**16.2**

Chris considers four used cars before buying the one with maximum expected utility. Pat considers ten cars and does the same. All other things being equal, which one is more likely to have the better car? Which is more likely to be disappointed with their car’s quality? By how much (in terms of standard deviations of expected quality)?

（1）Pat 更有可能买到更好的车。假设汽车质量服从某种分布，抽取n个样本并选择最大值，当n增加时，即可选择样本的范围增大时，最大值的期望也会增加。因此，Pat（n=10）的车的期望质量会高于 Chris（n=4）的车的期望质量。

（2）Pat更可能对车子的质量感到失望。

决策者选择车时，是基于带有噪声的效用进行选择的，而非车的真实质量。

设真实质量Q~N(μ, σ2)，带有噪声的效用S=Q+ϵ，其中随机误差ϵ~N(0,τ2)且独立于Q。

决策者选择样本中S最大的车，预期质量U=S，定义失望D=S-Q=ϵ，即质量的高估程度。

当样本数量n增加时，选中的车更可能因正向噪声误差（ϵ>0）而被高估，因此真实质量低于预期的概率更高。

（3）预期质量的标准差σU=

平均失望E[ϵ| 最大的S]=(1-k)E[Sn]，其中k=为信噪比参数，E[Sn]=cn×σU，c

n是标准正态分布下 n个样本最大值的期望系数。

对于标准正态分布：c4≈1.029，c10≈1.539

不妨设，即k=0.5

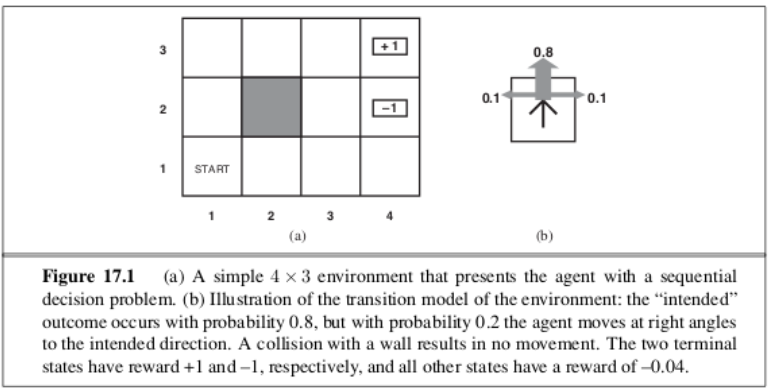
E[DPat]=(1-k)×c10×σU=0.7695σU

E[DChris]=(1-k)×c4×σU=0.5145σU

E[DPat]-E[DChris]=0.255σU

**17.1**

For the 4×3 world shown in Figure 17.1, calculate which squares can be reached from (1,1) by the action sequence [Right,Right,Right,Up,Up] and with what probabilities. Explain how this computation is related to the prediction task (see Section 14.2.1 for a hidden Markov model.)



（1）以下坐标表示均为(行序号， 列序号)

状态转移矩阵（每一列为列状态转移到行状态的概率分布）



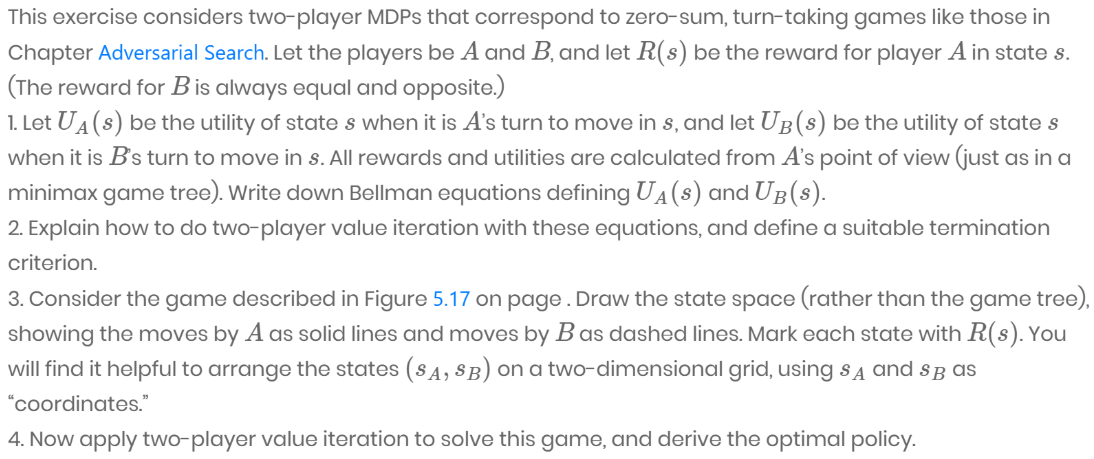


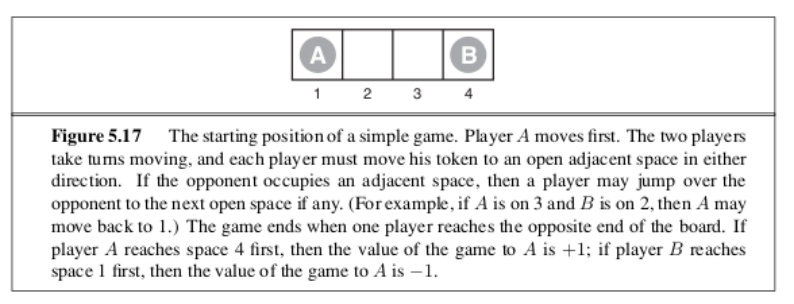
状态向量表（表格中为各步后处于各状态的概率分布）



（2）在隐马尔可夫模型（HMM）中，预测过程即将状态概率向量乘以转移矩阵。这里特殊的地方在于，Up和Right两种动作对应着不同的转移矩阵。

**17.9**





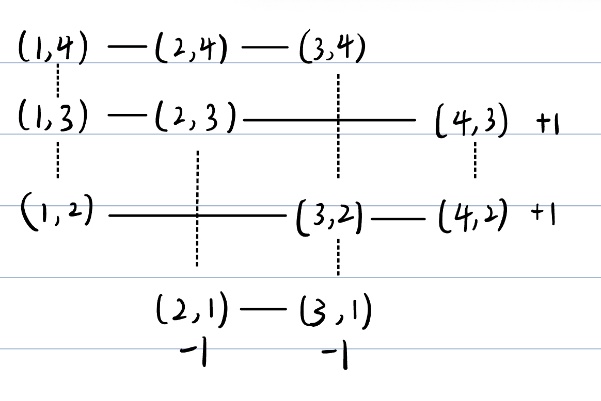
1. UA的贝尔曼方程为UA(s)=R(s)+

UB的贝尔曼方程为UB(s)=R(s)+

2. 价值迭代过程：进行值迭代时，需将每个方程转化为贝尔曼更新公式，并交替应用这些更新公式，对所有状态进行更新。

终止标准：**迭代过程在某个智能体的效用向量与其前一次的效用向量完全相同时终止。**

3. 状态空间如图：



4. 价值迭代过程如下表：

终止标准：**迭代过程在某个智能体的效用向量与其前一次的效用向量完全相同时终止。**



玩家A、B的最佳策略如下：



**17.19**

In the children's game of rock–paper–scissors each player reveals at the same time a choice of rock, paper, or scissors. Paper wraps rock, rock blunts scissors, and scissors cut paper. In the extended version rock–paper–scissors–fire–water, fire beats rock, paper, and scissors; rock, paper, and scissors beat water; and water beats fire. Write out the payoff matrix and find a mixed-strategy solution to this game.

收益矩阵：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **A vs B** | **R** | **P** | **S** | **F** | **W** |
| **R** | 0,0 | -1,1 | 1,-1 | -1,1 | 1,-1 |
| **P** | 1,-1 | 0,0 | -1,1 | -1,1 | 1,-1 |
| **S** | -1,1 | 1,-1 | 0,0 | -1,1 | 1,-1 |
| **F** | 1,-1 | 1,-1 | 1,-1 | 0,0 | -1,1 |
| **W** | -1,1 | -1,1 | -1,1 | 1,-1 | 0,0 |

考虑一个混合策略使得对手无论选哪个选项，期望收益都一样。

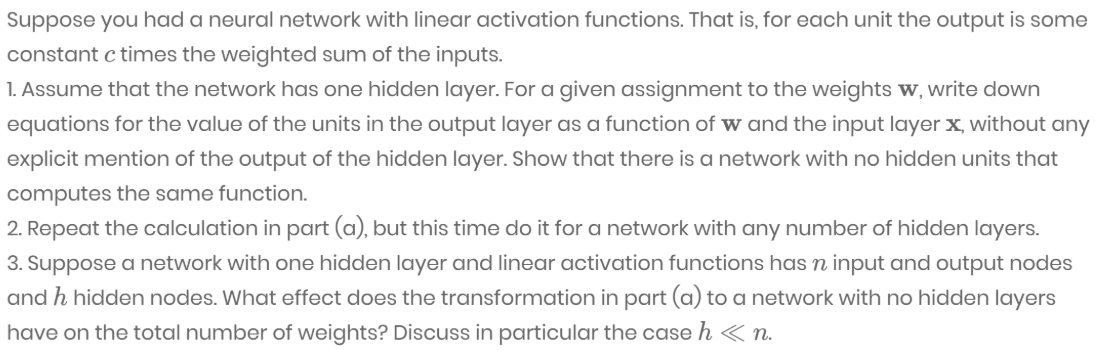
设 pR=pP=pS=a（由对称性），pF​=b，pW=c，且3a+b+c=1

由于B选哪个选项，期望收益都一样：b-c = -3a+c = 3a-b

解得a=1/9, b=c=1/3

即出剪刀石头布的概率均为1/9，出水或火的概率为1/3，是一个混合策略NASH均衡解

**18.29**



1. （1）设输入层有n各节点，x∈Rn；隐藏层有h个节点，权重矩阵W1属于Rh×n；输出层有m个节点，权重矩阵W2属于Rm×h

则隐藏层的输出为h=W1x，输出层输入为y=W2h=W2(W1x)

设W=W2W1，那么输出层（不显式包含隐藏层的输出的）方程为y=Wx

（2）证明：一个无隐藏层的网络（即单层感知机）可直接使用权重矩阵 W∈Rm×n计算 y=Wx，与上述函数相同。因此，存在无隐藏单元的网络可计算相同函数。

2. 同样使用1.中的简化方法可以直接将一个n层网络简化为一个d单层网络。即线性激活函数限制了神经网络只能表示线性函数。

3. 原始网络共有 2hn个权重；而简化后的网络有 n2个权重。当 h≪n时，原始网络的权重要少得多，因此能更简洁地表示输入/输出映射。因此这样的网络比简化网络学得更快。