FPT Education
FPT UNIVERSITY

Estimating Action

Objectives

- Understand action value and how to estimate the action value
- Learn some methods to estimate action value

Value of an Action

- The value of an action is the expected reward when that action is taken

$$q_*(a) \doteq \mathbb{E}[R_t \mid A_t = a]$$

- The action value q is not know → we estimate it !
- Imaging that you have k action, how do you calculate the action value?

Value of an Action

- Sample- Average method

$$Q_t(a) \doteq \frac{\text{sum of rewards when } a \text{ taken prior to } t}{\text{number of times } a \text{ taken prior to } t}$$

$$= \frac{\sum_{i=1}^{t-1} R_i}{t-1}$$

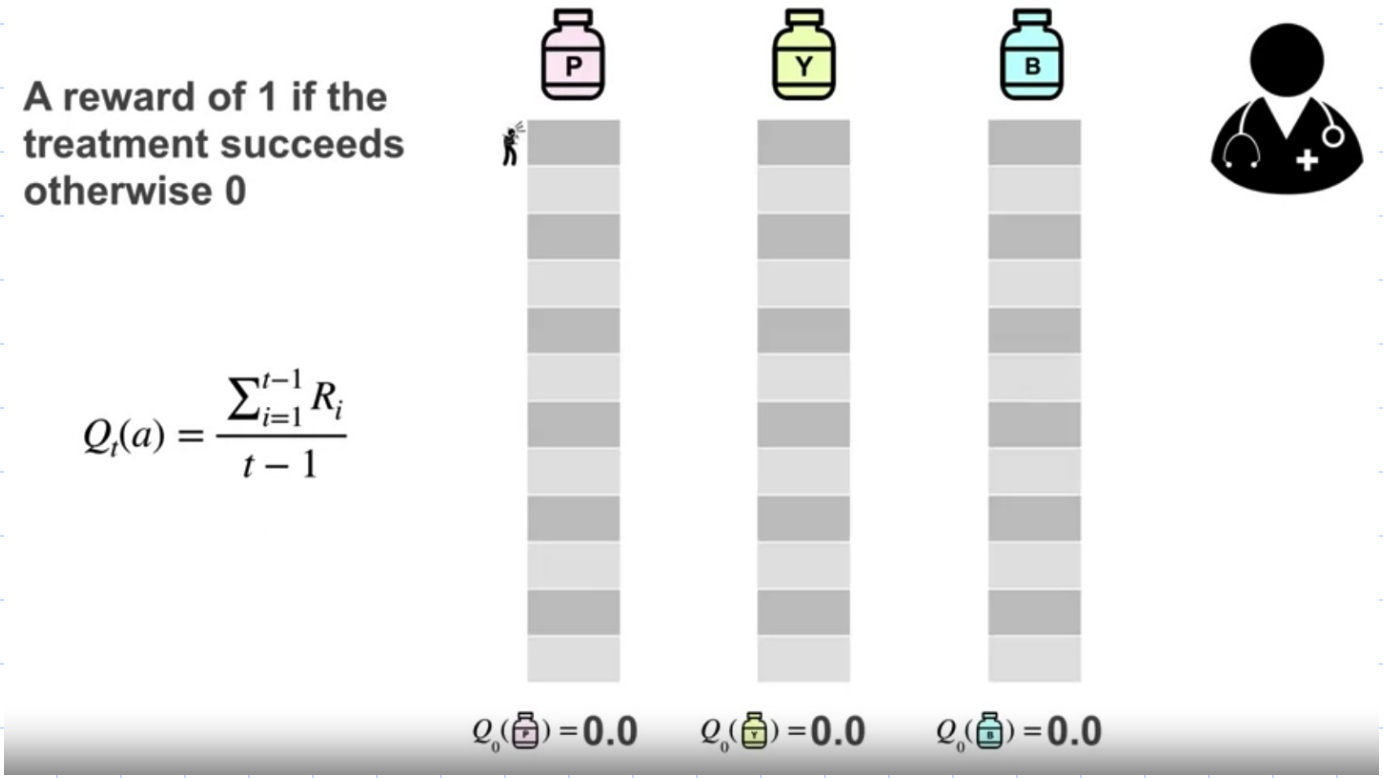
- The total reward for each action and divide it by the number of times that action has been selected.

Value of an Action

- Example- Clinical trials: A doctor must decide which of the three possible treatments to prescribe. If the patient gets better, the doctor records a reward of one. Otherwise, the doctor record

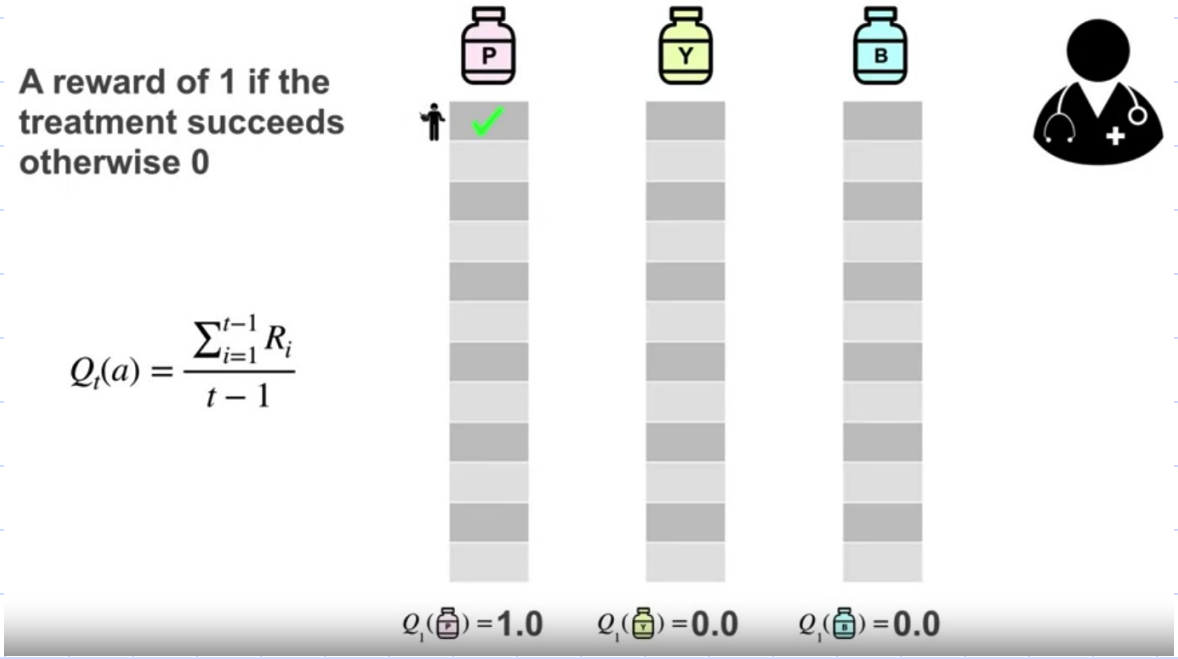
Value of an Action

- First action value for all treatment is zero



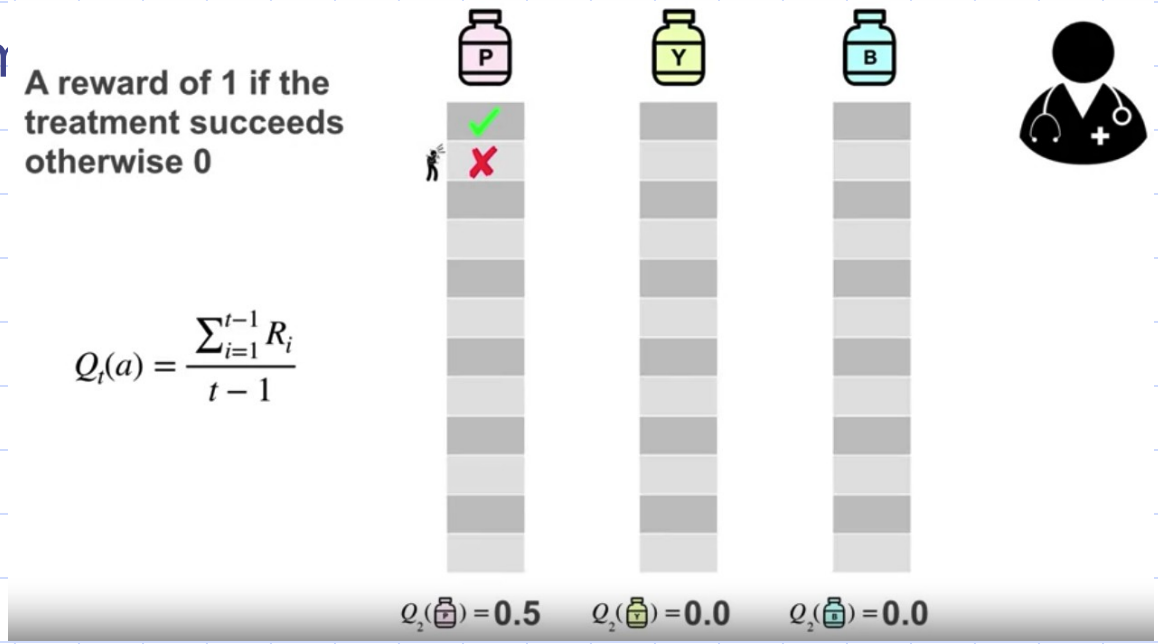
Value of an Action

- The patient reports feeling better. The doctor records a reward of one for that treatment and updates the estimate of the value.



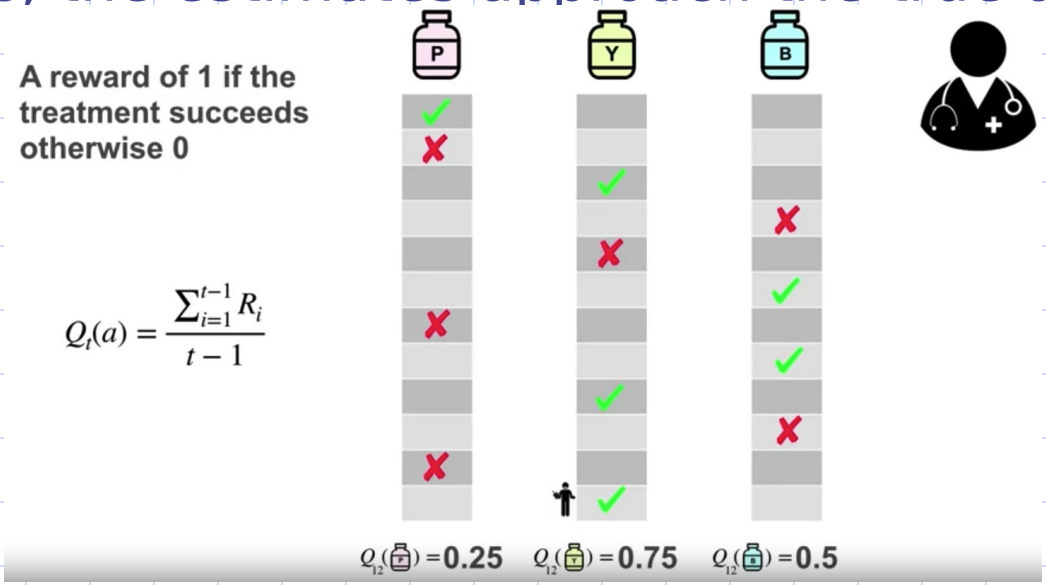
Value of an Action

- Second patient: The doctor randomly prescribes treatment P again. It fails, the doctor records and rewards zero, and updates the value estimate for treatn



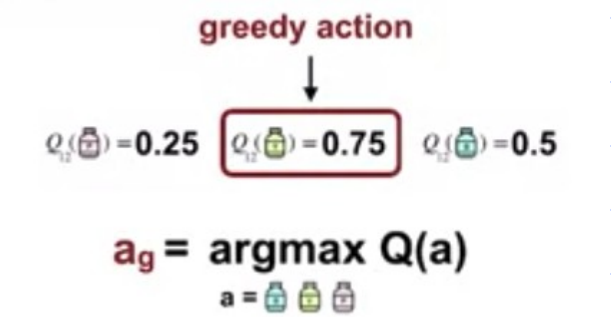
Value of an Action

- After each treatment has been tried a few times, we can calculate the estimated values from the observed data. As the doctor observes more patients, the estimates approach the true action values.



Value of an Action

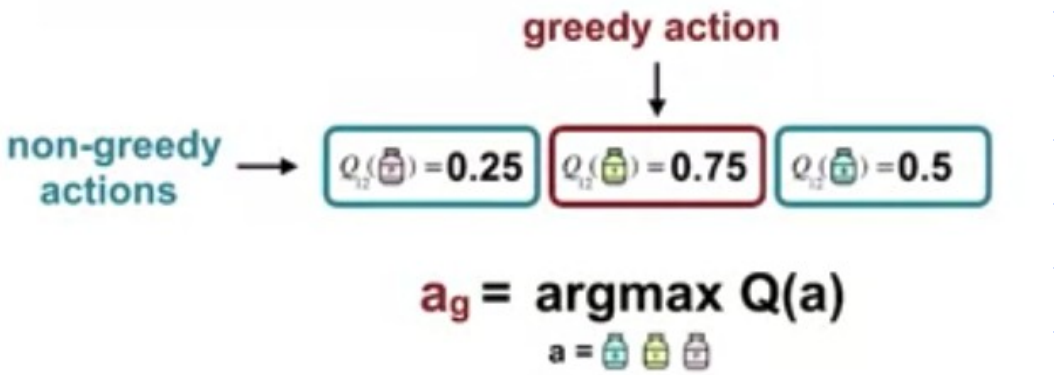
- Greedy method:



- The greedy action is the action that currently has the largest estimated value. Selecting the greedy action means the agent is exploiting its current knowledge.

Value of an Action

- Alternatively, the agent may choose to explore by choosing a non-greedy action. The agent would sacrifice immediate reward hoping to gain more information about the other actions.



Summary

- Understand action value and how to estimate the action value
- Learn some methods to estimate action value

Q & A