ゲーム AI への認知バイアス・生物学的制約の導入に向けた人間の行動選択に関する分析　正誤表

坂本　洸

ゲームプログラミングワークショップ (GPW-22) で発表した際の資料に誤りがありましたのでお詫びして訂正します.

- 表 2 について，データの載せ間違いがあったため， 訂正した表を以下に示します.
- 表 3 について，腕変更回数のカウント処理と， パターン 9，10 で行った特殊処理についてプログラムの実装ミスがありました. 実装を修正した上での訂正した表を以下に示します.
- パターン 9 と 10 について述べた 4.3【少ない結果から判断する】では，“人間はバンディットアルゴリズムよりも序盤に起きた偏った結果に引きずられやすい傾向がある” と主張していましたが， 実装の修正後では， 全体的にアルゴリズムも人間とはあまり差が見られない結果になっていました. そのため，結論も “序盤で 3 連続で当たる腕 A の選択率が，3 連続で外れる腕 B の選択率よりも高いという傾向はバンディットアルゴリズムでも見られる” となりました.
  表 3 の修正に合わせて，4.3【少ない結果から判断する】の内容を以下の通り訂正します.

表 2: 被験者の平均報酬，正解率,腕変更回数の平均と標準偏差

pattern	平均報酬	正解率	変更回数
1 (0.25：0.85)	0.75 ± 0.09	0.86 ± 0.11	6.6 ± 4.8
2 (0.55：0.45)	0.51 ± 0.09	0.63 ± 0.18	14.4 ± 12.8
3 (0.80：0.70)	0.77 ± 0.06	0.60 ± 0.30	8.1 ± 6.1
4 (0.20：0.30)	0.23 ± 0.06	0.54 ± 0.19	19.6 ± 13.6
5 (0.65：0.35)	0.60 ± 0.08	0.77 ± 0.16	10.8 ± 8.1
6 (0.05：0.10)	0.08 ± 0.05	0.62 ± 0.13	21.0 ± 15.4
7 (0.95：0.90)	0.93 ± 0.05	0.62 ± 0.36	8.2 ± 9.7
8 (0.50：0.50)	0.53 ± 0.05	(A:0.47 ± 0.26)	12.4 ± 10.6
9 (0.50：0.50)	0.50 ± 0.07	(A:0.77 ± 0.17)	10.4 ± 9.2
10 (0.35：0.55)	0.43 ± 0.08	(B:0.38 ± 0.22)	13.4 ± 11.3
11 (0.25：0.85)	0.71 ± 0.17	0.76 ± 0.14	3.3 ± 2.6
12 (0.30：0.20)	0.24 ± 0.11	0.50 ± 0.13	6.1 ± 2.0
13 (0.80：0.70)	0.72 ± 0.12	0.70 ± 0.23	3.7 ± 2.7
14 (0.50：0.50)	0.46 ± 0.12	(A:0.48 ± 0.20)	4.6 ± 2.5

表 3: 各バンディットアルゴリズムの平均報酬，正解率，腕変更回数の平均と標準偏差 (mean ± sd)

pattern	ε -greedy (ε =0.2)			UCB (c=$1/\sqrt{2}$)			Thompson-Sampling		
	平均報酬	正解率	変更回数	平均報酬	正解率	変更回数	平均報酬	正解率	変更回数
1	0.75 ± 0.09	0.84 ± 0.13	8.9 ± 3.7	0.81 ± 0.05	0.93 ± 0.04	4.8 ± 2.0	0.81 ± 0.06	0.93 ± 0.05	5.0 ± 3.0
2	0.51 ± 0.08	0.60 ± 0.31	9.6 ± 4.0	0.51 ± 0.07	0.64 ± 0.22	11.3 ± 4.2	0.51 ± 0.07	0.62 ± 0.24	15.5 ± 6.3
3	0.76 ± 0.07	0.62 ± 0.30	9.4 ± 4.0	0.76 ± 0.06	0.65 ± 0.21	10.5 ± 5.9	0.76 ± 0.06	0.64 ± 0.26	13.7 ± 6.8
4	0.26 ± 0.07	0.62 ± 0.31	10.4 ± 4.2	0.26 ± 0.07	0.63 ± 0.18	17.5 ± 4.8	0.26 ± 0.07	0.62 ± 0.20	17.5 ± 5.4
5	0.58 ± 0.10	0.75 ± 0.23	9.2 ± 3.8	0.60 ± 0.08	0.82 ± 0.13	8.7 ± 3.5	0.59 ± 0.08	0.80 ± 0.15	11.1 ± 5.8
6	0.08 ± 0.04	0.61 ± 0.28	13.0 ± 5.5	0.08 ± 0.04	0.57 ± 0.11	32.6 ± 6.2	0.08 ± 0.04	0.58 ± 0.15	21.4 ± 4.4
7	0.93 ± 0.04	0.60 ± 0.27	10.9 ± 5.2	0.93 ± 0.04	0.59 ± 0.15	24.4 ± 12.8	0.93 ± 0.04	0.62 ± 0.28	13.5 ± 7.2
8	0.50 ± 0.07	(A:0.50 ± 0.32)	9.7 ± 4.0	0.50 ± 0.07	(A:0.50 ± 0.24)	11.6 ± 4.2	0.50 ± 0.07	(A:0.50 ± 0.25)	16.2 ± 6.2
9	0.50 ± 0.07	(A:0.87 ± 0.09)	9.1 ± 3.6	0.50 ± 0.07	(A:0.81 ± 0.16)	7.7 ± 1.8	0.50 ± 0.07	(A:0.87 ± 0.10)	8.9 ± 4.7
10	0.39 ± 0.07	(B:0.17 ± 0.15)	9.1 ± 3.7	0.44 ± 0.07	(B:0.41 ± 0.20)	9.3 ± 2.4	0.41 ± 0.07	(B:0.26 ± 0.17)	13.0 ± 5.0
11	0.67 ± 0.21	0.70 ± 0.30	1.8 ± 1.6	0.74 ± 0.14	0.82 ± 0.13	2.3 ± 1.3	0.70 ± 0.16	0.76 ± 0.16	2.7 ± 1.5
12	0.25 ± 0.14	0.56 ± 0.32	2.4 ± 1.8	0.26 ± 0.14	0.57 ± 0.18	5.0 ± 1.7	0.25 ± 0.14	0.54 ± 0.18	4.2 ± 1.5
13	0.76 ± 0.14	0.53 ± 0.36	1.7 ± 1.7	0.76 ± 0.14	0.59 ± 0.26	3.9 ± 2.5	0.76 ± 0.14	0.56 ± 0.27	3.1 ± 1.8
14	0.50 ± 0.16	(A:0.51 ± 0.36)	1.8 ± 1.7	0.50 ± 0.16	(A:0.52 ± 0.24)	3.7 ± 1.7	0.50 ± 0.16	(A:0.50 ± 0.23)	3.6 ± 1.6

【序盤の偏った結果の影響】

最初に腕 A は 3 連続で当たり，腕 B は 3 連続で外れるという特殊な処理を施したパターン 9，10 について述べる．

パターン 9 はどちらも報酬確率が 0.5 であり，通常特殊な処理をしていない場合はパターン 8 のように各腕の選択率は 0.5 程度になると考える．パターン 9 は，特殊処理の後は通常の報酬確率に戻ったものの，腕 A の選択率は 0.77 とパターン 8 よりも高くなり，腕の変更回数は減少する結果になった．

パターン 10 は各腕の報酬確率が同じパターン 9 と異なり，報酬確率が 0.35 と低い腕 A が 3 連続で当たり，報酬確率が 0.55 で高い腕 B が 3 連続で外れるという特殊処理を行った．その結果，特殊処理後の報酬確率は腕 B の方が高いにもかかわらず，腕 B の選択率は 0.38 となり腕 A の選択率よりも低くなった．これらの傾向は人間だけでなく，各バンディットアルゴリズムでも見られた．

また，パターン 10 における被験者選択履歴では，後半（25 ステップ以降）で報酬確率の低い腕 A を選択し続けるという振る舞いは 20 試行中 6 試行確認できた．加えて，腕 A を選択し続けるような振る舞いでなく，腕を交互に選択するような振る舞いであったとしても終盤まで腕 A を多く選択する傾向が見られた．これは，最初は連続で当たった腕 A が当たりにくくなったとしても，自身の持つ，腕 A が当たりやすいという信念を修正できず固執してしまう「保守性バイアス」などに当てはまるのではないかと推測する．