

## Εργασία Εξαμήνου - Φάση Γ

Ημερομηνία Παράδοσης (11/1/2026)

### Blackjack

Το αντικείμενο του δημοφιλούς παιχνιδιού καρτών καζίνο Μπλάκτζακ είναι να αποκτήσετε κάρτες όπου το άθροισμα των οποίων είναι όσο το δυνατόν μεγαλύτερο, χωρίς να υπερβαίνει το 21. Όλες οι κάρτες με πρόσωπο μετράνε ως 10, ενώ ένας άσος μπορεί να μετρήσει είτε ως 1 ή ως 11. Το παιχνίδι ξεκινά με δύο φύλα που μοιράζονται τόσο ο ντίλερ όσο και ο παίκτης. Ένα από τα χαρτιά του ντίλερ είναι φανερό και το άλλο κρυμμένο. Αν ο παίκτης έχει 21 αμέσως (άσος και 10), τότε ονομάζεται natural. Επομένως ο παίκτης κερδίζει, εκτός αν ο ντίλερ έχει επίσης natural, οπότε το παιχνίδι είναι ισοπαλία. Εάν ο παίκτης δεν έχει natural, τότε μπορεί να ζητήσει επιπλέον κάρτες, μία προς μία (hits), μέχρι να σταματήσει (sticks) ή να υπερβεί το 21 (καίγεται και χάνει το παιχνίδι). Αν κάνει stick τότε έρχεται η σειρά του ντίλερ. Ο ντίλερ κάνει hit ή stick σύμφωνα με μια σταθερή στρατηγική: κάνει stick όταν έχει συγκεντρώσει 17 ή μεγαλύτερο του 17, διαφορετικά κάνει hit. Αν ο ντίλερ καεί, τότε κερδίζει ο παίκτης. διαφορετικά, το αποτέλεσμα — νίκη, ήττα ή ισοπαλία — καθορίζεται από το τελικό συγκεντρωτικό ποσό του οποίου είναι πιο κοντά στο 21.

Κάθε παιχνίδι Blackjack είναι ένα επεισόδιο. Οι ανταμοιβές +1, -1, και 0 δίνονται για νίκη, ήττα, και ισοπαλία, αντίστοιχα. Όλες οι ανταμοιβές μέσα σε ένα παιχνίδι είναι μηδενικές και δεν κάνουμε έκπτωση ( $\gamma = 1$ ). Επομένως, οι εν λόγω τερματικές ανταμοιβές είναι επίσης οι επιστροφές. Οι ενέργειες του παίκτη είναι είτε hit ή stick. Οι καταστάσεις εξαρτώνται από τις κάρτες του παίκτη και την κάρτα του ντίλερ. Υποθέτουμε ότι τα φύλα μοιράζονται από ένα άπειρο deck (δηλαδή, με αντικατάσταση), έτσι ώστε να μην υπάρχει κανένα πλεονέκτημα για την παρακολούθηση των καρτών που έχουν ήδη μοιραστεί. Αν ο παίκτης κατέχει άσο που θα μπορούσε να μετρήσει ως 11 χωρίς να καεί, τότε ο άσος λέγεται ότι είναι usable. Σε αυτή την περίπτωση πάντα υπολογίζεται ως 11, επειδή μετρώντας τον ως 1 θα κάνει το ποσό 11 ή λιγότερο, στην οποία περίπτωση δεν υπάρχει καμία απόφαση να παρθεί, επειδή, προφανώς, ο παίκτης θα πρέπει πάντα να κάνει hit. Έτσι, ο παίκτης λαμβάνει αποφάσεις με βάση τρεις μεταβλητές: το τρέχον άθροισμα του (12 – 21), το πρώτο φύλο του ντίλερ (1 – 10), και το αν κατέχει ή όχι έναν usable άσο. Αυτό κάνει συνολικά 200 καταστάσεις.

Θέλουμε να βρούμε την βέλτιστη πολιτική με εφαρμογή της μεθόδου Monte Carlo Exploring Starts (ES).

- Αρχική πολιτική του παίκτη είναι αυτή που κάνει stick μόνο όταν το άθροισμα είναι 20 ή 21.