

Zmija – učenje podsticajem

Ovaj projekat služi za treniranje agenta koji uči da sam igra igru i poboljšava svoj rezultat.

Kod za implementaciju same igre je preuzet sa github repozitorijuma:

<https://github.com/clear-code-projects/Snake>

Da bi kod radio, potrebne su sledeće biblioteke:

- numpy
- matplotlib
- pygame

Kod se sastoji od 5 fajlova:

- agent.py
- snake_game.py
- snake.py
- fruit.py
- settings.py

Settings fajl sadrži promenljive u kojima se mogu menjati veličine polja, kao i veličina table. Da bi se promenila veličina mreže (table) za igru potrebno je promeniti vrednost promenljive `cell_number`, a za promenu veličine polja se menja promenljiva `cell_size`. Ne preporučuje se promena `cell_size` jer može doći do cepanja grafike.

Fajlovi `fruit.py` i `snake.py` sadže klase za iscrtavanje objekata (zmije i jabuka) koji se koriste u igri. Klasa `Fruit` služi za nasumično iscrtavanje jabuke na igračkoj tabli. Klasa `Snake` služi za iscrtavanje igrača na tabli, kao i za ažuriranje grafike pri kretanju zmije po tabli.

`Snake game` fajl vezuje sve u celinu i služi za pokretanje same igre. Sadrži funkcije za iscrtavanje trave na tabli, funkciju za proveru kolizije sa ivicama table, koliziju sa samim sobom, kao i za kontroler.

Fajlovi `fruit.py`, `snake.py` i `snake_game.py` ne bi trebali da budu menjani.

Agent

Informacije:

- Nagrada ispred
- Nagrada iza
- Nagrada levo
- Nagrada desno
- Prepreka ispred

- Prepreka levo
- Prepreka desno

Potrebno je ukupno 7 bitova da bi predstavili sva stanja, pa je ukupan broj mogućih stanja $2^7 = 128$. Informacije služe da daju sledeće stanje u kome će se agent naći, nakon čega se preduzima najpovoljnija moguća akcija.

Akcije

Agent može da odabere između tri akcije:

- Nastavi pravo
- Skreni levo
- Skreni desno

Nakon odabira akcije, algoritam uzima informacije i računa nova stanja. Nakon svake preduzete akcije, agent dobija nagrade (ili kazne) u zavisnosti od preduzetog poteza.

Nagrade

Agent može da dobije sledeće nagrade:

- Ukoliko se agent kreće ka jabuci: +5
- Ukoliko se agent kreće od jabuke: -5
- Ukoliko agent pojede jabuku: +500
- Ukoliko agent udari u prepreku: -1000

Hiperparametri

Početna stopa učenja i parametar epsilon za epsilon-greedy politiku su 0.5 i 0.01.

Na tabli dimenzija 10 puta 10, najveći dobijeni rezultat je bio 31.

Na tabli dimenzija 15 puta 15, najveći dobijeni rezultat je bio 50.

Moguće je promeniti brzinu izvršavanja samog programa. Smanjenje brzine kojom se igra odvija se izvodi tako što se menja vrednost promenljive:

```
self.snake_game.clock.tick()
```

Preporučuje se vrednost između 120 i 300.

Pokretanje programa

Na Windows operativnom sistemu, program se pokreće iz command prompt konzole pomoću komande `python3 agent.py`. Ukoliko želite da samostalno igrate potrebno je pokrenuti komandu `python3 snake_game.py`. Neophodno je da potrebne biblioteke budu instalirane. Biblioteke se mogu instalirati komandom `pip3 install ime_biblioteke`.

Na GNU/Linux operativnim sistemima je potrebno dodati dozvole za izvršavanje. Ovo se postiže komandom `chmod +x ime_fajla.py`. Nakon promene dozvola, dovoljno je ukucati `python3 agent.py` u terminal i program će biti izvršen. Pokretanje samostalnog igranja se vrši izvršavanjem `python3 snake_game.py` u terminalu.

MacOS operativni sistem niko ne koristi, pa je besmisleno pisati ikakva uputstva.