

VILNIAUS UNIVERSITETAS
INFORMATIKOS INSTITUTAS
PROGRAMŲ SISTEMOS

Sustiprinto mokymosi taikymas žaidimo agento valdymo programos kūrimui

Application of reinforcement learning to the software development for game agent management

Bakalauro baigiamojo darbo planas

Atliko: Jokūbas Rusakevičius (parašas)

Darbo vadovas: vyresn. m.d. Virginijus Marcinkevičius (parašas)

Darbo recenzentas: (parašas)

Vilnius – 2020

Tyrimo objektas ir aktualumas

Kompiuterių pajėgumui ir atliekamų operacijų per sekundę skaičiui nuolatos didėjant – didėja ir lūkesčiai bei sprendžiamų uždavinių sudėtingumas. Dar reletyviai neseniai sudėtingiausios programos ir kompiuterių sprendžiami uždaviniai susidėjo iš skaičiuotuvo operacijų ar žinučių perdavimo. Tačiau technologijoms tobulėjant, kiekvienam žmogui kišenėje besinešiojant pirmųjų kompiuterių kaip „ENIAC“ [oCom] didį pajuokenčius kompiuterinius įrenginius, natūraliai didėja ir jiems keliama iššūkiai.

Šiais laikais kompiuteriai gali simuliuoti atominius sprogius, nuspėti orus ir atlikti kitas didžiulių skaičiavimo išteklių reikalaujančias užduotis [Ker]. Tačiau užduoties sudėtingumą gali lemti ne tik milžiniškų išteklių skaičiaus reikalavimas. 2016 metais matėme, kaip „Google’s AlphaGo“ nugalėjo pasaulio aukščiausio lygio „Go“ žaidėją ir čempioną Ke Jie [Moz17]. Autonomiškai gatvėmis važinėjantys automobiliai neišvengiamai artėja, o „Boston Dynamics“ robotai stebina savo galimybėmis [Dyn20].

Šie uždaviniai nėra trivialiai aprašomi ar išsprendžiami, jiems gali net neegzistuoti sprendimas. Tokiems uždaviniams spręsti yra naudojami mašininio mokymosi ar neuroninių tinklų metodai. Viena šių metodų atmaina yra „sustiprintas mokymas“ – agento atliekami veiksmai yra reguliariai vertinami ir atitinkamai agentas yra apdovanojamas arba baudžiamas.

Sustiprintas mokymas ir jam kuriamas agentas bus šio bakalauro darbo tiriamasis objektas.

Darbo tikslas, keliama uždavyniai ir laukiami rezultatai

Šio darbo **tikslas** – išanalizavus populiariausius sustiprinto mokymosi algoritmus, pritaikyti labiausiai tinkamą algoritmą parinktai eksperimentiniai aplinkai bei pagerinti gaunamus rezultatus pritaikius agento mokymosi gerinimo principus.

Darbui iškelti **uždaviniai**:

1. Paruošti eksperimentinę aplinką ir agentą.
2. Apmokyti agentą.
3. Palyginti gaunamą agento efektyvumą atliekant aplinkoje realizuotą užduotį, keičiant aplinkos bei agento mokymosi kriterijus.
4. Pateikti rekomendacijas agento apmokymui parinktoje aplinkoje.

Darbo metu laukiami **rezultatai**:

1. Paruošta eksperimentinė aplinka ir agentas.
2. Agentas yra apmokytas.
3. Pakeisti aplinkos ir agento mokymosi kriterijai, palyginti gaunami rezultatai ir taip pasiektas geriausias užduoties atlikimas.
4. Pateiktos rekomendacijos agento apmokymui parinktoje aplinkoje.

Tyrimo metodas

Darbui atlikti bus naudojami šie tyrimo metodai:

1. **Mokslinės literatūros analizė.**

2. Eksperimentas.

Kiekybinis metodas – apmokyti pirminius agentus, stebėti jų elgseną. Atliekamas, su keliais skirtingais kriterijais ir ieškoma tinkamiausio bei daugiausiai žadančio pradinio rezultato eksperimentui.

Kokybinis metodas – su gautais pradiniais agentais bei jų mokymosi kriterijais atliekamos gilesnės kriterijų bei agento mokymosi parametrų manipuliacijos bei stebimi gaunamų rezultatų pakitimai.

3. Gautų duomenų ir rezultatų analizė.

Numatomas darbo atlikimo procesas

Darbo procesas susidės iš kelių skirtingų dalių:

1. Visų pirma, eksperimentui atlikti bus paruošta eksperimentinė aplinka.
2. Agento paruošimas bus atliekamas pasirenkant pradinis, neutralius mokymosi kriterijus.
3. Atliekamas pirminis agento apmokymas, skirtas eksperimento bazei gauti bei įsitikinto aplinkos ir agento tarpusavio veikla.
4. Pradėjus pagrindinę eksperimento dalį bus ieškoma ir eksperimentuojama su metodais, kuriuos pritaikius būtų gaunami geresni nei aukščiau gaunami rezultatai.
5. Galutinė visų atliktų tyrimų ir gautų rezultatų analizė.
6. Pateikiamos rekomendacijos, kaip pagerinti agento apmokymą parinktai aplinkai ir jos keliamoms problemoms bei uždaviniams.

Darbui aktualūs šaltiniai

1. [Sch19] – viešai prieinama „OpenAI“ [Ope16] aplinka realizuojanti tradicinę Japonišką sandėlių prižiūrėjimo video žaidimą „Sokoban“. Tai yra parinkta aplinka bakalauro darbo eksperimento daliai.
2. [Gér17] – knyga – mašininio ir neuroninio mokymosi algoritmų rinkinys. Rašoma apie daugelį populiariausių ir plačiausia žinomų skirtingų mokymosi algoritmų bei principų. Pateikiami pritaikymo bei panaudojimo pavyzdžiai.
3. [Sze10] – klasikinė knyga apie sustiprintą mokymą. Šiame šaltinyje rašoma apie tai kas yra sustiprintas mokymas, kokie algoritmai egzistavo šaltinio leidimo metu bei kaip ir kada juos naudoti.
4. [Incb] – „Udemy“ [Ude] kursas apie dirbtinį intelektą bei pasiruošimą darbui su giliuoju sustiprintu mokymu. Kursas apžvelgia sustiprinto mokymosi algoritmų pritaikymo, programavimo ir dizaino principus.
5. [Inca] – „Udemy“ [Ude] kursas apie dirbtinio intelekto pritaikymo naudojantis giliuoju mokymusi ir neuroniniais tinklais įvaldymą.

Literatūra

- [Dyn20] Boston Dynamics. Boston dynamics. 2020. URL: <https://www.bostondynamics.com/about> (tikrinta 2020-03-19).
- [Gér17] A. Géron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2017. ISBN: 9781491962244. URL: <https://books.google.lt/books?id=bRpYDgAAQBAJ>.
- [Inca] Lazy Programmer Inc. Advanced ai: deep reinforcement learning in python. URL: <https://www.udemy.com/course/deep-reinforcement-learning-in-python/learn/lecture/6740866?start=0#overview> (tikrinta 2020-03-19).
- [Incb] Lazy Programmer Inc. Artificial intelligence: reinforcement learning in python. URL: <https://www.udemy.com/course/artificial-intelligence-reinforcement-learning-in-python/learn/lecture/6386364#overview> (tikrinta 2020-03-19).
- [Ker] Kate Kershner. What are supercomputers currently used for? URL: <https://computer.howstuffworks.com/supercomputers-used-for1.htm> (tikrinta 2020-03-19).
- [Moz17] Paul Mozur. Google's alphago defeats chinese go master in win for a.i. 2017. URL: <https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html> (tikrinta 2020-03-19).
- [oCom] History of Computers. History of computers. URL: <https://homepage.cs.uri.edu/faculty/wolfe/book/Readings/Reading03.htm> (tikrinta 2020-03-19).
- [Ope16] OpenAI. Getting started with gym. 2016. URL: <http://gym.openai.com/docs/> (tikrinta 2020-03-19).
- [Sch19] Max Schrader. Sokoban environment for openai gym. 2019-07-19. URL: <https://github.com/mpSchrader/gym-sokoban> (tikrinta 2020-03-19).
- [Sze10] Csaba Szepesvári. *Algorithms for Reinforcement Learning*, tom. 4. 2010-01. DOI: 10.2200/S00268ED1V01Y201005AIM009.
- [Ude] Udemy. Udemy. URL: <https://www.udemy.com/> (tikrinta 2020-03-19).