

Automatizarea deciziilor agenților într-un mediu virtual

Uicoabă Alexandru

Întrebările de cercetare

- ◆ Care sunt datele de care agentul are nevoie pentru a învăța?
- ◆ Care algoritmi de învățare automată sunt mai performanți pentru task-ul acesta?
- ◆ Poate un agent fi mai bun decât o persoană umană?



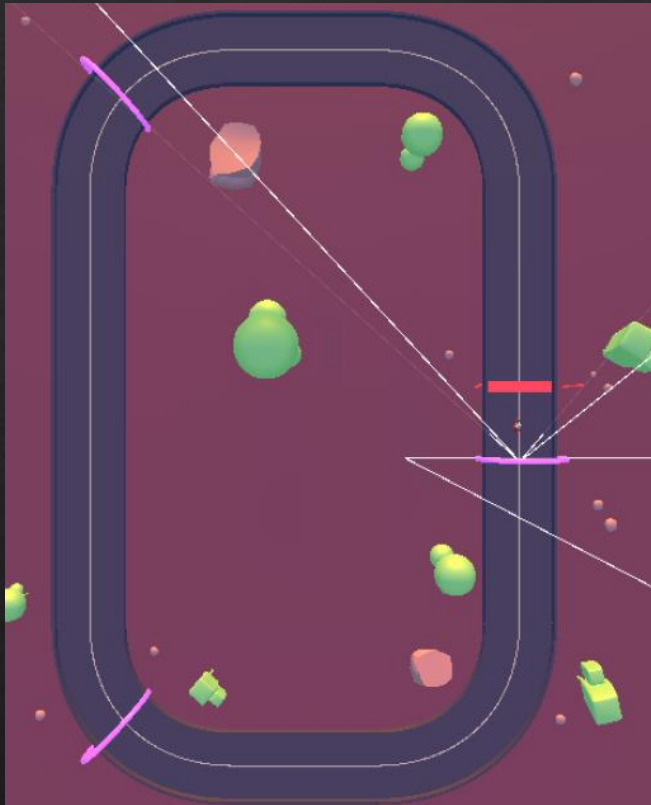
Objective

- ◆ Analiză teoretică
 - ◆ Soluții conexe
 - ◆ Platforme software necesare
- ◆ Realizare experiment
 - ◆ Adaptare mediu Unity
 - ◆ Realizare infrastructură
 - ◆ Antrenare agent



Mediul Unity

◆ Formă circuit



◆ Imaginea jucătorului (POV)

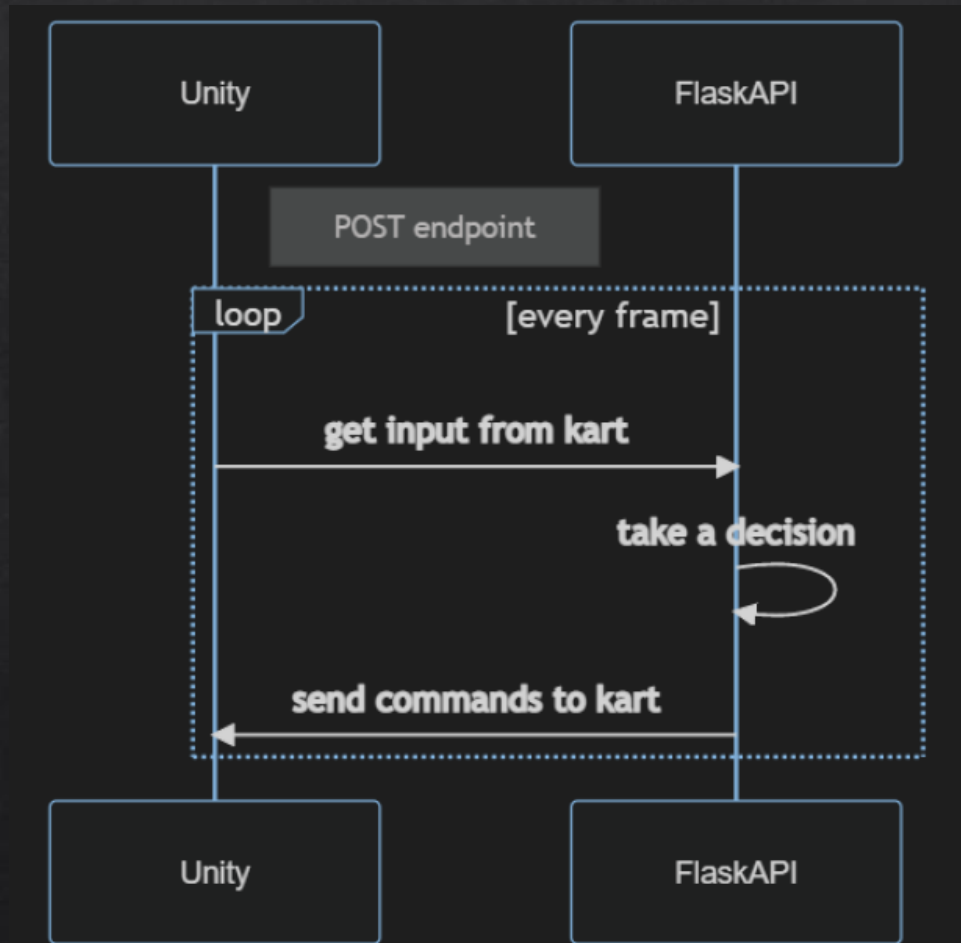


Model Agent

- ◆ Time
- ◆ xPos, yPos, zPos
- ◆ Sensors
- ◆ Zone
- ◆ MovingForward
- ◆ gameOver
- ◆ State



S1 - Flux de date învățare automată supravegheată



S1. Rezultate

	Utilizator uman	Random Forest	Decision Tree
Turul 1	27 secunde	29 secunde	29 secunde
Turul 2	27 secunde	28 secunde	28 secunde
Turul 3	28 secunde	27 secunde	28 secunde
Turul 4	26 secunde	27 secunde	29 secunde
Turul 5	26 secunde	28 secunde	28 secunde

S2 . Flux de date – Reinforcement Learning



S2. OpenAI Gym

◇ Structura pe directoare

```
main.py
setup.py
├── gym_examples
│   ├── __init__.py
│   └── envs
│       ├── custom_env.py
│       ├── __init__.py
│       ├── __pycache__
│       │   ├── custom_env.cpython-38.pyc
│       │   └── __init__.cpython-38.pyc
│       └── __pycache__
│           ├── __init__.cpython-38.pyc
```

◇ Recompense/Penalizări

Distanța până la obstacol	Recompensă
$dis \geq 5$	+ 10
$dis < 5 \ \&\& \ dis \geq 2.5$	-50
$dis < 2.5 \ \&\& \ dis \geq 1$	-1000
$dis < 1$	-5000
movingForward = true	+100
movingForward = false	-5000

S2. Rezultate

- ◆ După un număr ridicat de episoade agentul a reușit să facă task-ul
- ◆ Limitări
 - ◆ numărul mare de episoade necesare



Concluzii

- ◊ Utilizator uman ~ algoritmi de învățare automată supravegheați
- ◊ Învățare automată supravegheată
 - ◊ Circuit cunoscut
 - ◊ Colectare anterioară a datelor
 - ◊ Durează puțin antrenarea
- ◊ Reinforcement Learning
 - ◊ Circuit generat în mod aleatoriu
 - ◊ Număr mare de episoade
 - ◊ Durează mult antrenarea

Dezvoltări viitoare

- ◆ Circuit generat în mod aleatoriu
- ◆ Mai mulți agenți simultan
- ◆ Mai multe tururi de circuit
- ◆ Circuit mai complex
- ◆ Mai multe episoade pentru RF





Mulțumesc
pentru atenție!

Întrebări?