Answers for Assignment 5 [1]

1

Since we apply the value iteration algorithm (*VI*) to deterministic games, the algorithm becomes

`forall` s { `if` $G(s)$ `then` $L(s) = V(s)$ `else` $L(s) = 0$} (1)
`repeat` (2)
    `forall` s (3)
        `if` $s \in X$ `then` $L(s) = max_{s' \in N(s)} L(s')$ (4)
        `if` $s \in Y$ `then` $L(s) = min_{s' \in N(s)} L(s')$ (5)
`until ``changes are small''` (6)

The "loopy" algorithm is

`forall` s { `if` $G(s)$ `then` $L(s) = V(s)$ `else` $L(s) =$ "?"} (1)
`repeat` (2)
    `forall` $x \in X$ (3)
        `if` $\exists x' \in N(x).L(x') = +1$ `then` $L(x) = +1$ (4)
        `if` $\forall x' \in N(x).L(x') = -1$ `then` $L(x) = -1$ (5)
        `else` $L(x)$ `unchanged` (6)
    `forall` $y \in Y$ (7)
        `if` $\exists y' \in N(y).L(y') = -1$ `then` $L(y) = -1$ (8)
        `if` $\forall y' \in N(y).L(y') = +1$ `then` $L(y) = +1$ (9)
        `else` $L(y)$ `unchanged` (10)
`until ``no changes''` (11)

Line (1) initialises all non-leaf nodes with 0 in VI. Since a win for either X or Y yields values in $\{-1, 1\}$, the above initialisation indicates neither an X win nor a Y win and therefore has the same effect as the "loopy" initialisation to "?" for all non-leaf nodes.

Consider line (3) of VI. For $x \in X$, values taken by the expression $max_{x' \in N(x)} L(x')$ can only be from $\{-1, 0, 1\}$. If *any* $L(x') = 1$, we have $max_{x' \in N(x)} L(x') = 1$, and therefore $L(x) = 1$. This corresponds to line (4) of loopy.

If *every* $L(x') = -1$ then $max_{x' \in N(x)} L(x') = -1$ and $L(x) = -1$. This corresponds to line (5) in loopy. If the above conditions are not satisfied then at least one $L(x') = 0$, and $max_{x' \in N(x)} L(x') = 0$ and the value of $L(x)$ remains unchanged. This corresponds to line (6) in loopy.

We can similarly prove the case for $s \in Y$.

Since we consider deterministic games, any changes that occur are discrete, i.e. values may change only between $\{-1, 0, 1\}$. Hence, changes will be "small" only if there are no changes at all. Therefore, the termination condition corresponds to that in the "loopy" algorithm.

[1] Found an error? Please contact Kedar Swadi(kswadi@cs.princeton.edu)

2

In this game tree, X has 4 possible strategies :

A X1=L, X4=L
B X1=L, X4=R
C X1=R, X4=L
D X1=R, X4=R

Y also has 4 possible strategies :

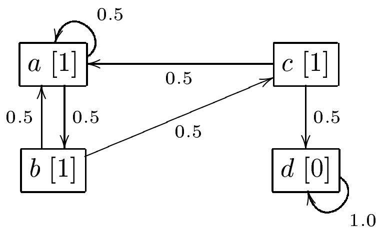
α Y2=L, Y3=L
β Y2=L, Y3=R
γ Y2=R, Y3=L
δ Y2=R, Y3=R

The tabular representation for the game tree is

	A	B	C	D
α	-1	4	5	5
β	-1	4	2	2
γ	2	2	5	5
δ	2	2	2	2

3

The figure below gives the 3-heads Markov chain. State a represents the state of just having started or just having gotten a tail. State b represents one head in a row, c represents 2 heads in a row, and d represents 3 heads in a row.



For each node, s, we define a *neighbour* set $N(s)$ to be the nodes reachable from s in one transition, and we associate a label $[s_c]$, the cost of going from s to any $t \in N(s)$.

We define $P[s \to t]$ as the probability with which we make a transition from s to $t \in N(s)$. We recursively define $E[s \to v]$, the expected number of steps taken to reach node v from node s.

$$E[s \to v] = s_c + \sum_{t \in N(s)} P[s \to t] \times E[t \to v]$$

Informally, the above equation reflects the fact that to go to v from s, we first need to take s_c steps to go to some t, a neighbour of s. This is done with the probability $P[s \to t]$. From each t, we take $E[t \to v]$ expected number of steps to reach the target node v. The sum over all t's of $P[s \to v] \times E[t \to v]$ gives us the expected number of extra steps we need to reach v from s.

In our example, s_c represents the number of coin flips required to go to any of the neighbours. s_c is 1 for all nodes other than nodes d, because once we get 3 heads, we perform no more coin flips. Depending on whether the flip results in a head or a tail, we move to one of the two neighbours of each of nodes a, b, and c. Since we get heads or tails with equal probability, we have

$P[a \to a] = P[a \to b] = P[b \to a] = P[b \to c] = P[c \to a] = P[c \to d] = 0.5$

In nodes d, we perform no flips and make no state changes. Therefore, $P[d \to d] = 1$. d is the "base case" for the above system, since $E[d \to d] = 0$.

The expected number of flips required to obtain three consecutive heads is given by $E[a \to d]$

We can write the required recursive equations as :

$E[a \to d] = 1 + 0.5E[a \to d] + 0.5E[b \to d]$

$E[b \to d] = 1 + 0.5E[a \to d] + 0.5E[c \to d]$

$E[c \to d] = 1 + 0.5E[a \to d] + 0.5E[d \to d]$

$E[d \to d] = 0$

This may be solved either by substitution or by making a matrix. In matrix form, we have

$$\begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix}$$

Trying to solve the problem, we solve the equation $E = C + PE$ by changing it into $(I - P)E = C$. However, we can see that this results in a matrix with the bottom row being all zeros. Since this matrix is singular, we cannot solve it as $E = (I - P)^{-1}C$.

To see why this problem occurs, notice that in this framework, we have $E[d \to d] = 0 + E[d \to d]$, which results in an under-specified system of equations. It does not account for the fact that $E[d \to d] = 0$.

Hence, we "plug in" this information and get a new equation:

$$\begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ 0 \end{pmatrix}$$

This presents us with an equation of the form $E = C + PE^*$. To recover the earlier form of $E = C + PE$, we notice that

$$\begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix}$$

Now, we may solve the equation as

$$\begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix}$$

to get

$$\begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix} = \left(I - \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right)^{-1} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

This gives the solution

$$\begin{pmatrix} E[a \to d] \\ E[b \to d] \\ E[c \to d] \\ E[d \to d] \end{pmatrix} = \begin{pmatrix} 14 \\ 12 \\ 8 \\ 0 \end{pmatrix}$$

The first row is the required answer.

We can use repeated substitution to get
$E[c \to d] = 1 + 0.5E[a \to d]$
$E[b \to d] = 1 + 0.5(1 + 0.5E[a \to d]) + 0.5E[a \to d]$
$\Rightarrow E[b \to d] = 1.5 + 0.75E[a \to d]$
$E[a \to d] = 1 + 0.5E[a \to d] + 0.5(1.5 + 0.75E[a \to d])$
$\Rightarrow E[a \to d] = 1 + 0.5E[a \to d] + (0.75 + 0.375E[a \to d])$
$\Rightarrow E[a \to d] = 1.75 + 0.875E[a \to d]$
$\Rightarrow E[a \to d] = 1.75/0.125 = 14$

Notice that the probability of flipping three coints and getting three heads is 1/8, so we'd expect to flip 3 coins 8 times (24 flips) to get 3 heads in a row this way (looking at the coins in non-overlapping groups of three).