

A. Επεξηγησιμότητα στη συνεργασία ανθρώπου-TN.

Οι μέθοδοι βαθιάς Ενισχυτικής Μάθησης μας δίνουν τη δυνατότητα να μελετήσουμε την από κοινού δράση ενός ανθρώπου και ενός πράκτορα TN σε πραγματικό χρόνο [1], και κατ' επέκταση να καταλάβουμε διάφορες πλευρές αυτής της συνεργασίας. Αυτό μας επιτρέπει τη δημιουργία μεθόδων TN που να προσφέρουν επεξηγησιμότητα και κατά συνέπεια βελτίωση της ποιότητας συνεργασίας.

Σκοπός της διπλωματικής είναι η μελέτη της συνεργασίας ανθρώπου-TN και ο σχεδιασμός κατάλληλων μεθόδων που μπορούν να υποστηρίξουν συνεργατική ευφυΐα μέσα από την επεξήγηση της συμπεριφοράς της ομάδας.

Στα πλαίσια της διπλωματικής:

- α) θα μελετηθούν μέθοδοι βαθιάς ενισχυτικής μάθησης,
- β) θα σχεδιαστεί κατάλληλη μέθοδοι για επεξηγησιμότητα,
- γ) θα γίνει συλλογή δεδομένων από πραγματικούς χρήστες για τον σχεδιασμό και την αξιολόγηση των μεθόδων.

Αναφορές

[1] Koutrintzes, D., Spatharis, .C, and Dagiloglou, M. 2024. Human-Aware Design for Transferring Knowledge During Human-AI Co-Learning. In HAI5.0: Embracing Human-Aware AI in Industry 5.0, Workshop at ECAI

B. Συνεργατικοί πράκτορες ενισχυτικής μάθησης

Η πολυπρακτορική ενισχυτική μάθηση (Multi-Agent Reinforcement Learning - MARL) μελετά πως πολλαπλοί πράκτορες μπορούν να μάθουν και να αλληλεπιδρούν σε κοινά περιβάλλοντα, είτε συνεργατικά είτε ανταγωνιστικά. Διάφορες έρευνες σε συνεργατικούς πράκτορες έχουν εστιάσει στην ανάπτυξη αλγορίθμων που μπορούν να σταθεροποιήσουν την μάθηση και να ενισχύσουν τον συντονισμό μεταξύ των πρακτόρων [1,2].

Σύμφωνα με πρόσφατες έρευνες [3], δύο ανεξάρτητοι πράκτορες αδυνατούν να μάθουν επιτυχημένες πολιτικές σε κάποια (συνεργατικά) περιβάλλοντα. Η παρούσα διπλωματική εργασία εστιάζει στη μελέτη πολυπρακτορικών σεναρίων συνεργασίας, με στόχο την εφαρμογή τεχνικών MARL για την εκπαίδευση έμπειρων πρακτόρων που θα μπορούν στη συνέχεια να μεταδίδουν τη γνώση τους σε νέους χρήστες. Συγκεκριμένα, θα εξεταστεί η δυνατότητα αξιοποίησης προ-εκπαιδευμένων πρακτόρων για την καθοδήγηση ανθρώπινων χρηστών σε κοινές εργασίες, διευκολύνοντας έτσι την διαδικασία μάθησης.

Συγκεκριμένα, στα πλαίσια της διπλωματικής:

- α) θα μελετηθούν μέθοδοι βαθιάς συνεργατικής πολυπρακτορικής ενισχυτικής μάθησης,
- β) θα τροποποιηθεί/σχεδιαστεί κατάλληλο περιβάλλον για πειραματισμό σε ψηφιακό περιβάλλον,

γ) θα γίνει συλλογή δεδομένων από πραγματικούς χρήστες για την αξιολόγηση των μεθόδων.

Αναφορές

[1] Rashid, T., Samvelyan, M., De Witt, C., Farquhar, G., Foerster, J., and Whiteson, J., 2020. Monotonic Value Function Factorisation for Deep Multi-agent Reinforcement Learning. In The Journal of Machine Learning Research (pp. 7234-7284)

[2] Yu, C., Velu, A., Vinitzky, E., Gao, J., Wang, Y., Bayen, A., and Wu, Y. 2022. The Surprising Effectiveness of PPO in Cooperative Multi-agent Games. In Proceedings of the 36th International Conference on Neural Information Processing Systems (NIPS '22)

[3] Koutrintzes, D., Spatharis, .C, and Dagiloglou, M. 2024. Human-Aware Design for Transferring Knowledge During Human-AI Co-Learning. In HAIL5.0: Embracing Human-Aware AI in Industry 5.0, Workshop at ECAI

Γ. Εκμάθηση από Προτιμήσεις σε Ρομποτικές Τροχιές

Συνοπτική περιγραφή:

Η εκμάθηση ανταμοιβής από ανθρώπινη εποπτεία είναι βασικό στοιχείο στην αλληλεπίδραση ανθρώπου-ρομπότ. Παραδοσιακές προσεγγίσεις, όπως το Bayesian Inverse Reinforcement Learning (BIRL)[1], βασίζονται σε κριτικές τύπου "καλό/κακό" για τροχιές. Ωστόσο, η εκμάθηση από προτιμήσεις (π.χ. "Η τροχιά Α είναι προτιμότερη από τη Β") προσφέρει έναν πιο φυσικό και διαισθητικό τρόπο έκφρασης για τους χρήστες.

Σκοπός αυτής της διπλωματικής εργασίας είναι η επέκταση της μεθόδου BLIND [1] ώστε να ενσωματώνει ανθρώπινες προτιμήσεις αντί για κριτικές για την εκμάθηση της συνάρτησης ανταμοιβής [2]. Η τροποποιημένη μέθοδος θα συγκριθεί με την αρχική εκδοχή του BLIND όσον αφορά την ποιότητα των παραγόμενων τροχιών και τη δυνατότητα γενίκευσης.

Στα πλαίσια της διπλωματικής:

α) θα μελετηθούν αλγόριθμοι εκμάθησης προτιμήσεων υπό το πρίσμα της ενισχυτικής μάθησης,

β) θα τροποποιηθεί ο υπολογισμός της συνάρτησης ανταμοιβής ώστε να υποστηρίζει προτιμήσεις τροχιών,

γ) θα πραγματοποιηθεί συγκριτική αξιολόγηση με βάση τις αποδόσεις των δύο μεθόδων σε κοινά σενάρια.

Αναφορές

[1] Quintero-Peña, C., Chamzas, C., Sun, Z., Unhelkar, V. and Kavraki, L. E., 2022. Human-Guided Motion Planning in Partially Observable Environments. In 2022 International Conference on Robotics and Automation (ICRA), (pp. 7226–7232)

[2]Hejna, J., and Sadigh, D. 2023. Inverse Preference Learning: Preference-based RL Without a Reward Function. In Proceedings of the International Conference on Neural Information Processing Systems (NIPS '23)

Δ. Συντονισμός Ανθρώπου-TN σε μέσω Ιεραρχικής Ενισχυτικής Μάθησης

Η παρούσα διπλωματική εργασία στοχεύει στην επέκταση σύγχρονων προσεγγίσεων επιρροής ανθρώπινης συμπεριφοράς μέσω offline ενισχυτικής μάθησης[1], εισάγοντας ένα ιεραρχικό πλαίσιο λήψης αποφάσεων. Σε αντίθεση με τις ήδη υπάρχουσες μεθόδους που επικεντρώνονται στην άμεση επίδραση πράξεων ή βραχυπρόθεσμων στρατηγικών, η διπλωματική επιδιώκει να επιτρέψει σε ευφυείς πράκτορες να επιδρούν σε πιο αφηρημένα επίπεδα, στοχεύοντας τους ανθρώπινους στόχους και προθέσεις κατά την εκτέλεση συνεργατικών εργασιών.

Η μεθοδολογία βασίζεται σε μια πολυεπίπεδη αρχιτεκτονική ενισχυτικής μάθησης, στην οποία μια ανώτερη πολιτική (high-level policy) αναλαμβάνει την πρόβλεψη και επιλογή των επιθυμητών ανθρώπινων στόχων, ενώ μια κατώτερη πολιτική (low-level policy) σχεδιάζει και εκτελεί τις αντίστοιχες ενέργειες ώστε να κατευθύνει τον άνθρωπο προς την επίτευξή τους. Στόχος είναι η ενσωμάτωση μηχανισμών πρόβλεψης ανθρώπινων στόχων[2] με ιεραρχική offline ενισχυτική μάθηση, αξιοποιώντας αποθηκευμένα δεδομένα από ανθρώπινες αλληλεπιδράσεις σε κατάλληλο ψηφιακό περιβάλλον.

Στα πλαίσια της διπλωματικής:

- α) θα μελετηθούν αλγόριθμοι offline και ιεραρχικής ενισχυτικής μάθησης,
- β) θα διεξαχθούν πειράματα σε ψηφιακό περιβάλλον,
- γ) θα γίνει αξιολόγηση της μεθοδολογίας μέσα από πειράματα με πραγματικούς χρήστες.

Αναφορές

[1] Hong, J., Levine, S., & Dragan, A. 2024. Learning to Influence Human Behavior with Offline Reinforcement Learning. In Proceedings of the International Conference on Neural Information Processing Systems (NIPS '24)

[2] Hoffman, G., Bhattacharjee, T., and Nikolaidis, S. 2024. Inferring Human Intent and Predicting Human Action in Human–Robot Collaboration. In Annual Review of Control, Robotics and Autonomous Systems