

准备好后再次尝试

获得的成绩 75% 最新提交作业的评分 75% 通过条件 80% 或更高

重试

您正在查看此评估的自动翻译版本
您可以根据需要切换回去，以源语言查看这些内容。即使您更改语言，也不会丢失任何进度。[显示英语（English）版本](#)

取消

1.

1 / 1个点

您正在使用强化学习来控制一个四脚机器人。机器人的位置就是它的 _____。

- ☐ 返回
- ☒ 国
- ☐ 行动
- ☐ 报酬

正确
好极了

2.

1 / 1个点

您正在控制一辆火星车。如果它进入状态 1（重大科学发现），你会非常非常高兴；如果它进入状态 2（小的科学发现），你会稍微高兴；如果它进入状态 3（漫游车永久损坏），你会不高兴。为了反映这一点，请选择一个奖励函数，以便

- ☐ $R(1) > R(2) > R(3)$ ，其中 $R(1)$ 、 $R(2)$ 和 $R(3)$ 均为负值。
- ☐ $R(1) < R(2) < R(3)$ ，其中 $R(1)$ 和 $R(2)$ 为负值， $R(3)$ 为正值。
- ☐ $R(1) > R(2) > R(3)$ ，其中 $R(1)$ 、 $R(2)$ 和 $R(3)$ 均为正数。
- ☒ $R(1) > R(2) > R(3)$ ，其中 $R(1)$ 和 $R(2)$ 为正值， $R(3)$ 为负值。

正确
干得好

3.

0 / 1个点

您正在使用强化学习来驾驶直升机。使用 0.75 的贴现因子，您的直升机从某种状态开始，第一步获得奖励-100，第二步获得奖励-100，第三步也就是最后一步获得奖励 1000（在这一步它达到了终点状态）。回报是多少？

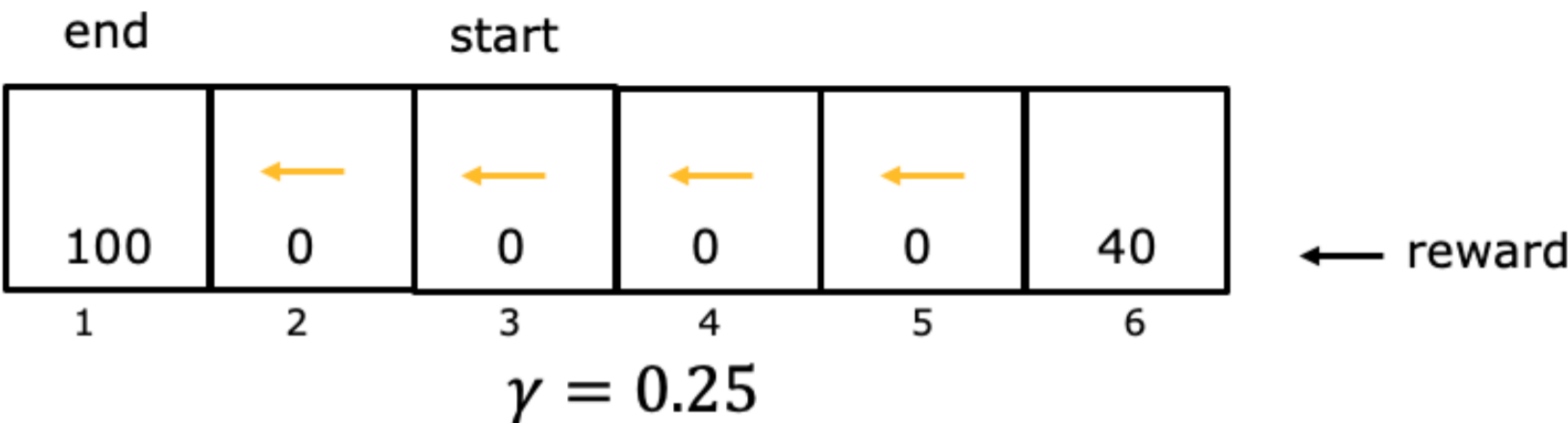
- ☐ $-100 - 0.25 \times 100 + 0.25^2 \times 1000$
- ☐ $-100 - 0.75 \times 100 + 0.75^2 \times 1000$
- ☒ $-0.75 \times 100 - 0.75^2 \times 100 + 0.75^3 \times 1000$
- ☐ $-0.25 \times 100 - 0.25^2 \times 100 + 0.25^3 \times 1000$

错误
记住，第一次奖励是不打折扣的。

4.

1 / 1个点

给定以下奖励和行动，计算状态 3 的收益，贴现率为 $\gamma = 0.25$ 。



- ☐ 0
- ☐ 25
- ☒ 6.25
- ☐ 0.39

正确
如果从状态 3 开始，奖励在状态 3、2 和 1 中，回报为 $0 + (0.25) \times 0 + (0.25)^2 \times 100 = 6.25$ 。