

数値最適化手法の コンピュータ将棋への応用

2012年6月16日

早稲田大学 数学教育学会講演会
保木邦仁

内容

- 思考型ゲーム概要
- コンピュータ将棋の現状
- 評価関数の最適化
 - ✓ K. Hoki, T. Kaneko, ACG13, 2011.
 - ✓ T. Kaneko, K. Hoki, ACG13, 2011.
- 力づく探索
 - ✓ K. Hoki and M. Muramatsu, Entertainment Computing, 2011.
- 合議法
 - ✓ 伊藤毅志ら、情報処理学会論文誌, 2011.
- まとめ

研究の背景ー思考型ゲームについて

✓ 思考型ゲームの種類

- 二人・情報公開型ゲーム(比較的簡単)
チェス、将棋、囲碁等
- 三人以上や情報非公開型ゲーム(より難しい)
カードゲーム、麻雀等

✓ ゲームはなぜ面白いのか

- 明確なルールが存在する
プログラム・アルゴリズムの性能を評価しやすい
- さまざまな難易度のゲームがある
- ゲーム自体が面白く興味を引きやすい

ゲームは「人間の知的活動をコンピュータで実現させる」
人工知能研究の良い題材

(簡単と言われている)2人・完全情報型ゲーム

	探索空間	現在の進捗状況
チェッカー	10^{30}	解かれた(引き分け)
オセロ	10^{60}	コンピューターが勝つ
チェス	10^{100}	コンピューターが勝つ
将棋	10^{200}	アマ6段以上
囲碁	10^{300}	アマ初段以上

2012年にもなって、今だにオセロすら解かれていない。

コンピュータ将棋はいつ人間を追い越すのか

コンピュータ将棋対人間

- 2007 Bonanza 対渡辺明竜王
 - ✓ コンピュータ側: Intel Xeon 2.66GHz 8 core
 - ✓ 人間側: 現在も竜王タイトルを保持
 - ✓ コンピュータ敗北
- 2010 あから対清水市代女流王将
 - ✓ コンピュータ側: 約200台の計算機使用
 - ✓ 人間側: 通算タイトル獲得数歴代1位
 - ✓ コンピュータ勝利
- 2012 ボンクラーズ対米長邦雄永世棋聖
 - ✓ コンピュータ側: 伊藤英紀氏(富士通)開発
 - ✓ 人間側: 元プロ棋士(68歳)
 - ✓ コンピュータ勝利

コンピュータはトッププレイヤーに未だ勝利していない

あから2010について

合議法の利用

- 約200台の計算機を使用
- 分散並列探索法＋合議法
4異種プログラム (Gekisashi, GPS Shogi, Bonanza, YSS) で多数決

表: 多数決による性能の向上。勝率は一手3秒 1,000局より計算

Player	勝率 (%)
多数決合議	73
Gekisashi	50
GPS Shogi	36
Bonanza	35
YSS	37

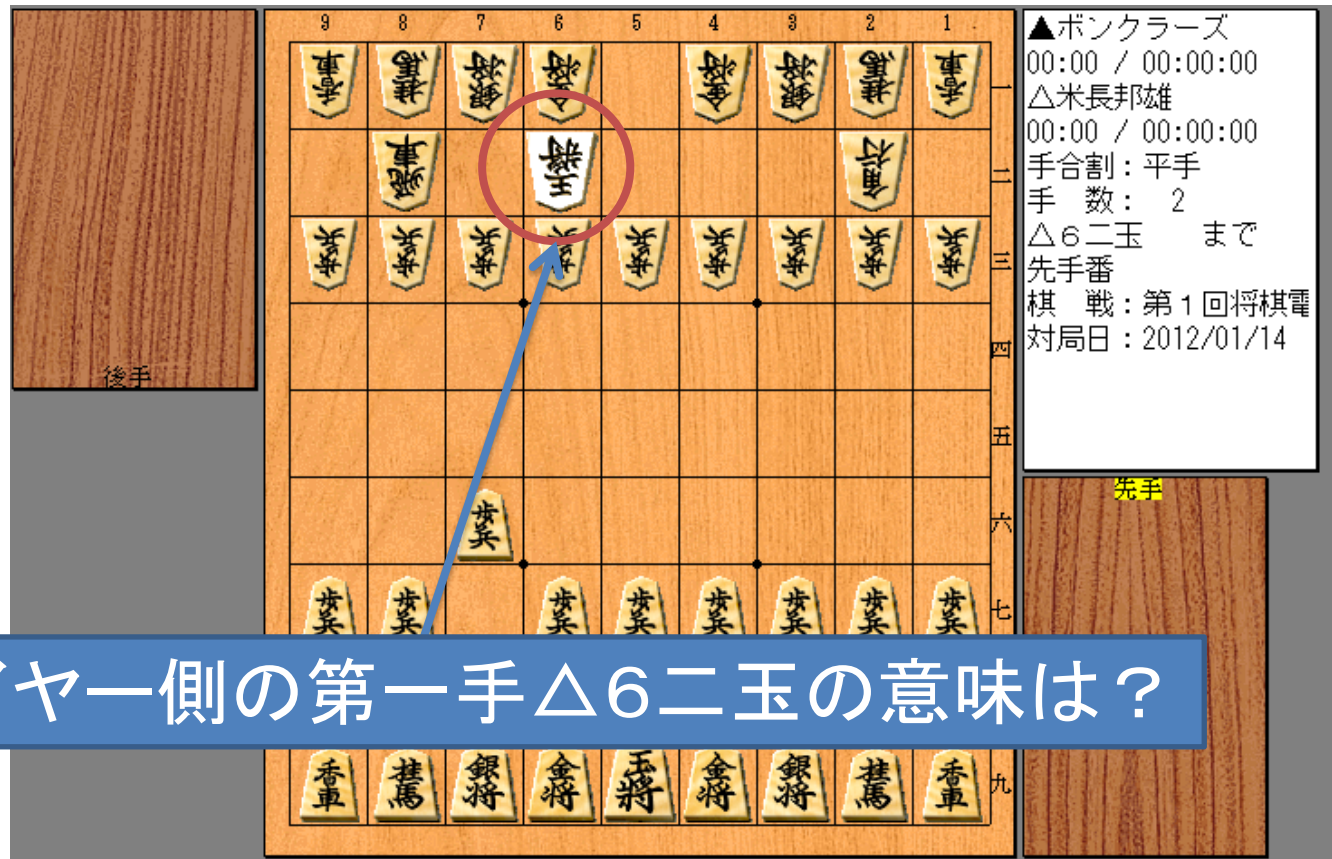
あから2010は清水女流王将に勝利した



ボンクラーズ対米長邦雄永世棋聖 (2012)

公式戦で初めて人間が対コンピュータ戦略をとる

- ボンクラーズは2011年コンピュータ将棋選手権で優勝
- Bonanza のソースコードを基にして作成された(といわれる)



異種格闘戦, 東京, 1976

レスリング (アントニオ猪木)
キックが得意

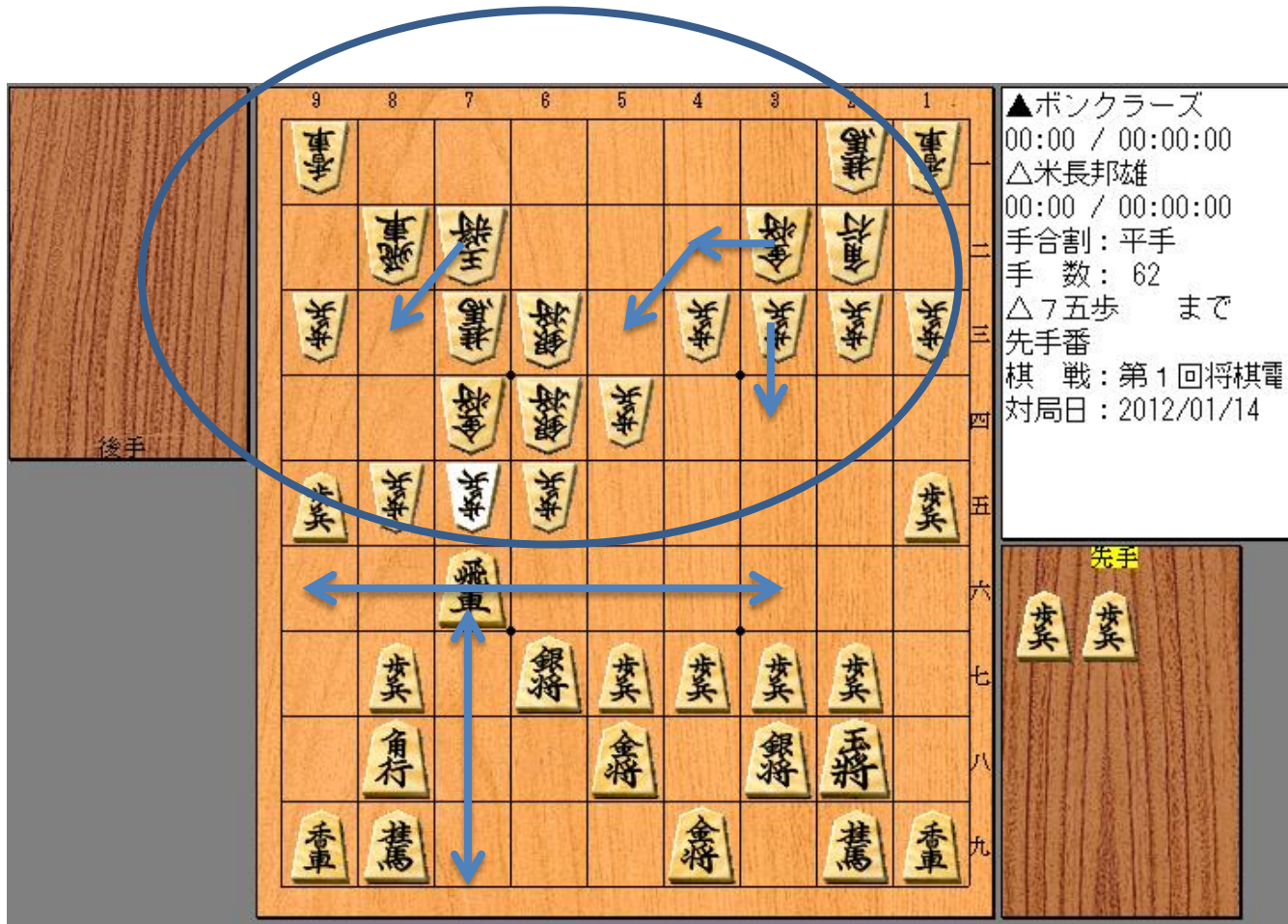


ボクシング (モハメド・アリ)
パンチが得意



図: アントニオ猪木は1ラウンド、ほとんど寝転がった

15ラウンド(最終ラウンド)まで決着つかず、引き分け



- 怪しげな駒の運びでインファイトを回避、防衛ラインを築く
- コンピュータは飛車を往復させて手待ちの繰り返し
- 人間側は引き分けにする権利を得ていたかのように見えたが、その後接近戦になった

コンピュータの勝利

コンピュータ将棋の主な技術

2002年 実現確率探索（激指）
DFPN（詰将棋）

2006年 評価関数の機械学習（Bonanza）
力づく探索（Bonanza）

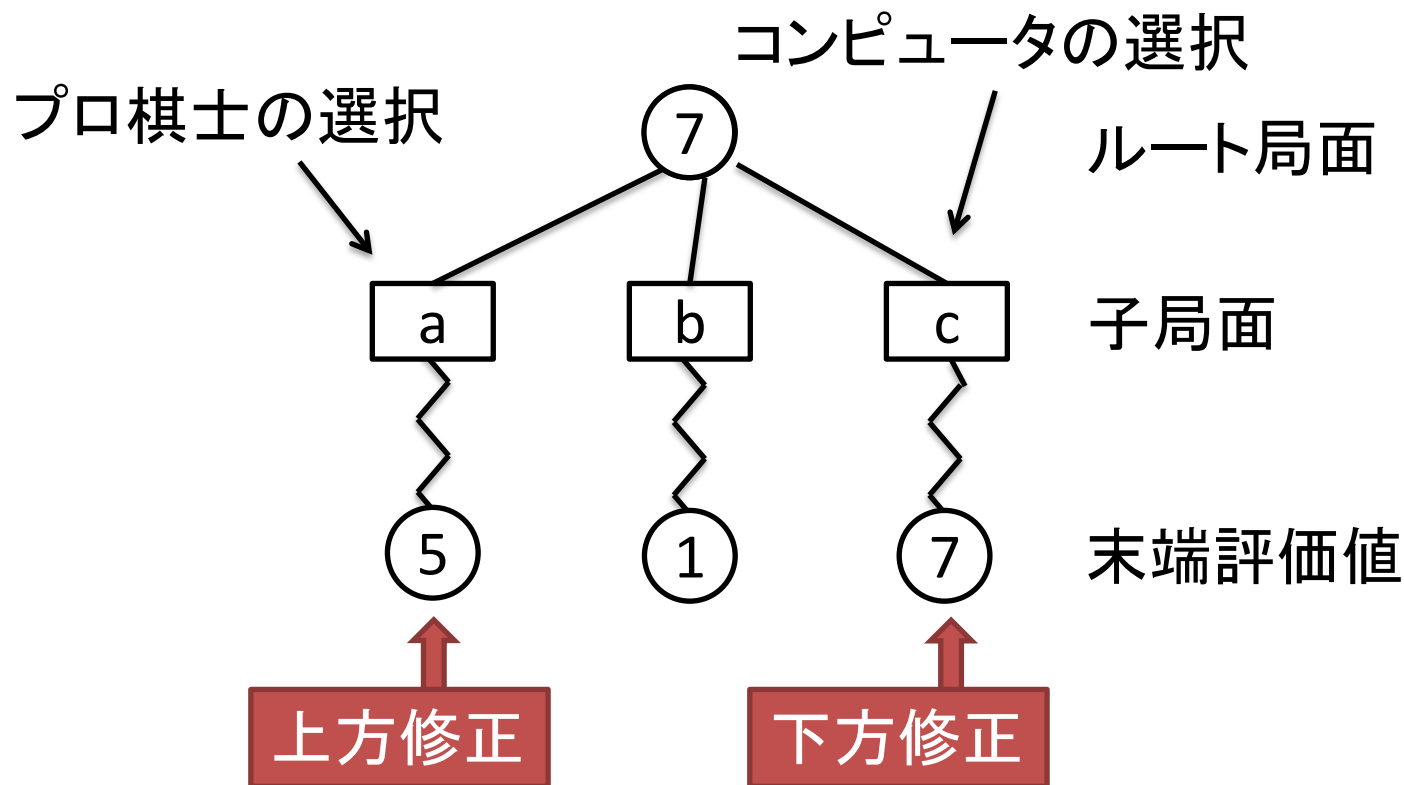
2009年 合議法（文殊）

2010年 分散並列探索の実用化（GPS 将棋）

2006年以降、数の暴力に頼った方法が
将棋でも成功をおさめている

テーマ1：評価関数の機械学習

最適化問題としてMinimax探索ごと評価関数の自動調整を行う



1990年代後半のコンピュータチェスにも同様な試みが見られるが、大規模な機械学習の試みは挑戦なされていなかった。

目的関数の概要

教師データとプログラムとの指し手一致度を測る目的関数 J を設計

$$J'(P_0, P_1, \dots, P_{N-1}, \mathbf{v}) = \sum_{i=0}^{N-1} l(P_i, \mathbf{v}) + \lambda M_1(\mathbf{v}) + M_2(\mathbf{v}) \quad (1)$$

P_i : 棋譜データ中の一局局面

$l(P_i, \mathbf{v})$: 損失関数

$\lambda M_1(\mathbf{v})$: 最適化を安定化させる
束縛条件

$M_2(\mathbf{v})$: 最適化問題の正則化項

$$l(P, \mathbf{v}) = \sum_{m=1}^M T[\xi(p_m, \mathbf{v}) - \xi(p_{m=0}, \mathbf{v})] \quad (2)$$

p_m : 局面 P を合法手 m で進めた子局面

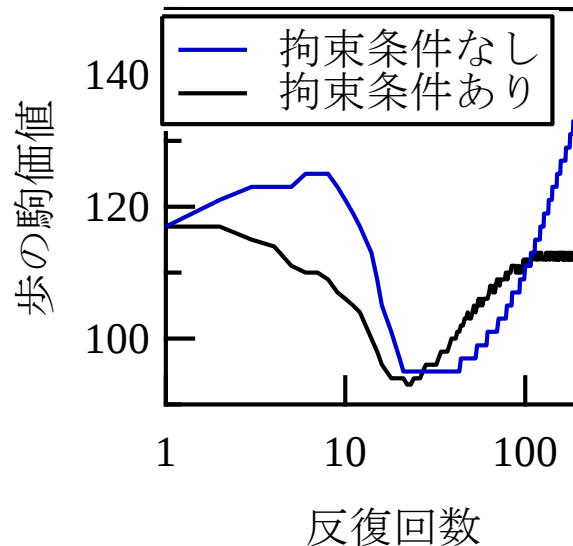
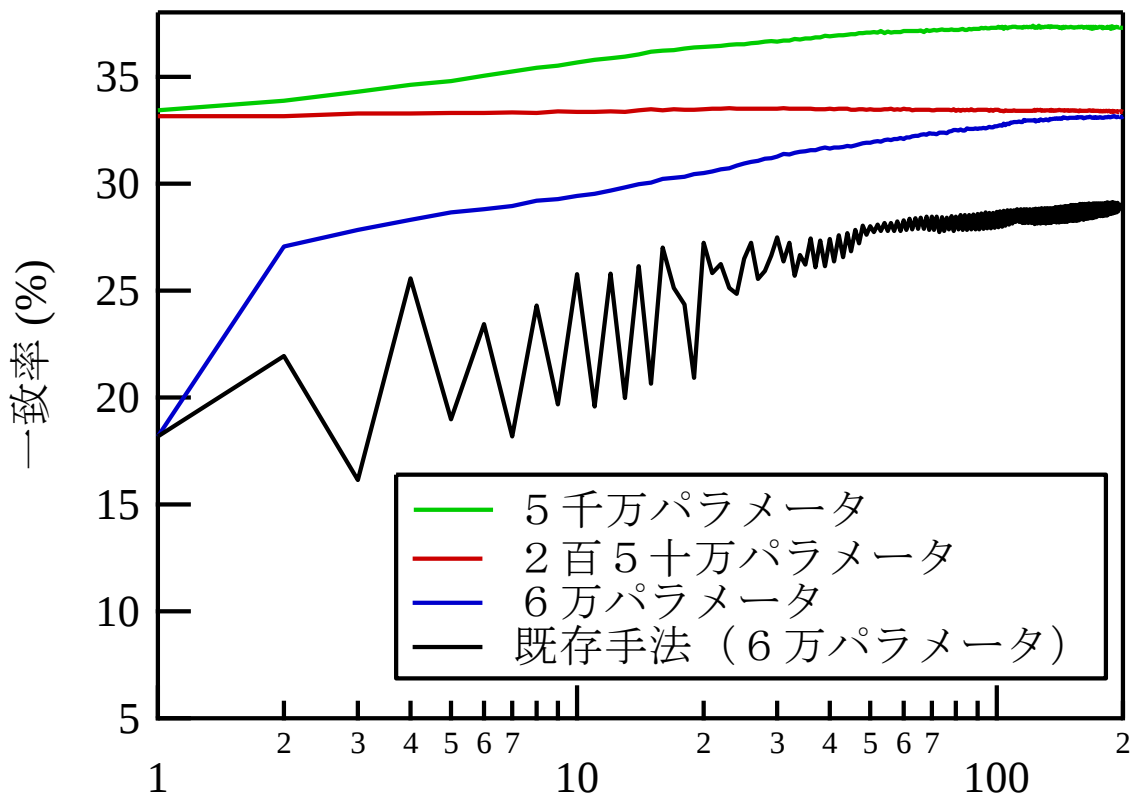
M : 局面 P での合法手の数

$m = 0$: 棋譜中で実際に指された手

$\xi(p_m, \mathbf{v})$: Minimax 探索の値

$T(x)$: 評価値の差を指し手一致度に変換する関数

大規模機械学習の将棋での試み

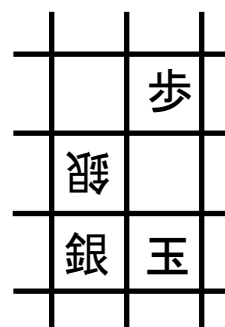


拘束条件がないと収束が困難

大規模な機械学習が安定して行われる



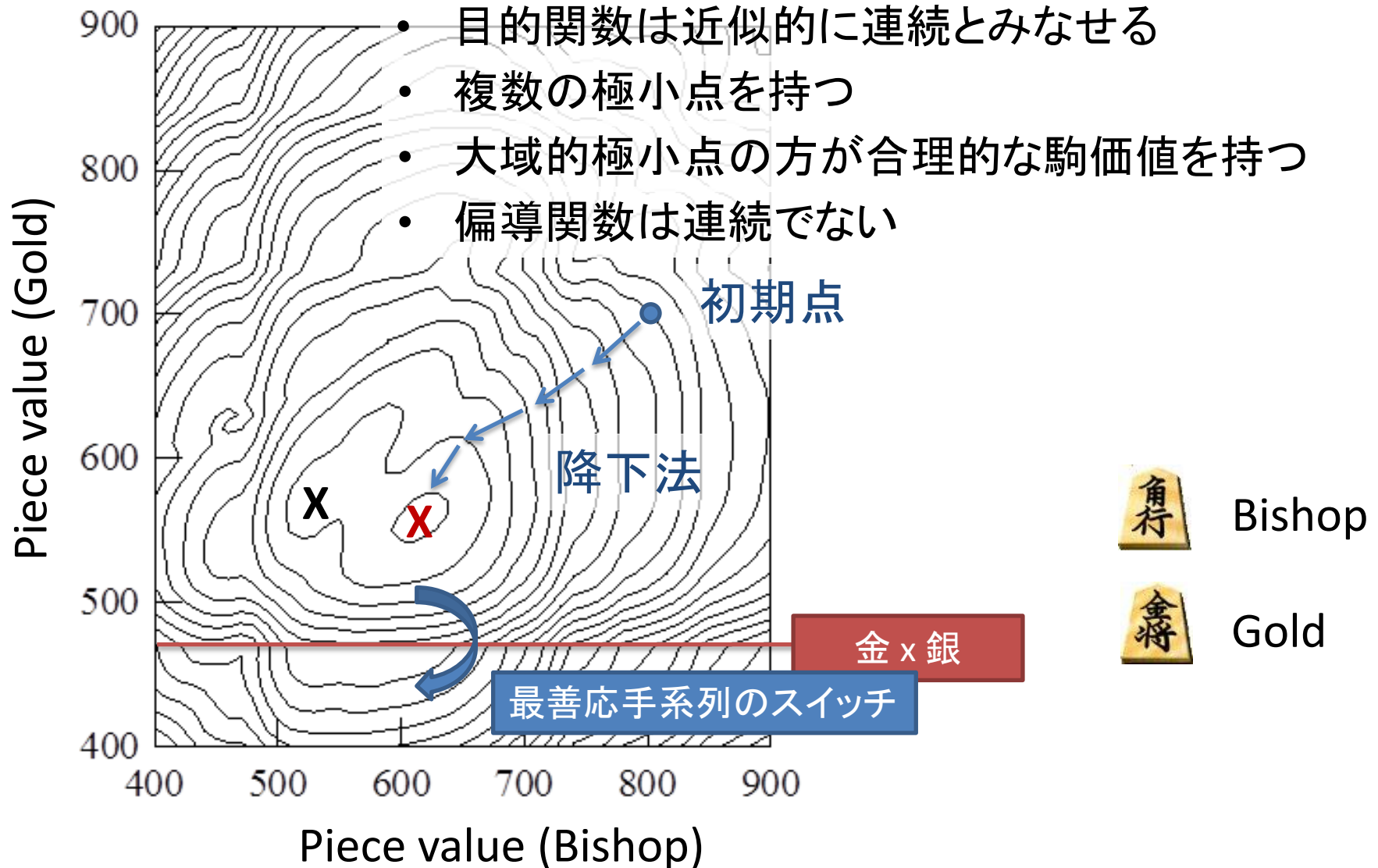
反復回数



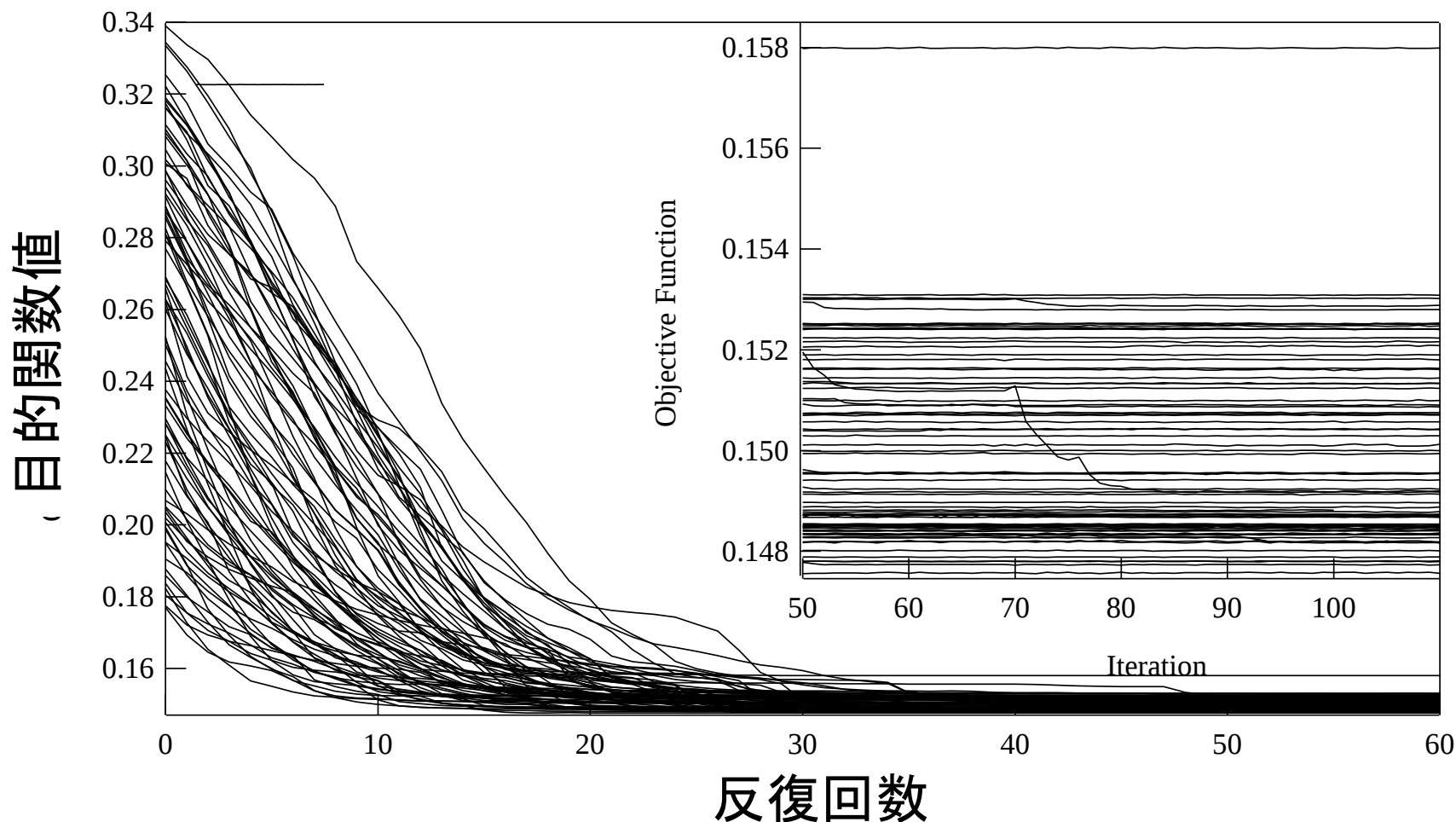
2次元モデルによる解析(駒価値学習)

目的関数の性質:

- 目的関数は近似的に連続とみなせる
- 複数の極小点を持つ
- 大域的極小点の方が合理的な駒価値を持つ
- 偏導関数は連続でない



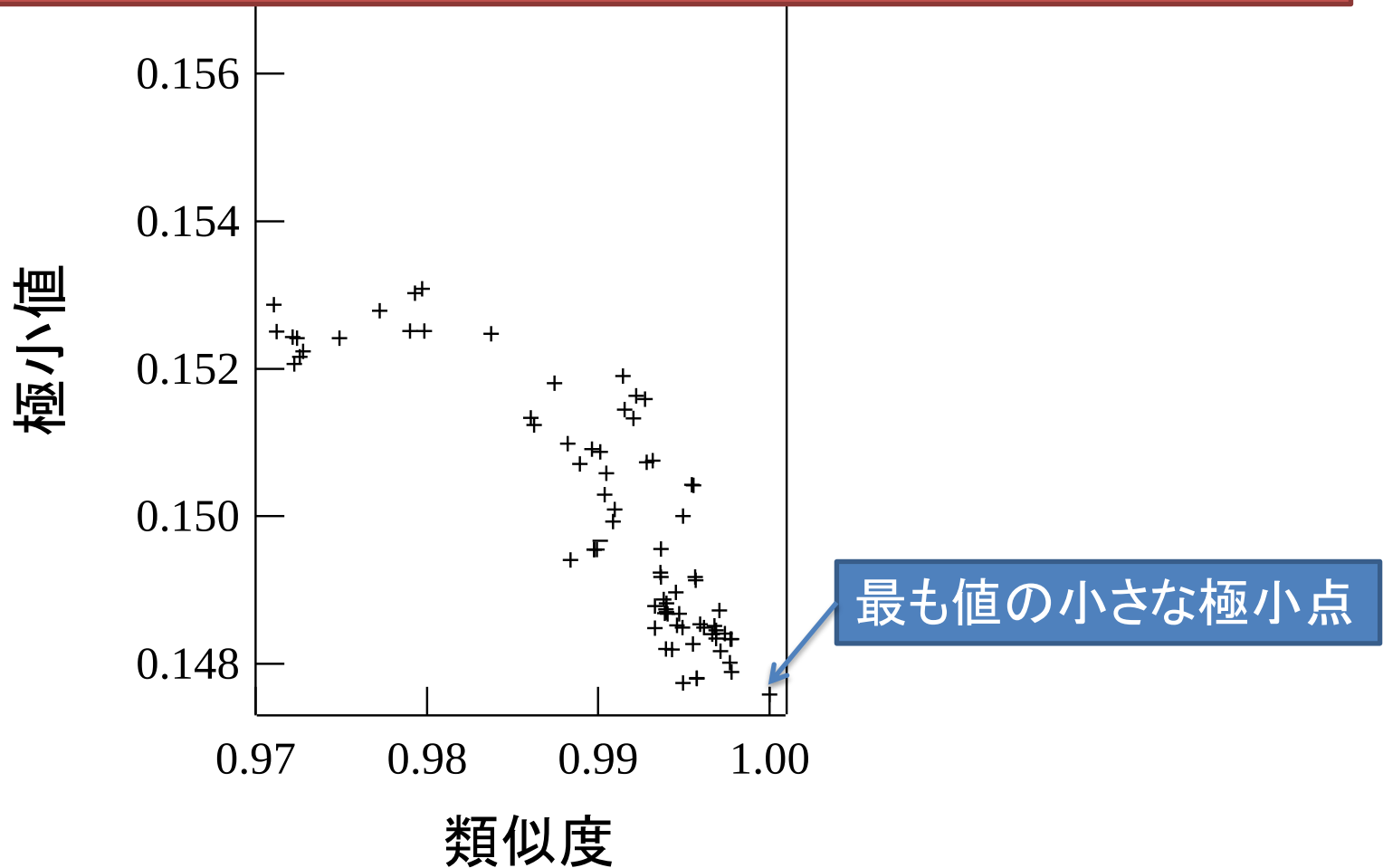
大域的収束性に関する数値的テスト(13次元)



- 全初期点(78個)で数値的収束が得られた
- 多数の極小点を持つ

極小点の解析

目的関数はずり鉢状、良い極小点は局在している。



テーマ2: 力づく探索の効率改善

コンピュータチェスの前向き枝刈法の性能評価

チェス: Crafty 使用

分岐因子の大きいチェス序盤の 2,014 局面で結果を平均

将棋: Bonanza 使用

分岐因子の大きい将棋終盤の 101 局面で結果を平均

表: 有効分岐因子 (枝刈り率)

	Chess Openings	Shogi Endgames
前向き枝刈なし	4.3 (100%)	6.7 (100%)
①Futility 枝刈	3.8 (88%)	5.1 (76%)
②Null move 枝刈	2.8 (65%)	4.5 (67%)
③LMR 法	2.4 (56%)	3.6 (54%)
①+②+③	2.1 (49%)	2.8 (42%)

合法手数が多い将棋の方が、チェスより多く枝を刈れる

探索精度の評価

Chess: Crafty

次の一手問題 2,180 題

Shogi: Bonanza

次の一手問題 7,607 題

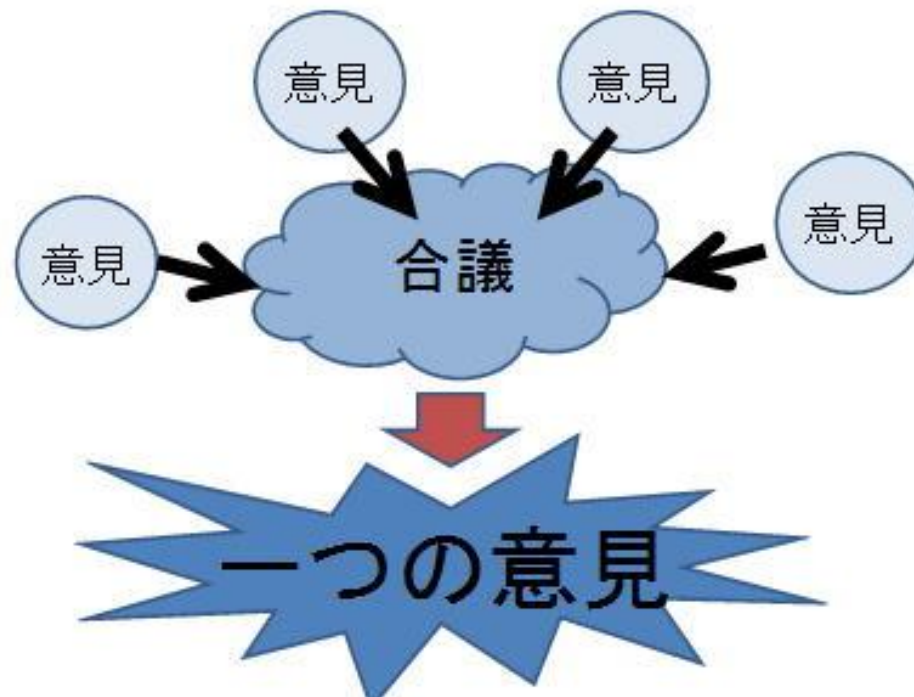
表: 次の一手問題正答率 (探索節点数 $\div 10^3$)

Standard depth	Chess (sd=10)	Shogi (sd=6)
前向き枝刈なし	81.3% (162,144)	43.8% (84,850)
①Futility 枝刈	78.0% (33,134)	43.3% (4,806)
②Null move 枝刈	80.1% (5,118)	43.6% (14,837)
③LMR 法	73.9% (2,350)	41.7% (5,807)
①+②+③	70.5% (507)	42.2% (507)

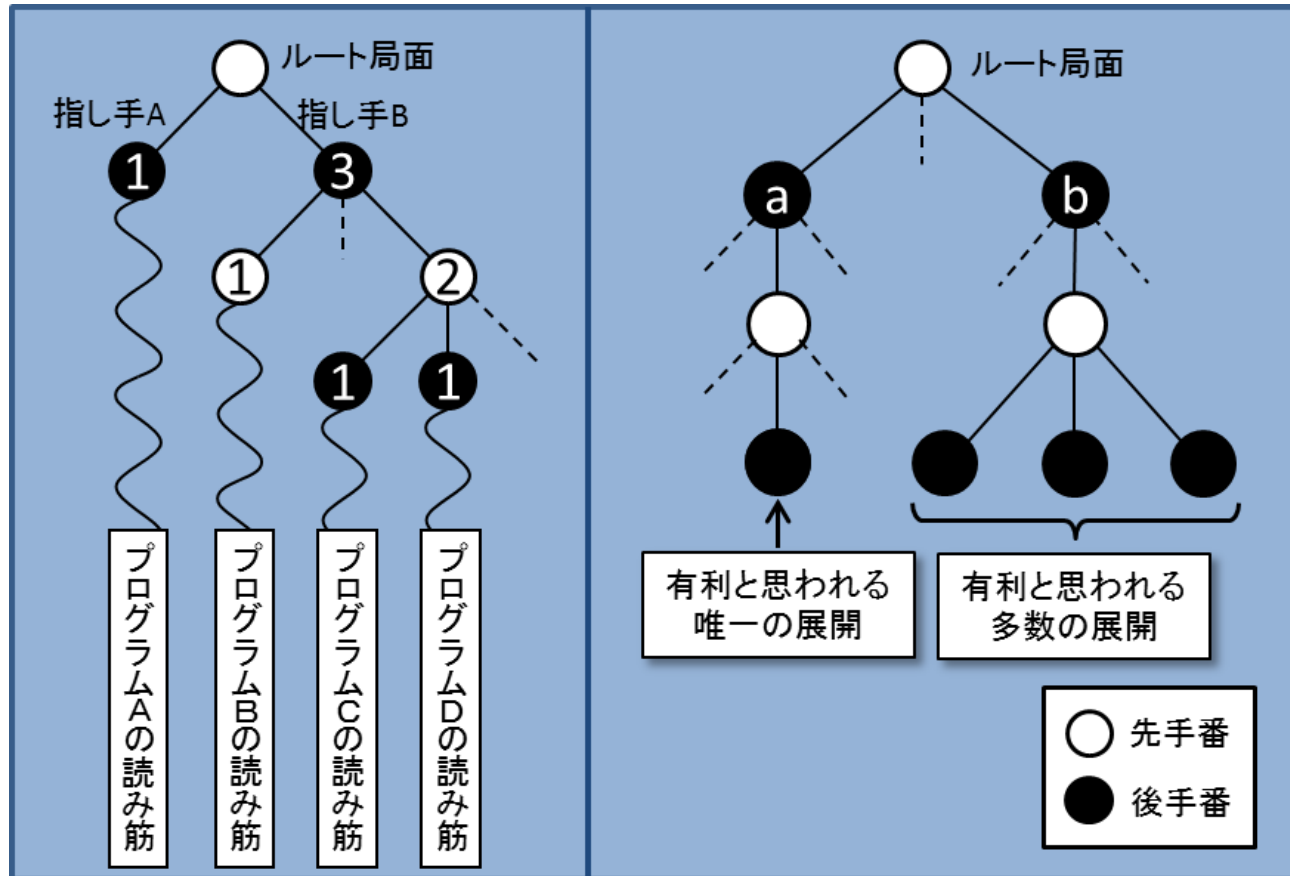
- 将棋では前向き枝刈により探索精度はさほど減少しない
- チェスの前向き枝刈の方法は、将棋ではより威力を発揮

テーマ3: 合議法

- フェイルセーフな分散並列環境の構築
- 複数プログラムの寄せ集めで強い人工知能作成



Minimax探索を行うプログラムの合議



合議法によって、ミニマックス探索の結果が安定化されるのではないか？

まとめ

表：今年のコンピュータ将棋選手権
結果

順位	
1	GPS 将棋
2	Puella α
3	ツツカナ
4	Ponanza
5	習甦
6	激指
7	YSS
8	Blunder

大量のデータを許容できる時間内に
できるだけ沢山処理する技術

- ・局面の深く広い探索
- ・大規模機械学習
- ・分散並列化

今年のコンピュータ将棋選手権では、渡辺竜王を苦しめたと言われている Bonanzaより強いプログラムが8個もあった！

トッププロにもう少しで追いつきそう