

法政大学学術機関リポジトリ

HOSEI UNIVERSITY REPOSITORY

PDF issue: 2024-07-28

Baldwin効果における実世界の性質の役割

田村, 太樹 / TAMURA, Daiki

(発行年 / Year)

2009-03-24

(学位授与年月日 / Date of Granted)

2009-03-24

(学位名 / Degree Name)

修士(工学)

(学位授与機関 / Degree Grantor)

法政大学 (Hosei University)

Baldwin 効果における実世界の 性質の役割

法政大学大学院 工学研究科 システム工学専攻

田村 太樹

目次

1. はじめに	3
2. 獲得形質遺伝の法則	4
3. Baldwin 効果.....	5
4. 従来の研究.....	7
5. 本研究.....	7
6. 逃避タスク	7
6.1. タスク	7
6.2. 追跡エージェント	8
6.3. 逃避エージェント	9
7. 進化計算のフローチャート	11
8. 進化による逃避エージェントの学習機構の獲得(GA)	12
9. 実世界を用いた処理の実現	13
9.1. 遺伝子	14
9.2. 状態構成回路	14
9.3. 行動構成回路	17
10. 学習機構の実現	19
11. シミュレーション	19
11.1. シミュレーションの設定	19
12. 結果	21
13. 考察	29
14. おわりに	32
参考文献	33
謝辞.....	34

1. はじめに

生物は、環境に適応するために学習を行い、その学習結果は進化に対して影響を与えている。この進化と学習の相互関係について様々な議論がなされてきた。その先駆けとして Lamarck は、獲得形質遺伝の法則を提案した[1][2]。

この獲得形質遺伝の法則は、親が学習により獲得した形質が遺伝により子へ受け継がれるというものである。しかし、これは学習により遺伝子が書き換えられるとは考え難く、現在は否定されている。

その後、Lamarck が提案した獲得形質遺伝の法則と同じような効果のみられる Baldwin 効果という新しい説が提案された。

Baldwin 効果とは、学習する能力にあたる形質は遺伝し、選択されてその結、獲得形質が遺伝するかのように見える現象である[3]。鈴木らによれば繰り返し囚人のジレンマゲームやクラシファイアシステムによる仮想空間などのシミュレーションでの Baldwin 効果については確認されているが、実世界での環境において Baldwin 効果が現れるかについては確認されていない[4-10]。

そこで本研究では、Baldwin 効果が実世界の性質に存在すると考え、その効果の役割について検証を行う。

具体的には、「ニュートンの3法則」及び、「光の逆二乗の法則」を考慮した実世界の性質考慮した環境において逃避行動を目的とするロボットについてシミュレーションを行い、進化させる。学習しやすさを適応度とし、Baldwin 効果において実世界の性質がどのような役割をもつか検証を行う。

2. 獲得形質遺伝の法則

獲得形質遺伝の法則とは、親が学習により獲得した形質が遺伝により子へ受け継がれていくというものであった。以下 Fig.1 によって説明を行う。

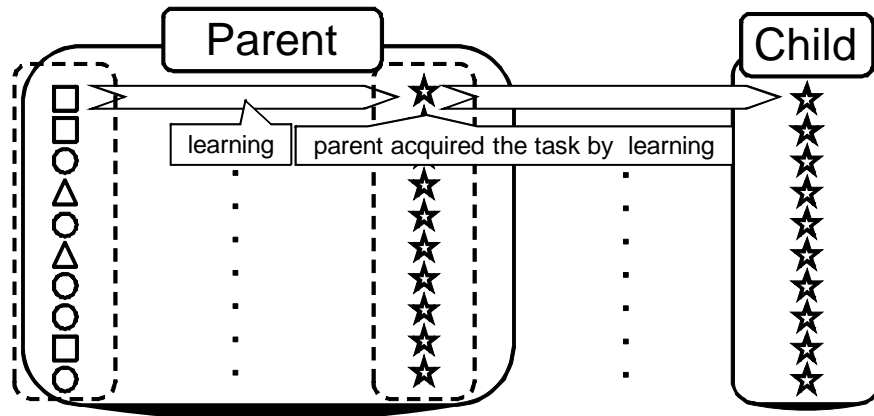


Fig.1 獲得形質遺伝の法則

親のある一個体を例に説明を行います。親が現在■という個体であるとします。この個体はあるタスクを学習によって獲得し、★という個体になります。そして、その獲得したタスクが遺伝子に書き込まれ子の個体に遺伝します。子は、親が学習して獲得して行うことができるようになったタスクを先天的に行うことができるというものです。

しかし、現在ではこの法則は、実現不可能であると考えられています。なぜならば、親が学習した結果が子に遺伝子を用いて受け継がれるとは考え難いからです。わかりやすくいうと、親が東大出身だからといって子は生まれながらに東大レベルの学力を持っているとは考えられないからです。

そこでこの法則と同様の効果がみられる Baldwin 効果が Baldwin によって提案された。

3. Baldwin 効果

Baldwin 効果とは学習する能力にあたる形質は遺伝し選択され、その結果獲得形質が遺伝するかのように見える現象である[3]。以下 Fig.2 と Fig.3 によって説明を行う。

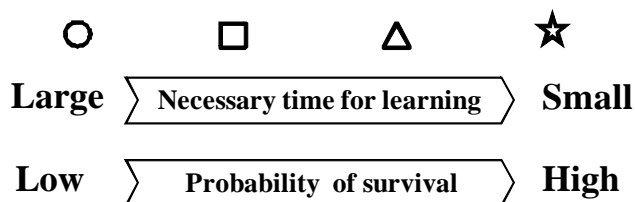


Fig.2 学習個体

Fig.2 では、Fig.3 で用いる個体の説明を行います。●の個体はあるタスクを獲得するために必要な学習が多い個体を表しています。そして、右の個体に進むにつれて必要な学習量が減少していきます。

また、あるタスクを獲得するために必要な学習量が少ない個体というのは、環境に適用しやすい個体と同位であると考えられています。そのため、自然淘汰などにより生き残っていく確率が高くなります。

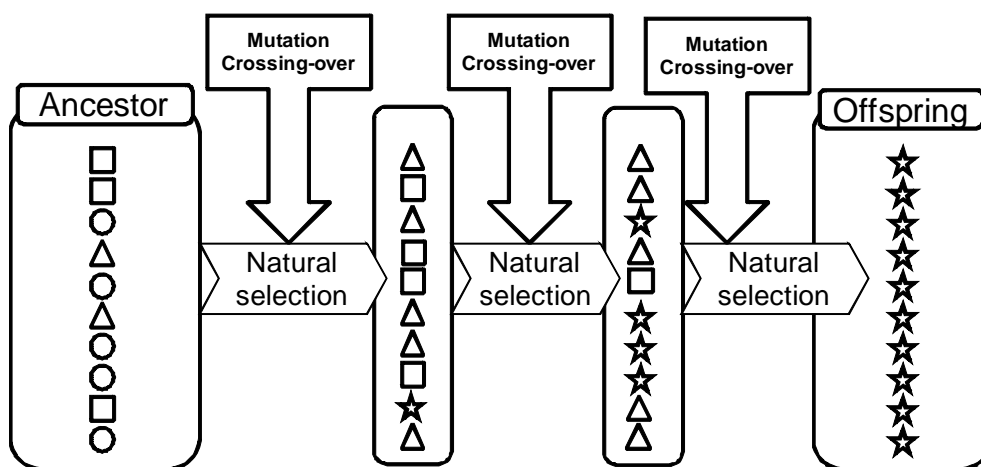


Fig.3 Baldwin 効果

Fig.3 は Baldwin 効果の概略を表しています。親の世代が図のような個体であるとして、親の世代では、自然淘汰により環境に適応しやすい個体が生き残っていきます。図の中では、あるタスクを獲得するために一番学習量の多い●の個体が減少していき、一番学習量の少ない▲の個体が増加していきます。ここで、突然変異や交叉によって★の個体が生まれる。この★の個体はほかのどの個体よりも学習量の少ない個体になります。★の個体の存在によって次は■の個体が減少し★の個体が増加していきます。これを繰り返すことにより、最終的に子の世代では★の個体が数多く残ることになります。

つまり、自然淘汰、突然変異や交叉によってあるタスクを獲得するために必要な学習量の少ない個体が残っていき、子の世代では、ほとんど学習の必要のない個体が残ることになります。

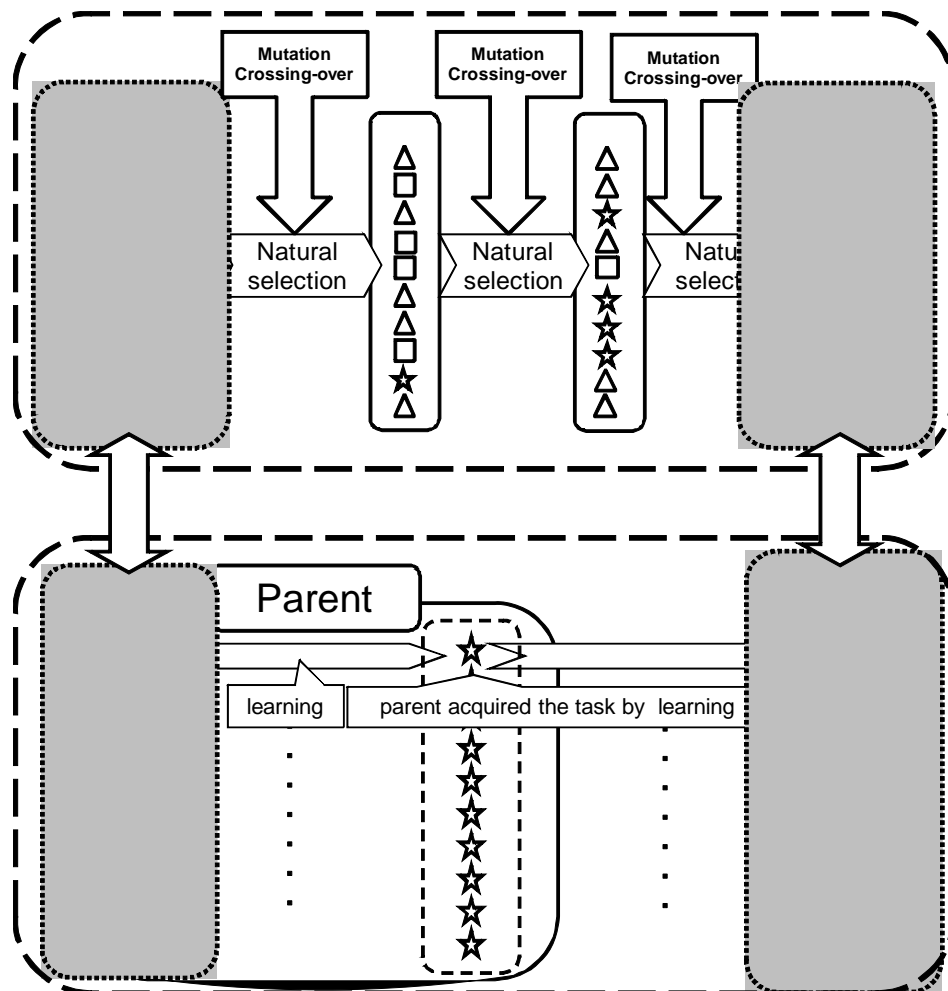


Fig.4 Baldwin 効果と獲得形質遺伝の法則

Fig.4 の示すとおり親から子の変化だけを見ると、Baldwin 効果は獲得形質遺伝の法則と同様に学習した結果が遺伝子に書き込まれ子に遺伝したかのように見ることができる。

4. 従来の研究

これまでの研究では、鈴木らによって繰り返し囚人のジレンマゲームや井上らによるクラシファイアシステムにおいて仮想環境での Baldwin 効果については確認がされてきた。

しかし、実世界の環境においてはまだ検証されていない。

5. 本研究

本研究では、Baldwin 効果が実世界の性質に存在すると考え、その効果の役割について検証を行う。

具体的には、「ニュートンの運動方程式」及び、「光の逆二乗の法則」を考慮した実世界の環境において逃避行動を目的とするロボットについてシミュレーションを行い、進化させる。そして、学習しやすさを適応度とし、Baldwin 効果において実世界の性質がどのような役割をもつか検証を行う。

6. 逃避タスク

6.1. タスク

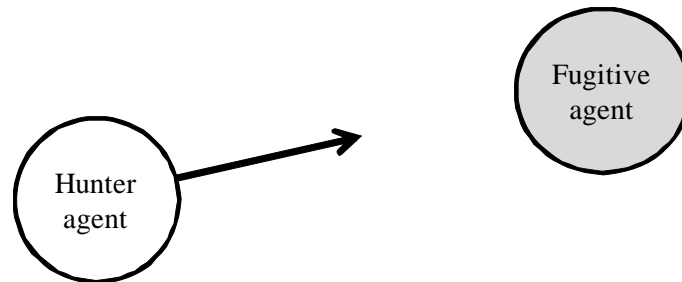


Fig.5 逃避行動

追跡エージェントは、常に逃避エージェントのいる方向に進むように設定されています。そして、設定された推進力で追跡し、その推進力は試行ごとに一定とします。

逃避エージェントは、ランダムに設定される推進力で追跡エージェントから逃げ、その逃げ方を学習と進化によって獲得していきます。

6.2. 追跡エージェント

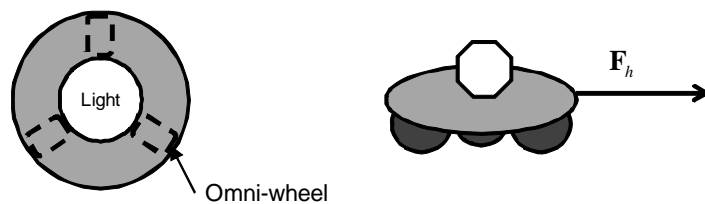


Fig.6 追跡エージェント

Fig.6 は追跡エージェントの概要図を表わしている。3つのオムニホイールによって全方向への移動を可能としている。そして、追跡エージェントの行動を行うための運動方程式は式(1)で表わされる。

$$\mathbf{F}_h = m \frac{d^2 \mathbf{x}_h}{dt^2} + c \frac{d\mathbf{x}_h}{dt} \quad (1)$$

追跡エージェントには、ライトを中央に搭載しておりそのライトの明るさをさまざまに設定することが可能である。

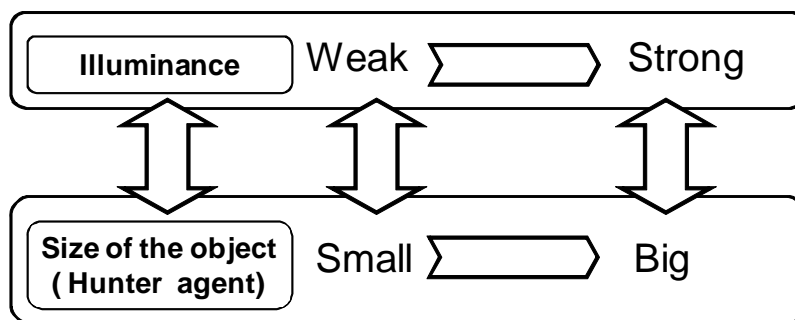


Fig.7 ライトの明るさと物体の明るさの対応

ライトの明かりは、物体の大きさを表すために用いており、Fig.7のように実世界の生物に対応させるとライトの明るさが物体の大きさを示していてライトの明るさが弱ければ物体の大きさは小さいと認識し、ライトの明るさが強ければ物体の大きさは大きいと認識するように設定されている。

6.3. 逃避エージェント

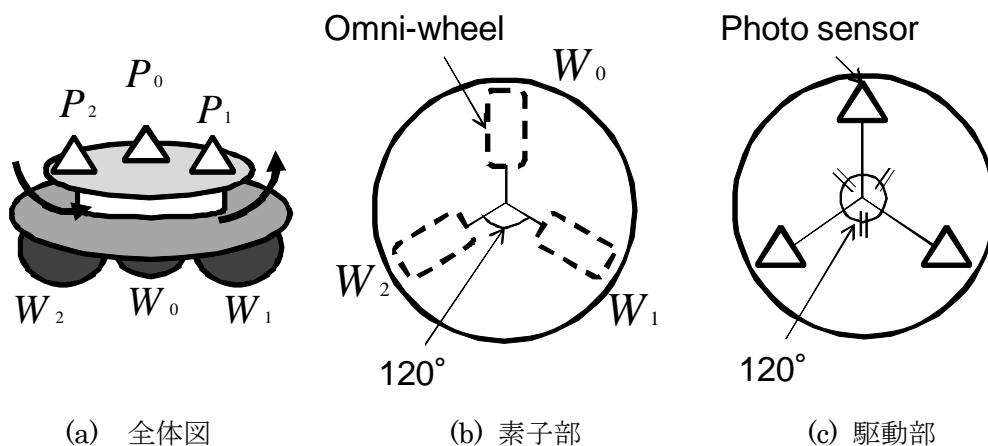


Fig.8 逃避エージェント

逃避エージェントは、3つの光センサを搭載しておりそれらによって光の強い方向を認識することができる。そして、これらの光センサは指向性を持ち、センサの正面($\theta = 0^\circ$)に光がある場合最も強く反応する。正面からの角度が大きくなるにつれて反応は、弱くなっていくものとする。また、感知可能な最大角度は左右 120° とする。式(3)は、センサの指向性を表している。 p_n は、光センサの出力、 k_{sensor} は光センサのゲイン、 I は照度を表している。

$$p_n = k_{sensor} \times \left(\frac{-\theta}{120} + 1 \right) \times I \quad (2)$$

また、式(4)は環境のモデルとしている光の逆2乗則を表している。 L が、光度を表し、 D がエージェント同士の距離を表している。

$$I = \frac{L}{D^2} \quad (3)$$

逃避エージェントは、追跡エージェントと同様にオムニホイールによって全方向に対して移動が可能であり、行動するための運動方程式は式(2)で表わされる。

$$\mathbf{F}_f = m \frac{d^2 \mathbf{x}_f}{dt^2} + c \frac{d\mathbf{x}_f}{dt} \quad (4)$$

しかし、実際はオムニホイルの力の関係を以下の図のように考える。

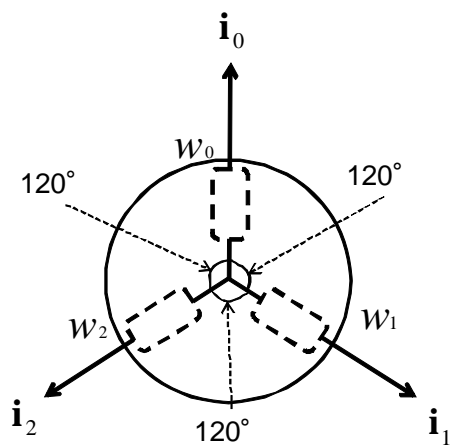


Fig.9 オムニホイルの力の関係

そして、Fig.8 より以下の式を用いて考えていく。

$$\mathbf{F}_f = w_0 \mathbf{i}_0 + w_1 \mathbf{i}_1 + w_2 \mathbf{i}_2 \quad (5)$$

\mathbf{i}_0 ; w_0 の単位ベクトル, \mathbf{i}_1 ; w_1 の単位ベクトル, \mathbf{i}_2 ; w_2 の単位ベクトル

7. 進化計算のフローチャート

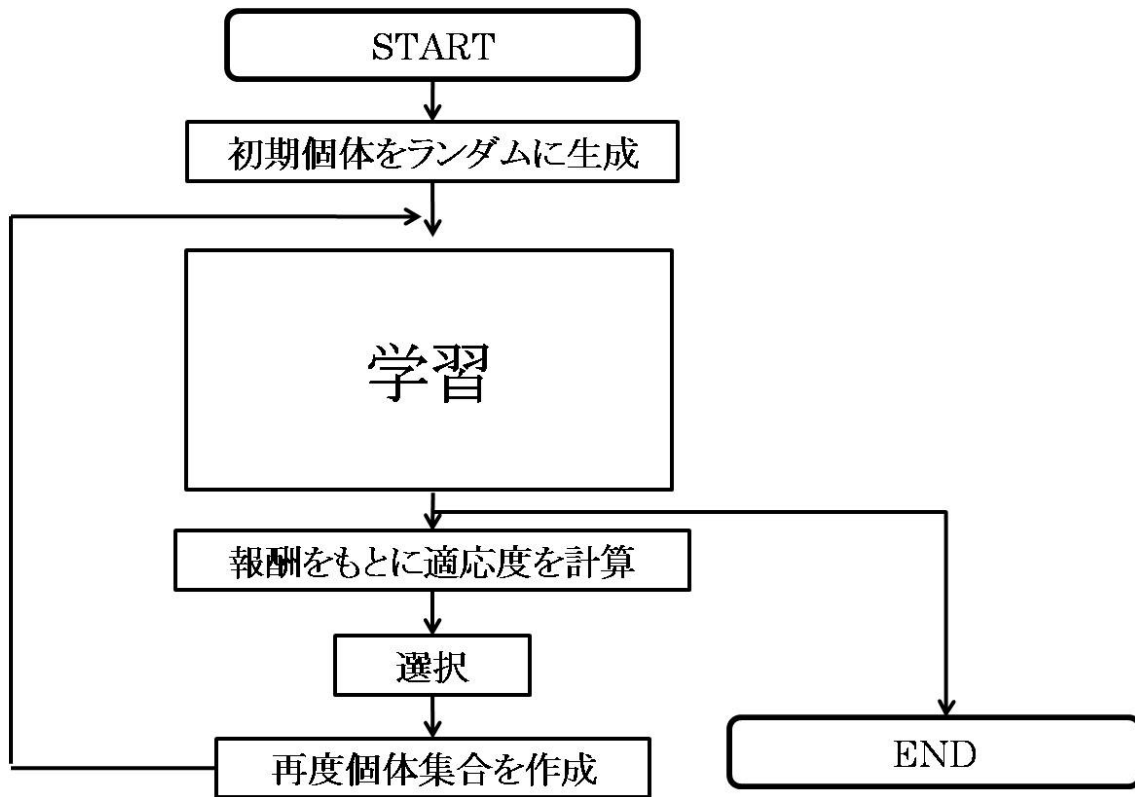


Fig.10 進化計算のフローチャート

このフローチャートに基づき、進化計算を行っていきます。まず、初期個体を生成し遺伝子を状態行動空間に変換、Q-tableの作成、Q学習を行う。そして、そこで得られた報酬を基に適応度を計算し、自然淘汰、交叉や突然変異により再度個体を生成する。

これを繰り返し行い、指定した回数行われたら終了となる。

8. 進化による逃避エージェントの学習機構の獲得(GA)

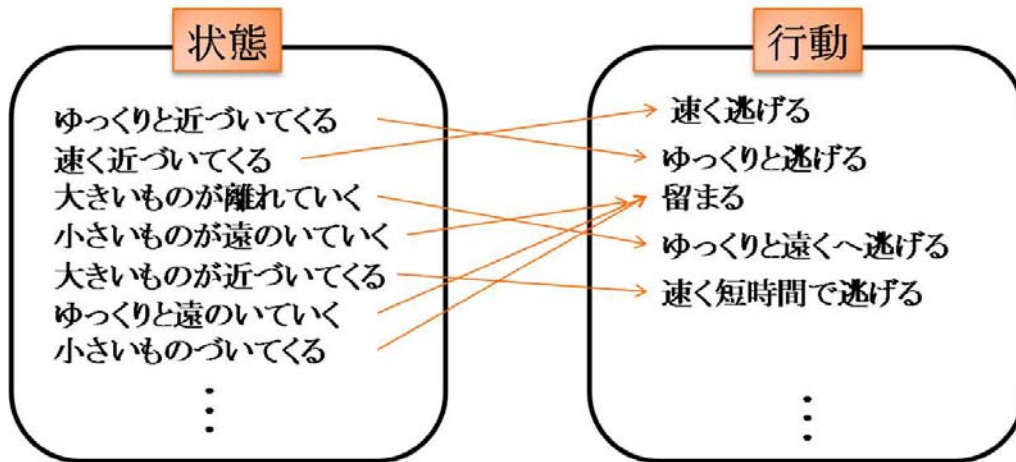


Fig.11 状態と行動の例

本研究は、逃避エージェントの状態と行動の学習をおこなっている。状態と行動の関係では Fig.8 のようなことが考えられ、状態と行動の組み合わせが多く考えられます。

たとえば、「ゆっくりと近づいてくる」という状態の場合には、「ゆっくりと逃げる」という行動をとり、また「大きいものが離れていく」という状態に対しては、「ゆっくりと遠くへ逃げる」という行動をとるということである。

そこで、この状態と行動を進化させることによって状態行動空間を縮退させる。これにより、学習量が減少する。この状態と行動を表すためにそれらを遺伝子で表わし、その遺伝子を用いて状態式と行動式を生成する。この状態式と行動式については章で詳しく説明する。

9. 実世界を用いた処理の実現

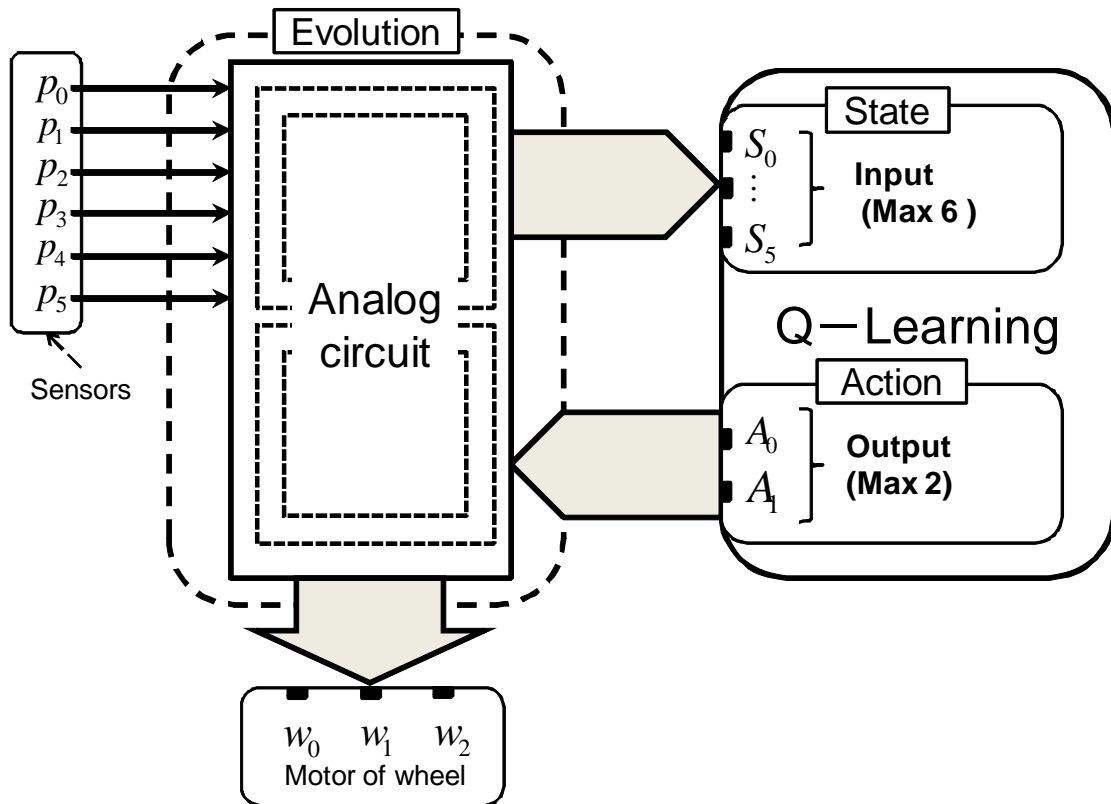


Fig.12 学習と進化の組織図

Fig.9 は学習と進化の関係を表しています。これは、入力として光素子センサの値を用いる。その入力値は、出力を p_0 , p_1 , p_2 , 出力の微分値を p_3 , p_4 , p_5 である。

この6つの入力値から状態式 S_n を最大で6つ, 行動式 A_n も同様に6つ生成する。しかし, 実際に出される行動式は最大で2つである。そして, この行動式は, A_n はホイールの出力を表しており, 各ホイールの出力 w_0 , w_1 , w_2 で表す。

9.1. 遺伝子

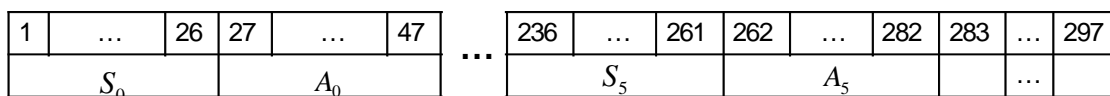


Fig.13 遺伝子の構成

遺伝子は、Fig.10のように設定されており、状態式 S_n と行動式 A_n が交互に生成されるようになっている。また、遺伝子の長さは、297個ある。

9.2. 状態構成回路

状態は、アナログ回路で構成すると考えこの回路を進化により獲得する。状態を構成する状態式 S_n は、遺伝子の値と光素子センサの値を用いた四則演算により構成する。 u_j は状態の項を表し、 $j(j=1,2,3,4,5,6)$ は項のまとまりを表す番号である。

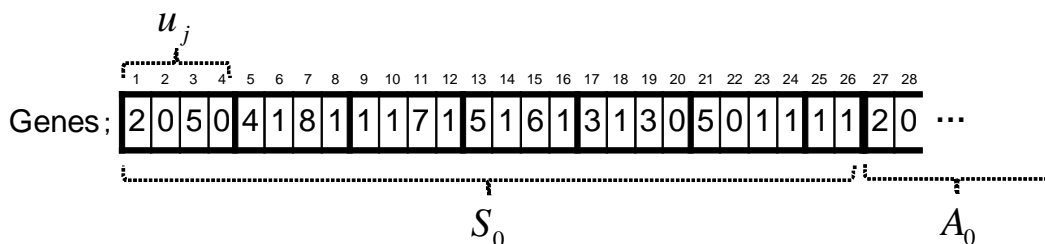


Fig.14 状態式の遺伝子構成

その具体的な生成方法を以下に示す。

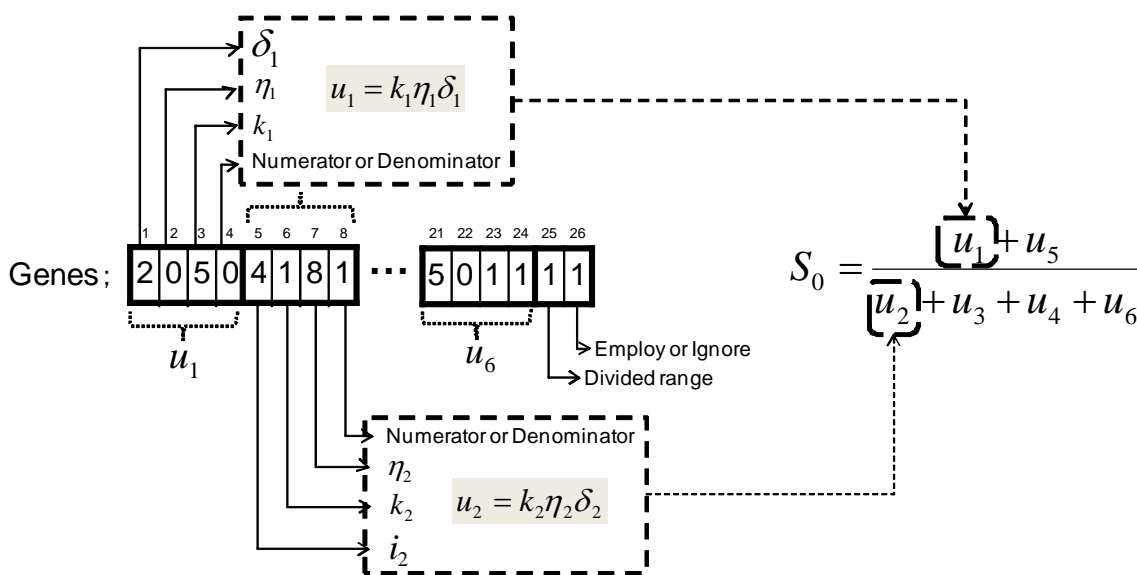


Fig.15 状態式の作成概要図

Fig.8 は状態式の作成概要図を表している。Fig.7 と Fig.8 より状態を構成する項は、式(6)で表される。

$$u_j = k_j \eta_j \delta_j \quad (6)$$

k_j : 増幅倍率, η_j : 増幅倍率の正負, δ_j : 光素子センサの出力値

それぞれの遺伝子型と表現型の対応関係を以下の表に示している.

Table.1 k_j の値

	k_j									
Genotype	0	1	2	3	4	5	6	7	8	9
Phenotype	0	1	2	3	4	5	6	7	8	9

Table.2 η_j の値

	η_j	
Genotype	0	1
Phenotype	1	-1

Table.3 δ_i の値

	δ_i					
Genotype	0	1	2	3	4	5
Phenotype	p_0	p_1	p_2	p_3	p_4	p_5

また, 分母か分子を決定する遺伝子型と表現型の対応関係を表したものを table.4 に示す.

Table.4 分母分子の選択

	4 th gene	
Genotype	0	1
Phenotype	Numerator	Denominator

そして, 式(5)と table.1 から table.4 を用いて状態式を構成していく. ここでは, Fig.7 の遺伝子を用いて具体的に状態式を式(7)として示す.

$$S_0 = \frac{5p_2 - 3p_3}{8p_4 - 7p_1 - 6p_5 + p_5} \quad (7)$$

そして最後に, それぞれ分割幅の遺伝子型と表現型を table.5, 生成された状態式 S_n の使用の有無を表した遺伝子型と表現型を table.6 に示す.

Table.5 分割幅

	25 th gene		
Genotype	0	1	2
Phenotype	Division range : 0.1	Division range : 1	Division range : 10

Table.6 状態式の使用の有無

	26 th gene	
Genotype	0	1
Phenotype	Ignore	Employ

9.3. 行動構成回路

行動の遺伝子を Fig.9 に示す. 行動式 A_n は, 状態式 S_n と同様に, 遺伝子の値と光素子センサの値を用いた四則演算により構成する. 行動式 A_n の項 u_j とする. j は状態式と同様に光素子センサの出力番号を示している.

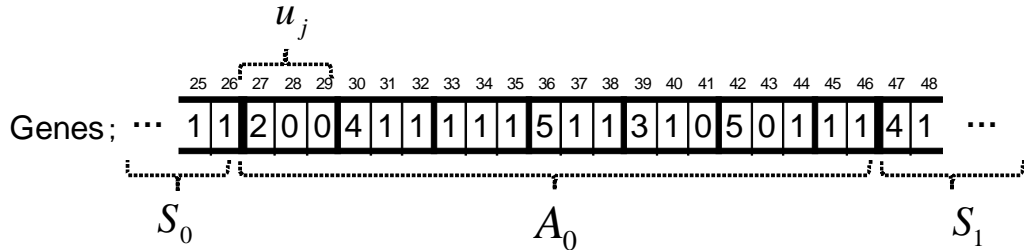


Fig.16 行動式の遺伝子構成

その具体的な生成方法を以下に示す.

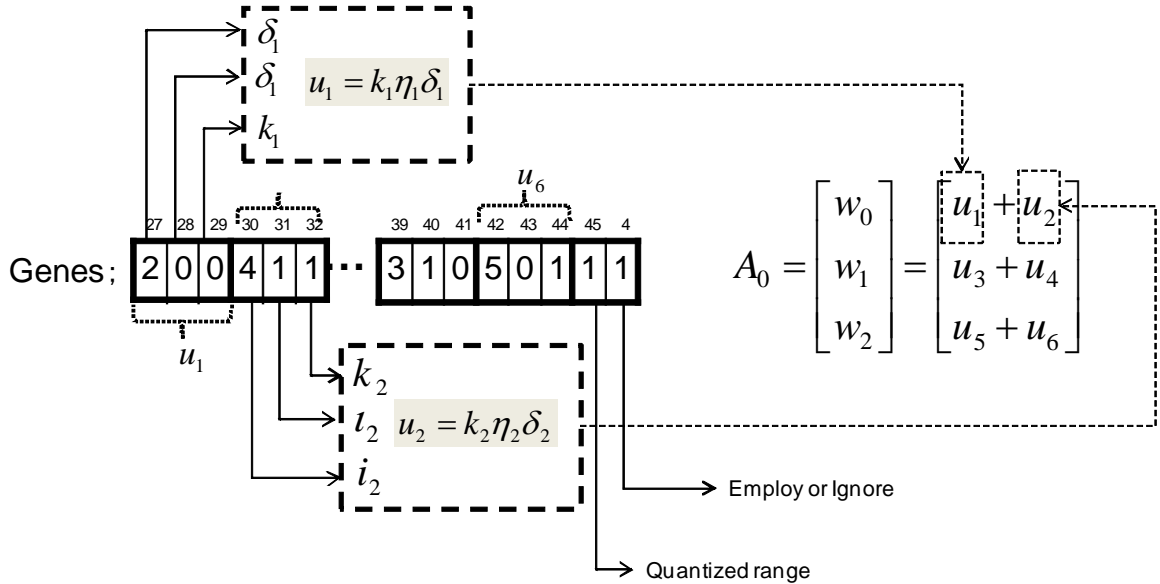


Fig.17 行動式の作成の概要図

Fig.9 と Fig.10 より, 行動式 A_n の項は, 状態式 S_n と同様に式(5)で表される.

また, 具体的な遺伝子の考え方は, 状態式 S_n と同様である. 以下には, 行動式 A_n の一般式を式(8)で表す. ここで w_0, w_1, w_2 は前述したとおり, 各ホイールの出力を示している.

$$A_n = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} u_1 + u_2 \\ u_3 + u_4 \\ u_5 + u_6 \end{bmatrix} \quad (8)$$

具体的に Fig.9 の遺伝子は式(8)を用いて計算すると式(9)で表される.

$$A_0 = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0p_2 + (-1)p_4 \\ (-1)p_1 + (-1)p_5 \\ 0p_5 + 1p_5 \end{bmatrix} = \begin{bmatrix} -p_4 \\ -p_1 - p_5 \\ p_5 \end{bmatrix} \quad (9)$$

最後に、行動式 A_n の分割幅の遺伝子型と表現型を table.7, 生成された行動式 A_n の使用の有無を表した遺伝子型と表現型を table.8 に示す.

Table.7 分割幅

	45 th gene		
Genotype	0	1	2
Phenotype	Division range : 0.1	Division range : 1	Division range : 10

Table.8 行動式 A_n の使用の有無

	46 th gene	
Genotype	0	1
Phenotype	Ignore	Employ

10. 学習機構の実現

本研究では強化学習として Q 学習[11][12]を用いる. Q 値の更新式を式(7)に示す.

$$Q(s,a) \leftarrow (1-\alpha) Q(s,a) + \alpha \{r(s,a) + \gamma \max_{a'} Q(s',a')\} \quad (10)$$

s =現在の状態 a =現在の行動 s' =次の状態

a' =次の行動 r =報酬 α =学習率 γ =割引率

また, 学習しやすさを適応度として式(8)で定義する. これは, 全試行で獲得された報酬の和を意味する.

$$\text{適応度} = \sum_{i=1}^{N_{\text{trial}}} \sum_{j=1}^{N_{\text{step}}} r_{ij} \quad (11)$$

N_{trial} =試行数 N_{step} =ステップ数 r_{ij} = i 試行 j ステップで得られた報酬

11. シミュレーション

11.1. シミュレーションの設定

逃亡ロボットをランダムに置き, 追跡ロボットは逃亡ロボットを中心とする半径 5m の円上にランダムに置く. 追跡ロボットの推進力の大きさ, 光の強さは試行毎に Table 1 内の値からランダムで選択され, 試行中は一定とする. 逃亡ロボットは 1s ごとに行動を決定する. 追跡ロボットはエネルギーを上限まで消費すると, 10s 間停止するものとする[13]. 1 試行は最大 300s とし, 10s 逃げ切る毎に正の報酬 100 を, 300s 逃げきるとさらに 1000 の報酬を与えられる. 逃亡ロボットが 300s 間追跡ロボットから逃げ切ることができれば, 次の試行に移る. 途中で試行を失敗した場合は, 負の報酬-100 を与え, 次の試行に移る.

Table.9 Q-table の GA パラメータ

ロボットの重量[kg]	10
タスク開始時の逃亡者と追跡者の距離[m]	5
追跡者の推進力[N]	5.2~5.4
逃亡者の推進力[N]	5
追跡者のライトの光量[cd]	10,20,30,40,50
追跡者のエネルギー量[J]	100

Table.10 エージェントの GA パラメータ

学習率	0.9	個体数	100
割引率	0.5	交叉確率	0.3
ランダムに行動選択する確率	0.1	突然変異確率	0.05

Table.11 報酬の設定

1試行の時間[s]	300
10s逃げ切る毎の報酬	100
捕まった時の報酬	-100

逃避エージェントが動いた場合のエネルギー計算

$$\text{エージェントが得る報酬} = \text{報酬} - \text{使ったエネルギー} \quad (12)$$

12. 結果

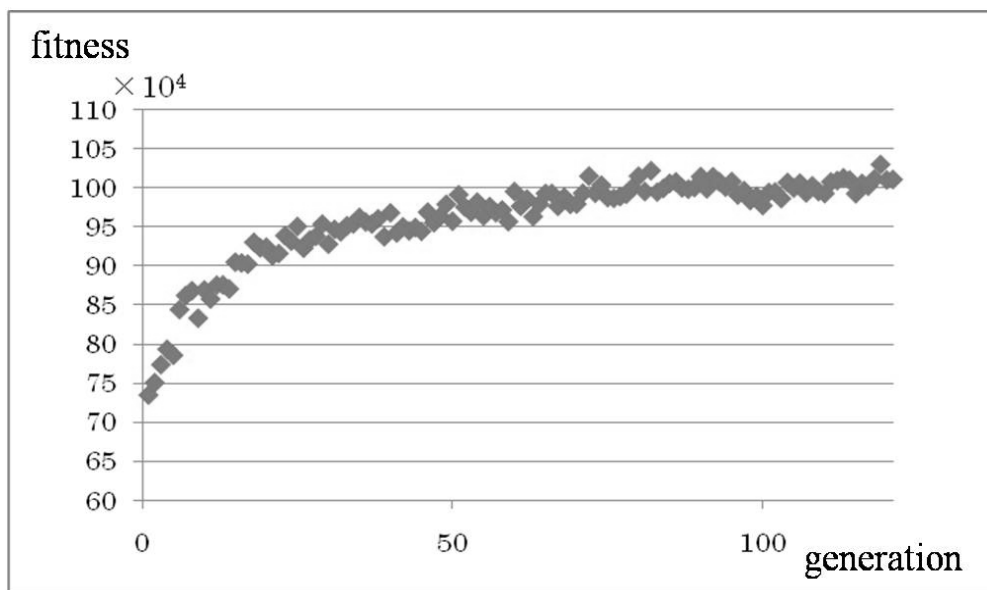


Fig.18 個体数 100 の適応度

この結果から、10 世代目の個体と 121 世代目の個体について具体的に示す。

Fig.12 と Fig.13 は、個体の 300 試行おこなったときに獲得した報酬の変化をグラフとして示したものである。

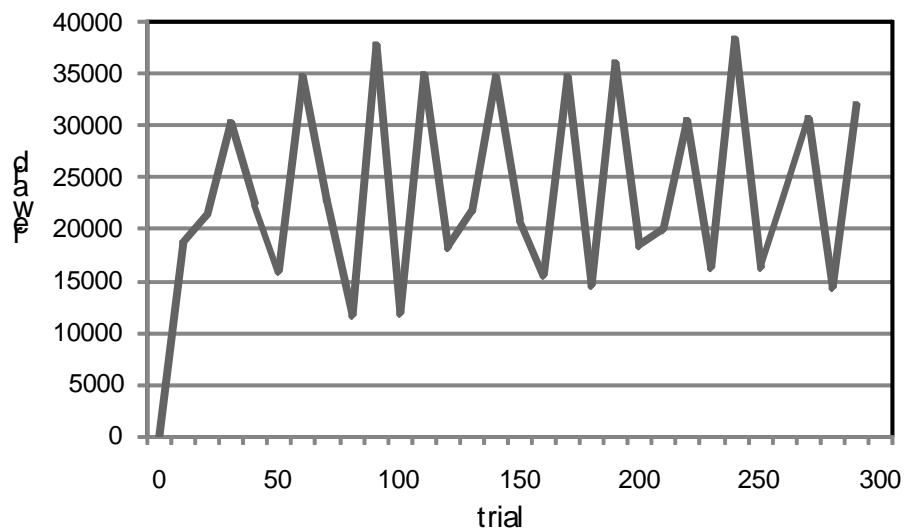


Fig.19 10 世代目の個体の試行数と報酬の関係

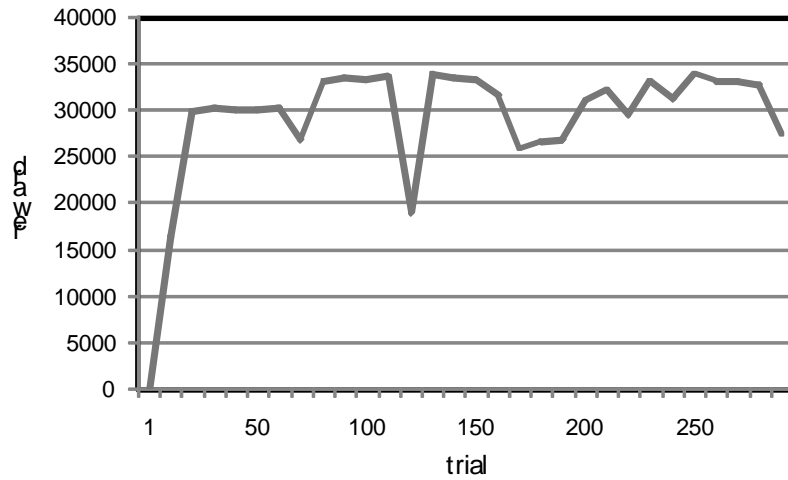


Fig.20 121 世代目の個体の試行数と報酬の関係

Fig.18 と Fig.20 を比較すると, 10 世代目の個体は全体として獲得している報酬が少なく, 進化がまだ十分ではないことが確認できる. また, 各試行について見ていくと獲得する報酬が安定していない. このことから進化が十分ではないことがうかがえる. 次に 121 世代目の個体は, Fig.12 に比べて全体として獲得している報酬が多い. 各試行について見ていくと, 最初は獲得している報酬が少ないが早い試行数で獲得する報酬が高くなり, 安定していることが確認できる. このことから, 進化が十分であるので学習にかかる時間が短くなっている.

以下の図は, 10 世代目と 121 世代目の個体の遺伝子を用いて動画を切り出したものを一定時間ごとに示したものである.

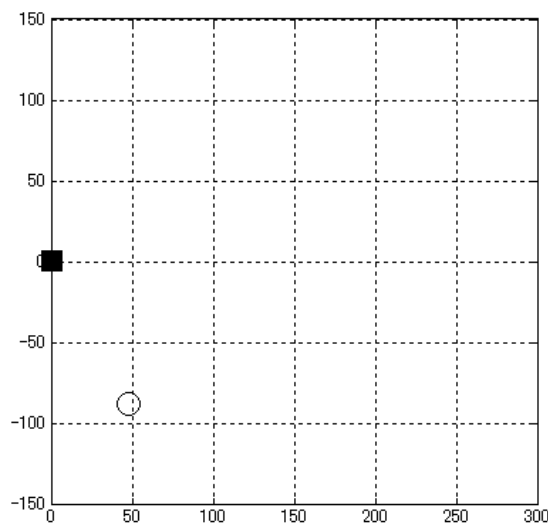


Fig.21 $t=0$ のとき

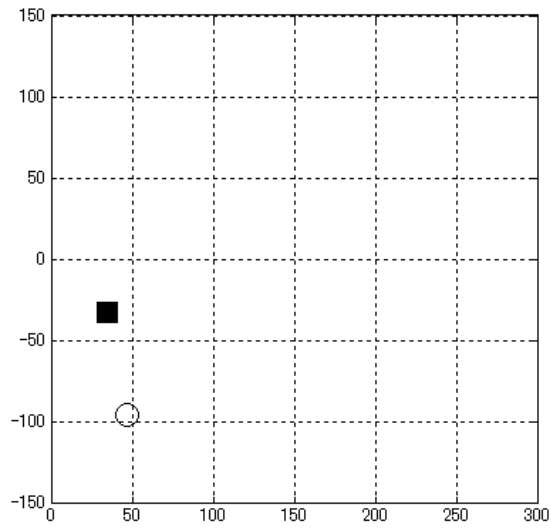


Fig.22 t=1

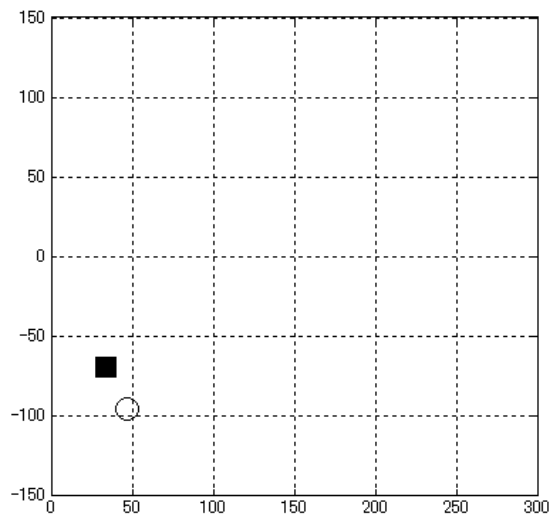


Fig.23 t=2 のとき

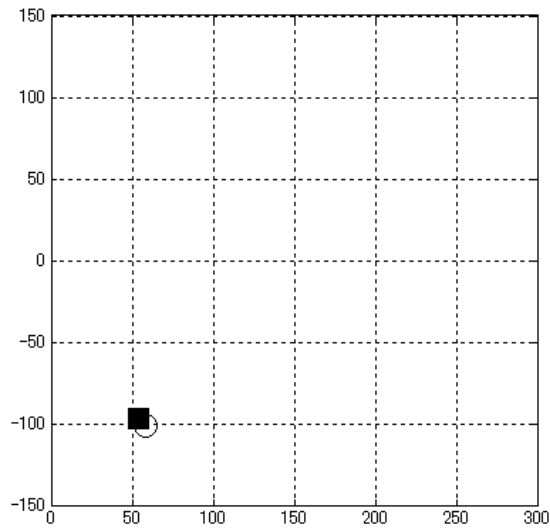


Fig.24 t=3

この Fig.21 から Fig.24 の動画のクリップより、逃避エージェントが追跡エージェントから逃げ切ることができていない。このことから、10 世代目でまだ十分に進化が行なわれていないために、学習に時間がかかり早い段階で学習を終えることができないことが確認できる。

次に 121 世代目の個体について以下に示す。

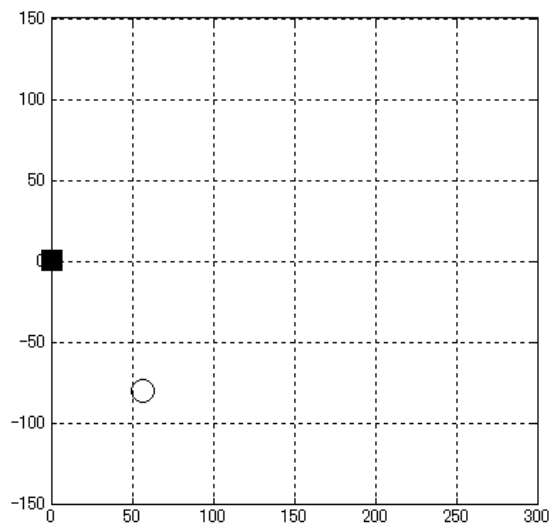


Fig.25 t=0

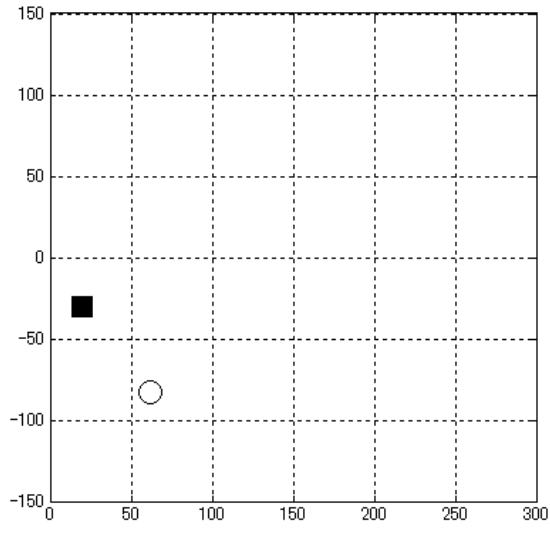


Fig.26 t=1

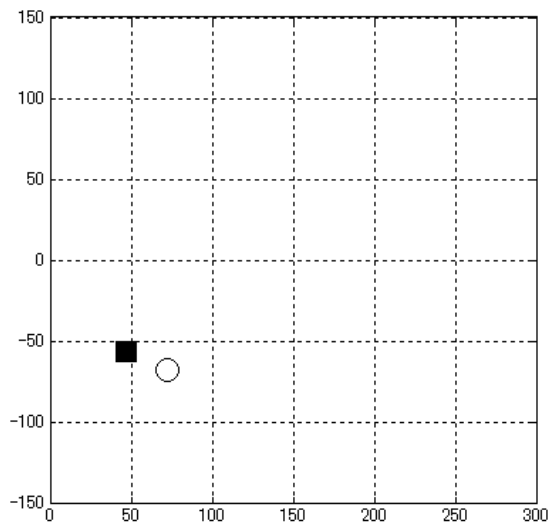


Fig.27 t=2

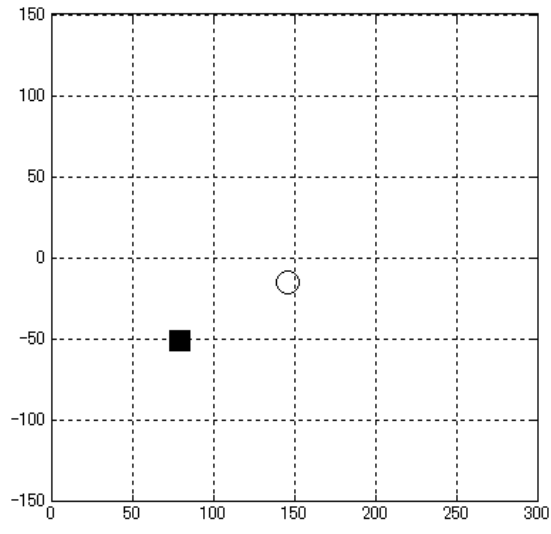


Fig.28 t=3

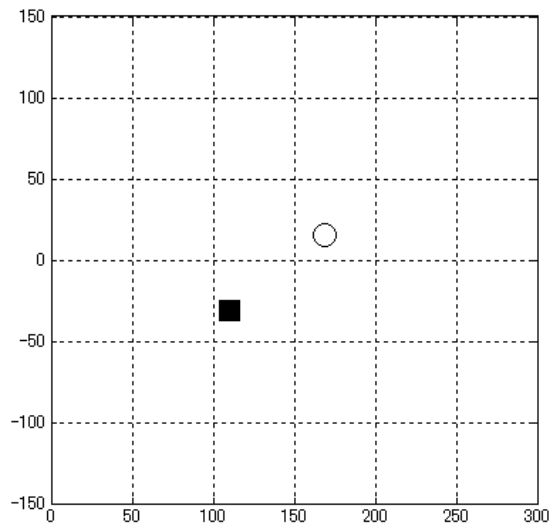


Fig.29 t=4

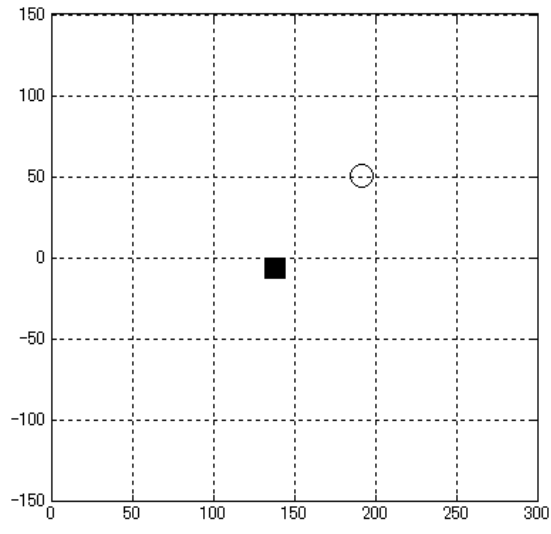


Fig.30 t=5

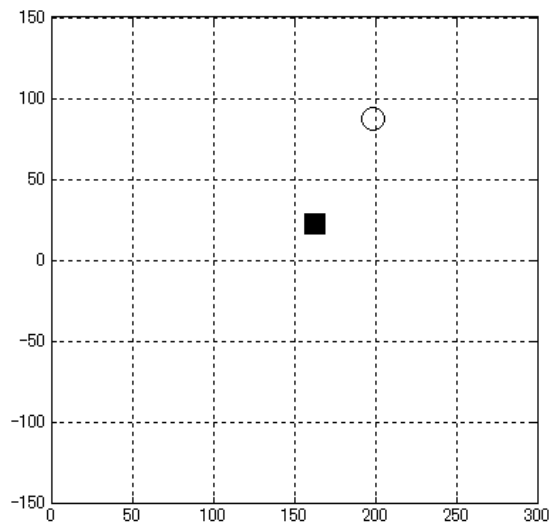


Fig.31 t=6

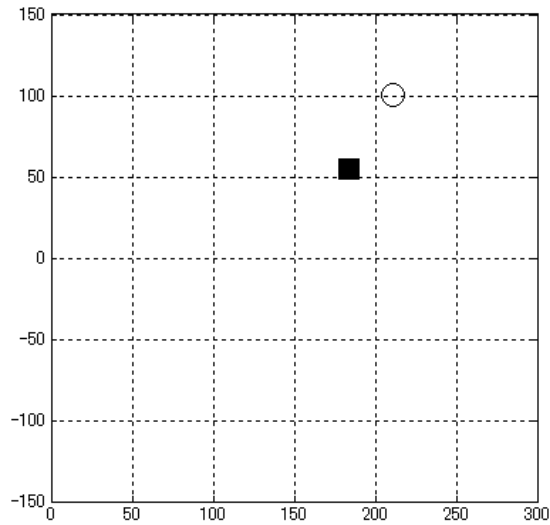


Fig.32 t=7

Fig.25 から Fig.31 の動画のクリップを見ると, うまく逃げ切れていることが確認できる. 特に, 逃避エージェントが追跡エージェントに追いつかれそうになってもしっかりと逃げ切ることが確認できる.

この世代において逃避エージェントが逃げ切ることができたのは, 十分に進化が行なわれているため早い段階で学習が完了しているからである. それは, Fig.11 ,Fig.12, Fig.13 からも確認できるとおり早い段階から高い報酬を獲得しているために適応度が 10 世代目よりも高くなっている.

10 世代目と 121 世代目の関係を表にまとめると以下のようなになる

Table.12 世代間比較

	学習速度	適応度
10 世代目	遅	低
120 世代目	早	高

13. 考察

次に、学習をほとんど行わなくても逃げ切ることが可能な理由を考察する。学習がほとんど必要のない世代の状態式を式(10)、式(11)として以下に示す。

$$S_{i_1} = \frac{6 \dot{p}_1 - p_1 - 3 \dot{p}_1}{-8 \dot{p}_3 - 2 \dot{p}_2} \quad (13)$$

$$S_{2i} = \frac{-3 \dot{p}_2 + 6 \dot{p}_1 - 7 \dot{p}_2}{-6 \dot{p}_2 + 6 \dot{p}_1 + 3 \dot{p}_1} \quad (14)$$

この2式に対して式(2)と式(3)から得られた式(12)、式(13)が得られる。

$$p_n = k \times \left(\frac{-\theta}{120} + 1 \right) \times I \quad (2)$$

$$I = \frac{a}{D^2} \quad (3)$$

$$P_n = \frac{a}{D^2} \quad (15)$$

$$\dot{P}_n = \frac{d}{dt} \left(\frac{a}{D^2} \right) \quad (16)$$

そして、この2式を式(10)、式(11)に代入すると式(14)、式(15)のようになる。

$$S_{i_1} = \frac{6 \dot{p}_1 - p_1 - 3 \dot{p}_1}{-8 \dot{p}_3 - 2 \dot{p}_2} = K \frac{6 \frac{d}{dt} \left(\frac{a}{D^2} \right) - \left(\frac{a}{D^2} \right) - 3 \frac{d}{dt} \left(\frac{a}{D^2} \right)}{-8 \frac{d}{dt} \left(\frac{a}{dD^2} \right) - 2 \frac{d}{dt} \left(\frac{a}{D^2} \right)} = K \cdot \frac{D}{\dot{D}} \quad (17)$$

$$S_{2'} = \frac{-3 \dot{p}_2 + 6 \dot{p}_1 - 7 \dot{p}_2}{-6 \dot{p}_2 + 6 \dot{p}_1 + 3 \dot{p}_1} = K \frac{-3 \left(\frac{a}{D^2} \right) + 6 \frac{d}{dt} \left(\frac{a}{D^2} \right) - 7 \frac{d}{dt} \left(\frac{a}{D^2} \right)}{-6 \frac{d}{dt} \left(\frac{a}{dD^2} \right) + 6 \frac{d}{dt} \left(\frac{a}{D^2} \right) + 3 \frac{d}{dt} \left(\frac{a}{dD^2} \right)} = K \cdot \frac{D}{\dot{D}} \quad (18)$$

ここで、 K は定数で D は追跡エージェントと逃避エージェントの距離を表し、 \dot{D} は2つのエージェントの相対速度を表している。この2式を検討すると生態心理学における接触までの残り時間を表す式と等価であることが確認できる。これは、121世代目の個体の動画を切り出した Fig.25 から Fig.32 からも、逃避エージェントが逃げるタイミングを獲得していることがあらわされている。

つまり、進化の過程において実世界の性質を取り入れることで逃げるタイミングを獲得していると考えられる。

進化前と進化後をより具体的に示すと、以下のとおりになる。

進化前は、Fig.33 で表される。

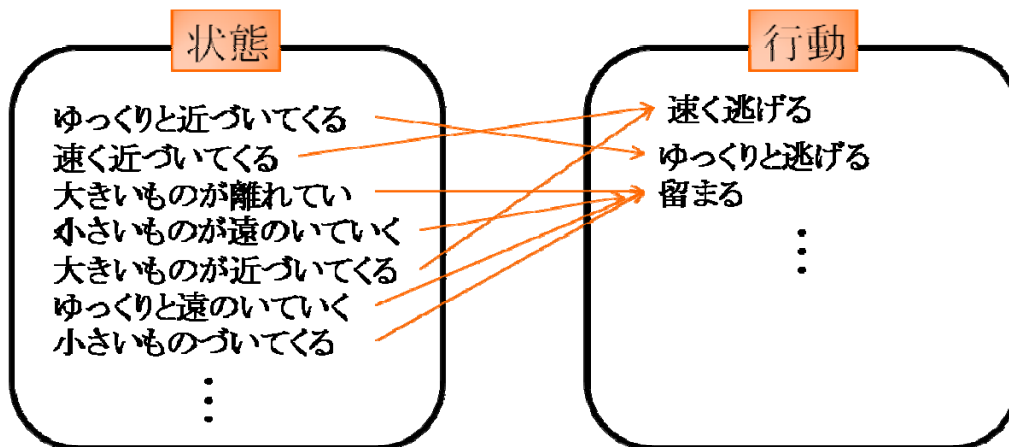


Fig.33 進化前の状態と行動

進化後の状態と行動は、Fig.34 のように示される。

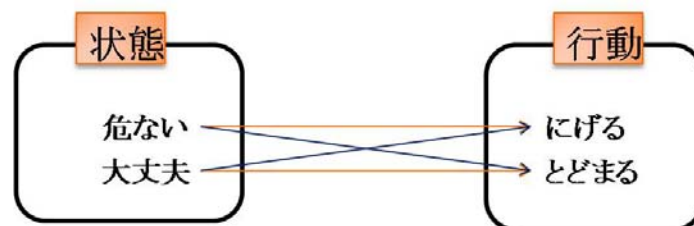


Fig.34 進化後の状態と行動

Fig.33 と Fig.34 を比較して、学習量が減少していることが確認できる。このことから Baldwin 効果における実世界の性質の役割は、学習量の減少にあることが確認された。

14. おわりに

獲得形質遺伝の法則は、親が学習した内容が遺伝によって遺伝子が書き換えられ、子に遺伝するという法則であった。そして、その子は、親が学習によって獲得したタスクを先天的に行うことが可能である。

また、本研究で提案した獲得形質遺伝の法則と同様の効果が表れる **Baldwin** 効果は、親が学習によって獲得したあるタスクが実世界の性質を取り込みながら進化することで、子は親と同じタスクを親よりも少ない学習量で獲得することが可能となる効果である。

そして、本研究では、**Baldwin** 効果における実世界の性質の役割について考え、例として実世界の性質であるニュートンの運動方程式と光の逆二乗則を用いた逃避行動を目的とするシミュレーションを行った。

その結果、実世界の性質を進化の過程で取り入れていくことでタイミングをとる接触までの残り時間 τ を獲得していることを確認した。

そして、それによって実世界の性質の役割が学習量の減少につながっていることを確認した。

参考文献

- [1] 金子隆一“生命誕生からラマルク, ダーウィン, 利己的遺伝子まで もっとわかる進化論”, 日本実業出版社, 1992
- [2] 駒井卓 著 “遺伝学に基づく生物の進化”, 培風舎, pp5, 1963
- [3] Baldwin, J, M “A new factor in evolution”, *America Naturalist*, Vol.30, pp. 441-451, 1896
- [4] G. E. Hinton, S. J. Nowlan, How learning can guide evolution, *Complex Systems*, Vol. 1, pp. 495-502, 1987.
- [5] Yusuke Watanabe, Reiji Suzuki, Takaya Arita, Language Evolution and the Baldwin Effect, *Artificial life and robotics*, Vol. 12, No. 1, pp. 65-69, 2008.
- [6] Reiji Suzuki, Takaya Arita, Repeated Occurrences of the Baldwin Effect Can Guide Evolution on Rugged Fitness Landscapes, *Proceedings of the 2007 IEEE Symposium on Artificial Life*, pp. 8-14, 2007.
- [7] Takaya Arita, Reiji Suzuki, Interactions between Learning and Evolution: The Outstanding Strategy Generated by the Baldwin Effect, *Proc. of artificial life VII*, pp. 196-205, 2000.
- [8] Reiji Suzuki, Takaya Arita, The Dynamic Changes in Roles of Learning through the Baldwin Effect, *Artificial life*, Vol. 13, No.1, pp. 31-43, 2007.
- [9] Reiji Suzuki, Takaya Arita, Simulations and Analyses for an Interaction between Learning and Evolution: The Baldwin Effect in the Iterated Prisoner’s Dilemma, *Journal of the Japanese society of the artificial intelligence*, Vol. 15, No. 3, pp. 495-502, 2007 (in Japanese).
- [10] Reiji Suzuki, Takaya Arita, How do roles of learning change through the course of evolution? , *Journal of information processing society of Japan*, Vol. 47, No. SIG 1, pp. 101-110, 2006 (in Japanese).
- [11] C.J.C.H WATKINS:”Technical Note: Q-Learning”, *Machine Learning*, Vol.8, pp.279-292, 1992
- [12] C. J. C. H. Watkins, “Learning from delayed rewards”, PhD thesis, King's College, University of Cambridge, May (1989).
- [13] 藤井宏行, 伊藤一之, 五福明夫, 実世界を用いた情報抽象化による実学習ー受動知能の進化的獲得ー, 第24回日本ロボット学会学術講演会予稿集, 講演番号 2D11, 2006

謝辞

3年間、伊藤先生には数多くの指導をしていただきとても感謝しております。

日本ロボット学会、A-ROBやSICEでは、最後の最後までご指導いただきました。文章の構成の仕方、書き方などの基礎から研究をする上でのものの考え方など多くのことを学ばせて頂きました。

また、就職活動においてもなかなか決まらず、心配していただいたこともありました。先生にご指導いただいた3年間で私は、かなりの成長を遂げることができました。感謝してもしきれません。

ありがとうございました。