

Universal Multiple-Octet Coded Character Set
 International Organization for Standardization
 Organisation internationale de normalisation
 Международная организация по стандартизации

























Doc Type: Working Group Document**Title: Proposal to add standardized variation sequences for chess notation****Source: Michael Everson and Garth Wallace****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2017-04-01****1. Introduction**

The orthodox chess pieces have been encoded since Unicode 1.1. Perhaps surprisingly, most currently available chess fonts do not make use of these characters: they either use ASCII and Latin-1 code positions, or they use Private Use Area code positions via the old MS Windows Symbols encoding mapped to U+F000..F0FF. It would appear that the primary reason chess fonts do not use UCS characters is that there is no standardized mechanism to use these characters to prepare chess diagrams, which require chesspieces to be displayed on both white and black squares. From time to time suggestions have been made regarding encoding a set of chess-characters-on-black but the response has normally been “use a higher-level protocol”. No robust and interchangeable protocol of that kind exists, and the older fonts have not been replaced. This document proposes that standardized variation sequences could solve the problem, making chess diagrams interchangeable, permitting the parsing of chess diagram data, and facilitating simple font changes.

























2. Current font implementations

There are many chess fonts available and have been since the early 1990s. One repository for these is at <http://www.enpassant.dk/chess/fonteng.htm>. A variety of different encodings for chessboard notation are used in those fonts. This lack of standardization does not benefit the chess community, and prevents easy interchange of chess problems. Most of the chess fonts used currently do not make use of the characters encoded at U+2654..265F.

























The Chess Leipzig font by prolific chess-font designer Armando H. Marroquin is a classic font inspired by a German chess book from the beginning of the 20th century. It uses PUA code positions with the MS Windows Symbols encoding, which maps them (or at least did in older environments) to ASCII characters, so that people could type them easily. Thus a white king on a white square would be “k”, a white king on a black square “K”, with the corresponding black kings on “l” and “L”, adjacent on the keyboard. Queen is on “q” and “Q” and adjacent “w” and “W”; rook is on “r” and “R” and adjacent “t” and “T”; bishop is on “b” and “B” and adjacent “v” and “V”; knight is on “n” and “N” and adjacent “m” and “M”; and pawn is on “p” and “P” and adjacent “o” and “O”.

	F06B (k)		F04B (K)		F06C (l)		F04C (L)
	F071 (q)		F051 (Q)		F077 (w)		F057 (W)
	F072 (r)		F052 (R)		F074 (t)		F054 (T)
	F062 (b)		F042 (B)		F076 (v)		F056 (V)
	F06E (n)		F04E (N)		F06D (m)		F04D (M)
	F070 (p)		F050 (P)		F06F (o)		F04F (O)

























The Chess Berlin font by Eric Bentzen is based on the familiar design from the East German “Sportverlag”, which published many popular chess books. It does the same thing, but the black bishop, white knight, and black knight are mapped to *different* PUA characters (and ASCII equivalents). A chart set in one of the fonts will be corrupted if the font is changed from Leipzig to Berlin or vice-versa. King is on “k” and “K” and adjacent “l” and “L”; white queen is on “q” and “Q” and black queen on adjacent “w” and “W”; white rook is on “r” and “R” and black rook on adjacent “t” and “T”; white bishop is on “b” and “B” and black bishop on adjacent “n” and “N”; white knight is on “h” and “H” and black knight on adjacent “j” and “J”; and white pawn is on “p” and “P” and black pawn on adjacent “o” and “O”.

	F06B (k)		F04B (K)		F06C (l)		F04C (L)
	F071 (q)		F051 (Q)		F077 (w)		F057 (W)
	F072 (r)		F052 (R)		F074 (t)		F054 (T)
	F062 (b)		F042 (B)		F06E (n)		F04E (N)
	F068 (h)		F04E (H)		F06A (j)		F04D (J)
	F070 (p)		F050 (P)		F06F (o)		F04F (O)

























Alastair Scott’s Chess font, based on a font called Cheq, follows the same encoding that the Chess Berlin font does.

	F06B (k)		F04B (K)		F06C (l)		F04C (L)
	F071 (q)		F051 (Q)		F077 (w)		F057 (W)
	F072 (r)		F052 (R)		F074 (t)		F054 (T)
	F062 (b)		F042 (B)		F06E (n)		F04E (N)
	F068 (h)		F04E (H)		F06A (j)		F04D (J)
	F070 (p)		F050 (P)		F06F (o)		F04F (O)

The Chess Utrecht font by Hans Bodlaender is modern and stylized, with a rather different encoding. White king is on “k” and “l” and black king on “K” and “L”; white queen is on “q” and “w” and black queen on “Q” and “W”; white rook is on “r” and “t” and black rook on “R” and “T”; white bishop is on “b” and “n” and black bishop on “B” and “N”; white knight is on “n” and “m” and black knight on “N” and “M”; and white pawn is on “p” and “o” and black pawn on “P” and “O”.

























	006B (k)		F06C (l)		F04B (K)		F04C (L)
	F071 (q)		F077 (w)		F051 (Q)		F057 (W)
	F072 (r)		F074 (t)		F052 (R)		F054 (T)
	0062 (b)		F06E (n)		F042 (B)		F04E (N)
	F06E (n)		F06D (m)		F04E (N)		F04D (M)
	F070 (p)		F06F (o)		F050 (P)		F04F (O)

The Chess Kingdom font, also by Armando H. Marroquin, uses the same encoding as his Chess Leipzig.


	F06B (k)		F04B (K)		F06C (l)		F04C (L)
	F071 (q)		F051 (Q)		F077 (w)		F057 (W)
	F072 (r)		F052 (R)		F074 (t)		F054 (T)
	F062 (b)		F042 (B)		F076 (v)		F056 (V)
	F06E (h)		F04E (H)		F06D (m)		F04D (M)
	F070 (p)		F050 (P)		F06F (o)		F04F (O)




The Chess Skak font by Egon Madsen uses actual ASCII characters (not MS Windows Symbol characters), Danish piece names Konge, Dronning, Tårn, Løber, Springer, Bonde inspire the mapping.

























Thus white king is on “k” and “i” and black king on “K” and “I”; white queen is on “d” and “e” and black queen on “D” and “E”; white rook is on “t” and “y” and black rook on “T” and “Y”; white bishop is on “l” and “o” and black bishop on “L” and “O”; white knight is on “s” and “w” and black knight on “S” and “W”; and white pawn is on “b” and “g” and black pawn on “B” and “G”.

	006B (k)		0069 (i)		004B (K)		F049 (I)
	0064 (d)		0065 (e)		0044 (D)		F045 (E)
	0074 (t)		0079 (y)		0054 (T)		F059 (Y)
	006C (l)		006F (o)		004C (L)		F04F (O)
	0073 (s)		0077 (w)		0053 (S)		F057 (W)
	0062 (b)		0067 (g)		0042 (B)		F047 (G)

The Chess Diagramm Pirat font by Klaus Wolf is encoded using the same PUA range that the MS Windows Symbol fonts are mapped to, but it simply puts them in code-chart order vis à vis ASCII (though there are no graphic characters in the range U+0082..008C).

	F072 (r)		F078 (x)		F07E (~)		F087
	F073 (s)		F079 (y)		F082		F088
	F074 (t)		F07A (z)		F083		F089
	F075 (u)		F07B (}{)		F084		F08A
	F076 (v)		F07C (l)		F085		F08B
	F077 (w)		F07D (})		F086		F08C

The 1Echecs font handles the problem in a completely different way. First, the white queen, white rook, and black rook are composed by using left- and right-half pieces (the font has these to construct some characters used in Fairy Chess). Apart from that, there is some mapping to French piece names (Roi, Dame, Tour, Fou, Cavalier, Pion). The shading for black squares is likewise implemented in two parts, so for the white king, for instance, the sequence F030  + F031  + F072  is used.

	F072 (r)		F030 F031 F072 (01r)		F052 (R)		F030 F031 F052 (01R)
	F064 F065 (de)		F032 F033 F064 F065 (23de)		F044 (D)		F032 F033 F044 (23D)
	F0E8 F0E9 (èé)		F034 F035 F0E8 F0E9 (45èé)		F0EA F0BF (êï)		F034 F035 F0EA F0BF (45êï)
	F066 (f)		F036 F037 F066 (67f)		F046 (F)		F036 F027 F046 (67F)
	F063 (c)		F038 F039 F063 (89c)		F043 (C)		F038 F039 F043 (89C)
	F070 (p)		F028 F029 F070 (>()xp)		F050 (P)		F028 F029 F050 (>()P)

3. Proposed variation sequences

Standardized variation sequences offer a solution to this glyph-level alignment ambiguity by using one variation selector, such as VS1 (U+FE00), to indicate pieces on a white square, and another, such as VS2 (U+FE01), to indicate pieces on a black square. Pieces for both chess and draughts are given, as are two Geometric Shapes which are to be used to represent the board squares. A font with appropriate entries in its Format 14 (Unicode Variation Sequences) ‘cmap’ subtable can enable these distinctions to be shown and preserved in plain text. For some applications “liga” or “rlig” tables may be used, since support for Format 14 tables is not universally implemented. Below is a complete list of the proposed sequences as they would appear in the StandardizedVariants.txt file.










































```
# Chesspiece on white versus Chesspiece on black variation sequences
25A1 FE00; White chessboard square; # WHITE SQUARE
25A8 FE01; Black chessboard square; # SQUARE WITH UPPER RIGHT TO LOWER LEFT FILL
2654 FE00; Chesspiece on white; # WHITE CHESS KING
2654 FE01; Chesspiece on black; # WHITE CHESS KING
2655 FE00; Chesspiece on white; # WHITE CHESS QUEEN
2655 FE01; Chesspiece on black; # WHITE CHESS QUEEN
2656 FE00; Chesspiece on white; # WHITE CHESS ROOK
```

```

2656 FE01; Chesspiece on black; # WHITE CHESS ROOK
2657 FE00; Chesspiece on white; # WHITE CHESS BISHOP
2657 FE01; Chesspiece on black; # WHITE CHESS BISHOP
2658 FE00; Chesspiece on white; # WHITE CHESS KNIGHT
2658 FE01; Chesspiece on black; # WHITE CHESS KNIGHT
2659 FE00; Chesspiece on white; # WHITE CHESS PAWN
2659 FE01; Chesspiece on black; # WHITE CHESS PAWN
265A FE00; Chesspiece on white; # BLACK CHESS KING
265A FE01; Chesspiece on black; # BLACK CHESS KING
265B FE00; Chesspiece on white; # BLACK CHESS QUEEN
265B FE01; Chesspiece on black; # BLACK CHESS QUEEN
265C FE00; Chesspiece on white; # BLACK CHESS ROOK
265C FE01; Chesspiece on black; # BLACK CHESS ROOK
265D FE00; Chesspiece on white; # BLACK CHESS BISHOP
265D FE01; Chesspiece on black; # BLACK CHESS BISHOP
265E FE00; Chesspiece on white; # BLACK CHESS KNIGHT
265E FE01; Chesspiece on black; # BLACK CHESS KNIGHT
265F FE00; Chesspiece on white; # BLACK CHESS PAWN
265F FE01; Chesspiece on black; # BLACK CHESS PAWN
26C0 FE00; Draughts piece on white; # WHITE DRAUGHTS MAN
26C0 FE01; Draughts piece on black; # WHITE DRAUGHTS MAN
26C1 FE00; Draughts piece on white; # WHITE DRAUGHTS KING
26C1 FE01; Draughts piece on black; # WHITE DRAUGHTS KING
26C2 FE00; Draughts piece on white; # BLACK DRAUGHTS MAN
26C2 FE01; Draughts piece on black; # BLACK DRAUGHTS MAN
26C3 FE00; Draughts piece on white; # BLACK DRAUGHTS KING
26C3 FE01; Draughts piece on black; # BLACK DRAUGHTS KING

```

The table below demonstrates an actual implementation—using an OpenType chess font with an appropriately-built Format 14 ‘cmap’ subtable—that uses VS1 and VS2 as described above for all of the eighteen characters in this proposal.

<i>Code</i>	<i>Char.</i>	<i>VS1</i>	<i>VS2</i>
25A1		—	—
25A8		—	
2654			
2655			
2656			
2657			
2658			
2659			
265A			
265B			
265C			
265D			
265E			
265F			
26C0			
26C1			
26C2			
26C3			

4. Game board borders

In addition to these characters, a chess font should contain eight Block Element characters, designed to match the width and height of the font’s board square. These can draw a border around the diagram. A

chess font may draw these as multiple rules if desired. To draw a box around a single row of half a chessboard with a knight on one square, the following characters would be used.

Top line: 2597 2581 2581 2581 2581 2596
 Middle line: 2595 25A1 FE00 25A8 FE01 265E FE00 25A8 FE01 258F
 Bottom line: 259D 2594 2594 2594 2594 2598



The eight Block Elements used for creating chessboard border rules are existing characters and do not require a variation-selector mechanism. They also map one-to-one to the ASCII and MS Windows Symbol characters used for creating border rules in pre-Unicode legacy fonts. See Figure 3.

2581	—	LOWER ONE EIGHTH BLOCK	(maps to LOW LINE _ in the Skak font)
258F		LEFT ONE EIGHTH BLOCK	(maps to VERTICAL LINE)
2594	—	UPPER ONE EIGHTH BLOCK	(maps to - HYPHEN-MINUS)
2595		RIGHT ONE EIGHTH BLOCK	(maps to \ REVERSE SOLIDUS)
2596	.)	QUADRANT LOWER LEFT	(maps to) RIGHT PARENTHESIS)
2597	.)	QUADRANT LOWER RIGHT	(maps to 9 DIGIT NINE)
2598	.)	QUADRANT UPPER LEFT	(maps to = EQUALS SIGN)
259D	.)	QUADRANT UPPER RIGHT	(maps to 0 DIGIT ZERO)

5. Figures

Figure 1. Diagram for the initial position in Turkish Draughts (Dama), set in Ludus in 24 points with 26-point leading using Variation Sequences.

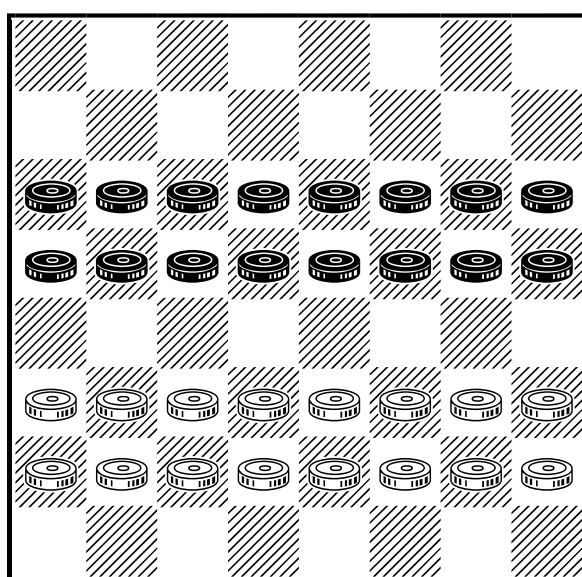
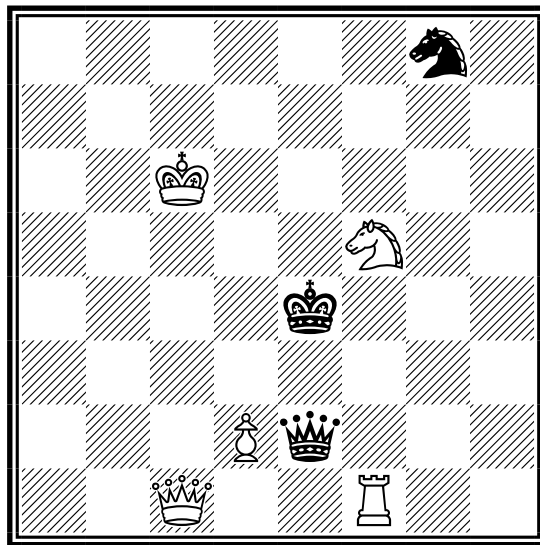


Figure 2. The example below demonstrates the Chess Condal font using Variation Sequences to represent a chessboard diagram, and also to represent a passage of text with inline chess characters without any variation selectors. The table itself is set in 24 points on 24-point leading.

White Pawn (Alice) to play, and win in eleven moves.

RED



WHITE

- | | |
|---|--|
| 1. Alice ♖ d2 meets Red Queen ♛ | 1. Red Queen ♛ e2–h5 |
| 2. Alice ♖ d2–d3 (<i>by railway</i>) | 2. White Queen ♛ c1–c4 (<i>after shawl</i>) |
| ♖ d3–d4 (<i>Tweedledum and Tweedledee</i>) | |
| 3. Alice ♖ d4 meets White Queen ♛ (<i>with shawl</i>) | 3. White Queen ♛ c4–e5 (<i>becomes sheep</i>) |
| 4. Alice ♖ d4–d5 (<i>shop, river, shop</i>) | 4. White Queen ♛ c4–f8 (<i>leaves egg on shelf</i>) |
| 5. Alice ♖ d5–d6 (<i>Humpty Dumpty</i>) | 5. White Queen ♛ f8–c8 (<i>flying from Red Knight</i>) |
| 6. Alice ♖ d6–d7 (<i>forest</i>) | 6. Red Knight ♞ g8–e7+ (<i>check</i>) |
| 7. White Knight ♞ f5 x e7 takes Red Knight ♞ | 7. White Knight ♞ e7–f5 |
| 8. Alice ♖ d7–d8 (<i>coronation</i>) | 8. Red Queen ♛ h5–e8 (<i>examination</i>) |
| 9. Alice ♖ becomes Queen ♛ | 9. Queens ♛ ♛ castle |
| 10. Alice ♛ d8 castles (<i>feast</i>) | 10. White Queen ♛ c8–a6 (<i>soup</i>) |
| 11. Alice ♛ takes Red Queen ♛ & wins | |

Figure 3. Examples of the Looking-Glass diagram set in plain text in an ASCII-encoded font (the display font on the left is Courier in 14 points on 17-point leading, tracked to give a more square impression; on the right the font is Chess Skak in 14 points on 17-point leading). Below them is the same diagram set in plain text in Chess Condal in 20 points on 20-point leading, displayed right-justified and encoded without variation selectors on the left, and with them on the right. And below that is the same again, set in Ludus in 20 points with 22-point leading, on the left force-justified with OpenType features turned off (with the font showing glyphs for the variation selectors) and on the right with OpenType features turned on. The tables set in Condal and Ludus are can still be read in plain text, though the result is not beautiful.

```

9  _ _ _ _ _ _ _ _ )
   | p F p F p F S F \
   | F p F p F p F p \
   | p F k F p F p F \
   | F p F p F s F p \
   | p F p F K F p F \
   | F p F p F p F p \
   | p F p g D F p F \
   | F p e p F t F p \
0  - - - - - - - - =

```

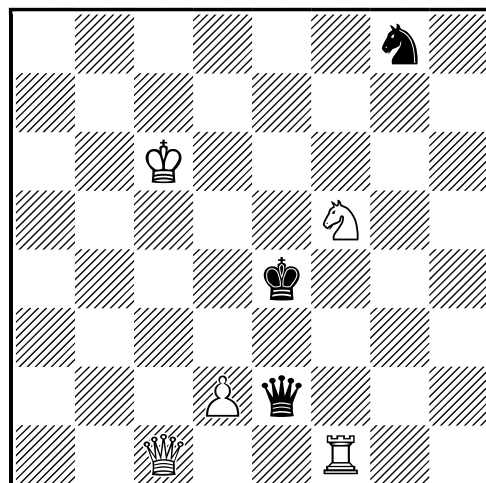
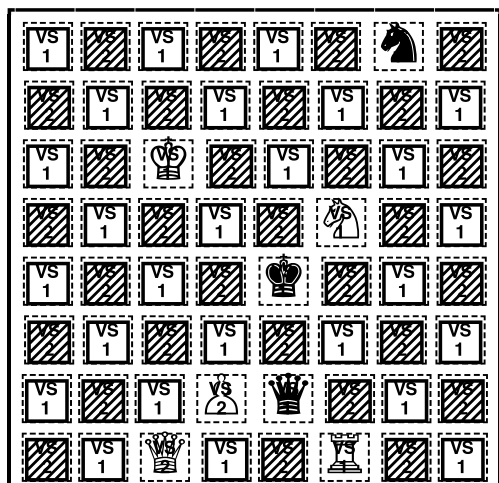
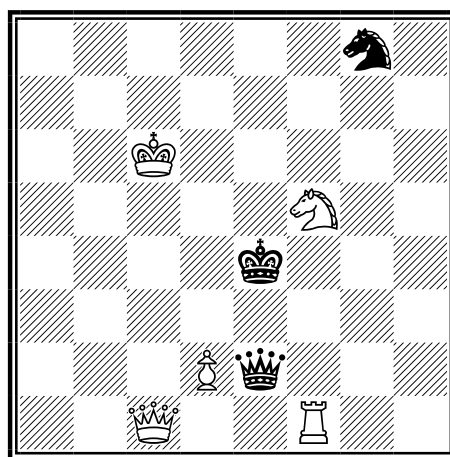
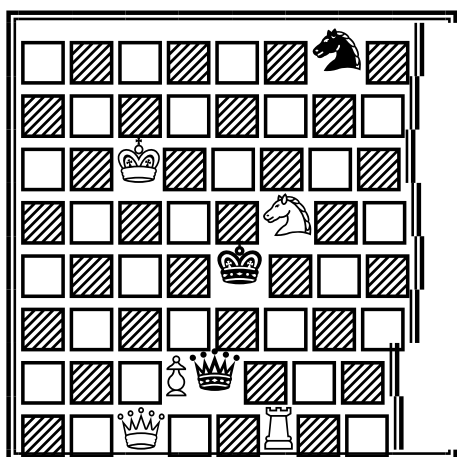
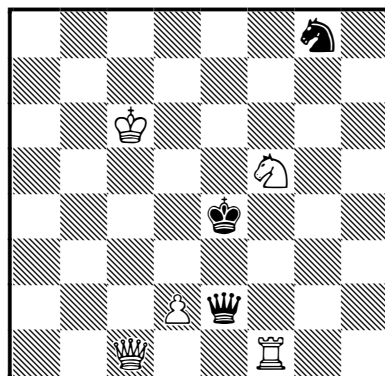
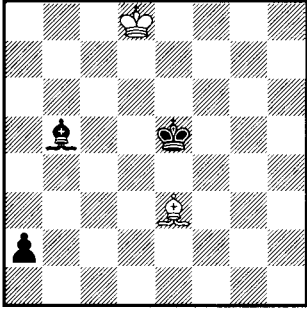


Figure 4. Above, two problems from *Enciklopedija Šahovskih Završnica*, 1993. Below, these have been re-set below using two fonts implementing the Variation Sequences, Ludus on the left (18pt on 18pt leading) and Chess Condal on the right (20pt on 20pt leading).

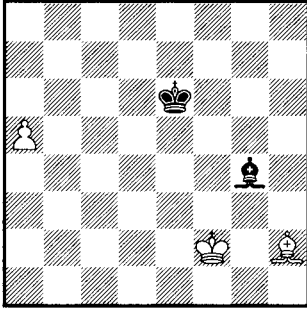
♞♚ **6/a**

1327. GRAAFLAND*
Tidskrift 1916



1. ? =

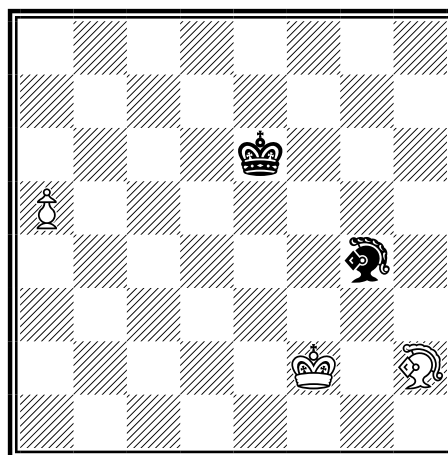
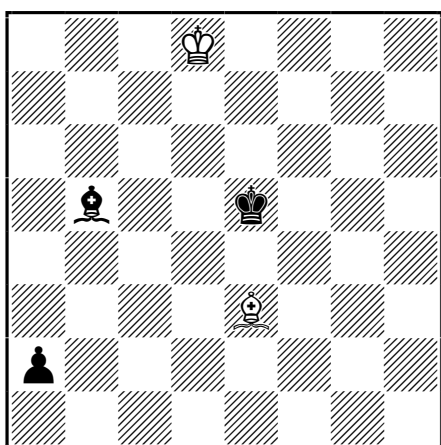
1328. BERGER
1922



1. ? =

1327. **1.** ♖h6 ♜f6 **2.** ♖d2 a1♚ **3.**
♖c3=

1328. **1.** a6 ♖f5 **2.** ♜f3 [2. ♜e3
♖h3 **3.** ♜f3 ♖f1 **4.** a7 ♖c4 △ ♖d5=]
♖d3! [2... ♜d5? **3.** a7 ♖e4 **4.** ♜e3+-]
3. a7 ♖c4 **4.** a8♚ ♖d5=



Interchange of data in this format will facilitate font change for presentation, but also provide a regularly formatted plain text which can be analysed; game-play positions could also be reciprocally generated from base data on moves and turns. This would be of great benefit to chess enthusiasts.

Figure 5. Diagram for a problem in Canadian Draughts. This larger-than-usual board size is no problem to set in plain text, displayed with variation sequences.

