Trends in
Cognitive Sciences

CellPress

Review

Looking deeper into the algorithms underlying human planning

Ionatan Kuperwajs[1,*], Evan M. Russek[1], Marcelo G. Mattar[2,3], Wei Ji Ma[2,3], and Thomas L. Griffiths[1,4]

Humans possess a remarkable ability to form sophisticated multi-step plans even in complex environments. In this review article, we consider efforts that attempt to characterize the mechanisms underlying human planning using a computational framework, primarily focusing on methods that search a tree of possible solutions. These studies range from experimental probes for heuristics that people employ while thinking ahead to normative models for reducing the computational costs of planning. Additionally, we examine the recent successes of artificial intelligence in the domain of planning and how these innovations can be applied to better understand human sequential decision-making. As examples, we highlight this approach in two tasks that require planning many steps into the future, namely 4-in-a-row and chess.

Highlights

People have the ability to form complex multi-step plans. Real-world environments often require mental simulation of the future outcomes associated with available actions in order to make good decisions.

Decision tree search has emerged as the primary computational framework for modeling human planning in cognitive science. As a result, various mechanisms involved in planning have been identified and normative approaches have emerged to address how people efficiently use cognitive resources while thinking ahead.

Artificial intelligence has developed powerful algorithms for solving planning problems in large state spaces and these innovations can be leveraged to more deeply understand how people plan in tasks where evaluating every course of action is intractable.

Planning is a hallmark of intelligence

In our everyday lives, we must constantly make decisions by **planning** (see Glossary), or mentally simulating different possible courses of action. Deciding which courses to take in college, choosing between different job opportunities, or even just figuring out the logistics of an upcoming trip are all settings where considering the consequences of our actions is beneficial. This process is made all the more demanding by the fact that we live in a complex world, one in which various events are interconnected throughout time with outcomes that are difficult to predict. Machines operating in such a world also need to solve planning problems, and research in artificial intelligence has resulted in algorithms that provide a rich source of hypotheses about human cognition.

A core challenge in the study of human planning is that planning is an internal and inherently unobservable process. Thus, while attempts to directly measure multi-step planning such as process-tracing paradigms exist, the primary method for inferring the algorithms underlying this process has been to fit a computational model to behavior and evaluate how closely the model predicts people's choices. Using this method, the dominant framework that has emerged for modeling planning is tree search. Broadly speaking, planning algorithms in this space can search much deeper ahead than a single step by constructing a **decision tree** that contains many actions leading to different state and reward trajectories. While traversing this tree leads to additional information about currently available actions, its size is exponential in the number of choice points, making it infeasible to evaluate every possible sequence. For example, if an agent has to make a sequence of N decisions with K options at each step, then the total number of sequences is K^N. Therefore, a growing body of literature has focused on characterizing the algorithms that allow people to plan efficiently.

Planning problems formalized as search over a decision tree have been similarly influential in artificial intelligence research. The primary distinction between approaches to planning in cognitive science and artificial intelligence is that algorithms in the latter field do not need to be constrained by the computational limitations of the human mind and can instead be engineered for task

[1]Department of Computer Science, Princeton University, Princeton, NJ, USA
[2]Center for Neural Science, New York University, New York, NY, USA
[3]Department of Psychology, New York University, New York, NY, USA
[4]Department of Psychology, Princeton University, Princeton, NJ, USA

*Correspondence: ikuperwajs@princeton.edu (I. Kuperwajs).

performance. Therefore, artificial intelligence research has fully embraced the challenge of developing powerful algorithms to solve a wide array of problems in large state spaces. Throughout the history of artificial intelligence, algorithms such as **heuristic search** and **Monte Carlo tree search** have been implemented to play zero-sum, two-player games like tic-tac-toe, chess, and Go. In recent years, modern techniques including **artificial neural networks** have further augmented the capabilities of these search methods to achieve superhuman performance on complex planning problems.

Here, we review the literature that aims to understand the cognitive mechanisms by which people plan sequences of actions and discuss how recent advancements in both computational modeling and data collection have started to change that understanding. We begin by focusing on the experimental tasks that have been leveraged to explain how people construct and navigate decision trees, resulting in evidence for a wide range of concepts such as arbitration and pruning during search. Since exhaustive search is often intractable, normative approaches to modeling human planning have emerged to more precisely specify the underlying computations and representations involved in this process by, for example, casting planning as a problem of efficiently using available cognitive or informational resources. Then, we discuss how innovations in artificial intelligence, including heuristic search and artificial neural networks, can be used to more deeply understand how people form multi-step plans. Specifically, we highlight two illustrative case studies: 4-in-a-row, a combinatorial game of intermediate complexity, and chess, a more complex game that has a rich history in experimental psychology. This approach, coupled with prior findings, presents a promising path forward for yielding more detailed characterizations of the cognitive processes underlying human planning.

Characterizing the cognitive mechanisms underlying human planning

To review progress that has been made thus far in understanding the cognitive mechanisms involved in human planning, we discuss experimental tasks as well as normative models that have provided insight into this process. We then turn our attention to representations beyond tree search that have been useful in characterizing how people plan.

Experimental tasks and heuristics

Perhaps the most influential sequential decision-making paradigm, and the logical starting point for reviewing experimental findings on human planning, is the two-step task [1] (for an overview of studies on the neural basis of planning, see Box 1). In this task, participants make a sequence of two binary choices. In the first decision stage, the choice between two stimuli leads probabilistically to another set of states that, following a second decision, yield a monetary reward. The reward probabilities fluctuate slowly, so participants have to constantly adapt the values they associate with the stimuli and adjust their decisions accordingly. Notably, this is a simple task in which **model-free reinforcement learning** and **model-based reinforcement learning** make different behavioral predictions. When a reward is received, a model-based agent has the capacity to take into account whether it arrived there through a common or rare transition, whereas a model-free learner does not. In the seminal work in which the task was introduced, it was found that people use a mixture of model-based and model-free learning [1].

Recent work has refined interpretation of behavior in the two-step task in at least two ways. First, several studies have challenged the assumption that behavioral signatures cleanly reflect either model-based or model-free strategies. Some have argued that what appears to be model-based behavior could, in fact, emerge from more sophisticated forms of model-free learning [2,3]. Conversely, other work has argued that apparent model-free behavior may reflect model-based planning under incorrect assumptions about the task structure [4]. Second, the

Glossary

Artificial neural networks: a class of models consisting of an architecture, which describes how different units are connected, and a learning algorithm, which is used to learn the appropriate connection weights for the model's parameters.
Best-first search: a heuristic search algorithm that iteratively selects a sequence of the most promising actions.
Decision tree: a representation of a planning problem where nodes denotes states and arrows possible actions that could be made.
Heuristic search: a class of algorithms that use a heuristic function to approximate the values of states and construct a partial decision tree.
Markov decision process: a framework for modeling sequential decision-making that defines how an agent interacts with the environment using a set of states, a set of actions, the transition probabilities between states conditioned on actions, and the rewards received when making an action in a given state.
Model-based reinforcement learning: learning a model of an environment's dynamics in order to plan.
Model-free reinforcement learning: learning directly from an environment via repeated association of actions with subsequent rewards.
Monte Carlo tree search: a heuristic search algorithm that estimates action values by averaging the returns of many simulated trajectories that are balanced between searching in promising and unexplored areas of the state space.
Planning: the process of making a decision by mentally simulating the future consequences associated with potential courses of action and selecting the one that maximizes expected value.
Successor representation: a representation of a planning problem that summarizes the long-range predictive relationships between states of an environment in order to balance efficiency and flexibility.

CellPress

Box 1. The neural basis of multi-step planning

Recent years have seen a surge in experimental studies aimed at understanding the neural mechanisms of multi-step planning in the brain. This line of work has shown that a wide range of neural structures that contribute to associative learning, food foraging, and spatial navigation are implicated in planning as well.

Evidence suggests that in the two-step task, people's arbitration between model-free and model-based strategies is influenced by dopamine precursors [87]. Another study dissociated neural correlates of reward and state prediction errors by extending the two-step task to include more second-level states and introducing third-level states that deterministically lead to reward [88]. In a two-player variant of the task, the transitions from the first- to second-level states were made by an adversarial computer agent. This allowed for the identification of neural correlates of the values of individual branching steps in a minimax decision tree [89], and further evidence for the neural substrates of planning in animals has been found in different adaptations of the task [90–92].

Historically, a cornerstone of the empirical study of sequential decision-making was Tolman's finding that rats navigating in a maze use data gathered from free exploration to build a mental map of their environment that they can subsequently use for efficient, goal-directed planning [93]. Following this, the activity of hippocampal place cells was decoded to determine what part of space is being represented on a moment-to-moment basis. When a rat pauses at a choice point in a maze, these representations sweep forward along the possible paths that the animal can take [94]. Furthermore, the spatial trajectories represented by these sweeps closely correspond to the rat's subsequent navigational behavior [95]. Several theoretical studies have proposed that this non-local hippocampal activity reflects simulation of future states used for planning [96,97], demonstrating that patterns in which non-local locations are re-activated can be explained as those which are most useful for computing decision values.

Recent work has begun to uncover relationships between this sort of non-local neural activity, associated with simulation, and planning behavior in humans. Both functional magnetic resonance imaging and magnetoencephalography (MEG) have revealed a relationship between the simulation of outcomes in the brain and people's decisions [5,98,99]. Replay-like patterns of sequential neural activity can be decoded from visual cortex and have been linked to both decision-time planning [100,101] and planning computations that occur at reward receipt [7]. These sequential neural activity patterns have been measured as unfolding rapidly, with lags between simulated steps occurring between 40 and 200 ms. Novel techniques for analyzing MEG have uncovered a slower process of step-by-step simulation, likely more aligned with our conscious experience of planning, on the order of 800 ms per state [6]. This process was found to occur in hippocampus and later consolidate in prefrontal cortex. Whether these fast and slow simulation processes subserve distinct computational roles in planning remains an open question.

conception of when model-based computations occur has been expanded. While simulation supporting planning often occurs during deliberation at the time of choice [5,6], it can also occur at the time of reward receipt [7] or even offline during rest periods between trials [8]. The timing and direction of planning can vary systematically with task demands and context [9,10].

Arbitration between model-based and model-free algorithms, as contextualized by the two-step task, is a solution to the problem of deciding which of these two algorithms should be used. To choose an action in a given state, a model-based system mentally simulates the consequences of possible actions multiple steps into the future using a decision tree, whereas the model-free system considers the outcome of actions taken in the same or similar states in past experience. These dual systems have been discussed under various names and implementations, including habitual and goal-directed control of learned behavioral patterns [11,12]. The model-based system is slow and computationally expensive, but can determine high-value actions from any state, including ones that have never been previously encountered. Alternatively, the model-free system is fast but needs previous experience to inform its policy. Note that these systems model choices and not response times, and, as such, speed is derived from the implementation of these algorithms. In order to combine information from these systems, people may utilize the uncertainty estimates provided by both systems [13,14] or balance the search time and accuracy associated with each system [15,16]. A related problem is how these systems can benefit from each other's computations, for which one solution is to allow the model-free system to learn not only from direct experience, but also from computations made by the model-based system [17,18].

Beyond the two-step task, the study of human planning has employed a wide range of experimental tasks, and we cover a non-exhaustive list here. While most research has relied on planning tasks of limited complexity compared to those used in artificial intelligence, this has facilitated the usage of computational frameworks, which in turn have led to substantial progress in understanding the algorithmic mechanisms that people use to plan. One idea that emerged in a series of papers is that planning can be conceptualized as probabilistic inference, where the goal is to infer a decision policy which could lead to a conditioned, imagined, high return [19–21]. This notion is also at the core of the active inference approach to planning, which frames decision-making as well as many aspects of cognition in general as problems of probabilistic inference [22]. Planning as inference has been tested in a decision task where participants had to make a series of choices between items embedded in a tree structure that they had previously ranked by desirability [23]. Behavior was captured by noisy evidence integration implementing probabilistic inference while treating each path through the decision tree as a competitor in a bounded accumulation process.

In another goal-directed decision-making task, participants were asked to make a sequence of multiple two-alternative choices by which they traversed a graph. Each transition incurred a reward, which could be either positive or negative, and the task was designed such that the optimal policy required taking large negative rewards to obtain positive future rewards. This revealed that people plan along multiple branches in a decision tree, but eliminate unpromising branches by pruning and decompose the task into a hierarchy of subtasks [24,25]. Human planning has also been studied in a fast-paced, dynamic environment where participants watched a triangular lattice of disks of different sizes scroll down a touchscreen and traced the most rewarding path [26]. Participants received a reward proportional to the size of all disks on that trajectory, and human behavior was found to be consistent with planning several steps into the future. Notably, participants in this task preferred to reduce their depth of computation or increase the recalculation period rather than sacrifice the precision of computation. This is one example of tradeoffs in planning computations, where resource limitations require balancing efficiency and accuracy.

Normative and optimal approaches

An outstanding problem that has not been directly addressed by the studies in the previous section is how people are able to plan well despite the huge number of actions and associated outcomes to evaluate under time and resource constraints. One way to frame the results covered thus far is in terms of experimental findings (Figure 1A). Arbitration between model-based and model-free reinforcement learning poses the question of whether to plan or rely on experience, and one proposal for how to implement a planner is via probabilistic inference. Heuristics such as pruning or depth-limited search serve to reduce the costs associated with planning, a dependence that has been cited as far back as one of the earliest attempts to replicate human-like intelligence in a computer [27,28]. Ultimately, these are a direct result of researchers proposing and testing different hypotheses derived from intuition across planning tasks as opposed to a more principled approach that could explain multiple findings.

In many domains, progress has been made by analyzing optimal solutions to a problem that a cognitive system is meant to solve [29,30]. For planning, this means addressing the metalevel problem, which is to determine whether and in which direction the tree should be expanded. Normative approaches to modeling human planning are fairly sparse, although there have been a number of recent attempts (Figure 1B). A commonality among several of these is that an agent maintains a belief over values associated with each action, and the effect of planning is to decrease uncertainty. This can in turn lead to better choices and higher rewards. Models designed in this manner are subtly distinct in the assumptions that are made in order to derive the optimal

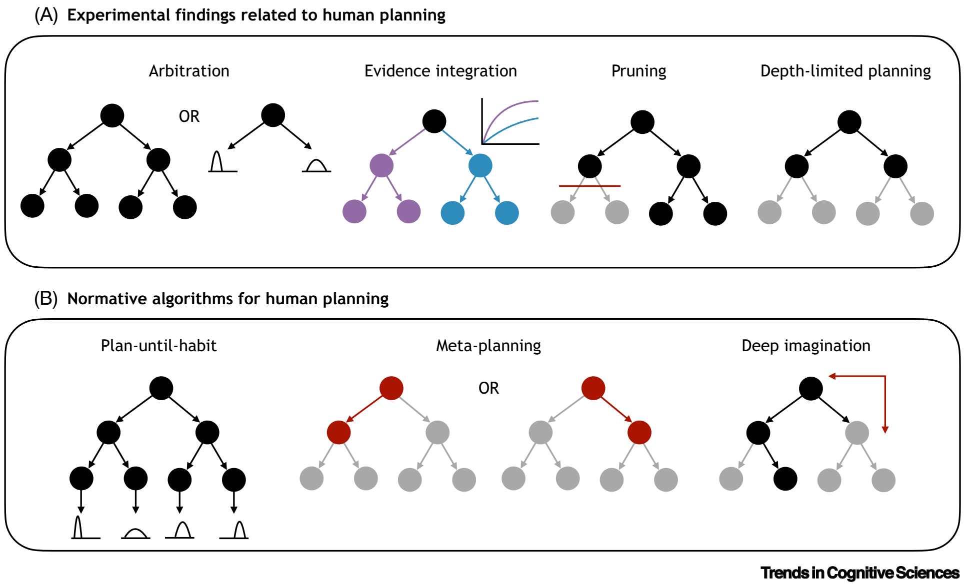

Figure 1. Mechanisms for human decision tree search. (A) Experimental findings related to human planning including arbitration between model-based and model-free reinforcement learning algorithms [1], treating paths through a decision tree as competitors in a bounded accumulation process [23], eliminating unpromising branches of the decision tree via pruning [24], and reducing the depth of search in a decision tree [26]. (B) Normative algorithms for human planning including the execution of forward search up to a certain depth in a decision tree before exploiting habitual values [32], solving the metalevel problem of determining in which direction to plan via resource rationality [36] or information sampling [37], and optimizing the tradeoff between breadth and depth in decision trees [39].

planning strategy and in the mathematical tools used to formalize the meta-planning problem. One example is the plan-until-habit scheme, which executes forward search up to some depth and then exploits heuristic values from a habitual system as proxies for consequences that may arise further into the future [31,32]. This framework is designed to optimally trade off speed and accuracy under the assumption that deeper planning leads to more accurate evaluations, but at the cost of slower decision-making. The critical value to be computed when deciding if to expand the decision tree in a certain trajectory is whether a new piece of information could change the agent's decision about what action to take and how much extra value is expected to be gained by that policy improvement. This results in an expansion metric that is cheap to compute, but relies on cached values that summarize past experiences at the frontier of the decision tree. The plan-until-habit scheme is close in form to arbitration between model-based and model-free reinforcement learning, and can explain several behavioral patterns in grid-world environments and reproduce results in the task from [24], namely the effect of time pressure on the depth of planning, the effect of reward magnitudes on the direction of planning, and the gradual shift from goal-directed to habitual behavior during training.

An alternative approach originates from the field of resource rationality, which strives towards optimality by deriving models of behavior that take into account which cognitive operations are available to people, how long they take, and how costly they are [33–35]. Using a process-tracing paradigm, the conceptual and technical tools of a **Markov decision process** can be leveraged to solve the sequential decision problem of constructing a decision tree [36]. A third framework starts by explicitly mapping the meta-planning problem onto one of information search, where the objective is to choose the single most rewarding option given a number of alternatives [37]. Both information search and planning are fundamentally about improving the selection of future actions, with the distinction that sampling information is an overt action in the real world while planning requires mental simulation. Thus, planning is a form of internal information search that combines past experiences with simulation, and tractable Bayesian models can be derived in this setting to decide which action to plan for. Note that neither of these specify a process-level model of human planning, which is an important direction for future work.

A final option is to avoid modeling the full metalevel problem and instead evaluate the effectiveness of a search strategy to determine in which conditions it would be favored. Recent work has tackled the breadth–depth dilemma, or the tradeoff between evaluating many different options as opposed to gaining more information about a smaller number of options when faced with a decision [38,39]. In large decision trees, the optimal policy is to allocate few samples per level so that deep levels can be reached, with the exception being poor environments and at low capacity where it is marginally better to broadly sample branches.

Representation

While we have focused almost exclusively on heuristics and normative algorithms that operate on decision trees until now, there are numerous recent findings that characterize an additional aspect of human planning: representation. This crucial component of planning is neglected in most prior studies, which assume complete and fixed task representations. However, efficient and flexible planning might also need to control these representations in order to quickly simplify and more easily reason about problems. One possibility is that people represent planning problems by breaking them down into smaller components. This decomposition can be based on clusters of states that support hierarchical planning [40] or on resource-rational tradeoffs [41]. Further evidence for hierarchical planning has been found via an experimental paradigm that uses program induction [42]. A related idea is that of value-guided construals, which characterize how an optimal cognitively limited decision-maker balances the complexity of a representation and its use for planning and acting [43]. This model posits that task representations can be controlled and that such control allows people to construct a simplified mental representation of a problem that is sufficient to solve it. Thus far we have discussed decision trees as the primary cognitive and computational representation of how people plan, but there may be tasks that require planning and not tree search.

An alternative for thinking about representation while planning is the **successor representation**, an algorithm that was initially introduced as a method for generalization in reinforcement learning [44]. The successor representation balances flexibility and efficiency via multi-step representation learning, storing long-term predictions about future events. When faced with a decision, how often successor states are expected to be visited can be combined with a reward function to evaluate an action. The successor representation can thus appropriately respond to distal reward changes and achieve some of the flexibility associated with model-based reinforcement learning without confronting the computational burden of fully planning [3,45]. This concept has been iterated on to make it more flexible in response to changes [46], enable learning of prospective and retrospective cognitive maps [47], and adaptively deploy forward and backward prediction [9]. The successor representation is closely related to work on task decomposition, as both involve simplifying planning over time through temporal abstraction [48]. More specifically, multiple temporal steps of prediction are accomplished in fewer steps of simulation by utilizing a type of task model that represents temporally abstract transitions, linking actions with events that occur over multiple future timesteps. Viewed through this lens, planning with the successor representation, where the expected future occupancy of all states is computed in a single step, and planning by simulating one transition at a time can be seen as two ends of a continuum defining the granularity of simulation. Intermediate forms of temporal abstraction, such as those explored in frameworks of task decomposition and hierarchical planning, strike a balance between these extremes. They allow simulation at an intermediate level of coarseness: larger than single-step transitions but more fine-grained than the successor representation, by grouping multiple low-level actions into higher-level, temporally extended transitions. Overall, the computational approach to studying planning in humans continues to be an active area of research, with contributions guided by disparate task environments, the interplay between heuristics and optimality, and how people represent problems while planning.

Using artificial intelligence to scale the study of human planning

Despite the ubiquity of sequential decision-making in naturalistic behavior, the study of the cognitive mechanisms underlying such decisions has been primarily limited to relatively simple tasks. The constrained laboratory studies that we have covered thus far are designed to test theories about planning by eliciting specific behaviors that are amenable to precise analysis and modeling. However, the majority of these experiments are not intrinsically motivating for the participants, are learned in short periods of time, and result in small data sets. As such, they cannot capture the complexity of the real world and limit the potential to analyze sophisticated behavior. To circumvent this, computational methods derived from artificial intelligence have recently begun to be combined with tasks where evaluating every course of action is intractable and large-scale data sets. Scaling task complexity in this manner has the added benefit of allowing more complex models to make differentiable predictions.

Signatures of complex planning

In order to study complex planning, the ideal task needs to be difficult enough that good decisions require thinking multiple steps ahead while simultaneously preserving tractability for computational modeling. Furthermore, it should be novel, have simple rules, and be engaging in order to encourage learning and motivation. Considering these competing desiderata, 4-in-a-row, a variant of tic-tac-toe in which two players alternate placing tokens on a 4×9 board with the objective of getting four tokens in a row in any orientation (Figure 2A), becomes an appealing candidate [49]. With approximately 1.2×10^{16} non-terminal states, 4-in-a-row has a state space complexity far beyond common cognitive science tasks, preventing any exhaustive search or brute force algorithms from being successful. Therefore, people as well as artificial agents who play the game need to address the challenge of efficient search. This is just one example of how cognitive science as a field has started shifting away from the traditional, reductive approach to science towards using more complex tasks [50–52]. More complex tasks like 4-in-a-row, as well as those aimed at exploring the intersection between planning and behaviors such as navigation [53] and puzzle solving [54], may result in more precise descriptions of human cognition while planning. In turn, these descriptions can be used to study biases apparent in psychopathology, for example, anxiety [55,56].

The computational cognitive model for human behavior in 4-in-a-row is adapted from the artificial intelligence literature, in particular, heuristic search. The model combines a heuristic evaluation function, which is a weighted linear combination of board features [57–59], with the construction of a decision tree via **best-first search** (Figure 2B). Best-first search iteratively expands nodes on the principal variation, or the sequence of actions that lead to the best outcome for both players given the current decision tree [60]. To allow the model to capture variability in human play and make human-like mistakes, Gaussian noise and feature dropout are added to the value function. For each move the model makes, it randomly drops some features from the heuristic function before it performs search. Such feature omissions can be interpreted as lapses of selective attention [61]. During search, the model also prunes the decision tree by removing branches with low heuristic value. Fitting the model to individual people's decisions requires a number of technical innovations, including estimating the log probability of a move in a given board position with inverse

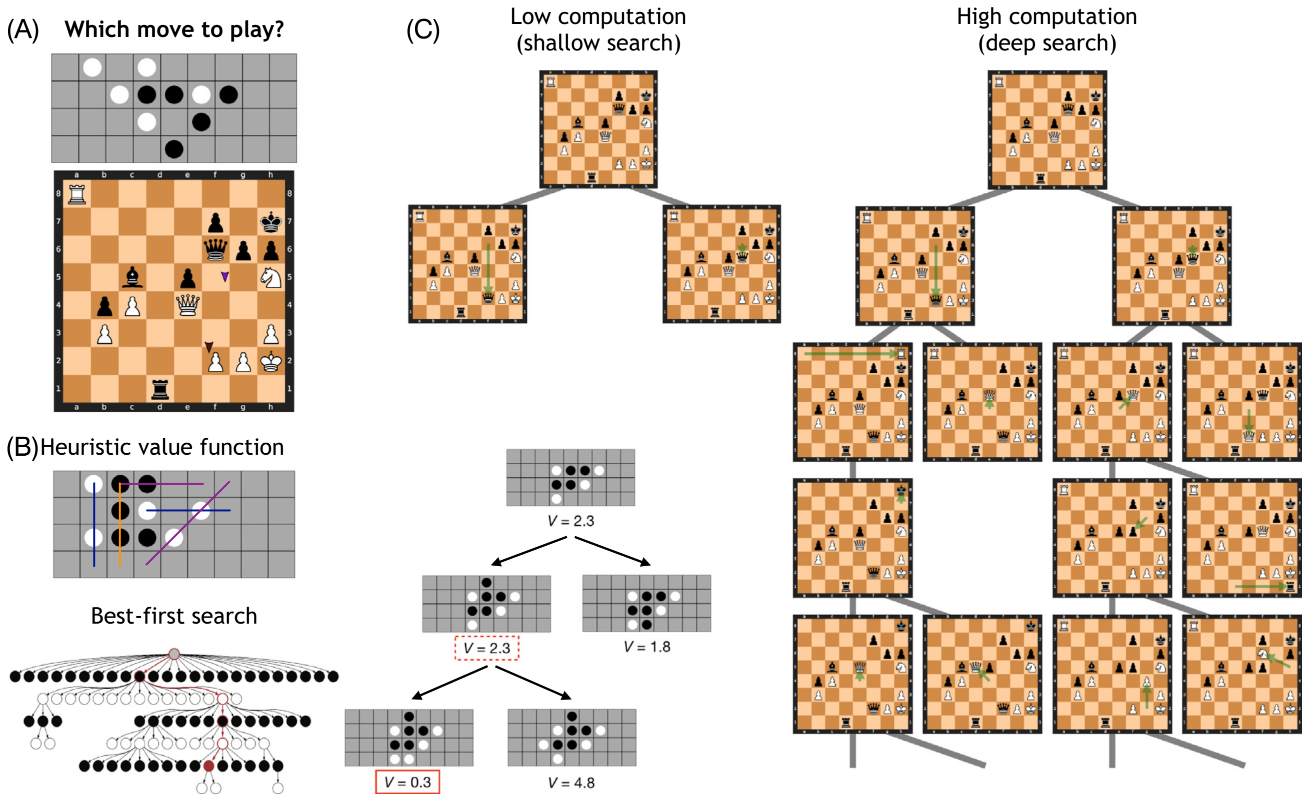

Figure 2. Computational accounts of human planning in complex tasks. (A) Example board positions in 4-in-a-row (top) and chess (bottom) where the primary task for a decision-maker is to select which move to play. (B) Illustration of the computational cognitive model for 4-in-a-row [49]. The heuristic value function uses features which are intermediate patterns to winning the game highlighted by different color lines (left, top). Features with identical colors are constrained to the same weights, and the evaluation is a sum over the counts of these features. The model then conducts best-first search to build a decision tree where red nodes indicate the sequence of highest-value moves for both players (left, bottom). As an example, black is to move in the root position at the top of the decision tree (right). After expanding the root node with two candidate moves for black and evaluating the resulting positions using the heuristic function, the algorithm selects the highest value node ($V = 2.3$) on the second iteration and expands it with two candidate moves for white. The algorithm evaluates the resulting positions, and backpropagates the lowest value ($V = 0.3$) since white is the opponent. This means that the value in the red solid unbroken box replaces the one in the red broken box and the root node is updated to the highest value among its children ($V = 1.8$). On the next iteration, the algorithm will again expand the child node with the highest value. (C) Illustration of a resource rational account of planning in chess [83]. In such a framework, the time individuals spend thinking should be sensitive to the benefits and costs of computation. Across board positions, the benefit of computation can be measured as the benefit of the deciding through a high computation, deep tree search (right) compared to a low computation, shallow tree search (left).

binomial sampling [62] and optimizing the log-likelihood function with Bayesian adaptive direct search [63].

With a process-level model that can accurately predict people's move choices given a board state, it now becomes possible to make progress towards understanding human decision-making in a complex planning task. A notable use case is to investigate how expert players differ from novices, resulting in robust evidence for increased planning depth with expertise [49]. Additionally, experts drop less board features and seem to memorize and reconstruct these features more accurately. Beyond the nature of expertise, 4-in-a-row can be leveraged to study the interplay between different reinforcement learning systems [64], comparisons between algorithms for human and machine planning [65], the continued development of model-based decision strategies from childhood into adulthood [66], and planning impairments following ventromedial prefrontal cortex lesions [67].

Given a baseline model in a complex task, a key issue that arises is improving the model to more accurately reflect human cognition. This is difficult because residuals between a model and behavior can often be noisy. Neural network techniques for guided model improvement have established themselves as an emerging field to solve this problem, providing another connection with artificial intelligence. The core concept is to train a network to predict human behavior in a particular task, and then identify deviations between the network and a cognitive model's predictions. This highlights situations in which the model requires novel mechanisms to explain human behavior because the network can detect patterns in the data without requiring human understanding of these patterns *a priori.* This method was pioneered to discover algorithms underlying human decision-making [68,69] and categorization [70], while a related line of work has started to develop recurrent neural networks for automated model discovery, thus far primarily in reinforcement learning environments [71–73]. In the domain of planning, this has been utilized to extend the 4-in-a-row model with mechanisms ranging from a simple opening bias to adjustments regarding endgame decision-making [74]. In sum, 4-in-a-row and its model provide a framework under which various disparate research questions related to complex planning can be pursued.

Machine and human planning in chess

Chess presents an intriguing case study in complex decision-making and multi-step planning, and artificial intelligence has been able to make significant progress on developing methods for planning in combinatorial games. In 1950, Claude Shannon published a groundbreaking paper describing how a machine or computer could be designed to play a reasonable game of chess [75]. His algorithm was based on a minimax procedure, which used an evaluation function of chess positions to select the best move for both players. Later, TD-Gammon was developed as the first program to play backgammon at human master level [76], and DeepBlue, the chess-playing computer that defeated reigning world champion Gary Kasparov, made use of alpha-beta pruning to decrease the number of nodes evaluated by its minimax algorithm [59]. This variation on heuristic search stops evaluating a move when at least one possibility has been found that proves the move to be worse than a previously examined move. Combining prior work on Monte Carlo tree search with artificial neural networks, AlphaGo became the first artificial agent to achieve superhuman performance in Go with a series of stunning victories against world champion Lee Sedol [77]. The main innovation behind AlphaGo is that it selected moves using deep convolutional neural networks to both evaluate positions and sample actions. Additionally, instead of starting from random network weights, it used weights that were pretrained on human experts as a starting point, iterating on previous work that aimed to predict human moves in large Go databases [78,79].

AlphaGo was improved upon with variants that do not rely on human data [80] or domain-specific knowledge [81]. These represent state-of-the art advances in artificial intelligence, namely those that utilize self-play reinforcement learning to create computer agents that solve complex planning problems at a level beyond human capabilities. Similar cutting-edge methods have been used in the chess community to create publicly available chess engines such as Stockfish, and there have been attempts to build such engines that are optimized to match human play rather than performance. The most prominent example in this space is Maia, a customized version of AlphaZero that is trained on human chess games [82]. Maia is a unified modeling approach that captures human style and improvement across different skill levels rather than simply matching aggregate human performance. Most important to this article, work in artificial intelligence has provided a starting point for developing cognitive models of human chess play. By pointing to technical components such as algorithms for heuristic search and artificial neural networks as well as domain-specific findings such as knowledge acquisition and concept discovery, the literature on machine planning shows that developing computational accounts of planning in large state spaces is viable.

Despite impactful early work focused on expertise, the promise of chess as a model system for cognitive science has not yet been fulfilled due to its intractability for computational modeling (for a brief history of work on the psychology underlying chess play, see Box 2). However, two recent developments have opened the door to more precise characterizations of human reasoning in chess. The popularity of online chess platforms has resulted in large-scale data sets, and modern chess engines provide better tools for developing detailed computational models of human decision-making and planning in this setting. One approach in this domain tested the hypothesis that people intelligently select the situations in which computational resources are spent [83]. Specifically, players seemed to spend more time thinking in board positions where planning was more beneficial, and this effect was greater in stronger players. Stockfish was used to estimate the benefit of applying planning computations for each board position occurring in 12.5 million games from the Lichess database. The benefit of computation is then the increase in board position advantage, where players can make the maximum utility move with no planning or perform a planning computation which leads to a more accurate utility function and then select the new maximum utility move (Figure 2C). Meanwhile, other studies have used large-scale chess data sets to develop algorithms for calculating the riskiness of each move in a chess game [84], investigate the kinds of decisions where people are likely to make errors [85], and provide evidence that people employ sophisticated learning algorithms [86]. This collection of results suggest that it is possible to use quantitative methods to investigate human cognition in a planning paradigm as complex as chess. Thus, chess is now returning to the forefront of research in both cognitive science and artificial intelligence.

Concluding remarks

Within cognitive science, planning has traditionally been investigated with constrained tasks which encourage mental simulation, and normative approaches to how people conduct decision tree search have been formalized only recently. Similar problems have been tackled in artificial

Box 2. Chess: the '*Drosophila* of psychology'

The psychology of chess has invariably been a topic of much interest, and was referred to by Chase and Simon as the '*Drosophila* of psychology' – a standard task environment around which knowledge and understanding can accumulate much like model organisms in biology. In 1946, de Groot proposed that strong chess players make moves by constructing a decision tree through an iterative deepening algorithm [102]. Experimentally, he presented players with preconfigured board positions and asked them to freely narrate their thought process while selecting a move. This showed no differences between stronger and weaker players. de Groot then conducted another experiment in which he instructed players to memorize and reconstruct chess positions, this time finding that stronger players were able to place more pieces correctly. In 1973, Chase and Simon added a control condition in which players were provided with scrambled and often illegal chess positions [103]. They found that players were better at reconstructing legal positions, leading to the hypothesis that people represent chess positions with an array of small patterns called chunks that allow them to compress information and avoid capacity limits.

Since these experiments, the explanation for the superior performance of experts in chess has been hotly contested. One line of thought is that this difference is primarily due to better pattern recognition. To support this hypothesis, another replication of the reconstruction experiment analyzed which specific features of a chess position players remember incorrectly [104]. Other studies found no difference in search between experts and novices [105] or used eye movements and visual search tasks to further validate that experts possess chess-specific improvements in performance [106]. Conversely, some experiments have shown that deeper search is a key factor for improved play in chess. To selectively impair search while leaving pattern recognition abilities intact, players were asked to make moves under time pressure [107] or while counting backwards [108], which affected experts more than novices. There have also been studies that directly investigated differences in search between experts and novices, finding that they do exist [109,110]. At least one intermediate proposal has been suggested, which is that improved search may be responsible for the development from novice to expert, but the step from expert to to Grandmaster level relies on pattern recognition [111].

However, developing computational cognitive models that accurately predict the moves of individual chess players has proven to be difficult [112,113]. Currently, there is still no process-level theory of human planning in chess.

Outstanding questions

Do prior results from constrained laboratory studies replicate in unconstrained settings? Massive data sets and novel computational methods can be applied to naturalistic tasks in order to address this.

How does planning interact with related fields in cognitive science, such as social cognition or working memory? For example, how do people use theory of mind to infer and incorporate the knowledge and strategies of others while thinking many steps ahead? What is the optimal algorithm under which to traverse a decision tree if people have a limited memory capacity rather than a simple global cost per computation?

How can we construct process-level models of human planning that scale to the complexity of real-world environments? The computational cognitive model for 4-in-a-row is a start, but descriptive models that incorporate a wider variety of heuristics from the literature are needed.

How general are the signatures of planning found in combinatorial games? 4-in-a-row and chess are two tasks that require forward thinking, but there are countless other planning tasks that may or may not share similar underlying principles.

Which constraints posed by the neuroscience of planning in animals apply to humans, and how do they modify the algorithms typically used to model planning? The study of multi-step planning in the brain has made great progress towards linking various neural structures with behavior.

How can additional tools from artificial intelligence be leveraged to better understand human planning? Heuristic search and artificial neural networks are examples of such instances, but contemporary methods such as large language models can also be used towards this goal.

intelligence, with the goal of creating agents that can achieve superhuman performance in complex environments. To illustrate how these innovations can be used to better characterize the algorithms underlying human planning, we examined studies across two tasks where evaluating every course of action is intractable. The increasing availability of massive data sets and the development of novel computational methods will enable key questions to be addressed moving forward, in particular in understanding how people plan in complex environments (see Outstanding questions).

Acknowledgments

This work was made possible by grants from the National Science Foundation (number IIS-2312373) and the NOMIS Foundation.

Declaration of interests

The authors declare no competing interests.

References

1. Daw, N.D. *et al.* (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69, 1204–1215
2. Akam, T. *et al.* (2015) Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLoS Comput. Biol.* 11, e1004648
3. Russek, E.M. *et al.* (2017) Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Comput. Biol.* 13, e1005768
4. Feher da Silva, C. and Hare, T.A. (2020) Humans primarily use model-based inference in the two-stage task. *Nat. Hum. Behav.* 4, 1053–1066
5. Doll, B.B. *et al.* (2015) Model-based choices involve prospective neural activity. *Nat. Neurosci.* 18, 767–772
6. Vikbladh, O.M. *et al.* (2024) Consolidation of sequential planning. *bioRxiv*, Published online November 1, 2024. https://doi.org/10.1101/2024.11.01.621475
7. Liu, Y. *et al.* (2021) Experience replay is associated with efficient nonlocal learning. *Science* 372, eabf1357
8. Momennejad, I. *et al.* (2018) Offline replay supports planning in human reinforcement learning. *eLife* 7, e32548
9. Sharp, P.B. and Eldar, E. (2024) Humans adaptively deploy forward and backward prediction. *Nat. Hum. Behav.* 8, 1726–1737
10. Nicholas, J. *et al.* (2025) Proactive and reactive construction of memory-based preferences. *Nat. Commun.* 16, 1618
11. Dickinson, A. (1985) Actions and habits: the development of behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–78
12. Dolan, R.J. and Dayan, P. (2013) Goals and habits in the brain. *Neuron* 80, 312–325
13. Daw, N.D. *et al.* (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711
14. Lee, S.W. *et al.* (2014) Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81, 687–699
15. Keramati, M. *et al.* (2011) Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7, e1002055
16. Kool, W. *et al.* (2017) Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychol. Sci.* 28, 1321–1333
17. Sutton, R.S. (1991) Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bull.* 2, 160–163
18. Gershman, S.J. *et al.* (2014) Retrospective revaluation in sequential decision making: a tale of two systems. *J. Exp. Psychol. Gen.* 143, 182
19. Todorov, E. (2008) General duality between optimal control and estimation. In *2008 47th IEEE Conference on Decision and Control*, pp. 4286–4292, IEEE
20. Botvinick, M. and Toussaint, M. (2012) Planning as inference. *Trends Cogn. Sci.* 16, 485–488
21. Solway, A. and Botvinick, M.M. (2012) Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychol. Rev.* 119, 120
22. Parr, T. *et al.* (2022) *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*, MIT Press
23. Solway, A. and Botvinick, M.M. (2015) Evidence integration in model-based tree search. *Proc. Natl. Acad. Sci. U. S. A.* 112, 11708–11713
24. Huys, Q.J.M. *et al.* (2012) Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Comput. Biol.* 8, e1002410
25. Huys, Q.-J. *et al.* (2015) Interplay of approximate planning strategies. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3098–3103
26. Snider, J. *et al.* (2015) Prospective optimization with limited resources. *PLoS Comput. Biol.* 11, e1004501
27. Newell, A. and Simon, H. (1956) The logic theory machine–a complex information processing system. *IRE Trans. Inform. Theory* 2, 61–79
28. Newell, A. *et al.* (1959) Report on a general problem solving program. In *IFIP Congress* (Vol. 256), pp. 64
29. Marr, D. (2010) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, MIT Press
30. Anderson, J.R. (2013) *The Adaptive Character of Thought*, Psychology Press
31. Keramati, M. *et al.* (2016) Adaptive integration of habits into depth-limited planning defines a habitual-goal–directed spectrum. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12868–12873
32. Sezener, C.A. *et al.* (2019) Optimizing the depth and direction of prospective planning using information values. *PLoS Comput. Biol.* 15, e1006827
33. Russell, S. and Wefald, E. (1991) Principles of metareasoning. *Artif. Intell.* 49, 361–395
34. Griffiths, T.-L. *et al.* (2015) Rational use of cognitive resources: levels of analysis between the computational and the algorithmic. *Topics Cognit. Sci.* 7, 217–229
35. Lieder, F. and Griffiths, T.L. (2017) Strategy selection as rational metareasoning. *Psychol. Rev.* 124, 762
36. Callaway, F. *et al.* (2022) Rational use of cognitive resources in human planning. *Na. Hum. Behav.* 6, 1112–1125
37. I. Kuperwajs, *et al.* Heuristics for meta-planning from a normative model of information search. PsyArXiv. Published online March 26, 2024. https://doi.org/10.31234/osf.io/qtk9j
38. Moreno-Bote, R. *et al.* (2020) Heuristics and optimal solutions to the breadth–depth dilemma. *Proc. Natl. Acad. Sci. U. S. A.* 117, 19799–19808
39. Mastrogiuseppe, C. and Moreno-Bote, R. (2022) Deep imagination is a close to optimal policy for planning in large decision trees under limited resources. *Sci. Rep.* 12, 10411
40. Tomov, M.-S. *et al.* (2020) Discovery of hierarchical representations for efficient planning. *PLoS Comput. Biol.* 16, e1007594
41. Correa, C.-G. *et al.* (2023) Humans decompose tasks by trading off utility and computational cost. *PLoS Comput. Biol.* 19, e1011087
42. Correa, C.-G. *et al.* (2025) Exploring the hierarchical structure of human plans via program generation. *Cognition* 255, 105990
43. Ho, M.-K. *et al.* (2022) People construct simplified mental representations to plan. *Nature* 606, 129–136

44. Dayan, P. (1993) Improving generalization for temporal difference learning: the successor representation. *Neural Comput.* 5, 613–624
45. Momennejad, I. *et al.* (2017) The successor representation in human reinforcement learning. *Nat. Hum. Behav.* 1, 680–692
46. Piray, P. and Daw, N.D. (2021) Linear reinforcement learning in planning, grid fields, and cognitive control. *Nat. Commun.* 12, 4942
47. Namboodiri, V.M.K. and Stuber, G.D. (2021) The learning of prospective and retrospective cognitive maps within neural circuits. *Neuron* 109, 3552–3575
48. Precup, D. and Sutton, R.S. (1997) Multi-time models for temporally abstract planning. *Adv. Neural Inf. Proces. Syst.* 10, 1050–1056
49. van Opheusden, B. *et al.* (2023) Expertise increases planning depth in human gameplay. *Nature* 618, 1–6
50. Schulz, E. *et al.* (2019) Structured, uncertainty-driven exploration in real-world consumer choice. *Proc. Natl. Acad. Sci. U. S. A.* 116, 13903–13908
51. Steyvers, M. and Schafer, R.J. (2020) Inferring latent learning factors in large-scale cognitive training data. *Nat. Hum. Behav.* 4, 1145–1155
52. Brändle, F. *et al.* (2023) Empowerment contributes to exploration behaviour in a creative video game. *Nat. Hum. Behav.* 7, 1481–1489
53. Fernandez Velasco, P. *et al.* (2025) Expert navigators deploy rational complexity–based decision precaching for large-scale real-world planning. *Proc. Natl. Acad. Sci. U. S. A.* 122, e2407814122
54. Olieslagers, J. *et al.* (2024) Backward reasoning as search through and-or trees in human problem solving. *PsyArXiv*, Published online November 19, 2024. https://doi.org/10.31234/osf.io/h9bsu
55. Dickson, J.-M. and MacLeod, A.K. (2004) Approach and avoidance goals and plans: their relationship to anxiety and depression. *Cogn. Ther. Res.* 28, 415–432
56. Sharp, P.-B. (2024) Anxiety involves altered planning. *Trends Cogn. Sci.* 29, 118–121
57. Sutton, R. and Barto, A. (2018) *Reinforcement Learning: An Introduction* (2nd edn), MIT Press
58. Bonet, B. and Geffner, H. (2001) Planning as heuristic search. *Artif. Intell.* 129, 5–33
59. Campbell, M. *et al.* (2002) Deep blue. *Artif. Intell.* 134, 57–83
60. Dechter, R. and Pearl, J. (1985) Generalized best-first search strategies and the optimality of a, A. *J. ACM* 32, 505–536
61. Treisman, A.M. and Gelade, G. (1980) A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136
62. van Opheusden, B. *et al.* (2020) Unbiased and efficient log-likelihood estimation with inverse binomial sampling. *PLoS Comput. Biol.* 16, e1008483
63. Acerbi, L. and Ma, W.J. (2017) Practical bayesian optimization for model fitting with bayesian adaptive direct search. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 1834–1844
64. Kuperwajs, I. *et al.* (2019) Prospective planning and retrospective learning in a large-scale combinatorial game. In *2019 Conference on Cognitive Computational Neuroscience*, pp. 13–16
65. Lin, D.-X. *et al.* (2024) Limitations in planning ability in AlphaZero. In *NeurIPS 2024 Workshop on Behavioral Machine Learning*
66. Ma, I. *et al.* (2022) Distinct developmental trajectories in the cognitive components of complex planning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44), pp. 1131–1137
67. Holton, E. *et al.* (2025) Disentangling the component processes in complex planning impairments following ventromedial prefrontal lesions. *J. Neurosci.* 45, e1814242025
68. Agrawal, M. *et al.* (2020) Scaling up psychology via scientific regret minimization. *Proc. Natl. Acad. Sci. U. S. A.* 117, 8825–8835
69. Peterson, J.-C. *et al.* (2021) Using large-scale experiments and machine learning to discover theories of human decision-making. *Science* 372, 1209–1214
70. Battleday, R.-M. *et al.* (2020) Capturing human categorization of natural images by combining deep networks and cognitive models. *Nat. Commun.* 11, 1–14
71. Dezfouli, A. *et al.* (2019) Models that learn how humans learn: the case of decision-making and its disorders. *PLoS Comput. Biol.* 15, e1006903
72. Miller, K. *et al.* (2024) Cognitive model discovery via disentangled RNNs. *Adv. Neural Inf. Proces. Syst.* 36, 61377–61394
73. Ji-An, L. *et al.* (2024) Discovering cognitive strategies with tiny recurrent neural networks. *bioRxiv*, Published online October 5, 2024. https://doi.org/10.1101/2023.04.12.536629
74. Kuperwajs, I. *et al.* (2023) Using deep neural networks as a guide for modeling human planning. *Sci. Rep.* 13, 20269
75. Shannon, C.-E. (1950) XXII. Programming a computer for playing chess. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 41, 256–275
76. Tesauro, G. *et al.* (1995) Temporal difference learning and TD-Gammon. *Commun. ACM* 38, 58–68
77. Silver, D. *et al.* (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 484–489
78. Stern, D. *et al.* (2006) Bayesian pattern ranking for move prediction in the game of Go. In *Proceedings of the 23rd International Conference on Machine Learning*, pp. 873–880
79. Clark, C. and Storkey, A. (2015) Training deep convolutional neural networks to play Go. In *International Conference on Machine Learning*, pp. 1766–1774, PMLR
80. Silver, D. *et al.* (2017) Mastering the game of Go without human knowledge. *Nature* 550, 354–359
81. Silver, D. *et al.* (2018) A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 1140–1144
82. McIlroy-Young, R. *et al.* (2020) Aligning superhuman AI with human behavior: chess as a model system. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1677–1687
83. Russek, E. *et al.* (2022) Time spent thinking in online chess reflects the value of computation. *PsyArXiv*, Published online October 12, 2022. https://doi.org/10.31234/osf.io/8j9zx
84. Holdaway, C. and Vul, E. (2021) Risk-taking in adversarial games: what can 1 billion online chess games tell us? *Proceedings of the Annual Meeting of the Cognitive Science Society* Vol. 43, 986–992
85. Anderson, A. *et al.* (2017) Assessing human error against a benchmark of perfection. *ACM Trans. Knowl. Discov. Data (TKDD)* 11, 1–25
86. Kuperwajs, I. *et al.* (2024) Learning from rewards and social information in naturalistic strategic behavior. *PsyArXiv*, Published online August 14, 2024. https://doi.org/10.31234/osf.io/d8zje
87. Wunderlich, K. *et al.* (2012) Dopamine enhances model-based over model-free choice behavior. *Neuron* 75, 418–424
88. Gläscher, J. *et al.* (2010) States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595
89. Wunderlich, K. *et al.* (2012) Mapping value based planning and extensively trained choice in the human brain. *Nat. Neurosci.* 15, 786–791
90. Miller, K.-J. *et al.* (2017) Dorsal hippocampus contributes to model-based planning. *Nat. Neurosci.* 20, 1269–1276
91. Groman, S.-M. *et al.* (2018) Chronic exposure to methamphetamine disrupts reinforcement-based decision making in rats. *Neuropsychopharmacology* 43, 770–780
92. Akam, T. *et al.* (2021) The anterior cingulate cortex predicts future states to mediate model-based action selection. *Neuron* 109, 149–163
93. Tolman, E.-C. (1948) Cognitive maps in rats and men. *Psychol. Rev.* 55, 189
94. Johnson, A. and Redish, A.D. (2007) Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27, 12176–12189
95. Pfeiffer, B.E. and Foster, D.J. (2013) Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* 497, 74–79
96. Mattar, M.G. and Daw, N.D. (2018) Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* 21, 1609–1617
97. Pezzulo, G. *et al.* (2019) Planning at decision time and in the background during spatial navigation. *Curr. Opin. Behav. Sci.* 29, 69–76

98. Wise, T. *et al.* (2021) Model-based aversive learning in humans is supported by preferential task state reactivation. *Sci. Adv.* 7, eabf9616
99. Russek, E.-M. *et al.* (2024) Heuristics in risky decision-making relate to preferential representation of information. *Nat. Commun.* 15, 4269
100. Eldar, E. *et al.* (2020) The roles of online and offline replay in planning. *eLife* 9, e56911
101. Wimmer, G.-E. *et al.* (2023) Distinct replay signatures for prospective decision-making and memory preservation. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2205211120
102. De Groot, A.-D. (2014) *Thought and Choice in Chess*, Vol. 4. De Gruyter Mouton
103. Chase, W.-G. and Simon, H.A. (1973) Perception in chess. *Cogn. Psychol.* 4, 55–81
104. Linhares, A. *et al.* (2012) Entanglement of perception and reasoning in the combinatorial game of chess: differential errors of strategic reconstruction. *Cogn. Syst. Res.* 13, 72–86
105. Gobet, F. and Simon, H.A. (1998) Expert chess memory: revisiting the chunking hypothesis. *Memory* 6, 225–255
106. Bilalić, M. *et al.* (2010) Mechanisms and neural basis of object and pattern recognition: a study with chess experts. *J. Exp. Psychol. Gen.* 139, 728
107. Holding, D.-H. (1992) Theories of chess skill. *Psychol. Res.* 54, 10–16
108. Holding, D.-H. (1989) Counting backward during chess move choice. *Bull. Psychon. Soc.* 27, 421–424
109. Saariluoma, P. (1992) Visuospatial and articulatory interference in chess players' information intake. *Appl. Cogn. Psychol.* 6, 77–89
110. Campitelli, G. and Gobet, F. (2004) Adaptive expert decision making: skilled chess players search more and deeper. *ICGA J.* 27, 209–216
111. Ericsson, K.-A. and Smith, J. (1991) *Toward a General Theory of Expertise: Prospects and Limits*, Cambridge University Press
112. Gobet, F. and Jansen, P. (1994) Towards a chess program based on a model of human memory. *Adv. Comput. Chess* 7, 35–60
113. Gobet, F. (1997) A pattern-recognition theory of search in expert problem solving. *Think. Reason.* 3, 291–313