

## 多智能体系统 2022 Assignment (一)

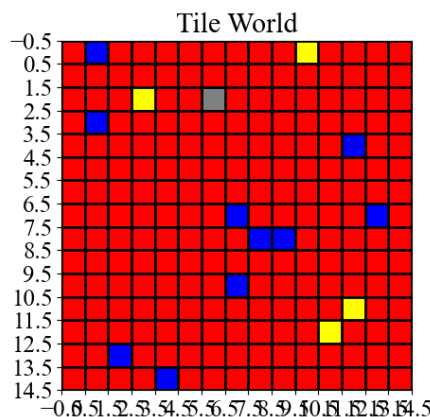
授课老师：章宗长

助教：王铖鹤、张福翔

姓名：思意

2022 年 4 月 9 日





**Figure 1.1** 运行过程模拟。图中蓝色节点代表障碍物，红色节点代表普通的格子，黄色节点代表洞穴，灰色节点代表 Agent。

## 第 1 章 作业说明

本次作业要求我们复现 [Commitment and Effectiveness of Situated Agents \[1\]](#) 论文中的实验与结果。具体的：论文中的实验环境为有固定障碍物的方格瓦片世界，环境中会随机的生成和湮灭洞穴（食物）。我们需要设计实用推理 (BDI) Agent 来尽可能地吃到更多的洞穴（食物）。实用推理 Agent 主要包括两个步骤：

- (1) 慎思过程 (deliberation)。Agent 通过感知外界信息调整信念状态 (Belief)，再通过信念状态确定自身意图 (Intention)。
- (2) 目标手段推理 (means-ends reasoning)。Agent 根据意图给出合理的规划 (Planning)。

在本实验中，我们探究 BDI-Agent 在方格瓦片世界中的表现。具体的：在每个时钟周期，Agent 感知环境状态，并且确定要到达的洞穴（意图）。本例中，我们使用 BFS 确定所有洞穴与 Agent 位置之间的最短距离，并且综合衡量每个洞穴的距离、得分、存在时间的因素，得到最优的洞穴意图。确定意图之后，根据 BFS 求取得到的路径信息，得到实现目标的规划（最短路径），并进行执行。在动作执行过程中，为了避免 Agent 盲目通往一个目标而忽略动态环境中目标洞穴的湮灭或者更好的洞穴的出现，我们需要 Agent 能够及时的重新思考自身的意图，否则就会导致 Agent 对目标的盲目承诺。在本实验中，我们主要探究得分率与 Agent 的规划时间、承诺属性（包括专一承诺 (Single Minded Commitment) Agent、坦率承诺 (Open Minded Commitment) Agent、以及使用自定义 Reconsider 的诸多策略的 Agent）的关系。

本次作业的主要亮点包括：

- (1) 依照论文原文完全复现了论文中的诸多细节设置，复现了文中提及的所有对比结果，与论文中保持一致。
- (2) 使用 Python 语言实现了简略易懂（核心类代码小于 100 行）、高度的参数可定制化（几乎所有环境参数都可以自定义）的实验环境，有利于实验设计与分析。
- (3) 仅依赖 Matplotlib 实现了模拟过程的可视化，如图 1.1。方便研究者分析 Agent 的具体运行过程。

## 第 2 章 实验对比

### 2.1 环境变化率的影响

我们首先复现文中 Agent 的得分随着环境变化率的影响。如图2.1所示。

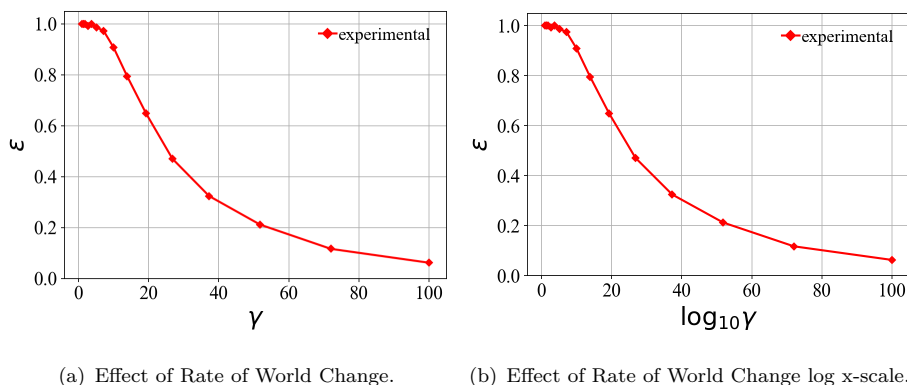


Figure 2.1 世界变化率的影响

### 2.2 规划时间的影响

Agent 的得分率随着规划时间的影响如图2.2所示。观察到随着规划时间的增加，得分率逐渐降低。<sup>1</sup>

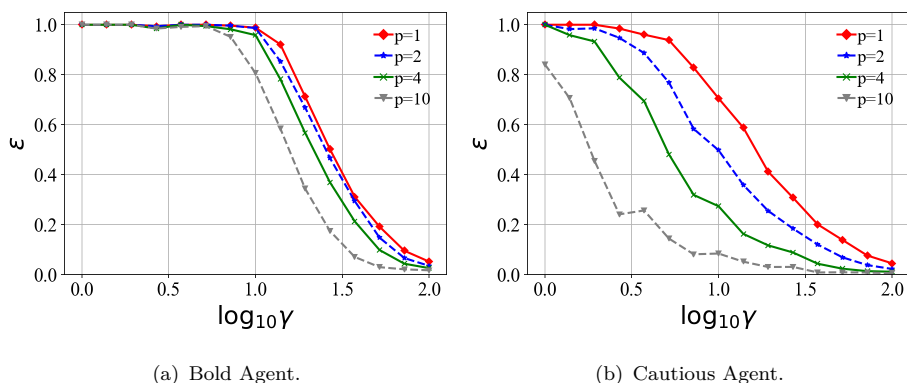


Figure 2.2 规划时间的影响

### 2.3 承诺属性的影响

我们复现论文中承诺属性的影响，具体的，Bold 的 Agent 除非到达意图否则不会重新慎思新的意图，Cautious 的 Agent 每次行动之后都会重新思考意图，Normal 的 Agent 每行动四次会重新思考一次意图。如图所示，在大多数情况下，Bold 的 Agent 能够取得更好地性能，因为他减少了很多的重新慎思的过程。但是，如图3(c)，当规划时间较短，并且世界变化速率较大时，我们发现 Normal 的 Agent 的得分率超过了 Bold 的 Agent。说明我们应该在 Bold 和 Agent 之间取得一个有效的折中能够获得更好地性能。

### 2.4 反应策略的影响

我们为 Agent 的提供更多重新慎思的触发条件 (时钟周期或者走完所有行动)。例如当 Agent 的目标已经不可实现时，即目标食物已经湮灭，Agent 就应该停下来重新慎思新的意图。具体的，我们比较了四种策略

<sup>1</sup> 我们此处与文中对比的略有不同，文中设置  $p=0.5, 1, 2, 4$ ；我们设置  $p=1, 2, 4, 10$  来使得曲线之间的距离更明显。

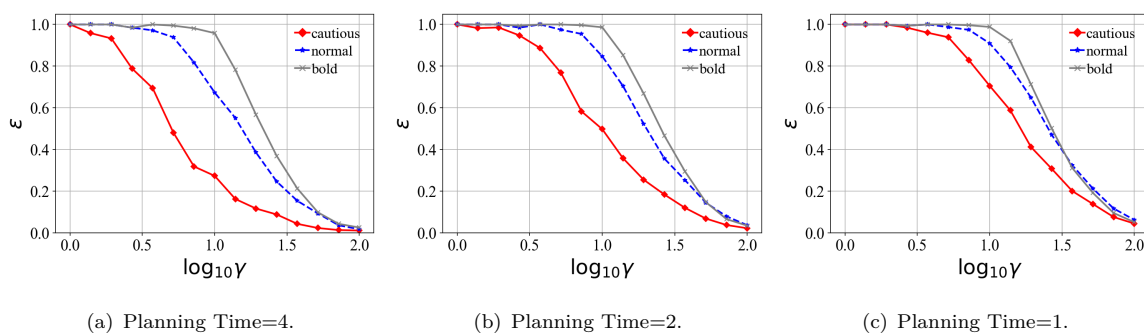


Figure 2.3 承诺属性的影响

- (1) 原始策略 (时钟周期或者走完所有行动)–Blind。
- (2) 当目标湮灭–Disapper。
- (3) 当目标湮灭或者新的洞穴出现–Any\_hole。
- (4) 当目标湮灭或者更近 (我们使用曼哈顿距离来近似真实的最近) 的洞穴出现–Nearer\_hole。

最终实验结果如图2.4所示, 可以看到使用 Near\_hole 或者 Disapper 的策略都相较于原来的 Blind 的策略有明显的提升。而使用 Any\_hole 策略的 Agent 的性能相较于 Blind 略有下降, 因为规划的次数变得太多, 同样与原论文中的结果相一致。

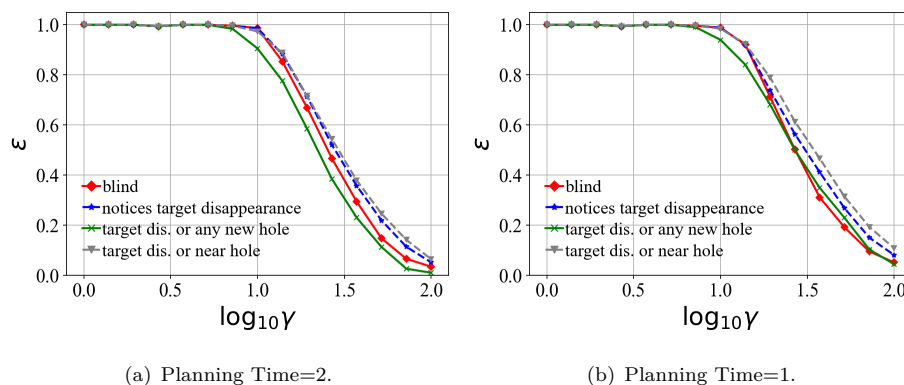


Figure 2.4 反应策略的影响

最后我们还设置  $p=1$ , 比较了使用 Disapper 策略下的 Agent 的承诺属性的影响, 如图2.5所示。可以看到 Bold 的 Agent 当新增了新的反应策略之后在任何情况下都要由于其他的 Agent。而使用 Blind 策略的图3(c)中的结果则显示 Normal 的 Agent 在世界变化率较高的情况下会超过 Bold。同样与原论文中的结果与分析一致。

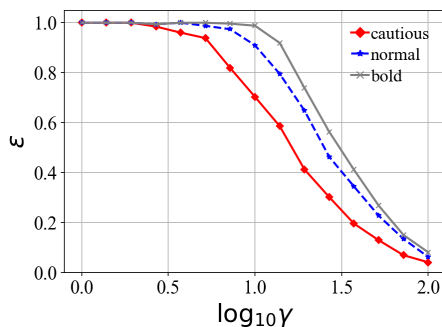


Figure 2.5 使用 Disapper 的反应策略后 Agent 承诺属性变化。

## 参考文献

- [1] D Kinny and M George. Commitment and effectiveness of situated agents. In *IJCAI-91*, pages 82–88. Citeseer, 1991.