

格闘ゲームキャラクタの模倣学習

服部裕介 田中彰人 星野准一
筑波大学大学院システム情報工学研究科

格闘ゲームのような対戦型のアクションゲームにおいて、コンピュータが操作するキャラクタ(以下 COM)の行動は変化に乏しく、プレイヤーは繰り返しプレイすることによって COM の行動パターンを憶えてしまい、ゲームに飽きてしまうという問題がある。そこで、本稿ではプレイヤーを模倣学習する手法を用いて、COM の行動パターンを拡張する手法を提案する。本手法では、プレイログを記録し、対戦相手であるプレイヤーの行動パターンを分析し、そのプレイヤーの行動パターンの一部を模倣することで COM の行動パターンを拡張する。本手法を用いることで、試合ごとに行動パターンを拡張し、成長していきける COM を生成することが可能となる。

The imitation learning of the fighting game character

Yusuke HATTORI Akihito TANAKA Junichi HOSHINO
University of Tsukuba, Graduate School of Systems and Information Engineering

In action game, the computer's behavior lacks diversity and human players are able to learn how the computer behaves by playing the same game over and over again. As a result, human players eventually grow tired of the game. Therefore, this paper proposes technique to expand the behavior pattern of COM. To this end, we use imitation learning. In this technique, play log is recorded. And COM expands the behavior pattern by imitating a part of player's behavior pattern. In the result, COM is able to expand the behavior pattern every match and grow up.

1. はじめに

格闘ゲームと呼ばれるゲームのジャンルはアーケードや家庭用ゲーム機などにおいて、長年にわたり人気を集めている。格闘ゲームとは、一対一でパンチやキック、その他特殊な技を駆使し、相手の体力をゼロにするか、一定時間内により多く相手の体力を減らした方が勝利する、というゲームである。

格闘ゲームのような多くの対戦型アクションゲームには、コンピュータが制御するキャラクタ(以下 COM)と対戦する対 COM 戦と、人間同士で対戦する対人戦といったモードがある。対人戦では対戦相手によって行動パターンが異なり、試合ごとに相手の技や戦い方が変化するのに対して、対 COM 戦では COM の行動パターンが有限状態マシンで設定されていたり、スクリプトなどによって予め書き込まれていたりする。そのため COM の行動パターンは変化に乏しい。そして、行動パターンに変化がないので、プレイヤーは COM の行動パターンを憶えてしまう。これは、プレイヤーが COM との対戦に飽きてしまう原因の一つとなっており、多くのプレイヤーは対 COM 戦よりも対人戦を好む傾向がある。しかし、対人戦を行いたいときにその場に相手がいるとは限らないという問題がある。そこで、対 COM 戦でも COM が特徴のある行動パターンをもち、試合をするごとに随時行動パターンを変化することができ

れば、その問題の解消に繋がると考えられる。

我々は、人のプレイしたプレイログからそのプレイヤーの特徴を抽出し、それを基に COM を制御する模倣学習を提案した[1]。この手法では、まず模倣対象となるプレイヤーのプレイログを取得する。そして、そのプレイヤーの行動パターンを分析し、特徴をプレイヤープロフィールとして記録する。そして、それを用いて COM の行動を制御することで、模倣対象であるプレイヤーの特徴を再現する。

本稿では、行動パターンを変化させることができる COM を生成するために、模倣学習の手法を用いて COM の行動パターンを拡張する手法を提案する。人間のプレイヤーが自分の戦い方を変化させる方法の一つに、相手の技を模倣することが考えられる。そこで、本手法では、COM の行動パターンを適宜変化させるために、模倣学習を応用して、対戦相手であるプレイヤーの行動パターンを分析し、特徴を抽出する。そして、得られた行動パターンから模倣する行動を選択、追加することで COM の行動パターンを拡張する。本稿では 2 章で概要について、3 章でベースとなる行動パターンの生成法について説明する。そして、4 章でベースとなる行動パターンを拡張する手法について説明し、実験結果を示す。

1.1 関連研究

ゲームを題材とした研究では、ゲームの対戦相手のモデル化に関する研究[2]が挙げられる。

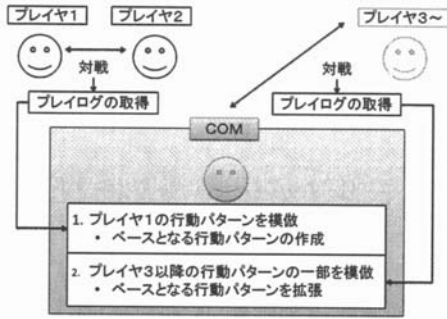


図1 模倣によるベース行動パターンの作成と拡張
Fig1 Creation and expansion of base behavior pattern by imitation

観測データから模倣学習する AI として[3][4][5][6][7][8]が挙げられる。Thurau らはニューラルガス法を用いて人間が操作したキャラクタの3次元経路を模倣し、従来のパスに沿った経路ではなく、より人間らしい移動経路を生成している。また、Bain らや Sammut らは、飛行機やコンテナクレーンなどを例にエキスパートが操縦した際のログから状況と行動をペアにしたルールを抽出し自動操縦を行った。Thurau らによるものは、経路を模倣するものであり、トポロジカルな手法を用いているため離散的な状態空間となる行動の選択の模倣には適していない。また、Bain らや Sammut らのものは決定木により状況と選択される行動の関係を表現しているため、離散的で非連続な状態空間を扱うことができるが、ある状態に対して、選択され得る行動は一つである。これらの研究はAIがエキスパートの操縦を観測することで適切な操縦技能を模倣することを目的としているため、ある状況に対して適切な一つの行動が選択されれば良い。しかし、本稿で扱うゲームのように一つの状況に対して複数の行動が考えられるような場合を再現することはできない。

このように模倣学習などの模倣に関連した研究がなされているが、多くのものはエキスパートの行動を模倣することによって、状況に対して適切な判断を行い、行動を選択することを目的としている。しかし、本稿で扱うCOMが選択する行動は決して適切である必要はなく、ゲームで対戦するキャラクタとして行動するために、特徴のある行動パターンを取得することを目的としている。

2. 概要

本手法を用いて、コンピュータが操作する格闘ゲームキャラクタの行動パターンを生成する流れを図1に示す。ま

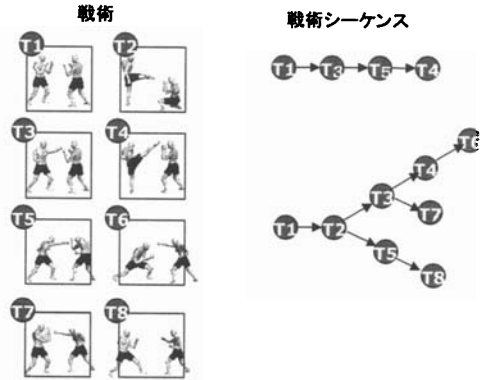


図2 戦術と戦術シーケンス
Fig2 Tactics and sequence of tactics

ず、従来の模倣学習法を用いて対戦相手のプレイヤーのプレイログからそのプレイヤーの行動パターンを抽出する。これがCOMのベースの行動パターンとなる。そしてベースの行動パターンを用いてCOMに別のプレイヤーと試合をさせる。COMは試合をするごとに各試合のプレイログを分析し、対戦相手である各プレイヤーの戦い方の一部を模倣することで行動パターンを拡張していく。このようにしてCOMは自身の行動パターンを拡張し続けることが可能となる。

本手法を用いることで、行動パターンを拡張することによって成長するCOMを生成することができる。自宅で対戦した友人やネットワーク上で対戦した相手、ゲームセンターで対戦した人のプレイログから様々なベースとなる行動パターンを作成する。そして、それぞれの行動パターンを適用したCOMが試合をするごとに、COMは対戦相手であるプレイヤーのプレイログを分析し、行動パターンを拡張していく。その結果、従来の作り込みのCOMではなく、試合を重ねるごとに、行動パターンを拡張し、柔軟に戦い方を変えていくことができるCOMを生成することが可能となる。

3. プレイヤーの模倣行動の生成手法

本章では、COMのベースとなる行動パターンの作成方法について説明する。プレイヤーの特徴を模倣するためには、プレイヤーの特徴をとらえ、その特徴を基に行動を選択する必要がある。格闘ゲームにおいてプレイヤーの特徴は、「どのような状況でどのような行動を選択するか」によって表される。そのため、この手法では毎時刻におけるプレイヤーの行動や間合い等のゲームを構成する状況をプレイヤー

タとして記録する。そして、記録した複数のプレイデータから戦術と戦術シーケンスを取得する。戦術とは「どのような状況でどのような行動をとったか」という情報である。また、格闘ゲームには、行動に決まった順序が現れることも多い。このような、行動の決まった順序は図2のようなシーケンスとして表すことができる。この手法では、このような戦術の組み合わせを戦術シーケンスと呼ぶ。

プレイヤーのプレイデータから得られた戦術、戦術シーケンスの集合を、そのプレイヤーのプレイヤープロフィールとしてCOMに適用することでそのプレイヤーを模倣した行動を生成する。

3.1 プレイデータ

模倣対象となるプレイヤーの行動パターンを得るために、プレイしている最中のゲームを構成する状況を図3のように毎時刻について記録する。格闘ゲームの状況を構成する要素については以下のようなものが挙げられる。「プレイヤーの行動」「相手の行動」「相手の行動のフェイズ」「相手との間合い」「プレイヤーの残り体力」「相手の残り体力」「残り時間」。

以上のように、プレイヤーは“行動のフェイズ”や“相手の行動”などの状況構成要素を基に行動を選択している。したがって、行動の状況構成要素をできるだけ多く記録し、COMが行動を選択する際の行動選択条件として用いれば、より高い再現性が期待できる。

3.2 戦術

プレイデータから、プレイヤーが操るキャラクターがおかれた状況と、その状況に対して選択した行動の対応関係を取

得し、これを戦術とする。戦術は図4のように戦術が選択された際に実行される選択行動と、行動選択直前の状況（プレイヤーの行動、相手の行動、相手の行動のフェイズ、間合い、プレイヤーの残り体力、相手の残り体力、残り時間）を戦術の選択条件として構成する。

プレイヤーによって個人差はあるが、戦術は10試合分のプレイデータから約400~600個取り出すことができた。

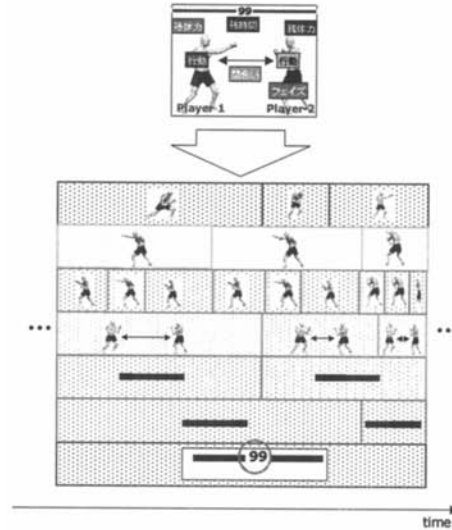


図3 プレイデータの取得

Fig3 Acquire play data

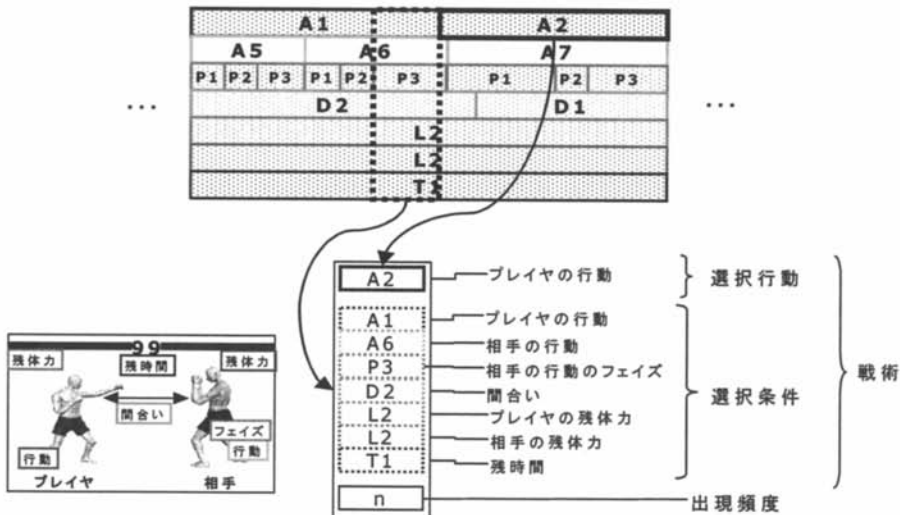


図4 戦術

Fig4 Tactics

3.3 戦術シーケンス

取得したプレイヤーデータから戦術を抜き出していくことで、プレイヤーの戦術の列を取得することができる。この戦術の列からプレイヤーの連続した行動を戦術のシーケンスとして切り取る。これを戦術シーケンスとする。

戦術シーケンスを複数取得すると、同じパターンのシーケンスが何度も観測されたり、同じ状況下でも異なる選択をしているシーケンスが見られる。人間は同じ状況下でも毎回同じ行動を選択するとは限らず、敢えて行動を変えてみるようなことがある。そのような頻繁に選択される行動や稀に選択される行動などの、プレイヤーの行動選択における確率的な要素を再現するために、各シーケンスの出現回数を取得する。戦術シーケンスの出現回数を基に確率的に戦術を選択することでプレイヤーの行動選択傾向を再現する。

3.4 プレイヤプロフィールの適用

この模倣手法では試合の状況をプレイヤーデータとして記録し、そこから戦術、戦術シーケンスを取得する。この戦術と戦術シーケンスをそのプレイヤーのプレイヤプロフィールとする。

このプレイヤプロフィールを適用する際は、戦術シーケンスを戦術グラフにしておく。戦術グラフとは戦術シーケンスをマージしたもので、途中まで同じ戦術で構成されている複数の戦術シーケンスは一つの戦術グラフにまとめることができる。この戦術グラフを探索し、複数の戦術シーケンスを確率的に出現させることでプレイヤーの行動パターンを再現することができる。

3.5 実験

以上の手法を用いてプロフィールを作成した。まず、模倣対象をCOMと対戦させ、模倣対象のプロフィールを作成し、観察によって模倣ができていないことを確認した。そして、複数人の模倣対象の比較を行い、視覚的な違いがあることを確認した。表1はある2つのプロフィールを比較

表1 プロフィールの比較

Table1 Compare with profiles

プレイヤー1	プレイヤー2
左ジャブ (41.4%)	しゃがみパンチ (36.1%)
ハイキック (20.1%)	左ジャブ (19.3%)
右ストレート (20.0%)	右ストレート (14.6%)

したものである。この結果によってプレイヤーごとに使用されるアクションの頻度に違いがあることがわかる。これらを戦術、戦術シーケンスによって再現することでプレイヤーの行動パターンの特徴をCOMに再現させる。

4. 相手の戦術シーケンスの模倣

本章では、戦術シーケンスを拡張することによって行動パターンを変化させることのできるCOMを生成するために、模倣行動を利用する手法を述べる。人間のプレイヤーが自分の戦い方を変化させる方法の一つに、相手の技を模倣することが考えられる。そこで、対戦相手であるプレイヤーのプレイヤプロフィールを、模倣学習を用いて取得し、その一部をCOM自身のプレイヤプロフィールに追加することで、COMの行動パターンを変化させる手法を提案する。

4.1 概要

概要を図5に示す。まず、COMに試合をさせ、その試合のプレイヤーデータを取得する。1~数試合、対戦相手のプレイヤーと試合をさせ、そのプレイヤーデータから対戦相手のプレイヤーのプレイヤプロフィールを作成する。プレイヤプロフィールを作成することで対戦相手のプレイヤーの戦い方の特徴を取得することができる。このようにして作成した対戦相手のプレイヤーのプレイヤプロフィールと、COMのプレイヤプロフィールを比較、評価し、追加する戦術シーケンスを決定する。

4.2 追加する戦術シーケンスの評価

人が相手の技を模倣する場合、その方法は複数考えることができる。そこで本稿でも、3つの評価手法を用意し、確率的に選択することにした。以下に評価手法の説明を記述する。

4.2.1 既に使用している戦術シーケンスと共通部分のある戦術シーケンス

既に使用している技のバリエーションとして使用できる技を探すという手法は、戦い方を工夫する上で良く利用される手法である。そこで本稿でもこのようなキャラクターの成長を再現するための戦術シーケンスの追加方法を提案する。本稿ではこの評価方法を「評価方法1」とする。

対戦相手のプレイヤーのプレイヤーデータから生成したプレイヤプロフィールに含まれる戦術シーケンスと、COMのプレイヤプロフィールに含まれる戦術シーケンスを比較して、同じ行動が長く続くほど評価値を高く設定する。これは、同じ行動が長く続くほど、特徴が似ていると考えられるためである。ただし、もともと長い戦術シーケンスでは同じ行動が長く続いても、全体から見るとそれほど特徴を表したものにしない場合がある。そのため、ある値を

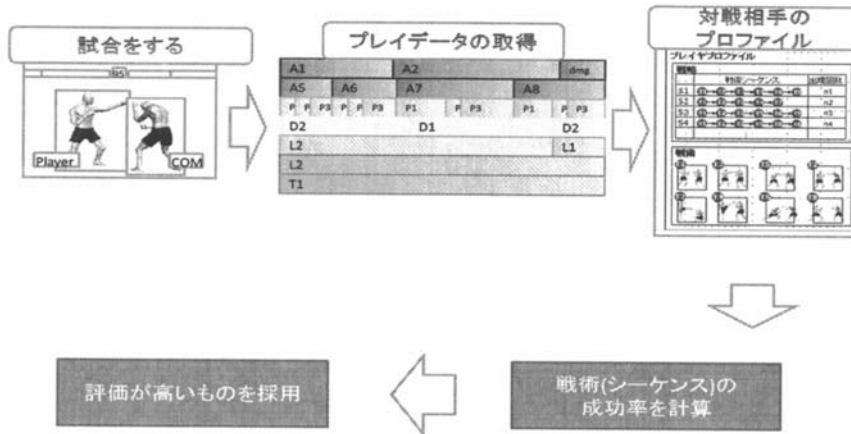


図5 相手の戦術シーケンスの模倣

Fig5 Imitation of enemy sequence of tactics

設定して戦術シーケンスの全長との比率が一定の値以上でなければならないという条件を設ける。この値によって、どのような戦術シーケンスを模倣に適したものとするかの基準が変わるため、キャラクタによってこの値を変更させることで、キャラクタごとに成長の特徴に変化をもたせることができる。以下に評価値を求めるための式を示す。

$$H = \begin{cases} a & (a \geq \frac{L_{\min}}{n}) \\ 0 & (a < \frac{L_{\min}}{n}) \end{cases} \quad (1)$$

Hは評価値、aは比較する2つの戦術シーケンスの共通部分の長さ、 L_{\min} は比較する2つの戦術シーケンスのうち、短い方の戦術シーケンスの長さ、nは閾値を決定するための定数である。

4.2.2 自分が使用する行動を含んだ戦術シーケンス

自分の使用する行動を効果的に使用している相手の戦い方を真似るとい手法も、人が他人の戦い方を模倣する際によく利用される手法である。本稿ではこの評価方法を「評価方法2」とする。

このような模倣の場合も効果的であるということが重要であるため、成功率を評価基準とした。その他、出現回数と戦術シーケンスのシーケンス長をもとに相手の戦術シーケンスから模倣する戦術シーケンスを決定する。模倣する戦術シーケンスを選択する際は、相手の戦術シーケンスの中から、自分が使用している行動を含んだ戦術シーケンスを抽出し、成功率、出現回数、シーケンス長をもとにレイヤの戦術シーケンスに含まれる戦術の行動のうち、COMのプレイヤープロフィールに存在しない行動の数、 W_s ,

模倣する戦術シーケンスを決定する。この3つの重みは、重みづけを調整することで、キャラクタごとに成長の特徴に変化をつけることができる。

$$H = \begin{cases} W_s S + W_n n + W_L L & (m \neq 0) \\ 0 & (m = 0) \end{cases} \quad (2)$$

Hは評価値、Sは成功率、nは出現回数、Lは戦術シーケンスのシーケンス長、mは現在評価している対戦相手のプレイヤーの戦術シーケンスに含まれる戦術の行動のうち、COMのプレイヤープロフィールに存在しない行動の数、 W_s , W_n , W_L はそれぞれにかかる重みを表す。

4.2.3 自分が使用しない行動を含んだ戦術シーケンス

自分の使用しない行動を効果的に使用している相手の戦いかたを真似るとい手法も、人が他人の戦い方を模倣する際によく利用される手法である。本稿ではこの評価方法を「評価方法3」とする。

このような模倣の場合も上記評価方法と同様に、成功率、出現回数、シーケンス長を重みに用いて評価値を決定する。上記評価手法では自分が使用する行動で構成された戦術シーケンスの中から模倣する戦術シーケンスを決定したが、この評価方法では自分の使用しない行動を含む戦術シーケンスから模倣する戦術シーケンスを決定する。

$$H = \begin{cases} W_s S + W_n n + W_L L & (m \neq 0) \\ 0 & (m = 0) \end{cases} \quad (3)$$

Hは評価値、Sは成功率、nは出現回数、Lは戦術シーケンスのシーケンス長、mは現在評価している対戦相手のプレイヤーの戦術シーケンスに含まれる戦術の行動のうち、COMのプレイヤープロフィールに存在しない行動の数、 W_s , W_n , W_L はそれぞれにかかる重みを表す。

4.3 戦術シーケンスの決定

以上の複数の評価方法を用いて対戦相手のプレイヤーから模倣する戦術シーケンスを決定する。各試合が終わるごとに対戦相手のプレイヤーの戦術シーケンスの模倣をするかどうか、また、どの評価方法を用いて相手から模倣する戦術シーケンスを決定するかを確率的に決定する。模倣をする確率と各模倣手法を選択する確率をキャラクタごとに調整することによって、キャラクタごとに成長の特徴に変化をつけることができる。

4.4 実験

以上の手法を用いて、あるプレイヤープロフィールを拡張させる実験を行った。拡張する前のプレイヤープロフィールは10試合分のプレイデータから取得し、戦術数が549個で、戦術シーケンス数は48個であった。このプレイヤープロフィールを用いてあるプレイヤーと10試合の試合を行い、追加された戦術、戦術シーケンスを調べた。今回は10試合の評価なので、相手のシーケンスを模倣する確率を高めに設定した。

表2に示す通り、10試合で戦術シーケンスが4個、戦術が21個それぞれ増加している。戦術が増加しているのは対戦相手のプレイヤーから模倣した戦術シーケンスの中に自分が使用していない戦術が含まれる場合に追加されるためである。

実験の結果、1, 2試合目では、それぞれ「評価方法1」、「評価方法2」が選択されたが、試合データが少なかったため模倣条件を満たせず戦術シーケンスは追加されなかった。3, 4試合目にはそれぞれ「評価方法2」「評価方法1」によって戦術シーケンスが追加された。3試合目に追加された戦術シーケンスは前進してジャブ(パンチ)を繰り出すというもので、「評価方法2」の特徴通り、もともと使用頻度が高かったジャブを使用したものであった。4試合目に追加された戦術シーケンスは、間合いを開けてからハイキックでふら付かせて接近して連続パンチで攻撃するという戦術シーケンスであった。もともと間合いを開けてハイキックで攻撃する特徴があったCOMの戦術シーケンスと似ていたため「評価方法1」で得られた。その後5, 6, 7, 8試合目は試合が追加されず、9, 10試合目にそれぞれ「評価方法3」「評価方法1」によって戦術シーケンスが追加された。9試合目で「評価方法3」で得られた戦術シーケンスはCOMがもともとは使用していなかったミドルキックを使用したものであり、「評価方法3」の特徴から得られた戦術シーケンスといえる。

このように相手の戦術シーケンスを模倣することによって行動パターンを拡張することができる。

表2 模倣結果

Table2 Result of imitation

	戦術	戦術シーケンス
模倣前	549	48
模倣後	570	52

5. おわりに

本稿では、ゲームログを用いたプレイヤーの模倣手法を応用し、COMの行動パターンを拡張する手法を提案した。この手法ではゲームログの情報をもとにした模倣学習を用いて、対戦したプレイヤーの特徴ある行動パターンを模倣することができる。この手法を用いることによって従来の作りこまれたCOMではなく、試合を重ねるごとに行動パターンを拡張し、柔軟に戦い方を変えていけるCOMを生成することができる。

今後は、対戦相手のプレイヤーから模倣する戦術シーケンスを決定する際の評価方法について更に検討を進めていきたい。対戦相手のプレイヤーの戦術シーケンスとCOMが既に使用している戦術シーケンスの類似度に基づいた判定ができると、より特徴を明確にしたCOMを作成できると思われるため、戦術シーケンス間の類似度を判定する手法を検討したい。

参考文献

- 1) 服部裕介, 田中彰人, 星野准一: 対戦型アクションゲームにおけるプレイヤーの模倣行動の生成. 第17回ゲーム情報学研究会, 情報処理学会, (2007)
- 2) 泰地真弘人, 池上高志: ゲームにおける学習プレイヤーのダイナミクス, 認知科学, Vol.6.No.1, pp.21-30(1999)
- 3) C. Sammut, S. Hurst, D. Kedzier, and D. Michie. Learning to fly. In Proceedings of the Ninth International Conference on Machine Learning, p.p385-393, (1992)
- 4) Van Lent M, Laird J. Learning Procedural knowledge through observation. In Proceedings of International Conference On Knowledge Capture. ACM Press, p.p179-186, (2001)
- 5) M. Bain, C.Sammut. A framework for behavioral cloning. In S. Muggeleton, K. Furukawa, D. Michie, editors Machine Intelligence 14, Oxford University Press, (1995)
- 6) C. Thureau, C. Bauckhage, and G. Sagerer. Learning Human-Like Movement Behavior for Computer Games. In Proc. Int. Conf. on the Simulation of Adaptive Behavior, pp.315-323. MIT Press (2004)
- 7) C. Thureau, C.Bauckhage, and G. Sagerer. Imitation learning at all levels of Game-AI. In Proc. Int. Conf. on Computer Games, Artificial Intelligence, Design and Education, pp.402-408 (2004)
- 8) C. Bauckhage, C. Thureau, and G. Sagerer. Learning Human-like Opponent Behavior for Interactive Computer Games. In B. Michaelis and G. Krell, editors, Pattern Recognition, Lecture notes in Computer Science 2781, pages 148-155. Springer-Verlag, (2003)