

# 実進化型2足歩行ロボットシステムへの先行評価の導入

Introducing Pre-evaluation into the Embodied-Evolution framework for a biped robot.

中井 淳一\*1      有田 隆也\*2  
Junichi Nakai      Takaya Arita

\*1 株式会社ビジネスデザイン研究所  
Business Design Laboratory Co., Ltd

\*2 名古屋大学大学院情報科学研究科  
Department of Complex Systems Science, Graduate School of Information Science, Nagoya University

“Embodied Evolution (EE)” is a methodology in evolutionary robotics, in which, without simulations on a host computer, real robots evolve based on the interactions with actual environment. However, we had to accept robot behavior with low fitness especially in the early generations when adopting the EE framework. We introduced pre-evaluation into the EE framework so as to restrain robot behavior whose fitness is predicted to be low. This paper reports on the introduction of pre-evaluation into the Embodied-Evolution framework for a biped robot in order to reduce the risk of falling.

## 1. はじめに

進化ロボティクス [Nolfi 2000] は進化計算によるロボット制御器の自動設計を目指す研究分野であり、近年のロボット応用範囲の広がりに応じて注目を集めている。従来は、ホスト計算機でのシミュレーションによる適応度評価に基づいて制御器を進化させ、それをロボットに転送して動作させる手法が主流であったが、シミュレーション環境に適応した制御が実世界のロボットで必ずしも同様に有効とは限らず、実環境とのギャップを生みがちであり、また、環境やタスクを変更するたびに、シミュレータを変更する必要がある。

Watson らは実進化型ロボットシステムを提案し [Watson 2002]、ホスト計算機なしに実環境でのロボット行動により適応度評価することで、問題解決を狙った (図 1)。制御器を構成するニューラルネットワークの結合重みを表す遺伝子セットはロボット行動によって評価され、他ロボットに遭遇したときに適応度に応じた確率で送信されることで、ロボットの行動は進化する。ただし、彼らの方式では、ロボット間の遭遇頻度に進化が強く依存してしまう。

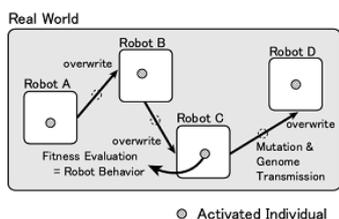


図 1: Watson らの実進化手法

臼井らは、各ロボットに複数の仮想個体を持たせ、各個体単体でも進化可能にすることで、ロボット間遭遇頻度に左右され

連絡先: 中井淳一, 株式会社ビジネスデザイン研究所, 〒460-0008 名古屋市中区栄 3-18-1 ナディアパーク デザインセンタービル 7F デザインラボ, nakai@business-design.co.jp, nakai@alife.cs.is.nagoya-u.ac.jp

ない実進化を実現した [Usui 2003] (図 2)。また、Elfwing らも臼井らと同様に複数の仮想個体をロボット内に持つ方式を提案しているが [Elfwing 2005]、汎用的なタスク実行を狙っておらず、生物的妥当性に重点を置いて設計されたものと考えられ、ロボット単体では進化することができず、ロボット遭遇回数に対する進化依存性は解決していない。

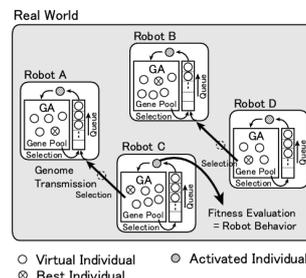


図 2: 臼井らの実進化手法

これらの研究では、ロボット行動だけで適応度評価するため、特に適応度の向上していない進化初期に弊害を生みやすい。また、時間的コストとロボットの疲労が増大する可能性があった。そこで、我々は、ロボット行動の前に先行評価し、低適応度とみなす個体による行動を回避することによる解決を狙っている [笠井 2007]。先行評価器はその精度を適応度として進化させ、制御器の進化系との共進化系を構成している。ただし、先行評価系の導入が制御器の進化に対して与える影響には未解明な点が多い。また、タスクによっては先行評価値が実評価値よりも高くなりすぎて進化が停滞するなど、この共進化系独自の問題があった。

そこで、我々は、複雑な挙動を示す笠井らのアーキテクチャを極めて単純化したモデルに変換した上で、性能評価を行った [中井 2008]。それにより、先行評価系を導入することで、基準適応度に達するまでに必要な実評価回数を削減可能なことを示した。また、先行評価系の環境個体の適正な進化のために、進化初期のランダム性に基づく高適応度の環境個体を遺伝子プールから徐々に排除して再評価を促すこと(「風化」)が重



価値とともに自己個体プールへ戻される．閾値  $K$  は自己個体プール内の自己個体の評価値の平均値に環境個体プール内の環境個体の適応度の平均値を掛けた値

$$K = \frac{(\sum_{i=1}^P F_{ri})(\sum_{i=1}^P F_{pi})}{P^2} \quad (10)$$

とする．先行評価値が閾値より大きい場合は実評価し，自己個体は実評価値と共に自己個体プールに戻し，環境個体は環境個体の適応度と共に環境個体プールに戻す．

偶発的な高適応度の自己個体や進化初期のランダム性による高適応度の環境個体がプール内に残るのを防ぐために「風化」が導入されている．これは，自己個体プール内の自己個体，環境個体プール内の環境個体の適応度を時間経過に応じて一定の割合  $w = 1\%$  で減少させるものである．

次に，自己個体プールからルーレット選択により 2 個体を選び，突然変異，交叉を行い，自己個体を作り，環境個体プールからも自己個体プールと同様の操作で，次に実行する環境個体を作る．突然変異率は 0.3 とした．この一連の流れを繰り返す．

#### 4. 評価実験

まず，先行評価導入前と後の自己個体プール内の平均実評価値の推移を図 5 に示す．同図より，平均実評価値が約 650 に到達するのに必要な実評価回数が先行評価導入前後で約 20 から約 10 に減少したことがわかる．これは，ロボットの疲労を軽減することだけでなく，先行評価の時間が実評価に比べて無視できることから，時間の削減をも意味している．

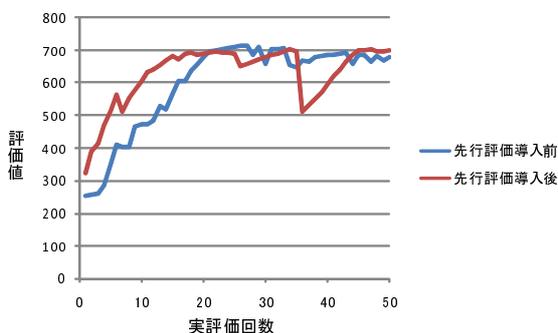


図 5: 実評価値の自己個体プール内平均の推移 (先行評価導入前後)

実評価値に関して，先行評価の導入前を図 6 に，導入後を図 7 に示す．両図には，さらに転倒した時点も示し，図 7 には，先行評価値，先行評価を行った時点と回数（実評価を回避したとき）も示す．導入前は実評価回数約 20 までに約 10 回転倒しているのに対し，導入後は実評価回数約 20 までに約 2 回しか転倒していない．さらに，導入前は徐々に実評価値が上昇しているのに対し，導入後は進化初期から高適応度の個体が多い．これは実評価回数約 20 の間に実評価を約 25 回避していることから，ロボットが転倒する恐れのある低適応度行動を先行評価系により回避しているためと思われる．今回のタスク設定では，転倒した場合としなかった場合の適応度の差が極端であるため，転倒する状況をうまく回避した場合，進化初期から実評価値が高くなる傾向がある．

図 8 に環境個体の適応度の推移を示す．全般に比較的高い値を維持して実環境をうまく再現しているように見える．進化

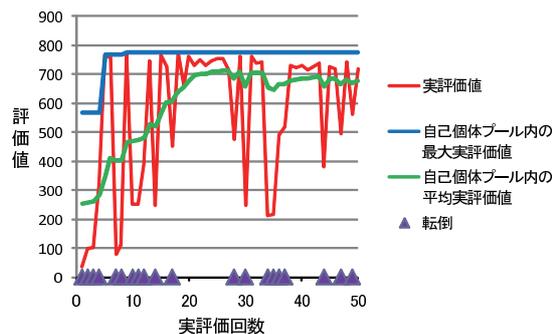


図 6: 評価値の推移 (先行評価導入前)

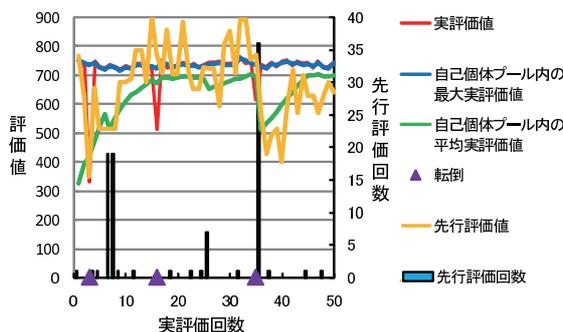


図 7: 評価値の推移 (先行評価導入後)

初期から高適応度となることが多く，ある程度上下を繰り返しているが，これは，実評価値と先行評価値の差を適応度としているので，特定の自己個体との適応度評価のみで高適応度となっても必ずしも環境を的確に表しているとは限らないということに起因すると考えられる．

また，自己個体の遺伝子データの推移を図 9 に，環境個体の遺伝子データの推移を図 10 に示す．より前に進むため，歩幅は，最も高い値である 0.05 が多くなる．歩幅を大きく保ったままの安定歩行のため，横に傾ける距離は約 0.02 に，高さは約 0.008 に収束している．低適応度行動となる個体の実評価を先行評価系が回避し，転倒する恐れのない比較的近い遺伝子データを適応度評価しているため，全体的に値が大きくは変化しない．ただし，その環境個体の遺伝子データ値に依存し続けず，自己個体の進化に従って変化しており，適切に共進化していると考えられる．

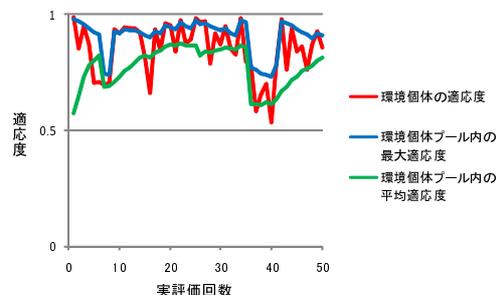


図 8: 環境個体の適応度の推移

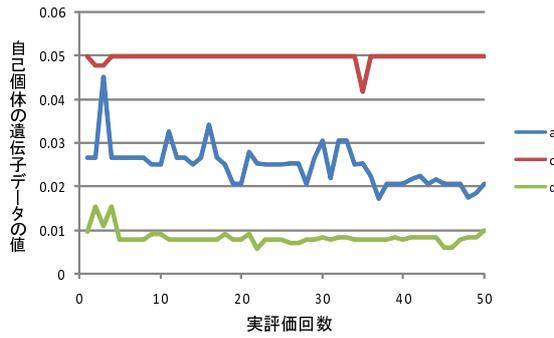


図 9: 自己個体の遺伝子データの推移

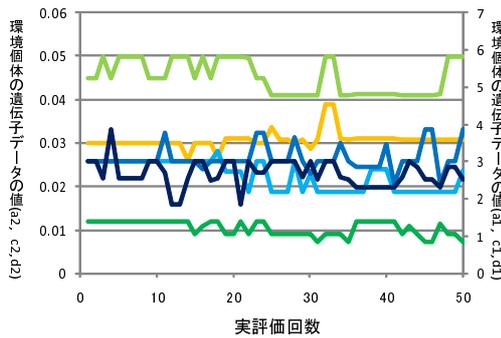


図 10: 環境個体の遺伝子データの推移

さらに、 $r = 18000$  のみ初期値を設定し、他の環境個体の遺伝子データ ( $a_1, a_2, c_1, c_2, d_1, d_2$ ) の初期値をまったくのランダムに生成した場合の環境個体の遺伝子データの推移を図 11 に、環境個体の適応度の推移を図 12 に示す。予想値を設定したときと比べて、値が大きく変動しているが、進化に従いある値に収束していることがわかる。環境個体の適応度もほぼ大きい値を保っており、環境個体が実進化系との共進化が実現されていることが分かる。

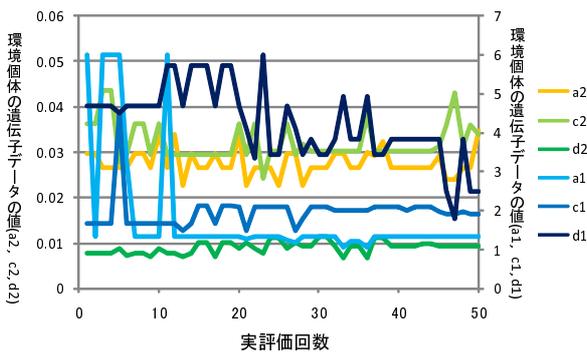


図 11: 環境個体の遺伝子データの推移 (初期値ランダム設定)

## 5. おわりに

実進化では、進化初期の低適応度個体もロボット行動で評価するため、収束に時間を要し、不測の事態も引き起こしかねない。そこで、ロボット行動の前に先行評価し、低適応度とみな

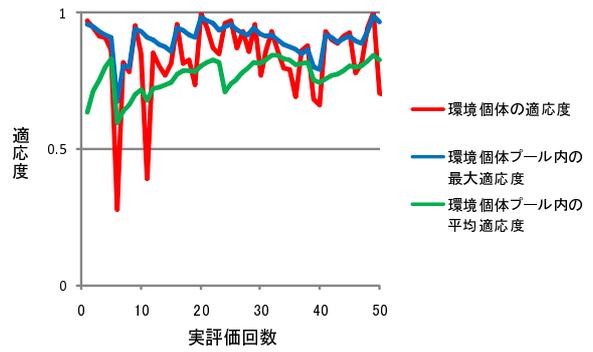


図 12: 環境個体の適応度の推移 (初期値ランダム設定)

す個体による行動を回避する方式を提案し、実進化 2 足歩行ロボットを用いて、その基本的評価を行った。その結果、歩行の自律的獲得の初期に頻出する転倒を削減しうることが示された。このことは、ロボット疲労の軽減だけでなく、時間的コストの軽減も意味する。提案手法では、先行評価器は、その精度を適応度してロボット行動の制御器との共進化系を構成することにより、作成のコストの軽減を狙っているが、環境個体のやや特異な進化の様子も観察された。事前に行動を選別すると同時に環境モデルを内部モデルとして作っていく処理は、人の行動獲得における脳の働きと類似した側面を持っており、人の知能の形成という観点からも興味深い。

## 参考文献

- [Nolfi 2000] Nolfi, S. and Floreano, D.: *Evolutionary Robotics*, MIT Press (2000).
- [Watson 99] Watson, R. A. and Ficici, S. G. and Pollack, J. B.: Embodied Evolution: Embodying an Evolutionary Algorithm in a Population of Robots, *Proc. of the 1999 Congress on Evolutionary Computation* (1999).
- [Watson 2002] Watson, R. A. and Ficici, S. G. and Pollack, J. B.: Embodied Evolution: Distributing an Evolutionary Algorithm in a Population of Robots, *Robotics and Autonomous Systems* (2002).
- [Usui 2003] Usui, Y. and Arita, T.: Situated and Embodied Evolution in Collective Evolutionary Robotics, *Proc. of the Eighth International Symposium on Artificial Life and Robotics* (2003).
- [笠井 2007] 笠井大輔, 有田隆也: 先行評価系を有する実進化型群ロボットシステム, 第 34 回知能システムシンポジウム講演論文集 (2007).
- [Elfwing 2005] Elfwing, S. and Uchibe, E. and Doya, K. and Christensen, H.: Biologically Inspired Embodied Evolution of Survival, *Proc. of the IEEE Congress on Evolutionary Computation 2005* (2005).
- [中井 2008] 中井淳一, 有田隆也: 実進化型ロボットシステムへの先行評価導入に関する基本的検討, 第 22 回人工知能学会全国大会論文集 (2008).