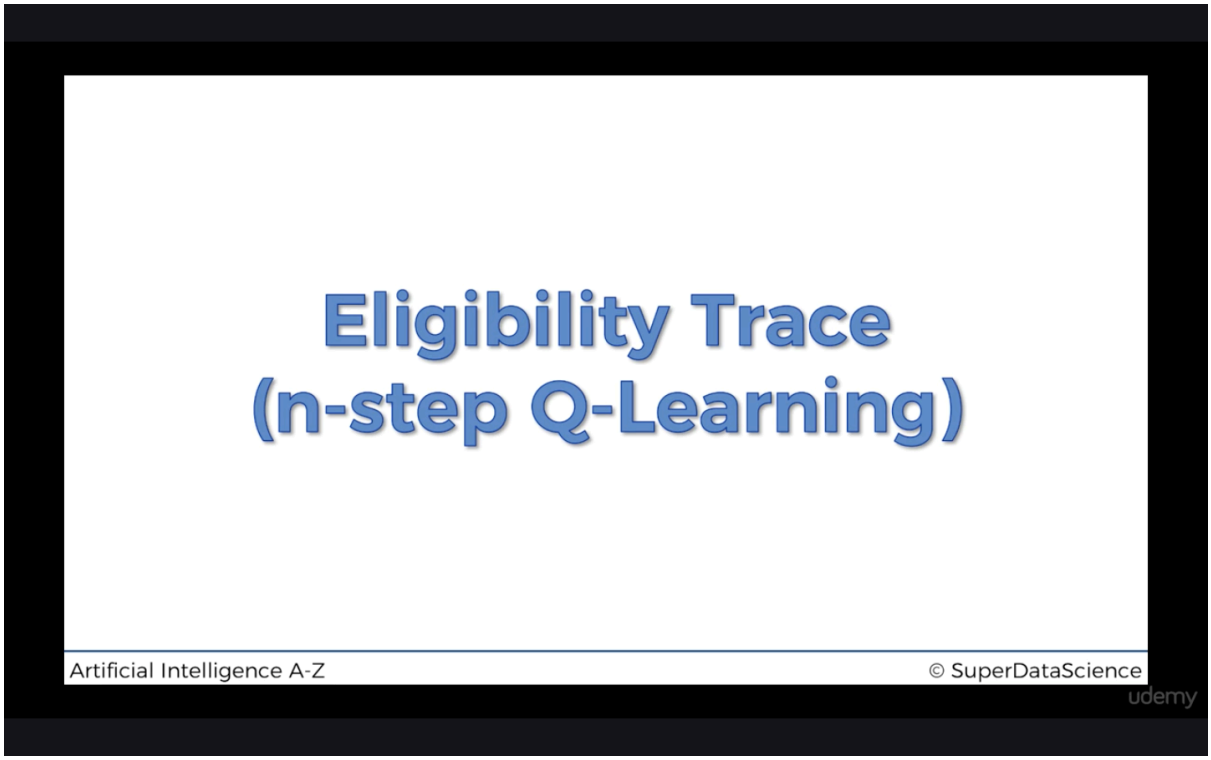
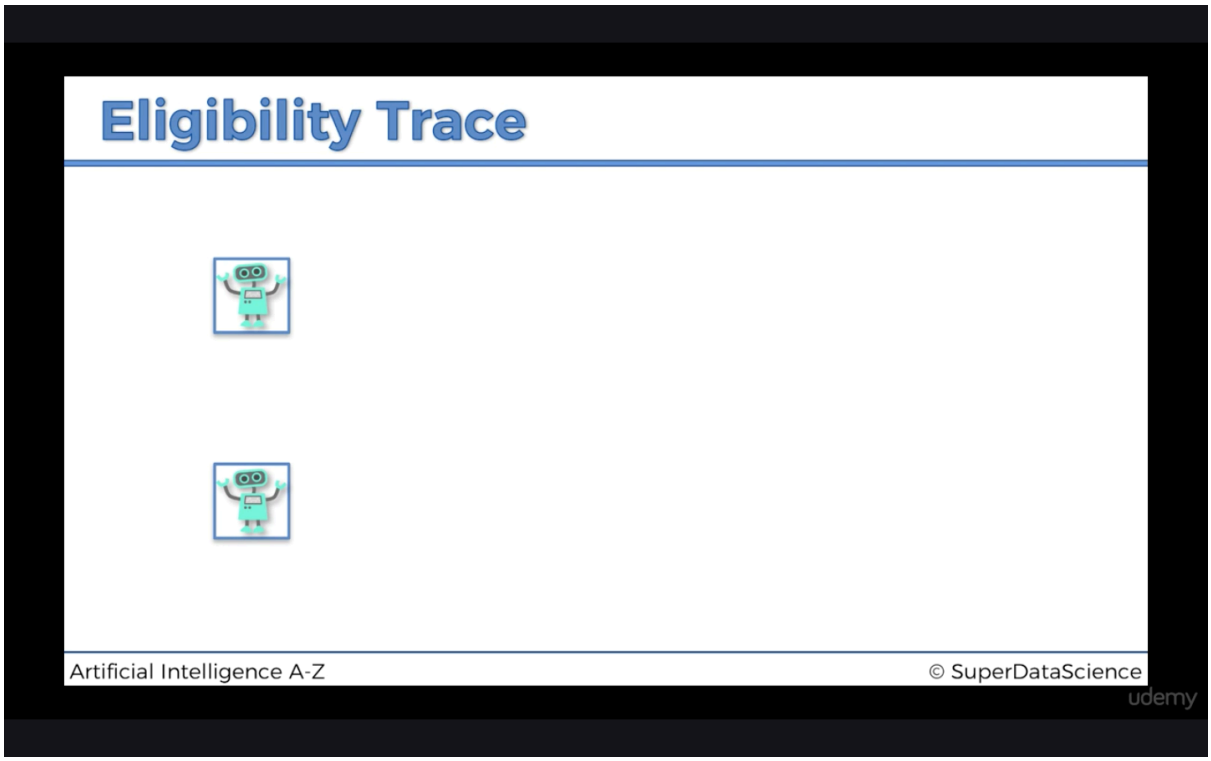
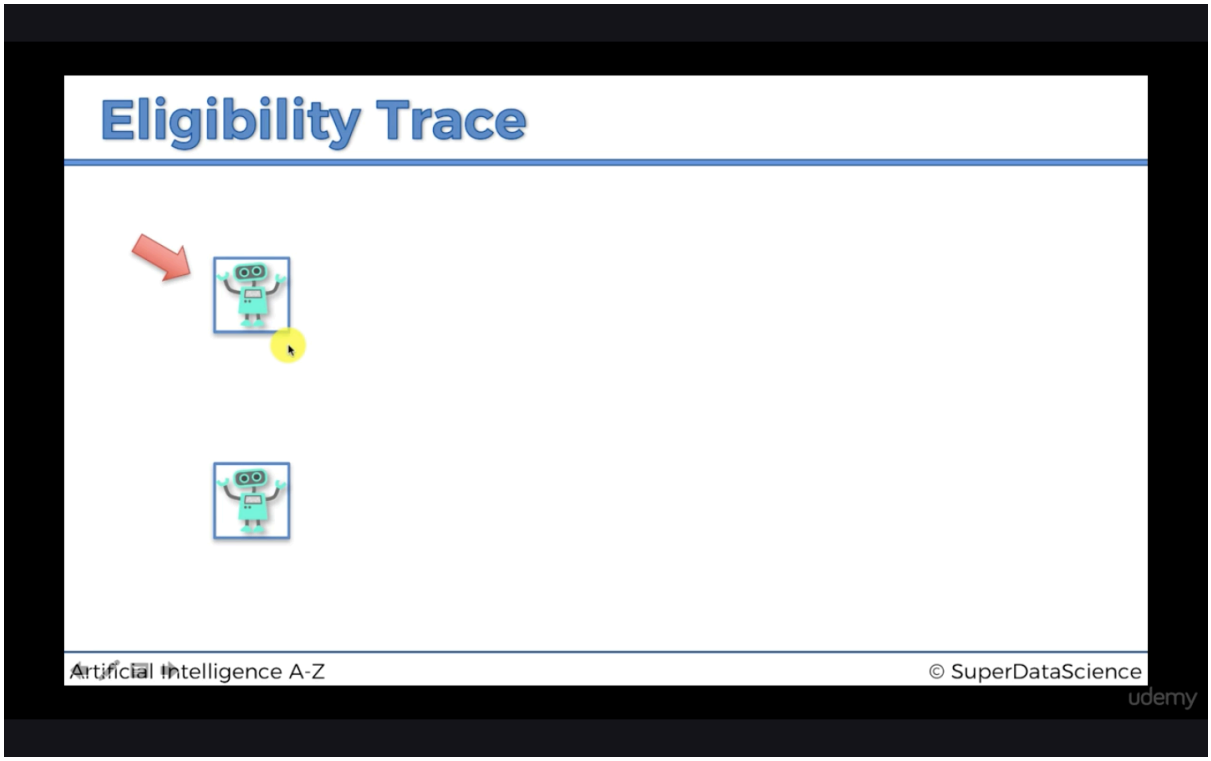
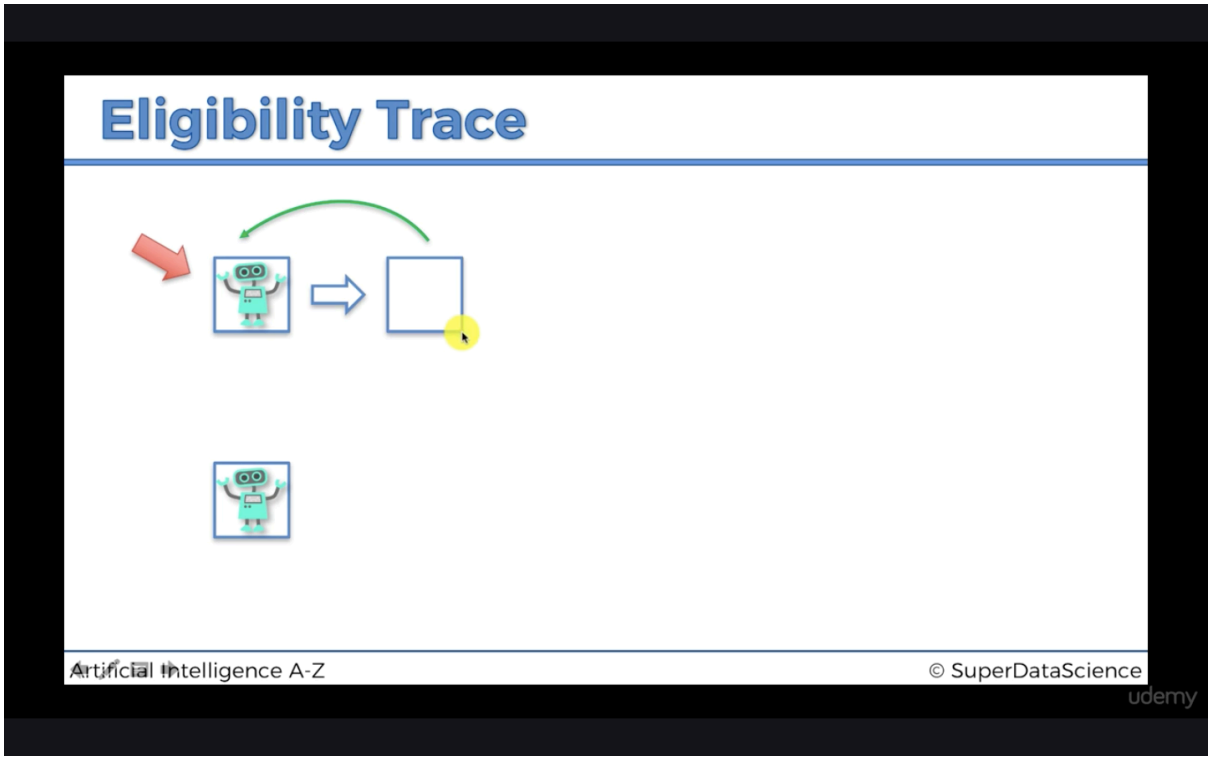
Eligibility Trace (n-step Q-Learning)





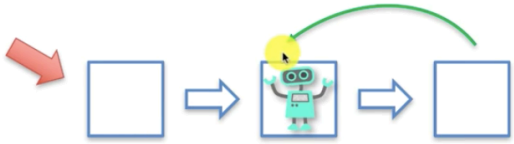
These two agents navigating the same environment. The first one is not going to work with eligibility trace but the second one will.



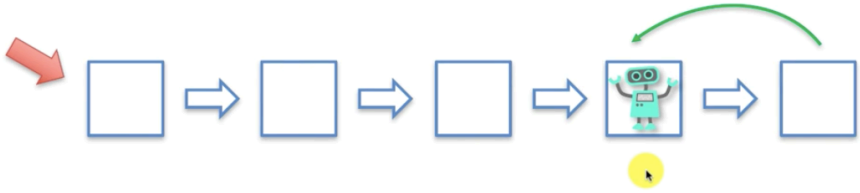


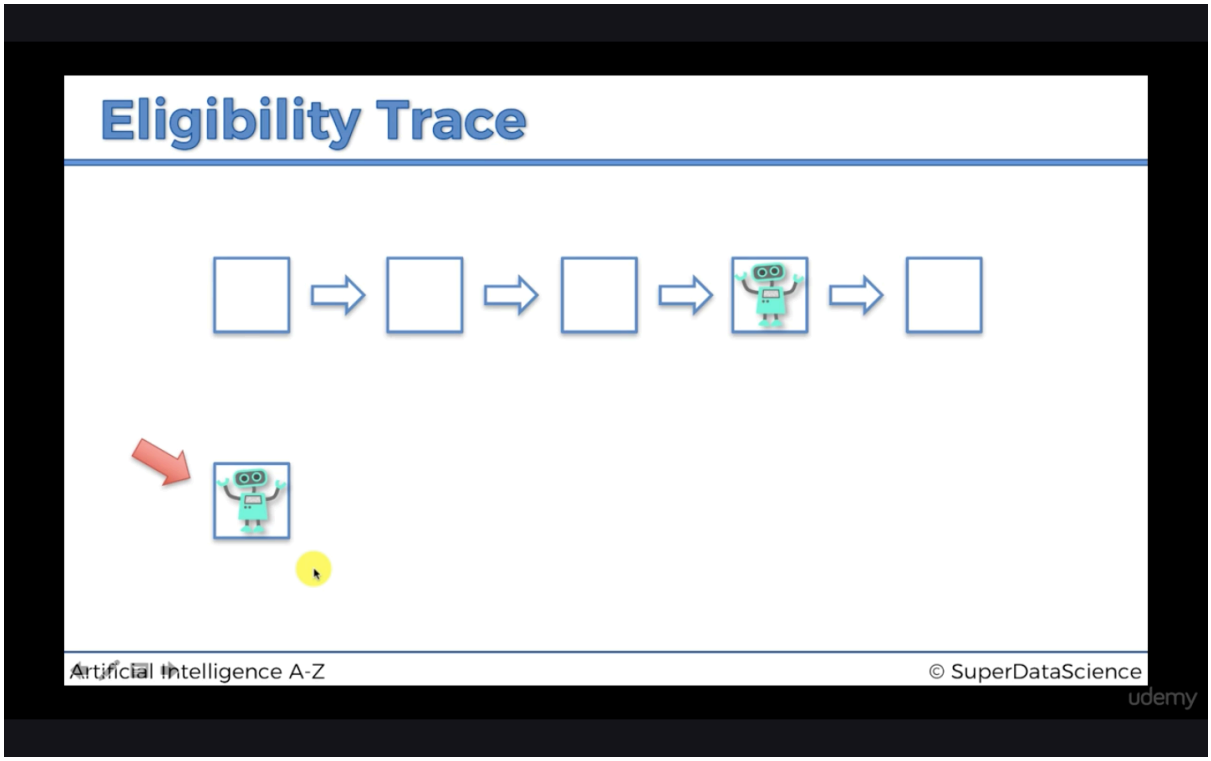
When the agent take action and goes to a new state, it gets a reward and put that reward through algorithms and then updates the neural network.

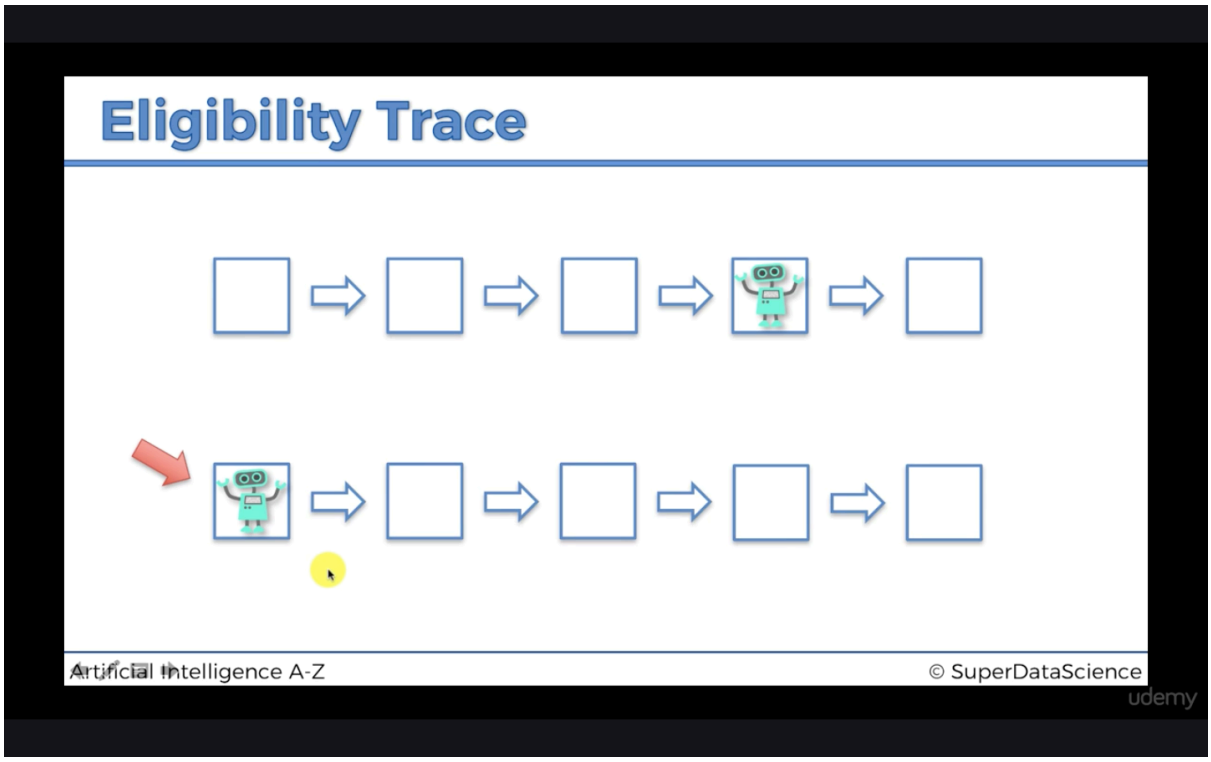
Eligibility Trace



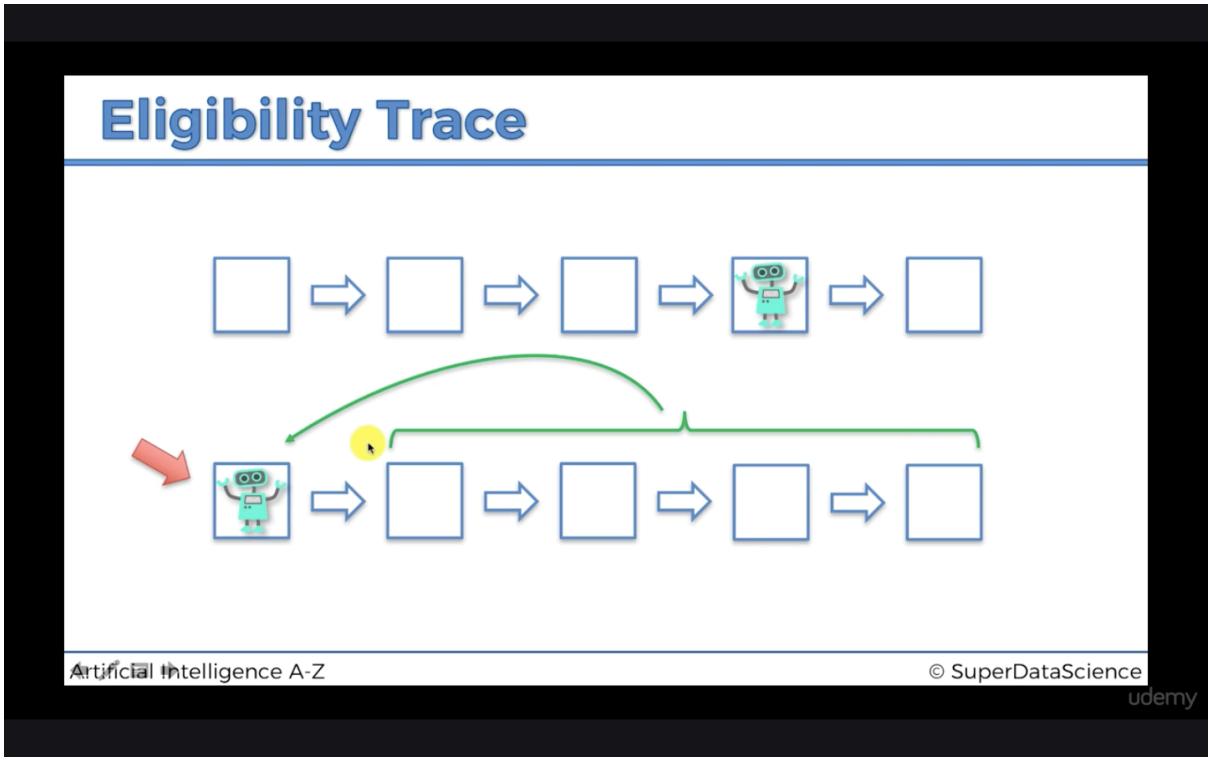
Eligibility Trace





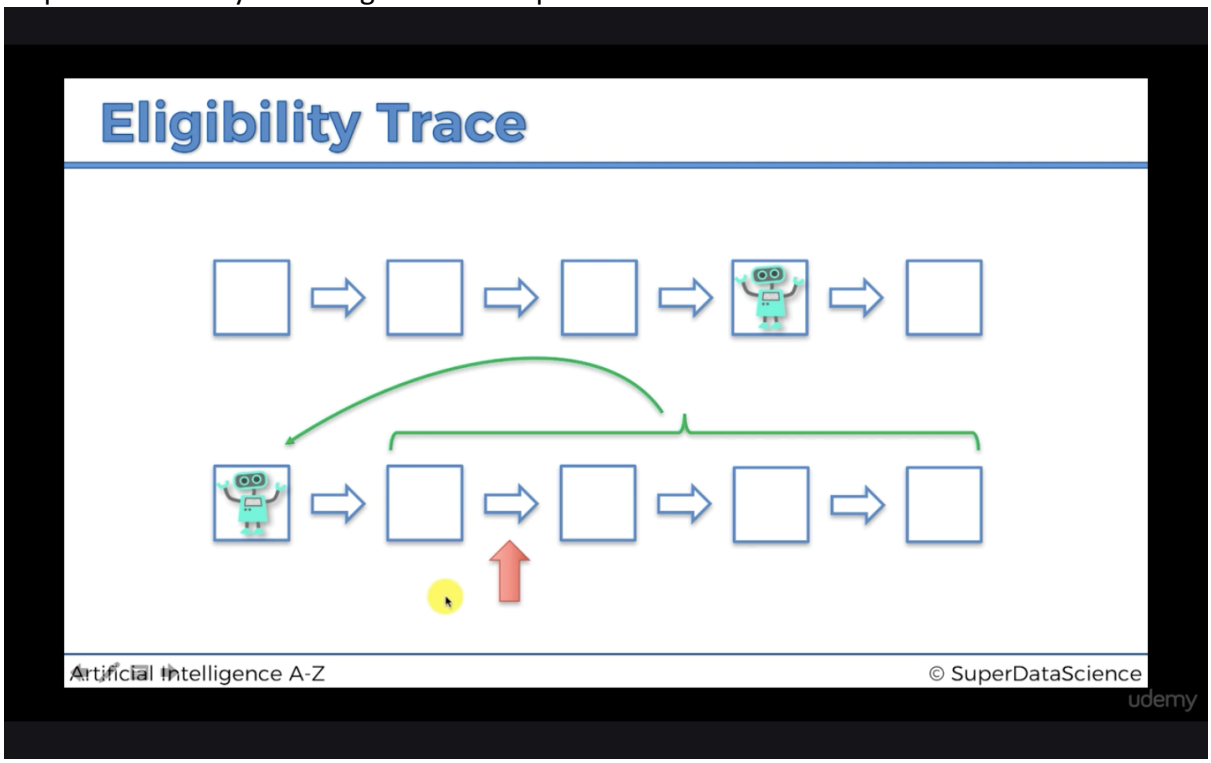


First it's going to take n steps which in here is 4 steps.



Afterward the agent calculates the reward of those steps it took and the agent put it into its network to learn from that. The reason why the second agent is more powerful is because it knows what's at the end and the first agent only acts only through the rewards that the each specific environment gives it.

In here not only we know the overall pattern which is the output of 4 steps but also we keep the trace of eligibility for example if we get a negative reward then we know which of these steps is most likely to be eligible for that punishment.



For instance, if this was the negative reward then it will indicate that this step is responsible for that. It might be positive two. It works both way.

Additional Reading

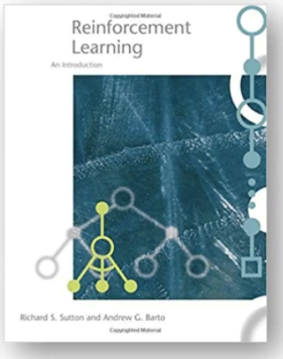
Additional Reading:

Reinforcement Learning: An Introduction

Richard S. Sutton and
Andrew G. Barto (1998)

Link:

https://mitpress.mit.edu/books/reinforcement-learning



Eligibility trace is chapter 7.

Additional Reading

Additional Reading:

Asynchronous Methods for Deep Reinforcement Learning

Volodymyr Mnih et al. (2016)

Link:

https://arxiv.org/pdf/1602.01783.pdf