

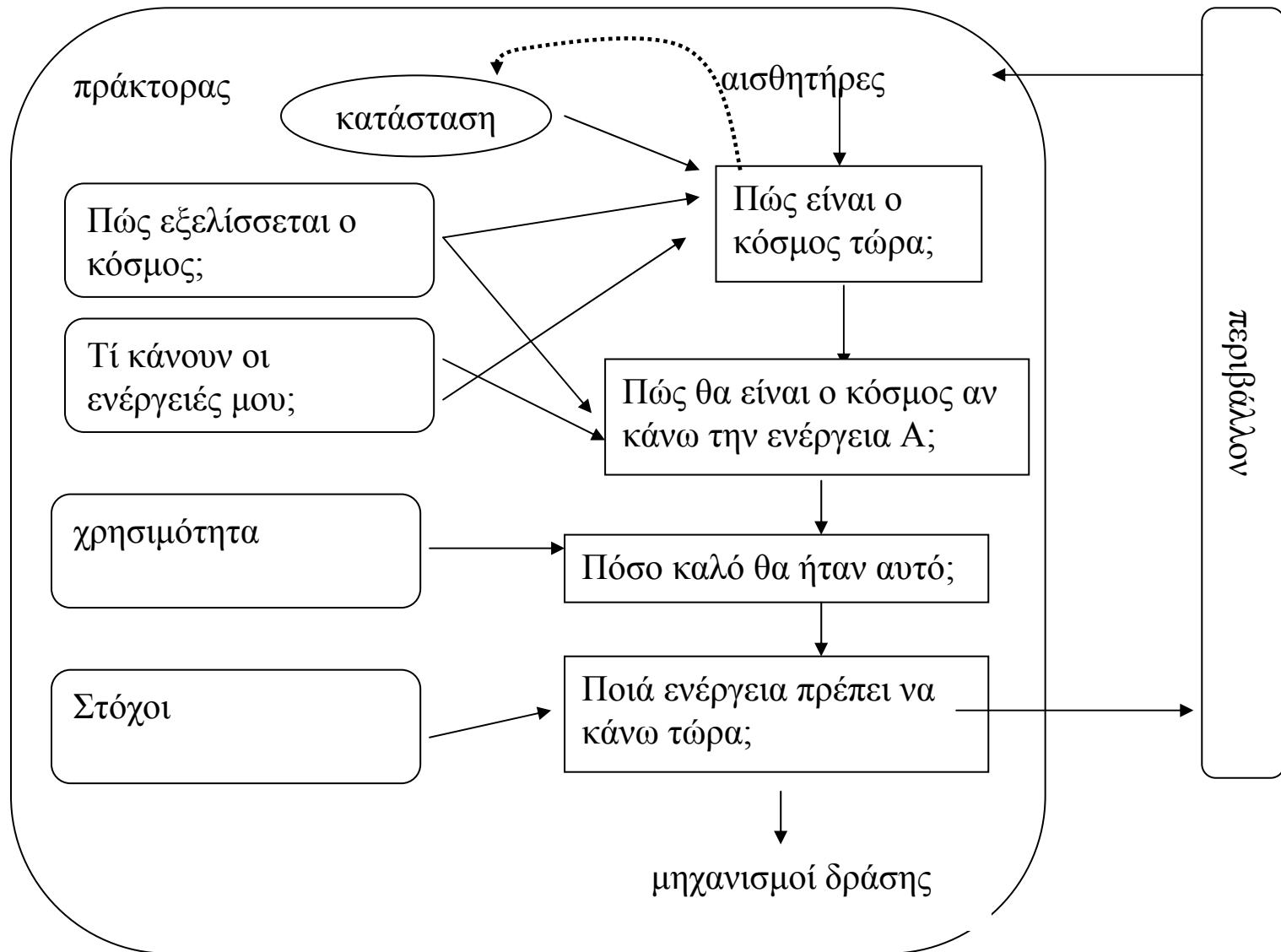
# Εισαγωγή στη Θεωρία Αποφάσεων

Λήψη απλών αποφάσεων για έναν πράκτορα

# Ορθολογικές αποφάσεις

- Ένας πράκτορας βασισμένος στη λογική
  - Έχει ένα στόχο (μια κατάσταση περιβάλλοντος που θέλει να πετύχει)
  - Καταστρώνει ένα σχέδιο δράσης (μια ακολουθία ενεργειών) επιλέγοντας στο πλάνο του ενέργειες από το ρεπερτόριό του, με κριτήριο το **αν αυτές επιτυγχάνουν ή όχι το στόχο.**

## Πράκτορες βασισμένοι στη χρησιμότητα



Η **κατάσταση** του πράκτορα είναι η αναπαράσταση των πιθανοτήτων για όλες τις δυνατές πραγματικές καταστάσεις του κόσμου.

# Αβεβαιότητα και ορθολογικές αποφάσεις

- **Πηγές αβεβαιότητας;**
  - Κάθε ενέργεια έχει **πιθανότητα** επιτυχίας και αποτυχίας.
    - Σφάλματα μηχανισμών δράσης, άλλοι πράκτορες...
  - Κάθε πλάνο έχει πιθανότητα επιτυχίας και αποτυχίας.
- Ένας πράκτορας που αποφασίζει υπό αβεβαιότητα πρέπει να έχει **προτιμήσεις** μεταξύ των αποτελεσμάτων των ενεργειών/πλάνων του.
- **Θεωρία χρησιμότητων:** πώς αναπαριστούνται προτιμήσεις και πώς χρησιμοποιείται αυτή η αναπαράσταση στη συλλογιστική ενός πράκτορα;
- **Θεωρία αποφάσεων = θεωρία χρησιμότητων + θεωρία πιθανοτήτων**
  - Ένας πράκτορας είναι ορθολογικός αν και μόνο αν επιλέγει την ενέργεια που παρέχει την **υψηλότερη αναμενόμενη χρησιμότητα** (meu)

# Θεωρία Προτιμήσεων/χρησιμοτήτων

- **Συνάρτηση χρησιμότητας:** Σε κάθε κατάσταση του κόσμου αποδίδεται μια αριθμητική τιμή μέσω της  $u : S \rightarrow \mathbb{R}$ , η οποία εκφράζει πόσο επιθυμητή είναι η κατάσταση. Η συνάρτηση χρησιμότητας ουσιαστικά εκφράζει τις **προτιμήσεις** του πράκτορα ανάμεσα στα αποτελέσματα των ενεργειών του.

$A \succ B$  ο πράκτορας προτιμά το A από το B αν και μόνο αν  $u(A) > u(B)$

$A \sim B$  ο πράκτορας είναι αδιάφορος μεταξύ A και B αν και μόνο αν  $u(A) = u(B)$

$A \# B$  ο πράκτορας προτιμά το A από το B ή είναι αδιάφορος

Δηλαδή  $A \# B \equiv (A \succ B) \vee (A \sim B)$

Αν οι ενέργειες του πράκτορα είναι ντετερμινιστικές, τα A και B είναι οι συγκεκριμένες, πλήρως προσδιορισμένες, καταστάσεις περιβάλλοντος που προκαλούνται από τις ενέργειές του.

## Μη ντετερμινιστικές ενέργειες: Λοταρίες

- Στη γενική περίπτωση οι ενέργειες του πράκτορα είναι μη ντετερμινιστικές οπότε τα  $A$  και  $B$  είναι **λοταρίες** (οι ντετερμινιστικές ενέργειες περιγράφονται επίσης ως λοταρίες που έχουν ένα μόνο αποτέλεσμα).
- Λοταρία  $L$  είναι μια κατανομή πιθανοτήτων  $\{p_1, \dots, p_n\}$  ως προς ένα σύνολο αποτελεσμάτων που μπορεί να προκύψουν  $\{C_1, \dots, C_n\}$ :

$$L = [p_1, C_1; p_2, C_2; \dots; p_n, C_n]$$

## Αρχή της μέγιστης αναμενόμενης χρησιμότητας (meu)

- Έστω  $A$  μια μη-ντετερμινιστική ενέργεια και  $result_i(A)$  οι δυνατές καταστάσεις αποτελέσματος της  $A$ , όπου ο δείκτης  $i$  κυμαίνεται μεταξύ των διαφόρων αποτελεσμάτων.
- Πριν την εκτέλεση της ενέργειας  $A$  ο πράκτορας αποδίδει πιθανότητα  $P(result_i(A) | do(A), E)$  σε κάθε αποτέλεσμα, με βάση τις μαρτυρίες  $E$  που διαθέτει σχετικά με τον κόσμο (η πρόταση  $do(A)$  σημαίνει ότι η ενέργεια  $A$  εκτελείται στην τρέχουσα κατάσταση).
- Η αναμενόμενη χρησιμότητα ( $eu$ ) της ενέργειας  $A$  με βάση τις μαρτυρίες  $E$  είναι

$$eu(A | E) = \sum_i P(result_i(A) | do(A), E) \cdot u(result_i(A))$$

- Αναμενόμενη χρησιμότητα λοταρίας  
Η χρησιμότητα μιας λοταρίας είναι το άθροισμα των γινομένων που προκύπτουν από τον πολλαπλασιασμό της πιθανότητας κάθε αποτελέσματος επί τη χρησιμότητα αυτού του αποτελέσματος:

$$u([p_1, C_1; \dots; p_n, C_n]) = \sum_i p_i u(C_i)$$

•

# Παράδειγμα

Έστω ότι μια διαδικασία απόφασης έχει τέσσερις καταστάσεις  $\{\Sigma 1, \Sigma 2, \Sigma 3, \Sigma 4\}$  με χρησιμότητες  $+1, -1, 0$  και  $+2$  αντίστοιχα. Στις καταστάσεις  $\Sigma 1$  και  $\Sigma 2$  είναι δυνατές τρεις ενέργειες,  $\alpha, \beta$ , και  $\gamma$ . Οι καταστάσεις  $\Sigma 3$  και  $\Sigma 4$  είναι τερματικές. Το μοντέλο μετάβασης προσδιορίζεται από λοταρίες ως εξής:

Στην κατάσταση  $\Sigma 1$  η ενέργεια  $\alpha$  έχει τα εξής αποτελέσματα  $[0.5, \Sigma 3; 0.3, \Sigma 2; 0.2, \Sigma 1]$ , η ενέργεια  $\beta$  έχει τα εξής αποτελέσματα  $[0.8, \Sigma 1; 0.2, \Sigma 4]$ , και η ενέργεια  $\gamma$   $[0.8, \Sigma 3; 0.2, \Sigma 4]$ .

Στην κατάσταση  $\Sigma 2$  η ενέργεια  $\alpha$  έχει αποτελέσματα  $[0.8, \Sigma 4; 0.2, \Sigma 2]$ , η ενέργεια  $\beta$   $[0.2, \Sigma 3; 0.8, \Sigma 2]$  και η ενέργεια  $\gamma$   $[0.5, \Sigma 1; 0.5, \Sigma 3]$ .

## Ερωτήσεις

- Τι προτιμήσεις έχει ένας πράκτορας σε αυτό το περιβάλλον;
- Ποια ενέργεια οφείλει να επιλέξει ένας πράκτορας σε αυτό το περιβάλλον, όταν βρίσκεται στην κατάσταση  $\Sigma 1$ , και όταν βρίσκεται στην κατάσταση  $\Sigma 2$ .