Example - Bart Simpson's Garden

Bart Simpson wants to grow vegies in his backyard. Over the years, he observed that current year's productivity can be assumed to depend only on last year's soil condition. The transition probabilities over a 1-year period from one productivity state to another is as follows:

(1) If he doesn't do anything about the soil, i.e. no fertilizing; and

$p^{NF} =$

		State of soil next year		
		Good	Fair	Poor
State of soil this year	Good	0.2	0.5	0.3
	Fair	0	0.5	0.5
	Poor	0	0	1

Example - Bart Simpson's Garden

(2) if he is keen on his garden and fertilize it.

$p^F =$

		State of soil next year		
State		Good	Fair	Poor
of soil	Good	0.3	0.6	0.1
this	Fair	0.1	0.6	0.3
year	Poor	0.05	0.4	0.55

Example - Bart Simpson's Garden

The productivity of the garden is measured by a reward structure:
(1) If Bart Simpson doesn't fertilize his garden; and

$$r^{NF} = \begin{array}{cccc} & \text{Good} & \text{Fair} & \text{Poor} \\ \text{Good} & 7 & 6 & 3 \\ \text{Fair} & 0 & 5 & 1 \\ \text{Poor} & 0 & 0 & -1 \end{array}$$

Example - Bart Simpson's Garden

(2) if he does.

$r^F =$

	Good	Fair	Poor
Good	6	5	-1
Fair	7	4	0
Poor	6	3	-2

Question: Help poor Bart with his backyard by formulating a LP model for an infinite-stage MDP problem, assuming a discounted rate of $\alpha = 0.6$.

Example - Bart Simpson's Garden

Solution:

Objective: maximise productivity.

Let f_i be the discounted productivity if period begins with state i.

Observe:

$$f_i = \max_{d \in F, NF} \left\{ r_{id} + \alpha \sum_{j=1}^{3} p_{ij}^{d} f_j \right\},$$

for $i =$ good, fair, and poor.

Using 1-Good, 2-Fair, 3-Poor: we have, for example,

$$\begin{aligned} f_1 &= \max \Big\{ r_{1,NF} + 0.6(0.2f_1 + 0.5f_2 + 0.3f_3), \\ &\qquad r_{1,F} + 0.6(0.3f_1 + 0.6f_2 + 0.1f_3) \Big\}. \end{aligned}$$

(Exercise: Figure out what f_2 and f_3 are!)

The rewards, i.e., $r_{1,F}$, $r_{1,NF}$, $r_{2,F}$, $r_{2,NF}$, $r_{3,F}$, $r_{3,NF}$ are:

$$r_{1,NF} = \sum_{j=1}^{3} p_{ij}^{NF} r_{ij}^{NF} = (0.2)7 + (0.5)6 + (0.3)3 = 5.3;$$

and similarly

$$\begin{aligned} r_{2,NF} &= 3; \\ r_{3,NF} &= -1; \\ r_{1,F} &= 4.7; \\ r_{2,F} &= 3.1; \text{ and} \\ r_{3,F} &= 0.4. \end{aligned}$$

Example - Bart Simpson's Garden

Thus, to find the optimal policy for Bart, we solve the following LP:

$$
\begin{array}{rcll}
\min z & = & f_1 + f_2 + f_3 & \\
f_1 & \geq & 4.7 + 0.12 f_1 + 0.3 f_2 + 0.18 f_3 & (1, F) \\
f_1 & \geq & 5.3 + 0.18 f_1 + 0.36 f_2 + 0.18 f_3 & (1, NF) \\
f_2 & \geq & 3.1 + 0.06 f_1 + 0.36 f_2 + 0.18 f_3 & (2, F) \\
f_2 & \geq & 3 + 0.3 f_2 + 0.3 f_3 & (2, NF) \\
f_3 & \geq & 0.4 + 0.03 f_1 + 0.24 f_2 + 0.33 f_3 & (3, F) \\
f_3 & \geq & -1 + 0.6 f_3 & (3, NF) \\
 & & f_1, f_2, f_3 \geq 0 &
\end{array}
$$

Example - Bart Simpson's Garden

Results: $f_1 = 10.204$, $f_2 = 6.780$, and $f_3 = 3.483$.

Conclusions: The constraints representing $(1, F)$, $(2, NF)$, and $(3, NF)$ are slack, meaning these decisions are not part of the optimal policy, and those representing $(1, NF)$, $(2, F)$, and $(3, F)$ are binding, meaning these decision are in fact part of the optimal policy. (Alternatively, we can look at the values of the optimal dual variables: those with respect to decisions $(1, NF)$, $(2, F)$, and $(3, F)$ have non-zero values in the optimal solution). Hence the optimal policy is: **No** fertilising is needed when the soil is in good condition, and **fertilising** is needed when soil is fair or poor conditions! (This does make a lot of sense, doesn't it?)