



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Επικοινωνιών, Ηλεκτρονικής & Συστημάτων Πληροφορικής

Εργαστήριο Διαχείρισης και Βέλτιστου Σχεδιασμού Δικτύων - NETMODE

Ηρώων Πολυτεχνείου 9, Ζωγράφου, 157 80 Αθήνα, Τηλ: 210-772.2503, Fax: 210-772.1452

e-mail: maglaris@netmode.ntua.gr, URL: <http://www.netmode.ntua.gr>

ΣΤΟΧΑΣΤΙΚΕΣ ΔΙΕΡΓΑΣΙΕΣ & ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ

(ΔΠΜΣ Επιστήμη Δεδομένων & Μηχανική Μάθηση)

ΕΝΔΕΙΚΤΙΚΕΣ ΑΣΚΗΣΕΙΣ, ΕΑΡΙΝΟ ΕΞΑΜΗΝΟ 2021

Value Iteration

Να εκτελέσετε τις πρώτες 3 επαναλήψεις του αλγορίθμου Value Iteration για το παρακάτω πρόβλημα βελτιστοποίησης: Ένα κινητό ξεκινάει από την κατάσταση S και κινείται στο πλέγμα που φαίνεται παρακάτω με σκοπό να μεγιστοποιήσει το κέρδος του. Η προσπάθειά του τερματίζεται είτε στην κατάσταση (1,3) είτε στην κατάσταση (2,3). Σε κάθε βήμα, παίρνει ή χάνει τους πόντους που αναγράφονται στο πλέγμα. Το κινητό μπορεί να κινηθεί πάνω, κάτω, αριστερά ή δεξιά, ενώ στην περίπτωση που η κίνησή του το οδηγεί έξω από το πλέγμα, το κινητό παραμένει στην ίδια κατάσταση. Οι κινήσεις του είναι στοχαστικές, δηλαδή με πιθανότητα 80% πραγματοποιεί την κίνηση που σκόπευε, ενώ με πιθανότητες 10% και 10% αντίστοιχα γλιστράει σε μία από τις δύο θέσεις που είναι κάθετες ως προς την κίνηση που σκόπευε να κάνει.

0, (2,1)	0, (2,2)	+5, (2,3)
S, 0, (1,1)	0, (1,2)	-5, (1,3)

Η βελτιστοποίηση γίνεται σύμφωνα με την εξίσωση Bellman:

$$V_{i+1}(s) = \max_a \left(\sum_{s'} P(s, a, s') (R(s, a, s') + \gamma V_i(s')) \right)$$

όπου V_i είναι η τιμή της Value Function στην επανάληψη i , P είναι ο πίνακας πιθανοτήτων μετάβασης της διαδικασίας, s είναι η τρέχουσα κατάσταση, s' η επόμενη κατάσταση, R το κέρδος του σε μία μετάβαση και με βάση μία ενέργεια a και γ το discount factor για το οποίο θα θεωρήσουμε ότι είναι ίσο με 0.9.

Ο πίνακας με τις επαναλήψεις του αλγορίθμου Value Iteration είναι ο παρακάτω:

s	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
$V_0(s)$	0	0	0	0	0	0
$V_1(s)$	0	0	0	0	4.0	0
$V_2(s)$	0	2.38	0	2.88	4.0	0
$V_3(s)$	0	2.38	0	3.14	4.57	0

Ενδεικτικοί Υπολογισμοί:

$$\begin{aligned}
 V_1((2,2)) &\equiv V_1(2,2) = \\
 &= \max_a \begin{cases} 0.8 \cdot (0 + 0.9 \cdot V_0(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_0(2,1)) + 0.1 \cdot (5 + 0.9 \cdot V_0(2,3)), a = \text{πάνω} \\ 0.8 \cdot (0 + 0.9 \cdot V_0(1,2)) + 0.1 \cdot (5 + 0.9 \cdot V_0(2,3)) + 0.1 \cdot (0 + 0.9 \cdot V_0(2,1)), a = \text{κάτω} \\ 0.8 \cdot (0 + 0.9 \cdot V_0(2,1)) + 0.1 \cdot (0 + 0.9 \cdot V_0(1,2)) + 0.1 \cdot (0 + 0.9 \cdot V_0(2,2)), a = \text{αριστερά} \\ 0.8 \cdot (5 + 0.9 \cdot V_0(2,3)) + 0.1 \cdot (0 + 0.9 \cdot V_0(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_0(1,2)), a = \text{δεξιά} \end{cases} \\
 &= \max_a \begin{cases} 0 + 0 + 0.5, a = \text{πάνω} \\ 0 + 0.5 + 0, a = \text{κάτω} \\ 0 + 0 + 0, a = \text{αριστερά} \\ 4 + 0 + 0, a = \text{δεξιά} \end{cases} = 4.0
 \end{aligned}$$

$$\begin{aligned}
 V_2(1,2) &= \\
 &= \max_a \begin{cases} 0.8 \cdot (0 + 0.9 \cdot V_1(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_1(1,1)) + 0.1 \cdot (-5 + 0.9 \cdot V_1(1,3)), a = \text{πάνω} \\ 0.8 \cdot (0 + 0.9 \cdot V_1(1,2)) + 0.1 \cdot (-5 + 0.9 \cdot V_1(1,3)) + 0.1 \cdot (0 + 0.9 \cdot V_1(1,1)), a = \text{κάτω} \\ 0.8 \cdot (0 + 0.9 \cdot V_1(1,1)) + 0.1 \cdot (0 + 0.9 \cdot V_1(1,2)) + 0.1 \cdot (0 + 0.9 \cdot V_1(2,2)), a = \text{αριστερά} \\ 0.8 \cdot (-5 + 0.9 \cdot V_1(1,3)) + 0.1 \cdot (0 + 0.9 \cdot V_1(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_1(1,2)), a = \text{δεξιά} \end{cases} \\
 &= \max_a \begin{cases} 2.88 + 0 - 0.5, a = \text{πάνω} \\ 0 - 0.5 + 0, a = \text{κάτω} \\ 0 + 0 + 0.36, a = \text{αριστερά} \\ -4 + 0.36 + 0, a = \text{δεξιά} \end{cases} = 2.38
 \end{aligned}$$

$$\begin{aligned}
 V_3(1,2) &= \\
 &= \max_a \begin{cases} 0.8 \cdot (0 + 0.9 \cdot V_2(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_2(1,1)) + 0.1 \cdot (-5 + 0.9 \cdot V_2(1,3)), a = \text{πάνω} \\ 0.8 \cdot (0 + 0.9 \cdot V_2(1,2)) + 0.1 \cdot (-5 + 0.9 \cdot V_2(1,3)) + 0.1 \cdot (0 + 0.9 \cdot V_2(1,1)), a = \text{κάτω} \\ 0.8 \cdot (0 + 0.9 \cdot V_2(1,1)) + 0.1 \cdot (0 + 0.9 \cdot V_2(1,2)) + 0.1 \cdot (0 + 0.9 \cdot V_2(2,2)), a = \text{αριστερά} \\ 0.8 \cdot (-5 + 0.9 \cdot V_2(1,3)) + 0.1 \cdot (0 + 0.9 \cdot V_2(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_2(1,2)), a = \text{δεξιά} \end{cases} \\
 &= 2.38
 \end{aligned}$$

$$\begin{aligned}
 V_3(2,2) &= \\
 &= \max_a \begin{cases} 0.8 \cdot (0 + 0.9 \cdot V_2(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_2(2,1)) + 0.1 \cdot (5 + 0.9 \cdot V_2(2,3)), a = \text{πάνω} \\ 0.8 \cdot (0 + 0.9 \cdot V_2(1,2)) + 0.1 \cdot (5 + 0.9 \cdot V_2(2,3)) + 0.1 \cdot (0 + 0.9 \cdot V_2(2,1)), a = \text{κάτω} \\ 0.8 \cdot (0 + 0.9 \cdot V_2(2,1)) + 0.1 \cdot (0 + 0.9 \cdot V_2(1,2)) + 0.1 \cdot (0 + 0.9 \cdot V_2(2,2)), a = \text{αριστερά} \\ 0.8 \cdot (5 + 0.9 \cdot V_2(2,3)) + 0.1 \cdot (0 + 0.9 \cdot V_2(2,2)) + 0.1 \cdot (0 + 0.9 \cdot V_2(1,2)), a = \text{δεξιά} \end{cases} \\
 &= 4.57
 \end{aligned}$$

Συντάχθηκε από τους υπεύθυνους εργαστηριακής υποστήριξης του μαθήματος
Νίκο Κωστόπουλο και Δημήτρη Πανταζάτο, Υποψήφιους Διδάκτορες Ε.Μ.Π.