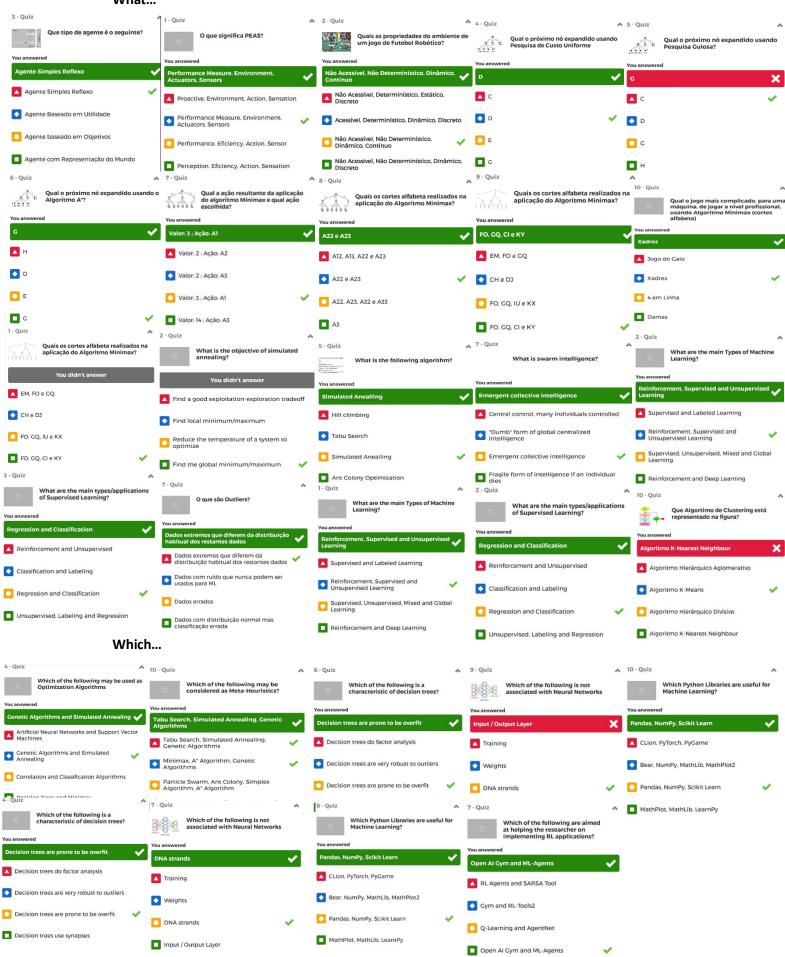
What...





The reward tells us the exact sequence of actions we should do to solve it

The greedy action selection strategy uses both exploit and explore (50/50)

The expected reward of a policy is the "value" of the policy

The multiplicative factor The reward factor

The penalty factor

