多目的ＧＡを用いたロボットの
行動決定ネットワークの進化的獲得

東原 智幸，渥美 雅保

創価大学工学部情報システム学科

概要

近年，ロボットの行動決定ニューラルネットワークの構造および結合荷重を，遺伝的アルゴリズム(GA)を用いて進化的に獲得する研究が行われている．しかし，複雑な環境に適応可能なネットワークを進化的に獲得するためには，進化の方向の鍵を握る適合度関数を適切に設定する必要がある点，また，ある環境に特化した適合度関数を用いる場合，その環境に対して過学習の状態となって汎化性が下がり，別の環境への適応能力が低下する点が問題となりうる．そこで本論文では，複数の目的関数を適合度関数として持つロボットが，それらに対するパレート最適解を与える行動決定ネットワークを進化的に獲得するための方法として，多目的遺伝的アルゴリズム MOGA をすべての目的をバランスよく達成することに重点をおいて拡張した GA を提案する．そして，本方法が上記の問題の解決に貢献することを経路追従タスク，及びゴミ拾いタスクを用いて示す．

Evolutionary Acquisition of Sensory-Action Network of Mobile Robot using Multiobjective Genetic Algorithm

Tomoyuki Higashihara，Masayasu Atsumi

Dept. of Information Systems Science, Faculty of Eng., Soka Univ.

Abstract

Recently, there have been a lot of studies about evolutionary acquisition of weights and/or structure of sensory-action neural network for mobile robots. In these studies, two problems should be solved to acquire well-evolved neural networks that are able to adapt themselves to complex environments. First one is the design problem of a suitable fitness function for a given environment, that is necessary to make robots adapt to the given environment. Second one is the over-fitting problem to a specific environment, that make adaptation capability of robots lack for other environments. In this paper, we propose the multiobjective genetic algorithm, that extends Fonseca's MOGA, to genetically acquire sensory-action network of mobile robot with multiple objectives as a set of fitness functions. The extension of MOGA is focused on evolving pareto-optimal solutions with balanced achievement of multiple objectives. We present how contributes our extended MOGA approach to solve the above problems by simulations using a corridor following task and a garbage collecting task.

情報処理学会知能と複雑系研究会研究報告No.111, 98-ICS-111, pp.1-6 (1998).

1　はじめに

ロボットの行動決定ニューラルネットワークの構造および結合荷重を，遺伝的アルゴリズム(GA)を用いて進化的に獲得する研究が行われている．例えば，[Law 94]の壁沿い行動のためのネットワークの獲得，[Nolfi 96]によるゴミ拾いロボットの行動決定ネットワークの獲得などがある．しかし，複雑な環境に適応可能なネットワークを進化的に獲得するためには，進化の方向の鍵を握る適合度関数を適切に設定する必要がある点，また，ある環境に特化した適合度関数を用いる場合，その環境に対して過学習の状態となって汎化性が下がり，別の環境への適応能力が低下する点が問題となりうる．そこで本論文では，複数の目的関数を適合度関数として持つロボットが，それらに対するパレート最適解を与える行動決定ネットワークを進化的に獲得するための方法として，多目的遺伝的アルゴリズム MOGA をすべての目的をバランスよく達成することに重点をおいて拡張したGA(以下，拡張多目的GA)を提案する．そして，本方法が上記の問題の解決に貢献することを経路追従タスク，及びゴミ拾いタスクを用いて示す．

2　拡張多目的ＧＡに基づくロボットの行動決定ネットワーク獲得モデル

本モデルは，ロボットの行動決定ネットワークを拡張多目的GAを用いて進化させるエージェント集団から構成される．

2.1　多目的ＧＡ

多くの競合する目的の最適化問題は，それら多目的をある有用な関数形にスカラー化することにより，解くことができる．しかし多くの場合において，スカラー化関数は既知ではなく，スカラー化のためには，設計者の問題に対する洞察が必要とされる．多目的GAは，多目的最適化問題に対するパレート最適解の集合を，GAの枠組みで求めるための方法である．具体的な手法としては，［Goldberg 89］で，従来のランキング選択を，他の解に支配されていないという関係に基づくランキング選択に拡張して，その選択のもとで遺伝的に解く方法が提案されている．また，［Fonseca 93］は，Goldberg のこのような考えをさらに拡張した，多目的ランキング(ある世代における個体群の中の個体XiがPi個の個体に支配されているときのXiの多目的ランキングをPi+1とする)を用いた多目的GA(MOGA)と呼ばれる方法を提案した．本研究では，この多目的ランキングを拡張した拡張多目的GAを用いることとする．

2.2　エージェントの構成

エージェントの構成を図１に示す．各エージェントは，ある範囲内にある障害物までの距離を測定する赤外線，ソナーセンサーにより計測されたデータをもとに行動決定を行う．エージェントはNomadロボットであり，その行動決定ネットワークは入力層，中間層，出力層のフィードフォワード3層構造になっている．入力は赤外線，ソナーセンサーの値を0～1 に正規化しネットワークに入力する．出力は，ネットワークが前進速度0～1，回転角度-0.5～0.5を出力し，実行部でそれぞれ，0～15インチ/second，-90°～+90°に変換されて行動を行う．

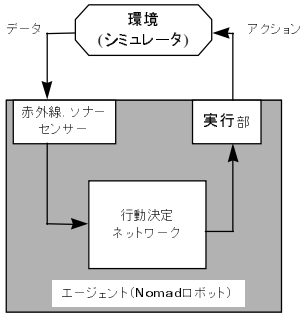


図１ エージェントの構成

2.3　行動決定ネットワークの染色体表現と遺伝操作

ネットワークをGAで扱うために，Node-based Encodingを使う．図2は，簡単な3層構造ネットワークの表現型と染色体表現(番号の列)を示したものである．

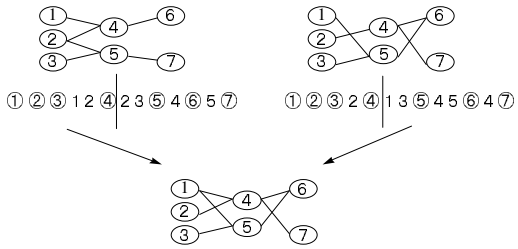


図2 ネットワークの表現型と染色体表現及び交叉の例

交叉は，２個体から同じ交叉点ノードを選び，それぞれの個体から，交叉点の前の部分と後ろの部分を受け取り新しい個体を生成する．突然変異は，リンクの削除または付加を行う．

2.4　拡張多目的ＧＡに基づく行動決定ネットワークの進化

Fonsecaらの多目的GAは，パレート最適解を一様に求めることを目的としている．その結果，ある目的に対する評価だけが高い解も，それが非支配解であれば，高いランキングが与えられる．我々の多目的GAは，すべての目的をバランスよく達成する解を求めることを目的とする．そのために，

拡張多目的ランキングでは，多目的のバランスを規定する制約と，多目的達成水準を規定する制約を導入する．前者を「バランス制約」，後者を「区切りランク」と呼ぶ．

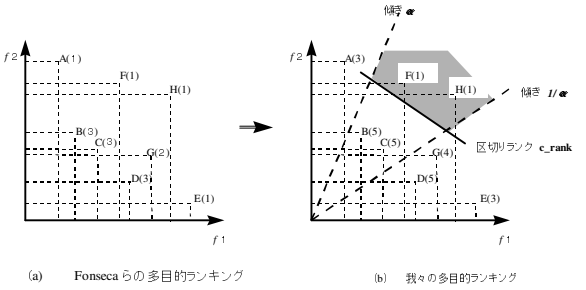
拡張多目的ランキングに基づく進化の方法を次に示す．

step 1 各エージェント毎に評価を行う．

step 2 拡張多目的ランキングを決める．(2 目的最大化問題(f_1, f_2)の多目的ランキングを図３にて説明する)

(1) まず，図３－(a)のようにFonseca らの多目的ランキング付けを行う．このランクを r_x とする．

(2) 図３－(b)のようにバランス制約を与える傾き α ～$1/\alpha$ の領域にある個体の中で，区切りランク ***c_rank***(図３では *c_rank*=2)より小さい個体群(図３では，点F，点H)は，ランクをそのままとし，その他の個体のランクを，r_x+*c_rank* とする．



(a) Fonseca らの多目的ランキング　　(b) 我々の多目的ランキング

図3 拡張多目的ランキング

step 3 拡張多目的ランキングの結果に基づき遺伝操作を行う．

遺伝操作の戦略として，２つの戦略を考える．

(戦略1) 全ての個体でランダムに単純な交叉を行い，その後，突然変異を行う．

(戦略2) ランクによる集団分割を行い，上位ランク集団は無条件で残すエリート選択，下位ランク集団はエリートとの交叉後，突然変異を行う．

行動決定ネットワークの進化は，*step3*の交叉，突然変異に2.3 の方法を用いることによりなされる．

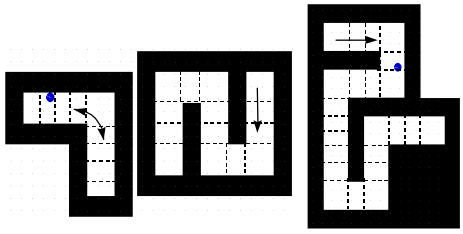
経路追従タスクでは戦略1，ゴミ拾いタスクでは戦略2 を使う．

3　タスクについて

本研究では，経路追従問題とゴミ拾いタスクの２つのタスクを使い実験を行う．

3.1　経路追従タスク

経路追従タスク(図4，●はロボット)は，与えられた経路を，壁にぶつかることなくできるだけ遠くまで移動するタスクである．実験では，トレーニング環境で進化を行ったエージェントが，テスト環境で良い行動(遠く，衝突は少なく)を取れるかどうかを確認する．トレーニング環境は，複数(左回りと右回り)の要素環境からなり，各々の環境でうまく行動することを目的とする．評価は，トレーニング環境，テスト環境の経路をいくつかの領域に分割し，その到達領域と衝突回数を使って行う．

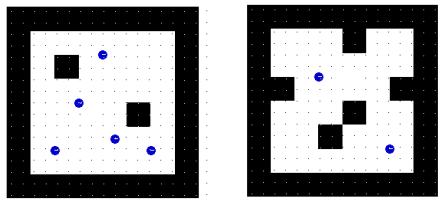


トレーニング環境　テスト環境Ⅰ　テスト環境Ⅱ

図4　経路追従タスク

3.2　ゴミ拾いタスク

ゴミ拾いタスク(図5)は，障害物にできるだけ衝突せずに，多くのゴミを拾うタスクである．環境は，障害物(■)，壁，ゴミ(●)からなる．図２の環境(ａ)は，障害物２個，ゴミ５個，環境(ｂ)は障害物５個，ゴミ２個の環境である．障害物にぶつからないこと，ゴミを多く拾うことが，それぞれ目的となる．



環境(ａ)　　環境(ｂ)

図5　ゴミ拾いタスク

4　経路追従タスクを用いた実験

4.1　実験の枠組み

トレーニング環境における右回り(f_R)と左回り(f_L)の２つの評価関数(適合度関数)は，それぞれ次の式で与えられる．

$$f = \frac{field}{MAX} \times \left(1 - \frac{\text{衝突回数}}{\text{行動回数}}\right) \times \left(2 - \frac{\text{step数}}{\text{行動回数}}\right)$$

field　：エージェントの到達領域

MAX　：最大領域数(９領域)

行動回数：ロボットの行動できる回数

step 数　：行動回数または最終領域到達回数

遺伝操作は，簡易交叉，突然変異，再生(ルーレット選択)を用いて，パラメータは人口 50 個体，世代数 30 世代，突然変異率 0.01 とした．エージェントの行動回数はトレーニング環境 15 回，テスト環境Ⅰ 40 回，テスト環境Ⅱ 80 回，行動決定ネットワークの入力ノード数は 18(赤外線 9，ソナー9)，中間ノード数 18，出力ノード数 2 とする．

4.2　実験結果

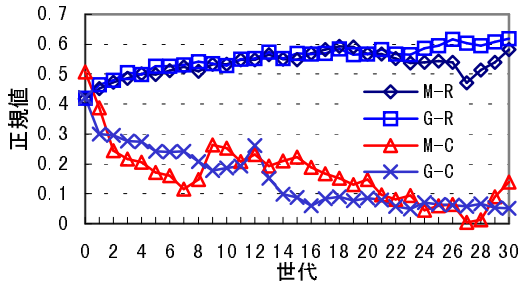


図6　拡張多目的ＧＡ(M)と

単純なＧＡ(G)の進化の比較

図6，表1はα=4，c_rank=１の拡張多目的GA(M)と単純なGA(G)の世代毎の進化の様子と３世代毎にトレーニングで評価の高い方から10個体のテストの結果を示したものである．単純なGAでは，適合度をf_R+f_Lとしている．Rは到達領域番号（トレーニング環境では1~9，テスト環境Ⅰでは1~14，Ⅱでは1~23の範囲である），Cは衝突回数（トレーニング環境では0~14，テスト環境Ⅰでは0~39，Ⅱでは0~79の範囲である）を表す．図6の縦軸はR，Cを0~1に正規化した値であり，右回りと左回りの値の平均を表す．MとGを進化の点で比較すると，図6からRのグラフがほぼ重なっており，CではMもGも0に近づいているので，同様の進化を行っていると考えられる．表1の灰色の部分はよいパフォーマンスを示した個体(領域の約８割以上到達した個体)で，テストⅠ，Ⅱともに灰色になっている個体は，よりよい個体といえる．また，□で囲んでいる部分は，テストⅠ，Ⅱともに最終領域に達した個体で，Mでは２個体獲得とGより多く獲得できた．

4.3　考察

トレーニング環境においては，拡張多目的GAと単純なGAとでは，ほぼ同じ行動を獲得した．また，テスト環境でも，拡張多目的GAと単純なGAは，汎化性において同等であった．このタスクにおいては，単純なGAで複数の目的関数f_Rとf_Lのスカラー化をf_R+f_Lとで与えることは合理的であると考えられ，また，それは，拡張多目的GAでα=4とαを大きく設定する場合とほぼ同じ効果を生む．また，ほとんどの個体は，単に壁沿い行動を獲得しているのみのようである．これらが，拡張多目的GAに基づく方法が，単純なGAに基づく方法を有意に上回ることのできない原因と考えれられる．

5　ゴミ拾いタスクを用いた実験

5.1　実験の枠組み

ゴミ拾いタスクでは，図5の環境(ａ)，環境(ｂ)2つの環境のそれぞれで進化を行う．評価関数(適合度関数)としては，

f_1=(拾ったゴミの数)／(初期のゴミの数)

f_2＝１―(衝突回数)／(行動回数)

の2つを用いる．

実験では，次の4タイプの個体集団について比較を行う．集団1と集団2は，α＝2，c_rank=10としてf_1とf_2を使いランク付けを行ったあと，戦略2に従い遺伝操作を行う．

［集団1(Gx_3)］ランクのいい方から，1/3の個体は無条件で残し，残り2/3の個体群は無条件で残した1/3の個体群と１点交叉を行う．その後，交叉でできた個体を突然変異率0.001で突然変異させる．

［集団2(Gx_5)］ランクのいい方から，1/5の個体は無条件で残し，残り4/5の個体群は無条件で残した1/5の個体群と１点交叉を行う．その後，交叉でできた個体を突然変異率0.001で突然変異させる．

また，集団3と集団4は単純なGAを用いて進化させる集団で，交叉率0.9，世代ギャップ0.9，突然変異率0.001として，適合度は次のようにする．

［集団3(S1)］適合度を0.2×f_1+0.8×f_2とする．

［集団4(S2)］適合度を0.8×f_1+0.2×f_2とする．

GAパラメータは，人口50，世代20，エージェントの行動回数は100回，行動決定ネットワークの入力ノード数は9(赤外線4，ソナー5)，中間ノード数18，出力ノード数2と

表　１　　テスト環境Ⅰ、Ⅱにおけるシミュレーション結果

する.

5.2 実験結果

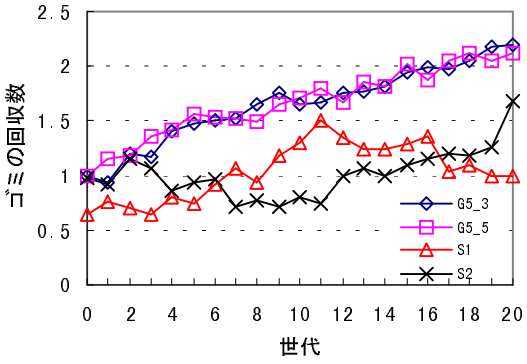


図7 環境(a)におけるゴミの平均回収数

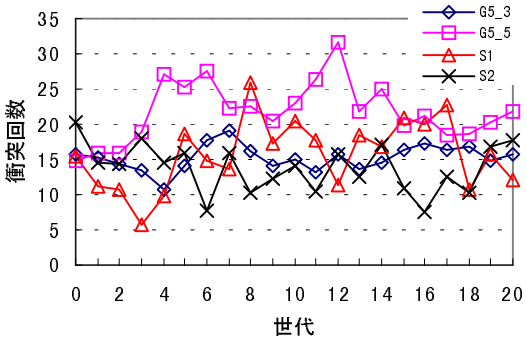


図8 環境(a)における平均衝突回数

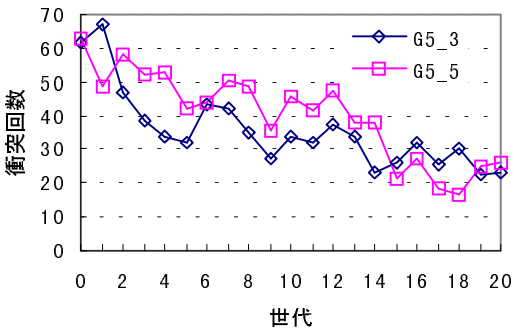


図9 環境(a)におけるゴミ1個当たりの衝突回数

図7, 図8は環境(a)における世代ごとのゴミの平均回収数と平均衝突回数を示したもので, G5_5 は集団2, G5_3 は集団1である. 図7からわかるように, G5_5, G5_3 ともにいくら衝突してもゴミ集めに集中すればよいS2より, ゴミを多く集めていることがわかる. 衝突回数では, G5_3 が衝突回避重視のS1と比べ, 悪くなっているとはいえない. 図9は, 環境(a)でのG5_5とG5_3の世代ごとのゴミ1個当たりの衝突回数を示したもので, いずれの戦略も衝突回数が減っている.

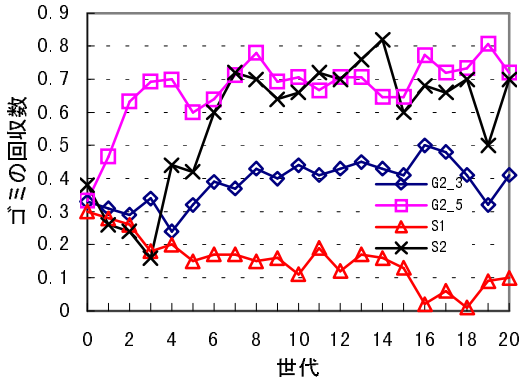


図10 環境(b)におけるゴミの平均回収数

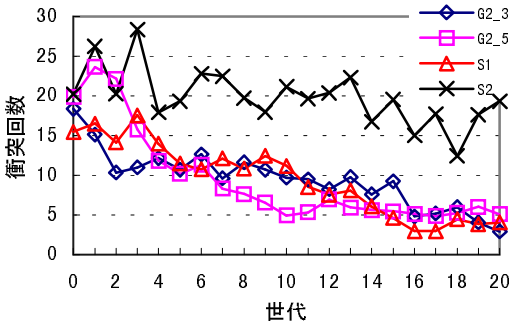


図11 環境(b)における平均衝突回数

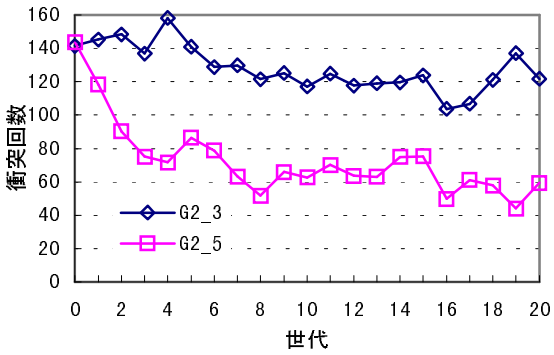
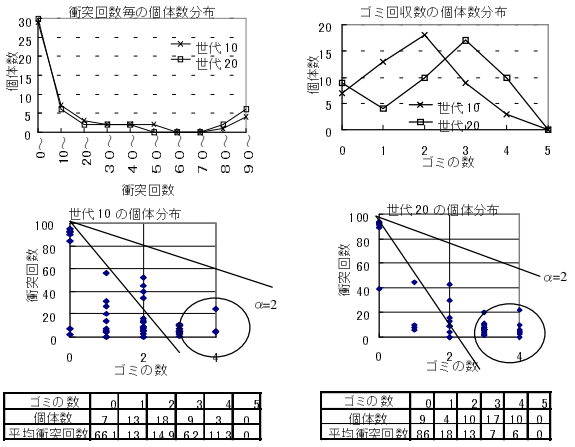


図12 環境(b)におけるゴミ1個当たりの衝突回数

図10, 図11は環境(b)における世代ごとのゴミの平均回収数と平均衝突回数を示したもので, G2_5は集団2, G2_3は集団1である. ゴミの回収数では, G2_5はゴミ回収数を

重視するS2と同程度の回収数を達成し，かつ進化の速度が速いことがわかる．衝突回数では，衝突回数を重視しないS2以外は，ほぼ順調に進化している．図12は，環境(b)でのG2_5とG2_3の世代ごとのゴミ1個当たりの衝突回数を示したもので，G2_5の方がよい結果が得られた．



ゴミの数	0	1	2	3	4	5
個体数	7	13	18	9	3	0
平均衝突回数	66.1	13	14.9	6.2	11.3	0

ゴミの数	0	1	2	3	4	5
個体数	9	4	10	17	10	0
平均衝突回数	86	18	13	7	6	0

図13　10世代と20世代の個体分布

図13は，環境(b)における10世代と20世代の個体の分布である(○は，次の世代に残る個体)．10世代はゴミを2個拾う個体が1番多いが，20世代では3個または4個ゴミを拾う個体が増えて進化がいい方向に進んでことが分かる．衝突回数の分布については，変化が見られない．

	環境(a) ⟶		環境(b)	
	ゴミの数	衝突回数	ゴミの数	衝突回数
S2	2.4	34.2	1.1	35.7
S1	2.7	39.2	1	43.8
G5_3	2.4	27	1.1	38.3
G5_5	2.3	30.6	0.7	36

	環境(b) ⟶		環境(a)	
	ゴミの数	衝突回数	ゴミの数	衝突回数
S2	0.7	38.1	1.7	52
S1	0.2	4.46	1.3	8.1
G2_3	0.7	19	1.9	16
G2_5	0.9	12.4	1.5	15.5

表2　他の環境でのテストの結果

表2は，環境(b)で進化を行ったベスト個体を環境(a)で，環境(a)で進化を行ったベスト個体を環境(b)でテストした結果の表である．テストでは，Gx_3が他の集団よりよい行動を行っている．

5.3　考察

環境(a)では，ゴミが多く障害物が少ないため，f2のウェイトを高くした適合度関数を持つ集団4(S2)が得意とする環境であり，環境(b)は，障害物が多いため，f1のウェイトを高くした適合度関数を持つ集団3(S1)が得意とする環境であると考えられる．実験の結果，図7，図8より，環境(a)においては，S1よりS2が若干であるがよい性能を示している．一方，環境(b)においては，S1はS2より衝突回数は，はるかに少なくなっているが，反面ゴミ回収に失敗している．これらは，複数の目的関数のスカラー化の難しさを示しているといえる．拡張多目的GAに基づく集団Gx_3，Gx_5をS1，S2と比較すると，図7，図8より環境(a)においては，特にゴミの回収数においてS1，S2を大きく上回り，また環境(b)においては，S2と同程度もしくは若干上回る性能を示していることがわかる．これらより，拡張多目的GAに基づく方法が，適合度関数の設定の難しさの解消に貢献するということができる．

表2より，Gx_3は，進化した環境においてS2のゴミ回収数と同じでいい結果を出している．また，Gx_3のテストでは，環境(b)においてS2と同程度，環境(a)においては最もいい結果になっている．これは，次の世代に残すエリートの中に環境に特化していない個体も入る可能性がGx_5より大きいためであると考えられる．このことから，汎化性を考えて進化を行う場合は，集団1で用いた方法が有効だと推測される．しかし，Gx_3は図10からもわかるように，エリートの中に環境に特化していない個体が残る可能性が高い分，進化を行う環境での結果があまりよくならない場合もありうる．

6　むすび

本論文では，拡張多目的GAによりロボットの行動決定ネットワークの構造を進化的に獲得する方法を提案した．そして，本方法が，複雑なタスクに対する適合度関数の適切な設定に役立ちうること，即ち，複雑なタスクに対するロボットの行動決定ネットワークの進化的獲得に有用となりうることを確かめた．また，新しい環境への適応能力が単純なGAで進化させるより，高くなりうることを確かめた．

今後の課題として，任意の環境に対して高い適応能力を備えた行動決定ネットワークを進化的に獲得する手法の開発を目指したい．

参考文献

[Goldberg 89]Goldberg, D. E.: *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison Wesley (1989).

[Fonseca 93] Fonseca, C.M. and Fleming, P.J.: Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization, Proceedings of the Fifth International Conference on Genetic Algorithms, pp.416−423 (1993).

[Nolfi 96] Nolfi, S.: Evolving Non-Trivial Behaviors on Real Robots: A Garbage Collecting Robots, Institute of Psychology, National Research Council, TR96-04(1996).

[Law 94] Law, D. Miikkukainen, R. : Grounding Robotic Control with Genetic Neural Networks, Dept. of Computer Sci. The University of Texas, TRAI94-223(1994).