

F-039

缶蹴りゲームにおける鬼の戦略学習 Strategy Learning in Kan-Keri Game

糸山 隆司†
ITOYAMA Takashi

六沢 一昭‡
ROKUSAWA Kazuaki

1. はじめに

「缶蹴りゲーム」は、鬼が守る缶を蹴る遊びである。この缶蹴りゲームは、「鬼ごっこ」と「隠れん坊」の要素を持つ。本稿は、缶蹴りゲームにおいて鬼に戦略を学習させるシステムについて述べる。

2. 扱う世界

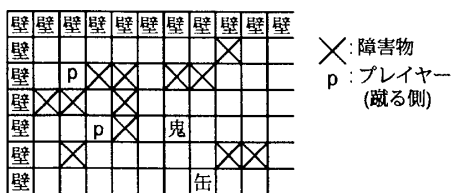
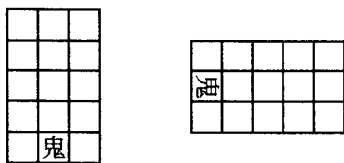


図 1: 仮想世界

本システムが仮定する世界を図1に示す。この世界には鬼とプレイヤー、缶、壁、障害物が存在する。



(a)鬼が上向き
(b)鬼が右向き

図 2: 視界

視界 鬼とプレイヤーは、3×5の視界を持っている(図2)。(視界外から)鬼の視界内にプレイヤーが入ったことを、プレイヤーを発見したと呼ぶ。

行動 鬼とプレイヤーの行動は、前進、右を向く、左を向く、待機のいずれかである。

プレイヤーの行動決定 プレイヤーは缶の方向と自分の周囲8マスの状況から、各行動を実行する確率を算出し、その確率に従って行動を決定する。

試合終了 プレイヤーが缶の位置に移動した場合、プレイヤーの勝ちとなる。一方、プレイヤーを発見した鬼がプレイヤーよりも早く、缶の位置に移動すると、そのプレイヤーを捕まえたことになる。すべてのプレイヤーを捕まえたら鬼の勝ちとなる。鬼かプレイヤーのどちらかが勝ったら一試合終了とする。

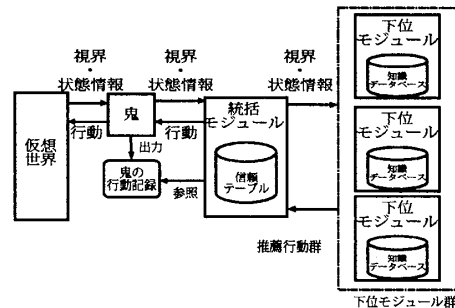


図 3: システム構成

表 1: 一つの「行動」の構成

視界	状態情報	行動	行動決定方法
p	a	前進	参照:A
q	b	待機	参照しない

3. システム構成

下位モジュール 自分の知識データベースを参照して鬼の行動を決定し、統括モジュールにその行動を推薦する。また、知識データベースを更新する。

統括モジュール 下位モジュール群から推薦された行動群から一つを選び、鬼が行う行動を決定する。また信頼テーブルを更新する。

鬼の行動記録 一つの試合における鬼の「行動」列と鬼の勝敗が記録されている。一つの行動は、視界と鬼の状態情報、その時に行った鬼の行動、行動決定方法からなる(表1)。行動決定方法は、信頼テーブルを参照(選択下位モジュール)あるいは、参照しないである。

信頼テーブル 鬼の視界と、その視界における下位モジュールそれぞれに対する信頼度を一組とする知識を蓄積する。(表2)

知識データベース 鬼の視界と状態情報、(その時における)行動、そしてその行動の評価点を一組とする知識を蓄積している。(表3)

鬼の状態情報は、缶から鬼の現在地までの距離、残りプレイヤーの人数、鬼がプレイヤーを発見しているかの三つのいずれかである。下位モジュールによって、学習及び行動決定に用いる鬼の状態情報が異なる。

†千葉工業大学 大学院 情報科学研究科 情報科学専攻
‡千葉工業大学 情報科学部 情報工学科

表 2: 信頼テーブルの構成

視界	各下位モジュールの信頼度		
	A	B	C
p	a	b	c
q	d	e	f

表 3: 知識データベースの構成

視界	状態情報	行動	行動の評価点
p	a	右を向く	b
p	a	前進	d
q	c	待機	e

4. 鬼の行動の決定

鬼の行動を決定するのは統括モジュールである。統括モジュールは、各下位モジュールからの推薦行動群から一つを選択しそれを鬼の行動とする。

4.1 下位モジュールの行動決定

知識データベースは、ある視界と状態情報における行動と、その行動が良いか/悪いかを示している。これが下位モジュールの知識である。

下位モジュールは、知識データベースの知識の中から、与えられた鬼の視界及び鬼の状態情報と一致する知識を抜き出し、抜き出された知識群の中からもっとも評価点の高い知識の行動を推薦行動とする。

一致する知識が知識データベースに存在しない場合は、乱数を用いて決定した行動を推薦行動とする。

4.2 統括モジュールの行動決定 (鬼の行動の決定)

信頼テーブルは、ある視界における下位モジュールの信頼度を示す。これが統括モジュールの知識である。

統括モジュールは、下位モジュール群から推薦行動群を受け取るとまず多数決を試みる。同じ行動が二つ以上あったら、その行動を選択する (例えば、前進、右を向く、前進。ならば、前進を選択する)。推薦された行動が全て異なる (例えば、左を向く、待機、前進) 場合は、信頼テーブルを参照して、与えられた視界と一致する知識を抜き出し、もっとも信頼度の高い下位モジュールを選択し、その下位モジュールの行動を鬼の行動とする。一致する知識がない場合は、乱数を用いて下位モジュールを選択し、その下位モジュールの行動に決定する。

5. 学習

5.1 統括モジュールの学習

鬼の行動記録を参照して信頼テーブルを更新することが統括モジュールの学習である。鬼の行動記録中のすべての「行動」のうち、信頼テーブルを参照して決定した「行動」のそれぞれに対して以下の処理を行う。

「行動」の視界と一致する知識を信頼テーブルから抜き出し、抜き出された知識における該当する下位モジュール

に対する信頼度に勝敗点 (勝ちプラス点、負け-マイナス点) を加点する。一致する知識が信頼テーブルに存在しない場合は、行動の内容を新規の知識として追加する。

5.2 下位モジュール群の学習

鬼の行動記録を参照して知識データベースを更新することが下位モジュールの学習である。鬼の行動記録中の「行動」の視界、状態情報と一致する知識を知識データベースをから抜き出し、抜き出された知識における行動の評価点に勝敗点を加点する。一致する知識が知識データベースに存在しない場合は、行動の内容を新規の知識として追加する。

6. 実験結果

10×10サイズの仮想世界で、プレイヤー二人として学習を行わせた。n 試合 (n=10,20,50,...10000) 学習させた鬼に、1000 試合毎に蹴りゲームを行わせることを1セットとして、それを10セット行った。その時の鬼の勝率と知識量を図4と図5に示す。

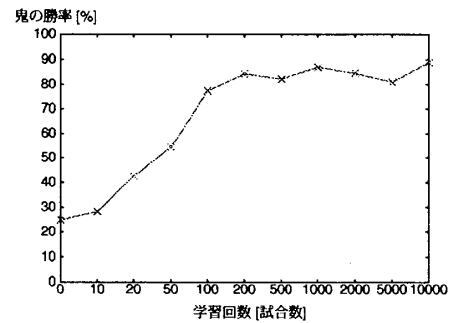


図 4: 鬼の勝率の推移

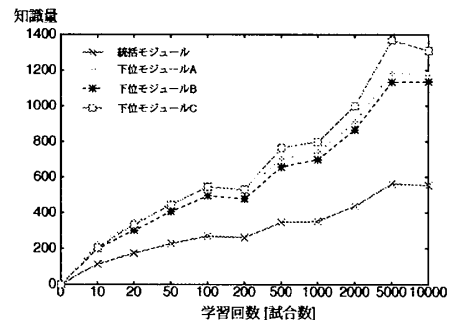


図 5: 各モジュールの知識量の推移

図より、下位モジュールの学習回数が増えると、鬼の勝率も上昇することが確認できた。しかし、学習回数が100試合を越えた辺りからは、知識量は増加するが、鬼の勝率の上昇率が鈍くなっている。これは追加される、より良い知識が少なくなった為であると予想される。

7. おわりに

「缶蹴りゲーム」における鬼に戦略を学習させるシステムを作成した。学習を重ねることで鬼の勝率が増加することが確認できた。