

ՊԱՇՏՈՆԱԿԱՆ ԸՆԴԴԻՄԱԽՈՍԻ

ԿԱՐԾԻՔ

Ե.13.04 «Հաշվողական մեթոդների, համալիրների, համակարգերի և ցանցերի մաթեմատիկական և ծրագրային ապահովում» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի հայցման համար ներկայացված Միքայել Էդուարդի Սամվելյանի «խորքային բազմագործակալ ամրապնդմամբ ուսուցման արդյունավետ մեթոդների մշակում և գնահատում» թեմայով ատենախոսության վերաբերյալ:

Վերջին տարիներին կտրուկ աճել է հետաքրքրությունը ամրապնդմամբ ուսուցման մեթոդների նկատմամբ: Մեկ գործակալով ամրապնդմամբ ուսուցման ալգորիթմների միջոցով հաջողվել է ստանալ համակարգչային ծրագրեր, որոնք սովորում են խաղալ բազմաթիվ սեղանի խաղեր և համակարգչային խաղեր: Մասնավորապես հիշարժան է գո խաղի աշխարհի չեմպիոնին հաղթած AlphaGo ծրագրի մշակումը Google DeepMind ընկերության հետազոտողների կողմից՝ խորը ուսուցման մեթոդների կիրառմամբ: Համեմատաբար քիչ են ուսումնասիրված բազմագործակալ միջավայրերում ամրապնդմամբ ուսուցման ալգորիթմները, ինչին էլ նվիրված է աշխատանքը: Թեման արդիական է, ունի կարևոր տեսական և կիրառական նշանակություն: Բազմագործակալ միջավայրերում համագործակցային ալգորիթմները, որոնք հանդիսանում են այս աշխատանքի ուսումնասիրության հիմնական առարկաները, կարող են կիրառվել մեծ թվով ռոբոտների, անօդաչու թռչող սարքերի համատեղ աշխատանքը արդյունավետ կազմակերպելու նպատակով:

Միքայել Սամվելյանի ատենախոսական աշխատանքը բաղկացած է ներածությունից, չորս գլուխներից, եզրակացությունից և գրականության ցանկից:

Ներածությունում համառոտակի ներկայացված է ատենախոսության թեման, նրա արդիականությունը, գիտական նորույթը և ստացված արդյունքների գործնական նշանակությունը:

Առաջին գլխում նկարագրված է բազմագործակալ ամրապնդմամբ ուսուցման ինդիքը, դրա տարատեսակները, տրված են անհրաժեշտ սահմանումներն ու նշանակումները, համառոտ նկարագրված են ամրապնդմամբ ուսուցման դասական ալգորիթմները, խորը ուսուցման մեթոդները և դրանց կիրառությունը ամրապնդմամբ ուսուցման ալգորիթմներում: Առաջին գլխի վերջին ենթագլխում տրված է համագործակցային ամրապնդմամբ ուսուցման ինդիքի ձևակերպումը որպես ապակենտրոնացված մասնակի դիտարկելի Մարկովյան որոշումների գործընթաց և նկարագրված են այդ ինդիքի համար մշակված հայտնի ալգորիթմները՝ IQL, COMA, VDN, նշված են դրանց որոշ թերություններ:

Գլուխ 2-ում մշակված է QMIX ալգորիթմը, որը հանդիսանում է VDN ալգորիթմի ընդլայնում և ուղղում է նրա որոշ թերություններ: Ամենաեական տարբերությունը, թերևս, QMIX մեթոդում Q_{tot} կենտրոնացված արժեքային ֆունկցիայի կառուցման ճկունությունն է, որն ապահովվում է խորը նեյրոնային ցանցերի կիրառման շնորհիվ: Այդպիսի ցանցերի օգնությամբ առանձին գործակալների Q_a արժեքային ֆունկցիաների միավորումը ընդհանուր արժեքային ֆունկցիայի մեջ կարող է թուլացնել համակարգի կայունությունը, ինչից խուսափելու համար նեյրոնային ցանցի պարամետրերը որոշվում են այսպես կոչված հիպերցանցերի միջոցով՝ բավարարելով որոշ սահմանափակումներ: 2.6 ենթագլխում ցույց է տրված QMIX ալգորիթմի առավելությունը VDN և IQL ալգորիթմների նկատմամբ՝ StarCraft II համակարգչային խաղի մի շարք բազմագործակալ սցենարներում:

Երրորդ գլխում մշակված է բազմագործակալ ամրապնդմամբ ուսուցման առաջին բազմաֆունկցիոնալ հենանիշ միջավայրը՝ SMAC-ը, որը նախատեսված է գործակալների համագործակցության մի շարք մեխանիզմների առկայությունը ստուգելու համար: Այն հիմնված է StarCraft II խաղի վրա և թույլ է տալիս տարբեր գիտական խմբերին աշխատել միևնույն չափանիշներով, կատարել տարբեր ալգորիթմների արդարացի համեմատություններ բազմազան խաղային սցենարներում:

Չորրորդ գլխում դիտարկված է գործակալների կողմից միջավայրի համաձայնեցված ուսումնասիրության խնդիրը որոշակի բազմագործակալ սցենարներում: Մշակված է ամրապնդմամբ ուսուցման ալգորիթմ՝ MAVEN, որտեղ բոլոր գործակալների արժեքային ֆունկցիաները ունեն ընդհանուր հատված, որը կախված է միջավայրի սկզբնական վիճակից: Ցույց է տրված, որ MAVEN ճարտարապետության օգնությամբ QMIX ալգորիթմը կարողանում է հայտնաբերել օպտիմալ ռազմավարություն SMAC հենանիշի այնպիսի սցենարներում, որոնցում QMIX ալգորիթմը իր նախնական տեսքով չի աշխատում:

Ատենախոսությունում նկատվել են հետևյալ թերությունները.

1. QMIX ալգորիթմում հիպերցանցերի կիրառությունը ունի երկակի ազդեցություն. մի կողմից այն ապահովում է խառնիչ ցանցի կշիռների ոչ բացասական լինելը, մյուս կողմից՝ առաջացնում է ցանցի կախվածություն միջավայրի վիճակից: 2.3 ենթագլխում նկարագրված QMIX-NS մեթոդի հետ համեմատությունը ցույց է տալիս հիպերցանցերի կարևորությունը միջավայրի վիճակից կախվածության մասով, սակայն բաց է մնում մոնոտոնության երաշխավորման անհրաժեշտության հարցը: Կարծում եմ, որ լրացուցիչ փորձը առանց հիպերցանցի խառնիչ ցանցի օգնությամբ, որը որպես հավելյալ մուտք կստանար վիճակը նկարագրող վեկտորը, լրացուցիչ կպարզաբաներ այդ բաղադրիչի կարևորությունը:

2. Աշխատանքի շարադրանքի հերթականությունը ժամանակագրական է, ըստ արդյունքների ստացման հերթականության, ինչի հետևանքով, օրինակ, QMIX ալգորիթմի վերաբերյալ փորձերի արդյունքներ նշված են ինչպես 2.6, այնպես էլ 3.6 ենթագլուխներում: Կարծում եմ ատենախոսության շարադրանքը կլիներ էապես ավելի սահուն, եթե նախ ներկայացվեր SMAC հենանիշը, ապա հերթականորեն նոր մշակված ալգորիթմները, իսկ բոլոր փորձերի արդյունքները շարադրվեին բացառապես SMAC հենանիշի շրջանակներում:

3. Չորրորդ գլխի փորձարկումներում ուշագրավ է այն փաստը, որ IQL ալգորիթմի արդյունքները բավականաչափ երկար ուսուցման արդյունքում մոտենում են առավել բարդ ալգորիթմների լավագույն արդյունքներին: Դա կարող է վկայել այն մասին, որ IQL-ի արդյունավետությունը հնարավոր է բարձրացնել առավել հզոր օպտիմիզացիայի ալգորիթմների կիրառմամբ՝ առանց կառուցվածքային բարդություններ ներմուծելու: Հասկանալի է, որ նման հետազոտությունները կպահանջեն հսկայական հաշվողական ռեսուրսներ, սակայն կարծում եմ, որ այս երևույթի ուսումնասիրությունը կարող է հետագա աշխատանքների համար հետաքրքիր ուղղություն լինել:

Կարծում եմ, որ նշված թերությունները չեն արժեզրկում ատենախոսությունում ստացված արդյունքները: Աշխատանքի երկրորդ և չորրորդ գլուխների բովանդակությունը ներկայացվել է համապատասխանաբար ICML և NeurIPS գիտաժողովներում, որոնք հանդիսանում են ոլորտի ամենահեղինակավոր երկու հարթակները: Ատենախոսությունում մշակված QMIX ալգորիթմի հաջողության մասին է վկայում նաև այն փաստը, որ այլ հեղինակների կողմից արդեն մշակվել են այդ ալգորիթմի մի քանի ընդլայնված տարբերակներ: SMAC հենանիշը նկարագրող հոդվածը ներկայացվել է ամրապնդմամբ ուսուցման խնդիրներով զբաղվող մասնագիտական համայնքի համար կենտրոնական համարվող DeepRL աշխատաժողովում: Չնայած որ SMAC-ի վերաբերյալ հոդվածի առաջին տարբերակը հանրությանը հասանելի էր դարձել ընդամենը մեկ տարի առաջ, այն արդեն գրավել է

գիտական համայնքի ուշադրությունը, և այսօր հասանելի են մի շարք հետազոտական խմբերի կողմից արված աշխատանքներ, որոնց փորձերը իրականացվել են SMAC-ի շրջանակներում:

Սեղմագիրը համապատասխանում է ատենախոսության բովանդակությանը: Ատենախոսությունն իրենից ներկայացնում է ամբողջական ավարտուն գիտական աշխատանք, որը բավարարում է ՀՀ Բարձրագույն Որակավորման Հանձնաժողովի կողմից թեկնածուական ատենախոսությունների նկատմամբ ներկայացվող բոլոր պահանջներին, իսկ նրա հեղինակը՝ Միքայել Էդուարդի Սամվելյանը, արժանի է տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի շնորհմանը:

Պաշտոնական ընդդիմախոս,
ԵՊՀ ԻԿՄ դիսկրետ մաթեմատիկայի և
տեսական ինֆորմատիկայի ամբիոնի դասախոս, Ֆ.մ.գ.թ.



Յ. Յ. Խաչատրյան

30 մարտի 2020թ.