

統数研共同研究集会「経済物理学とその周辺」24度第一回研究会

2012年8月27日～28日

キャノングローバル研究所

進化型IPDを用いた最適戦略の探索

糸井良太, 田中美栄子

鳥取大学大学院工学研究科

情報エレクトロニクス専攻

目次

1. はじめに
2. 目的
3. 囚人のジレンマ
4. 進化型IPDモデル
5. 考察
6. まとめ

はじめに

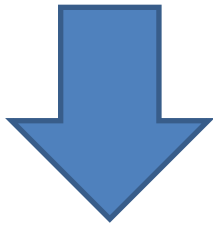
- 二酸化炭素排出問題
- 核兵器保有問題
- 価格競争問題 etc

対立構造問題における疑問

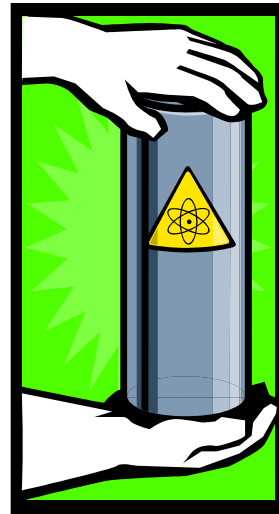
『自己利益を高める行動選択』とは？

はじめに ～核兵器保有問題～

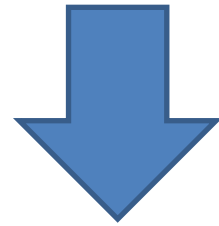
核を持たない



リスクが高い



核を持つ



リスクが低い

『核を持たない = 利益が高い』ならば

高い利益を得るにはどうすれば良いか

最適な行動選択が必要

目的

最適な行動選択
を発見する

四人のジレンマ

囚人のジレンマとは



個人にとって最適な振る舞いが
社会全体にとって損となるモデル

		プレイヤーB	
		協力	裏切り
プレイヤーA	協力	R, R	S, T
	裏切り	T, S	P, P

$$\text{※ } S + T < 2R, S < P < R < T$$

繰り返し囚人のジレンマ ～アクセルロッドの実験～

複数の異なる**戦略**(行動選択法)を用いて

囚人のジレンマゲームを200回繰り返して得点を競う大会を開催



協力的な戦略が良い結果を残す

しかし、用いられた戦略は100個にも満たない

➡ もっと多くの条件で実験を行うべき

進化型IPDモデル

進化型IPDモデル ～どんなモデルか？～

生物進化の要素を取り入れたモデル

- 新しい戦略が**突然変異**によって出現

➡ 戦略の自動生成

- 戦略の優劣により**人口(同じ戦略に所属する人の数)**が変化

➡ 強い戦略の人口は増加する



強い戦略が繁栄し生き残る

進化型IPDモデル ～戦略の扱い～

**過去の相手と自分の行動から次手を決定
(履歴)**

履歴1の一例

直前の
相手の行動

裏切り

協力

次の行動

履歴2の一例

1手前の
自分の行動

直前の
相手の行動

裏切り

協力

直前の
相手の行動

裏切り

協力

裏切り

協力

裏切り

協力

裏切り

協力

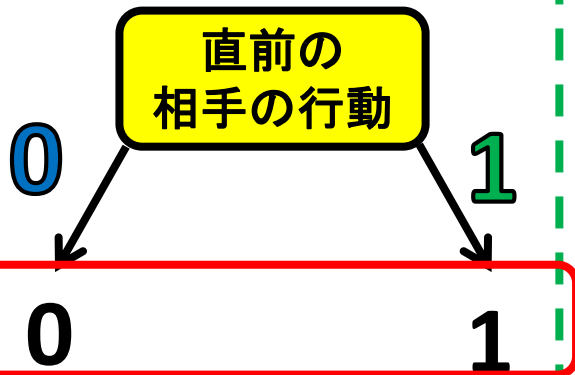
裏切り

協力

進化型IPDモデル ～戦略の2値化～

過去の相手と自分の行動から次手を決定
(履歴)

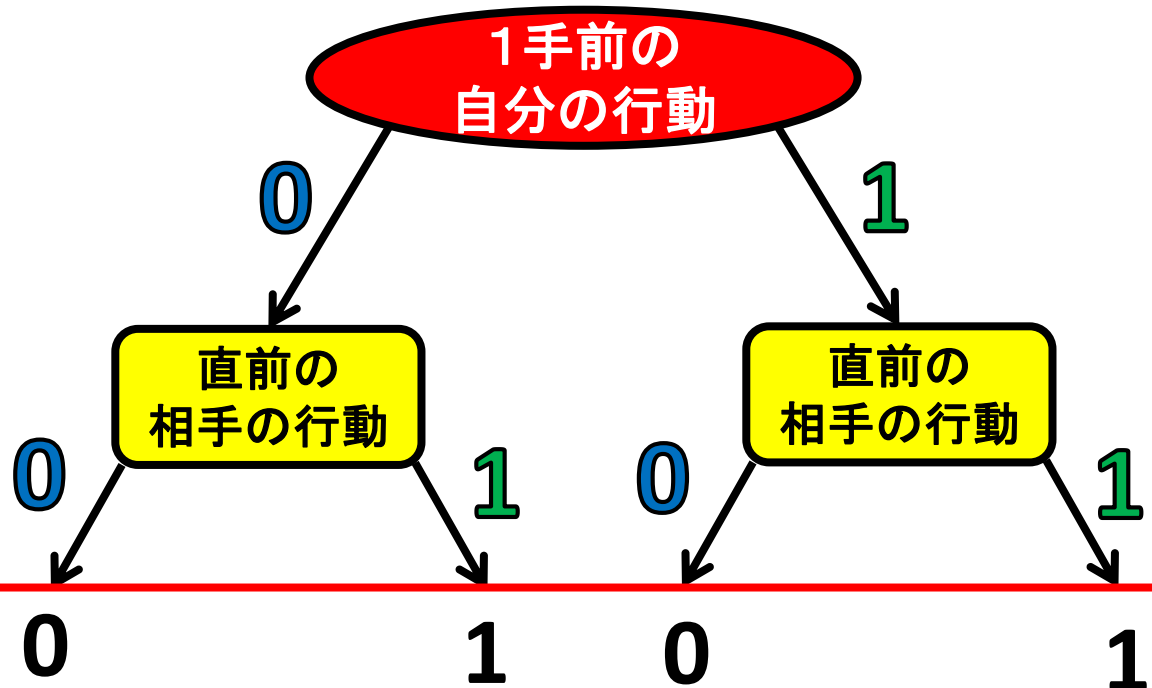
履歴1の一例



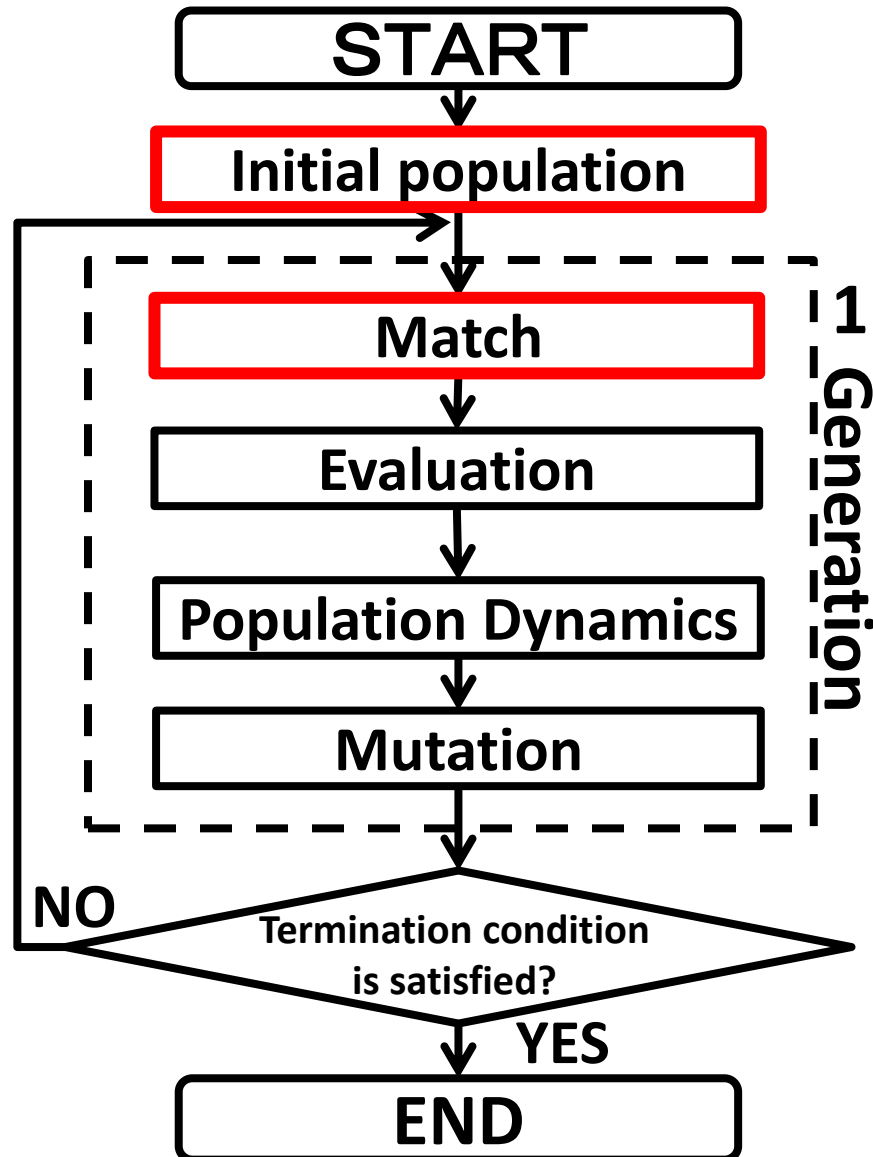
↑
戦略名



履歴2の一例



進化型IPDモデルのフローチャート



対戦の流れ ～総当たり～

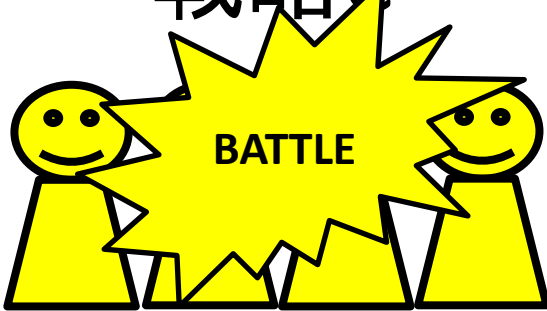
戦略A



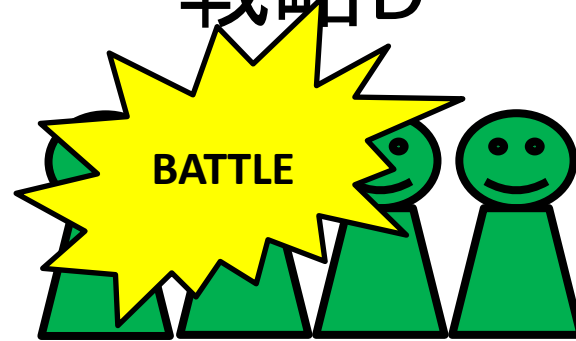
戦略B



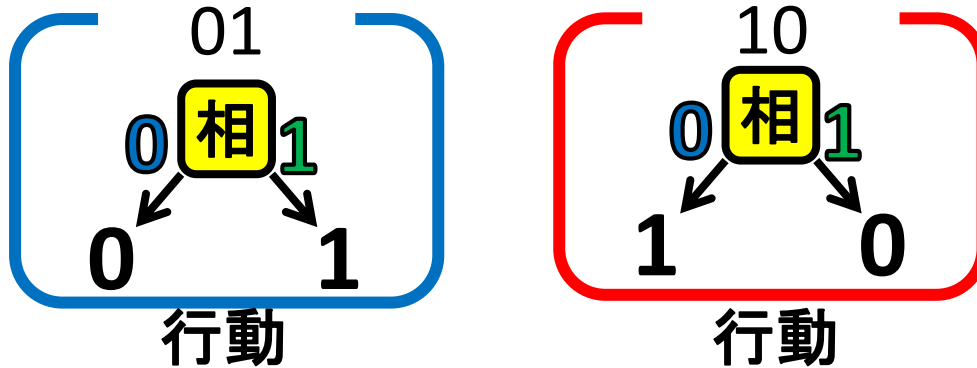
戦略C



戦略D



対戦の流れ

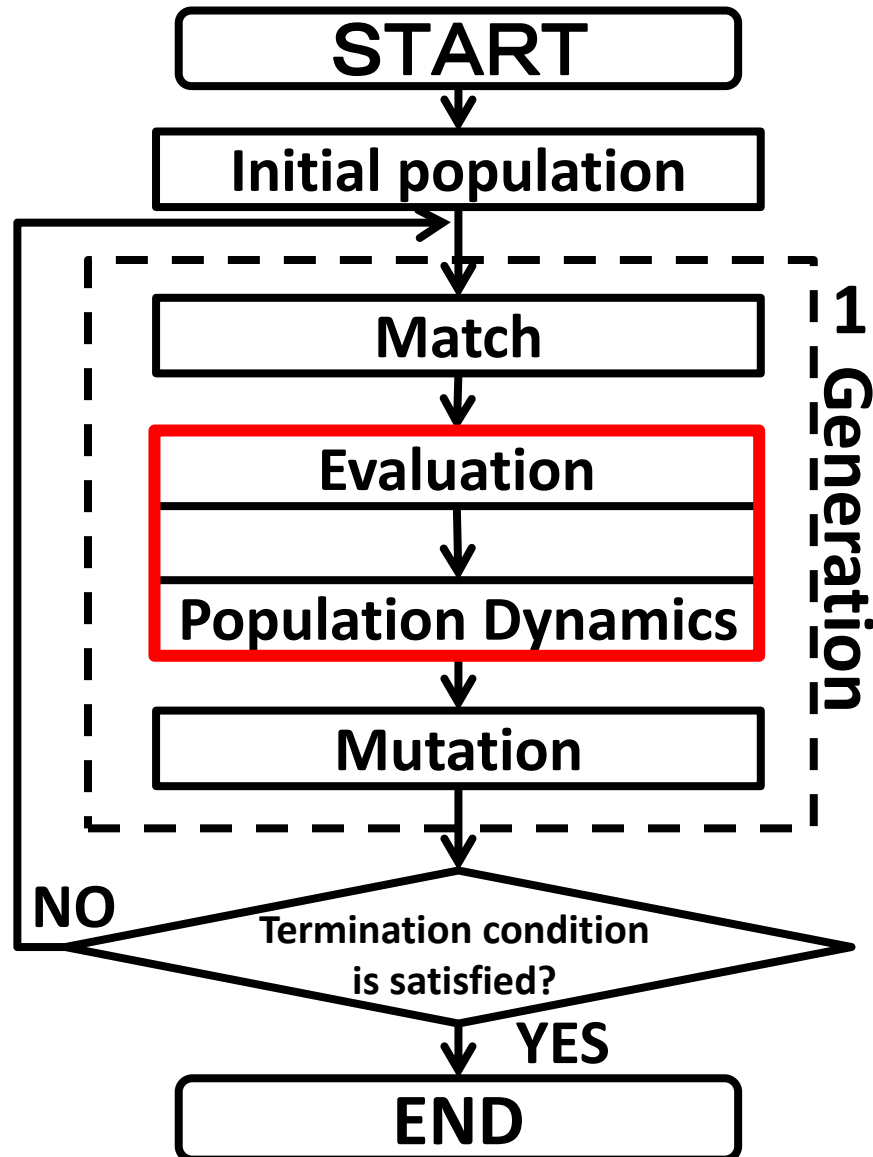


得点

得点

3	1		1	3
0	1		0	5
5	0	行動失敗	0 → 1	0
3	1		1	3
0	1		0	5
0	1 ← 0	行動失敗	0	5
1	0		0	1

進化型IPDモデルのフローチャート



世代交代 ～人口の更新～

世代交代には以下の式を用いる

$$x_i(t+1) - x_i(t) = \alpha(S_i(t) - S(t))x_i(t)$$

戦略*i*の獲得利得

全体利得平均

$$S_i(t) = \sum_j \frac{g_{i,j}(t)x_j(t)}{c_{i,j}}$$

$$S(t) = \sum_i S_i(t)x_i(t)$$

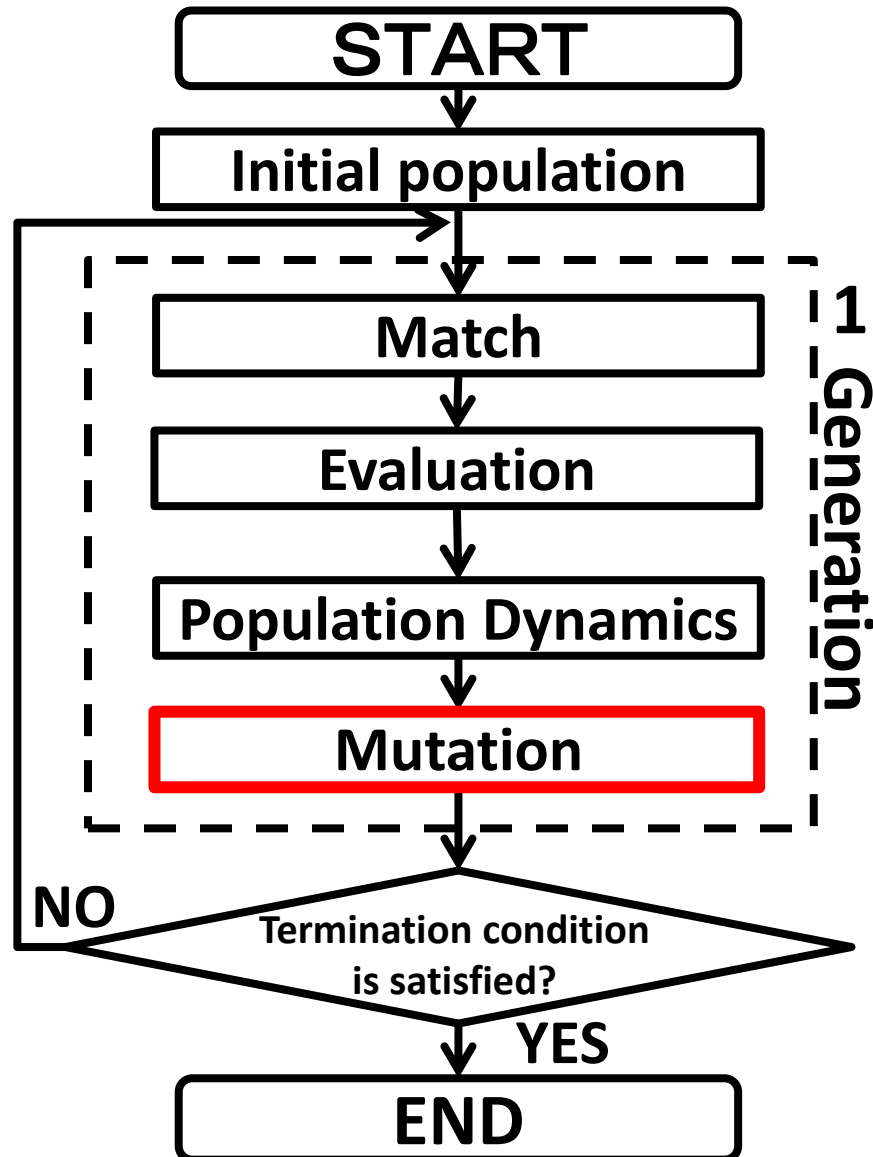
x_i : 戦略*i*の人口比率

$g_{i,j}$: 戦略*i*が戦略*j*から得た利得

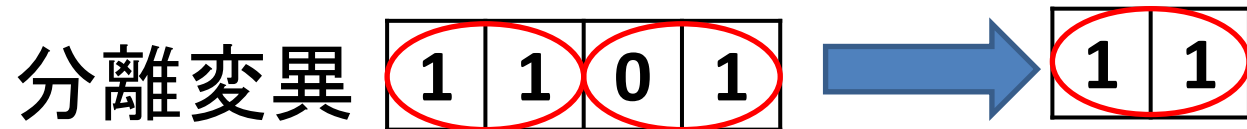
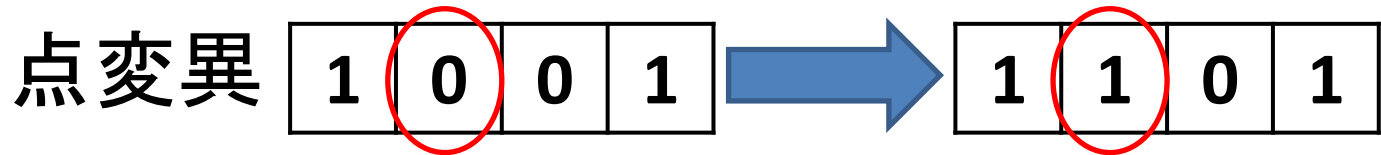
$c_{i,j}$: 戦略*i*が戦略*j*と対戦した回数

α : 定数

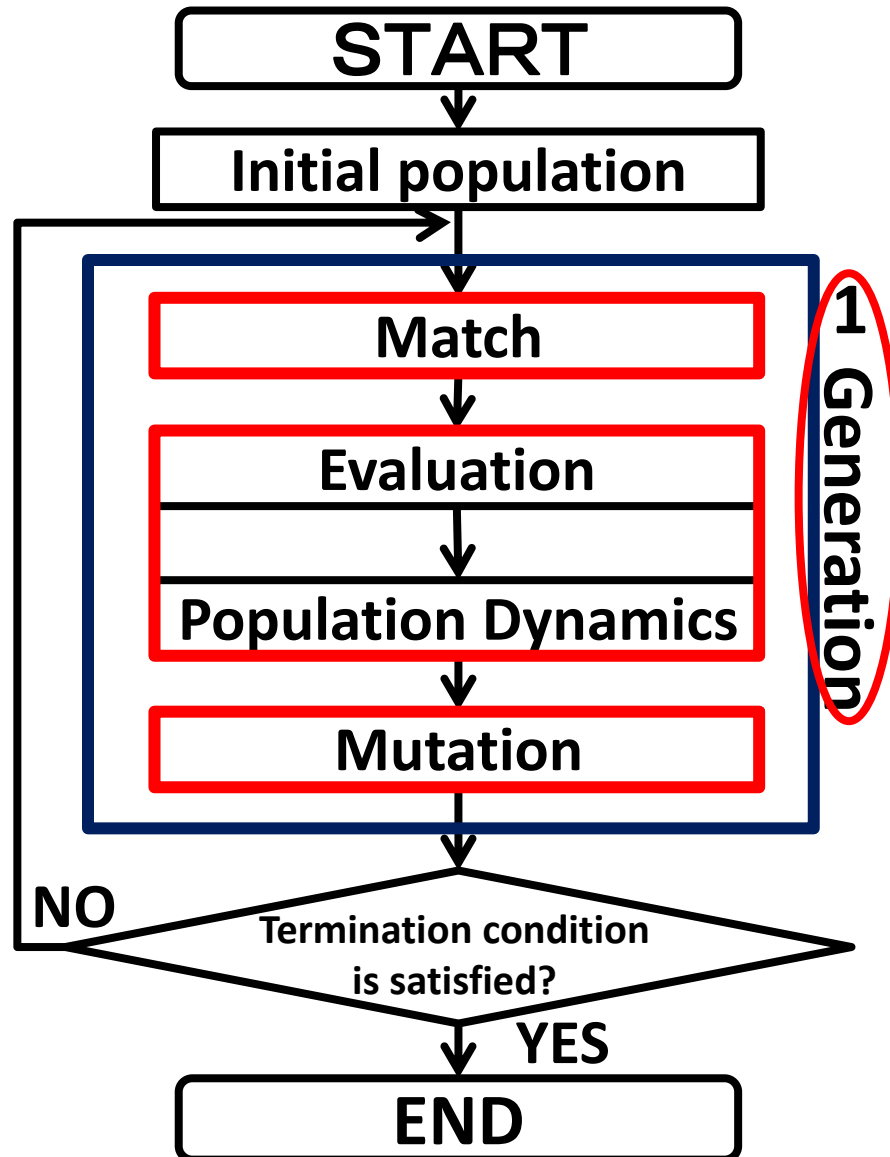
進化型IPDモデルのフローチャート



進化型IPDモデル ～突然変異～



進化型IPDモデルのフローチャート



実験条件

二つの条件で実験を行う

条件1

初期戦略	初期人口比率
00	25%
01	25%
10	25%
11	25%

条件2

初期戦略

初期人口比率

ランダム

共通の条件

点変異 : 2×10^{-5}

複写変異 : 1×10^{-5}

分離変異 : 1×10^{-5}

ノイズ発生率 : 1%

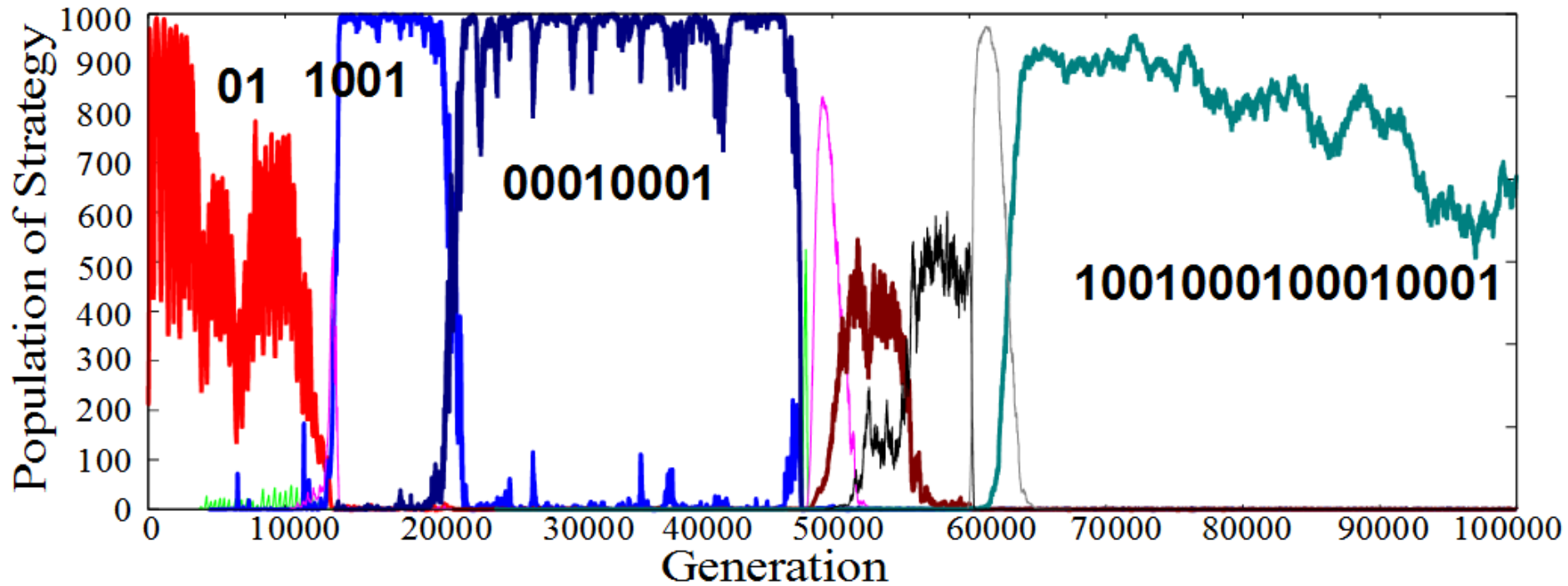
対戦終了確率 : 0.01%

利得表

自分, 相手	協力	裏切り
協力	3, 3	0, 5
裏切り	5, 0	1, 1

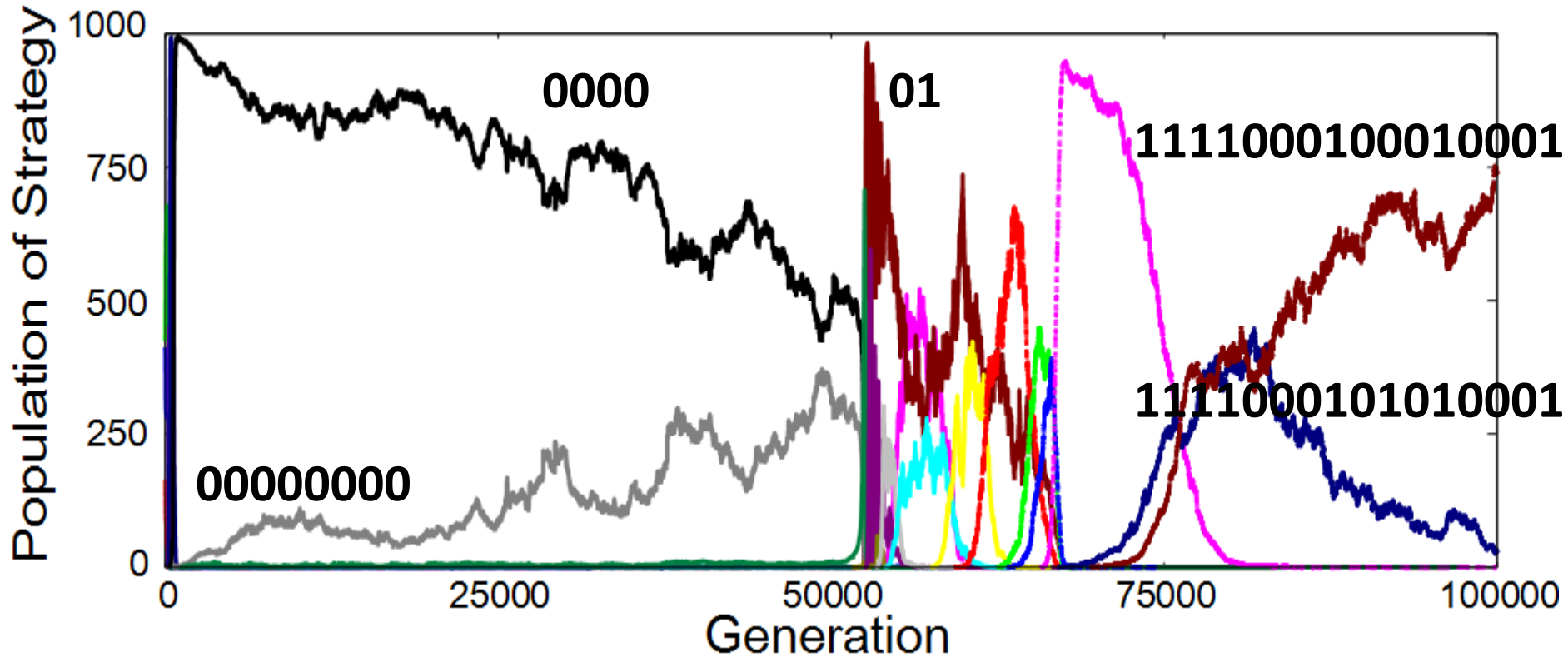
実験結果

シミュレーション結果①(初期戦略固定)



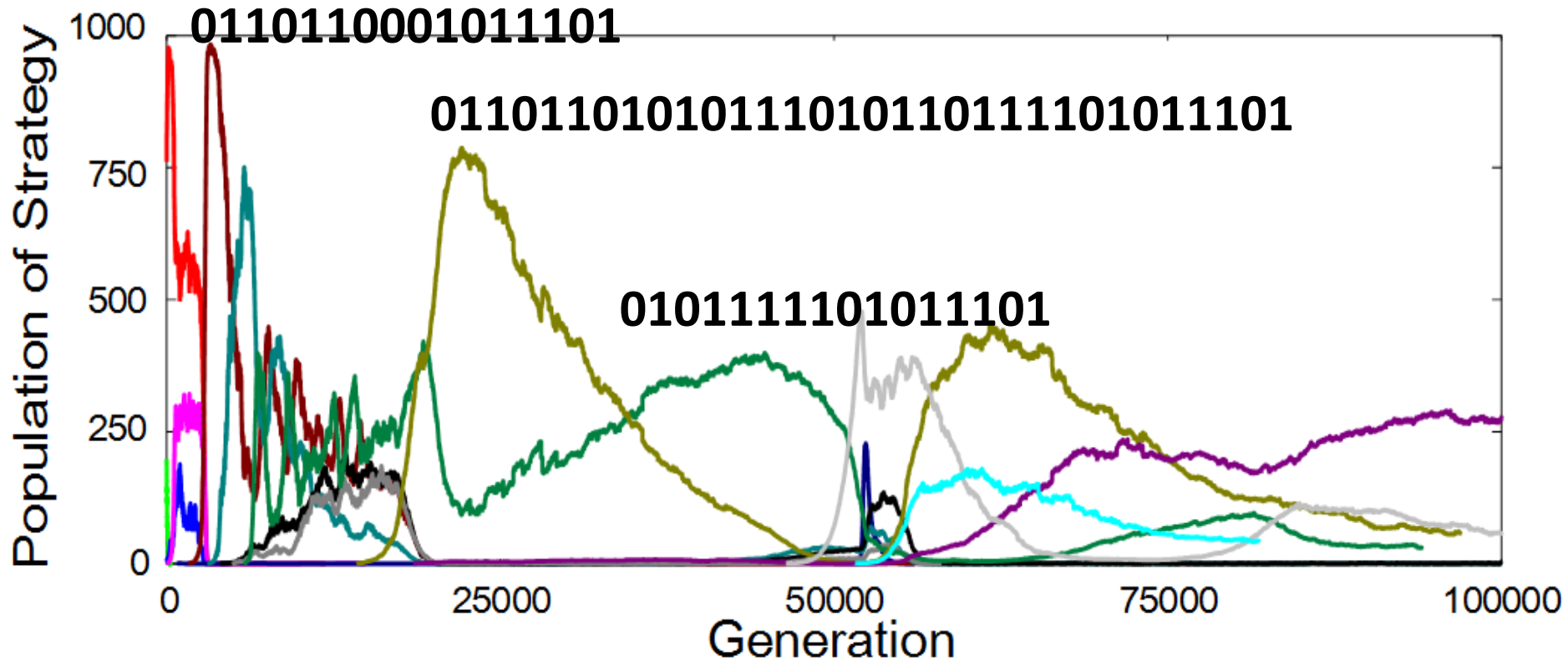
単純な戦略からより複雑な戦略へと進化する

シミュレーション結果②(初期戦略固定)



単純な戦略からより複雑な戦略へと進化する

シミュレーション結果③(初期戦略ランダム)



強い戦略は常に移り変わる

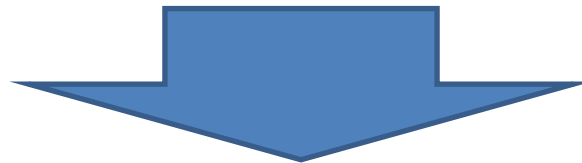
考察

戦略の利得による評価

強い戦略とは？



平均値以上の利得を得る戦略



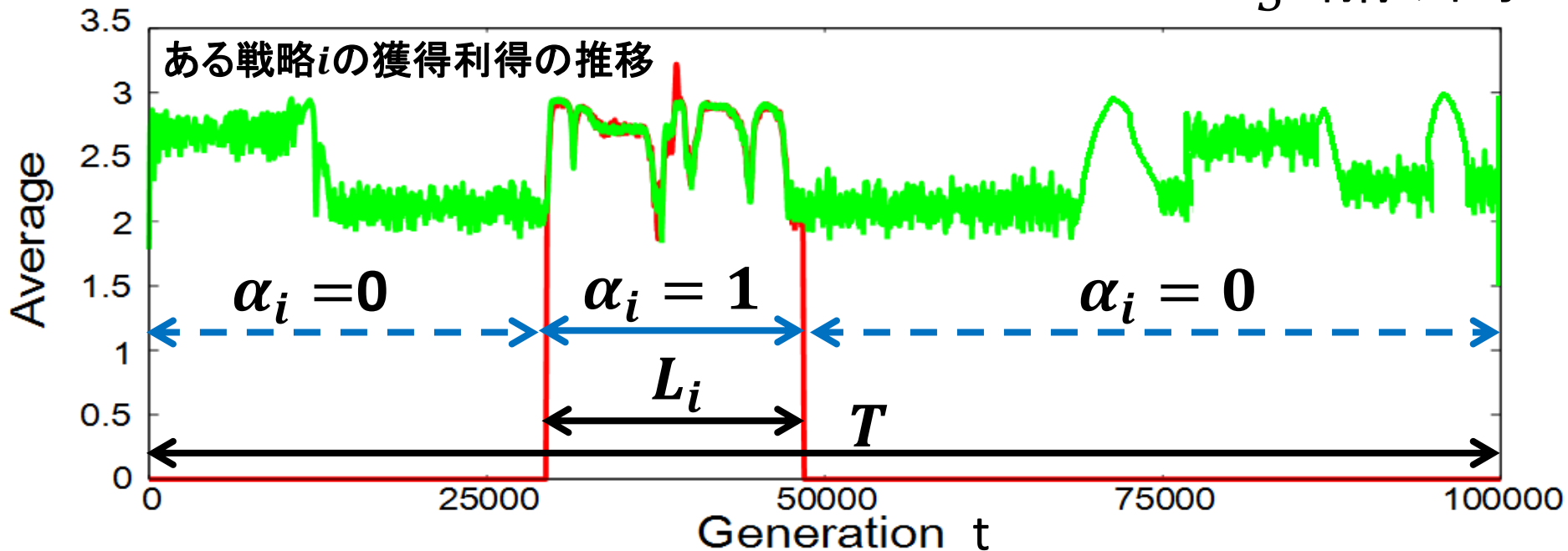
戦略の利得と平均値との差が評価の基準

戦略*i*の評価値*W_i*の求め方

$$W_i = \frac{1}{L_i} \sum_{t=1}^T \alpha_i(t) (S_i(t) - S(t))$$

$$\alpha_i(t) = \begin{cases} 0 & (\text{Strategy } i \text{ is DEAD}) \\ 1 & (\text{Strategy } i \text{ is ALIVE}) \end{cases}$$

— S_i 戦略*i*の利得
— S 利得の平均



戦略の評価結果

初期戦略固定

Strategy	W_i
1101 1001	0.123
0101 1001	0.077
1101 0110	0.064
1010 0011	0.050
1101 0100	0.047
1001 0001 0001 0001	0.041
0001 1011	0.040
0100 1001	0.039
1101 0111	0.029
1001 1011 1001 1011	0.028

初期戦略ランダム

Strategy	W_i
1011	0.078
0000 0011	0.070
1101 1010	0.059
1001 1001	0.049
1101 1011 1101 1011	0.045
0001 0011	0.038
1101 0101 0001 1001	0.036
1101 1101 0000 0111	0.032
1000 0000 0100 0001	0.029
1111 0101 0101 1110	0.027

$W_i \geq 0$ を満たす戦略を対象に分析を行う

分析対象戦略数

初期戦略固定

シミュレーション45回で出現した戦略

8153(重複あり)

$W_i \geq 0$ を満たした戦略の数

340(重複あり)

初期戦略ランダム

シミュレーション50回で出現した戦略

11753(重複あり)

$W_i \geq 0$ を満たした戦略の数

785(重複あり)

戦略の分析 ~文字列分析~

1001

0000 1111

1111 0000 1010 0101

戦略の分析 ~文字列分析~

戦略の長さを合わせる

1001 1001 1001 1001 1001 1001 1001 1001

0000 1111 0000 1111 0000 1111 0000 1111

1111 0000 1010 0101 1111 0000 1010 0101

戦略の分析 ~文字列分析~

各配置に'1'が使われる確率を求める

1001 1001 1001 1001 1001 1001 1001 1001

0000 1111 0000 1111 0000 1111 0000 1111

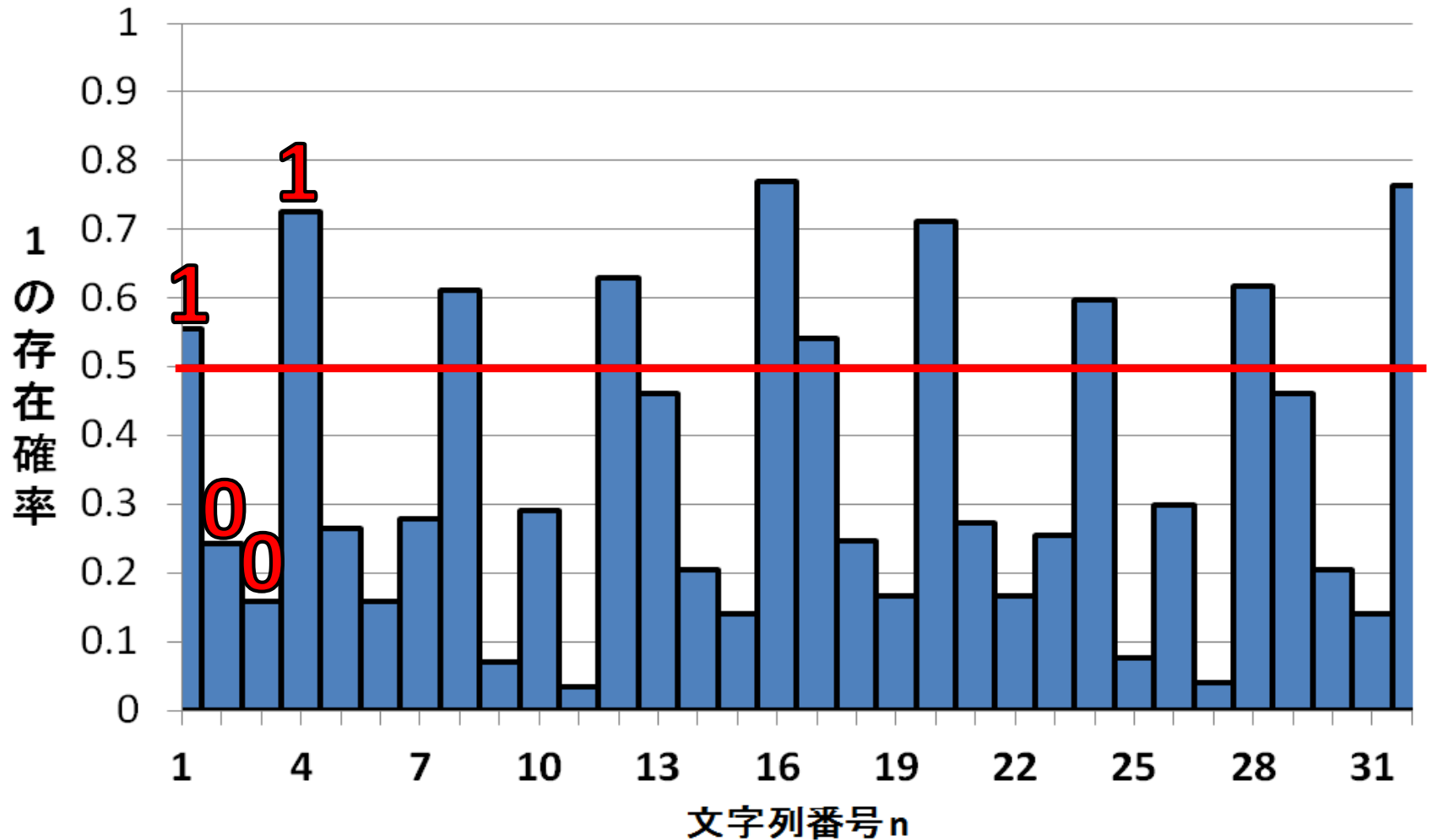
1111 0000 1010 0101 1111 0000 1010 0101

0.66

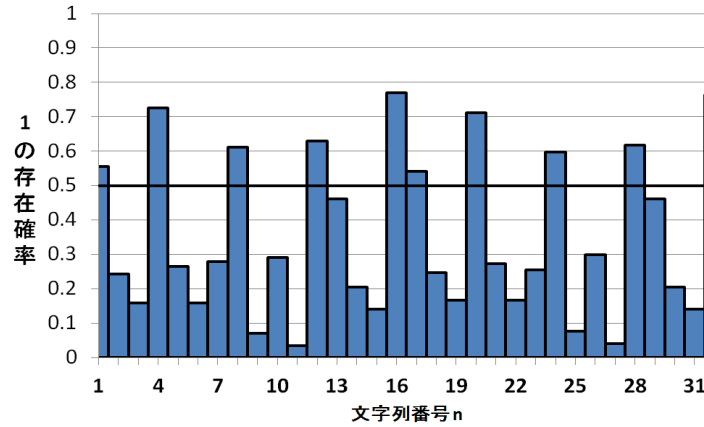
0.33

0.00

分析結果(初期戦略固定)



最も多かった戦略



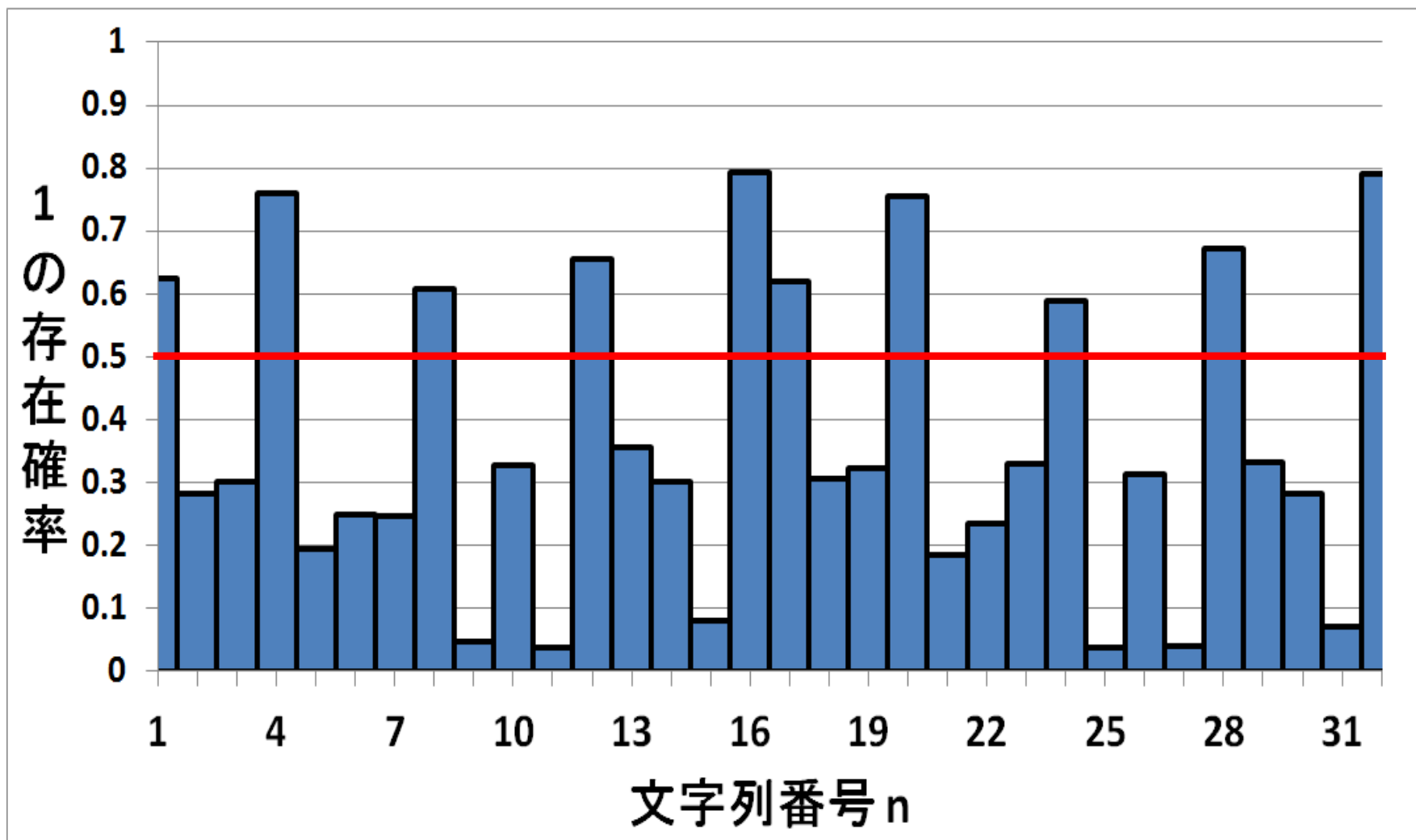
1001 0001 0001 0001 **1001 0001 0001 0001**

同じ文字列

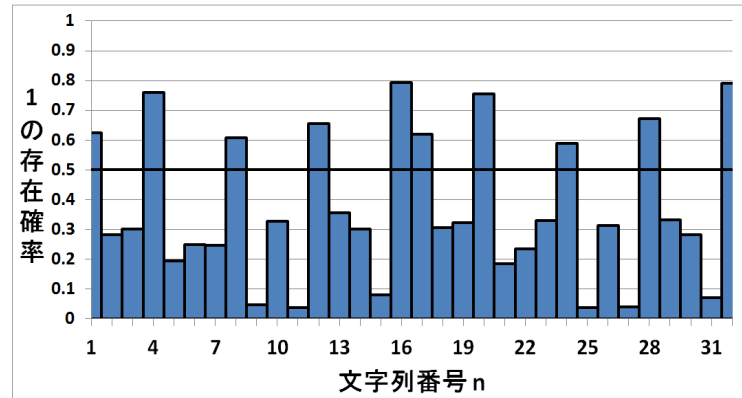


1001 0001 0001 0001

分析結果(初期戦略ランダム)



最も多かった戦略



1001 0001 0001 0001 **1001 0001 0001 0001**

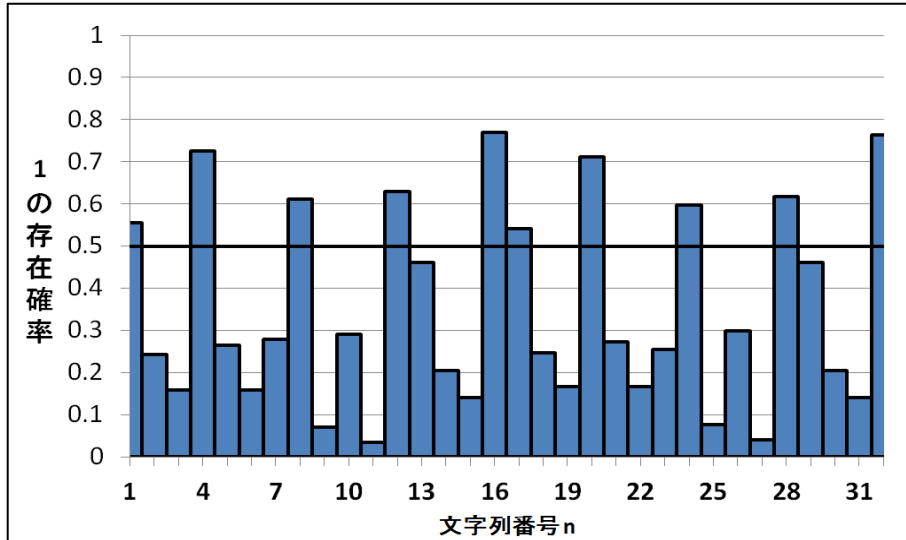
同じ文字列



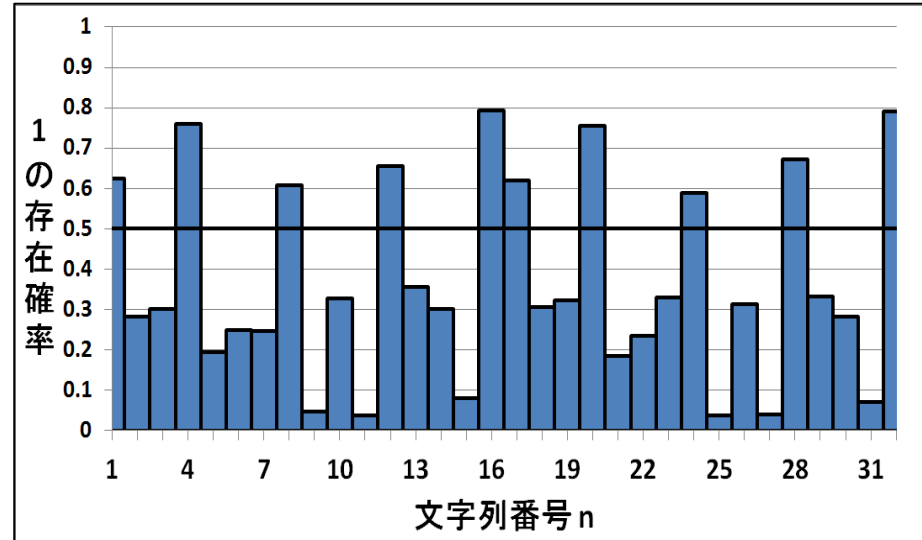
1001 0001 0001 0001

分析結果 まとめ

初期戦略固定



初期戦略ランダム

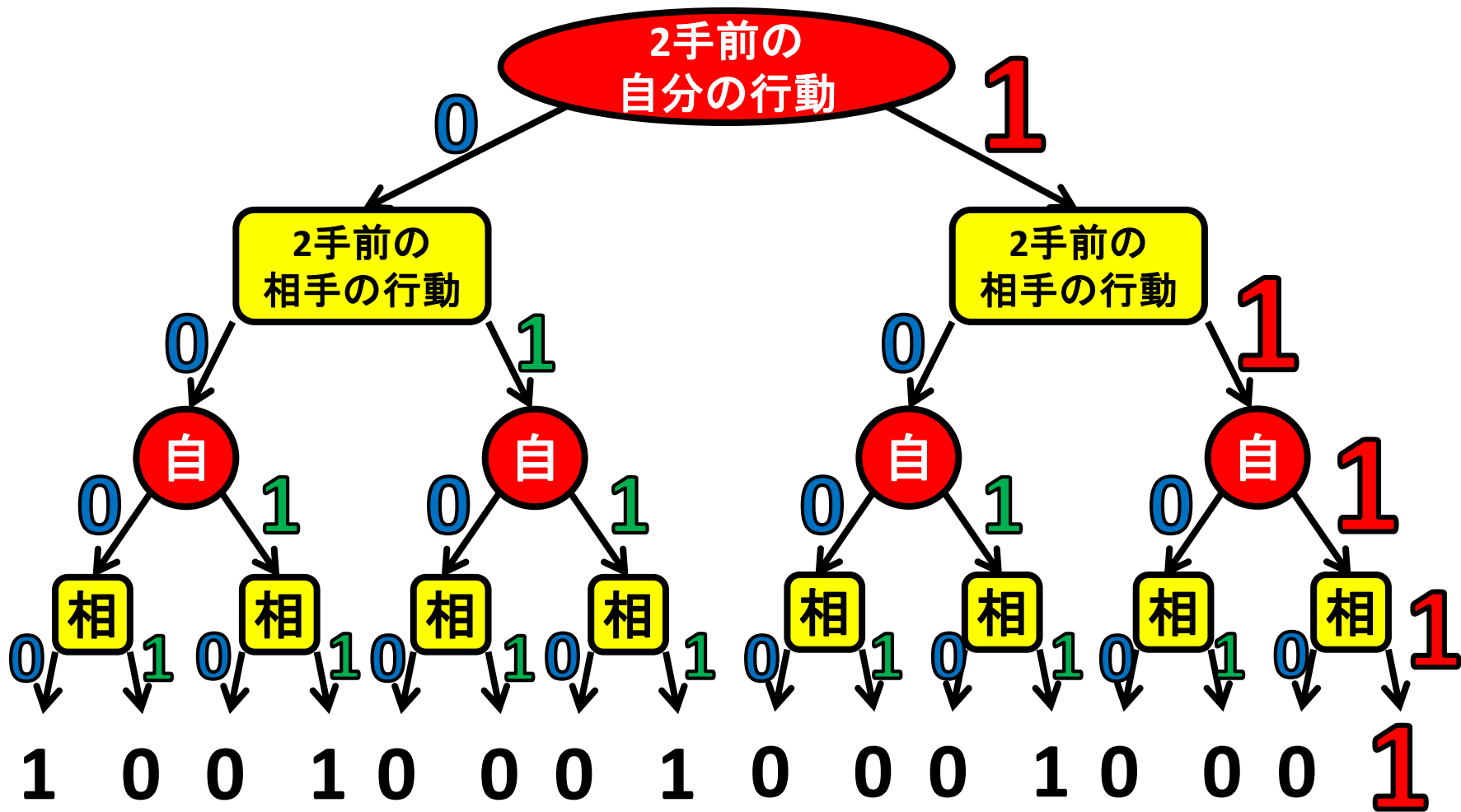


初期戦略に関わらず

1001 0001 0001 0001

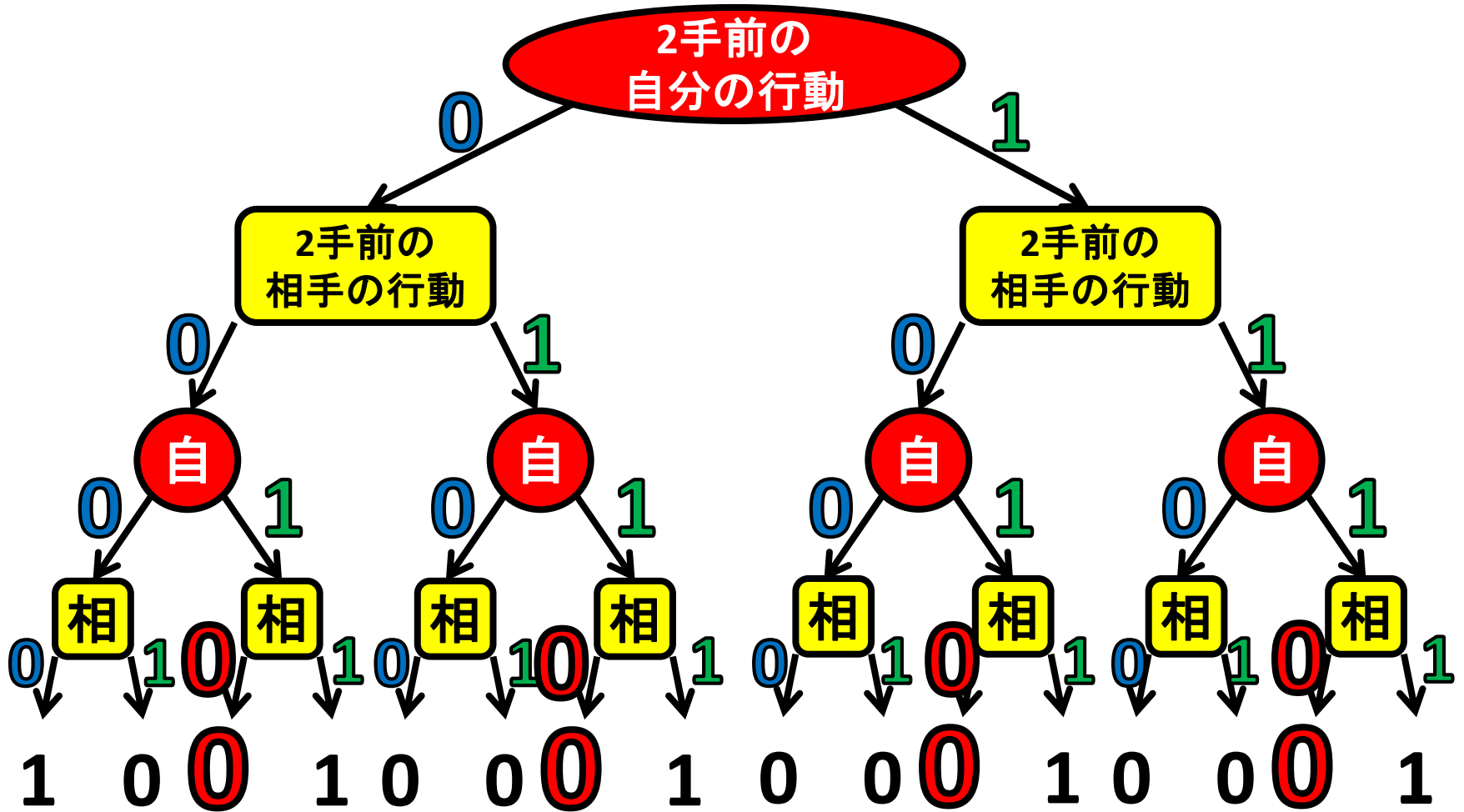
構造の戦略が良い結果を残した

[1001 0001 0001 0001]の特徴①



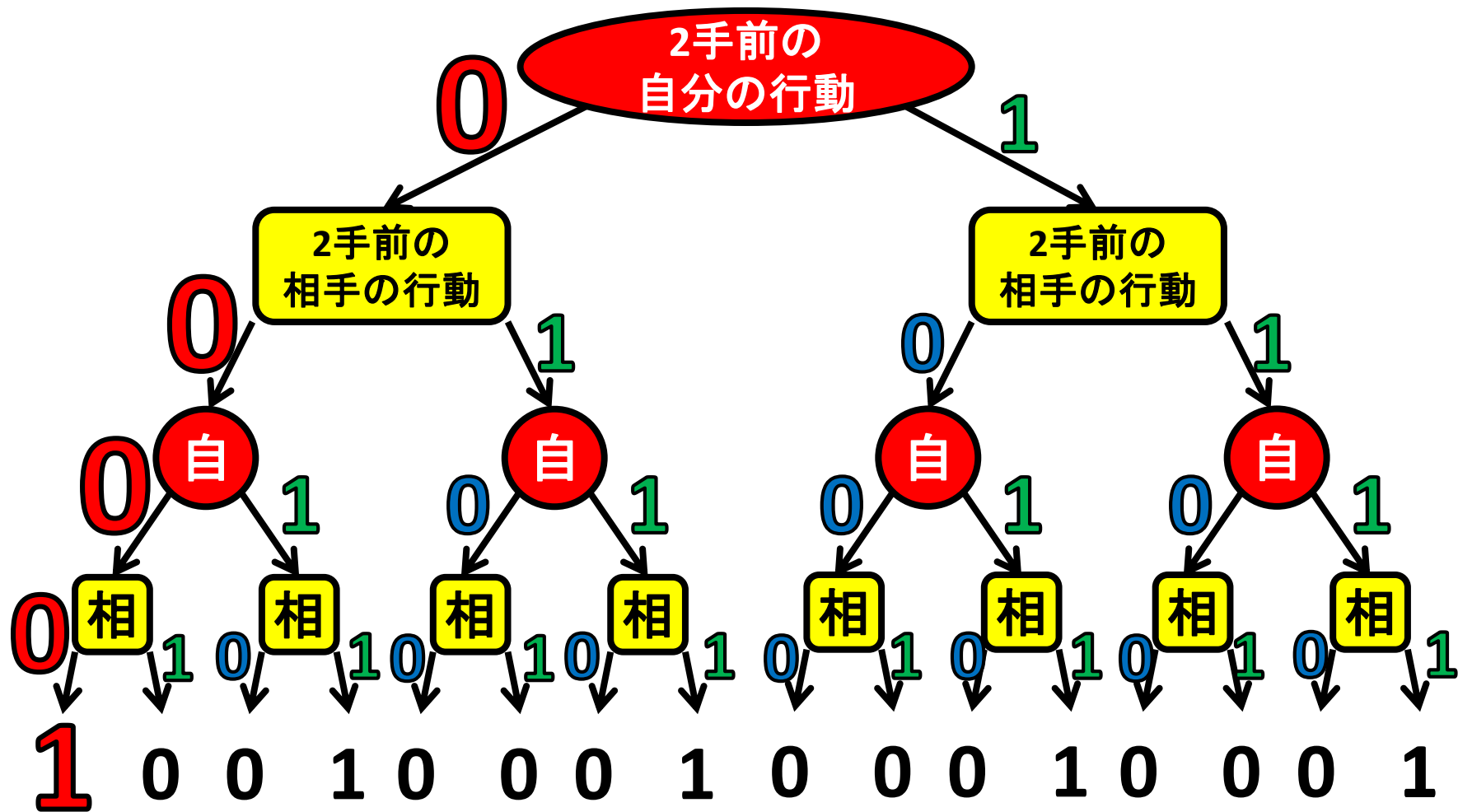
自分からは裏切らない

[1001 0001 0001 0001]の特徴②



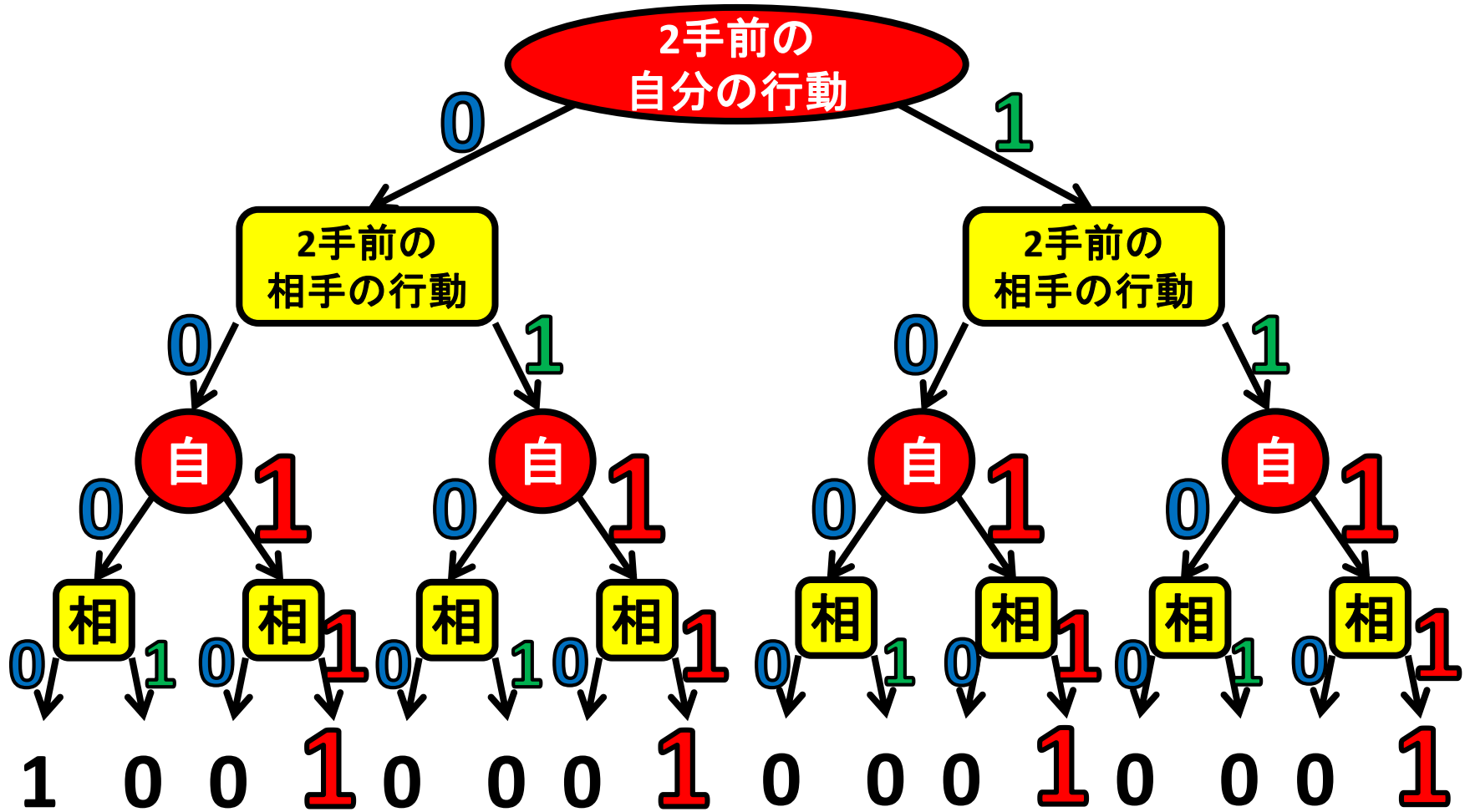
裏切られたら 即座に報復する

[1001 0001 0001 0001]の特徴③



裏切り合いの関係が続くと関係修復を図る

[1001 0001 0001 0001]の特徴④



協力関係成立時には、協力関係の維持を行う

まとめ

研究目的

最適な行動選択を発見する

- **進化型IPDモデル**を用いて様々な戦略を出現, 競わせた
- 出現した戦略を**評価、分析**を行った

まとめ

研究結果

強い戦略に共通する特徴を発見した

- 自分からは**裏切らない**
- 裏切られたら**即座に報復**
- 裏切り合いが続くと、**関係修復を図る**
- 協力関係成立時には、**協力関係の維持を行う**

まとめ

今後の課題 実験条件を変えて行う

- より長い戦略が出現できる環境
- 利得表の変更
- 各パラメータの変更

ご清聴ありがとうございました