Advanced AI

Tuur Vanhoutte

September 22, 2021

Contents

1 Reinforcement learning in context

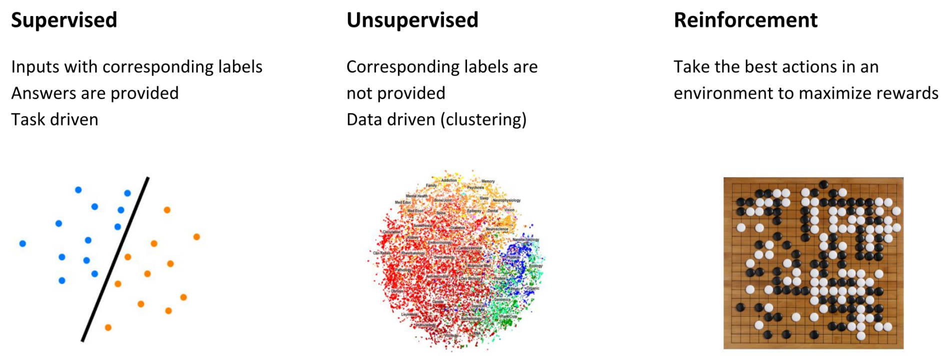

Figure 1: Overview learning algorithms

1.1 What is Reinforcement Learning

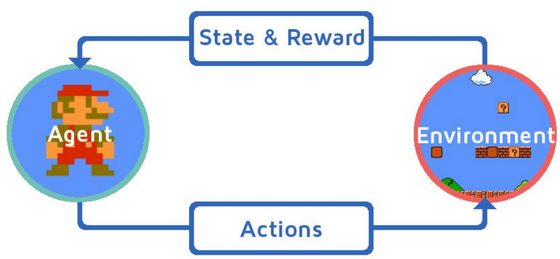

Figure 2: An agent interacts with the environment, changes its state, and gets rewards

1.1.1 Examples

- http://www.youtube.com/watch?v=vZ7eHwVhe9c
- https://deepmind.com/blog/article/alphago-zero-starting-scratch
- https://www.youtube.com/watch?v=4MlZncshy1Q
- https://www.youtube.com/watch?v=eG1Ed8PTJ18
- http://www.youtube.com/watch?v=tlOIHko8ySg
- https://www.youtube.com/watch?v=W_gxLKSsSIE
- https://www.youtube.com/watch?v=ZBFwe1gF0FU
- https://www.youtube.com/watch?v=VCdxqn0fcnE
- https://www.youtube.com/watch?v=opsmd5yuBF0
- https://www.youtube.com/watch?v=gn4nRCC9TwQ
- http://www.youtube.com/watch?v=2cjkKnAxCug
- http://www.youtube.com/watch?v=kopoLzvh5jY

1.1.2 Reinforcement learning vs Deep learning

Definition 1.1 ***Deep learning*** *seeks to iteratively minimize a certain loss function that indicates how accurate the functional representation of a system is.*

It responds to ***patterns*** *in the data. It typically assumes the data it works with is* ***independent and identically distributed (IDD)****, and with a* ***stationary distribution***

Definition 1.2 ***Reinforcement learning*** *seeks to iteratively maximize a certain notion of a numerical reward obtained through continued interaction with its environment.*

It learns to make ***sequential decisions****. There is no need for the data to have* ***IDD*** *properties. The agent* ***can deal with non-stationarity*** *in its environment during interaction*

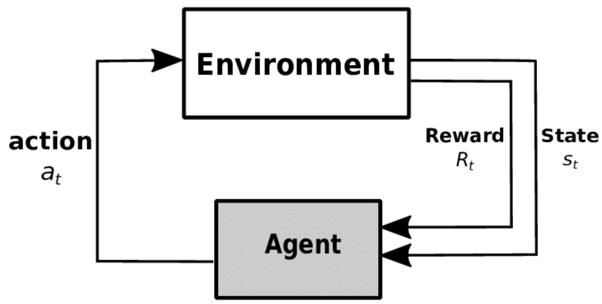

Figure 3: Reinforcement Learning terminologies

Definition 1.3 (Agent) *The learner and decision maker*

Definition 1.4 (Environment) *Where the agent learns and decides what actions to take*

Definition 1.5 (Action) *A set of actions which the agent can perform*

Definition 1.6 (State) *How the agent perceives the environment*

Definition 1.7 (Reward) *For each action selected by the agent, the environment provides a reward. Usually a scalar value.*

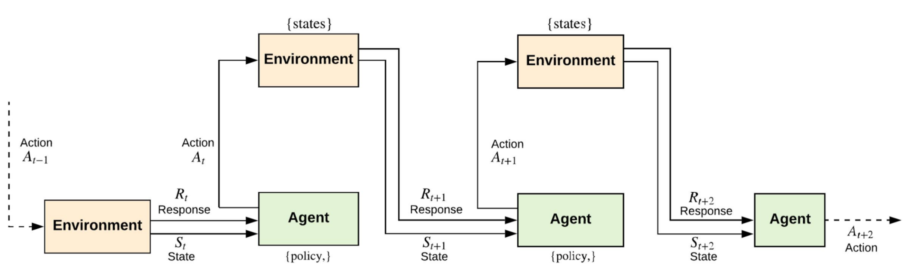

Figure 4: Reinforcement Learning terminologies

1. The agent executes an action in the environment
2. The environment responds with a reward and updates the state
3. The agent decides the best action based on the new state and the reward it received
4. Repeat . . .

Definition 1.8 (Policy) *The decision-making function (control strategy) of the agent, which represents a mapping from situations to actions. 'Brain of the agent', the strategy your agent will follow.*

Definition 1.9 (Value function) *Mapping from states to real numbers, where the value of a state represents the long-term reward achieved starting from that state, and executing a particular policy.*

Definition 1.10 (Function approximator) *Refers to inducing a function from training examples. Standard approximators include decision trees, neural networks, and nearest-neighbor methods*

Definition 1.11 (Markov Decision Process (MDP)) *A probabilistic model of a sequential decision problem, where states can be perceived exactly, and the current state and action selected determine a probability distribution on future states. Essentially, the outcome of applying an action to a state depends only on the current action and state (and not on preceding actions or states).*

Definition 1.12 (Model) *The agent's view of the environment, which maps state-action pairs to probability distributions over states. Note that not every reinforcement learning agent uses a model of its environment*

1.2 Reinforcement learning taxonomy

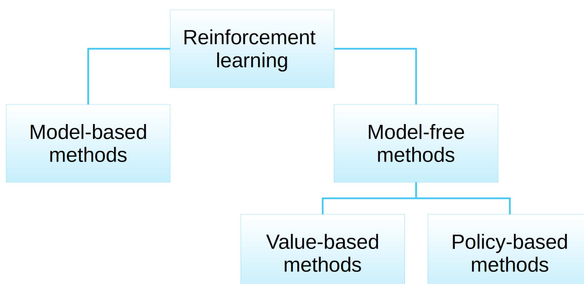

Figure 5: Two RL categories: model-based methods & model-free methods

1.2.1 Model based methods

- We can (partly) calculate the optimal actions using the model directly.
- The model may be known or learned.
- Can train from simulated experiences.
- Model-based RL has a strong advantage of being sample efficient.
- Usually computationally more complex.
- We will **not** focus on these methods in detail

`https://medium.com/@jonathan_hui/rl-model-based-reinforcement-learning-3c2b6f0aa323`

1.2.2 Model-free

- The model is (partly) ignored.
- Needs no accurate representation of the environment in order to be effective.
- Computationally less complex.

- Actual experiences need to be gathered in order for training, which makes exploration more dangerous.
- Cannot carry an explicit plan of how environmental dynamics affects the system

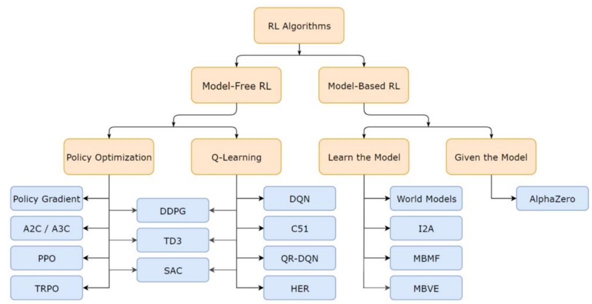

Figure 6: Model based vs model-free

Figure 7: Model based vs model-free in terms of sample efficiency

2 OpenAI Gym

Gym is a toolkit for developing and comparing reinforcement learning algorithms. It supports teaching agents everything from walking to playing games like Pong or Pinball.

```
# installation
pip install gym
```

Documentation: `https://gym.openai.com/docs/`