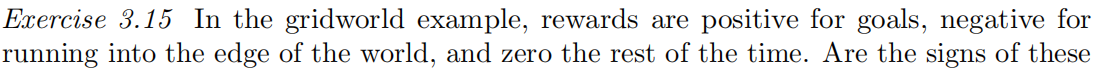
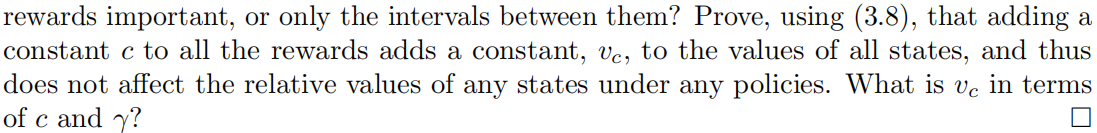


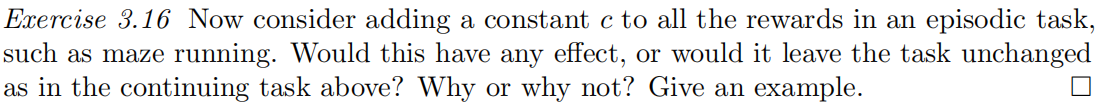
解：

解：是rewards的符号重要。

加上一个常数c后，

因为 所以 , = +

由此可见，对每个收益值都加上一个常数c后就等于对每个状态价值都加上一个常数



解：在分幕式任务中，

加上一个常数c后，

如果加上常数c后，收益值的符号发生了改变，这会影响到智能体的动作选择策略，就可能与分幕式任务的最终目标背道而驰；

而不会改变收益值的符号时，智能体能够在有限时间T内完成任务，但是状态价值的增量会影响终结时间T的大小，例如：负收益值从-1到-0.1，正收益值从10到10.9，可能会导致本来5分钟完成的走迷宫，这次需要花10分钟才能走完，终结时间T的增加也是分幕式任务的不利影响。

策略提升

TC无偏