

一般化ジャンケンと 繰り返しジャンケンAI

研究室勉強会

2021年12月23日

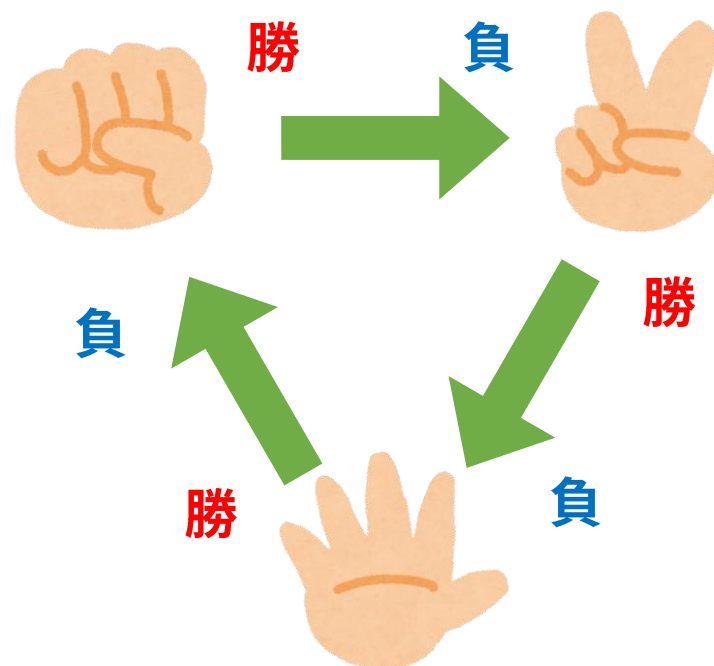
工学院大学

はじめに

ジャンケン

じゃんけん（漢字表記：石拳、両拳、雀拳）は、3種類の指の出し方（グー・パー・チョキ）で**三すくみ**の関係を構成し、その相性により勝敗を決める遊戯である。（Wikipediaより）<https://w.wiki/3Nw8>

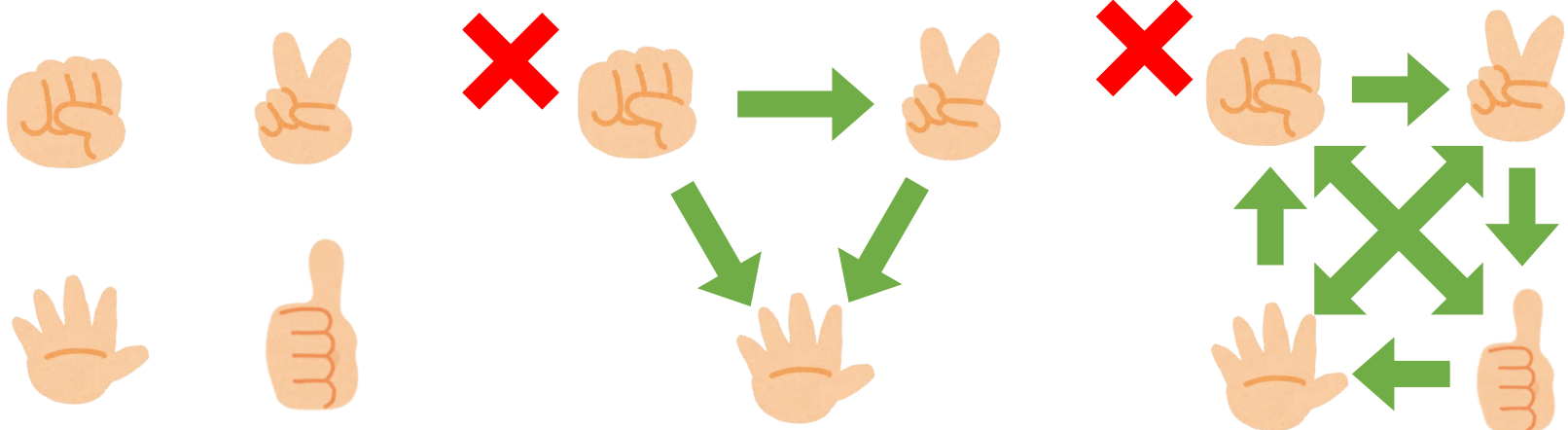
以降は2人でジャンケンする場合を考える



手数が4以上のジャンケンは作れるでしょうか？

ルール

1. 4つの異なる手から構成される
2. ある手の上位 / 下位互換となる手(無駄な手)が存在しない
3. 異なる手が出た場合は必ず一方が勝ち、もう一方が負けになる（＝あいこになるのは同じ手が出た場合のみ）



目的

- 4手以上のジャンケンができるかを考える.
- 7手ジャンケンの深層強化学習エージェントがどのような振る舞いをするか確認する.

目次

- 一般化ジャンケン
 - 手数が4以上のジャンケンは作れるか
 - ジャンケンのルールの面白さとは
- 繰り返しジャンケンAI
 - 7手ジャンケンを実装
 - Random vs α Random vs DQN (深層強化学習)

<https://book.mynavi.jp/ec/products/detail/id=90706>

<https://www.morikita.co.jp/books/mid/001771>

<https://www.saiensu.co.jp/search/?isbn=4910054691214&y=2021>

2021/12/23

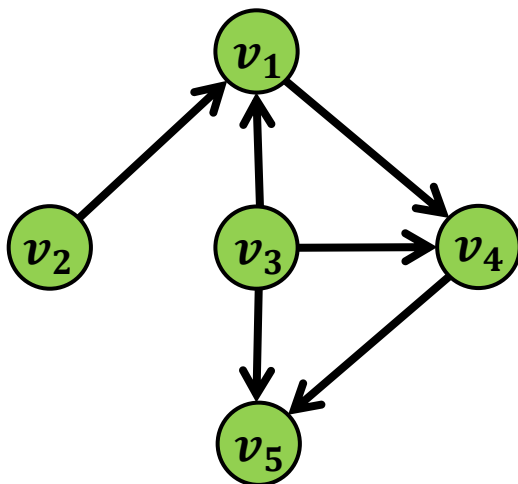
工学院大学



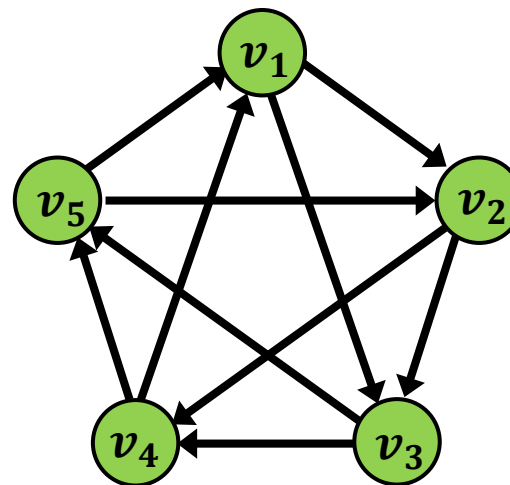
有向グラフ

■ $G = (V, E)$

- 有向グラフ（左側）に関して
 - $V(G) = \{v_1, v_2, v_3, v_4, v_5\}$
 - $E(G) = \{(v_1, v_4), (v_2, v_1), (v_3, v_1), (v_3, v_4), (v_3, v_5), (v_4, v_5)\}$



完全グラフ



手数 $n = 1, 2, 3$ の場合

- 無駄な手のないジャンケンは存在するか？

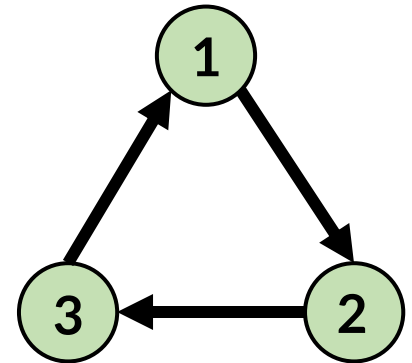
$n = 1$: 存在する



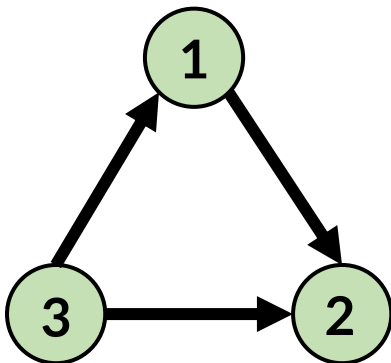
$n = 2$: 存在しない



$n = 3$: 存在する



$n = 3$ で無駄な手の存在するジャンケン



- 手1と手3はどちらも手2に勝つ
- しかし手3は手1に勝つので、手3を出すくらいならば手1を出した方が得である。
- 手1は無駄な手であり、手3は手1を「優越する」という。

手数4の無駄な手のないジャンケンは存在しない

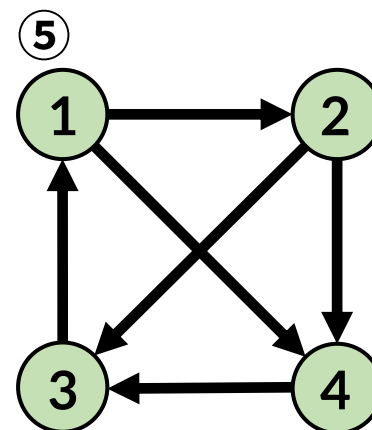
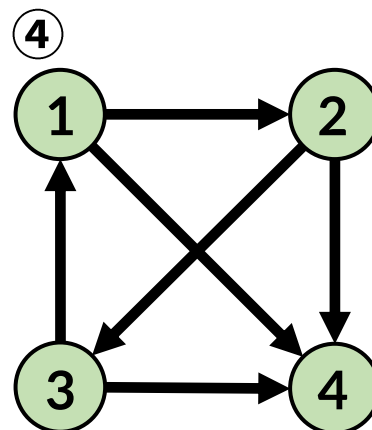
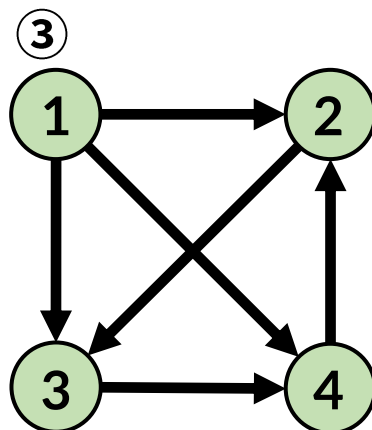
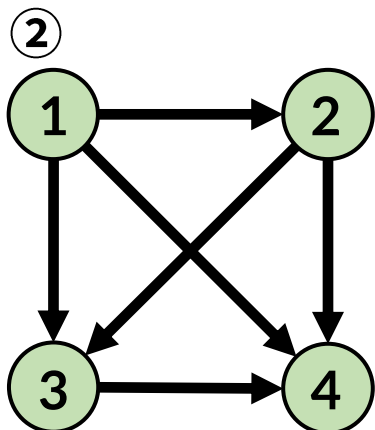
■ 勝敗の総和は6勝6敗 → 各手max=3で6を配分

勝つ手数

	手1	手2	手3	手4
①	3	3	0	0
②	3	2	1	0
③	3	1	1	1
④	2	2	2	0
⑤	2	2	1	1

- ①のパターンは作れない
- ②③④は全勝 or 全敗の手がある
- ⑤に関して

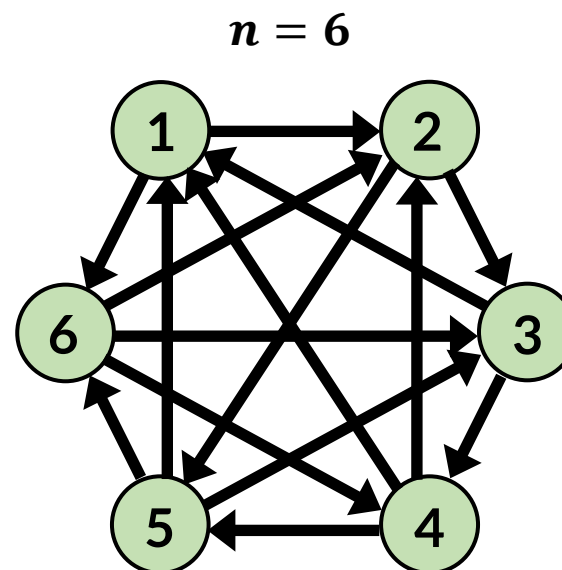
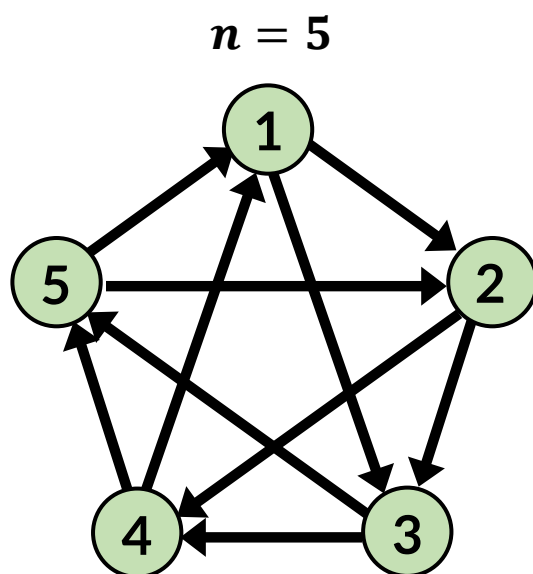
- 手2と手4はどちらも手3に勝ち、手1に負ける
- しかし手2は手4に勝つので、手4を出すくらいならば手2を出した方が得である。
- 手4は無駄な手であり、手2は手4を優越する。



手数 $n = 5, 6$ の場合は存在する

- 無駄な手のないジャンケンは存在するか？

ex)

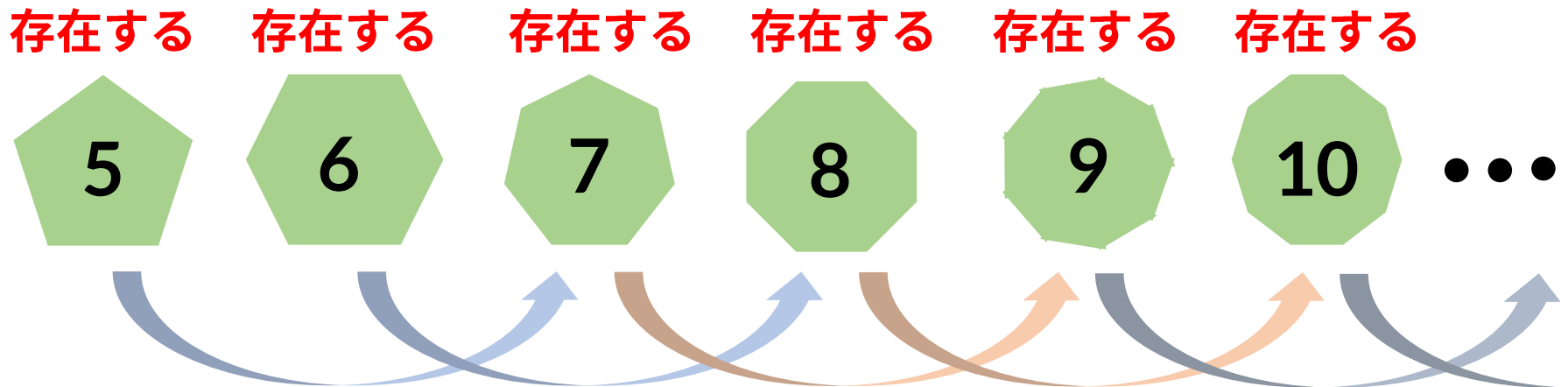


$n = 5$ の例：無駄な手がないようにそれぞれ2勝2敗

$n = 6$ の例： \forall 2勝3敗 / 3勝2敗する手 defeat \exists 3勝2敗 / 2勝3敗する手

手数 $n \geq 7$ の場合

- 無駄な手のないジャンケンは**存在する**！
- $n \leq 6$ については前ページまでで示している($n = 2, 4$ の場合は無駄な手のないジャンケンは存在しない)
- 補題：手数が n のジャンケンにおいて、無駄な手のないジャンケンがないならば、手数が $n + 2$ のジャンケンにも同様のことが言える。
- $n = 5, 6$ で無駄な手のないジャンケンができることを示しているので、7以上でも無駄な手のないジャンケンはできる。



無駄な手のないジャンケンは**存在する**！

■ 補題と証明 [伊藤, 2010, pp.135-140]

補題1

トーナメント $G = (V, E)$ がジャンケンとして無駄な手を持たない必要十分条件は、任意の節点 $x \in V$ から任意の節点 $y \in V$ へ長さ2の有向路が存在する。

補題2 (補題1を使って証明できる)

節点数 n のトーナメント $T_n = (V = \{1, 2, \dots, n\}, E)$ に対して、節点数 $n + 2$ のトーナメント $T_{n+2} = (V', E')$ を次のように構築する。

$$V' = V \cup \{n + 1, n + 2\}$$

$$E' = E \cup \{(n + 1, n + 2)\} \cup \{(i, n + 1), (n + 2, i) | i \in V\}$$

このとき、 T_n に無駄な手がないならば、 T_{n+2} にも無駄な手はない。

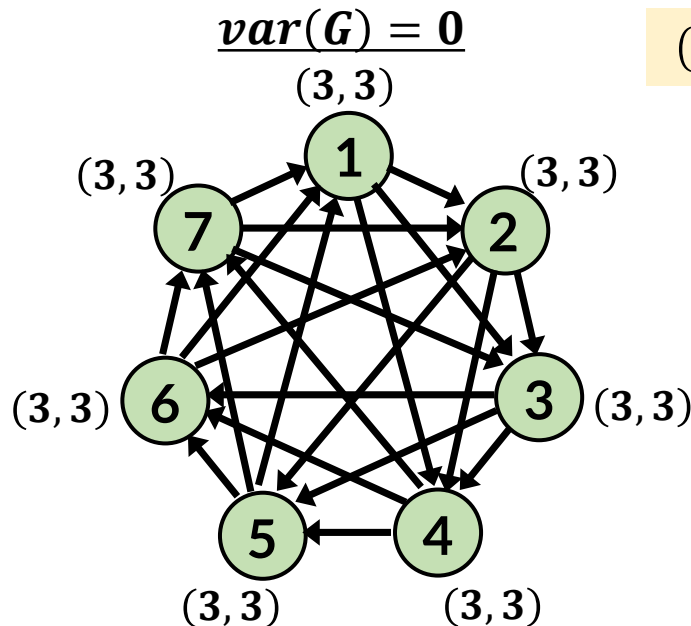
- $n \leq 6$ については前ページまでで示している。 ($n = 2, 4$ の場合は無駄な手のないジャンケンは存在しない)
- $n \geq 7$ については $n = 5, 6$ に補題2を適用していくことで、無駄な手のないジャンケンが存在することを示せる。

ジャンケンの不規則性: 面白いジャンケンとは

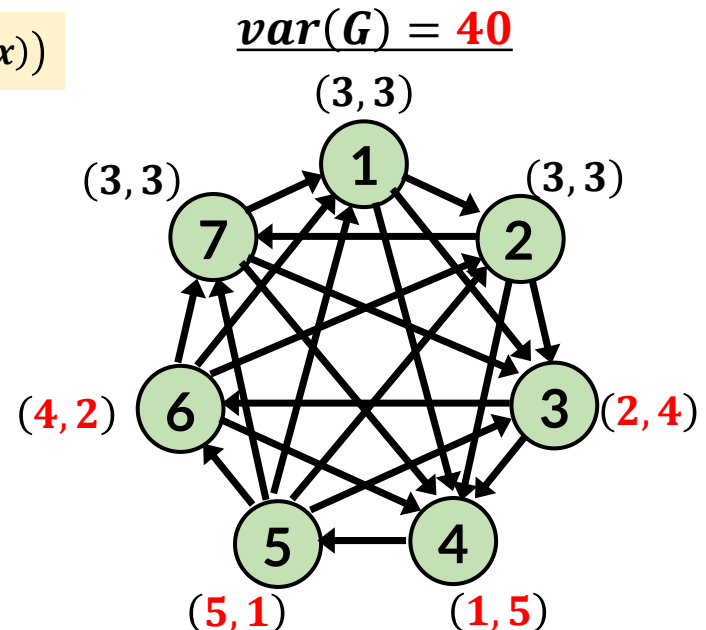
- 勝ち負けのばらつきを表す指標

$$\text{var}(G) = \sum_{x \in V} (d^+(x) - d^-(x))^2$$

$d^+(x) = |\{y \in V | (x, y) \in E\}|$ (出次数), $d^-(x) = |\{y \in V | (y, x) \in E\}|$ (入次数)

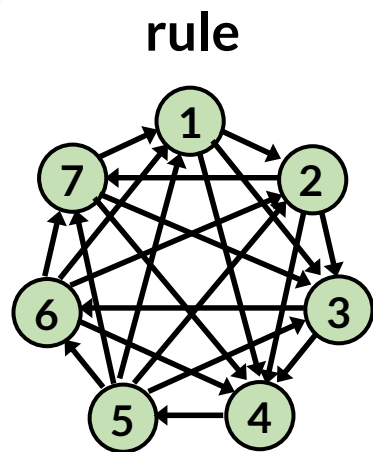


$(d^+(x), d^-(x))$



難しいことはコンピュータにやらせよう

class RockPaperScissors



Agent2の手

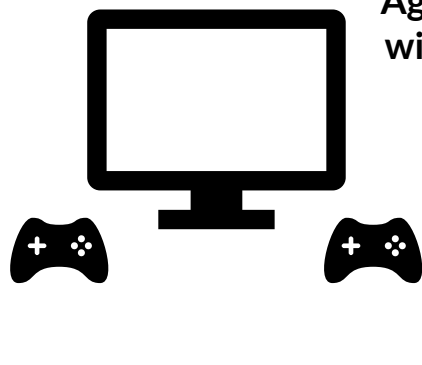
	1	2	3	4	5	6	7
1	0	1	1	1	-1	-1	-1
2	-1	0	1	1	-1	-1	1
3	-1	-1	0	1	-1	1	-1
4	-1	-1	-1	0	1	-1	-1
5	1	1	1	-1	0	1	1
6	1	1	-1	1	-1	0	1
7	1	-1	1	1	-1	-1	0

Agent1の手

Agent1の勝敗テーブル
win: 1, draw: 0, lose: -1



Agent1



Agent2

```
1 for i in range(100):
2     game.play(False)
```

RandomAgent1 win: 1-2
RandomAgent1 win: 4-1
RandomAgent2 win: 6-1
draw: 2-2
RandomAgent1 win: 4-0
draw: 5-5
RandomAgent2 win: 4-3
RandomAgent1 win: 5-6
RandomAgent2 win: 6-4
draw: 2-2
RandomAgent1 win: 2-3
RandomAgent2 win: 5-4
RandomAgent2 win: 6-4
RandomAgent2 win: 6-4
draw: 0-0
RandomAgent2 win: 6-5
RandomAgent2 win: 6-1
RandomAgent2 win: 0-4
RandomAgent2 win: 2-1

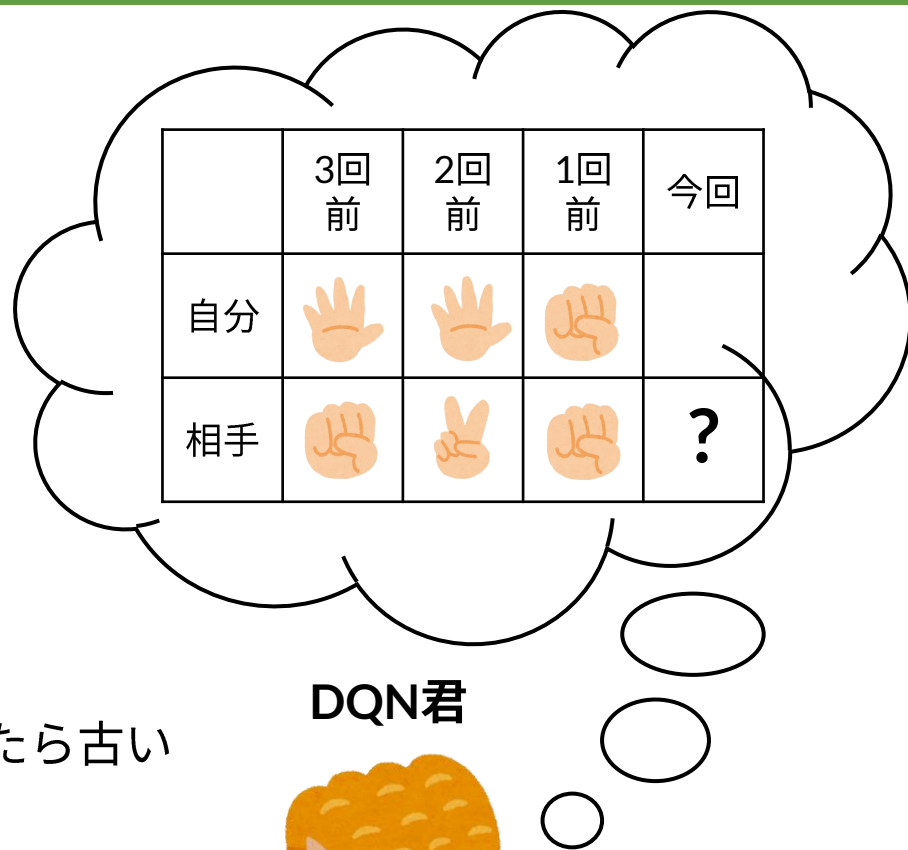
ジャンケンAgent

状態 s ：過去 n 回の対戦の履歴

行動 a ：次に何を出すか

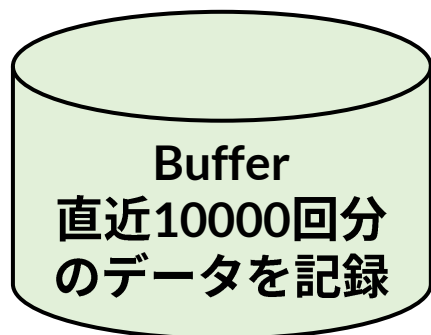
報酬 r ：勝ち1，引き分け0，負け-1

次の状態 s' ： a を出した後の状態



Replay Buffer：

対戦したらデータを記録し，一杯になったら古いデータから新しいデータに上書き



DQN君

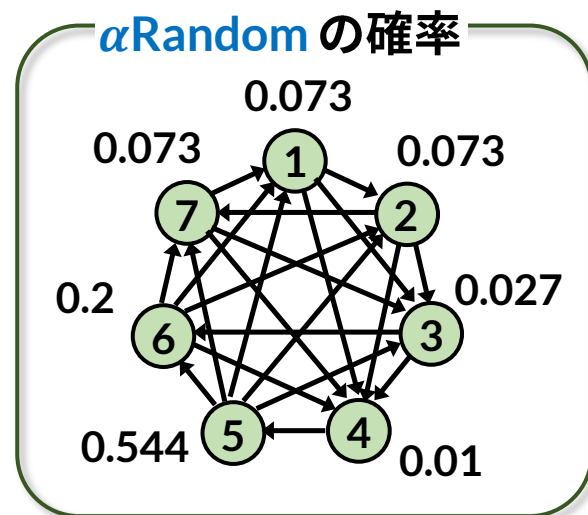
Agentは3種類

■ Random Agent

- 状態によらずランダム(等確率)に手を出す

■ α Random Agent

- 状態によらずランダムに手を出すが、手を出す確率を制御
- ジャンケンのグラフの出次数が大きいものに確率を大きく振る



■ DQN

- 後述

10万回対戦させた場合

	前者の勝ち	後者の勝ち	あいこ
Random vs Random	42769	42905	14326
Random vs α Random	25171	60825	14004
α Random vs α Random	32331	32398	35271

Q学習

■ 学習

行動価値 $Q(s, a)$ を

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha \left(r + \gamma \max_{a' \in \mathcal{A}(s')} Q(s', a') \right)$$

によって更新する．（ α ：学習率， γ ：割引率）

■ Q値からの行動選択： ϵ -greedy方策，ボルツマン方策など

- ϵ -greedy方策

$$\pi(a|s) = \begin{cases} 1 - \epsilon + \frac{\epsilon}{|\mathcal{A}(s)|} & (a = \operatorname{argmax}_a Q(s, a)) \\ \frac{\epsilon}{|\mathcal{A}(s)|} & (\text{otherwise}) \end{cases}$$

Deep Q-Network (DQN)

■ Q関数をNNで近似

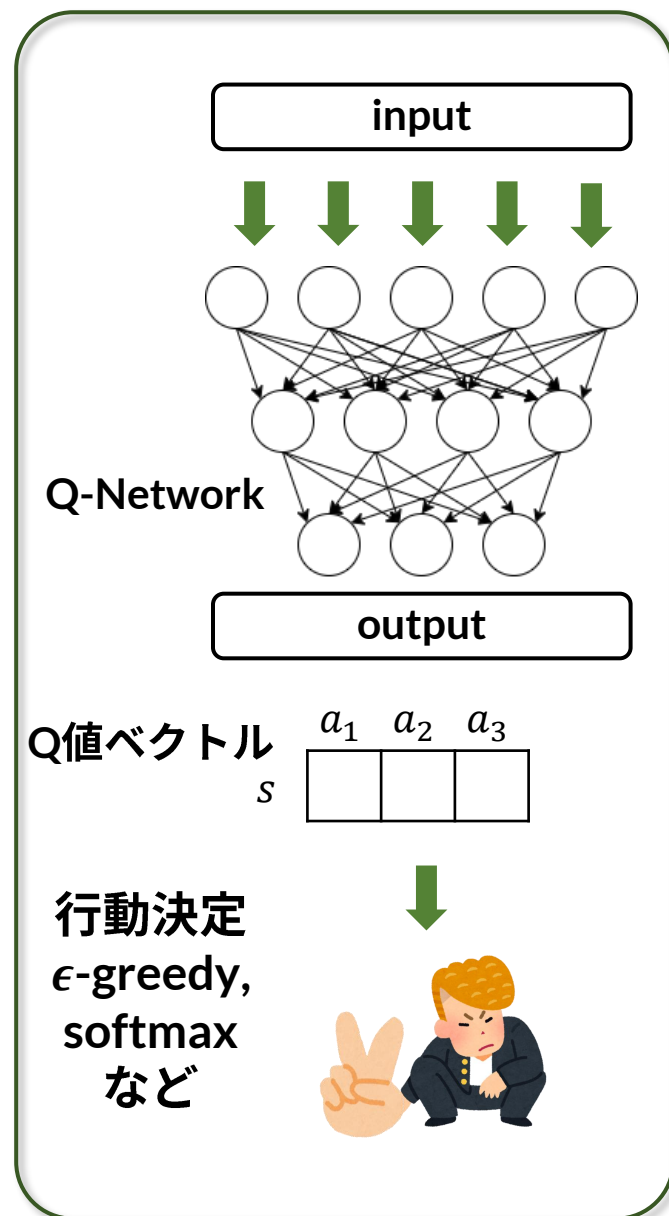
- 深層学習モデルを使った強化学習＝深層強化学習(deep reinforcement learning)
- TD誤差により学習する

$$\left(\underbrace{r + \gamma \max_{a'} Q(s', a'; \theta)}_{\text{教師}} - \underbrace{Q(s, a; \theta)}_{\text{Q-Network の出力}} \right)^2$$

■ Agentの状態＝DQNへの入力

(履歴数3, 3手ジャンケンの場合)

3回前						2回前						1回前					
自分		相手				自分		相手				自分		相手			
0	0	1	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0



実装

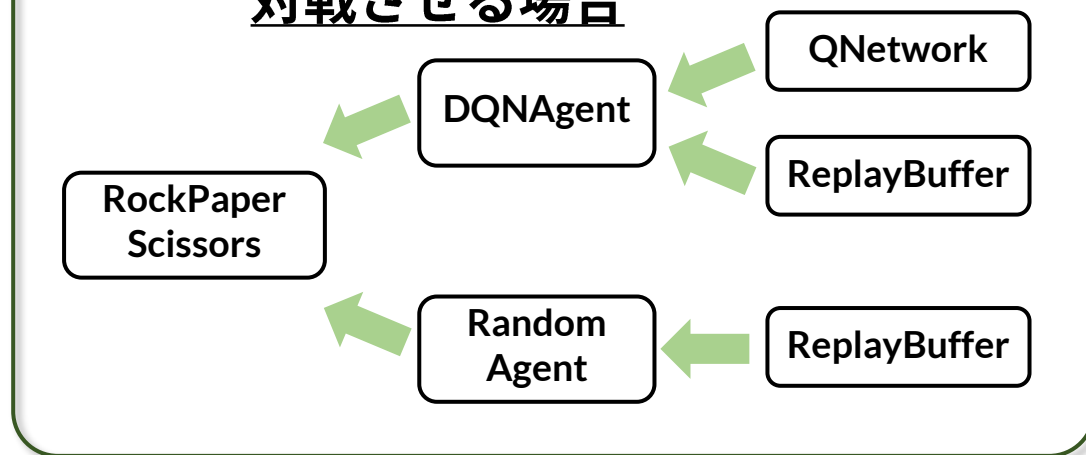
- ライブラリ

PyTorch version 1.10.0

- プログラム

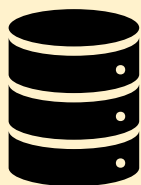
https://colab.research.google.com/drive/188NVk4NFIVtWK0BhHz7M_wxJOYIPfUeV?usp=sharing

DQNAgentとRandomAgentを 対戦させる場合



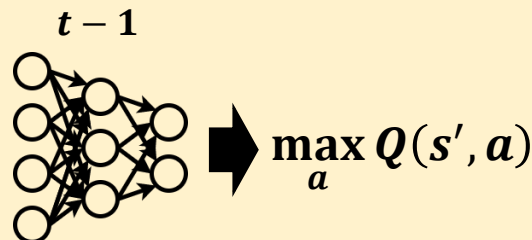
実装時の工夫 [小川, 2018, pp.125-126]

Experience Replay



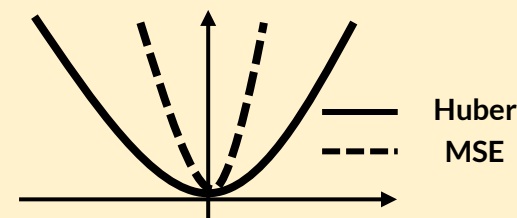
ある程度過去の記録を貯めておき、Q-Netの学習時にここからサンプリング

Fixed Target Q-Network



教師信号の $\max_a Q(s', a)$ を計算する際に、少し古いQ関数を使用し学習を安定させる

Huber関数

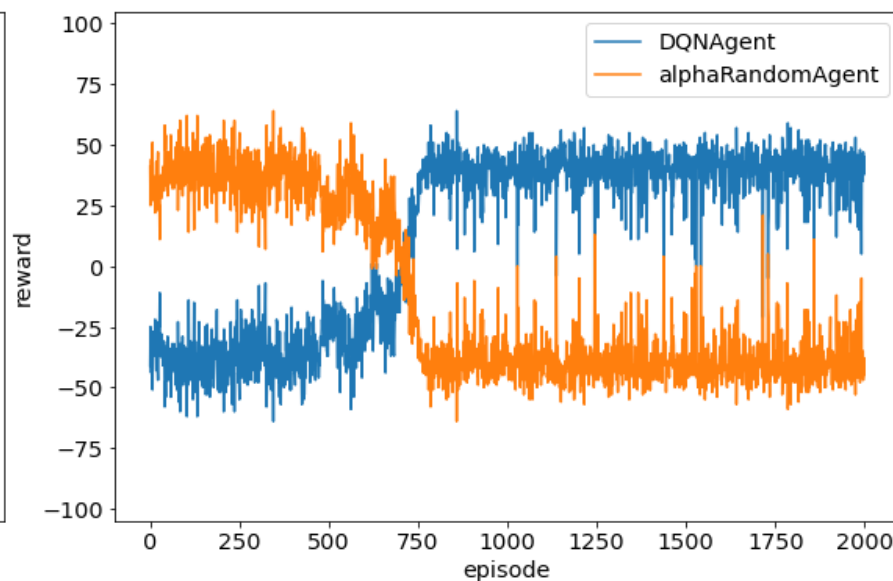
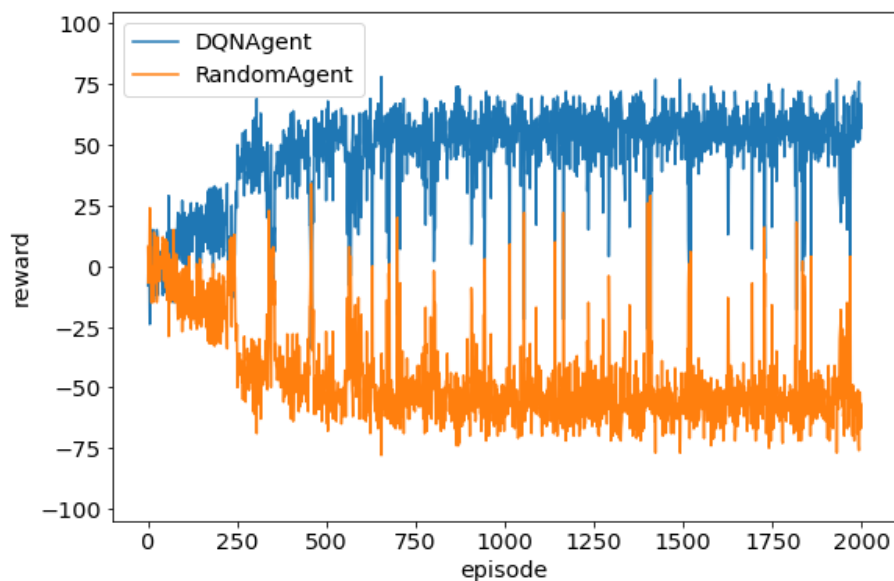


二乗誤差関数よりも誤差関数の出力値を抑え、学習を安定させることができる

<https://pytorch.org/docs/stable/generated/torch.nn.HuberLoss.html>

学習中の報酬の変化

- 1episodeあたり100回繰り返しジャンケンする。
 - 1episodeの総報酬 $\in [-100, 100]$
 - 報酬0以上であれば相手に勝ち越せている。



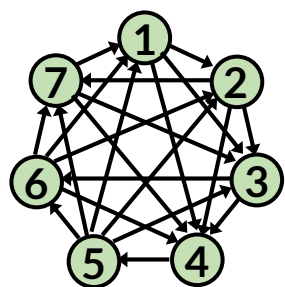
Random vs DQN

■ 学習後に10万回繰り返しジャンケン

- Randomは各手を等確率（1/7）に出すため、手5を出し続ければ7割以上の確率で勝てる

■ 結果

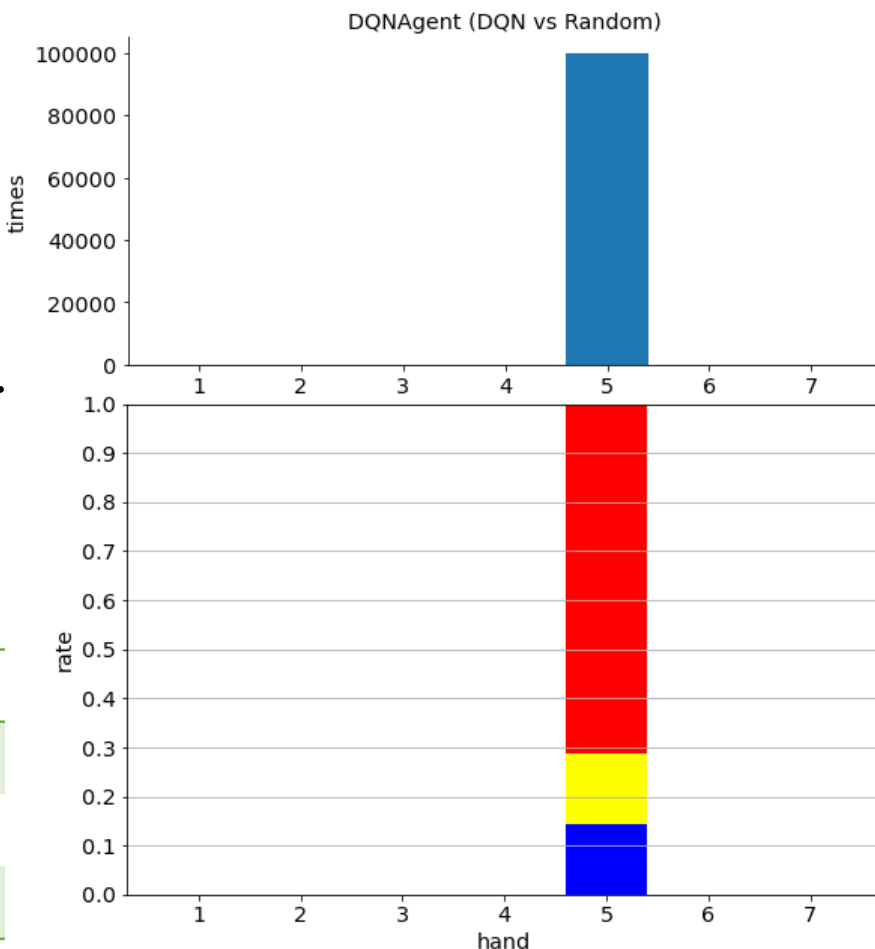
- DQNが手5を出し続けて7割以上勝った。



勝者	回数
DQN	71326
引き分け	14314
Random	14360

上：出した手の回数
下：各手の勝敗の割合

赤：DQNの勝ち
黄：引き分け
青： α Randomの勝ち



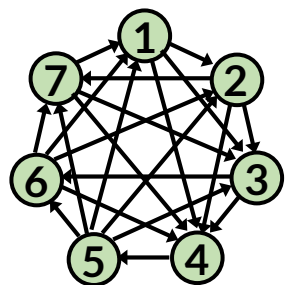
α Random vs DQN

■ 学習後に10万回繰り返しジャンケン

- α Randomは5割以上の確率で手5を出すので、手4を出し続けていれば半分以上は勝てる。
- また α Randomはほぼ4を出さないで、手5を出せばほとんど負けない。

■ 結果

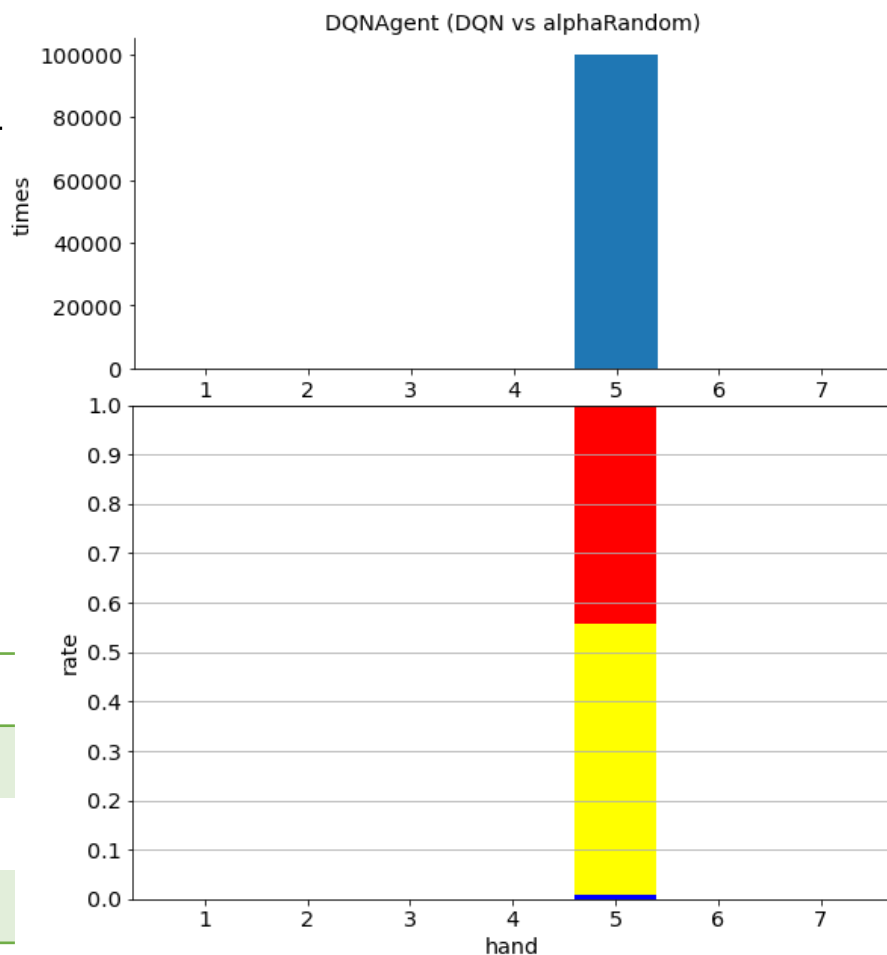
- 手5のみを出し続けて4割以上勝った



勝者	回数
DQN	44386
引き分け	54603
α Random	1011

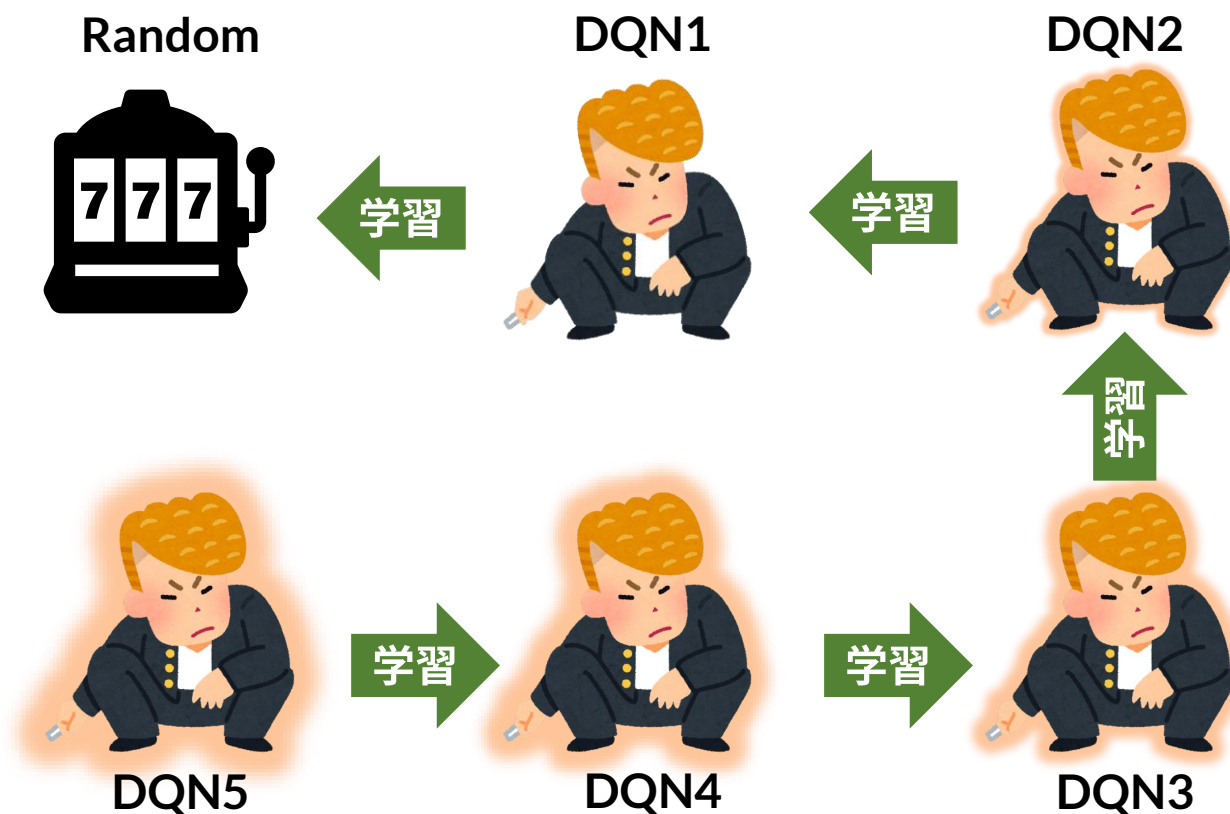
上：出した手の回数
下：各手の勝敗の割合

赤：DQNの勝ち
黄：引き分け
青： α Randomの勝ち

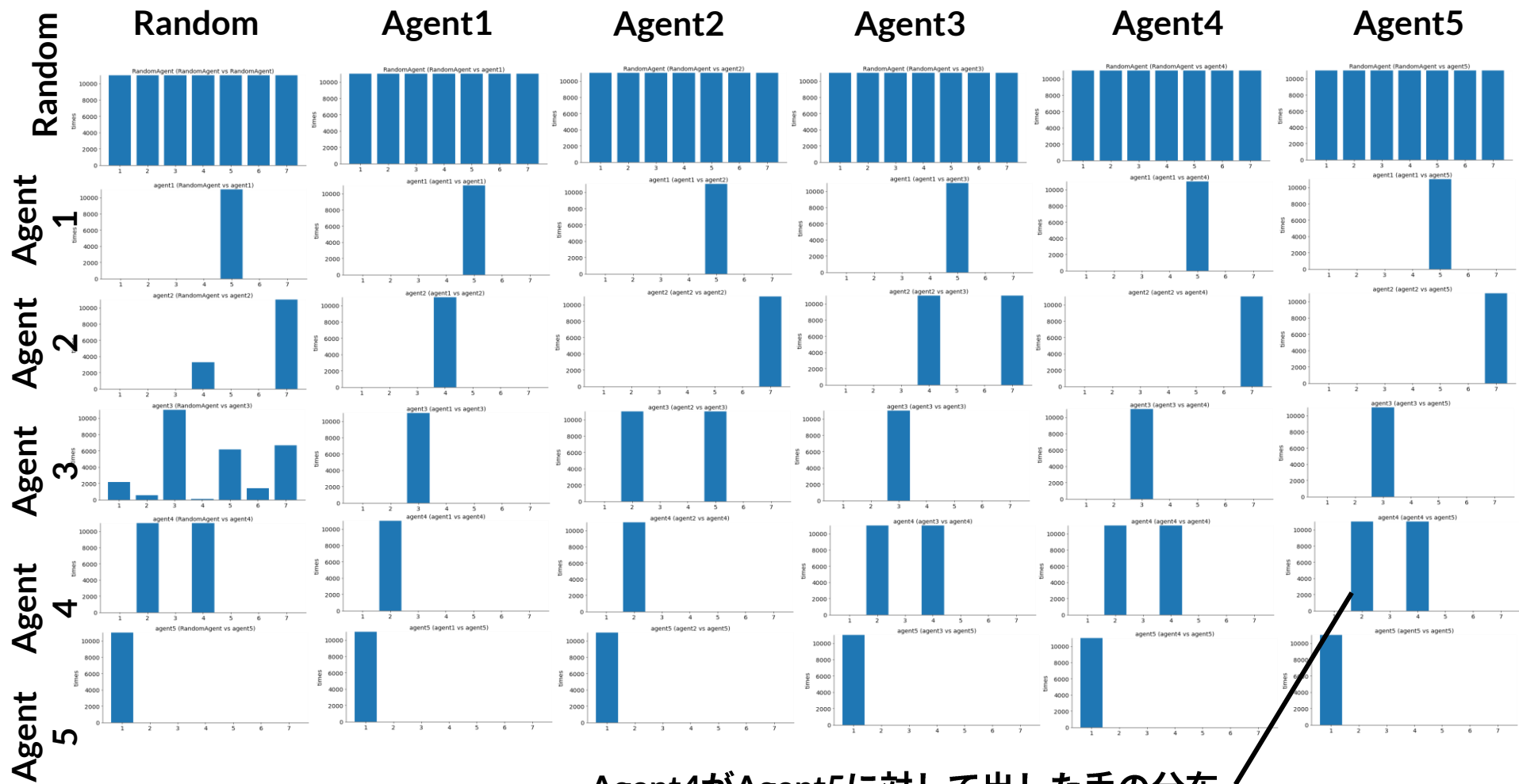


Random相手では直近の対戦履歴はあまり関係なさそう...

- 学習させたDQNAgentで，別のDQNAgentを学習させる



すべての手を使ったのはRandomとAgent3のみ



Agent4がAgent5に対して出した手の分布

それぞれ10万回ずつ勝負

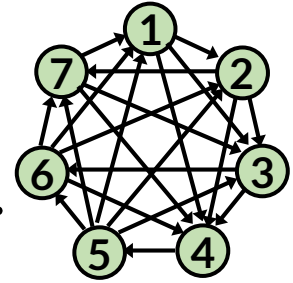
勝率（上段）と負率（下段）

	Random	Agent1	Agent2	Agent3	Agent4	Agent5
Random	0.4306	0.1421	0.4362	0.5246	0.5528	0.4275
Agent1	<u>0.7149</u>	0.0000	0.0000	1.0000	1.0000	1.0000
Agent2	0.4220	<u>1.0000</u>	0.0000	0.0000	0.0000	1.0000
Agent3	0.3332	0.0000	<u>1.0000</u>	0.0000	0.4000	0.0000
Agent4	0.3045	0.0000	1.0000	<u>0.6000</u>	0.0000	0.0000
Agent5	0.4296	0.0000	0.0000	1.0000	<u>1.0000</u>	0.0000

	Random	Agent1	Agent2	Agent3	Agent4	Agent5
Random	0.4272	0.7149	0.4220	0.3332	0.3045	0.4296
Agent1	0.1421	0.0000	1.0000	0.0000	0.0000	0.0000
Agent2	0.4362	0.0000	0.0000	1.0000	1.0000	0.0000
Agent3	0.5246	1.0000	0.0000	0.0000	0.6000	1.0000
Agent4	0.5528	1.0000	0.0000	0.4000	0.0000	1.0000
Agent5	0.4275	1.0000	1.0000	0.0000	0.0000	0.0000

※下線は学習したAgent

まとめ



- 4手以上の無駄の手のないジャンケンが存在するかを考えた.
 - $n \neq 2, 4$ であれば無駄な手のないジャンケンは存在する.
- ジャンケンの不規則性について
 - 同じ手数でも異なるパターンのジャンケンができる場合もある
- DQNエージェントの振る舞い
 - 7手すべては使わず，ある程度出す手を絞った戦略をとる

感想

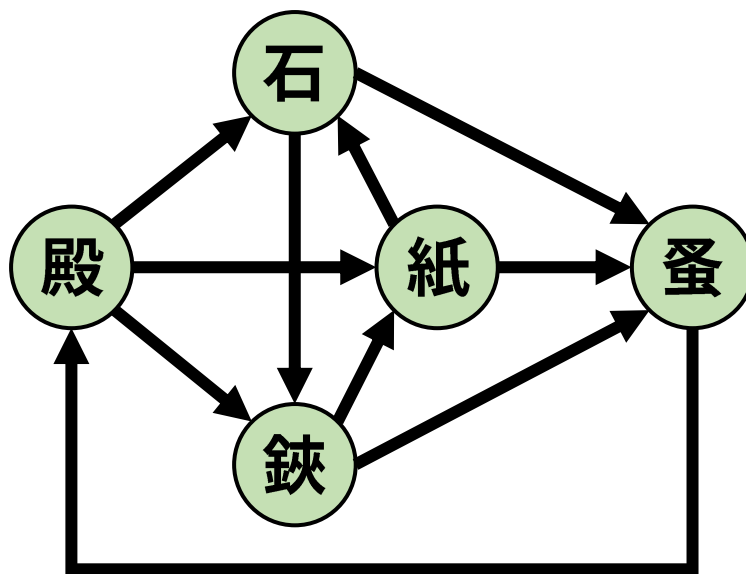
今回は時間がなくて出来なかったが，考慮する履歴数が大きくor小さくなるとどのような戦略になるのか気になった．

参考文献

- 田地宏一 (2013) 手数の多いじゃんけん. 日本数学コンクール: <http://www.aip.nagoya-u.ac.jp/public/math-con/old/2013/index.html>
- 伊藤大雄 (2010) パズル・ゲームで楽しむ数学. 森北出版. 6章.
<https://www.morikita.co.jp/books/mid/001771>
- 安藤清, 土屋守正, 松井泰子 (2013). 例題で学ぶグラフ理論. 森北出版.
<https://www.morikita.co.jp/books/mid/005281>
- 伊藤大雄, 永持仁 (1995) ジャンケンのトーナメント表現と意味のある拡張
<https://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/pdf/0906-3.pdf>
- 小川雄太郎 (2018) つくりながら学ぶ！ 深層強化学習 PyTorchによる実践プログラミング. マイナビ出版. <https://book.mynavi.jp/ec/products/detail/id=90706>
- 浅田稔 et al. (2016) これからの強化学習. 森北出版.
<https://www.morikita.co.jp/books/mid/088031>
- 解良貴 (2019) Q-Network を用いた繰り返しじゃんけん AI における割引率の特性解析. 工学院大学大学院修士学位論文.

蚤殿（のみとの）ジャンケン

- 手数5で構成できる不規則性（後述）最大のじゃんけん
 - 石，紙，鋏の関係は普通のじゃんけんと同じ
 - 殿は石，紙，鋏に勝つが，蚤にだけ負ける
 - 蚤は石，紙，鋏に負けるが，殿にだけ勝つ



「無駄な手」の定義 [伊藤, 2010, p.135]

ジャンケンを表現したトーナメント $G = (V, E)$ とある手 (ノード) $x \in V$ に対して, ある $y \in V$ が存在し,

$$(y, x) \in E \text{ かつ } (x, z) \in E \Rightarrow (y, z) \in E (\forall z \in V)$$

であるとき, x を **無駄な手** といい, y は x に **優越する** という.

