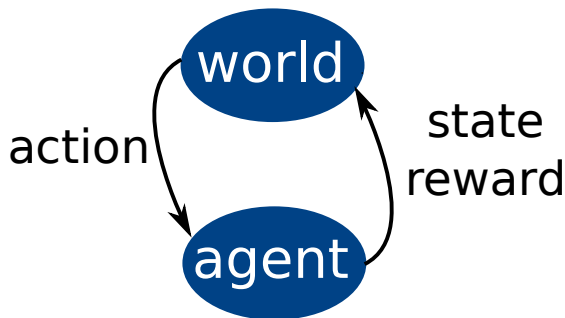




Your marketing dollars go far

Philippe Desjardins-Proulx, Xiao Wu, Harsha Gopianandan

Reinforcement Learning



A reinforcement learning agent is a smart program capable of adapting to the world (i.e.: your product). It takes feedback and reward (e.g.: user A bought X when Y) and then try to take the best action (i.e.: user A should be shown X).

How does it work?

First, you setup the agent! For example, let's say you run a website selling electronics and trying to maintain a loyal userbase. You want to optimize the ads you are showing: do you show a laptop, a desktop computer, a tablet. Which brand? Do you show a samsung tablet to someone who bought a samsung laptop? You also want to optimize how many ads to show, and type of ads. Being too aggressive might annoy your users, and perhaps you need to be more aggressive at some time, e.g. less aggressive with ads after a sale.

How does it work?

To setup the agent, you need to determine its goal: in this case selling as many electronics as possible. The agent's action: in this case showing ads of different size and of different products. And what the agent can see: in this case what a user saw, searched, and ultimately what he bought. The agent will learn how to optimize its actions to optimize its goal.

Teaching the agent

After the agent is setup, you can send information so it can adapt:

```
$ curl --data "history=lenovo_lap&bought=samsung_tab&ad=samsung_lap" www.conquer.ai/api/v0/  
$ curl --data "history=ipad&ad=apple&bought=macbookair" www.conquer.ai/api/v0/  
$ curl --data "history=ipad&ad=samsung&bought=nothing" www.conquer.ai/api/v0/
```

In this case we tell the agent that after seeing a large ad of a Lenovo laptop, the user bought a Samsung tablet, and that another user bought Apple products, and a third user was shown a Samsung ad and bought nothing. With this information, the agent is learning the ads to show.

Querying the agent

Then, you can query the agent to ask what product to show to a user:

```
$ curl --data "history=lenovo_lap&search=lap" www.conquer.ai/api/v0/  
$ { "type": "laptop", "brand": "Apple", "computer": "MacBook Pro" }
```

Not the whole story...

A feature that is hard to show in simple example is that reinforcement learning agent can think in terms of complex sequences of event. For example, showing the ad of an expansive product might yield a bit less money now but could yield greater long-term rewards if the user think of your company when he's thinking of buying an expansive computer.

For Engineers & Machine Learning Specialists

Three Paradigms of Machine Learning:

Supervised learning:

$$\{(\mathbf{x}, y)_i\}_{i=1}^n \mapsto (\hat{f}(\mathbf{x}) \mapsto y).$$

Unsupervised learning:

$$\{(\mathbf{x})_i\}_{i=1}^n \mapsto (\hat{f}(\mathbf{x})).$$

Reinforcement learning:

$$\langle S, A, T, R \rangle.$$

\mathbf{x} = input/independent var.

y = output/dependent var.

n = number of data points

\hat{f} = estimated function

S = states (environments)

A = actions

T = transitions

R = rewards

Supervised Learning \approx Classification, Regression.

Unsupervised Learning \approx Clustering.