

2-108 不完全情報ゲームにおける意志決定プロセス

Decision Making Process in Game with Imperfect Information

○高橋千晴 (筑波大) 鬼沢武久 (筑波大)

Chiharu TAKAHASHI, Onisawa Lab., Institute of Engineering Mechanics and Systems, University of Tsukuba,
1-1-1, Tennodai, Tsukuba 305-8573, Japan
Takehisa ONISAWA, University of Tsukuba

This paper aims at the construction of a playing system of Seven-card-Stud Poker game and that plays a bluff. This system makes its strategy, for the assessment of the comparative strength of a poker hand, a decision making on amount of a bet and whether to drop or not, by using the fuzzy inference to consider the tendency of each opponent player. Subjects experiments are performed to analyze the behavior of the presented system. The experimental results show that a system adapts its strategy to opponent players and sometimes plays a bluff, and that human players are sometimes deceived by this system's bluff.

Key Words: fuzzy inference, game with imperfect information, decision making, bluff

1. はじめに

ゲーム理論⁽¹⁾では、プレーヤが常に理性的で合理的な行動をとるとする仮定のもとで、与えられた条件下で自分の利益を最大にする戦略を理論的に論じている。しかし人間プレーヤの場合、常に理性的で合理的な戦略をとるとは限らないため、対戦相手に応じた戦略やヒューリスティックな戦略をたてていくことが重要である。

ゲームプログラミングの研究には、チェスや将棋など完全情報ゲームを題材としているものと、ポーカーやブリッジなどの不完全情報ゲームを題材にしているものがある。前者のゲームでは相手の手が完全に分かっているものの、場合の数が非常に多いため必勝法は見つかっていない。そのため、強いアルゴリズム開発を目指した研究が多い⁽²⁾⁽³⁾。後者のゲームでは、プレーヤごとに知ることのできる情報とできない情報があるため相手プレーヤの手の強さを推論する必要がある⁽⁴⁾⁽⁵⁾。そのため、完全情報ゲームと比べると難しさがあり、チェスや将棋のシステムと比べるとまだ弱い。

そんな中で、セブencardスタッドポーカー⁽⁶⁾を題材にして、手役の強弱、賭け金の多少などの言語表現を用いて、自分の手の優劣推定、ドロップの判断、賭ける金額の判断を推論してゲームを行うシステムを開発している研究がある⁽⁷⁾⁽⁸⁾⁽⁹⁾。これらの研究では、優劣推定の際、対戦相手の傾向に応じて推論を行うために変形ファジィルールを用いてその有効性を示しているものの、伏せられたカードを開いていくプロセスまではふれられておらず、また、ポーカーでの常套手段であるハットリについては考慮されていない。

本研究では従来の研究を拡張して、カードの配りはじめから伏せられたカードを開いて、ショウダウンにいたるまでのプロセス、および、カードを開くかけひきの場面ではハットリを意図的に行う戦略を考慮したプレイングシステムを構築する。そして、本システムと人間プレーヤとでポーカーゲームを行う被験者実験で得られたデータを分析し、本システムの有効性を検討する。

2. システムの構成

2.1 全体構成 システムの全体の流れ図を図1に示す。本システムは大きく分けると、優劣推定、ドロップの判断、賭け金の決定、開くカードの選択、ルールの修正の5つからなる。

ゲーム中の実際の意志決定は優劣推定、ドロップの判断、賭け金の決定、開くカードの選択、それぞれで行われる。最

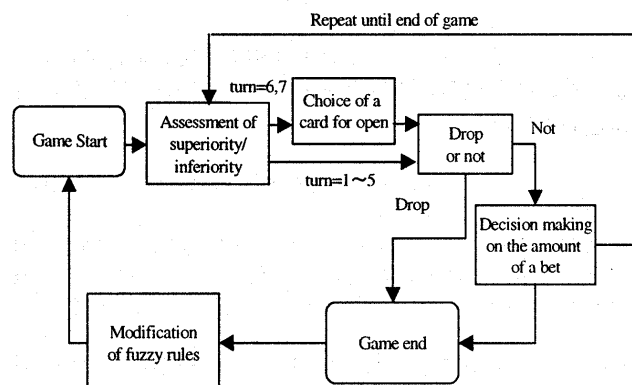


Fig. 1 Flow chart of a playing system

初に裏向きのカード (ホールカード) 2 枚と表向きのカード (アップカード) 1 枚を配る。ここで、最初のベッティングインターバル (お金を賭ける、またはドロップをするかの決定) となる。先攻と後攻の賭け金と同じになれば、次のカードを配り、2 ターン目となる。「ターン」とは、ある 1 回のゲームにおいてシステムが優劣推定を行っているのが何回目のベッティングインターバルなのかを示す値である。4 ターン目までは表向きに、5 ターン目は裏向きにカードを配る。ベッティングインターバルごとに優劣推定、ドロップの判断、賭け金の決定を行う。開くカードの選択は 6 ターン目と 7 ターン目のみで、どのホールカードを開くかというかけひきが行われる。7 ターン終わった時点で両プレーヤともドロップしていない場合は、残りの 1 枚のホールカードを開いて手役の強さを勝負する (ショウダウン)。

2.2 優劣推定 7 枚配られるカードのうち 3 枚のホールカードは配られた本人しか見ることができないため、見えている他のカードや賭け金の額、ゲームの進行状況などから、自分の手役が相手に比べて優勢なのか劣勢なのかを推定しなくてはならない。本研究ではこの推定を優劣推定と呼ぶ。

2.2.1 優劣推定の流れ まず、相手の見えているカードの情報から相手手役の期待値に対するシステムの手役の強さを求める。次に、求められた手役の強さ、賭け金、ターンなどの情報をもとにして、ファジィ推論によって状況ごとの相手プレーヤの傾向に応じた相手手役の強さの期待値に対するシステムの手役の強さの調整度を求める。この調整度に応じてシ

システムの手役の強さの調整を行い、優劣推定とする。

2.2.2 相手手役の期待値に対するシステムの手役の強さ

システムの手番になったら、見えていない相手のカードに考え得る全てのカードの組み合わせを代入し、対戦相手がどの手役を持っているか、それぞれの手役ごとにできている確率を調べる。それにより、相手手役の累積分布関数 $F(x)$ が求まる。ナッシングからロイヤルストレートフラッシュまでの手役にそれぞれ 0 から 9 までの点数を割り当て、対戦相手の手役の点数の期待値を計算する。

この全数探索の際に、対戦相手の手役の点数の期待値を計算する。まず、対戦相手の手役として可能な手役が n 種あるとする。その n 種の手役の点数をそれぞれ $s_i (i=1, 2, \dots, n)$ とする。点数が s_i である手役ができる確率を $p_1(s_i)$ とすると、対戦相手の手役の点数の期待値は、(1)式で求めることができる。

$$hum_expect = \sum_{i=1}^n s_i p_1(s_i) \quad (1)$$

システムの手役点数を com とおき、「相手手役の期待値に対するシステムの手役の強さ」を exp とすれば、 exp は式(2)のように定めることができる。

$$exp = F(com) - F(hum_expect) \quad (2)$$

この値が大きいくほど、システムの手役が優勢である確率が高いことになる。この値は期待値の調整度を求めるファジィ推論の前件部に用いられ、相手の傾向などによって調整される。

2.2.3 手役の強さの調整度の推論 If “相手手役の期待値に対するシステムの手役の強さ is $\tilde{A}$ ” and “正規化した対戦相手の賭け金 is $\tilde{B}$ ” and “7で正規化したターン is $\tilde{C}$ ” then “基準値の調整度 is $\tilde{D}$ ”

$\tilde{A}$: Weak, Strong(図 2(a)) $\tilde{B}$: Few, Middle, Much(図 2(b))
 $\tilde{C}$: Early, Middle, End(図 2(c)) $\tilde{D}$: Negative, Positive(図 2(d))の推論ルールを利用して手役の強さの調整度の推論を行う。表 1 に推論ルールを示す。ターン数は 1 から 7 の整数値を取ることとなるが、ルール中では 0 から 1 の範囲に正規化して用いる。対戦相手の賭け金についても、場に出ている合計賭け金で割ることによって 0 から 1 に正規化する。ルール中のそれぞれの要素に対するファジィ集合を図 2 に示す。

表 1 のファジィルールは後件部が確定していない。その代わり、ルールごとに傾向値というパラメータを設定している。表 1 中に $t_{ij} (i=1, 2, j=1, 2, 3)$ と示してあるのが傾向値である。傾向値は $[0, 100]$ の実数であり、対戦相手の傾向に応じて変更される。この表は、 $t_{ij}/100$ の確率で Positive を、 $1-t_{ij}/100$ の確率で Negative を取ることを意味する。 $\tilde{C}$ = Middle, End についても表 1 と同様に表現される。なお、傾向値の初期値は 50 とする。

まず、状況 k に適合したルールの数を k とする。このとき、それぞれのルールの後件部を Positive か Negative のいずれかに定めるとすると、 2^k 通りの後件部の組み合わせが考えられる。各ルールの後件部は $t_{ij}/100$ の確率で Positive をとり、 $1-t_{ij}/100$ の確率で Negative をとる。それぞれの後件部の組み合わせを $c_i (i=1, 2, \dots, 2^k)$ とし、後件部の組み合わせが c_i となる確率を $p_2(c_i)$ と表す。後件部の組み合わせが c_i となったときに Mamdani の方法⁽¹⁰⁾を用いてファジィ推論し、重心法で非ファジィ化することで得られる推論結果を $r(c_i)$ とする。変

Table 1 Fuzzy rules with changeable consequent parts

$\tilde{D}$ ($\tilde{C} = \text{Early}$)		$\tilde{B}$		
		Few	Middle	Much
$\tilde{A}$	Weak	Positive t_{11} Negative	Positive t_{12} Negative	Positive t_{13} Negative
	Strong	Positive t_{21} Negative	Positive t_{22} Negative	Positive t_{23} Negative

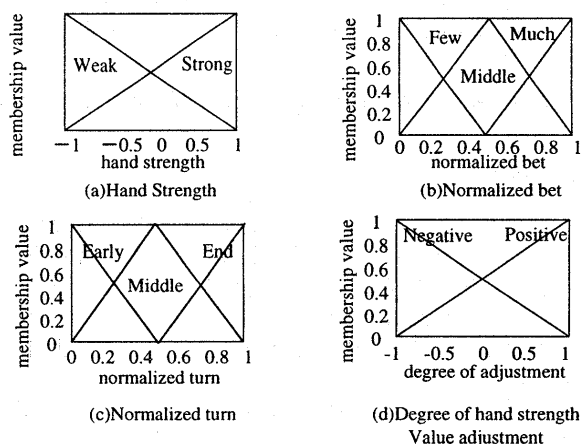


Fig. 2 Membership functions for adjusting hand strength value

形のファジィルールを用いたときの推論結果を $infer$ とすると、式(3)で求めることができる。

$$infer = \sum_{i=1}^{2^n} r(c_i) p_2(c_i) \quad (3)$$

このような変形ファジィルールを採用する理由は、後件部が確定しているファジィルールでは対戦相手に作戦が読まれてしまう恐れがあるためである。

この調整度の値を用いて、2.2.2 によって求めた手役の強さの期待値を調整する。調整は、人間プレーヤの実際の手役点数を hum としたとき、ある特定の状況において $F(com) - F(hum)$ のとりうる値の範囲を調べ、その範囲内で行う。累積分布関数が求めれば $F(com)$ は特定の値となる。このように特定の値に決まった $F(com)$ の値を p とすると、 $F(com) - F(hum)$ は $[p-1, p]$ の範囲をとることになる。 $F(com) - F(hum_expect)$ を(4)式によって相手の傾向やゲーム状況に応じた値へ調整し、 $[p-1, p]$ の範囲内で相対的な優劣推定値へ変更する。調整後の値を優劣推定結果として出力する。

$$result = \begin{cases} exp + (p - exp) \times infer & (0 \leq infer \leq 1) \\ exp + \{exp - (p - 1)\} \times infer & (-1 \leq infer \leq 0) \end{cases} \quad (4)$$

2.3 開くカードの選択 6ターン目と7ターン目はお互いのプレーヤがホールカードのうちから 1 枚を選択し、そのカードを表にしてベッティングインターバルを行う。そのため、どのカードを開くかというかけひきが重要になる。5ターン目で 7 枚のカード全てが出そろうので、このときの手役がそのゲームにおける手役となる。通常、5ターン目で劣勢ならば

ドロップをするというのが妥当な戦略であるが、表になっているカードによってはゲームを続行することで、相手にハッタリをきかせることが可能となる場合もある。また自分が優勢である場合も、強いカードを見せるか、弱いふりをするかなどの戦略がある。

本システムでは、開くカードの選択において、ポーカーゲームの常套手段であるハッタリをシステムに取り入れる。開くカードの選択は以下の手順により行う。

2.3.1 ポテンシャルの算出 開くカードを選択するために、あるホールカード1枚を表にしたと仮定し、残りのホールカードに考えられる全てのカードの組み合わせを代入して、みかけの手役の強さの期待値（ポテンシャルと呼ぶ）を求める。ポテンシャルを求める際に、対戦相手の手役として可能な手役が m 種あるとする。その m 種の手役の点数をそれぞれ $u_i (i=1,2,\dots,m)$ とする。点数が u_i である手役ができる確率を $p_3(u_i)$ とすると、対戦相手のポテンシャルは(5)式で求めることができ、このポテンシャルの計算をすべてのホールカードに対して行う。

$$potential = \sum_{i=1}^m u_i p_3(u_i) \quad (5)$$

2.3.2 実際に開くカードを決定 2.2.3 の優劣推定において、自分が優勢であると判断していて、かつ、自分の手役がすでにアップカードで完成している場合は、乱数を用いてランダムに開くカードを決定する。それに対し、隠れているカードを開くと自分の手役が相手に見えるようになる場合は、対戦相手に悟られないために強く見えるカードは隠しておくのが一般的な戦略であると考えられる。しかし、毎回同じルールで開いていたのでは、ハッタリをする場合に見抜かれてしまう恐れがある。そのため、乱数により 80% の確率で最もアップカードが弱く見えるカードを、残りの 20% の確率で次に弱く見えるカードを開く。

劣勢の場合は、明らかに自分が弱いという状況ではハッタリを行っても最終的に大きな損失になってしまう。よって、2.3.1 で算出したポテンシャルが、実際の手役点数より大きく上回る場合はそのカードを開き、そのようなカードがない場合はランダムに 1 枚を開き、次のドロップの判断部によって必ずドロップを選択する。

2.4 ドロップの判断 優劣推定の結果が優勢である場合はドロップしない。劣勢と判断した場合のみドロップの判断に入る。ドロップの判断は図 3 のような流れで行われる。ドロップのファジールールは表 2、ドロップの意図のメンバーシップ関数は図 4 に示す。ドロップの推論ルールは If “正規化した対戦相手の賭け金 is $\tilde{E}$ ” and “正規化したターン is $\tilde{F}$ ” then “ドロップをする意図 is $\tilde{G}$ ”

$\tilde{E}$: Few, Middle, Much(図 2(b)) $\tilde{F}$: Early, Middle, End(図 2(c))
 $\tilde{G}$: ZO, PS1, PS2, PB1, PB2(図 4) で、以下の三つの知識に基づいている。

1. ゲームの序盤では賭け金も少なく、その後のカード次第で強くなる可能性があるためあまりドロップは考えない
2. ゲームの中盤では相手の賭け金が大きくなってきたらドロップを少し強く考え始める
3. ゲームの終盤で、劣勢と推定した場合は強くドロップを考える

ただし、カードを開いていく第 6 ターンと第 7 ターンで劣勢と推定した場合は、2.3.1 により求めたポテンシャルの値が

大きければドロップは行わないものとし、そうでない場合は必ずドロップを行う。

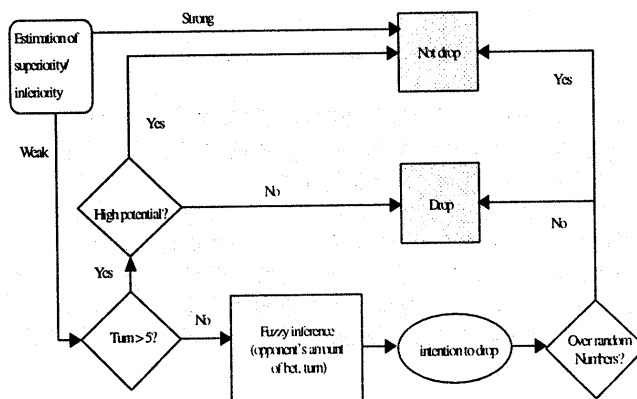


Fig. 3 Flow chart of decision making on drop

Table 2 Fuzzy rules for making decision on drop

$\tilde{G}$		$\tilde{F}$		
		Early	Middle	End
$\tilde{E}$	Few	ZO	PS1	PB1
	Middle	ZO	PS2	PB2
	Much	ZO	PS2	PB2

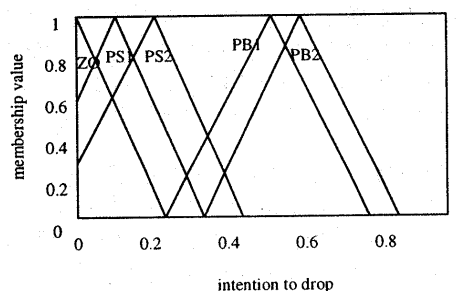


Fig. 4 Membership function to decide drop or not

2.5 賭け金の決定 ポーカーのルールによって定められる賭け金の上限の値を *upper*、下限の値を *lower* と表す。これらの値は、各ベッティングインターバルごとに異なってくる。本システムでは、賭け金の決定を簡単に扱うために $a \in [0,1]$ なる実数値 a を用い、0 を *lower* に、1 を *upper* に対応させる。システムの賭け金を C_{bet} とおき、式(6)によって求めることにする。これにより、どのベッティングインターバルにおいても a を 1 つ選ぶことによって賭け金を決定することができる。

$$C_{bet} = lower + a \times (upper - lower) \quad (6)$$

ただし、 a は実数であるので、 C_{bet} が整数とならない場合は小数点以下を切り捨てることで整数化する。

賭け金の決定は優劣推定の結果に応じて行うこととする。優劣推定の結果が優勢である場合は賭け金を多く賭け、劣勢である場合は賭け金を少なく賭けることが、通常は妥当な戦略であると考えられるが、常に優劣推定の結果と賭け金の額が対応していたのでは、対戦相手にシステムの戦術を見抜かれる恐れがあるため、賭け金の決定も優劣推定の結果のみからではなく、乱数を用いて決定される賭け金に不確定性を持

たせる。そして、優劣推定値である *result* を用いて、パラメータ *a* を式(7)により決定する。

$$a = 0.5 \times \text{result} + 0.5 + \text{noise} \quad (7)$$

ここで、*result* は $[-1,1]$ の範囲の実数であり、*noise* は平均 0、分散 0.1 の正規乱数である。計算結果が 1 以上の場合と 0 以下の場合は、それぞれ計算結果を 1、0 であるとする。

また、5 ターン目以降に劣勢でハッキリをする場合は、実際の優劣推定値に関わらず、ポテンシャルの値にふさわしい額をかけなければならない。そのため、ポテンシャルを 9 で割ることによってポテンシャルを正規化する。ポテンシャルを正規化した値を式(7)の *result* の代わりに用いる。このとき *noise* = 0 とする。

2.6 ルールの修正 ルールの修正部では、優劣推定で用いられている、手役の強さの調整度を求めるルール中の傾向値を修正する。

一回のゲームが終わるごとに、優劣推定の結果と実際のカードの強さから、ファジィ推論によって傾向値の修正量を求める。しかし、ポーカーは不確定なゲームであるため、1 回の対戦だけから相手の傾向を読み取ることは難しい。そこで、ファジィ推論によって求めた修正量を記録しておき、その 5 回分の平均と最新の修正量をもとに傾向値を実際にどれだけ変更するかを決定する。

2.6.1 傾向値の修正量 1 回の対戦が終わるごとに、傾向値の修正量を求める。修正量は、優劣推定を行った際に、最も適合度が高かったファジィルールに対して求めるものとし、1 回の優劣推定につき、1 つ求める。修正量を求める推論ルールは、

If “優劣推定で推定したカードの強さは $\tilde{H}$ ” and “実際のカードの強さは $\tilde{I}$ ” then “傾向値の修正量は $\tilde{J}$ ”
 $\tilde{H}, \tilde{I}$: Weak, Strong (図 2(a)) $\tilde{J}$: N, Z, P (図 5) であるとし、修正量を求めるためのファジィルールと、傾向値の修正量を表すファジィ集合をそれぞれ表 3、図 5 に示す。

修正量の値は、優劣推定で推定したカードの強さと実際のカードの強さとの差に応じて推論する。優劣推定の推定値が実際の強さの値よりも低い場合は正の修正量を推論し、高い場合は負の修正量を推論する。また、推定したカードの強さが実際のカードの強さと同じくらいであれば修正量はほぼ 0 であるとする。このような知識から、表 3 のファジィルールは設定されている。

Table 3 Fuzzy rules for modification of tendency value

$\tilde{J}$		$\tilde{I}$	
		Weak	Strong
$\tilde{H}$	Weak	Z	P
	Strong	N	Z

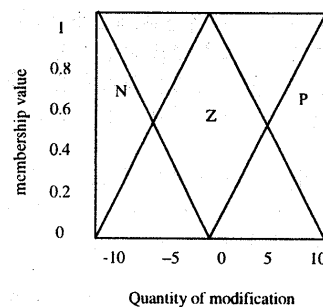


Fig. 5 Membership function for calculation of modification of tendency value

2.6.2 実際に傾向値を変更 2.6.1 で求めた修正量を記録しておき、5 回分の修正量の平均と最新の修正量を用いて傾向値の変更量をファジィ推論によって推論する。最新の値のみでなく、過去 5 回の平均も考慮することにより、少し長い期間での対戦相手の傾向をとらえる。この傾向値の変更により、各対戦相手の状況ごとの傾向を学習する。

傾向値を変更するための推論ルールは、If “5 回分の修正量の平均は $\tilde{K}$ ” and “最新の対戦の修正量は $\tilde{L}$ ” then “傾向値の実際の変更量は $\tilde{M}$ ”

$\tilde{K}, \tilde{L}, \tilde{M}$: N, Z, P (図 5) で示され、そのファジィルールを表 4 に示す。

Table 4 Fuzzy rules of a quantity of modification

$\tilde{M}$		$\tilde{L}$		
		N	Z	P
$\tilde{K}$	N	N	N	Z
	Z	N	Z	P
	P	Z	P	P

3. 被験者実験

3.1 実験方法 本システムの有効性を調べるために被験者実験を行う。被験者とシステムが 1 対 1 で、100 回のポーカーの対戦を行い、利得や勝利数などの対戦成績や、ハッタリがゲーム結果に大きく影響しているいくつかの具体的な事例について考察を行う。被験者は男性 4 名、女性 4 名の 8 名で、全ての被験者はポーカーのルールが分かる程度の初心者である。実験開始時の利得ポイントはシステム、人間プレーヤともに 100000 ポイントである。

本実験で用いたカードのデータは、全てショウダウンした場合は 50 勝 50 敗になるように設定してある。そして、全ての被験者に対し、配るカードの種類、順序は同じにしている。これは、全被験者が同一の条件で実験を行うことにより、被験者ごとの傾向を出やすくするためである。ただし、このことは被験者には伝えていない。

3.2 実験結果

3.2.1 対戦成績 各被験者に対するシステムの最終利得と勝利数を表 5 に示す。

Table 5 Experimental results

Subject No.	Gain of a system	The number of system's win times
1	33150	47
2	60432	41
3	28680	45
4	-65568	52
5	91898	50
6	-63772	47
*7	-100000	11
**8	100000	19

*被験者7は、システムの所持金がマイナスになってしまったため、29ゲーム目でゲームオーバー

**被験者8は、人間プレーヤの所持金がマイナスになってしまったため、45ゲーム目でゲームオーバー

利得によるシステムの人間プレーヤに対する対戦成績は5勝3敗であり、システムはおおよそ人間の初心者レベルと同じ強さを持っていると言える。表5から、勝利数はあまり変わらなくても、被験者によってシステムの最終利得が大きく異なっていることがわかる。これは、終盤ターンになるほど賭け金が大きくなっていくため、かけひきが成功するか否かで大きく所持金が動くためである。つまり、システムが対戦相手の戦略をうまくつかんでいる場合は大きく勝ち越すことができるが、優劣推定を誤ったり、被験者のハッタリにだまされたりした場合は、逆に被験者に大勝されてしまうことがある。これはスタッドポーカーの特徴と考える。最終利得の大きなプレーヤは、賭け金が大きくなってきた場合にうまくかけひきを利用している、と考えることができる。

3.2.2 ハッタリの効果 本システムは、ゲーム理論では最適戦略とされていないハッタリを戦略の一種として採用している。本システムの行ったハッタリに対する人間プレーヤの対応について、いくつか具体例を示す。

表6は、被験者6の96ゲーム目の例で、本当は劣勢である

Table 6 An example of bluff (with weak hand)

turn			1	2	3	4	5	6
human bet			10 2	15 2	20 42	50 125	15 73	Drop
card data								
system bet			12	17	62	175	88	
card data								
result value			0.004	0.322	0.243	-0.001	-0.051	-0.131

Table 7 An example of bluff (with strong hand)

turn			1	2	3	4	5	6	7
human bet			10 CALL	10 CALL	50 CALL	100 81	512 512	1000 1420	7000 5015
card data									
system bet			10	10	50	181	1024	2420	12015
card data									
result value			-0.675	-0.703	-0.667	0.455	0.317	0.4	0.363

とわかっているのに、システムが強いように見せてハッタリに成功した事例である。第5ターンの優劣推定値は-0.051でシステム劣勢となっている。実際、このままショウダウンをすれば、人間プレーヤが3と9のツーペア、システムがJのワンペアで人間プレーヤの勝ちとなる。しかし、アップカードの数字が9~Qまで並んでいるため、システムは高いポテンシャルを自分は持っていると判断し、戦略としてハッタリを選ぶ。

この場合は、ハッタリで自分の手を強く見せるために、システムはJを開いてワンペアを見せている。通常、強いカードはとっておくものだと思っていた被験者6は、「ワンペアという手役のカードをわざわざ見せるということは、ストレートができていたのを隠しているのだと思った」とゲーム後に答えており、ドロップをしてしまっている。ハッタリを行うことにより勝つことができた事例である。

次に、表7は被験者7の25ゲーム目において、自分の強さを隠すことで相手から多くの利得をあげた事例である。被験者7のカードの開き方から、自分をストレートができていように見せようというハッタリの戦略が見て取れる。システムの手役はフラッシュであるが、ストレートにもなるカードの組み合わせであり、優勢であると判断している。しかし、システムはカードを開く時に、自分が大きく優勢であることを悟られないようにフラッシュにもストレートにもならない順でカードを開いている。これにより、対戦相手をドロップさせることなく、大きな金額をシステムは得ている。

システムのハッタリが失敗した場合が表8の事例で、被験者1の90ゲーム目である。被験者1の手役は4のワンペア、システムの手役はナッシングである。第5ターンにおける優劣推定値は-0.419とかなり大きく負になっているが、ストレートのように強い手役に見えるポテンシャルであるため、システムはハッタリで一番強く見えるカードを開いている。しかし、被験者もドロップをせずにワンペアを見せて賭けてきているので、結局次のターンでシステムがドロップをしている。劣勢ですぐにドロップをしなかったために、結果的に損失が大きくなってしまった例である。

Table 8 An example of bluff ended in failure

turn			1	2	3	4	5	6
human bet			10 2	34 CALL	102 19	344 56	1144 807	3000
card data	♠ 10	♣ 4	♥ K	♠ A	♥ 6	♠ 5	♥ 4	♥ 4
system bet			12	34	121	400	1951	Drop
card data	♠ 9	♥ 8	♥ 7	♠ 2	♠ J	♠ Q	♠ 3	♠ 9
result value			-0.248	-0.365	-0.356	-0.368	-0.419	-0.589

4. まとめと今後の課題

本研究では、従来のセブスタッドポーカープレイングシステムを、開くカードの選択のセクションを加えた7ターンルールに拡張し、システムの取り得る戦略としてハッタリを意図的に行うルールを設けた。被験者実験を行った結果、システムはポーカー初心者の人間プレーヤーとほぼ同じくらいの強さを発揮していることが確認された。また、自分が優勢であると推論している時にシステムが強いカードを隠したり、弱いと推論しているのに強気に出ると、人間プレーヤーも時にハッタリにあうことがあったり、逆に、人間プレーヤーがシステムのハッタリを見抜いて、システムがかけひきに失敗をする事例を示した。

人間プレーヤーがシステムを破産させたり、かなり大きな利得をあげたプレーヤーがいたが、システムがこのように大敗してしまう原因に、ポテンシャルや優劣推定の算出に期待値を用いていることがあげられる。ストレートのように、完成する確率は低いが非常に強い手役ができている場合、システムの優劣推定にはあまり反映されていない。実際に、人間プレーヤーに大敗したゲームは、ほとんどの人間プレーヤーの手役がストレートであった。

今後の課題としては、ポテンシャルや優劣推定の場面において、期待値のみではなく連続した数字や同じマークがそろっている場合は、対戦相手が特に高いポテンシャルを持っている可能性があるという知識を推論に加えることがあげられる。

また、「対戦相手がドロップをしやすいプレーヤーであるかどうか」や、「どの程度のポテンシャルを高いとみなすか」とい

う判断を対戦相手ごとに行うために、新しい傾向値を設けてファジィ推論を行ったり、被験者の優劣推定における傾向値の推移とハッタリのだまされやすさに相関があるのかを調査することも今後の課題である。

参考文献

- (1) John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, John Wiley & Sons, Inc., New York, 1944
- (2) 松原仁, コンピュータ将棋の進歩, 共立出版, 1996
- (3) 松原仁, 滝沢武信, コンピュータ将棋はどのようにしてアマ4段まで強くなったか, 人工知能学会誌, vol.16, no.3, pp.378-383, 2001
- (4) 安藤剛寿, 関谷好之, 上原貴夫, "コンピュータブリッジのビッドにおけるパートナーシップ", 電子情報通信学会論文誌, vol.J81-D-II, no.10, pp.2366-2375, 1998
- (5) D.Billings, D.Papp, J.Schaeffer, and D. Szafron, "Opponent Modeling in Poker", *Proceedings of AAAI-98/IAAI-98*, 1998, pp.493-499
- (6) デヴィッド・バーレット, トランプゲーム百科, 社会思想社, pp.393-435, 1988
- (7) S. Kazami and T. Onisawa, Adaptive Fuzzy System for Playing Incomplete Information Game, Proc. of the Fourth Asian Fuzzy Systems Symposium, Vol.1, pp.611-616, 2000
- (8) 風見覚, 鬼沢武久, 不完全情報ゲームプレイングシステムの行動とその考察, 第10回インテリジェント・システム・シンポジウム講演論文集, pp.257-260, 2000
- (9) 鬼沢武久, 風見覚, 高橋千晴, 不完全情報ゲームプレイングシステムの構築—スタッドポーカーを例にして—, 日本ファジィ学会誌, vol.15, no.1, 2003 (掲載予定)
- (10) 本多中二, 大里有生, ファジィ工学入門, 海文堂, 1989