# Off-Policy vs On-Policy

- **Off-Policy:** using a different policy for acting and training.

For instance we can use $\epsilon$-greedy policy for training and a greedy policy (that always selects best action) for acting.

- **On-Policy:** Using the same policy for acting and training

Glossary ——//———//———//———//—

## Epsilon-greedy strategy

- Common exploration strategy used reinforcement learning that involves balancing exploration and exploitation.

- Choses the action with the highest expected reward with a probability of $1 - epsilon$

- Chooses a random action with a probability of epsilon

- Epsilon is typically decreased over time to shift focus towards exploitty

## Greedy Strategy

- Always chooses exploitation.

- Does not include exploration

- Can be disadvangeous in environments with uncertainty.

——//——— //——— //——— //—