MCTS

测验, 4 个问题

1 point

1。

What is true about planning in RL?

- [ ] Planning does not make use of Dynamic Programming
- [x] Planning is computationaly intensive.
- [x] Planning allows to *compute* (contrast with learning) the best possible action.
- [ ] For planning we do not need to explore – we are already given all the knowledge we would need for learning optimal decision making.

---

1 point

2。

What are the differences between model-free and model-based settings?

- [x] In a model-based setting we can find out which (reward, next state) pairs are possible given current state and action.
- [x] In a model-free setting we know nothing about environment dynamics. Optimisation of agent decisions is based solely on sample based experiences of the world.
- [ ] In a model-based setting we know nothing about environment dynamics. Agent's learning is formulated as optimisation of some parametric model.
- [ ] In a model-free setting an agent's learning is formulated in way that is not related to any of parametric models.
- [ ] In a model-free setting we know which (reward, next state) pairs are possible given current state and action.

MCTS

测验, 4 个问题

1 point

3。

What are the different types of planning?

- [x] Decision time planning - planning starts after the agent falls into new state, this planning is performed to select the optimal decision only in current state.
- [x] Background planning - the approach of learning from environment model samples with any model-free method.
- [ ] Decision time planning - the approach of learning correct action decisions from environment model samples with any model-free method.
- [ ] Background planning - planning starts in the background after the agent falls into new state, this planning is performed to select the optimal decision only in current state.

---

1 point

4。

What are the ideas behind the strong planning algorithm?

- [ ] Use binary search to select the best action.
- [x] Stop expanding the search tree as soon as the prespecified depth is reached.
- [ ] Save immediate rewards following each particular action and each particular state.
- [x] Use heuristics to guide the search tree growth.
- [ ] Replace MDP with its deterministic version with replacing all the transitions with only the most probable ones.
- [x] Approximate the returns from the leaves of a search tree till the end of episode.

- [x] 我（**伟臣 沈**）了解提交不是我自己完成的作业 将永远不会通过此课程或导致我的 Coursera 帐号被关闭。

了解荣誉准则的更多信息

Submit Quiz

MCTS

测验, 4 个问题