

# ゲームプログラミング コンピュータが将棋を指す仕組み

2008年9月 保木邦仁

## 概要

現在の幅広い研究対象の中の一つとして、ゲームがあげられる。西洋では 1950 年頃からチェスを指す計算機に関する研究報告がある。ゲームは人工知能研究におけるモデル系として都合の良いこともあるが、強いプログラムを作成する事への純粋な知的好奇心も研究者達の興味をそそった一因のようである。本講演では、コンピュータが将棋型ゲームで思考する仕組みについて理解り易く概説する。

# 自己紹介

## コンピュータ将棋プログラム Bonanza

- Bonanza 製品版を販売
- 大和証券杯特別対局  
新聞一面カラー  
全国版テレビニュース
- NHK衛星第2「運命の一手」  
1時間に及ぶドキュメンタリー
- 角川書店新書  
ボナンザVS勝負脳



一般の方々にアピール

何故こんな事になってしまったのだろう。

- 強い将棋プログラムを無料で公開
- プログラムの動作がユニーク

将棋の指し方は  
殆どしりません

# ゲームプログラミング

オセロ・将棋等の思考型ゲーム

- ルールが明確
- 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

「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

ゲームプログラムと密接な関係がある認知科学，情報科学研究分野

- 機械学習
- グラフ探索
- ニューラルネットワーク

# ゲームの場合の数

- チェッカー            10の30乗            コンピュータ
- オセロ                10の60乗            コンピュータ
- チェス                10の120乗           コンピュータ
- 将棋                    10の220乗           アマ5段強
- 囲碁                    10の360乗           アマ1級

必勝法が求まるのはチェッカーまで

オセロ・チェスは力任せの探索(全幅探索)と、  
簡単な評価関数で人間よりは強くなる

将棋や囲碁はどうすればよいのだろうか？

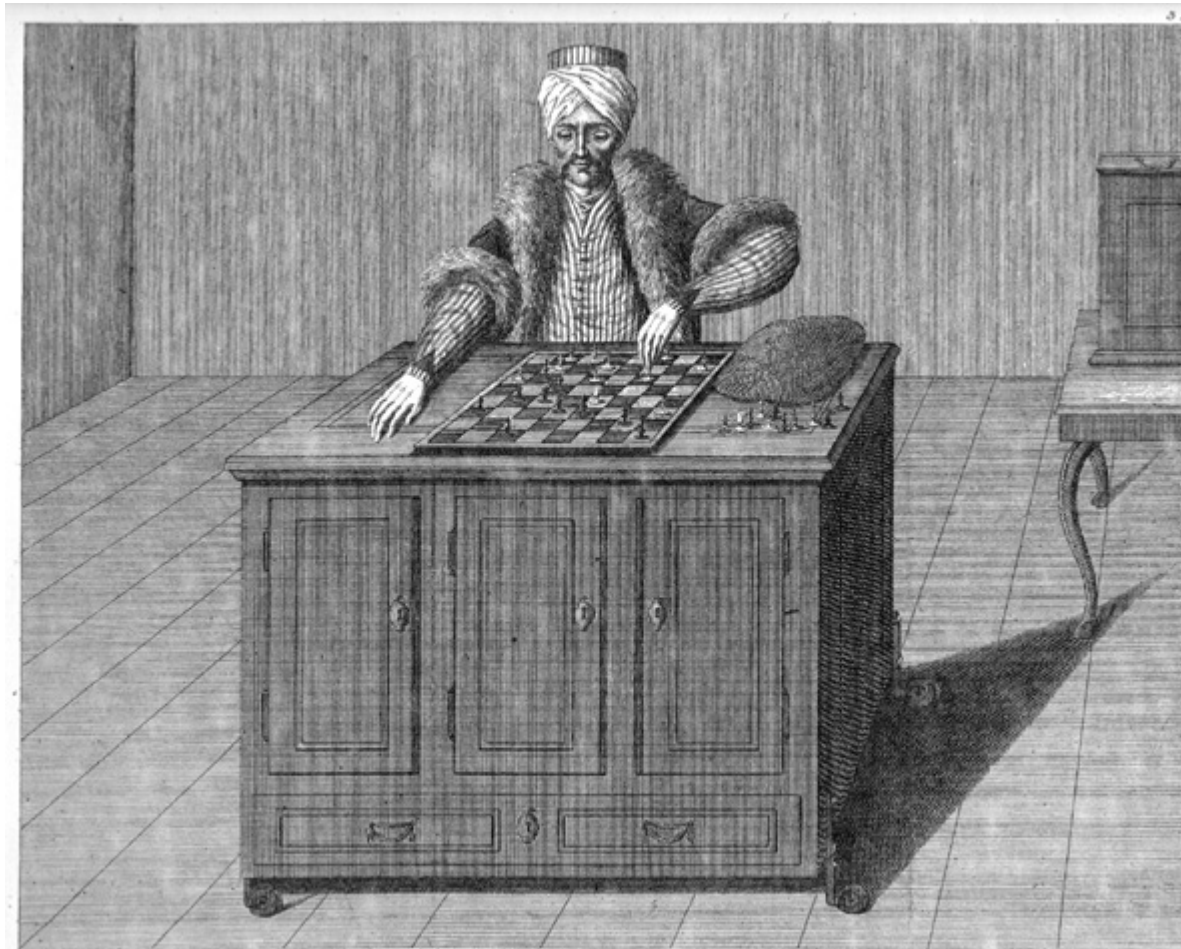
人間にはいつ勝てるのか？

# コンピュータチェスの歴史

- 古来から西洋では、チェスは知性の象徴と考えられていた
- 人工知能研究やコンピュータ技術発展の初期の頃(1950年)から熱心に研究される
- 1997年、IBMのDeep Blueが世界チャンピオンに勝利

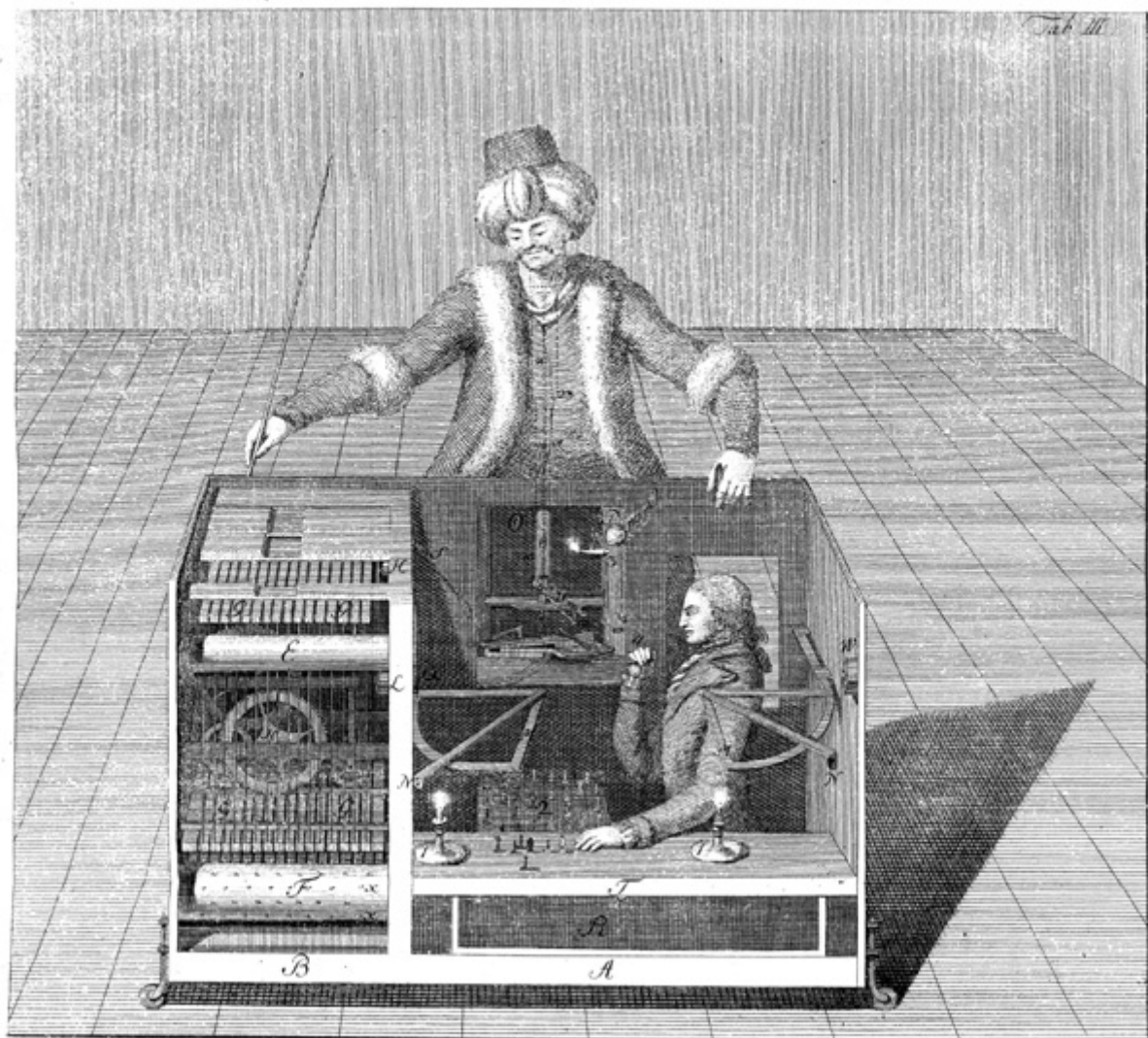


# The Turk (1770)

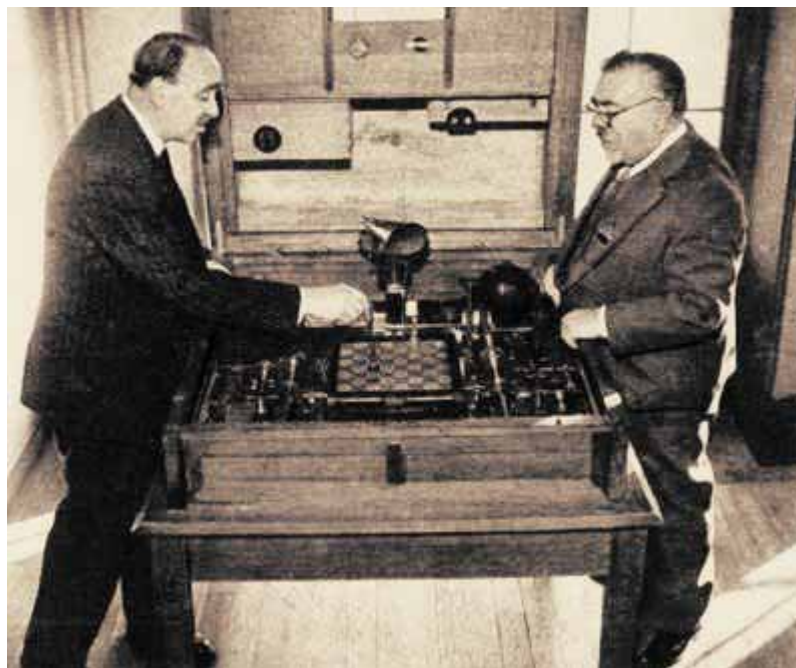


*W. de Kempelen del. Chr. à Machel excud. Basilea. P. G. Ritz. sc.*  
*Der Schachspieler im Spiele begriffen. | Le Joueur d'Échecs tel qu'on le voit pendant le jeu.*





# El Ajedrecista (1912)



# Deep Blue (1996)



# オセロ・将棋等のゲームでの思考法

ゲームの解(ゲーム値・最適戦略)を求めることが一つの目標

オセロ・将棋のようなゲームの:

- ゲーム値(勝ち・負け・引き分け)の意味
- 最適戦略とは何なのか

# 戦略形ゲーム

## 支配戦略と支配戦略均衡

- 2人の店主が原価10の商品を売る
- 低価格で沢山売れる
- 客は安い方の店から商品を買う

1 \ 2	11	15	20
11	(90,90)	(180,0)	(180,0)
15	(0,180)	(80,80)	(160,0)
20	(0,180)	(0,160)	(70,70)

相手プレイヤーの行動基準がどうであろうとも、  
支配戦略をとる方が有利なのは明らか

# パレート最適性と支配戦略均衡 (電気屋の価格競争)

ドンキ \ コジマ	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、なにを基準に  
「最適な戦略」とするか不明瞭な場合がある

# ナッシュ均衡

他プレイヤーのある戦略に対する最適反応戦略を考える

1 \ 2	11	15	20
11	(9,9)	(18,0)	(18,0)
15	(0,18)	(8,8)	(16,0)
20	(0,18)	(0,16)	(7,7)

- ナッシュ均衡戦略を支配する戦略はない
- 支配戦略均衡はナッシュ均衡
- 戦略の組がナッシュ均衡を形成する場合、各プレイヤーは戦略変更の積極的な理由がない

# ナッシュ均衡の性質 1

## 非合理的なプレイヤーに対する不安

1 \ 2	戦略 A <sub>2</sub>	戦略 B <sub>2</sub>
戦略 A <sub>1</sub>	(5,5)	(0,1)
戦略 B <sub>1</sub>	(0,1)	(1,1)

- 戦略の組  $(A_1, A_2)$  はナッシュ均衡を形成
- プレイヤー 2 が戦略 B<sub>2</sub> を選らんでしまった場合、プレイヤー 1 の利得も減少する。



# ナッシュ均衡の性質2

## チキンレース

ジム \ ジョン	A <sub>2</sub> , ハンドル切る	B <sub>2</sub> , ハンドル切らない
A <sub>1</sub> , ハンドル切る	(チキン, チキン)	(チキン, 勝ち)
B <sub>1</sub> , ハンドル切らない	(勝ち, チキン)	(死亡, 死亡)

戦略の組(A<sub>1</sub>, B<sub>2</sub>)と(B<sub>1</sub>, A<sub>2</sub>)はナッシュ均衡. しかし,  
これを入れ替えた組(A<sub>1</sub>, B<sub>2</sub>)と(B<sub>1</sub>, A<sub>2</sub>)はナッシュ均衡ではない.

このゲームではナッシュ均衡が複数存在し,  
ナッシュ均衡戦略は戦略の指針となっていない

# 2人ゼロ和ゲーム

利得の和が常に一定

1 \ 2	戦略 A <sub>2</sub>	戦略 B <sub>2</sub>
戦略 A <sub>1</sub>	(1,-1)	(0,0)
戦略 B <sub>1</sub>	(0,0)	(-1,1)

以下のように簡略化して利得行列を書くことが多い

1 \ 2	戦略 A <sub>2</sub>	戦略 B <sub>2</sub>
戦略 A <sub>1</sub>	1	0
戦略 B <sub>1</sub>	0	-1

# 鞍点の良い性質 1

2人ゼロ和ゲームのナッシュ均衡は鞍点と呼ばれる

1 \ 2	戦略 A <sub>2</sub>	戦略 B <sub>2</sub>
戦略 A <sub>1</sub>	0	5
戦略 B <sub>1</sub>	-5	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

# 鞍点の良い性質 2

## 鞍点の交換可能性

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( $A_1, A_2$ )と( $C_1, A_2$ )はナッシュ均衡を形成  
ナッシュ均衡戦略を交換した組もナッシュ均衡を形成, 利得が等しい

# マックスミニ戦略 (鞍点が無い場合)

## 保証水準を最大にする戦略

1 \ 2	戦略 A <sub>2</sub>	戦略 B <sub>2</sub>	戦略 C <sub>2</sub>
戦略 A <sub>1</sub>	0	2	-6
戦略 B <sub>1</sub>	1	0	3
戦略 C <sub>1</sub>	6	-3	0

- 一般にマックスミニ値の方が小さいか等しい
- マックスミニ値は0, ミニマックス値が2
- マックスミニ戦略 (B<sub>1</sub>) をとると保証水準 (0) を確実に達成可能
- しかし, ミニマックス値のほうが大きく,  
それを達する戦略を狙わなかったことを後悔するかもしれない

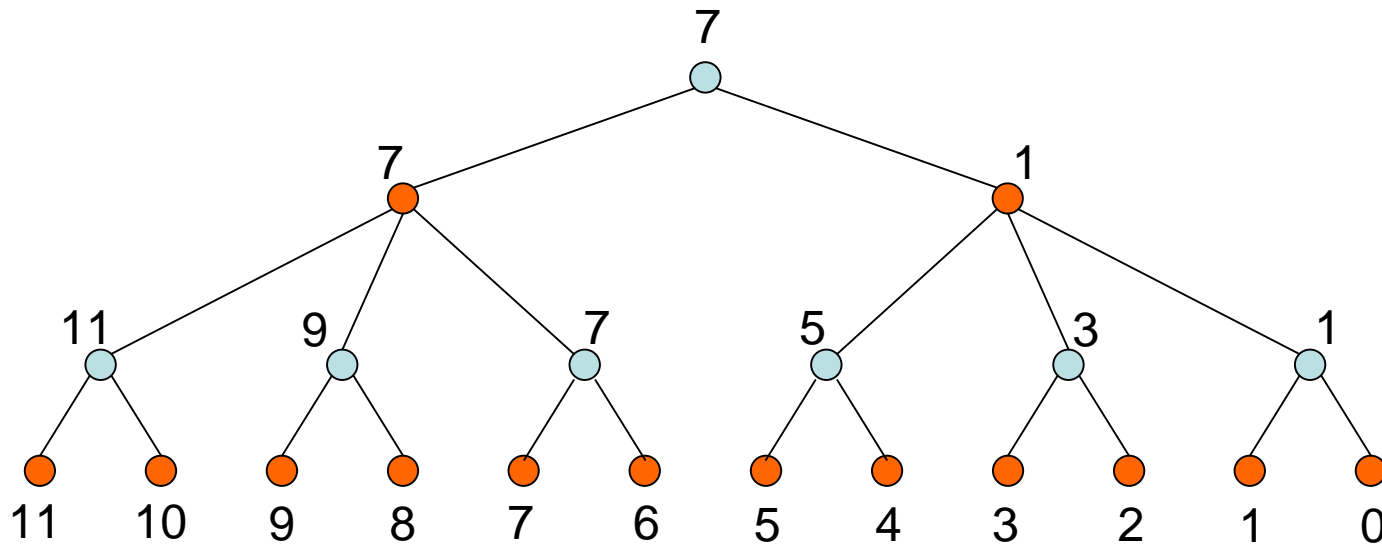
# 鞍点が存在する場合

## 2人ゼロ和ゲームの最適戦略とゲーム値

1 \ 2	戦略 A <sub>2</sub>	戦略 B <sub>2</sub>	戦略 C <sub>2</sub>
戦略 A <sub>1</sub>	0	1	6
戦略 B <sub>1</sub>	4	2	3
戦略 C <sub>1</sub>	5	0	0

- マックスミニ値とミニマックス値が一致(2)
- マックスミニ戦略とミニマックス戦略は鞍点を生成
- マックスミニ戦略とミニマックス戦略が複数存在する場合, どの戦略の組も鞍点を生成し利得は変わらない

# 2人ゼロ和, 展開形ゲーム



- 局所戦略の組み合わせを一つの戦略と考えることにより, 展開形ゲームは戦略形ゲームに置き換えられる.
- 鞍点が存在し, 最適戦略は再帰的に求められる.
- 大抵のゲームでは終局までノードを展開することが不可能. 適当な局面評価を行い, 最適な戦略を近似計算.

# ゲーム木の探索

## 2人ゼロ和ゲームの解を求めるアルゴリズム

- Minimax 探索

探索時間は大体  $n^m$  に比例  
(局所戦略の数  $n$ , 探索の深さ  $m$ )

$N = 80, m = 100$  の場合,  $n^m = 10^{190}$

- Alpha-Beta 探索

探索時間は大体  $nm/2$  に比例

$N = 80, m = 100$  の場合,  $n^m = 10^{95}$



# 評価関数

- 局面の良し悪しを“**適当に**”見積もる関数
  - 精度よりも, 実行速度重視
  - Minimax 探索等とともに用いる
- 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

# 局面評価の機械学習

- TD-Gammon (G. Tesauro)  
思考部の自己学習  
Temporal difference + neuron network
- Logistello (M. Buro)  
辺のパターンの重み学習  
最小二乗法
- GPS 将棋 (金子ら)  
序盤の駒組みを棋譜から学習  
親子, 兄弟モデルを用いた線形回帰
- Bonanza (保木)  
ゲーム木の探索結果を最適制御  
強いプレイヤーと出来るだけ同じ手をさすようにする

# 最適制御理論

最大(小)化問題として力学系の制御方針を推論する

- ラグランジュ形式の解析力学
- パルス整形による化学反応制御
- 最小燃料のロケット弾道
- 工場の消費電力
- 池の魚に与える餌

$$J = \int_0^T l(x, u, t) dt$$

$t$ : 時間に関する数(離散も可)

$x(t)$ : 系の状態

$u$ : 制御変数

$t$  を手数,  $x(t)$  を minimax 探索の指し手で発展していく局面,  $u$  を特徴ベクトルとみなし, 最適制御理論に基づいた将棋プログラムの機械学習をおこなう

# Minimax 探索結果の最適制御

最適制御法に基づき、棋譜の指し手と minimax 探索が良く一致する  
特徴ベクトル  $\nu$  を求める

局面の優劣を表す最適な数値を求める

指し手の一致度を測る目的関数  $J'$  を以下のように設計する

$$J'(P_0, P_1, K, P_{N-1}, \nu) = \sum_{i=0}^{N-1} l(P_i, \nu)$$

$P_i$ : サンプルされた棋譜中の一局面

$l(P_i, \nu)$ : 棋譜中の手とプログラムの指し手一致度

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

$p_m$ : 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

$\xi(p_m, \nu)$ : minimax 探索の評価値

$T(x)$ : 評価値の差を、指し手一致度に変換する関数

# 結果

駒割 (交換値)

106 272 279 304 323 363 415 428 527 617 698 700 854

歩 と 香 桂 成香 成桂 成銀 銀 金 角 馬 飛 竜

持ち駒の数

歩 : 27 33 21 6 -8 -17 -23...  
香 : 28 39 51 63  
桂 : 22 12 -15 -48  
銀 : 37 28 -2 -51  
金 : 31 21 -4 -39  
角 : 28 9  
飛 : 59 45

角・馬の動ける升の数

角 : -55 -25 -7 0 8 14 6 9  
馬 : -28 -12 -3 1 8 10 16 11

香・飛・竜縦方向に動ける升の数

香 : x -7 2 6 18 25 27 24  
飛 : -61 -43 -22 -9 -2 9 12 17  
竜 : -45 -26 -17 -13 -1 1 10 8

飛・竜横方向に動ける升の数

飛 : -72 -47 -18 4 10 28 21 27  
竜 : -35 -23 -16 -10 -4 4 11 12

### 王が8八にいる時の味方の金・と

-67 -60 -45 -27 -30 -51 -61 -73 -67  
 -57 -59 -28 -4 6 -13 -48 -82 -62  
 -18 -26 -26 0 14 -5 -27 -51 -40  
 -57 -35 -20 3 -7 -8 -15 -49 -64  
 -55 -36 -13 -18 -13 -23 -17 -35 -77  
 -57 -16 -5 -8 -12 -25 -37 -54 -68  
-34 -3 13 19 -25 -39 -55 -73 -92  
-97 K 30 -16 -6 -51 -36 -85-150  
-40 34 -49 23 -40 -17-109-123-118

### 王が9九にいる時の味方の金・と

-69 -45 -17 8 -13 -38 -70 -79 -75  
 -21 -55 11 29 22 -12 -55 -80 -75  
 -12 -29 -8 28 37 0 -36 -56 -65  
 14 -35 -7 33 9 -5 -33 -56 -83  
 -81 -49 -9 -11 -14 -25 -39 -56-101  
 -59 -25 2 -28 -37 -57 -63 -56 -72  
 14 36 17 -19 -49 -74 -80-108 -75  
45 27 18 1 -32 -91-123-127-138  
 K 40 31 0 -14 -56-139-133-116

### 王が8八にいる時の敵の金・と

-106 -25 -5 2 -42 -11 -20 -41 -67  
 -53 -7 10 -4 5 -3 8 -18 -72  
 -62 -26 -18 9 -21 4 -11 -28 -85  
 -50 -14 9 -9 -9 -20 -13 -28 -86  
 -12 24 25 4 -20 -27 -30 -47 -63  
 110 186 144 49 2 -29 -28 -36 -40  
 450 450 450 149 27 -12 -22 -25 -34  
 450 K 450 156 15 -24 -40 -31 -44  
 112 450 212 63 -19 -52 -69 -59 -72

### 王が9九にいる時の敵の金・と

-52 -8 15 7 -47 -34 -32 -57-177  
 -54 14 30 5 -2 -16 6 -49 -98  
 -80 -11 -1 12 -17 -1 -22 -31 -79  
 -58 -9 -8 -10 -25 -30 -28 -40-130  
 -26 -7 -6 -8 -32 -46 -42 -85-182  
 128 120 35 -15 -7 -57 -56 -62-116  
 320 271 206 65 11 -18 -45 -36 -68  
 318 444 214 106 14 -31 -47 -55 -77  
 K 448 207 51 -8 -61 -91 -99 -96

### 王が8八にいる時の味方の歩

x	x	x	x	x	x	x	x	x
-1	-39	-44	-52	-36	-49	-45	-67	-47
16	-31	-42	-13	0	-3	-32	-30	5
<u>21</u>	<u>21</u>	<u>15</u>	<u>21</u>	<u>25</u>	<u>21</u>	<u>14</u>	<u>23</u>	<u>14</u>
2	-12	1	4	4	-6	-1	-5	4
7	-12	2	11	0	-5	0	-8	2
-5	7	3	0	-3	-3	-6	3	2
62	K	48	10	-3	-19	-34	-27	-79
37	<u>-97</u>	15	13	-5	-23	-31	-24	-60

### 王が9九にいる時の味方の歩

x	x	x	x	x	x	x	x	x
-48	-86	-48	11	-7	-40	-35	-49	-35
28	14	-16	12	19	0	-60	-25	-10
<u>58</u>	<u>29</u>	<u>39</u>	<u>26</u>	<u>38</u>	<u>22</u>	<u>8</u>	<u>24</u>	-1
-7	-6	7	2	-5	-6	-5	-5	-4
-5	-13	0	3	0	-11	4	-16	-1
23	9	2	2	-8	2	-8	-1	8
72	-3	40	26	-10	-37	-49	-40	-99
K	36	67	33	-10	-39	-49	-56	-90

### 王が8八にいる時の敵の歩

-48	-111	-57	-38	-21	-25	-29	-60	-56
8	-52	-41	-6	-3	-2	-24	-24	2
-8	-3	-7	-7	-2	-1	-4	1	1
3	-9	-3	-1	-1	-1	4	-13	1
19	4	4	15	-3	-2	0	-11	-19
<u>86</u>	<u>156</u>	<u>82</u>	<u>31</u>	<u>27</u>	11	3	17	-20
<u>115</u>	<u>0</u>	<u>77</u>	<u>50</u>	4	-13	-18	-21	-80
<u>83</u>	K	<u>84</u>	-19	-4	-39	-25	-29	-86
x	x	x	x	x	x	x	x	x

### 王が9九にいる時の敵の歩

-101	-97	-49	-4	0	-11	-26	-48	-69
-24	-81	-25	-2	7	-2	-21	-32	-27
-30	-9	-13	-16	-10	-5	-10	2	3
-18	-5	1	-8	-4	-12	-9	-24	-8
1	8	-4	0	-3	-12	-19	-20	-38
<u>72</u>	<u>108</u>	<u>31</u>	<u>44</u>	15	10	-13	7	-76
<u>200</u>	<u>199</u>	<u>76</u>	<u>53</u>	3	-16	-35	-43	-136
<u>0</u>	<u>200</u>	<u>111</u>	<u>36</u>	-4	-54	-32	-35	-87
K	x	x	x	x	x	x	x	x

### 88王の時の味方の駒2つの位置

#### 金の位置の得点

+8	-1	+1	-7	+1	+1
王	銀	<u>-33</u>	王	金	+1
+5	-14	<u>+1</u>	-2	+9	<u>-96</u>

#### 歩の位置の得点

+1	+1	-2	+1	+1	-1
王	銀	<u>-16</u>	王	金	+2
<u>+72</u>	+19	<u>+52</u>	<u>-107</u>	-6	<u>-42</u>

### 78王の時の味方の駒2つの位置

#### 金の位置の得点

+4	+6	+23	-1	+9	+18
王	銀	<u>-12</u>	王	金	+10
+18	+4	<u>+33</u>	+2	-18	<u>-28</u>

#### 歩の位置の得点

+0	+0	+0	+12	+6	+4
王	銀	+4	王	金	+12
<u>+32</u>	+21	<u>+32</u>	+25	+4	<u>-42</u>



# プロ棋士と Bonanza の思考内容

	9	8	7	6	5	4	3	2	1	
	▲	桂						桂	▲	一
			▲	王		▲	▲			二
			▲	▲	▲		▲	▲		三
	▲				▲					四
			▲			▲				五
▲	▲		▲	▲	▲	▲				六
	▲		▲	▲	▲		▲		▲	七
	▲	▲	▲	▲				▲		八
	▲	▲						▲	▲	九

【▽ 4四飛まで】

プロ棋士は 7七歩, 4五歩, 9五歩に着目

伊藤毅志, 松原仁, ライエル・グリーンベルゲン  
 将棋の認知科学的研究, コンピュータ将棋の進歩5  
 松原仁編, 共立出版, 2005年

指し手	探索ノード数	割合
7七歩	17,330,794	17.8%
2六飛	41,907,150	43.1%
7九玉	13,377,128	13.8%
2五飛	5,513,094	5.7%
9五歩	3,494,148	3.6%
他の手合計	11,878,409	16.2%

図の局面におけるBonanzaの思考．基準深さ12，全探索ノード数約1億，最善手は 7七歩（後手歩1枚有利）．

# 対竜王戦，中盤から終盤へ

	9	8	7	6	5	4	3	2	1	
▲	龍王	龍王						龍王	龍王	二
△							金将	金将	車	三
▲	歩			歩		歩	相	銀		四
△			歩	歩	歩	銀	歩	歩	歩	五
▲										六
△			歩		歩			歩		七
▲	歩		桂			飛	歩	銀	歩	八
△							金	銀	香	九
▲	王						金	桂	王	

▲ 歩

▲ 歩

▲ 車

▲ 後手

▲ ボナンザ  
00:00 / 00:00:00

△ 渡辺明竜王  
00:00 / 00:00:00

手合割：平手

手数：71

▲ 6四歩 まで

後手番

棋戦：大和証券杯\*

対局日：2007/03/21

▲ 先手

▲ 香車

▲ 歩

# 対竜王戦，投了図

▲ ボナンザ  
00:00 / 00:00:00  
△ 渡辺明竜王  
00:00 / 00:00:00  
手合割：平手  
手数：113  
▲ 投了 まで  
後手番  
棋戦：大和証券杯  
対局日：2007/03/21

先手  
飛車 金将 銀将  
銀将 香車 香車  
歩兵 歩兵 歩兵 歩兵 歩兵

後手  
飛車 金将 銀将  
銀将 香車 香車  
歩兵 歩兵 歩兵 歩兵 歩兵