

# 多阶段决策实验指导语

在此实验的每个试次中，你将会处于这样一个设定之中：

1. 每一时刻你将在某一个“地点”
2. 此时你将可以使用“b”，“n”，“h”三种按键来“走”到其他“地点”
3. 你的目的就是到达目标“地点”

请注意：

1. 在不同“地点”使用相同按键，所到达的“地点”可能不同
2. 在同一“地点”使用相同按键，所到达的“地点”也有可能不同，但只有较小几率出现“走错”的情况，即在同一“地点”使用同样按键，较大概率会到一个稳定的“地点”。
3. 在不同试次间，“地图”是一致的，即同一“地点”同一按键的结果是一样的（如果不考虑“走错”的情况）

奖励计算方法：

根据完成任务的步数计算奖励，每个试次初始奖励是20，每走一步就会减少1，直到减少到1。