

1. 写出一个智能代理的例子，可以来自虚构的世界(书、戏剧、电影、电视节目等)，也可以来自现实世界。
写几行文字来描述这个例子，并说:

- (一)使用何种传感器;
- (b)它可以采取什么行动;和
- (c)它在什么环境下运作。

2. 根据前面的问题对环境进行分类。

您应该按照我们在讲座中描述的方式对环境进行分类——可访问的或不可访问的、静态的或动态的，等等。解释为什么你会给出这样的答案。

3. 提出另一个在不同环境中操作的代理示例。对环境也进行分类。

4. 图 1(下一页)显示了位于环境中的代理。我们可以认为环境 E 由 36 个州组成:

$$E_0, E_0, \\ E_{0,1}, \dots$$

其中每个 e 的下标表示网格中的一个正方形。

因此，坐在左下角的施动者处于状态 $e_{0,0}$ ，而如果行动者在目标，在右上角，它将处于状态 $e_{5,5}$ 。 $e_{0,0}$ 是环境的初始状态。

填充的方格表示障碍——代理程序不能处于这些状态。

代理可以向北、南、东或西移动，我们可以这样写:

$$\alpha_n, \alpha_s, \alpha_e, \alpha_w$$

这些都有你所期望的效果。如果代理处于状态 $e_{0,0}$ 采取 α 行动 n ，它将以 $e_{0,1}$ 的状态结束，而如果行动者处于 $e_{3,2}$ 的状态并采取行动 αe ，它将以 $e_{4,2}$ 的状态结束。如果代理尝试移动到网格之外，那么它不会移动(例如，如果代理在 e 中 $_{0,5}$ 然后试着求 α_n 然后保持在 $e_{0,5}$)。

如果座席进入状态 $e_{5,5}$ ，然后它将获得 10 个奖励。如果它进入状态 $e_{1,4}$ ，用一个黑圈标记，得到-10 的奖励(即损失)。

(a)写下从 $e_{0,0}$ 到 $e_{1,4}$ 的 agent 运行。(b)考虑下列控制程序:

```
While(不是状态  $e_{5,5}$ ){
    随机选择一个  $\alpha_n$  或  $\alpha_e$  (每个概率为 0.5)执行所选择的动作
}
```

写下代理在执行此程序时可能执行的两次运行。

(c)如果代理程序执行，它能到达 $e_{5,0}$? 为什么?

(d)代理运行该程序所能获得的最大和最小奖励是什么?为什么?

(e)代理运行上述程序时，获得-10 奖励的可能性有多大?

(提示:有一个精确的概率，代理将到达相关的状态。你应该算出如何计算这个。

5. 假设讲座中的真空世界(即讲座 1 的幻灯片中的真空世界)包含了 agent 必须避免的障碍，并且 agent 有一个传感器来检测这些障碍。也就是说，有些方块因为无法到达而被阻挡(如墙)，而其他方块则有障碍物(如人的脚)。为代理写一个新的控制程序。

解释一下，如果有的话，你可以保证你的解决方案能使房间变得干净，并证明你的答案是正确的。

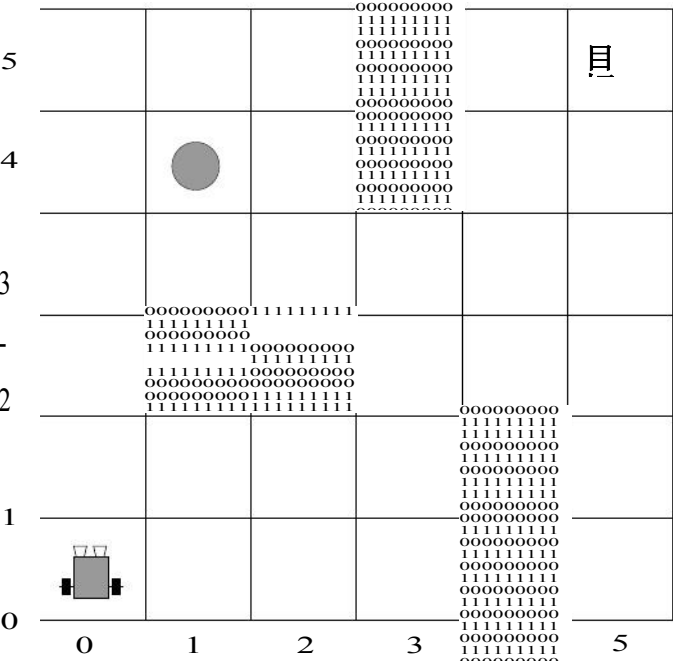


图 1:

6. 假设真空世界代理的传感器——在讲座中看到灰尘的那个，在问题 5 中看到障碍物的那个——现在是嘈杂的，所以它们只能给出 80% 的正确答案。

(你可以理解为，如果有灰尘或障碍物，相关的传感器会在 80% 的时间里说它们在那里。你可以假设当没有灰尘或障碍物时，传感器总是正确地报告这一点。) 这如何改变基于逻辑的控制程序? 假设 **agent** 试图移动到一个有障碍物的正方形中，它不会移动。

解释一下，如果有的话，你可以保证你的解决方案能使房间变得干净，并证明你的答案是正确的。

第 1 课回答 (Version 1)

1.一个例子是踢足球的机器人，一个通讯机器人。

- (a)它有许多不同种类的传感器:麦克风、相机、触觉传感器、声纳(用来探测前面的东西)、wifi，所以可以从....网站获取输入信息
- (b)发出噪音，眼睛变色来表达情感，能握东西的手，能自由移动四肢的马达....
- (c)在物理世界中运作。我们可以特别考虑一下机器人世界杯的环境。

2.机器人幼崽的实际环境是:

- 部分可达:nao 机器人只“看到”环境的一小部分——在它(相当狭窄的)视野内，在相机和麦克风的范围内。它也可能不得不决定从它得到的输入中处理什么——例如，它可能会占用太多的资源来处理它接收到的所有声音输入。
- 不确定性:电机中的错误可能意味着某个特定动作没有成功执行。此外，如果你只是以一定的粒度来感知/建模世界，那么你无法察觉的差异可能会对你的行动结果产生影响。
- 片段式或顺序式——取决于你考虑的粒度级别。如果你将锦标赛视为一个情节，你可能会认为你在锦标赛中的行为不太可能影响它在未来任何锦标赛中的表现。尽管你也可能会说，你的所作所为会影响别人对你的看法，所以这可能意味着他们对你的表现会有所不同。如果你认为游戏是一系列的情节，那么环境就是连续的，因为机器人做出的选择(特别是当你考虑到它如何与团队互动以及它对团队做出的假设时)会影响它在游戏后期的表现。或者你可以把机器人看作是根据经验来学习的东西，或者是根据经验来建立对手和队友的模型，然后是顺序的。
- 动态:有许多其他的代理和过程作用于环境。
- 连续:电机可以通过任何角度转动，并可以以许多不同的方式组合。触摸传感器(原则上)可以记录连续的压力范围。但你可能会认为它是离散的。

重要的是要证明你的答案，并明确你的假设!

3.帮助照顾老人的陪伴机器人。它的环境是物理世界，特别是人的家。

- 不可接近:机器人不知道其他房间里发生了什么，但也不能确定人的心理或健康状况。
- 不确定性:通常他们行为的结果也可能取决于个人。例如，想想交流行为:机器人可能会发出声音，目的是让人高兴起来，但这可能并不总是有预期的效果。
- 顺序:在决定做什么时，需要考虑过去和未来的行动。例如，当决定是否给电池充电时，它应该考虑这是否会阻止未来任何重要的行动。或者，当决定是否提醒某人吃药时，它可能会想它已经提醒过这个人多少次了，它通常会让这个人多么恼火。
- 动态:人也会对环境产生影响，但也会对中央供暖系统、天气等因素产生影响。
- 连续:机器人可以发出任意数量的噪音或动作，可以记录任意数量的输入。但实际上很可能被视为离散的。

4. (一)一种可能性:

$$\frac{e_0}{4} \alpha_n e_0, 1, \alpha_n, e_0, 2, \alpha_n, e_0, 3, \alpha_n, e_1, 3, \alpha_n, e_1,$$

另一种可能性:

$$\frac{e_0}{4} \alpha_n e_0, 1, \alpha_n, e_0, 2, \alpha_n, e_0, 2, \alpha_n, e_0, 2, \alpha_n, e_0, 3, \alpha_n, e_1, 3, \alpha_n, e_1,$$

6.当传感器的报告准确率只有 80%时，真空吸尘器就无法“看到”那里的灰尘，这意味着它可能会使房间变脏。重复检查房间(就像上面的程序一样)意味着，程序运行的时间越长，看不到灰尘也就越不吸灰尘的概率趋于零，但对于任何有限的运行都是非零的。你可以考虑改变程序，让代理总是做“Suck”，而不用费心去检查它是否认为自己是肮脏的。

如果真空没有“看到”障碍物，并试图在它不应该移动的时候移动它，那么障碍就不是什么大问题，如果真空不移动，所有的一切都会发生(我们忽略了在撞上障碍物时造成伤害的可能性)。真空甚至可以计算出它无法移动的障碍。