Practical Capacities, Know How, and Skill: The View from Reinforcement Learning

May 16, 2025

We know a lot. We know how to tie knots. We know how to navigate across rooms, buildings, and cities. We (or some of us) know how to make cheese. We know, or can come to know, how to ride a bicycle or play the piano.

We have all sorts of know-how, and we have it in abundance. And we are not alone in this: rats also know, in their own way, how to navigate across rooms, buildings, and cities and how to find cheese; with a little ingenuity, they might be taught to play some rudimentary piano. Ants know how to find their way home after finding food, and bees know how to communicate the location of pollen to one another.

Clearly, these capacities are diverse. But do they share a common nature? Is there a common explanation of their possession or exercise?

According to *intellectualism*, there is. Intellectualists hold that knowing how to do something is knowing that something is true (namely, that such-and-such is a way to do that thing).

Intellectualism is the view that

Intellectualists, however, would deny the rats, ants, and bees. For according to intellectualists, to know how to do something is to know that something is true (namely, that such-and-such is a way to do that thing). But rats, ants, and bees have no such knowledge. According to intellectualism, they therefore lack know-how. According to intellectualists, the rats do not, after all, know how to navigate rooms or find cheese (let alone play the piano).

1 Sharpening the question

In this section, we lay out some of the key positions and distinctions that structure the debate on practical capacities and know-how.

Intellectualism, as Ryle originally characterizes it [CITE: Ryle], is the view that knowing how to do something consists in the possession of some relevant propositional knowledge. Ryle was concerned to refute the view that, for example, my knowing how to play chess consists in my knowing various propositions, such as that it is good to get one's queen out early. Contemporary advocates of intellectualism have clarified the kind of propositional knowledge at issue, giving us the following stardard statement of intellectualism (φ ranges over actions):

> INTELLECTUALISM: knowing how to φ consists in knowing, of some way w of performing actions, that w is a way to φ.
>
> [CITE: Stanley and Williamson; Pavese]

Proponents of INTELLECTUALISM often add that the way w must be represented under a *practical mode of presentation*. Practical modes of presentation are supposed to be a species of Fregean modes of presentation. Modes of presentation support a fine-grained notion of mental content of the kind suitable for psychological and rational explanation [CITE: Fodor, Burge, Rescorla, Peacocke]. Practical modes of presentation index representational content to the exercise of practical capacities [CITE: Pavese]. They are typically invoked to explain why knowing how to φ (usually) enables one to φ. While we (unlike some critics of intellectualism [CITE: Noe]) do not find the notion of practical modes of presentation irredeemably obscure, they will not be our concern in what follows, and so we do not explain them any further.

[Should probably say a bit more about what a "way" is supposed to be here, especially since that *will* become relevant later.]

INTELLECTUALISM is a strong view. Requiring propositional representation for know-how is already a substantial commitment, implying as it does that minds incapable of entertaining propositions lack know-how. But INTELLECTUALISM goes far beyond. It requires a propositional representation (i) that w is a way of doing φ, where (ii) w is grasped under a specific mode of presentation (viz. practical), and (iii) this propositional representation must amount to knowledge. Each of (i)–(iii) strengthens the already-strong view that know-how requires propositional representation.

In addition, INTELLECTUALISM does not merely require propositional knowledge for knowing-how. It holds that propositional knowledge is constitutive of knowledge-how. A constitutive claim is typically stronger than a modal one [CITE: Fine]: [example here?].

Not all who identify as intellectualists would subscribe to INTELLECTUALISM as defined above. Some eschew practical modes of presentation [CITE!]. Others do not require knowledge but instead (true) belief [CITE!]. Yet others [CITE: Bengson and Moffett] hold that while know-how does not consist in propositional knowledge, it requires "reasonable mastery of the concepts in a correct and complete conception of a way of φ-ing." All of these views, however, maintain that propositional (or conceptual)[1] representations are necessary for know-how: a creature incapable of entertaining propositions cannot know how to do things.

Thus, while not all intellectualist views are committed to the claim that know-how consists in propositional knowledge, they *are* committed to the necessity of propositional representations for the possession of know-how. In fact, they are committed to a slightly stronger view, which will be my main target in what follows. They are committed to the following:

> Propositional representations are necessary to explain the possession and exercise of knowledge-how.

[1] [Connect propositional and conceptual representation]

This strengthening highlights that the issue is not whether a knower-how must be capable of propositional thought, but whether a capacity for propositional thought plays any explanatory role with respect to know-how. The issue is not whether

If one wished to deny, for whatever reason, that lower animals *really* possess know-how, one would be left with humans, and humans clearly do possess propositional capacities. But the question would remain: must these propositional capacities be implicated in any meaningful way in the possession or exercise of human know-how? That is the question I take to be at the center of the debates concerning know-how and practical capacities.

- introduce weaker views, that don't depend on propositional representation (way-representationalism etc.) - comment on the question of whether representations are involved at all—how does this connect to the original debate? - will probably have to say something about know-how vs practical capacities. - debate usually poorly framed; it's not propositions vs disposition. Question is whether and if so which mental capacities explain know-how/practical capacities.

2 RL and Marr's Levels

Contemporary reinforcement learning models are computational models. These models are presented at various levels of explanatory grain, following Marr's three levels [CITE: Marr 1982, Niv and Langdon 2016]. In most studies, the model specifies various quantities computed by the agent. For example, many models posit that agents maintain an estimate of action values (a Q-function). In addition, a model may posit that the agent estimates the variability in outcome attaching to actions, or the uncertainty associated with each state. In any case, a central part of most reinforcement learning models is a specification of the functions computed by the agent.

More ambitiously, the model may specify how the agent computes these values or functions. For example, the model may specify that action values are learned according to one of various iterative algorithms for value learning, such as Q-learning [CITE: Watson and Dayan 1992] or SARSA [CITE: Sutton and Barto]. We must be careful not to read too much into the choice of algorithm. In some cases, the choice of algorithm is a core commitment of the model. In others, it is a mere convenience.

Indeed, some studies are aimed at determining which of several competing algorithms are used by a given agent [CITE]. In such cases, it is reasonable to take a realist stance on specific aspects of the algorithm. For example, estimating the value of an action using SARSA involves sampling a "next action." The Q-learning algorithm is exactly the same, except that instead of sampling a next action, the agent considers the action with the highest estimated value among possible next actions. And the Expected-SARSA algorithm is like Q-learning, except that instead of considering the maximal value achievable by the next action, it considers the expected value of the next action. Thus, where SARSA samples, Q-learning takes a $\max$ operation, and Expected-SARSA takes an expectation. These algorithmic details entail behavioral and computational differences and require different cognitive capacities (for example, Expected-SARSA, but not the others, requires the agent to take an expectation over actions, and hence to maintain a probability distribution over actions). Determining which algorithm an agent uses is therefore a reasonable experimental goal. In such contexts, the success of a Q-learning model over a SARSA model provides *prima facie* evidence that the agent implements the Q-learning algorithm, and in particular that the agent computes a maximum operation.

In other contexts, however, this realist interpretation is unwarranted. For example, many studies target the question whether the agent employs model-based or model-free reinforcement learning [CITE: Daw, Niv, and Dayan, Drummond and Niv, Momennejad et al.]. In model-based reinforcement learning, the agent has access to a model of the causal or statistical structure of its environment. Typically, the agent learns this model over the course of its interactions with the environment and uses it to plan; the model itself is learned from experience, usually through association. In model-free reinforcement learning, the agent lacks a model. Instead, it (usually) caches estimates of the value of different actions and uses these action values to choose actions.

Both model-based and model-free reinforcement learning can be implemented via a wide range of algorithms (the three mentioned in the previous paragraph are all model-free algorithms; see [CITE: Sutton and Barto] for a small taste of the diversity of reinforcement learning algorithm). In studies designed to tease apart model-based and model-free methods, experimenters sometimes choose specific algorithmic implementations of each method in order to derive behavioral predictions. However, no effort is made to comprehensively search over different model-based and model-free algorithms to find the best fit. It is assumed (reasonably) that the behavioral differences predicted by model-based and model-free methods are robust to the choice of underlying algorithm. Indeed, in other such studies, no algorithm is proposed, and the behavioral differences between model-free and model-based reinforcement learning are instead characterized qualitatively. For example, a hallmark of model-free learning is its insensitivity to *outcome devaluation*: model-free learners will continue to pursue actions that have led to reward in the past, in spite of the fact that they now lead to undesirable (or not desirable) consequences. This difference arises from the structure of model-free and model-based learning and does not require experimenters to choose specific algorithmic implementations of either kind of learning. That a model-free method is a better fit than a model-based one on a given task thus lends virtually no support to a realist interpretation of the distinctive features of the chosen model-free algorithm (if any).[2]

The takeaway is that there is no automatic inference from the use of a particular algorithm in a reinforcement learning model to the psychological reality of the processes postulated by that algorithm. This is not to say that such inferences are never warranted: often they are. But the warrant depends on the explanatory use of the algorithm's features. If the distinctive features of the Q-learning algorithm play a role in explaining the behavioral or neural data, then it is reasonable to take a realist stance toward these features.

The two explanatory ambitions I have discussed thus far—ascribing the computation of a function and describing how that function is computed—correspond to Marr's computational and algorithmic levels of explanation.[3]

[2] Does it provide *any* support? Perhaps. But not enough to put much weight on the algorithmic details.

[3] The nomenclature is unfortunate: all three levels are computational in the sense of describing com-

Finally, and most ambitiously, cognitive scientists may seek to identify neural correlates of key algorithmic quantities or operations. Indeed, much of the early enthusiasm for reinforcement learning models owes to the discovery of specific neural mechanisms realizing a key quantity at the heart of many model-free algorithms: the *temporal difference* (TD) error.

The TD error is the difference between the agent's estimates of an action's value, $Q(a)$ and a bootstrapped estimate $B(a)$ of that value (in reality, action value estimates are indexed to the current state s; we suppress the state parameter for simplicity). $Q(a)$ is the agent's estimate of the value of action a at a given time t. The bootstrapped estimate is like the agent's estimate, except that it incorporates feedback from the environment. By incorporating this feedback, the bootstrapped estimate is statistically less biased than the original estimate. Most model-free reinforcement learning algorithms therefore push the agent's estimate $Q(a)$ in the direction of the bootstrapped estimate. To do so, they compute $B(a) - Q(a)$, the (signed) distance between the current estimate and the bootstrap. This difference is the TD error. It is an *error* insofar as the bootstrap estimate is statistically closer to the true value of the action than the estimate $Q(a)$.[4] Intuitively, a positive TD error indicates that the action is turned out to be better than expected; the good news causes the learner to revise its estimate upwards (and conversely in the case of a negative TD error).

The TD error is perhaps the most fundamental algorithmic idea in reinforcement learning. It allows for a simple iterative computation of action values (and from there of optimal policies) that relies only on locally available information: the reward obtained at that time step. As Sutton and Barto explain in their popular textbook,

> If one had to identify one idea as central and novel to reinforcement learning, it would undoubtedly be *temporal difference* (TD) learning.
>
> [CITE: Sutton and Barto: 119]

In particular, TD errors are at the heart of Q-learning and the actor-critic algorithms, arguably the two most influential model-free reinforcement learning algorithms.

[[Explain connection between TD error and domapine; neural realization of actor-critic algorithm]]

If there is evidence of brain regions implementing reinforcement learning algorithms, that is of course reason to take reinforcement learning models realistically. Unfortunately, however, the neurological evidence regarding the implementation of these algorithms is not as clear-cut as one might have hoped. The brain, as it turns out, is a complicated organ. This complexity gives rise to several difficulties: [draft]

- It is difficult to get very reliable evidence about the function of a given neural population.

putational processes and providing computational explanations. "Functional" might be a better term for the so-called computational level. And I am not sure that the algorithmic and implementational levels can always be cleanly distinguished; for one thing, what looks like implementation at one level is often algorithmic at another (see [CITE: Rueckl 1991: Connectionism and the notion of levels] for elaboration). But the distinction remains useful, and the terminology has stuck, so we follow the literature.

[4] To be precise, unless the agent's estimate $Q(a)$ is already accurate, the expected value of the bootstrap estimate is closer to the true action value than $Q(a)$ is.

- It is difficult to unambiguously interpret such evidence as one can get (even if we had noiseless data about the firing pattern of a given neuronal cluster, it may be very difficult to know why they exhibit this pattern, or what that pattern is for).
- The brain probably does not implement "pure" reinforcement learning. That is, if some form of reinforcement learning is implemented in the brain, it likely does not have the form of the "textbook" presentations of reinforcement learning. For example, the brain may be computing several reward signals at once, tracking different values and playing different computational roles. These refinements can be incorporated into the general machinery of reinforcement learning, but doing so might require conceptual developments in reinforcement learning itself (e.g. consider the introduction of task construals)

These points do not support an anti-realist stance about the use of reinforcement learning models. But they do show that one has to exercise cautious judgement in making inferences from the use of reinforcement learning in psychology and neuroscience to the existence of a given computational structure in a mind or brain.

3 RL and representation

3.1 Representation of actions

An RL agent must take an action at each time step. (We continue to use the anthropomorphic term "action" without thereby ascribing agency to RL systems in any interesting sense.[5] Recall also that an action, in the RL framework, can be just about anything.) Learning to choose good actions is the main point of RL algorithms. We therefore consider the question whether the agent must represent its own actions. As we will see, in some but not all cases, a representation of actions is necessary for learning.

Before getting into it, let us clarify the question. What would it mean for an RL agent to represent its actions? Representations are commonly understood as states with content. While there is no consensus regarding the nature and function, contents are minimally seen as providing *veridicality conditions* for mental states and as accounting for *rational relations* among mental states. Veridicality conditions are, roughly speaking, conditions that the world, or some part of the world, must meet to satisfy a representation with that content. For example, the belief that the mean temperature in Los Angeles in July is 95F has as content the proposition that the mean temperature in Los Angeles in July is 95F. This content determines conditions on the world which must be met if the belief is to be true. Namely, the content determines that the belief is true if and only if the mean temperature in Los Angeles in July is 95F.

[not sure I want to make the point about rational relations, actually. The idea is that content is invoked in psychological explanations (e.g. belief-desire psychology, but also elsewhere). Maybe note in passing that this role requires a relatively fine-grained notion of content, because of Frege cases.]

An action representation is unlikely to possess the same kind of content as a belief. For one thing, the content of an action representation specifies an action type, not a

[5]Though see [CITE: Butlin 2020, 2024] for an argument that RL systems are agents in a non-trivial sense.

state of the world. [Actually, what's the truth here? Do action representations represent actions, or states of affairs in which the actions are performed (by the agent?)?] More importantly, an action representation is connected to action in a more direct way than beliefs are. An action representation functions to initiate action. When all goes well (in particular, when downstream systems "cooperate"), an action representation issues in action. By contrast, a belief, even a belief that it would be good to do a given action, has no such direct connection to action. [Well, we have to be careful: an action representation doesn't function to issue in action whenever it's tokened—only when tokened "decisively" (e.g. when passed downstream to generate the action). But the point is that there is such a think as tokening an action representation decisively, whereas there is no such thing for belief.]

[Also lots to say here about how actions in RL differ from action representations in other domains, e.g. motor control/motor representations.]

[...]

The point of most RL algorithms is to learn a good policy. ...

3.2 Representations of Value