices. These divergent objectives have led to a divide in the complexity of tasks studied, both in perception and cognition. The purpose of this paper is to explore the middle ground: are there tasks that are reasonably attractive to both fields and could provide fertile ground for synergy?

In perception, object recognition has emerged as a central task for aligning the studies of biological and machine vision. For biological vision, object recognition is a core task used to understand representations of visual stimuli in the ventral stream and specifically the infero-temporal cortex [8,9]. In machine vision, object recognition has proven to be an equally fruitful test bed [10,11]. This alignment of tasks has led to great synergy, for example in the form of comparisons between cortical activity and the activity of units in trained neural networks [12,13].

In this paper, we focus on planning, which we roughly define as any cognitive process in which the decision-maker mentally simulates future states, actions or outcomes in a decision tree. These decision trees can often become exponentially large, and strategies for efficiently searching decision trees and making fast and accurate decisions are of interest for artificial intelligence and cognitive psychology. In the artificial intelligence of planning, the game of chess has long taken center stage, from Shannon [14] to Deep-Blue [4] to AlphaZero [7]. The cognitive psychology of chess has lagged behind. Chase and Simon advocated for chess as a “standard task environment” similar to *Drosophila* in genetics [15]. Researchers have attempted to measure depth of thought in chess through clever experimental methods such as board reconstruction [16], “thinking aloud” [15,17], playing under cognitive load [18] or time pressure [19–21]. Overall, however, the prominence of chess as an experimental paradigm in cognitive science has declined. One potential reason for this decline is the difficulty of developing a computational model for human chess play, and a simpler game is needed.

In the following sections, we will reflect on the notion of task complexity and review sequential decision-making tasks that have been used in computational cognitive science to test hypotheses about planning. We will then examine candidate tasks for aligning human and machine planning, and set an agenda for how such alignment could be achieved.

Defining task complexity

The complexity of a decision-making task can be defined in multiple ways. We focus on *state space complexity*, the number of different states that a decision-maker can reach in any sequence of actions starting from the initial state.¹ We choose this metric since it is often easy to compute or approximate, and in practice closely matches with intuitive notions of task difficulty. In two-player

¹ Traditionally, state space complexity refers to the number of all possible game states, we count only non-terminal states.

games, state space complexity is often correlated with game tree complexity [22]. Using state space complexity, however, ignores important elements of difficulty, such as how easily people can learn the task rules, or whether the task requires managing uncertainty, learning models of one's opponent, and potential symmetries in the state or action space. A more principled complexity metric measures the size of a compressed state space, with abstracted states and actions, which preserve as much as possible the task structure [23,24]. Therefore, the exact numbers should be taken as only a rough quantification of the intuitive notion of complexity (Table 1).

Tasks used to study human planning

The study of human planning is rich and has used many experimental tasks. A core challenge across tasks is that researchers can only observe people's decisions, and not the cognitive processes (planning algorithms) that generated them. A principled method to infer planning algorithms is by fitting a computational model to individual player's choices. Here, we review a selection of tasks that have been used to study human planning within a computational framework. The tasks are roughly sorted by increasing state space complexity.

Solway and Botvinick task (Figure 1a). [27**] This paper used a task in which participants choose between up to 4 consumer items that they previously ranked by desirability. Crucially, they presented items in two groups on either side of a computer screen, and participants reported their choice sequentially, by first selecting a group and afterwards an item within that group. This task is equivalent to navigating a decision tree with up

to 3 internal nodes, and Solway and Botvinick found that people's behavior can be captured by noisy evidence integration, which treats each path through the decision tree as a competitor in a bounded accumulation process.

Daw et al. task (Figure 1b) [28] In the two-step task, participants make a sequence of two binary choices. In the first decision stage, the participant might choose between stimuli A1 and A2. Stimulus A1 usually (on 70% of trials, labelled "common" transitions) leads to the state B, and stimulus A2 to state C. However, occasionally (on 30% of trials, "rare" transitions), A1 leads to C, or A2 to B. In the second stage, participants make another choice between two stimuli, which yields a monetary reward with some probability. The reward probabilities fluctuate slowly, so participants have to constantly adapt the values they associate with the stimuli and adjust their policy accordingly.

The two-step task has state space complexity 3 (1 first-level state and 2 second-level states). This simplicity is intentional, as the two-step task is the simplest task that distinguishes model-free and model-based learning. (However, it has been argued that sophisticated model-free learning can masquerade as model-based learning in this task [39].) Specifically, when an agent receives a reward at a second-level state, a model-based learner takes into account whether it arrived there through a "common" or "rare" transition, whereas a model-free learner does not. This task has been used to demonstrate that people use a mixture of model-based and model-free learning [40,28], that the relative usage of each system depends on the reliability of their respective predictions [41] in an on-line cost-benefit analysis [42], and that people's arbitration between these systems changes under cognitive load [43] or when they receive a dopamine precursor [44].

Gläscher et al. task (Figure 1c) [29]. Their task preceded the two-step task and differed from it in two ways: first, there are 4 second-level states, and each combination of a first-level choice and a "common" or "rare" transition leads to a different state. Second, instead of a second-level state leading to a reward with some probability, the choice leads probabilistically to a third state, and then deterministically to a reward. These modifications allowed Gläscher et al. to dissociate neural correlates of reward prediction errors and state prediction errors.

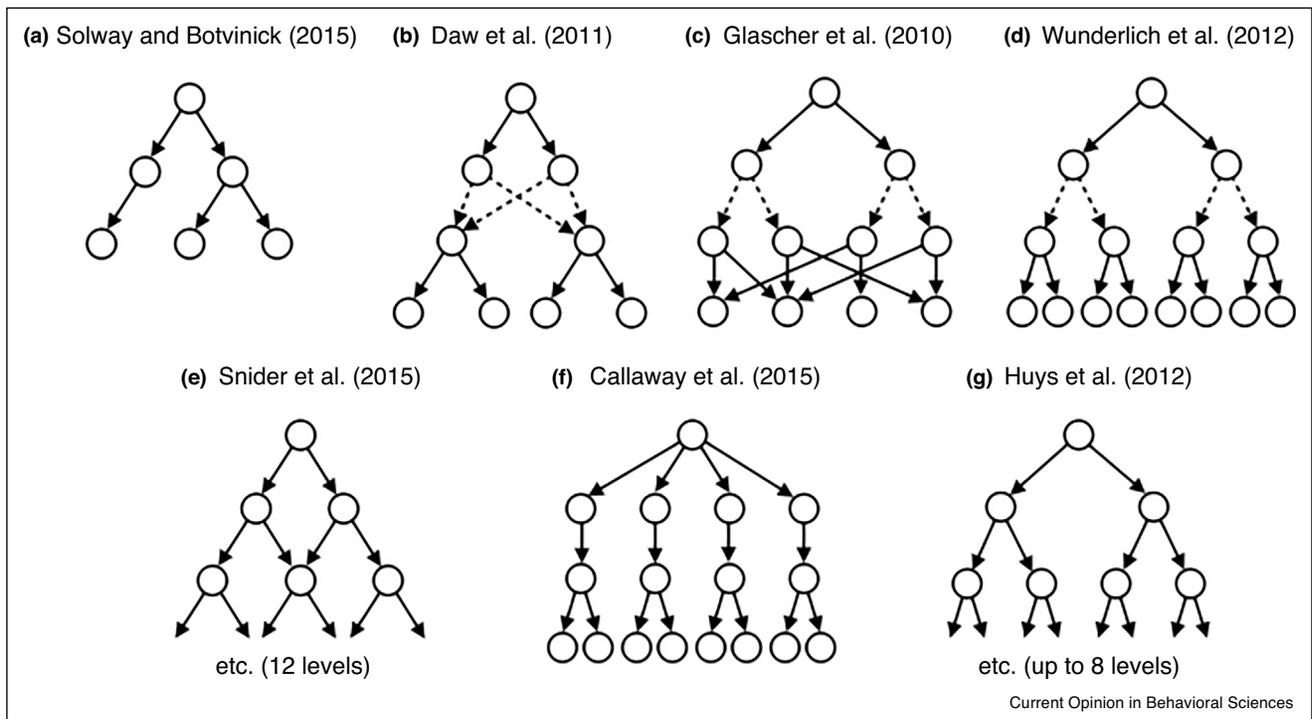
Wunderlich et al. task (Figure 1d) [30] In this variant of the two-step task, the transitions from the first to second level were made by an adversarial computer agent, creating a two-player game. This allowed them to study the computational processes underlying forward planning by searching for neural correlates of values of individual branching steps in a minimax decision tree.

Table 1

Tasks and their complexities. Artificial intelligence studies tend to focus on complex tasks, whereas cognitive scientists favor simplicity. State space complexity for chess and go are taken from [25] and [26]

Task	State space complexity	Human cognition studies	AI studies
Solway and Botvinick task	3	[27**]	
Daw et al. "two-step" task	3	[28]	
Gläscher et al. task	5	[29]	
Wunderlich et al. task	7	[30]	
Callaway et al. task	≤ 20	[31]	
Snider et al. task	66	[32**]	
Huys et al. task	2 to 128	[33] [34]	
Four-in-a-row	1.2×10^{16}	[35**]	[36]
Chess	$\sim 10^{47}$	[16] [15] [37]	[14] [4] [7]
Go	2.1×10^{170}		[6]

Figure 1



Decision trees in some tasks used to study human planning. Each node represents a task state. Solid arrows indicate transitions that are under the subject's control, dashed arrows transitions that are not.

Callaway et al. task (Figure 1e) [31,45**] To more directly measure the planning process, they employed a process tracing method inspired by the “Mouselab” paradigm [46]. Participants navigate a directed graph, in which each node is associated with a reward. Although all nodes and edges of the graph are always visible to the participant, rewards are only revealed when the participant clicks or hovers over the corresponding node. The sequence in which participants choose to reveal rewards provides insight in the cognitive process by which participants plan their actions. Although the paradigm scales in principle to arbitrary graphs, Callaway et al. study graphs with up to 20 nodes.

Snider et al. task (Figure 1f) [32**] Participants watch a triangular grid of 12 rows of variable-sized disks scroll down a touchscreen, and at each downward step, participants control whether the grid moves right or left by pointing with a stylus. Thus, the participant traces out a trajectory through the grid of disks, and they receive reward proportional to the size of all disks on that trajectory. To plan optimally, participants have to navigate a decision tree with 66 internal nodes, making this one of the most complex tasks in which human planning has been quantitatively modeled.

Huys et al. task (Figure 1g) [33] Participants make a sequence of up to 8 binary decisions, through which they

traverse a graph of 6 states. Each transition incurs a reward that can be either positive or negative, and the task is designed such that the optimal policy requires taking large negative rewards to obtain later positive rewards. This task has been used to study how people prune decision trees following a large loss [33], how they decompose a task into a hierarchy of subtasks [34], and to develop planning algorithms that optimally trade off speed and accuracy [47].

Candidate tasks for aligning human and machine planning

Most cognitive science studies use tasks with single-digit state space complexity, and studies with more complex tasks are relatively rare. By contrast, AI researchers study games many orders of magnitude more complex. To connect these fields, we argue for tasks with an intermediate level of complexity. In particular, we propose a task that we have studied ourselves, namely four-in-a-row.

Even apart from the objective of aligning fields, cognitive scientists might have a good reason to study tasks of intermediate complexity: there likely is a categorical difference in algorithms people use for simple and complex tasks. When the number of states in a task exceeds 10-20, it becomes implausible that people are able to represent all states in working memory and exhaustively

search the tree [48]. Instead, working memory constraints may force people to use tree search algorithms with constant memory and computational time requirements, such as Monte Carlo Tree Search [49] or heuristic search [50]. This possibility should make such tasks interesting to cognitive science.

Four-in-a-row. To study human planning with exponentially large decision trees, we studied human decision-making and learning in two-player deterministic game. In this game, two players compete to create four-in-a-row on a 4-by-9 board (Figure 2a). This game, part of the family of (m, n, k) games [36] is considerably more complex than most cognitive science tasks, as it contains 1.1812×10^{18} non-terminal states.

We were able to predict individual participants' choices on individual moves with a computational model that combines a feature-based heuristic for state evaluation, a best-first search algorithm for planning, and feature drop-out as a mechanisms for lapses of attention (Figure 2b). Moreover, this model could generalize from predicting people's in-game decisions to predicting their choices on a two-alternative forced-choice task, board evaluations, response times and eye movements. Finally, by fitting the model parameters to data from participants in consecutive sessions, or when with limited allotted thinking time, we found large and robust effects of expertise and thinking time on people's policy, and in particular the estimated size of their mental decision trees (Figure 2c-d).

Mazes and grid worlds. Mazes and grid worlds are an important collection of tasks used to study human and machine planning. These tasks have been used to study human [51] and animal reinforcement learning [52], or to develop and benchmark novel reinforcement learning algorithms [53,54]. Planning has been studied in the motor domain using a space similar to a grid world [55].

However, grid worlds are far from a representative sequential decision-making task, as the decision tree contains many transpositions (different action sequences leading to the same state) and cycles. Therefore, these tasks often have relatively low state space complexity and lend themselves uniquely to be solved by dynamic programming algorithms [53] that use memory proportional to the number of states. Since we focus on tasks for which forward planning at decision time is necessary to curtail the complexity, mazes and grid worlds may not be an ideal unifying task.

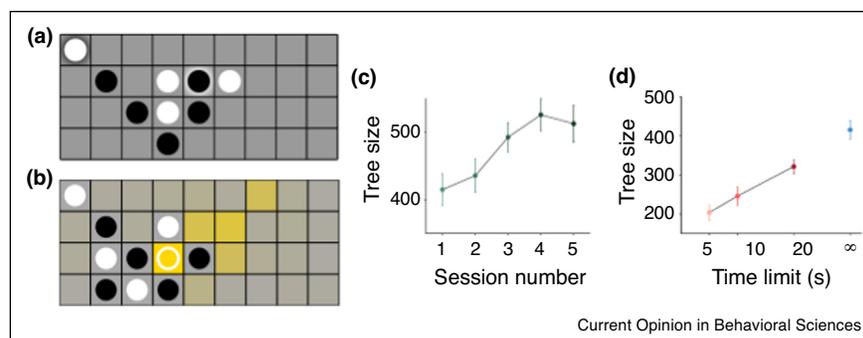
Atari games. Another recent proposal for a task to study deep neural networks and their relation to human intelligence are Atari games and in particular the game *Frostbite* [2*]. Deep Q-learning networks can learn these games [5] and outperform human players, but training them requires an exceedingly large amount of simulated games. One reason for this slow learning is that these networks are often initialized randomly, whereas people approach Atari games with strong inductive biases [56]. Although we are excited to see a convergence of cognitive and AI research on Atari games, as a paradigm to study forward planning they might prove too difficult.

What it means to align human and machine intelligence

Similarities and differences between human and machine intelligence have been much discussed [2*,13,57,58]. Here, we focus on what we believe to be necessary conditions for alignment between studies of human and machine planning in a given task, in particular what requirements an AI algorithm needs to satisfy to be a successful cognitive model.

We believe that understanding human cognition requires a computational model which makes predictions for individual participants' choices on single trials, in multiple

Figure 2



Four-in-a-row. (a) Example board position. Black moves first, and the goal is to get four in a row. (b) Board position with overlaid in false color a map of the probabilities that a heuristic search model fitted to one individual's play predicts for that individual's next move. (c-d) The heuristic search model allows us to estimate the size of the tree built by human players as they learn across sessions (c) or have different time limits (d). Panels (B-D) adapted from [35**].

experimental conditions. A computational model may take the form of an explicit algorithm that people mentally execute, or a neural network. In either case, the model should include a description of how its policy changes after each experienced episode.

The main requirement for a computational model to be successful is that, when presented with the same stimuli that a participant experienced in a given condition in a task, its trajectory through policy space should match that participant's trajectory. This requirement encompasses several conditions that are difficult for AI algorithms to satisfy:

1. The model's trajectory should start from a policy that matches people's inductive biases in people's decision-making [2*,56,59,60], and converge to one that matches people's behavior including their task performance.
2. The model should improve on the task at the same rate as human participants.
3. Finally, the model needs to satisfy these requirements for each individual participant in different experimental conditions, forcing it to match not just one trajectory but also capture individual differences in participant's trajectories.

One method to constrain individual variability is to construct models with a small number of parameters, which limits the set of policies the model can express to a low-dimensional manifold. This is the strategy employed by traditional computational cognitive science: the model is carefully chosen so that all participant's policies lie on its manifold, and policy trajectories can be translated to trajectories in a small space of parameters.

Another method would be to use infinitely expressible models such as deep neural networks, but include hyperparameters in their dynamics or initial policy, to match different people's trajectories with minimal adjustments.

Although developing an AI algorithm that satisfies these conditions on even a single task will be challenging, doing so would enable a "cognitive psychology of artificial intelligence" [61*], in which researchers subject AI algorithms to new psychological experiments. For example, one can characterize changes in the AI's policy trajectory in response to manipulations of the input data or constraints on its use of memory or computational time. These changes can then be compared to the changes in humans doing the same experiments. We believe that this approach will deepen understanding of the correspondence between human and machine intelligence.

Conclusion

We have argued that to align studies of artificial intelligence with studies of human intelligence, there needs to be a focus on finding tasks that are of interest to both fields. In doing so, both fields might have to compromise a little. Cognitive scientists might have to go out of their comfort zone in terms of experimental control and theoretical tractability: suitable tasks are likely orders of magnitude higher in complexity than common in cognitive science, and human behavior is likely far from optimal. AI researchers might have to settle for tasks that might not be the most challenging in terms of achieving high task performance but that offer more opportunities for systematic analysis of learned policies and linking those to human cognition. While chess might be too complex from both an experimental and modeling point of view, simpler tasks such as four-in-a-row might fit the bill. Once a suitable task has been chosen, we have laid out an agenda for using experimental manipulations in that task to reach a deeper understanding of the correspondence between human and machine intelligence.

Acknowledgements

This work was supported by grants IIS-1344256 from the National Science Foundation and R01MH118925-01 from the National Institutes of Health.

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. Turing AM: **Computing machinery and intelligence**. *Mind* 1950, **49**:433-460.
2. Lake BM, Ullman TD, Tenenbaum JB, Gershman SJ: **Building machines that learn and think like people**. *Behav Brain Sci* 2017, **40**.
This paper outlines challenges to building artificial systems that mimic human intelligence. They emphasize causal models of the world, learning from few examples, and rapid generalization.
3. Tesauro G: **Temporal difference learning and td-gammon**. *Commun ACM* 1995, **38**:58-68.
4. Campbell M, Hoane AJ Jr, Hsu F-h: **Deep blue**. *Artif Intel* 2002, **134**:57-83.
5. Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, Riedmiller M: *Playing atari with deep reinforcement learning*. *arXiv preprint arXiv:1312.2013*:5602.
6. Silver D, Huang A, Maddison CJ, Guez A, Sifre L, Van Den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M et al.: **Mastering the game of go with deep neural networks and tree search**. *Nature* 2016, **529**:484-489.
7. Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T et al.: *Mastering chess and shogi by self-play with a general reinforcement learning algorithm*. *arXiv preprint arXiv:1712.01815*. 2017.
8. Riesenhuber M, Poggio T: **Models of object recognition**. *Nature Neurosci* 2000, **3**:1199.
9. DiCarlo JJ, Zoccolan D, Rust NC: **How does the brain solve visual object recognition?** *Neuron* 2012, **73**:415-434.
10. Krizhevsky A, Sutskever I, Hinton GE: **Imagenet classification with deep convolutional neural networks**. *In Advances in neural information processing systems* 2012:1097-1105.

11. LeCun Y, Bengio Y, Hinton G: **Deep learning**. *Nature* 2015, **521**:436.
12. Yamins DL, DiCarlo JJ: **Using goal-driven deep learning models to understand sensory cortex**. *Nature Neurosci* 2016, **19**:356.
13. Kriegeskorte N: **Deep neural networks: a new framework for modeling biological vision and brain information processing**. *Annu Rev Vision Sci* 2015, **1**:417-446.
14. Shannon CE: **Xxii. programming a computer for playing chess**. *The London, Edinburgh, and Dublin Philosophical Magazine J Sci* 1950, **41**:256-275.
15. Chase WG, Simon HA: **Perception in chess**. *Cognit Psychol* 1973, **4**:55-81.
16. de Groot AD: **Het Denken van den sckaken**. Noord-Holland. Uitgev. Maatschappij 1946.
17. Campitelli G, Gobet F: **Adaptive expert decision making: Skilled chess players search more and deeper**. 2004.
18. Holding DH: **Counting backward during chess move choice**. *Bull Psychonomic Soc* 1989, **27**:421-424.
19. Chabris CF, Hearst ES: **Visualization, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors**. *Cognit Sci* 2003, **27**:637-648.
20. Burns BD: **The effects of speed on skilled chess performance**. *Psychol Sci* 2004, **15**:442-447.
21. Van Harreveld F, Wagenmakers E-J, Van Der Maas HL: **The effects of time pressure on chess skill: an investigation into fast and slow processes underlying expert performance**. *Psychol Res* 2007, **71**:591-597.
22. Allis LV et al.: **Searching for solutions in games and artificial intelligence**. Rijksuniversiteit Limburg; 1994.
23. Botvinick MM: **Hierarchical reinforcement learning and decision making**. *Curr Neurobiol* 2012, **22**:956-962.
24. Sanborn, S, Bourgin, DD, Chang, M, Griffiths, TL: 2018. Representational efficiency outweighs action efficiency in human program induction. arXiv preprint arXiv:1807.07134.
25. Chinchalkar S: **An upper bound for the number of reachable positions**. *ICGA J* 1996, **19**:181-183.
26. Tromp J: **The number of legal go positions**. In *International Conference on Computers and Games*. Springer 2016:183-190.
27. Solway A, Botvinick MM: **Evidence integration in model-based tree search**. *Proc Natl Acad Sci* 2015, **112**:11708-11713.
This paper studies the dynamics of reward-based, goal-directed decision-making in a multi-step choice task, and find that human behavior can be understood in terms of evidence integration. Their model poses that people accrue noisy evidence for different trajectories through a decision tree, and make decisions when that evidence clearly favor one choice.
28. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ: **Model-based influences on humans' choices and striatal prediction errors**. *Neuron* 2011, **69**:1204-1215.
29. Gläscher J, Daw N, Dayan P, O'Doherty JP: **States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning**. *Neuron* 2010, **66**:585-595.
30. Wunderlich K, Dayan P, Dolan RJ: **Mapping value based planning and extensively trained choice in the human brain**. *Nature Neurosci* 2012, **15**:786.
31. Callaway F, Lieder F, Krueger PM, Griffiths TL: **Mouselab-mdp: A new paradigm for tracing how people plan**. In *The 3rd multidisciplinary conference on reinforcement learning and decision making* 2017.
32. Snider J, Lee D, Poizner H, Gepshtein S: **Prospective optimization with limited resources**. *PLoS Comput Biol* 2015, **11**:e1004501.
By comparing human behavior with optimal models in a complex sequential decision-making task, the authors were able to identify people's planning strategies and in particular estimate their depth of computation.
- They found that people's behavior was consistent with a complete, brute-force exploration of all possible paths in decision tree up to a resource-limited finite depth.
33. Huys QJ, Eshel N, O'Nions E, Sheridan L, Dayan P, Roiser JP: **Bonsai trees in your head: how the Pavlovian system sculpts goal-directed choices by pruning decision trees**. *PLoS Comput Biol* 2012, **8**:e1002410.
34. Huys QJ, Lally N, Faulkner P, Eshel N, Seifritz E, Gershman SJ, Dayan P, Roiser JP: **Interplay of approximate planning strategies**. *Proc Natl Acad Sci* 2015, **112**:3098-3103.
35. van Opheusden B, Galbiati G, Bnaya Z, Li Y, Ma WJ: **A computational model for decision tree search**. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* 2017:1254-1259.
This paper introduced a sequential decision task of a level of complexity that is unusually high for cognitive science (10¹⁶), and showed that people's choices matched those of a best-first search algorithm with attentional oversights and pruning. We used this model to show that tree size decreased with time pressure and increased with expertise.
36. Beck J: *Combinatorial games: tic-tac-toe theory, volume 114*. Cambridge University Press; 2008.
37. Holding DH: *The psychology of chess skill*. Lawrence Erlbaum; 1985.
39. Akam T, Costa R, Dayan P: **Simple plans or sophisticated habits?. State, transition and learning interactions in the two-step task**. *PLoS Comput Biol* 2015, **11**:e1004648.
40. Daw ND, Niv Y, Dayan P: **Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control**. *Nature Neurosci*. 2005, **8**:1704.
41. Lee SW, Shimojo S, O'Doherty JP: **Neural computations underlying arbitration between model-based and model-free learning**. *Neuron* 2014, **81**:687-699.
42. Kool W, Cushman FA, Gershman SJ: **When does model-based control pay off?** *PLoS Comput Biol* 2016, **12**:e1005090.
43. Otto AR, Gershman SJ, Markman AB, Daw ND: **The curse of planning: dissecting multiple reinforcement-learning systems by taxing the central executive**. *Psychol Sci* 2013, **24**:751-761.
44. Wunderlich K, Smittenaar P, Dolan RJ: **Dopamine enhances model-based over model-free choice behavior**. *Neuron* 2012, **75**:418-424.
45. Callaway F, Lieder F, Das P, Gul S, Krueger PM, Griffiths TL: **A resource-rational analysis of human planning**. In *Proceedings of the 40th Annual Conference of the Cognitive Science Society*. 2018.
This paper employs a process tracing paradigm that forces participants to reveal which states they consider in a sequential decision-making task with an information-gathering component. Based on this data, they propose a resource-rational model for human planning, in which people make rational meta-decisions about which nodes in a decision tree to explore, and when to stop planning.
46. Payne JW, Bettman JR, Johnson EJ: **Adaptive strategy selection in decision making**. *J Exp Psychol: Learn Mem Cognit* 1988, **14**:534.
47. Sezener CE, Dezfouli A, Keramati M: **Optimizing the depth and the direction of prospective planning using information values**. *PLoS Comput Biol* 2019, **15**:e1006827.
48. Miller GA: **The magical number seven, plus or minus two: Some limits on our capacity for processing information**. *Psychol Rev* 1956, **63**:81.
49. Kocsis L, Szepesvári C: **Bandit based monte-carlo planning**. In *European conference on machine learning, Springer* 2006:282-293.
50. Pearl J: *Heuristics: intelligent search strategies for computer problem solving*. 1984.
51. Simon DA, Daw ND: **Neural correlates of forward planning in a spatial decision task in humans**. *J Neurosci* 2011, **31**:5526-5539.
52. Pfeiffer BE, Foster DJ: **Hippocampal place-cell sequences depict future paths to remembered goals**. *Nature* 2013, **497**:74.

53. Sutton RS, Barto AG: *Reinforcement learning: An introduction, volume 1*. Cambridge: MIT Press; 1998.
 54. Mattar MG, Daw ND: **Prioritized memory access explains planning and hippocampal replay**. *Nature Neurosci* 2018, **21**:1609.
 55. Diamond JS, Wolpert DM, Flanagan JR: **Rapid target foraging with reach or gaze: The hand looks further ahead than the eye**. *PLoS Comput Biol* 2017, **13**:e1005504.
 56. Dubey, R., Agrawal, P., Pathak, D., Griffiths, T.L., and Efros, A.A. (2018). Investigating human priors for playing video games. arXiv preprint arXiv:1802.10217.
 57. Hassabis D, Kumaran D, Summerfield C, Botvinick M: **Neuroscience-inspired artificial intelligence**. *Neuron* 2017, **95**:245-258.
 58. Marblestone AH, Wayne G, Kording KP: **Toward an integration of deep learning and neuroscience**. *Front Comput Neurosci* 2016, **10**:94.
 59. Colunga E, Smith LB: **From the lexicon to expectations about kinds: A role for associative learning**. *Psychol Rev* 2005, **112**:347.
 60. Feinman R, Lake BM: (2018). Learning inductive biases with simple neural networks. arXiv preprint arXiv:1802.02745.
 61. Ritter S, Barrett DG, Santoro A, Botvinick MM: **Cognitive psychology for deep neural networks: A shape bias case study**. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70* 2017:2940-2949. JMLR. org..
- This paper proposed an agenda of interrogating deep neural networks as if they are human subjects. Specifically, as humans learn, they exhibit inductive biases; for example, they prefer to categorize objects according to shape rather than color. The authors find that deep neural networks trained on ImageNet vary widely in their ability to reproduce this bias.