

Πολυπρακτορικά Συστήματα

Αναφορά 2ης Εργασίας Προγραμματισμού

Ομάδα εργασίας:

Λυτρίδης Νικόλαος
Σούρσος Πέτρος

Υπό αβεβαιότητα δημιουργία συνασπισμών για την εκτέλεση εργασιών

Στο παιχνίδι συμμετέχουν συνολικά 12 πράκτορες (6 ηλεκτρολόγοι 6 υδραυλικοί) τριών δυνατών υπό-τύπων ο καθένας (καλός , κακός , μέτριος). Ο κάθε πράκτορες γνωρίζει τους κύριους τύπους αλλά δεν έχει γνώση των υπό -τύπων των άλλων πρακτόρων οι οποίοι καθορίζουν την αποδοτικότητα των συνασπισμών. Το παιχνίδι χωρίζεται σε γύρους , με πεπερασμένο πλήθος. Η υλοποίηση του πράκτορά μας έγινε με σκοπό να παίρνει τις καλύτερες δυνατές αποφάσεις , όσον αφορά τις προτάσεις που κάνει, και τις προτάσεις που του γίνονται. Για να παίρνει τις πιο συμφέρουσες προτάσεις ο πράκτορες αξιολογεί την αναμενόμενη γι' αυτόν αξία των συμβολαίων που προτείνει ή του προτείνονται, λαμβάνοντας υπ ' όψιν τις πιθανότητες επιτυχίας των εργασιών που αντιστοιχούν σε αυτές.

Για τη λύση προσπαθήσαμε εξ αρχής να μοντελοποιήσαμε το Πρόβλημα (της ακολουθιακής λήψης αποφάσεων υπό αβεβαιότητα) ως ένα Bayesian Colitional Formation Model με **N** πράκτορες **T** τύπους , **B** πεποίθηση για κάθε πράκτορα όσο αφορά τον τύπο των άλλων πρακτόρων , **Pr(o/t)** η πιθανότητα να πετύχει ο συνασπισμός δεδομένου των τύπων, **A** οι ενέργειες που μπορεί να κάνει ένας πράκτορας σε ένα συνασπισμό, **O** το αποτέλεσμα του συνασπισμού δεδομένου των ενεργειών και των τύπων, και **R(o)** το ποσό που θα αποκομίσει ο συνασπισμός σε επιτυχία.

Για το δικό μας Πρόβλημα γίνονται οι εξής παραδοχές, για να ενεργεί ο πράκτορας μας ανάλογα, βάσει και των δεδομένων της άσκησης:

- Οι πράκτορες είναι συνολικά 12.
- Οι τύποι είναι τρεις.
- Οι πεποιθήσεις βασίζονται στο κατά πόσο οι πράκτορες με τη συμμετοχή τους σε συνασπισμούς συμβάλουν στη περάτωση της εργασίας.
- Οι ενέργειες **A** ουσιαστικά σε εμάς δεν είναι τίποτα άλλο πέρα απ το ποια εργασία θα επιλέξει ένας πράκτορας(δηλαδή δεν πρόκειται για κάποια ενέργεια εντός του συνασπισμού).
- Το αποτέλεσμα **O** είναι ουσιαστικά η περάτωση ή όχι της εργασίας στον συνασπισμό που συμμετέχουμε.
- **R(o)** είναι το ποσοστό που θα πάρει ο δικός μας πράκτορας απ τον σύνολο της ανταμοιβής.

Βάσει των παραπάνω ορίζουμε αναμενόμενη αξία ενός συνασπισμού να είναι:

$$expVi = \max [\sum (B(tc) * (\sum Pr(o/tc) * R(o)))]$$

όπου :

tc : το σύνολο των υπό-τύπων του συνασπισμού

B(tc): Είναι η πεποίθηση για τους τύπους αυτούς

R(o): η απολαβή του πράκτορα

Pr(o/t) η πιθανότητα να πραγματοποιηθεί ο συνασπισμός βάση των υπό-τύπων (σύμφωνα με τις προδιαγραφές της άσκησης)

Για να μπορεί να υπολογίζει το πρόγραμμά μας την εξίσωση αυτή χρειάζεται να γίνεται κάποια ενημέρωση στις πεποιθήσεις. Αυτό το πετυχαίνουμε με ένα διάνυσμα **array b** μεγέθους όσο και το σύνολο των πρακτόρων .

Αρχικά είναι αρχικοποιημένο για όλους στο 0. Έπειτα σύμφωνα με τις εν δυνάμει συνασπισμούς κάθε στοιχείο του πίνακα σε κάθε γύρω αυξάνεται (μειώνεται) κατά μία μονάδα για κάθε συνασπισμό που απέδωσε (δεν απέδωσε). Έπειτα υπολογίζεται ο μέσος όρος όλων των πεποιθήσεων και για κάθε πράκτορα ανάλογα με την πεποίθησή μας γι αυτό σε σχέση με τον μέσο όρο του προσδίδουμε έναν υπό-τύπο. (Εδώ βάλαμε αυθαίρετα αν $B[i] > M.O+2$ τότε καλός, $B[i] > M.O-2$ μέτριος , αλλιώς κακός).

Μετά από έναν αριθμό επαναλήψεων (7,8) είδαμε πως με αυτόν τον τρόπο αξιολογούμε τους πράκτορες με επιτυχία κοντά 72% (8 στους 11 dummies).

Τα προβλήματα της παραπάνω μεθόδου έγκειται στο ότι για έναν συνασπισμό που περιέχει και καλό και κακό παίκτη τότε αν το αποτέλεσμα είναι κακό αυτό διαστρεβλώνει την εικόνα μας για τον καλό παίκτη προσδίδοντάς του αρνητική βαθμολογία.

Σε όλη τη παραπάνω διαδικασία την εξίσωσή αξίας την παίρνουμε υπ όψιν όσον αφορά την επιλογή συνασπισμών μετά τον εκατοστό γύρω ώστε να διαμορφωθούν καλές πεποιθήσεις για όλους τους πράκτορες.(παίζοντας όπως και ο dummy ως εκείνο το σημείο). Από εκεί και πέρα όταν είμαι προτείνων διαλέγω με τη λογική:

“ Απ' τους διαθέσιμους τους πιο ικανούς”

(δηλ. Αν χρειάζομαι 1 υδραυλικό και έχω 2 διαθέσιμους καλό και μέτριο θα πάρω τον καλό)

και όταν είμαι προτεινόμενος:

“Απληστος αλλά όχι τζογαδόρος”

(δηλ. Θα συνυπολογίσω την αξία πολλαπλασιάζοντας για κάθε δυνατή συμμετοχή το ποσοστό μου με την πιθανότητα επιτυχίας και διαλέγοντας το μέγιστο).

Για την εκπόνηση της εργασίας μελετήθηκε σχετική βιβλιογραφία:

- Chalkiadakis, G., and Boutilier, C. 2004. Bayesian reinforcement learning for coalition formation under uncertainty. In 3rd AAMAS
- Chalkiadakis, G., & Boutilier, C. (2008). Sequential decision making in repeated coalition formation under uncertainty. In Proceedings of the seventh international joint conference on autonomous agents and multiagent systems (AAMAS'08) (pp. 347–354)