

JTC1/SC2/WG2 N4719R
2016-07-11

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 encode symbols for Go game notation
Source: Andrew West
Status: Individual Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Date: 2016-07-11

1. Introduction

The Unicode and ISO/IEC 10646 standards include four characters used for marking stones in the game of Go (known as Weiqi in Chinese and Baduk in Korean):

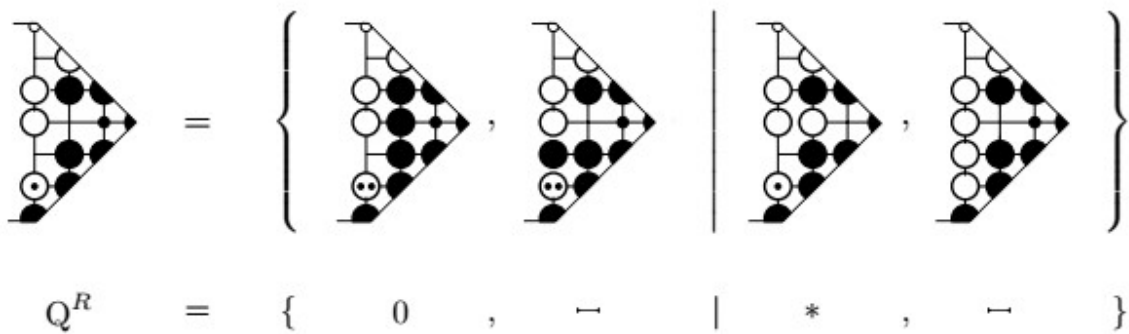
- U+2686 ◯ WHITE CIRCLE WITH DOT RIGHT
- U+2687 ◐ WHITE CIRCLE WITH TWO DOTS
- U+2688 ● BLACK CIRCLE WITH WHITE DOT RIGHT
- U+2689 ◐ BLACK CIRCLE WITH TWO WHITE DOTS

As discussed on the Unicode mailing list (see thread starting at <http://www.unicode.org/mail-arch/unicode-ml/y2016-m03/0022.html>), these four characters are used in mathematical research into the game of Go (see Fig. 1), but do not seem to be attested in actual Go game notation.

However, there are four unencoded symbols that are widely used in English, Japanese and Chinese sources for the notation of Go games, and these are proposed for encoding in this document.

Fig. 1: Elwyn Berlekamp and Yonghoan Kim, "Where Is the "Thousand-Dollar Ko"?"; *Games of No Chance* vol. 29 (1996) p. 206

Values must also be computed for those regions that do not appear verbatim in the book. Such calculations are done as described in Chapter 2 of *Mathematical Go*. As an example of such a calculation, we consider the region Q. We have $Q = \{Q^L \mid Q^R\}$, where Q^R is the position after White has played there and Q^L is the position after black has played there. After appropriate changes of markings (which affects only the integer parts of the relative score), we may continue as follows:



Although 0 and $-$ are both formal Black followers of Q^R , the miny option is dominated and we have

$$Q^R = \{0 \mid *, -\}.$$

The two formal White followers are * and $-$, which are incomparable. However, White's move from Q^R to $-$ reverses through a hot position to 0, and we

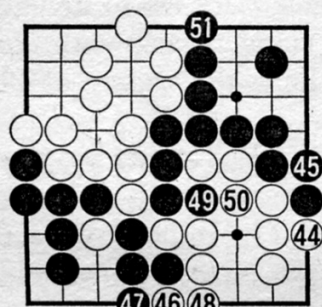
2. Circled Triangles and Squares

In commentary on Go game positions, it is customary to mark a particular stone or group of stones under discussion with a triangle. When an additional stone or group of stones needs to be marked for discussion in the same diagram in which a stone of the same colour marked with a triangle already occurs, then it is marked with a square. Fig. 2 below shows white and black stones marked with triangles and squares, both within the game diagram, and in text discussion of the game diagram.

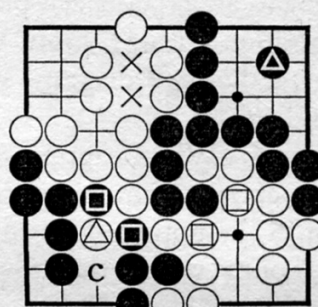
Fig. 2: Kaoru Iwamoto, *Go for Beginners* (Penguin Books, 1976) pp. 18-19

18 *Go for Beginners*

Dia. 8 At the end of this diagram all the boundaries are completed. Both players recognize that there is nothing more they can do to enlarge their own territories or to reduce their opponent's, so the game is over. Let's count and see who has won.



Dia. 8



Dia. 9

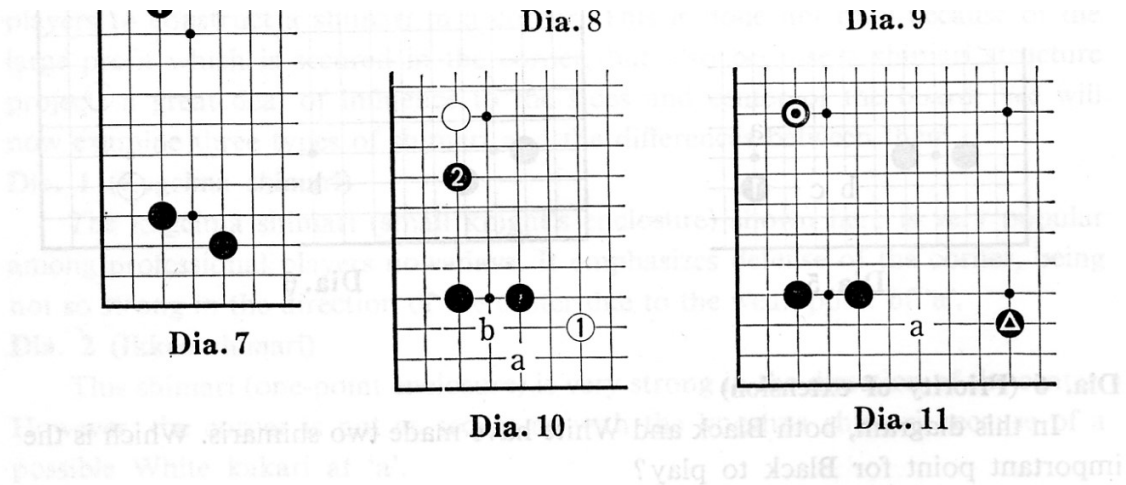
Demonstration Game 19

it from the board as his prisoner, leaving himself seven vacant points here. Again there is a diagonal break between the two stones marked \blacksquare , and again that is all right.

The rules of go also state that one point is to be subtracted from a player's territory for every stone lost, so we have to figure another point for White \triangle . Instead of subtracting one point from White's score for this stone, it is easier to add a point to Black's territory, so we shall follow that course and count eight points for Black in the lower left corner: six vacant points, one that becomes vacant when White \triangle is removed, and one more 'prisoner point'. That is, Black gets two points for \triangle .

Further examples of go stones marked with a triangle or square from books published in Japan and China are shown in Fig. 3 through Fig. 8 below. Usage of these symbols on internet websites is shown in Fig. 8.

Fig. 3: Haruyama and Nagahara, *Basic Techniques of Go* (Ishi press, 1973) p. 58



Dia. 10 (Weakness in the corner)

From the last two diagrams we saw that the ikken shimari is very strong towards the center. But because of the weakness at White 1, this shimari is not so strong with respect to the corner. After Black extends to 2, White will aim at 'a', to which Black will respond with 'b'. Depending upon the situation, White might choose to play at 'b' instead of 'a'.

Dia. 11 (Double-wing formation)

When Black has made an ikken shimari, the extensions of Black \blacktriangle and \bullet are ideal placements. This configuration is known as the double-wing formation.

Fig. 4: James Davies, *Tesuji* (Ishi Press, 1975) p. 89

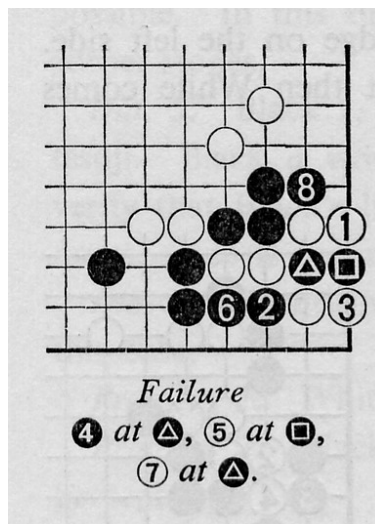


Fig. 5: James Davies, *Life and Death* (Ishi Press, 1975) p. 9

Dia. 6. Two-stage ko, (favorable to Black).

White 1 starts a two-stage ko. While the ko is being fought at 1 and 2, it is direct for Black, (all he has to do is to capture White \triangle), and indirect for White, (he has to play *a* and then *b* to win it). If White ignores a ko threat, however, the ko will move over and be fought at \triangle and *a*, and in that stage it will be direct for White and indirect for Black.

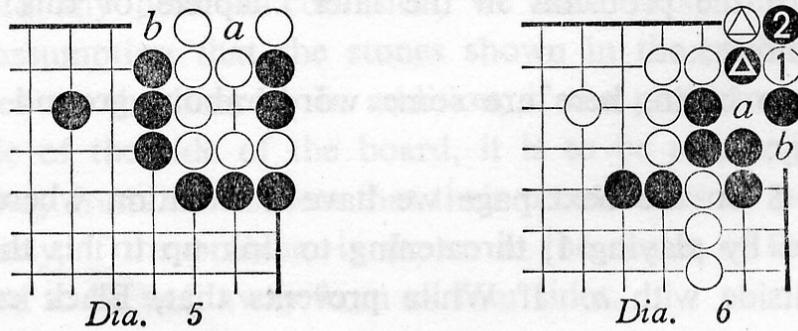
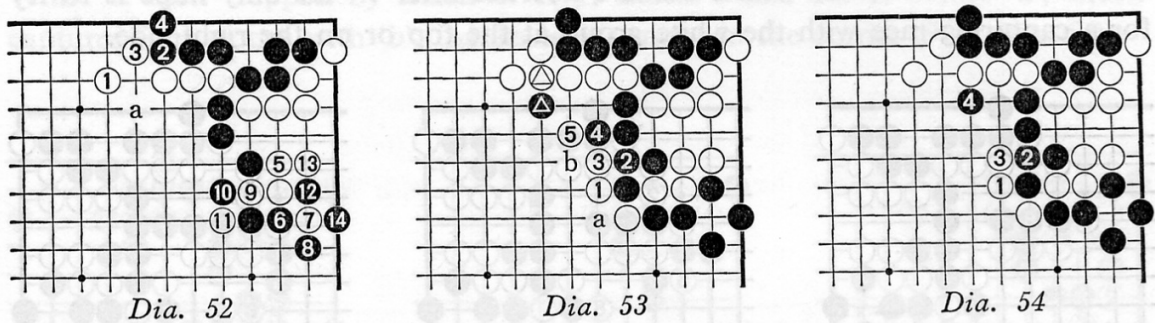


Fig. 6: Kaoru Iwamoto, *Dictionary of Basic Joseki* (Ishi Press, 1977) vol. 3 p. 31

Dia. 53 (the meaning of the peep). If the \triangle \triangle exchange has been made, there is no problem. White captures the black group with 1 to 5. Trying to escape by cutting at 'a' and 'b' is of course useless. This is why Black omits \triangle .

Dia. 54 (tricked). By omitting the peep, Black is able to escape by attaching at 4. There is now no way White can avoid collapse.



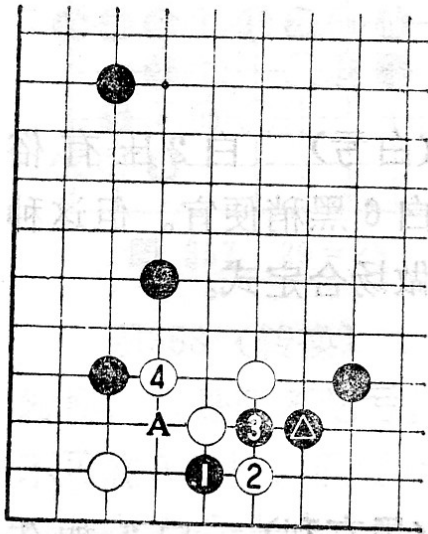


图 351

图351 (手筋?) 黑▲尖时, 白如应一手是可以确保活棋的, 但在实战中也常脱先。

黑1托, 是普通场合的手筋, 白如A退, 白形不好。但白2扳、4虎好, 抵挡住黑1的进攻。

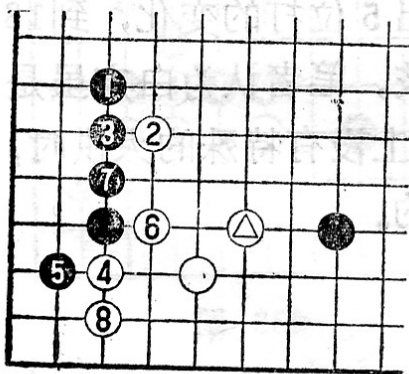
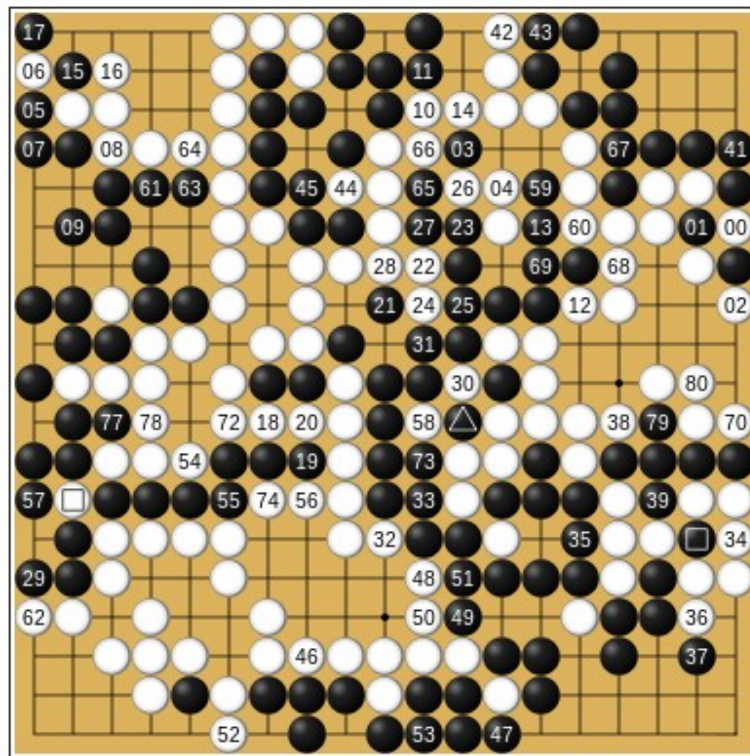





图352

图352 (白好形) 黑1拆二, 在本形不适用, 白2点后到6时因有▲关系黑不能反打, 到白8后白形很挺, 黑处低位, 和图347相比得失可以了然。



Fig. 8: AlphaGo vs Lee Sedol Game 5
 (https://en.wikipedia.org/wiki/AlphaGo_versus_Lee_Sedol)







Moves 200-280 (240 at ,
 275 at , 276 at )

Sometimes a stone is marked with a circle or a filled circle (see Fig. 3), but these cases can be represented using existing characters such as U+25CE ☉ BULLSEYE, U+25C9 ● FISHEYE, U+1F78A ◎ WHITE CIRCLE CONTAINING BLACK SMALL CIRCLE, and U+1F789 ● EXTREMELY HEAVY WHITE CIRCLE.

In order to represent the usage of triangles and squares to mark stones in Go notation, I propose encoding four symbol characters in the Geometric Shapes Extended block

Table 1: Proposed Characters for Marking Go Stones

Code Point	Glyph	Character Name
1F7D5		CIRCLED TRIANGLE
1F7D6		NEGATIVE CIRCLED TRIANGLE
1F7D7		CIRCLED SQUARE
1F7D8		NEGATIVE CIRCLED SQUARE

3. Unicode Properties

```
1F7D5;CIRCLED TRIANGLE;So;0;ON;;;;N;;;;;
1F7D6;NEGATIVE CIRCLED TRIANGLE;So;0;ON;;;;N;;;;;
1F7D7;CIRCLED SQUARE;So;0;ON;;;;N;;;;;
1F7D8;NEGATIVE CIRCLED SQUARE;So;0;ON;;;;N;;;;;
```


4. Proposal Summary Form

**SO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.
See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title: *Proposal to encode symbols for Go game notation*

2. Requester's name: *Andrew West*

3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*

4. Submission date: *2016-07-11*

5. Requester's reference (if applicable):

6. Choose one of the following:

This is a complete proposal: YES

(or) More information will be provided later:

B. Technical – General

1. Choose one of the following:

a. This proposal is for a new script (set of characters): NO
Proposed name of script:

b. The proposal is for addition of character(s) to an existing block: YES
Name of the existing block: *Geometric Shapes Extended*

2. Number of characters in proposal: *4*

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large collection)
 C-Major extinct D-Attested extinct E-Minor extinct
 F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided? YES

a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? YES

b. Are the character shapes attached in a legible form suitable for review? YES

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? *Andrew West*

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): *Andrew West*

6. References:

a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? YES

b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? YES

7. Special encoding issues:

Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? YES

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain		NO
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents:		YES
	Unicode mailing list	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:		NO
4. The context of use for the proposed characters (type of use; common or rare) Reference:		common
5. Are the proposed characters in current use by the user community? If YES, where? Reference:		YES
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:		NO
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		NO
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:		NO
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:		NO
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:		NO
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:		NO
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)		NO
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		NO