

№14 Дәріс. Көп агентті нықталандырып оқыту

Дәрістің мақсаты:

Көп агентті нықталандырумен оқыту жүйелерінің (Multi-Agent Reinforcement Learning, MARL) негізгі ұғымдарын, олардың ерекшеліктерін, өзара әрекеттесу механизмдерін және қолдану салаларын қарастыру. Агенттердің бір-бірімен ынтымақтаса немесе бәсекелесе отырып үйрену процесін түсіндіру.

Кілт сөздер:

Көп агентті жүйе, нықталандырумен оқыту, орта, агент, саясат (policy), ынтымақтастық, бәсекелестік, координация, MARL, ортаның динамикасы, командалық оқыту.

Дәріс жоспары:

1. Көп агентті нықталандырумен оқытуға кіріспе
2. Бір агентті және көп агентті оқытудың айырмашылығы
3. MARL жүйесінің негізгі компоненттері
4. Агенттердің өзара әрекеттесу түрлері
5. Ынтымақтастық және бәсекелестік сценарийлер
6. Көп агентті оқыту әдістері
7. MARL қолдану салалары
8. Қиындықтары мен болашағы
9. Қорытынды

Кіріспе

Көп агентті нықталандырумен оқыту (MARL) — бірнеше интеллектуалды агенттердің ортада өзара әрекеттесу арқылы тәжірибе жинап, мінез-құлық стратегияларын үйрену процесі. Әрбір агент ортадан ақпарат алып, шешім қабылдайды және соның негізінде марапат немесе жаза алады.

Мұндай жүйелер бір агентті нықталандырумен оқытудан әлдеқайда күрделі, себебі әр агенттің әрекеті орта динамикасына ғана емес, басқа агенттердің шешімдеріне де тәуелді. Бұл жағдай «орта тұрақсыздығы» деп аталатын құбылысты тудырады, яғни әр агенттің оқу процесі басқа агенттердің мінез-құлқына байланысты өзгеріп отырады.

1. Көп агентті нықталандырумен оқытуға кіріспе

MARL жүйелерінде агенттер бір уақытта ортада әрекет етеді. Әрқайсысының өз мақсаты мен стратегиясы болады. Агенттердің бірлескен әрекеті кейде ортақ мақсатқа жетуді жеңілдетсе, кейде керісінше — бәсекелестік тудырады.

Көп агентті жүйелер табиғи ортадағы көптеген процестерге ұқсайды: адамдар арасындағы әлеуметтік қатынастар, жануарлар тобының мінез-құлқы, нарықтық экономикалық агенттердің шешімдері және автономды

көліктердің қозғалысын үйлестіру сияқты күрделі сценарийлерді модельдеуге мүмкіндік береді.

2. Бір агентті және көп агентті оқытудың айырмашылығы

Бір агентті жүйеде агент тек ортамен әрекеттеседі және орта мінез-құлқы тұрақты болып саналады. Ал MARL жүйесінде орта динамикасы басқа агенттердің әрекеттеріне байланысты өзгеріп отырады. Бұл оқытуды әлдеқайда күрделі етеді, себебі агентке тек орта емес, өзге агенттердің мінез-құлқын да болжауға тура келеді.

3. MARL жүйесінің негізгі компоненттері

Көп агентті нықталандырумен оқыту жүйесі келесі компоненттерден тұрады:

- Агенттер (Agents): өз бетімен шешім қабылдайтын бірліктер.
- Орта (Environment): агенттердің әрекет ететін кеңістігі.
- Күйлер (States): орта мен агенттердің ағымдағы жағдайы.
- Әрекеттер (Actions): агенттердің қабылдай алатын шешімдері.
- Марапат (Reward): әр агенттің әрекетінің тиімділігін бағалайтын сигнал.

4. Агенттердің өзара әрекеттесу түрлері

Агенттер арасындағы өзара әрекеттесу үш негізгі түрге бөлінеді:

1. Бәсекелестік (Competitive): әр агент өз пайдасын барынша арттыруға тырысады.
2. Ынтымақтастық (Cooperative): барлық агенттер ортақ мақсатқа жұмыс істейді.
3. Аралас (Mixed): кейбір агенттер бірігіп әрекет етеді, ал басқалары өз мүддесін көздейді.

Мысалы, шахмат немесе футбол ойынында агенттер бәсекелеседі, ал роботтар тобы бір жүктемені көтеру үшін ынтымақтаса әрекет етеді.

5. Ынтымақтастық және бәсекелестік сценарийлер

Ынтымақтастық жүйелерде агенттер ортақ марапат алады, яғни барлығының табысы жалпы нәтижеге байланысты. Мұнда үйлестіру (координация) маңызды рөл атқарады.

Бәсекелестік жүйелерде агенттердің мақсаты қарама-қайшы, яғни бір агенттің ұтысы екіншісінің ұтылуы арқылы анықталады. Мұндай жағдайларда агент қарсыласының стратегиясын болжауға үйренуі тиіс.

6. Көп агентті оқыту әдістері

MARL әдістері негізінен екі бағытқа бөлінеді:

- Орталықтандырылған оқыту (Centralized Training): барлық агенттердің ақпараттары біріктіріліп, жалпы саясат үйретіледі.
- Децентрализованный оқыту (Decentralized Learning): әр агент өз тәжірибесі негізінде үйренеді.

Кейбір заманауи әдістер бұл тәсілдерді біріктіреді, яғни оқыту орталықтандырылған түрде жүргізіліп, ал қолдану (policy execution) децентрализованный түрде іске асады.

7. MARL қолдану салалары

Көп агентті нықталандырумен оқыту көптеген нақты салаларда кеңінен қолданылады:

- Робототехника: бірнеше роботтың үйлестірілген қозғалысын басқару.
- Энергетика: электр желісін оңтайландыру және жүктемені бөлу.
- Транспорт: автономды көліктердің қозғалысын үйлестіру.
- Экономика және нарықтық модельдеу: агенттердің мінез-құлқын болжау.
- Компьютерлік ойындар: көп ойыншы қатысатын стратегияларды дамыту.

8. Қиындықтары мен болашағы

Қиындықтары:

- Орта тұрақсыздығы (әр агенттің үйренуі басқа агентке әсер етеді).
- Есептеу қиындығы (күй мен әрекет кеңістігі өте үлкен).
- Ынтымақтастықтағы стратегияларды үйлестірудің қиындығы.

Болашағы:

- Нақты өмірлік жүйелерді басқаруда MARL маңызды құралға айналу.
- Болашақта көп агентті жүйелер интеллектуалды қалаларда, әскери стратегияларда, логистикада және автономды көліктер желісінде кеңінен қолданылмақ.

Қорытынды

Көп агентті нықталандырумен оқыту — жасанды интеллекттің ең күрделі әрі болашағы зор бағыттарының бірі. Бұл әдіс бір-бірімен өзара әрекеттесетін бірнеше агенттің ортада үйрену процесін модельдейді. Ынтымақтастық пен бәсекелестік сияқты күрделі мінез-құлық үлгілерін үйрету арқылы MARL жүйелері шынайы өмірдегі көпкомпонентті процестерді тиімді басқаруға мүмкіндік береді.

Бақылау сұрақтары:

1. Көп агентті нықталандырумен оқыту бір агентті RL-ден несімен ерекшеленеді?
2. Агенттердің өзара әрекеттесу түрлерін атаңыз.
3. Ынтымақтастық және бәсекелестік сценарийлердің айырмашылығы неде?
4. MARL жүйелерінің негізгі қиындықтары қандай?
5. MARL қай салаларда тиімді қолданылады?

Пайдаланылған әдебиеттер:

1. Sutton, R. S., & Barto, A. G. *Reinforcement Learning: An Introduction*. MIT Press, 2018.
2. Busoniu, L., Babuska, R., De Schutter, B., & Ernst, D. *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC Press, 2010.
3. Zhang, K., Yang, Z., & Başar, T. *Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms*. arXiv:1911.10635, 2019.

4. Silver, D. et al. *Deterministic Policy Gradient Algorithms*. ICML, 2014.
5. Shoham, Y., Powers, R., & Grenager, T. *Multi-Agent Reinforcement Learning: A Critical Survey*. Technical Report, Stanford University, 2003.