将棋における合議アルゴリズム-局面評価値に基づいた指し手の選択-

杉　山　卓　弥[†1]　小　幡　拓　弥[†2]　斎　藤　博　昭[†1]
保　木　邦　仁[†3]　伊　藤　毅　志[†2]

コンピュータ将棋において，複数のクライアントの意見から 1 つの指し手を選択する合議アルゴリズムが注目されている．非常にシンプルな手法ながら，この合議アルゴリズムによって，思考エンジンの性能改善が報告されている．本研究では，各クライアントの指し手に加えて，局面の優劣評価の値も利用する新しい合議法について報告する．正規乱数を用いて局面評価値を修正し生成された複数クライアントから，最大の評価値を付けたクライアントの意見を採用する楽観的合議法を提案し，その有効性と仕組みについて議論する．

Consultation Algorithm in Brain Game
-A Move Decision Based on the Positional Evaluation Value of Each Player -

Takuya Sugiyama,[†1] Takuya Obata,[†2] Hiroaki Saito,[†1]
Kunihito Hoki[†3] and Takeshi Ito[†2]

A recently proposed consultation algorithm, which selects a single move decision from those of many players, has attracted a great deal of attention. In a simple way, the consultation algorithm improves the performance of computer Shogi engines. In this paper, we propose a new strategy to a move decision concerning positional evaluation values, in addition to move decisions of many players. Furthermore, the performance and mechanism of the proposed consultation algorithm are clarified.

1. はじめに

「三人寄れば文殊の知恵」を検証するようなグループによる問題解決は，古くから認知科学分野で行われてきた．1932 年に Shaw は，人間を題材に個人の問題解決能力とグループとしての問題解決能力を比較した実験を行った[1)]．その後の研究においても様々な課題を用いて，グループの問題解決が個人よりも優れたパフォーマンスを示すことを確認した．

しかし，Shaw らの研究は，知的資源の量に対する効率を考慮したものではなかった．そこで Lorge らは，メンバーの持つ様々な知的資源を考慮した時，その総和以上の結果がグループで得られるかについて議論した[2)]．すなわち，グループによる議論が，個々の問題解決能力を上回る結論を導きうるのかという点について議論した．彼らは，グループメンバーの一人が課題を正しく解ければグループ全体の解となると考え，それを基準にグループの問題解決が個々の問題解決能力を上回りうるかを検証した．個人の正解確率を p (簡易化のため，個人間で一定とする) として，グループの人数を n とするとき，グループの正解率の理論値 P は，$P = 1 - (1 - p)^n$ となる．彼らは，それまでの Shaw らの研究で扱った課題を用いて検証した結果，殆どが理論値を有意に下回るか，同じ程度であることを示した．

機械学習の分野にグループによる問題解決という考え方を展開した例として，近年，Boosting 法が注目されている．これは，正答率の低い分類器を集めて高性能の分類器を構成する手法であり，画像認識等の分野での応用が期待されている[3)4)]．

思考ゲームの手の選択にこの考えを用いた研究では，1985 年頃から行われた Althofer らの「3-Hirn システム」が挙げられる[5)]．彼らは，2 台のチェスプログラムが思考した別々の指し手を候補手とし，それを十分に強い人間が選択するという手法を用いることにより，本来のチェスプログラムよりレーティングが 200 程度上昇することを示した．しかし,「3-Hirn システム」では，最終的な手の選択は人間が行っていたため，システム単独での思考は行えなかった．

†1 慶應義塾大学大学院理工学研究科
Keio University Science and Technology
†2 電気通信大学情報工学科
The University of Electro-Communications Department of Computer Science
†3 東北大学大学院理学研究科
Tohoku University Graduate School of Science

2009 年 5 月に開催された第 19 回コンピュータ将棋選手権においては，コンピュータのみで様々な候補手の中から最終的な手を選択する「合議アルゴリズム」を利用した「文殊」が第 3 位という優秀な成績をおさめた[6]．ここで用いた合議アルゴリズムは，局面評価値に乱数を加えた複数の Bonanza に次の一手を思考させ，指し手の多数決を行うものであった．

この合議アルゴリズムは，疎結合マルチプロセッサを活用する手法として，その有効性や実用性が期待される．ムーアの法則に基づいて 2000 年代まで，CPU の性能は発展を続けてきたが，近年，その伸びに陰りが見え始めており，マルチプロセッサ化が進んでいる．ソフトウェアの分野でも並列化の試みが盛んに行われているのは周知の事実である．将棋においても並列処理により効率化を図るのは重要ではあるが，一般に将棋のようなゲーム木の探索を並列化することは容易ではない．このような密結合マルチプロセッサの概念から離れ，疎結合マルチプロセッサとして実装できる手法の一つとして合議アルゴリズムは新しい方向性を示していると言える．

「文殊」の採用した合議手法は単純多数決であったが，複数の思考システムから一つの意見を選択する手法には，まだ工夫の余地があると考えられる．そこで，本研究では，局面評価値に注目することにした．今回は，その中でもシンプルな「最も高い評価値 (楽観的評価値)，低い評価値 (悲観的評価値) を出したクライアントの指し手を選択する合議アルゴリズム」を提案する．本報告では，Bonanza を用いて行った実験結果をもとに考察していく．

2. 関 連 研 究

将棋における合議アルゴリズムは，2009 年 3 月に５五将棋を題材に塙らにより発表され[7]，同年 6 月に小幡らによって本将棋においても有効性が示された[8]．小幡らの示した合議システムは，乱数を用いた単一プログラムが返す指し手を単純に多数決する合議手法である．

合議を行うためには，異なる思考を持つクライアントが必要不可欠である．彼らの研究では，様々なプログラムを合議させることが困難であったため，単一の思考プログラムの評価関数にあらかじめ設定した乱数を加えることで複数のクライアントを作り出した．乱数を加えることで単体としては従来のプログラムに比べて弱くなってしまう傾向にあるが，その単体を合議させることによりグループとしての強さを持つことを期待し，勝率の変化を検証している．

クライアントとして 2009 年 2 月に公開された将棋プログラム「Bonanza version 4.0.4」[9] を用い，その評価関数から出力される評価値に分散 D を用いた正規分布 $N(0,D^2)$ に従う乱数を加えることで，複数の Bonanza を作り出している．正規乱数列は Bonanza の起動時に生成し，ハッシュキーによって参照しているため，トランスポジションテーブルの矛盾は起こらない．このようにして，複数の指し手と評価値を生成する M 個の Bonanza を起動し，各々の探索による最適な手の多数決をとり，最も多い手を最終的な指し手としている．

彼らは，様々な設定で提案手法と手を加えていないオリジナルの Bonanza との自己対戦で実験を行っている．D の値を 4，8，16 の 3 通り，M の値を歩の交換値 (202) の 1/8(25)，1/4(50)，1/2(101)，1/1(202) の 4 通りで行い，先後 500 局ずつ計 1000 局の対局で勝率のデータを得ている．各設定で最も勝率の良かったものを以下の表 1 に示す．なお，勝率は合議側の視点である．

表 1　単純多数決を用いた合議アルゴリズムの最高勝率

M	D	探索ノード数	定跡	勝率
16	50	200000	有	57.65 %
8	50	400000	有	57.42 %

どの設定においても，57%台という結果となった．この実験により，適切な D と M を用いることでオリジナルのプログラムに勝ち越すことから強さが向上している可能性が示唆された．

彼らの研究では，多数決のみによる指し手の選択であったが，本研究では各クライアントが出力する指し手の評価値を考慮に入れてさらなる勝率の向上を試みた．そこでシンプルな手法である「最も高い評価値 (楽観的評価値)，低い評価値 (悲観的評価値) を出したクライアントの指し手を選択する合議アルゴリズム」についての実験，検証を行った．

3. 実　　験

3.1 方　　法

実験には，将棋プログラム「Bonanza version 4.0.4」を用い，単一プログラムを用いて複数の指し手を生成する方法は，従来の小幡らの研究を参考にした．Bonanza の評価関数から出力される評価値に分散 D を用いた正規分布 $N(0,D^2)$ に従う乱数を加えることで，複数の Bonanza を作る．正規乱数列は Bonanza の起動時に生成し，ハッシュキーによって参照しているため，トランスポジションテーブルの矛盾は起こらない．こ

表 2　最高の評価値で指し手を選択する合議 (楽観的合議) の実験結果

M / D	25 (歩の交換値 / 8)	50 (歩の交換値 / 4)	101 (歩の交換値 / 2)	202 (歩の交換値)
1	49.35%	50.71%	40.93%	34.40%
4	54.59%	56.35%	55.72%	46.94%
8	55.27%	58.71%	58.94%	47.19%
16	58.43%	60.20%	56.38%	52.51%
32	57.96%	**61.71**%	57.16%	51.91%
64	60.30%	61.39%	-%	-%

表 3　最低の評価値で指し手の選択をする合議 (悲観的合議) の実験結果

M / D	25 (歩の交換値 / 8)	50 (歩の交換値 / 4)	101 (歩の交換値 / 2)	202 (歩の交換値)
1	49.35%	50.71%	40.93%	34.40%
4	43.61%	39.86%	33.80%	21.66%
8	41.15%	34.47%	26.51%	12.24%

のようにして，複数の指し手と評価値を生成する M 個の Bonanza を起動し，各々の探索で発見した次の一手の中で評価値を考慮の上，指し手を選択する．実験では，序盤は Bonanza の定跡を使用し，定跡から外れた段階から合議を開始することにした．すべての設定において，同じ乱数シードを用いているため，定跡による不公平は一切ないものとする．これによって，1000 局程度の自己対戦では，棋譜の重複が殆ど無くなることが，従来の研究で示されている．また，マシンスペックなどの実験環境の影響を極力少なくするため 1 クライアントで一手あたり 20 万ノードの手を読む仕様とした．20 万ノードの手の探索時間は現在普及している標準的なプロセッサ上で約 1 秒である．上述の D と M を変化させ，オリジナルの Bonanza との 1000 局 (先後 500 局ずつ) の自己対戦を行い，その勝率を調べた．なお，千日手と 256 手以上の対局は引き分けとし，勝率は (勝利数)/(1000-引き分け) として算出する．

3.2　結　　果

3.2.1　楽観的評価値の指し手を選択する合議実験

楽観的評価値の指し手を選択する合議は，最も高い評価値を出したクライアントの手を指し手とする手法である (以下，楽観的合議)．但し，評価値が同じで候補手が異なる場合，それらの中からランダムに指し手を選択することにする．楽観的合議の実験データを表 2 に示す．なお表中の数値はすべて合議側の視点である．参考値として，二項分布を用いた仮説検定によれば，1000 局中での勝率が，有意水準 5%では 52.7%，有意水準 1%では 53.7%以上で有意に強いと言える．

D の値は，小幡らが基準とした「Bonanza version 4.0.4」における歩の交換値 (202) を参考にし，25(1/8)，50(1/4)，101(1/2)，202(1/1) を用いた．また，M の値は，1，4，8，16，32，64 とした．なお，M=64 では，D=25，50 のみのデータとする．ちなみに，M=1 は，Bonanza に単に乱数を加えたもののオリジナルに対する勝率を表している．合議なし (M=1) の場合は，いずれもオリジナルの Bonanza とほぼ同等の勝率であるのに対し，合議ありの場合は，常にオリジナルの Bonanza に対し，大きく勝ち越す結果となった．特に，D=50，M=32 における結果では，小幡らの従来の研究に比べても最高の勝率 61.71%を得た．

3.2.2　悲観的評価値の指し手を選択する合議実験

悲観的評価値の指し手を選択する合議は，最も低い評価値を出力したクライアントの手を指し手とする手法である (以下，悲観的合議)．3.2.1 で得られたデータの裏付けとして，悲観的合議の実験を自己対戦にて行った．指し手の選択以外は，楽観的合議の場合と同じである．

悲観的合議の実験データを表 3 に示す．なお表中の数値はすべて合議側の視点である．

当初の予測通り，D と M の値が大きくなるほど，勝率が下がることを結果として得られた．

4. 考　　察

結果として，楽観的な評価値の指し手を選ぶ本研究の合議手法には有意な効果が認められた．この手法も従来の合議手法と同様にアルゴリズムは非常に単純である．従って，分散メモリ型の並列探索が可能な点と，様々なプログラムへの応用も容易に行えるという利点がある．また，コンピュータリソースを十分に活用できる点からもこれからの時代に不可欠な手法の一つとなると期待できる．

4.1　楽観的合議の挙動

多数決合議と楽観的合議の指し手の違う確率を表 4

に示す．設定は，D=50，M=32，探索ノード数 20 万ノード，定跡有りとした．なお，総合議手数は先後 500 局ずつシミュレーションした場合の合議手数である．

表 4　多数決合議と楽観的合議の指し手の違う確率

総合議手数	相違数	確率
35180	5054	14.37%

多くの場合は多数決合議と楽観的合議は同じ結果となるが，約 7 手に 1 手の割合で各合議の手が異なることが分かる．試合の展開としては各合議で大きく異なってくることが伺える．

多数決合議との比較に重点を置いて楽観的合議の挙動例を何点か挙げる．

4.1.1　例　　1

図 1 に，多数決合議と楽観的合議の指し手が異なる局面の一例を示す．

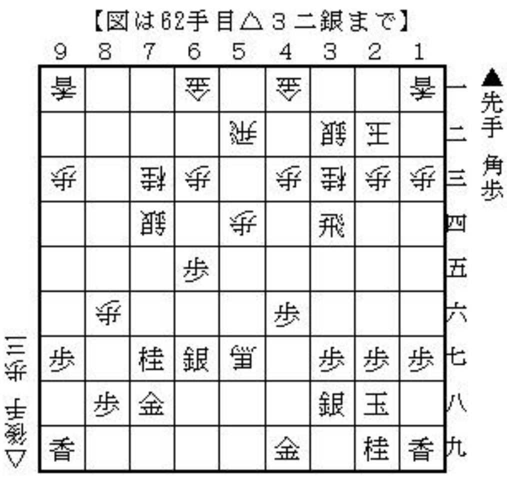

図 1　多数決合議と楽観的合議の指し手が異なる局面 A

後手は馬を作っており，先手は▲３六飛等とゆっくり相手を伺っていると△３五歩～△４六馬～△７六歩等，着実に攻められるため，一気に敗勢に変わってしまう．従って，▲５三歩△同飛▲８二角等と攻めたい局面である．

表 5 に，D=50，M=1000，20 万ノードの探索で得られる候補手を示す．

▲５三歩の頻度は 43 と少ないが，正の評価値を与える唯一の指し手であった．多数決合議では最も頻度が高い▲３六飛を，楽観的合議では最大評価値が一番大きい▲５三歩を選択する確率が高い．実際に，▲３六飛と▲５三歩を指した局面から楽観的合議でシミュレーションを行ったところ，前者は負け，後者は勝ちという結果になった．この局面では，▲５三歩が好手であり，楽観的合議の方が良い手を選択する一例と言える．

表 5　1000 回の探索により得られた候補手の頻度 (局面 A)

候補手	頻度	最大評価値	最小評価値	平均
▲３六飛	934	0	-299	-67.77
▲５三歩	43	53	-124	-3.58
▲４五歩	13	-68	-292	-220.69
▲５五歩	6	-54	-198	-108.00
▲６六角	4	-80	-112	-94.75

表 5 で示されている平均値は 200 程度のばらつきがある．評価関数に与えた標準偏差 50 の正規乱数では，このばらつきは証明できない．そこで，探索の不安定性と正規乱数の関係を検証した．

ここでは，Late Move Reduction(以下，LMR) と Move Ordering から来る不安定性に着目する．LMR とは深さの短縮を行う手法である．短縮を行う展開局面の決定は Move Ordering に依存し，乱数が探索結果に大きな影響を与える可能性がある．

図 2 に，LMR が各候補手の評価値の分布に与える影響を示す．探索は D=50，M=1000，基準深さ 7 の条件で行った．また，探索に不安定性を与えるもう一つの要因であるトランスポジションテーブルは使用しなかった．

主に得られる候補手は，LMR 無しでは好手の▲５三歩，有りでは▲３六飛である．一方，興味深いことに，最大の評価値を与える候補手はいずれの場合も▲５三歩である．

この局面では，基準深さ 7 の探索時間は LMR により半分以上削減される (表 6)．にもかかわらず，LMR 有りの楽観的合議は，LMR 無しと同じ手を選択する．

表 6　局面 A における 1 クライアントあたりの思考時間 (深さ 7)

LMR 有	LMR 無	時間比
3.09 秒	7.09 秒	1 : 2.29

また，楽観的合議と多数決合議が選択する候補手の，探索深さ依存性を検証した (表 7)．ただし，D=50，LMR 有り，トランスポジションテーブル無しとした．中の数値は頻度を表し，最下段では楽観的合議の最終的な指し手を表している．

表 7 の結果から，楽観的合議では深さ 7 で▲５三歩を選択する．一方，多数決合議で同じ結論を得るには深さ 8 の探索が必要になる．合議を全く行わない場合においては，探索の基準深さが 9 に達してもこの好手

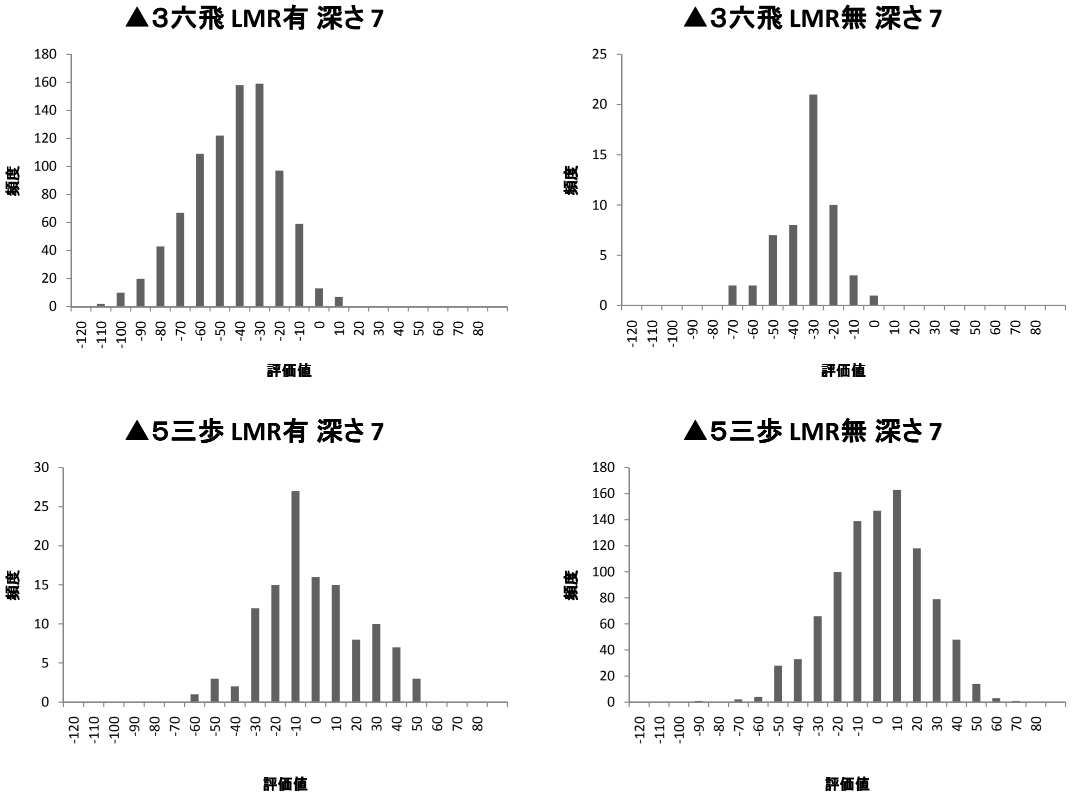

図 2　局面 A の候補手と頻度の Late Move Reduction 依存性

表 7　深さ別の 1000 回の探索における候補手の比較

候補手 / 深さ	4	5	6	7	8	9
a. ▲３六飛	8	691	921	899	158	25
b. ▲５三歩	0	51	18	92	842	975
c. ▲４五歩	16	21	3	0	0	0
d. ▲５五歩	962	236	56	0	0	0
e. ▲６六角	11	1	9	2	0	0
f. ▲５八銀	3	0	0	0	0	0
指し手	d	a	a	**b**	b	b

にたどり着かない可能性が 2～3%あると言える．

4.1.2　例　　　2

図３に，多数決合議と楽観的合議の指し手が異なる局面の一例を示す．

局面Ｂは角を咎められているが，先手はこれから▲同角～▲７一銀～▲６二銀成と攻めを続けたいところである．相手の守りは薄いが先手の持ち駒がないため，▲５七角で攻めは切れてしまい，相手の△３七角や△６五歩等で形勢が悪くなってしまう可能性がある．ここでは▲同角が好手である．

この局面Ｂから D=50，M=1000，20 万ノード探索で一手に対しての合議を行ったところ，分布として表

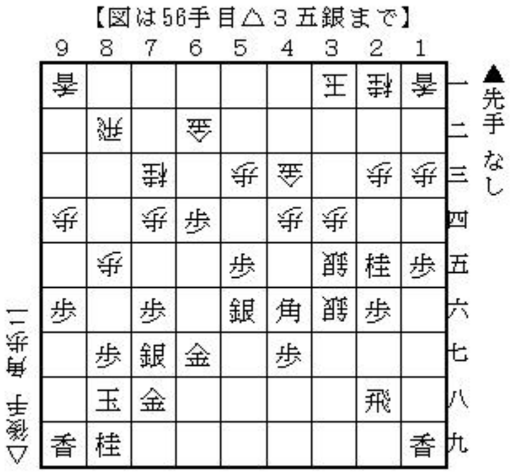

図 3　多数決合議と楽観的合議の指し手が異なる局面 B

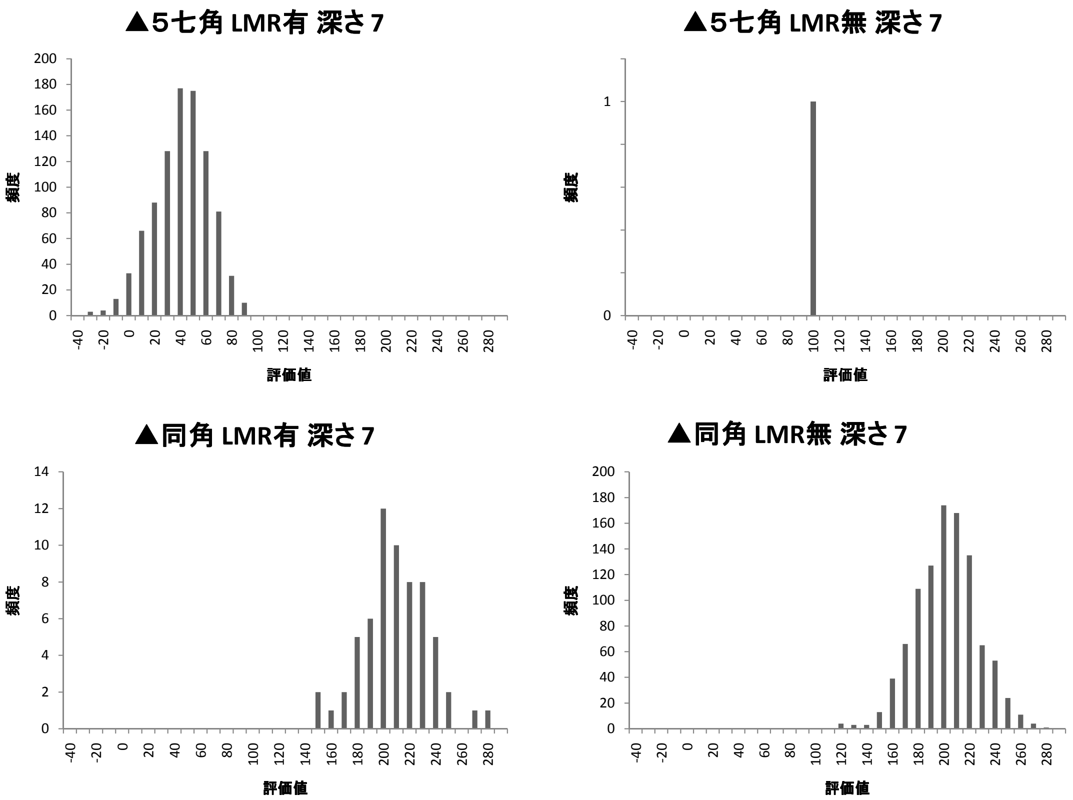

図 4　局面 B の候補手と頻度の Late Move Reduction 依存性

8 のような結果が得られた．

表 8　1000 回の探索により得られた候補手の頻度 (局面 B)

候補手	頻度	最大評価値	最小評価値	平均
▲5七角	948	96	-28	36.07
▲同角	52	259	-17	183.79

1000 回の探索を行った結果，▲5七角，▲同角の 2 通りの手が候補手となった．▲同角の手は例 1 同様な頻度差であるが，多数決合議と楽観的合議の指し手の評価値平均の差はより激しい．図 4 に LMR が各候補手の評価値の分布に与える影響を示す．局面 B では，20 万ノードは深さ 7 が限界であるため，探索は，D=50，M=1000，基準深さ 7 の条件で行った．また，探索に不安定性を与えるもう一つの要因であるトランスポジションテーブルは使用しなかった．

図 4 から，多数決合議では，LMR の有りの場合▲５七角を，LMR 無しの場合▲同角を候補手とする．しかし，楽観的合議は LMR の有無にかかわらず▲同角が候補手となり，この局面においても，例 1 と同様に思考時間を削減しつつ (表 9)，好手を選択することができる．

表 9　局面 B における 1 クライアントあたりの思考時間 (深さ 7)

LMR 有	LMR 無	時間比
0.98 秒	2.53 秒	1 : 2.58

また，楽観的合議と多数決合議が選択する候補手の，探索深さ依存性を検証すると表 10 のような結果が得られた．但し，D=50，LMR 有り，トランスポジションテーブル無しとした．中の数値は頻度を表し，最下段では楽観的合議の選択した候補手を表す．

表 10　深さ別の 1000 回の探索における候補手の比較

候補手 / 深さ	4	5	6	7	8	9
a. ▲5七角	865	754	988	945	925	0
b. ▲3五角	0	0	0	55	75	1000
c. ▲6八角	5	18	3	0	0	0
d. ▲1三桂	130	164	6	0	0	0
e. ▲1三桂成	0	64	3	0	0	0
指し手	a	a	a	**b**	**b**	b

表 10 の結果から，楽観的合議では深さ 7 で▲同角

を選択する．一方，多数決合議で同じ結論を得るためには深さ 9 の探索が必要である．この例においては，楽観的合議を行うことで多数決合議よりも深さ 2 を補っていることができると言える．また，基準深さ 9 に達した場合，多数決合議は合議なしと結果が全く変わらない．

5. 今後の展望

LMR の有無による検証で楽観的合議の果たしている一側面を示すことが出来た．しかし，今回の実験では，まだ「Bonanza version 4.0.4」だけの議論に留まっている．他のバージョンや，他プログラムでも同様の実験結果が得られるのか検証する必要があると考えている．今後は，様々な側面からの検証を重ねて，本手法の有効性について，さらに明らかにしていく必要があるだろう．

また，対戦における勝率データも，Bonanza による自己対戦のみからしか得られなかったが，以後検証を進めていく上で異種プログラムとの対戦や合議が必要不可欠な実験となるであろう．現在，多数決合議の研究においては GPS[★1] や YSS[★2] などの強豪プログラムの合議のデータが得られている．早急に楽観的合議においても結果をまとめていきたい．

6. おわりに

合議アルゴリズムの最大の利点は，実装の容易さ，構造のシンプルさである．どのプログラマでも気軽に取り入れることができるほか，コンピュータのリソースを充分に利用することができより良いパフォーマンスを発揮できる点で非常に良い手法だと言える．現在 CPU のクロック性能は伸び悩む状況であり，単にコア数が増えていくばかりである．本研究で提案した楽観的合議は，そのような状況の中，分散メモリ型の並列探索が可能であり，各コアで評価値と指し手のみを統括プログラムに受け渡すだけで良い．また，合議アルゴリズムは今までの研究とは違った方向性を持っている．これまでに提案されてきたコンピュータ将棋の手法はほぼすべて,「探索」,「評価関数」の改善についてである．しかし合議アルゴリズムは，どちらにも当てはまらない新しい方向の手法である．

★1 東京大学大学院総合文化研究科の教員・学生が開催しているゲームプログラミングセミナー（Game Programming Seminar = GPS）のメンバーが中心になって開発が行われているソフトウェア．第 19 回世界コンピュータ将棋選手権優勝．

★2 山下 宏氏が作成した将棋プログラム．第 7 回，第 14 回，第 17 回世界コンピュータ将棋選手権優勝．

また，今回提案した楽観的合議は，合議の手法のバリエーションとしての可能性を示すことができたと考えている．これまで有効性が示されてきた多数決合議だけでなく，様々な合議の可能性も模索してみる必要があるだろう．以後，多数決と楽観的合議の融合や新たな合議手法も検討して，より効果的な合議手法を突き詰めていきたい．

謝辞 本研究は社団法人情報処理学会から，共同研究による助成を受けている．

参 考 文 献

1) 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).
2) Lorge, I. and Solomon, H. : Two Models of Group Behavior in the Solution of Eureka-Type Problems, Psychometrica, 20, pp.139-148(1955).
3) YoavFreund and Robert E. Schapire : A decision-theoretic generalization of on-line learning and an application to boosting, Journal of Computer and System Sciences, 55(1), pp.119-139(1997).
4) 金森 敬文, 畑埜 晃平, 渡辺 治, 小川 英光 : ”ブースティング - 学習アルゴリズムの設計技法”, 森北出版 (2006).
5) Althofer, I. and Snotzke, R.G. : Playing Games with Multiple Choice System, Computer and Games, pp.142-153(2002).
6) 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).
7) 塙 雅織，伊藤 毅志 : 思考アルゴリズムにおける最適合議システム，第 3 回エンターテイメントと認知科学シンポジウム，pp.72-75(2009).
8) 小幡 拓弥, 塙 雅織, 伊藤毅志 : 思考ゲームによる合議アルゴリズム 〜単純多数決の有効性について〜, 情報処理学会ゲーム情報学研究会報告, 2009-GI-22 No.2(2009).
9) Bonanza は著者 (保木) が作成した将棋プログラム．http://www.geocities.jp/bonanza_shogi/ にてソースファイルが公開されている．市販版も存在する (マグノリア，2008).