

第4回 全脳アーキテクチャ シンポジウム

～ 脳に学ぶ統合を問い直す ～

2019年6月26日 13:00 開始

主 催：NPO法人 全脳アーキテクチャ・イニシアティブ
パナソニック株式会社



■ Animal-AI

触ってみました。



■ 自己紹介

株式会社TOPWELL

稗田 雅昭 (AI担当)



TOPWELL
Make ALL Happy

■ 環境(v0.5時点)

- ①Python3.6
- ②必要なライブラリをインストール
(animalai、animalai-train、tensorflow)

後はファイル式をダウンロードするだけで実行できます。

<https://github.com/beyretb/AnimalAI-Olympics>

※実行ファイル(env)はOS毎に用意されています。

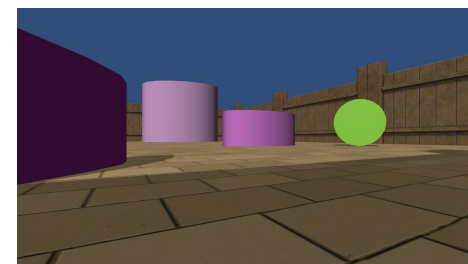
OS	Environment link
Linux	download v0.5
MacOS	download v0.5
Windows	download v0.5

■ 手動制御

ダウンロードしてきた実行可能ファイルを実行すると、キーボードで操作できる状態で起動します。
※左図中央の水色の動物を操作できます。



キーボードキー	アクション
W	エージェントを前に移動
S	エージェントを後退させる
A	エージェントを左に曲げる
D	エージェントを右に向ける
C	カメラを切り替える
R	環境をリセットする



■ オブジェクトの種類:エサ

・良いエサ



取るとポイントがプラスされ、次のイベントが始まる。

GoodGoal : 動かない

GoodGoalMove : 動く (物にぶつくと止まる)

GoodGoalBounce : 動く (物にぶつくと跳ね返る)

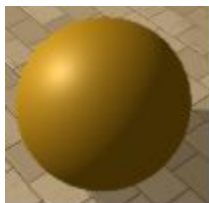
・悪いエサ



取るとポイントが減られ、次のイベントが始まる。

※動きは「良いエサ」と同様3種類

・良いエサ(マルチ)

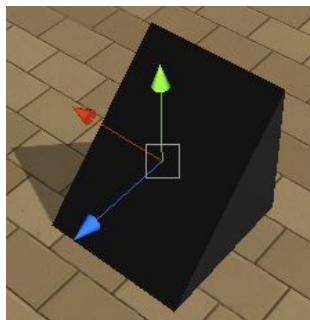
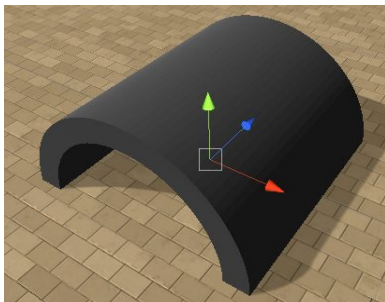


取るとポイントがプラスされる。全てを取り終わると次のイベントが始まる。

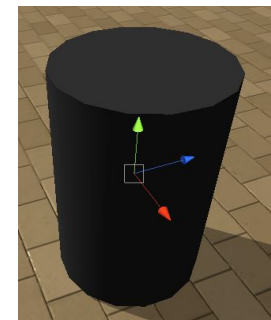
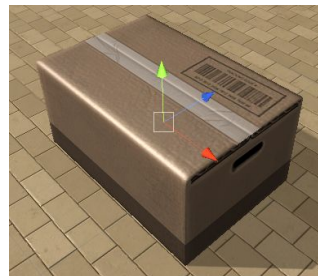
※動きは「良いエサ」と同様3種類

■ オブジェクトの種類:障害物

・動かない障害物(例)

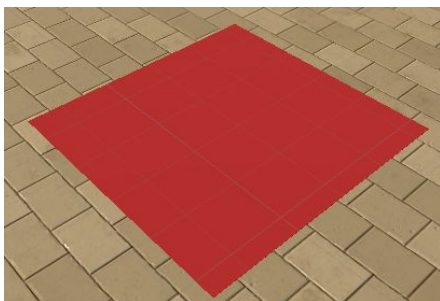


・動かせる障害物(例)



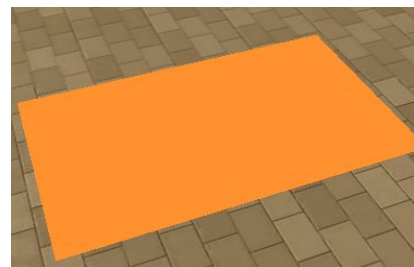
■ オブジェクトの種類:床

・デスゾーン



踏むとポイントが減られ、次のイベントが始まる。

・ホットゾーン



踏むと一定時間毎にポイントが減られる。

■ アリーナ設定ファイル

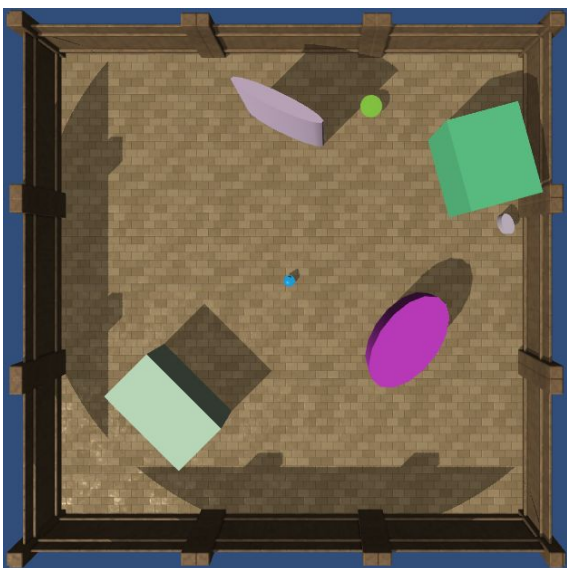
設定ファイルを指定して起動すると、オブジェクトの配置を変更できます。

```
python visualizeArena.py configs/configExample.yaml
```

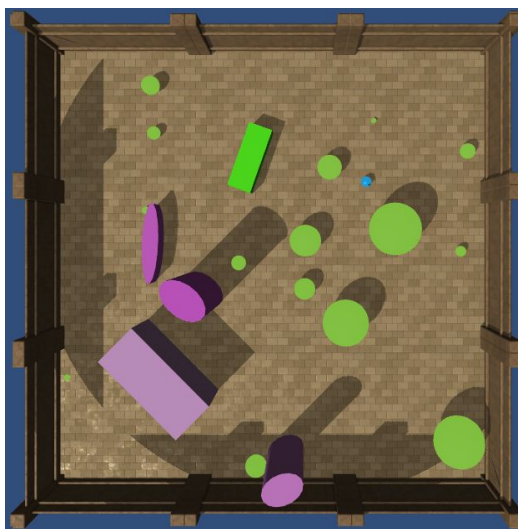
※オブジェクト毎に下記を指定可能

positions: 初期位置 sizes: サイズ rotations: 向き colors: 色

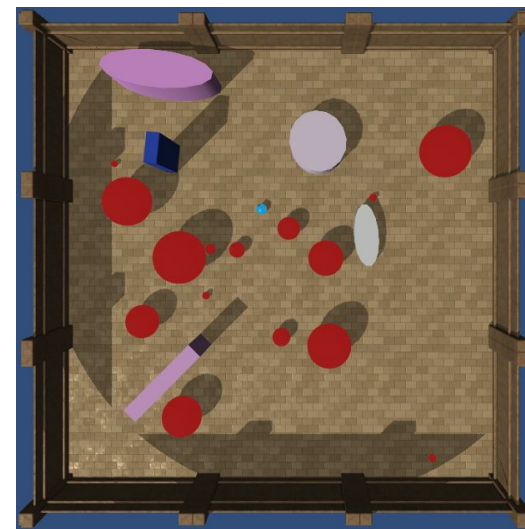
・初期設定



・変更後1



・変更後2



■ トレーニング

トレーニングエージェントは2つ用意されています。

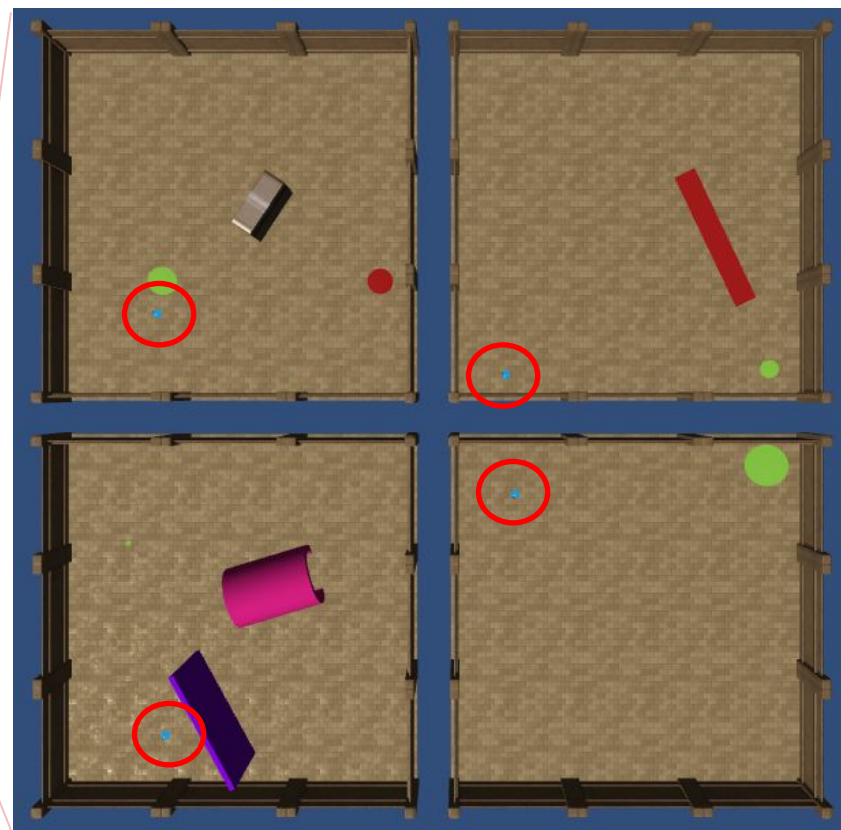
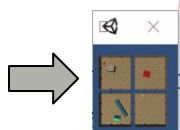
ML-Agentsだと4つ同時にトレーニングできるとのことなので、そちらを使用しました。

- ・ML-Agents (4つ同時に学習可)
- ・Gym

実行すると設定情報がログに出力されて、小さいウィンドウが起動します。

...よく見ると、水色の丸が動いています。

```
管理者: コマンドプロンプト - python trainMLAgents.py
batch_size: 64
beta: 0.01
buffer_size: 2024
epsilon: 0.2
gamma: 0.99
hidden_units: 256
lambda: 0.95
learning_rate: 0.0003
max_steps: 5.0e6
normalize: False
num_epoch: 3
num_layers: 1
time_horizon: 128
sequence_length: 64
summary_freq: 1000
use_recurrent: False
summary_path: ./summaries/train_example-1_Learner
memory_size: 256
use_curiosity: True
curiosity_strength: 0.01
curiosity_enc_size: 256
model_path: ./models/train_example/Learner
```



■ トレーニング結果

トレーニングを実行して暫くすると結果のログが出力されます。

平均報酬 (Mean Reward) : 最初はマイナスが多いですが、学習が進むとプラスが多くなっていきます。

```
Example/Learner
Learner: Step: 1000 Mean Reward: -0.496 Std of Reward: 1.642.
Learner: Step: 2000 Mean Reward: -0.638 Std of Reward: 0.888.
Learner: Step: 3000 Mean Reward: 0.628 Std of Reward: 1.702.
Learner: Step: 4000 Mean Reward: -0.300 Std of Reward: 1.131.

Learner: Step: 5000 Mean Reward: 0.126 Std of Reward: 2.122.
Learner: Step: 6000 Mean Reward: 0.051 Std of Reward: 1.879.
Learner: Step: 7000 Mean Reward: -0.442 Std of Reward: 1.116.
Learner: Step: 8000 Mean Reward: -0.179 Std of Reward: 1.642.
Learner: Step: 9000 Mean Reward: 0.084 Std of Reward: 1.338.

Learner: Step: 10000 Mean Reward: 0.831 Std of Reward: 2.259.
Learner: Step: 11000 Mean Reward: 1.678 Std of Reward: 1.801.
Learner: Step: 12000 Mean Reward: -0.508 Std of Reward: 0.983.
Learner: Step: 13000 Mean Reward: -0.021 Std of Reward: 2.022.
Learner: Step: 14000 Mean Reward: 1.481 Std of Reward: 2.259.

Learner: Step: 15000 Mean Reward: -0.465 Std of Reward: 1.757.
Learner: Step: 16000 Mean Reward: 0.104 Std of Reward: 1.626.
Learner: Step: 17000 Mean Reward: 1.191 Std of Reward: 2.141.
Learner: Step: 18000 Mean Reward: 1.022 Std of Reward: 2.104.
Learner: Step: 19000 Mean Reward: 0.933 Std of Reward: 2.091.

Learner: Step: 20000 Mean Reward: 1.620 Std of Reward: 2.072.
Learner: Step: 21000 Mean Reward: 1.586 Std of Reward: 2.025.
Learner: Step: 22000 Mean Reward: 0.923 Std of Reward: 2.029.
Learner: Step: 23000 Mean Reward: -0.022 Std of Reward: 1.784.
Learner: Step: 24000 Mean Reward: 2.048 Std of Reward: 1.941.
```

■ 所感

サンプルが用意されているので、AIを勉強中の方でもトレーニングや設定変更ができます。

Pythonに慣れている人でしたらトレーニングの改善も問題なくできると思われます。

株式会社TOPWELL
稗田 雅昭



TOPWELL
Make ALL Happy