

# 正誤情報

このたびは森北出版株式会社発行の書籍をお買い求めいただき、誠にありがとうございました。下記の書籍につきまして誤りのある箇所がございましたので、お詫びし訂正させていただきます。

2020年5月8日 森北出版株式会社 生産マネジメント部

## タイトル

# これからの強化学習

## 正誤対象

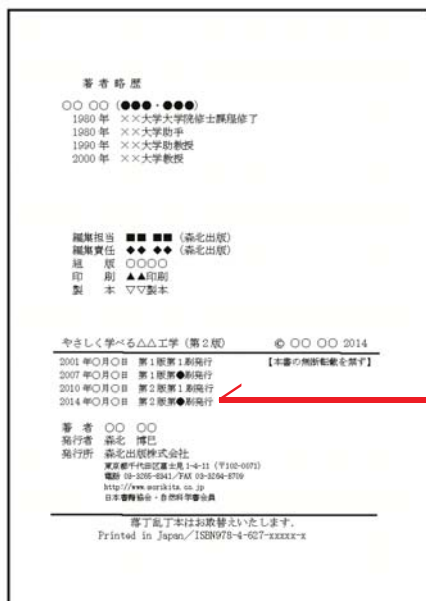
お手持ちの書籍の刷数をお調べのうえ、下の表をご覧ください。正誤表内の一番左に「対応刷数」という列がございます。該当する刷数の訂正情報をご参照下さい。

なお、刷数につきましては下記「刷数の調べ方」をご参照ください。

お持ちの本の刷数					
1,2	対応刷数	2	より	6	までをご参照ください
3	対応刷数	3	より	6	までをご参照ください
4	対応刷数	4	より	6	までをご参照ください
5,6	対応刷数	6	を		ご参照ください
それ以降		現在把握している訂正情報はございません			

## 刷数の調べ方

本の一番後ろのページ(広告等除く)に下図のようなページがございます。ご参照いただき、お持ちの本の刷数をお調べください。



日付の最も新しい行に記載された数字がお持ちの本の刷数となります

対応 刷数	頁	行数, 図・ 表・式番号	誤	正
4	6	1.1.2 5,8,9行目	… $p_k$ …	… $p_i$ …
4	6	1.1.2 8行目	…腕 $k$ を…期待値が $R_{p_k}$ で…	…腕 $i$ を…期待値が $R_{p_i}$ で…
4	6	1.1.2 9行目	…腕 $k$ を…	…腕 $i$ を…
3	9	下から 5行目	Bubeck らの説明図によれば, …	Bubeck らの説明図によれば, …
3	11	8行目	$\mu_i + U_i$ が最大の腕 $i$ を選ぶ	<u><math>x_i = \mu_i + U_i</math></u> が最大の腕 $i$ を選ぶ
3	23	式 (1.2.13)	… $[G_{t+1} S_t = s]$	… $[G_t S_t = s]$
3	23	式 (1.2.16) 1行目	… $[G_{t+1} S_t = s]$	… $[G_t S_t = s]$
3	24	式 (1.2.18) 1行目	… $[G_{t+1} S_t = s] = \dots$	… $[G_t S_t = s] = \dots$
3	25	式 (1.2.23)	… $[G_{t+1} S_t = s, A_t = a]$	… $[G_t S_t = s, A_t = a]$
3	25	式 (1.2.25)	… $\{\Xi = (s, a, s') \mid \forall s \in \mathcal{S}, \forall a \in \mathcal{A}, \forall s' \in \mathcal{S}\}$	… $\{\Xi = (s, a, s') \mid s \in \mathcal{S}, a \in \mathcal{A}, s' \in \mathcal{S}\}$
3	25	式 (1.2.26)	… $\{\Xi = (s, a, s') \mid \forall a \in \mathcal{A}, \forall s' \in \mathcal{S}\}$	… $\{\Xi = (s, a, s') \mid a \in \mathcal{A}, s' \in \mathcal{S}\}$
3	26	式 (1.2.27)	… $\{\Xi = (s, a, s') \mid \forall s' \in \mathcal{S}\}$	… $\{\Xi = (s, a, s') \mid s' \in \mathcal{S}\}$
3	26	式 (1.2.31)	… $\{\Xi = (s, a, s', a', s'') \mid \forall a \in \mathcal{A}, \forall s' \in \mathcal{S}\}$	… $\{\Xi = (s, a, s', a', s'') \mid a \in \mathcal{A}, s' \in \mathcal{S}, a' \in \mathcal{A}, s'' \in \mathcal{S}\}$
3	26	式 (1.2.32)	… $\{\Xi = (s, a, s', a', s'') \mid \forall s' \in \mathcal{S}\}$	… $\{\Xi = (s, a, s', a', s'') \mid s' \in \mathcal{S}, a' \in \mathcal{A}, s'' \in \mathcal{S}\}$

3	27	図 1.2.5	右のように修正	
3	30	9 行目	… $R_{t_k}, R_{t_k+1}, \dots$ がわかっているので…	… $R_{t_k+1}, R_{t_k+2}, \dots$ がわかっているので…
3	31	式 (1.3.11)	… $\left( r(s, a, s') + \sum_{a' \in \mathcal{A}(s')} \gamma P(a' s') Q^\pi(s', a') \right)$	… $\left( r(s, a, s') + \sum_{a' \in \mathcal{A}(s')} \gamma \pi(a' s') Q^\pi(s', a') \right)$
3	31	図 1.3.1	右のように修正	
3	32	式 (1.3.15)	… $= 0 + \gamma V^{\pi_1}(s_1)$	… $= -1 + \gamma V^{\pi_1}(s_1)$
3	32	式 (1.3.17)	… $= 0 + 0.8 \cdot 1 = 4$	… $= 0 + 0.8 \cdot 5 = 4$
3	32	式 (1.3.18)	… $= 0 + 0.8 \cdot 4 = 3.2$	… $= -1 + 0.8 \cdot 4 = 2.2$
3	33	式 (1.3.22)	… $= 0.5(-10 + 0.8V^{\pi_2}(s_1)) + 0.5(1 + 0.8V^{\pi_2}(s_4))$	… $= 0.5(-100 + 0.8V^{\pi_2}(s_1)) + 0.5(5 + 0.8V^{\pi_2}(s_4))$
3	33	式 (1.3.25)	… $\approx -121.7$	… $\approx -121.8$
4	33	下から 2 行目	…Sarsa は、時刻 $S_t$ で行動…	…Sarsa は、状態 $S_t$ において行動…
3	33	最下行	… $R_t$ と次状態 $S_{t+1}$ を…	… $R_{t+1}$ と次状態 $S_{t+1}$ を…
3	34	17 行目	… $P(a s')$ のかわりに、…	… $\pi(a s')$ のかわりに、…
3	34	式 (1.3.28)	… $+ \gamma Q(s_2, a_1) - \dots$	… $+ \gamma Q(s_3, a_1) - \dots$

3	35	5行目	… $\pi_1$ では…	… $\pi_2$ では…
3	36	9行目	図 1.3.1(b)は, …	図 1.3.2(b)は, …
3	36	式 (1.3.35)	$\dots + \gamma \max_{a' \in \mathcal{A}} Q^*(s', a')$	$\dots + \gamma \max_{a' \in \mathcal{A}(s')} Q^*(s', a')$
3	37	式 (1.3.36)	$\dots + \gamma \max_{a' \in \mathcal{A}(s')} Q(S_{t+1}, a')$	$\dots + \gamma \max_{a' \in \mathcal{A}(S_{t+1})} Q(S_{t+1}, a')$
3	37	式 (1.3.37)	$\dots + \gamma \max_{a' \in \mathcal{A}(s')} Q(S_{t+1}, a') \dots$	$\dots + \gamma \max_{a' \in \mathcal{A}(S_{t+1})} Q(S_{t+1}, a') \dots$
3	37	式 (1.3.38)	$r(s_1, a_1, s_2) + \gamma \max_{a' \in \mathcal{A}(s_2)} Q(s_2, a') \dots$	$r(s_1, a_1, s_3) + \gamma \max_{a' \in \mathcal{A}(s_3)} Q(s_3, a') \dots$
3	38	式 (1.3.39)	$\dots = r(s_1, a_1, s_2) + \gamma \max_{a' \in \mathcal{A}(s_2)} Q(s_2, a')$	$\dots = r(s_1, a_1, s_3) + \gamma \max_{a' \in \mathcal{A}(s_3)} Q(s_3, a')$
3	59	6行目	…が, <u>この観測は15%の確率で誤っている</u> としよう. これらの <u>情報化</u> から, …	…が, 15%の確率で <u>逆方向からの音が観測される</u> としよう. これらの <u>不確実な情報</u> から, …
3	83	1~4行目を差し替え, 6行目冒頭 9文字削除 (式(2.1.32)は変わらず) 脚注も削除	一般に…  …をもつ. ここで…	演算 $\max_{a_i}$ を含まないベルマンオペレータ $B^{\pi}$ は, 価値関数に状態遷移確率という重みで重み付き和をとって, 報酬 $r_{t+1}$ を足すという価値関数に関して線形な演算である. そのため, パラメータに関して線形な関数近似器を用いて価値関数 $V$ を近似し, $B^{\pi} V$ と $V$ の間のマハラノビス距離を小さくするようパラメータを勾配法を用いて更新する場合, そのマハラノビス距離の計量のとり方によらず, パラメータの更新則は以下のような線形な更新式で表現される. $\theta_{t+1} = \theta_t - \alpha_t (A\theta_t - b) \quad (2.1.32)$ ここで…
3	84	14行目	ここで, $R_{t,T}^{\lambda}$ は…	ここで, $R_{t+1,T}^{\lambda}$ は…
3	84	15行目	…) $R_t^{(n)}$ を用いて…	…) $R_{t+1}^{(n)}$ を用いて…
3	84	式 (2.1.39)	$R_{t,T}^{\lambda} = (1-\lambda) \sum_{n=1}^{T-t} \lambda R_t^{(n)} + \lambda^{T-t} R_t^{(T-t)}$	$R_{t+1,T}^{\lambda} = (1-\lambda) \sum_{n=1}^{T-t} \lambda^{n-1} R_{t+1}^{(n)} + \lambda^{T-t} R_{t+1}^{(T-t+1)}$

3	84	式 (2.1.40)	$R_t^{(n)} = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^{(n-1)} r_{t+n-1} + \gamma^n \hat{V}^\pi(s_{t+n-1}   \theta)$	$R_{t+1}^{(n)} = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots + \gamma^{(n-1)} r_{t+n} + \gamma^n \hat{V}^\pi(s_{t+n}   \theta)$																														
3	85	14 行目	$\sum_{k=1}^T \sum_{t=1}^T f_{t,k}$ の恒等式と...	$\sum_{k=1}^T \sum_{t=k}^T f_{t,k}$ の恒等式と...																														
3	85	17 行目	$\sum_{k=1}^T \sum_{t=1}^T f_{t,k}$ を用いれば...	$\sum_{k=1}^T \sum_{t=k}^T f_{t,k}$ を用いれば...																														
6	117	下から 7 行目	... $P$ が $T^{4/5}$ に比例する) ...	... $P$ が $T^{-4/5}$ に比例する) ...																														
3	193	4 行目	...ところで, 価値関数を時間方向に...	...ところで, 価値関数を時間逆方向に...																														
3	194	下から 5~4 行目	...(SPGPs)を用いた[16].	...(SPGPs)[16]を用いる方法などが考えられる.																														
3	194 ~ 195	194 ページ 下から 3 行 目~195 ペ ージ 2 行目	「ここで扱う状態の次元は 30... ..文献[14]を参照されたい。」の 5 行を削除する.																															
3	195	10 行目	...ここでは SPGPs を用いて...	...ここでは上述の SPGPs を用いて...																														
3	195	14 行目	...Integrals (PI2)を用いた.	...Integrals (PI <sup>2</sup> )を用いた[15].																														
3	196	6 行目	...軌道ベースまわりの価値関数...	...軌道まわりの価値関数...																														
3	198	文献[19]	Koji Ishihara, Jun Morimoto ...	K. Ishihara and J. Morimoto ...																														
3	260	図 3.7.2	(図の上部) ※破線で分けられた左右の 1 行目を入れ替える NIPS 版 Deep Q-Network[20]      Nature 版 Deep Q-Network[21]	Nature 版 Deep Q-Network[21]      NIPS 版 Deep Q-Network[20]																														
2	263	図 3.7.4	右図のように (棒グラフ下の数字を取る).	<table border="1"> <caption>Figure 3.7.4: Performance metrics for five games across four configurations.</caption> <thead> <tr> <th>Game</th> <th>wR-wT</th> <th>wR-woT</th> <th>woR-wT</th> <th>woR-woT</th> </tr> </thead> <tbody> <tr> <td>Breakout</td> <td>1.3</td> <td>1.0</td> <td>0.1</td> <td>0.1</td> </tr> <tr> <td>Enduro</td> <td>1.2</td> <td>1.0</td> <td>0.1</td> <td>0.1</td> </tr> <tr> <td>RiverRaid</td> <td>1.8</td> <td>1.0</td> <td>0.7</td> <td>0.4</td> </tr> <tr> <td>Seaquest</td> <td>3.5</td> <td>1.0</td> <td>1.2</td> <td>0.4</td> </tr> <tr> <td>SpaceInvaders</td> <td>1.3</td> <td>1.0</td> <td>0.4</td> <td>0.4</td> </tr> </tbody> </table>	Game	wR-wT	wR-woT	woR-wT	woR-woT	Breakout	1.3	1.0	0.1	0.1	Enduro	1.2	1.0	0.1	0.1	RiverRaid	1.8	1.0	0.7	0.4	Seaquest	3.5	1.0	1.2	0.4	SpaceInvaders	1.3	1.0	0.4	0.4
Game	wR-wT	wR-woT	woR-wT	woR-woT																														
Breakout	1.3	1.0	0.1	0.1																														
Enduro	1.2	1.0	0.1	0.1																														
RiverRaid	1.8	1.0	0.7	0.4																														
Seaquest	3.5	1.0	1.2	0.4																														
SpaceInvaders	1.3	1.0	0.4	0.4																														

2

289

図 4.1.2

右図のように (SNc/VTA の右の○を●にする).

