

2021/01/27

ターン制コマンドバトルゲーム における強化学習による ゲームAIの学習

made by **mrK**

背景：ゲームのAI戦闘

ゲームAIの活用例

- ・ 戦闘の自動化

AIが行動を選択

- ・ 戦略に沿った行動

プレイヤーの命令に従う



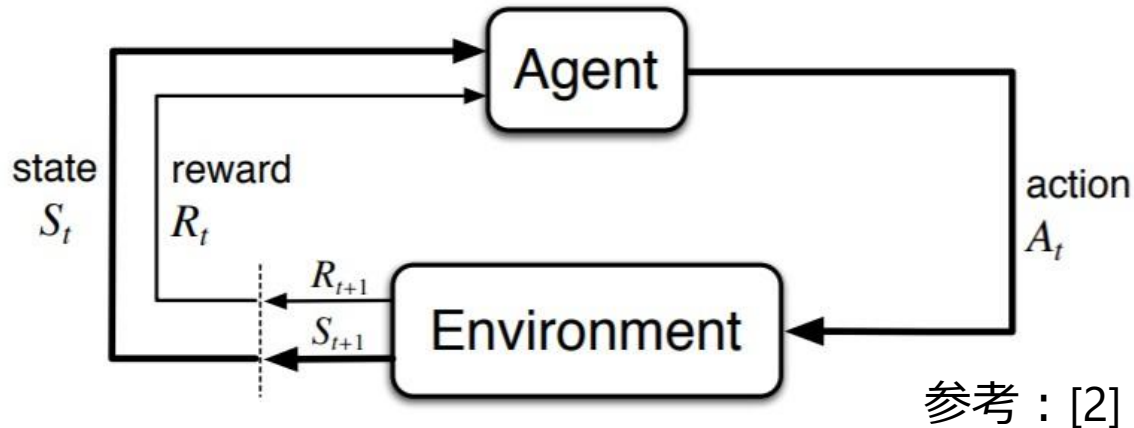
参考：[1]

ゲームAIの作成 ▶▶▶ 強化学習が必要

[1] SQUARE ENIX (2016), “ドラゴンクエストIV 導かれし者たち”,
<http://www.dragonquest.jp/30th/history/4.html>

目的：学習効率の検証

強化学習によるAIの作成



Q学習 ▶▶▶ 効率的な学習方法を評価

- [2] Richard S. Sutton, et al.(2018), "Reinforcement Learning: An Introduction",
<http://incompleteideas.net/book/bookdraft2018jan1.pdf>

原理：Q学習



Q値の更新式

$$Q(s, a) \leftarrow Q(s, a) + \alpha \left[r + \gamma \max_{a'} Q(s', a') - Q(s, a) \right]$$

(α : 学習率 γ : 割引率)

実験仕様

・状態

プレイヤー・敵のHP割合を10分割 ▶ **100通り**

・行動

物理攻撃・回復・光攻撃 ▶ **3通り**

・報酬

プレイヤーが勝利 ▶ **+1**

プレイヤーが敗北 ▶ **-1**

・ステータス

	player	enemy
HP	138	240
MP	123	40
ATK	125	107
DEF	69	110
AGT	82	55

実験内容

共通

1000回のQ学習 ▶▶▶ 1000回の模擬バトル

- **学習回数と勝率の相関**

固定条件 : $\alpha = 0.1, \gamma = 0.9$

- **学習率と勝率の相関**

固定条件 : $\gamma = 0.9$ 可変条件 : $\alpha = 0.001, 0.01, 0.1, 0.5$

- **割引率と勝率の相関**

固定条件 : $\alpha = 0.1$ 可変条件 : $\gamma = 0.1, 0.5, 0.9, 0.99$

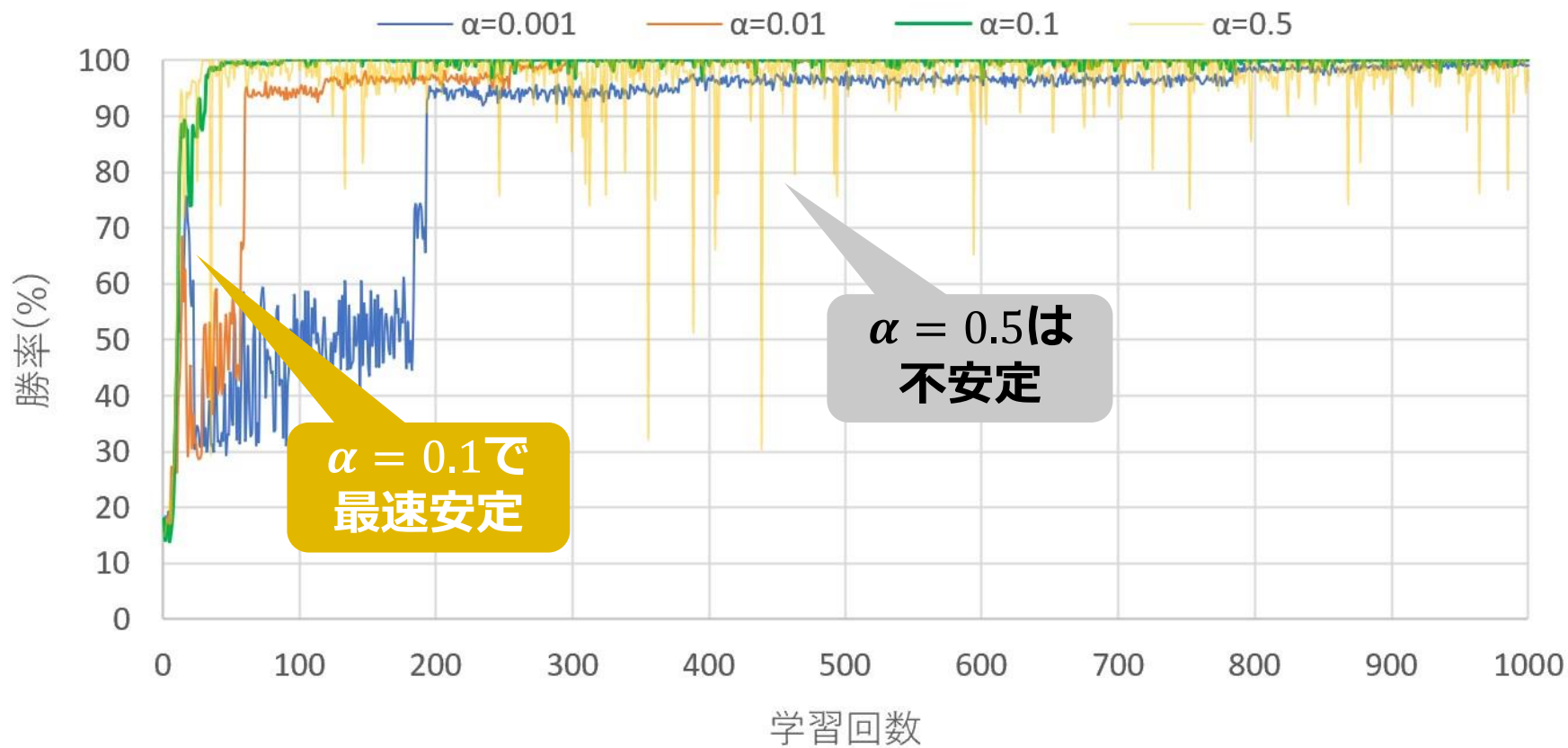
実験結果

学習回数と勝率の相関



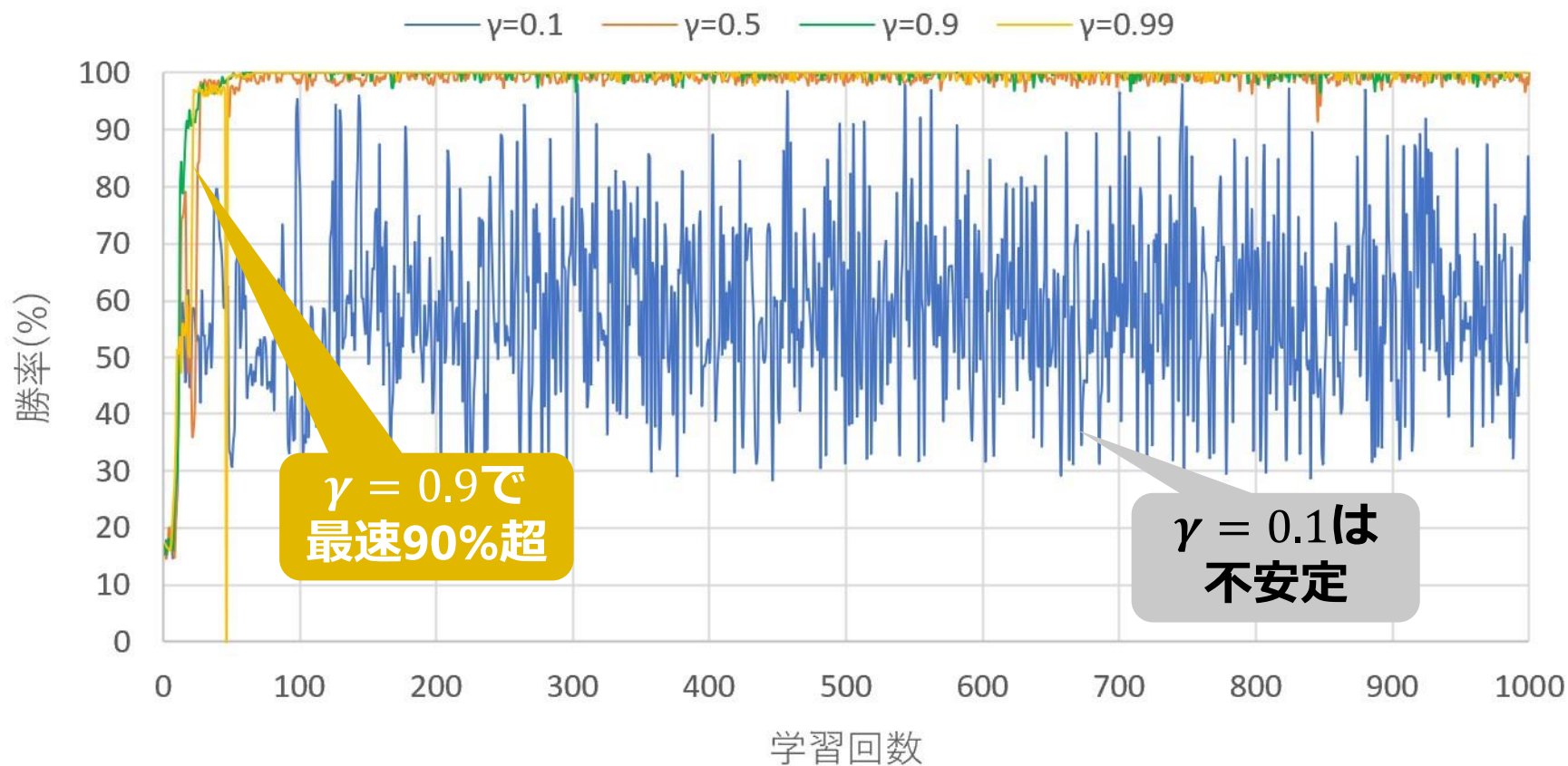
実験結果

学習率と勝率の相関



実験結果

割引率と勝率の相関



考察：Q値と最適行動

Q値が最大の行動をまとめた表

Q値最大の行動		enemy.hp(%)									
		0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
player.hp(%)	0-10	1	1	0	0	0	※	※	※	※	※
	10-20	2	1	1	2	0	1	※	※		
	20-30	1	1	1	1	1	0	※	※		
	30-40	1	1	1	1	1	1	1	1	※	※
	40-50	1	1	1	1	1	1	1	1	※	※
	50-60	1	1	1	1	1	1	1	1	※	※
	60-70	1	1	1	2	1	1	1	1	2	※
	70-80	2	2	1	2	2	1	1	1	2	※
	80-90	2	2	2	1	2	1	1	2	※	※
	90-100	2	2	2	2	2	2	2	2	2	2

回復優先

攻撃優先

0：物理攻撃 1：回復 2：光攻撃 ※：random(未学習)

考察：ゲーム調整

敵のステータス変化と勝率

Q学習後の勝率(%)		enemy.atk							
		107	117	127	137	147	157	167	177
enemy.hp	240	99	99.4	100	100	99.3	98.1	63.7	0
	250	100	100	100	100	99.8	98	17.1	0
	260	100	100	100	100	98.3	0.4	68.4	0
	270	100	100	100	100	100	97.2	0	0
	280	100	100	99.6	98.8	62.9	96.6	0	0
	290	100	100	100	100	100	96.3	0	0
	300	100	100	100	100	99.8	97.4	0	0
	310	100	100	100	100	99.8	91.3	0	0
	320	100	100	100	100	99.9	73.6	0	0
	330	100	100	100	100	100	97.7	0	0
	340	100	100	100	100	99.9	1.8	0	0
	350	100	100	100	100	99.9	33.3	0	0
	360	100	100	100	99.9	82.6	2.4	0	0
	370	100	100	100	100	98.9	0.2	0	0
	380	100	100	100	99.9	50.1	0	0	0
	390	0	100	100	100	99.5	0	0	0
	400	0	100	100	10.8	0	0	0	0
	410	0	0	0	0	0	0	0	0
	420	0	0	0	0	0	0	0	0

敵HP・ATKを変化
▶勝率が急激に減少

学習後の勝率が0%
▶ゲームとして不成立

ゲームバランスの
調整に応用

結論

AIの学習効率を向上させるには

- ・ **学習回数**の増加
- ・ 適切な**学習率**・**割引率**の設定
 $\alpha = 0.1, \gamma = 0.9$ によるQ学習 ▶ 最適

ゲームAIの応用分野

- ・ **ゲームバランス**の調整
敵のステータスと勝率を観測