情報処理学会論文誌　数理モデル化と応用　Vol. 3　No. 2　11–21 (Mar. 2010)

多人数不完全情報ゲームの簡略化評価値による探索を用いた終盤データベースの構築

西　野　順　二[†1]　　西　野　哲　朗[†1]

本論文では簡略化した評価値を導入することで，不決定性を持つため探索がしにくい多人数ゲームの終盤データベースを構築した．さらに，これを用いて多人数不完全情報ゲームの意思決定を行うプレイヤモデルへ適用し，コンピュータ大貧民大会サーバを用いた実験によりデータベースの有効性を示した．多人数ゲームには自己の判断によって利得を制御できず，第三者の合理的でない判断によって左右される不決定という状態を持つ．これに対して大貧民のサブセットである単貧民化を行って手の縮約を施し局面を限ったうえで，シングルトンにより単純化した評価を用いることで最終 10 枚の終盤データベースの構築を行い，3 人の場合で 38%，4 人で 25%の場合について必勝手を発見した．不完全情報の局面を必勝手に帰着することで着手決定を行うプレイヤモデルを構築し，対戦実験によりパフォーマンスの向上がみられ，有効なモデルであることを示した．

An Endgame Database Construction for an N-person Imperfect Information Game Based on Singleton Leaf Value

Junji Nishino[†1] and Tetsuro Nishino[†1]

In this paper, we introduce a new endgame database making method for n-person games using simplified singleton values as search tree leaf values. We show a player model for an n-person imperfect information game using this endgame database which was pre-made. Multiplayer games have tie-breaking nodes in which the winner can not be determined by his own decision, thus searching was difficult to approach these games. Tanhinmin, that is a small sized Daihinmin game is introduced to make the endgame database, which use a singleton value model and shrinking method for game situations. As a result of making endgame database in 10 cards for 3 or 4 persons, 38%, 25% of game moves are revealed to be the winning moves. The result of experiments using Daihinmin game tournament server system show good performance of the model.

1. はじめに

利害関係のあるもの同士での意志行動決定は，ゲーム理論によってモデル化され，関与するプレイヤの数に応じてクラス分けがされている．

本論文では，3 人以上の多人数ゲームに焦点をあて，終盤局面の事前探索によるデータベース化した知識を構築するため評価値の簡略化を導入する．構築した終盤データベースを最適行動を目指した自動プレイヤ，とくにトランプゲームに代表される不完全情報多人数ゲームへの応用を試みる．

ゲーム情報学や人工知能分野において，2 人で同一の資源を取り合うような問題，いわゆる 2 人ゼロ和有限確定完全情報の展開形ゲームについては，チェスや将棋，囲碁などを例題として様々に研究されており，多くの知見が得られている．とくに完全情報である場合は Min-Max 法により，ゲーム木探索で必勝・必敗，ゲームによっては引き分けが，原理的に初手から確定するという特徴がある．

一方，多くの社会活動では 3 者以上が関わる多人数ゲームであることも多いが，そのクラスが多数にわたることなどから，2 人ゲームに比べてモデル化や分析・研究が進んでいない分野である．トランプ（カード）ゲームは，手近な有限不完全情報不確定ゲームとして日本でも身近な対象である．しかし，その不完全情報性や多人数ゲームであること，などからプレイアルゴリズムに関する研究はほとんどなされておらず，ゲームプログラムの多くはヒューリスティックな規則に基づいて作られている．多人数ゲームでは 2 人ゼロ和に相当するプレイヤの利得が定和であっても，必勝，必敗のほか引き分けではない不決定状態になるという特性がある．不決定は自己の判断が自己の利得には無関係であり，他者の利得にのみ影響する現象である．このため，完全情報であっても従来の意味でのゲーム木探索を一意に行うことができず，計算量にかかわらず，最適着手を決められない場合が発生する．

多人数ゲーム木での着手決定には，Max-Min アルゴリズムの自然な拡張である Max^n アルゴリズム[1] があり，その解は均衡解の 1 つとなることが知られている．これは，線形順序を仮定しないベクトル評価値の比較に基づいており，比較不能のタイブレークを行う節点が存在する．タイブレークの選び方によって容易に，他の均衡解へ大きく移動してしまうという解の不安定さを本質的に持っている．最悪値のリスクを軽減させるよりロバストな

†1 電気通信大学
University of Electro Communications

Soft-Max^n [2] も提案されているが，効率良い探索を行うためには相手の選好モデルを仮定しなければならない．序盤を含むゲーム全体では，UCT を用いたオンラインのモンテカルロ探索法によるプレイヤの探求も行われ顕著な成果をあげている[3]．これに対し本論文では読みきることが可能な終盤に着目し，タイブレークへの対応のため Soft-Max^n 法よりは簡易なシングルトン評価値への変換を導入することを試みる．

終盤データベースは，終盤において駒の減少や局面の縮小をともなうチェスやチェッカープレイヤ[4] などの 2 人ゼロ和ゲームで，主として計算量負荷を減らすことを目的に詰めの定跡として利用されてきた．多人数ゲームを含む非ゼロ和ゲームでは，プレイヤアルゴリズムにおける枝刈りの効率化のため，事前の計算により終盤データベースを構築し，末端局面とその評価値を動的にハッシュテーブルに格納して利用する手法が提案されている[5]．本論文では，これとは異なった評価値の簡略化を提案し，変更を加えることで多人数ゲームの終盤データベースの構築を試みる．

以下では，多人数ゲームの探索と終盤データベース構築アルゴリズムを検討し，国内で代表的なトランプゲームである大貧民のコンピュータプレイヤの構築を例として，多人数不完全情報ゲームに探索による終盤データベースを利用するモデルを提案する．

2. 多人数不完全情報ゲームの探索

2.1 多人数展開型ゲームと均衡解

多人数展開型ゲームは戦略と利得がすべて与えられたならば，均衡解を持ち，Max^n アルゴリズムにより求めることができる[1]．ゲームの均衡解とは，直近で損をしない戦略の組であり，どのプイレヤがいかに手を変えてもそれ以上の利得が得られない状態である．展開型においても各節点での選択が戦略であり，どの節点においても手を変えるかぎり利得が下がる状態が均衡解である．同時に着手をする戦略型ゲームでは，Nash の均衡解によって解が求まる．順に着手をする展開型ゲームでは局所最適な解となる．形式的に表現すれば，各葉節点での利得をプレイヤそれぞれの利得の組としてベクトル $v = (v_1, \cdots v_i \cdots v_n)$ で与えると，アルゴリズムの目的はすべてのプレイヤが利己的かつ合理的に戦略決定したときの戦略の組すなわち m 個の各局面における最適着手の組 $\vec{s} = (s_1, \cdots s_j \cdots s_m)$ を求めることである．このときに得られる利得ベクトル $v^o = (v_1^o, \cdots v_i^o \cdots v_n^o)$ が，ゲーム木全体の利得となる．Max^n で得られた戦略は均衡解となることが保証され，プレイヤ i が最適戦略から 1 つの局面 j で異なった着手 $\vec{s}' = (s_1, \cdots s_j' \cdots s_m)$ をとるとき，得られる利得ベクトル $v' = (v_1', \cdots v_i' \cdots v_n')$ に対し，$v_i' \leq v_i^o$ となる．

Max^n アルゴリズムの概略は以下のとおりである．プレイヤ i の手番節点では，v_i を最大化する手 s を選ぶ．ただし最大値をとる v_i が複数ある場合には任意の 1 つを選んでかまわないものとする．手 s によって導かれる子節点の評価値をこの手番節点の評価値として採用する．これを葉から順に行い根まで行うことで Max^n アルゴリズムが完成する．ここで得られた手の組 $\vec{s_i}$ がプレイヤ i の戦略である．

均衡解が得られる証明が参考文献 1) にあるが直観的には以下のように説明できる．アルゴリズムの定義から，根の評価値を与える探索の主経路で考えるとどのプレイヤの節点での手すなわち戦略を変更しても，$v \leq v^o$ がなりたち，評価値は少なくとも増加はしないことが保証され，このため均衡解の定義を満たす．

2.2 多人数ゲーム木探索の不決定性

2 人ゲームにおいて利得のゼロ和あるいは定和が仮定できることの意義は，各節点での評価値が 1 つの変数で表現できることにある．各プレイヤの利得 x，y に対して定和であるとき定数 C を用いて $x + y = C$ が成り立つ．このため $y = C - x$ であり，ゲーム木全域のすべての利得を x のみの一次元で表すことができ，線形順序によって比較可能となる．max-min 戦略によってゲーム木全体の代表利得を算出できることは，すべてこの利得の線形比較に基づいている．ひいては $\alpha\beta$ 法による探索枝刈りにも通ずる．

3 人以上の多人数ゲームにおいては，たとえ利得の総和が一定であったとしても，すくなくとも自由度が 2 次元以上となり，利得はベクトルとなるため線形な比較は不可能となる．このため，max-min のように全域にわたる比較によって木全体の解を決定することができない．ゲーム木の解が求まらないという性質，すなわち不決定性は本質的に多人数ゲームにおける最適行動の決定の難しさを引き起こしている．

従来の研究では，多人数版の枝刈り法の検討が行われる一方で，根源的な不決定性の問題にはあまり触れられてこなかった．

プレイヤ i の視点に固定すれば，$v = (v_1, \cdots v_i \cdots v_n)$ と $v' = (v_1', \cdots v_i' \cdots v_n')$ の大小を，$v_i < v_i'$ で比較判定することができ，Max^n 法で使用されている．しかし v_i の値が少ない整数値のみをとる場合，v_i と v_i' が等しくなる場合が多くなりプレイヤ i の視点では決められなくなる．

たとえば，3 人ゲーム木の葉ノードでの利得を各プレイヤの順位を基準にして，2，1，0 点が割り当てられるとする．このとき各葉ノードでの各プレイヤの利得の合計は一定である．葉ノードでのプレイヤ A，B，C の利得をそれぞれ v_A，v_B，v_C とすると，この利得をベクトルで (v_A, v_B, v_C) のように表せ，$v_A + v_B + v_C = 3$ であり，その組合せの種類は

3 の階乗で 6 通りである．

こうして利得および評価値がベクトルで表されるとして，決定ができないことは，簡単な例で示すことができる．プレイヤ A の局面で選択肢が p，q の 2 つありそれぞれの利得ベクトルが $g(p) = (0, 1, 2)$ と $g(q) = (0, 2, 1)$ とする．このとき，どちらも A の利得は 0 であるから，プレイヤ A は p，q の選択肢についてはまったく無差別であり，合理的にどちらかを選ぶ決定はできない．しかしなおプレイヤ B，C の利得はプレイヤ A の選択にかかっている．ここでこの A の手番のノードの評価値を，仮に $g(A) = <(0, 1, 2), (0, 2, 1)>$ という 2 項組で表しておくことができるが，一方に定めることはできない．

ここでさらに，探索木を遡った 1 つ上のノード，プレイヤ C の局面を考える．ここで選択肢 r，s があり，それぞれの選択の後先の A の局面になるとする．r を選んだときの A_r の局面の評価値が，$g(A_r) = <(0, 1, 2), (0, 2, 1)>$ であるとする．一方，s を選んだときの A ノード A_s の評価値が，$g(A_s^1) = (0, 1, 2)$ または $(1, 0, 2)$ であるならば，C は選択肢 s を選べば確実に利得 2 となる．結局 C ノードの評価値は $g(C) = g(A_s^1)$ を採用し，$(0, 1, 2)$ または $(1, 0, 2)$ のいずれかに確定することができる．

A_r の手番ノードのように，一般に多人数ゲームでは自己の利得からは決定できないが，他者の利得を変化させる複数の選択肢が現れる．一方，運が良いときには，C の手番ノードのように一意に自己の利得を決定する手を選ぶことができる場合もある．このように多人数ゲーム木の探索は，2 人ゲームの最適着手決定を包含し，多くの場合に不決定性があるという特徴を持つ．

一般に不決定な状態のほうが起こりやすく，たとえば選択肢 s を選んだときの A ノードの評価値が，$g(A_s^2) = (2, 1, 0)$ または $(1, 2, 0)$ ならば，C の得る利得は選択肢 r を選んだときの利得 2 もしくは 1 よりも確実に低くなる．このため C は選択肢 r を選ぶことになる．選択肢 r を選ぶとすれば，この C ノードの評価値も不確定な値 $g(C) = g(A_r) <(0, 1, 2), (0, 2, 1)>$ となる．

2.3　Max^n 探索の解の不安定性

Max^n 探索で均衡解が求まるが，最良の手が他にも存在する可能性があり，他の最良手の発見は一般に困難である．とくに自己の評価値が同じ複数の手の選択肢がある節点で，そのタイブレークのしかたによって相異なる複数の均衡解戦略が自然に得られ，評価値も全体戦略も大きく変化することがある．

たとえばプレイヤ B のある節点で選択可能な 2 つの手 a，b それぞれに対して $a(0, \underline{1}, 2)$ と $b(2, \underline{1}, 0)$ という評価ベクトルが与えられているとする．このとき手 a と手 b は，プレイヤ B にとってどちらも評価点 1 が与えられて無差別であり決定不能なタイ節点となっている．仮に，a を選んだとして，ゲーム木の親節点プレイヤ A の手番に評価値 $(0, 1, 2)$ を戻すと，プレイヤ A の利得は 0 で，a は悲観的な節点となり他の節点との関係の中で選ばれない可能性が高い．同じ節点でプレイヤ B が $b(2, 1, 0)$ をタイブレークとして選んだとすると，親節点でのプレイヤ A の評価値は 2 となり逆に選ばれる可能性が高くなる．

一方，プレイヤ B の別のある節点で選択可能な 2 つの手 c，d の評価ベクトルが $c(\underline{0}, \underline{1}, 2)$ と $d(\underline{0}, \underline{1}, 0)$ であったとする．このときはプレイヤ B がどちらの手を選んでも，親節点でのプレイヤ A の利得は 0 となって変わらない．第三者であるプレイヤ C の利得は 0 と 2 と大きなふれ幅を持つが，B がどちらを選んだとしてもなお A に選ばれる可能性が低い手である．このように評価に影響が少なく手の選択ブレが親節点に影響を与えないため，探索が不安定とならずに全体として同じ結果に収束するものもある．選択肢の各評価ベクトルがすべて同一であったり，優越するものが 1 種類であったりすればこれも安定である．

また，必ず Max^n でタイブレークが必要なく安定に解が求まるクラスがある．すなわち，j 番節点の評価ベクトルを v^j とし，そこでのプレイヤ i の評価値を v_i^j としたとき，プレイヤ i のすべての評価値が互いに相異なり線形順序を与えられるとき，すべての節点でタイブレークが起こらず，安定に Max^n を適用することができることになる．

2.4　Soft-Max^n

このように多人数ゲームの探索では，本質的に非決定なタイブレーク節点があり，他プレイヤの非論理的な選好に依存するため必勝手順を与えられない．

そこで，タイブレーク探索で親ノードに返す値を，複数のベクトル値の集合とする，Soft-Max^n [2] が提案されている．すべての可能性を伝播させることで，探索のロバストネスを上げリスクを減らす手法である．しかし，Soft-Max^n は，従来の手法に比べよりロバストな解を与えているが，依然として自身の最終利得は他者の合理的判断基準の存在しない「きまぐれ」に依存している．このため，同時に相手モデルを併用することでその効果を高めることが提案されている．しかしながら，一般に相手モデルを作成することは負担が多く実装，使用には大きな制限がある．

2.5　シングルトン評価による探索アルゴリズム

前節までで見たように 3 人以上のゲームでは，あるノードの戦略を選択するプレイヤにとっては戦略の合理的な決定ができない不決定の場面が発生する．また，あるパターンにおける必勝手順を探索するうえで，自身の必勝手がないノードにおける他のプレイヤの評価値は必ずしも重要とはいえない．

そこで，葉ノードでのベクトル値利得をより簡略化し，1 位以外をゼロとして $\{(1,0,0),(0,1,0),(0,0,1)\}$ の 3 種類としたシングルトン評価方式を導入する．この 3 種類の利得はすなわち，その葉ノードでプレイヤ A，B，C が勝利することと対応している．よってさらに簡略記号化して単に $\{a,b,c\}$ と書くこととする．

この利得表現は 2 位以下は無視することになる．しかし，たとえば 5 人ゲームで先に 2 人がアガリとなった残りの 3 人の中であれば，ここで 1 位になることは全体としては 3 位になることである．このためサブグループ内で 1 位を目指すことは必要かつ有効な目標である．

探索木の各ノードの評価値を伝搬することで，統合された節点の利得は，3 種類の葉ノードの値 $\{a,b,c\}$ のべき集合の要素すなわち，$\{a,b,c,ab,bc,ca,abc\}$ のいずれかとなる．ここで ab，bc，ca，abc はそれぞれに含まれる文字のどちらかが勝利するが，その勝敗の決定は他者に無差別にゆだねられていることを意味している．たとえば ab ならば，C の手番で無差別に a，b のどちらかが選ばれる．一般則として，$c > ca$，$bc > abc > a$，b，ab の線形な順序に則るものとすれば，max-min と同じく各節点の利得を一意に定めることができるようになる．たとえば，ある C プレイヤ節点の選択肢が 4 つあり，評価値が c，bc，abc，ab であるなら，順序に基づき $MAX_{singleton}(c,bc,abc,ab)=c$ として選択肢 4 つを統合した節点評価値を c とすることができる．C プレイヤの場合単一の c が自己にとって最も良いため，1 つでもこれを与える選択肢があれば他はいっさい無関係となる．ca と bc のように同順位にあるもの同士となったときは，どちらとも決められないため，ca と bc からは abc，a と b からは ab のような融合ルールを適用するものとする．節点の選択肢とその利得の組合せと統合した総合評価を**表 1** に示す．左辺の必須要素があるときには，他の評価値に優先して総合値に影響をあたえる．他の評価値はそれらのベキ集合のいずれかでよい．すなわち，一部のみであっても同様の結果となる．

葉節点が a，b，c のシングルトン評価値による利得であるときに，木を遡った節点の統合利得が ab などのベキ集合要素で表されることは，探索木を順にたどることで明らかとなる．

ある C の手番ノードでとれる戦略 $p,q,r,\cdots$ が複数あり，それらの結果はそれぞれ a，b，c のいずれかとする．C は自身にとって最も有利な手を選ぶと仮定すると，選択肢全体の総合評価は a，b，c，ab のいずれかになる．ab になるのは手番の選択肢の利得が a と b ばかりのときである．

1 つ上の B の手番のノードでは，選択肢それぞれの評価は先の C のノードの評価値 a，b，c，ab のいずれかの組合せとなる．たとえば B が戦略 p，q を選択したときのノードの値が，$g(p)=c$，$g(q)=ab$ であるとすると，B は自身が勝つ可能性のある ab の値を持つ戦略 q を採択するのが合理的である．選択肢の値のありうる組合せについて考えると，表 1 に示した評価値の融合規則を B の手番に読みかえて適用し，結局 a，b，c，ab，ca の 5 種類となる．

表 1　プレイヤ C 節点での手と総合評価
Table 1　Value unifier operation at player C's node.

必須要素	子節点の評価の組合せ		総合評価
c		→	c
c	a, b, ab, abc, ac, bc	→	c
ac	a, b, ab, abc	→	ac
bc	a, b, ab, abc	→	bc
ac, bc	a, b, ab, abc	→	abc
abc	a, b, ab	→	abc
a		→	a
b		→	b
a, b		→	ab
ab	a, b	→	ab

さらに 1 つ上に遡った A の手番では，a，b，c，ab，ca の様々な組合せ評価を A の利益について行い，結局 a，b，c，ab，bc，ca，abc 7 通りとなる．評価 abc とは，A の手番で選べる戦略が ab，ca の 2 つとなったときの総合評価である．この 2 つは A にとっては無差別であるためどちらとも決定できず，節点の評価としては abc とすることになる．

$Soft-Max^n$ 法も可能性のある手の評価を組み合わせて伝播させる点で同様である．ここで提案するシングルトン評価の利点は，組合せが 3 人ゲームで 7 種類，4 人ゲームで 15 種類のみに限定されることである．ab，bc からなる B プレイヤ節点での評価を統合した abc に置き換えず，そのまま $\{ab,bc\}$ として組として持つ方法も考えられるが，この場合 $\{a,b,c\}$ のベキ集合のベキ集合となり，3 人ゲームであっても評価値の組合せは 255 種類に増大してしまう．こうした増大を起こさないことはゲーム木として探索するうえで重要である．

3. 終盤データベースを用いた大貧民プレイヤモデル

本論文で提案するモデルは，多人数ゲームの多くに適用可能であるが，ここではトランプゲーム大貧民を対象例として，終盤データベースを用いたプレイヤモデルの構築を説明する．

大貧民もしくは大富豪は日本で著名なカードゲームで，多くのローカルルールがあるものの基本形は最もよく知られたゲームの 1 つである．トランプを切って配ることにより，対戦相手の状態を完全には知ることができず，多人数不完全情報ゲームに分類される．

以降では，とくに電気通信大学で開催されているコンピュータ大貧民大会のルールを規定

として，プレイヤの構築を行う．

3.1　多人数ゲームとしての大貧民関連研究

大貧民は日本国内で広く知られているトランプカードゲームである．52 枚を 3–5 人に配布する組合せは大きく，コンピュータプレイヤを構築するのは容易ではない．また多人数ゲームの特性上プレイヤ同士の一対比較はできず，単体としてのプレイヤプログラムの強さを評価することも難しい．2006 年に電気通信大学でコンピュータプレイヤ相互の競技会[6),7)]が開催され，プログラムの強さを評価する枠組みとして有効であった．

現状のプレイヤプログラムの多くはヒューリスティックに基づいたルールベースによる着手戦略が用いられている．対象とする大貧民の終盤データベース探索は，すべてのプレイヤの手が互いに明らかであり 1 枚だけのプレイに制限した完全情報化した大貧民ゲームであると近似できる．日本で身近なカードゲームである大貧民のプレイヤプログラムの終盤での読みきりデータベースに寄与することを目指し，その分析が行われている[8),9)]．

カードゲームとしての大貧民としては，古くは 1980 年代に有澤が特殊な 55 枚の組合せによる 5 人自動プレイの実験的分析について調べている[10)]．コンピュータプレイヤをいくつかの戦略の組合せでアドホックに構成し，順位の変遷について実験対戦の結果に基づいて分析した結果，同じ戦略の有効性が順位によって変わることを示した．手持ちの札の状態により戦略をメタに切り替える必要性を示している．

多人数ゲームとしての大貧民について，完全情報としても探索結果は不確定となる．後藤らは多人数ゲームにおける葉ノードでの評価を順位とし，不確定部分を列挙したノード値を用いる探索を提案している[11)]．この中では，2 人から 6 人まで全員に 5 枚までの同じカード組合せを配り，いくつかの枝刈り法を比較している．これと比して本研究は，実際に配られる可能性のある組合せのバリエーションから，終盤データベースを作ることを目指して完全探索することを目的とする点が異なっている．

多人数ゲームの探索法として 1986 年に Luckhardt らが提案した Min-Max 法の自然な拡張である，Max^n 探索[1)] がある．4 人将棋プログラムの実装のために選択的 M^3 サーチ[12)]が提案されており，基本構造は Max^n 探索と同様であるが，ヒューリスティックに基づく選択的探索を行って効率化をはかっている．大川らの 4 人将棋でのアルゴリズム[13)] もベクトル値によるカット配列法を提案し探索に枝刈りを導入している．

探索の効率化の観点からは，ダイアモンドゲームを対象とした 3 人ゲームの探索でその順位を評価値とした「早い者勝ちゲーム」としての分析[14)] や，同じくダイアモンドゲームを対象に，最良優先探索の提案[15)] がある．これらは，探索の効率化に主眼がおかれ，同一の初期状態からのみ実験が行われている．しかし，そもそも問題が可解な状態であるかという判断が，多人数ゲームでは必要である．不安定な木を枝刈りして探索しても，そこで得られた戦略が有効であるとはいいがたい．

3.2　大貧民のルール

電気通信大学の大貧民大会公式ルールでは，5 人のプレイヤでゲームを行う．使用カードはジョーカを含む 53 枚で，席順により 10，10，11，11，11 の不均等に分配される．直前の試合順位に応じてカード交換を行う．カード交換は 5 位，4 位のプレイヤはそれぞれ前回 1 位と 2 位のプレイヤに手札の強いものから 2 枚，ないし 1 枚を渡し，受け取った 1 位と 2 位のプレイヤは手を見たうえで戦略的に不要なカード 2 枚ないし 1 枚を渡すことで実施する．カードの強さは，$3, 4 \cdots 12, 13, \mathrm{A}$ の順で 2 が最強，3 が最弱である．ジョーカは 2 より強いカードとして使用でき，スペードの 3 はジョーカにだけ強いカードとして使用することができる．

自分の手番では場のカードより強いカードを出すことができ，また出すカードがない場合および戦略的に任意にパスをすることができる．ただし 1 度パスをしたらそのターンが終了して場がクリアされるまでは，カードを出すことはできずパスするのみとなる．

出し方は，1 枚，同ランク 2 枚組み，3 枚組み，4 枚組み，同スートの階段 3 枚以上ができる．このときジョーカはワイルドカードとして希望のカードに代えて使用することができる．8 切りあり，アガリ札の制約なし，同じスートが続いて出るとシバリとなりターン終了までそのスートしか出せない．4 枚組もしくは 5 枚以上の階段によって革命となり，それ以降 1 ゲーム終了までカード強さが逆転する．

手札をより早く出しきったものからアガリとなり，早い順に 5 位までの順位が決定する．

大貧民のゲーム木の節点では，以下の変数の組によって状態表現できる．

$(h_1, ..h_n, i, P, T)$

ここで，h_n はプレーヤ n の所持するカード，i は現在の手番プレーヤ，P はパスベクトル $(x_1, ..x_n)|x_n \in \{0, 1\}$ で，そのターンにすでにパスしたプレイヤ n を $X_n = 1$ で示す．T はテーブルに出ている場のカード/組．

ここでプレーヤ i は自己の手 h_i 以外の $h_1..h_n$ の個々の内容を知ることができない．この意味で不完全情報ゲームである．ただし $|h_j|$ すなわち各プレイヤの持つ手 h_j の枚数と，その和集合 $h_1 \cup h_2.. \cup hn$ は知ることができる．

3.3　着手決定アルゴリズム

提案モデルによるプレイヤは，ベースとなるオリジナルプレイヤをもとに，終盤のみ提案

アルゴリズムを加えたプレイヤモデルとなっている．オリジナルプレイヤは，現況を判別する規則から着手を決定するルールベース型プレイヤで，2007 年度コンピュータ大貧民大会に出場したものである．以後オリジナルプレイヤをベースプレイヤと呼ぶ．

提案モデルの着手決定アルゴリズムを以下に示す．

（1） カード全体の残枚数 N および人数 M が終盤データベースの形式にマッチするまで，ベースプレイヤのルールベースによる着手を行う．

（2） (N, M) が条件に適合し，終盤データベースと符合する場合以下の手順に従う．

（a） 自手以外のカード $h_2, ..h_5$ をランダムに配布し，パターン $P_i = (h_1, h_2^i, ..h_5^i)$ を作成する．

（b） 配布パターン P_i を 3.5 節で述べる縮約型 Q_i に縮約変形する．

（c） Q_i を終盤データベースから検索し必勝最適手 x_i を求める．

（d） 最適手 x_i を縮約型から自手に含まれる本来のカード $c_k \in h_1$ に還元し，その件数 W_{c_k} を集計する．

（e） L 個の配布パターン $P_1, ..P_L$ を生成し，集計値 W_{c_k} の最大値を与える手 c_k を着手とする．

本論文では，$N = 10$，$M \in \{2, 3, 4\}$ である．終盤データベースはパターンの縮約を用いて，以下に示すように数百万程度の比較的少数になっている．一方，自手以外のカードの分布は不完全情報のため不明であり，その組合せは自分をのぞく 3 人に 8 枚が行き渡っているとして，$53!/(53-8)!/2!/3!/3!$ で約 5,000 億通り以上となる．すべてを実時間で生成，検査することはほぼ不可能であるため，一定時間内に適合するパターンをランダムに可能なかぎり生成し，それらの集計を行うことで着手を決定するものとする．本システムでは，Intel Core2Duo2.53 GHz の計算機で，5 秒間でパターンの生成数 L は最大 329,412 であった．

3.4　終盤データベースの構築：残 10 枚の組合せ

終盤データベースは，ゲーム後半で局面が十分狭まり，完全探索が可能な状態の解をあらかじめ求めておくものである．

多人数不完全情報ゲームである大貧民では，前述のとおり全探索は一般にできない．しかしながら，後半局面ではそれぞれのプレイヤの手札がある程度推論できるなど，完全情報に近い状態となる．たとえば，2 人だけが残った状況では，それまでの出現カードを記憶しておくことで，相手の持ち札を知ることができ，完全情報ゲームとなる．

本論文で提案するプレイヤのために，以下の仮定をおいて終盤データベースを構築する．ゲームは完全情報でかつ 1 枚出しのみしか行わないものとし，8 はすでにないものとする．このような 1 枚出し限定の完全情報多人数ゲーム化した大貧民を単貧民と呼ぶ．

枚数については，たとえば 5 人でのゲーム終盤に 2 人が残ったとすると，2 人合わせて最大で 22 枚である．実際にはここまでにそれぞれ何枚かは減らしている可能性が高いため，本論文では合計 10 枚以下の完全探索を行って利用することにした．3 人，4 人で合計で 10 枚を分けあって持っている状況下での全探索を行った．

3.5　単 貧 民 化

単貧民は，大貧民のルールをベースに，手札を公開して 1 枚プレイのみで行うゲームである．完全情報多人数ゲームとなり，ある種の詰め大貧民ととらえることもできる．

大貧民の基本ルールは，各プレイヤが順に手札を出してゆき，最初に手札を出しきったプレイヤが勝利し，以下，手札のなくなった順に順位を付けるというものである．ただし出せる手札は場札より強いカードに限られ，出せない場合にはパスとなる．1 人を除いて他の全員がパスしたとき，場が流れて最後のプレイヤが次のリードを自由に出すことができる．通常のルールでは 3 が最も弱く数字の大きさに比例して強く 13 に相当する K の上として，A があり，さらに 2 が最も強いとしている．本論文中ではこれらの強弱の順に番号を振り直し，1 が最弱で 13 が最強であるとする．

詰め大貧民では，あたえられた配布状態から，最も良い手を考えることを目的とする．ここでは詰め大貧民の例として残 10 枚のみを対象とし，相対的に強度番号を振り直して，強さ 1 から 10 のカード 10 枚が 4 人に配られている次の状況を考える．

$$[[1, 3, 10], [2], [4, 6, 8], [5, 7, 9]]$$

プレイヤ 1 は，1，3，10 の 3 枚のカードを持っている．どのカードからプレイすればよいかを求める．ここでは 3 からプレイするのが正解で，すると 2 番はパス，3 番，4 番がそれぞれ 1 枚ずつ出したところで 10 を出して場を流し，最後に 1 を出して上がることができる．

一方，最初に 1 を出してしまうと 2 番のプレイヤが上がってしまうし，最初に 10 を出すと全員パスののち，1 を出せば 2 番プレイヤが勝利，3 を出せばその後 3 番と 4 番プレイヤどちらかが勝利するまで 2 人だけが出し合う試合となってしまう．

一般に知られた大貧民のヒューリスティックではより小さいカードから出すのが良いとされるが，このように最適手順を詰め問題として考えると直感と異なる場合がある．

大貧民では上がった順位によって次回ゲームのカード配布で優越されるという決まりがある．このため，順位がそのまま利得となると考えるのが自然である．すると 3 人プレイヤでは $(0, 1, 2)$，$(0, 2, 1)$，$(1, 0, 2)$，$(1, 2, 0)$，$(2, 0, 1)$，$(2, 1, 0)$ と，6 種類の可能性がある．

一方，探索木の葉節点はこれよりもはるかに多くなる．たとえば 9 枚を 3 人に均等に 3

枚ずつ配ったとき初手はパスを含め 4 通りとしてパスは連続せず 2 巡につき 1 枚ずつは減らすと仮定して計算すると，$4^3 \times 3^3 \times 2^3 = 331{,}776$ 通りの葉節点がある．10 枚で連続したパスも許せば組合せ数と葉節点の数はさらに増える．しかし各葉に割り当てられる利得ベクトルは 6 種類しかないため，各所で同じ値が使用され，非常に多い割合でタイブレークが発生しやすいといえる．

3.6　手の同値構造

単貧民は着手を 1 枚出しのみに制限した大貧民である．大貧民のルールのもとで 1 枚のみのゲーム単貧民を行う場合には，異なる配布組合せどうしであってもカードの強弱の相対関係とゲームの進展から等価性を考えることができる．終盤データベースを構築するにあたり，手パターン数の縮約が可能となる単貧民状態[8] を仮定した．

複数のカードの配布が以下に示す強弱の等価性により同一視できるとき，その代表パターンの集合を縮約空間 S と呼ぶことにする．パターン数の縮約手続きは，カード配布全体の集合 P から，縮約空間 S への写像 $g()$ として定義される．

カード x，y の強弱関係を与える関数を $f(x,y)$ とすれば，$f(x,y) = f(g(x), g(y))$ を保存する写像を選ぶことができる．ここではすでに提案されている以下の 3 種類の変換[8] により探索すべきパターン全体を縮約する．

(1)　全体のシフト
$(c_1, ..c_n) \rightarrow (c_1 - s, .., c_n - s)$
すべてのカードの強さに対して，同数 s だけ増減してもゲームは変わらない．

(2)　強さの差の整理
例 $(1,3,4,7) \rightarrow (1,2,3,4)$
連続するカードの強さの差が 1 となるよう規格化できる．

(3)　連続する自手の同値化
例：手 A$(1,4)$ と手 B$(2,3)$ との試合は，手 A'$(1,4)$ と手 B'$(2,2)$ の試合と等価となる．上記の整理を行ったのち自己でのみ連続するカードは同じ強さとしてまとめられる（これはペア出しを許さないため）．

これらにより等価なパターンを削減し計算量を削減した．

3.7　3 人 10 枚

以上の準備のもと，3 人に 10 枚を配布した結果について探索を行う．

3 人に 10 枚を配布する組合せは，スートを考慮せずカード強さについて同じと見なせるカードが最大 10 枚として，10 枚の中での強さの組合せが 2^9 通り，その 10 枚の 3 人への配り方について 3^{10} 通りあり，全体で $2^9 \times 3^{10} = 30{,}233{,}088$ 通りである．ただし 1 枚も配られない場合も含んでいる．これに対して前述の手の構造に基づいて縮約を行うと，1,428,867 通りに縮約することができた．

表 2　3 人 10 枚の探索結果
Table 2　winning patterns for 10 cards, 3 players.

必勝手	:	パターン数
先手 1 を出す	:	116,856
先手 2 を出す	:	137,173
先手 3 を出す	:	116,956
先手 4 を出す	:	73,429
先手 5 を出す	:	33,244
先手 6 を出す	:	8,085
先手 7 を出す	:	979
先手 8 を出す	:	66
先手 9 を出す	:	2
先手小計	:	486,790
合計	:	1,252,189

探索の結果，先手必勝 621,368 通り，先手勝不確定 807,499 通りであった．先手の勝ちが不確定なものは必敗と不確定の両方を含んでいる．完全情報を仮定すると，3 人であっても約 43%のパターンで先手が勝つことが分かった．

これらの中には，自手にカードが 1 枚しかなく，戦略が自明なものも含んでいる．自手にカードが 2 枚以上あり戦略が自明ではない組合せでは，全体で 1,252,189 通り，そのうち先手必勝が探索できたパターンが，486,790 通りで，全体の 38%が解けることが分かっている．

先手の手による勝ちパターン数を**表 2** に示す．

ここでは，$([1,3,5,5,5],[2,4],[2,2,6])$ 先手 5，のように，下 2 つの 1，3 をおいて先に 5 から出さなければならないパターンも多数発生した．

先手 2 の場合には，2 の下に 1 がある場合が 4,159 パターン，残り 157,352 パターンは 2 が手中の最弱カードであり，1 は他のプレイヤが持っていた．3 以上の先手の場合には，最弱でない場合が多く含まれるようになることが分かった．

先手着手に 7 から 9 などの強いカード使いが少ないのは，事前のパターン縮約規則 2 および規則 3 によってカード分布が圧縮され，そもそも強いカードを本質的に含むパターン自体が少ないためである．

3 人の場合，実際のゲームは完全情報ではなく，10 枚全体の種類は捨て札の追跡により

表 3　4 人 10 枚の探索結果
Table 3　winning patterns for 10 cards, 4 players.

手	:	パターン数
先手 1 を出す	:	11,040
先手 2 を出す	:	13,116
先手 3 を出す	:	15,274
先手 4 を出す	:	17,572
先手 5 を出す	:	20,068
先手 6 を出す	:	21,551
先手 7 を出す	:	21,536
先手 8 を出す	:	18,092
先手 9 を出す	:	11,280
先手 10 を出す	:	13,104
先手勝小計	:	162,633
全自 2 枚上	:	637,020
先手勝合計	:	344,133
全合計	:	818,520

事後に判明する．よって相手 2 人にどのように分かれているかは不明であるが，これらのデータを終盤データベースとしてオンラインで利用することで，不完全とはいえ，より有効な着手を選択する手がかりとすることができる．

3.8　4 人 10 枚

プレイヤ数 4 人に対して 10 枚を配布した組合せすべてについても勝敗探索を行った．4 人に配布することで，同じ強さとしてまとめられるひと続きのカードが手にある可能性が少なくなり，組合せの縮約の効果が薄くなる．このため，あらためて縮約しても組合せを減らすことはせず，全体を対象として探索を行った．各初手ごとに分けた結果を**表 3** に示す．

10 枚を 4 人の手に 1 枚以上配布する組合せは 818,520 通りであり，解が求まるものは 344,133 通り（42%）であった．ただしこのうちには自手にカードが 1 枚しかなく戦略が自明なものを含んでいる．自手が 2 枚以上あり戦略が自明ではない組合せは全体で 637,020 通りあり，先手必勝手順が探索できたパターンは 162,633 通り（25%）であった．

4.　プレイヤにおける終盤データベースの効果

本論文で提案するクライアントモデルの有効性を検証するため，大貧民大会サーバ 2008 年度版を用い，複数のクライアントのパフォーマンス比較実験を行った．

ゲームのルールは使用した大貧民サーバ（tndhm 2008 年版）のとおりである．

4.1　実験の環境と設定

実験に使用したのは，終盤データベース使用型（以下，提案モデルと呼ぶ）のほか，コンピュータ大貧民サーバに付属するデフォルトクライアント（以下，デフォルトと呼ぶ），2007 年出場に用いたルールベース型ベースオリジナルプログラム（以下，ベースと呼ぶ），の 3 種類である．

カードを出す順番のもととなる席順は，ゲーム順位とは無関係に一定間隔ごとにシャッフルされる．一般には強いプレイヤの直後の席は有利で，直前の席は不利とされているが，席順のシャッフルによりこの影響はほとんどないものと見なすことができる．各回とも，ゲーム順位とは無関係にダイヤの 3 を持つプレーヤからスタートする．

前回の試合順位により，次試合の最初にカード交換が行われ，1 位と 5 位で 2 枚ずつ，2 位と 3 位で 1 枚ずつ，低位の側は強い順に渡し，高位の client は，全体のバランスと戦略から選んで返す．

試合数は 5,000 回とした．試合ごと 1 位から順に 5，4，3，2，1 点を与え，5,000 試合後の得点を評価するため最高は 25,000 点，最低 5,000 点，平均 15,000 点である．

4.2　実験の結果

ベースに対して，終盤のみ変更することの効果を見るため，提案モデルとベースの 1 対 4 での直接対戦を行った．結果は**図 1** に示すとおり，提案モデルの優勢であった．

第 3 のプレイヤとして，デフォルトクライアントを用い，これに対する優勢度の違いをみるため，提案モデル対デフォルト 4 つ，および，ベース対デフォルト 4 つの対戦を行った．それぞれ結果を**図 2** と，**図 3** に示す．

また，提案モデルとベースの試合にデフォルトクライアント 3 つを加えた 3 種類のプレイヤからなる試合を行った．その結果を**図 4** に示す．

最後に，提案モデル対ベース 3 つの試合において，1 つだけ第三者のデフォルトクライアントを加えたときの結果を**図 5** に示す．

5.　終盤データベースを用いるモデルの特性

概して，終盤データベースを用いるモデルがベースに比べて高得点をあげ，その優位性を示している．

提案モデルは，図 1 に示すように，ベースプログラムとの 4 対 1 の直接対戦では大きく優位である．

序盤中盤は同じ思考であっても，終盤データベースを用いることで，最終的な勝敗を有利

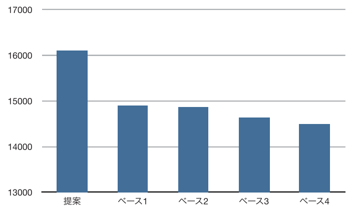

図 1　5,000 試合，提案モデル対ベース x4 直接対戦
Fig. 1　Proposed model player vs. base palyer x 4, 5,000 matches.

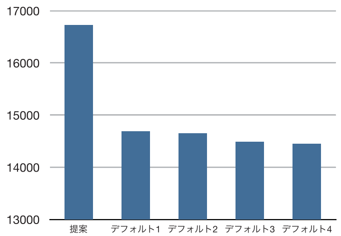

図 2　5,000 試合，提案モデル対デフォルト x4
Fig. 2　Proposed model player vs. default palyer x 4, 5,000 matches.

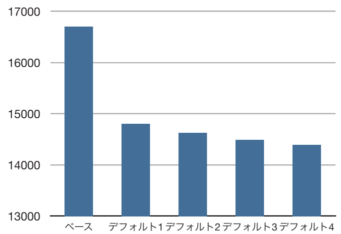

図 3　5,000 試合，ベース対デフォルト x 4
Fig. 3　base player vs. default palyer x 4, 5,000 matches.

にすすめている．このことから，本論文で提案した簡略化をほどこして構築した終盤データベースが，多人数不完全情報ゲームに適用できることが分かった．

図 4 や図 5 での，提案モデルとベースとの得点差が減っており，デフォルトクライアントが加わった場合にはその有利さが低くなる傾向がみられる．

図 2 および図 3 では，デフォルトクライアントとの 4 対 1 直接対戦をそれぞれ行い，提

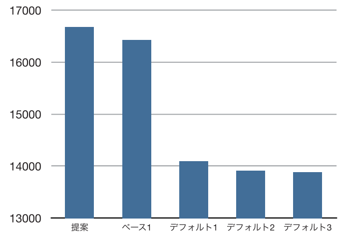

図 4　5,000 試合，提案モデル対ベース，デフォルト x3
Fig. 4　Proposed model player vs. base player, default palyer x 3, 5,000 matches.

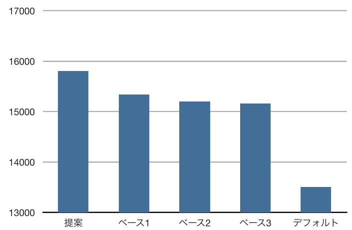

図 5　5,000 試合，提案モデル対ベース x3，デフォルト
Fig. 5　Proposed model player vs. base player x 3, default palyer, 5,000 matches.

案モデルの総利得がベースの総利得をやや上回っている．1 対 1 の相性であれば提案モデルはベースより明らかに優位であるから，三つどもえになったときの二者の関係に第三者の存在が影響する結果となっている．第三者であるデフォルトクライアント自体は，提案プログラムにも，ベースプログラムにも勝てないものの，提案方式対ベース方式において，ベース方式を優位に持ち上げる触媒として働いているといえる．

6. ま と め

本論文では，不決定性のある多人数ゲームにおいて，$Soft-Max^n$ 探索の考え方をもとに，評価値をシングルトンによる単純化した終盤データベースの構築を行うことを提案した．さらにこの終盤データベースを用いた多人数不完全情報ゲームのプレイヤモデルへの応用により終盤データベースの有用性を示した．大貧民プレイヤに適用した実験により，終盤データベースを用いたプレイヤモデルは，用いないルールベース型のベースプレイヤより強くなり，作られた終盤データベースの有効性を示した．

終盤として最終 10 枚を仮定し，データベースを構築するための単純化した終盤モデルとして完全情報多人数ゲーム単貧民を導入した．これによりゲーム状態の縮約，シングルトン評価値による探索を可能としたうえで，3 人および 4 人に 10 枚が配られた状況について終盤データベースを構築した．

完全情報として全探索し勝敗と必勝手を明らかにした．探索に際して，大貧民のルールに即した手の構造に基づいて等価な手を定義し，パターンの縮約を行い，3 人については 1,428,867 通りに減らすことができた．4 人に対しては組合せ数が 818,520 通りと 3 人に比べて少なく，手の散布が多様でパターンの縮約の余地も少ないため，そのまま全体の組合せに対して探索を行った．3 人 4 人それぞれ，38%，25%の配布パターンについて，同値による非決定な節点を含まない単一の確定的な解を求めることができた．

単貧民型の終盤データベースを用いることで，ベースプレイヤに対するパフォーマンスが向上したが，第三者としてデフォルトクライアントを加えるとその優位性が減少することを見いだした．多人数ゲームの本質として，自己の利得を合理的に予測できない不決定局面がある．試合参加するプレイヤの組合せによって，この不決定性に似た触媒的な働きがあることが分かった．同時に試合を行うプレイヤの組合せの選び方自体，いわゆる面子の揃え方も重要であることが分かった．

今回作成した終盤データベースは，完全情報の単貧民を仮定し，かつ，シングルトン評価による探索の簡略化を行ったものである．その組合せから確率的に選ぶことで，不完全情報ゲームへの対応も行うことができた．これらの工夫は計算量を少なくして全探索する効果は高かったが，実際のゲームでは終盤になっても提案モデルでの仮定に反するペア出しや 8 切りも発生する．

今後はこうしたペアを含む手に対して，どの程度シングルトン方式が有用であるかの定量評価や，ペアなども考慮した終盤データベースとの比較，不完全情報への異なるアプローチも検討する必要がある．本手法で提案したモデルとアルゴリズムは，大貧民に限らず終盤データベースを構成できる多人数ゲームに適用可能である．他のゲーム類や，展開型多人数ゲームとして定式化された各種の問題への適用可能性も期待される．

参考文献

1) Luckhardt, C.A. and Irani, K.B.: An algorithmic solution of N-person games, *AAAI-86*, pp.158–162 (1986).
2) Sturtevant, N. and Bowling, M.: Robust Game Play Against Unknown Opponents, *AAMAS'06*, ACM, pp.713–719 (2006).
3) Sturtevant, N.: An Analysis of UCT in Multi-player Games, *Computers and Games*, Lecture Notes in Computer Science, Vol.5131, pp.37–49, Springer (2008).
4) Björnsson, Y., Schaeffer, J. and Sturtevent, N.: Imperfect Information Endgame Databases, *Advances in Computer Games 11*, Lecture Notes in Computing Science, pp.11–22, Springer (2005).
5) Sturtevant, N.: Leaf-value tables for pruning non-zero-sum games, *IJCAI'05: Proc. 19th international joint conference on Artificial intelligence*, San Francisco, CA, USA, pp.317–323, Morgan Kaufmann Publishers Inc. (2005).
6) 大久保，小林，本多，眞鍋，青木，柿下，小松原，西野：第 1 回コンピュータ大貧民大会（UECda-2006）の報告，情報処理学会ゲーム情報学研究報告，Vol.GI-17, pp.25–32 (2007).
7) 西野哲朗：第 1 回 UEC コンピュータ大貧民大会（UECda-2006）の実施報告，情報処理学会誌，Vol.48, No.8, pp.884–888 (2007).
8) 西野順二：大貧民における手の構造，情報処理学会ゲーム情報学研究報告，Vol.GI-17, pp.33–39 (2007).
9) 西野順二：単貧民における多人数完全情報展開型ゲームの考察，第 12 回ゲームプログラミングワークショップ，pp.66–73 (2007).
10) 有澤　誠：3.3 パズルとゲーム，55 枚大貧民の実験的分析，プログラミングレクリエーション（2），近代科学社 (1982).
11) 後藤，乾，小谷：多人数ゲームの順位を決定するゲーム木探索，第 7 回ゲームプログラミングワークショップ，pp.109–115 (2002).
12) 橋本，平沢，梶原，佐々木，飯田：四人将棋プログラムの基本的アルゴリズム，情報処理学会ゲーム情報学研究報告，Vol.GI-1, pp.99–106 (1999).
13) 大川，桜井，小谷，辻：多人数ゲームにおける枝刈りと四人将棋への応用，情報処理学会ゲーム情報学研究報告，Vol.GI-7, pp.73–80 (2002).
14) 梶浦正浩，安西祐一郎：早い者勝ち n 人ゲームの木探索，電子情報通信学会論文誌，Vol.J77-D-I, No.3, pp.253–260 (1994).
15) 篠原拓嗣，石田　亨：N 人ゲームにおける最良優先探索，情報処理学会論文誌，Vol.43, No.10, pp.2981–2989 (2002).

(平成 21 年 8 月 13 日受付)
(平成 21 年 12 月 14 日再受付)
(平成 22 年 1 月 13 日採録)

西野　順二（正会員）

昭和 42 年生．平成 7 年東京工業大学大学院総合理工学研究科システム科学専攻博士課程退学．同年福井大学情報工学科助手，平成 11 年同知能システム工学科助手，平成 13 年電気通信大学システム工学科助手，平成 21 年同助教．平成 14 年人工知能学会研究奨励賞，平成 17 年度 IPA 認定天才プログラマー，平成 19 年度山下記念研究賞．ファジィシステムおよび知的情報処理の研究に従事．日本知能情報ファジィ学会会員．

西野　哲朗（正会員）

昭和 34 年生．昭和 57 年早稲田大学理工学部数学科卒業．昭和 59 年同大学院理工学研究科博士前期課程修了．同年日本アイ・ビー・エム（株）入社．昭和 62 年東京電機大学理工学部情報科学科助手．平成 4 年北陸先端科学技術大学院大学助教授．平成 6 年電気通信大学電気通信学部助教授．平成 18 年同教授，現在に至る．理学博士．回路計算量理論，量子計算量理論，計算論的学習理論等の研究に従事．平成 7 年情報処理学会 Best Author 賞，平成 10 年人工知能学会研究奨励賞，平成 14 年電子情報通信学会ソサイエティ論文賞，平成 15 年船井情報科学振興賞，平成 20 年 IBM Faculty Award 各受賞．日本ソフトウェア科学会，人工知能学会，日本数学会，ACM，IEEE，EATCS 各会員．