同一チェスプログラムグループによる合議制における投票結果の分散パターンの調査

曾根 彰吾[†1]　長東 薫[†1]　由井薗 隆也[†2]　飯田 弘之[†1]

近年，様々な研究により，複数のプログラムによる合議システムが将棋やチェスなどにおいて有効であることはよく知られている．しかし，なぜ合議制による集団が性能を向上させるのかという疑問に明確な答えは出ていない．この問題の答えを探るために,「集団の知恵」という考え方に着目し，合議アルゴリズムにおける集約性についての提案と考察を行う．その為に，全く同一のチェスプログラムから成る単純合議アルゴリズムを用いて実験を行い，グループサイズが大きいほどそれよりも小さいものと比べて勝率が高くなることがわかった．また，その実験における各プログラムの投票結果を 4 つのパターンへと分類し，投票がどのように集約され，また分散されたのかを分析した．その結果から，合議制として適したグループサイズは，十分な大きさを持ち，かつ奇数の集団ではないかと考えた．

Investigation of the dispersion patterns of the voting results in Council System with Homogeneous Chess Program Group

SHOGO SONE[†1]　KAORU NAGATSUKA[†1]
TAKAYA YUIZONO[†2]　HIROYUKI IIDA[†1]

It is well known that the council system of three or more programs is a powerful evaluation tool used in various social sciences, and well-adapted for computer games such as chess and shogi. However, no one has a clear answer why performance is improved by the group voting system. In order to find the answer to this question, we focused on the concept "The Wisdom of Crowds." We proposed the Aggregation in Council Algorithm, and analyzed the results of thousands of chess games. Using majority voting consisting of iterations of the same program to test the effect of group size on performance, we classified the outputs of each program into four categories and investigated how the voting is gathered or divided. As a result, we propose that the condition of optimal group size in council system may have enough engines and odd numbers groups.

1. はじめに

「三人寄れば文殊の知恵」という諺があるように，合議（合意形成プロセス）を経て，集団で意思決定することは有用であろうと考えられる．そして昨今，合議制の利点をコンピュータに適用する研究が展開され，優れた成果がもたらされている．2009 年 5 月に開催された，第 19 回世界コンピュータ将棋選手権では，6 個の将棋プログラムから成る合議システム「文殊」が第 3 位という優秀な成績をおさめた[1]．また，2010 年 10 月には合議アルゴリズムが用いられた「あから 2010」が清水市代女流王将と対局し，勝利をおさめた．このように，合議が有用であることは十分に確かめられている．しかし，なぜ個人が集団と成ることで性能が向上するのかという疑問には明確な答えは出ていない．

この問題を考えるに当たり , Surowiecki の「集団の知恵」という考え方に着目した[2]．Surowiecki は多くの実例や社会科学の研究結果を上げ,「集団は優れた知力を発揮し，その集団の知恵は集団の中でいちばん優秀な個人の知力よりも優れている」と述べている．しかし，全ての集団がいつも効果的に働くとは限らない．そこで，Surowiecki は集団の知恵が成立する条件として，意見の多様性，独立性，分散性，集約性の 4 つを挙げた．意見の多様性とは，各人が固有の情報を有しており , 集団の多様性が高い状態を表す．独立性とは，個人の意見は他者に流されることがあってはならないということである．分散性とは，個々がそれぞれ専門分野や得意とする分野を持ち，判断するということである．最後の集約性とは，意見をまとめあげるシステムが存在することを指す．これらの観点から見ると，コンピュータ科学の分野では主に意見の多様性と集約性についての研究が見られる．Marcolino らは囲碁の合議制研究を通して，性能は低いが多様なエージェントから成る集団の方が，強いエージェントの複製で構成された均一なチームよりも性能が向上する可能性があると述べている[3]．また，合議システム文殊で行われている乱数合議では，各プログラムの多様性と性能について言及され，杉山らにより提案された楽観的合議というアルゴリズムは，合議の集約部に工夫を

†1 北陸先端科学技術大学院大学 情報科学研究科
Japan Advanced Institute of Science and Technology School of Information Science
†2 北陸先端科学技術大学院大学 知識科学研究科
Japan Advanced Institute of Science and Technology School of Knowledge Science

したものとなっている[4][5]．この様に，合議の仕組みとその性能について多くの研究が成されているが，そもそも，なぜ合議が生まれるのかということについて深く考えられてはいない．本来，チェスや将棋，囲碁などの二人零和有限確定完全情報ゲームにおいて，どのような局面でも最善手を打つことができるのであれば，次の手は常に決定的と成るはずである．この場合，合議制を用いても結果は変わらないため，有効性が見られることはない．つまり，合議システムが有効に働くには，各プログラムによる投票結果にある程度の差がでなければならない．故に，合議制の有効性を議論するには投票結果がどのように分かれるのか，また，合議制の集団の大きさと意見の分かれ具合やその合議制の性能差との間にどのような関係があるのか，その集約部分を調べる必要があると考えた．

本論文では，合議制がなぜ有効であるかを探るために全く同じチェスプログラムによる単純合議制を作成し，使用するチェスプログラムには一切の操作や差別化を行わず，全く同じ条件下での実験を行うことで，チェスプログラムの種類による違いの影響を抑えることを試みる．そしてまず，合議制のプログラム数が合議制の性能にどのように関係するのかを調査し，個々の試合の投票結果を調べることで投票結果の分散パターンを調査し，全く同じチェスプログラムによる合議制の集約性や性能がプログラム数とどのような関係があるのかを調べる．

2. 関連研究

1932 年，Shaw は個人の問題解決能力と比較するためにグループによる協力パズル問題解法についての研究を行った[6]．そして，これらのグループの中には明確なアイディアのコミュニケーションがあり，個人と比べて正答率が非常に高く成ることが保証されることがわかった．Shaw によればこの現象の主な理由として，個人にて起こり得る誤りがグループと成ることによって修正されるからではないかと述べている．そして Althöfer らによってチェスプログラムによる合議制の検証が行われ，その有効性について実証された[7]．また，合議による意見の集約に関して研究されたものとして，Althöfer らによって行われた 3-Hirn がある[8]．3-Hirn とは，2 つのチェスプログラムが出力した候補手を，十分に強い人間のチェスプレイヤーが選択するという手法のことであり，この合議手法により元のプログラムより強くなることが知られている．

そして，Spoerer らによって，Stockfish，TogaII，Bobcat の 3 つのチェスプログラムによる合議制の性能分析の研究が行われた[9]．この研究では，合議アルゴリズムに単純合議制が用いられ，合議グループと Stockfish 単体の試合を合議グループ側の各プログラムの探索深さを変更して検証した．その際，投票の意見がわかれた時に次の手を決定できるように，3 つのプログラムの中で一番強いものをリーダーとして設定し，多数決で決まらない場合はリーダーを含むグループの手を指すように設定された．その結果，3 つのプログラムによる合議制は有効に働くことがわかった．さらに，合議制が有効に働く条件として，グループメンバーは，ほぼ同じ強さであるべきである，そして，グループのリーダーの候補手が合議時に否決される割合はメンバーの強さに依存するという 2 つの条件がわかった．さらに，Spoerer らによって同一のチェスプログラムグループによる合議制の性能分析が行われた[10]．その結果，グループサイズが大きくなるほど勝率も増加することがわかった．

その他，合議制の性能分析における研究として乱数合議や楽観的合議などがある[4][5]．乱数合議とは，単一の将棋プログラムを局面評価値に乱数を加えて複製し，複数のプログラムの合議とすることで多様性の増した将棋プログラムの集団による合議制を作成する手法である．この乱数合議では，与えられる乱数の標準偏差が大きくなるほど，各局面での意見が分かれることが確認されている．そして，この乱数の標準偏差と合議制の性能を比較した実験も行われた．また楽観的合議とは，複数のプログラムから出力される候補手の局面評価値を比べ，その中で最も高い局面評価値を出力したプログラムの手を選ぶアルゴリズムのことである．この楽観合議では多数決を全く取らないので，たとえ一番高い評価値を出力したプログラムが一つであっても，そのプログラムの候補手を選択する．

また，合議制の数学的な分析として，佐藤らにより合議制の数学的モデルが提案され，合議アルゴリズムが成功する為の必要十分条件が導き出された[11]．

3. 単一プログラム対合議グループによる実験

3.1 同一チェスプログラムグループによる合議

実験には，十分に強いチェスプログラム「Stockfish 2.2.2 64-bit」と「Gull II b2 64-bit」を用いる．Stockfish 2.2.2 64-bit のエロレーティングは 3163，Gull II b2 64-bit のエロレーティングは 3006 である．そして，これらを用いた同一チェスプログラムから成るグループによる単純合議システムを作成する．そして，単一のチェスプログラム S_1，G_1 と n 個から成るチェスプログラムグループ S_n，G_n をそれぞれ自己対戦させ，グループサイズの増加と共に勝率がどのように変化するのかを調べた．ここで使用されるすべてのチェスプログラムの探索の深さは 11 に固定し，時間的な制約などは設けず，使用する CPU コア数は 1 とする．これは，時間制限により探索が打ち切られることで投票結果が変わることを防ぐためである．また、各チェスプログラムに対して一切の変更は行われない．チェスの試合管理には「Winboard 4.7.3」を用い，各試合の序盤は Winboard に用意されている opening book を使用した．チェスプログラムによる探索は

少々，同じような序盤を生成する傾向があるため，多様な試合局面を作成するために無作為に序盤の局面を作成できる opening book が使用される．この opening book を使用している間，合議は行われない．

この条件のもとで，単一プログラムとグループとの対戦を 1000 局（先手後手 500 局ずつ）の自己対戦を行い，次の式により勝率を計算した．

$$\text{winning\_rate} = (\text{wins} + 0.5 * \text{draws})/1{,}000 .$$

3.2 単純合議制

本実験における合議は多数決による単純合議制を用いる．この単純合議の手順は小幡らによって次のように表される[12].

(1) グループ内の n 個のプログラムにより n 個の候補手を算出する．

(2) 各候補手の合計数を計算する．

(3) もし最多数候補がある場合，その手をグループの手として選択する．

(4) もし最多数候補がない場合，リーダーの選択した候補手を選択する．

(4) において，通常はグループで最も強いプログラムがリーダーとして選ばれる．しかし，本論文では同じ強さを持つ単一のプログラムを使用しているため，グループに最初に追加されたプログラムをリーダーとして合議を行う．

3.3 単一プログラムに対するグループの勝率

図 1 に，Stockfish における単一プログラムに対する各プログラムグループ S_n の勝率を，図 2 に Gull における単一プログラムに対する各プログラムグループ G_n の勝率を示す．

図 1 より，Stockfish では，例えば s3 対 s1 でのグループ側の勝率が 56.6%，s10 対 s1 での勝率は 60.4%など，全てのグループにおいて単一のプログラムに対して勝ち越す結果となった．また図 2 より，Gull においても同様に，g3 対 g1 でのグループ側の勝率が 54.7%，g10 対 g1 での勝率は 58.6%と全てのグループにおいて単一のプログラムに対して勝ち越す結果となった．しかし，これらの結果からは誤差の影響も考えられ，グループサイズと勝率の関係ははっきりとは見られない．そこで，大きなグループが他のグループに対して有意であるか確認するため，グループ同士によるリーグ戦を行う．

3.4 グループ同士によるリーグ戦

グループサイズと勝率の関係を調べるため，Stockfish，Gull それぞれの場合で，単一プログラムと合議グループの自己対戦によるリーグ戦を行った．実験条件は 4.1 と同様であり，個々の試合数は 1,000 局（先手後手 500 局ずつ）

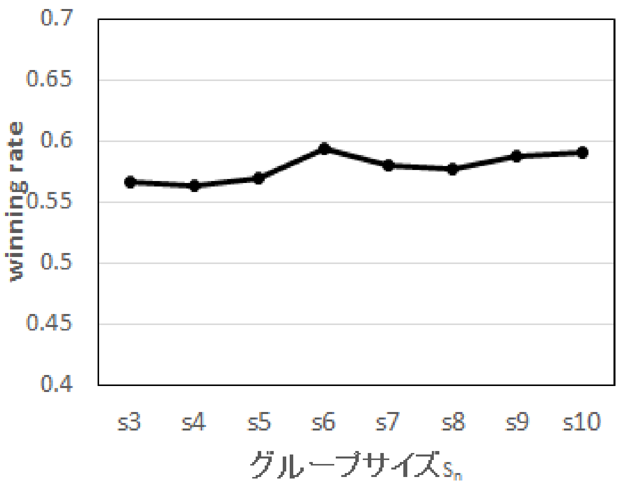

図 1　単一プログラムに対するグループ Sn の勝率 (Stockfish 2.2.2 64-bit)

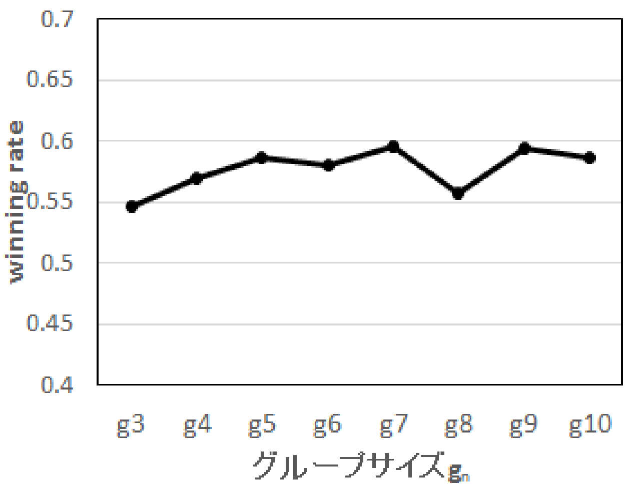

図 2 単一プログラムに対するグループ Gn の勝率 (Gull II b2 64-bit)

行った．それぞれの実験結果の詳細を表 1，表 2 に示す．

表 1，表 2 からもわかるように，Stockfish，Gull どちらのチェスプログラムにおいてもグループサイズが大きいほうのグループの勝率が高くなっている．ただし，s7 対 s6 や，g5 対 g4 の様に，チームサイズの差が少ない対戦では勝率に有意な差があるとは言えない．

さらに，表 1，表 2 のそれぞれのチェスプログラムグループのスコアの合計を求め，その値を各プログラムグループが行った全試合数 5,000 で割ったトータルスコアをグラフにしたものが図 3，図 4 である．これにより，チェスプログラムグループのサイズが大きくなるほど，グループサイズがそれよりも小さいものと比べて勝率が高くなっていくことがわかった．

4. 候補手の集約

4.1 投票結果が同数の場合

単純合議制による候補手の集約の際に，票の別れ方によって答えが一意に決まらないことがある．例えば，チェスプログラムが 4 つのグループを考えた場合，意見の分かれ方は(4)，(3,1)，(2,2)，(2,1,1)，(1,1,1,1)である．ここで，多数決で決められるのは(4)，(3,1)，(2,1,1)の 3 パターンの場合であり． (2,2)と(1,1,1,1)のパターンは最多数候補が同数であるため，次の一手を一つに決めることができない．このような意見が分かれる場面が性能とどのような関係があるのかを調べるため，投票結果を更に細かく分類する．

4.2 過半数，全一致，同数，全不一致

各チェスプログラムの投票結果の分類として次の表 3 のように分類した．

表３　投票パターンの分類

	状態 X
過半数	全一致
	その他（過半数）
同数	全不一致
	その他（同数）

多数決で決まる過半数状態を更に細分化し，全一致の場合とその他の場合とで分け，また，同数となり多数決では決められない状態も，全不一致の場合とその他の場合とで分類する．

先のグループ同士によるリーグ戦より，チェスグループ S_n，G_n の各 1000 試合の棋譜，投票結果などはすでに得られている．その為，この各状態 X が個々の試合中に何度現れたのかを集計し，以下の計算式で出現率を求める．opening book を使用している間は合議が行われないため，合議が行われた総局面は全試合の総局面から opening book の総局面で引かれた値となる．

$$\text{状態 X 出現率} = \frac{\text{状態 X の総出現数}}{\text{全チェス試合の総局面} - \textbf{opening book}\text{ の総局面}}$$

表 4，表 5 はそれぞれ，Stockfish と Gull の自己対戦によるリーグ戦における全一致，その他（過半数），全不一致，その他（同数）の総数と合議が行われた総局面の一覧である．ここで，s3 と g3 のグループにおいて，その他（同数）の総数が 0 なのは，3 つのプログラムによる単純投票において全不一致を除く同数となる状態が起こりえないためである．

本実験では，全く同一のプログラム，同じ探索の深さや探索時間を使用している．それにもかかわらず，s7 や s8，g7 や g8 大きなプログラムグループにおいて非常に低確率ではあるが全不一致が起こっている．

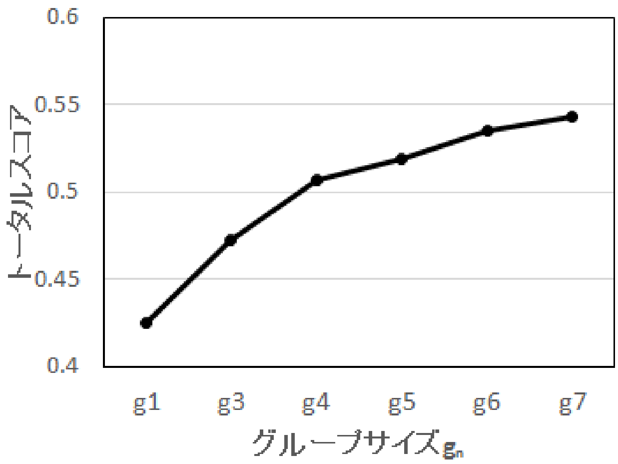

図 3　各グループのリーグ戦によるトータルスコア (Stockfish 2.2.2 64-bit)

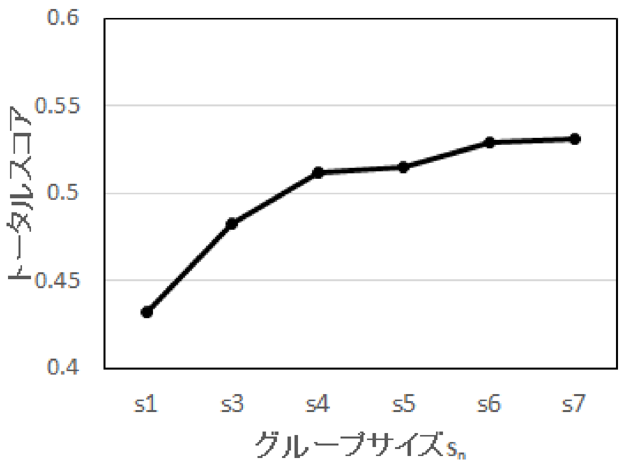

図 4　各グループのリーグ戦によるトータルスコア (Gull II b2 64-bit)

表１　グループ S_n 同士による自己対戦の実験結果 (Stockfish 2.2.2 64-bit)

	s7	s6	s5	s4	s3	s1	Total
s7		501	515	512.5	540.5	584	2653
s6	499		519.5	512.5	539.5	572	2642.5
s5	485	480.5		504	531	574.5	2575
s4	487.5	487.5	496		517	568.5	2556.5
s3	459.5	460.5	469	486		540.5	2412.5
s1	416	428	425.5	431.5	459.5		2160.5

表２　グループ Gn 同士による自己対戦の実験結果 (Gull II b2 64-bit)

	g7	g6	g5	g4	g3	g1	Total
g7		503	522.5	530.5	562	595.5	2713.5
g6	497		523	529.5	543	580	2672.5
g5	477.5	477		502	552.5	587	2596
g4	469.5	470.5	498			569	2532.5
g3	438	457	447.5	474.5		546.5	2363.5
g1	404.5	420	413	431	453.5		2122

また，全一致の総数はどちらのチェスプログラムでもグループサイズが大きくなるにつれて減少し，その他（過半数）の総数は，グループサイズが大きくなるにつれて増加している．一方，同数状態については，全不一致も含めてグループサイズが大きくなるにつれて減少している．

次に，Stockfish と Gull の投票結果パターンの状態出現率を計算し図示したものを図 5 と図 6 に示す．

どちらのプログラムの場合でも，グループサイズが大きくなるにつれて全一致と全不一致の確率は単調に減少している．それに対し，その他（過半数）はグループサイズが偶数の時は谷に，奇数の時は山にと増減を繰り返しながら全体としては増加している．それとは逆に，その他（同数）は偶数の時は山に，奇数の時は谷にとその他（過半数）とは逆の挙動を示しながら減少している．

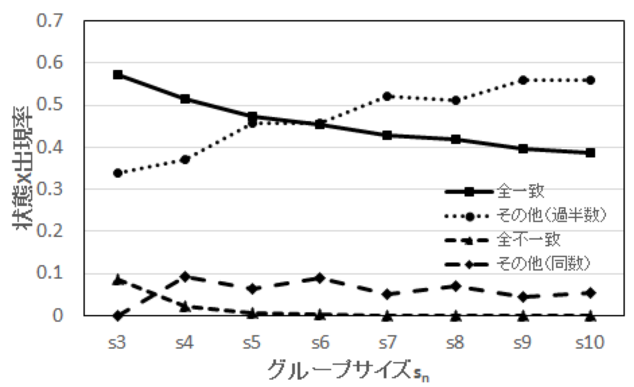

図５　各投票結果パターンの出現率
(Stockfish 2.2.2 64 bit)

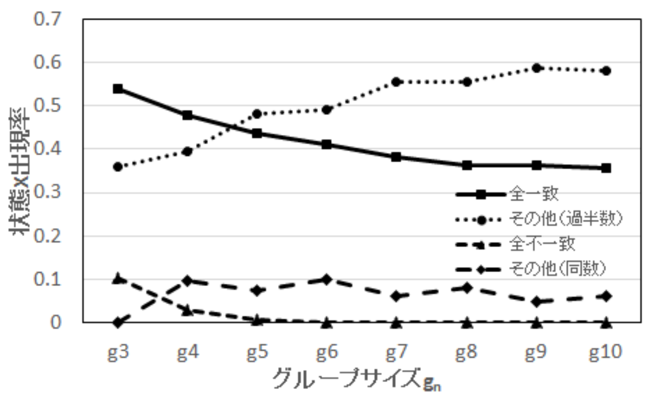

図６　各投票結果パターンの出現率
(Gull II b2 64-bit)

表 4　各チェスエンジングループにおける投票結果パターンの総出現数　(Stockfish 2.2.2 64-bit)

	全一致	その他（過半数)	全不一致	その他(同数)	合議が行われた局面の総数
s3	36840	21759	5579	0	64178
s4	34213	24536	1443	6193	66385
s5	31612	30603	319	4315	66849
s6	28167	28309	59	5595	62130
s7	28515	34801	21	3375	66712
s8	27216	33321	2	4493	65032
s9	26248	37151	1	2933	66333
s10	25212	36637	0	3592	65441

表 5　各チェスエンジングループにおける投票結果パターンの総出現数　(Gull II b2 64-bit)

	全一致	その他(過半数)	全不一致	その他(同数)	合議が行われた手の総数
g3	38467	25718	7417	0	71602
g4	34007	28157	2079	6939	71182
g4	30605	33780	493	5115	69993
g6	27964	33414	144	6785	68307
g7	27767	40292	43	4444	72546
g8	25984	39688	12	5667	71351
g9	25126	40635	3	3412	69176
g10	24965	40562	0	4255	69782

5. 考察

単一のプログラムと複数のプログラムによる合議制による対戦結果より, 全く同じプログラムを用いた場合でも合議制が有効であることが確認された. また, グループサイズが大きくなるにつれて, それより小さいグループに対して全体の勝率が上回っていることがわかる. なぜ, この様なことが起こるのかということについて考えるに当たり, まずは, 例え同じプログラムであっても同じ局面で違う候補手を出力することがあるという事について考察する. チェスプログラムでは, そのプログラムの評価関数や探索部に乱数がかけられていることが多い. これは同じ局面で同じ手を打ってしまうことで, 対戦相手にとって分かりやすい打ち筋とならないように工夫するためであることが主な理由である. 本研究では, 探索時間を無制限, 探索深さを固定とすることでこれらの影響による出力の変化は少ないと見られる. 故に, 同じ局面で違う候補手が出力される要因の一つは, この様なチェスプログラムの乱数構造によるものであると考えられる. この様な出力される候補手の変化は, 同じプログラム, 同じ局面にて出力される候補手が好手になるか悪手になるかということにも関係してくる. よって, 合議アルゴリズムを用いることにより, 1 つのプログラムがたまたま選んでしまう悪手を, 複数のプログラムで打ち消すことができるのではないかと直観的に考えることができる. さらに, 合議制のグループサイズが大きくなると, この効果を受ける機会が単純に増えるため, 勝率が上がるような効果が得られたのではないかと考えた.

また, 合議による投票結果が分かれた局面について, そのような局面は評価の分かれる難しい局面だったのではないかと推測することができる. 実際, この考えをもとに, 合議システム文殊は合議結果が 6 個中 5 個以上一致しなかった場合, 探索延長を 3 段階まで延長するような機能が持たされている[1]. 本研究の結果から, グループサイズが大きい場合がそれより小さい場合よりも有意に働くというのは, グループサイズが大きくなるにつれて多数決による決定が増え, 投票結果が同数と成るような状態が減ることによって, 難しい局面においてグループサイズの効果によってより良い手が選ばれるようになるため, 合議制による性能が向上したのではないかと考えた.

また, プログラムの種類によらず, 合議のグループのサイズが大きくなるにつれて全一致と全不一致の確率が減少するのは, 単純にグループサイズが増える

ことでこれらが起こりにくく成るためであろうとかんがえる．そしてそれらの分が，多数決による決定とその他（同数）の対応するそれぞれへと振り分けられるのではないかと考えた．以上の直観的により，もしも投票結果が同数となるような状態をなるべく少なくすることで合議の性能が向上するのであれば，その他（同数）の出現率が，グループサイズが偶数の時は高く，奇数の時は低くなるという結果から，合議制として適したグループサイズは十分な大きさを持ち，かつ，グループサイズが奇数のグループの時ではないかと考えた．

6. まとめ

本論文では，チェスにおける単純合議制にて，プログラムの設定による違いを持たせない全く同一のプログラムによって構成された集団であっても効果的であることが明確になった．また，合議グループのサイズが増加するにつれて，合議グループの性能も向上することが分かった．また，合議グループの各チェスエンジンが出力する投票結果を 4 つのパターンに分類し分析することで，グループサイズの増加に伴い各パターンの出現率が特徴的な挙動を見せることを示した．以上の結果より，合議制として適したグループサイズは十分な大きさを持ち，かつ，グループサイズが奇数のグループの時ではないかと提案する．

しかし，本研究のデータのみからは，集団の勝率と投票結果パターンとの間に関係があるのかは断定できない．その為，例えば投票結果が同数時には多数決で決定できるまで再投票を行うと勝率はどうなるのか， というような評価実験による検証が必要である．また，それぞれの投票結果パターンの局面を更に細分化し，詳細に分析することでどのような局面で意見が分かれやすいのか，またどのような別れ方に成るのかを分析する．さらに，異なったチェスプログラムグループによる合議制にて同様の分析を行うことで，合議制の有効性についてより深く研究する．

参考文献

1) 伊藤毅志: 合議アルゴリズム「文殊」単純多数決で勝率を上げる新技術, 情報処理, Vol.50, No.9, Sep.2009, pp.887-894 (2009)
2) James Surowiecki: The Wisdom of Crowds, New York, Doubleday, Anchor (2004)
3) Marcolino, L.S. Jiang, A.X. and Tambe, M.: Diversity beats strength? - Towards forming a powerful team, In 15th International Workshop on Coordination, Organizations, Institutios and Norms, COIN 2013 (2013)
4) 小幡拓弥, 杉山卓弥, 保木邦仁, 伊藤毅志: 将棋における合議アルゴリズム：既存プログラムを組み合わせて強いプレイヤを作れるか？, 第 14 回ゲーム・プログラミング ワークショップ 2009, pp.51-58 (2009)
5) Sugiyama, T. Obata, T. Hoki, K. and Ito, T.: Optimistic Selection Rule Better Than Majority Voting System, In Computer and Games 2010, LNCS 6515, pp.156-165, Springer, (2010)
6) Shaw, M.E.: Comparison of Individuals and Small Groups in the Relational Solution of Complex Problems, American Journal of psychology, 44, pp.491-504 (1932)
7) Althöfer, I.: Improved game play by multiple hints, Theoretical Computer Science, Vol.313, Issue.3, pp.315-324 (2004)
8) Althöfer, I.: Selective trees and majority systems: two experiments with commercial chess computers, Advances in Computer Chess 6, pp.37-59 (1991)
9) Spoerer, K, T. Okaneya, T. Ikeda, K. and Iida, H.: Further Investigations of 3-Member Simple Majority Voting for Chess, 8th International Conference (CG 2013), pp.199-207 (2014)
10) Spoerer, K, T. Sirivichayakul, T. and Iida, H.: Homogeneous Group Performance in Chess, 4th International Conference on Electrical Engineering and Informatics (ICEEI 2013)
11) Sato, Y. Cincotti, A. and Iida, H.: An analysis of voting algorithm in games, In Computer Games Workshop, at European Conference on Artificial Intelligence (ECAI), pp.102-113 (2012)
12) Obata, T. Sugiyama, T. Hoki, K. and Ito, T.: Consultation Algorithm for Computer Shogi: Move Decisions by Majority, The 7th International Conference on Computers and Games (CG 2010), vol.6515, pp.156-165 (2010)