A Method for Illustrating Shogi Postmortems Using Results of Statistical Analysis of Kifu Data

Nobutane Hanayama, Ryuichi Nogami

A Method for Illustrating Shogi Postmortems Using Results of Statistical Analysis of Kifu Data

Nobutane Hanayama[1*] and Ryuichi Nogami[1]

[1] Shobi University, Kawagoe 350-1110, Japan
nob-hanayama@jcom.home.ne.jp

Abstract. Kifu, which is a Japanese term for a game record for a shogi games, is considered as a special type of multi-dimensional time series data, because it consists of items indicating who moved a piece and where it was moved on a grid diagram. Because of its complexity, however, it is not easy for amateur or non-players of shogi to grasp or understand overall shogi games from it. In this study we suggest averages, variances, skewness of row numbers where pawns are put, numbers of times gold and silver generals and nights are moved, which are easily calculated and understandable without comprehensive knowledge of shogi game, as features of position of shogi games. And the usability of those features is shown by a result of discriminant analysis for winner of game based on those features. Further, a software for illustrating shogi postmortems using discriminant scores obtained from discriminant analysis of Kifu data are shown.

Keywords: Discriminant analysis of winner of shogi game· Feature of position of shogi game· Statistical indicator of position of shogi game

1 Introduction

Kifu, which is a Japanese term for a game record for a shogi games, is considered as a special type of multi-dimensional time series data, because it consists of items indicating who moved a piece and where it was moved on a grid diagram. In Kifu each move of piece is noted like “the first move, (3, 6), pawn”. Because of its complexity, however, it is not easy for amateur or non-players of shogi to grasp or understand overall shogi games from it. In this study we suggest to consider averages, variances and skewness of row numbers where pawns (FUs) are put, and times gold and silver generals (KIs and GIs) and nights (KEs) are moved as features of position of shogi games.

Shogi game has been mainly studied in the field of artificial intelligence (AI) as related to the chess game study. In this field, “evaluation functions” [1] are frequently used when AIs evaluate which player or which of human player or computer is superior and determine the next move at a certain position. However, evaluation functions require us comprehensive knowledge of shogi game when we make AIs calculate from “features”, “material” or “mobility” for examples. On the other hand, average, variance and skewness of row numbers where pawns (FUs) are put are simple statistical indicators, which can be easily calculated. So in this study, we propose that average, variance and skewness of row numbers where pawns (FUs) are put, times gold and silver generals (KIs and GIs) and nights (KEs), which are all easily calculated and understandable without comprehensive knowledge of shogi game, as features of shogi games (Section 2). And we show the usability of proposed features by a result of discriminant analysis of winner/looser based on kifu data (Section 3).

So far, the aim of shogi game study concerning AI is to make computer software win against first-rate professional shogi players [2]. However, the strengthening promotion committee "Akara", which was a computer software developed by a project team organized by IPSJ (the information society of Japan) or "Ponanza" produced great results [3], which means that the shogi game study concerning AI has achieved most of its aims [4]. Hence, studies of methods concerned with entertainment, methods for conducting a post-mortem for an example, may become more important [5] recently. So in this study, we suggest a method for illustrating shogi postmortems using discriminant scores obtained from discriminant analysis of Kifu data, and show its usability by introducing a software incorporated with the suggested method (Section 3).

2 Average, Variance and Skewness of Row Numbers Where Pawns (FUs) Are Put

Let $J_i^{(k)}$ indicate number of pawns (FUs) at the i -th move ($i = 1, 2, \cdots$), where it is for the white player (the first move) if $k = 1$ and for the black player (the defensive move) if $k = 2$, and let $h_{i,j}^{(k)}$ indicate row number of the j -th pawn (FU) at the i -th move. Then average, variance and skewness of row numbers where pawns (FUs) are put at the i -th move, $\bar{h}_i^{(k)}$, $\tilde{h}_i^{(k)}$ and $\ddot{h}_i^{(k)}$ are calculated by

$$\bar{h}_i^{(k)} = \frac{1}{J_i^{(k)}} \sum_j^{J_i^{(k)}} h_{i,j}^{(k)} \, , \; \tilde{h}_i^{(k)} = \sum_j^{J_i^{(k)}} \frac{(h_{i,j}^{(k)} - \bar{h}_i^{(k)})^2}{J_i^{(k)} - 1} \, , \; \tilde{h}_i^{(k)} = \frac{1}{J_i^h} \sum_j^{J_i^h} (h_i^{(j)} - h_i^a)^2 \Big/ h_i^{v3/2} \, . \quad (1)$$

In the above, though averages $\bar{h}_i^{(k)}$. variances $\tilde{h}_i^{(k)}$ and skewnesses $\ddot{h}_i^{(k)}$ are indicators calculated for i , that is, every move, whole game are analyzed rather than every move in this study. Hence, let's consider every 10 moves as "phase". In addition, numbers of moves from the first move to the end/resigned move are different among games. So let's consider $\bar{h}_i^{(k)}$, $\tilde{h}_i^{(k)}$ and $\ddot{h}_i^{(k)}$ calculated for 40 moves from the first move to the end/resigned move as features of whole game, that is, moves from the first to the 10th as the first phase, moves from the 11th to the 20th as the second phase, moves from the 21th to the 30th as the third phase and moves from the 31th to the 40th as the fourth phase and

$$\bar{p}_l^{(k)} = \sum_{i=1+(l-1)\times 10}^{l\times 10} \frac{\bar{h}_i^{(k)}}{10} \, , \; \tilde{p}_l^{(k)} = \sum_{i=1+(l-1)\times 10}^{l\times 10} \frac{\bar{h}_i^{(k)}}{10} \; \ddot{p}_l^{(k)} = \sum_{i=1+(l-1)\times 10}^{l\times 10} \frac{\ddot{h}_i^{(k)}}{10} \quad (2)$$

for $l = 1,2,3,4$, as features for whole games. In addition to the above, consider numbers of times gold and silver generals (KIs and GIs) and nights (KEs) are moved calculated by

$$\begin{aligned} g_l^{(k)} &= count(< r_{1+(l-1)\times 10}^{(k)}, r_{1+(l-1)\times 10+1}^{(k)}, \cdots, r_{1+l\times 10}^{(k)} >, \text{"gold"}) \\ s_l^{(k)} &= count(< r_{1+(l-1)\times 10}^{(k)}, r_{1+(l-1)\times 10+1}^{(k)}, \cdots, r_{1+l\times 10}^{(k)} >, \text{"silver"}) \\ n_l^{(k)} &= count(< r_{1+(l-1)\times 10}^{(k)}, r_{1+(l-1)\times 10+1}^{(k)}, \cdots, r_{1+l\times 10}^{(k)} >, \text{"night"}) \end{aligned} \quad (3)$$

as features of position of shogi games, where $r_i^{(k)}$ indicates the i-th character and $< s_1, s_2, \cdots >$ is a concatenation of strings $s_1, s_2, \cdots$ and $count(< s_1, s_2, \cdots >,$ $\text{"string"})$ indicates the number of characters in string "string" .

3 Result of Discriminant Analysis of Winner/Looser Based on Kifu Data

Now let us consider a discriminant function:

$$z_m = \sum_{l=1}^{4}\sum_{k=1}^{2}(a_l^{(\bar{p})}\bar{p}_{l,m}^{(k)} + a_l^{(\tilde{p})}\tilde{p}_{l,m}^{(k)} + a_l^{(\ddot{p})}\ddot{p}_{l,m}^{(k)} + a_l^{(g)}g_{l,m}^{(k)} + a_l^{(s)}s_{l,m}^{(k)} + a_l^{(n)}n_{l,m}^{(k)}) + c \;. \quad (4)$$

where $a_{l,m}^{(\cdot)}$ are discriminant coefficients c is a constant and z_m are discriminant scores. In the discriminant analysis using the function (4), the white player (the first move) is discriminated as a winner when $z_m \geq 0$ and the black player (the defensive move) is discriminated as a winner when $z_m < 0$.

Table 1 indicates the accuracy of discriminant function (4) estimated based the ki-hu data for 50 shogi games ($M = 50$) form the 14th Young Lion Tournament is shown, Because the number of parameters to be estimated is $4 \times 6 \times 2 = 28$, which is around 3/5 of the number of games, discriminant coefficients are estimated with the forward selection method by $F = 2$. As indicated on the table, the estimated discriminant function is significant to some extent because the percentage of correct classifications is 84 %, and it is found that the estimated discriminant function is significant to some extent.

Table 1. Accuracy of discriminant analysis with the forward selection method by $F = 2$

Percentage of correct classifications	84.00%
Percentage of incorrect classifications	16.74%
Mahalanobis square distance	372.03%
Correlation ratio	49.21%

Table 2 indicates the estimates of the discriminant coefficients estimated with with the forward selection method by $F = 2$. Descriptions for selected variables on the table are corresponding to formulas (3) as follows:

times black moves silver in the 4th phase : $g^{(2)}_{4,m}$,

skewess of pawn (FU) for black in the 3rd phase : $\ddot{p}^{(2)}_{3,m}$,

times white moves nights (KEs) in the 2nd phase : $n^{(1)}_{2,m}$,

variance of pawns (FUs) for black in the 2nd phase : $\tilde{p}^{(2)}_{2,m}$,

variance of pawns (FUs) for white in the 2nd phase : $\tilde{p}^{(1)}_{2,m}$,

times white moves nights (KEs) in the 4th phase : $n^{(1)}_{4,m}$.

Table 2. Estimates of the discriminant coefficients estimated with with the forward selection method by $F = 2$

Selected variable	Discriminant coefficient	Standard discriminant coefficient	F-value	p-value	Judge
Times black moves silver in the 4th phase	0.96	0.04	15.60	0.00	[**]
Skewess of pawn for black in the 3rd phase	1.54	0.29	3.60	0.06	[ ]
Times white moves nights in the 2nd phase	-20.04	-4.78	10.43	0.00	[**]
Variance of pawns for black in the 2nd phase	6.11	1.77	9.64	0.00	[**]
Variance of pawns for white in the 2nd phase	4.71	0.49	4.70	0.04	[*]
Times white moves nights in the 4th phase	-0.54	-0.11	2.28	0.14	[ ]
Constant	-6.70				

The followings are found from the above results:

- the probability that white players (the first move) win games increases if numbers of times black players (the defensive move) move silver generals (KIs) from the 31^{th} to the 40^{th} moves, skewness of row numbers where pawns (FUs) are put for black players (the defensive move) from the 21^{th} to the 30^{th} moves, variances of row numbers where pawns (FUs) are put for black players (the defensive move) from the 11^{th} to the 20^{th} moves and variances of row numbers where pawns (FUs) are put for white players (the first move) from the 11^{th} to the 20^{th} moves increase,
- the probability that black players (the defensive move) win games increase if numbers of times white players (the first move) move nights (KEs) from the 11^{th} to the 20^{th} moves and numbers of times white players (the first move) move nights (KEs) from the 31^{th} to the 40^{th} moves increase.

4 A Method for Conducting Shogi Postmortems Using Discriminant Scores Obtained from Discriminant Analysis of Kifu Data

In this section, we suggest a method for conducting shogi postmortems using discriminant scores obtained from discriminant analysis of Kifu data, and show the usability of the proposed method by introducing a software incorporated with the suggested method. Discriminant scores are calculated for every game and not ones indicating which player is superior at a certain position. However, in our data used for discriminant analysis, explanatory valuables are given by phase ($l = 1, 2, 3, 4$). So, we calculate discriminant scores separately by phase $l = 1, 2, 3, 4$ and game $m = 1, \cdots, 50$ like

$$z_{l,m} = a_{l,m}^{(\bar{p})} \bar{p}_l^{(k)} + a_{l,m}^{(\tilde{p})} \tilde{p}_l^{(k)} + a_{l,m}^{(\ddot{p})} \ddot{p}_l^{(k)} + a_{l,m}^{(g)} {}_m g_l^{(k)} + a_{l,m}^{(\mathrm{s})} {}_m s_l^{(k)} + a_{l,m}^{(\mathrm{n})} {}_m n_l^{(k)} + cl/4, \tag{5}$$

and consider $z_{l,m}$ as scores indicating which player is superior, that is, white player (the first move) is superior if $z_{l,m} \geq 0$ or black player (the defensive move) is superior if $z_{l,m} < 0$ at phase $l = 1, 2, 3, 4$. Further, we divide $z_{l,m}$ into ones concerning movement of pawns (FUs):

$$z_{l,m}^{(p)} = a_{l,m}^{(\bar{p})} \bar{p}_l^{(k)} + a_{l,m}^{(\tilde{p})} \tilde{p}_l^{(k)} + a_{l,m}^{(\ddot{p})} \ddot{p}_l^{(k)} \tag{6}$$

and one concerning movement of gold, silver generals (KIs, GIs) and nights (KEs) :

$$z_{l,m}^{(gsn)} = a_{l,m}^{(g)} {}_m g_l^{(k)} + a_{l,m}^{(\mathrm{s})} {}_m s_l^{(k)} + a_{l,m}^{(\mathrm{n})} {}_m n_l^{(k)} , \tag{7}$$

and use $z_{l,m}$, $z_{l,m}^{(p)}$ and $z_{l,m}^{(gsn)}$ as indicators for illustrating shogi postmortems.

Fig 1 indicate the images of developed software for illustrating shogi postmortems using $z_{l,m}$, $z_{l,m}^{(p)}$ and $z_{l,m}^{(gsn)}$. In the figure, the final match between Satoshi Murayama vs. Manabu Senzaki at the 14th Young Lion Tournament on the 22nd in October, 1991 is illustrated. Fig. 1 illustrates that the black player (the defensive move) was superior to the white player before his 20th move, and Fig 2 is illustrating that the white player became superior after the 60th move because of his way of moving pawns, from which the significance of the suggested method for illustrating shogi postmortems may be shown.

5 Concluding Remarks

In this article, average, variance and skewness of row numbers where pawns (FUs) are put, in addition to times gold and silver generals (KIs and GIs) and nights (KEs) were proposed as features of shogi games and their significance as features was shown by the result of discriminant analysis of actual Kifu data base on the proposed features. In addition, a software developed for illustrating shogi postmortems was introduced and usability of proposed features was shown.

Sawa and Ito [6] showed what kinds of features determined people's impression for play style base on their result of statistical analysis of Kifu data. In their study they use the numbers of times players move gold and silver generals (KIs and GIs) and nights (KEs). In this their study has something to do with our study. However, our way of studying is different from theirs in the meaning that we consider the numbers of times players move gold and silver generals (KIs and GIs) and nights (KEs) as a time series by dividing games into phases consisting of 10 moves.

The aim of this study illustrate important point or move determining winners of shogi games, average, variance and skewness of row numbers where pawns (FUs) are put and times gold and silver generals (KIs and GIs) and nights (KEs) are moved may be too simple to determine winner of games[7]. So our next aim is to seek other features to determine winners of games and incorporate such features into our software for illustrating shogi postmortems.

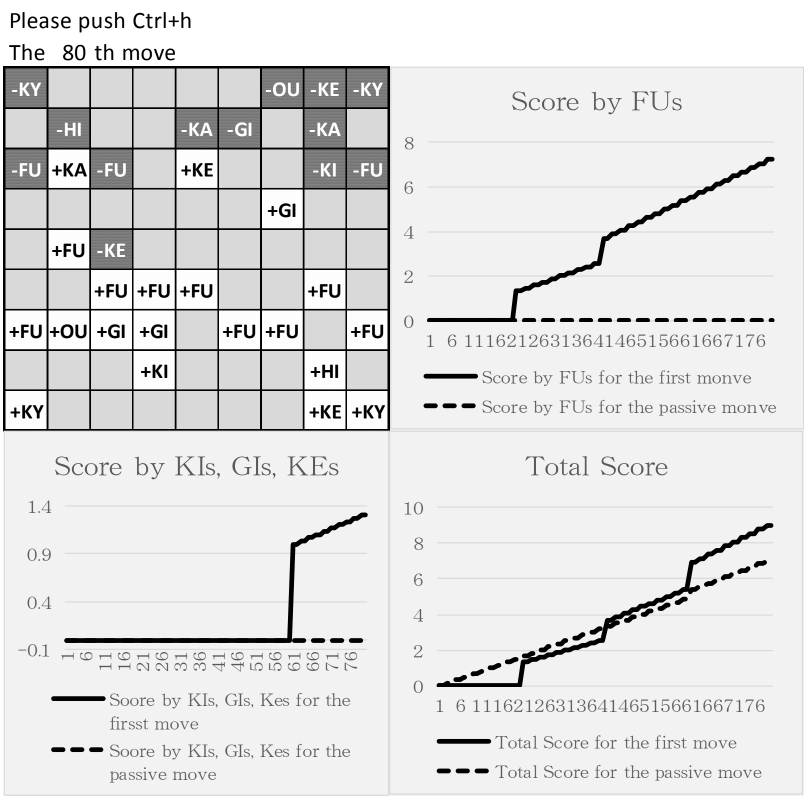

Fig 1. Illustration of the final match between Satoshi Murayama vs. Manabu Senzaki at the 14th Young Lion Tournament on the 22nd in October, 1991 at the 80th move

References

1. Kunihito Hoki and Tomoyuki Kaneko. “Large-Scale Optimization for Evaluation Functions with Minimax Search”. Journal of Artificial Intelligence 49, pp. 527-568 (2014).
2. Hitoshi Matsubara ed. “A special issue for the way of winning since 2010”. Journal of Information Processing, Vol.52、No.2, p.152-190 (2011) [in Japanse].
3. Takenobu Takizawa. “Contemporary Computer Shogi (May 2013)”. IPSJ SIG Technical Reports, Vol. 30, 2013, pp. 1 – 8 [in Japanse].
4. Yoshiyuki Kotani. “Looking Back on the 3rd Shogi Dennou-sen : 3. An Objective Analysis on the Strength of Computer Shogi - Did It Reach to the Human Top Player?”. IPSJ Magazine 55(8), 851-852 (2014) [in Japanese].
5. Kouki Kajiwara, Ryo Miura, Miroyuki Tarumi. “Searching Positions from Shogi Game Recores unsing Jargous”. The proceeding of Enternaimnet Computing Symposeum 2016, November 3-5, 2016, Osaka [in Japanese].
6. Nobushige Sawa, Takeshi Ito. “Statistical analysis of Elements of Play Style in SHOGI (Japanese Chess)”. Game Informatics, 26(3), 1-8 (2011) [in Japanese].
7. Yoshikuni Sato and Daisuke Takahashi. “Learning Weights of Training Data by Large Game Results”. The proceeding of game programing workshop (6), 22-29 (2012) [in Japanese].