

コンピュータ将棋の技術と展望

囲碁将棋から学ぶゲーム情報学
公開講座
保木邦仁
2012年12月8日

自己紹介

- ・名前 保木邦仁 (生まれ 北海道東区)
- ・年齢 36
- ・職業 電気通信大学 特任助教
- ・専門 2010年頃まで化学, 以降ゲーム情報学
- ・コンピュータ将棋プログラム Bonanza を作っています

内容

- ・将棋と関係するゲーム理論概略
- ・チェス・将棋の思考アルゴリズム
- ・コンピュータ将棋対人間の歴史

将棋と関係するゲーム理論概略

ゲームの完全解明 (神の一手?) は
究極的な目標の一つ

囲碁・将棋のようなゲームで

フォン・ノイマン
(ミニマックス定理)

- ・ゲーム値(勝ち・負け・引き分け)
- ・最適な戦略

とはどのようなものなのだろうか。

ジョン・ナッシュ
(ナッシュ均衡)

戦略形ゲーム—支配戦略と支配戦略均衡

できるだけ売上を増やしたい店長ゲーム

- ・他店と競争しなければならない
- ・過去のデータから、値段設定に対する売上は大体想像可能
 - 低価格で沢山売れる
 - 客は安い方の店から商品を買う

| 自店 \ 他店 | 7 | 9 | 11 |
|---------|-----------|-----------|-----------|
| 7 | (900,900) | (1800,0) | (1800,0) |
| 9 | (0,1800) | (800,800) | (1600,0) |
| 11 | (0,1800) | (0,1600) | (700,700) |

相手プレイヤーの行動基準がどうであろうとも
支配戦略(7)をとるのが良い

パレート最適性(囚人のジレンマ)

支配戦略のみで説明できない場合(その1)

| 自店 \ 他店 | 通常営業 | 一斉値下げ |
|---------|--------------|--------------|
| 通常営業 | (+1千万, +1千万) | (倒産, 2千万) |
| 一斉値下げ | (2千万, 倒産) | (-1千万, -1千万) |

支配戦略均衡: (一斉値下げ, 一斉値下げ)?

パレート最適: (通常営業, 通常営業)

ゲームの性質によっては
何が最善なのかははっきりしない場合がある

ナッシュ均衡(最適反応)

将棋で重要!

支配戦略のみで説明できない場合(その2)

| | | |
|-------------------|-------------------|-------------------|
| 2 \ 1 | 戦略 A ₁ | 戦略 B ₁ |
| 戦略 A ₂ | (2,2) → (0,0) | (0,0) |
| 戦略 B ₂ | (0,0) ← (1,1) | (1,1) |

支配戦略均衡: 無し
ナッシュ均衡: (2,2) と (1,1)

支配戦略均衡よりも適応範囲が広い

ナッシュ均衡の良い性質

- 各プレイヤーは戦略変更の積極的な理由がない
- 支配戦略均衡はナッシュ均衡

先ほどの支配戦略均衡の例

| | | | |
|---------|-----------|-----------|-----------|
| 自店 \ 他店 | 7 | 9 | 11 |
| 7 | (900,900) | (1800,0) | (1800,0) |
| 9 | (0,1800) | (800,800) | (1600,0) |
| 11 | (0,1800) | (0,1600) | (700,700) |

- ナッシュ均衡戦略を支配する戦略はない

ナッシュ均衡の良くない性質1

非合理的なプレイヤーに対する不安

| | | |
|-------------------|-------------------|-------------------|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ |
| 戦略 A ₁ | (2,2) → (0,1) | (0,1) |
| 戦略 B ₁ | (0,1) ← (1,0) | (1,0) |

戦略の組(A₁, A₂)が唯一のナッシュ均衡

プレイヤー2が戦略Bを選んでしまった場合にプレイヤー1も戦略Bを選べばよかったと後悔

ナッシュ均衡の良くない性質2 チキンレース

| | | |
|----------|------------|-----------|
| ジョン \ ジム | ハンドル切る | ハンドル切らない |
| ハンドル切る | (チキン, チキン) | (チキン, 勝ち) |
| ハンドル切らない | (勝ち, チキン) | (死亡, 死亡) |

戦略の組(切る, 切らない)と(切らない, 切る)はナッシュ均衡

相手がどっちの均衡を目指すのか不明な場合
ナッシュ均衡は戦略決定の指針とならない

2人ゼロ和ゲーム

将棋で重要!

利得の和がゼロ

| | | |
|-------------------|-------------------|-------------------|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ |
| 戦略 A ₁ | (1,-1) | (0,0) |
| 戦略 B ₁ | (0,0) | (-1,1) |

以下のように簡略化して利得行列を書く

| | | |
|-------------------|-------------------|-------------------|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ |
| 戦略 A ₁ | 1 | 0 |
| 戦略 B ₁ | 0 | -1 |

ゼロ和の場合の ナッシュ均衡の更に良い性質1

| | | |
|-------------------|-------------------|-------------------|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ |
| 戦略 A ₁ | 0 → 5 | 5 |
| 戦略 B ₁ | -5 ← 0 | 0 |

他のプレイヤーが非合理的な戦略を選んでも
自分の利得が減少することはない

ゼロ和の場合の ナッシュ均衡の更に良い性質2

| | | | |
|-------------------|-------------------|-------------------|-------------------|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ | 戦略 C ₂ |
| 戦略 A ₁ | 0 | 1 | 1 |
| 戦略 B ₁ | -1 | -1 | -1 |
| 戦略 C ₁ | 0 | 0 | 0 |

複数の戦略の組(A1, A2)と(C1, A2)はナッシュ均衡を形成
均衡戦略を交換した組もまた均衡を形成し利得が等しい

ミニマックスとマックスミニ戦略

保証水準を最大にする戦略

| | | | | |
|-------------------|-------------------|-------------------|-------------------|----|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ | 戦略 C ₂ | |
| 戦略 A ₁ | 0 | 1 | -6 | -6 |
| 戦略 B ₁ | -1 | 0 | 3 | -1 |
| 戦略 C ₁ | 6 | -3 | 0 | -3 |
| | 6 | 1 | 3 | |

- 一般にマックスミニ値 ≤ ミニマックス値
- プレイヤー1はミニマックス値を狙うと戦略 B
- プレイヤー2がマックスミニ値を狙うと予想すると戦略 A

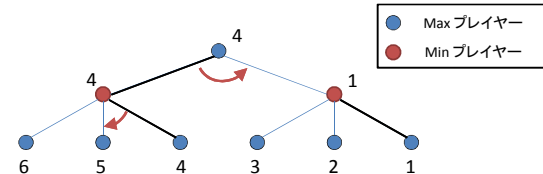
ゼロ和の場合の ナッシュ均衡の更に良い性質3

| | | | | |
|-------------------|-------------------|-------------------|-------------------|----|
| 1 \ 2 | 戦略 A ₂ | 戦略 B ₂ | 戦略 C ₂ | |
| 戦略 A ₁ | 0 | 1 | -6 | -6 |
| 戦略 B ₁ | 1 | 1 | 3 | 1 |
| 戦略 C ₁ | 6 | -3 | 0 | -3 |
| | 6 | 1 | 3 | |

将棋で重要!

- マックスミニ値とミニマックス値が一致
- マックスミニ戦略とミニマックス戦略は均衡点を形成

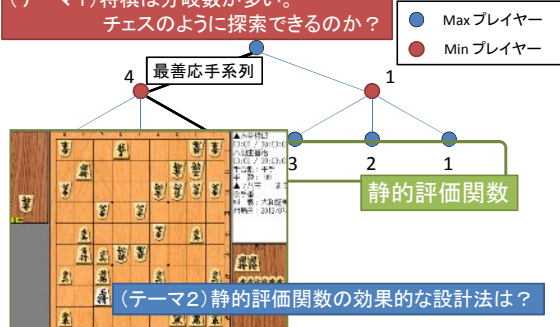
展開型ゲームの良い性質 将棋で重要!



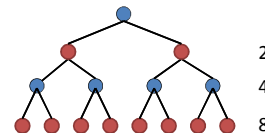
- 展開型ゲームは標準型ゲームに置き換えることが可能
- ナッシュ均衡戦略を再帰的に求めることが可能
- ミニマックス値(4)がこのようなゲームの解と考えられる
 - 最適反応戦略
 - 不合理なプレイヤーに対しても損をしない
 - マックスミニ値と等しい
 - どの均衡戦略が複数あっても値は同じ
 - 他の戦略に支配されない

チェス・将棋の思考アルゴリズム

(テーマ1) 将棋は分岐数が多い。
チェスのように探索できるのか?

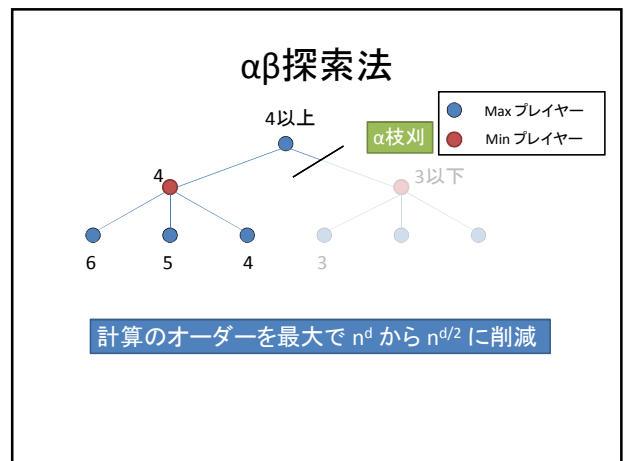
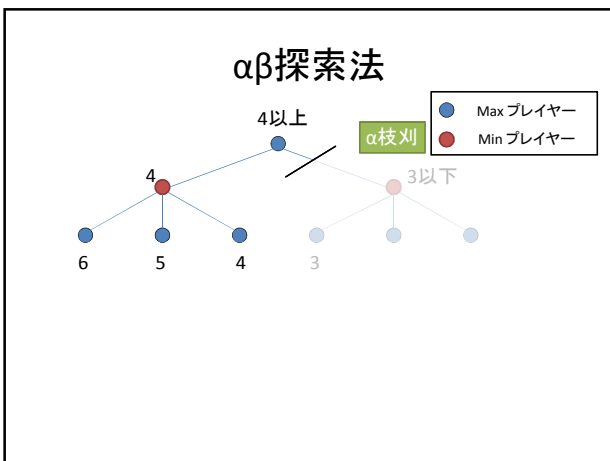
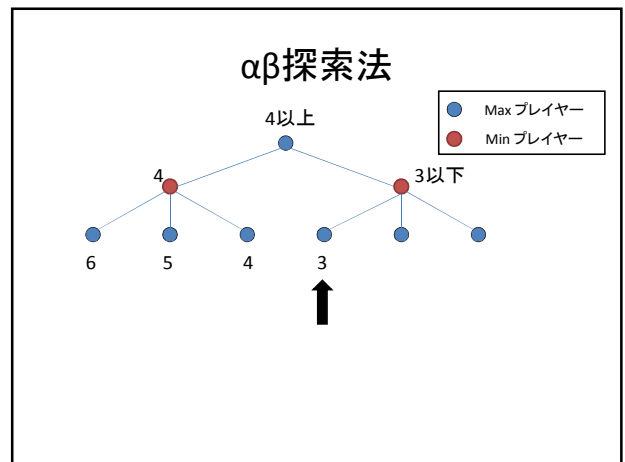
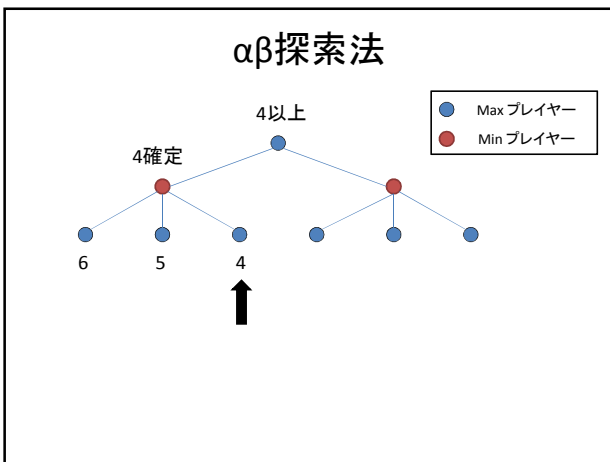
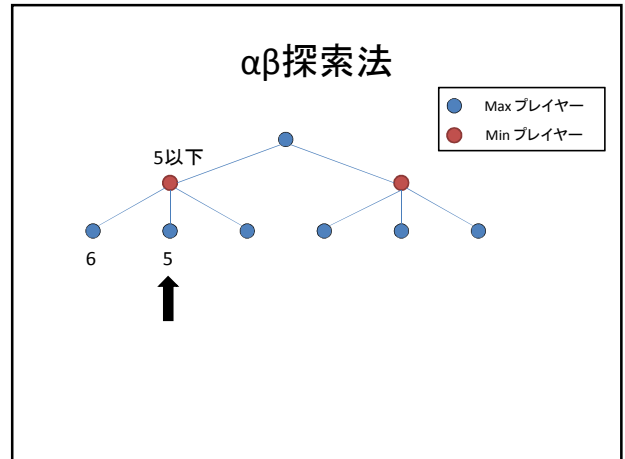
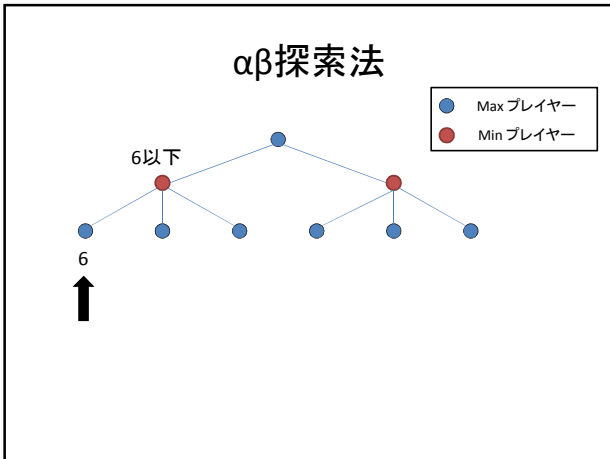


かづく探索の効率改善

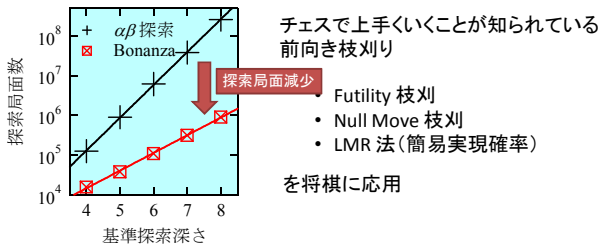


将棋の合法手数を持ち駒ルールのため平均 80手
末端局面数は 80^d (d は探索深さ)

- 枝刈によって計算量を削減
- αβ 枝刈
 - 前向き枝刈



将棋ゲーム木の前向き枝刈り



チェスで上手くいくことが知られている前向き枝刈り

- Futility 枝刈り
- Null Move 枝刈り
- LMR 法 (簡易実現確率)

を将棋に応用

- 1秒程度の時間で、深さ8の全幅探索相当の計算が可能
- これはコンピュータの長所で、人間にはとても無理

将棋の局面評価法

局面の良し悪しを“適当に”見積もる関数
ゲーム中の局面の特徴を、重みを付けて足し合わせる

- チェス：駒割り・機動性・中央制圧度
- オセロ：合法手の数・辺、隅の形
- 将棋：局面の評価が大変困難といわれていた。

2005年ごろから評価関数の大規模な自動学習が成功

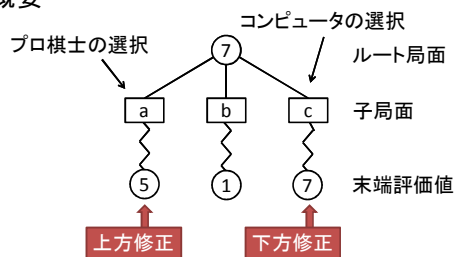
2009年コンピュータ将棋選手権

| 順位 | プログラム名 |
|----|---------|
| 1 | GPS将棋 |
| 2 | 大槻将棋 |
| 3 | 文殊 |
| 4 | KCC将棋 |
| 5 | Bonanza |

1位から5位まで、この自動学習法を採用
コンピュータが一層強くなった。

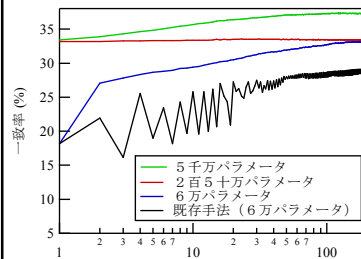
評価関数の教師付き機械学習

概要



性質の良い目的関数を設計して
ミニマックス探索ごと自動調整

大規模機械学習の将棋での試み



大規模な機械学習が安定して行われる



現在の機械学習の問題点

- 人間熟達者の棋譜から学習
 - 人間を超えることができるのか？
- 棋譜に表れにくい状況
 - 入玉型
 - 不思議で怪しい駒組み

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: 現役時代トッププレイヤー
 - ✓ 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 |

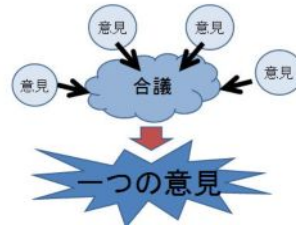
あから2010は清水女流王将に勝利した



T. Obata, T. Sugiyama, K. Hoki, T. Ito, CG2010

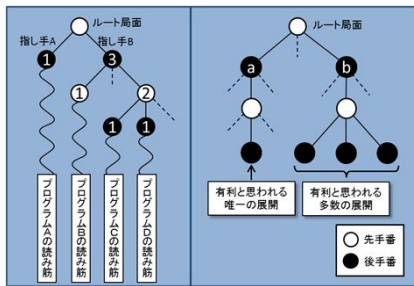
合議法について

- 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



電通大 伊藤毅志助教との共同研究

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な？

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 のソースコードを参考にして作成された(といわれる)



人間プレイヤー側の第一手△6二玉の意味は？

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

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 | 習経 |
| 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予選敗退！
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て追いつきそう