Tab 1

Reading list for Rich's RL approach to AI

Oct 2, 2023

At the end of this post there is a numbered list of the 10 most-important background readings for understanding my view of AI. Immediately below I comment on the readings, referencing them by number.

Reading [1] is the RL textbook. This is the best starting point and reference. It is also the best reference for deeply understanding the reasons behind the core algorithms. These particularly fascinate me, but I don't think we all have to understand them to the same degree.

What do we all have to understand? I think we have to understand the goal. I seek an architecture for AI that can operate online with limited computational resources, while interacting in real-time with a complex, open-ended environment. An agent design that could operate a robot with visual and other sensors and with a large multi-dimensional motor output---though we may choose to operate it in simulation or as a purely video-conferencing agent. Key here is what we call "the big-world perspective"---that the totality of the environment, which includes people and other agents, is much more complex than the mind of our AI agent. Thus, though we start with discrete states and tabular agents, any method we take seriously must be amenable to approximation. The big-world perspective implies that learning cannot be done once in advance and then never again. The environment is big and probably changing, so the agent must continually adapt to the part of the world that it is in now.

Reading [2] is the original temporal-difference (TD) paper. It introduces TD(lambda)---the one algorithm that I feel we all must understand completely. Of course TD(lambda) is also presented in [1], but it is spread across Chapters 6 and 12, and besides, sometimes going back in time, here to 1988, is a fun way to learn. But yeah, this reading is optional.

The next three readings are essential bits that are not in the RL textbook. Reading [3] is the key IDBD paper on step-size optimization. Reading [4] is the options paper on temporal abstraction. Reading [5] is the best first paper on GVFs.

Reading [6] is the Alberta Plan. After [1-5] this plan should make sense to you. The one missing bit from [1-5] is advanced planning. Basic planning (tabular, one step) is covered in [1], Chapter 8, while advanced planning does not have a really good reference. The best currently available is [7], the STOMP paper, which comes close to being a complete tabular-ish prototype AI. Also on planning I recommend the video [8] and, if you want more, consider the unpublished paper [9] and the Linear Dyna paper [10].

[1] Reinforcement Learning: An Introduction. Sutton & Barto (2018). The linked online version is slightly more up to date than the printed version.

[2] Learning to predict by the methods of temporal differences. Sutton (1988). I also very much like this video on TD learning.

[3] Adapting bias by gradient descent: An incremental version of delta-bar-delta. Sutton (1992)

[4] Between MDPs and semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning. Sutton, Precup, & Singh (1999)

[5] Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. Sutton, Modayil, Delp, Degris, Pilarski, White, & Precup (2011)

[6] The Alberta plan for AI research. Sutton, Bowling, & Pilarski (2022)

[7] Reward-respecting subtasks for model-based reinforcement learning. Sutton, Machado, Holland, Timbers, Tanner, & White (2023)

[8] Gaps in the Foundations of Planning with Approximation (video). Sutton (2021)

[9] Toward a New Approach to Model-based Reinforcement Learning. Sutton (2020)

[10] Dyna-style planning with linear function approximation and prioritized sweeping. Sutton, Szepesvari, Geramifard, Bowling (2008)