Reinforcement learning and linguistic convention

Paul Christiano Follow
Dec 4, 2015 · 3 min read

Existing machine learning techniques are most effective when we can provide concrete feedback—such as prediction accuracy or score in a game. For the purpose of AI control, I think that this is an important and natural category, and I really need a term for it.

I've been referring to this regime as "supervised" learning, but that's an extremely unusual use of the term and probably a mistake. I plan to use the term "reinforcement learning" in the future.

Sutton and Barto, who I will take as authorities, define reinforcement learning informally as:

> *learning what to do—how to map situations to actions—so as to maximize a numerical reward signal.*

If we construe "actions" broadly, so as to include cognitive actions (like forming or updating a representation), then this is exactly what I want to talk about. It's not clear whether Sutton and Barto mean to include supervised learning as a special case, but Barto at least writes: "it it is possible to convert any supervised learning task into a reinforcement learning task."

This category is much broader than what people normally mean when they talk about "reinforcement learning." It's probably even broader than what Sutton and Barto have in mind. But still, I think "reinforcement learning" captures what I want relatively well, and is definitely better than any other existing term. So I'm going to run with it.

From my perspective, the key (and almost only) assumption of reinforcement learning is that the objective-to-be-optimized comes in the form of numerical rewards that are **actually provided to the algorithm**. This is required for many modern methods (e.g. anything using gradient descent), even those that we don't normally think of as RL.

Clarifications

- The term “reinforcement learning” is often meant to be exclusive of supervised learning or other learning problems that fit into a narrower framework. I definitely *don’t* mean to use it in this narrower sense. I am using the term to refer to a very broad category that intentionally subsumes most modern machine learning.

- Reinforcement learning can be coupled with reward engineering. For example, supervised learning fits into this definition of reinforcement learning, since we can use the data distribution and loss function to define rewards.

- Reinforcement learning often refers to sequential decision problems, but I don’t mean to make this restriction. So far, I think these are generalizations that researchers in RL would agree with (though they’d likely consider sequential problems most interesting).

- Reinforcement learning implies an interaction between an agent and an environment, but I don’t mean to make any assumptions on the nature of the environment. The “environment” is just whatever process computes the rewards and observations. It could be anything from a SAT checker, to a human reviewer, to a board game, to a rich and realistic environment.

- Reinforcement learning often focuses on choosing actions, but I want to explicitly include cognitive actions. Some of these are clear fits—e.g. allocating memory effectively. Others don’t feel at all like “reinforcement learning”—e.g. learning to form sparse representations. But from a formal perspective a representation is just another kind of output, and the reinforcement learning framework captures these cases as well.

Supervised learning

Supervised learning seems like an important special case.

From the perspective of AI control, I think the most important simplifying assumption is that the information provided to the algorithm does not depend on its behavior. So there is no need for explicit exploration, data can be reused, and so on.