1 Lawler's 'Introduction to Stochastic Processes' Chapter 4 solutions [1]

4.1 [**0**,2.5,5,7.5,**10**,8.6,7.2,5.8,4.4,**3**,**0**] Optimal stopping is in the bold points. Notice that it's clear we should stop at the maximum of the array, and so it splits the problem into two smaller problems (left and right). If a points is smaller than the average of its neighbors, there is no stopping there. If it's not, then at least one of the neighbors must also be a no-stop, and so on. Since the problem is small I used this method of deduction rather than follow the algorithm given in the chapter.

4.2 a. Expectation is $\frac{5}{6} \cdot 7 = 5\frac{5}{6}$. b. We will use the same recursive method time and again in these exercises. We can write

$$V = \frac{12}{36}x_{12} + \frac{22}{36}x_{11} + \frac{30}{36}x_{10} + x_9 + \frac{40}{36}x_8 + \frac{30}{36}x_6 + \frac{20}{36}x_5 + \frac{12}{36}x_4 + \frac{6}{36}x_3 + \frac{2}{36}x_2 +$$
$$(\frac{5}{6} - \frac{1}{36}x_{12} - \frac{2}{36}x_{11} - \frac{3}{36}x_{10} - \frac{4}{36}x_9 - \frac{5}{36}x_8 - \frac{5}{36}x_6 - \frac{4}{36}x_5 - \frac{3}{36}x_4 - \frac{2}{36}x_3 - \frac{1}{36}x_2)V$$

Where $x_{12}, ..., x_2$ are $\{0, 1\}$ indicator variables for stopping on a roll i of the two dices. The coefficients in the first line are the multiple of probability times revenue if chosen to stop, and in the second line the probabilities.

Changing wings we get

$$V = \frac{12x_{12} + 22x_{11} + 30x_{10} + 36x_9 + 40x_8 + 30x_6 + 20x_5 + 12x_4 + 6x_3 + 2x_2}{6 + x_{12} + 2x_{11} + 3x_{10} + 4x_9 + 5x_8 + 5x_6 + 4x_5 + 3x_4 + 2x_3 + x_2}$$

Notice a convenient property that the ratio of the variables between nominator and denominator coefficients is monotone increasing. This condition holds throughout the exercise, and it means we can only consider monotone rules, that is some roll value that is a threshold for stopping above it, and no stopping below it. In this case we stop for 8 and above, with expectation $6\frac{2}{3}$.

4.3 a. [**0**, **2**, **4**, 6.25, **10**, 7.25, **6**, **4**, **3**, **3**, **0**]

b. [**0**, **2**, 4.11,6.7,**10**,7.6,**6**, 4.554,3.588, **3**,**0**]

c. The stopping rule is same as a. Notice that as we add more costs and discount factors, the stopping rule is monotone increasing (set-wise). So we only need to consider whether at the two non-stopping points something is changed, but it isn't.

4.4. a. Given costs, the change in the recursive formula is you need to split the V coefficient in rhs to $(V - 1)$ and $(V - 2)$ coefficients accordingly, and change wings again, then again check monotonically where to stop. (It's a bit similar to h-index calculation - you need to see where the total ratio meets the individual ratio of some indicator). The only non-stops are 2,3 dice roll sums.

b. Given a discount factor, the V in rhs needs to be multiplied by it. The only non-stops are 2,3,4 dice roll sums.

c. The only non-stops are at 2,3 dice roll sums.

4.5. Since the stopping rule of stopping everywhere is optimal, it is better than any other stopping rule, specifically the 4 rules where you move only in one of the nodes. This yields a set of inequalities that reduces to $\alpha \leq \frac{4}{7}$. For B, D even making one step and reaching w.p 1 to C with such discount factor is worse than staying, so we stay there for sure. We always stay at the max. So we reduced to a case where only A may move, but this is already answered for with $\alpha = \frac{4}{7}$. So this is our maximal α value.

4.6 a. Optimal stopping is at $5, 6$ dice rolls. Which gives $11\frac{1}{3}$ expected revenue.

b. We arrive at the parametric equation

$$V = \frac{-5r + (36+r)x_6 + (25+r)x_5 + (16+r)x_4 + (9+r)x_3 + (4+r)x_2}{1 + x_6 + x_5 + x_4 + x_3 + x_2}$$

We need to solve $\frac{-r+86}{5} \geq 4 + r$, $\frac{-2r+77}{4} \leq 9 + r$, which yields $6\frac{5}{6} \leq r \leq 11$, so the minimal r cost for such optimal stopping is $6\frac{5}{6}$.

4.8. Notice that every round w.p $\frac{11}{12}$ we win an average of extra 600, and w.p. $\frac{1}{12}$ get 0. So the equation is $V = \frac{11}{12}(600 + V)$ and so the threshold value to stop is $V = 6600$.

Now we can use dynamic programming to go back and calculate expected revenue for any value we currently have in the game. If we have some value of 6600 or above, we get this value (because we stop). For 6500, we get w.p $\frac{1}{12}$ any of the $6500 + 100i$ values with $1 \leq i \leq 11$. And so on. A computer program calculates it and gets that the expected value of the game is 3133.7606770342763.

4.9 a. $\frac{(n+1)^2+(n-1)^2}{2} = n^2 + 1 > n^2$.

b. Since it's infinite and not finite, the aim to find a single 'optimal' rule is futile. It's clear that the rules are uniquely determined by the first point of stopping, and that it's monotonically increasing to move it further and further. But it doesn't mean that moving it 'infinitely' further and removing the point of stopping altogether is best - it is not.

References

[1] Gregory F Lawler. *Introduction to stochastic processes*. CRC Press, 2006.