Reinforcement Learning: Entry survey

- In a partially-observable scenario, can reinforcement be used to learn MDP rewards?
  Answer: Yes, as you explore previously observed locations, you can calculate MDP rewards for the already known part of the environment.

- How can we improve MDP by using the plan-execute cycle?
  Answer: As in the previous answer, after each exploration (execute) you can recalculate MDP (plan).

Exit survey

- What's the difference between MDPs and Reinforcement Learning?
  Answer: RL works in partially observable situations where goal and penalty locations are unknown.

- What is the dilemma between exploration and exploitation?
  Answer: In RL, unknown locations must be explored well. For instance it would be a pity if some part of the environment that contained a large reward was not explored so it wasn't included in the search space. On the other hand, you don't want to waste too much time and energy exploring the environment without collecting the rewards.