

実世界における基本行為のコミットメント戦略について

Commitment Strategies for Basic Actions in Real World

高田司郎 *1

Shiro TAKATA

新出尚之 *2

NIDE, Naoyuki

*1 近畿大学理工学部

KINKI UNIVERSITY Faculty of Science and Engineering

*2 奈良女子大学研究院生活環境科学系

Faculty, Division of Human Life and Environmental Sciences, Nara Women's University

In real world, a robot to do plans must adjust actuators according to sensor error and coefficient change of rolling friction. If an atomic action falls into a misstep state, the action would be got back to the successful state so that he/she maintains an intention of a basic action using a commitment strategy and promotes simplification of his/her decision-making. We confirm the validity of an adjustment of actuators to come off a goal state using a single-minded commitment strategy.

1. はじめに

我々は、BDI モデルを用いた実世界における合理的エージェントの学習と意思決定に関する研究課題に取り組んでいる [新出 12, 高田 13]。一般的に、BDI モデルの制御サイクルは、環境認知-計画-プラン本体の基本行為の実行を繰り返す。しかも基本行為は実行制御装置にて BDI モデルの制御サイクルと独立して実行しているため、その基本行為が成功したか失敗したかは次の制御サイクルの環境認知まで判断できない。

そこで、複数の動作で構成される基本行為、たとえば、両輪をモーターで回すことで「直進する」というロボットの基本行為の実行は、実行制御装置に環境との相互作用能力がない場合は、実行中に失敗して転倒してもモーターを 5 秒間回転するという基本行為が終了するまで実行を継続してしまい、その後の BDI アーキテクチャ制御サイクルにおける意思決定が複雑になる [藤田 12]。そこで本稿では、サッカーロボットをテストベッドとして、BDI モデルの制御サイクルとは独立に実行制御装置に環境との相互作用能力を持たせ、環境の外乱に対して頑健に（反射的行為として）基本行為をコミットメント戦略に基づいて遂行する方式を提案する。

2. 基本行為のコミットメント戦略

2.1 基本行為

基本行為とは、プランのプラン本体に記述されそのまま実行できる行為であり、たとえば本稿では、「パスする」「レシュープする」「シュートする」などを指す。

2.2 コミットメント戦略

事前にプランされた基本行為を正にこれから実行するという現在指向的意図（以下、意図と略す）の持続と破棄に関する「コミットメント戦略」 [Rao 91] について紹介する。

- Blind(盲目的) エージェントは、意図はすでに実現されていると信じるまで、その意図を持続する。

- Single-minded(一意専心) 意図はすでに実現されていると信じるか、もしくは、その意図の実現が可能であると信じなくなるまで、その意図を持続する。
- Open-minded(心の広い) その意図を形成した欲求を実現するという状況でなくなるまで、意図を持続する。つまり、その意図はすでに実現されたと信じるか、その欲求を実現する理由がなくなったので取り下げるまで、その意図を持続する。

BDI モデルを用いた合理的エージェントはこのようなコミットメント戦略に基づいて意図の持続や破棄を行う。本稿では、意図を実現する動作パターンが見つからなくなるまでその意図を持続する「Single-minded」を採用する。たとえば、基本行為「ドリブルする」という意図が実現されたと信じるか、もしくは、その意図を実現する「動作パターン」がなくなるまで、その意図を持続する。

また、コミットメント戦略を用いれば、以下のように、基本行為レベルや目標レベルでの再考慮も可能となる。たとえば、基本行為が失敗したときは、その基本行為を再実行するか、またはその基本行為が実行できなければ、現在目指している目標を達成する別の手段を目的-手段推論して、その目標を持続することができる。さらに、この目標の実現手段がなくなれば、目標レベルで再考慮し、意図を実現する別の目標を熟考して、その目標を達成するよう意図を持続する。このような再考慮は意図を持たない（たとえば目標だけを持つ）エージェントで実現することは難しい。

3. サッカーロボットの概要

3.1 使用機器と環境

本稿では、図 1 のレゴマインドストーム NXT、図 2 の ROBOCUP JUNIOR 公式赤外線発光ボール、図 3 のフィールドを使用する。

3.2 サッカーロボットの構成

図 4 にシステム構成を示す。PC 上の BDI エージェントは、Web カメラと QPToolkit[QPT] を用いて、NXT の座標と角度（向き）およびボールの座標を、例えば 200ms サイクルで取得

連絡先: 高田司郎, 近畿大学理工学部, 〒 577-8502 大阪府東大阪市小若江 3-4-1, 06-6721-2332, shiro@info.kindai.ac.jp



図 1: NXT 図 2: 赤外線ボール 図 3: フィールド

し、環境情報として更新する。これは、NXT にはカメラが搭載されておらず、また、赤外線センサを用いて赤外線ボールを認識できる距離は 30cm 内であるためである。PC 上の BDI エージェントは、この環境情報を適時観測することで意思決定を行い、NXT に行うべき行為を目標として指示する。これにより NXT は、装備されたセンサを用いて取得した部分情報に適応して与えられた目標を実現するために、合理的に動作パターンを実行する BDI ロボットとなっている。

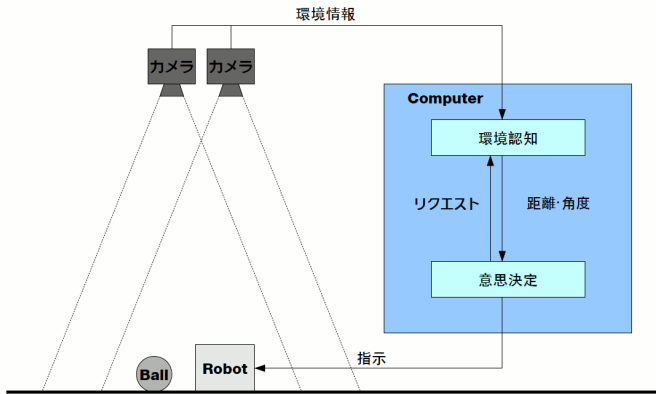


図 4: システム構成図

3.3 BDI エージェントの概要

サッカーロボット間の共同意図を実現するための基本行為レベルの意思決定は、PC 上の BDI エージェントが行う。特に本稿では、2 台の NXT が共同意図「パスまたはドリブルを行った後シュートする」を実現する。

3.3.1 プランとしての状態遷移図

BDI エージェントが共同意図「パスまたはドリブルを行った後シュートする」を実現するために選択するプランである状態遷移図を図 5 に示す。ボールに近い NXT が、味方の NXT にパスかドリブルを行い、味方の NXT がボールをレシーブ後、シュートするというプランである。

3.3.2 NXT に送信するコマンド

表 1 は、NXT に送信するコマンド一覧表である。例えば、味方の NXT に dribble を指示する場合、環境情報から NXT 自身の向きと位置、および味方 NXT の位置を得て、NXT 間の距離と味方 NXT の方向として絶対的方向（北を 0 度として左回りを正とする角度）を計算し、コマンド dribble, 距離, 絶対的方向 を NXT に送信する。

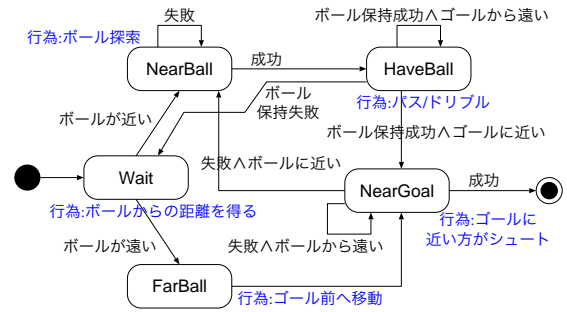


図 5: 状態遷移図

表 1: NXT に送信するコマンド一覧表

コマンド	指示する目標	付与パラメータ
search	ボールを探索	ボールの方向と距離
dribble	ボールを保持し移動	移動点の方向と距離
pass	味方にパス	味方の方向と距離
receive	パスをレシーブ	味方の方向と距離
shoot	ゴールへシュート	ゴールの方向と距離

3.4 BDI ロボットの概要

BDI エージェントから送られた指示を、BDI ロボットの目標として達成する NXT について述べる。

3.4.1 NXT の立脚性

NXT は立脚性に基づいて基本行為を実行する。つまり、自らなんらかの動作を行なうことで表 2 のセンサを用いて測定した環境情報に適応して、表 3 の基本行為を実行する。

表 2: センサの種類と取得情報

センサの種類	設置箇所	取得情報
赤外線	前面	赤外線ボールの探知
超音波	前面	ボールや障害物との距離測定
コンパス	上部前向	ロボットの向き測定

表 3: 取得した環境情報に適応する基本行為

基本行為	使用センサ	適応する環境情報
shoot	コンパス	回転方向
dribble	コンパス	回転方向
pass	コンパス	回転方向
receive	全て	回転方向、ボール位置と把持状態
search	全て	回転方向、ボール位置と向き正面

3.4.2 コンパスセンサの誤差調整

NXT のコンパスセンサを用いた回転には以下の問題点がある。

- NXT のコンパスセンサはモーターや蛍光灯などが発生する磁気によって誤差が生じる
- NXT の回転時はコンパスセンサも回転するため、NXT が静止した後でない限り NXT の方向が測れないため機敏な動きができない

表 4 は、Web カメラと QPToolkit から取得した方向（北を 0 度として左回りを正とする角度）とコンパスセンサから取得

した方向である。本稿では、QPToolkitの認識は正確であると仮定し、このような表に基づいて、コンパスセンサの誤差調整を行う。たとえば、コンパスセンサから18°を取得した際は、コンパスセンサの誤差-2°を調整して20°と見做す。これは、NXTからPCへはbluetooth通信することができないためNXTからQPToolkitに問い合わせができないこと、また、立脚性に基づいて得た環境情報を用いて即座に動作を行わせたいためである。

表 4: カメラとコンパスセンサの方向の差異

カメラ	0°10°20°30°40°50°...310°320°330°340°350°
コンパス	4°12°18°26°33°42°...332°338°344°350°358°

このため、以下の方法で、実行時の環境に応じた表4のような方向の差異表を、初期設定として求めておく。

1. NXTを真北に向ける
2. BDIエージェントから送られたQPToolkitを用いて得た角度とコンパスセンサから得た角度を表に記憶する
3. NXTを10°程度、左回りする
4. 上記2.から3.を36回繰り返す。

3.4.3 摩擦係数の変化への対応

図3のフィールドはベニヤ板を用いて製作しているため、場所により表面の摩擦係数は異なる。回転などの動作に必要なのは、3.4.4で述べるモーターの予測回転時間である。そこで、摩擦係数の絶対値は必要はなく、以下の方法にて、初期値を1.0とした最新の相対値を求めればよい。ただし、下記は左回転時の計算方法であり右回転時は左右のモーターの回転を逆とする。

1. 相対摩擦係数 μ の初期値を1.0とする。
2. 右のモーターを前進方向、左モーターを後進方向に t 秒間まわして左回転した前後のコンパスセンサ値を誤差調整した後の回転角度 k を求める。
3. BDIエージェントからの指示された方向回転を実現するために予測したモーター回転時間を t' として回転した前後のコンパスセンサ値を誤差調整した後の回転角度 k' を求める。
4. 相対摩擦係数 $\mu' = \mu * ((k' * (t/t'))/k)$ を計算する。
5. $t = t'$, $k = k'$ として、方向回転する際は、3.に戻る。

3.4.4 モーター回転時間の予測

上記の相対摩擦係数の計算式を用いると、回転角度 k' だけ回転するためのモーターの予測回転時間は $t' = t * (k'/k) * (\mu/\mu')$ である。

3.4.5 意図の形成と実行

NXTの意図の形成と実行サイクルは下記の通りである。

1. 意図の形成 環境やBDIエージェントからの指示に適応して選択される未来指向的意図を表5に示す。

表 5: 環境に適応して形成される意図(, は and 条件)

環境	形成される意図
BDI エージェントからの指示	その指示
「ボール不保持」, 「光が強い」	近くを search
「ボール不保持」, 「光が弱い」	遠くを search
「ボール不保持」, 「光が弱い」, 「物体が近い」	横に移動
「ボール保持」	前に移動

2. 熟考 形成した意図の中から優先順位を決定して、次に実行すべき現在指向的意図を表6に基づいて熟考する。たとえば、receive 実行中で「ボール不保持」かつ「光が強い」場合は「近くを search」を優先するのは、正面30cm内にボールを検知した状況であり、receiveを中断して「近くを search」を実行すれば積極的にボールを正面に検知して把持を容易にすることができるため、receiveの成功率を上げることができるという信念である。

表 6: 現在指向的意図の熟考(, は and 条件、/ は or)

実行中の意図	環境	優先意図
receive	「ボール不保持」, 「光が強い」	近くを search
receive	「ボール不保持」, 「物体が近い」	横に移動
dribble	「ボール不保持」, 「光が強い」	近くを search
shoot/pass	「ボール不保持」	前に移動

3. 意図の実行 意図を実行する。また、実行中に新しい指示が送られてきたときに動作途中であっても動作を中断して1.へ移行する。つまり、直近の指示を最優先の意図として形成する。
4. 意図の継続と破棄 single-minded コミットメント戦略に従って、意図の継続と破棄処理を行う。
5. 1.へループする

4. 基本行為の実装と実験

NXTの基本行為を下記のように実装した後、BDIエージェントから指示を送って基本行為に関する実験を行う。

4.1 基本行為の実装

まず、基本行為に含まれる動作パターンとして下記の実装する。

- 回転 全ての基本行為にて使用される動作パターンである。与えられた回転角度を k' として3.4.4に記述したモーター予測回転時間 t' を求め、その時間分だけモーターを回して回転する。回転後は3.4.3に記述した相対摩擦係数 μ' を更新する。その後すぐ、その基本行為を実現する動作パターンを実行する。つまり、回転誤差の調整は直後には行わず前進などの動作パターン実行中に微調整する。
- 前進 目標地点までの距離を入力としてモーター予測回転時間(回転と同様に前回の前進を参考にして計算する)を計算する。前進の実行は1秒単位に分割して行う。たとえば、予測時間が5.6秒とすると6回に分割して行う。ただし最後の回は0.6秒だけ実行する。1サイクル単位に

新しい指示がないかチェックする。また、コンパスセンサ測定値を誤差調整した後、付与された目標地点の角度とのズレに応じて微調整を行う。たとえば、左に向き過ぎていれば、左のモーターのレベルを0.1秒間だけ上げることで右を向くように微調整する。

- 蹴る dribble 時には進行しながらアームを挙げることでボールを押し出す。

これらの動作パターンを用いて、下記の基本行為を実装する。

- shoot ゴールに向かって回転した後、蹴るを実行する。
- dribble ボールを保持して目標地点に向かって回転した後、前進を実行する。
- pass 回転した後、蹴るを実行する。
- receive 味方の方向に回転した後、待機。近くにボールを検知したら search して把持する。
- search まずボールの方向に回転して、ボールまで移動する。その後、左右に少しずつ回転することで、超音場センサを用いてボールを正面に検知すれば、回転を停止する。

4.2 基本行為の実験

表3に挙げた基本行為をそれぞれ20回ずつ実験するために、BDIエージェントは基本行為の指示をそれぞれ20回送信する。成功かどうかの判断は表7に示す。たとえば、dribbleの成功/失敗はBDIエージェントがNXTとボールの位置を1秒単位に計測して、成功条件の領域にNXTが入れば成功、入ることが不可能と判断すれば失敗と見做す。

表7: 実験の成功判断

基本行為	成功の条件
shoot	ボールをゴールできた場合
dribble	ボールが目標位置から半径5cm以内
pass	味方が receive できた場合
receive	ボールを把持できた場合
search	ボールを探索できた場合

4.3 実験結果

基本行為が成功するかどうかの実験結果を表8に掲載する。

表8: 基本行為の成功率

基本行為	成功率
shoot	距離が50cm以内であれば90%、1m以内は50%
dribble	80%、ボールの離してしまう失敗が多い
pass	65%、パスがズレて味方が把持できない
search	80%、ボールに当たって移動させてしまう

目標地点への回転は、ほぼ成功している。ほとんどの失敗は、NXTのボールの扱い方、つまり把持や蹴るなどである。

5. 考察

5.1 微調整の有効性

実験結果から、コンパスセンサの誤差調整や相対摩擦率の計算に基づくNXTの回転に関しては、ほぼ成功している。つ

まり、これら微調整により基本行為の成功率は向上している。実際、これら調整を行わないで実験を行うと成功率は50%以下となる。このことより、これら基本行為の微調整や表6を用いた熟考などにより、single-mindedコミットメント戦略を用いた意図の継続率が増し、基本行為の成功率が向上したと考えられる。また、BDIエージェントは、図5をプランとして実行するため、基本行為が成功裏に終了すれば、状態遷移図に従って共同意図を実現するために有効な次の基本行為を容易に選択することができる。しかし、基本行為の実行途中で失敗すると次状態には遷移するが、何が原因で失敗したのか分からず、ほとんどの場合は失敗のループを嵌ってしまい、共同意図を実現する意思決定は容易でなくなる。

5.2 今後の課題

基本行為の失敗の主な原因は、ボールの扱いにある。この対処としては、NXTの形状を修正する必要がある。たとえば、動作できる自由度を下げること、予期せぬ動作をさせないということなどが考えられる。

また本稿では、コンパスセンサによる誤差は、初期処理においてカメラとコンパスセンサの方向の差分を10°単位に作成した表の按分値から求めた。また、相対摩擦係数は、局所的に連続変化していることを仮定している。どちらの誤差も、場所によって非連続的に変化する場合も考慮し、場所情報を入力パラメータとした回帰学習を行う必要がある。今後の課題である。

6. おわりに

実世界ロボットが、意図を実現するには、センサー誤差や摩擦係数の変化に適応してアクチュエータを調整する必要がある。特に、基本行為が失敗しそうに至った時に、single-commitmentコミット戦略を用いることで、アクチュエータを調整した動作パターンを実行することができれば、基本行為を成功に導く可能性が高まることをNXTを用いて提示した。つまり、アクチュエータの微調整が実世界における基本行為のコミットメント戦略の継続に有効であることを示すと共に、意図を成功裏に実現することができれば、以後の基本行為の意思決定が簡素化できることを提示した。また、今後の課題にて自由度を下げた形状や機械学習の可能性について述べた。

参考文献

- [QPT] QPToolkit Web カメラを使ったかんたん位置測定, <http://kougaku-navi.net/QPToolkit/>
- [Rao 91] Rao, A. S. and Georgeff, M. P.: Modeling Rational Agents within a BDI-Architecture, in *Proc. of International Conference on Principles of Knowledge Representation and Reasoning*, pp. 473-484 (1991)
- [高田 13] 高田 司郎, 新出 尚之, 濱砂 幸裕, 波部 斉, 藤田 恵: アトラクター状態を用いた実世界における基本行為の学習について, 情報処理学会研究報告 2013-MPS-92 (2013), 掲載予定
- [新出 12] 新出 尚之, 高田 司郎, 藤田 恵: 連続した状態空間での合理的エージェントの行為を扱う論理モデルの試み, in *Proc. of JAWS2012* (2012)
- [藤田 12] 藤田 恵, 片山 寛子, 新出 尚之, 高田 司郎: 実世界の多様性に適応したBDIロボットについて, 情報処理学会論文誌数理モデル化と応用, Vol. 5, No. 1, pp. 50-64 (2012)