作业题目

1: GRIDWORLD

An agent moving in the 4 x 4 grid world.

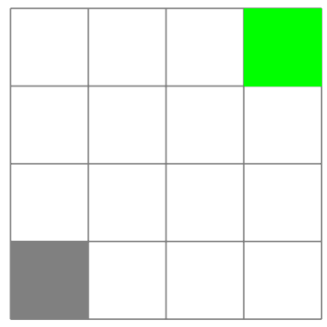
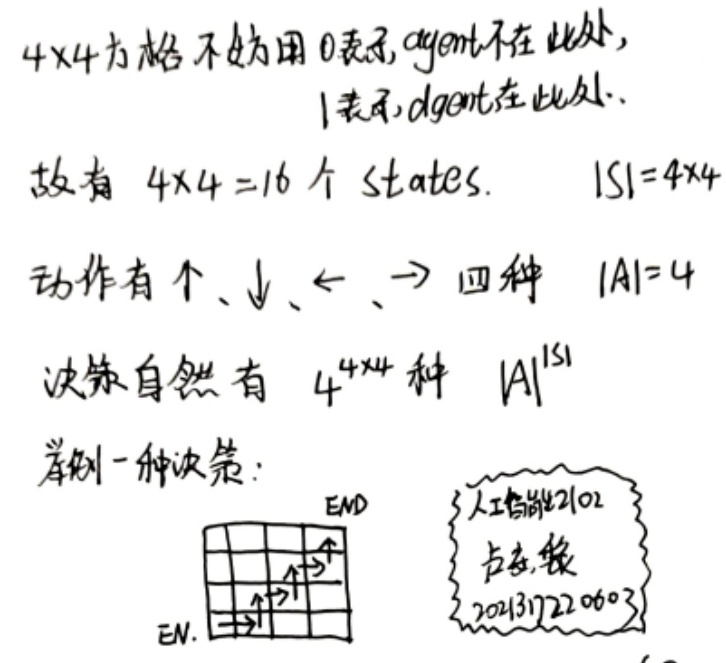


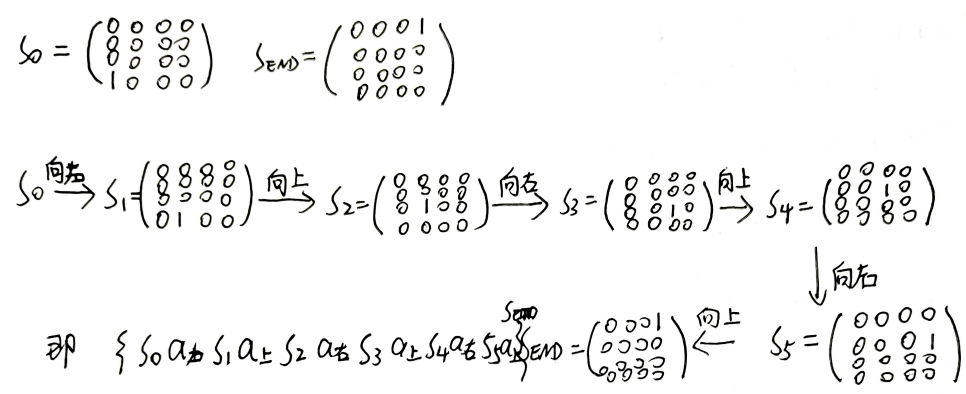
图. Possible actions in each cell: ↑, ↓, ←, →. Cell with gray entrance and green means ending.

- The number of States?
- The number of Actions?
- The number of policies?
- Please give one policy?
- Please encode each state.

作答：



决策举例子并编码



作业

在强化学习中，观测和状态的区别

- 观测是智能体感知到的环境信息，而状态是环境的一个特定描述，它包含了智能体需要的所有信息
- 观测可能是不完整或含有噪声的，而状态通常是完整的且精确的。
- 由于状态可能不完全可见或不可观测，智能体可能需要通过观测来近似或推断状态。

强化学习和有监督学习的差别

① 强化学习是基于试错学习的方式，智能体通过与环境的交互来学习适合的策略，不需要显式的标记好的数据集。

有监督学习：有监督学习是基于标记好的数据集进行学习的方式，算法通过输入和对应的输出来学习模型，需要大量的带标签数据作为输入。

② 在强化学习中，智能体的目标是通过与环境的交互学习到一个最优策略，使得长期累积的奖励最大化

在有监督学习中，算法的目标是学习到一个能够准确预测或分类输入数据的模型，使得模型的预测结果与真实标签尽可能一致。

③ 强化学习通常用于序贯决策问题，如游戏、机器人控制、自动驾驶等。

有监督学习广泛应用于分类、回归等静态问题，如图像分类、语音识别、自然语言处理等领域。