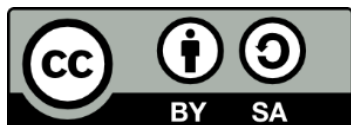


ΘΕΩΡΙΑ ΠΑΙΓΝΙΩΝ

Ενότητα 9: Απείρως επαναλαμβανόμενα παίγνια

Ρεφανίδης Ιωάννης
Τμήμα Εφαρμοσμένης Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Μακεδονίας» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Infinitely repeated games

Απείρως επαναλαμβανόμενα παιχνίδια

Γενικά (1/2)

- Όταν ένα παιχνίδι επαναλαμβάνεται άπειρες φορές, δεν μπορούμε να θεωρήσουμε ως συνολικό όφελος το άθροισμα για τα επί μέρους παιχνίδια.
- Υπάρχουν διάφοροι τρόποι αντιμετώπισης των απειροαθροισμάτων, όπως:
 - Μέσος όρος
 - Προεξόφληση μελλοντικών αποδόσεων
- Θα χρησιμοποιήσουμε την τελευταία μέθοδο, σύμφωνα με την οποία:
 - Υπάρχει ένας παράγοντας προεξόφλησης (discount factor) $\delta < 1$.
 - Η σημερινή αξία μιας απόδοσης u_t τη χρονική στιγμή t είναι $\delta_t u_t$.
 - Αν θεωρήσουμε σταθερή μέση απόδοση u σε κάθε χρονική στιγμή, η συνολική απόδοση του παιχνιδιού είναι:

$$u_{ολ} = \sum_{t=0}^{\infty} u \cdot \delta^t = u \cdot \sum_{t=0}^{\infty} \delta^t = \frac{u}{1-\delta}$$

Γενικά (2/2)

- Ο παράγοντας προεξόφλησης μπορεί να ερμηνευτεί με πολλούς τρόπους:
 - Εάν το όφελος είναι χρηματικές αποδόσεις, τότε ο παράγοντας προεξόφλησης μπορεί να θεωρηθεί ότι σχετίζεται με το επιτόκιο της τράπεζας.
 - Εάν δεν είναι γνωστό το πλήθος των επαναλήψεων του παιχνιδιού, ο παράγοντας προεξόφλησης είναι η πιθανότητα που δίνουμε κάθε φορά στο να υπάρξει ακόμη μια επανάληψη.
 - Γενικότερα ένας παίκτης προτιμά μια άμεση απόδοση έναντι μιας μελλοντικής απόδοσης.

Στρατηγική ενεργοποίησης (1/5)

- Έστω το παιχνίδι ΔΦ (στην απλή του μορφή, με αποδόσεις τέτοιες ώστε μεγαλύτεροι αριθμοί να είναι καλύτεροι) το οποίο παίζεται απείρως επαναλαμβανόμενα.
- Έστω το παρακάτω ζεύγος στρατηγικών:
 - Οι παίκτες συμφωνούν να παίξουν (ΔΟ, ΔΟ) για πάντα. Εάν όμως κάποιος παίκτης ομολογήσει, στο εξής και οι δύο θα επιλέγουν Ο για πάντα.
- Η παραπάνω στρατηγική (κοινή και για τους δύο παίκτες) ονομάζεται στρατηγική ενεργοποίησης (trigger strategy).
- Η στρατηγική χωρίζεται σε δύο μέρη:
 - Στο μέρος πρώτο, όπου κανείς παίκτης δεν ομολογεί, κάθε παίκτης ενδεχομένως θα είχε την τάση να ομολογήσει.
 - Στο δεύτερο μέρος, όπου και οι δύο παίκτες ομολογούν, κανείς παίκτης δεν έχει την τάση (από μόνος του) να μην-ομολογήσει.

Στρατηγική ενεργοποίησης (2/5)

- Έστω ότι ύστερα από t-1 γύρους, όπου οι δύο παίκτες τήρησαν τη συμφωνία, είμαστε στην επανάληψη t.
- Έστω ότι ένας παίκτης, π.χ. ο A, σκέφτεται να ομολογήσει.
 - Το γεγονός ότι έχουν προηγηθεί t-1 γύροι δεν αλλάζει ουσιαστικά σε τίποτα τη συμπεριφορά των παικτών στον t-1 γύρο. Το παιχνίδι που ξεκινά στον t γύρο είναι ίδιο με αυτό που ξεκίνησε στον πρώτο γύρο, εφόσον οι παίκτες τηρούν τη συμφωνία.
- Το άμεσο όφελός του θα είναι $\delta^t \cdot 15$ (αντί για $\delta^t \cdot 14$ αν τηρούσε τη συμφωνία).
- Το όφελός του από τους επόμενους γύρους θα είναι:

$$\delta^{t+1} \cdot 10 \cdot \frac{1}{1-\delta}$$

	B		
A		O	ΔO
	O	10, 10	15, 0
	ΔO	0, 15	14 14

Προσοχή:
Μεγαλύτερες αποδόσεις θεωρούνται καλύτερες.

Στρατηγική ενεργοποίησης (3/5)

- Εάν δεν έσπαγε τη συμφωνία, τότε το όφελός του για τους επόμενους (άπειρους) γύρους θα ήταν:

$$\delta^t \cdot 14 \cdot \frac{1}{1-\delta}$$

- Θα συνέφερε τον A να σπάσει τη συμφωνία, εάν ίσχυε:

$$\delta^t \cdot 14 \cdot \frac{1}{1-\delta} < \delta^t \cdot 15 + \delta^{t+1} \cdot 10 \cdot \frac{1}{1-\delta}$$

ή

$$\delta < \frac{1}{5}$$

	B		
A		O	ΔO
	O	10, 10	15, 0
	ΔO	0, 15	14 14

Στρατηγική ενεργοποίησης (4/5)

- Άρα, για $\delta > 1/5$ δεν συμφέρει κανέναν παίκτη να σπάσει τη συμφωνία.
- Σε αυτή την περίπτωση, η στρατηγική ενεργοποίησης είναι σημείο ισορροπίας, αφού κανείς παίκτης δεν τολμά να την αλλάξει μεμονωμένα!
- Γενικότερα, η στρατηγική ενεργοποίησης επιβραβεύει την καλή συμπεριφορά ή, ισοδύναμα, τιμωρεί την κακή συμπεριφορά, όταν ο παράγοντας προεξόφλησης τείνει στη μονάδα, όσο δηλαδή αυξάνει η βαρύτητα του μέλλοντος.
 - Αντίστοιχη συμπεριφορά δεν υπάρχει στα πεπερασμένα παιχνίδια, γιατί το μέλλον είναι πεπερασμένο.

	B		
A		O	ΔO
	O	10, 10	15, 0
	ΔO	0, 15	14 14

Στρατηγική ενεργοποίησης (5/5)

- Η στρατηγική ενεργοποίησης που είδαμε δεν είναι η μοναδική. Μια παραλλαγή της θα μπορούσε να είναι η εξής:
 - Αρχικά οι παίκτες παίζουν εναλλάξ (Ο,ΔΟ) και (ΔΟ,Ο). Εάν κάποια στιγμή ένας παίκτης αθετήσει τη συμφωνία, στη συνέχεια και οι δύο παίκτες παίζουν (Ο,Ο) για πάντα.
- Η παραπάνω στρατηγική βέβαια δεν αποτελεί σημείο ισορροπίας Nash, μιας και οι παίκτες κερδίζουν κατά μέσο όρο 7.5 ανά γύρο, λιγότερο από ότι κερδίζουν στο σημείο ισορροπίας Nash.
- Ανάλογα με τις επιμέρους αποδόσεις του παιχνιδιού γύρου, κάποια εκ των δύο στρατηγικών ενεργοποίησης που είδαμε (στην προκειμένη περίπτωση η πρώτη) έχει καλύτερη συνολική απόδοση.

		B	
		Ο	ΔΟ
A	Ο	10, 10	15, 0
	ΔΟ	0, 15	14 14

Επιεικής στρατηγική

- Η επιεικής στρατηγική είναι παραλλαγή της στρατηγικής ενεργοποίησης: Η «τιμωρία» λήγει μετά από κάποιον αριθμό γύρων.
 - Οι παίκτες συμφωνούν να παίξουν (ΔΟ, ΔΟ) για πάντα. Εάν όμως κάποιος παίκτης ομολογήσει, ο άλλος παίκτης επιλέξει Ο για τους επόμενους Τ γύρους.
- Για να αποτελεί η παραπάνω στρατηγική σημείο ισορροπίας, πρέπει να επιλέξουμε το Τ με τέτοιο τρόπο ώστε κανέναν παίκτη να μην τον συμφέρει να αθετήσει τη συμφωνία.
- Μετά από (αρκετές...) πράξεις βρίσκεται η αναγκαία συνθήκη για να αποτελεί η παραπάνω στρατηγική σημείο ισορροπίας.

		B	
		O	ΔO
A	O	10, 10	15, 0
	ΔO	0, 15	14 14

Το κοινό θεώρημα (1/2)

- Το ερώτημα που τίθεται για τα απείρως επαναλαμβανόμενα παιχνίδια είναι το εξής:
 - Πόσοι (και ποιοι) είναι εκείνοι οι συνδυασμοί στρατηγικών που οδηγούν σε ισορροπία;
- Η απάντηση είναι: Πάρα πολλοί.
- Θα αποδείξουμε για το $\Delta\Phi$.
- Ορίζουμε ως κύκλο συμπεριφοράς μια ακολουθία γύρων, όπου οι δύο παίκτες για T_1 γύρους επιλέγουν $(\Delta O, \Delta O)$, στη συνέχεια για T_2 γύρους επιλέγουν (O, O) , στη συνέχεια για T_3 γύρους επιλέγουν $(O, \Delta O)$ και τέλος για T_4 γύρους επιλέγουν $(O, \Delta O)$.
 - Πρέπει να ισχύει $T = T_1 + T_2 + T_3 + T_4 > 0$

Το κοινό θεώρημα (2/2)

- Ένας κύκλος συμπεριφοράς χαρακτηρίζεται *ατομικά λογικός* (individually rational) εάν το όφελος για κάθε έναν από τους δύο παίκτες για όλο τον κύκλο είναι μεγαλύτερο ή ίσο από ό,τι αν επέλεγαν συνέχεια $(0,0)$.
- Το κοινό θεώρημα (folk theorem): Από κάθε ατομικά λογικό κύκλο μπορεί να κατασκευαστεί μια τέλεια ισορροπία υποπαιγνίου, εάν συνδυάσουμε τον κύκλο με την στρατηγική ενεργοποίησης και έναν συντελεστή προεξόφλησης κοντά στη μονάδα.
 - Φυσικά ο συντελεστής προεξόφλησης είναι μια «φυσική σταθερά» του προβλήματος και δεν είναι στο χέρι μας να τον επηρεάσουμε.
- Όσο πιο κοντά στη μονάδα είναι ο συντελεστής προεξόφλησης, τόσο περισσότερους συνδυασμούς στρατηγικών που να αποτελούν ισορροπία μπορούμε να κατασκευάσουμε.
 - Αν λάβουμε υπόψη και την επιεική στρατηγική, οι συνδυασμοί μπορεί να γίνουν άπειροι.

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

