

修 士 論 文

キングメーカーが存在する
展開型ゲームの均衡点

Equilibrium Points of
Games in Extensive Form with
Kingmakers

指導教員

近山 隆 教授



東京大学大学院 工学系研究科
電気系工学専攻

氏 名

37-106498 水野 悠

提 出 日

平成 24 年 2 月 8 日

概要

プレイヤーの数が3人以上の多人数ゲームを行うコンピュータープレイヤーには、 \max^n アルゴリズムが用いられることが多い。 \max^n アルゴリズムが正しく動作することの根拠はゲーム理論の「部分ゲーム完全均衡点」である。

ゲーム理論の分野では利得が実数であるゲームを主な対象としており、ほとんどのゲームで、部分ゲーム完全均衡点が一意に定まる。しかし、現実によく存在する「多人数2値ゲーム」では利得が2値で表され、ほとんどの場合「キングメーカーノード」が含まれる。このようなゲームに対して、部分ゲーム完全均衡点は無数に存在し、その多くは合理的でないものである。

本論文では、「多人数2値ゲーム」を対象とし、既存手法である部分ゲーム完全均衡点で導かれる合理的でない戦略の組を排除し、合理的な戦略の組のみを導く、既存の均衡点よりも強い条件により定義される均衡点を提案した。

目次

第 1 章	序論	1
1.1	背景	1
1.2	本研究の目的	1
1.3	本研究の貢献	2
1.4	本論文の構成	2
第 2 章	関連研究	3
2.1	ゲーム理論における展開形ゲームの解	3
2.1.1	ゲーム理論	3
2.1.2	戦略形ゲーム	4
2.1.3	ナッシュ均衡点	5
2.1.4	展開形ゲーム	10
2.1.5	部分ゲーム完全均衡点	13
2.1.6	均衡点の精緻化	15
2.2	コンピュータプレイヤーの用いる探索手法	16
2.2.1	探索手法	16
2.2.2	Mini-max アルゴリズム	17
2.2.3	Max ⁿ アルゴリズム	20
2.2.4	Max ⁿ アルゴリズムの応用手法とその他の探索アルゴリズム	23
第 3 章	2 値ゲーム	26
3.1	2 値ゲームとキングメーカーノード	26
3.2	2 値ゲームでの既存手法の挙動	28
3.3	利得実数ゲームで用いられている方法	30
第 4 章	提案手法	32
4.1	kmn- η 則	32
4.2	2 値ゲームの均衡点	34
4.2.1	一段階予測均衡点	34
4.2.2	部分ゲーム完全一段階予測均衡点	37
4.3	同値変形	40
4.4	ゲームへの適用	41

4.4.1	3人2値ゲーム	41
4.4.2	4人以上2値ゲーム	43
第5章	実験と評価	45
5.1	ダイヤモンドゲーム	45
5.2	各種実験と結果	47
5.2.1	実験1:戦略が一意に定まるノードの数	47
5.2.2	実験2:ゲーム木の同値変形	48
5.2.3	実験3:既存手法と提案手法の対戦	48
5.3	考察	50
第6章	結論	52
6.1	まとめ	52
6.2	今後の課題	52

目次

2.1	予測の連鎖とナッシュ均衡点	6
2.2	展開形ゲーム	10
2.3	チェーンストアゲーム	13
2.4	部分ゲーム完全均衡点	15
2.5	全探索可能な場合の mini-max アルゴリズム	18
2.6	全探索不可能な場合の mini-max アルゴリズム	20
2.7	全探索可能な場合の \max^n アルゴリズム	22
2.8	全探索可能な場合の \max^n アルゴリズム	23
3.1	必勝ノードの例	27
3.2	キングメーカーノードの例	28
3.3	キングメーカーノードを含むツリーでの部分ゲーム完全均衡点	29
3.4	部分ゲーム完全均衡点の例	30
4.1	どの戦略を選んでも利得が 0 と推論してしまう例	32
4.2	どの戦略を選んでも利得が 1 と推論してしまう例	33
4.3	kmn- η 則による自然な選択好順	34
4.4	一段階予測に基づく戦略変更の動機	36
4.5	一段階予測均衡点	37
4.6	2つの一段階予測均衡点	38
4.7	同値変形	41
4.8	戦略が定まるノードでの同値変形	42
4.9	一段階予測均衡点の同値変形の例	42
4.10	3人2値ゲームの同値変形後	43
4.11	4人2値ゲームの同値変形例 1	43
4.12	4人2値ゲームの同値変形例 2	44
4.13	4人2値ゲームの同値変形例 3	44
5.1	ダイヤモンドゲーム	45
5.2	ダイヤモンドゲームのゲーム木	46
5.3	kmn- η 則に従うゲーム木の同値変形	49
5.4	部分ゲーム完全一段階予測均衡点をとるゲーム木の同値変形	49

表目次

2.1	戦略形ゲームの例 1 : 囚人のジレンマ	5
2.2	戦略形ゲームの例 2 : じゃんけん	5
5.1	戦略が一意に定まらないノードの数	47
5.2	同値変形によるノードの数の削減	48
5.3	相手の戦略を知らない場合の各戦略の平均利得	50
5.4	相手の戦略を知る場合の各戦略の平均利得	50

第1章 序論

1.1 背景

思考ゲームを扱うコンピュータゲームプレイヤーは人工知能の一分野として研究が行われている。もっとも研究が進んでいるのは2人零和完全情報確定ゲームの分野である。主なゲームの例として、チェス、オセロ、囲碁、将棋、バックギャモンなどが挙げられる。これらのゲームはコンピュータゲームプレイヤーが人間のトッププロを破るほどの実力となるほど研究が進められている。

一方、プレイヤーの数が3人以上の多人数ゲームでは、2人零和ゲームの手法を単純に拡張しただけでは適用が困難である場合が多く、研究が進んでいない部分が多い。

多人数ゲームのコンピュータゲームプレイヤーでは、 \max^n アルゴリズムが広く用いられており、この手法が各ノードでの最善手を導くことを前提として議論や手法の提案・改良がなされることが多い。この \max^n アルゴリズムとは、各プレイヤーが自分の利得が最大である子ノードを選ぶことを仮定して他のプレイヤーの行動を予測する手法である。 \max^n アルゴリズムが正しいとされる根拠は、ゲーム理論の分野の部分ゲーム完全均衡点である。ゲーム理論の分野では、利得が実数のゲームを主な対象としており、部分ゲーム完全均衡点は多くのゲームで合理的なゲームの解を一意に導く。

しかし現実の世界には、各プレイヤーの目的が勝利することのみであり、利得が2値である「多人数2値ゲーム」が多く存在する。「多人数2値ゲーム」に対しては、ほとんどの場合、部分ゲーム完全均衡点が無数に存在し、その多くは合理的でないものになってしまう。これは、多人数2値ゲームに多くの、どのような戦略をとっても利得が変わらない「キングメーカーノード」が含まれ、既存の均衡点ではキングメーカーノード上での戦略を一意に定めることが出来ないためである。

1.2 本研究の目的

本論文では「多人数2値ゲーム」を対象とし、既存手法である部分ゲーム完全均衡点で導かれる合理的でない戦略の組を排除し、合理的な戦略の組のみを導く方法を提案することを目的とした。

この目的を達成するために、各プレイヤーに子ノード選択の自然な優先順位を持たせることができる「 $kmn-\eta$ 則」と、既存の均衡点よりも強い条件持つ「一段階予測均衡点」を定義し、その合理性を示した。

また「多人数2値ゲーム」の解析を容易にするため、ゲームの「本質」を変化させないまま、ゲーム木を単純化するための「同値変形」の手法を提案した。

1.3 本研究の貢献

本研究で定義した「 $kmn-\eta$ 則」「一段階予測均衡点」により、既存手法ではゲームの解として導かれてしまう、合理的でない戦略の組を排除し、合理的な戦略の組のみをゲームの解として導くことが可能となった。このことにより、現実のゲームに対する根拠のある探索手法の導かれたといえる。

「同値変形」により、ゲーム木の「本質」を最低限の大きさのゲーム木で表現することが可能になり、ゲームの性質の直感的な理解、解析が容易になった。

また、特に 3 人で行う 2 値ゲーム「3 人零和 2 値ゲーム」に対して提案手法を適用することにより、3 人中 2 人のプレイヤーの全ての手番ノードでの振る舞いを一意に定めることが可能であり、ゲームの「本質」は、1 つのキングメーカーノードと 2 つのリーフであることを示した。

1.4 本論文の構成

本論文の構成は以下の通りである。

- 第 2 章では、関連研究である、ゲーム理論における展開形ゲームの解と、ゲームプレイヤーの探索手法について述べる。
- 第 3 章では、本論文で対象としている 2 値ゲームと、2 値ゲームに既存手法を適用した場合の挙動について述べる。
- 第 4 章では、3 つの提案手法について、その前提と定義について詳しく述べる。
- 第 5 章では、提案手法について行った評価実験と、その結果について述べる。
- 第 6 章では、本論文のまとめと、今後の課題について述べる。

第2章 関連研究

2.1 ゲーム理論における展開形ゲームの解

ゲーム理論とは現実世界に起こる様々なゲームの状況を解析する学問である [28] . 現実に存在するゲームを対象とすることも可能である . ゲームだけでなく , 様々な人間や組織の行動のモデルとして適用が可能である .

2.1.1 ゲーム理論

この節ではゲーム理論の基本的な知識 , 用語について述べる .

ゲームには何人かの「プレイヤー」が参加する . 各プレイヤーの目的は , ゲームの結果によって決定する自分の「利得」の最大化である . それぞれのプレイヤーは定められた「ルール」の範囲で自分の行動を選択することが出来る . この行動をどのように選ぶかを「戦略」と呼び , 各プレイヤーの戦略のとり方によってゲームの結果が変化し , それぞれのプレイヤーの利得が決定する .

ここで「共有知識」の概念について述べる . 「共有知識」とは「広く知れ渡ったことがら」の意味で使われ , 次のように定義されている .

定義 2.1.1 (共有知識). 事象 X が n 人のプレイヤー $1, \dots, n$ の共有知識であるとは ,

- 事象 X をプレイヤー a が知っている , ということをプレイヤー b が知っている , ということをプレイヤー c が知っている , ということを...

という命題を任意の回数繰り返し , a, b, c, \dots に $1, 2, 3, \dots$ を任意の順番で当てはめても真になることである .

ゲームのルールは全プレイヤーの「共有知識」である .

また , 各プレイヤーは自分の利得を最大化するために , 自分の戦略をどのようにとれば良いかについて , 十分な論理的思考を持って考察し , 最良な戦略を必ずとることが仮定されている . このことを各プレイヤーは「合理的」であるという . また , 各プレイヤーは自分が相手の立場であればどのような戦略をとるかについても十分な論理的思考を持って考察することができる . このことを各プレイヤーは「理性的」であるという . 各プレイヤーが合理的かつ理性的であることは , 全プレイヤーの共有知識である .

2.1.2 戦略形ゲーム

ゲーム理論の重要な概念である均衡点について説明するため、まず、戦略形ゲームの定義と用語を導入する。

戦略形ゲームは様々なゲームの基本となる構造をもったゲームで、「標準形ゲーム」とも呼ばれる。各プレイヤーが取ることのできる戦略から 1 つを選択し、その戦略の組み合わせによって、各プレイヤーへの利得が決定する。

戦略形 n 人ゲーム G は以下のように 3 つの要素によって表される。

$$G = (N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N})$$

各要素の意味は以下の通り。

- $N = \{1, \dots, n\}$ はプレイヤーの集合
- S_i はプレイヤー i が選択可能な戦略の集合
- f_i は直積集合 $S = S_1 \times \dots \times S_n$ 上の実数値関数。プレイヤー i の利得関数。

これらゲームの定義は共有知識である。

ゲームのプレイは次の手順で行われる。各プレイヤー $1, \dots, n$ は他のプレイヤーの戦略の選択を知らずに、それぞれ自分の戦略 $s_1 \in S_1, \dots, s_n \in S_n$ を選択する。その結果、プレイヤー i は利得 $f_i(s_1, \dots, s_n)$ を得る。

各プレイヤーが確定的にいずれかの戦略を選ぶことを「純戦略」と呼ぶ。以上に述べた戦略形ゲームの定義は純戦略の範囲でのものである。純戦略に対して、各プレイヤーがある確率分布に従って戦略を選ぶことを「混合戦略」と呼ぶ。以下に混合戦略の範囲での戦略形ゲームの定義を述べる。

戦略形 n 人ゲーム $G = (N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N})$ の混合拡大は以下のように 3 つの要素によって定義される。

$$G = (N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N})$$

- $N = \{1, \dots, n\}$ はプレイヤーの集合
- Q_i はプレイヤー i の混合戦略 q_i の集合。混合戦略 q_i は S_i 上の任意の確率分布である。
- F_i はプレイヤー i の期待利得関数。 F_i は直積集合 $Q = Q_1 \times \dots \times Q_n$ 上の実数値関数であり、 $q = (q_1, \dots, q_n) \in Q$ を入力とし、各プレイヤーが混合戦略 q_1, \dots, q_n をとったときのプレイヤー i の期待利得を返す。 $q_j(s_j)$ を混合戦略 q_j が純戦略 s_j に付与する確率として、次の式で定義される。

$$F_i(q_1, \dots, q_n) = \sum_{s_1 \in S_1} \dots \sum_{s_n \in S_n} \left\{ \prod_{j=1}^n q_j(s_j) \right\} f_i(s_1, \dots, s_n)$$

このゲームの定義は全プレイヤーの共有知識である。

標準形ゲームの例を 2 つ示す。表 2.1 に示したゲームは「囚人のジレンマ」とよばれる有名な標準形ゲームである。内容は、共犯である二人の囚人が犯行について自白を迫られている場面であり、2

表 2.1: 戦略形ゲームの例 1 : 囚人のジレンマ

		2 の戦略	
		s_{21} (黙秘)	s_{22} (自白)
1 の戦略	s_{11} (黙秘)	$f_1 = -3, f_2 = -3$	$f_1 = -10, f_2 = -1$
	s_{12} (自白)	$f_1 = -1, f_2 = -10$	$f_1 = -5, f_2 = -5$

表 2.2: 戦略形ゲームの例 2 : じゃんけん

		2 の戦略		
		s_{21} (グー)	s_{22} (チョキ)	s_{23} (パー)
1 の戦略	s_{11} (グー)	$f_1 = 1/2, f_2 = 1/2$	$f_1 = 1, f_2 = 0$	$f_1 = 0, f_2 = 1$
	s_{12} (チョキ)	$f_1 = 0, f_2 = 1$	$f_1 = 1/2, f_2 = 1/2$	$f_1 = 1, f_2 = 0$
	s_{13} (パー)	$f_1 = 1, f_2 = 0$	$f_1 = 0, f_2 = 1$	$f_1 = 1/2, f_2 = 1/2$

人の囚人はそれぞれ「自白」「黙秘」のいずれかを行う。2人の囚人の行動によってそれぞれの刑期が決まる。

- 両方が黙秘をすると懲役は3年。
- 片方のみが自白をすると、自白したものは懲役が1年。黙秘をしたものは懲役が10年。
- 両方が自白をすると懲役は5年。

囚人はお互いの行動を知ることが出来ない。

上に述べたゲーム理論の用語をこの囚人のジレンマに当てはめると、2人の囚人はそれぞれプレイヤー1とプレイヤー2であり、プレイヤー*i*はそれぞれ戦略 s_{i1} (黙秘) か戦略 s_{i2} (自白) を個別に選ぶ。その結果、プレイヤー*i*は表 2.1 に示す利得 f_i を得る。

2つ目の例、表 2.2 は2人で行うじゃんけんを戦略形ゲームとして表したものである。プレイヤー*i*はそれぞれ戦略を $\{s_{i1}, s_{i2}, s_{i3}\} = \{\text{グー}, \text{チョキ}, \text{パー}\}$ から個別に選び、その結果、プレイヤー*i*は表 2.2 に示した利得を得る。勝ったプレイヤーの利得は1、負けたプレイヤーの利得は0、同点ならばそれぞれ1/2の利得を得る。

2.1.3 ナッシュ均衡点

ナッシュ均衡点とは、Nash[12, 11] によって定義された「ゲームの解」の概念である。

ゲームにおいて自分の利得を最適化するためには、他のプレイヤーがどのような戦略をとるのかを予測し、それに対して自分の戦略を決定することが重要である。しかし、戦略決定の際に互いの戦略が相互に依存する状況で、相手の戦略完全に予測することは本質的に不可能である。このことを説明する、Morgenstern が示した有名な例である「シャーロック・ホームズとモリアティ教授の追跡

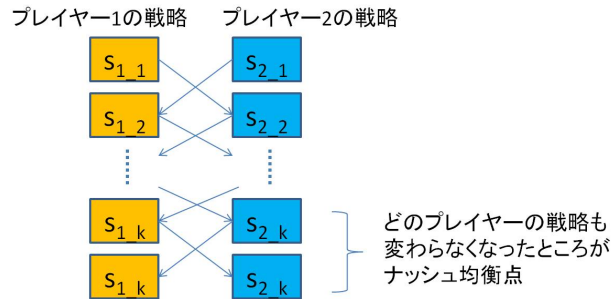


図 2.1: 予測の連鎖とナッシュ均衡点

ゲーム」について述べる [10, 28] シャーロック・ホームズとモリアティ教授はともに高度な推論能力を持つ。シャーロック・ホームズがモリアティ教授の行動を A と予想し、 A に対する最善な行動 B をとろうとする。モリアティ教授は、シャーロック・ホームズのこのような推論を予測することができるので、 B に対する最善な行動 C をとろうとする。シャーロック・ホームズはこの推論をさらに予測して、 C に対して最善な行動 D をとろうとする。以後、2 人の間で予測の連鎖が際限なく続き、お互いは相手の行動を完全に予測することが出来ない。

ナッシュ均衡点の概念はこのような予測の連鎖が止まる戦略の組を、合理的・理性的なプレイヤーたちが整合的に満たす「ゲームの解」とするものである。このゲームの解である戦略の組のことをナッシュ均衡点と呼ぶ。プレイヤーが 3 人以上の場合も同様である。

このことを模式的に示したのが図 2.1 である。矢印は相手の戦略に対する最善の戦略を示している。相手の戦略に対する最善の戦略を、既にお互いのプレイヤーがとっている状態がナッシュ均衡点である。

ナッシュ均衡点の定義について述べるため「最適応答」という概念について述べる。 $G = (N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N})$ を純戦略の範囲の戦略形 n 人ゲームとして、プレイヤーの戦略の組を $s = (s_1, \dots, s_n)$ と表し、 s から第 i 成分 s_i を除いた $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ をプレイヤー i 以外の $n-1$ 人のプレイヤーの戦略の組と呼ぶ。プレイヤー i の戦略 s_i と、その他の $n-1$ 人のプレイヤーの戦略の組 s_{-i} を並べると n 人のプレイヤーの戦略の組 s を表す。すなわち $s = (s_i, s_{-i})$ である。この表記を用いて、戦略の組 s に対するプレイヤー i の利得 $f_i(s)$ を $f_i(s_i, s_{-i})$ と表記する。

定義 2.1.2 (最適応答). プレイヤー i の戦略 $s_i \in S_i$ が他の $n-1$ 人のプレイヤーの戦略 $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n)$ に対する最適応答であるとは、

$$f_i(s_i, s_{-i}) = \max_{t_i \in S_i} f_i(t_i, s_{-i}) \quad (2.1)$$

であることをいう。

戦略の組 s_{-i} に対するプレイヤー i の最適応答の全体を $B_i(s_{-i})$ と表す。

他の $n-1$ 人のプレイヤーの戦略 s_{-i} が予想などで定まったときに、自分の利得を最大化する戦略 s_i が最適戦略である。

例として図 2.1 に示した囚人のジレンマにおける、プレイヤー 2 の戦略 s_{21} (黙秘) に対するプレイヤー 1 の最適応答を考える。 $f_1(s_{11}, s_{21}) = -3$ と $f_1(s_{12}, s_{21}) = -1$ を比べて $f_1(s_{11}, s_{21}) < f_1(s_{12}, s_{21})$ なので、 s_{21} (黙秘) に対するプレイヤー 1 の最適応答は s_{12} (自白) である。同様に考えて、プレイヤー 2 の戦略 s_{22} (自白) に対するプレイヤー 1 の最適応答も s_{12} (自白) である。プレイヤー 1 のそれぞれの戦略に対するプレイヤー 2 の最適応答を考えると、 s_{11} (黙秘) に対するプレイヤー 2 の最適応答は s_{22} (自白)、 s_{12} (自白) に対するプレイヤー 2 の最適応答も s_{22} (自白) である。

図 2.2 に示したじゃんけんの場合、プレイヤー 2 の戦略 s_{21} (グー) に対するプレイヤー 1 の最適応答は s_{13} (パー)、プレイヤー 2 の戦略 s_{22} (チョキ) に対するプレイヤー 1 の最適応答は s_{11} (グー)、プレイヤー 2 の戦略 s_{23} (パー) に対するプレイヤー 1 の最適応答は s_{12} (チョキ) である。

ナッシュ均衡点は以下のように定義される。

定義 2.1.3 (ナッシュ均衡点). 戦略形 n 人ゲーム G において、プレイヤーの戦略の組 $s^* = (s_1^*, \dots, s_n^*)$ がナッシュ均衡点であるとは、全てのプレイヤー $i (i = 1, \dots, n)$ に対して、戦略 s_i^* が他のプレイヤーの戦略の組 s_{-i}^* に対する最適応答であることをいう。

この定義を言い換えると、戦略の組 $s^* = (s_1^*, \dots, s_n^*)$ がナッシュ均衡点であるとは、全てのプレイヤー $i (i = 1, \dots, n)$ に対して、

$$f_i(s_i^*) \geq f_i(s_i, s_{-i}), \forall s_i \in S_i \quad (2.2)$$

が成立することである。この式は、ナッシュ均衡点においてはどのプレイヤーにとっても「他のプレイヤー誰も戦略を変えない場合、自分が戦略を変えても自分の利得が増加することがないので、戦略を変更する動機を持たない」ということを示している。

例として図 2.1 に示した囚人のジレンマにおけるナッシュ均衡点を求める。プレイヤー 2 の戦略 s_{21}, s_{22} のどちらに対してもプレイヤー 1 の最適応答は s_{12} (黙秘) であり、プレイヤー 1 の戦略 s_{11}, s_{12} のどちらに対してもプレイヤー 2 の最適応答は s_{22} (黙秘) であるので、ナッシュ均衡点である戦略の組 s は $s = (s_{12}, s_{22})$ (黙秘, 黙秘) のみであることが分かる。

一方、図 2.2 に示したじゃんけんの場合、プレイヤー 2 の戦略 s_{21}, s_{22}, s_{23} のそれぞれに対するプレイヤー 1 の最適応答はそれぞれ s_{13}, s_{11}, s_{12} であり、プレイヤー 1 の戦略 s_{13}, s_{11}, s_{12} のそれぞれに対するプレイヤー 2 の最適応答はそれぞれ s_{22}, s_{23}, s_{21} であるので、お互いの戦略が相手の戦略に対する最適応答になっている戦略の組は存在せず、純戦略の範囲ではナッシュ均衡点は存在しない。

ここまでは純戦略の範囲でのナッシュ均衡点の定義を述べた。純戦略の範囲ではナッシュ均衡点は必ずしも存在するとは限らないが、混合戦略の範囲では少なくとも 1 つの均衡点が存在することが示されている [11]。混合戦略の範囲でのナッシュ均衡点の定義を以下に述べる。それぞれの概念は純戦略の範囲の場合と同じである。

$G = (N, \{Q_i\}_{i \in N}, \{F_i\}_{i \in N})$ を戦略形 n 人ゲームの混合拡大として、プレイヤーの混合戦略の組を $q = (q_1, \dots, q_n)$ と表し、 q から第 i 成分 q_i を除いた $q_{-i} = (q_1, \dots, q_{i-1}, q_{i+1}, \dots, q_n)$ をプレイヤー i 以外の $n-1$ 人のプレイヤーの戦略の組と呼ぶ。プレイヤー i の戦略 q_i と、その他の $n-1$ 人のプレイヤーの戦略の組 q_{-i} を並べると n 人のプレイヤーの戦略の組 q を表す。すなわち $q = (q_i, q_{-i})$ である。この表記を用いて、戦略の組 q に対するプレイヤー i の利得 $F_i(q)$ を $F_i(q_i, q_{-i})$ とも表記する。

定義 2.1.4 (混合戦略の範囲での最適応答). 戦略の組 q_{-i} に対するプレイヤー i の最適応答の全体の集合を $B_i(q_{-i})$ と表し,

$$B_i(q_{-i}) = \{q_i \in Q_i | F_i(q_i, q_{-i}) = \max_{r_i \in Q_i} F_i(r_i, q_{-i})\}$$

写像 $B_i(\cdot)$ はプレイヤー i の「最適応答対応」と呼ばれ, 直積集合 $Q_1 \times \cdots \times Q_{i-1} \times Q_{i+1} \times \cdots \times Q_n$ から集合 Q_i への点対集合写像である. すなわち, 集合 $Q_1 \times \cdots \times Q_{i-1} \times Q_{i+1} \times \cdots \times Q_n$ に含まれる 1 つの要素 q_{-i} (プレイヤー i 以外のプレイヤーの戦略の組) を入力とし, 集合 Q_i の部分集合 (プレイヤー i の戦略のうち q_{-i} に対する最適応答になっているもの全ての集合) を出力とする写像である. さらに, 戦略の組 $q = (q_1, \dots, q_n)$ に対して集合

$$B(q) = B_1(q_{-1}) \times \cdots \times B_n(q_{-n})$$

を定義する. 写像 $B(\cdot)$ は戦略形ゲーム G の最適応答と呼ばれ, 直積集合 $Q = Q_1 \times \cdots \times Q_n$ からそれ自身への点対集合写像である. すなわち, 集合 $Q = Q_1 \times \cdots \times Q_n$ に含まれる 1 つの要素 q (全プレイヤーの戦略の組) を入力とし, 集合 Q の部分集合 (各プレイヤーの戦略が, 自分以外のプレイヤーの戦略の組に対する最適応答になっているような, 全プレイヤーの戦略の組の全ての集合) を出力とする写像である.

定義 2.1.5 (混合戦略の範囲でのナッシュ均衡点). 戦略形 n 人ゲームにおいて混合戦略の組 $q^* = (q_1^*, \dots, q_n^*)$ が

$$q^* \in B(q^*)$$

を満たすとき, q^* はナッシュ均衡点であるという.

図 2.2 に示したじゃんけんの混合戦略の範囲での最適応答対応およびナッシュ均衡点を求める. プレイヤー i がとる混合戦略 q_i が純戦略 s_{i1}, s_{i2}, s_{i3} のそれぞれに付与する確率を p_{i1}, p_{i2}, p_{i3} と表す. $0 \leq p_{i1}, p_{i2}, p_{i3} \leq 1$, $p_{i1} + p_{i2} + p_{i3} = 1$ が成り立つ. プレイヤー 2 が決定した q_2 すなわち p_{21}, p_{22}, p_{23} に対して, プレイヤー 1 が混合戦略を q_1 を用いたときのプレイヤー 1 の期待利得 F_1 は

$$F_1(q_1, q_2) = p_{11} \cdot \left(\frac{p_{21}}{2} + p_{22}\right) + p_{12} \cdot \left(\frac{p_{22}}{2} + p_{23}\right) + p_{13} \cdot \left(\frac{p_{23}}{2} + p_{21}\right) \quad (2.3)$$

で求めることが出来る. 式 (2.3) の右辺の括弧内をそれぞれ A_1, A_2, A_3 とおく. すなわち

$$A_1 = \left(\frac{p_{21}}{2} + p_{22}\right), \quad A_2 = \left(\frac{p_{22}}{2} + p_{23}\right), \quad A_3 = \left(\frac{p_{23}}{2} + p_{21}\right) \quad (2.4)$$

であり, 式 (2.3) は

$$F_1(q_1, q_2) = p_{11} \cdot A_1 + p_{12} \cdot A_2 + p_{13} \cdot A_3 \quad (2.5)$$

となる. この式よりプレイヤー 2 の混合戦略 q_2 に対するプレイヤー 1 の最適応答対応 $B_1(q_2)$ は次の通りになる.

- A_1, A_2, A_3 を比べて, 最大のものが 1 つだけの場合.
 - A_1 が最大の場合, $p_{11} = 1$, $p_{12} = p_{13} = 0$ とする戦略.

- A_2 が最大の場合, $p_{12} = 1, p_{11} = p_{13} = 0$ とする戦略.
- A_3 が最大の場合, $p_{13} = 1, p_{11} = p_{12} = 0$ とする戦略.

それぞれの最適応答対応はこの戦略 1 通りのみ.

- A_1, A_2, A_3 を比べて, 最大のもの 2 つが等しく残りの 1 つが異なる場合

- $A_1 = A_2$ が最大の場合, $p_{11} + p_{12} = 1, p_{13} = 0$ とする戦略.
- $A_1 = A_3$ が最大の場合, $p_{11} + p_{13} = 1, p_{12} = 0$ とする戦略.
- $A_2 = A_3$ が最大の場合, $p_{12} + p_{13} = 1, p_{11} = 0$ とする戦略.

それぞれの式を満たす混合戦略全てが最適応答となる.

- $A_1 = A_2 = A_3$ の場合

- 任意の戦略が最適応答になる. (ただし条件より $0 \leq p_{11}, p_{12}, p_{13} \leq 1, p_{11} + p_{12} + p_{13} = 1$)

プレイヤー 1 の戦略 q_1 に対するプレイヤー 2 の最適応答対応 $B_2(q_1)$ も同様に求まる.

ゲーム G の最適応答 $B = B_1 \times B_2$ に対して, $q \in B(q) = B_1(q_2) \times B_2(q_1)$ となる戦略の組 $q = (q_1, q_2)$ を考えると, q_1 が $p_{11} = p_{12} = p_{13} = 1/3$, q_2 が $p_{21} = p_{22} = p_{23} = 1/3$ である場合のみが, この式を満たしていることが分かる. 従ってこの戦略の組 $q = (q_1, q_2)$ がじゃんけんの唯一のナッシュ均衡点である.

ここでゲーム理論の重要な概念である戦略の「支配」について述べておく.

定義 2.1.6 (強支配). プレイヤー i の 2 つの戦略 s_i と t_i に関して, プレイヤー i 以外の $n-1$ 人のプレイヤーの, 全ての戦略の組 $s_{-i} \in S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$ に対して $f_i(s_i, s_{-i}) > f_i(t_i, s_{-i})$ が成立するとき, 戦略 s_i が戦略 t_i を強支配するという.

定義 2.1.7 (弱支配). プレイヤー i の 2 つの戦略 s_i と t_i に関して以下の 2 つの条件が満たされるとき, 戦略 s_i が戦略 t_i を弱支配するという.

- プレイヤー i 以外の $n-1$ 人のプレイヤーの, 全ての戦略の組 $s_{-i} \in S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$ に対して $f_i(s_i, s_{-i}) \geq f_i(t_i, s_{-i})$ が成立する.
- プレイヤー i 以外の $n-1$ 人のプレイヤーの, 少なくとも 1 つの戦略の組 $t_{-i} \in S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_n$ に対して $f_i(s_i, s_{-i}) < f_i(t_i, s_{-i})$ が成立する.

プレイヤー i のある戦略 t_i が他の戦略 s_i によって強支配されているとき, 明らかにプレイヤー i は戦略 t_i を選ぶ動機を持たないといえる. また, t_i が他のプレイヤーの戦略の組 s_{-i} に対する最適応答になることは, 最適応答の定義よりありえない. 戦略 t_i を選択するより s_i を選択した方がプレイヤー i の利得が大きくなるからである. したがって, ナッシュ均衡点である戦略の組 s が t_i を含むこともない.

プレイヤー i のある戦略 t_i が他の戦略 s_i によって弱支配されているときは, 他のプレイヤーの戦略によっては, 戦略 t_i を選ぶ動機も選ばない動機もない状況が起こり得るが, 戦略 t_i を選ぶ動機を

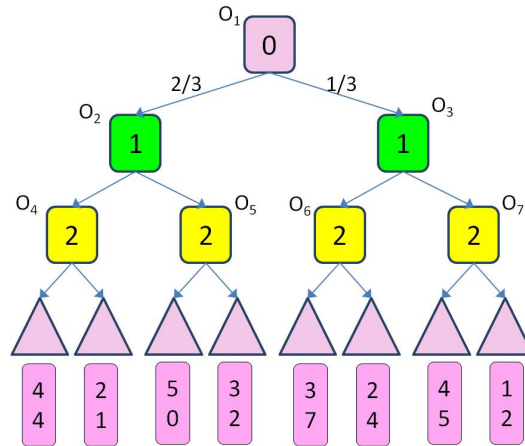


図 2.2: 展開形ゲーム

失うような他のプレイヤーの戦略の組が少なくとも1つは存在する。 t_i が他のプレイヤーの戦略の組 s_{-i} に対する最適応答になることはありうる。この時、必ず戦略 s_i も s_{-i} に対する最適応答になっている。

2.1.4 展開形ゲーム

展開形ゲームの定義・用語について述べる。

展開形ゲームとは各プレイヤーが順に「手番」を行うことによって進んでいくゲームのことである [7]。各手番では手番プレイヤーが何らかの選択を行い、選択の違いによって次に行われる手番が決まり、終局面に到達することで各プレイヤーの利得が決定する。将棋やチェスをはじめとするボードゲームや、現実の社会でみられるゲーム的状况がこの展開形ゲームに分類される。

図 2.2 にその例を示す。図の意味については以降に各要素とともに述べる。

展開形ゲーム Γ は5つの要素の組を用いた表現形式で表され、

$$\Gamma = (K, P, p, U, h)$$

で定義される。各要素について以下に述べる。なお、この定義は全プレイヤーの共有知識である。

ゲーム木 K

ゲーム木 K はゲームの手番の構造を表す木構造の有向グラフである。ゲームの局面がノードとして表されており、その局面の手番プレイヤーが手を選ぶことはノードから出たエッジの先にあるノードのいずれかを選ぶことに相当する。ノード x に対し、エッジの先にあるノードを「子ノード」または「選択肢」と呼び、 $A(x)$ で表す。逆に、子ノードから見た、このノード x を「親ノード」と呼ぶ。

また、同じ親ノードを持つノード同士を「兄弟ノード」、あるノードの親ノード、親ノードの親ノード...をまとめて「先祖ノード」、あるノードの子ノード、子ノードの子ノード...をまとめて「子孫ノード」と呼ぶ。

ゲームの終局面に相当するノードは「リーフノード」、「リーフ」または「頂点」と呼ばれ、その全体の集合を W で表す。リーフノードからは、それに続く局面がないため、エッジは出ていない。このリーフノードを除いたノードを「中間ノード」または「手番」と呼び、その全体の集合を X で表す。単に「ノード」という場合もこの中間ノードを指すことが多い。また、ゲームの開始局面に相当するノードは特に「ルートノード」または単に「ルート」と呼ばれる。ルートノードは1つのゲーム木に1つだけ存在する。

本論文の図では四角でノードで、三角でリーフを表す。図 2.2 の展開形ゲームでは例えばノード O_2 の子ノードは O_4 と O_5 、ノード O_2 の兄弟ノードは O_3 、ノード O_2 の親ノードは O_1 である。ゲームの木 K は親ノードを上、子ノードを下に表すことが多く、エッジの向きが通常上から下であることからエッジを矢印でなく直線のみで表すことが多い。

ルートノードにおいて手番プレイヤーが子ノードのいずれかを選び、次はその子ノードの手番プレイヤーがさらに子ノードを選ぶ...という作業を、選んだ子ノードがリーフノードになるまで繰り返すことが、実際のゲームのプレイに相当する。プレイ中に、ノード x が子ノードとして選ばれることを「ノード x に到達する」と表現する。

プレイヤー分割 P

プレイヤー分割 $P = [P_0, P_1, \dots, P_n]$ はゲーム木 K のノードの全体の集合 X 上の1つの分割である。集合 $P_i (i = 1, \dots, n)$ はプレイヤー i の手番の全体を表す。プレイヤーの集合を $N = 1, \dots, n$ と表す。

集合 P_0 は偶然手番の全体である。偶然手番ではどのプレイヤーの意志とも関係せず手を選ばれる。具体例としては、サイコロを振ったり、トランプを引く、といったものである。

各ノードはいずれかのプレイヤーの手番が偶然手番であり、手番を持たないプレイヤーはいないので、 P は以下の式を満たす。

- $X = P_0 \cup P_1 \cup \dots \cup P_n$
- $i \neq j$ のとき $P_i \cap P_j = \phi$
- $P_i \neq \phi, i = 1, \dots, n$

図 2.2 の展開形ゲームのプレイヤー分割は $P = [P_0, P_1, P_2]$ 、 $P_0 = \{O_1\}$ 、 $P_1 = \{O_2, O_3\}$ 、 $P_2 = \{O_4, O_5, O_6, O_7\}$ である。

偶然手番の確率分布族 p

ゲーム木に含まれる全ての偶然手番 $x \in P_0$ に対して、 x の子ノード $A(x)$ 上の1つの確率分布 p_x があらかじめ定められている。確率分布 p_x が各子ノード $e \in A(x)$ に付与する確率を $p_x(e)$ とする

とき,

$$\sum_{e \in A(x)} p_x(e) = 1, 0 \leq p_x(e) \leq 1$$

が成り立つ．偶然手番 $x \in P_0$ の確率分布 p_x の族を p と表す．

図 2.2 の展開形ゲームでは p の唯一の要素は p_{O_1} , すなわち $p = \{p_{O_1}\}$ である． p_{O_1} が各子ノードの付与する確率は $p_{O_1}(O_2) = 2/3$, $p_{O_1}(O_3) = 1/3$ である．

情報分割 U

本論文で対象としている「完全情報ゲーム」では情報分割 U は本質的な意味を持たないので，簡潔な説明のみに留める．

他のプレイヤーの手番や偶然手番で，どの子ノードを選ばれたかを知ることが出来ない「不完全情報ゲーム」では，各プレイヤーの手番は「その集合のうちいずれかのノードに到達したことは分かるが，それがどのノードなのかは分からない」というノードの集合を表す「情報集合」という概念が必要になる．情報分割 U はこの情報集合がどのような集合によって構成されているかを表している．

各プレイヤーの手番や偶然手番でどの子ノードを選ばれたかを知ることが出来る「完全情報ゲーム」では，各プレイヤーは各情報集合は 1 つ 1 つ ノードからなり，情報分割 U は本質的な意味を持たない．

利得関数 h

ゲーム木に含まれる全てのリーフノード $w \in W$ のそれぞれに対して，各プレイヤーの利得である利得ベクトル $(h_1(w), \dots, h_n(w))$ が定められている．ここで h_i はプレイヤー i の利得である．利得関数 h はリーフノード $w \in W$ に対して，次式に示すように利得ベクトルを対応させる関数である．

$$h(w) = (h_1(w), \dots, h_n(w))$$

本論文では図 2.2 のように各リーフの下に， h がそのリーフに対応させる利得ベクトルを示している．上の数字がプレイヤー 1 の利得，下の数字がプレイヤー 2 の利得である．プレイヤーが 3 人以上の場合も同様に上から i 番目の数字がプレイヤー i の利得を示す．

各プレイヤーの利得の和が常に 0 になる「零和ゲーム」で，プレイヤーの数が 2 人の場合，プレイヤー 1 の利得が定まればプレイヤー 2 の利得も定まるため，利得ベクトルの代わりにプレイヤー 1 の利得のみを用いる場合が多い．このときは利得関数 h は，リーフノード $w \in W$ に対して，次式に示すようにプレイヤー 1 の利得を返す関数である．

$$h(w) = h_1(w)$$

以上が，展開形ゲーム Γ の 5 つの要素である．展開形ゲームに関するその他の用語，基本知識について以下に述べる．

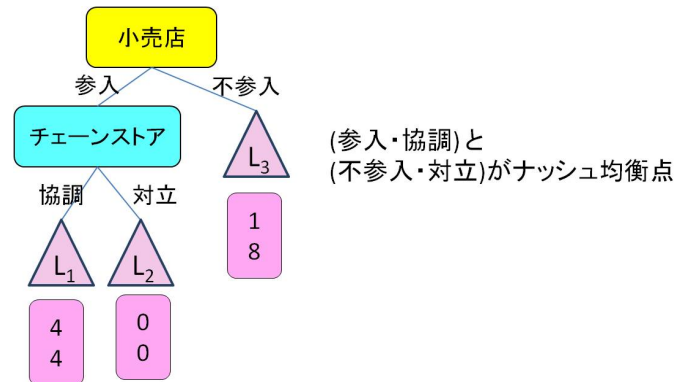


図 2.3: チェーンストアゲーム

展開形ゲームにおいて「戦略」は、自分の手番ノードでどのように子ノードを選択するかを計画のことを意味する。ここでは完全情報ゲームにおける「純戦略」と「行動戦略」について述べる。純戦略とは自分の手番ノードのそれぞれに対し、それぞれ1つの子ノードを確定的に対応させる戦略のことである。行動戦略とは自分の手番ノードのそれぞれに対し、「局所戦略」を対応させる戦略である。局所戦略とは自分の手番ノードの1つに対し、各子ノードを選ぶ確率を示す確率分布を対応させる戦略である。すなわち、行動戦略は自分の手番ノードで確率的に子ノードを選ぶ戦略である。

完全情報ゲームにおいてはゲーム木 K の部分木を1つの展開形ゲームと見ることができる。これを「 Γ の部分ゲーム」という。 Γ そのものも Γ の部分ゲームに含む。特にゲーム木の部分木がそれ自身しか存在しない部分ゲームを「極小」とであるという。極小な部分ゲーム木は、ノードを1つとその子ノードであるリーフのみを含むゲーム木であるともいえる。

2.1.5 部分ゲーム完全均衡点

展開形ゲームのナッシュ均衡点が「不自然な」解となる例が存在する。図 2.3 に示したのは「チェーンストアゲーム」という展開形ゲームでその例を示す。

内容はチェーンストア(チェーン店であるスーパーマーケット)が存在する街に同業種の小売店が参入しようとしている状況である。参加するプレイヤーは、プレイヤー1が「小売店」、プレイヤー2が「チェーンストア」である。各リーフの利得ベクトルの上の数字が小売店の利得、下の数字がチェーンストアの利得を示す。

小売店の戦略は「参入」と「不参入」、チェーンストアの戦略は「協調」と「対立」である。小売店が「不参入」を選択した場合リーフ L_3 に到達する。 L_3 はチェーンストアが街での売上を独占し、小売店は別の業種で利益を得た状態を示している。小売店が「参入」、チェーンストアが「協調」を選択した場合リーフ L_1 に到達する。 L_1 はチェーンストアと小売店が価格協定を結び、街での利益を分け合った状態を示している。小売店が「参入」、チェーンストアが「対立」を選択した場合リー

フ L_2 に到達する． L_2 はチェーンストアと小売店が対立し，価格競争を行った結果，どちらも利益を得られなかった状態を示している．

小売店のそれぞれの戦略に対するチェーンストアの最適応答は「参入」に対する最適応答は「協調」，「不参入」に対する最適応答は「協調」または「対立」である．また，チェーンストアのそれぞれの戦略に対する小売店の最適応答は「協調」に対する最適応答は「参入」，「対立」に対する最適応答は「不参入」である．

したがって，お互いの戦略が相手の戦略の最適応答になっている戦略の組は「参入，協調」と「不参入，対立」である．この 2 つの戦略の組がこのゲームのナッシュ均衡点である．

しかし，この 2 つのナッシュ均衡点のうち「不参入，対立」は以下に述べるように「不自然な」解である．小売店が「参入」を選んだ場合について考えると，チェーンストアは「各プレイヤーは自分の利得を最大化しようとする」というゲーム理論の仮定に反する手をとることになる．

チェーンストアが「対立」を選択することは，小売店が「不参入」を選ぶ動機を持たせるための脅しであるといえる．しかし実際にチェーンストアの手番にくと，チェーンストアは「参入」をとるはずであるので，小売店はチェーンストアは「参入」をとるものと考えてよいことになる．このチェーンストアによる「対立」の選択のような「選択されると他のプレイヤーの戦略の選択に影響するが，合理的な観点からは実際にはとられない戦略」は「信憑性のない脅し」と呼ばれる．

一方，ナッシュ均衡点「参入，協調」は信憑性のない脅しを含まない自然な解である．このような信憑性のない脅しを含まない戦略の組のみを解として導くために，ナッシュ均衡点より定義の条件を強くした「部分ゲーム完全均衡点」の概念が定義された [15] ．

定義 2.1.8 (部分ゲーム完全均衡点)．展開形ゲーム Γ の行動戦略の組 $b^* = (b_1^*, \dots, b_n^*)$ が部分ゲーム完全均衡点であるとは， b^* が Γ の全ての部分ゲームに対してナッシュ均衡点を導くことである．

この部分ゲーム完全均衡点をゲームの解とすることにより，ゲームツリー全体でナッシュ均衡点となっても，部分ゲームでナッシュ均衡点となっていない戦略の組は解から除かれる．

チェーンストアゲームの部分ゲーム完全均衡点を求めると，図 2.4 で示す「参入，協調」のみが解であることが分かる．この均衡点は，チェーンストアの手番をルートととする部分ゲームにおいてもナッシュ均衡点となっている．

完全情報有限ゲームである展開形ゲーム Γ には少なくとも 1 つの部分ゲーム完全均衡点が存在することが証明されている．部分ゲーム完全均衡点を求める方法は「後ろ向き帰納法」と呼ばれる．その手順は以下の通りである．

1. 全ての「極小である， Γ の部分ゲーム」でのナッシュ均衡点を求める．極小とは自分自身のみを部分ゲームとして持つゲームのことである．
2. 1 で求めたナッシュ均衡を前提として，全ての「自分自身と極小なゲームのみを部分ゲームとして持つ， Γ の部分ゲーム」でのナッシュ均衡点を求める．
3. 以下同様に，既に求めたナッシュ均衡点を前提として，「自分自身と，既にナッシュ均衡点を求めた部分ゲームのみを持つ， Γ の部分ゲーム」でのナッシュ均衡点を求めることを繰り返す．

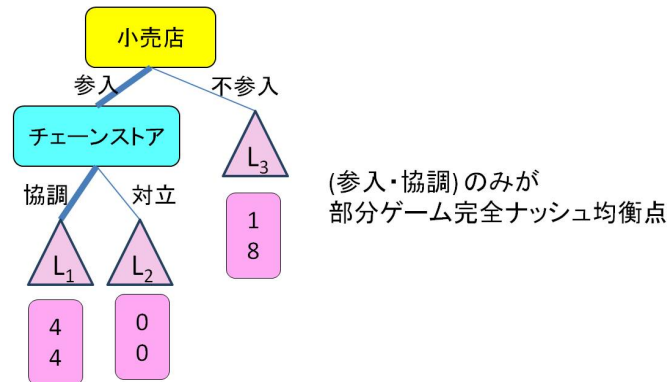


図 2.4: 部分ゲーム完全均衡点

4. Γ 全体でのナッシュ均衡点が求まれば終了する．そのナッシュ均衡点が Γ の部分ゲーム完全均衡点である．

リーフに近いノードから順に，そのノードでの最適応答を求めることを繰り返している，ということも出来る．

2.1.6 均衡点の精緻化

前節ではナッシュ均衡点が展開形ゲーム上で起こす問題を解決する概念として，部分ゲーム完全均衡点が定義されたと述べた．その他にも，既存の均衡点では問題が起こる状況の解決や特定の目的のために新たな均衡点を定義する研究は「均衡点の精緻化 [25]」と呼ばれている．様々な形で均衡点の精緻化が行われている．

いくつかの例を簡潔に述べる．

- 完全均衡点 [16]
ナッシュ均衡点や部分ゲーム完全均衡点では問題が起こる，真の部分ゲームを含まないゲーム木に対し，問題なく均衡点を導くため定義された．プレイヤーが微小な確率 ε で選択を誤る「変動ゲーム」上でナッシュ均衡点を求め， $\varepsilon \rightarrow 0$ としたときに，連続的に収束した均衡点を完全均衡点と定義している．
- 逐次均衡点 [6]
求めるのが困難である完全均衡点の定義を緩め，比較的容易に計算出来るようにした均衡点．
- 強完全均衡点 [14, 28]
いかなる変動ゲーム列に対しても安定に収束する完全均衡点．

本論文の 4 章で提案する「一段階予測均衡点」「部分ゲーム完全一段階予測均衡点」も，均衡点の精緻化の 1 つであるといえる．

詳細については 3 章や 4 章で述べるが、簡潔に説明しておく、既存の部分ゲーム完全均衡点が「多人数 2 値ゲーム」では無数に導かれ、その多くは合理的でない戦略の組である。提案手法ではそれらの合理的でない戦略の組を排除し、合理的な戦略の組のみを導く均衡点である。

2.2 コンピュータープレイヤーの用いる探索手法

現実に存在するゲームを対象に「強いコンピュータープレイヤー」を作ろうとする研究が幅広く行われている。

2.2.1 探索手法

コンピュータープレイヤーが使う探索手法の基礎について述べる。

展開形ゲームで各プレイヤーが行うことは「自分の手番で手を選択すること」である。したがって強いコンピュータープレイヤーを作るために行うべきことは「自分の手番でどのように手を選べば、到達するリーフで高い利得を得ることが出来るか」を考えることである。

一般的には「探索」と呼ばれる手法が用いられる。プレイヤー i の手番である局面 x での手を選ぶための「探索」は以下のように行われる。まず、局面 x をグラフ木のルートノードとみなし、この局面 x でとることの出来る手と、それぞれの手をとった後の局面を調べ、それらをノード x からエッジがつながる子ノードとしてグラフ木に追加する。この「あるノード x' でとることの出来る手を調べ、それぞれの手をとった後の局面を x' の子ノードとしてグラフ木に追加する」ことを「ノード x' を展開する」という。追加されたノードが終局面を表す場合、そのノードはリーフノードとなる。また、リーフノードを展開することは出来ない。追加された x の子ノードを順に展開し、さらに追加された x の子孫ノードも再帰的に展開を行い、ゲーム木を構成する。

構成したゲーム木に含まれる、プレイヤー x の手番ノードでは将来そこに到達した時にどのような手をとるかを決めておくことができるが、他のプレイヤーの手番ノードに関しては、どのような手をとるかが分からないので、なんらかの指針に従って予測を行う。構成したゲーム木の末端から順にとられる手を予測していくと、ルートノード x でそれぞれの手を選べば、ゲームのプレイがどのように進み、どのような局面に到達するかを予測することが出来る。この到達すると予測した局面のうち、最もプレイヤー i にとって望ましい局免へつなげる手を選ぶことがノード x における最善手とすることが出来る。

ゲーム理論における展開形ゲームとの違いとして、ゲーム木の全体が全プレイヤーの既知のものであるかどうかが挙げられる。ゲーム理論の均衡点は、各プレイヤーの共有知識であるゲームのルールにゲーム木の全体が含まれており、各プレイヤーがゲーム木の全体の構造を初めから知っていることを前提として定義されている。しかし現実のゲームでは、各ノードでとることが出来る手とそれぞれの手をとった場合の子ノードを求めるためのルールは、各プレイヤーの共有知識であるが、ゲーム木の全体は共有知識ではない。これは現実に存在するゲームを表すゲーム木は、現実的な時間や計算資源の量では全体を処理出来ないほど巨大であることが多いからである。無限の時間と計算資源があればゲーム木の全体を求めることは可能である。また、現実的な時間と計算資源の量で全体

を求められるほどの大きさのゲーム木に関しても、実際のゲームのルールとゲーム理論の分野の共有知識は本質的に同じであるといえる。この「現実的な時間と計算資源の量で全体を求められるほどの大きさのゲーム木」で表されるゲームを「全探索可能なゲーム」、そうでないものを「全探索不可能なゲーム」と呼ぶことにする。あるゲームが全探索可能か全探索不可能であるかは、計算機の性能や計算時間の制約によって変わる可能性がある。

また、ゲーム理論の主要な目的は、全プレイヤーの戦略が合理的である「ゲームの解」を求めることであった。ゲーム理論の前提を現実のゲームに当てはめると「各プレイヤーは、他のプレイヤーがとろうとしている戦略が分かる」という状況にあたる。これに対し、現実のゲームを行うコンピュータプレイヤーは他のプレイヤーがどのような戦略をとるかは知らず、他プレイヤーの振る舞いをなんらかの指針にしたがって予測している。

以下の節では探索手法の基本的なものである mini-max アルゴリズムと \max^n アルゴリズムについて述べる。Mini-max アルゴリズムは 2 人零和ゲームに対して用いることのできるアルゴリズム、 \max^n アルゴリズムは一般の n 人ゲームに対して用いることのできるアルゴリズムである。どちらも確定ゲーム・不確定ゲームの両方に対して用いることが出来るが、不完全情報ゲームに対して用いることは出来ない。Mini-max アルゴリズムは \max^n アルゴリズムの特別な場合とみなすことも出来るが、利得に関して若干の違いがあり、その違いが拡張・応用の容易さに大きな影響を与えている。詳しくは 2.2.4 で述べる。それぞれのアルゴリズムに関して、全探索可能なゲームに対して用いる場合と全探索不可能なゲームに対して用いる場合の方法について述べる。

2.2.2 Mini-max アルゴリズム

2 人零和ゲームでは、片方のプレイヤーの利得が定まればもう片方のプレイヤーの利得も定まるため、利得ベクトルの代わりに、プレイヤー 1 の利得のみを用いる場合が多い。プレイヤー 1 の利得のみを用いて、2 人零和ゲームの局面における最善手を求める探索手法が mini-max アルゴリズムである。

全探索が可能な場合と可能でない場合の mini-max アルゴリズムについてそれぞれ述べる。

全探索が可能な場合

プレイヤー 1 は「プレイヤー 2 は利得を最小化するように行動する」という合理的な仮定、プレイヤー 2 は「プレイヤー 1 は利得を最大化するように行動する」という合理的な仮定にそれぞれ基づき、与えられた局面における最善手を求める。このことから、プレイヤー 1 の手番ノードは「max ノード」、プレイヤー 2 の手番ノードは「min ノード」と呼ばれる。

手順について述べる。

1. 与えられた局面をルートノードとする。ルートノードを展開する。
2. 全ノードを展開するまで、各ノードの展開を繰り返す。
3. 各ノードの持つ利得を以下の手順で計算する。

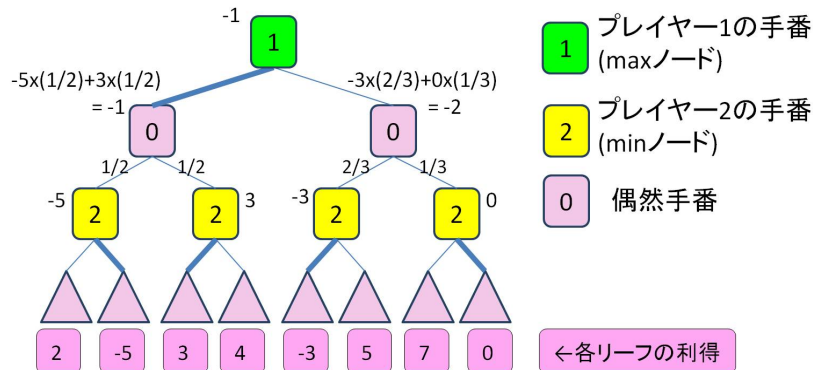


図 2.5: 全探索可能な場合の mini-max アルゴリズム

- リーフの持つ利得は利得関数により求まる。
 - リーフ以外のノード x が持つ利得は全ての子ノードの持つ利得が求まってから求める。子ノードらが持つ利得のうち、 x が max ノードのときは最大のもの、 x が min ノードのときは最小のものを、ノード x の持つ利得とする。この時選ばれた子ノードが、ノード x でプレイヤーが選択する手を示す。
 - なお、ノード x が偶然手番の場合は、各子ノード e の持つ利得にノード x で子ノード e が選ばれる確率 $p_x(e)$ をそれぞれかけたものの和を、ノード x の持つ利得とする。
4. ルートノードの持つ利得が求まるまで 3 を繰り返す。ルートノードで選ばれた子ノードを、与えられた局面における最善手とする。

図 2.5 に全探索が可能な木に対して mini-max アルゴリズムを適用した例を示す。図に示したゲーム木は全てのノードが 2 つづつの子ノードを持っている。各リーフの利得を最初に求め、図の下にあるノードから順にそのノードでの最善手を定めることができる。太い線で示したエッジが各ノードで選ばれる最善手である。ルートノードでの最善手は左のノードを確率 1 で選ぶことであると求めることができる。

全探索が不可能な場合

全探索が不可能ほど大きなゲームを扱う場合、「評価値」という実数値を用いる。評価値はあるノードにおいて、各プレイヤーがそれぞれどれだけ有利であるかを表す数値である。自分の評価値が高いノードほど、そのノードからプレイを行って、リーフに到達した時に得られる利得が高くなるであろうということを意味し、各プレイヤーに選択好順を与えることになる。そのノードからプレイを行った際に得られる利得のある種の「近似」とも言える。

評価値は、ゲームのノードまたはリーフを入力として評価値を出力する関数である「評価関数」によって求められる。例えば、ノード $x \in X$ における各プレイヤー $1, \dots, n$ の評価値 v_1, \dots, v_n は評

価関数 E によって次式に示すように求められる .

$$E(x) = (v_1, \dots, v_n)$$

全プレイヤーの評価値をまとめた n 次のベクトル $v = (v_1, \dots, v_n)$ を「評価値ベクトル」と呼ぶことにする . 2 人零和ゲームの場合 , $v_1 = -v_2$ が常に成立するので , 評価値ベクトルではなく , 片方のプレイヤー (通常はプレイヤー 1) の評価値のみを明示的に用いることが多い .

全探索が不可能な 2 人零和ゲームでは「評価値を用いた mini-max アルゴリズム」が , 与えられたノードの最善手を探索に用いられる . 利得の場合と同様 , 評価値ベクトルではなくプレイヤー 1 の評価値のみを用いる . プレイヤー 1 は「プレイヤー 2 は評価値を最小化するように行動する」という仮定 , プレイヤー 2 は「プレイヤー 1 は評価値を最大化するように行動する」という仮定に基づき , 与えられた局面における最善手を求める .

評価値を用いた場合の mini-max アルゴリズムは , 全探索可能な場合とほぼ同じである . 異なる点は以下の二つである .

- 全てのノードを展開することが出来ないので , 可能な限りのノードの展開を行う .
- 各ノードの持つ利得を順に求めるのではなく , 各ノードのもつ評価値を順に求めていく .

手順について述べる .

1. 与えられた局面をルートノードとする . ルートノードを展開する .
2. 可能な限り各ノードの展開を繰り返す . 終状態であるリーフ , 及び , 展開されずに残ったノードを以降の説明でリーフと呼ぶ .
3. 各ノードの持つ評価値を以下の手順で計算する .
 - リーフの持つ評価値は評価関数により求まる .
 - リーフ以外のノード x が持つ評価値は全ての子ノードの持つ評価値が求まってから求める . 子ノードらが持つ評価値のうち , x が max ノードのときは最大のものを , x が min ノードのときは最小のものをを選び , ノード x の持つ評価値とする . この時選ばれた子ノードが , ノード x で手番プレイヤーが選択する手を示す .
 - なお , ノード x が偶然手番の場合は , 各子ノード e の持つ評価値にノード x で子ノード e が選ばれる確率 $p_x(e)$ をそれぞれかけたものの和を , ノード x の持つ評価値とする .
4. ルートノードの持つ評価値が求まるまで 3 を繰り返す . ルートノードで選ばれた子ノードを , 与えられた局面における最善手とする .

2 での展開は , あらかじめ決められた探索時間によって制限されることが多い . ノードを展開する順番に関しては , 幅優先探索を行いながらルートノードに近いノードから徐々に深いノードへと展開を行う方法が一般的である . 各ノードの評価値をもとにした最良優先探索で動的に展開順を決めていく方法もある [5] .

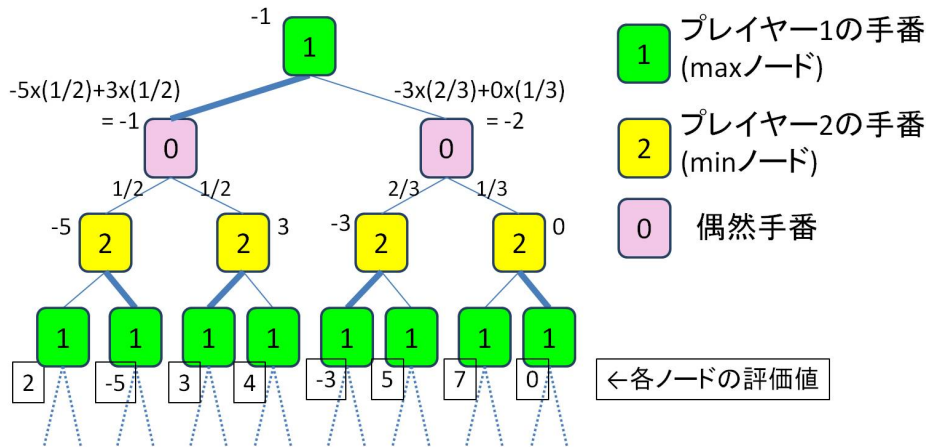


図 2.6: 全探索不可能な場合の mini-max アルゴリズム

図 2.6 に全探索が不可能な木に対して mini-max アルゴリズムを適用した例を示す．図に示したゲーム木は全てのノードが 2 つづつの子ノードを持っている．末端のノードの評価値を評価関数を用いて最初に求め、図の下にあるノードから順にそのノードでの最善手を定めることができる．太い線で示したエッジが各ノードで選ばれる最善手である．ルートノードでの最善手は左のノードを確率 1 で選ぶことであると求めることができる．

mini-max アルゴリズムの応用

展開形ゲームの探索において、最善手としてとられないことがないノードをそれ以上探索しないようにすることを「枝刈り」という．枝刈りで不必要なノードの探索にかかる時間を節約し、その分の時間で最善手となる可能性のあるノードをより深く探索することができる．

mini-max アルゴリズムの改良として、は α - β 法 [3] という効率的な枝刈りの手法が広く知られている．この手法は、評価値が 1 つの値である性質を利用するものである．

その他、mini-max アルゴリズムの改良・拡張にあたる多くの手法が提案・研究されている．

2.2.3 Max^n アルゴリズム

max^n アルゴリズム [9] とは展開形ゲームの局面での最善手を求めるアルゴリズムである．ゲーム理論の文脈では「均衡点」について議論されることが多いのに対し、コンピュータプレイヤーの探索手法に関する文脈では「均衡点」という概念に触れず、この「 max^n アルゴリズムが最善手を導く」とことを前提として議論されることが多い．

全探索可能なゲーム木と不可能なゲーム木に対する max^n アルゴリズムについて述べる．

全探索が可能な場合

全探索可能な場合の \max^n アルゴリズムは「自分以外のプレイヤーは自分の利得を最大にするように行動する」という仮定と、「自分以外のプレイヤーも、他のプレイヤーに対して自分と同じ仮定をおいている」という仮定に基づき、与えられた局面における最善手を求める。

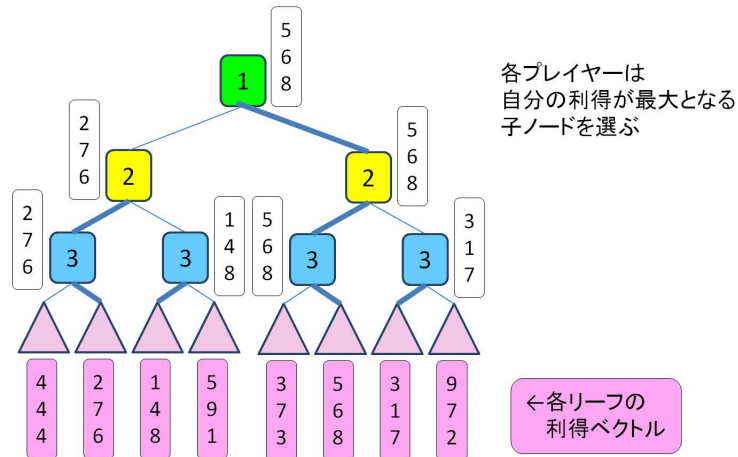
その手順について述べる。

1. 与えられた局面をルートノードとする。ルートノードを展開する。
2. 全ノードを展開するまで、各ノードの展開を繰り返す。
3. 各ノードの持つ利得ベクトルを以下の手順で計算する。
 - リーフの持つ利得ベクトルは利得関数により求まる。
 - リーフ以外のノード x が持つ利得ベクトルは、全ての子ノードの持つ利得ベクトルが求まってから求める。子ノードらが持つ利得ベクトルのうち、ノード x の手番のプレイヤーの評価値が最も大きい利得ベクトルを 1 つ選び、ノード x の持つベクトルとする。この選ばれた子ノードが、ノード x でプレイヤーが選択する手を示す。
 - なお、ノード x が偶然手番の場合は、各子ノード e の持つ利得ベクトルにノード x で子ノード e が選ばれる確率 $p_x(e)$ をそれぞれかけたものの和を、ノード x の持つ利得ベクトルとする。
4. ルートノードの持つ利得ベクトルが求まるまで、3 を繰り返す。ルートノードで選ばれた子ノードを、与えられた局面における最善手とする。

図 2.7 に全探索が可能な木に対して \max^n アルゴリズムを適用した例を示す。図に示したゲーム木は全てのノードが 2 つづつの子ノードを持っている。各リーフの利得ベクトルを最初に求め、図の下にあるノードから順にそのノードでの最善手を定めることができる。太い線で示したエッジが各ノードで選ばれる最善手である。ルートノードでの最善手は右のノードを確率 1 で選ぶことであると求めることができる。

与えられたノードにおける、全探索が可能な場合の \max^n アルゴリズムで各ノードに対して予測された最善手と、2.1.5 で述べた後ろ向き帰納法で求まる部分ゲーム完全均衡点はほとんどの場合全く同じものであり、 \max^n アルゴリズムが「正しい」ことの根拠になっている。

部分ゲーム完全均衡点と \max^n アルゴリズムが各ノードに予測する戦略が、同じにならないのは、プレイヤー i のある手番ノード x の子ノードのうち、プレイヤー i の利得が最大であるものが複数存在する場合である。プレイヤー i の利得が最大である子ノードらを y_1, \dots, y_k とする。部分ゲーム完全均衡点となる戦略の組は、後ろ向き帰納法を行っている際に、ノード x での局所戦略を y_1, \dots, y_k 上の任意の確率分布に定め、それを前提にノード x の先祖ノードでも最適応答を定めていった戦略の組全てである。一方 \max^n アルゴリズムでは、何らかの指針に従ってノード x でとられる戦略を一意に決め、それを前提にノード x の先祖ノードでも定めていく。どのような指針に従っているかは 2.2.4 節で述べる

図 2.7: 全探索可能な場合の \max^n アルゴリズム

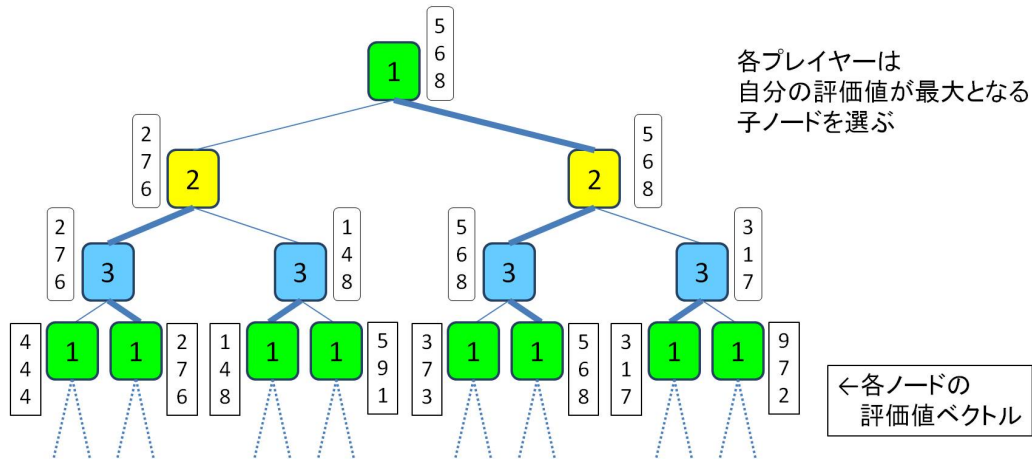
全探索が不可能な場合

全探索が不可能な場合の \max^n アルゴリズムは「自分以外のプレイヤーは自分の評価値を最大にするように行動する」「自分以外のプレイヤーも自分と同じ仮定を用いている」という仮定に基づき、与えられた局面における最善手を求める。全探索が可能な場合は各ノードの持つ利得ベクトルを順に求めるのに対し、この全探索が不可能な場合は、各ノードの持つ評価値ベクトルを順に求める。

可能な限り展開したゲーム木で、後ろ向き帰納法を用いて部分ゲーム完全均衡点を求めることに相当する。利得ベクトルの代わりに、末端となったノードから評価関数で求めた評価値ベクトルを用いる。

その手順について述べる。

1. 与えられた局面をルートノードとする。ルートノードを展開する。
2. 可能な限り各ノードの展開を繰り返す。終状態であるリーフ、及び、展開されずに残ったノードを以降の説明でリーフと呼ぶ。
3. 各ノードの持つ評価値ベクトルを以下の手順で計算する。
 - リーフの持つ評価値ベクトルは評価関数により求める。
 - リーフ以外のノード x が持つ評価値ベクトルは、全ての子ノードの持つ評価値ベクトルが求まってから求める。子ノードらが持つ評価値ベクトルのうち、ノード x の手番のプレイヤーの評価値が最も大きい評価値ベクトルを 1 つ選び、ノード x の持つベクトルとする。この時選ばれた子ノードが、ノード x で手番プレイヤーが選択する手を示す。
 - なお、ノード x が偶然手番の場合は、各子ノード e の持つ利得ベクトルにノード x で子ノード e が選ばれる確率 $p_x(e)$ をそれぞれかけたものの和を、ノード x の持つ利得ベクトルとする。

図 2.8: 全探索可能な場合の \max^n アルゴリズム

4. ルートノードの持つ評価値ベクトルが求まるまで、3を繰り返す。ルートノードで選ばれた子ノードを、与えられた局面における最善手とする。

2 展開の順番は mini-max アルゴリズムと同様、幅優先探索を行うのが一般的である。mini-max アルゴリズム同様、最良優先探索を行う方法もある。

図 2.8 に全探索が可能な木に対して \max^n アルゴリズムを適用した例を示す。図に示したゲーム木は全てのノードが 2 つつつの子ノードを持っている。末端のノードで評価値ベクトルを評価関数を用いて最初に求め、図の下にあるノードから順にそのノードでの最善手を定めることができる。太い線で示したエッジが各ノードで選ばれる最善手である。ルートノードでの最善手は右のノードを確率 1 で選ぶことであると求めることができる。

前節で述べた mini-max アルゴリズムこの \max^n アルゴリズムの特別な場合であるということも出来る。しかし、mini-max アルゴリズムは 1 つの実数である利得または評価値を用いるのに対し、 \max^n アルゴリズムが実数の n 次ベクトルである利得ベクトルまたは評価値を用いる。この 2 つの性質の差による影響は大きく、枝刈りを初めとする手法の改良の容易さに大きな差が生じている。

2.2.4 \max^n アルゴリズムの応用手法とその他の探索アルゴリズム

多人数ゲームに対する \max^n アルゴリズムの応用である手法について述べる。ほとんどの応用手法が全探索が不可能なゲーム木に対して用いられるアルゴリズムである。

\max^n アルゴリズムでも mini-max アルゴリズムと同様に、様々な到達することがないノードをそれ以上探索しないようにする枝刈りの手法が研究されている [9][22]。1 つの実数である評価値として用いる mini-max アルゴリズムでは α - β 法 [3] に代表される効率的な枝刈りの手法を適用できるが、 n 次の評価値ベクトルを用いる \max^n アルゴリズムには α - β 法を適用することができず、mini-max アルゴリズムアルゴリズムほど効率的な枝刈りを行えない。

展開形ゲームでは直前に手番を終えたプレイヤーの評価値が増加する傾向があり、多人数ゲームではその傾向が顕著である [13]。このため多人数ゲームの最良優先探索では十分な深さの探索ができない。この問題を解消するのが手番を考慮した最良優先 \max^n アルゴリズム [17] である。このアルゴリズムでは同じプレイヤーの手番のみを評価値の比較の対象とすることにより、単純な最良優先探索より深くまで探索を行うことを可能にしている。

Max^n 以外の多人数ゲームの探索手法の提案・研究もなされている。他のプレイヤーの振る舞いを max^n とは違った仮定で予想している。こちらも、ほとんどの手法が全探索が不可能なゲーム木に対して用いられるアルゴリズムである。

Paranoid アルゴリズム [22][18] は他のプレイヤーの振る舞いへの予測が max^n と異なる。ルートノードの手番プレイヤー i が、子孫ノードでの他プレイヤーの振る舞いを予測する際、 max^n アルゴリズムでは「プレイヤー j の手番ノードでは j の評価値が最大の子ノードが選ばれる」としていたのに対し、paranoid アルゴリズムでは「プレイヤー j の手番ノードでは i の評価値が最小の子ノードが選ばれる」と予測する。なお、どちらのアルゴリズムでもプレイヤー i の手番ではプレイヤー i の評価値が最大の子ノードを選ぶ。 max^n アルゴリズムは n 次の評価値ベクトルを使うのに対し、paranoid プレイヤー i の利得のみを使うことである。この性質により paranoid アルゴリズムには α - β 法をはじめとする、mini-max で用いられる手法の多くを適用することが可能で、このことが paranoid アルゴリズムの利点となっている。 Max^n アルゴリズムが「各プレイヤーはおおのこの得する手をとる」と予測するのに対し、paranoid アルゴリズムは「各プレイヤーが自分の邪魔をする」と予測していることに相当する。

Offensive アルゴリズム [27][26] は、自分はターゲットと定めたプレイヤー t の評価値を最小化する手を選び、他のプレイヤーの振る舞いの予測は max^n アルゴリズム同様「プレイヤー j の手番ノードでは j の評価値が最大の子ノードが選ばれる」とするアルゴリズムである。Offensive アルゴリズムが単体で使われることは少なく、次に述べる MP-MIX アルゴリズムの一部として用いられる。

MP-MIX アルゴリズム [27][26] は状況に応じて max^n アルゴリズム、paranoid アルゴリズム、offensive アルゴリズムを使い分ける手法である。基本的には max^n アルゴリズムを用いる。自分の順位が 2 位以下で、1 位と 2 位のプレイヤーの評価値の差が閾値 T_o 以上の時は、1 位のプレイヤーをターゲットとした offensive アルゴリズムを用いる。自分の順位が 1 位で、2 位のプレイヤーとの評価値の差が閾値 T_p 以下の時は、paranoid アルゴリズムを用いる。 T_o, T_p はあらかじめ定めておいた閾値と呼ばれるパラメーターである。実際のゲームに当てはめると、独走しているプレイヤーがいるときはそのプレイヤーの邪魔をし、自分が独走しているときは他のプレイヤーから邪魔されることを想定する、ということに相当する。

零和 2 人ゲームでは「自分の評価値を最大化する」ことが「相手の利得を最小化する」ことと同値であるので、各プレイヤーがとる「最善手」とはどういう意味をを持つのが比較的容易に定められる。一方、多人数ゲームでは「自分の評価値を最大化する」ことが必ずしも「他のプレイヤー全員の利得を最小化する」ことにつながらず、時には自分の利得を減らしてでも他のプレイヤーの利得を小さくしようとするプレイヤーがいてもおかしくない。こういったことから、多人数ゲームでは「他のプレイヤーがどのような振る舞いをするか」をモデリングすることが重要である。

多くのモデリングの研究は特定のゲームを対象としている場合が多い。 Soft-max^n アルゴリズム

では [20] では Spades というゲームを主な対象とし、既にとられた他のプレイヤーの選択を参照し、あらかじめ用意しておいた、評価値ベクトルの選択好順の異なるモデルのどれに当てはまるかを考え、モデリングを行う。同様に prob-maxⁿ アルゴリズム [21] でも Spades を主な対象とし、既にとられた他のプレイヤーの選択から他プレイヤーが各子ノードを選択する確率分布をモデリングする。その他、ポーカーを行うプレイヤーのモデリング [2][1] や 4 人チェスを行うプレイヤーのモデリング [8] の研究も行われている。

UCT アルゴリズム [19] では評価関数を使わず、あるノード x 以降各プレイヤーがランダムに手の選択を行ったときに得られる利得の平均をノード x の評価値とし、評価値が最大となる可能性の高いノードから優先的に探索を行っていく。他のプレイヤーも自分と同じ方法で探索を行っていると仮定して各ノードでの振る舞いを予測している。零和 2 人ゲームで用いられる UCT アルゴリズム [4] で用いられる、1 つの実数である評価値を n 次の実数ベクトルに変更した自然な拡張であるといえる。評価関数を作るのが難しいゲームに対して有用な手法であり、Catan というルールが比較的複雑なボードゲームのプレイヤーにこの UCT アルゴリズムを用いた研究がなされている [23]。

第3章 2値ゲーム

本論文が対象としている「2値ゲーム」について述べる。

3.1 2値ゲームとキングメーカーノード

ゲーム理論では基本的に各プレイヤーの利得を任意の「実数」であるとしている。これを実際のゲームに当てはめると、ゲームが終わった時点での「得点」や「儲け」について考えているといえる。一方、実際のゲームでは「勝敗」のみが意味を持ち、「得点」や「儲け」は意味を持たないことも多々ある。特に将棋やチェスをはじめとボードゲーム類は、そもそも「得点」の概念がないものも多く存在する。

このようなゲームでは、どの手をとっても自分が勝つ可能性はないが、自分のとる手によって勝者が決まってしまう、というような局面が存在する。「勝敗」よりも「得点」が重要な意味を持つゲームでは、自分が1位になることがなくなっても、少しでも得点を大きくなる手を選ぼうという動機を持つが、「勝敗」のみが意味を持つゲームでは自分の勝利の可能性がなくなれば、どの手を選ぶ動機もなくなってしまう。こういった局面を「キングメーカーノード」と名付け、以下に定義した。

まず「勝敗」のみが意味を持つゲームでは各プレイヤーの利得が「勝ち」を意味する「1」と「負け」を意味する「0」の2値しかとらないものとする。この利得が1,0の2値をとるゲームを「2値ゲーム」と名づける。2値ゲームでは「プレイヤー*i*が勝利する」「勝つ」という言葉を「リーフでプレイヤー*i*の利得が1となる」という意味で、「プレイヤー*i*が敗北する」「負ける」という言葉を「リーフでプレイヤー*i*の利得が0となる」という意味で用いることにする。なおこれ以降の議論で扱うゲームは基本的に、勝者が必ず1人のゲーム、言い換えると、各プレイヤーの利得の合計が1の定和ゲームとする。また、偶然手番を含まない、確定ゲーム、かつ、全プレイヤーが同じ情報を知っている完全情報ゲームとする。すなわち、以降で扱うのは「多人数・確定・完全情報・定和2値ゲーム」である。2値ゲームに対して、利得が実数であるゲームを「利得実数ゲーム」と呼ぶことにする。

2値ゲームは利得実数ゲームに比べて、利得がとりうる値の可能性ははるかに少なく、一見簡単になったようにも見える。しかし実際は、以下に述べる、複数の子ノードの利得が同じである「キングメーカーノード」の存在により、多くのノードで各プレイヤーがどのように手をとるべきかを、一意に定めることが出来ず、ゲームの解を求める問題は難しくなっている。

確定定和2値ゲームにおいて、「この手番に到達すれば、プレイヤー*i*の勝利が確定するノード」を「プレイヤー*i*の必勝ノード」と呼ぶことにする。どのプレイヤーが勝者であるかに特に注目しないときは「必勝ノード」と呼ぶことにする。定義は以下のとおりである。

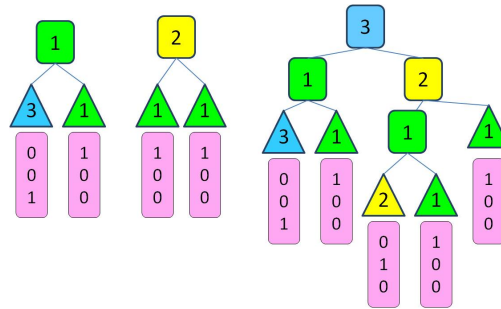


図 3.1: 必胜ノードの例

定義 3.1.1 (必胜ノード). まず「プレイヤー i の利得が 1 であるリーフ」を、プレイヤー i の必胜ノードとする。また、再帰的に、ノード x とその子ノードの集合 $A(x)$ について「 $A(x)$ が全てプレイヤー i の必胜ノードである」または「ノード x の手番がプレイヤー i かつ、 $A(x)$ にプレイヤー i の必胜ノードが含まれる」という状態のとき、ノード x をプレイヤー i の必胜ノードとする。

直感的な意味は、「自分の手番ノード x の子ノードに自分の勝利があれば、必ずその子ノードを選べば良いので、 x も自分の必胜ノード」、「他人の手番ノード x の子ノードが全て自分の必胜ノードであれば、どれを選んでも自分が勝つので、 x も自分の必胜ノード」ということである。あるいは「ノード x 以降到達する可能性のあるノードにおいて、プレイヤー i が正しい手をとれば、他のプレイヤーがどのような手をとってもプレイヤー i が勝利するノード x 」ということも出来る。

必胜ノードの例を図 3.1 に示す。図中の三角のノードは、中の数字のプレイヤーの利得が 1、その他のプレイヤーの利得が 0 であるリーフを表している。示された 3 つのゲーム木はいずれもプレイヤー 1 の必胜ノードである。

ほとんどの場面で、プレイヤー i の必胜ノードはプレイヤー i の利得が 1 であるリーフと同様に扱うことが出来る。今後は基本的に、「プレイヤー i の必胜ノード」と、「プレイヤー i の利得が 1 であるリーフ」を特に区別なく使う図中に出てくる、数字 i が書かれた三角のノードは、「プレイヤー i の必胜ノード」および「プレイヤー i の利得が 1 であるリーフ」の意味で用いる。

続いて、キングメーカーノードを定義する。

定義 3.1.2 (キングメーカーノード). プレイヤー i の手番ノード x とその子ノードの集合 $A(x)$ について、「 $A(x)$ は全て、プレイヤー i 以外のプレイヤーの必胜ノード」かつ「 $A(x)$ に 2 人以上のプレイヤーの必胜ノードが含まれる」とき、ノード x を「プレイヤー i のキングメーカーノード」という。

手番プレイヤーに特に注目しないときは「キングメーカーノード」と呼ぶことにする。また、キングメーカーノードの手番プレイヤーは「キングメーカー」と呼ぶ。図 3.2 にキングメーカーノードの例を示す。

実際のゲームにおけるキングメーカーノードは「自分は勝つ可能性がないが、自分の選択によって勝者が変わるノード」、「自分が勝利出来ないが、勝利者を決める権利を有するノード」に相当する。

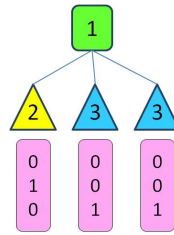


図 3.2: キングメーカーノードの例

なお、キングメーカーノードの定義からも分かるが、2 人ゲームではキングメーカーノードは発生しない。

3.2 2 値ゲームでの既存手法の挙動

キングメーカーノードを含む 2 値ゲームの部分ゲーム完全均衡点は、一意に定まらず無数に存在する。この中には、既存の均衡点の定義では「合理的である」とされているが、各プレイヤーの戦略について考察すると「合理的でない」戦略の組が多く含まれている。

後ろ向き帰納法で部分ゲーム完全均衡点を求めることを考える。キングメーカーノードでは、複数の子ノードのうちどれを選んででも手番プレイヤーの利得が変わらないため、それぞれの子ノードを任意の確率分布で選ぶ戦略が全てナッシュ均衡点となる。キングメーカーノードを子孫ノードに持つノードでは、キングメーカーノードでとられる戦略を前提に、最適応答を選ぶ。このようなノードにおいても、子孫のキングメーカーノードで選ばれた戦略によっては、複数の子ノードのうちどれを選んででも手番プレイヤーの利得が変わらず、それぞれの子ノードを任意の確率分布で選ぶ戦略が全てナッシュ均衡点となる場合が多々ある。キングメーカーノードやその先祖ノードで、最適応答が無数に存在することから、ゲーム全体に対する部分ゲーム完全均衡点は無数に存在することになる。

図 3.3 にキングメーカーノードを含んだゲームの例を示す。このゲームの部分ゲーム完全均衡点を求める。各プレイヤーが戦略を決定することは、自分の手番においてそれぞれの子ノードを選ぶ確率を決定することと等価である。それぞれのノードで左の子ノードを選ぶ確率を p_1, p_2, p_3 とおいた。従って、それぞれのノードで右の子ノードが選ばれる確率は $1 - p_1, 1 - p_2, 1 - p_3$ である。 p_1, p_2, p_3 は確率を表すので、 $0 \leq p_1, p_2, p_3 \leq 1$ の式を満たす。

プレイヤー 1 の手番ノードに注目すると、このノードはキングメーカーであり、プレイヤー 1 のどちらの子ノードをとっても利得に変化はない。従って、プレイヤー 1 はどちらの子ノードを選ぶ動機も存在せず、 p_1 を $0 \leq p_1 \leq 1$ の範囲で任意に決めたものが全て、このノードをルートとする部分ゲームのナッシュ均衡点になる。

プレイヤー 2 の手番ノードに関しても同様であり、 p_2 を $0 \leq p_1 \leq 1$ の範囲で任意に決めたものが全て、このノードをルートとする部分ゲームのナッシュ均衡点になる。

プレイヤー 3 の手番に注目すると、プレイヤー 3 の利得は左の子ノードを選べば p_1 、右の子ノード

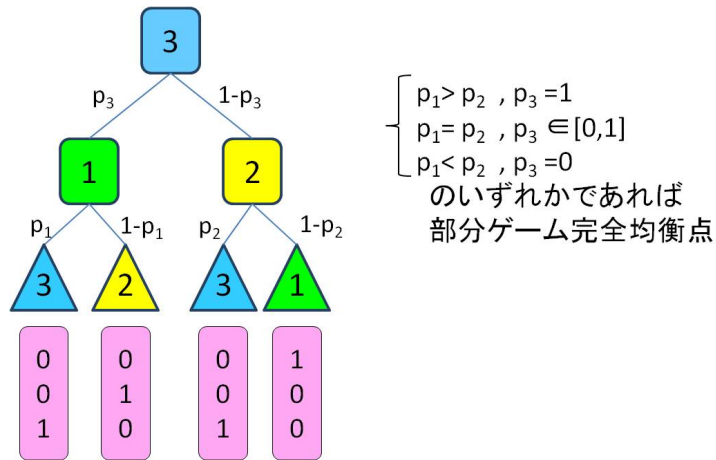


図 3.3: キングメーカーノードを含むツリーでの部分ゲーム完全均衡点

ドを選ばば p_2 である．この大きい方を確率 1 で選択することが最善手となる． p_1, p_2 が同じ値であれば， p_3 をどのような値にしても最適応答である．まとめると，プレイヤー 3 の最適応答は以下の通りである．

- p_1, p_2 に対する予測が $p_1 > p_2$ である時は，常に左の子ノードを選ぶ戦略，すなわち $p_3 = 1$ とする戦略が最適応答である．
- p_1, p_2 に対する予測が $p_1 < p_2$ である時は，常に右の子ノードを選ぶ戦略，すなわち $p_3 = 0$ とする戦略が最適応答である．
- p_1, p_2 に対する予測が $p_1 = p_2$ である時は，任意の確率で子ノードを選ぶ戦略が全て最適となる．すなわち $0 \leq p_3 \leq 1$ である任意の p_3 が全て最適応答となる．

以上よりゲーム全体での部分ゲーム完全均衡点は次の 3 つの条件のいずれかを満たす任意の p_1, p_2, p_3 の組である．

$$0 \leq p_2 < p_1 \leq 1, \quad p_3 = 1 \tag{3.1}$$

$$0 \leq p_1 = p_2 \leq 1, \quad 0 \leq p_3 \leq 1 \tag{3.2}$$

$$0 \leq p_1 < p_2 \leq 1, \quad p_3 = 0 \tag{3.3}$$

図 3.4 に式 (3.1), 式 (3.2), 式 (3.3), の条件を満たす部分ゲーム完全均衡点の例を示す．

ここまでは，キングメーカーノードを含む 2 値ゲームに対して，既存手法の均衡点は無数に導かれてしまうことを述べた．同様の問題は 2 値ゲームのみでなく，利得が実数であるゲームや，全探索が不可能なゲームで評価値を使っている場合にも起こりうる．問題が起こるのは，プレイヤー i の手番ノードの子ノードのうち，プレイヤー i の利得や評価値が最大のものが複数存在する場合である．

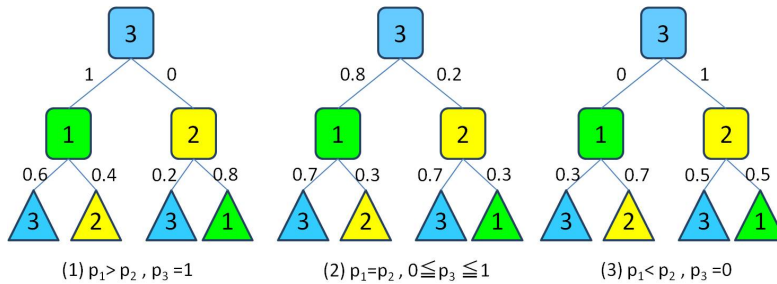


図 3.4: 部分ゲーム完全均衡点の例

しかし、利得や評価値が実数をとる場合、利得が 2 値しかとらない場合に比べて、等しくなることははるかに少なく、また等しくなったとしても、どちらを選んでも、先祖ノードでの戦略決定に与える影響がほとんどない、といった状況もしばしば起こる。このことから本論文では、既存手法では無数に均衡点が導かれてしまう状況が発生しやすい 2 値ゲームを対象としている。

また、2 値ゲームでも、プレイヤーが 2 人の場合は、キングメーカーノードが存在しないので、部分ゲーム完全均衡点を求めると、必ずルートノードでの勝者が決まる。複数の子ノードが同じ利得を持つことは多々あるが、どのような局所戦略をとってもルートノードでの勝者は同じである。

3.3 利得実数ゲームで用いられている方法

2 値ゲームについて扱っている既存手法は少ないが、研究が進んでいる評価値を用いる手法や利得が実数であるゲームで、利得が最大となる子ノードが複数存在する場合に行われている方法を紹介する。以下では、プレイヤー i のある手番ノード x の子ノードのうち、子ノード y_1, \dots, y_k で等しくプレイヤー i の評価値 (もしくは利得) が最大であるとする。

まず、ゲームプレイヤーの探索手法で用いられている例を述べる。ノード x の持つ評価値ベクトルを何らかの指針で一意に定めてしまう場合が多い。

最も単純な方法は、 y_1, \dots, y_k の中で最初に評価値ベクトルが計算された子ノードを必ず選ぶとする方法である [22]。定和ゲームで枝刈りを行う際に有効な場合がある。

次に単純な方法は、 x の持つ評価値ベクトルは y_1, \dots, y_k の持つ評価値ベクトルの平均であるとする方法である [9]。「プレイヤー i がノード x において y_1, \dots, y_k を等しい確率で選ぶ」とし、他のプレイヤーもプレイヤー i の振る舞いをそのように予測していることに相当する。

実験結果から経験則的にうまくいく手法として示されているのが、プレイヤー j がプレイヤー i の振る舞いを予測する際「 y_1, \dots, y_k のうちプレイヤー j の評価値を最小化する」とする方法である [18]。人間が実際に Hearts というゲームを行う際にもこのような予測をしているとしており、Hearts を使った実験でこの方法がうまく行くことが示されている。このように予測することは、望ましくない状況である「プレイヤー i が自由に j の評価値を変化させることが出来る局面」になることを避

けることに相当する。

その他の方法として, Soft-maxⁿ[20] では, 各ノードが評価値ベクトルを持つのではなく, 到達する可能性のあるリーフの持つ評価値ベクトルの集合を持っている。

ノード x では y_1, \dots, y_k の全てが選ばれる可能性があるとして, それぞれのノードの持つ評価値ベクトルの集合をノード x が持つとする。プレイヤー i の手番ノード x' での子ノードの選択を予測・決定する際は, 各子ノード y'_1, \dots, y'_k の持つ集合 s'_1, \dots, s'_k を比べる。ある子ノード y'_l のもつ集合 s_l に含まれるどの評価値ベクトルのプレイヤー i の評価値よりも, 別の子ノード y'_m のもつ集合 s_m に含まれるどの評価値ベクトルのプレイヤー i の評価値が小さいとき, ノード y'_l はノード y'_m に支配されているとして, プレイヤー i はノード x' で y'_l を選ぶ動機を持たない。ノード x' では他の子ノードに支配されていない子ノードが全て選ばれる可能性があるとして, それらのノードの持つ評価値ベクトルの集合の和をノード x' の持つ集合とする。

最終的にノードが持つと求められた評価値ベクトルの集合が, 順戦略によって到達しうるリーフの持つ評価値ベクトルの集合である。この集合を用いて, 他プレイヤーがどのような集合の選択好順をもつかのモデリングを行い, 各ノードでの具体的な振る舞いを予測し, 自分の手番ノードでの振る舞いを決定している。

また, ゲーム理論の分野ではノード x でプレイヤー i が y_1, \dots, y_k に対してどのような確率分布を持つ局所戦略をとっても最適応答であり, それぞれの局所戦略に対して部分ゲーム完全均衡点が定まるとしている場合が多い。

経路完全部分ゲーム完全均衡点 [24] では, このようなノード x ではどちらかの手を確率 1 でとることを信憑性のある振る舞いとして認め, 親ノードがプレイヤー i の利得が大きくなるような戦略をプレイヤー i がノード x で手を選ぶとしている。

本論文はこの経路完全部分ゲーム完全均衡点と似たアプローチをとっていると言えるが, 経路完全部分ゲーム完全均衡点が主な対象としている利得が実数のゲームより, はるかに利得が同じ子ノードが発生する可能性の高い 2 値ゲームを本論文では対象としており, 2 値ゲーム特有の問題についても解決出来るような手法を提案している。

第4章 提案手法

2値ゲームに対して，合理的なゲームの解である戦略の組を導くための手法として，各プレイヤーに自然な選択好順を持たせるための「 $kmn-\eta$ 則」と，既存の均衡点よりも強い条件で定義される「一段階予測均衡点」を提案する．また，2値ゲームの分析に有用な手法である「同値変形」を提案する．

4.1 $kmn-\eta$ 則

2値ゲームに対して，無数に導かれる部分ゲーム完全均衡点に含まれる，合理的でない「選択好順」で子ノードが選択されている戦略の組を排除するための手法である「 $kmn-\eta$ 則」を定義する．選択好順とは，プレイヤーが選択すべき子ノードの優先順位のことである．

キングメーカーノードを含むゲーム木において，後ろ向き帰納法で一段階予測均衡点を求めている際，子孫ノードでとられる戦略の予測によっては「自分の戦略をどのように選んでも利得が0になるので，どのような戦略をとっても利得は変わらない」という状況になりうる．例えば図 4.2 に示すゲーム木でプレイヤー 1 のキングメーカーノードでとられる戦略を「プレイヤー 3 の必勝ノードをとる確率は0」としたとする．この時プレイヤー 3 は自分の手番ではどのような戦略をとっても期待利得は0であるので，どのような戦略をとっても最適応答となる．しかし，プレイヤー 3 がこのように手を選ぶのは以下の観点から不自然であるといえる．プレイヤー 3 の選択できる2つの子ノードを比べると，右の子ノードを選ぶとプレイヤー 3 の期待利得は必ず0になるのに対し，左の子ノードを選ぶとプレイヤー 3 の期待利得は0より大きくなる可能性がある．したがって，プレイヤー 3 は左の子ノードを確率1で選ぶという戦略をとることが自然である．

また，図 4.1 のゲーム木についてプレイヤー 1 と 3 が左の子ノードを選ぶ確率をそれぞれ p_1, p_3 とおくと，プレイヤー 3 の期待利得 h_3 は $h_3 = p_3 \cdot p_1$ であるので， $p_3 = 1$ とする戦略によるプレイ

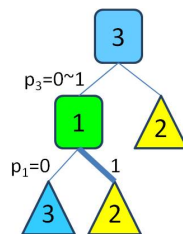


図 4.1: どの戦略を選んでも利得が0と推論してしまう例

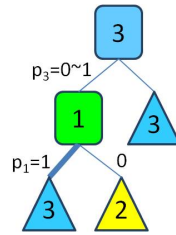


図 4.2: どの戦略を選んでも利得が 1 と推論してしまう例

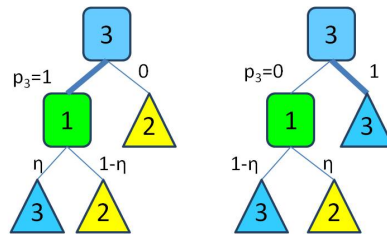
プレイヤー 3 の利得 $h_3|_{p_3=1}$ と, $0 \leq p_3 < 1$ とする戦略によるプレイヤー 3 の利得 $h_3|_{0 \leq p_3 < 1}$ との関係は $h_3|_{p_3=1} \geq h_3|_{0 \leq p_3 < 1}$ となる. 等号が成立するのは $p_1 = 0$ の場合のみである. すなわち, $p_3 = 1$ とする戦略はその他の戦略を弱支配していることが分かる. この観点からもプレイヤー 3 は左の子ノードを確率 1 で選ぶという戦略をとることが自然であるといえる.

次に図 4.2 に示した, 「自分の戦略をどのように選んでも利得が 1 になるので, どのような戦略を選んでも変わらない」という状況について考える. 後ろ向き帰納法で, プレイヤー 1 のキングメーカーノードでとられる戦略を「プレイヤー 3 の必勝ノードをとる確率は 1」としたとする. この時プレイヤー 3 は自分の手番ではどのような戦略をとっても期待利得は 1 であるので, どのような戦略をとっても最適応答となる. しかしこのプレイヤー 3 の戦略は先述の例と同様に 2 つの観点から, 合理的でない戦略である.

プレイヤー 3 の選択できる 2 つの子ノードを比べると, 右の子ノードは常に利得が 1 になるのに対し, 左の子ノードは利得が 1 より小さくなることもあるので, プレイヤー 3 は確率 1 で右の子ノードを選ぶ戦略をとるべきである. また, プレイヤー 1, 3 が左の子ノードを選ぶ確率をそれぞれ p_1, p_3 とおくと, $p_3 = 0$ とする戦略はその他の戦略を弱支配していることが分かり, プレイヤー 3 は $p_3 = 0$ である戦略をとることが自然であるといえる. $p_3 = 0$ とする戦略とその他の戦略の期待利得が同じになるのは, $p_1 = 1$ の場合のみである.

以上の 2 つの例でも見たように, 「各プレイヤーが到達することが望ましいと考えるノードの優先順位」すなわち「ノードの選択好順」は「自分の利得が必ず 1 となるノード」, 「自分の利得が 0 から 1 の間になるノード」, 「自分の利得が必ず 0 となるノード」の順であるとするのが合理的である. それぞれの性質を持つノードについて考えると, 選択好順は「自分の必勝ノード」, 「誰の必勝ノードでもなく, 子孫に自分の必勝ノードを含むノード」, 「他人の必勝ノード, および, 誰の必勝ノードでもなく, 子孫に自分の必勝ノードを含まないノード」の順であると言い換えることが出来る. この選択好順は自分の手番ノードでの選択のみでなく, 他のプレイヤーの手番ノードで選択される際にも望ましい順を示している. この選択好順に従った, 戦略の選択を各プレイヤーに行わせるために, 次に定義する $kmn-\eta$ 則を導入する.

定義 4.1.1 ($kmn-\eta$ 則). 子ノード $A(x)$ にプレイヤー i の必勝ノードが含まれるキングメーカーノード x に関して, プレイヤー i はこのキングメーカーノード x でとられる戦略を「自分の必勝ノードが選ばれる確率 p_{win-i} は, η を微小な正の数として $\eta \leq p_{win-i} \leq 1 - \eta$ である」と予測するものと

図 4.3: $kmn-\eta$ 則による自然な選択好順

する．この条件を「 $kmn-\eta$ 則」と呼ぶ．

なお「 kmn 」とは King Maker Node の頭文字をとったものである．

現実のゲームにこの $kmn-\eta$ 則が適用された場合，その意味するところは「他プレイヤーのキングメーカーノードでは，自分を勝たせる手を必ずとろうとしたり，必ずとらないようにしようとはしない」と予測することである．あるいは「他プレイヤーのキングメーカーノードでは，微小な確率で手がすべて戦略とは異なる手をとる可能性がある」と考えることに相当するとも言える．

この $kmn-\eta$ 則により先述の合理的な選択好順が導かれる．図 4.1 と図 4.2 で見た例に関して，プレイヤー 3 は図 4.3 で示すように，合理的な戦略をとるようになる．図 4.3 のそれぞれのツリーでプレイヤー 3 は，期待利得が大きくなる子ノードを確率 1 で選ぶ戦略をとることになり，いずれも合理的な選択好順に一致している．

4.2 2 値ゲームの均衡点

ナッシュ均衡点よりも強い条件によって定義される均衡点として「一段階予測均衡点」を提案する．また，ナッシュ均衡点の改良として，部分ゲーム完全均衡点が定義されたのと同様，一段階予測均衡点に対しても，その改良である「部分ゲーム完全一段階予測均衡点」を定義することを提案する．

4.2.1 一段階予測均衡点

2 値ゲームに対して，無数に導かれる部分ゲーム完全均衡点のうち，合理的でない戦略の組の例と，それらの戦略の組を含まない，より強い条件で定義される一段階予測均衡点について述べる．

3.2 節で述べたように 2 値ゲームに対して，既存手法の均衡点「部分ゲーム完全均衡点」は無数に導かれる．この「部分ゲーム完全均衡点」となっている戦略の組では，どのノードにおいても他のプレイヤーの戦略に対する最適応答である戦略がとられており，各プレイヤーは「他のプレイヤーも，そのプレイヤー以外の戦略に対する最適応答をとっている」ことを知っているため「戦略を変えても利得が増えることはないので，誰も戦略を変える動機はない」と推論している．

しかし，キングメーカーノードを含むゲーム木の部分ゲーム完全均衡点では「現在とっている戦略と異なる戦略に変えても利得が減ることもないので，戦略を変えない動機もない」という状況が

多く存在する．あるノードで戦略 b_i をとっているプレイヤー i にとって「自分の利得が変化しない範囲で戦略を b'_i に変更すると，他のプレイヤーが戦略を変更する動機を持つようになる」という状況があり得ることが分かる．この「戦略を変更する動機を持ったプレイヤー」が，実際に戦略を変更することによりプレイヤー i の利得が増加するならば，プレイヤー i は戦略を現在の b_i のままにしておくより， b'_i に変更すべきである．以上の話をまとめて，次のような概念を導入する．

部分ゲーム完全均衡点である戦略の組 $b = (b_1, \dots, b_n)$ において，プレイヤー i が部分ゲーム完全均衡点の範囲で戦略を b'_i に変えると，その変更に対して，少なくとも 1 人のプレイヤー j が戦略を b'_j に変更する動機を持ち，その変更の結果，プレイヤー i の利得が初めの状態 $h_i(b)$ より大きくなる場合「 b においてプレイヤー i は一段階予測に基づく戦略変更の動機を持つ」という．定義は以下の通りである． $B_i(b_{-i})$ は先述のとおり， $n-1$ 人の戦略の組 b_{-i} に対するプレイヤー i の最適応答の集合である． b_{-ij} は戦略の組 b から第 i 番目の要素と第 j 番目の要素を除いたものである．また b' は戦略の組 $b' = (b'_1, \dots, b'_n)$ を意味する．

定義 4.2.1 (一段階予測に基づく戦略変更の動機)．部分ゲーム完全均衡点である行動戦略の組 $b = (b_1, \dots, b_n)$ に関して，以下の 2 つの条件を満たす，プレイヤー i の行動戦略 b'_i が存在するとき「 b においてプレイヤー i は一段階予測に基づく戦略変更の動機を持つ」という．

- b'_i は $B_i(b_{-i})$ の要素である．すなわち $b'_i \in B_i(b_{-i})$ が成立する．
- $j = 1, \dots, i-1, i+1, \dots, n$ に対して b'_j を次のように定めたとき，式 $h_i(b') > h_i(b)$ が成立する．
 - $b_j \in B_j(b'_i, b_{-ij})$ であれば b'_j は $b'_j = b_j$
 - $b_j \notin B_j(b'_i, b_{-ij})$ であれば b'_j は $B_j(b'_i, b_{-ij})$ の任意の要素

1 つ目の条件は，プレイヤー i が戦略の変更をするのは最適応答の範囲である，すなわち，プレイヤー i が戦略を変更し，他のプレイヤーがもし戦略を変更しなかったとしても，戦略の組が部分ゲーム完全均衡点となっていることを示す．

2 つ目の条件は，プレイヤー i の戦略変更に対し，戦略を変更をする動機があるプレイヤーは最適応答である戦略の 1 つへと変更し，戦略を変更をする動機がないプレイヤーは戦略を変更しないとき，プレイヤー i の利得が増加することを示す．

例として，図 4.4 にて一段階予測に基づく戦略変更の動機概念を説明する．左のゲーム木で各プレイヤーがとっている戦略 $p_1 = 0.6, p_2 = 0.2, p_3 = 1$ は部分ゲーム完全均衡点となっている．ここでプレイヤー 1 が戦略を $p_1 = 0.6$ から $p_1 = 0.1$ へと変更することを考える．まず，プレイヤー 1 はどのような戦略をとっても自分の利得 $h_1 = 0$ が変化することはないので，プレイヤー 1 のこの戦略 $p_1 = 0.1$ は他のプレイヤーの戦略 $p_2 = 0.2, p_3 = 1$ に対する最適応答である．このプレイヤー 1 の戦略変更に対し，プレイヤー 2 の現在とっている戦略 $p_2 = 0.2$ は $p_1 = 0.1, p_3 = 1$ に対する最適応答の 1 つなので，戦略を変更する動機を持たず，そのままの戦略をとる．プレイヤー 3 は現在とっている戦略 $p_3 = 1$ は $p_1 = 0.6, p_2 = 0.2$ に対する最適応答でなくなり， $p_3 = 0$ が最適応答になるので，動機を持って $p_3 = 0$ に戦略を変更する．この結果，各プレイヤーの戦略は $p_1 = 0.1, p_2 = 0.2, p_3 = 0$ となり，

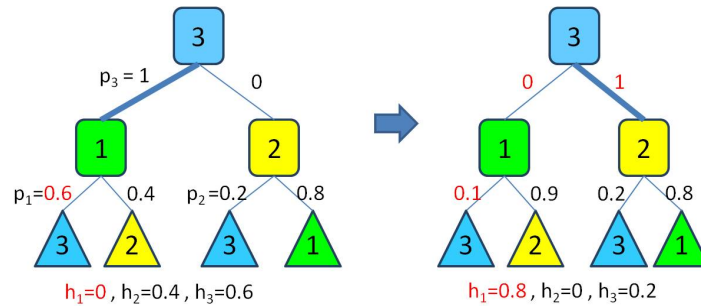


図 4.4: 一段階予測に基づく戦略変更の動機

再び部分ゲーム完全均衡点となる．プレイヤー 1 の期待利得 h_1 の変化を見ると，戦略変更後のほうが増加していることが分かる．以上の議論より，プレイヤー 1 は初めの状態 $p_1 = 0.6, p_2 = 0.2, p_3 = 1$ において，戦略を $p_1 = 0.1$ に変更する動機を持つので，一段階予測に基づく戦略変更の動機を持つことが分かった．

図 3.4 のゲーム木の (1) $p_1 > p_2, p_3 = 1$ である部分ゲーム完全均衡点では， $p_2 \neq 0$ の場合はプレイヤー 1 のみが $0 \leq p_1 < p_2$ とする一段階予測に基づく戦略変更の動機を持つ． $p_2 = 0$ の場合はどのプレイヤーも一段階予測に基づく戦略変更の動機を持たない．ただし， $k_{mn}-\eta$ 則を適用すると， $p_2 = 0$ の場合でもプレイヤー 1 のみが $0 \leq p_1 < \eta$ とする一段階予測に基づく戦略変更の動機を持つ．

図 3.4 のゲーム木の (3) $p_1 < p_2, p_3 = 0$ である部分ゲーム完全均衡点では， $p_1 \neq 0$ の場合はプレイヤー 2 のみが $0 \leq p_2 < p_1$ とする一段階予測に基づく戦略変更の動機を持つ． $p_1 = 0$ の場合はどのプレイヤーも一段階予測に基づく戦略変更の動機を持たない．ただし， $k_{mn}-\eta$ 則を適用すると， $p_1 = 0$ の場合でもプレイヤー 2 のみが $0 \leq p_2 < \eta$ とする一段階予測に基づく戦略変更の動機を持つ．

図 3.4 のゲーム木の (2) $p_1 = p_2, 0 \leq p_3 \leq 1$ である部分ゲーム完全均衡点では， $p_1 = p_2 \neq 0$ の場合はプレイヤー 1 が $0 \leq p_1 < p_2$ ，プレイヤー 2 が $0 \leq p_2 < p_1$ ，とする一段階予測に基づく戦略変更の動機を持つ． $p_1 = p_2 = 0$ の場合はどのプレイヤーも一段階予測に基づく戦略変更の動機を持たない．

各プレイヤーが一段階予測に基づく戦略変更の動機に従う場合，戦略の読み合いが止まるのは，どのプレイヤーも一段階予測に基づく戦略変更の動機を持たないような戦略の組のみである．このような戦略の組を「一段階予測均衡点」と定義する．

定義 4.2.2 (一段階予測均衡点). 部分ゲーム完全均衡点である行動戦略の組 $b = (b_1, \dots, b_n)$ において，どのプレイヤーも一段階予測に基づく戦略変更の動機を持たないとき， b は一段階予測均衡点であるという．

比較としてナッシュ均衡点の定義の 1 つを改めて述べておく．

定義 4.2.3 (ナッシュ均衡点). 部分ゲーム完全均衡点である行動戦略の組 $b = (b_1, \dots, b_n)$ において，どのプレイヤーも (最適応答に基づく) 戦略変更の動機を持たないとき， b はナッシュ均衡点であるという．

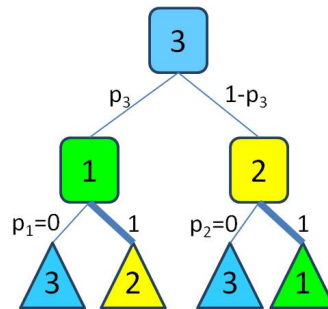


図 4.5: 一段階予測均衡点

一段階予測均衡点とナッシュ均衡点の違いは戦略変更の動機として「最適応答に基づく戦略変更の動機」のみについて考えているか、それに加えて「一段階予測に基づく戦略変更の動機」について考えているかどうかである。

図 4.5 に示したゲーム木について、 $kmn-\eta$ 則が成り立つとすると、今までに見てきた一段階予測に基づく戦略変更の動機に関する議論より、一段階予測均衡点は $p_1 = p_2 = 0, 0 \leq p_3 \leq 1$ のみであることが分かる。

4.2.2 部分ゲーム完全一段階予測均衡点

一段階予測均衡点である戦略の組でも、合理的なゲームの解となっていない場合がある。その例と、解決するためのより強い条件で定義される部分ゲーム完全一段階予測均衡点について述べる。

図 4.6 に示したゲーム木 (1) とゲーム木 (2) は構造が同じであり、それぞれ異なった一段階予測均衡点である戦略の組を表している。太線のエッジは親ノードにおいてその子ノードが確率 1 で選ばれることを示しており、どの子ノードにつながるエッジも太線になっていないノードではどのような戦略をとっても期待利得が変化しないことを意味している。2つのゲーム木の違いはプレイヤー 2 の手番での戦略の違いである。それに伴い、プレイヤー 1 の左側の手番での戦略に違いが生じている。

ゲーム木 (1) ではプレイヤー 3 がキングメーカーになり、ゲーム木 (2) ではプレイヤー 1 がキングメーカーになる。しかし、ゲーム木 (2) に示した一段階予測均衡点にはルートノードのプレイヤー 3 の手番から見ると問題がある。ゲーム木 (2) で、仮にプレイヤー 3 がゲームのプレイで左の子ノードを選択し、左のプレイヤー 1 の手番ノードに到達したとする。このとき、このノード以降の部分ゲームでは、一段階予測均衡点の観点から、プレイヤー 2 は先ほどまでの「プレイヤー 1 の必勝ノードを選ぶ」から「プレイヤー 3 の必勝ノードを選ぶ」へと戦略変更を行うことになる。従って、プレイヤー 2 が最初にとっていた「プレイヤー 1 の必勝ノードを選ぶ」という戦略は「信憑性のない脅し」であることになる。よってプレイヤー 2 はゲーム木 (1) で示した戦略をとることが合理的であることが分かるので、プレイヤー 3 はルートノードで左の子ノードを選択することを避ける動機を持たない。ルートノードで右の子ノードを選ぶと自分の負けが確定してしまうので、当然左の子ノードを選ぶべきである。

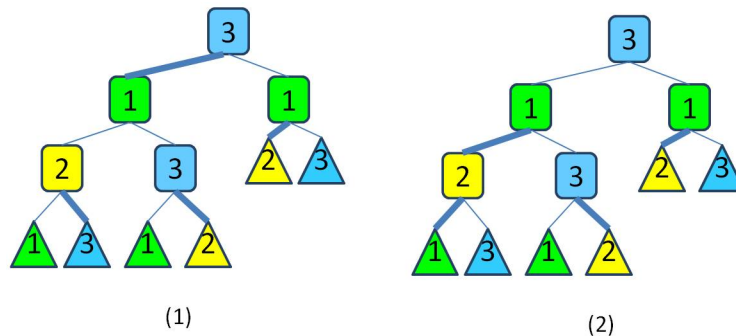


図 4.6: 2 つの一段階予測均衡点

一方、ゲーム木 (1) に示した一段階予測均衡点ではどのプレイヤーの戦略も信憑性のない脅しにはなっていない。また、このゲーム木のどの部分ゲームを見ても、各ノードでとられている戦略は一段階予測均衡点になっていることが分かる。このような戦略の組を「部分ゲーム完全一段階予測均衡点」として定義する。

定義 4.2.4 (部分ゲーム完全一段階予測均衡点). 戦略の組 $b = (b_1, \dots, b_n)$ が、展開形ゲーム Γ の任意の部分ゲームに対して一段階予測均衡点を導くとき、 b を Γ の「部分ゲーム完全一段階予測均衡点」とであるという。

逆に、部分ゲーム完全一段階予測均衡点でない一段階予測均衡点は、必ず信憑性のない脅しである戦略を含んでしまい、ゲームの解として合理的でないものになってしまう。具体的には、部分ゲームで一段階予測均衡点となっていない戦略の組が全て、信憑性のない脅しになっている。

比較のため部分ゲーム完全均衡点の定義を再び述べておく。

定義 4.2.5 (部分ゲーム完全均衡点). 戦略の組 $b = (b_1, \dots, b_n)$ が、展開形ゲーム Γ の任意の部分ゲームに対してナッシュ均衡点を導くとき、 b を Γ の「部分ゲーム完全均衡点」とであるという。

ナッシュ均衡点の問題を取り除いたのが部分ゲーム完全均衡点であったように、一段階予測均衡点の問題を取り除いたのが部分ゲーム完全一段階予測均衡点である。

以上の議論より、多人数 2 値ゲームの均衡点となるのは部分ゲーム完全一段階予測均衡点のみであることが分かる。

確定 2 値ゲームの部分ゲーム完全一段階予測均衡点の具体的な求め方を以下に述べる。各リーフが持つ利得ベクトルは、全ていずれかのプレイヤーの勝利ノードとなっている。すなわち、いずれかのプレイヤーの利得のみが 1 であり、他のプレイヤーの利得は必ず 0 である。注目するのは、各ノードで「どのプレイヤーの必勝ノードを選ぶと」部分ゲーム完全一段階予測均衡点になるかということである。

ゲーム木の全体は共有知識として与えられている。リーフの持つ利得ベクトルは利得関数により求まる。それぞれいずれかのプレイヤーの必勝ノードになっている。

まず、内部ノードのそれぞれでとられる戦略を求めていく。全ての子ノードでとられる戦略が求まったものから順に求める。プレイヤー i の手番ノード x でとられる戦略と、 x がどのプレイヤーの必勝かは以下の手順で求める

1. x の子ノード $A(x)$ にプレイヤー i の必勝ノード y が含まれる場合、その子ノード y を確率 1 で選ぶ戦略が最適応答である。 $A(x)$ にプレイヤー i の必勝ノードが複数含まれる場合、それらのノードを任意の確率で選ぶ戦略は全て最適応答である。どちらの場合も、 x はプレイヤー i の必勝ノードとなり、このノード x に到達した場合に勝利する可能性のあるプレイヤーはプレイヤー i のみである。
2. 1 に当てはまらず、 x の子ノード $A(x)$ にプレイヤー i が勝利する可能性のあるノード y が含まれる場合、その子ノード y を確率 1 で選ぶ戦略が最適応答である。 $A(x)$ にプレイヤー i が勝利する可能性のあるノードが複数含まれる場合、それらを y_1, y_2, \dots として、それらのノードを任意の確率で選ぶ戦略は全て最適応答である。このノード x に到達した場合に勝利する可能性のあるプレイヤーはプレイヤー i と、 y または y_1, y_2, \dots で勝利する可能性のあるプレイヤー全員である。
3. 1, 2 のどちらにも当てはまらず、 x の子ノード全てにおいて勝利する可能性のあるプレイヤーが一致する場合、各ノードを任意の確率で選ぶ戦略は全て最適応答である。このノード x に到達した場合に勝利する可能性のあるプレイヤーは子ノードに到達したときに勝利する可能性のあるノードと同じである。
4. 1, 2, 3 のいずれにも当てはまらないのは、 x の子ノード $A(x)$ にプレイヤー i が勝利する可能性のあるノードが 1 つも含まれず、かつ、子ノードの選択によって勝利する可能性のあるプレイヤーが変化する場合である。すなわち、ノード x がキングメーカーノードである場合である。この場合、ノード x での最適応答は、ゲーム木を親ノード、祖父ノード、... と先祖ノードを順に辿って決定することになる。

ノード x に到達した場合に勝利する可能性のあるプレイヤーは子ノードのそれぞれで勝利する可能性のあるプレイヤー全員であると一時的に定めておく。親ノード、祖父ノード、... と、各ノードのそれぞれでの最適応答が順に定まっていく先祖ノードの中に、次の条件を満たすプレイヤー j の手番ノード z が存在したとき、ノード x での最適応答と、勝利する可能性のあるプレイヤーを変更する。「ノード z でプレイヤー i が勝利する可能性がある」かつ「ノード x でプレイヤー j が勝利する可能性がある」かつ「ノード x で選ぶ子ノードの変更によっては、ノード x でプレイヤー j が勝利する可能性をなくすことが出来る」がノード z の条件である。

この条件を満たす z が存在した場合、ノード x での戦略を「ノード x でプレイヤー j が勝利する可能性をなくすことが出来る」ものに変更する。ノード x での戦略の変更は、 x の先祖かつ z の子孫であるノードでの戦略の変更に影響は与えないが、各ノードでの勝利する可能性のあるプレイヤーには影響を与える。条件を満たす子ノードが複数ある場合は、さらに先祖のノードで条件を満たすノードがあるかを調べる必要がある。条件を満たすノードがある度に、ノ

ド x で選ぶことの出来る子ノードが限定されていく。ルートノードでの戦略が定まっても、ノード x で選ぶことの出来る子ノードが複数ある場合は、それらを任意の確率分布で選ぶ戦略全てがノード x での最適応答となる。

ルートノードでの最適応答が求まるまで、上の手順を繰り返す。各ノードで最適応答となっている戦略の組全てが、部分ゲーム完全一段階予測均衡点となる。この手順の計算量は、最悪の場合でも (ノード数) $\cdot (n - 1)$ となる。 n はプレイヤーの人数である。

なお、(4) で戦略を変更する必要があるプレイヤー i の手番ノード x に関して、戦略を変更する前の一時的な戦略では、プレイヤー i が一段階予測による戦略変更の動機を持っている。プレイヤー i が一段階予測による戦略変更の動機を持たなくなる戦略が、変更した後の戦略である。

4.3 同値変形

定和 2 値ゲームのゲーム木を「本質」を変えない範囲でできるだけ単純化する手法である「同値変形」について述べる。

定和 2 値ゲームでは 1 人のプレイヤーのみがゲームで勝利することができる。そこで「どのプレイヤーの勝利ノードに到達する可能性があるか」と「それを決めることができる、キングメーカーであるプレイヤーは誰か」がゲーム木の「本質」であると考え、ゲーム木に含まれるノードの数を減らし単純化することができる。これらの単純化を「同値変形」と呼ぶことにする。同値変形の目的はゲーム木を、その本質を表す最低限の大きさのゲーム木へと単純化することで、その本質を容易に理解出来るようにすることと、ゲーム木の解析を容易にすることである。

ゲーム木に対し、その本質を変えず、以下の 4 つの同値変形を行うことが出来る。

1. あるノード x が子ノードを 1 つしか持たない場合、 x の親ノード z から x につながるエッジを、 x の子ノード y へと直接つなぎ替え、子ノードと x を削除する。図 4.7 の (1) に例を示す。
2. あるノード x と、その子ノード y の手番プレイヤーが同じ場合、 x から y の子ノードらへ直接エッジをつなぎ、ノード y を削除する。「ノード x とノード y の子ノードのどれに到達するかは手番プレイヤーが選べる」というゲーム木の本質を残したまま、ゲーム木を単純化している。図 4.7 の (2) に例を示す。
3. あるノードの子ノードに同じプレイヤーの必勝ノードが複数ある場合、それらの子ノードは本質的には同じものであるため、1 つを残して削除する。図 4.7 の (3) に例を示す。
4. あるノードの子ノードに同じプレイヤーの手番、かつ、全く同じプレイヤーらが勝利する可能性があるキングメーカーノードが複数ある場合、それらの子ノードは本質的には同じものであるため、1 つを残して削除する。図 4.7 の (4) に例を示す。

それぞれの変形を順不同で何度繰り返しても、ゲーム木の本質は失われない。

また、各プレイヤーが $k_{mn-\eta}$ 則や部分ゲーム完全一段階予測均衡点などに従い、一部のノードで戦略が一意に定まる場合、上記の 4 つの同値変形に加えて次の同値変形を行うことが出来る。それ

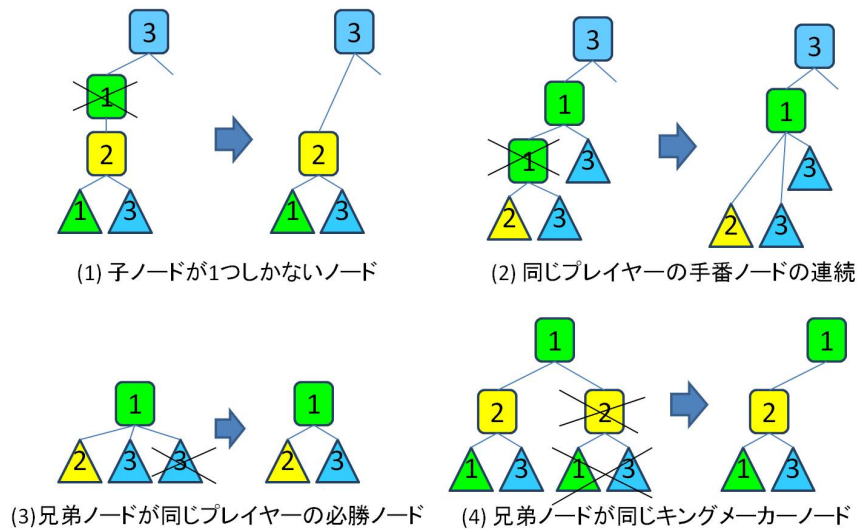


図 4.7: 同値変形

それぞれのノードに対して一意に定めた戦略がとられる場合のゲーム木の本質を変えずに、ゲーム木を単純化することに相当する。

1. あるノード x においてとられる戦略が、子ノード y を確率 0 で選ぶものである場合、 y とその子孫ノード全てを削除する。「ノード y とその子孫ノードへゲームのプレイで到達することはない」というゲーム木の本質を表現したまま、ゲーム木を単純化している。図 4.8 の (1) に例を示す。
2. あるノード x においてとられる戦略が、子ノード y を確率 1 で選ぶものである場合、 x の親ノード z から x につながるエッジを y へと直接つなぎ替え、 x 、 x の y 以外の子ノード、その子孫ノードを削除する。「プレイでノード x に到達した場合、必ずノード y へ到達する」というゲーム木の本質を表現したまま、ゲーム木を単純化している。図 4.8 の (2) に例を示す。

上で述べた 4 つの同値変形と、この 2 つの同値変形のそれぞれを順不同で何度繰り返しても、ゲーム木の本質は失われない。

一段階予測均衡点で一部のノードに戦略が一意に定められたゲーム木に同値変形を行った例を図 4.9 示す。

4.4 ゲームへの適用

4.4.1 3人2値ゲーム

全探索が可能な 3 人 2 値定和ゲームについて考える。

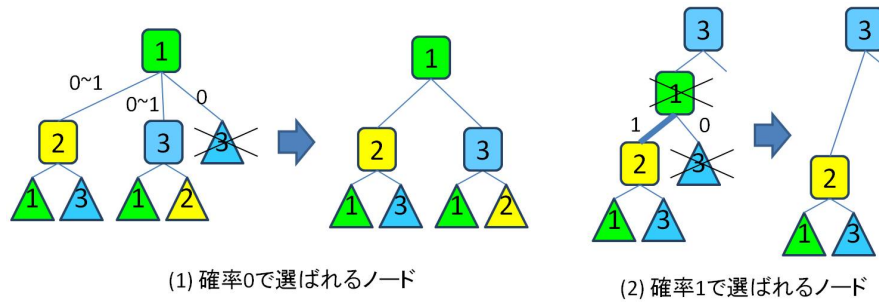


図 4.8: 戦略が定まるノードでの同値変形

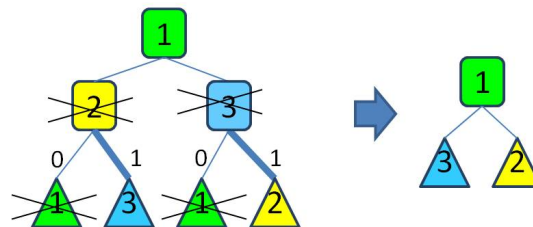


図 4.9: 一段階予測均衡点の同値変形の例

部分ゲーム完全均衡点の範囲で考えると、キングメーカーノードでは各子ノードを任意の確率分布で選択する局所戦略が全て最適応答となる。この決められた戦略に対して、各プレイヤーは手番ノードにおいて最適応答となる戦略をとることになる。ゲーム木全体では無数の戦略の組が部分ゲーム完全均衡点となる。

各プレイヤーが $k_{mn-\eta}$ 則とし、部分ゲーム完全一段階予測均衡点である戦略の組を求める。

ルートノードがいずれかのプレイヤー i の必勝ノードである場合は、プレイヤー i の最適応答は自分の手番で必ず「自分の必勝ノードである子ノードを選ぶ戦略」を選択することであるので、必ずプレイヤー i が勝利する。キングメーカーノードが存在していたとしても、プレイでキングメーカーノードに到達することはない。従って、キングメーカーノード上の戦略をどのように決定してもプレイには関係しない。

一方、ルートノードが必勝ノードでない場合は必ずプレイで一度はキングメーカーノードに到達することになる。3人2値ゲームにおいて、各プレイヤーが $k_{mn-\eta}$ 則に従うとし、部分ゲーム完全一段階予測均衡点を求めると、求まった均衡点の全てで3人のうち2人のプレイヤーのキングメーカーノードでは必ず戦略が一意に定まる。この時定まる戦略はどちらかの子ノードを確率1で選択するものである。戦略が一意に定まらないキングメーカーノードは全て同じプレイヤーの手番ノードとなる。

3人2値ゲームで、各プレイヤーが $k_{mn-\eta}$ 則に従い、部分ゲーム完全一段階予測均衡点である戦略の組がとられるゲーム木を考え、同値変形を適用すると、図 4.10 に示すように1つのキングメー

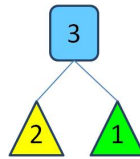


図 4.10: 3 人 2 値ゲームの同値変形後

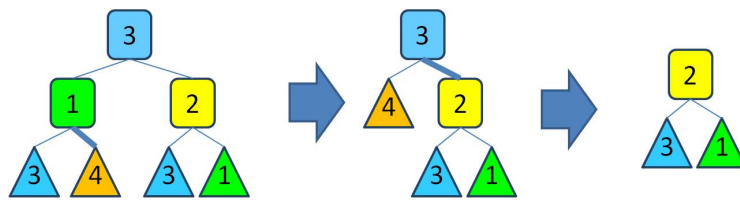


図 4.11: 4 人 2 値ゲームの同値変形例 1

カーノードと 2 つのリーフからなるゲーム木へと単純化される．このキングメーカーノードでとられる戦略を決定することは出来ないが，元のゲーム木で勝ち目のあるプレイヤーが 2 人しかおらず，それが誰なのかを求めることが出来る．元のゲーム木でも，キングメーカーが 1 人のプレイヤーであったことが分かる．

なお，3 人 2 値ゲームの場合これらの同値変形の結果は，1 人だけ戦略が一意に定まらず，キングメーカーノードが残ったプレイヤー i が， i 以外の 2 人のうち片方のプレイヤー j を勝たせる戦略を選ぶ確率が全てのキングメーカーで等しいと仮定したときの，ゲームのプレイの様子を表しているともいえる．

4.4.2 4 人以上 2 値ゲーム

4 人以上で行う 2 値ゲームに対して提案手法の適用を行った場合，3 人 2 値ゲームとは違い，複数のキングメーカーが残る場合もあり，同値変形の結果は様々である．いくつかの例を示す．

図 4.11 は，3 人 2 値ゲーム同様，キングメーカーが 1 人だけ残った例である．図 4.12 は，2 人のキングメーカーノードが残り，プレイヤー 1 とプレイヤー 2 の戦略によって，プレイヤー 3 の戦略が決定する場合を示している．図 4.12 では，プレイヤー 3 の左の子ノード以降は，プレイヤー 1 とプレイヤー 2 の手番によって，1 つのキングメーカーノードが構成されているような状態である．この場合も，プレイヤー 1 とプレイヤー 2 の戦略に対する最適応答をプレイヤー 3 はとることになる．

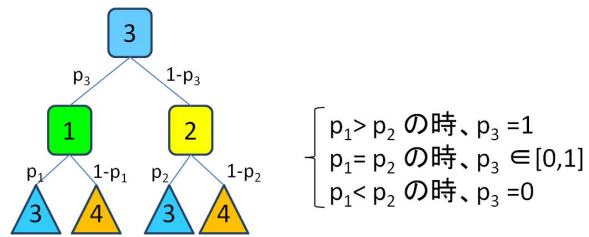


図 4.12: 4人2値ゲームの同値変形例 2

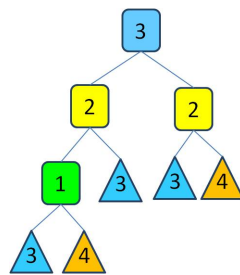


図 4.13: 4人2値ゲームの同値変形例 3

第5章 実験と評価

提案手法を現実のゲームであるダイヤモンドゲームへと適用し、その効果を評価した。

5.1 ダイヤモンドゲーム

現実に存在する3人2値定和ゲームの例として3人ダイヤモンドゲームを用いて実験を行った。ダイヤモンドゲームは英語では Chinese Checker と呼ばれ、各プレイヤーが順番に自分のコマを動かして、最も早く自分のコマを全てゴールに到達させたプレイヤーが勝ち、というゲームである。通常は3人のプレイヤーで行われる。完全情報・確定の性質を持つので、多人数ゲームの例としてしばしば研究の対象にされている。

ダイヤモンドゲームには盤面の大きさが様々なものが存在し、大きさによって各プレイヤーのコマの数が変わる。今回は全探索が可能なサイズで実験を行うため、図 5.1 に示した最小の大きさのダイヤモンドゲームを用いた。各プレイヤーは1つずつのコマを動かして、ゴールに最も早く到達させることを目指す。

詳細なルールを述べる。参加するプレイヤーはプレイヤー a 、プレイヤー b 、プレイヤー c の3人で、手番のこの順番に行う。図 5.1 の図中の丸は各コマが移動するマスを示している。初期局免では、各プレイヤーのコマは赤いスタートマスに置かれている。各プレイヤーは自分の手番になると、自分のコマを左右または斜めに隣合っているマスに移動させる。この際、隣合ったマスに他のプレイヤーのコマがいる場合は、そのコマを越えた先にマスを移動させることができる。コマを越えたマスにさらに他のコマがいるときや、盤面の外に出てしまうときには飛び越すことが出来ない。順番



図 5.1: ダイヤモンドゲーム

表 5.1: 戦略が一意に定まらないノードの数

	全ノード数	戦略が一意に定まらないノード数		うち、戦略の違いにより利得に変化があるノード数	
		割合	割合	割合	割合
既存手法	84	84	100%	84	100%
kmn- η	84	59	70%	49	70%
kmn- η , 均衡点	84	46	55%	6	7%

5.2 各種実験と結果

提案手法である「kmn- η 則」「部分ゲーム完全一段階予測均衡点」「同値変形」による効果を調べるため、ダイヤモンドゲームを用いて次の実験を行った。

1 つ目の実験は、各種手法を用いることで戦略が一意に定まらないノードの数をどれだけ減らせるかを調べるものである。2 つ目の実験は、同値変形や、他の手法と同値変形の組み合わせによってどれだけゲーム木を単純化できるかを調べるものである。3 つ目の実験は、部分ゲーム完全一段階予測均衡点に従った戦略をとるプレイヤーと既存手法を用いるプレイヤーをダイヤモンドゲームで戦わせ、得られる利得の違いを調べるものである。

5.2.1 実験 1: 戦略が一意に定まるノードの数

各種手法をダイヤモンドゲームのゲーム木に用いた場合、どれだけノードで戦略を一意に定めることが出来るかを調べた。

戦略が一意に定まらないノードでは、利得が最大である子ノードらを任意の確率分布で選ぶ戦略が全て最適応答となっており、合理的でない戦略がとられている場合も多い。従って、戦略が一意に定まらないノードが少ないことは、手法をコンピュータープレイヤーへ応用したとき、合理的な戦略の組がとられる可能性がより高くなることを表す指標となる。

それぞれのノードの数を調べたのは以下の 3 つの場合である。

- 既存手法である部分ゲーム完全均衡点である戦略をとる場合
- 各プレイヤーが kmn- η 則に従って戦略をとる場合
- 各プレイヤーが kmn- η 則に従い、かつ、部分ゲーム完全一段階予測均衡点である戦略をとる場合

それぞれの場合を「既存手法」、「kmn- η 」、「kmn- η , 均衡点」と表記する。

表 5.1 にその結果を示す。戦略が一意に定まらないノードと、その全ノードに対する割合を示す。また、戦略が一意に定まらないノードのうち、戦略の変化により利得が変わってしまうノードの数と、その全ノードに対する割合を示した。

表 5.2: 同値変形によるノードの数の削減

	ノード数	リーフ数	キングメーカーノードの数
元のゲーム木	84	112	26
同値変形	72	85	19
$k_{mn-\eta}$, 同値変形	9	10	5
$k_{mn-\eta}$, 均衡点, 同値変形	1	2	1

5.2.2 実験 2: ゲーム木の同値変形

ダイヤモンドゲームの元のゲーム木や、各種手法によりいくつかのノードにおける戦略を一意に決定した後のゲーム木に対して、同値変形を適用した場合にゲーム木がどれだけ単純化されるかについて調べた。

実験 1 同様、同値変形によりノードが少なくなる手法ほど、手法をコンピュータープレイヤーへ応用したとき、合理的な戦略の組がとられる可能性がより高くなることを表す指標となる。

調べたのは次の 4 つの場合である。

- 元のゲーム木
- 元のゲーム木に対し同値変形を行う場合
- 各プレイヤーが $k_{mn-\eta}$ 則に従うとした場合に、到達される可能性のあるノードからなるゲーム木に同値変形を行う場合
- 各プレイヤーが $k_{mn-\eta}$ 則に従い、部分ゲーム完全一段階予測均衡点である戦略をとる際に、到達される可能性のあるノードからなるゲーム木に同値変形を行う場合

それぞれの場合を「元のゲーム木」、「同値変形」、「 $k_{mn-\eta}$, 同値変形」、「 $k_{mn-\eta}$, 均衡点, 同値変形」と表記する。

表 5.2 にその結果を示す。

また、「 $k_{mn-\eta}$, 同値変形」、「 $k_{mn-\eta}$, 均衡点, 同値変形」のそれぞれで、同値変形によって単純化されたゲーム木を図 5.1 と図 5.1 に示す。

5.2.3 実験 3: 既存手法と提案手法の対戦

4 つの戦略をとるプレイヤーでダイヤモンドゲームで対戦を行い、得られる利得を調べた。

4 つの戦略は以下のとおりである。

- RNDall
 全ての子ノードを等しい確率で選ぶ戦略。
 \max^n アルゴリズムでは、どのノードに対しても根拠を持った戦略を一意に定義できず、全ての手が戦略の一つとなりうるので、全ての手を等しい確率で選ぶとしたもの。

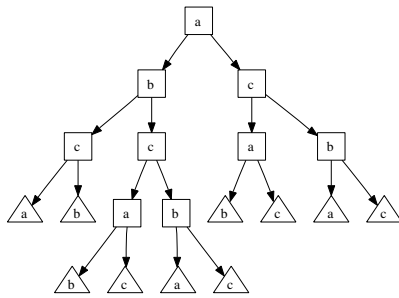
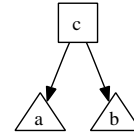
図 5.3: $kmn-\eta$ 則に従うゲーム木の同値変形

図 5.4: 部分ゲーム完全一段階予測均衡点をとるゲーム木の同値変形

- RNDeta

$kmn-\eta$ 則によって導かれる選択好順に従い、勝つ可能性があるノードを等しい確率で選ぶ戦略。自分の勝利ノードを子孫に持つ子ノードと持たない子ノードが存在する場合、持つ子ノードを選ぶ戦略は持たない子ノードを選ぶ戦略を弱支配するので、持つ子ノードを選んだ方が良く、という方針に従う。自分の勝利ノードを子孫に持つかに関して、複数の子ノードの条件が同じ場合は、それらを等しい確率で選ぶ。

- eqprob

「他のプレイヤーも自分と同じ戦略をとる」と仮定して他のプレイヤーの振る舞いを予測し、期待利得が最大であると予測されるノードを等しい確率で選ぶ戦略。期待利得が最大であるノードが 1 つだけの時はそれを確率 1 で選ぶ。他のプレイヤーの振る舞いを指針を決めて予測し、それに対する最適応答をとる戦略である。

- 1EP

$kmn-\eta$ 則に従った予測を行い、全プレイヤーが部分ゲーム完全一段階予測均衡点に従うと予測し、自分もそれに従う戦略。期待利得が最大の子ノードが 1 つしかないときはその子ノードを確率 1 で選び、部分ゲーム完全一段階予測均衡点に従う子ノードが複数ある場合はそれらを等しい確率で選ぶ。部分ゲーム完全一段階予測均衡点でも戦略が一意に定まらない場合は、各プレイヤーの勝利ノードを等しい確率で選ぶ。

対戦は「戦略 1EP を使うプレイヤーが 1 人と、他の戦略を使うプレイヤーが 2 人」「1EP を使うプレイヤーが 2 人と他の戦略を使うプレイヤーが 1 人」で行い、プレイヤーの各手番順を 1 回ずつ行った時に得られる期待利得の平均をとった。

各プレイヤーが他のプレイヤーの戦略を知らず、他のプレイヤーが自分と同じ戦略を用いていると仮定して手を選ぶ場合の結果を表 5.3 に、各プレイヤーが他のプレイヤーの戦略を知っている場合の結果を表 5.4 に示す。表 5.3 は現実のゲームで対戦したときの結果、表 5.4 はゲーム理論の分野でゲームの解を求めたときの結果であるといえる。

表 5.3: 相手の戦略を知らない場合の各戦略の平均利得

戦略	平均利得	戦略	平均利得
1EP 1 人	0.374	RNDall 2 人	0.313
1EP 2 人	0.421	RNDall 1 人	0.157
1EP 1 人	0.347	RNDeta 2 人	0.327
1EP 2 人	0.359	RNDeta 1 人	0.281
1EP 1 人	0.213	eqprob 2 人	0.394
1EP 2 人	0.333	eqprob 1 人	0.333

表 5.4: 相手の戦略を知っている場合の各戦略の平均利得

戦略	平均利得	戦略	平均利得
1EP 1 人	0.615	RNDall 2 人	0.192
1EP 2 人	0.431	RNDall 1 人	0.139
1EP 1 人	0.449	RNDeta 2 人	0.276
1EP 2 人	0.458	RNDeta 1 人	0.083
1EP 1 人	0.472	eqprob 2 人	0.264
1EP 2 人	0.417	eqprob 1 人	0.167

5.3 考察

実験 1 では「既存手法」, 「 $k_{mn}-\eta$ 」, 「 $k_{mn}-\eta$, 均衡点」の順に, 戦略が一意に定まるノードの数が増加するという, 理論通りの結果が得られた.

実験 2 では「元のゲーム木」, 「同値変形」, 「 $k_{mn}-\eta$, 同値変形」, 「 $k_{mn}-\eta$, 均衡点, 同値変形」の順にノードの数が増えていくという, 理論通りの結果が得られた. 「 $k_{mn}-\eta$, 同値変形」で大きくノードの数が増えたのは, ダイヤモンドゲームは手番を順に行うゲームであるので, 子ノードがどれも同じキングメーカーノードであるので 1 つ残してそれらを削除する, という変形が多く行われたと考えられる. また, 「 $k_{mn}-\eta$, 均衡点, 同値変形」では, 各プレイヤーが $k_{mn}-\eta$ 則に従い部分ゲーム完全一段階予測均衡点に従うとすると, キングメーカーノードが 1 つあるだけの木まで単純化できることを実験で確認することができた.

実験 3 の結果の 1 つ目の表 5.3 を見ると, 2 つの組み合わせを除き 1EP が得た利得が他の戦略を上回っている. ランダムを中心とする戦略に対して 1EP のほうが強い戦略であると言える.

しかし, 「お互いの戦略を知らない時の戦略 1EP が 1 人, eqprob が 2 人」では eqprob に下回り, 「お互いの戦略を知らない時の戦略 1EP が 2 人, eqprob が 1 人」では同点となっていた. この原因を調べるために実際にプレイされたゲーム木を見ると, 戦略 1EP を用いるプレイヤー a が「プレイヤー b はプレイヤー c を勝たせないように振る舞う」と予測して戦略を決めているが, プレイヤー b が実際は戦略 eqprob を用いているためにこの予測が間違っており, プレイヤー a は c の得る期待利得が予測より大きく自分の得る期待利得がその分小さくなっている子ノードを選んでしまう, という状況が見られた. a と b の戦略が逆の場合にも同様の状況が見られた. また, 戦略 1EP を用いるプレイヤーは多くのノードで子ノードを等しい確率で選ぶので, eqprob に比べ, ある部分木でのミスが全て最終的に得られる利得へと影響してしまうといった欠点が見られた.

一方実験 3 の 2 つ目の表 5.4 では, 全ての場合において, 戦略 1EP を用いるプレイヤーが他のプレイヤーの戦略を上回っている. 1 つ目の表 5.3 で示した結果よりも大きく, 1EP が他の戦略を上回っている. 1EP の戦略は「自分の戦略の変更によって, 他のプレイヤーに戦略を変えさせることにより自分の利得を上げる」というものであるため, お互いの戦略を知っている場合には 1EP を用いるプレイヤーの意図通りの戦略を他のプレイヤーが取り, 1EP を用いているプレイヤーの利得が大きくなる, という理論通りの結果となった.

これらの実験の結果から，提案手法はゲーム理論の分野で，合理的な戦略の組であるゲームの解を求める際には有用であったが，現実のゲームでのコンピュータプレイヤーを作る際は，単純に提案手法を探索手法に当てはめるだけでは，既存の方法を用いているプレイヤーより弱くなってしまうことが分かった．

第6章 結論

6.1 まとめ

本稿では、ゲーム理論の観点から2値ゲームの解析に有用ないくつかの概念を示した。

各プレイヤーが合理的な選択好順で子ノードの選択を行うようにするために、従うべき法則として $kmn-\eta$ 則をその根拠とともに導入した。また、一段階予測に基づく戦略変更の動機に従うことが合理的であることを示し、この動機によって導かれる部分ゲーム完全一段階予測均衡点を定義した。2値ゲームに対して、既存手法である部分ゲーム完全均衡点は無数に存在し、合理的でない戦略の組も多く含んでいたが、 $kmn-\eta$ 即と部分ゲーム完全一段階予測均衡点によって、合理的な戦略の組のみが導かれるようになった。

また、2値ゲームの本質を最小限の大きさのゲーム木で表現することを可能にする同値変形の手順を示した。この最小限の大きさのゲーム木により、ゲームの解析が容易になり、ゲームの本質を直感的に理解することができるようになった。

3人等和2値ゲームについて考察を行い、 $kmn-\eta$ 即と部分ゲーム完全一段階予測均衡点によって2人のプレイヤーの全てのノードで戦略が一意に導かれることを示した。このゲーム木に同値変形を行うと、1つのキングメーカーノードと2つのリーフからなるゲーム木にまで単純化でき、ゲームの本質は1人のキングメーカーであることを示した。

最後に実際のゲームである3人ダイヤモンドゲームを用いた実験を行い、導入した概念による効果の実例を示した。

6.2 今後の課題

今後の課題としては、全探索が不可能で評価値ベクトルを用いなければならないゲームや利得が実数のゲームにおいて、この部分ゲーム完全一段階予測均衡点が有効な働きをするのかどうかの検証や、そのための均衡点の概念の改良が考えられる。

また、現実のゲームにおいてはこの部分ゲーム完全一段階予測均衡点に従った戦略をとるだけでは、強いコンピューターゲームプレイヤーにはならないことが実験的に示されたので、現実のゲームで強いプレイヤーを作るための、均衡点の探索手法への適用の方法を改良する必要がある。

また、全探索不可能な2値ゲームでは、評価値を用いて \max^n アルゴリズムを用いる方法が一般的であるが、探索中にリーフ近くのキングメーカーノードに到達してしまった場合、最善手を一意に定めることが出来ない。ゲーム木の末端では部分ゲーム完全一段階予測均衡点に従う \max^n を用いることで、この問題を解決することが出来るかの検証も今後の課題として考えられる。

参考文献

- [1] D. Billings, A. Davidson, J. Schaeffer, and D. Szafron. The challenge of poker. *Artificial Intelligence*, Vol. 134, No. 1-2, pp. 201–240, 2002.
- [2] D. Billings, D. Papp, J. Schaeffer, and D. Szafron. Opponent modeling in poker. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, pp. 493–499. JOHN WILEY & SONS LTD, 1998.
- [3] D.E. Knuth and R.W. Moore. An analysis of alpha-beta pruning* 1. *Artificial intelligence*, Vol. 6, No. 4, pp. 293–326, 1975.
- [4] L. Kocsis and C. Szepesvári. Bandit based monte-carlo planning. *Machine Learning: ECML 2006*, pp. 282–293, 2006.
- [5] R.E. Korf and D.M. Chickering. Best-first minimax search. *Artificial Intelligence*, Vol. 84, No. 1-2, pp. 299–337, 1996.
- [6] D.M. Kreps and R. Wilson. Sequential equilibria. *Econometrica: Journal of the Econometric Society*, pp. 863–894, 1982.
- [7] H.W. Kuhn. Extensive games and the problem of information. *Contributions to the Theory of Games*, Vol. 2, No. 28, pp. 193–216, 1953.
- [8] U. Lorenz and T. Tscheuschner. Player modeling, search algorithms and strategies in multi-player games. *Advances in Computer Games*, pp. 210–224, 2006.
- [9] C. Luckhardt and K. Irani. An algorithmic solution of n-person games. In *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI)*, pp. 158–162, 1986.
- [10] O. Morgenstern. *Wirtschaftsprognose*. J. Springer, 1928.
- [11] J. Nash. Non-cooperative games. *The Annals of Mathematics*, Vol. 54, No. 2, pp. 286–295, 1951.
- [12] J.F. Nash, et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, Vol. 36, No. 1, pp. 48–49, 1950.
- [13] D.S. Nau. The last player theorem. *Artificial Intelligence*, Vol. 18, No. 1, pp. 53–65, 1982.

- [14] A. Okada. On stability of perfect equilibrium points. *International Journal of Game Theory*, Vol. 10, No. 2, pp. 67–73, 1981.
- [15] R. Selten. Spieltheoretische behandlung eines oligopolmodells mit nachfragerträgeit: Teil i: Bestimmung des dynamischen preisgleichgewichts. *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics*, Vol. 121, No. 2, pp. 301–324, 1965.
- [16] R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International journal of game theory*, Vol. 4, No. 1, pp. 25–55, 1975.
- [17] T. SHINOHARA and T. ISHIDA. Best-first Search for N-player Games. *Transactions*, Vol. 43, No. 10, pp. 2981–2989, 2002.
- [18] N. Sturtevant. A comparison of algorithms for multi-player games. *Computers and Games*, pp. 108–122, 2002.
- [19] N. Sturtevant. An analysis of UCT in multi-player games. *Computers and Games*, pp. 37–49, 2008.
- [20] N. Sturtevant and M. Bowling. Robust game play against unknown opponents. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, p. 719. ACM, 2006.
- [21] N. Sturtevant, M. Zinkevich, and M. Bowling. Prob-Maxⁿ: Playing N-Player Games with Opponent Models. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, Vol. 21, p. 1057. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006.
- [22] N.R. Sturtevant and R.E. Korf. On pruning techniques for multi-player games. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, pp. 201–208. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2000.
- [23] I. Szita, G. Chaslot, and P. Spronck. Monte-Carlo Tree Search in Settlers of Catan. *Advances in Computer Games*, pp. 21–32, 2010.
- [24] T. Tranæs. Tie-breaking in games of perfect information. *Games and Economic Behavior*, Vol. 22, No. 1, pp. 148–161, 1998.
- [25] E. Van Damme. *Stability and perfection of Nash equilibria*. Springer, 1991.
- [26] I. Zuckerman and A. Felner. The mp-mix algorithm: Dynamic search strategy selection in multi-player adversarial search. *Computational Intelligence and AI in Games, IEEE Transactions on*, No. 99, pp. 1–1, 2011.

-
- [27] I. Zuckerman, A. Felner, and S. Kraus. Mixing search strategies for multi-player games. In *Proceedings of the 21st international joint conference on Artificial intelligence*, pp. 646–651. Morgan Kaufmann Publishers Inc., 2009.
- [28] 岡田章. ゲーム理論. 有斐閣, 1996.

発表文献

1. 水野 悠, 鶴岡 慶雅, 近山 隆. キングメーカーノードを含むゲーム木の均衡点. 情報処理学会第27回ゲーム情報学研究会, 小金井, 2012年3月.

謝辞

本研究を進めるに当たり多くの方々にお世話になりました。

指導教員である近山隆教授には研究の内容・論文作成・発表などについて、多岐に渡るご指導をしていただきました。研究の方針や問題解決のためのアイデアについて、多くのアドバイスをいただきました。

研究室で行っている AI 勉強会では、鶴岡慶雅准教授，三輪誠先輩，浦晃先輩，矢野友貴先輩を初めとするメンバーの方々から，研究内容について様々なご意見・アドバイスをいただきました。AI 勉強会のメンバーとの議論によって研究の質を高めることができました。

田浦健次朗准教授には卒論生のころから研究の進め方や発表に関して，様々な指導をしていただきました。普段の研究室生活では近山・鶴岡研究室，田浦研究室のみなさまにお世話になりました。

皆様，本当に有難うございました。心よりお礼申し上げます。

平成 24 年 2 月 8 日