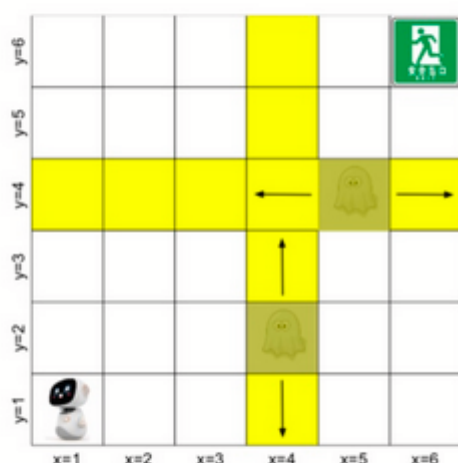
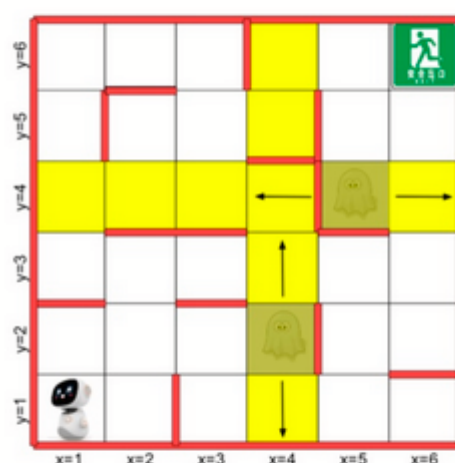


人工智能技术基础课后作业（一）——智闯鬼屋

东北大学信息学院智能感知与机器人研究所计划搭建一个智能机器人仿真测试平台。该平台模拟一个鬼屋游戏，鬼屋中有两个幽灵在游荡。智能机器人从左下角房间($x = 1, y = 1$)出发，以到达右上角的绿色出口($x = 6, y = 6$)为目标，如在移动中遇到幽灵，则视为失败。



问题 A



问题 B 和问题 C

幽灵行为模式：幽灵沿着各自所在的黄色通道每轮移动一格，初始位置和初始方向随机设定，幽灵可以穿过鬼屋内部墙壁，但遇到外墙后会反向前进。

机器人运动模式：机器人每轮可以向当前房间的四邻域房间移动一格，也可以选择停留在当前房间，但不能穿墙。

机器人感知模式：机器人能够判断是否可以从当前房间移动到某一个邻域房间，机器人不知道幽灵的具体位置，但在一定范围内能够感知到幽灵带来的

寒意，两个幽灵带来的寒意可以叠加。单个幽灵带来的寒意浓度与机器人和幽灵之间的**棋盘距离**¹成反比，具体为：

棋盘距离	1	2	≥ 3
寒意浓度	2	1	0

问题 A：假设鬼屋内部没有墙壁，如左图所示，请为该机器人设计一套智能行动规则，并给出机器人成功到达出口过程的路线。

问题 B：假设鬼屋内存在墙壁，如右图所示，请为该机器人设计一套智能行动规则，尽可能快速地到达目的地，画出两条以上的成功路径，评估到达目的地的成功率和平均时间。

问题 C：假设机器人预先知道鬼屋内所有墙壁的位置以及幽灵的初始位置与初始方向，请设计一个搜索算法找到最快路径，并展示结果。

答题要求：在回答上述问题时，请明确写出机器人的行动规则，根据行动规则移动的推理过程、幽灵的初始状态以及问题要求的其他回答的计算过程。

提交形式：将回答的文字、图表以 pdf 文档形式提交，如有代码或演示视频等附件请与 pdf 文档放入同一个压缩包。压缩包要求.rar 或.zip 格式，不得有密码，压缩包文件名格式为“学号-姓名”，例如“2020011-张三.rar”

¹ **棋盘距离**指目标只能横向或纵向移动时，两个格子间的最短移动距离，格子 (x_1, y_1) 与格子 (x_2, y_2) 的棋盘距离为： $|x_1 - x_2| + |y_1 - y_2|$

一.前提

AI的任务是设计Agent程序，实现的是把感知信息映射到行动的Agent函数。

二.任务环境

2.1 PEAS描述

- Performance性能：到达出口的成功率，到达出口的行动步数
- Environment环境：6X6方格内，存在两个随机行动的幽灵，存在围墙
- Actuators执行器：在自身四邻域内移动，也可以不行动，不能穿墙
- Sensors传感器：通过曼哈顿距离感知幽灵，感知可以叠加

2.2 性质

2.2.1 完全可观察的vs部分可观察的

如果Agent的传感器在每个时间节点上都能获取环境的完整状态，这个任务环境就是完全可观察的。否则，则是部分可观察的。

如果传感器能够检测所有与行动决策相关的信息，该任务环境就是有效完全可观察的。
如果Agent根本没有传感器，环境则是无法观察的。

2.2.2 单Agent vs 多Agent

区分两者的关键在于Agent B 行为的性能度量最大化是否需要依赖于Agent A的行为。

竞争性的多Agent环境

Agent B想要最大化自己的性能度量，就需要最小化Agent A的性能度量。

合作性的多Agent环境

Agent B想要最大化自己的性能度量，就需要最大化Agent A的性能度量。

部分合作部分竞争的多Agent环境

上述两种情况都会发生。

2.2.3 确定的 vs 随机的

如果环境的下一个状态完全取决于当前状态和Agent执行的动作，则该环境是确定的；否则，是随机的。

不确定与随机的区别

环境不确定是指 环境不是完全可观察的或不确定的，行动后果可能有多种，但与概率无关。

环境随机是指后果是不确定的并且可以用概率来量化。

2.2.4 片段式 vs 延续式

片段式是指 当前决策不会影响到未来的决策。

延续式是指 当前决策会影响到所有未来的决策。

2.2.5 静态 vs 动态

如果环境在Agent计算的时候会变化，该环境是动态的，否则是静态的。

如果环境本身不随时间变化而变化，但Agent的性能评价随时间变化，则环境是半动态的。

2.2.6 离线 vs 连续

离线和连续的使用场景有：

环境的状态

时间的处理方式

Agent的感知信息和行动

2.2.7 已知 vs 未知

Agent的知识状态。

如果环境是未知的，Agent需要学习环境是如何工作的，以便做出好的决策。

问题A：部分可观察，环境随机，延续式，动态，离散的，已知环境

问题B：部分可观察，环境随机，延续式，动态，离散的，已知环境

问题C：完全可观察，环境确定，延续式，动态，离散的，已知环境

3 Agent的结构

3.1 基于模型的反射Agent

定义

世界模型：

知识一：世界是如何独立于Agent而发展的信息

知识二：Agent自身的行动如何影响世界

使用世界模型的Agent称为基于模型的Agent

通过将当前的感知信息与过去的内部状态结合起来更新当前状态

针对部分可观测环境

缺点

部分可观察环境中的Agent不能精确定当前状态

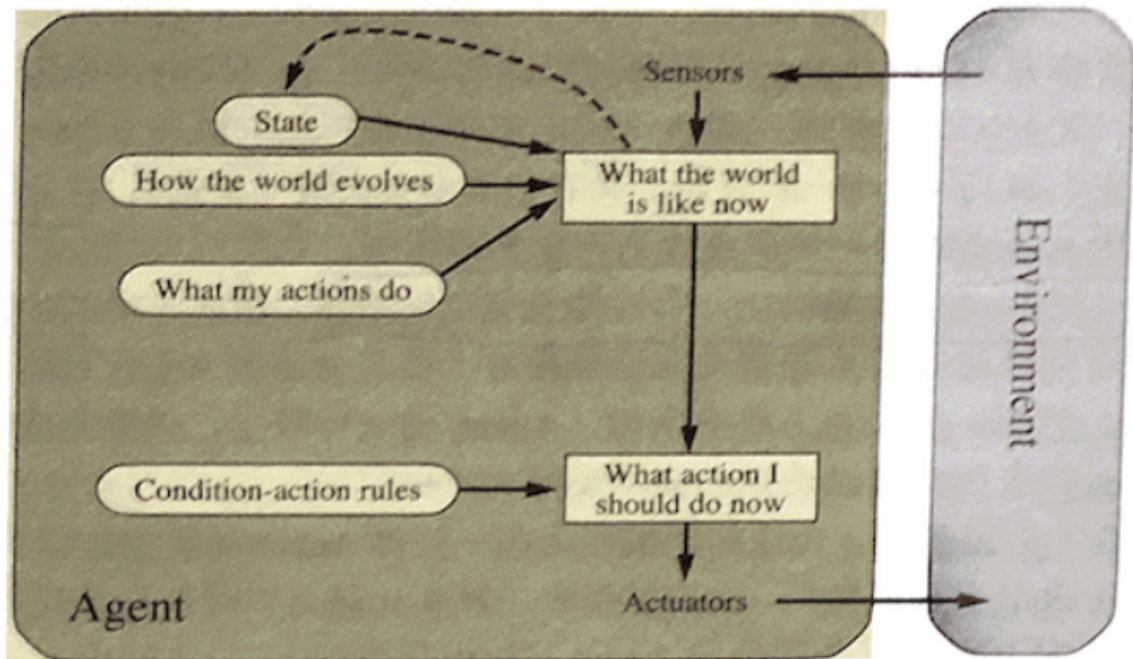


图 2.11 基于模型的反射 Agent

```

function MODEL-BASED-REFLEX-AGENT(percept) returns an action
    persistent: state, the agent's current conception of the world state
                 model, a description of how the next state depends on current state and action
                 rules, a set of condition-action rules
                 action, the most recent action, initially none

    state ← UPDATE-STATE(state, action, percept, model)
    rule ← RULE-MATCH(state, rules)
    action ← rule.ACTION
    return action
  
```

图 2.12 基于模型的反射 Agent。它使用内部模型跟踪世界的当前状态。它按照与反射 Agent 相同的方式选择行动。 csdn.net/greenqingqings

3.2 基于效用的Agent

定义

理性的基于效用的Agent 选择 期望效用最大化的行动，Agent在给定每个结果的概率和效用下，期望得到的平均效用。

方法

Step 1: 使用关于世界的模型，以及对各个世界状态的偏好程度进行度量的效用函数。

Step 2: 选择可以取得最佳期望效用的行动。

Step 3: 通过结果的概率来确定权值，最佳期望效用是通过计算所有可能结果状态的加权平均值得到的。

适用情况

- 1、当多个目标互相冲突时，只有其中一些目标可以达到时，效用函数可以在它们之间适当的折中。
- 2、当Agent有几个目标，但没有一个有把握达到时，效用函数可以在它们之间适当的折中。

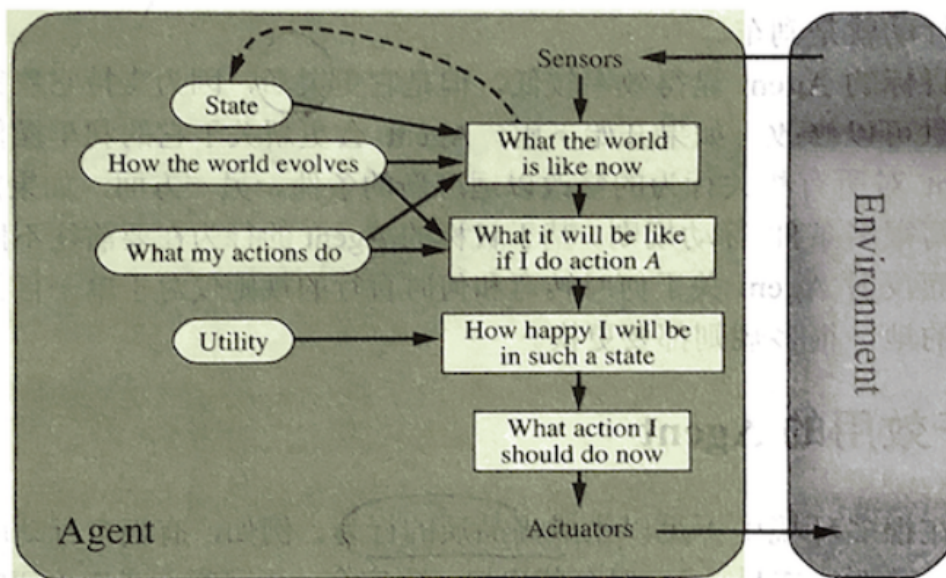


图 2.14 基于模型和效用的 Agent。它使用了关于世界的模型，连同对一个度量对各个世界状态的偏好程度的效用函数。接着它选择导致最佳期望效用的行动，最佳期望效用是通过计算所有可能结果状态的加权平均值得到的，其权值由结果的概率确定。