**Vilniaus Užupio gimnazija**

**Oskaras Klimašauskas, Paulius Kurlavičius.**

Objektų skatinamasis mokymas naudojant neuroninio tinklo ir genetinio algoritmo sąsąją.

Kūrybinis/Tiriamasis darbas

Darbo vadovas

Vaidas Raižys

ĮVADAS

Pirmieji neuroniniai tinkai jau buvo sukurti ir sėkmingai apmokyti XX a. viduryje, tačiau jų pritaikymo galimybės buvo labai robota dėl mokymo duomenų stokos ir prastų kompiuterio skaičiavimo pajėgumų. Reikėjo palaukti 60 metų, kad kartu su interneto proveržiu masiniai duomenys būtų lengvai pasiekiami visur – prasideda mašininio mokymo laikai. Dar kiek palaukus ir padobulėjus kompiuterio procesoriams ir vaizo plokštėm, atsirado galimybė tiksliai atkurti realiaus gyvenimo scenarijus kompiuterio ekrane, juose patalpinant objektus sąveikaujančius su aplinka. Taip su kompiuterio pagalba, atliekant didelius skaičiavimus, objektai „išmoksta“ prisitaikyti prie aplinkos pagal mūsų norus.

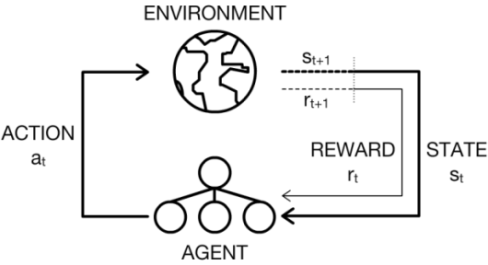
Mes naudodamiesi Unity žaidimų varikliu sukūrėme supaprastintą grafinę aplinką mūsų objektams, kurioje stengtasi simuliuoti kosminę erdvę tarp planetos Žemės ir mėnulio, su tikslu, kad objektai iš pastovios pradinės vietos išmoks pasiekti taikinį, kurio vieta keisis priklausomai nuo žaidėjo įvesties. Tikslą pasiekiamė naudodamiesi daugeliu internetiniu resursų, taip parašydami savo supaprastintą neuroninio tinklo architektūros sudarymo biblioteką ir algoritmus, paremtus skatinamuoju mokymu (angl. *Reinforcement learning*), skirtus objektų apmokymui.

Grafinės aplinkos kūrimas ir skatinamasis apmokymas

Saktinamasis mokymas is mašininio mokymo disciplina tyrinėjanti kaip objektai (angl. *Agents* arba *Software agents*) turėtų priimti sprendimus ir judėti sukurtoje aplinkoje, kad gautų kuo didesnės skaitmeninės vertės „apdovanojimą“. Šis mokymo būdas yra vienas iš trijų populiariausių mašininio mokymo būdų kartu su prižiūrimuoju mokymu (angl. *Supervised learning)* ir neprižiūriamuoju mokymu (angl. *Unsupervised learning)*.

**Skatinamasis mokymas susideda iš keturių esminių elementų:**

1. **Objektas.** Programa, kuri yra ”treniruojama“, kad atliktų kūrėjo paskirtą specifinę užduotį
2. **Aplinka.** Pasaulis, realus arba virtualus, kuriame objektas atlieka veiksmus.
3. **Veiksmas.** Veiksmas atlikas objekto, kuris lemia aplinkos būsenos pasikeitima.
4. **Apdovanojimas.** Veiksmo įvertinimas, kuris gali būti teigiamas arba neigiamas.



**Pavyzdžiui:**

**Objektas.** Progrmam kontroliuojanti robotą.

**Aplinka.** Realus pasaulis.

**Veiksmas.** Vienas iš keturių veiksmų: pirmyn (1), atgal (2), kairėn (3), dešinėn (4).

**Apdovanojimas.** Teigiamas kai robotas sėkmingai pasiekia tikslą. Neigiamas jei jis nukrenta, per ilgai užtrunka, ne ten nuvažiuoja.

Šiame pavyzdyje robotas gali save „išmokyti“ važiuoti efektyviau ir tiksliau reaguojant į savo gautas apdovanojimo vertes.

**Mūsų sukurto proceso aprašymas:**

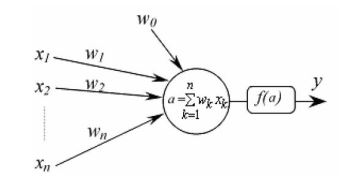
Sukūrėme virtualią aplinka naudodamiesi Unity žaidimų varikliu, kuri susideda iš komponentų: pradinio objektų iniciavimo taško (Žemė ekrano apačioje), tikslo (kita planeta viršuje), kurio vieta priklauso nuo stebėtojo įvesties, objekto (raketos). Simuliacijos pradžioje yra inicijuojama populiacija originalaus objekto kopijų (generacija), sukuriamas ir kiekvienam objektui priskiriamas neuroninis tinklas su randominėmis reikšmėmis. Objektai atlieka veiksmus pagal neuroninių tinklų gautas išvestis, po kiek sekundžių generacija nutraukiama, surušiuojami objektai apdovanojimo vertės tvarka, kuri labiausiai priklauso nuo pasiekto atstumo iki tikslo. Geriausiai pasirodžiusiųjų neuroniniai tinklai yra sumaišomi (angl. *crossover),* pradedama nauja generacija, su viltimi, kad objektai atlikts savo užduotį geriau nei praeitą kartą.

Dirbtiniai neuroniniai tinklai

Dirbtiniai neuroniniai tinklai (DNT) (angl. *artificial neural networks*) – tai informacijos apdorojimo struktūros, netiksliai imituojančios, kai kuriuos gyvųjų organizmų smegenyse vykstančius informacijos apdorojimo procesus. DNT sudaromi iš daugelio tarpusavyje sujungtų labai paprastų skaičiavimo elementų. Šie elementai jungiami vieni su kitais įvairaus stiprumo jungtimis.

**Perceptrono (DNT vieno neurono) architektūra:**

Labiausiai paplitusi ir paprasčiausia dirbtinio neurono schema pavaizduota paveiklėlyje apačioje. Čia *x1, ..., xn* žymi neurono įvestis. Jungčių „stiprumas“ nustatamas pagal svorių turimą vertę (Realieji skaičiai), kurie yra pažymėti *w1,...,wn*. Žymėjimas w0 reiškia slenksčio vertę (angl. *bias,* dažniausiai skirtas apsisaugoti 0 išvesties). Neurono viduje atliekama įvesčių ir svorių sandaugos suma, kuri po to perleidžiama pro aktyvacijos funkciją (pati paprasčiausia yra vadinama *sigmoid,* paverčianti visas vertes atitinkamai nuo 0 iki 1), skirta „sureguliuoti“ gautą rezultatą, o *y* – neurono išvestis.



**DNT mokymas:**