

Chinese Chess File Format

Drafted by Huayong Yang

with input from Peter Sung, Xi Chen, Tie Zeng, and other people

(adapted from ICCS system, USCF notation system, and Smart-Go format)

[HY: my comments are enclosed in "[" and "]", beginning with "HY:". In the following, I explain what would be in the file line by line. OK, now the file starts with ...]

FORMAT format, piece_symbols

[HY: the format refers to the move format, i.e., one of the following four (taking one common move for example:)

ICCS (ICCS format): H2-E2
AXF (or conventional format): C2=5
LALG (long algebraic format): Ch2-e2
SALG (short algebraic format): Che2

I think AXF is the same as the one used in XiangQi Review, which is also referred to as XQR notation. For definition of XQR notation, see Appendix A in the end of this file. The long/short algebraic formats are adapted from the counterpart in chess. See relevant chess notation documents for more information. If nothing is specified, then the default "ICCS" is assumed (this is reasonable because the format H2-E2 doesn't use any piece symbols, thus avoids the trouble that the "piece_symbols" part in FORMAT line has to deal with.)

If AXF, LALG, or SALG format is used, then the remark symbol (! for "good", ? for "questionable", + for "check", !! for "excellent", !? for "uncertain", ?? for "very bad", ++ for "checkmate") may be used following the move, e.g., C2=5!!!.

The "piece_symbols" part in the FORMAT line refers to the piece symbols to be used. The defaults are those used on ICCS, i.e., K for King, R for Rook, N for Knight, C for Cannon, G for Guard, M for Minister, P for Pawn. If one chooses to use symbols other than the defaults, he should indicate that like this:

FORMAT AXF, K(K) R(C) N(H) C(P) G(A) M(E) P(S)

which means he is using AXF format and the piece symbols specified in brackets. That is, he uses which corresponds to symbol: default symbol: K K C R H N P C A G E M S P

In other words, he uses C for Chariot (rook), H for Horse (knight), P for Pao (cannon), A for Assistant (guard), E for Elephant (minister), and S for Soldier (pawn).

Some people like using lower cases for black pieces and upper for red. However, to achieve flexibility, I'd better not rely on this (but I will accept that form).

There should be **at most** one FORMAT line in a file. If the FORMAT line is missing, the default "ICCS" is assumed.]

```
GAME game_number, game_title
```

[HY: the game_number numbers the game in a file. This allows multiple games in one file. For example,

```
GAME 1, ppfast.wshen.1
GAME 2, wshen.ppfast.1
```

NOTE: every game must start with this line. The game_number and/or game_title may be omitted, however. The following are two examples:

```
GAME opening study
GAME
```

]

```
TIME time
RESULT ?-? (reason)
```

[HY: the time format will follow what ICCS does. ICCS uses something like R+Resign to indicate the result. This is fine except for two occasions where the game is ajourned or a draw. Then it is not clear to me if R or B should be used (talking about symmetry!) So I think the format

```
1-0 (Resign) (if red wins by resignation)
0.5-0.5 (if it is a draw)
0-0 (if it is ajourned, or incomplete)
0-1 (Checkmate) (if black wins by checkmate)
```

is better. It is also useful to indicate the points awarded to both players, in particular, in such a case where a win gets 3 pts, a draw 1 pts, and a loss 0 pts.

In case of a draw, 0.5-0.5 can be shortened to 0.5 and the reason can be omitted. Actually the reason "Adjourned" can be left out too, I think. But it is clearer if

it stays.

The "reason" can be the following:

```
Resign
Checkmate
Forfeit on time
Draw
Adjourned
```

]

```
RED name; rating; title
BLACK name; rating; title
```

[HY: the reason I use ";" as the delimiter is that we may need the comma to separate last name and first name. If one field is missing, use blank instead. For example,

```
RED Hu, Ronghua; ; GM
```

In the above, Hu Ronghua's rating is left as blank.]

```
DATE xxxx-xx-xx
PLACE place
EVENT event
```

[HY: DATE and PLACE are just like the ones used in ICCS files. I add one more item EVENT. The event could be something like "US-CA-TW Leitai".]

[HY: so much for the game header. To summarize, a complete header has the form:

```
GAME xx, xxxxxxxx
TIME xxxx xxxx xxxx xxxx
RESULT xx-xx (xxxx)
RED xxxx; xxxx; xx
BLACK xxxx; xxxx; xx
DATE xxxx-xx-xx
PLACE xxxxxxxxxxxxxxxxxxxxxxxx
EVENT xxxxxxxxxxxx
```

The GAME line is required; others are optional.

Following the game header is the game body, which has the form:

```

SETUP{ ... }
START{ ... }

```

The SETUP portion is optional. More on the game body below. For the time being, let's look at the basic layout of a file:

```

FORMAT xxx

```

```

GAME
START{ ... }

```

```

GAME
START{ ... }

```

That example file contains two game records. The following is a concrete example file, followed by detailed explanations. *NOTE*: some specifications have been changed since the file is written.]

```

----- game file -----

```

```

FORMAT LALG

```

```

GAME 1, Wind Screen Horse #3

```

```

RED C Ng

```

```

BLACK C Ng

```

```

DATE 1994-03-20

```

```

DIAG{  a b c d e f g h i

```

```

1 2 3 4 5 6 7 8 9

```

```

9 r n m g k g m n r 9

```

```

8 . . . . . 8

```

```

7 . c . . . . c . 7

```

```

6 p . p . p . p . p 6

```

```

5 ----- 5

```

```

4 ----- 4

```

```

3 P . P . P . P . P 3

```

```

2 . C . . . . C . 2

```

```

1 . . . . . 1

```

```

0 R N M G K G M N R 0

```

```

9 8 7 6 5 4 3 2 1

```

```

a b c d e f g h i

```

```

Red to move (Move 1)

```

```

}

```

```

SETUP{

```

```

Ra0; Nb0; Mc0; Gd0; Ke0; Gf0; Mg0; Nh0; Ri0; Cb2; Ch2; Pa3;

```

```

Pc3; Pe3; Pg3; Pi3; pa6; pc6; pe6; pg6; pi6; cb7; ch7; ra9;

```

```

nb9; mc9; gd9; ke9; gf9; mg9; nh9; ri9; MOVE 1, RED }

```

```

START{

```

```

1. Ch2-e2 Nh9-g7 2. Nh0-g2 Ri9-h9 3. Ri0-h0 Pg6-g5 4. Rh0-h6 Nb9-c7

```

```

5. Pc3-c4 ...

```

COM{ The 7th soldier variation is one the largest, and most complex opening in the central cannon vs. wind-screen horse opening. As the next few moves will show, it can lead to:...}

5. ... Ch7-i7 6. Rh6-g6 ...

COM{ There is also a possibility of exchanging rooks here. But of course, when the rooks are exhcnges that early in the game, the game will become more subdued. }

6. ... Ci7-i8 7. Nb0-c2 Ci8-g8 8. Rg6-f6 Gd9-e8 9. Nc2-d4 Mc9-e7
10. Cb2-c2 Rh9-h4 11. Ra0-b0 ...

DIAG{

```

a b c d e f g h i
1 2 3 4 5 6 7 8 9
9 r . . . k g m . . 9
8 . . . . g . c . . 8
7 . c n . . . n . . 7
6 p . p . p R . . p 6
5 -----p-r-- 5
4 ----P-N----- 4
3 P . . . P . P . P 3
2 . . C . C . N . . 2
1 . . . . . . . . 1
0 . R M G K G M . . 0
9 8 7 6 5 4 3 2 1
a b c d e f g h i
Black to move (Move 11)

```

}

COM{ see Diagram. Red horse patrolling the river is a very powerful piece.

Black must remove this horse from the center of the board, thereby distrubing it in move 10. Red see to exchange black's cannon in return. }

11. ... Cb7-b8 (1) 12. Pg3-g4 ... # if R87, see variation

12. ... Rh4-h5 (2) 13. Pc4-c5 Pg5xg4 14. Pc5xc6 Nc7-a8 15. Cc2-d2
Rh5-c5

16. Ng2-e1 Rc5-c3 17. Ne1-c2 Rc3-d3 18. Gf0-e1 ...

COM{ Both sides have 1 soldier across the river.

However, red's 3th horse can switch to the 7th file easily whereas black's

3th horse does not have a good position. Red's position is better.}

VAR[1]{

11. ... Ra9-b9 12. Pg3-g4 Rh4xg4 (3) 13. Ng2-f4 Cb7-b4 (4) 14. Mg0-i2 Rg4-h4
15. Ce2-g2 Ng7-h5 16. Cg2xg8 Pg5-g4 17. Rf6-f8 Nh5xf4 18. Nd4-f5 Rh4-h2
19. Nf5xe7 Rh2xc2 20. Cg8-g7 Rb9-b5 21. Ne7-g8 Ke9-d9 22. Cg7xg9 Kd9-d8
23. Ng8-h6 # Red sacrifices a piece to gain attack.

}

VAR[3]{

12. ... Rh4-h5 13. Nd4xc6 Pg5xg4 14. Nc6-a7 Nc7-b5 15. Rb0xb5 Rh5xb5
16. Na7xb9 Cb7-b8 17. Ce2-d2 Pg4-g3 18. Ng2-i1 Rb5-b3 19. Cd2-d8 ...

COM{ Red destroys black's defence by exchanging one of black's annons which form an integral part of the wind-screen horse defence. }

```

19. ... Cg8xd8 20. Nb9xd8 Rb3xe3 21. Gf0-e1 Re3-c3 22. Cc2-b2 Rc3xc0 (5)
23. Rf6-g6 # Red gains material.
}
VAR[5]{
22. ... Cb8-b6 23. Rf6-f4 Rc3xc0 24. Nd8-b7 ...
COM{ Red's pieces can organize attack easily. }
}
VAR[4]{
13. ... Rg4xg0 14. Ce2-g2
}
VAR[2]{
12. ... Rh4xg4 13. Ng2-f4 Rg4xg0 14. Ce2-g2 ...
COM{ forking black's pieces on 7th file.}
}
}
}
----- end of game file -----

```

The above game file demonstrates almost all major parts of a game body:

```
DIAG{ something }
```

[HY: the "something" is really arbitrary, although most likely it is a diagram to enhance the readability of the file. This part can be virtually anywhere in the game body, but not inside a unit. A "unit" refers to a keyword, a move, and SETUP{ }.]

```
SETUP{ piece(location); ... ; MOVE move_number,
RED|BLACK }
```

[HY: "piece" is one of those defined in the piece_symbols part in the FORMAT line. Use upper case for red and lower case for black. Alternatively, one can do this: SETUP{ RED: Ra0, Nb0, Mc0, Gd0, Ke0, MOVE 1; BLACK: Nb9, Mc9, Gd9, Ke9, Gf9 } without relying on cases. The move number indicates the number of the move to be played. And the RED or BLACK that is associated with the move number is the one to make his half-move. Note the difference between "move" and "half-move". If the SETUP{ } component is missing, the normal board position is assumed.]

```
START{ ... }
```

[HY: inside the brackets there could be moves, variations, and comments. See below.]

Comments in the game:

[HY: there are two types of comments: one is one-line comments, preceded by # and ended by a newline; another is multi-line comments, enclosed by /* and */. In the above file I used COM{ and } to enclose multi-line comments, but now I feel like

using `/*` and `*/`. The reason is, with multi-line comments, we have to take care of escaping `}` or `*/` in the comments (and I'd like to use backslash `\` to escape them). It seems to me `}` is more likely to show up in people's comments than `*/`. Thus using `*/` reduces the occurrence of the annoying backslashes. By the way, I will use `\` to escape any `}`'s inside the normal `{` and `}`. And a space should be left between the last character of the comments and `*/` so that one does not have to escape `\` too! Anything lead by `#` or enclosed by `/*` and `*/` are assumed to be the comments made on the move right before them (same as ICCS file). There is a tricky point to be noted. For example,

```
START{
some_moves
12. redmove12 blackmove12
VAR[5]{
some_moves
15. redmove15 ...
VAR[7]{
15. ... blackmove15
some_moves
19. redmove19 blackmove19
}
/* something */
}
```

} then the "something" comments are made on blackmove15 in VAR[5], not on blackmove19 in VAR[7] (because "something" is outside VAR[7]{...}) nor on blackmove12.]

Moves in the game:

[HY: the format is `move_number. red_move black_move` where `red_move` and `black_move` will follow the format defined in the `FORMAT` line (in the case of the above file, the format is defined as `LALG`, or long algebraic.), e.g., `23. Rf6-f4! Rc3xc0` (note that remark symbols can be used for every half-move).

If a move is incomplete (due to an immediate comment or variation, etc.), the unplayed part is filled by "...", e.g., `12. ... Rh4xg4`

If a move is followed by `(a_number)`, which is a combination of `"(`, a number, and `)"`, like `12. Pg3-g4 Rh4xg4 (3)`, then it indicates the move has a variation, numbered 3 in this game. Two things should be noted in this case. Firstly, a variation number is unique in a game (but not necessarily in a file). I once used a form like `(move_number, variation_number)` (e.g., `(12,1)`), but it soon turned out to be a bad idea because there could be the same `move_number` in different variations. Secondly, the first move in a variation is to replace the move right before the variation number. For instance,

```
12. Pg3-g4 Rh4xg4 (3)
...
```

VAR[3]{
 12. ... Rh4-h5 13. Nd4xc6 Pg5xg4 14. Nc6-a7 Nc7-b5 15. Rb0xb5 Rh5xb5 in this case,
 Rh4-h5 replaces Rh4xg4, or we can say Rh4-h5 "varies" the second half of move 12.

If a move has more than one variations, the numbers are listed in the brackets
 (and). I used to support the range expression (e.g. (1-4)), but now I feel it is
 not worth the trouble.

A move can be followed immediately by its variations (not shown in the above file).
 In that case, I think the form

12. Pg3-g4 Rh4xg4 {12. ... Rh4-h5 13. Nd4xc6 Pg5xg4 } { ... } would be fine.
 Note that the keyword VAR is dropped and a variation can be immediately followed by
 another variation.]

Variations in the game:

[HY: a variation is enclosed by { and }. If there is no keyword VAR in front of {,
 then it is assumed to "vary" the move right preceding it. If VAR is used, it must be
 followed by "[", a variation number, and "]". Other than that, a variation is just
 like a game body (hence recursive relationship).]

Appendix A: (XiangQi Review notation, forwarded by Stephen Leary)

Each half move is defined using 4 digits. The first digit tells the piece to be
 moved. The second digit is the file which the piece is on (numbered 1 to 9 from each
 player's right to left). In the case of more than one piece of the same kind on the
 same file, the front (+), back (-), or middle (=). However, this does not apply if
 there are two guards or bishops on the same file, as when one
 moves, there is no ambiguity when using the standard notation. The third character
 tells the direction of the move with reference to the player's own baseline. It can
 be up (+), down (-), or horizontal (=). The fourth digit will be the file the piece
 will end up on, or the number of steps it moves up or down.

If there are 2 sets of 2 pawns on the same file, these pawns are labeled with the
 letters A, B, C, D. They are labeled from RED's point of reference from the RIGHT,
 from North to South. So if black has 2 pawns on his 3rd file and 2 pawns on his 5th
 file, they would be labeled from red's viewpoint. A pawns would be the northmost
 pawn on the 5th file. B would be other 5th file pawn.
 C would be northmost pawn on 3rd file (which is red's 7th file). D would be the
 south pawn on the 3rd file. If you have 4 or 5 pawns on one file, they are labeled
 the same, from red's viewpoint, A-D. However, once the pawns move and there is only
 3 pawns left on a file, then you revert to the regular (+), (-), or (=) notation
 when one of them moves. You only use the letters when you have more than 3 pawns on
 a file, or 2 sets of 2 pawns on 2 files.

[HY: according to Stephen Leary, the 3-2 situation is definitely "2 sets of stacked
 pawns", so one would have to assign letters to all 5. However, he said he was not
 100% sure.]