

# Στοχαστικές Στρατηγικές

Τμήμα Μαθηματικών, ΑΠΘ

## 3<sup>η</sup> ενότητα: Στοχαστικά προβλήματα διαδρομής – Μεθοδολογία (2)

**Παπάνα Αγγελική**

Μεταδιδακτορική ερευνήτρια, ΑΠΘ & Πανεπιστήμιο Μακεδονίας

E-mail: angeliki.papana@gmail.com, agrapana@auth.gr

Webpage: <http://users.auth.gr/agrapana>

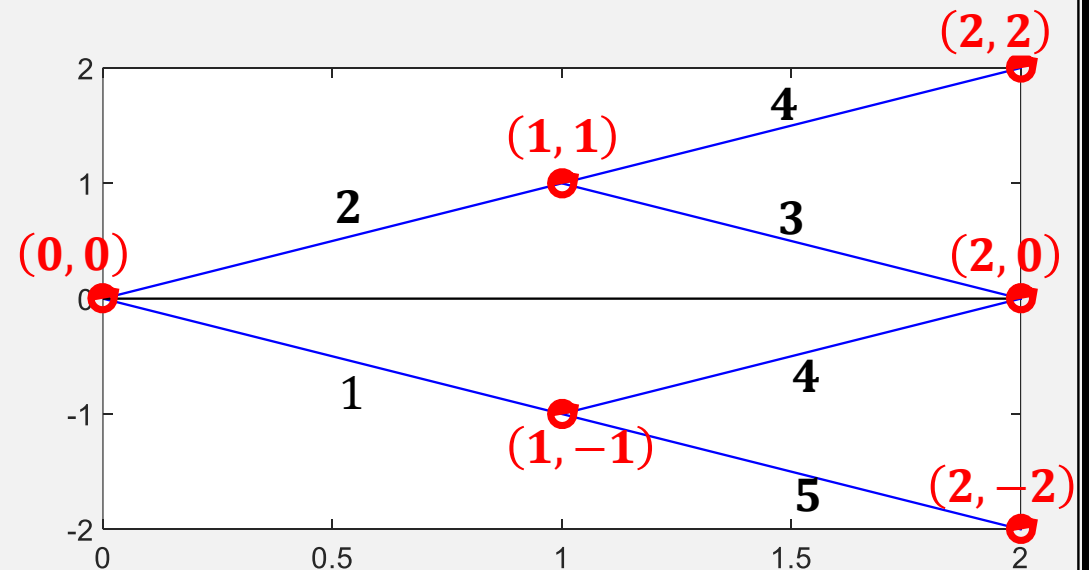
## Μέθοδος του ελέγχου ανοιχτού κύκλου

Σε αυτήν την μέθοδο οι οδηγίες δίνονται σαν μια ακολουθία. Για παράδειγμα η οδηγία **ΑΔ** σημαίνει ότι όταν βρίσκεσαι στον κόμβο  $(x, y)$ , πήγαινε προς τα **αριστερά** και μετά στον επόμενο κόμβο πάνε προς τα **δεξιά**.

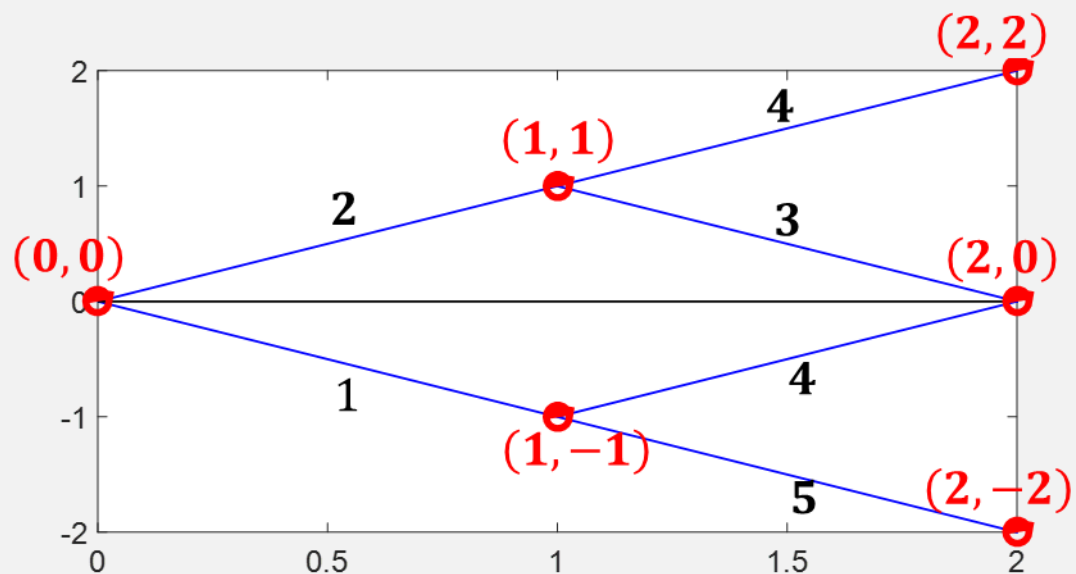
Σε ένα πλέγμα δεδομένων της δοθείσας μορφής, υπάρχουν **4** δυνατές οδηγίες:

**1) ΑΑ 2) ΑΔ 3) ΔΑ 4) ΔΔ**

Από αυτές πρέπει να βρούμε ποια είναι η βέλτιστη δυάδα εντολών.



Θεωρούμε ότι η εντολή εκτελείται στο **πρώτο βήμα** με πιθανότητα  $p_1 = 0.2$  και στο **δεύτερο βήμα** με πιθανότητα  $p_2 = 0.3$ . Οι τιμές στα τόξα παριστάνουν **κέρδος**. Να λυθεί η άσκηση με την **μέθοδο ελέγχου ανοιχτού κύκλου**.



**E(AA) =**

$$\begin{aligned} & p_1 p_2 [a(0,0) + a(1,1)] + \\ & p_1 (1 - p_2) [a(0,0) + \delta(1,1)] + \\ & (1 - p_1) p_2 [\delta(0,0) + \alpha(1, -1)] + \\ & (1 - p_1) (1 - p_2) [\delta(0,0) + \delta(1, -1)] \end{aligned}$$

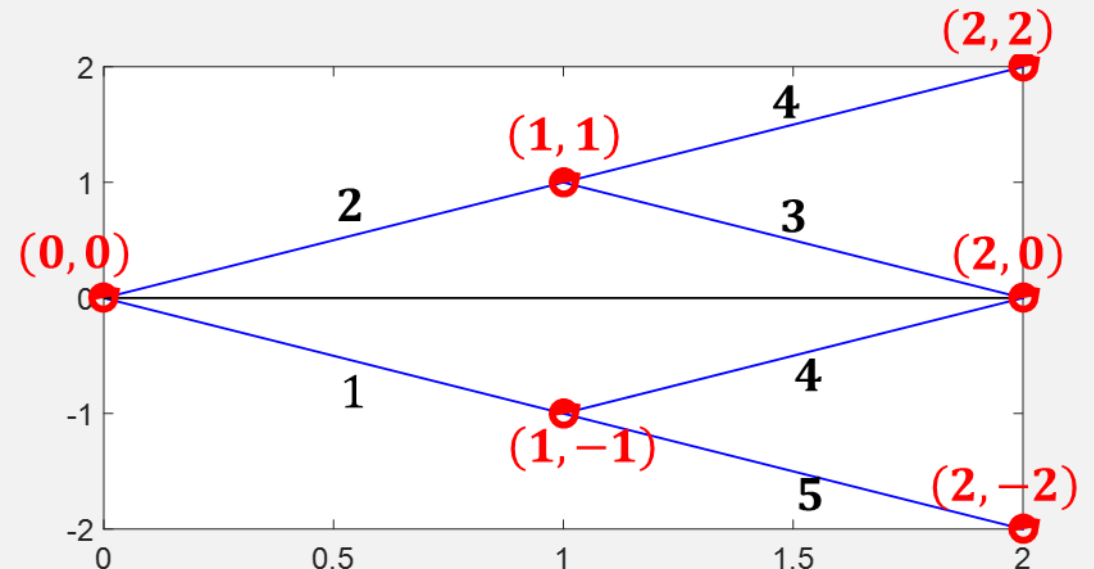
$$\begin{aligned} &= 0.2 * 0.3(2 + 4) + \\ & 0.2 * 0.7(2 + 3) + \\ & 0.8 * 0.3(1 + 4) + \\ & 0.8 * 0.7(1 + 5) \\ &= \mathbf{5.62} \end{aligned}$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε δέχομαι στο 1<sup>ο</sup> βήμα, δέχομαι στο 2<sup>ο</sup>

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(A\Delta) =$$

$$p_1 p_2 [a(0,0) + \delta(1,1)] +$$

$$p_1 (1 - p_2) [a(0,0) + a(1,1)] +$$

$$(1 - p_1) p_2 [\delta(0,0) + \delta(1, -1)] +$$

$$(1 - p_1) (1 - p_2) [\delta(0,0) + a(1, -1)]$$

$$= 0.2 * 0.3(2 + 3) +$$

$$0.2 * 0.7(2 + 4) +$$

$$0.8 * 0.3(1 + 5) +$$

$$0.8 * 0.7(1 + 4)$$

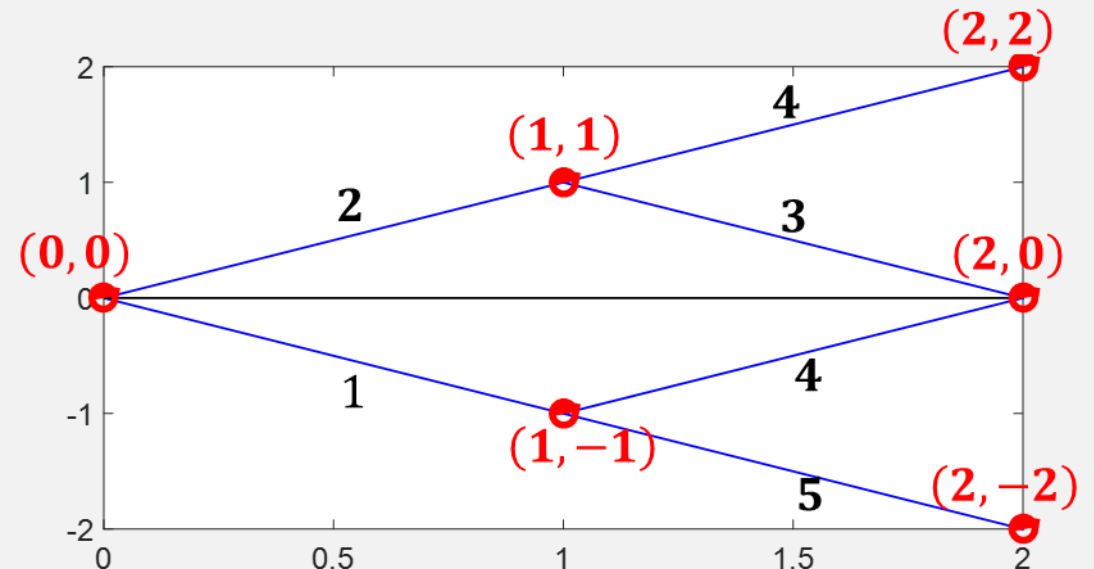
$$= 5.38$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε δέχομαι στο 1<sup>ο</sup> βήμα, δέχομαι στο 2<sup>ο</sup>

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(\Delta A) =$$

$$p_1 p_2 [\delta(0,0) + \alpha(1, -1)] +$$

$$p_1 (1 - p_2) [\delta(0,0) + \delta(1, -1)] +$$

$$(1 - p_1) p_2 [\alpha(0,0) + \alpha(1, 1)] +$$

$$(1 - p_1) (1 - p_2) [\alpha(0,0) + \delta(1, 1)]$$

$$= 0.2 * 0.3(1 + 4) +$$

$$0.2 * 0.7(1 + 5) +$$

$$0.8 * 0.3(2 + 4) +$$

$$0.8 * 0.7(2 + 3)$$

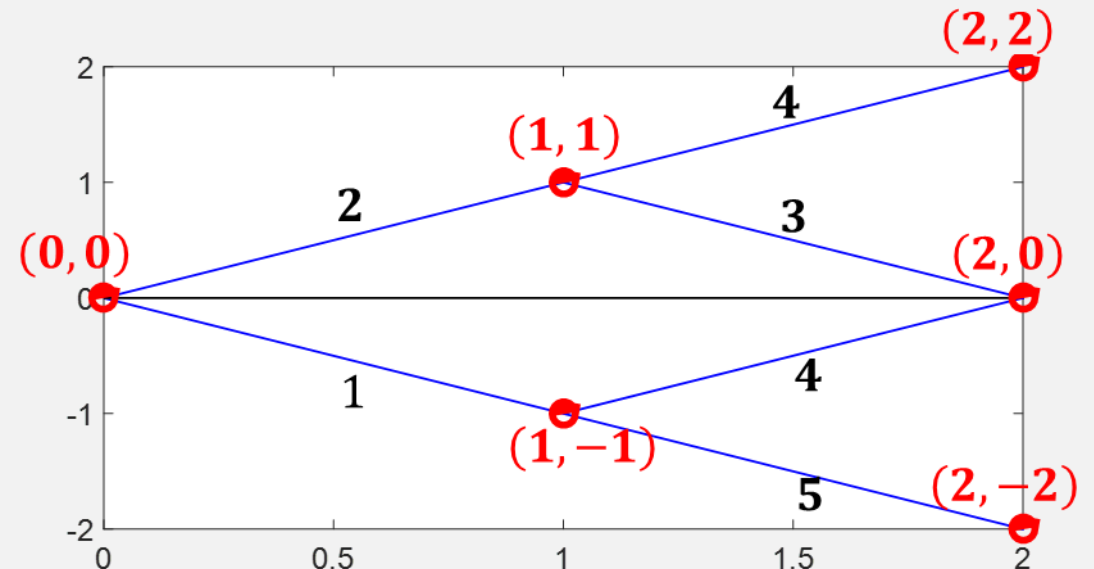
$$= 5.38$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε δέχομαι στο 1<sup>ο</sup> βήμα, δέχομαι στο 2<sup>ο</sup>

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(\Delta\Delta) =$$

$$p_1 p_2 [\delta(0,0) + \delta(1,-1)] +$$

$$p_1 (1 - p_2) [\delta(0,0) + \alpha(1,-1)] +$$

$$(1 - p_1) p_2 [\alpha(0,0) + \delta(1,1)] +$$

$$(1 - p_1) (1 - p_2) [\alpha(0,0) + \alpha(1,1)]$$

$$= 0.2 * 0.3(1 + 5) +$$

$$0.2 * 0.7(1 + 4) +$$

$$0.8 * 0.3(2 + 3) +$$

$$0.8 * 0.7(2 + 4)$$

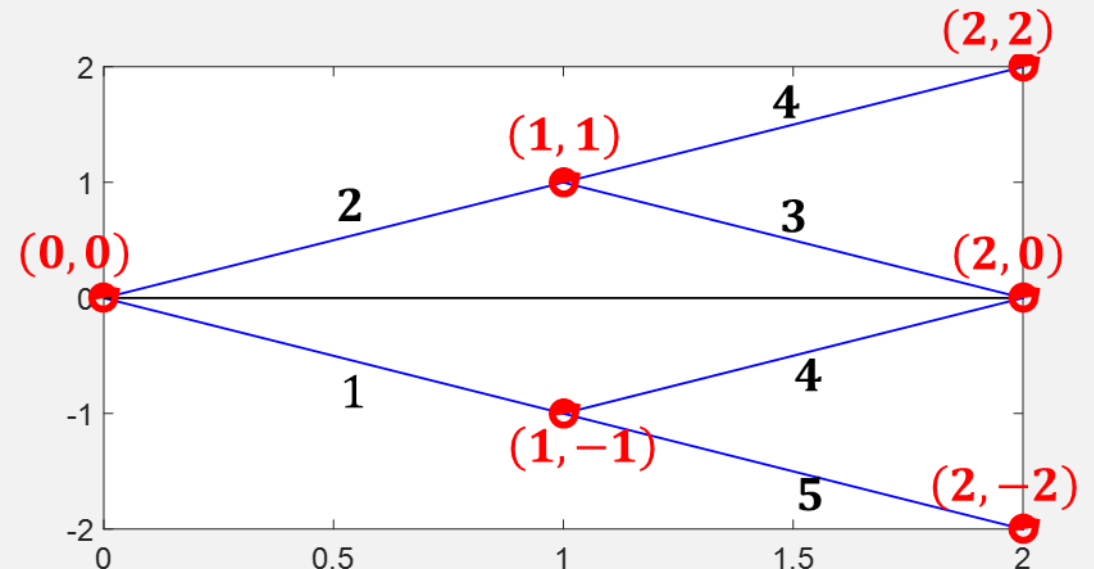
$$= 5.62$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε δέχομαι στο 1<sup>ο</sup> βήμα, δέχομαι στο 2<sup>ο</sup>

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

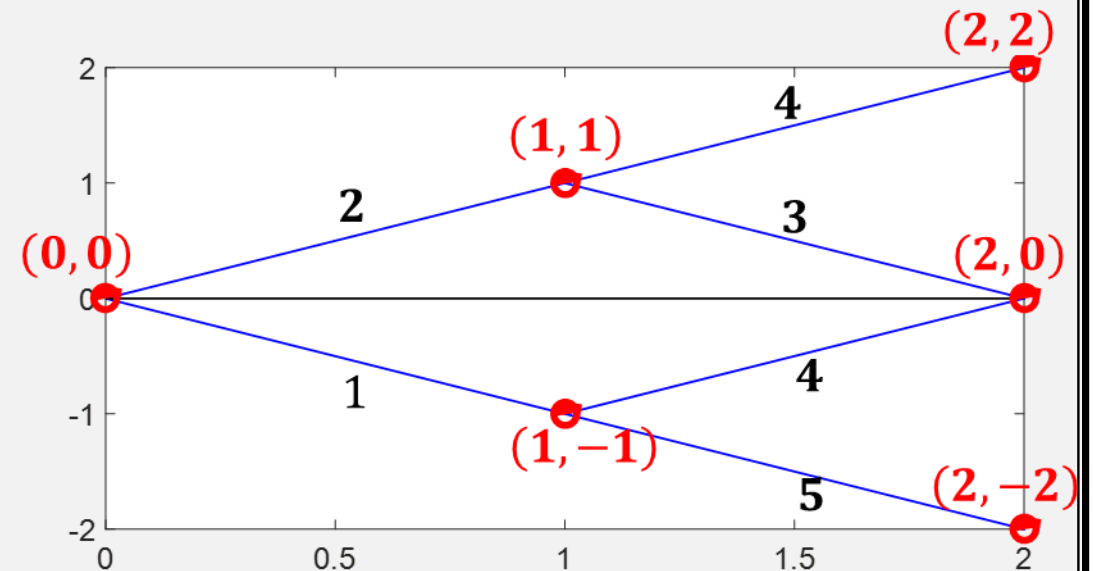


Άρα με βάση το μέγιστο αναμενόμενο κέρδος:

$$\max\{E(\mathbf{AA}), E(\mathbf{A}\Delta), E(\mathbf{\Delta A}), E(\mathbf{\Delta\Delta})\} =$$

$$\max\{5.62, 5.38, 5.38, 5.62\} = 5.62$$

Η βέλτιστη δυάδα εντολών είναι η  $\mathbf{AA}$  και η  $\mathbf{\Delta\Delta}$ .





Να βρεθεί η διασπορά του κέρδους της βέλτιστης πολιτικής ΔΔ.

Τι εκφράζει;

Το αναμενόμενο κέρδος για το ΔΔ υπολογίστηκε ως εξής:

$E(\Delta\Delta) =$

$$0.2 * 0.3 (1 + 5) + 0.2 * 0.7 (1 + 4) + 0.8 * 0.3 (2 + 3) + 0.8 * 0.7 (2 + 4) = \\ 0.06 * 6 + 0.14 * 5 + 0.24 * 5 + 0.56 * 6 = 5.62$$

$x_i$	$p_i$
6	0.06
5	0.14
5	0.24
6	0.56

Η διασπορά του κέρδους της πολιτικής  $\Delta\Delta$  δίνεται ως:

$$\sigma^2 = \sum_{i=1}^4 (x_i - \bar{x})^2 p_i$$

όπου  $\bar{x} = \mathbf{E}(\Delta\Delta)$

$x_i$	$p_i$
6	0.06
5	0.14
5	0.24
6	0.56

Άρα

$$\begin{aligned}\sigma^2 &= (6 - 5.62)^2 0.06 + \\ &\quad (5 - 5.62)^2 0.14 + \\ &\quad (5 - 5.62)^2 0.24 + \\ &\quad (6 - 5.62)^2 0.56 = \mathbf{0.2356}\end{aligned}$$

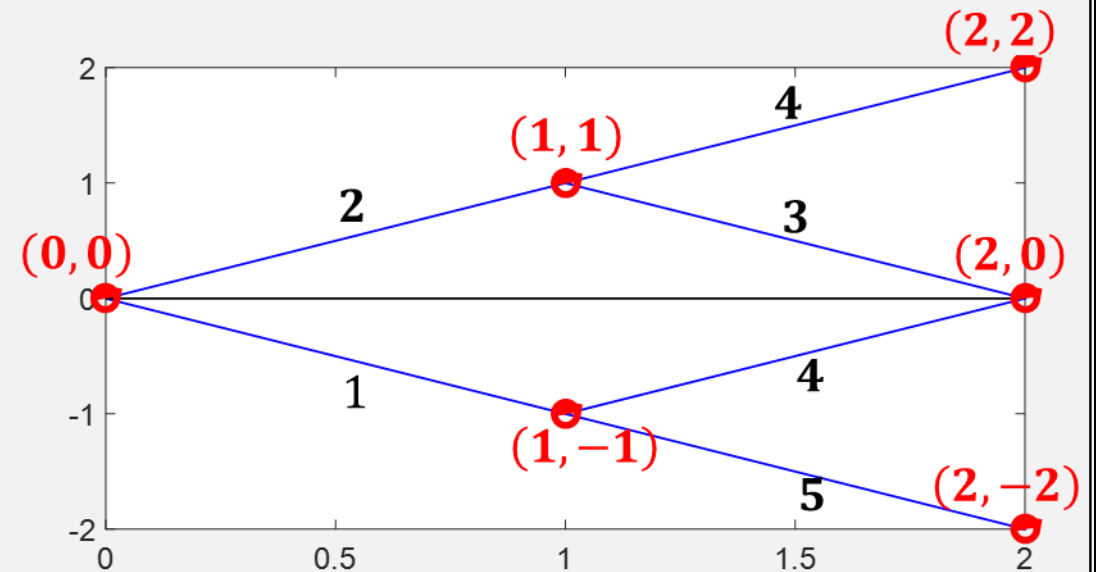
$\sigma = 0.4854$  τυπική απόκλιση

$x_i$	$p_i$
6	0.06
5	0.14
5	0.24
6	0.56

Μας δείχνει μια αναμενόμενη απόκλιση από την μέση τιμή των κερδών που περιμένουμε ακολουθώντας την βέλτιστη πολιτική.

## Άσκηση

Μια εντολή εκτελείται στο **πρώτο βήμα** με πιθανότητα **0.2** και στο **δεύτερο βήμα** με πιθανότητα **0.3**. Για κάθε **αλλαγή κατεύθυνσης** υπάρχει **κόστος 1**. Οι τιμές στα τόξα παριστάνουν κέρδος. Να λυθεί η άσκηση με την **μέθοδο ελέγχου ανοιχτού κύκλου**.



$$E(AA) =$$

$$p_1 p_2 [a(0,0) + a(1,1)] +$$

$$p_1 (1 - p_2) [a(0,0) + \delta(1,1) - 1] +$$

$$(1 - p_1) p_2 [\delta(0,0) + \alpha(1, -1) - 1] +$$

$$(1 - p_1) (1 - p_2) [\delta(0,0) + \delta(1, -1)]$$

$$= 0.2 * 0.3(2 + 4) +$$

$$0.2 * 0.7(2 + 3 - 1) +$$

$$0.8 * 0.3(1 + 4 - 1) +$$

$$0.8 * 0.7(1 + 5)$$

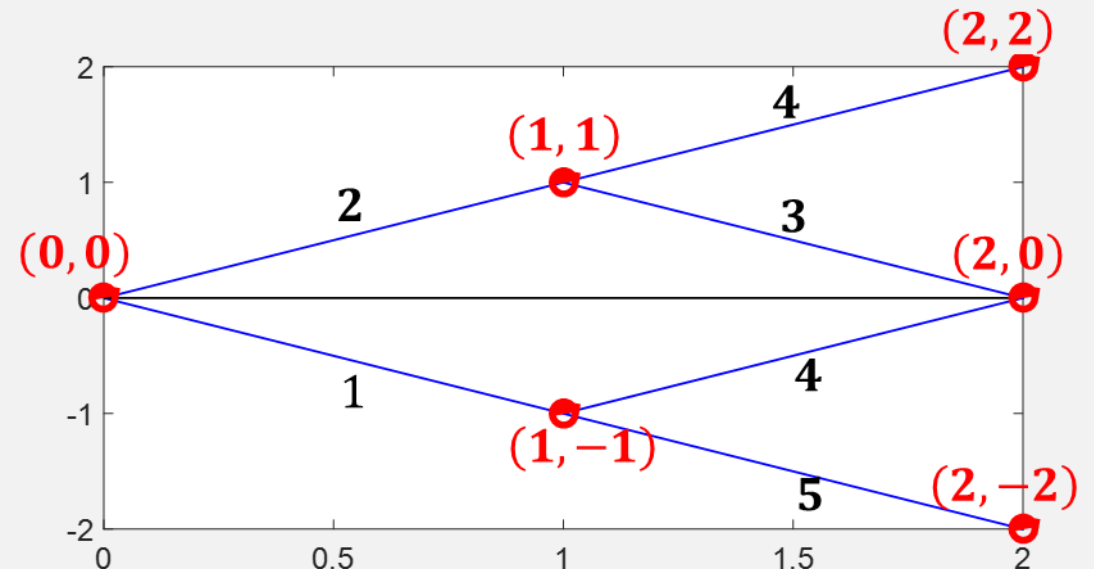
$$= 5.24$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε την δέχομαι στο 1<sup>ο</sup>, τη δέχομαι στο 2<sup>ο</sup> βήμα

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(\Delta) =$$

$$p_1 p_2 [a(0,0) + \delta(1,1) - 1] +$$

$$p_1 (1 - p_2) [a(0,0) + \alpha(1,1)] +$$

$$(1 - p_1) p_2 [\delta(0,0) + \delta(1, -1)] +$$

$$(1 - p_1) (1 - p_2) [\delta(0,0) + \alpha(1, -1) - 1]$$

$$= 0.2 * 0.3 (2 + 3 - 1) +$$

$$0.2 * 0.7 (2 + 4) +$$

$$0.8 * 0.3 (1 + 5) +$$

$$0.8 * 0.7 (1 + 4 - 1)$$

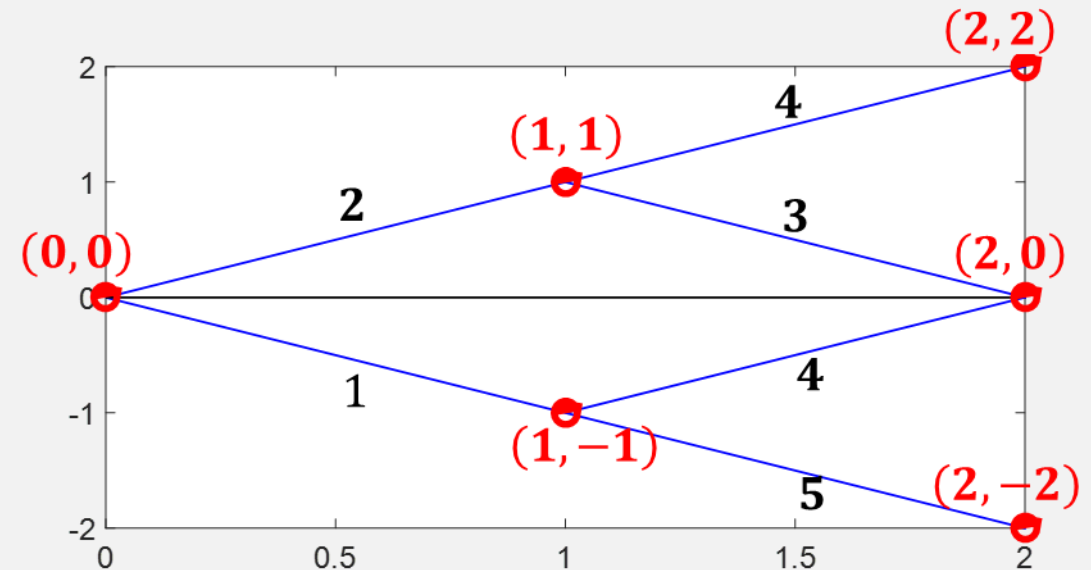
$$= \mathbf{4.76}$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε την δέχομαι στο 1<sup>ο</sup>, τη δέχομαι στο 2<sup>ο</sup> βήμα

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(\Delta A) =$$

$$p_1 p_2 [\delta(0,0) + \alpha(1, -1) - 1] +$$

$$p_1 (1 - p_2) [\delta(0,0) + \delta(1, -1)] +$$

$$(1 - p_1) p_2 [\alpha(0,0) + \alpha(1,1)] +$$

$$(1 - p_1) (1 - p_2) [\alpha(0,0) + \delta(1, 1) - 1]$$

$$= 0.2 * 0.3 (1 + 4 - 1) +$$

$$0.2 * 0.7 (1 + 5) +$$

$$0.8 * 0.3 (2 + 4) +$$

$$0.8 * 0.7 (2 + 3 - 1)$$

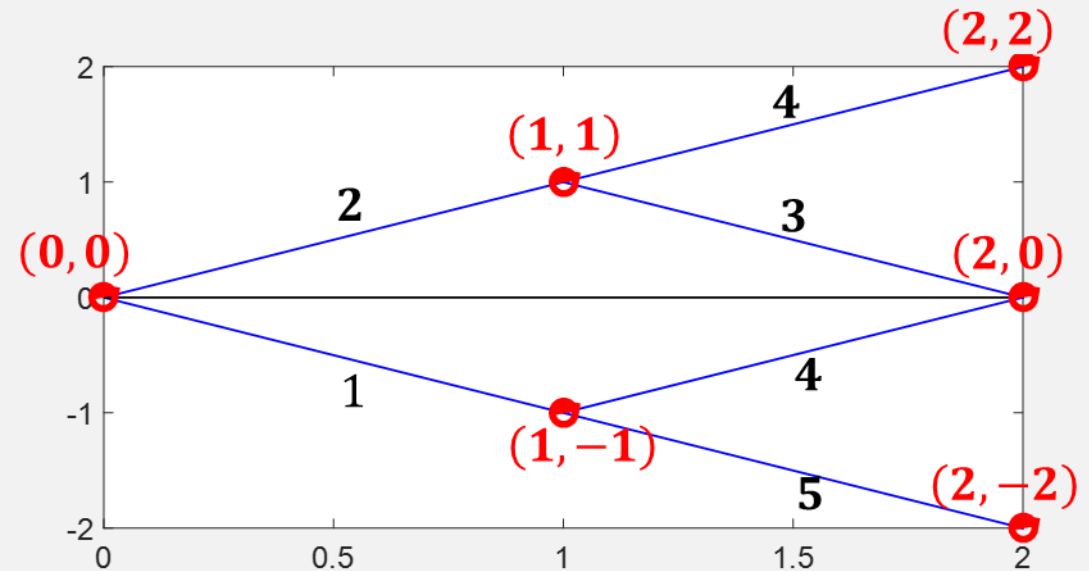
$$= \mathbf{4.76}$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε την δέχομαι στο 1<sup>ο</sup>, τη δέχομαι στο 2<sup>ο</sup> βήμα

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(\Delta\Delta) =$$

$$p_1 p_2 [\delta(0,0) + \delta(1,-1)] +$$

$$p_1 (1 - p_2) [\delta(0,0) + \alpha(1,-1) - 1] +$$

$$(1 - p_1) p_2 [\alpha(0,0) + \delta(1,1) - 1] +$$

$$(1 - p_1)(1 - p_2) [\alpha(0,0) + \alpha(1,1)]$$

$$= 0.2 * 0.3 (1 + 5) +$$

$$0.2 * 0.7 (1 + 4 - 1) +$$

$$0.8 * 0.3 (2 + 3 - 1) +$$

$$0.8 * 0.7 (2 + 4)$$

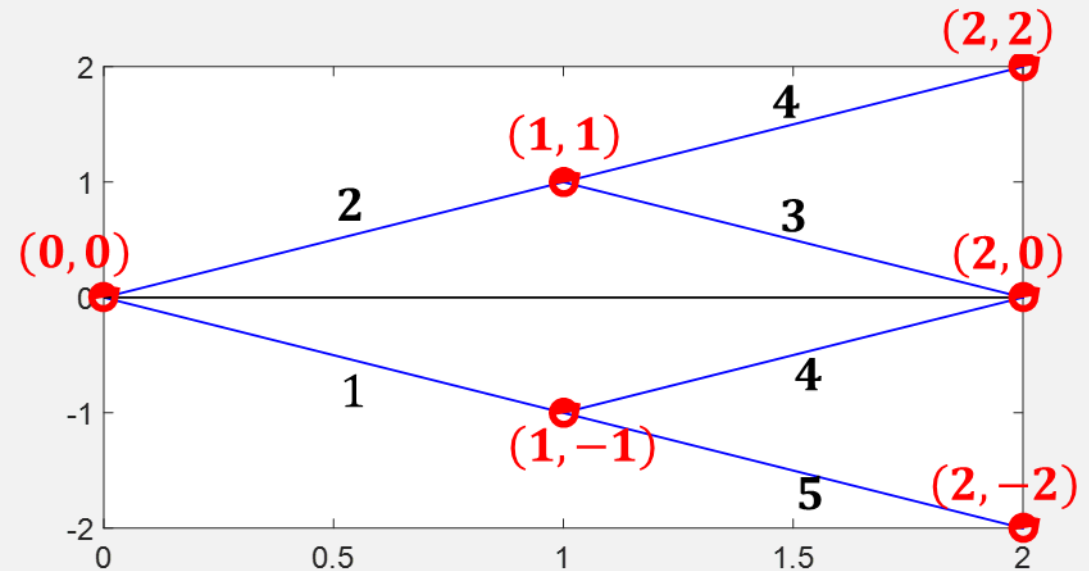
$$= \mathbf{5.24}$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε την δέχομαι στο 1<sup>ο</sup>, τη δέχομαι στο 2<sup>ο</sup> βήμα

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

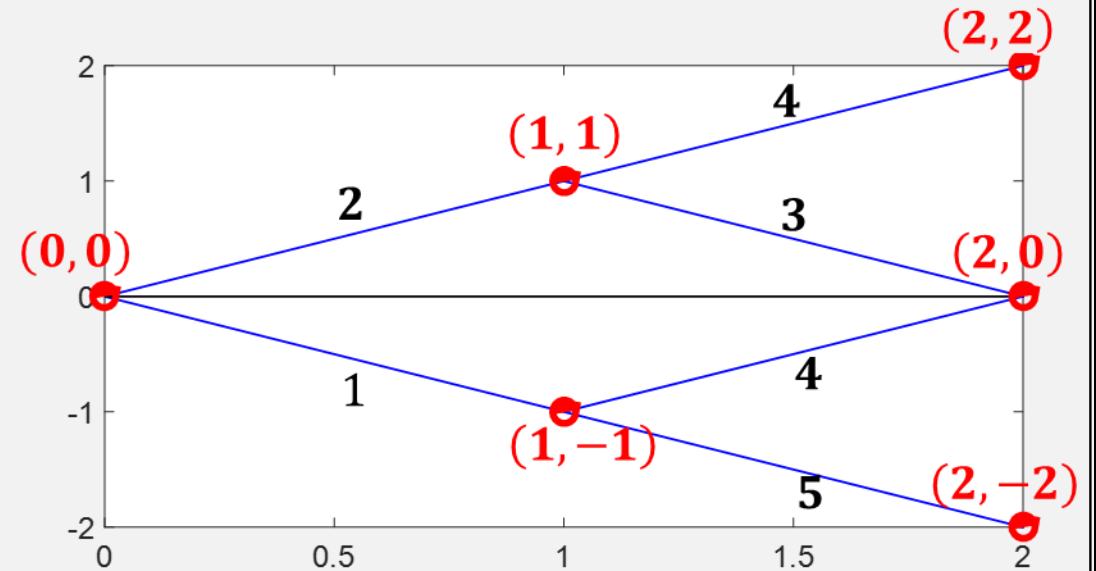




Άρα

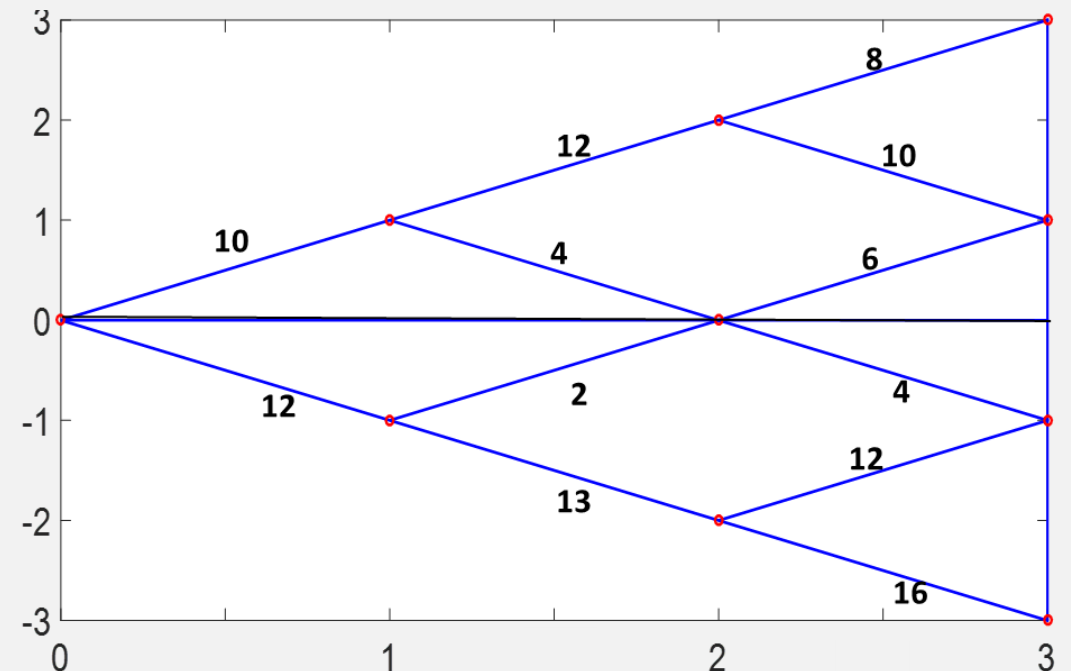
$$\max\{E(AA), E(A\Delta), E(\Delta A), E(\Delta\Delta)\} = \max\{5.24, 4.76, 4.76, 5.24\} = 5.24$$

Η βέλτιστη δυάδα εντολών είναι η **AA** και η **ΔΔ**.



**Άσκηση** Στο δικτυωτό του σχήματος, οι τιμές στα τόξα παριστάνουν **κόστος**. Σε κάθε κόμβο, ακολουθείται με **πιθανότητα  $p(x, y)$**  η οδηγία που δίνεται, διαφορετικά πάμε προς την αντίθετη κατεύθυνση με **πιθανότητα  $q(x, y)$** . Να βρεθεί η βέλτιστη πολιτική εντολών με το **ελάχιστο αναμενόμενο κόστος** με την **μέθοδο ανοιχτού κύκλου**.

		$p(x, y)$			$q(x, y)$				
		x	0	1	2	x	0	1	2
y	x	0	1	2	y	x	0	1	2
-2	-2	-	-	$\frac{1}{4}$	-2	-	-	$\frac{3}{4}$	
-1	-1	-	$\frac{3}{4}$	-	-1	-	$\frac{1}{4}$	-	
0	0	$\frac{3}{4}$	-	$\frac{1}{2}$	0	$\frac{1}{4}$	-	$\frac{1}{2}$	
1	1	-	$\frac{1}{2}$	-	1	-	$\frac{1}{2}$	-	
2	2	-	-	$\frac{3}{4}$	2	-	-	$\frac{1}{4}$	



Υπάρχουν 8 δυνατές διαφορετικές οδηγίες:

1) ΑΑΑ, 2) ΑΑΔ, 3) ΑΔΑ, 4) ΔΑΑ, 5) ΑΔΔ, 6) ΔΑΔ, 7) ΔΔΑ, 8) ΔΔΔ

Το **αναμενόμενο κόστος** για την οδηγία **ΑΑΑ** είναι:

$$\begin{aligned} E(\mathbf{ΑΑΑ}) = & p(0,0)p(1,1)p(2,2)\{a(0,0) + a(1,1) + a(2,2)\} + \mathbf{ΑΑΑ} \\ & p(0,0)p(1,1)q(2,2)\{a(0,0) + a(1,1) + \delta(2,2)\} + \mathbf{ΑΑΔ} \\ & p(0,0)q(1,1)p(2,0)\{a(0,0) + \delta(1,1) + a(2,0)\} + \mathbf{ΑΔΑ} \\ & q(0,0)p(1,-1)p(2,0)\{\delta(0,0) + a(1,-1) + a(2,0)\} + \mathbf{ΔΑΑ} \\ & p(0,0)q(1,1)q(2,0)\{a(0,0) + \delta(1,1) + \delta(2,0)\} + \mathbf{ΑΔΔ} \\ & q(0,0)p(1,-1)q(2,0)\{\delta(0,0) + a(1,-1) + \delta(2,0)\} + \mathbf{ΔΑΔ} \\ & q(0,0)q(1,-1)p(2,-2)\{\delta(0,0) + \delta(1,-1) + a(2,-2)\} + \mathbf{ΔΔΑ} \\ & q(0,0)q(1,-1)q(2,-2)\{\delta(0,0) + \delta(1,-1) + \delta(2,-2)\} = \mathbf{ΔΔΔ} \end{aligned}$$

$$\begin{aligned}
E(\text{AAA}) &= \frac{3}{4} \frac{1}{2} \frac{3}{4} \{10 + 12 + 8\} + && \text{AAA} \\
&\frac{3}{4} \frac{1}{2} \frac{1}{4} \{10 + 12 + 10\} + && \text{AA}\Delta \\
&\frac{3}{4} \frac{1}{2} \frac{1}{2} \{10 + 4 + 6\} + && \text{A}\Delta\text{A} \\
&\frac{1}{4} \frac{3}{4} \frac{1}{2} \{12 + 2 + 6\} + && \Delta\text{AA} \\
&\frac{3}{4} \frac{1}{2} \frac{1}{2} \{10 + 4 + 4\} + && \text{A}\Delta\Delta \\
&\frac{1}{4} \frac{3}{4} \frac{1}{2} \{12 + 2 + 4\} + && \Delta\text{A}\Delta \\
&\frac{1}{4} \frac{1}{4} \frac{1}{4} \{12 + 13 + 12\} + && \Delta\Delta\text{A} \\
&\frac{1}{4} \frac{1}{4} \frac{3}{4} \{12 + 13 + 16\} = 24.625 && \Delta\Delta\Delta
\end{aligned}$$

Ομοίως υπολογίζουμε το **αναμενόμενο κόστος** για κάθε μία από τις υπόλοιπες 7 δυνατές εντολές.

Όταν υπολογίσουμε όλα τα αναμενόμενα κόστη για όλες τις δυνατές εντολές, μπορούμε να αποφασίσουμε για την βέλτιστη σειρά εντολών.

## Παρατήρηση

Τα **αναμενόμενα κόστη** που προκύπτουν από την **μέθοδο ελέγχου ανοιχτού κύκλου** είναι μεγαλύτερα συγκριτικά με αυτά που προκύπτουν με την **μέθοδο του βέλτιστου με επαναπληρόφρηση ελέγχου**, καθώς οι πληροφορίες που έχουμε στην περίπτωση του βέλτιστου με επαναπληρόφρηση ελέγχου είναι περισσότερες.

## Μέθοδος του ελέγχου ανοιχτού κύκλου με επαναπληρόρηση

Η μέθοδος αυτή είναι συνδυασμός των δύο προηγούμενων μεθόδων.

Στον πρώτο κόμβο εφαρμόζουμε την μέθοδο **ελέγχου ανοιχτού κύκλου** και έτσι παίρνουμε την πρώτη απόφαση.

Στον δεύτερο κόμβο εφαρμόζουμε την μέθοδο **ελέγχου ανοιχτού κύκλου** για την διαδρομή που απομένει και έτσι παίρνουμε την δεύτερη απόφαση, κ.ο.κ.

## Μέθοδος του ελέγχου ανοιχτού κύκλου με επαναπληρόρηση

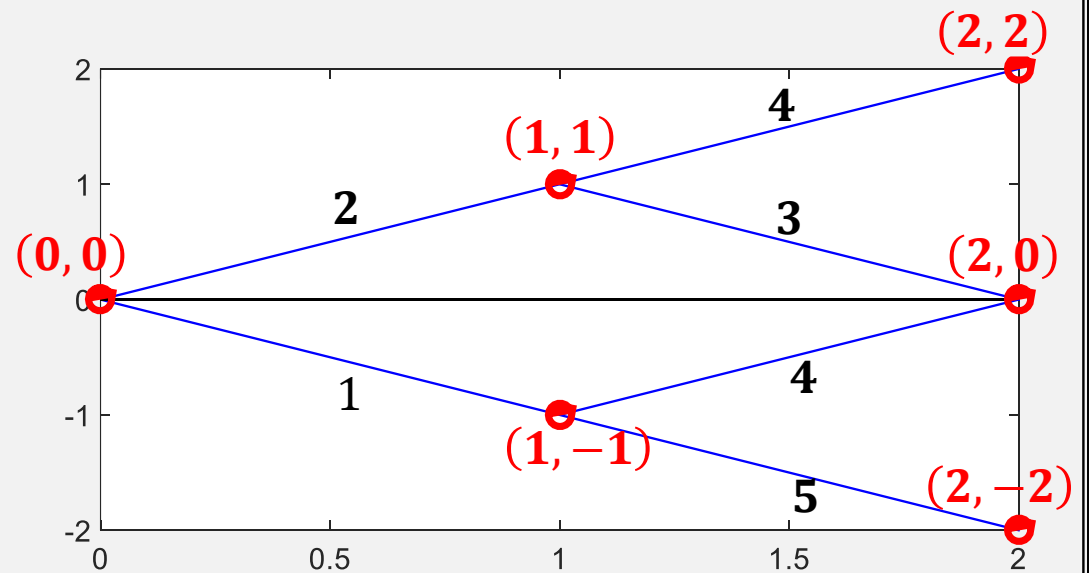
Έστω το παρακάτω δικτυωτό, όπου οι τιμές στα τόξα είναι **κέρδος**. Θεωρούμε ότι η εντολή εκτελείται στο **πρώτο βήμα** με πιθανότητα  $p_1 = 0.2$  και στο **δεύτερο βήμα** με πιθανότητα  $p_2 = 0.3$ .

Λύση

### ■ Κόμβος (0,0)

Υπάρχουν **4** δυνατές οδηγίες από τον αρχικό κόμβο (0,0):

**1) AA 2) AD 3) DA 4) DD**





$E(AA) =$

$$\begin{aligned} & p_1 p_2 [a(0,0) + a(1,1)] + \\ & p_1 (1 - p_2) [a(0,0) + \delta(1,1)] + \\ & (1 - p_1) p_2 [\delta(0,0) + \alpha(1, -1)] + \\ & (1 - p_1) (1 - p_2) [\delta(0,0) + \delta(1, -1)] \end{aligned}$$

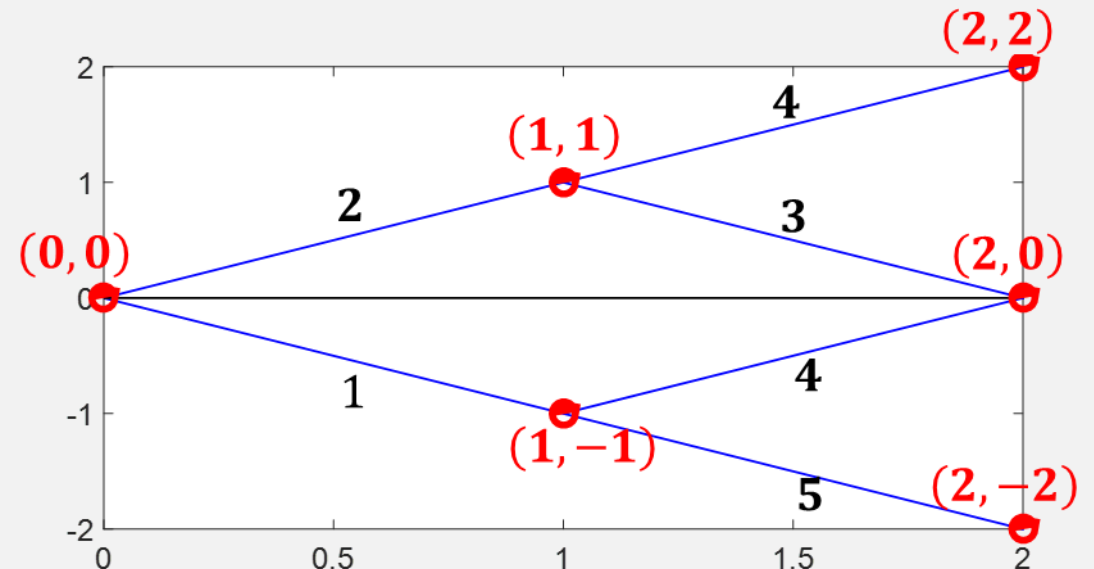
$$\begin{aligned} &= 0.2 * 0.3 * (2 + 4) + \\ & 0.2 * 0.7 * (2 + 3) + \\ & 0.8 * 0.3 * (1 + 4) + \\ & 0.8 * 0.7 * (1 + 5) \\ &= \mathbf{5.62} \end{aligned}$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε την δέχομαι στο 1<sup>ο</sup>, τη δέχομαι στο 2<sup>ο</sup> βήμα

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



$$E(A\Delta) =$$

$$p_1 p_2 [a(0,0) + \delta(1,1)] +$$

$$p_1 (1 - p_2) [a(0,0) + a(1,1)] +$$

$$(1 - p_1) p_2 [\delta(0,0) + \delta(1, -1)] +$$

$$(1 - p_1) (1 - p_2) [\delta(0,0) + a(1, -1)]$$

$$= 0.2 * 0.3 * (2 + 3) +$$

$$0.2 * 0.7 * (2 + 4) +$$

$$0.8 * 0.3 * (1 + 5) +$$

$$0.8 * 0.7 * (1 + 4)$$

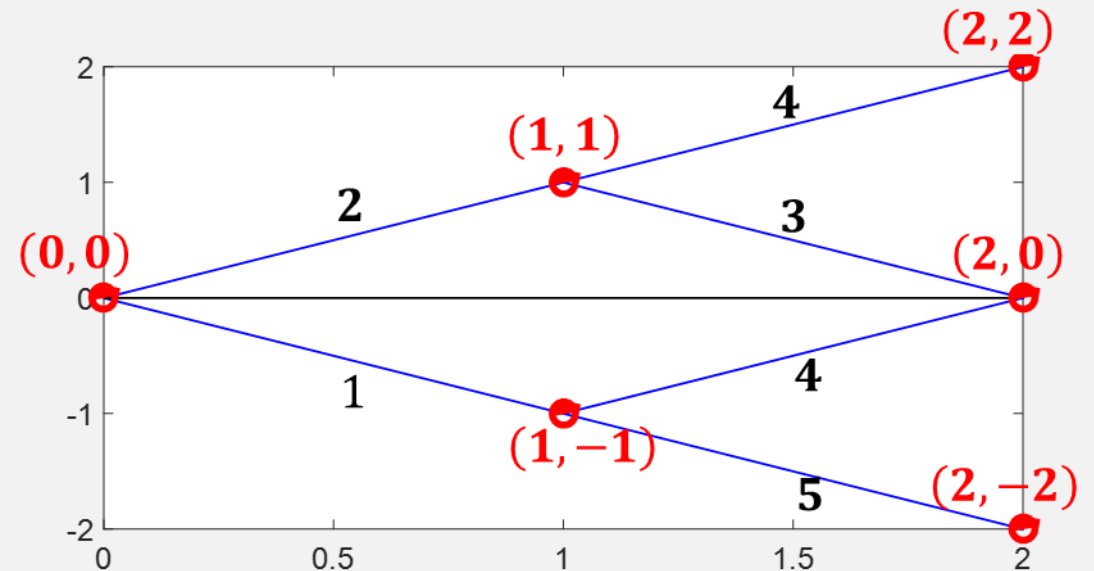
$$= \mathbf{5.38}$$

Δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα

Δέχομαι την εντολή στο 1<sup>ο</sup> βήμα, όχι στο 2<sup>ο</sup>

Δε την δέχομαι στο 1<sup>ο</sup>, τη δέχομαι στο 2<sup>ο</sup> βήμα

Δε δέχομαι τις εντολές στο 1<sup>ο</sup> και 2<sup>ο</sup> βήμα



Ομοίως υπολογίζουμε

$$E(\Delta A) = 5.38, E(\Delta \Delta) = 5.62$$

- Για τον κόμβο  $(1, 1)$ :

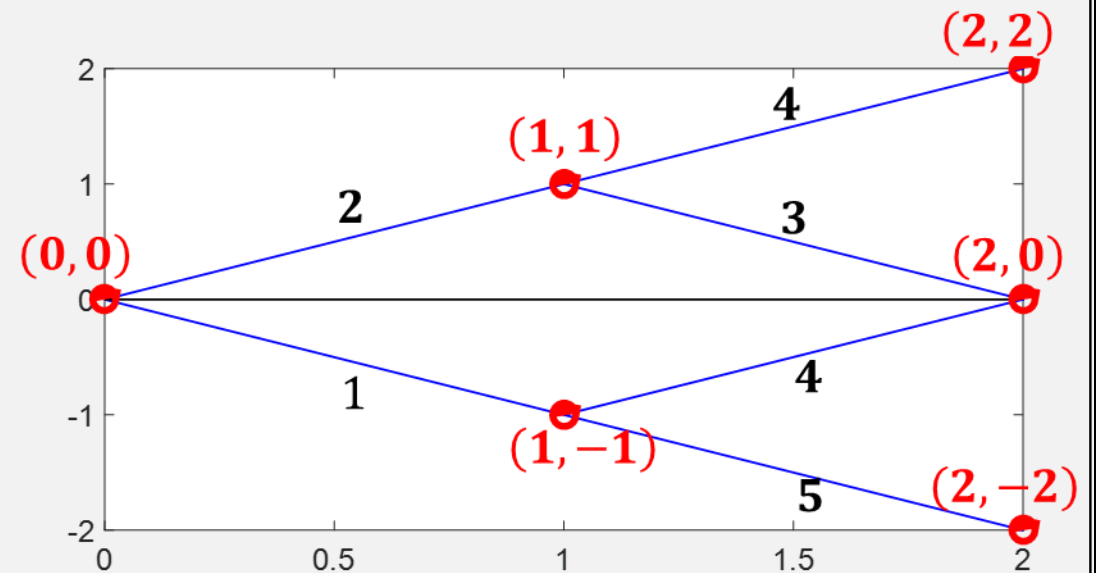
$$\begin{aligned} E(A) &= p_2 a(1,1) + (1 - p_2) \delta(1,1) \\ &= 0.3 * 4 + 0.7 * 3 = 3.3 \end{aligned}$$

$$\begin{aligned} E(\Delta) &= p_2 \delta(1,1) + (1 - p_2) \alpha(1,1) \\ &= 0.3 * 3 + 0.7 * 4 = 3.7 \end{aligned}$$

- Για τον κόμβο  $(1, -1)$ :

$$E(A) = p_2 a(1, -1) + (1 - p_2) \delta(1, -1) = 0.3 * 4 + 0.7 * 5 = 4.7$$

$$E(\Delta) = p_2 \delta(1, -1) + (1 - p_2) \alpha(1, -1) = 0.3 * 5 + 0.7 * 4 = 4.3$$



- Για τον κόμβο  $(0, 0)$  βρήκαμε:

$$E(AA) = 5.62, E(A\Delta) = 5.38$$

$$E(\Delta A) = 5.38, E(\Delta\Delta) = 5.62$$

Άρα οι βέλτιστες εντολές είναι

**AA** και **ΔΔ**.

- Για τον κόμβο  $(1, 1)$ :

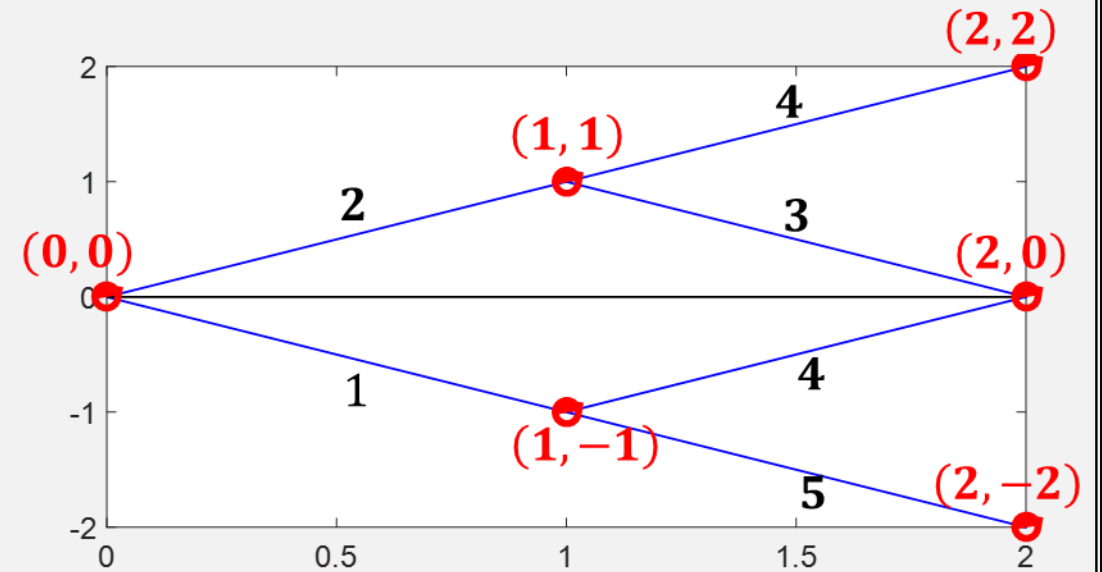
$$E(A) = 3.3, E(\Delta) = 3.7$$

Άρα η βέλτιστη εντολή είναι πάνε **δεξιά (Δ)**.

- Για τον κόμβο  $(1, -1)$ :

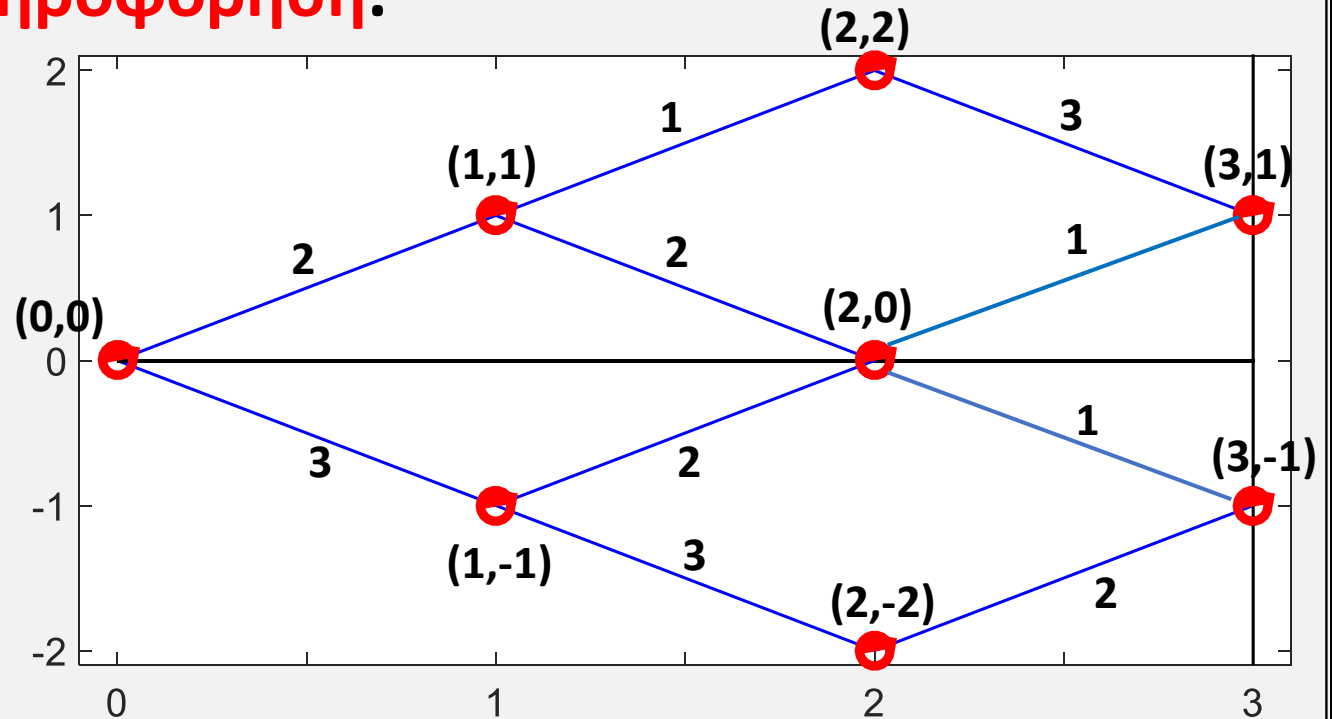
$$E(A) = 4.7, E(\Delta) = 4.3$$

Άρα η βέλτιστη εντολή είναι πάνε **αριστερά (A)**.



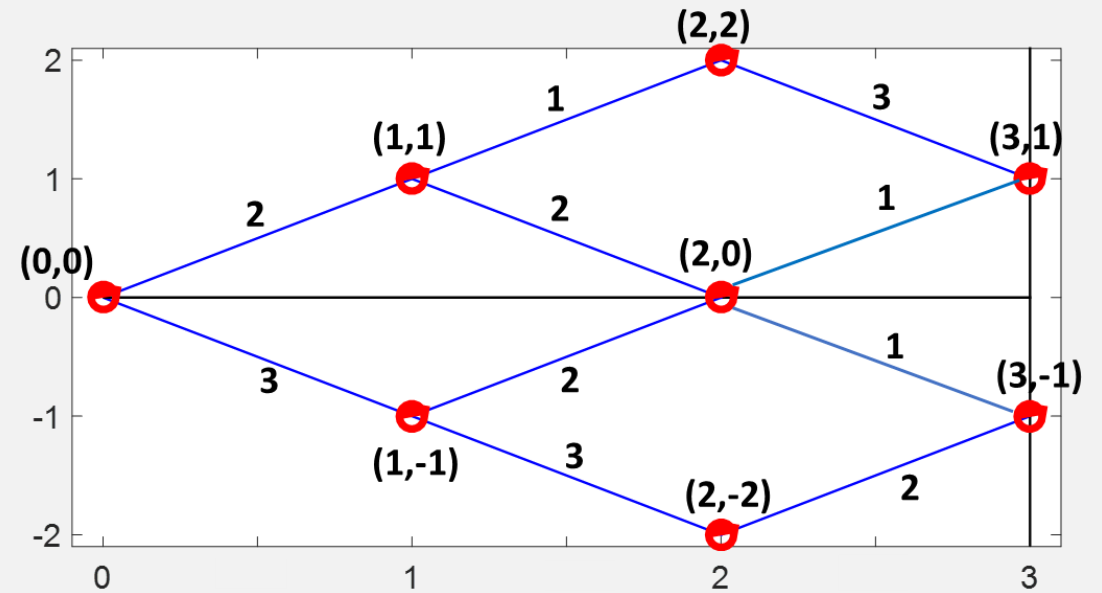
## Άσκηση

Έστω το παρακάτω δικτυωτό, όπου οι τιμές στα τόξα είναι **κόστος**. Θεωρούμε ότι η εντολή εκτελείται σε **κάθε βήμα** με πιθανότητα  $p = 0.4$ . Να βρεθεί η βέλτιστη πολιτική εντολών με την **μέθοδο ανοιχτού κύκλου με επαναπληρόρηση**.



Κόμβος (0,0):

Πιθανές εντολές	Δυνατές εντολές
ΑΑΑ	Χ
ΑΑΔ	ΑΑΔ
ΑΔΑ	ΑΔΑ
ΑΔΔ	ΑΔΔ
ΔΑΑ	ΔΑΑ
ΔΔΑ	ΔΔΑ
ΔΑΔ	ΔΑΔ
ΔΔΔ	Χ

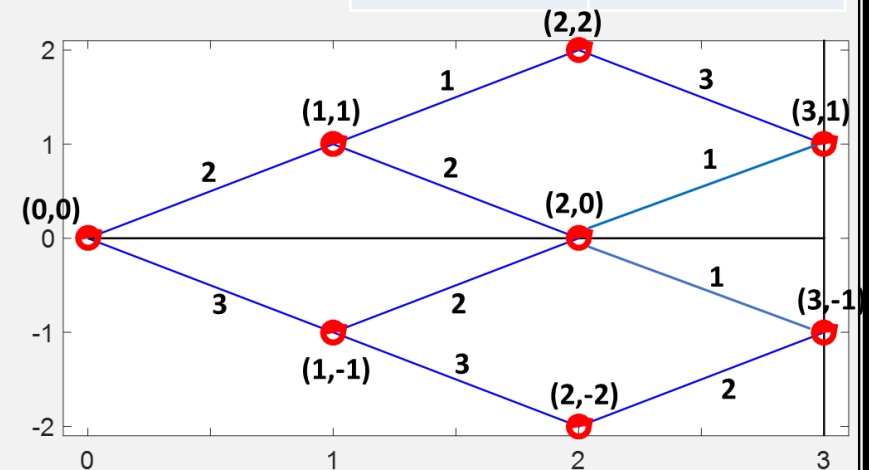


$$\begin{aligned}
 E(\text{ΑΑΔ}) &= p p \mathbf{1} (a(0,0) + a(1,1) + \delta(2,2)) + \text{ΑΑΔ} \\
 &+ p (1-p)(1-p)(a(0,0) + \delta(1,1) + a(2,0)) + \text{ΑΔΑ} \\
 &+ p (1-p)p(a(0,0) + \delta(1,1) + \delta(2,0)) + \text{ΑΔΔ} \\
 &+ (1-p)p(1-p)(\delta(0,0) + a(1,-1) + a(2,0)) + \text{ΔΑΑ} \\
 &+ (1-p)p (1-p)(\delta(0,0) + \alpha(1,-1) + \delta(2,0)) + \text{ΔΑΔ} \\
 &+ (1-p)(1-p) \mathbf{1} (\delta(0,0) + \delta(1,-1) + \alpha(2,-2)) \text{ΔΔΑ}
 \end{aligned}$$

Πιθανές Εντολές	Δυνατές κινήσεις
ΑΑΑ	Χ
ΑΑΔ	ΑΑΔ
ΑΔΑ	ΑΔΑ
ΑΔΔ	ΑΔΔ
ΔΑΑ	ΔΑΑ
ΔΑΔ	ΔΑΔ
ΔΔΑ	ΔΔΑ
ΔΔΔ	Χ

$$\begin{aligned}
 &= 0.4 * 0.4 * 1 * (2+1+3) + 0.4 * 0.6 * 0.4 * (2+2+1) + \\
 &0.4 * 0.6 * 0.6 * (2+2+1) + 0.6 * 0.4 * 0.4 * (3+2+1) + \\
 &0.6 * 0.6 * 1 * (3 + 3 + 2) + 0.6 * 0.4 * 0.6 * (3+2+1) \\
 &= \mathbf{6.48}
 \end{aligned}$$

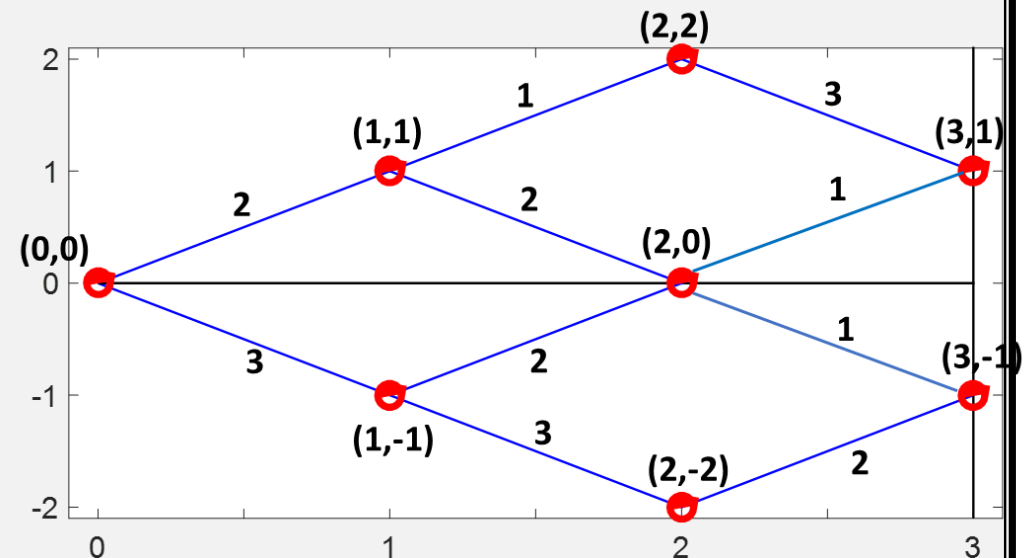
κ.ο.κ. για όλες τις δυνατές εντολές



## Κόμβος (1,1):

$$\begin{aligned} E(\mathbf{A}\Delta) &= p\mathbf{1}(a(1,1) + \delta(2,2)) + \\ &\quad (1-p)p(\delta(1,1) + \delta(2,0)) + \\ &\quad (1-p)(1-p)(\delta(1,1) + \alpha(2,0)) = \\ &= 0.4 * 1(1 + 3) + 0.6 * 0.4(2 + 1) + \\ &\quad 0.6 * 0.6(2 + 1) \\ &= \mathbf{3.4} \end{aligned}$$

Δυνατές εντολές	Δυνατές κινήσεις
ΑΑ	Χ
ΑΔ	ΑΔ
ΔΑ	ΔΑ
ΔΔ	ΔΔ



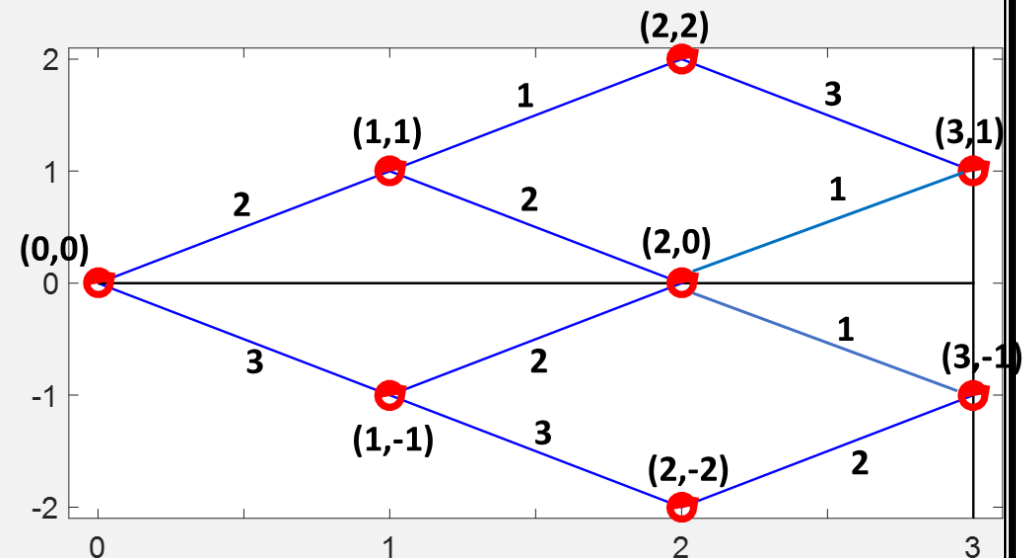


## Κόμβος (1,1):

$$\begin{aligned} E(\Delta A) &= pp(\delta(1,1) + a(2,0)) + \\ & p(1-p)(\delta(1,1) + \delta(2,0)) + \\ & (1-p)1(a(1,1) + \delta(2,2)) = \end{aligned}$$

$$\begin{aligned} &= 0.4 * 0.4(2 + 1) + 0.4 * 0.6(2 + 1) + 0.6 * 1(1 + 3) \\ &= \mathbf{3.6} \end{aligned}$$

Δυνατές εντολές	Δυνατές κινήσεις
ΑΑ	Χ
ΑΔ	ΑΔ
ΔΑ	ΔΑ
ΔΔ	ΔΔ

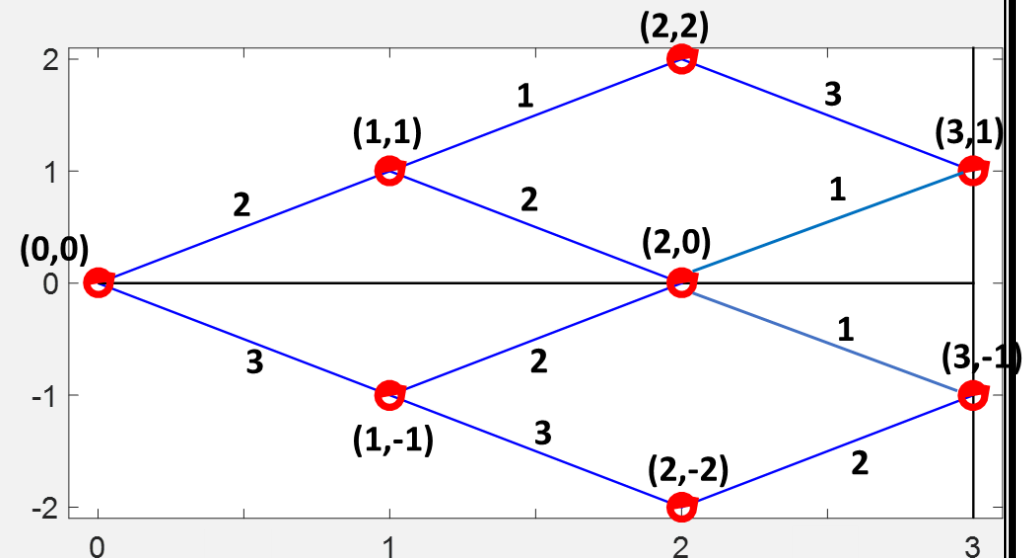


## Κόμβος (1,1):

$$\begin{aligned} E(\Delta\Delta) &= p p (\delta(1,1) + \delta(2,0)) + \\ &\quad p(1-p)(\delta(1,1) + a(2,0)) + \\ &\quad (1-p) \mathbf{1} (a(1,1) + \delta(2,2)) = \end{aligned}$$

$$\begin{aligned} &= 0.4 * 0.4(2 + 1) + 0.4 * 0.6(2 + 1) + 0.6 \\ &* 1(1 + 3) = \mathbf{3.6} \end{aligned}$$

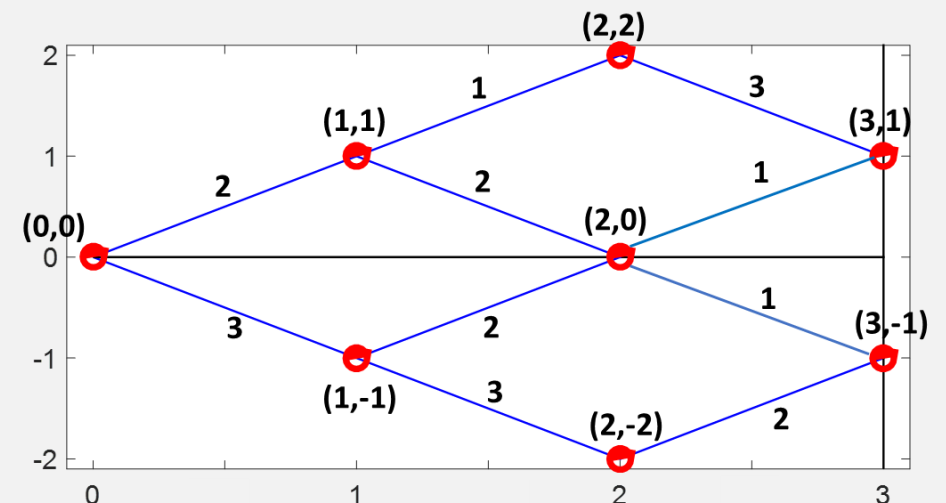
Δυνατές εντολές	Δυνατές κινήσεις
ΑΑ	Χ
ΑΔ	ΑΔ
ΔΑ	ΔΑ
ΔΔ	ΔΔ



## Κόμβος (1,-1):

$$\begin{aligned} E(\text{AA}) &= pp(a(1, -1) + a(2,0)) + \\ &\quad p(1 - p)(a(1, -1) + \delta(2,0)) + \\ &\quad (1 - p)\mathbf{1}(\delta(1, -1) + \alpha(2, -2)) = \\ &= 0.4 * 0.4(2 + 1) + 0.4 * 0.6(2 + 1) + 0.6 * \mathbf{1}(3 + 2) \\ &= \mathbf{4.2} \end{aligned}$$

Δυνατές εντολές	Δυνατές κινήσεις
AA	AA
AΔ	AΔ
ΔA	ΔA
ΔΔ	X

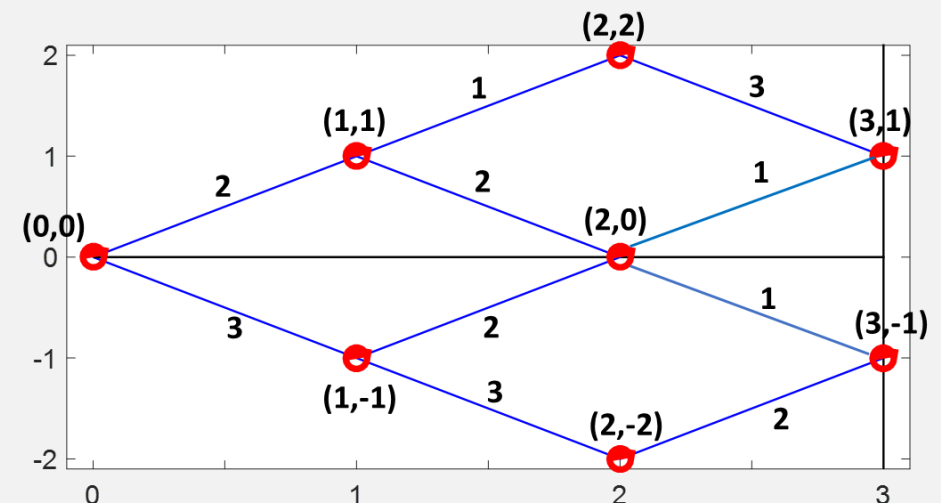


## Κόμβος (1,-1):

$$\begin{aligned} E(\mathbf{A}\Delta) &= pp(a(1, -1) + \delta(2,0)) + \\ &\quad p(1 - p)(a(1, -1) + \alpha(2,0)) + \\ &\quad (1 - p)\mathbf{1}(\delta(1, -1) + \alpha(2, -2)) = \\ &= 0.4 * 0.4(2 + 1) + 0.4 * 0.6(2 + 1) + 0.6 * \mathbf{1}(3 + 2) \\ &= \mathbf{4.2} \end{aligned}$$

$$\begin{aligned} E(\mathbf{\Delta A}) &= p\mathbf{1}(\delta(1, -1) + \alpha(2, -2)) + \\ &\quad (1 - p)p(a(1, -1) + \alpha(2,0)) + \\ &\quad (1 - p)(1 - p)(a(1, -1) + \delta(2,0)) = \\ &= 0.4 * \mathbf{1}(3 + 2) + 0.6 * 0.4(2 + 1) \\ &\quad + 0.6 * 0.6(2 + 1) = \mathbf{3.8} \end{aligned}$$

Δυνατές εντολές	Δυνατές κινήσεις
ΑΑ	ΑΑ
ΑΔ	ΑΔ
ΔΑ	ΔΑ
ΔΔ	Χ



## Κόμβος (2,2)

$$E(\Delta) = 1 * \delta(2,2) = 3$$

Δυνατές εντολές	Δυνατές κινήσεις
A	X
Δ	Δ

## Κόμβος (2,0)

$$E(A) = pa(2,0) + (1 - p)\delta(2,0) \\ = 0.4*1 + 0.6*1 = 1$$

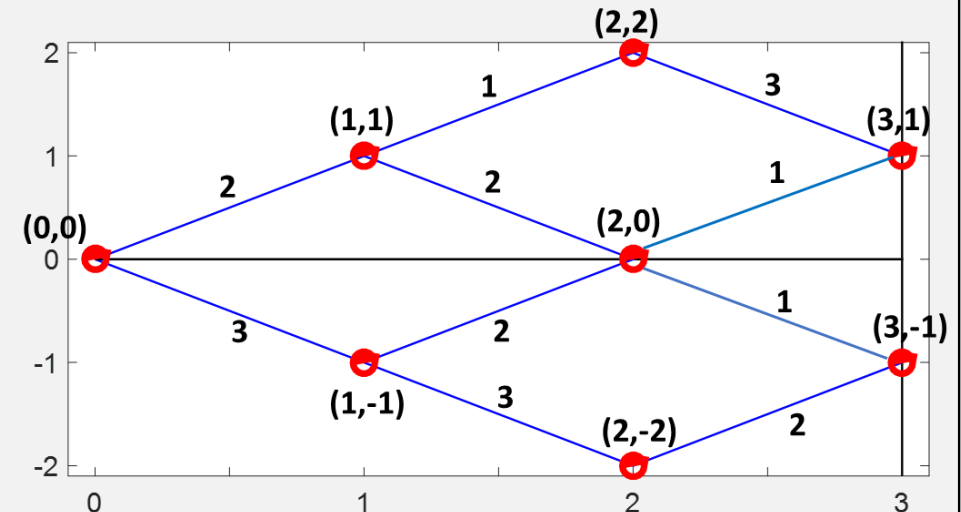
Δυνατές εντολές	Δυνατές κινήσεις
A	A
Δ	Δ

$$E(\Delta) = p\delta(2,0) + (1 - p)a(2,0) = \\ 0.4*1 + 0.6*1 = 1$$

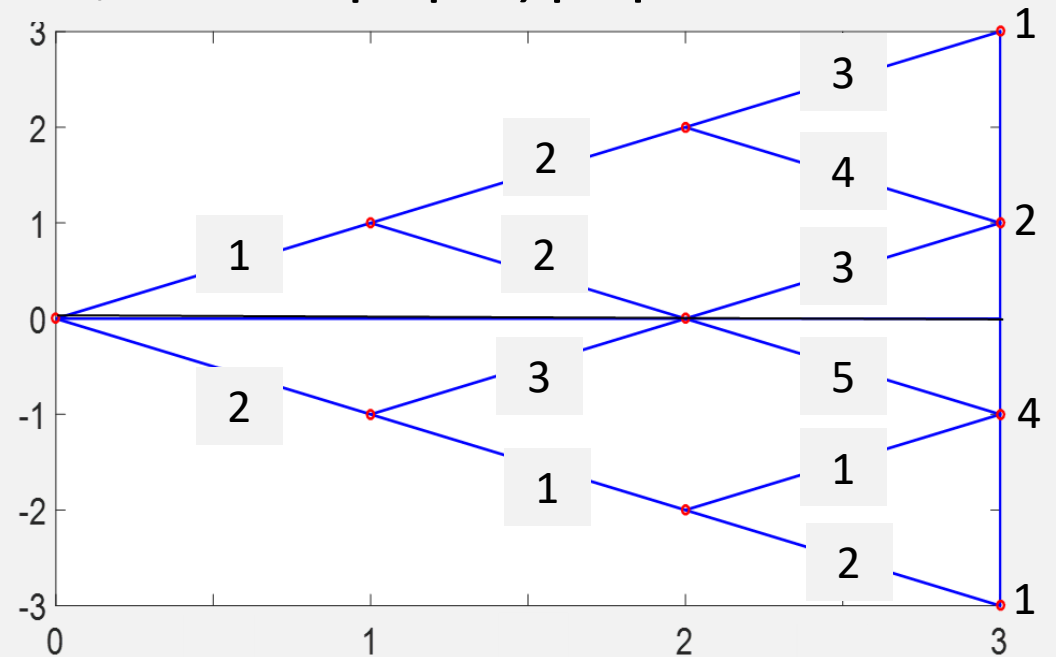
## Κόμβος (2,-2)

$$E(A) = 1 * a(2, -2) = \\ 1*2 = 2$$

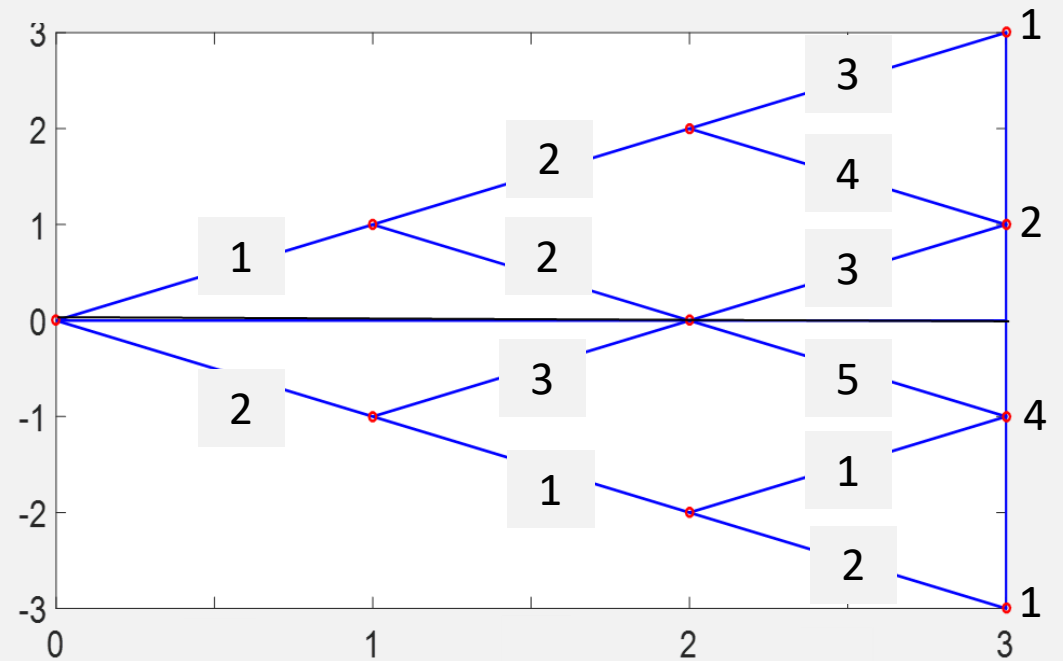
Δυνατές εντολές	Δυνατές κινήσεις
A	A
Δ	X



**Άσκηση** Έστω το παρακάτω δικτυωτό, όπου οι τιμές στα **τόξα** και στους **τελικούς κόμβους** είναι **κέρδος**. Θεωρούμε ότι μια εντολή εκτελείται με πιθανότητα  $p = 0.6$ . Θεωρούμε ως προϋπόθεση το γεγονός ότι όταν εκτελούνται διαδρομές με δύο τόξα προς τα δεξιά, το κέρδος της διαδρομής είναι 0 και δεν δίνονται εντολές για τέτοιες διαδρομές. Να βρεθούν με την **μέθοδο ανοιχτού κύκλου**, οι διαδρομές μεγίστου και ελαχίστου κέρδους.



Εντολές κίνησης	Δυνατές κινήσεις
ΑΑΑ	ΑΑΑ
ΑΑΔ	ΑΑΔ
ΑΔΑ	ΑΔΑ
Χ	ΑΔΔ
ΔΑΑ	ΔΑΑ
Χ	ΔΔΑ
Χ	ΔΑΔ
Χ	ΔΔΔ



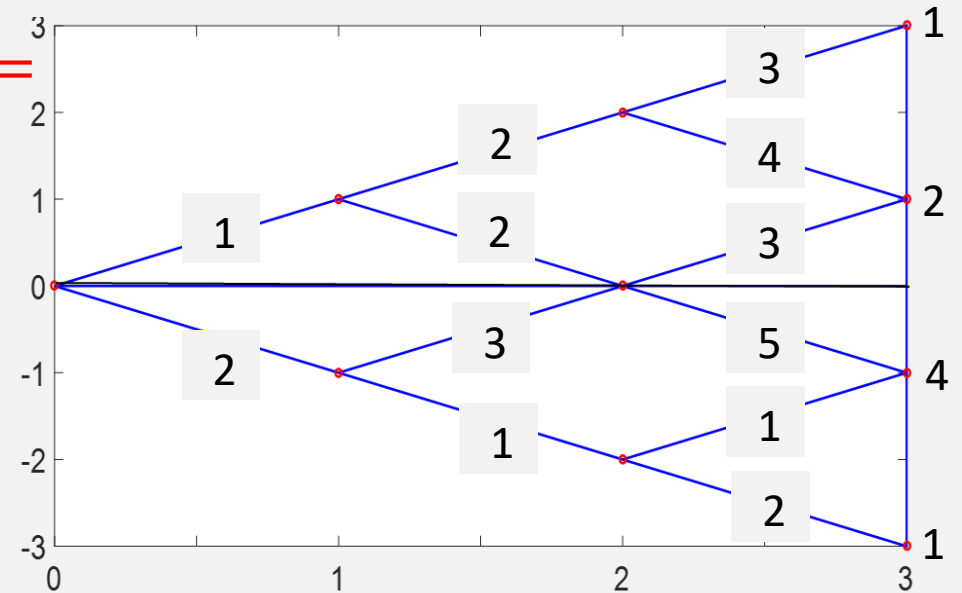
Οι εντολές κίνησης που δίνονται είναι:

**ΑΑΑ, ΑΑΔ, ΑΔΑ, ΔΑΑ** (αποκλείονται οι ΑΔΔ, ΔΑΔ, ΔΔΑ, ΔΔΔ)

$$\begin{aligned} E(\mathbf{A}\mathbf{A}\mathbf{A}) = & ppp(a(0,0) + a(1,1) + a(2,2) + 1) + \\ & ppq(a(0,0) + a(1,1) + \delta(2,2) + 2) + \\ & pqp(a(0,0) + \delta(1,1) + \alpha(2,0) + 2) + \\ & qpp(\delta(0,0) + \alpha(1,-1) + \alpha(2,0) + 2) \end{aligned}$$

$$+pqq * 0 + qrp * 0 + qqr * 0 + qqq * 0 =$$

$$=5.4$$



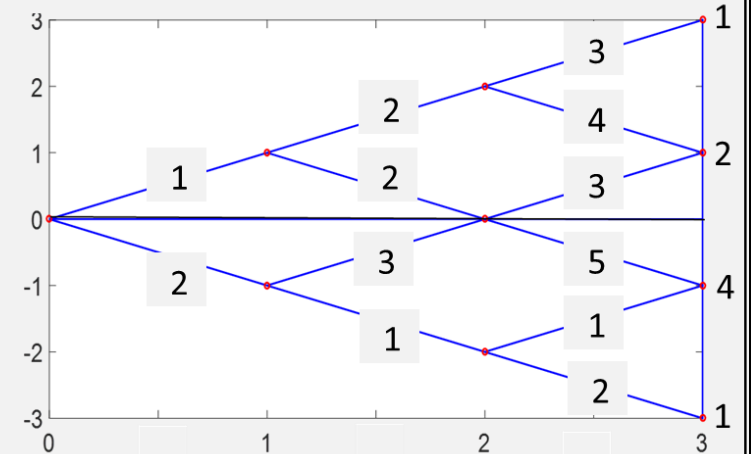


Ομοίως

$$\begin{aligned} E(\mathbf{A}\mathbf{A}\mathbf{A}) &= ppp(a(0,0) + a(1,1) + \delta(2,2) + 2) + \\ &\quad ppq(a(0,0) + a(1,1) + a(2,2) + 1) + \\ &\quad pqq(a(0,0) + \delta(1,1) + a(2,0) + 2) + \\ &\quad qpq(\delta(0,0) + a(1, -1) + a(2,0) + 2) = \mathbf{4.68} \end{aligned}$$

$$\begin{aligned} E(\mathbf{A}\mathbf{D}\mathbf{A}) &= ppp(a(0,0) + \delta(1,1) + a(2,0) + 2) + \\ &\quad pqp(a(0,0) + a(1,1) + a(2,2) + 1) + \\ &\quad pqq(a(0,0) + a(1,1) + \delta(2,2) + 2) + \\ &\quad qqp(\delta(0,0) + a(1, -1) + a(2,0) + 2) = \mathbf{4.56} \end{aligned}$$

$$\begin{aligned} E(\mathbf{D}\mathbf{A}\mathbf{A}) &= ppp(\delta(0,0) + a(1, -1) + a(2,0) + 2) + \\ &\quad qpp(a(0,0) + a(1,1) + a(2,2) + 1) + \\ &\quad qpq(a(0,0) + a(1,1) + \delta(2,2) + 2) + \\ &\quad qqp(a(0,0) + \delta(1,1) + a(2,0) + 2) = \mathbf{4.8} \end{aligned}$$



Άρα το μέγιστο αναμενόμενο κέρδος είναι **5.4** και το ελάχιστο **4.56**.

## 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.