Στοχαστικές Στρατηγικές

Τμήμα Μαθηματικών, ΑΠΘ

3η ενότητα: Στοχαστικά προβλήματα διαδρομής – Μεθοδολογία (1)

Παπάνα Αγγελική

Μεταδιδακτορική ερευνήτρια, ΑΠΘ & Πανεπιστήμιο Μακεδονίας
E-mail: angeliki.papana@gmail.com, agpapana@auth.gr
Webpage: http://users.auth.gr/agpapana

Μέθοδοι ελέγχου

Έστω ότι έχουμε ένα στοχαστικό πρόβλημα διαδρομής με την έννοια ότι σε κάθε κόμβο υπάρχει αβεβαιότητα για το γεγονός αν ακολουθείται η βέλτιστη διαδρομή παρά την γνώση ποια είναι αυτή.

Δηλαδή σε κάθε κόμβο υπάρχει μια πιθανότητα $\boldsymbol{p(x,y)}$ να ακολουθήσουμε την εντολή που μας δόθηκε ως προς τη διαδρομή που πρέπει να ακολουθήσουμε και μια πιθανότητα $\boldsymbol{q(x,y) = 1 - p(x,y)}$ να μην την ακολουθήσουμε.

Αυτό το γεγονός διαφοροποιεί σε πολλά σημεία και τα κριτήρια με τα οποία θα δοθούν οι λύσεις στο πρόβλημα.

π.χ. Έστω ότι θεωρούμε σαν κριτήριο **την ελαχιστοποίηση του αναμενόμενου κόστους**. Υπάρχουν πολλές πολιτικές για να πραγματοποιηθεί αυτό.

1η πολιτική: Μέθοδος του βέλτιστου με επαναπληροφόρηση ελέγχου

Σε κάθε κόμβο που είναι δεδομένο ότι βρισκόμαστε, βρίσκουμε ποια είναι η διαδρομή με το ελάχιστο αναμενόμενο κόστος. Την στιγμή της απόφασης έχουμε την επιπλέον πληροφορία σε ποιον ακριβώς κόμβο βρισκόμαστε.

2η πολιτική: Μέθοδος ελέγχου ανοιχτού κύκλου

Βρίσκουμε όλες τις δυνατές διαδρομές και επιλέγουμε εκείνη με το ελάχιστο αναμενόμενο κόστος.

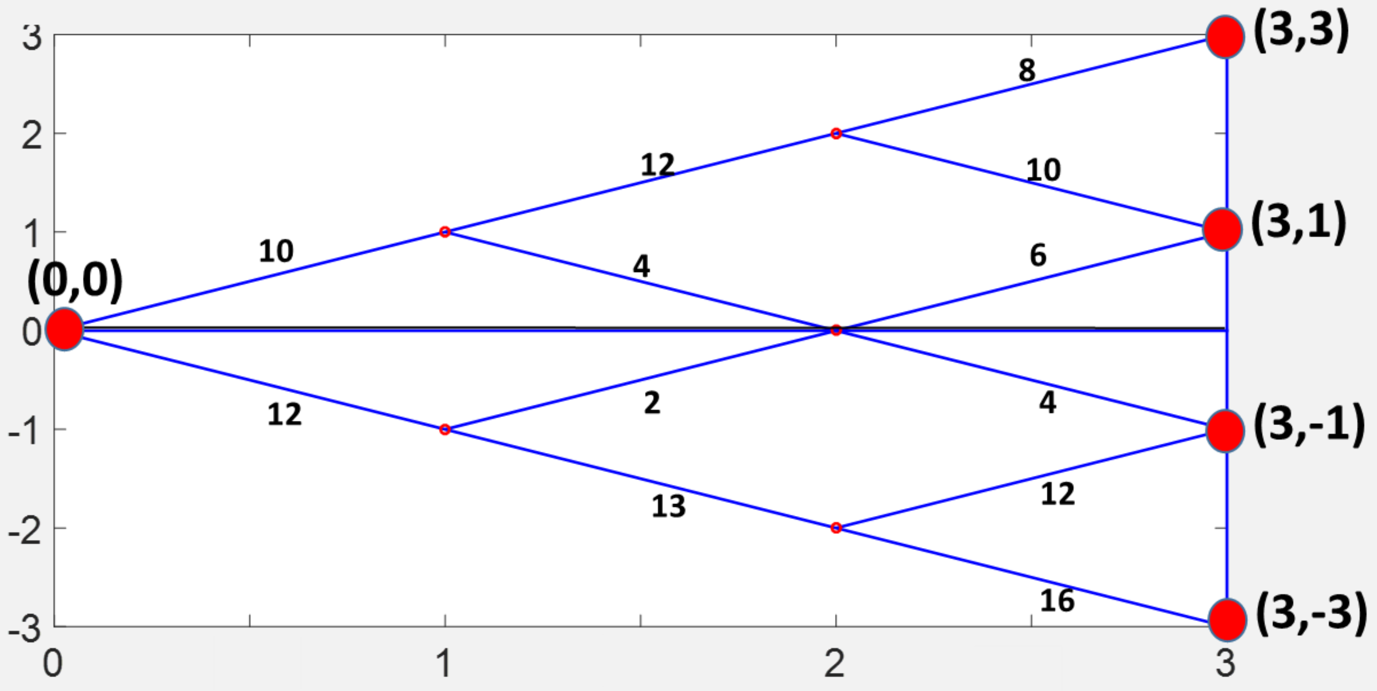
3η πολιτική: Μέθοδος ελέγχου ανοιχτού κύκλου με επαναπληροφόρηση

Είναι συνδυασμός των δυο πρώτων μεθόδων. Σε κάθε κόμβο, προσδιορίζουμε την βέλτιστη πολιτική με την μέθοδο ελέγχου ανοιχτού κύκλου.

Μέθοδος του βέλτιστου με επαναπληροφόρηση ελέγχου

Παράδειγμα

Να βρεθεί η βέλτιστη πολιτική ώστε να πάμε από το σημείο (0,0) σε οποιοδήποτε σημείο πάνω στην ευθεία Β ($x = 3$). Θεωρούμε ότι σε κάθε κόμβο υπάρχει πιθανότητα $p(x, y)$ να ακολουθήσει κάποιος τις οδηγίες μας και μια πιθανότητα $q(x, y)$ να ακολουθήσει την αντίθετη διαδρομή.



Το πρόβλημα

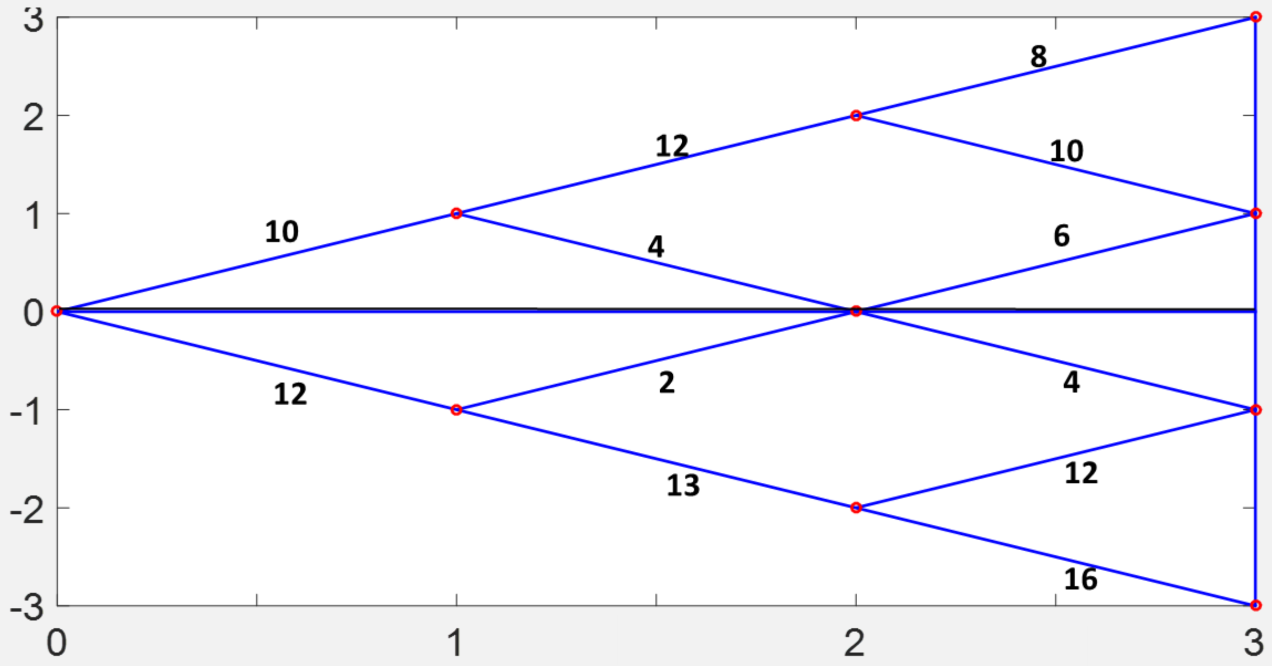
Ποιες πρέπει να είναι οι οδηγίες μας ώστε το αναμενόμενο κόστος να είναι ελάχιστο. Τα $p(x,y)$ και $q(x,y)$ δίνονται.

$p(x,y)$

y \ x	0	1	2
-2	–	–	¼
-1	–	¾	–
0	¾	–	½
1	–	½	–
2	–	–	¾

$q(x,y)$

y \ x	0	1	2
-2	–	–	¾
-1	–	¼	–
0	¼	–	½
1	–	½	–
2	–	–	¼



Μέθοδος του βέλτιστου με επαναπληροφόρηση ελέγχου

Η **βέλτιστη συνάρτηση** ορίζεται ως:

$$f(x,y) = \{\text{Το ελάχιστο αναμενόμενο κόστος από τον κόμβο } (x,y) \text{ μέχρι την ευθεία Β}\}$$

Έστω $\boldsymbol{a}\ (\boldsymbol{x},\boldsymbol{y})$: απόσταση **κόμβου** $(\boldsymbol{x},\boldsymbol{y})$ και $(\boldsymbol{x}+\mathbf{1},\boldsymbol{y}+\mathbf{1})$,

$\boldsymbol{\delta}\ (\boldsymbol{x},\boldsymbol{y})$: απόσταση **κόμβου** $(\boldsymbol{x},\boldsymbol{y})$ και $(\boldsymbol{x}+\mathbf{1},\boldsymbol{y}-\mathbf{1})$

Αν βρισκόμαστε στον **κόμβο** $(\boldsymbol{x},\boldsymbol{y})$ και δώσουμε οδηγία ότι πρέπει να πάμε στον επόμενο κόμβο, δηλ. στον **κόμβο** $(\boldsymbol{x}+\mathbf{1},\boldsymbol{y}+\mathbf{1})$ ή στον **κόμβο** $(\boldsymbol{x}+\mathbf{1},\boldsymbol{y}-\mathbf{1})$, τότε το **αναμενόμενο κόστος** αυτής της απόφασης είναι διαφορετικό και προκύπτει με βάση τον ορισμό της μέσης τιμής.

Εντολή κίνησης προς αριστερά

Αν βρισκόμαστε στον **κόμβο (x, y)** και δώσουμε οδηγία ότι πρέπει να πάμε στον **κόμβο $(x+1, y+1)$**, τότε το **αναμενόμενο κόστος** αυτής της απόφασης είναι:

κίνηση προς τα αριστερά — κίνηση προς τα δεξιά

$$\boxed{p(x,y)\,\{a\,(x,y)\ + f(x+1, y+1)\} + q(x,y)\,\{\delta\,(x,y)\ + f(x+1, y-1)\}}$$

Εντολή κίνησης προς δεξιά

Αν βρισκόμαστε στον **κόμβο (x, y)** και δώσουμε οδηγία ότι πρέπει να πάμε στον **κόμβο $(x+1, y-1)$**, τότε το **αναμενόμενο κόστος** αυτής της απόφασης είναι:

κίνηση προς τα δεξιά — κίνηση προς τα αριστερά

$$\boxed{p(x,y)\,\{\delta\,(x,y)\ + f(x+1, y-1)\} + q(x,y)\,\{\alpha\,(x,y)\ + f(x+1, y+1)\}}$$

Σύμφωνα με την **αρχή της βελτιστοποίησης**, οι οδηγίες που πρέπει να δώσουμε στον **κόμβο (x, y)** πρέπει να είναι τέτοιες ώστε το αναμενόμενο κόστος να είναι ελάχιστο.

Επομένως:

$$f(x,y) = \min \left\{ \begin{array}{c} p(x,y)[a(x,y) + f(x+1,y+1)] + q(x,y)[\delta(x,y) + f(x+1,y-1)] \\ p(x,y)[\delta(x,y) + f(x+1,y-1)] + q(x,y)[\alpha(x,y) + f(x+1,y+1)] \end{array} \right\}$$

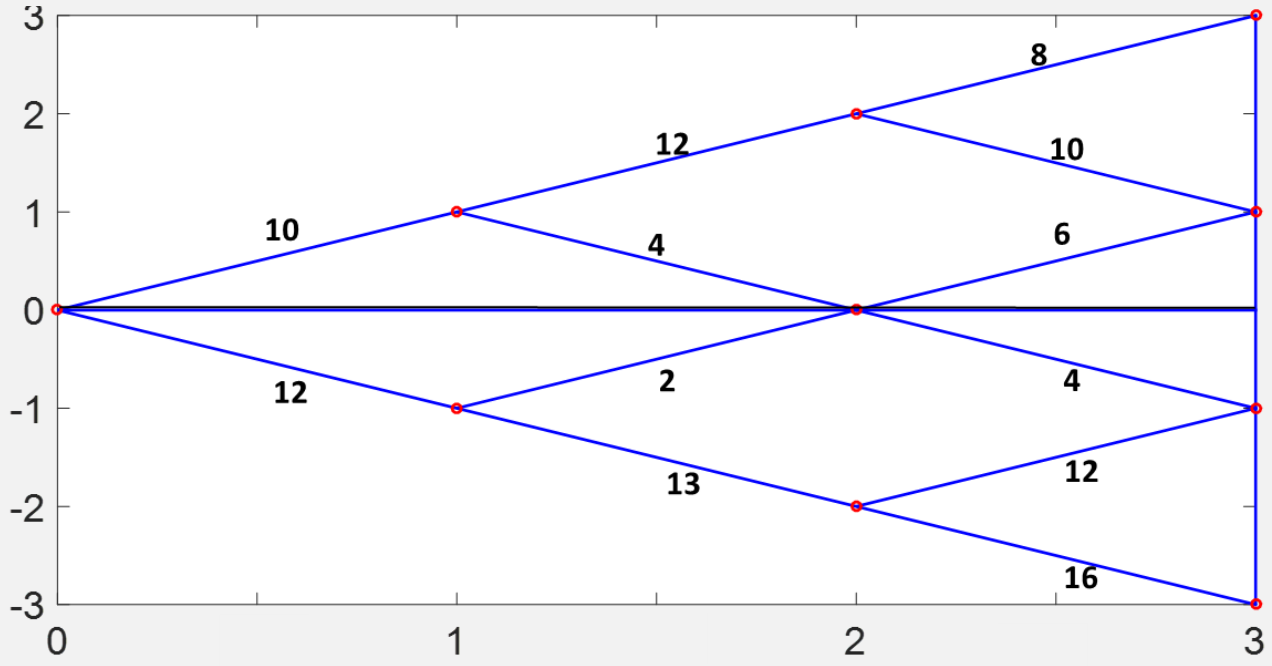
Οριακές συνθήκες

$$f(3,3) = 0$$

$$f(3,1) = 0$$

$$f(3,-1) = 0$$

$$f(3,-3) = 0$$



Τώρα θα υπολογίσουμε την τιμή της βέλτιστης αναμενόμενης συνάρτησης για όλους τους κόμβους.

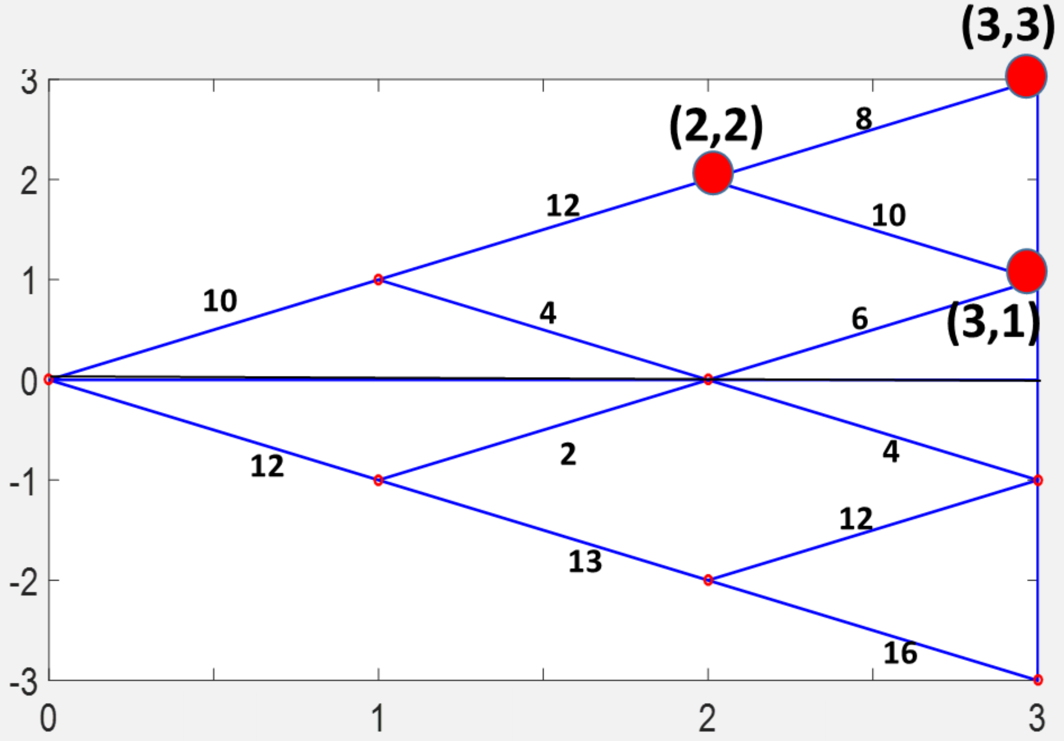
- Για $\boldsymbol{x = 2}$

$$f(2,2) = \min \begin{Bmatrix} p(2,2)\,(\alpha\,(2,2) + f(3,3)) \; + \; q(2,2)\,(\delta\,(2,2) + f(3,1)) \\ p(2,2)\,(\delta\,(2,2) + f(3,1)) \; + \; q(2,2)\,(\alpha\,(2,2) + f(3,3)) \end{Bmatrix}$$

$$= \min \begin{Bmatrix} \frac{3}{4}\{8 \; + 0\} + \frac{1}{4}\{10 \; + 0\} \\ \frac{3}{4}\{10 + 0\} + \frac{1}{4}\{8 + 0\} \end{Bmatrix} = \min\left\{\frac{17}{2}, \frac{19}{2}\right\} = \frac{17}{2}$$

$\boldsymbol{p(x,y)}$

y \ x	0	1	2
-2	-	-	¼
-1	-	¾	-
0	¾	-	½
1	-	½	-
2	-	-	¾



- Για $\boldsymbol{x = 2}$

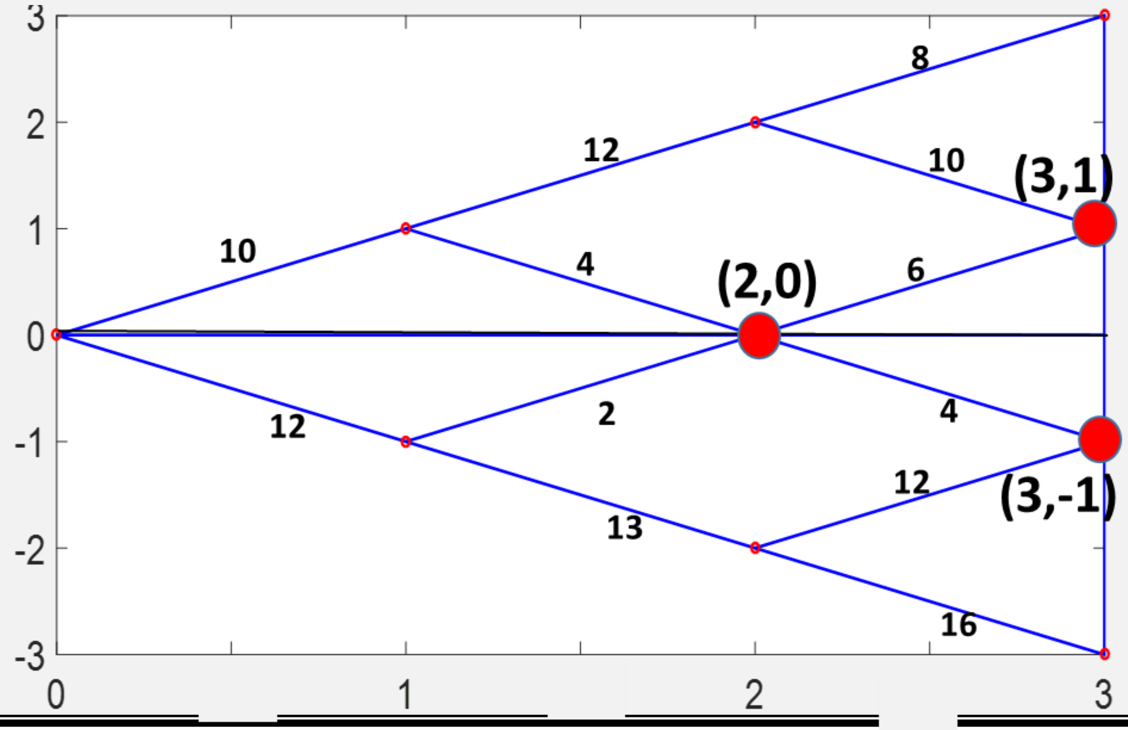
$$f(2,0) = \min \begin{Bmatrix} p(2,0)\,(\alpha\,(2,0) + f(3,1)) \quad + \; q(2,0)\,(\delta\,(2,0) + f(3,-1)) \\ p(2,0)\,(\delta\,(2,0) + f(3,-1)) + q(2,0)\,(\alpha\,(2,0) + f(3,1)) \end{Bmatrix}$$

$$= \min\{\frac{1}{2}\{6 \; + 0\} + \frac{1}{2}\{4 \; + 0\}, \frac{1}{2}\{4 + 0\} + \frac{1}{2}\{6 + 0\}\}$$

$$= \min\{5{,}5\} = 5$$

$\boldsymbol{p(x,y)}$

y \ x	0	1	2
-2	-	-	¼
-1	-	¾	-
0	¾	-	½
1	-	½	-
2	-	-	¾



- Για $\boldsymbol{x} = \mathbf{2}$

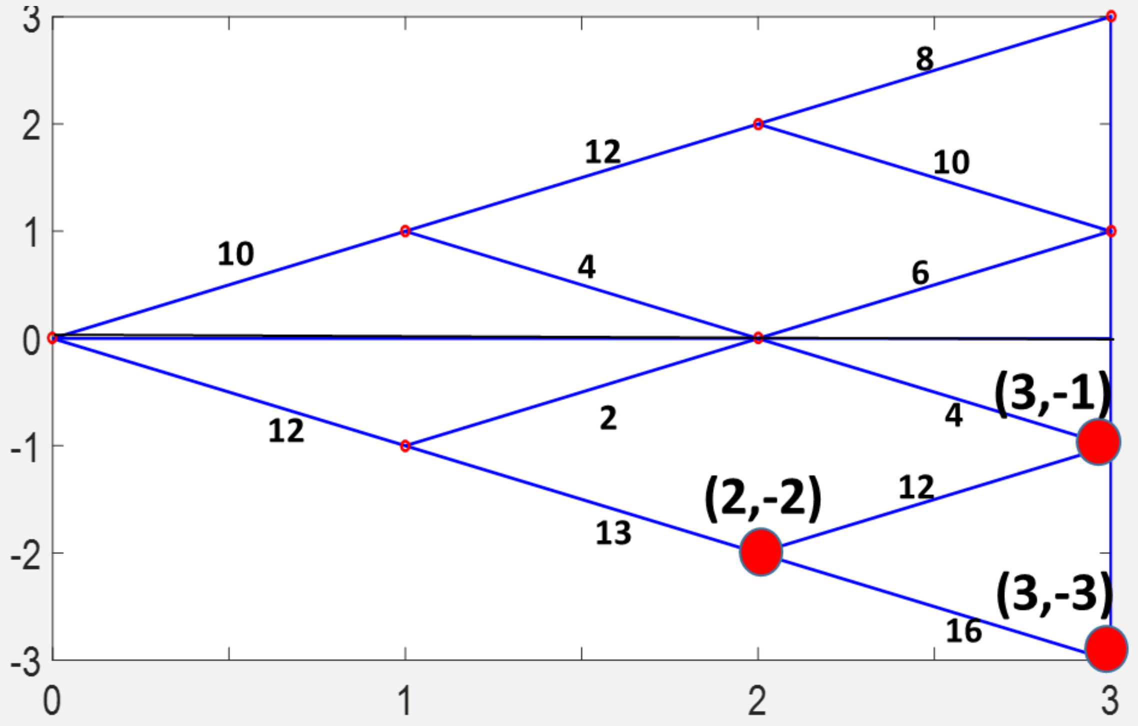
$$f(2,-2) = \min\begin{Bmatrix} p(2,-2)(a\,(2,-2) + f(3,-1)) \; + \; q(2,-2)(\delta\,(2,-2) + f(3,-3)) \\ p(2,-2)(\delta\,(2,-2) + f(3,-3)) + q(2,-2)\alpha\,(2,-2) + f(3,-1) \end{Bmatrix}$$

$$= min\{\frac{1}{4}\{12 \; + 0\} + \frac{3}{4}\{16 \; + 0\}, \frac{1}{4}\{16 + 0\} + \frac{3}{4}\{12 + 0\}\}$$

$$= min\{15, 13\} = 13$$

$\boldsymbol{p(x,y)}$

y \ x	0	1	2
-2	-	-	¼
-1	-	¾	-
0	¾	-	½
1	-	½	-
2	-	-	¾



- Για $x = 1$

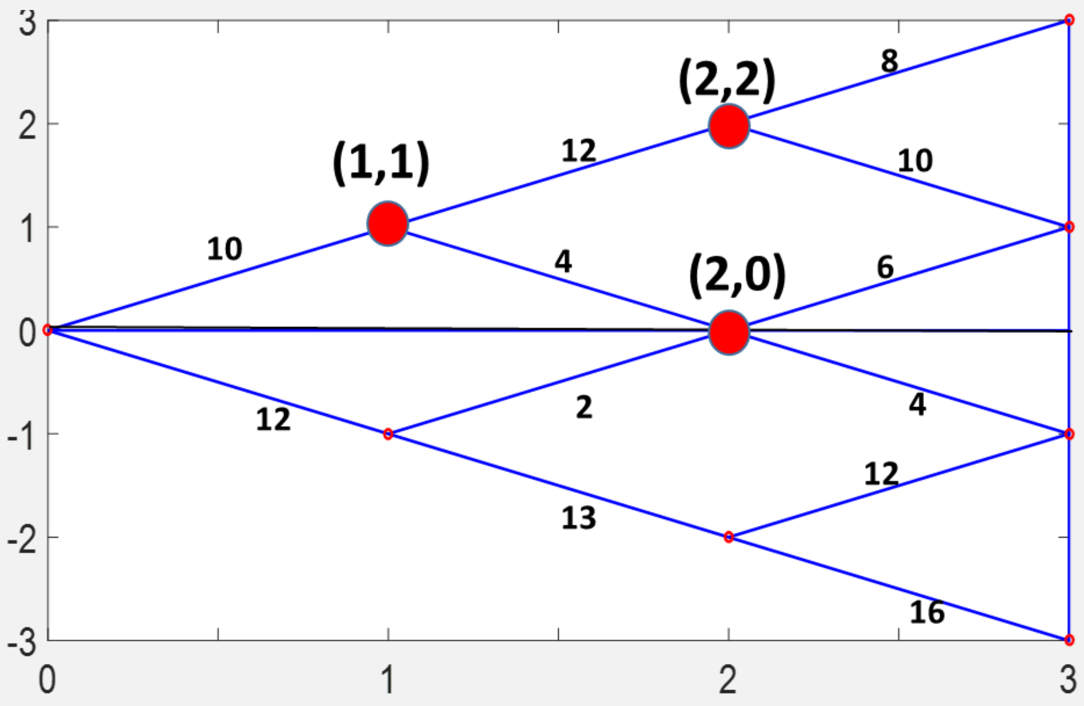
$$f(1,1) = \min \begin{Bmatrix} p(1,1)\,(\alpha\,(1,1) + f(2,2)) + q(1,1)\,(\delta\,(1,1) + f(2,0)) \\ p(1,1)(\delta\,(1,1) + f(2,0)) + q(1,1)(\alpha\,(1,1) + f(2,2)) \end{Bmatrix}$$

$$= min\{\frac{1}{2}\left\{12 \ + \frac{17}{2}\right\} + \frac{1}{2}\{4 \ + 5\}, \frac{1}{2}\{4 + 5\} + \frac{1}{2}\left\{12 + \frac{17}{2}\right\}\}$$

$$= min\left\{\frac{59}{4}, \frac{59}{4}\right\} = \frac{59}{4}$$

$p(x,y)$

y \ x	0	1	2
-2	-	-	¼
-1	-	¾	-
0	¾	-	½
1	-	½	-
2	-	-	¾



- Για $\boldsymbol{x = 1}$

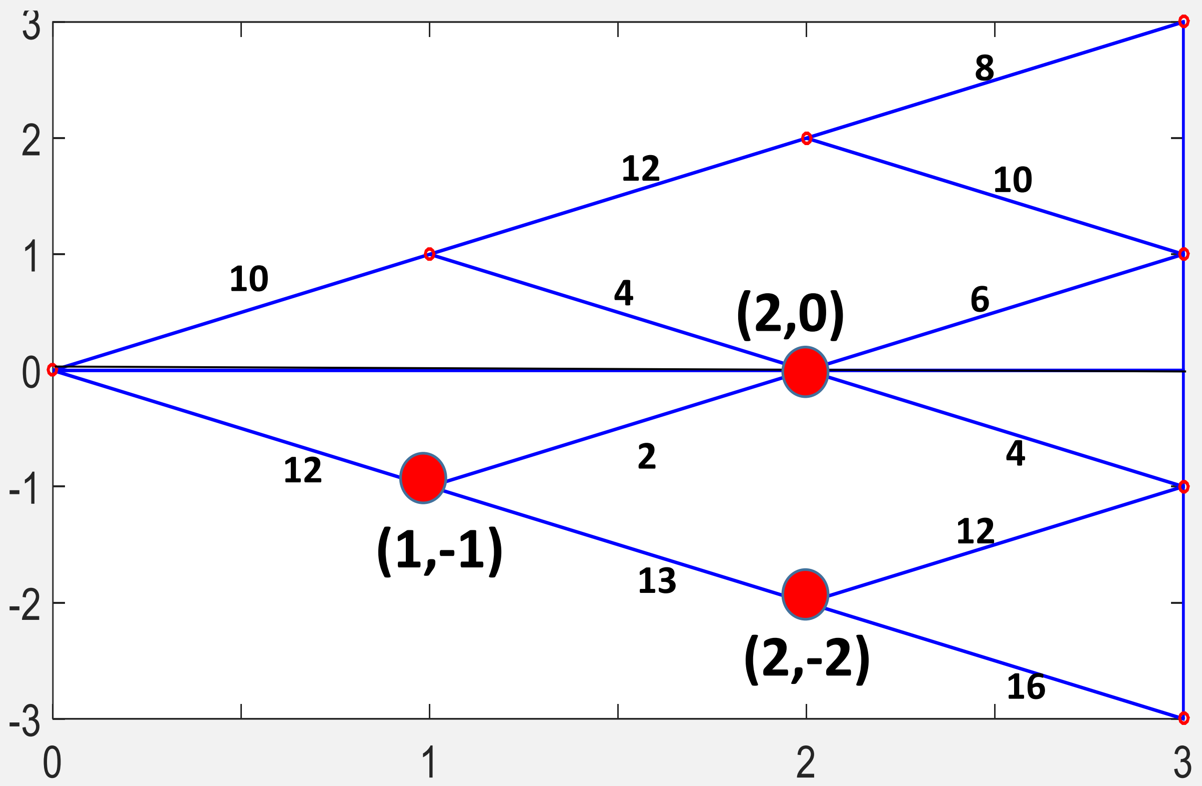
$$f(1,-1) = \min\begin{Bmatrix} p(1,-1)\,(\alpha\,(1,-1) + f(2{,}0)) + q(1,-1)\,(\delta\,(1,-1) + f(2,-2)) \\ p(1,-1)(\delta\,(1,-1) + f(2,-2)) + q(1,-1)(\alpha\,(1,-1) + f(2,-0)) \end{Bmatrix}$$

$$= min\{\frac{3}{4}\{2\ +5\} + \frac{1}{4}\{13\ +13\}, \frac{3}{4}\{13+13\} + \frac{1}{4}\{2+5\}\}$$

$$= min\left\{\frac{47}{4}, \frac{85}{4}\right\} = \frac{47}{4}$$

$p(x,y)$

y \ x	0	1	2
-2	-	-	¼
-1	-	¾	-
0	¾	-	½
1	-	½	-
2	-	-	¾



- Για $\boldsymbol{x = 0}$

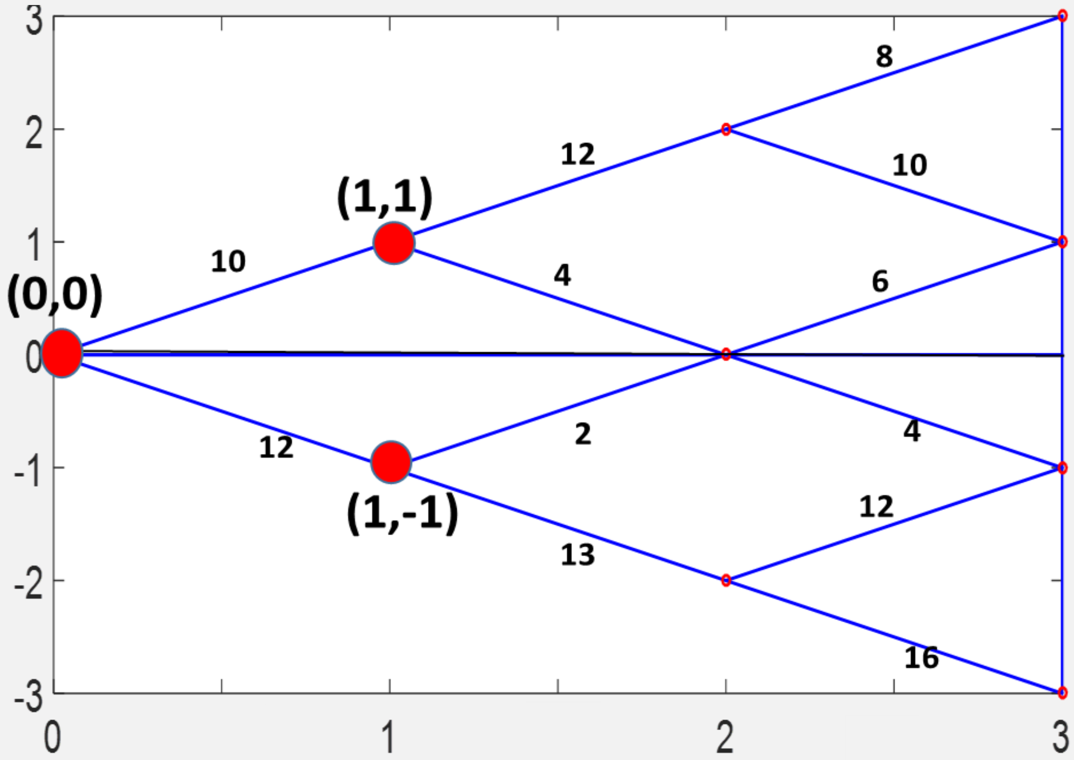
$$f(0,0) = \min\begin{Bmatrix} p(0,0)\,(a\,(0,0) + f(1,1)) + q(0,0)\,(\delta\,(0,0) + f(1,-1)) \\ p(0,0)(\delta\,(0,0) + f(1,-1) + q(0,0)(\alpha\,(0,0) + f(1,1)) \end{Bmatrix}$$

$$= min\{\frac{3}{4}\left\{10 + \frac{59}{4}\right\} + \frac{1}{4}\left\{12 + \frac{47}{4}\right\}, \frac{3}{4}\left\{12 + \frac{47}{4}\right\} + \frac{1}{4}\left\{10 + \frac{59}{4}\right\}\}$$

$$= min\left\{\frac{49}{2}, 24\right\} = 24$$

$\boldsymbol{p(x,y)}$

y \ x	0	1	2
-2	-	-	¼
-1	-	¾	-
0	¾	-	½
1	-	½	-
2	-	-	¾



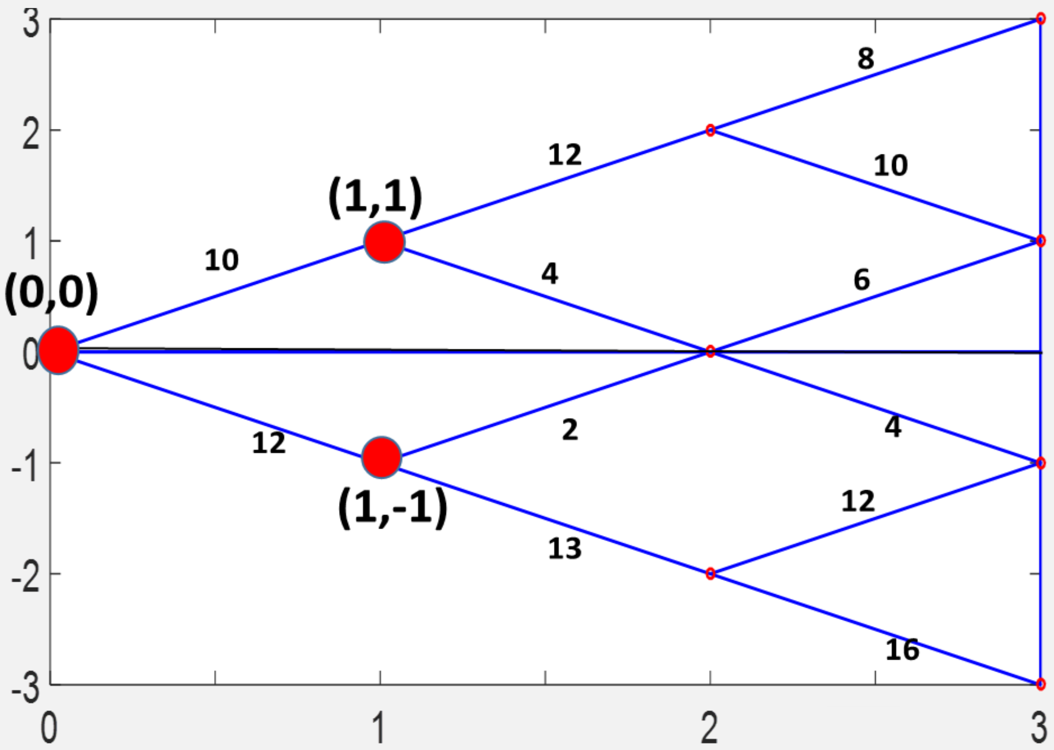
Βέλτιστη πολιτική εντολών

Κόμβος $(0,0)$

$$f(0,0) = \min\left\{\begin{matrix}\text{εντολή κίνησης: αριστερά} \\ \text{εντολή κίνησης: δεξιά}\end{matrix}\right\} = min\left\{\frac{49}{2}, 24\right\} = 24$$

Επομένως η βέλτιστη εντολή όταν βρισκόμαστε στον κόμβο $(0,0)$ είναι:

κίνηση προς τα **δεξιά**

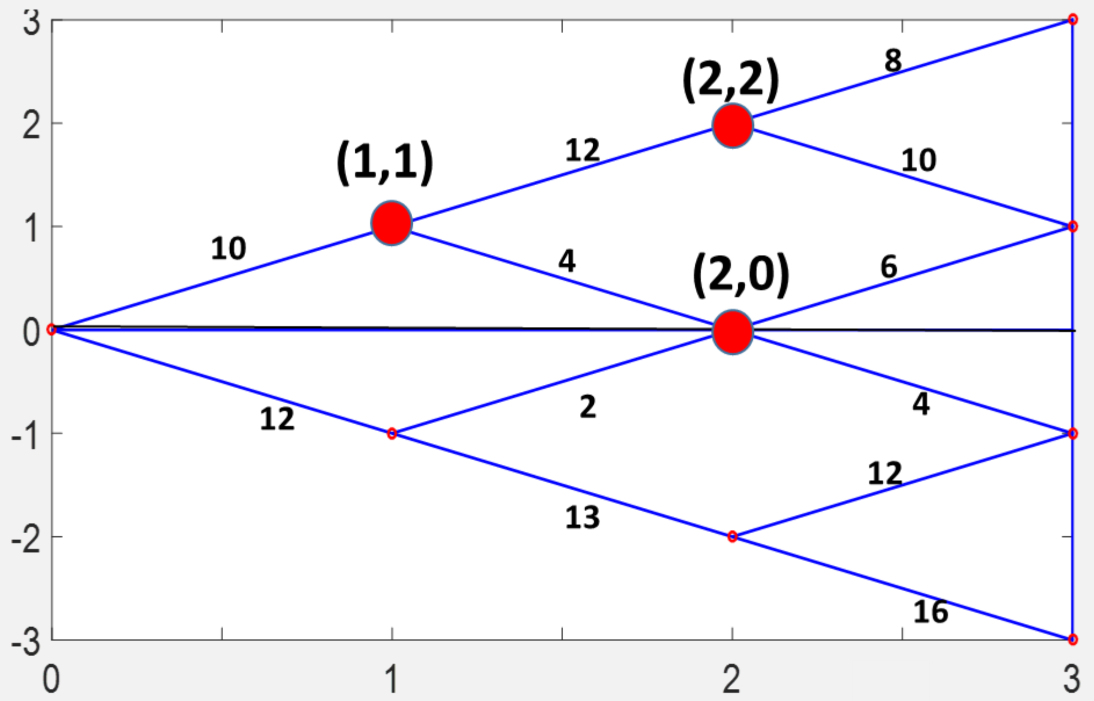


Κόμβος $(1,1)$

$$f(1,1) = \min\left\{\begin{matrix}\text{εντολή κίνησης: αριστερά} \\ \text{εντολή κίνησης: δεξιά}\end{matrix}\right\} = min\left\{\frac{59}{4}, \frac{59}{4}\right\} = \frac{59}{4}$$

Επομένως η βέλτιστη εντολή όταν βρισκόμαστε στον κόμβο $(1,1)$ είναι:

κίνηση προς τα **αριστερά** ή **δεξιά**

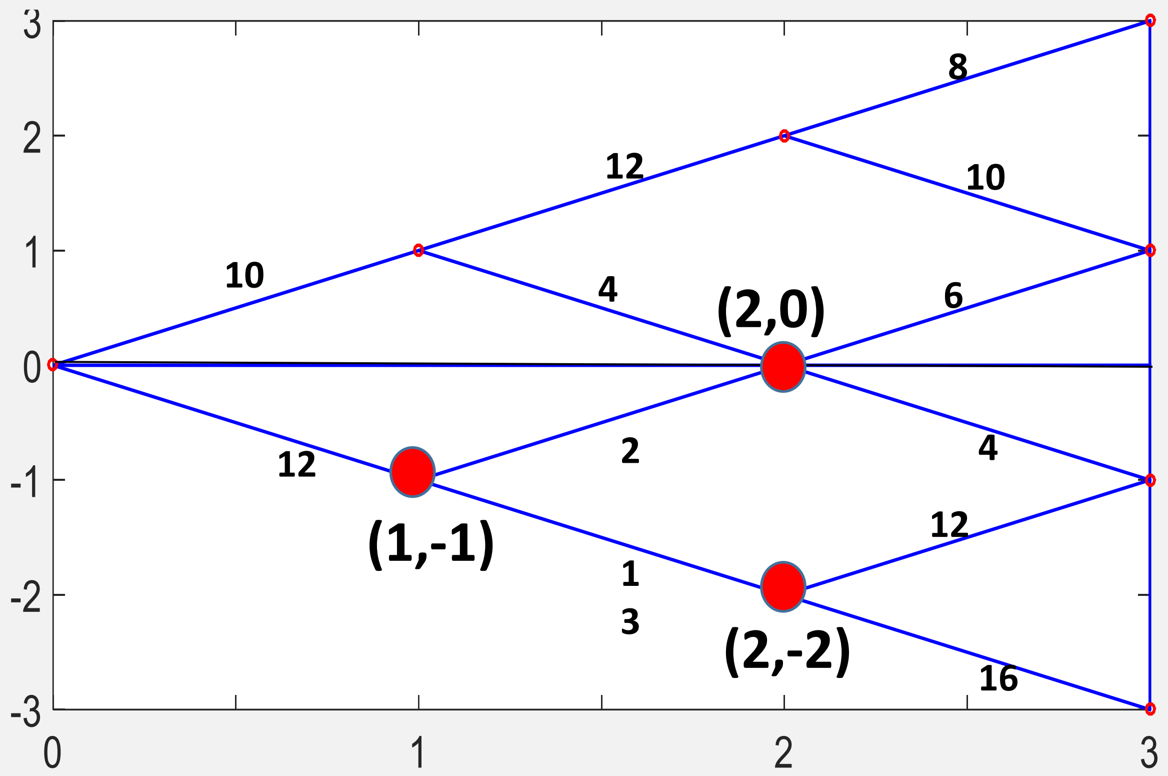


Κόμβος $(1, -1)$

$$f(1,-1) = \min\left\{\begin{matrix}\text{εντολή κίνησης: αριστερά} \\ \text{εντολή κίνησης: δεξιά}\end{matrix}\right\} = min\left\{\frac{47}{4}, \frac{85}{4}\right\} = \frac{47}{4}$$

Επομένως η βέλτιστη εντολή όταν βρισκόμαστε στον κόμβο $(1, -1)$ είναι:

κίνηση προς τα **αριστερά**



Βέλτιστη πολιτική εντολών

$(\mathbf{0}, \mathbf{0})$ $\rightarrow$ εντολή κίνησης προς αριστερά

$(\mathbf{1}, \mathbf{1})$ $\rightarrow$ εντολή κίνησης προς αριστερά ή δεξιά

$(\mathbf{1}, -\mathbf{1})$ $\rightarrow$ εντολή κίνησης προς αριστερά

$(\mathbf{2}, \mathbf{2})$ $\rightarrow$ εντολή κίνησης προς αριστερά

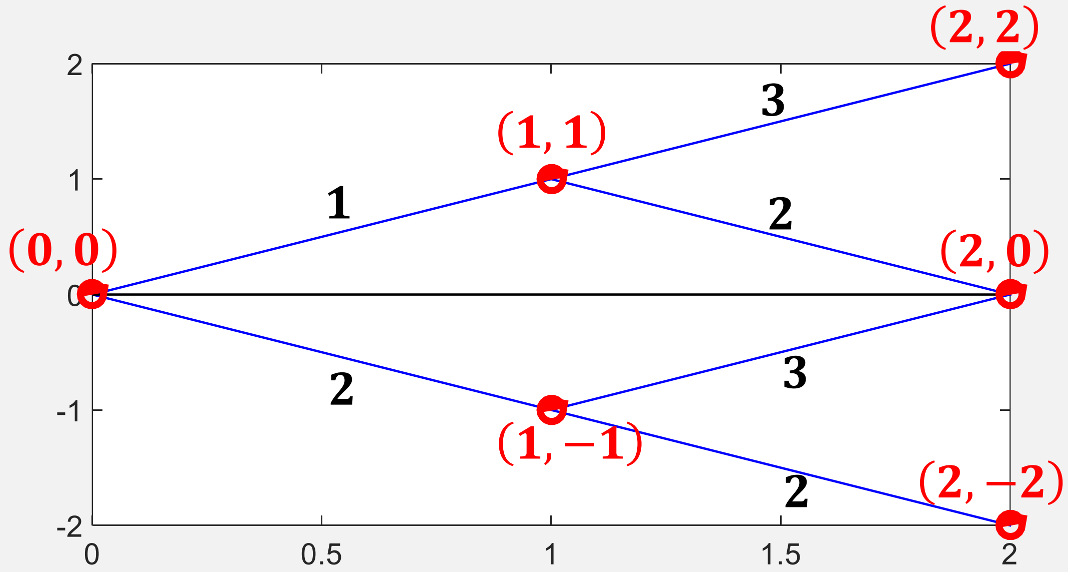
$(\mathbf{2}, \mathbf{0})$ $\rightarrow$ εντολή κίνησης προς αριστερά ή δεξιά

$(\mathbf{2}, -\mathbf{2})$ $\rightarrow$ εντολή κίνησης προς δεξιά

Άσκηση

Δίνουμε από ένα κέντρο διοίκησης εντολές σε ένα κινητό. Η πιθανότητα να γίνει **δεκτή η οδηγία κίνησης** είναι p. Αν **δεν γίνει δεκτή η οδηγία**, με πιθανότητα $1-p$, μπορεί να ακολουθήσει την **αντίθετη διαδρομή** με **πιθανότητα q** ή να **παραμείνει στην ίδια θέση** με **πιθανότητα $1-q$** και **κόστος μ**.

Να λυθεί το πρόβλημα με την μέθοδο του βέλτιστου με επαναπληροφόρηση ελέγχου για $p=\frac{2}{3}$, $q=\frac{1}{4}$ και $\mu=3$.



Ζητείται η **βέλτιστη πολιτική εντολών**.

Εφαρμόζουμε την προς τα πίσω μέθοδο.

Βέλτιστη συνάρτηση

$$f(x,y) = \{\text{Το ελάχιστο αναμενόμενο κόστος από τον κόμβο } (x,y) \text{ μέχρι το τέλος}\}$$

Επαναληπτική σχέση

$$f(x,y) = \min\left\{\begin{matrix}\text{εντολή κίνησης: αριστερά} \\ \text{εντολή κίνησης: δεξιά}\end{matrix}\right\}$$

Ελάχιστο κόστος για εντολή κίνησης προς αριστερά

Η εντολή γίνεται δεκτή με πιθανότητα p

$$\boldsymbol{p} \ \ (a\,(x,y) \ + f(x+1,y+1)) \ \ +$$

Η εντολή δεν γίνεται δεκτή, με πιθανότητα 1-p

Κίνηση δεξιά, με πιθανότητα q

Μένει στην ίδια θέση, με πιθανότητα 1-q

$$(\mathbf{1}-\boldsymbol{p})\{\,\boldsymbol{q}\,\big(\delta\,(x,y) \ + f(x+1,y-1)\big) \ + \ (\mathbf{1}-\boldsymbol{q})\,(f\,(x,y)+\boldsymbol{\mu}) \ \}$$

Ελάχιστο κόστος για εντολή κίνησης προς δεξιά

Η εντολή γίνεται δεκτή με πιθανότητα p

$$\boldsymbol{p} \ \ (\delta\,(x,y)\ + f(x+1, y-1)) \ \ +$$

Η εντολή δεν γίνεται δεκτή, με πιθανότητα 1-p

Κίνηση αριστερά, με πιθανότητα q

Μένει στην ίδια θέση, με πιθανότητα 1-q

$$(\mathbf{1}-\boldsymbol{p})\{\,\boldsymbol{q}\,\big(\alpha\,(x,y)\ + f(x+1, y+1)\big)\ +\ (\mathbf{1}-\boldsymbol{q})\,(f\,(x,y)+\boldsymbol{\mu})\ \ \}$$

Επαναληπτική σχέση

$$f(x,y) =$$

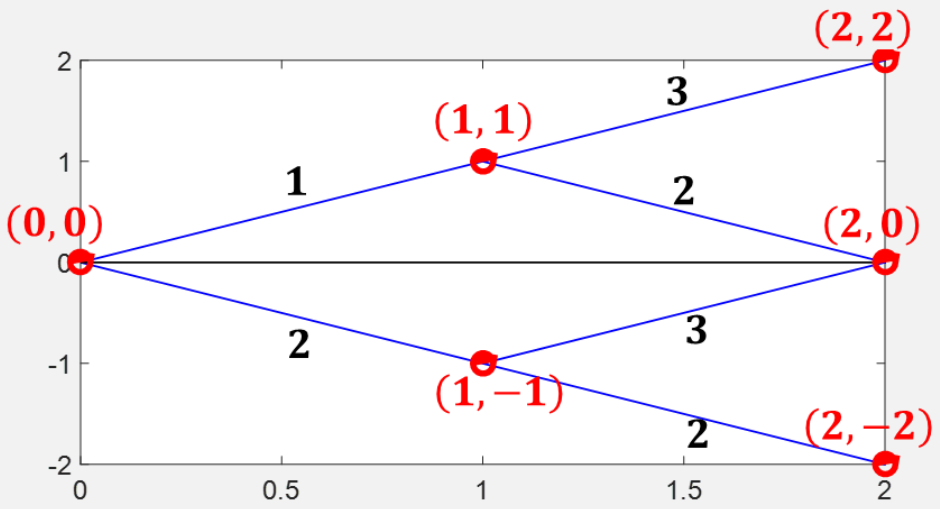
$$\min\begin{cases} p\{a\,(x,y) \; + f(x+1,y+1)\} + (1-p)\left[q\big(\delta\,(x,y) \; + f(x+1,y-1)\big) + (1-q)(f(x,y)+\mu)\right], \\ p\{\delta\,(x,y) \; + f(x+1,y-1)\} + (1-p)\left[q\big(\alpha\,(x,y) \; + f(x+1,y+1)\big) + (1-q)(f(x,y)+\mu)\right] \end{cases}$$

Οριακές συνθήκες

$$f(2,2) = 0$$

$$f(2,0) = 0$$

$$f(2,-2) = 0$$



- Για $x = 1$

$p = \frac{2}{3}, q = \frac{1}{4}, \mu = 3$

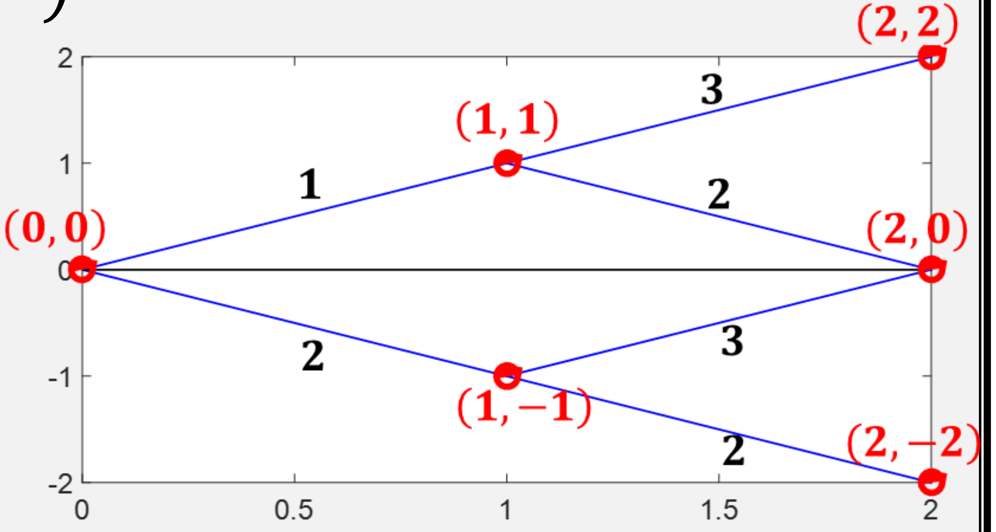
$f(1,1) =$

$$= \min \begin{Bmatrix} p(a(1,1) + f(2,2)) + (1-p)[q\big(\delta\,(1,1) + f(2,0)\big) + (1-q)(f(1,1)+\mu)] \\ p(\delta(1,1) + f(2,0)) + (1-p)[q\big(\alpha\,(1,1) + f(2,2)\big) + (1-q)(f(1,1)+\mu)] \end{Bmatrix}$$

$$= min \begin{Bmatrix} \frac{2}{3}(3+0) + \frac{1}{3}[\frac{1}{4}(2+0) + \frac{3}{4}(f(1,1)+3)] \\ \frac{2}{3}(2+0) + \frac{1}{3}[\frac{1}{4}(3+0) + \frac{3}{4}(f(1,1)+3)] \end{Bmatrix}$$

$$= min \begin{Bmatrix} \frac{35}{12} + \frac{f(1,1)}{4} \\ \frac{28}{12} + \frac{f(1,1)}{4} \end{Bmatrix}$$

$$= \frac{28}{12} + \frac{f(1,1)}{4}$$

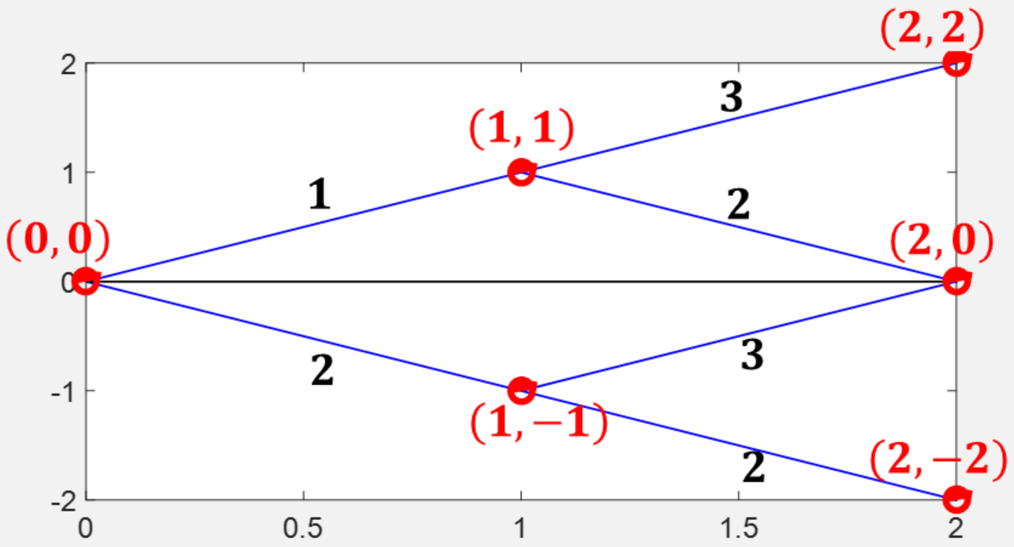


Επομένως βρήκαμε ότι:

$$f(1,1) = \frac{28}{12} + \frac{f(1,1)}{4}$$

και πρέπει να λύσουμε την εξίσωση πρώτου βαθμού που προέκυψε ώστε να υπολογίσουμε το $f(1,1)$.

$$f(1,1) = \frac{28}{9}$$



Ομοίως υπολογίζουμε το $f(1,-1)$

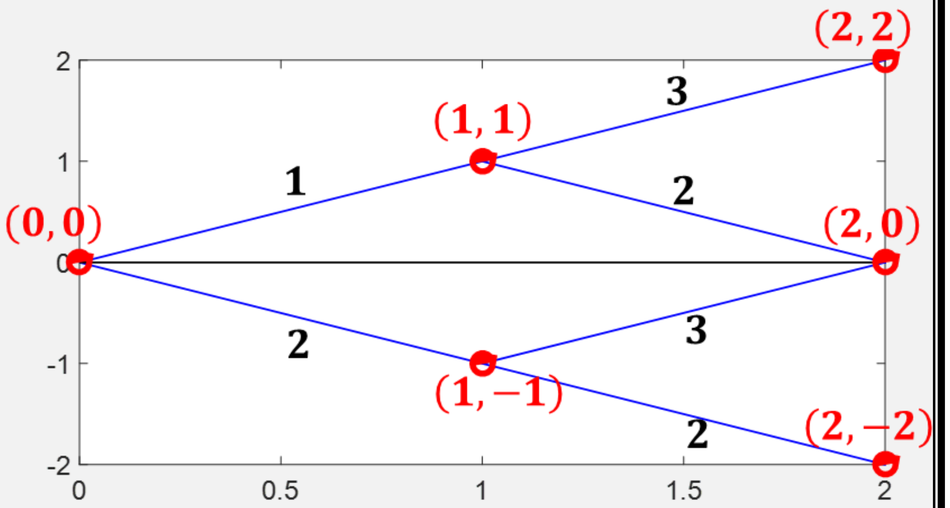
$p=\frac{2}{3}, q=\frac{1}{4}, \mu=3$

$$f(1,-1) = \min\left\{\begin{matrix} p(a(1,-1)+f(2,0)) + (1-p)[q\big(\delta(1,-1)+f(2,-2)\big)+(1-q)(f(1,-1)+\mu)] \\ p(\delta(1,-1)+f(2,-2)) + (1-p)[q\big(\alpha(1,-1)+f(2,0)\big)+(1-q)(f(1,-1)+\mu)] \end{matrix}\right\}$$

$$= \min\left\{\begin{matrix} \frac{2}{3}(3+0)+\frac{1}{3}[\frac{1}{4}(2+0)+\frac{3}{4}(f(1,-1)+3)] \\ \frac{2}{3}(2+0)+\frac{1}{3}[\frac{1}{4}(3+0)+\frac{3}{4}(f(1,-1)+3)] \end{matrix}\right\} = \cdots..$$

Οπότε βρίσκουμε:

$$f(1,-1) = \frac{28}{9}$$



- Για $\boldsymbol{x = 0}$

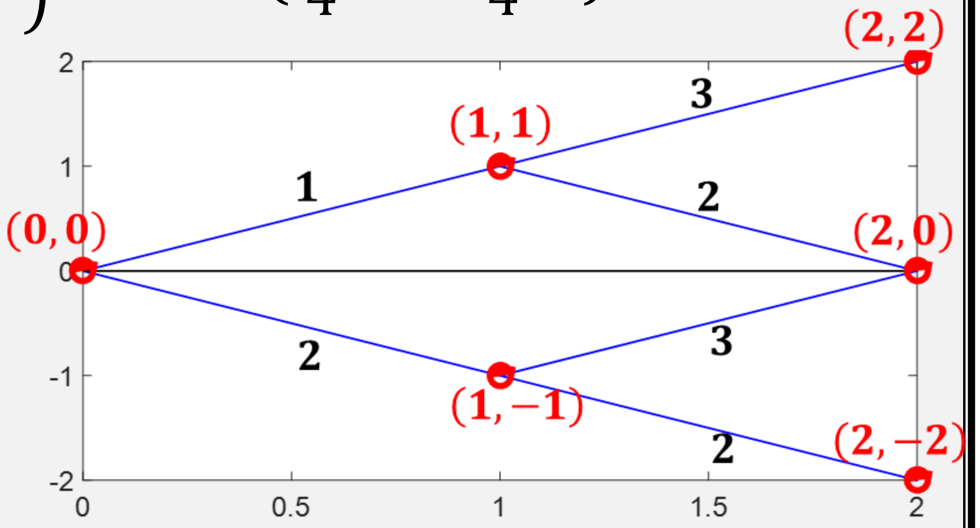
$\boldsymbol{p = \frac{2}{3}, q = \frac{1}{4}, \mu = 3}$

$$f(0,0) = \min\begin{Bmatrix} p(a(0,0) + f(1,1)) + (1-p)[q(\delta(0,0) + f(1,-1)) + (1-q)(f(0,0)+\mu)] \\ p(\delta(0,0) + f(1,-1)) + (1-p)[q(\alpha(0,0) + f(1,1)) + (1-q)(f(0,0)+\mu)] \end{Bmatrix}$$

$$= \min\begin{Bmatrix} \frac{2}{3}(1+\frac{28}{9}) + \frac{1}{3}[\frac{1}{4}\left(2+\frac{28}{9}\right) + \frac{3}{4}(f(0,0)+3)] \\ \frac{2}{3}(2+\frac{28}{9}) + \frac{1}{3}[\frac{1}{4}\left(1+\frac{28}{9}\right) + \frac{3}{4}(f(0,0)+3)] \end{Bmatrix} = \min\begin{Bmatrix} \frac{47}{12} + \frac{f(0,0)}{4} \\ \frac{18}{4} + \frac{f(0,0)}{4} \end{Bmatrix}$$

$$= \frac{47}{12} + \frac{f(0,0)}{4}$$

Οπότε βρίσκουμε: $f(0,0) = \frac{47}{9}$



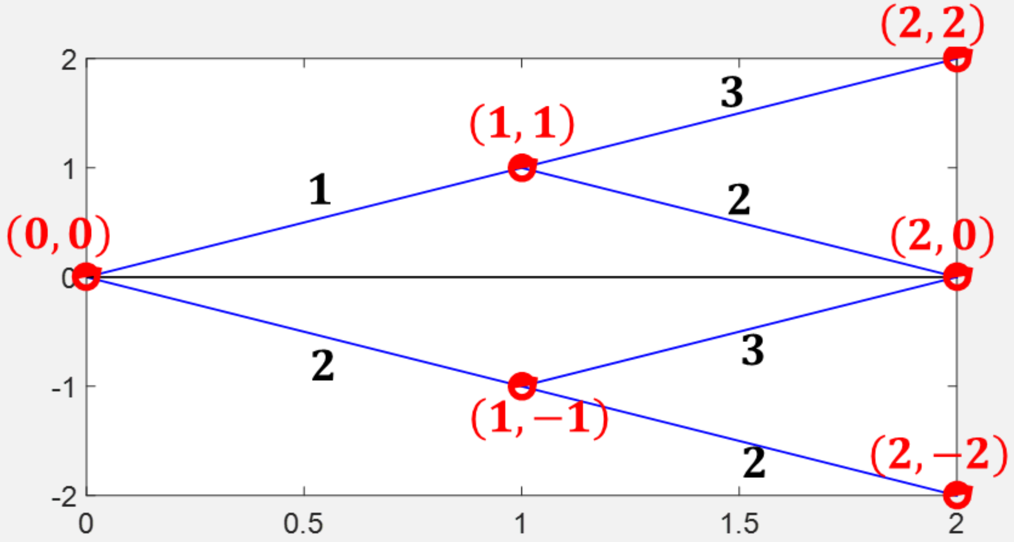
Η βέλτιστη πολιτική δίνεται για κάθε κόμβο χωριστά και όχι συνολικά.

Βρήκαμε:

$$f(0,0) = min\begin{Bmatrix} \frac{47}{12} + \frac{f(0,0)}{4} \\ \frac{18}{4} + \frac{f(0,0)}{4} \end{Bmatrix} = \frac{47}{12} + \frac{f(0,0)}{4}$$

Επομένως:

$(0,0) \quad \rightarrow$ κίνηση προς τα αριστερά



Βρήκαμε:

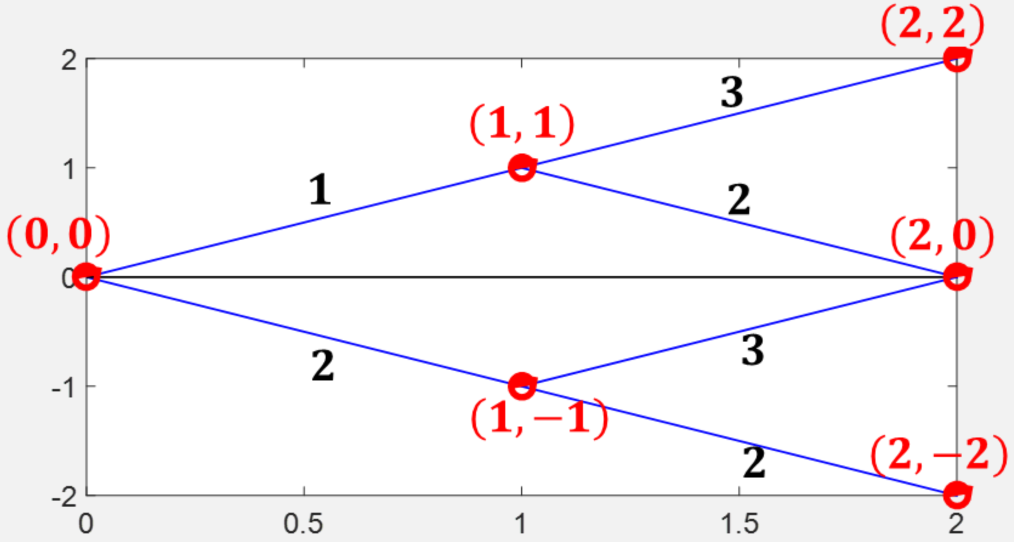
$$f(1,1) = min \begin{Bmatrix} \dfrac{35}{12} + \dfrac{f(1,1)}{4} \\ \dfrac{28}{12} + \dfrac{f(1,1)}{4} \end{Bmatrix} = \frac{28}{12} + \frac{f(1,1)}{4}$$

Επομένως:

$(1,1) \quad \rightarrow$ κίνηση προς τα δεξιά

Ομοίως προκύπτει:

$(1,-1) \quad \rightarrow$ κίνηση προς τα δεξιά



Άσκηση

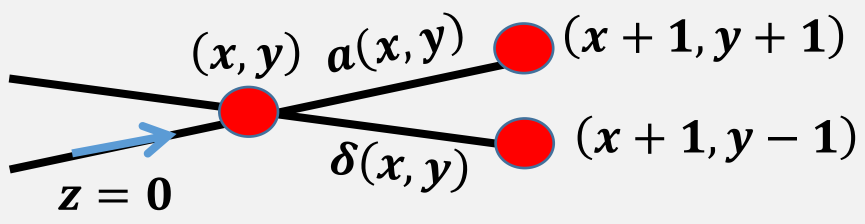
Έστω ένα κινητό δέχεται εντολές για να κινείται σε ένα δικτυωτό. Σε κάθε κόμβο, η εντολή να **πάμε αριστερά** έχει πιθανότητα $\boldsymbol{p_1}$ να εκτελεστεί. Αν **δεν** εκτελεστεί (με πιθανότητα $\mathbf{1} - \boldsymbol{p}_1$), τότε είτε **πάμε προς τα δεξιά** με πιθανότητα $\boldsymbol{q_1}$ είτε **παραμένει στην ίδια θέση** με πιθανότητα $\mathbf{1} - \boldsymbol{q_1}$. Αν η εντολή είναι να **πάμε προς τα δεξιά**, τότε οι αντίστοιχες πιθανότητες (των p_1, q_1) είναι $\boldsymbol{p_2}$, $\boldsymbol{q_2}$. Για κάθε αλλαγή κατεύθυνσης υπάρχει κόστος $\boldsymbol{k}$. Να βρεθεί η βέλτιστη πολιτική (βέλτιστες εντολές) με την μέθοδο του βέλτιστου με επαναπληροφόρηση ελέγχου.

Βέλτιστη συνάρτηση

$f(x, y, z) =$ {Το ελάχιστο αναμενόμενο κόστος από τον κόμβο (x, y) μέχρι το τέλος, όπου

$z = 0$ αν το τελευταίο βήμα προς τον κόμβο (x, y) έγινε προς τα αριστερά και

$z = 1$ αν έγινε προς τα δεξιά}



Βέλτιστη συνάρτηση

$f(x, y, z) =$ {Το ελάχιστο αναμενόμενο κόστος από τον κόμβο (x, y) μέχρι το τέλος,
όπου $z = 0$ αν το τελευταίο βήμα προς τον κόμβο (x, y) έγινε προς τα αριστερά και
$z = 1$ αν έγινε προς τα δεξιά}

Επαναληπτική σχέση

$$f(x, y, 0) =$$

$$\min \left\{ \begin{array}{c} \overbrace{p_1(\alpha(x, y) + f(x+1, y+1, 0)) +}^{\text{κίνηση προς τα αριστερά}} \\ (1 - p_1)[q_1[\delta(x, y) + f(x+1, y-1, 1) + k] + (1 - q_1) \overbrace{f(x, y, 0)}^{\text{παραμένει στην ίδια θέση}}] \\ \overbrace{p_2(\delta(x, y) + f(x+1, y-1, 1) + k) +}^{\text{κίνηση προς τα δεξιά}} \\ (1 - p_2)[q_2 \underbrace{[\alpha(x, y) + f(x+1, y+1, 0)]}_{\text{κίνηση προς τα αριστερά}} + (1 - q_2) \underbrace{f(x, y, 0)}_{\text{παραμένει στην ίδια θέση}}], \end{array} \right\}$$

εντολή κίνησης προς τα αριστερά

εντολή κίνησης προς τα δεξιά

Βέλτιστη συνάρτηση

$f(x, y, z) =$ {Το ελάχιστο αναμενόμενο κόστος από τον κόμβο (x, y) μέχρι το τέλος,
όπου $z = 0$ αν το τελευταίο βήμα προς τον κόμβο (x, y) έγινε προς τα αριστερά και
$z = 1$ αν έγινε προς τα δεξιά}

Επαναληπτική σχέση

$$f(x, y, 1) =$$

$$\min \begin{Bmatrix} \overbrace{p_1(\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y}) + \boldsymbol{f}(\boldsymbol{x}+\mathbf{1}, \boldsymbol{y}+\mathbf{1}, \mathbf{0}) + \boldsymbol{k})}^{\text{κίνηση προς τα αριστερά}} + \\ (1-p_1)[q_1[\underbrace{\boldsymbol{\delta}(\boldsymbol{x},\boldsymbol{y}) + \boldsymbol{f}(\boldsymbol{x}+\mathbf{1}, \boldsymbol{y}-\mathbf{1}, \mathbf{1})}_{\text{κίνηση προς τα δεξιά}}] + (1-q_1)\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y},\mathbf{1})] \\ p_2(\boldsymbol{\delta}(\boldsymbol{x},\boldsymbol{y}) + \boldsymbol{f}(\boldsymbol{x}+\mathbf{1}, \boldsymbol{y}-\mathbf{1}, \mathbf{1})) + \\ (1-p_2)[q_2[\underbrace{\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y}) + \boldsymbol{f}(\boldsymbol{x}+\mathbf{1}, \boldsymbol{y}+\mathbf{1}, \mathbf{0}) + \boldsymbol{k}}_{\text{κίνηση προς τα αριστερά}}] + (1-q_2)\boldsymbol{f}(\boldsymbol{x},\boldsymbol{y},\mathbf{1})], \end{Bmatrix}$$

παραμένει στην ίδια θέση

εντολή κίνησης προς τα αριστερά

εντολή κίνησης προς τα δεξιά

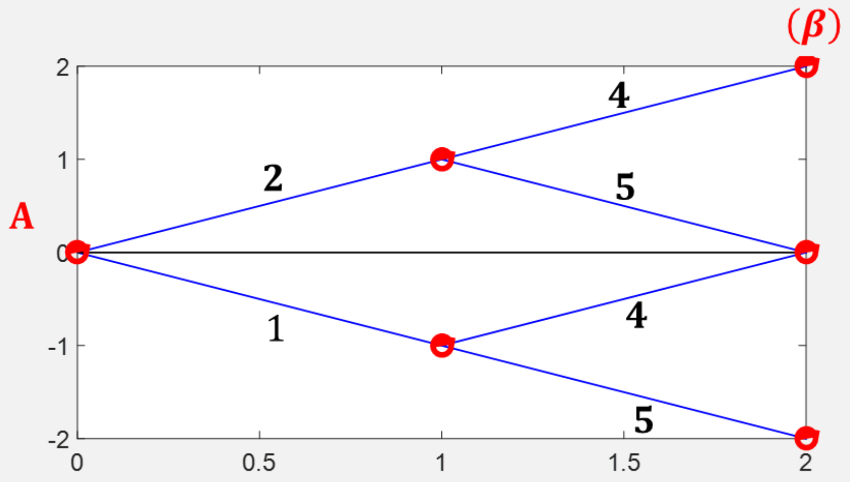
παραμένει στην ίδια θέση

Άσκηση

Δίνεται το παρακάτω δικτυωτό. Να βρεθεί με τη **μέθοδο του βέλτιστου με επαναπληροφόρηση ελέγχου**, το πρόβλημα διαδρομής μέγιστου κέρδους από το σημείο Α ως την ευθεία (β).

Μια εντολή εκτελείται με πιθανότητα 0.1. Όταν δίνεται εντολή για διαδρομή που δεν έχει αλλαγές κατεύθυνσης, προστίθεται **κόστος 1**.

Οι τιμές στα τόξα παριστάνουν **κέρδος**.

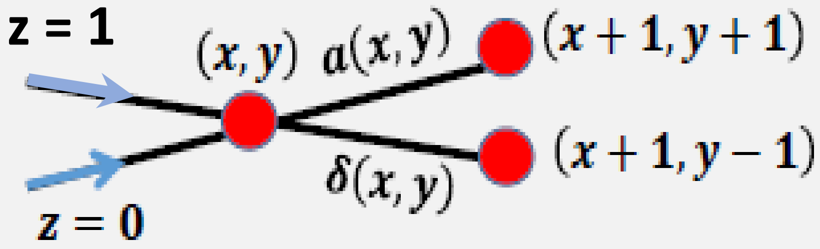


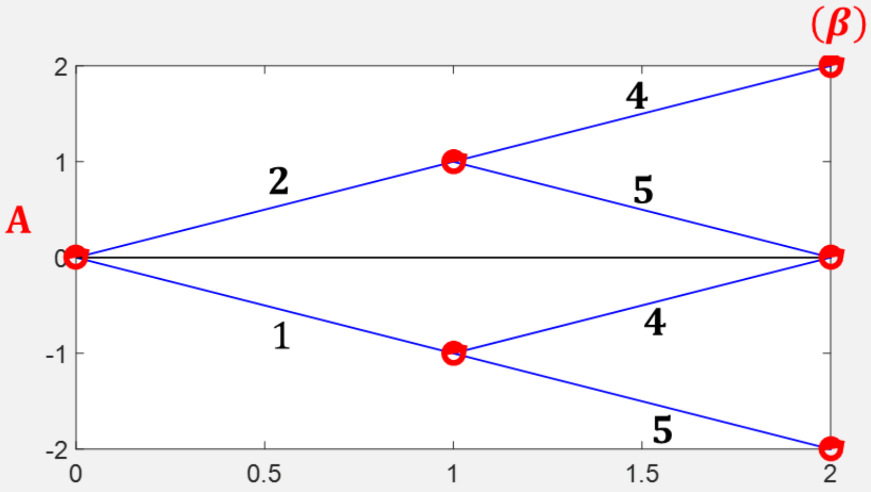
Βέλτιστη συνάρτηση

$f(x,y,z)$={Το μέγιστο αναμενόμενο κέρδος από τον κόμβο (x,y) μέχρι το τέλος, όπου

z=0 αν το τελευταίο βήμα προς τον κόμβο (x,y) έγινε προς τα αριστερά

z=1 αν το τελευταίο βήμα προς τον κόμβο (x,y) έγινε προς τα δεξιά}

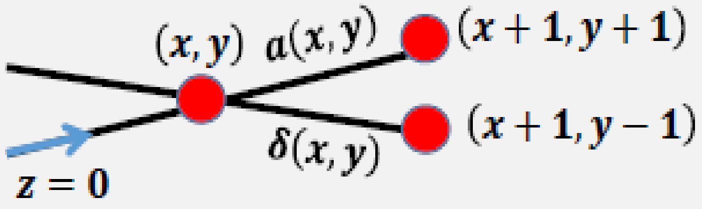




Επαναληπτική σχέση

$$f(x, y, \mathbf{0}) =$$

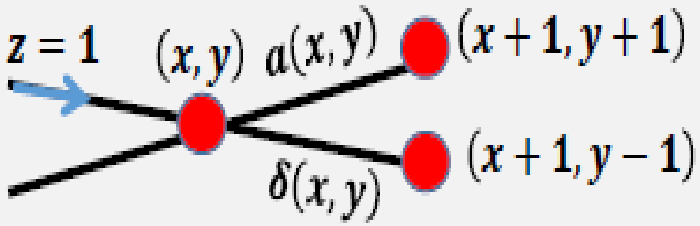
$$\max \begin{Bmatrix} p[a(x,y) + f(x+1, y+1, \mathbf{0}) - \mathbf{1}] + (1-p)[\delta(x,y) + f(x+1, y-1, \mathbf{1})] \\ p[\delta(x,y) + f(x+1, y-1, \mathbf{1})] + (1-p)[a(x,y) + f(x+1, y+1, \mathbf{0}) - \mathbf{1}] \end{Bmatrix}$$



Επαναληπτική σχέση

$$f(x, y, \mathbf{1}) =$$

$$max \begin{cases} p[a(x,y) + f(x+1, y+1, \mathbf{0})] \ +(1-p)\ [\delta(x,y) + f(x+1, y-1, \mathbf{1}) - \mathbf{1}] \\ p\,[\delta(x,y) + f(x+1, y-1, \mathbf{1}) - \mathbf{1}] \ +(1-p)\ [a(x,y) + f(x+1, y+1, \mathbf{0})] \end{cases}$$

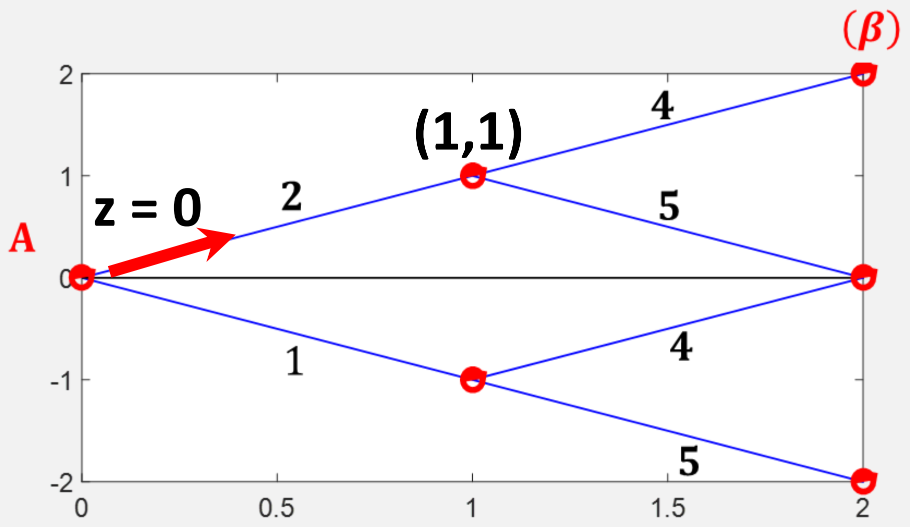


Οριακές συνθήκες

$$f(2,2,0) = 0$$
$$f(2,2,1) = 0$$

$$f(2,0,0) = 0$$
$$f(2,0,1) = 0$$

$$f(2,-2,0) = 0$$
$$f(2,-2,1) = 0$$



- Για $x = 1$

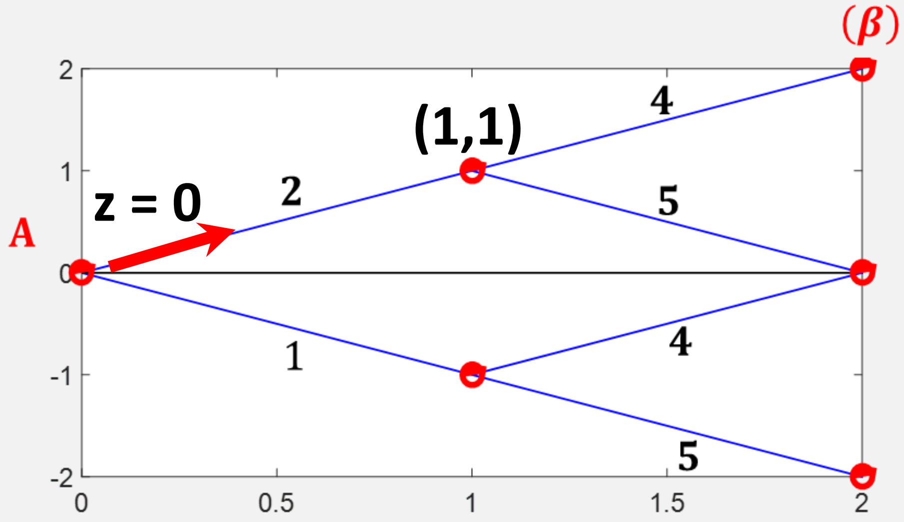
$$f(1,1,0) = max \begin{Bmatrix} p(a(1,1) + f(2,2,0) - 1) \ + (1-p)(\delta(1,1) + f(2,0,1)) \\ p\ (\delta(1,1) + f(2,0,1)) \ + (1-p)\ (\alpha(1,1) + f(2,2,0) - 1) \end{Bmatrix}$$

$$= \max \begin{Bmatrix} 0.1(4+0-1) + 0.9(5+0) \\ 0.1(5+0) + 0.9(4+0-1) \end{Bmatrix}$$

$$= \max \begin{Bmatrix} \mathbf{4.8} \\ 3.2 \end{Bmatrix} = \mathbf{4.8}$$

$f(1,1,1)$

δε γίνεται



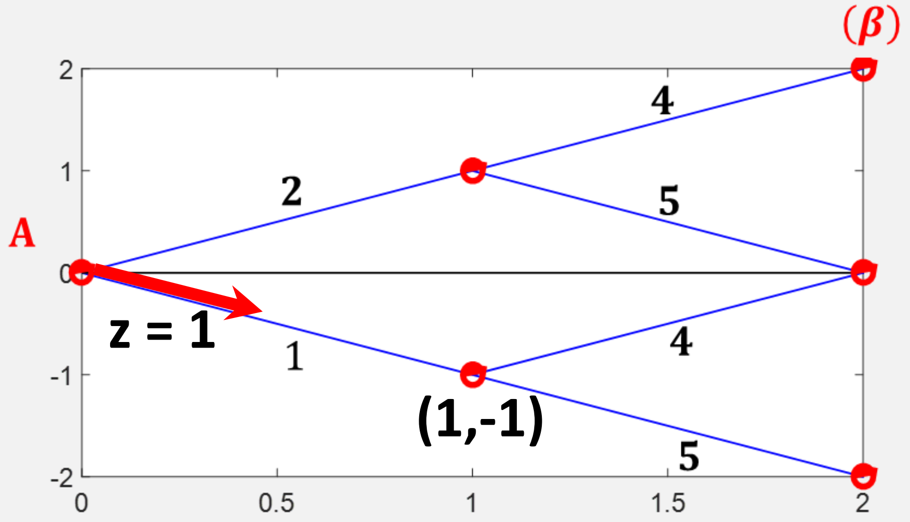
$$f(1,-1,1) = max\begin{Bmatrix} p(a(1,-1) + f(2,2,0)) \ + (1-p)(\delta(1,-1) + f(2,-2,1) - 1) \\ p\ (\delta(1,-1) + f(2,-2,1) - 1) \ + (1-p)\ (\alpha(1,-1) + f(2,0,0)) \end{Bmatrix}$$

$$= \max\begin{Bmatrix} 0.1(4+0) + 0.9(5+0-1) \\ 0.1(5+0-1) + 0.9(4+0) \end{Bmatrix}$$

$$= \max\begin{Bmatrix} \mathbf{4} \\ \mathbf{4} \end{Bmatrix} = \mathbf{4}$$

$f(1,-1,0)$

δε γίνεται

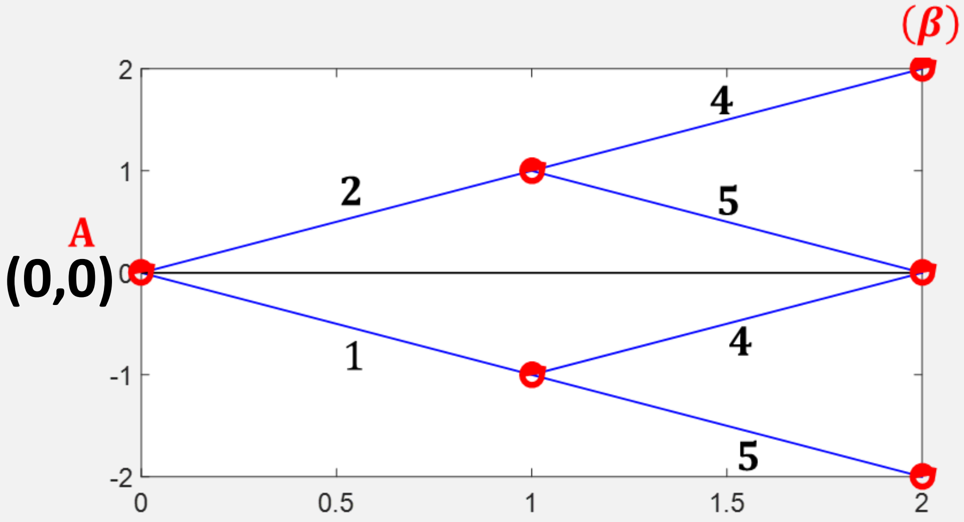


- Για $x = 0$

$$f(0,0,-) = max\begin{Bmatrix} p(a(0,0) + f(1,1,0)) \ + (1-p)(\delta(0,0) + f(1,-1,1)) \\ p\ (\delta(0,0) + f(1,-1,1)) \ + (1-p)\ (\alpha(0,0) + f(1,1,0)) \end{Bmatrix}$$

$$= \max\begin{Bmatrix} 0.1(2+4.8) + 0.9(1+4) \\ 0.1(1+4) + 0.9(2+4.80) \end{Bmatrix}$$

$$= \max\begin{Bmatrix} 5.18 \\ \mathbf{6.62} \end{Bmatrix} = \mathbf{6.62}$$



Βέλτιστη πολιτική εντολών

$$f(0,0,-) = \max\begin{Bmatrix} 5.18 \\ \mathbf{6.62} \end{Bmatrix} = \mathbf{6.62}$$

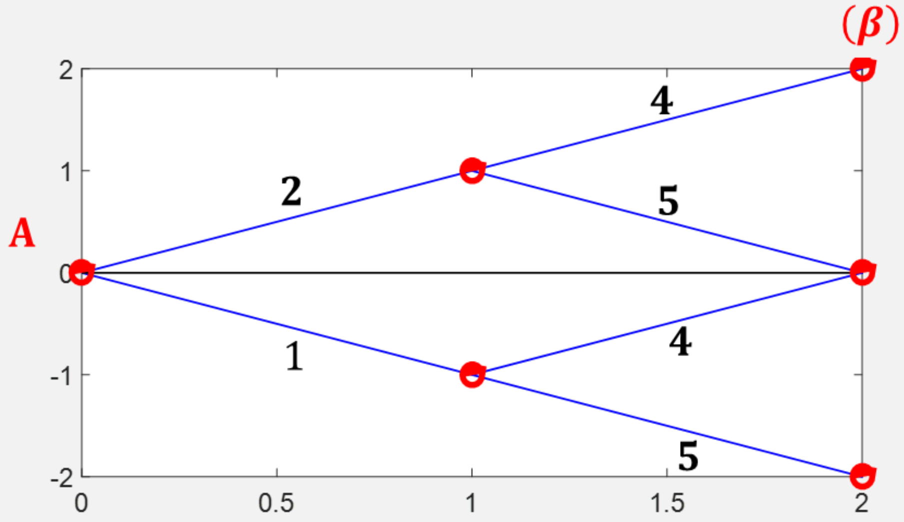
αρα εντολή κίνηση προς τα δεξιά

$$f(1,1,0) = \max\begin{Bmatrix} \mathbf{4.8} \\ 3.2 \end{Bmatrix} = \mathbf{4.8}$$

αρα εντολή κίνηση προς τα αριστερά

$$f(1,-1,0) = \max\begin{Bmatrix} \mathbf{4} \\ \mathbf{4} \end{Bmatrix} = \mathbf{4}$$

αρα εντολή κίνηση προς τα αριστερά ή δεξιά



Βιβλιογραφία

1) Π.-Χ. Βασιλείου (2001) Εφαρμοσμένος Μαθηματικός Προγραμματισμός, Εκδόσεις Ζήτη.
2) Π.-Χ. Βασιλείου, Γ. Τσακλίδης, Ν. Τσάντας (1998) Ασκήσεις στην Επιχειρησιακή Έρευνα, Εκδόσεις Ζήτη.