

Основи машинског и дубоког учења

Тема:
Аутономна вожња формула имплементирана
као
видео игра коришћењем учења подстицајем

Студенти:
Михаило Максимовић 602/2018
Ђорђе Гачић 626/2018

Професор:
Др Владимир Миловановић
Асистент:
Никола Радовановић



Коришћене технологије

- Unity
- ML-Agents
- Python 3.7.9
- C#
- Latex
- www.overleaf.com

Github:

https://github.com/djoto/ML-Agents-_Unity_FormulaRacing_project



Мотив и идеја

- Мотив: аутономна вожња и њена све већа популарност
- Идеја: тренирати модел формуле тако да буде равноправан са играчем који мануелно управља

Изглед графичких компоненти - формула

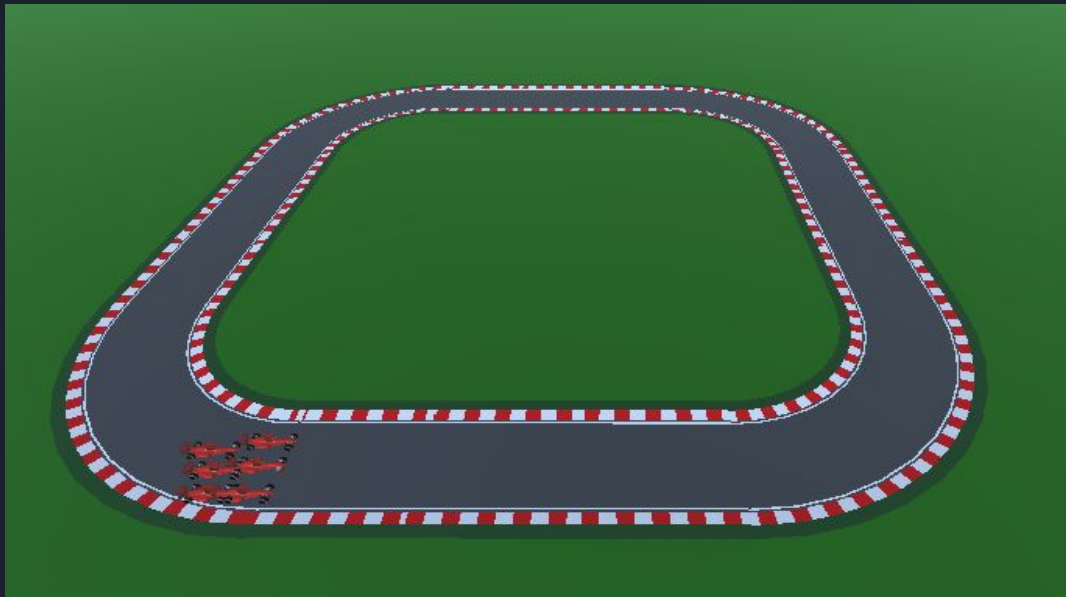


Извор:

<https://assetstore.unity.com/packages/3d/vehicles/land/formula-one-car-25693>

Изглед графичких компоненти

- стаза на лево

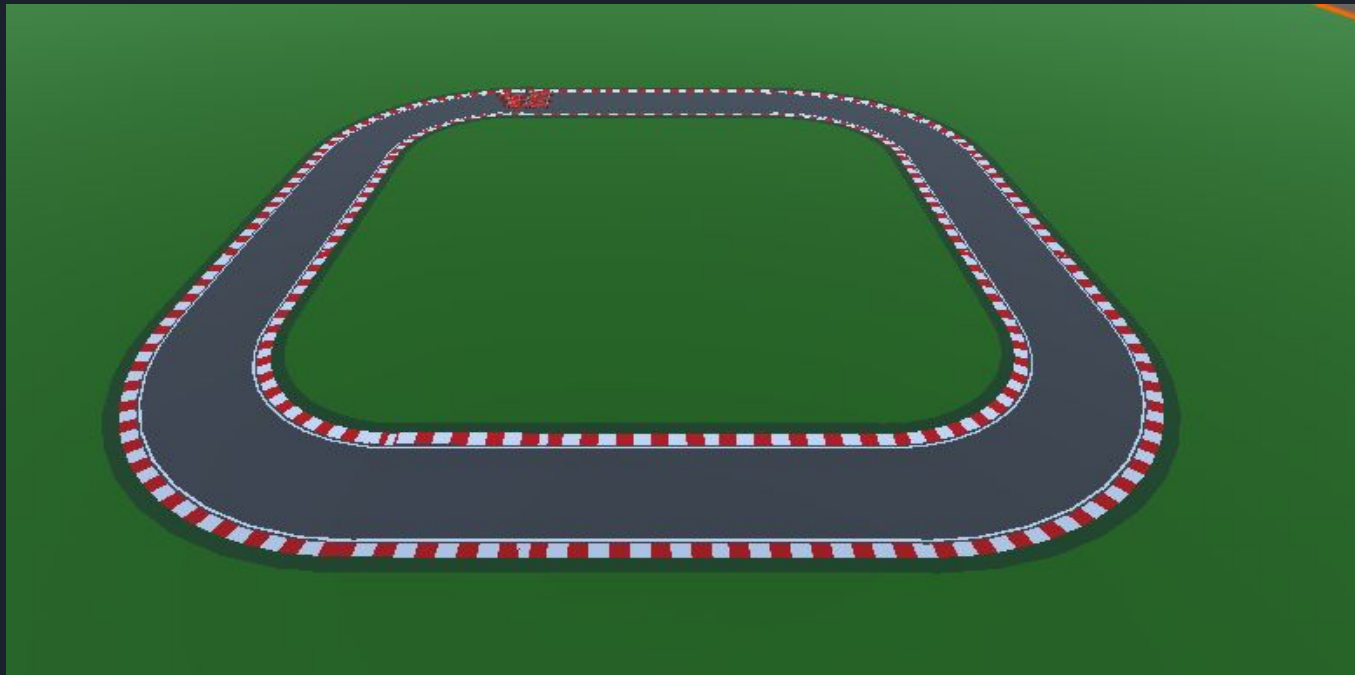


Извор:

<https://assetstore.unity.com/packages/3d/environments/roadways/modular-lowpoly-track-roads-free-205188>

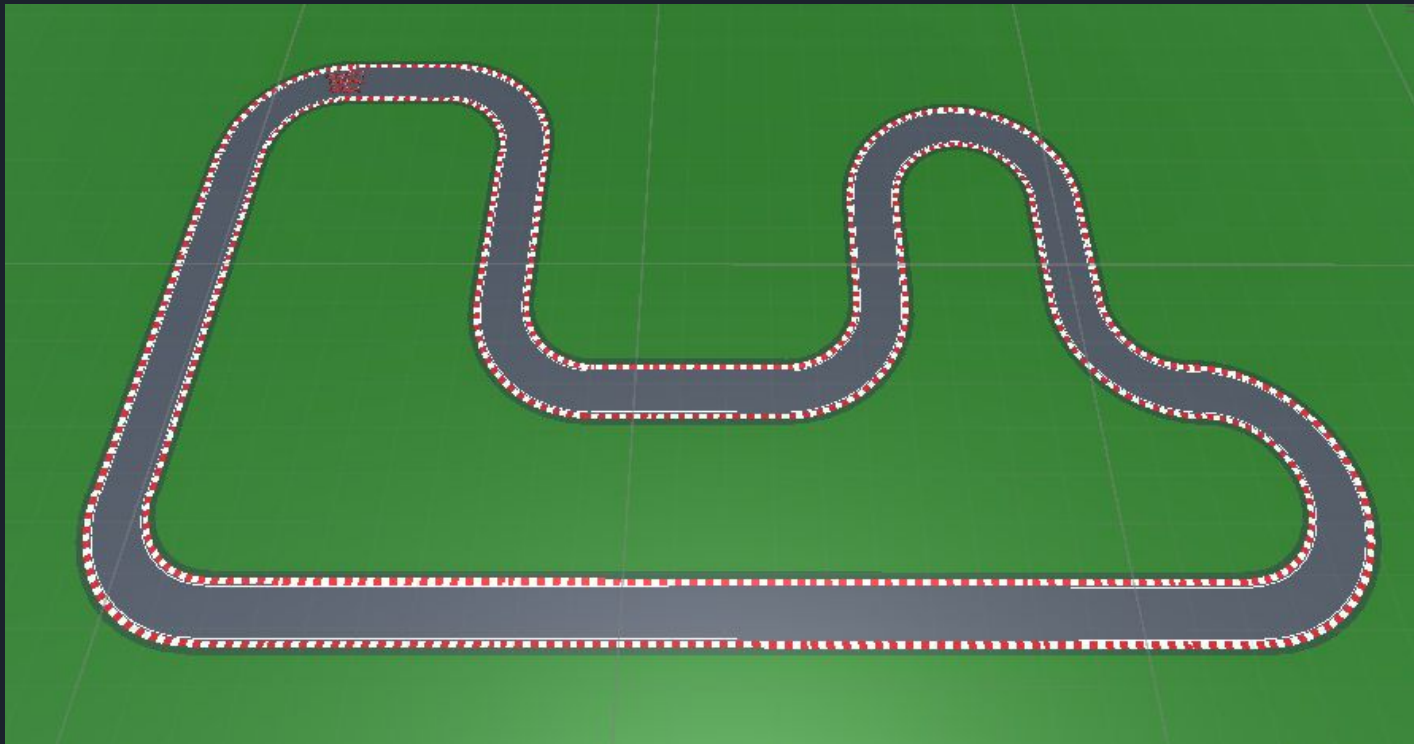
Изглед графичких компоненти

- стаза на десно



Изглед графичких компоненти

- стаза са комбинованим скретањима

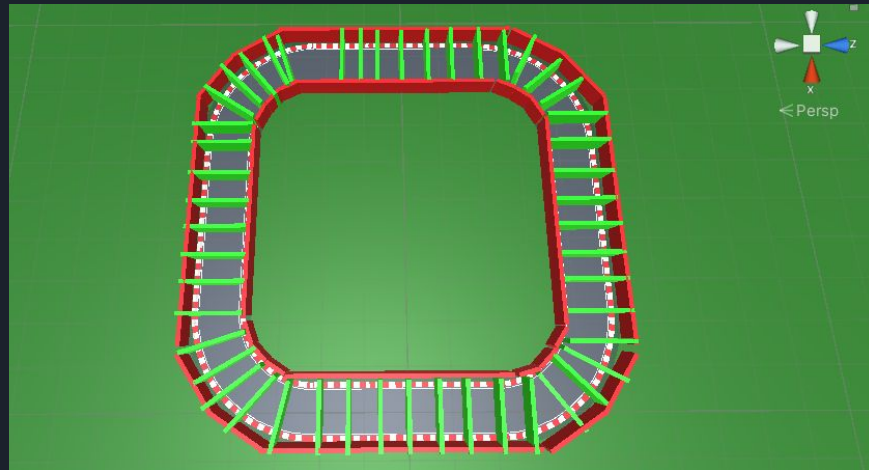
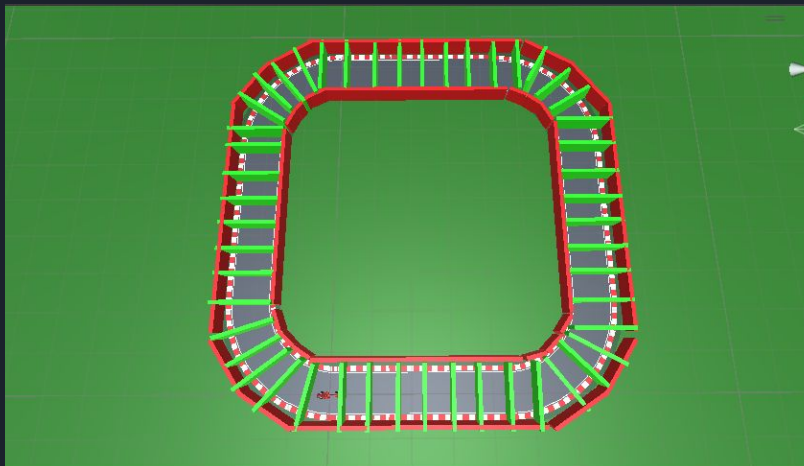


Графичке компоненте

- невидљиви зидови и контролне тачке (checkpoints)

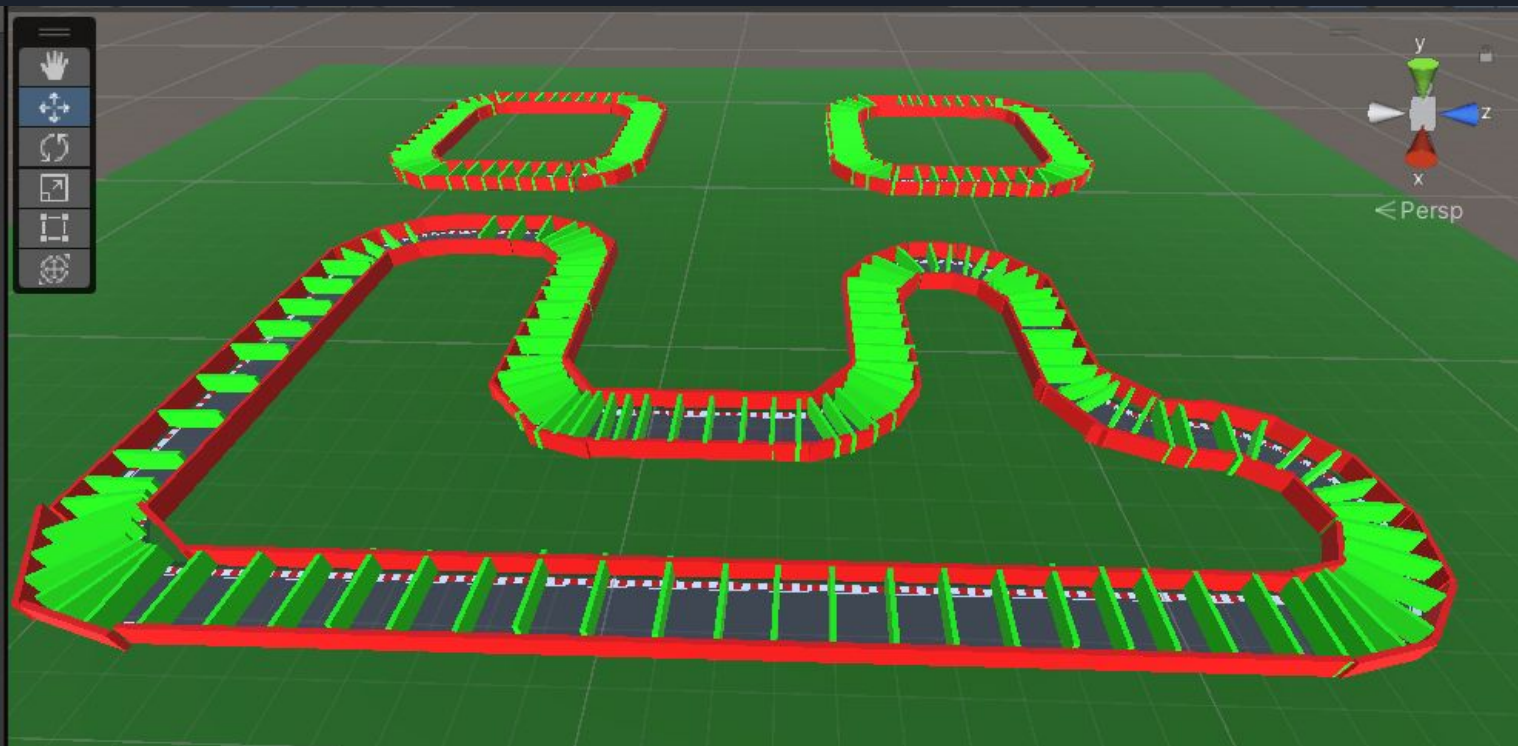
- C#
- Датотека CheckpointSingle.cs
- Датотека TrackCheckpoints.cs

Потребно за учење подстицајем !!!



Графичке компоненте

- невидљиви зидови и контролне тачке (checkpoints)






Функциональность модели формуле

- C#
- Датотека carDriver.cs
- Датотека carDriverPlayer.cs

УЧЕЊЕ ПОДСТИЦАЈЕМ






Учење подстицајем


- тренирање модела формуле

- C#
 - Датотека carDriverAgent.cs
 - Датотека CheckpointSingle.cs
 - Датотека TrackCheckpoints.cs
- Циљ: научити модел формуле да прелази стазу како би била (приближно) равноправна са играчем који мануелно управља формулом



Учење подстицајем стања и акције

- Број стања: 6
 1. убрзавање
 2. успоравање
 3. стајање у месту
 4. скретање на десно
 5. скретање на лево
 6. кретање право
- Број акција: 2
 1. промена брзине
 2. промена правца кретања



Учење подстицајем тренирање модела

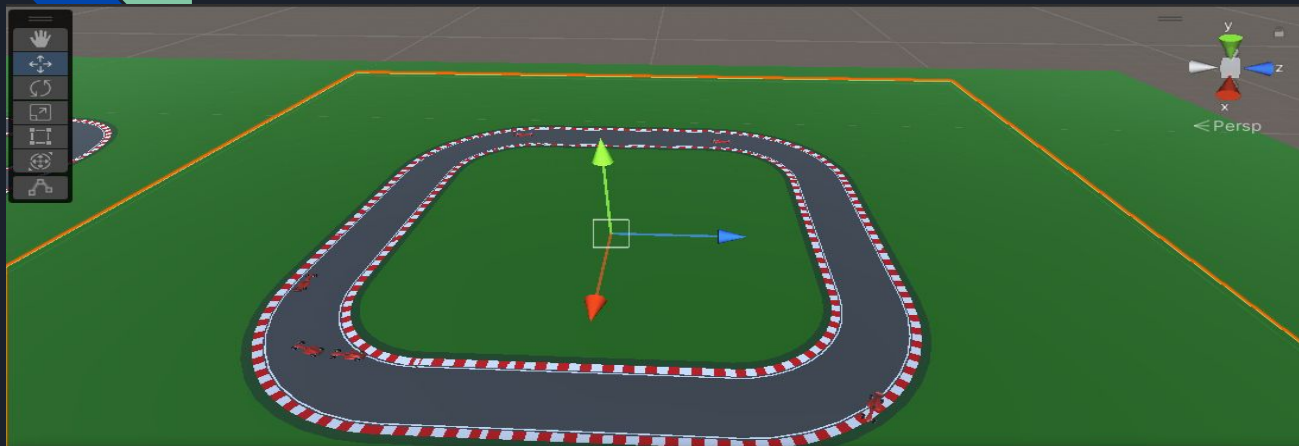
- Тренирање на стази са скретањима на лево
- Тренирање на стази са скретањима на десно
- Тренирање на стази са комбинованим скретањима (најзахтевније)

Циљ: максимизација укупне награде

Конфигурациона датотека за учење подстицајем: FormulaRacing.yaml

Научени модел тј. “мозак” алгоритма смешта се у FormulaRacing.onnx датотеку

Учење подстицајем приказ у току тренирања



- Видимо како се
укупна награда
временом
повећава

Anaconda Prompt (miniconda3) - mlagents-learn FormulaRacing.yaml --run-id=formulaDesno --initialize-from=formula

```
c:\users\mihai\miniconda3\lib\site-packages\mlagents\trainers\torch\networks.py:91: UserWarning: Creating a tensor from a list of tensors with the tensor constructor. Please consider converting the list to a single numpy.ndarray with numpy.array() before converting to a tensor. (Triggered by call to torch.tensor(...), at c:\users\mihai\miniconda3\lib\site-packages\mlagents\trainers\torch\networks.py:91)
enc.update_normalization(torch.as_tensor(vec_input))
```

```
[INFO] FormulaRacing. Step: 10000. Time Elapsed: 44.649 s. Mean Reward: 6.809. Std of Reward: 3.591. Training.
[INFO] FormulaRacing. Step: 20000. Time Elapsed: 70.622 s. Mean Reward: 8.468. Std of Reward: 1.440. Training.
[INFO] FormulaRacing. Step: 30000. Time Elapsed: 99.644 s. Mean Reward: 8.815. Std of Reward: 0.550. Training.
[INFO] FormulaRacing. Step: 40000. Time Elapsed: 128.496 s. Mean Reward: 8.344. Std of Reward: 1.725. Training.
[INFO] FormulaRacing. Step: 50000. Time Elapsed: 157.066 s. Mean Reward: 17.786. Std of Reward: 11.939. Training.
[INFO] FormulaRacing. Step: 60000. Time Elapsed: 186.060 s. Mean Reward: 17.111. Std of Reward: 14.052. Training.
[INFO] FormulaRacing. Step: 70000. Time Elapsed: 214.714 s. Mean Reward: 9.607. Std of Reward: 2.726. Training.
[INFO] FormulaRacing. Step: 80000. Time Elapsed: 243.482 s. Mean Reward: 9.571. Std of Reward: 2.629. Training.
[INFO] FormulaRacing. Step: 90000. Time Elapsed: 272.486 s. Mean Reward: 20.917. Std of Reward: 13.037. Training.
[INFO] FormulaRacing. Step: 100000. Time Elapsed: 301.117 s. Mean Reward: 35.773. Std of Reward: 33.080. Training.
[INFO] FormulaRacing. Step: 110000. Time Elapsed: 329.519 s. Mean Reward: 52.500. Std of Reward: 32.463. Training.
[INFO] FormulaRacing. Step: 120000. Time Elapsed: 358.171 s. Mean Reward: 68.722. Std of Reward: 61.528. Training.
[INFO] FormulaRacing. Step: 130000. Time Elapsed: 386.835 s. Mean Reward: 62.119. Std of Reward: 81.575. Training.
```



Демонстрација

- Приказ мануелног управљања
- Приказ наученог модела
- Приказ процеса учења



Хвала на пажњи!