

多自由度ロボットの運動学習のためのイベント駆動型制御とその表現

○ 山口明彦 (京都大学)* 川嶋宏彰 (京都大学) 松山隆司 (京都大学)

Event-driven control and its representation for multi-DOF robot to learn motion

* Akihiko YAMAGUCHI (Kyoto Univ.), Hiroaki KAWASHIMA (Kyoto Univ.),
Takashi MATSUYAMA (Kyoto Univ.)

Abstract— In this paper, we propose a control model and its representation for multi-degree-of-freedom (multi-DOF) robot to learn hole-body motion, such as hopping and vaulting horse. This control model is constructed with the idea that the temporal relation between local control signal and the changing of dynamical constraint of the body, such as landing or taking off, is an essential element in generating motion. Being an intuitive model, it can reflect qualitative knowledge of human, learn quantitative parameters, and also optimize qualitative knowledge itself. We take simulations of hopping task on trampoline to verify our methods.

Key Words: Event-driven control, EventGraph, motion learning, dynamic environment

1. 定性的な運動知識をもとにした運動学習

多自由度ロボットに何らかの運動を学習させるときの課題として、実現させたい運動をどのように与えるか(目標運動の記述方法)、運動を生成するための制御則をどのように表現するか(制御則の表現方法)、目標運動の記述から制御則の表現をどのように獲得するか(制御則の探索・最適化方法)の3点が主にあげられる。ここでは、歩行や跳躍、宙返り(Fig.1)といったある程度複雑な全身運動、及び跳び箱運動のような高次の運動を想定しており、動的な環境、すなわちロボットの行動によって状態が変わる、動特性を持つ環境下(例えばトランポリンやロイター板¹などで、静的な環境も含む)で行われる運動を対象にしている。

従来の運動学習では、制御則の表現(つまり探索空間)に目標関節角の時系列データを用いて(例えば[3]),与えられた何らかの評価関数(目標運動の記述)を最大化するために、強化学習[4]や遺伝アルゴリズム[5]などを適用することが多い。しかし、解に近い初期値を用いているにもかかわらず学習ステップ数が多いなど、探索空間がしばしば膨大になるという問題がある。本研究では、探索空間の構成、すなわち学習すべき運動の表現に焦点を絞る。そして、

仮説◇ 力学的拘束条件の変化といったイベントと局所的な制御信号の間の時間関係、及び運動に存在するイベントの構造が、運動の実現において重要な要素である

という経験的事実から来る仮説に基づき、イベント駆動型制御という制御モデル(制御信号の生成モデル)を提案する。イベント駆動型制御とは、身体のある関節に対する局所的な制御信号を出力するタイミングを、ほかの局所的な制御信号やイベントとの時間関係から決定して出力する制御モデルである。ここで局所的な制御信号とは、少数のパラメータとそれらで決定される関数によって表現される制御トルクであり、これをプリミティブと呼ぶ。プリミティブは探索空間を圧縮する役割を果たしている。

つまりイベント駆動型制御は、人間の定性的な運動知識、すなわちある運動を実現するとき身体を制御して行く手順を、直接的に反映させられる制御モデルであると言える。例えば跳び箱運動について、人間は制御の手順を書き下すことができるが、これをイベントと局所的な制御信号の関係として表すことはさほど困難ではないだろう。もちろん、定量的なパラメータ(時

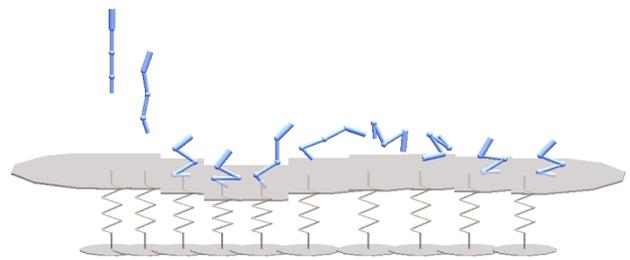


Fig.1 Somersault of multi-DOF robot on a trampoline.

間的距離やプリミティブのパラメータ)を正確に知ることはできないが、これは人間でも個人差があり、ロボットが学習によって最適化するのが適当である。本稿では、与えられた定性的な運動知識をもとにした学習を実現し、定性的な運動知識そのものの獲得に関する基礎的検討を行う。

2. イベント駆動型制御

2.1 仮説◇の検討：イベント駆動型制御の有効性

前節で述べた仮説◇の物理的根拠は、力学的拘束条件が変化する(時間的)前後で、拘束されている部位に加える力が運動に及ぼす影響が、大きく異なる点にある。例えば歩行では、地面に支持脚が着いた状態で地面を蹴らなければ意味が無い。さらに、動的環境下では環境の状態が変動するため、単なる前後ではなく、時間間隔も重要になる。例えば動的環境のひとつであるロイター板を用いて跳躍することを考えたとき、ロイター板を踏むというイベント(拘束条件が変化)の直後にロイター板を蹴っても、あまり高くは跳べない。もっとも高い跳躍ができ、このことはイベントと足を蹴るという行動の間の時間関係が重要であることを示している。足を蹴るという行動は、各関節に加える局所的な制御信号(トルク)で構成されるため、仮説◇につながっている。なお、着地イベントの前に脚を曲げるように、行動がイベントの前に来ることがあることを強調しておく。

ここで、イベントの定義を「ロボットが実現した軌道から安定して抽出される点(瞬間)」として拡張するために、Kuniyoshiらが行った、「roll-and-rise motion」

*現在、奈良先端科学技術大学院大学

¹跳び箱などで使われる踏切板;ジャンプ台

(起き上がり行動)に関する研究 [1][2] から得られる知見を以下にまとめる。

- K1 身体の力学的拘束条件を利用すれば、常に精密な制御をしなくても、ツボ (knacks) を押えた制御で大域的にはタスク (運動) が実現できる
- K2 タスクを成功させるために重要なのは、トータルエネルギーや関節角ではなく、局所的なエネルギーを出力するタイミングである

K2 のタイミングとは、動作全体を 3 区間に分けたときの、エネルギー投入重心時刻について述べたものであり、区間と区間の境界が本研究で言うイベントに一致すると考えられる。例えば足を完全に振り上げたときなどである。このようなイベントはロボットが実現した軌道から安定して抽出可能であると考えられ、これを基準に局所的な行動を起こすタイミングを定めることで、より安定な運動が実現できると考えられる。

イベントの構造が重要であることは、跳び箱運動と言った複雑な運動を例に考えると直ちに理解される。ロイター板を使った跳躍の例でも、着地イベントがあり、ロイター板を蹴った後、ロイター板から離れるというイベントがあるが、この前後関係は変わらない。跳び箱運動のように複雑な動作の場合だと、より複雑なイベントの構造が存在し、それらを参照することで安定に運動を生成できると考えられる²。

2.2 プリミティブの出力タイミングの決定

前節では陽に触れなかったが、局所的な制御信号 (プリミティブ) を出力するタイミングを決める際に、イベントだけでなく、ほかのプリミティブとの時間関係を用いることも有効である。ロイター板を使った跳躍の例で述べたように「ロイター板を蹴る」という行動は複数のプリミティブから構成される。つまりこれらのプリミティブは蹴るという行動としてまとまっており、その間に時間関係が存在することを意味する。なお、ほかのプリミティブとは、同じ関節に対するほかのプリミティブ以外に、異なる関節に対するプリミティブも含んでいる。

よって、イベント駆動型制御では、プリミティブを出力するタイミングを、イベント及びほかのプリミティブとの時間関係を用いて決定し、制御信号を出力する。時間関係とは二項関係であるから、イベント駆動型制御とは、イベントやプリミティブと言った要素とそれらの二項関係によって構成される運動のモデルに従い、制御信号を生成する制御モデルである。

2.3 プリミティブの形状

Kuniyoshi らの研究から得た知見より、運動タスクを実現する上では (局所的な) 制御信号の波形よりもむしろそのエネルギーが重要なファクタとなっており、細かな形状は運動の実現に大きく影響しないと考えられる。また、エネルギー投入のタイミングが重要であるから、信号のピークを明示できることが望ましい。これらのことから、本研究では Fig.2 に示すような釣鐘型のトルク波形を用いることにする (これ以外の定義も可能である)。この波形を特徴付けるパラメタは、 Δt_b , Δt_f , $peak$ であり、具体的には

$$T(\tau) = \begin{cases} peak \cdot \exp\left(-\frac{\tau^2}{2\Delta t_b^2}\right) & (\tau < 0) \\ peak \cdot \exp\left(-\frac{\tau^2}{2\Delta t_f^2}\right) & (\tau \geq 0) \end{cases} \quad (1)$$

によってトルクが計算される (τ はピークからの時間差)。

²ただし、跳び箱運動のような高次の運動においては、力学的拘束条件の変化と言ったイベントの構造は不変ではない。例えばロイター板まで助走するとき、ステップ数の調整が行われるが、これによってイベントの数が変化する。

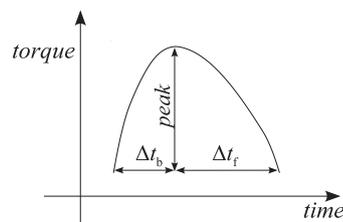


Fig.2 Control primitive.

3. イベント駆動型制御の表現

3.1 EventGraph

イベント駆動型制御は二項関係が基本となっているから、その表現にはグラフ構造が適していると考えられる。このグラフ構造を EventGraph と名付ける。EventGraph のエッジは要素間の時間関係を表現することになる。ところで、プリミティブは局所的な制御信号であるから、区間幅 (時間幅) をもつ。また、2.3 項で述べたように釣鐘型の形状でピークの位置を明示するため、プリミティブは開始点、ピーク点、終了点の 3 つの点 (これを「制御点」と呼ぶことにする) から構成される。これより、エッジがつなぐのはプリミティブの個々の制御点及びイベント (点) である。

プリミティブとプリミティブの時間関係から出力するタイミング (時間軸上の位置) を決定するとき、どちらのプリミティブを基準とするか (すなわち依存関係) を定めておく必要がある場合がある。この関係は、基本的にはイベントからエッジを辿ることであるが、グラフがループしている場合などは知ることができない。そこで、エッジに方向を持たせることで表現する。以上の議論から、EventGraph はプリミティブ、イベント、エッジから構成される有効グラフである。

3.2 EventGraph に基づく制御信号の生成

EventGraph が既存のイベント-アクション型の表現と異なっている特徴のひとつに、イベントに関連づけられたアクションが、イベントよりも時間的に前に実行されるという表現も許容していることがあげられる。このため「イベントの待機 → (遅延) → 動作」という実行の流れでは制御信号を生成できない。イベントが発生するよりも前にプリミティブを出力しなければならないから、イベントの発生時刻を「予測」する必要がある。これはイベントとイベントの時間関係に基づいて行われる。この予測は、一般的にイベントが近付く程正確になるはずだから、予測の情報が更新されたときに、各プリミティブの生起すべき時刻を更新する必要がある。この更新は、グラフ構造を辿って伝搬される。このように、イベントの予測と更新の伝播が、EventGraph に基づいて制御信号を生成するアルゴリズムの中核となる。

4. イベント駆動型制御に基づく運動学習とその評価

イベント駆動型制御に基づく運動学習とは、与えられた評価関数を最大にする EventGraph を求めることである。この学習は、

1. グラフの構造を固定したパラメタ (プリミティブの形状など) の最適化
2. グラフの構造自体の最適化

の 2 段階に分けられる (Fig.3 参照)。

本研究の最終的な目標は、多様な動的環境下においてこれらの最適化を実現し、多自由度ロボットに運動を学習させることである。そこで本節では、人間と類似の身体性をもったロボットに、人間の定性的な運動知識に基づいて EventGraph の構造を与え、(1) のグラ

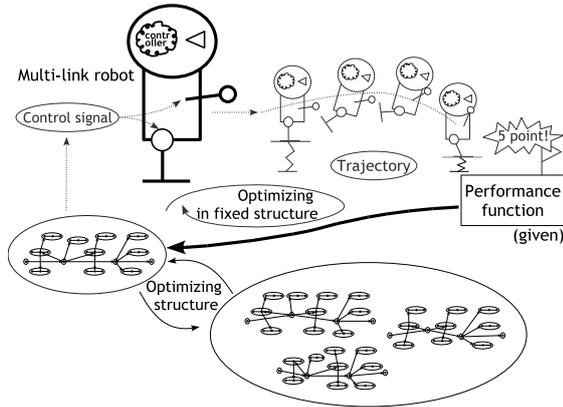


Fig.3 Motion learning of multi-link robot in a dynamic environment.

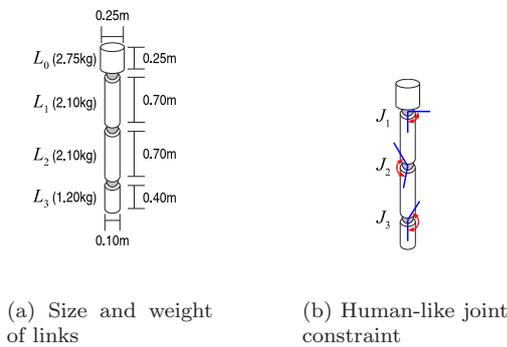


Fig.4 Four-link robot structure.

グラフ構造を固定した最適化を実現することで、イベント駆動型制御が制御モデルとして有効に機能することを明らかにする。また、(2)のグラフ構造自体の最適化についての基礎的な検討として、ロボットの身体性とグラフ構造をともに変化させた場合のタスクの達成度を調べる。

具体的な実験としては、トランポリン上で、4リンクの多自由度ロボット (Fig.4) に跳躍運動を学習させる。トランポリンは動的な環境であり、このため解析的に制御をすることが困難である³。

なお本実験は、Featherstone の順動力学学アルゴリズム [7] をもとに、トランポリン (平面で大幅に変形しない) と多自由度ロボットの動力学シミュレーション環境を構築して行った。

4.1 グラフ構造を固定したパラメタ最適化の手法

「構造を固定した最適化」は、EventGraph のエッジやノードの数、連結関係を固定し、個々のエッジのメトリックやプリミティブのパラメタをひとまとめにしたパラメタベクトルを X とおいたとき、

$$X^* = \arg \max_X \text{Evaluate}(\text{EventGraph}) \quad (2)$$

のように定式化される。

最適化手法としては、微分情報を用いられないこと、また、構造の最適化も視野に入れていることから遺伝

³ハイブリッドシステムの枠組でこの問題を解決し、フィードバック制御によってロボットに運動を生成させる手法などが提案されている [6] が、フィードバック制御では平衡状態を目標とするため、複雑な運動実現には無理がある。

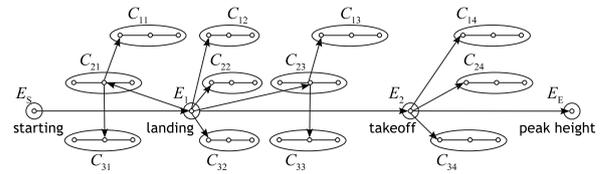


Fig.5 An EventGraph for trampoline hopping.

アルゴリズム (GA) を用いた。GA の構成は、世代交代モデルには Minimal Generation Gap [8]、交叉オペレーションにシンプレックス交叉 [9] を用いた。

4.2 グラフ構造を固定したパラメタ最適化の評価実験

ここでは 4.1 項で述べた、パラメタ最適化を行い、解の収束性を調べることによって最適化が成功するか評価する実験を行う。達成すべき運動 (タスク) を「一定の高さ⁴から落下させ、トランポリン上で跳躍し、できるだけ高い位置まで跳ぶ」という設定にした。

EventGraph の構造は人間の定性的な運動知識に基づいて Fig.5 のように与え、これを固定して 4.1 項で述べたパラメタ最適化を行う。「定性的な運動知識」と EventGraph の関係は、次のようなものである。

運動開始 (イベント E_S) 後、着地 (E_1) の前に足を曲げ (制御 C_{11}, C_{21}, C_{31})、着地の直後に足をふんばり (C_{12}, C_{22}, C_{32})、トランポリンが十分に沈みきってから足に力を加えて (C_{13}, C_{23}, C_{33}) 跳躍する。離陸 (E_2) したらバランスを取る制御 (C_{14}, C_{24}, C_{34}) をする。高さがもっとも高くなったら (E_E) 終了。

評価関数は、跳躍後の高さと同跳躍後のバランスで定義する。

$$\text{Evaluate} = \frac{5}{2} \cdot \text{after-height} + 2(1 + R_{33})^2 + \frac{4}{1 + |\omega_0|} + \sum_{i=1}^3 (1 + \cos q_i)^2 + \frac{2}{(1 + \dot{q}_i^2)} \quad (3)$$

ここで after-height は跳躍後の高さを、 R_{33} はエージェントの姿勢行列の (3,3) 要素を⁵、 ω_0 はリンク L_0 の角速度ベクトル、 q_i は関節 J_i の角度をそれぞれ表す。

実験結果と考察 1 回の学習を「GA の世代が 400 交代するまで」と定義して⁶、学習を 15 回行った。各世代において個体もつ評価値から最大のものを選び、15 回の学習での平均と分散を世代番号に対してプロットしたグラフを Fig.6 に示す。この結果は、最終的には収束する方向に向かっていることを示しており、学習が毎回のタスクで成功していると判断される。

学習によって得られたタスクの様子を、Fig.7 に示す。この図からも運動を実現できていることが確認され、イベント駆動型制御の有効性を確かめられたと言える。

4.3 EventGraph の構造最適化の基礎的な検討

4.2 項の実験で用いた EventGraph (Fig.5) は、人間の定性的な運動知識に基づくものであり、必ずしもロボットに最適ではない。したがって、この定性的な運動知識 (= グラフの構造) そのものも最適化する必要がある。以下では、そのための基礎的な検討を行う。

注目すべきであるのは、イベントの数及び種類、プリミティブの数、及びそれらの連結関係である。イベントについては、人間が最適に与えられると仮定して、残りの項目のうち、運動にもっとも大きく影響を与えるプリミティブの数について考察する。

⁴具体的には、トランポリンからロボットの頂点までの距離が 2.7m。

⁵Z-Y-X オイラー角表現における x 軸回転を φ 、 y 軸回転を θ としたとき $R_{33} = \cos \theta \cos \varphi$ である。

⁶1 個体にひとつの EventGraph が割り当てられており、評価値はタスクを 1 回行って計算する。また、1 回の世代交代において 8 個の子孫を生成している。「400 回」は、予備実験に基づく判断。

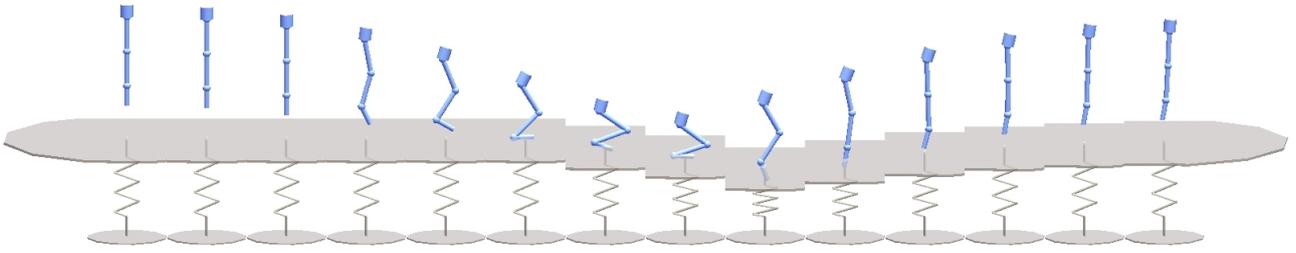


Fig.7 A series of images in trampoline hopping motion (0.1s per frame).

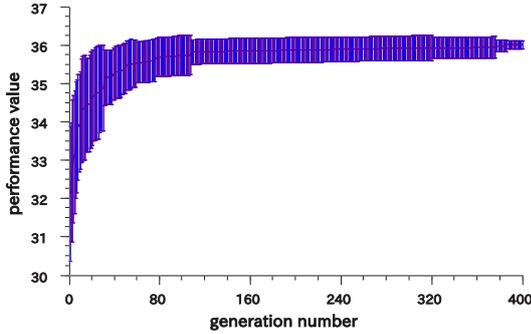


Fig.6 Result of learning trampoline hopping.

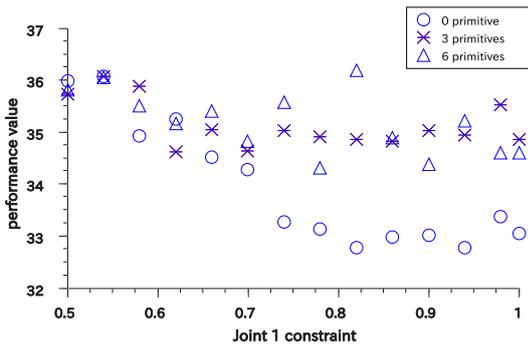


Fig.8 J_1 constraint – performance-value graph for 0 primitive, 3 primitives, and 6 primitives.

Fig.5のEventGraphの、 E_2 に接続されているプリミティブの数を0個、3個、6個に変化させ、それぞれについて関節 J_1 の可動範囲(関節制約)を $\pi/2$ から π まで変化させて、その各々についてパラメタ最適化を行うことにより、タスクの困難さ(可動範囲が大きい程困難になる)、プリミティブの数に対する達成度(学習後の評価関数の値)を調べた。結果をFig.8に示す。

このグラフより、可動範囲が小さい領域ではいずれのプリミティブの数でも十分学習できているが、可動範囲が大きい領域では、プリミティブが0個の場合達成度が低い。また、プリミティブの数を増やすと達成度は高くなるが、プリミティブが6個の場合などはバラツキが大きく十分学習できていない(あるいは局所解に陥っている)。以上より、プリミティブの数は細かな制御と学習の早さのトレードオフで決まると言える。

5. 結論と今後の展望

本研究では、動的な環境下で多自由度ロボットに運動学習をさせるため、イベントと局所的な制御信号の時間関係が重要であるという仮説に基づき、イベント

駆動型制御を設計し、その表現としてEventGraphを提案した。また、シミュレーションによる実験により動的環境下における運動学習に対する有効性を示し、構造最適化の基礎的な検討として、プリミティブの数の選び方について指標を見出した。これらのことから、イベント駆動型制御に基づく運動学習は有効な手段となり得ると考えられる。さらに、紙面の都合上省略するが、提案手法の応用として多自由度ロボットにトランポリン上で宙返りを学習させた結果(Fig.1)から、手法の汎用性も期待される。

今後の課題としては、(1) イベントの抽出をどうするか、(2) 構造最適化のアルゴリズム(例えば遺伝的プログラミング)、(3) プリミティブの形状、(4) より高速な学習などがあり、検討を進めている。

一度最適化されたグラフの構造は環境や身体の状態の変動に対して大きく変化しないため、異なる環境下や身体性が異なるロボット(ただしリンクの構造は同じという想定)に対してもグラフ構造を固定した最適化のみ行えばよく、効率的であると考えられる。

グラフの構造を含めた最適化を実現することで、未知の動的環境への適応や、人間とは全く異なる身体性をもつロボットの運動学習につながるだろうと期待される。また、グラフ構造に対する演算(ノードの追加や削除、グラフ同士の連結など)を制御信号に反映させることが実現できれば、高次の運動(跳び箱運動など)を動的に生成することが可能になると考えられる。謝辞 本研究の一部は、科学研究費補助金18049046の補助を受けて行った。

参考文献

- [1] Y. Kuniyoshi, Y. Ohmura, K. Terada, A. Nagakubo, S. Eitoku and T. Yamamoto: "Embodied basis of invariant features in execution and perception of whole body dynamic actions - knacks and focuses of roll-and-rise motion", The Second International Workshop on Man-Machine Symbiotic Systems, pp. 289-301 (2004).
- [2] 國吉, 大村, 寺田, 長久保: "等身大ヒューマノイドロボットによるダイナミック起き上がり行動の実現", 日本ロボット学会誌, 23, 6, pp. 706-717 (2005).
- [3] 村田, 浅井, 佐久間, 小林: "強化学習による多自由度2足歩行ロボットの制御", The 19th Annual Conference of the Japanese Society for Artificial Intelligence (2005).
- [4] 木村, 山下, 小林: "強化学習による4足ロボットの歩行動作獲得", 電気学会電子情報システム部門誌, 122-C, 3, pp. 330-337 (2002).
- [5] 吉田, 水内, 稲葉, 國吉, 井上: "脊椎構造を持つ人間型ロボットにおけるGAによる歩行動作の自動獲得", 日本機械学会ロボティクス・メカトロニクス講演会'02講演論文集, pp. 2P2-L04 (2002).
- [6] M.Ishikawa, A.Neki, J.Imura and S.Hara: "Energy preserving control of a hopping robot based on hybrid port-controlled hamiltonian modeling", IEEE Conference on Control Applications (CCA2003), No. CF002507 (2003).
- [7] R. Featherstone: "Robot Dynamics Algorithms", Kluwer Academic Publishers (1987).
- [8] 佐藤, 小野, 小林: "遺伝的アルゴリズムにおける世代交代モデルの提案と評価", 人工知能学会誌, 12, 5, pp. 734-744 (1997).
- [9] 樋口, 筒井, 山村: "実数値GAにおけるシンプレックス交叉の提案", 人工知能学会論文誌, 16, 1, pp. 147-155 (2001).