CS3061 Artificial Intelligence

Submit to Blackboard by Monday, March 5th (23:59)

Recall from lecture[1] that Sam is either fit or unfit

$$S \;=\; \{\text{fit, unfit}\}$$

and has to decide whether to exercise or relax

$$A \;=\; \{\text{exercise, relax}\}$$

on the basis of the following (probability, reward)-matrices $(p(s,a,s'), r(s,a,s'))$ for row s, column s' in table with corner a

exercise	fit	unfit
fit	.99, 8	.01, 8
unfit	.2, 0	.8, 0

relax	fit	unfit
fit	.7, 10	.3, 10
unfit	0, 5	1, 5

The γ-discounted value of (s,a) is

$$\lim_{n\to\infty} q_n(s,a)$$

where

$$\begin{aligned} q_0(s,a) &:= p(s,a,\text{fit})r(s,a,\text{fit}) + p(s,a,\text{unfit})r(s,a,\text{unfit}) \\ V_n(s) &:= max(q_n(s,\text{exercise}), q_n(s,\text{relax})) \\ q_{n+1}(s,a) &:= q_0(s,a) + \gamma\big(p(s,a,\text{fit})V_n(\text{fit}) + p(s,a,\text{unfit})V_n(\text{unfit})\big). \end{aligned}$$

In particular, $\gamma = 0.9$ leads to the following $q_n(s,a)$ for $n = 0, 1, 2$

	exercise	relax	π
fit	8, 16.955, 23.812	10, 17.65, 23.685	relax, relax, exercise
unfit	0, 5.4, 10.017	5, 9.5, 13.55	relax, relax, relax

Your task is to write a program that given

a positive integer n, a γ-setting G $(0 < G < 1)$, and a state s

returns the values

$$q_n(s,\text{exercise}) \quad \text{and} \quad q_n(s,\text{relax})$$

for $\gamma = G$. You may use any of the following programming languages

Prolog, Java, Python

but be prepared to demonstrate your program on Tue, March 6 (noon-1, LG 12, O'Reilly) or Wed, March 7 (10-11, LB04; on your machine).

[1] It may help to read Poole & Mackworth, 9.5 Decision Processes.