

ロボット行動学習への事例ベースの適用

頼光正典[†], 松田憲幸[†], 瀧 寛和[†], 安部憲広[‡]

[†]和歌山大学大学院システム工学研究科, 和歌山市栄谷 930, tel:0734-57-8126

E-mail: s011065@sys.wakayama-u.ac.jp (頼光), matsuda@sys.wakayama-u.ac.jp,
taki@sys.wakayama-u.ac.jp

, [‡]九州工業大学大学院情報工学研究科, 福岡県飯塚市川津 680-4, tel: 0948-29-7769

E-mail: abe@sein.mse.kyutech.ac.jp

あらまし 本論文は、リアルタイム環境下での学習に事例ベースを適用する方法について論じている。ロボットの行動学習は、行動主義的な人工知能の典型的な問題であり、強化学習、GA や帰納推論などを利用した学習方法が適用されている。一方、事例ベース推論は、獲得・経験した事例を多くの処理を必要とせず蓄積・適用が可能である点からリアルタイムに適した行動学習の方法である。学習内容では事前に知識を準備している強化学習には劣るが、事例ベースの利用には、事前知識なしで行動の事例を簡単に獲得できる利点がある。本研究では、「事例の一般化度合い」、「事例の問題カバレッジ」と「事例の個数」の観点からロボットの行動学習への適用方法について論じる。また、評価には、ロボカップのシミュレーションリーグのプラットフォームを利用している。

キーワード 事例ベース推論, ロボット工学, 機械学習, 行動学習

An Approach of Robot Behavior Learning using Case-based Reasoning

Masanori Yorimitsu[†], Noriyuki Matsuda[†], Hirokazu Taki[†], Norihiro Abe[‡]

[†]Department of Systems Engineering, Wakayama University,

[†]930 Sakaedani, Wakayama 640-8510, Japan, tel:0734-57-8126, E-mail:

s011065@sys.wakayama-u.ac.jp (Yorimitsu), matsuda@sys.wakayama-u.ac.jp,
taki@sys.wakayama-u.ac.jp

[‡]Department of Information Engineering, Kyushu Institute of Technology
680-4 Kawazu Iizuka-Shi Fukuoka, 820-8502, Japan, tel: 0948-29-7769,

E-mail: abe@sein.mse.kyutech.ac.jp

Abstract This paper described an application method of case-based reasoning for a robot action learning in real-time environment. The robot action learning is a typical problem of behavior-based AI. Reinforcement learning, GA and induction are adopted for this problem. On the other hand, the case-based reasoning is suitable for real-time action learning, because it is not necessary for the method to process row acquired cases. The efficiency of the knowledge extracted by this method is not better than the reinforcement learning. But the latter method needs the prior knowledge, and the former one has the advantage that it doesn't need that knowledge. We discuss the application methodology of case-based robot action learning in view of "generalization ratio of cases", "covering ratio of the problem space with cases" and "the number of cases". We use the Robocup simulator for evaluation of this method.

key words Case-based Reasoning, Robotics, Machine Learning, Behavior-based Learning

1. はじめに

ロボットの用途は、従来の産業分野だけでなく、AIBO[1]に見られるようなエンタテインメント分野などにも広がっている。利用される局面が広がるに従って、その利用方法は多種多様である。つまり、予めその利用環境を詳細に見極めることは非常に難しいと言える。このような状況への対応として、ロボットの行動学習に対する研究が進められている。一方、人工知能の研究スタイルの1つとして、従来の「記号处理的AI」のように知識を記号を利用して、明確な形で表現する表象主義的なAIから、環境とは独立した身体を持つエージェントを考え、環境からエージェントへの刺激と反射行動から知的な行動を獲得する「行動主義的AI」が独立した方法として確立した[2]。このパラダイムでは、多様な環境情報とエージェントの内部状態（過去の環境情報や行動履歴を含む）の条件と、ロボットの実行可能な行動を結びつけることにより、行動知識を獲得する。行動学習の方法として、種々の条件下での行動が適切な場合に、報酬を与えることで、その行動を強化する（環境に適したと判断し、利用しやすくする）強化学習[3]がもっとも有効な方法として利用されている。強化学習では、学習対象の知識（ルール）を予め準備し、そのルールへの報酬と、行動系列（ルールの実行系列）への報酬により、学習対象のルールの利用に関する知識（ルールの優先順位やルール系列）が獲得される。非常に簡便な方法であり、ロボット（エージェント）の行動に対して、適切な報酬を与えることで学習が進められる。報酬は即座に、知識の利用に反映されるため、リアルタイムの学習に適している。また、有限の知識を準備しているので、行動の推論に対する計算量も予め制限できる。ただ、学習に利用する事前知識を準備しなくてはならない。

強化学習の他に、利用される学習方法には、GA

（遺伝的アルゴリズム）[4]がある。GAは、単純な遺伝子操作（交差、突然変異など）と、その遺伝子の発現状態（ここでは、ロボットの行動）の環境適合度による選択（適者生存）により、行動を学習する方法である。進化の過程が進むに従って、行動の優れたロボット（エージェント）が生まれてくる。この学習は、遺伝子の能力が高水準に収束するまで、多くの遺伝子の能力を調べるのに多くの計算量が必要である。計算量の関係からリアルタイムの学習には向かないため、学習過程はオフライン（ロボットの行動外）で行われる。

本研究では、ロボットの行動学習に、リアルタイムで知識を獲得でき、事前知識を必要とせず、また、有限の知識で行動でき、より良い行動を残していく方法として、行動学習に利用できる事例ベース推論[5]を活用した学習方法を提案する。

2. リアルタイム性と事例ベース推論

リアルタイム性の観点から、事例ベース推論の利点と、欠点への対応について考察する。事例ベースの特徴を説明した後、ロボット行動への利用上の課題と対応について説明する。

2.1 事例ベース推論の特徴

事例ベース推論は、多くの事例を蓄積し、ある問題に対して、類似事象を持つ事例を検索し、その問題に適合するように修正した後、問題解決に修正した事例を適用する枠組みである。図

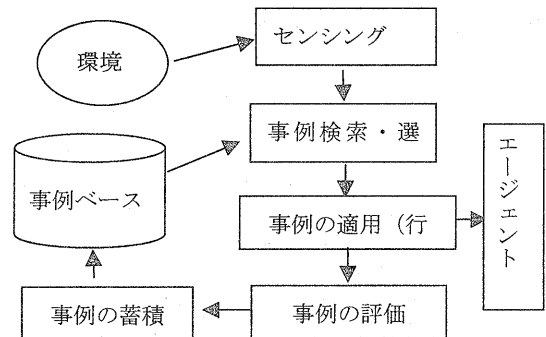


図1. 行動学習向き事例ベース推論

1は、ロボットの行動学習に事例ベース推論を適用した場合の事例の蓄積と適用に関する構成を示している。ロボットの置かれた環境から、行動に必要な情報をセンシングし、その環境に適合する行動を決定するために、過去の有効な類似の事例を検索・選択する。その事例を適用するか、または、ランダムな行動を選択し、ロボット（エージェント）に行動を指示する。ロボット行動後に、その事例（または、ランダム行動）の効果を評価した後、事例ベースに事例を蓄積する。事例ベース推論では、事例が増えるほど、多くの問題に対応できる能力に対処できる可能性は高くなるが、効率的な事例選択が必要となり、事例適用までの計算量が増える。事例ベース推論の活用で、最も難しい点は、事例の修正である。既知の事例を利用する場合には、目標とする問題に合うように事例を修正しなければならない。適切な修正を行うには、事例と適用分野に対する十分な知識（事前知識）が必要である。

問題解決能力の向上には、事例の増加が必要であり、リアルタイム性に反する。事例の適用には、事前知識が必要であり、多くの事前知識があるなら事例は不要である。事例選択から適用までの時間が長くなると環境が変化してしまい、ロボットの行動と環境の状況に差異が生じ、行動の環境適用能力が低下する。

2.2 リアルタイム性を持つ事例ベース

事例ベース推論をリアルタイムでの行動学習に利用するために、下記の機能を持つ枠組みを検討した。

- (1) 事例の個数の制限
- (2) 修正不要な事例表現
- (3) 単純な評価（報酬基準）を適用できる事例獲得系

事例の個数を制限するには、「事例ベースの事例個数を制限」「制限個数内で、より良い事例との入れ替え」「不要事例の消去機能」「事例の適

用範囲の動的変更」による対応が考えられる。修正不要な事例表現を実現するには、ロボットの行動の単純化と行動パラメータ学習の分離（有効なパラメータ値を1つだけ利用する）することで、できるだけ変更することなしに事例を利用できるようにすることが考えられる。事例を蓄積すべきかどうかを判断するための事例評価方法を考えた場合に、事例の単純化とその評価も単純化しなければならない。ロボットの行動学習において、ロボットの行動系を細かな部分機能に分解し、各機能がどのような条件で起動されるかを表現することで、その行動規則を単純化することが考えられる。

3. ロボットの行動系

3.1 ロボットの行動環境

ロボットの行動環境として、RoboCupのシミュレーションリーグ（サッカーサーバー）[6]（図2）をプラットフォームとして利用した。ただ、利用する機能（関数）、シミュレータの関数群ではなく野田氏のサッカークライアント用ライブラリ [7]を利用している。ロボットは、環境から「ボールの有無」「ボールの方向」「ボールとの相対距離」「ゴールの有無」「ゴールの方向」「ゴールとの相対距離」が得られ、行動可能なオペレーションとしては、「移動方向+移動距離+移動速度」「回転方向+回転角度」「ボールを蹴る（シュート、キック）+キック量」を利用することにした。

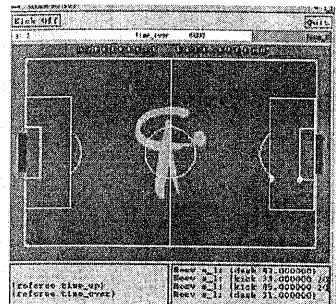


図2. RoboCup サッカーサーバー

3.2 ロボットの行動系の設計

ロボットの行動系は図3のように、各行動機能がセンサー入力から直接情報を得られる構造にしている（これは、将来、サブサンプリング・アーキテクチャに拡張を想定している）。これは、リアルタイム性を事例ベース推論で実現するために、「ロボットの行動の単純化と行動パラメータ学習の分離」に適している、また、「事例の単純化とその評価の単純化」に対応できるような設計となっている。図3では、センサー部、行動パラメータ決定部（回転、移動、キック）、行動選択部（協調・競合）、行動結果評価部で構成され、センサー入力に対応する事例の選択により、行動が決定される。ロボットの実行した行動は、各行動個別の評価により、その行動事例が評価され、評価基準が満たされると事例ベースに登録される。対応する事例が無い場合には、行動はランダムに決定される。また、一定の確率で、事例が存在する場合にもランダムな行動をとることができる。

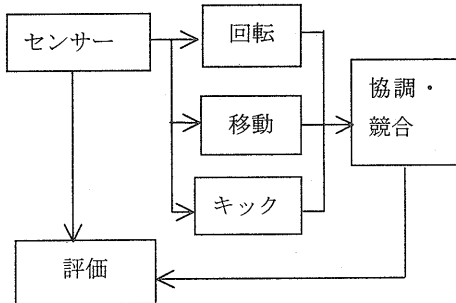


図3. ロボットの行動系の構造

3.3 事例の構造

事例は、「センサー入力と行動内容」で表現される。行動内容に依存して、センサー入力の項目は選択される。ボールへの接近動作では、ボールとの相対距離が重要となる。また、ボールに向かうための回転行動では、ボールとの相対角度が重要である。また、行動

内容は、オペレーション（行動）とその度合い（速度、回転量など）で表現される。

事例の構造：

[環境条件] [行動内容] [行動量（度合い）]

3.4 事例の評価

個々の事例の行動が異なっているため、行動部分の効果について事例は評価される。たとえば、回転行動により、ボールを正面に捉える行動では、ボールとの相対角度の変化（少ないほど良い）でその事例が評価される。

3.5 事例の適用

事例ベースに蓄積されている事例とまったく同じ条件の環境に遭遇することは稀である。そのため、通常事例は、環境に合わせて、修正した後、適用されるのが一般的な事例ベース推論の利用方法である。しかし、事例の修正には、多くの知識と計算時間が必要であるため、本システムでは、ロボットの行動を規定している部分の事例修正は行わず、行動とその度合い（速度、回転量など）は、事例の持つデータをそのまま適用する。このことにより、ロボットの環境適用能力は下がるが、事例の適用に必要な時間は節約できる。

3.6 事例の消去

蓄積された事例の数は、事例検索と選択の時間に影響する（正の相関を持つ）。その為、あまり有効でない事例は、消去されるべきである。本システムでは、その事例を適用したが、ロボットの行動評価が悪い場合（効果が有効でない場合）には、その事例が消去されるようにしている。事例に対して、その行動評価により、正負の評価点を与え、その累積和がある値以下になると事例は消去される。

4. 学習実験

4.1 実験の目的

ロボットの行動学習実験の主たる目的は、
<1> 事例ベース推論で行動学習が行えることの実証、
<2> リアルタイム性を持たせた学

習の可能性の実証である。他の学習方法（強化学習や GA の応用）との比較による学習効率の向上は目的としていない（今後の課題とした）。

4.2 実験環境

RoboCup のシミュレーションリーグのプラットフォームにおいて、1 エージェントがボールをゴールするまでの時間を計測することで学習効果を調べた。ロボットの初期状態では、事例を持たず、リアルタイムに学習しながら行動することとした。また、学習で蓄積される事例の条件空間上での分布も計測した。

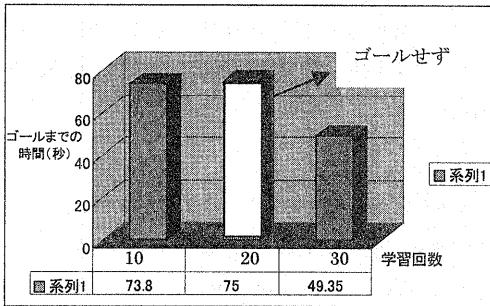


図 4. 学習回数とゴール時間

4.3 実験結果と評価

全体的評価として、正事例のみを用いた場合の学習回数毎のゴールまでの時間を計測した（図 4）。1 学習回数は、5 分または、ゴールするまでとし、このデータは学習後に学習処理を停止して（事例の追加なしに）、計測したものである。なお、学習回数 20 回のは、5 分以内に学習が完了しなかった。学習回数 10 回の事例数は回転 5、移動 2、キック 19 個、環境適合性が失われた 20 回目では、事例が増加して、それぞれ (21,2,2) 個。30 回目では、それぞれ (1,5,2) 個となり、事例の消去効果により不要事例が減少した。事例の分布の例を図 5 に示す。回転の条件でその分布を見るとボールとの相対角度が小さいときは、ボールとの距離があっても正事例

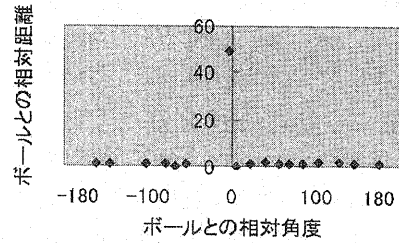


図 5. 事例と学習空間の例

になるが、ボールと相対角度を持つときは、ボールが近くないとボールが見つからないため、行動がとれないことを示している。

5. 事例の適用方法の拡張

この実験から、事例の学習空間のカバーがロボット行動に影響していることがわかるので、この点の拡張を試みている。また、学習空間をカバーしやすくするために、現在は固定となっている事例の適用範囲（条件の一般化の度合い）を可変にすることを試みている。

5.1 学習空間と事例によるカバー

図 6 は、学習空間（センサーの条件が作る空間の模式図）を表現している。通常は、センサーの信号から得られる特徴の種類で空間の次元が決まる。この図は、ボールの位置の

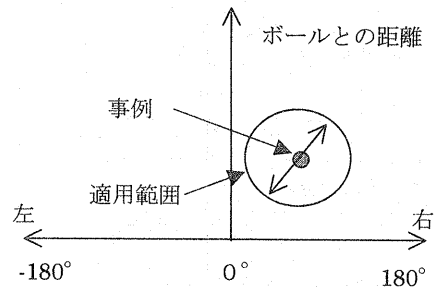


図 6. 事例の適用範囲

方向を横軸にして、縦軸には、ボールとの相対距離を示している。実験では、非常に多くの事例を獲得しているが、この図のように、各事例の適用範囲を広く取ることで、少ない事例で学習空間をカバーできる。また、行動

の精度を挙げるためには、カバー範囲を狭くして、多くの有効事例で学習空間を構成するのが適切である。

5.2 事例の適用範囲の動的変更

学習が開始されたときは、事例の数は少ないので、少ない事例で学習空間をカバーし、事例が増えてきた段階で、その事例の適用範囲を狭くすることができる。これは、事例の一般化による行動効果の増減に影響する。ある事例の適用範囲に、新たな事例（ランダムな行動による事例）が行われ、新事例の方が評点が高い場合には、新事例と旧事例を交換できる。

5.3 ランダム性と事例の一般化効果

この拡張をまとめると、ランダムな行動をとるか、効果が下がるが、事例の適用条件を緩和して利用すべきかのトレードオフを調べることになる（他にも、事例の個数による計算時間の増大による環境適応性の低下と事例の数の増大による適切な行動を含む可能性とのトレードオフもある。図7は、事例の適用

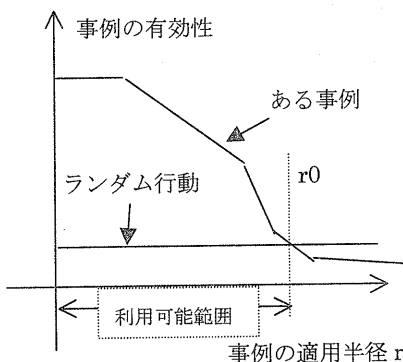


図7. 事例の適用範囲と有効性

半径（条件の空間、学習空間での適用範囲）と行動評価の有効性の関係を示している。ランダム行動の有効性と事例の有効性の交点の半径は、トレードオフ半径 r_0 である。事例の適用半径 r が r_0 以下では、事例の適用が有利であり、 r_0 を超える範囲では、ランダムな行

動の方が有利となる。

現在、本システムは、この拡張機能を搭載中であり、その実験を実施中である。

6. まとめ

事例ベース推論をロボットの行動学習に利用する方法に関して、リアルタイム性を中心に論じた。事例ベース推論の利用は、事前知識を必要としないことから、強化学習より効率落ち、事例の改良を行う機能が無いことから GA よりも学習効率は低いが、事前知識を準備せずにリアルタイム性を持たせられる利点がある。また、実験を通じて、事例の個数と適用範囲のトレードオフの必要性が明確となり、その拡張方法について説明した（この拡張については、実験中である）。

参考文献

[1] AIBO のホームページ：

<http://www.world.sony.com/JP/Electronics/aibo/>

[2] R.A.Brooks : New Approaches to Robotics, Science, Vol.253, September 1991, pp.1227-1232

[3] 浅田稔: 強化学習の実ロボットへの応用とその課題、人工知能学会誌、Vol.12, No.6, pp.831-836 (1997)

[4] 伊庭斉志: 遺伝的アルゴリズムの基礎、オーム社(1994)

[5] H. Taki, S. Hori, and N. Abe: "Case-Based Quality Management System Using Expectation Values", In Proceedings of the third International Conference on Case-Based Reasoning, ICCBR-99, pp.572-580 (1999)

[6] サッカーサーバーマニュアル

<http://www.ita.tutkie.tut.ac.jp/watta/RoboCup/jmanual.html>

[7] サッカークライアント用ライブラリ

<http://ci.etl.go.jp/noda/soccer/>