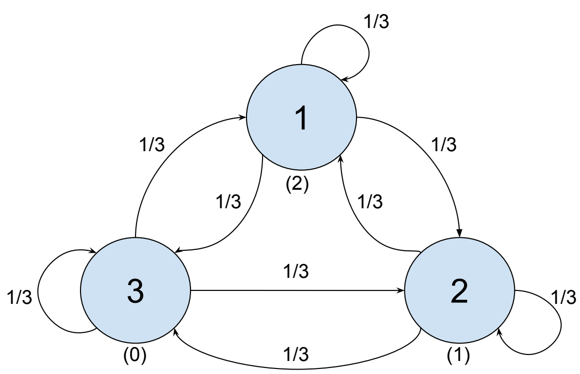
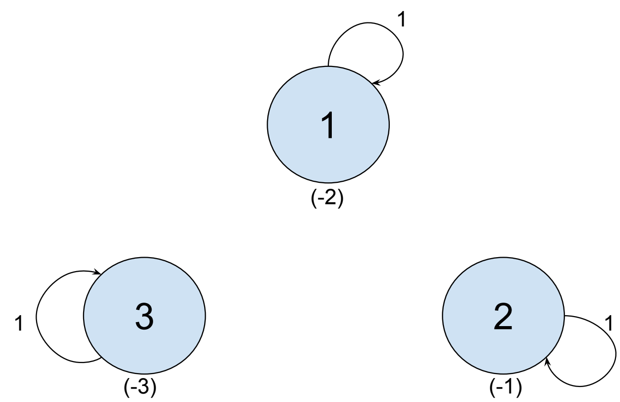
Ειδικά Θέματα Επιχειρησιακής Έρευνας
Ενδεικτικές λύσεις 4ης Σειράς Ασκήσεων

---

1. **Θεωρήστε τη Μαρκοβιανή διαδικασία απόφασης με καταστάσεις $S = \{1, 2, 3\}$, αποφάσεις $A = \{a_1, a_2\}$ και πιθανότητες μετάβασης οι οποίες δίδονται από τα διαγράμματα που ακολουθούν**



(α′) Απόφαση a_1



(β′) Απόφαση a_2

Το άμεσο κόστος $c(i, a)$ για την κατάσταση i και απόφαση a δίδεται σε παρενθέσεις κάτω από την κατάσταση i στο διάγραμμα της απόφασης a.

(α′) **Βρείτε τη βέλτιστη πολιτική σύμφωνα με το κριτήριο του μέσου συσσωρευμένου κόστους για ένα χρονικό ορίζοντα 4 βημάτων (βήμα 0 έως και 3).**

Η εξίσωση δυναμικού προγραμματισμού για το κριτήριο αυτό είναι

$$\begin{cases} V_{n+1}(1) = \min\left\{2 + \frac{1}{3}\left[V_n(1) + V_n(2) + V_n(3)\right], -2 + V_n(1)\right\} \\ V_{n+1}(2) = \min\left\{1 + \frac{1}{3}\left[V_n(1) + V_n(2) + V_n(3)\right], -1 + V_n(2)\right\} \\ V_{n+1}(3) = \min\left\{\frac{1}{3}\left[V_n(1) + V_n(2) + V_n(3)\right], -3 + V_n(3)\right\} \end{cases} \tag{1}$$

για $n = 0, 1, \ldots$ και $V_0(1) = \min\{2, -2\}, V_0(2) = \min\{1, -1\}, V_0(3) = \min\{0, -3\}$.

Όταν απομένουν $n = 0$ βήματα πριν το τέλος του χρονικού ορίζοντα, δηλαδή βρισκόμαστε στο 4ο βήμα, η βέλτιστη απόφαση είναι η a_2 ανεξάρτητα από την κατάσταση, και $V_0(1) = -2, V_0(2) = -1, V_0(3) = -3$.

Για να βρούμε τις βέλτιστες αποφάσεις στο βήμα 3 (όπου απομένει 1 βήμα πριν το τέλος) υπολογίζουμε την (1) για $n = 0$ και βρίσκουμε $V_1(1) = \min\{0, -4\} = -4(a_2), V_1(2) = \min\{-1, -2\} = -2(a_2), V_1(3) = \min\{-2, -6\} = -6(a_2)$, όπου στις παρενθέσεις δίδεται η βέλτιστη απόφαση για κάθε κατάσταση.

Για τις βέλτιστες αποφάσεις στο βήμα 2 (2 βήματα πριν το τέλος) υπολογίζουμε την (1) για $n = 1$: $V_2(1) = \min\{-2, -6\} = -6(a_2), V_2(2) = \min\{-3, -3\}(a_1/a_2), V_2(3) = \min\{-4, -9\} = -9(a_2)$.

Τέλος, στο 1ο βήμα (3 πριν το τέλος) έχουμε: $V_3(1) = \min\{-4, -8\} = -8(a_2), V_3(2) = \min\{-5, -4\} = -5(a_1), V_3(3) = \min\{-6, -12\} = -12(a_2)$.

Συνοψίζοντας, η βέλτιστη πολιτική λαμβάνει πάντοτε την απόφαση a_2 εκτός στο 1ο βήμα όταν αρχική κατάσταση είναι η 2, όπου λαμβάνεται η απόφαση a_1. Παρατηρήστε ότι είναι εξίσου βέλτιστο να παρθεί η απόφαση a_1 και η a_2 στο 2ο βήμα στην κατάσταση 2.

(β′) **Δώστε τις εξισώσεις δυναμικού προγραμματισμού για το κριτήριο του μέσου υποτιμούμενου κόστους με συντελεστή υποτίμησης $\beta = 3/4$.**

Η εξίσωση δυναμικού προγραμματισμού για αυτό το κριτήριο είναι

$$\begin{cases} V(1) = \min\left\{2 + \frac{1}{4}\left[V(1) + V(2) + V(3)\right], -2 + \frac{3}{4}V(1)\right\} \\ V(2) = \min\left\{1 + \frac{1}{4}\left[V(1) + V(2) + V(3)\right], -1 + \frac{3}{4}V(2)\right\} \\ V(3) = \min\left\{\frac{1}{4}\left[V(1) + V(2) + V(3)\right], -3 + \frac{3}{4}V(3)\right\} \end{cases} \tag{2}$$

Βρείτε τη βέλτιστη πολιτική σύμφωνα με αυτό το κριτήριο.

Για την εύρεση της βέλτιστης πολιτικής θα υπολογίσουμε τη λύση της (2) με τη μέθοδο βελτίωσης πολιτικής αρχίζοντας (αυθαίρετα) με την πολιτική όπου σε όλες τις καταστάσεις λαμβάνεται η απόφαση a_2.

1η επανάληψη: το σύστημα εξισώσεων που αντιστοιχεί στην αρχική πολιτική είναι

$$\begin{cases} V(1) = -2 + \frac{3}{4}V(1) \\ V(2) = -1 + \frac{3}{4}V(2) \\ V(3) = -3 + \frac{3}{4}V(3) \end{cases} \Leftrightarrow \begin{cases} V(1) = -8 \\ V(2) = -4 \\ V(3) = -12 \end{cases}$$

Τώρα βρίσκουμε τις καλύτερες αποφάσεις για κάθε κατάσταση. Για την κατάσταση 1, η a_1 δίνει $2 + \frac{1}{4}\left[V(1) + V(2) + V(3)\right] = -4 > -8 = -2 + \frac{3}{4}V(1)$, άρα στην επόμενη επανάληψη θα χρησιμοποιήσουμε την a_2 (στην κατάσταση 1).

Στην κατάσταση 2 έχουμε $1 + \frac{1}{4}\left[V(1) + V(2) + V(3)\right] = -5 < -4 = -1 + \frac{3}{4}V(2)$, άρα στην επόμενη επανάληψη θα χρησιμοποιηθεί η a_1.

Στην κατάσταση 3 έχουμε $\frac{1}{4}\left[V(1) + V(2) + V(3)\right] = -6 > -12 = -3 + \frac{3}{4}V(3)$, άρα στην επόμενη επανάληψη θα χρησιμοποιηθεί η a_2.

2η επανάληψη: η πολιτική που προέκυψε στην προηγούμενη επανάληψη, λαμβάνει τις αποφάσεις a_2, a_1 και a_2 στις καταστάσεις 1,2 και 3 αντίστοιχα. Άρα λύνουμε το σύστημα:

$$\begin{cases} V(1) = -2 + \frac{3}{4}V(1) \\ V(2) = 1 + \frac{1}{4}\left[V(1) + V(2) + V(3)\right] \\ V(3) = -3 + \frac{3}{4}V(3) \end{cases} \Leftrightarrow \begin{cases} V(1) = -8 \\ V(2) = -\frac{16}{3} \\ V(3) = -12 \end{cases}$$

Τώρα, για την κατάσταση 1 έχουμε $2 + \frac{1}{4}[V(1) + V(2) + V(3)] = -\frac{13}{3} > -8 = -2 + \frac{3}{4}V(1)$, άρα στην επόμενη επανάληψη θα χρησιμοποιήσουμε a_2.

Στην κατάσταση 2 έχουμε $1 + \frac{3}{4}V(2) = -5 > -\frac{16}{3} = 1 + \frac{1}{4}[V(1) + V(2) + V(3)] =$, άρα στην επόμενη επανάληψη θα χρησιμοποιηθεί η a_1.

Στην κατάσταση 3 έχουμε $\frac{1}{4}[V(1) + V(2) + V(3)] = -\frac{19}{3} > -12 = -3 + \frac{3}{4}V(3)$, άρα στην επόμενη επανάληψη θα χρησιμοποιηθεί η a_2.

Παρατηρούμε ότι η πολιτική που προκύπτει είναι η ίδια που βρήκαμε στην προηγούμενη επανάληψη, άρα $V(1) = -8, V(2) = -\frac{16}{3}, V(3) = -12$ αποτελεί λύση της (2) και η βέλτιστη πολιτική είναι λήψη απόφασης a_2, a_1 και a_2 στις καταστάσεις 1,2 και 3 αντίστοιχα.

(γ′) **Δώστε τις εξισώσεις δυναμικού προγραμματισμού για το κριτήριο του χρονικά μέσου κόστους.**

Η εξίσωση δυναμικού προγραμματισμού για αυτό το κριτήριο είναι

$$\begin{cases} h(1) + \gamma = \min\left\{2 + \frac{1}{3}[h(1) + h(2) + h(3)], -2 + h(1)\right\} \\ h(2) + \gamma = \min\left\{1 + \frac{1}{3}[h(1) + h(2) + h(3)], -1 + h(2)\right\} \\ h(3) + \gamma = \min\left\{\frac{1}{3}[h(1) + h(2) + h(3)], -3 + h(3)\right\} \end{cases}$$

ως προς τους άγνωστους $h(1), h(2), h(3), \gamma$. Εφόσον υπάρχει ένας παραπάνω βαθμός ελευθερίας για τους 3 πρώτους άγνωστους, θέτουμε αυθαίρετα $h(3) = 0$ και λαμβάνουμε 3 εξισώσεις με 3 άγνωστους:

$$\begin{cases} h(1) + \gamma = \min\left\{2 + \frac{1}{3}[h(1) + h(2)], -2 + h(1)\right\} \\ h(2) + \gamma = \min\left\{1 + \frac{1}{3}[h(1) + h(2)], -1 + h(2)\right\} \\ \gamma = \min\left\{\frac{1}{3}[h(1) + h(2)], -3\right\} \end{cases} \tag{3}$$

Βρείτε τη βέλτιστη πολιτική σύμφωνα με αυτό το κριτήριο.

Μπορούμε να μαντέψουμε τη βέλτιστη πολιτική για το κριτήριο του μέσου χρονικού κόστους, παρατηρώντας τα εξής. Το μικρότερο δυνατό κόστος σε ένα βήμα χρεώνεται όταν στην κατάσταση 3 λαμβάνεται η απόφαση a_2. Μάλιστα, η απόφαση a_2 οδηγεί πάλι στην κατάσταση 3, οπότε το καλύτερο δυνατό είναι η Μαρκοβιανή διαδικασία να μένει πάντα στην 3 όπου θα λαμβάνεται η a_2. Το μέσο χρονικά κόστος που προκύπτει είναι -3. Στις υπόλοιπες καταστάσεις, τα κόστη που χρεώνονται είναι μεγαλύτερα από -3, άρα αρκεί στις καταστάσεις αυτές να ληφθεί η απόφαση a_1 ώστε η διαδικασία κάποτε να επισκεφτεί την κατάσταση 3 και να μείνει εκεί για πάντα.

Στη συνέχεια θα επιβεβαιώσουμε ότι η βέλτιστη πολιτική είναι αυτή που λαμβάνει τις αποφάσεις a_1, a_1, a_2 στις καταστάσεις 1,2,3 αντίστοιχα. Για αυτόν το σκοπό, αρκεί να δείξουμε ότι η λύση του συστήματος

$$\begin{cases} h(1) + \gamma = 2 + \frac{1}{3}[h(1) + h(2)] \\ h(2) + \gamma = 1 + \frac{1}{3}[h(1) + h(2)] \\ \gamma = -3 \end{cases} \Leftrightarrow \begin{cases} h(1) = 14 \\ h(2) = 13 \\ \gamma = -3 \end{cases}$$

ικανοποιεί το σύστημα (3). Αυτό όντως ισχύει, αφού $-2+h(1)=12>11=2+\frac{1}{3}\left[h(1)+h(2)\right]$, $-1+h(2)=12>10=1+\frac{1}{3}\left[h(1)+h(2)\right]$ και $\frac{1}{3}\left[h(1)+h(2)\right]=9>-3$.

2. **Εργάζεστε σε μια γραμμή συναρμολόγησης ενός εργοστασίου όπου κάθε λεπτό έχετε στη διάθεσή σας έναν τυχαίο αριθμό Μη Συναρμολογημένων Αντικειμένων (ΜΣΑ) που πρέπει να συναρμολογήσετε. Κάθε λεπτό μπορείτε να επιλέξετε μεταξύ του να αναπαυθείτε (Α) ή να εργαστείτε (Ε) για το επόμενο λεπτό. Κάθε φορά που εργάζεστε, ο αριθμός των διαθέσιμων ΜΣΑ στο επόμενο λεπτό μειώνεται κατά ένα με πιθανότητα 1/2 εφόσον υπάρχουν διαθέσιμα ΜΣΑ, ή μένει σταθερός με πιθανότητα 1/2 (εάν δεν υπαρχουν διαθέσιμα ΜΣΑ τότε η πιθανότητα αυτή είναι 1). Όταν αναπαύεστε ο αριθμός των ΜΣΑ είτε αυξάνει κατά ένα με πιθανότητα 1/2, είτε μένει σταθερός με πιθανότητα 1/2. Εάν επιλέξετε να αναπαυθείτε όταν ο αριθμός των ΜΣΑ είναι 2, ο αριθμός τους παραμένει 2 (δηλαδή, δεν αυξάνεται) στο επόμενο λεπτό με πιθανότητα 1. Συνεπώς, ο αριθμός των ΜΣΑ σε κάθε λεπτό μπορεί να είναι 0, 1 ή 2.**

   Όταν ο αριθμός των ΜΣΑ είναι 0 ή 1, το κόστος της απόφασης Α είναι -1, ενώ όταν τα ΜΣΑ είναι 2, το κόστος αυτό είναι 2. Η απόφαση Ε σας δημιουργεί κόστος 1 για οποιοδήποτε αριθμό ΜΣΑ.

   (α΄) **Για το παραπάνω πρόβλημα, γράψτε την εξίσωση δυναμικού προγραμματισμού σύμφωνα με το κριτήριο του μέσου υποτιμούμενου κόστους, με συντελεστή υποτίμησης $\beta=1/2$.**

   Η περιγραφή του προβλήματος οδηγεί στην ακόλουθη εξίσωση.

$$\begin{cases} V(0)=\min\left\{-1+\frac{1}{4}\left[V(0)+V(1)\right],1+\frac{1}{2}V(0)\right\} \\ V(1)=\min\left\{-1+\frac{1}{4}\left[V(1)+V(2)\right],1+\frac{1}{4}\left[V(1)+V(0)\right]\right\} \\ V(2)=\min\left\{2+\frac{1}{2}V(2),1+\frac{1}{4}\left[V(2)+V(1)\right]\right\} \end{cases} \tag{4}$$

   (β΄) **Βρείτε τη βέλτιστη πολιτική σύμφωνα με το παραπάνω κριτήριο.**

   Για την εύρεση της βέλτιστης πολιτικής, λύνουμε την (4) ακολουθώντας τη μέθοδο βελτίωσης πολιτικής όπου αρχίζουμε από την πολιτική όπου ανάπαυση αποφασίζεται μόνο όταν δεν υπάρχουν ΜΣΑ, δηλαδή λαμβάνονται οι αποφάσεις A,E,E στις καταστάσεις 0,1 και 2 αντίστοιχα.

   1η επανάληψη: Το σύστημα εξισώσεων που αντιστοιχεί στην αρχική πολιτική είναι:

$$\begin{cases} V(0)=-1+\frac{1}{4}\left[V(0)+V(1)\right] \\ V(1)=1+\frac{1}{4}\left[V(1)+V(0)\right] \\ V(2)=1+\frac{1}{4}\left[V(2)+V(1)\right] \end{cases} \Leftrightarrow \begin{cases} V(0)=-1 \\ V(1)=1 \\ V(2)=\frac{5}{3} \end{cases}$$

   Στην κατάσταση 0, έχουμε $1+\frac{1}{2}V(0)=\frac{1}{2}>-1=-1+\frac{1}{4}\left[V(0)+V(1)\right]$ άρα χρησιμοποιούμε A στην επόμενη επανάληψη.

Στην κατάσταση 1, έχουμε $-1+\frac{1}{4}\left[V(1)+V(2)\right]=-\frac{1}{3}<1=1+\frac{1}{4}\left[V(1)+V(0)\right]$ άρα χρησιμοποιούμε A στην επόμενη επανάληψη.

Στην κατάσταση 2, έχουμε $2+\frac{1}{2}V(2)=\frac{17}{6}>\frac{10}{6}=1+\frac{1}{4}\left[V(1)+V(2)\right]$ άρα χρησιμοποιούμε E στην επόμενη επανάληψη.

2η επανάληψη: το σύστημα που αντιστοιχεί στην πολιτική που λαμβάνει την απόφαση A, A και E στην κατάσταση 0,1 και 2, είναι:

$$\begin{cases} V(0)=-1+\dfrac{1}{4}\left[V(0)+V(1)\right] \\ V(1)=-1+\dfrac{1}{4}\left[V(1)+V(2)\right] \\ V(2)=1+\dfrac{1}{4}\left[V(2)+V(1)\right] \end{cases} \Leftrightarrow \begin{cases} V(0)=-\dfrac{5}{3} \\ V(1)=-1 \\ V(2)=1 \end{cases}$$

Στην κατάσταση 0, έχουμε $1+\frac{1}{2}V(0)=\frac{1}{6}>-\frac{5}{3}=-1+\frac{1}{4}\left[V(0)+V(1)\right]$ άρα χρησιμοποιούμε A στην επόμενη επανάληψη.

Στην κατάσταση 1, έχουμε $1+\frac{1}{4}\left[V(1)+V(0)\right]=\frac{1}{3}>-1=-1+\frac{1}{4}\left[V(1)+V(2)\right]$ άρα χρησιμοποιούμε A στην επόμενη επανάληψη.

Στην κατάσταση 2, έχουμε $2+\frac{1}{2}V(2)=\frac{5}{2}>1=1+\frac{1}{4}\left[V(2)+V(1)\right]$ άρα χρησιμοποιούμε E στην επόμενη επανάληψη.

Παρατηρούμε ότι η πολιτική που προκύπτει είναι η ίδια που βρήκαμε στην προηγούμενη επανάληψη, άρα $V(0)=-\frac{5}{3}, V(1)=-1, V(2)=1$ αποτελεί λύση της (4) και η βέλτιστη πολιτική είναι να εργάζεστε μόνο στην κατάσταση 2.

3. **Λύστε το σύστημα εξισώσεων**

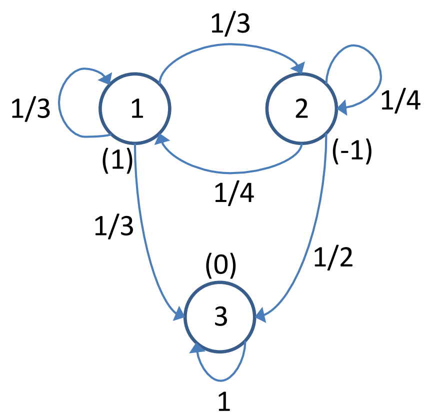
$$x=\min\left(1+\frac{x+y}{6},-1+\frac{y}{2}\right)$$
$$y=\min\left(-1+\frac{x+y}{8},1+\frac{x}{2}\right)$$

ως προς x, y.

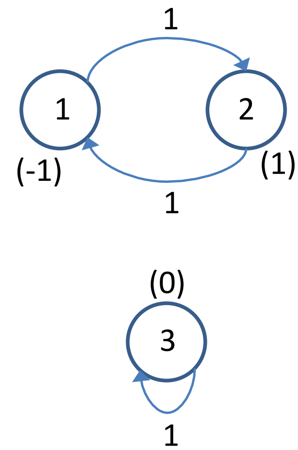
Παρατηρούμε ότι οι εξισώσεις μοιάζουν με την εξίσωση δυναμικού προγραμματισμού για το κριτήριο του υποτιμούμενου κόστους για μια Μαρκοβιανή διαδικασία απόφασης. Οπότε, μπορούμε να δοκιμάσουμε να τις λύσουμε με τη μέθοδο της βελτίωσης πολιτικής.

Για να μπορέσουμε να μαντέψουμε μια ¨καλή' αρχική πολιτική, κατασκευάζουμε την Μαρκοβιανή διαδικασία απόφασης της οποίας η εξίσωση δυναμικού προγραμματισμού αντιστοιχεί στις παραπάνω εξισώσεις. Θεωρήστε σύνολο καταστάσεων $S=\{1,2,3\}$, αποφάσεων $A=\{a_1,a_2\}$ και τις πιθανότητες μετάβασης και άμεσα κόστη (στις παρενθέσεις) που δίδονται από τα διαγράμματα στο Σχήμα 1. Η εξίσωση δυναμικού προγραμματισμού για το κριτήριο του υποτιμούμενου κόστους με συντελεστή υποτίμησης 1/2, είναι:

$$\begin{cases} V(1)=\min\left\{1+\dfrac{1}{2}\left[\dfrac{1}{3}V(1)+\dfrac{1}{3}V(2)+\dfrac{1}{3}V(3)\right],-1+\dfrac{1}{2}V(2)\right\} \\ V(2)=\min\left\{-1+\dfrac{1}{2}\left[\dfrac{1}{4}V(1)+\dfrac{1}{4}V(2)+\dfrac{1}{2}V(3)\right],1+\dfrac{1}{2}V(1)\right\} \\ V(3)=\min\left\{\dfrac{1}{2}V(3),\dfrac{1}{2}V(3)\right\} \end{cases} \tag{5}$$



(α′) Απόφαση a_1



(β′) Απόφαση a_2

Σχήμα 1: Μαρκοβιανή διαδικασία απόφασης στην άσκηση 3.

Παρατηρήστε ότι η τελευταία εξίσωση δίνει $V(3) = 0$, άρα αντικαθιστώντας στις υπόλοιπες λαμβάνουμε

$$\begin{cases} V(1) = \min\left\{1 + \frac{1}{2}\left[\frac{1}{3}V(1) + \frac{1}{3}V(2)\right], -1 + \frac{1}{2}V(2)\right\} \\ V(2) = \min\left\{-1 + \frac{1}{2}\left[\frac{1}{4}V(1) + \frac{1}{4}V(2)\right], 1 + \frac{1}{2}V(1)\right\} \end{cases} \tag{6}$$

οι οποίες είναι ισοδύναμες με το αρχικό σύστημα εξισώσεων με την ταύτιση $x = V(1), y = V(2)$.

Τώρα, από το Σχήμα 1 παρατηρούμε ότι η λήψη των αποφάσεων a_2 στην κατάσταση 1 και a_1 στην κατάσταση 2 φαίνεται να δημιουργεί μικρό υποτιμούμενο κόστος μιας και στις καταστάσεις 1,2 έχει άμεσο κόστος -1 και ποτέ δεν καταλογίζεται το ανώτατο κόστος 1. Ελέγχουμε εάν η πολιτική αυτή είναι βέλτιστη.

Το σύστημα που αντιστοιχεί σε αυτή την πολιτική είναι:

$$\begin{cases} V(1) = -1 + \frac{1}{2}V(2) \\ V(2) = -1 + \frac{1}{2}\left[\frac{1}{4}V(1) + \frac{1}{4}V(2)\right] \end{cases} \Leftrightarrow \begin{cases} V(1) = -\frac{22}{13} \\ V(2) = -\frac{18}{13} \end{cases}$$

Τώρα στην κατάσταση 1 έχουμε $1 + \frac{1}{2}\left[\frac{1}{3}V(1) + \frac{1}{3}V(2)\right] = \frac{19}{39} > -\frac{22}{13} = -1 + \frac{1}{2}V(2)$ και στην 2: $1 + \frac{1}{2}V(1) = \frac{2}{13} > -\frac{18}{13} = -1 + \frac{1}{2}\left[\frac{1}{4}V(1) + \frac{1}{4}V(2)\right]$. Άρα $x = V(1) = -\frac{22}{13}, y = V(2) = -\frac{18}{13}$ είναι η λύση των εξισώσεων.

4. **Ένας ληστής στην Άγρια Δύση κάθε ημέρα έχει την επιλογή είτε να ληστέψει την (ημερήσια) ταχυδρομική άμαξα από όπου θα πάρει λάφυρα αξίας 2, είτε να μην τη ληστέψει. Εάν ληστέψει τότε μπορεί να πιαστεί και να οδηγηθεί στη φυλακή είτε να μην πιαστεί με ίση πιθανότητα. (Εάν δε ληστέψει τότε δεν κινδυνεύει να πιαστεί.) Εάν οδηγηθεί φυλακή τότε θα μείνει εκεί για πάντα και κάθε ημέρα φυλάκισης του κοστίζει 1 μονάδα.**

(α′) **Δώστε μια Μαρκοβιανή διαδικασία απόφασης που να περιγράφει τα παραπάνω.**

Η διαδικασία απόφασης δίδεται στο Σχήμα 2, όπου οι καταστάσεις Ε, Φ αντιστοιχούν στον ληστή να είναι ελεύθερος ή στη φυλακή αντίστοιχα. Η απόφαση Λ αντιστοιχεί στη ληστεία και Δ στο να μην ληστέψει.

(α′) Απόφαση Λ(=Ληστεύει) (β′) Απόφαση Δ(=Δεν ληστεύει)

Σχήμα 2: Μαρκοβιανή διαδικασία απόφασης στην άσκηση 4.

(β′) **Βρείτε την πολιτική που ελαχιστοποιεί το μέσο κόστος που συσσωρεύεται στη διάρκεια 4 ημερών.**

Η εξίσωση δυναμικού προγραμματισμού για αυτό το κριτήριο είναι:

$$\begin{cases} V_{n+1}(\mathrm{E}) = \min\left\{-2 + \frac{1}{2}V_n(\mathrm{E}) + \frac{1}{2}V_n(\Phi), V_n(\mathrm{E})\right\} \\ V_{n+1}(\Phi) = 1 + V_n(\Phi) \end{cases} \tag{7}$$

για $n = 0, 1, \ldots$, με $V_0(\mathrm{E}) = \min\{-2, 0\}, V_0(\Phi) = 1$.

Παρατηρήστε ότι στην κατάσταση Φ και οι δύο αποφάσεις είναι εξίσου βέλτιστες και $V_n(\Phi) = n+1$, από τη 2η εξίσωση.

Στο 4ο βήμα (0 βήματα πριν το τέλος του χρονικού ορίζοντα) στην κατάσταση Ε η είναι βέλτιστο να ληφθεί η απόφαση Λ.

Στο 3ο βήμα (1 βήμα πριν το τέλος) στην κατάσταση Ε έχουμε $-2 + \frac{1}{2}V_0(\mathrm{E}) + \frac{1}{2} = -\frac{5}{2} < -2 = V_0(\mathrm{E})$, άρα η βέλτιστη απόφαση είναι Λ και $V_1(\mathrm{E}) = -\frac{5}{2}$.

Στο 2ο βήμα (2 βήματα πριν το τέλος) στην κατάσταση Ε έχουμε $-2 + \frac{1}{2}V_1(\mathrm{E}) + 1 = -\frac{9}{8} > -\frac{5}{2} = V_0(\mathrm{E})$, άρα η βέλτιστη απόφαση είναι Δ και $V_1(\mathrm{E}) = -\frac{5}{2}$.

Στο 1ο βήμα (3 βήματα πριν το τέλος) έχουμε $-2 + \frac{1}{2}V_2(\mathrm{E}) + \frac{3}{2} = -\frac{7}{4} > -\frac{5}{2} = V_2(\mathrm{E})$, άρα η βέλτιστη απόφαση είναι Δ.

Συνοψίζοντας, η βέλτιστη πολιτική είναι ο ληστής να δοκιμάσει να ληστέψει και στις δύο τελευταίες ημέρες.

(γ′) **Θεωρήστε ότι ο ληστής δίνει περισσότερο βάρος στο σήμερα από ότι στο αύριο με την εξής έννοια: ένα κόστος ή λάφυρο το οποίο αποκτάται την επόμενη ημέρα υποτιμάται κατά ένα συντελεστή $\beta \in (0,1)$, σε σχέση με αν αυτό λαμβάνονταν σήμερα. Δείξτε ότι η πολιτική που ελαχιστοποιεί το μέσο συσσωρευμένο υποτιμούμενο κόστος είναι ο ληστής να μη ληστεύει καθόλου εάν $\beta \geq 0.8$, ενώ εάν $\beta \leq 0.8$ η βέλτιστη πολιτική είναι πάντα να ληστεύει.**

Η εξίσωση δυναμικού προγραμματισμού για το κριτήριο του υποτιμούμενου κόστους με συντελεστή υποτίμησης β είναι:

$$\begin{cases} V(\mathrm{E}) = \min\left\{-2+\beta\left[\frac{1}{2}V(\mathrm{E})+V(\Phi)\right], \beta V(\mathrm{E})\right\} \\ V(\Phi) = 1+\beta V(\Phi) \end{cases}$$

$$\Leftrightarrow \begin{cases} V(\mathrm{E}) = \min\left\{\frac{5\beta-4}{2(1-\beta)}+\frac{\beta}{2}V(\mathrm{E}), \beta V(\mathrm{E})\right\} \\ V(\Phi) = \dfrac{1}{1-\beta} \end{cases} \tag{8}$$

Η βέλτιστη απόφαση (στην Ε) είναι να ληστέψει εάν και μόνο εάν

$$V(\mathrm{E}) = \frac{5\beta-4}{2(1-\beta)}+\frac{\beta}{2}V(\mathrm{E}) \text{ και } V(\mathrm{E}) \leq \beta V(\mathrm{E})$$
$$\Leftrightarrow V(\mathrm{E}) = \frac{5\beta-4}{(1-\beta)(2-\beta)} \text{ και } V(\mathrm{E}) \leq 0$$
$$\Leftrightarrow \frac{5\beta-4}{(1-\beta)(2-\beta)} \leq 0 \Leftrightarrow \beta \leq 0.8\,.$$