

Précis of *Cognition as a sequential decision problem*

Frederick Callaway

We must be prepared to accept the possibility that what we call “the environment” may lie, in part, within the skin of the biological organism. —Herb Simon (1955)

How can we build theoretically satisfying and practically useful models of the human mind? Historically, there have been two broad approaches. The *rational* approach, exemplified by Anderson (1990), focuses on characterizing the problems people have to solve and the optimal solutions to those problems. Under the assumption that the mind is well adapted to its environment, these optimal solutions then serve as models of cognition. Rational models are satisfying because they tell us *why* the mind works the way it does, and they are useful because they allow us to make generalizable predictions about how people will behave in new environments (i.e., rationally). However, by construction, such models don't explain *how* the mind achieves the rational ideal, and a growing list of systematic cognitive biases (Kahneman, 2011) draws their predictive utility into question.

In contrast, the *mechanistic* approach focuses on identifying the cognitive processes underlying behavior, often with an emphasis on explaining the behavioral idiosyncrasies that rational models gloss over. This approach can potentially tell us how the mind actually works, and it can produce extremely accurate models. However, lacking the optimality constraint, there is an enormous space of possible mechanistic models, and they often have many free parameters that are tuned for specific experimental setups. We are thus left wondering why this specific model fit the data best, and whether it would continue to make good predictions in a slightly different context.

Although the rational and mechanistic approaches have traditionally been viewed as conflicting, the past decade has seen a resurgence of an old idea (Simon, 1955): rationality can be seen as a property of cognitive mechanisms themselves. Specifically, a cognitive mechanism is rational if it makes optimal use of limited cognitive resources. Going under various names—cognitively bounded rational analysis (Howes et al., 2009), computational rationality (Lewis et al., 2014; Gershman et al., 2015), and resource-rational analysis (Griffiths et al., 2015; Lieder & Griffiths, 2020) to name a few—this view suggests that we should not expect people to be rational in the traditional sense of taking actions that maximize expected utility (Von Neumann & Morgenstern,

1944). Instead, we should expect people to select actions using mental strategies that strike a good tradeoff between the utility of the chosen action and the cognitive cost of making the decision.

But what defines a “good” tradeoff between action utility and cognitive cost? And how can we identify mental strategies that achieve such a tradeoff? In my dissertation, I suggest answers to these questions based on a key insight: *a rational mental strategy is one that optimally solves the sequential decision problem posed by one’s internal computational environment*. Under this view, cognition is a problem of stringing together a series of basic cognitive operations, or “computations”, in the service of choosing what to do in the world. An optimal cognitive process strings those basic operations together in such a way that maximizes the difference between the utility of the ultimate behavior and the total cost of all the cognitive operations that support the behavior.

The dissertation consists of six chapters. In the first chapter, I introduce the idea that cognition is a sequential decision problem, and provide theoretical context motivating my formalization of this idea. In the second chapter, I give a complete specification of that formalization using the framework of *metalevel Markov decision processes*. In the next three chapters, I illustrate how my colleagues and I have applied the framework to understand how people think about the present (attention), the past (memory), and the future (planning). In each case, we use process-tracing data to reveal in which ways people’s cognitive processes are consistent with and different from optimal cognitive processes. Taken together, the results suggest that people’s mental strategies are well-adapted to the limitations posed by their cognitive architectures, often in ways that could not be revealed without considering the sequential nature of cognition. In the final chapter, I identify ways in which the framework could be further developed.

CHAPTERS 1 AND 2: METALEVEL MARKOV DECISION PROCESSES

What does it mean to say that cognition is a sequential decision problem? One way to understand this claim is as an analogy between the type of problems posed by our external, physical environments and the type of problems posed by our internal, mental environments. To make things concrete, consider the problem facing a delivery robot, illustrated in Figure 1A. Completing the delivery will require visiting a sequence of locations before arriving at the final destination. And at each location, the robot will need to decide where to go next. Thus, the robot faces a sequential decision problem. Figure 1B illustrates how this type of problem is often modeled in artificial intelligence research: a *Markov decision process*, or MDP. At each time step, an agent (here, the robot) takes an *action* (e.g., driving forward). This action causes the environment to enter a new *state* (e.g., one where the robot is in a new location). Additionally, the agent receives a *reward*, a number that captures how good or bad the immediate consequences of the action are. For example, the delivery robot might receive a large positive reward for reaching the destination and a small negative reward every time it moves (capturing the desire to conserve battery life). The robot’s

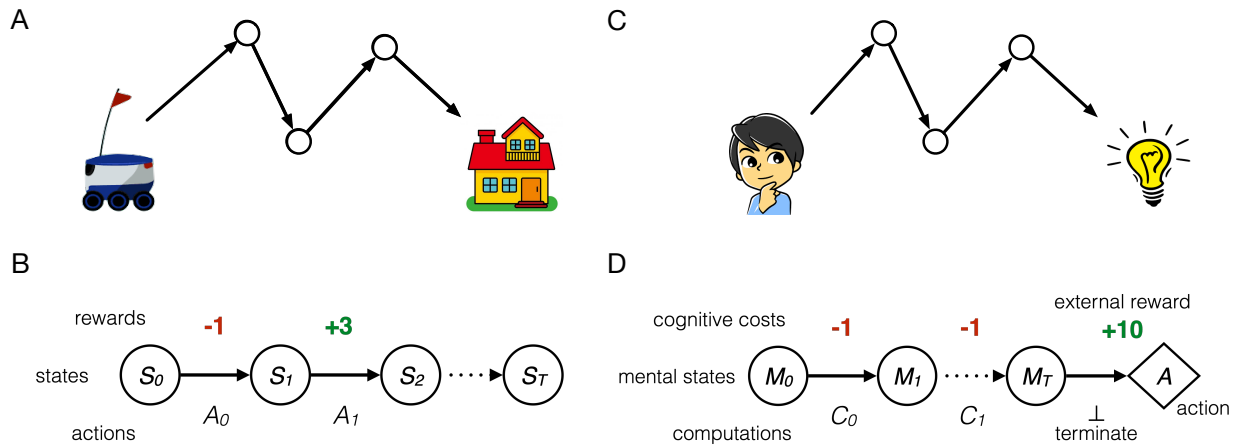


Figure 1: Sequential decision problems posed by external and internal environments. External environments (A) are often modeled by a Markov decision process, or MDP (B). An MDP is defined by a set of possible states the world can be in, a set of actions the agent can perform, a reward function that encodes the agent’s goals, and a transition function describing how actions change the state (arrows). An internal, cognitive environment (C) can be modeled as a *metalevel* MDP (D). A metalevel MDP is defined by a set of mental states and a set of computational operations that change that mental state (as specified in the metalevel transition function). The metalevel reward function captures both the cost of executing those operations as well as the external utility of the behavior that the computation produces.

goal is to maximize the total reward received.

Figure 1C illustrates a seemingly very different type of situation: a person trying to come up with a solution to a difficult problem. However, as the diagram suggests, the two cases actually share the same basic structure. Both involve an extended interaction between an agent and an environment; but whereas the robot is interacting with an *external* environment, the thinker is interacting with an *internal* environment: their own mind. Just as the robot makes several moves, and visits several locations before reaching the destination, the thinker has several thoughts, and enters several mental states before discovering the solution. Indeed, as illustrated in Figure 1D, this problem can be modeled in precisely the same way as the delivery problem. However, now the actions correspond to *computations* and the states correspond to *mental states*. Thinking changes one’s mental state just as moving changes one’s physical state; and it also incurs a cost—at the very least, thinking takes time. Because this MDP describes the metacognitive problem of how to interact with one’s own mind, it is called a *metalevel* MDP (Hay, 2016).

The power of identifying this parallel between external and internal environments is that it allows us to leverage existing knowledge about solving MDPs (a substantial chunk of AI research) to build rational mechanistic models of cognition. By modeling the internal environment associated with some cognitive system as a metalevel MDP, we can identify the optimal cognitive process as the optimal policy for that MDP. In the next three chapters, I show how this provides insights about three domains of cognition.

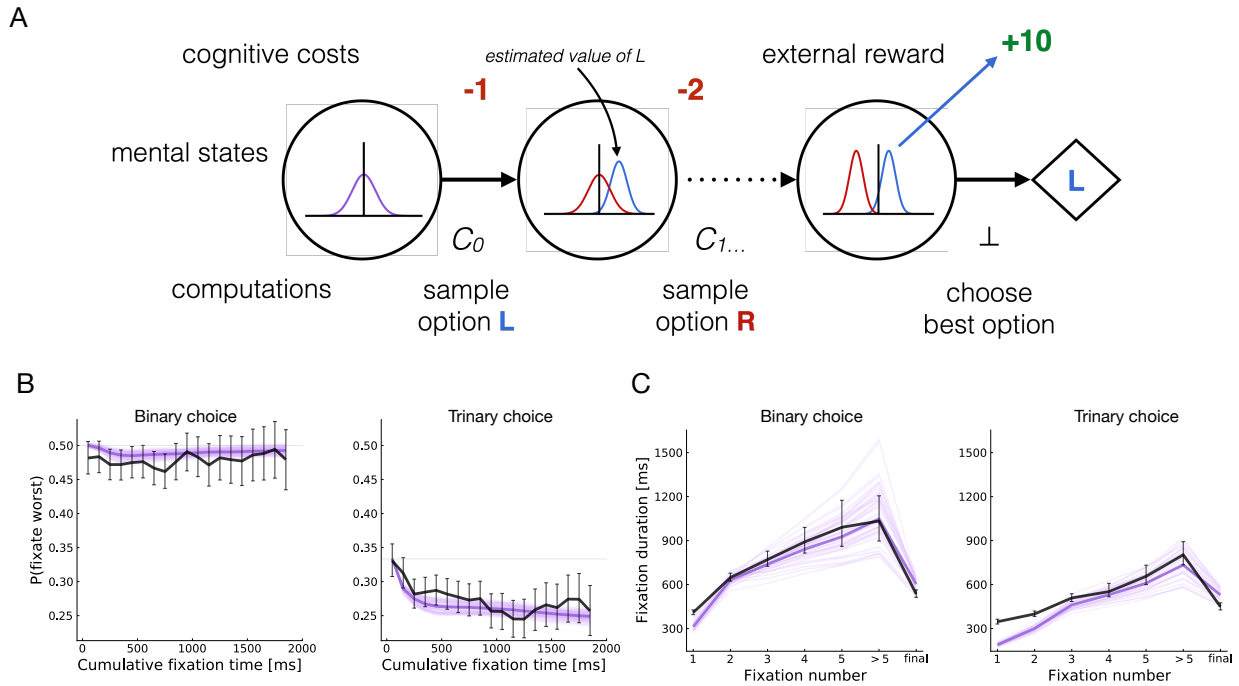


Figure 2: Attention (A) Meta-level MDP. The mental state encodes a distribution over the value of each option in the choice set. A computation corresponds to sampling the value of an option and updating its estimated value by Bayesian inference. **(B)** Both people (black) and the model (purple) are equally likely to attend to the better or worse option in binary choice (left). In trinary choice however (right), they are less likely to attend to the worst option in the set as the decision progresses. **(C)** Both people and the model show increasing fixation durations over the course of a decision.

CHAPTER 3: ATTENTION

Consider the problems faced by a diner at a buffet or a shopper at a supermarket shelf. They are presented with a number of options and must evaluate them until they identify the most desirable one. Previous work suggests that these decisions are made by integrating noisy evidence about the value of each alternative, which is sampled over time (Ratcliff & McKoon, 2008; Milosavljevic et al., 2010; Usher & McClelland, 2001). Moreover, this process is guided by visual attention, such that the evidence is strongest for the item currently being looked at; as a result, what we look at has consequences for what we choose (Shimojo et al., 2003; Armel et al., 2008; Krajbich et al., 2010). This raises an important question: **How do we decide what to pay attention to when making decisions?**

This problem is naturally cast as a metalevel MDP (Figure 2A). The mental states correspond to the agent's current estimates of (and uncertainty about) the value of each item. The computations correspond to attending to a given option, drawing a noisy sample of its value, and integrating that information into the corresponding value estimate by Bayesian inference. This sampling and updating process is encoded by the transition function. Finally, the reward function imposes a

fixed cost for each sample, an additional switching cost for saccades, and a reward equal to the true value of the chosen option.

Solving the metalevel MDP yields an optimal policy for allocating attention when making decisions. We compare this policy with human attention allocation using two datasets in which participants chose between junk food snacks (either two or three per trial) while their gaze was recorded with an eye tracker (Krajbich et al., 2010; Krajbich & Rangel, 2011). We can simulate this kind of data from the model by assuming that the attended item is fixated (100ms for each sample drawn). In the dissertation, I conduct a thorough comparison of optimal and human attention allocation in simple choice. Here, I highlight two of the most striking findings.

Figure 2B shows the probability of fixating on the worst option in the choice set over the course of a decision. When choosing between two items (left), both people (black) and the model (purple) are just about equally likely to fixate on the better and worse item throughout the decision. When choosing between three items, however, they are both decreasingly likely to look at the worst option as time goes on. Why do we see this pattern? Intuitively, when considering more than two options, one should focus on the top two contenders, as refining the value estimate for these options is most likely to change one's mind about which option is best—indeed, this is exactly what the optimal policy does. When there are only two options, they are by definition both in the top two, and thus there is no asymmetry. Furthermore, this effect emerges over time because attention must be allocated based on one's *estimates* of value, which only become related to the items' true value after some information has already been sampled.

To highlight a less intuitive prediction, Figure 2C shows the duration of fixations over the course of a decision. Although the model tends to underpredict the duration of the first two fixations in the three-item case, it captures well three key patterns: (a) the final fixation is shorter, (b) later (but non-final) fixations are longer and (c) fixations are substantially longer in the two-item case. The first occurs because final fixations are cut off when a choice is made, as in previous evidence accumulation models (Krajbich et al., 2010). The latter two patterns are unique predictions of the optimal model. They arise because more evidence is needed to alter beliefs when their precision is already high; this occurs late in the trial, especially in the two-item case where samples are split between fewer items.

Critically, these qualitative differences in the model predictions in binary and trinary choice are not the result of model fitting. The same set of parameters are used to generate predictions in each case, and the key qualitative differences are highly robust to the values of those parameters. This illustrates the power of rational models to generalize.

This work was published in *PLoS Computational Biology* (Callaway et al., 2021).

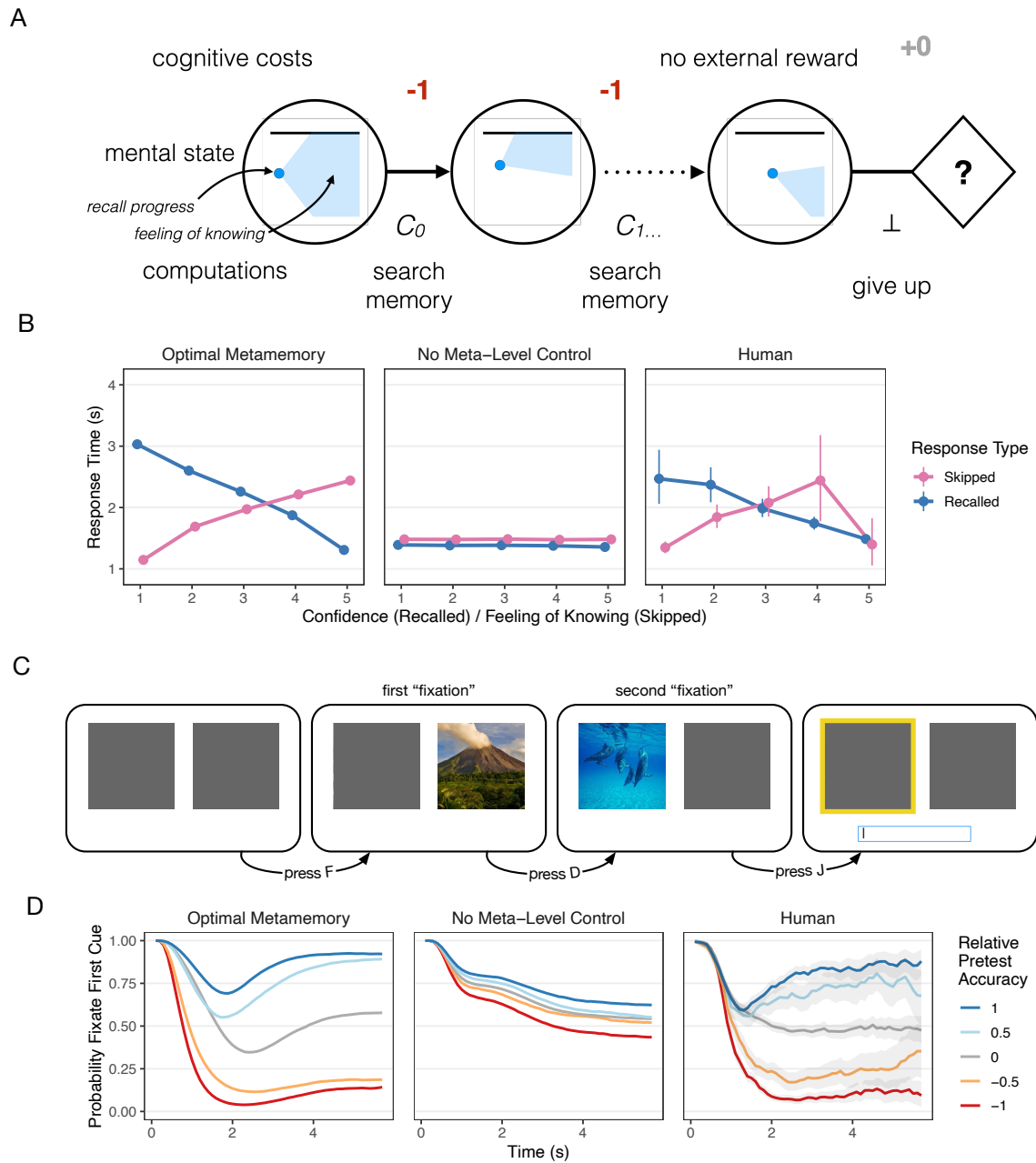


Figure 3: Memory (A) Meta-level MDP. The mental state encodes the current recall progress and an estimate of memory strength (rate of progress). A computation corresponds to searching for the target memory, generating recall progress. (B) Reaction time as a function of metamemory judgment (feeling of knowing for skip trials, confidence for recall trials), separately for trials in which participants correctly recalled the target vs. skipped without responding. The models' metamemory judgments are made based on the inferred memory strength at the end of the trial. (C) Multiple-cued recall task. Participants were presented with two images on each trial and could recall the word associated with either of them. Only one cue was visible at a time; participants could flip between them with the D and F keys. At any point they could press J or K to select an image for recall. (D) Timecourse of attention to the first cue, split by its memory strength (operationalized by accuracy in the pretest phase).

Consider next the all-too-familiar situation of running into someone whose name you cannot recall. If it feels as though the name is about to come to you, you may pause before saying hello. But every moment you delay only exacerbates the awkwardness of the situation. **How do people know when to stop searching for a memory?**

Most empirical work on metacognition in the domain of memory (or “metamemory”) has focused on how people are able to monitor their memory states (Reder & Ritter, 1992; Eakin, 2005) and on evaluating the accuracy of those judgments (Hart, 1965; Vesonder & Voss, 1985; Dunlosky & Lipko, 2007). Less emphasis, however, has been placed on understanding the function of metamemory judgments (Schwartz & Metcalfe, 2017). In a highly influential paper, Nelson & Narens (1990) proposed that the function of metacognitive systems is to allow effective control of ongoing cognition. For example, they outlined a theory in which a dynamically updated feeling of knowing is used to inform the decision of when to terminate an unsuccessful recall attempt. However, despite this early progress, previous work has not proposed a computational model of how these feeling of knowing estimates might be dynamically generated, nor of how they could be used to control recall efforts.

In this chapter, I provide such a model, formalizing the decision of when to terminate a memory search as a metalevel MDP (Figure 3A). Here, the mental states capture the amount of progress one has made towards recalling a memory as well as a metacognitive estimate of the rate of progress (that is, a “feeling of knowing” Hart, 1965). The computations correspond to continuing to search for the target. The transition function describes how recall progress noisily accumulates through search. Finally, the reward function encodes the value of recalling the item (when a threshold level of progress is reached) and a fixed cost for each moment spent searching.

Solving this metalevel MDP reveals that it is optimal to quickly abandon an attempt to recall a weak memory, one for which recall progresses slowly. The model thus captures the consistent empirical finding that people search longer before giving up on memories that they report having a higher feeling of knowing for Nelson, 1984; Nhouyvanisvong & Reder, 1998; Gruneberg et al., 1977; Lachman et al., 1979. In one striking demonstration of this phenomenon, Costermans et al. (1992) found opposite relationships between judgments of memory strength and response time when the item was recalled vs. not recalled: people gave high confidence judgments for items that were recalled quickly but low feeling-of-knowing judgments for items that were skipped quickly. As shown in Figure 3B, the optimal model reproduces this pattern, while a model without metalevel control (but the same underlying memory recall process) fails to capture either effect.

One weakness with the above finding, and indeed with many empirical results in metamemory, is that we observe at most one single metacognitive decision in each trial, the decision to terminate search. To provide a stronger test of the sequential search-evaluate loop, my colleagues

and I designed a modified cued-recall paradigm in which two candidate memories could be recalled on each trial (Figure 3C). By tracking attention to the corresponding cues using a keypress-contingent display, we could observe this metacognitive process unfolding even before a memory was recalled or given up on. Interestingly, Figure 3D shows that, contrary to the decision-making case studied in the previous chapter, the optimal policy preferentially attends to cues associated with stronger memories even when there are only two options. Participants showed the same pattern. This provides strong evidence that people can estimate the strength of a memory before recalling it, and use that information to guide their recall efforts. To our knowledge, it is also the first quantitative demonstration of a metamemory process unfolding over time.

CHAPTER 5: PLANNING

Finally, consider the problem faced by a traveler in an unfamiliar country, deciding which cities to visit. Geographical concerns limit which cities they can travel between, so their first stop will shape the whole trip. But there are far too many possible routes to consider every one. **How do we know which hypothetical courses of action to evaluate, and which to ignore?**

How people solve problems that require thinking multiple steps ahead has been a focal question since the very earliest attempts to understand the human mind in computational terms (Newell & Simon, 1956; Newell et al., 1972). Both then and now, most work on human planning has focused on identifying the heuristics people use to circumvent the intractability of exhaustive planning. For example, people might limit the depth of their search (MacGregor et al., 2001; Keramati et al., 2016; Krusche et al., 2018; Snider et al., 2015), “prune” away initially unpromising courses of action (Huys et al., 2012, 2015), or avoid planning altogether by relying on habit or “memoization” (Huys et al., 2015; Kool et al., 2017). However, the questions of why people use those particular heuristics and of predicting which of the many possible heuristics people will employ in any particular case, have been relatively unexplored.

To provide a normative theory of planning under computational constraints, we can model planning—specifically, decision tree search—as a metalevel MDP (Figure 4A). A mental state corresponds to a partially constructed decision tree, which represents possible sequences of actions and outcomes as a tree-structured graph. A computation corresponds to *node expansion*; this operation determines the cost or reward for visiting a state, integrates that value into the total value of the path leading to that state, and adds the immediate successors of the target state to the search frontier, that is, the set of nodes that can be expanded on the next iteration. The transition function encodes these dynamics—critically, it also encodes prior information the agent may have about where large rewards are likely to be found. Finally, the metalevel reward function assigns a fixed cost for each node expansion operation; and when the agent terminates planning, they receive a reward equal to the expected value of the external rewards that will be gained by executing

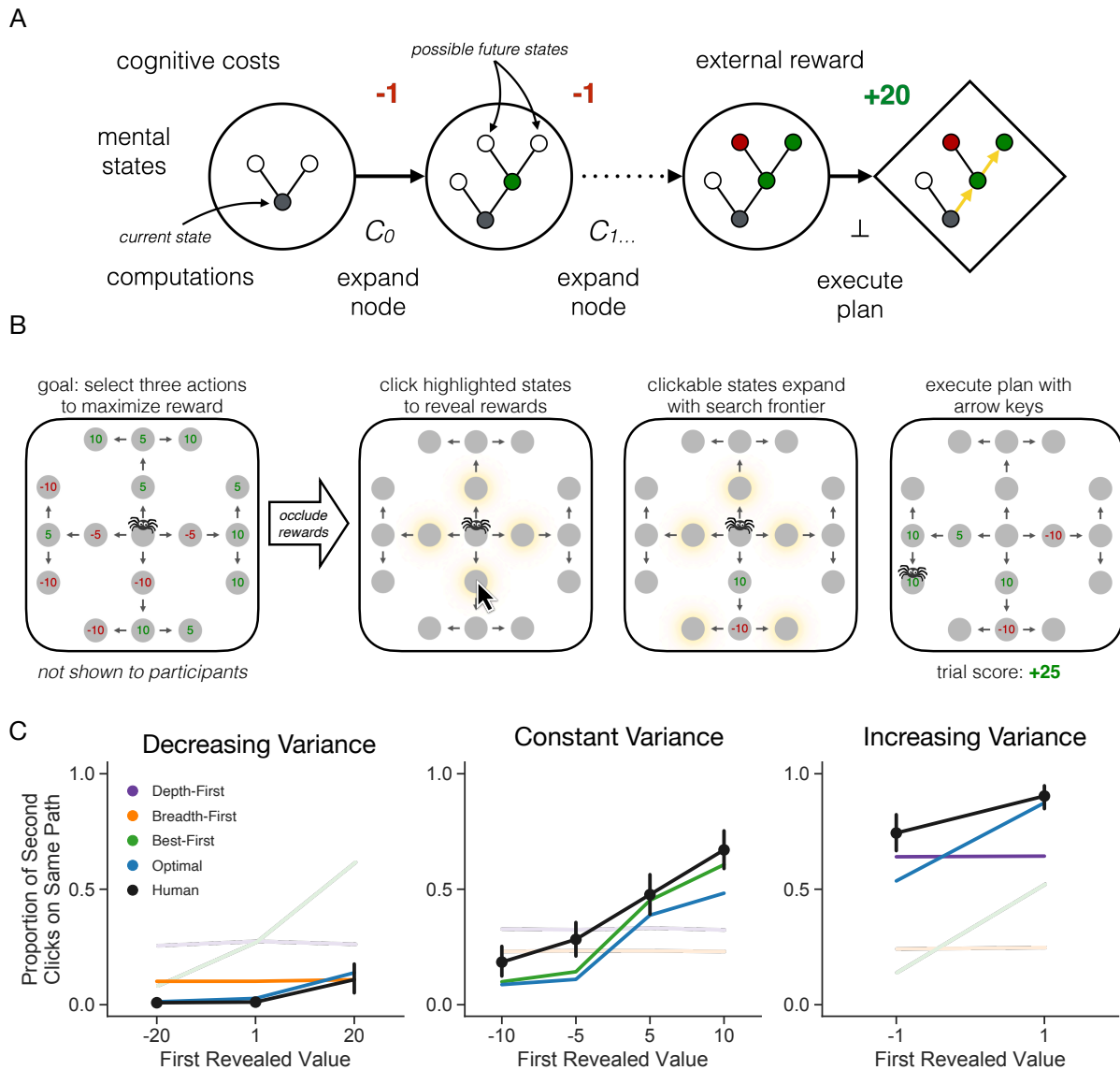


Figure 4: Planning (A) Meta-level MDP. The mental state is a decision-tree, which represents possible future actions (edges), states (nodes), and rewards (green/red). A computation corresponds to considering a future state and expanding the decision tree. **(B)** Experimental task. The internal process of expanding a decision tree is externalized by forcing participants to click to reveal the reward to be gained at each state. **(C)** Adaptation of planning strategy in environments with different reward structures (see main text). Each panel shows the probability of making a second click on the same path as the first click, depending on the value revealed by that first click. Models that achieve high reward in the respective environment are highlighted.

a plan chosen with the current decision tree.

Solving this metalevel MDP yields an optimal planning algorithm. But how can we compare this algorithm to human planning, given that the latter takes place entirely inside a person’s head? To circumvent this challenge, my colleagues and I designed a task that makes people’s planning directly observable (Figure 4B). Inspired by the classic Mouselab paradigm (Payne et al., 1988), our task externalizes planning operations as information-gathering actions. Specifically, participants must click future states to see what reward they would gain if they visited that state. The sequence of clicks thus reveals the order in which the participant considered each state. This allows us to evaluate candidate models at the level of individual node expansion operations, providing a stronger and more objective test than is possible with previous approaches based on model comparison over the actions people take (e.g., Huys et al., 2015; Van Opheusden et al., 2017) or verbal reports of their planning process (e.g., De Groot, 1965; Newell et al., 1972).

While I have emphasized adaptation to internal environments, our cognitive processes are also shaped by the external environment. To investigate this kind of adaptation, we constructed three environments with different reward distributions. In the “constant variance” environment, all states had the same reward distribution. In the other two environments, most states had small rewards and extreme rewards could only be found in one state on each path: the first state in the “decreasing variance” environment and the last state in the “increasing variance” environment. We designed these environments such that the optimal planning algorithm resembled a different classical algorithm in each case: breadth-first for decreasing variance, best-first for constant variance, and depth-first for increasing variance. Our participants appeared to adapt their planning strategies accordingly, as revealed by the frequency with which they continued down the first-clicked path with their second click (Figure 4C), doing so rarely in the decreasing variance case (like breadth-first), frequently in the increasing variance case (like depth-first), and only after revealing a large reward in the constant variance case (like best-first). Although one heuristic model in each condition resembles the behavioral pattern, only the optimal model could capture behavior in all three conditions.

So far, I have focused on the cases where people resemble the optimal model. However, we also found systematic deviations from optimal planning. Most notably, when we removed the search-frontier constraint (highlighted states in Figure 4B) in Experiments 3 and 4, we found a strong bias towards considering states in the order in which they would be traversed. This suggests that our metalevel MDP does not capture all the factors constraining human planning in naturalistic settings. In particular, in many cases people may not be able to sample arbitrary future states, and when they can they may have access to a generative model that makes it easier to simulate in temporal order than to reason backwards from effect to cause. If people’s planning algorithms are adapted to the naturalistic case, we would expect to see discrepancies when these important constraints are removed. This highlights another key strength of rational models: the cases where

the model is “wrong” can be just as—if not more—informative than the cases where it’s right.

This work was published in *Nature Human Behavior* (Callaway et al., 2022).

CHAPTER 6: CONCLUSION

In the final chapter, I discuss related theoretical approaches and identify four key directions in which the metalevel MDP framework could be extended in future work: (1) metalevel reinforcement learning, (2) incomplete or imperfect metacognition, (3) interleaved computation and action, and (4) optimization of the metalevel MDP itself. I then reflect on the role of frameworks in cognitive science, focusing on the ways in which they constrain our thinking, and emphasizing that these constraints are both a blessing and a curse.

My dissertation presents a general framework for modeling cognition as a sequential decision problem: metalevel Markov decision processes. It shows how the framework can be applied to derive rational mechanistic models in three different domains: attention, memory, and planning. And in each case, we find that human behavior showed substantial qualitative alignment with the optimal metalevel policy. Taken together, the results suggest that human cognitive processes are well-adapted to the internal environments in which they operate. More importantly, by formally characterizing the problems posed by those mental environments, and their optimal solutions, we develop a richer understanding of human cognition in each of these domains.

REFERENCES

- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Psychology Press.
- Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, 3(5), 396–403.
- Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling. *PLOS Computational Biology*, 17(3), e1008863.
- Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., & Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nat Hum Behav*, 6(8), 1112–1125.
- Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(1), 142–150.
- De Groot, A. D. (1965). *Thought and Choice in Chess*. The Hague: De Gruyter Mouton.
- Dunlosky, J. & Lipko, A. R. (2007). Metacomprehension: A Brief History and How to Improve Its Accuracy. *Curr Dir Psychol Sci*, 16(4), 228–232.
- Eakin, D. K. (2005). Illusions of knowing: Metamemory and memory under conditions of retroactive interference. *Journal of Memory and Language*, 52(4), 526–534.
- Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278.
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2), 217–229.
- Gruneberg, M. M., Monks, J., & Sykes, R. N. (1977). Some methodological problems with feeling of knowing studies. *Acta Psychologica*, 41(5), 365–371.
- Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of educational psychology*, 56(4), 208.
- Hay, N. (2016). Principles of Metalevel Control.
- Howes, A., Lewis, R. L., & Vera, A. (2009). Rational Adaptation Under Task and Processing Constraints: Implications for Testing Theories of Cognition and Action. *Psychological Review*, 116(4), 717–751.

- Huys, Q. J. M., Eshel, N., O’Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLOS Computational Biology*, 8(3), e1002410.
- Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., Dayan, P., & Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences of the United States of America*, 112(10), 3098–103.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Keramati, M., Smittenaar, P., Dolan, R. J., & Dayan, P. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal-directed spectrum. *Proceedings of the National Academy of Sciences*, 113(45), 12868–12873.
- Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychological Science*, 28(9), 1321–1333.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Krajbich, I. & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33), 13852–13857.
- Krusche, M. J. F., Schulz, E., Guez, A., & Speekenbrink, M. (2018). Adaptive planning in human search. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- Lachman, J. L., Lachman, R., & Thronesbery, C. (1979). Metamemory through the adult life span. *Developmental Psychology*, 15(5), 543–551.
- Lewis, R. L., Howes, A., & Singh, S. (2014). Computational rationality: Linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science*, 6(2), 279–311.
- Lieder, F. & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43.
- MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(1), 176–201.
- Milosavljevic, M., Malmaud, J., & Huth, A. (2010). The Drift Diffusion Model can account for the accuracy and reaction time of value-based choices under high and low time pressure. 5(6), 437–449.

- Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95(1), 109–133.
- Nelson, T. O. & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. In G. H. Bower (Ed.), *Psychology of Learning and Motivation*, volume 26 (pp. 125–173). Academic Press.
- Newell, A. & Simon, H. (1956). The logic theory machine—A complex information processing system. *IRE Transactions on Information Theory*, 2(3), 61–79.
- Newell, A., Simon, H. A., et al. (1972). *Human Problem Solving*, volume 104. Prentice-hall Englewood Cliffs, NJ.
- Nhouyvanisvong, A. & Reder, L. M. (1998). Rapid feeling-of-knowing: A strategy selection mechanism. In *Metacognition: Cognitive and Social Dimensions* (pp. 35–52). Thousand Oaks, CA, US: Sage Publications, Inc.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), 534–552.
- Ratcliff, R. & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput*, 20(4), 873–922.
- Reder, L. M. & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, memory, and cognition*, 18(3), 435.
- Schwartz, B. L. & Metcalfe, J. (2017). Metamemory: An Update of Critical Findings. In *Learning and Memory: A Comprehensive Reference* (pp. 423–432). Elsevier.
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6(12), 1317–1322.
- Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
- Snider, J., Lee, D., Poizner, H., & Gepshtein, S. (2015). Prospective Optimization with Limited Resources. *PLOS Computational Biology*, 11(9), e1004501.
- Usher, M. & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108(3), 550–592.
- Van Opheusden, B., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2017). A computational model for decision tree search. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Vesonder, G. T. & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language*, 24(3), 363–376.

Von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ, US: Princeton University Press.