シミュレーションに基づくスイス式トーナメント組み合わせ方式の評価

香山 健太郎
独立行政法人 通信総合研究所
kayama@crl.go.jp

概要

本稿では、シミュレーションによってスイス式トーナメントの性能を評価した。評価の指標としては、実力上位者が上位に入る率のほか、実際の対戦で勝った方が上位になる率や対戦相手の勝数の合計の差が小さくなる度合などの公平感を表すものをも用いた。対戦方式としては、通常のスイス式の他、変形スイス式やそれらの変種を用意した。また、参加選手数、回戦数、参加選手の強さの分布も多数用意した。その結果、ある条件下で、回戦数を増やしても正しい結果が得られる可能性が下がるという逆転現象が起こること、参加者を強さ順に並べることはむしろ性能悪化の危険があること、対戦相手の勝数の合計 (ソルコフ) の差が小さくなるような当て方が良いこと、通常スイス式と変形スイス式とではの 1 回戦程度の差があることがわかった。

Evaluation of Pairing Systems on Swiss-system Tournament by Simulation

Kentaro KAYAMA
Communications Research Laboratory

Abstract

In this paper I evaluated paring systems on Swiss-system tournament by simulation. As evaluation items not only the rate that stronger players are ranked by matches but also fairness that the winner precedes the loser and the bias of opponents is little is used. Various pairing system which are normal Swiss, modified Swiss, and variants of them are prepared. The large number of entry players and rounds and the distribution of the strength of the players are also prepared.

As a result, these conclusions are acquired: the accuracy is sometimes reduced according to the numbers of rounds, the player entry order according to their strength leads inaccurate result, the good pairing system is to reduce the bias of opponents, and the difference of normal Swiss pairing system and modified Swiss is about one round.

1 はじめに

多数の選手が参加する大会において、優勝者だけでなく上位の数選手、あるいは全選手の順位を決定しようとした場合、その対戦形式としてスイス式トーナメントが用いられることが多い。

しかし、スイス式トーナメントにおいてどの程度正確に順位が決定できるのか、という解析はまだあまり進んでいない。瀧澤ら [1] は第 12 回世界コンピュータ将棋選手権の対戦結果を用いて、総当たり式・通常スイス式・変形スイス式、および選手の並び方の順序による結果の違いを調べている。また、橋本ら [2] は、総当たり式・通常スイス式・ランダムスイス式・W杯予選方式、および選手の強さ分布による結果の違いをシミュレーションで調べている。しかし、指標として上位入賞回数のみを用いていること、試合数を固定していることなど、試合方式の良し悪しを詳細に分析するためのデータ量としてはまだ不足している。

そこで、本稿では、順位決定の正確さを測る様々な指標を用いて対戦方式を評価する。また、正確さだけでなく参加選手の公平感についての指標も導入する。評価する対象としては、スイス式の組み合わせ決定法を様々に変えるほか、参加選手数、回戦数、参加選手の強さの分布も変化させる。そして、それぞれの場合において 1000 回のシミュレーションを行って指標の変化を見る。

2　問題の定式化

2.1　組み合わせの評価法

組み合わせ評価の問題は、正しい順位を少ない対戦数でどれだけ正確に評価できるかということになる。しかし、勝敗の結果は基本的には実力に基づくものの最終的には確率的に決まるため、何をもって正しい順位とするかは難しい。

そこで、本発表では、2.3 節で述べるような指標を用意し、多方面から各対戦法を評価することとした。

2.2　シミュレーションの概要

本実験では、1 回の試行ごとに各選手のレーティングの点数を決定し、それをもとに全選手間での総当たり対戦表を作成する。そこから、以下の 3 種類の順位を求める。

順位 A レーティング点の順位

順位 B 総当たりの場合の順位

順位 C 各選手 n 回戦分だけを抜き出して求めた順位

そして、これらの差がどの程度のものになるか、およびどの程度公平感を感じられる結果になるかを調べた。

なお、順位決定にはスイス式で一般に行われている方法、すなわち、

1. 勝数
2. ソルコフ (対戦相手すべての勝数の合計)
3. SB(勝った相手の勝数の合計)
4. ミディアム (SB から最も勝数の多い相手、および最も少ない相手を除いたもの)
5. 直接対決

の順に考慮する方法を用いた。なお、シミュレーションでは引き分けは発生しないものとした。

そして、参加選手数、回戦数、参加選手の強さの分布、対戦方式を適宜変更して各 1000 回の試行を行い、次節で述べる指標をもとに得られる結果の性質を調べた。

2.3　評価の指標

今回は、以下のような評価の指標を用意した。そして、指標群 A を最も重視することにした。

指標群 A 上位選手の顔ぶれの一致度

指標群 B 順位の一致度

指標群 C 結果の公平感

それぞれの具体的な内容を表 1 に示す。

2.4　スイス式の組み合わせ方法

スイス式の組み合わせ方式としては、次のもの、およびその混合方式を採用した。

NS その直前の対戦までの結果を使う方式 (通常スイス式, Normal Swiss)

MS 2 戦前までの結果を使う、ただし 2 戦目の場合は初戦が上位者勝ちと仮定した結果を使う方式 (変形スイス式, Modified Swiss)

RA 完全にランダムに当てる方式

スイス式では、一般的に勝数ごとにグループ分けし、そのグループ内で閉じた対戦が組めない場合は下のグループ内での並び順が上の選手から当たるように、という方式を取っている。このグループの分け方について、

1. 勝数ごとにグループ分け
2. ソルコフごとにグループ分け

の 2 つを試した。また、グループ分けをせず、

3. 上位から組める順に対戦
4. 順位上位からソルコフの小さい選手に当てていく
5. ソルコフ上位からソルコフの小さい選手に当てていく

という方式を試した。

また、グループ内での並び順について、

A. 選手番号順 (通常のスイス式)

B. ソルコフ降順

C. ソルコフ昇順

を、同グループ内での当て方について、

a. グループを並び順に従ってほぼ 2 等分し、お互い上から当てていく (A,B,C,D の順なら A-C,B-D)

b. グループの上から順に当てていく (A,B,C,D の順なら A-B,C-D)

という方式を試した。

表 1: 評価の指標の具体的内容

指標群 A	上位選手の顔ぶれの一致度
A-1(n)	順位 A で n 位までの選手が順位 C で n 位に入った回数
A-2(n)	順位 B で n 位までの選手が順位 C で n 位に入った回数
指標群 B	順位の一致度
B-1a	順位 A と順位 C との差の絶対値の総和
B-2a	順位 A で 5 位以内の選手の順位 C の和
B-3a	順位 C で 5 位以内の選手の順位 A の和
B-1b,2b,3b	以上の順位 A を順位 B に置き換えたもの
指標群 C	結果の公平感
C-1a	順位 A で 10 位以内の選手の直接対決数
C-1b	順位 B で 10 位以内の選手の直接対決数
C-2	順位 C で 10 位以内の選手の直接対決数
C-3	順位 C で 10 位以内の選手のソルコフの分散
C-4	順位 C において下位の選手が上位の選手に勝った対戦数

2.5 参加選手のモデル

勝敗を決定するためのレーティング点数分布は、まずベースとなる点数として等幅 (乱数含まず)・擬似正規分布の 2 種類を用意した。さらに、選手番号の付け方として、点数順、ランダム、点数順に並べた上でその点数に乱数値を加算、の 3 種類を用意した。点差は、順位 1 つで平均 100 点差および 500 点差のものを基本とし、それ以外の点差も適宜調べた。また、選手数は、コンピュータ将棋選手権 2 次予選の参加者数である 24 を基本とし、12,16,20,22,26,28,32,40,48,64,72 の場合も適宜調べた。

擬似正規分布によるレーティング点数決定には以下の式を用いた。

$$R = \frac{Nd}{4}\sum_{i=1}^{n} rand(1)$$

ただし、R は任意の参加者の点数、N は選手数、d は平均点差、$rand(1)$ は 0 から 1 までの乱数とする。

2.6 勝敗の決定方法

各対戦の勝敗の決定は、以下のレーティング点差に基づく勝敗確率計算式 (簡易版) を用いた。

$$P_A = \begin{cases} 1 & (R_A - R_B \geq 400) \\ \frac{R_A - R_B + 400}{800} & (-400 < R_A - R_B < 400) \\ 0 & (R_A - R_B \leq -400) \end{cases}$$

ただし、P_A : 選手 A が勝つ確率, R_A : 選手 A のレーティング, R_B : 選手 B のレーティング とする。

3 シミュレーション結果とその分析

それぞれの組み合わせ方法、参加選手のモデルについて、4 回戦から 20 回戦程度までのシミュレーションを各 1000 回行い、2.3 節で述べた指標がどのように変化するかを調べた。

3.1 各指標の基本的性質

まず、様々な対戦方式・参加者分布・回戦数において、2.3 節で述べた各指標がどのように変化する傾向があるか調べた。

その結果、回戦を増やすと、全体的に以下のような性質があることがわかった。

- 上位選手の顔ぶれの一致度 (指標群 A) は振動しつつ増加する
- 順位の一致度 (指標群 B) は振動しつつ減少する
- 上位直接対決数 (C-1a, C-1b, C-2) は単調に増加するが、通常スイス式の場合少し振動する
- ソルコフ分散 (C-3) は特殊な形のグラフとなる
- 逆転対戦数 (C-4) は単調に減少する

図 1 に、参加者分布を実際の大会での分布に比較的近いと思われる平均 100 点差の擬似正規分布に ±100 点の一様乱数を加えたもの (3.4 節の記法で MN-100) にした際の代表的な指標のグラフを示す。

(a) 指標 A-2(5): 順位 B で 5 位までの選手が順位 C で 5 位以内に入った割合

(b) 指標 B-1a: 順位 A,C の差の絶対値の総和

(c) 指標 C-2: 順位 C で 10 位以内の選手の直接対決数

(d) 指標 C-3: 順位 C で 10 位以内の選手のソルコフの分散

なお、横軸は回線数を示す。また、グラフに示した対戦方式は 2.4 節で示したもののうち

RR 総当たり結果

NS-1Aa 通常スイス式 (一般型)

NS-1Ba 通常スイス式 (同勝数グループ内はソルコフ降順に並べる)

NS-2Aa 通常スイス式 (ソルコフごとにグループ分け)

NS-3 通常スイス式 (上位から組める順に対戦)

NS-5 通常スイス式 (ソルコフ上位と下位から当てる)

MS-1Aa 変形スイス式 (一般型)

RA ランダム方式

である。

3.2 対戦方式と性能の関係

方式別では次のような傾向があることがわかった。方式は 2.4 節で定義した記号で表している。

- 通常スイス式 (NS)、変形スイス式 (MS) とも、性質は似ている。
- 1A, 1B は似た性質を示し、指標群 A, B では良い値を示すが C-4 ではあまり良くない。また、回戦を増やしても性能が落ちる危険がある。
- 1C, 2A は似た性質を示し、全ての指標で良い値を示す。また、回戦を増やしたとき性能が落ちる危険が少ない。
- 3 は 1A, 1B と似た性質を示すが、回戦数が少ないときに性能が大幅に落ちる。
- 4, 5 は似た性質を示し、ソルコフ分散 (C-4) のみ良い値を示すが、それ以外の指標ではあまり良くない。また、回戦を増やしても性能が落ちる危険がある。
- ランダム方式 (RA) は各指標であまり良い値を示さない。

3.3 ソルコフ利用方式の性能詳細比較

3.2 節で述べたとおり、組み合わせの際のグループの並び順をソルコフ昇順にすること (NS-1Ca, NS-1Cb)、あるいはグループ分け自体を勝数別ではなくソルコフ別にすること (NS-2A) は有力である。

ただし、ソルコフ別にグループ分けをすることは、参加者にとって対戦がわかりにくくなる場合があり、不公平感を引き起こす恐れがある。

そこで、グループの並び順をソルコフ昇順にする方式 NS-1Ca, NS-1Cb について、いくつかの参加者数・分布別・回戦別に通過率を通常型スイス式 (NS-1Aa) と比較した場合の優劣を調べた。また、一部を変形スイス式とした場合の優劣についても調べた。その結果、NS-1Cb はやや性能が劣るものの、NS-1Ca は NS-1Aa とほぼ遜色ない結果が得られた。

ここで、指標 A-2(5) について、横軸を NS-1Aa、縦軸を NS-1Ca または NS-1Cb としたものを図 2 に示す。

3.4 参加者の分布による性質

参加者の分布については、平均点差が開くほど各指標の性質は良くなる。しかし、同一点差の分布については、回戦数を増やしたときレーティング点数順位・総当たり順位が上位の選手が上位入賞する回数が減る現象 (以降逆転現象と呼ぶ) が、特に選手番号が点数順に並んでいるときに多く発生した。また、逆転現象は点差が開くほど多くなる傾向があることもわかった。

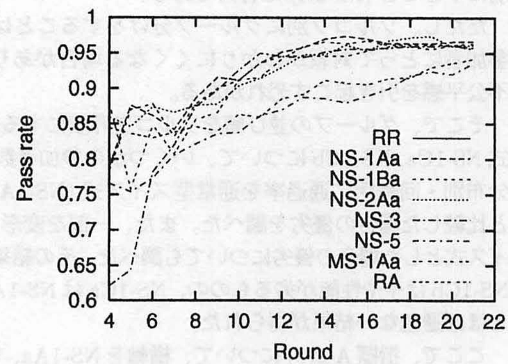
参加 24 人、NS-1Aa および NS-1Ca 方式の場合のグラフを図 3 に示す。なお、分布を表す記号については、1 文字目が A なら等幅・N なら擬似正規分布であり、2 文字目が S なら点数順、R ならランダム、M なら点数順に並べた上で −100〜＋100 の一様乱数を加えていることを示す。また、数字は平均点差を示す。

3.5 上位選手の顔ぶれの一致度 (指標群 A) の性質

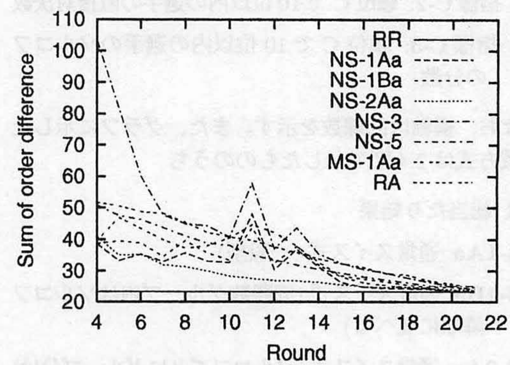
上位選手の顔ぶれの一致度 (指標群 A) については、これまでの図で示したように、回戦数を増やしても逆転現象が起こることがある。分布を 3.1 節の例で用いた MN-100 としたとき、上位 n 位に点数 n 位までが入った率 (指標 A-2(n)) について、参加 24 人、NS-1Aa および NS-1Ca 方式の場合のグラフを図 4 に示す。なお、逆転現象が起こる回戦・順位 (上位何位までを通過とするか) については、法則性は発見できなかった。

3.6 通常スイス式と変形スイス式の性能差

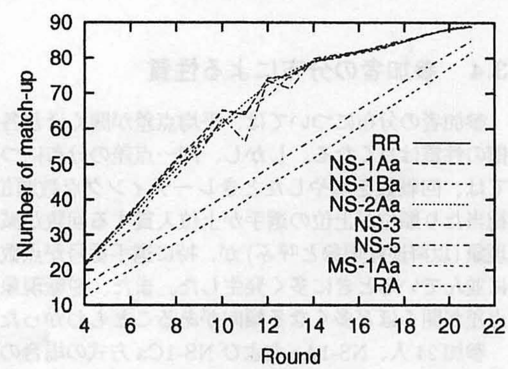
ほぼ全ての方式で回戦数を増やせばレーティング点数順位・総当たり順位通りに実際の順位が並ぶ可能性



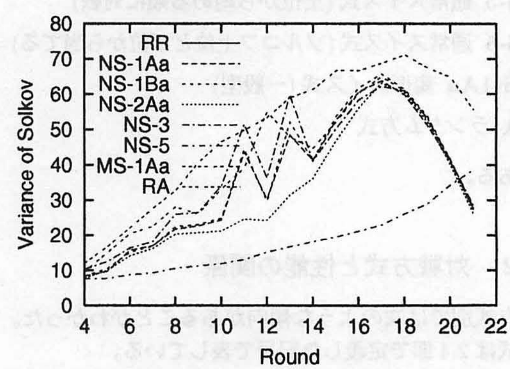
(a) A-2(5) 通過率



(b) B-1a 順位 A,C の差の絶対値の総和

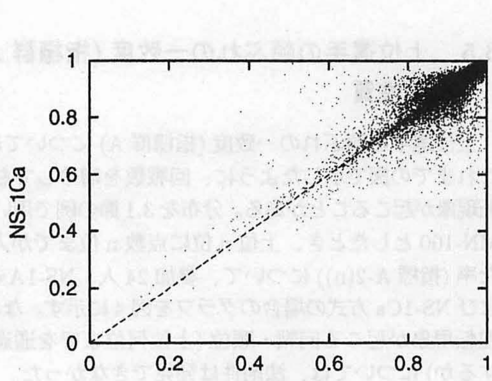


(c) C-2 直接対決数

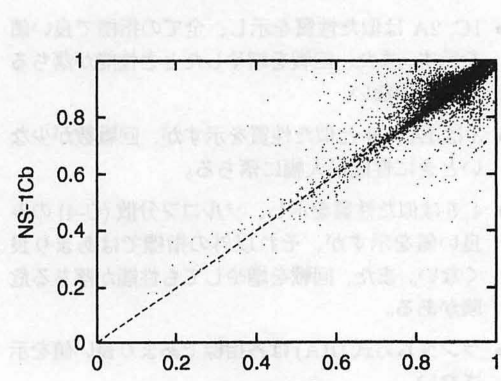


(d) C-3 ソルコフの分散

図 1: 各指標・対戦方式の傾向

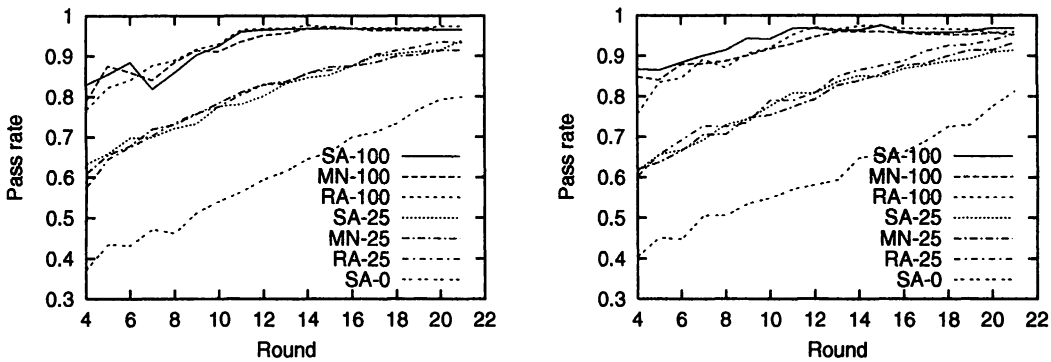


(a) NS-1Aa と NS-1Ca の通過率



(b) NS-1Aa と NS-1Cb の通過率

図 2: ソルコフを用いた方式の性能評価



(a) NS-1Aa 方式の場合　　(b) NS-1Ca 方式の場合

図 3: 参加者の分布による上位 5 人の通過率

表 2: 変形スイス式の性能

	M0	M1	M2
1Aa 方式・上位 3 選手	1.231	0.923	0.872
1Aa 方式・上位 5 選手	1.128	0.885	0.346
1Ca 方式・上位 3 選手	0.795	0.705	0.474
1Ca 方式・上位 5 選手	0.500	0.179	0.128

が増えるが、通常スイス式と変形スイス式ではどの程度性能の差があるかを調べた。ここでは、1Aa および 1Ca 方式について、通常 (NS)6-11 回戦を基準とし、その上位 3 選手の通過率 (A-2(3)) および上位 5 選手の通過率 (A-2(5)) の値を初めて上回るのに必要な余分な回戦数を、

M0 最終戦まで変形スイス式

M1 最終戦のみ通常スイス式、それまでは変形スイス式

M2 最終戦とその直前のみ通常スイス式、それまでは変形スイス式

の 3 方式について、分布は MN-25,MN-50,MN-75,MN-100、参加者数は 20,24,32、および参加者数 24 で分布 AS-100 の計 13 種の例について調べた。その結果を図 5 と表 2 に示す。

この結果、通常スイス式と変形スイス式の差はほぼ 1 回戦程度であることがわかった。

4 まとめ

本稿では、スイス式をベースとした組み合わせ決定方式について、様々な条件でシミュレーションを行い、評価の際に用いるべき指標は何か、および各条件によってその指標がどのように変化するかを述べた。

その結果、以下のことがわかった。

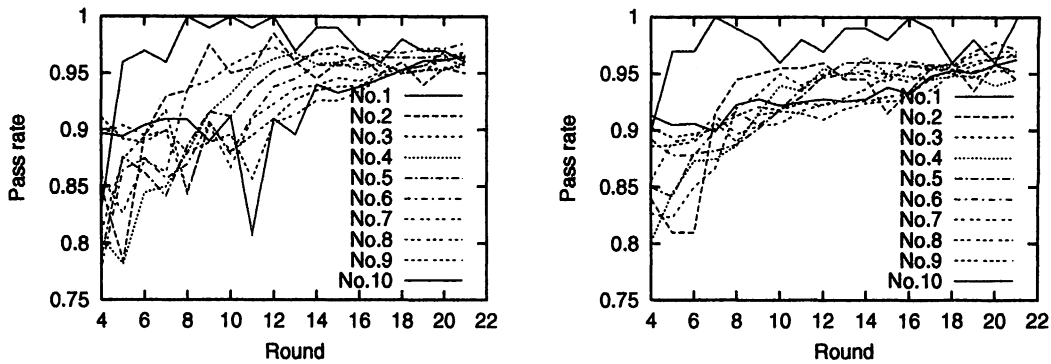
- ある条件下で、回戦数を増やしても正しい結果が得られる可能性が下がるという逆転現象が起こること
- 参加者を強さ順に並べることは、むしろ性能悪化の危険があること
- 対戦決定の際のグループ分けのときの並べ方はソルコフ昇順が良いこと
- 通常スイス式と変形スイス式とではの 1 回戦程度の差があること

今後の課題としては、統計的な精度を求めること、今回わかった性質を数学的に説明することあるいはなんらかの法則性を発見すること、およびスイス式以外の様々な方式を試して性能を評価することがあげられる。

参考文献

[1] 瀧澤武信, 柿木義一: 世界コンピュータ将棋選手権における対戦組み合わせシステムの有効性, 第 7 回ゲーム・プログラミングワークショップ 2002, pp. 93-100, 2002.

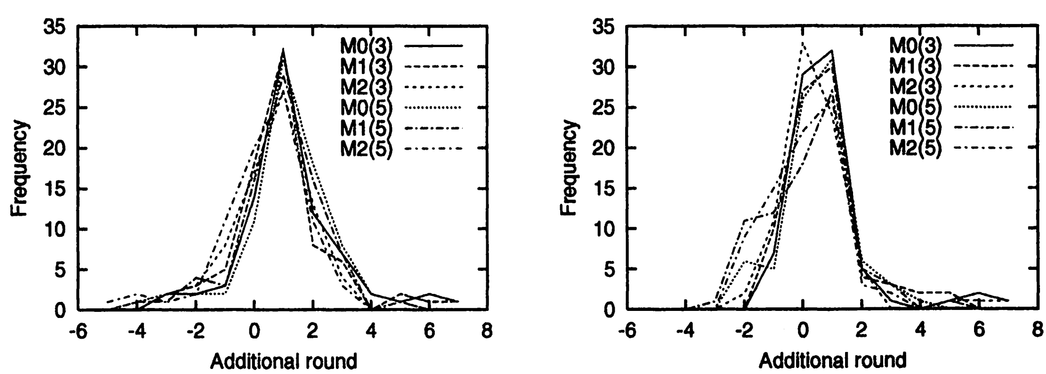
[2] 橋本剛, 長嶋淳, 飯田弘之: シミュレーションによる大会方式検証の提案ー世界コンピュータ将棋選手権を題材にしてー, 第 7 回ゲーム・プログラミングワークショップ 2002, pp. 101-108, 2002.



(a) NS-1Aa 方式の場合

(b) NS-1Ca 方式の場合

図 4: 回戦数と通過率 A-2(n) の関係



(a) 1Aa 方式の場合

(b) 1Ca 方式の場合

図 5: 変形スイス式の性能