

## Memory, Reinforcement Learning, and Temporally Extended Agency

This paper is about the role of memory in temporally extended intentional agency. An agent engages in temporally extended intentional agency if they intentionally pursue a single goal persistently across time (for instance, taking half a day to assemble a complicated piece of IKEA furniture). Depending on how one conceives of the way in which agency extend through time, one will conceive differently of the roles played by memory in enabling agency to extend.

We find two different conceptions of temporally extended intentional agency in contemporary neuroscience, psychology, and philosophy. On the one hand, some people think that temporarily extended agency is reducible to a series of proximal decisions (Sripada, 2025). In this paper, I focus on recent computational accounts of action selection and control in Reinforcement Learning (RL) models in cognitive neuroscience (Niv, 2009; Collins, 2024). According to this proximal conception, an agent's practical decision is always a decision about what to do at the most immediate choice point. The shape of a long-term project is constituted by the way in which the stable shape of the reward landscape shapes the trajectory of in-the-moment choices.

On the other hand, some people think that temporally extended agency can be controlled by long-term intentions to do something in the distal future (Bratman, 1987). According to this distal conception, an agent can make future-directed decisions and thereby form intentions that control actions by their temporally distal goals. The shape of a long-term project can be caused by a stable, persisting intention formed by making a future-directed decision. The stability of the intention can at times be at odds with proximal values and rewards.

The proximal and distal conceptions of temporally extended intentional agency imply different conceptions of memory. According to the proximal conception, decisions about what to do are always about imminent choices, and memory plays an important role only by providing the decision-maker with information about the world. Semantic and episodic memory helps the agent build a cognitive map of the choice space (Matar & Daw, 2018). The map represents values and facts. Memory thus helps the agent calculate possible cumulative future rewards given choices made here and now (Shohamy & Daw, 2015; Gershman & Daw, 2017). The stability of temporally extended agency is partly explained by a stable cognitive map (i.e., a stable space of options with stable assignments of values and possibilities) and memory mechanisms to maintain and retrieve the map.

By contrast, according to the distal conception, decisions about what to do can be about actions in the future, and memory can play a role in retaining a representation of the intended future action (Grünbaum & Kyllingsbæk, 2020; Oren et al., 2021). Future-directed intentions are sticky. Once formed, the intention is not reevaluated at each new choice point. Memory plays an important role in shielding the agent from in-the-moment motivational fluctuations and temptations. In this paper, I argue that semantic and episodic forms of memory are not

sufficient for satisfying this additional role of memory in retaining a representation of and commitment to a future action. I argue that, given their reliance on only semantic and episodic forms of memory, recent RL versions of the proximal account of agency are unable to explain everyday cases of resisting temptation, familiar from the planning theory (Bratman, 1987; Holton, 2009).

My argument proceeds as follows. First, I describe a simple case of an agent resisting the temptation of postponing the performance of a costly action. I consider an agent who must perform a costly action before Friday but on each weekday Monday to Thursday has the option of delaying the action. As time approaches Friday, the risk of not performing the action (due to unforeseen events) increases. Psychological studies have shown that, in such situations, agents would be able to plan and execute the action early in the week (Sheeran et al., 2024).

Second, I introduce the basic logic of RL models of temporally extended intentional agency. My focus is on intentional action. Consequently, I will only discuss so-called model-based accounts of action selection (Daw & Dayan, 2014). I provide some details on how current RL models conceive of episodic memory as providing information about the possible rewards to be obtained by different possible actions.

Third, I review three different ways in which an RL account might explain the ability to plan and execute the costly action early in the week. According to the first RL explanation, an agent's true preferences are revealed by the actions they select (Redish, 2013). Consequently, performing the action early in the week just demonstrates that it was in fact not costly (compared to other available options). According to the second RL explanation, if agent's compute the increasing risk involved in postponing the performance, in the long run the cost of postponing might be too high. The agent might adopt some kind of meta-control policy that will allow it to perform the action early in the week (Kool et al., 2018). According to the third RL explanation, agents can simulate performing the early action which could enable them to build so-called "successor representations" (Momennejad, 2020). This might enable the agent to chunk state-action trajectories. The main argument of this paper is that all reviewed RL explanations are unable to provide a satisfactory explanation of our ability to plan and later intentionally perform the costly action at a time when it can easily be postponed. The crux of the argument is centred on RL models' proximal conception of temporally extended agency and their commitment to semantic and episodic roles of memory.

I conclude the paper by briefly sketching an alternative conception of the role of memory in temporally extended intentional agency. If RL models with their proximal conception of action selection and control are unable to explain everyday temptation cases, we have reason to opt for a more distal account of temporally extended agency. I briefly suggest a possible alternative role for memory on such an account.