

# 強化学習を用いたスネークゲームの経路探索手法の検討

---

東京情報大学 総合情報学科 情報システム学系4年

知能情報学 システム研究室 永井ゼミ

J19068 王建州

# 目次

---

- 研究動機や目的
- 類似研究の現状
- 探索アルゴリズム
- 使用するソフトウェア
- 環境や評価
- 実行結果
- 学習中の様子
- 観察結果
- 今後の計画
- 参考文献

# 研究の動機や目的

---

## 動機

- AI技術は社会に多くの応用があります
- ゲームAIの限界はどこまであるのか

## 目的

- 特定の環境で報酬を最大化
- DQNを利用して探索アルゴリズム、探索効率が上がりたい。

# 類似研究の現状

---

1) <https://www.cs.unh.edu/~ruml/cs730/paper-examples/merrill-2011.pdf>

スネーク: 人工知能コントローラー

2) [https://www.cse.ust.hk/~rossiter/independent\\_studies\\_projects/snake\\_ai/snake\\_ai\\_report.pdf](https://www.cse.ust.hk/~rossiter/independent_studies_projects/snake_ai/snake_ai_report.pdf)

スネークとその AI の実装

# 探索アルゴリズム

---

## 1) 最短経路(BFS)

幅優先探索のアルゴリズムを使用すると、グラフの任意の2つの頂点間の最短経路を見つけること。可能なすべてのノードを同時に探索します。BFSアルゴリズムの結果は高精度です。

## 2) 人間

キーボードでスネークを操作して、自己決定によって四つの方向変えられる。

## 3) ランダム探索

エージェントはランダムパスを選択して、目的に到達する。パスの選択は不特定の確率、順序が決まっていないデータまたは情報の集まり。

# 探索アルゴリズム

---

## 4) 強化学習DQN

最初のデータは不明。リプレイから経験を積み、サンプリングされた経験を使用してQネットワークを更新する。エージェントはネットワークを利用して、トレーニングを行うことで、目的に到達するまでパフォーマンスが高くなる。次の動きごとに、エージェントはその確率を計算し、一番いい行動を選択し実行する。

states	action			
	←	→	↑	↓
1	0	0.2	0	1
2	0.3	0	0.98	0
3	-0.5	1	-0.43	0.91
4	1	-0.76	-0.5	-0.42
⋮	⋮	⋮	⋮	⋮
11	0.6	-0.1	0.22	-2.9

# 使用するソフトウェア

---

## 1)プログラミング言語

—python

## 2)ライブラリ

—pygame

## 3)ソフトウェア

—visual studio code

—pycharm

# 環境や評価方法

---

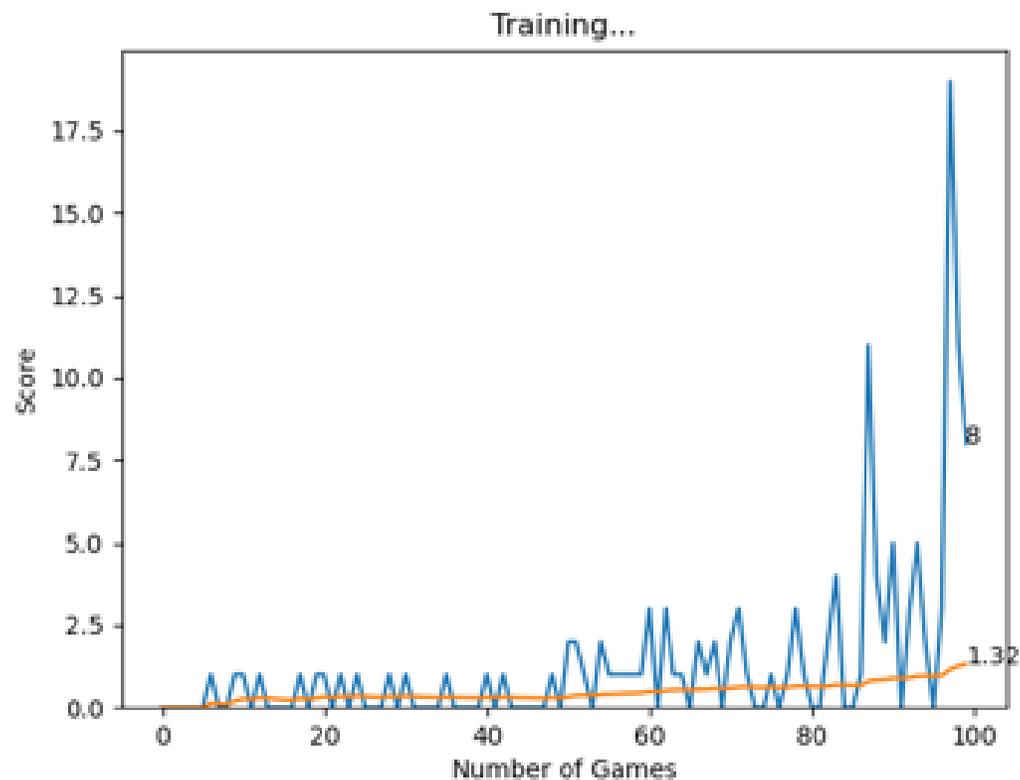
## 環境

- ヘビは境界を越えることができない
- ヘビの頭は移動方向、上、下、左、右を制御する

## 評価方法

- 壁にぶつかると負ける
- ヘビは、食べ物を食べたり、環境の中を移動したりしている間、体の一部や壁を避けなければならない
- ヘビは食べ物を食べる必要があり、ヘビが食べ物を食べるたびに、ヘビの体は一定の要因で成長する

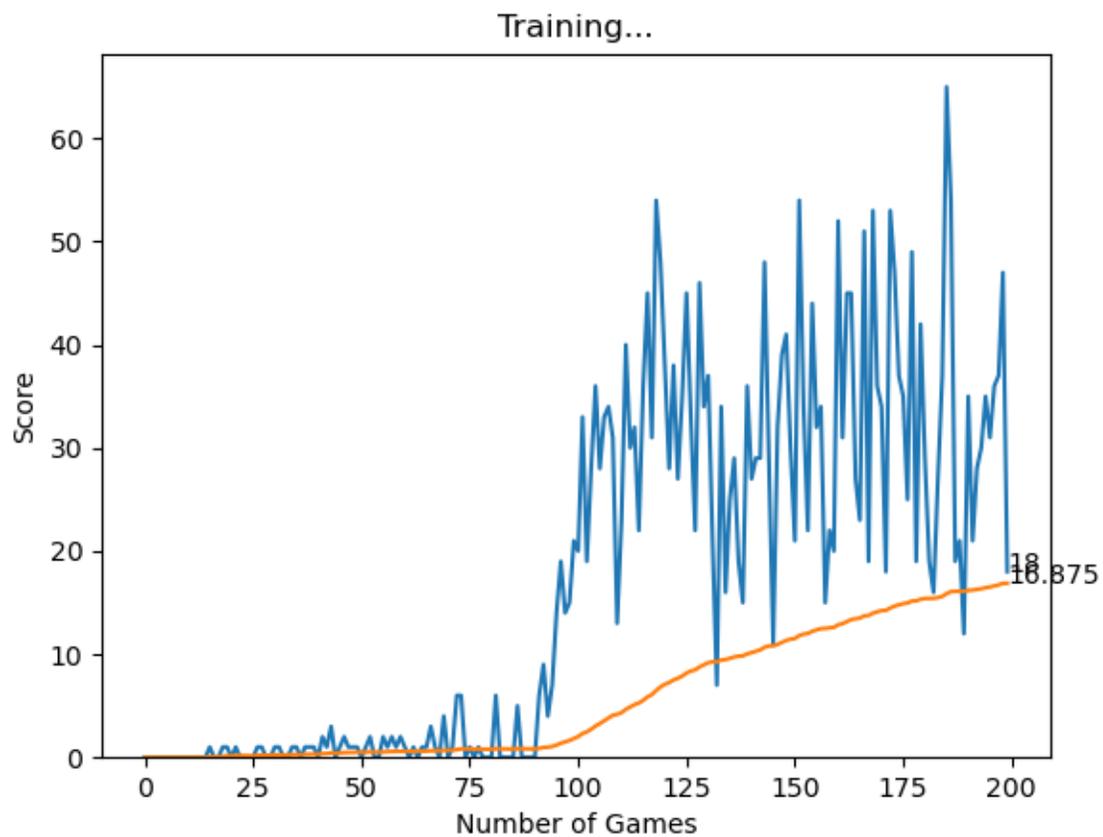
# 実行結果



```
Figure(640x480)
Game 77 Score 0 Record: 3
Figure(640x480)
Game 78 Score 1 Record: 3
Figure(640x480)
Game 79 Score 3 Record: 3
Figure(640x480)
Game 80 Score 1 Record: 3
Figure(640x480)
Game 81 Score 0 Record: 3
Figure(640x480)
Game 82 Score 0 Record: 3
Figure(640x480)
Game 83 Score 2 Record: 3
Figure(640x480)
Game 84 Score 4 Record: 4
Figure(640x480)
Game 85 Score 0 Record: 4
Figure(640x480)
Game 86 Score 0 Record: 4
Figure(640x480)
Game 87 Score 1 Record: 4
Figure(640x480)
Game 88 Score 11 Record: 11
Figure(640x480)
Game 89 Score 4 Record: 11
Figure(640x480)
Game 90 Score 2 Record: 11
Figure(640x480)
Game 91 Score 5 Record: 11
Figure(640x480)
Game 92 Score 0 Record: 11
Figure(640x480)
Game 93 Score 3 Record: 11
Figure(640x480)
Game 94 Score 5 Record: 11
Figure(640x480)
Game 95 Score 2 Record: 11
Figure(640x480)
Game 96 Score 0 Record: 11
Figure(640x480)
Game 97 Score 3 Record: 11
Figure(640x480)
Game 98 Score 19 Record: 19
Figure(640x480)
Game 99 Score 11 Record: 19
Figure(640x480)
Game 100 Score 8 Record: 19
Figure(640x480)
```

100ゲームを行い、最高のスコアは19である。

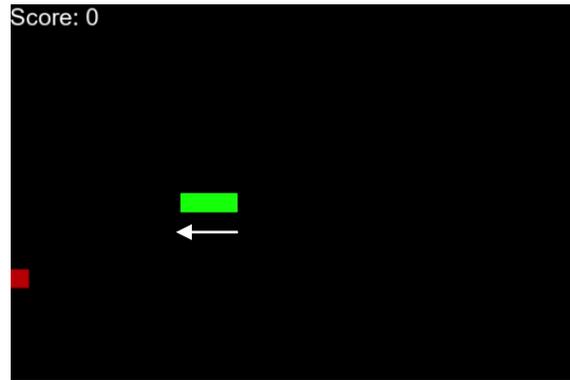
# 実行結果



```
Game 177 Score 25 Record: 54
Figure(640x480)
Game 178 Score 49 Record: 54
Figure(640x480)
Game 179 Score 19 Record: 54
Figure(640x480)
Game 180 Score 42 Record: 54
Figure(640x480)
Game 181 Score 29 Record: 54
Figure(640x480)
Game 182 Score 19 Record: 54
Figure(640x480)
Game 183 Score 16 Record: 54
Figure(640x480)
Game 184 Score 27 Record: 54
Figure(640x480)
Game 185 Score 37 Record: 54
Figure(640x480)
Game 186 Score 65 Record: 65
Figure(640x480)
Game 187 Score 54 Record: 65
Figure(640x480)
Game 188 Score 19 Record: 65
Figure(640x480)
Game 189 Score 21 Record: 65
Figure(640x480)
Game 190 Score 12 Record: 65
Figure(640x480)
Game 191 Score 35 Record: 65
Figure(640x480)
Game 192 Score 21 Record: 65
Figure(640x480)
Game 193 Score 28 Record: 65
Figure(640x480)
Game 194 Score 30 Record: 65
Figure(640x480)
Game 195 Score 35 Record: 65
Figure(640x480)
Game 196 Score 31 Record: 65
Figure(640x480)
Game 197 Score 36 Record: 65
Figure(640x480)
Game 198 Score 37 Record: 65
Figure(640x480)
Game 199 Score 47 Record: 65
Figure(640x480)
Game 200 Score 18 Record: 65
Figure(640x480)
```

200ゲームを行い、最高のスコアは65である。

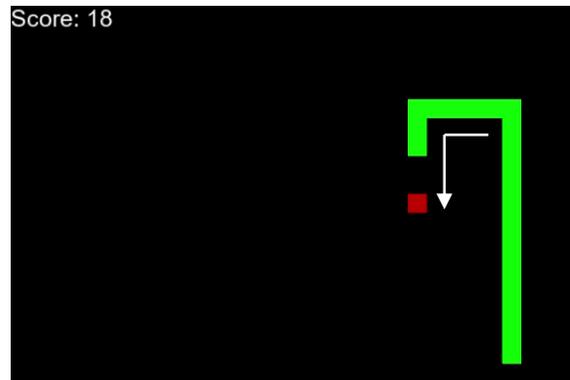
# スネークの様子



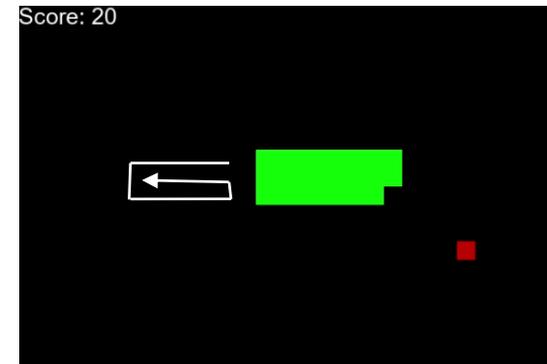
- 初めてスネークは環境を調べる。



- 環境を慣れて、リンゴを探す。



- リンゴの方向に向いて、スネーク頭の方角に変えていく。



- この場合、スネークは自分の体を食べたら、ゲーム終了する。

# 学習中のスネークエージェント

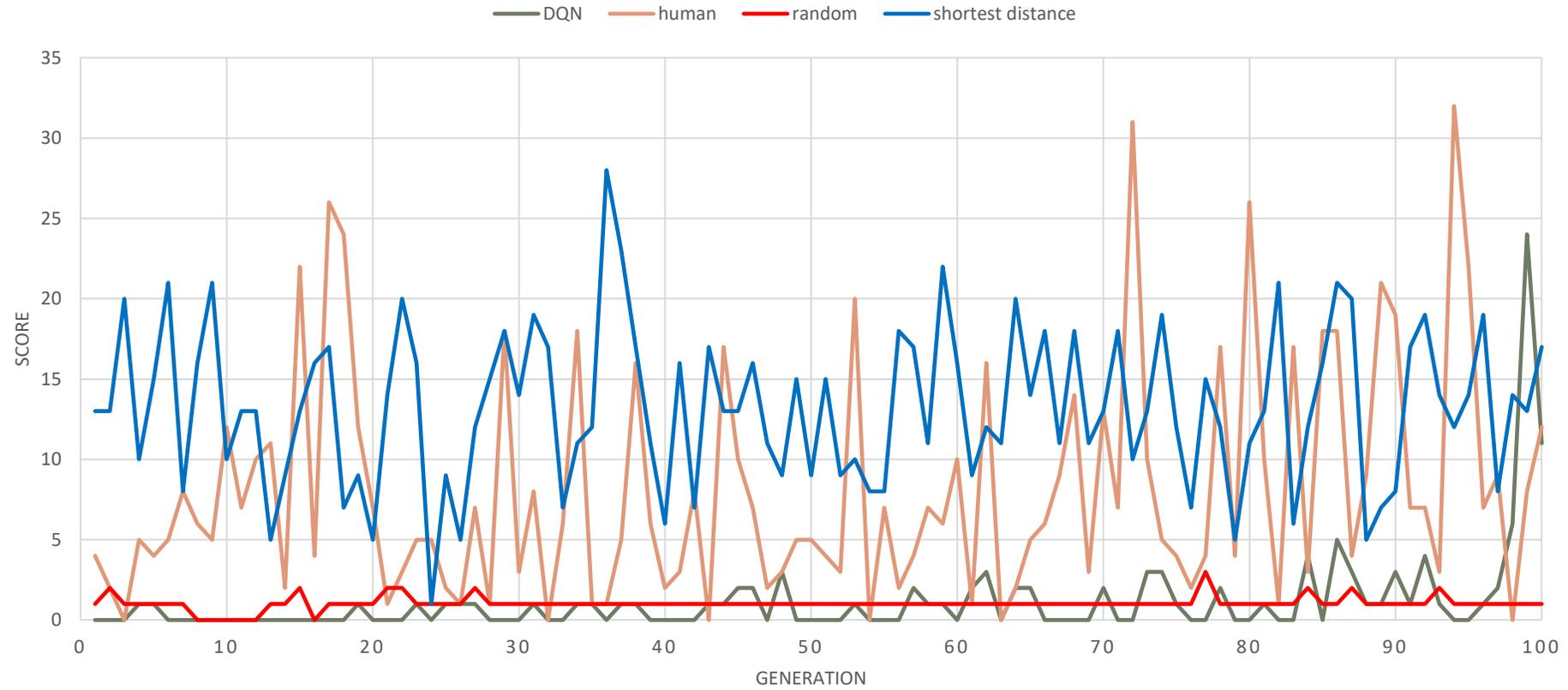
---



# 観察結果

## 100世代の比較

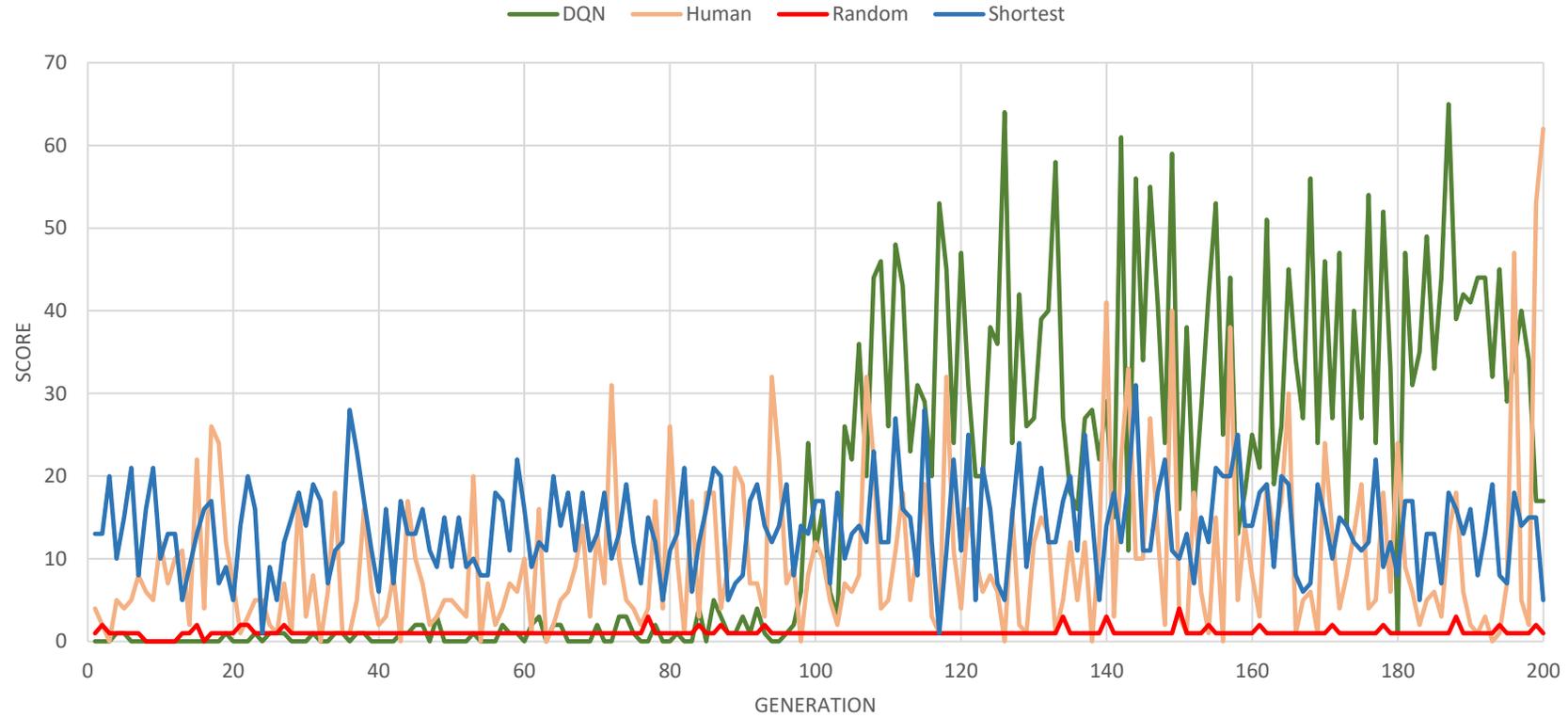
### COMPARISON OF 100 GENERATION



# 観察結果

## 200世代の比較

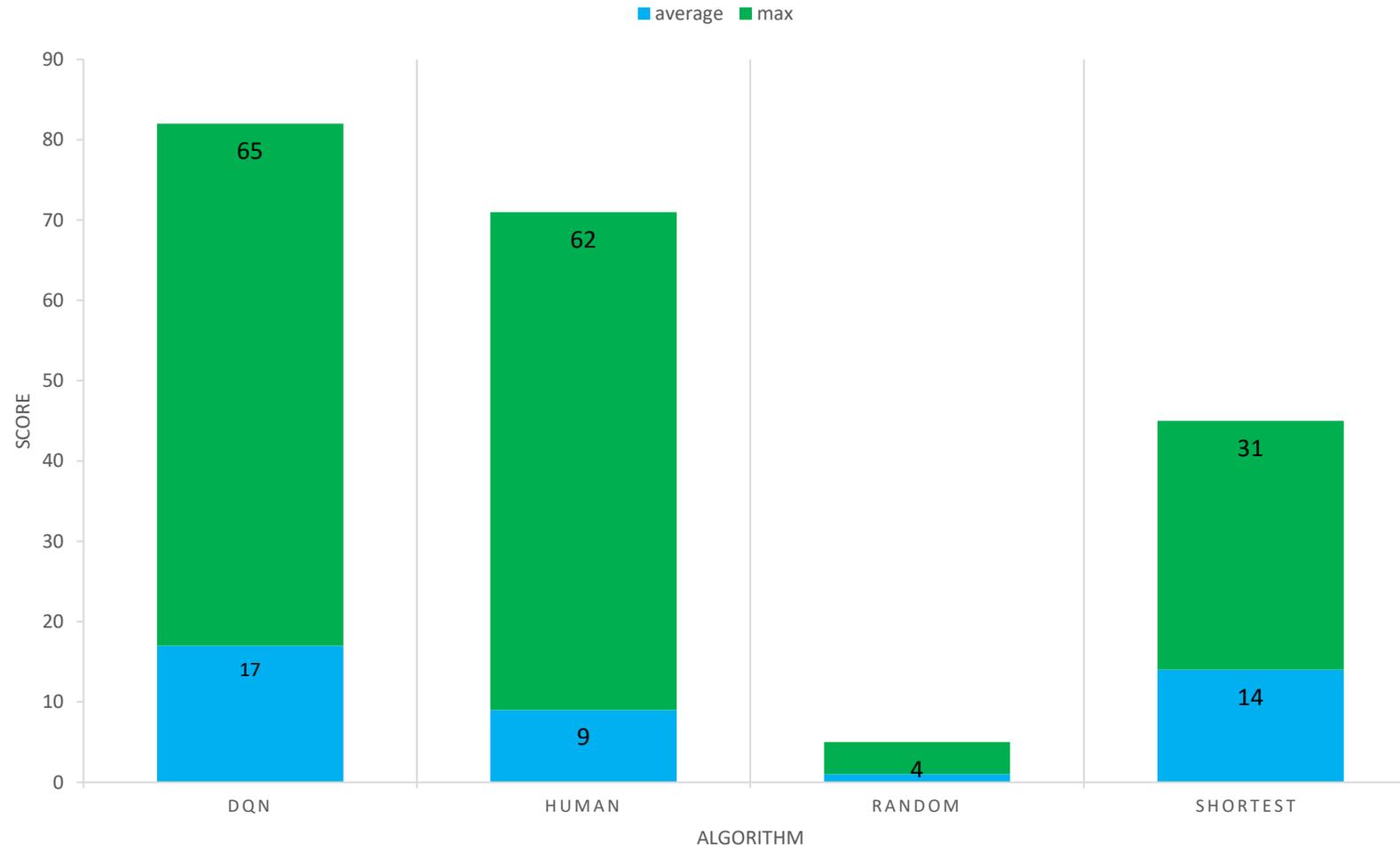
### COMPARISON OF 200 GENERATION



# 観察結果

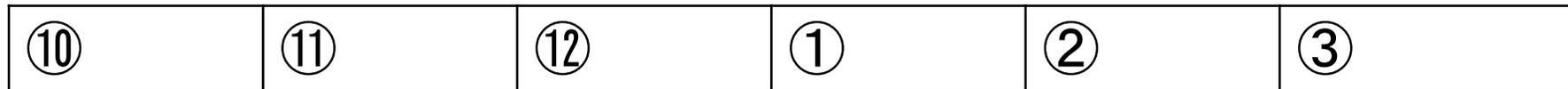
## 探索アルゴリズムのスコア

### SCORE OF SNAKE GAME ALGORITHM



# 研究計画

---



# 参考文献

[1] 遺伝的アルゴリズム, Wikipedia

<https://ja.wikipedia.org/wiki/%E9%81%BA%E4%BC%9D%E7%9A%84%E3%82%A2%E3%83%AB%E3%82%B4%E3%83%AA%E3%82%BA%E3%83%A0>

[2] Mastering the game of Go with deep neural networks and tree search (深層ニューラルネットワークと木探索により囲碁を究める)

<https://storage.googleapis.com/deepmind-media/alphago/AlphaGoNaturePaper.pdf>

[3] Playing atari with deep reinforcement learning (深い強化学習でアタリをプレイする)

V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller.  
arXiv preprint arXiv:1312.5602, 2013

[4] Proximal policy optimization algorithms (近接ポリシーの最適化アルゴリズム)

J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. arXiv preprint arXiv:1707.06347, 2017.