

Thinking Fast and Slow with Model-Based Reinforcement Learning

The book "Thinking Fast and Slow" explains two systems that govern the way we think. System 1 (habits) is fast, intuitive, emotional, and often without our conscious control; System 2 is slower, more deliberate, and logical and deals with conscious mind activities such as self-control, decisions (improvisation). Using our habits in situations where we do not have much knowledge can lead to bad results, but if we want to plan every action, we need much time. Therefore, we always use both.

This thesis aims to mimic the two systems in the field of model-based reinforcement learning (MBRL). MBRL refers to a branch of RL that aims to learn a model that approximates the dynamics of the environment. This model can generate augmented virtual measurements that can be used either to train a policy (background planning) or to plan the most optimal sequence of actions (decision-time planning). Background planning is very similar to our first system in that the computation of optimal actions at deployment is fast. Decision planning, on the other hand, is competent in unknown situations and acts without learning, making it analogous to our second system.

Tasks

- Train a model of the dynamic environment
- Design a method to estimate the uncertainty of the model's prediction (How familiar is the state we are in?).
- Design a hybrid approach of background planning and decision planning (For example, if there is not enough information to make a rational decision, use decision-time planning, otherwise, use background planning).
- Fine-tune the background planning using decision planning to fit new situations (create a new habit using your prior knowledge)

References

1. D. Kahneman "Thinking, Fast and Slow"
2. T. Anthony et al. "Thinking Fast and Slow with Deep Learning and Tree Search"
3. I. Mordatch and Jessica Hamrick "Tutorial on Model-Based Methods in Reinforcement Learning"
4. R. S. Sutton et al. "Reinforcement Learning: An Introduction"

