L2/04-029

ISO/IEC JTC 1/SC 2/WG 2

PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

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

(Please read Principles and Procedures Document for guidelines and details before filling this form.)
See http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html for latest Form.
See http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html for latest Principles and Procedures document.
See http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest roadmaps.

(Form number: N2352-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09)

A. Administrative

1. Title: Proposal to encode five additional CJK symbols and marks

2. Requester's name: Andrew C. West
3. Requester type: Individual contribution
4. Submission date: 19th January 2004
5. Requester's reference (if applicable):
6. (Choose one of the following:)
   This is a complete proposal: Complete Proposal
   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
   b. The proposal is for addition of character(s) to an existing block: Yes
   Name of the existing block: CJK SYMBOLS AND PUNCTUATION
   Note the CJK SYMBOLS AND PUNCTUATION block is full, and so it is suggested that a new SUPPLEMENTARY CJK SYMBOLS AND PUNCTUATION block be assigned at U+31CO..31XF.
2. Number of characters in proposal: 5
3. Proposed category (see section II, Character Categories): C
4. Proposed Level of Implementation (1, 2 or 3)
   (see clause 14, ISO/IEC 10646-1: 2000): 2
   Is a rationale provided for the choice? No
   If Yes, reference:
5. 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 ISO/IEC 10646-1: 2000? Yes
   b. Are the character shapes attached in a legible form suitable for review? Yes
6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?
   Andrew C. West
   If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:
   http://uk.geocities.com/BabelStone1357/Fonts/CJK.zip
7. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? No
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? Yes
8. 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)?

No_____

9. 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/Public/UNIDATA/UnicodeCharacterDatabase.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? No__
If YES explain _____
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.)? Yes_____
If YES, with whom? Richard Cook_____
If YES, available relevant documents: _____
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? No_____
Reference: _____
4. The context of use for the proposed characters (type of use; common or rare) Publishing and textual criticism_____
Reference: _____
5. Are the proposed characters in current use by the user community? Yes_____
If YES, where? Reference: Typeset reprints of pre-modern texts_____
6. After giving due considerations to the principles in *Principles and Procedures document* (a WG 2 standing document) must the proposed characters be entirely in the BMP? Yes_____
If YES, is a rationale provided? No_____
If YES, reference: _____
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? Yes_____
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? No_____
If YES, is a rationale for its inclusion provided? _____
If YES, reference: _____
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? No___
If YES, is a rationale for its inclusion provided? _____
If YES, reference: _____
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? Yes_____
If YES, is a rationale for its inclusion provided? Yes_____
If YES, reference: _____
11. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)? No_____
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: _____
12. Does the proposal contain characters with any special properties such as

```
  control function or similar semantics?                    No_____________
   If YES, describe in detail (include attachment if necessary) _____________
13. Does the proposal contain any Ideographic compatibility character(s)? No_
   If YES, is the equivalent corresponding unified ideographic character(s)
   identified?                                                  _____________
    If YES, reference: ____________________________________________________
```

ADDITIONAL INFORMATION

Outline of Additional Information

1. Proposed Characters
2. Ideographic White Square
3. Ideographic Black Square
4. Ideographic Iteration Marks

1. PROPOSED CHARACTERS

The five proposed characters are shown in Table 1.

Table 1 : Proposed Characters

Proposed Code Point	Representative Glyph	Proposed Character Name	Character Properties
31C0	□	IDEOGRAPHIC WHITE SQUARE	Lo;0;L;;;;;N;;;;;
31C1	■	IDEOGRAPHIC BLACK SQUARE	Lo;0;L;;;;;N;;;;;
31C2	〻	CHINESE IDEOGRAPHIC ITERATION MARK-1	Lm;0;L;;;;;N;;;;;
31C3	〆	CHINESE IDEOGRAPHIC ITERATION MARK-2	Lm;0;L;;;;;N;;;;;
31C4	〇	CHINESE IDEOGRAPHIC ITERATION MARK-3	Lm;0;L;;;;;N;;;;;

2. IDEOGRAPHIC WHITE SQUARE

When transcribing a Chinese text (printed book, paper manuscript, silk roll, bamboo slips, oracle bone, stone or bronze inscription, etc.) for publication, it is a widespread convention to represent any ideograph that cannot be recognized, either due to physical damage to the text or due to the poor quality of the printing or hand-writing, with a hollow square that takes up the same area as an ideograph. These hollow square glyphs are very common in modern printed texts, whether scholarly transcriptions of ancient texts (see Example 2-1) or typeset reprints of popular literature (see Example 2-2).

Less frequently, in scholarly transcriptions of ancient texts, a missing or unclear ideograph may be represented by means of a full-sized hollow square with a single diagonal line or two crossed diagonal lines inside. This may be the case in transcriptions of oracle bone texts, where a plain hollow square could be misread as the oracle bone form of the ideograph dīng 丁 (a square). Hollow squares with one or two diagonal lines may be considered to be glyph variants of the more usual empty hollow square.

The hollow square glyph may also be used in other contexts as a substitute for a Chinese ideograph, for example to replace the ideograph for an obscene word in a censored edition of novel, or to represent a syllable in dialect

speech that has no written form (see Example 2-3).

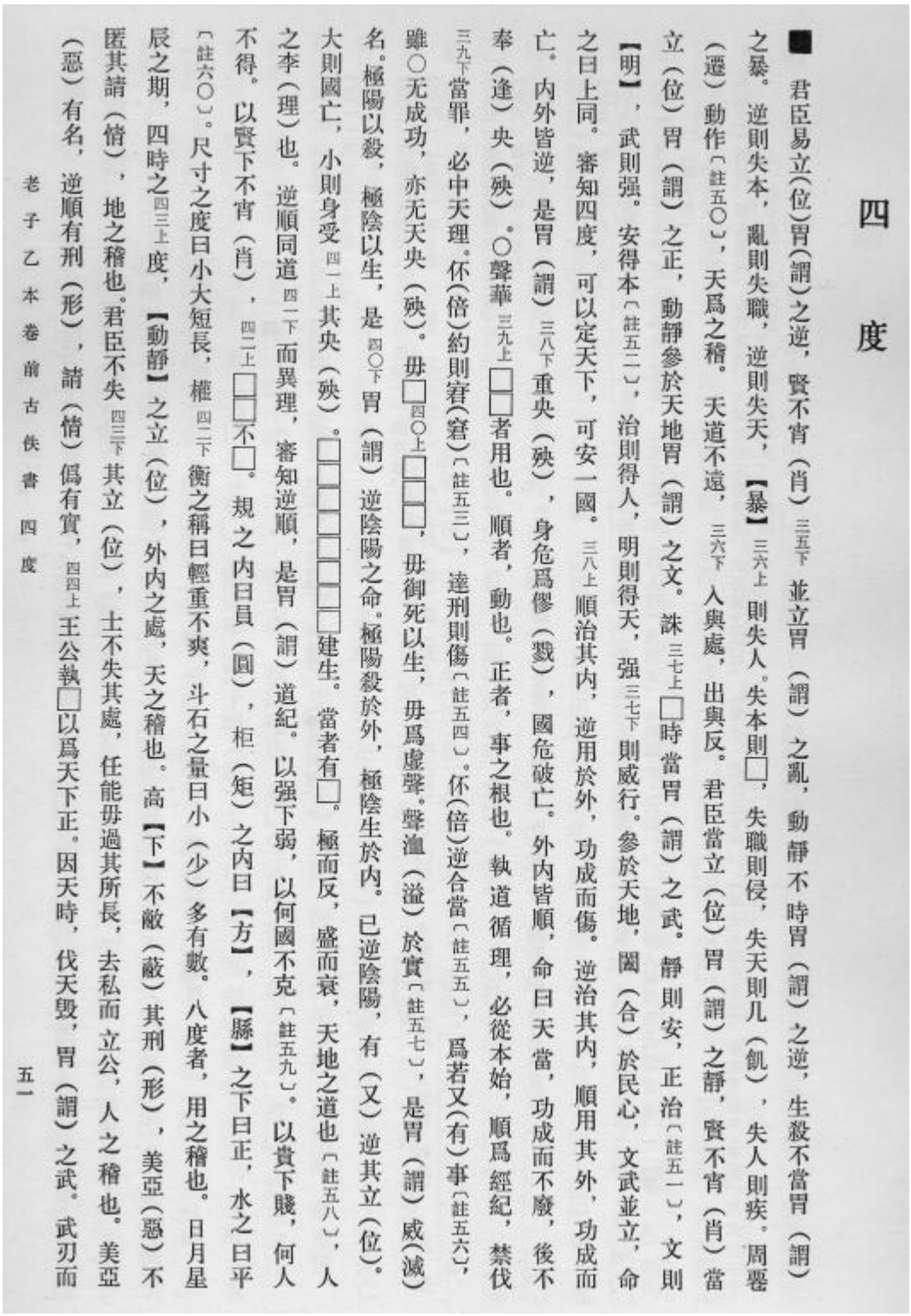
When a missing ideograph can be inferred from context or another text, the inferred ideograph is sometimes placed within a full-sized hollow square. This usage could probably be represented by means of U+20DE [COMBINING ENCLOSING SQUARE], and so no COMBINING IDEOGRAPHIC WHITE SQUARE character is proposed in this document.

There are already a number of white square characters encoded in Unicode (U+25A1 □ [WHITE SQUARE]; U+25FB □ [WHITE MEDIUM SQUARE]; U+25FD □ [WHITE MEDIUM SMALL SQUARE]). However the size of these characters varies from font to font, and they do not have the correct character properties, so I believe that it would be useful to have a dedicated Unicode character with the correct character properties to represent the missing ideograph glyph. This would greatly facilitate operations such as collating and searching, and would avoid the current situation where a missing ideograph may be represented by any one of the Unicode white square symbols, or even by the ideograph U+56D7 □ (a rare ideograph, mostly used as a radical = U+2F1E). Although I have myself used U+56D7 □ to represent missing ideographs in the past, as it is guaranteed to display at the correct size for ideographic text, I believe that the use of this character for this function is wholly inappropriate. Note in particular the glyph shape of U+56D7, which is not perfectly square in most Chinese fonts, with the two vertical strokes extending down below the lower horizontal stroke (cf. Examples 2-1 through 2-3 where the missing ideograph is represented by a completely square symbol).

The existing Unicode character U+3013 〓 [GETA MARK] is a Japanese syllable used to represent a missing kanji ideograph, but whilst the geta mark has very similar semantics to the proposed Ideographic White Square character, they are completely different characters, and neither could really be considered to be a glyph variant of the other. Moreover, if the Chinese hollow square character were to be unified with the geta mark, it would mean that in plain text hollow squares may be rendered as the geta mark, which is completely unknown to general Chinese readers. Unifying the two disimilar glyphs in a single Unicode character would merely cause confusion and annoyance in Chinese and Japanese readers alike.

Example 2-1

This is a page from a scholarly transcription of the silk manuscripts found in a Han dynasty tomb (dateable to 168 B.C.) in Changsha. Due to the condition of the silk, much of the text is illegible or has completely decayed away. Each missing ideograph is represented in the transcription by a full-sized hollow square.



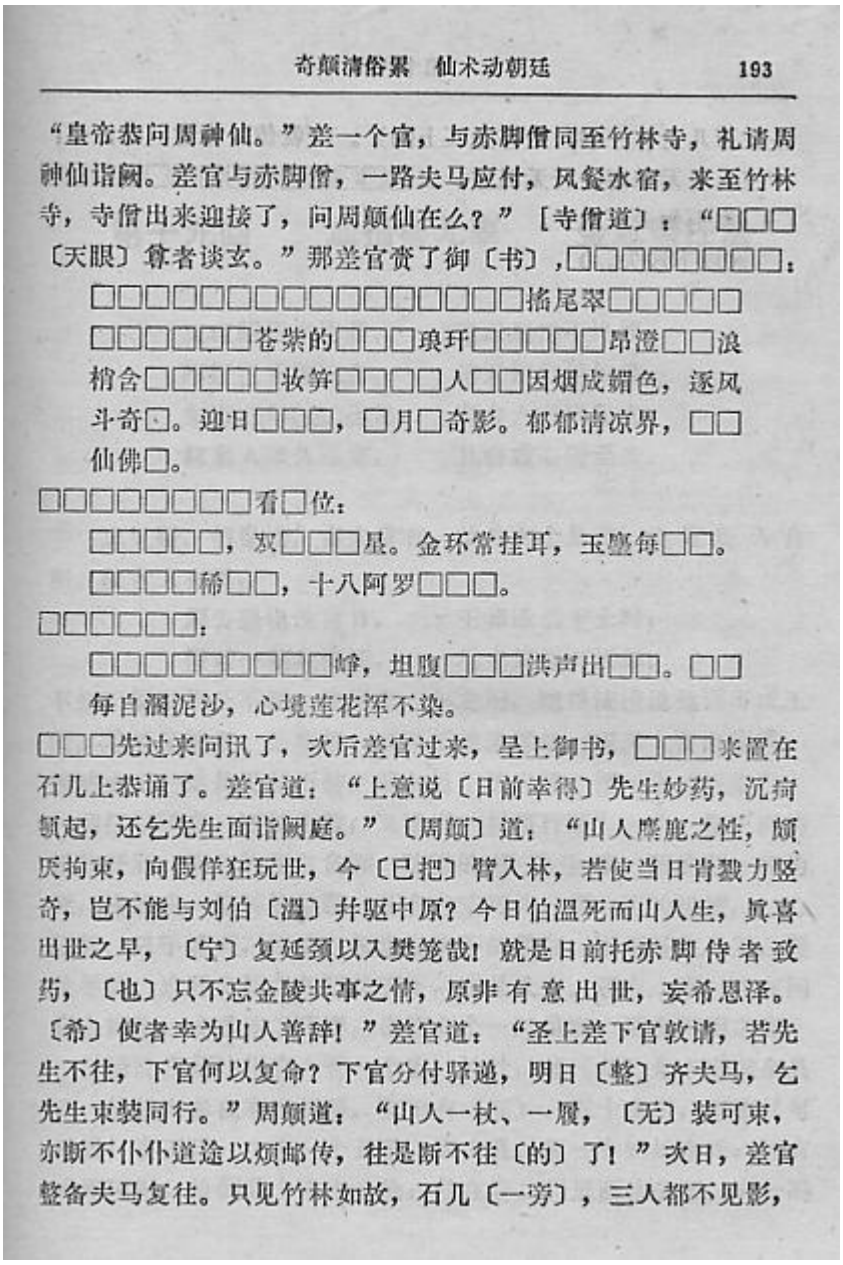
四度

■君臣易立(位)胃(謂)之逆，賢不宵(肖)三五下並立胃(謂)之亂，動靜不時胃(謂)之逆，生殺不當胃(謂)之暴。逆則失本，亂則失職，逆則失天，【暴】三六上則失人。失本則□，失職則侵，失天則几(飢)，失人則疾。周罌(還)動作〔註五〇〕，天爲之稽。天道不遠，三六下人與處，出與反。君臣當立(位)胃(謂)之靜，賢不宵(肖)當立(位)胃(謂)之正，動靜參於天地胃(謂)之文。誅三七上□時當胃(謂)之武。靜則安，正治〔註五一〕，文則【明】，武則强。安得本〔註五二〕，治則得人，明則得天，强三七下則威行。參於天地，閩(合)於民心，文武並立，命之曰上同。審知四度，可以定天下，可安一國。三八上順治其內，逆用於外，功成而傷。逆治其內，順用其外，功成而亡。內外皆逆，是胃(謂)三八下重央(殃)，身危爲僇(戮)，國危破亡。外內皆順，命曰天當，功成而不廢，後不奉(逢)央(殃)。○聲華三九上□□者用也。順者，動也。正者，事之根也。執道循理，必從本始，順爲經紀，禁伐三九下當罪，必中天理。伓(倍)約則窘(窘)〔註五三〕，達刑則傷〔註五四〕。伓(倍)逆合當〔註五五〕，爲若又(有)事〔註五六〕，雖○无成功，亦无天央(殃)。毋□四〇上□□□，毋御死以生，毋爲虛聲。聲溫(溢)於實〔註五七〕，是胃(謂)威(滅)名。極陽以殺，極陰以生，是四〇下胃(謂)逆陰陽之命。極陽殺於外，極陰生於內。已逆陰陽，有(又)逆其立(位)。大則國亡，小則身受四一上其央(殃)。□□□□□□□□□建生。當者有□。極而反，盛而衰，天地之道也〔註五八〕，人之李(理)也。逆順同道四一下而異理，審知逆順，是胃(謂)道紀。以强下弱，以何國不克〔註五九〕。以貴下賤，何人不得。以賢下不宵(肖)，四二上□□不□。規之內曰員(圓)，柜(矩)之內曰【方】，【縣】之下曰正，水之曰平〔註六〇〕。尺寸之度曰小大短長，權四二下衡之稱曰輕重不爽，斗石之量曰小(少)多有數。八度者，用之稽也。日月星辰之期，四時之四三上度，【動靜】之立(位)，外內之處，天之稽也。高【下】不敵(蔽)其刑(形)，美亞(惡)不匿其請(情)，地之稽也。君臣不失四三下其立(位)，士不失其處，任能毋過其所長，去私而立公，人之稽也。美亞(惡)有名，逆順有刑(形)，請(情)偽有實，四四上王公執□以爲天下正。因天時，伐天毀，胃(謂)之武。武刃而

Source : Mawangdui Hanmu Boshu 馬王堆漢墓帛書 (Beijing : Wenwu Chubanshe, 1980) vol.1 page 51.

Example 2-2

This is a page from a typeset reprint of a late Ming (1368-1644) collection of vernacular short stories that is only known from a single poor-quality woodblock edition. Full-sized hollow squares are substituted for illegible ideographs. Where the missing ideographs can be inferred from context, they are given within tortoise shell brackets [U+3014 and U+3015].



奇颠清俗累　仙术动朝廷　193

"皇帝恭问周神仙。"差一个官，与赤脚僧同至竹林寺，礼请周神仙诣阙。差官与赤脚僧，一路夫马应付，风餐水宿，来至竹林寺，寺僧出来迎接了，问周颠仙在么？"〔寺僧道〕："□□□〔天眼〕尊者谈玄。"那差官赍了御〔书〕，□□□□□□□□：

□□□□□□□□□□□□□□□□摇尾翠□□□□□

□□□□□□苍紫的□□□琅玕□□□□□昂澄□□浪

裙含□□□□□妆笋□□□□人□□因烟成媚色，逐风斗奇□。迎日□□□，□月□奇影。郁郁清凉界，□□仙佛□。

□□□□□□□□看□位：

□□□□□，双□□□星。金环常挂耳，玉麈每□□。

□□□□稀□□，十八阿罗□□□。

□□□□□□：

□□□□□□□□□□峥，坦腹□□□洪声出□□。□□每自溷泥沙，心境莲花浑不染。

□□□先过来问讯了，次后差官过来，呈上御书，□□□来置在石几上恭诵了。差官道："上意说〔日前幸得〕先生妙药，沉疴顿起，还乞先生面诣阙庭。"〔周颠〕道："山人麋鹿之性，颇厌拘束，向假佯狂玩世，今〔已把〕臂入林，若使当日肯戮力竖奇，岂不能与刘伯〔温〕并驱中原？今日伯温死而山人生，真喜出世之早，〔宁〕复延颈以入樊笼哉！就是日前托赤脚侍者致药，〔也〕只不忘金陵共事之情，原非有意出世，妄希恩泽。〔希〕使者幸为山人善辞！"差官道："圣上差下官敦请，若先生不往，下官何以复命？下官分付驿递，明日〔整〕齐夫马，乞先生束装同行。"周颠道："山人一杖、一履，〔无〕装可束，亦断不仆仆道途以烦邮传，往是断不往〔的〕了！"次日，差官整备夫马复往。只见竹林如故，石几〔一旁〕，三人都不见影，

Source : Sanke Pai'an Jingqi 三刻拍案惊奇 (Beijing : Beijing Daxue Chubanshe 北京大学出版社, 1987) page 193.

Example 2-3

This is part of a page from a modern treatise on Chinese dialects by Yuan Jiahua 袁家骅. In this work full-sized hollow squares are substituted for dialectal words that have no written form when transcribing examples of dialect speech into Chinese ideographs.

57. “去北京”一类的句子

普通话“我上北京去”(或“我到北京去”)一类的句子，闽南话都说成“我去北京”，这一点跟粤方言是一致的。例如：

厦门：汝去□□？[li˥˧ kʻi˨˩˥ toʔ˥ loʔ˥ (→troʔ˥)]（你上哪儿去？）

我去北京。[gua˥˧ kʻi˨˩˥ pak˥˩ kiŋ˥]（我上北京去）

潮州：汝去底块？[lɯ˥˧ kʻɯ˨˩˥ ti˨˩ ko˨˩]（你上哪儿去？）

我去汕头。[ua˥˧ kʻɯ˨˩˥ sũã˧˨ tʻau˥]（我上汕头去）

浙南：汝去□□？[dɯ˥˧ kʻɯ˨ toʔ˥˨ kɐʔ˥]（你上哪儿去？）

我去杭州。[gua˥˧ kʻɯ˨ xaŋ˧˩ tɕiu˥]（我上杭州去）

海南：汝去□？[duʔ˨ xu˧ ɗe˧]（你上哪儿去？）

我去街。[guaʔ˨ xu˧ koi˧]（我上街去）

当第一人称代词作主语时，还有一种说法，就是用“来去”代替“去”。例如：

厦门：我来去食饭。[gua˥˧ lai˧˨ kʻi˨˩˥ tsiaʔ˥˩ pŋ̍˧]（我吃饭去）

潮州：我来去睇戏。[ua˥˧ lai˥˩ kʻɯ˨˩˥ tõĩ˥˨ hi˨˩]（我看戏去）

“来去”含有“将要”的意思，表示一种意向，现在正开始行动。“我来去食饭”，意思是“我现在正要去吃饭”。

Source : Hanyu Fangyan Gaiyao 汉语方言概要 (Beijing : Wenzi Gaige Chubanshe 文字改革出版社, 1989) page 279.

---

3. IDEOGRAPHIC BLACK SQUARE

The Ideographic Black Square is proposed for use in representing a square black mark that may occur in traditional manuscript or woodblock editions of Chinese texts, known in Chinese as a mòdīng 墨釘 (also written 墨丁) "ink nail".

In ancient manuscript texts a square black mark may be used to indicate the start of a text. See Example 2-1 above for an example of the representation of such a mark in a scholarly transcription of a Han dynasty silk manuscript from Mawangdui.

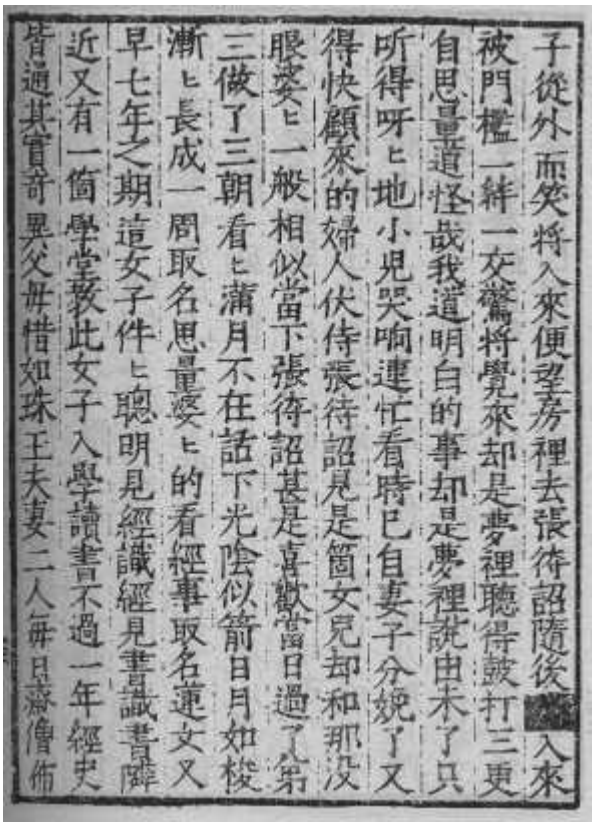
However, the **moding** mark is most frequently encountered in woodblock editions of Chinese texts in place of an expected ideograph (see Example 3-1). This comes about when the blocks for a woodblock edition of a book are recarved, as explained below :

A woodblock edition is produced by pasting face down onto a block of wood a sheet of paper on which is written the text to be printed. The black ink of the text shows up in reverse through the thin paper, and the engraver excavates the white areas of the pasted block, leaving the text standing out in relief. The remaining paper is scrubbed off, and the block smoothed down, ready for printing. When an edition is first printed, the text is written out carefully by hand onto sheets of paper by a scribe specifically for use by the engraver. However, when an edition is to be reprinted (because the old blocks have become worn out or because another publisher wants to print a pre-existing edition) it was common practice to simply unbind a copy of the existing printed text, and engrave the new blocks from the existing printed pages. If an ideograph in the engraver's copy of the existing edition was unclear or illegible, either due to physical damage to the copy or due to it having been poorly printed (quite likely when the old blocks had become worn down, thus necessitating their replacement), then the engraver would engrave around the square area nominally occupied by the unclear/illegible ideograph, thus leaving a square block of wood standing in relief where the unclear/illegible ideograph was. When printed, this produced a square black mark (i.e. the **moding**). When several unclear/illegible ideographs occured together, then a rectangular black column corresponding to several **moding** occur (see Example 3-2).

Whilst it may be argued that the **moding** mark could be represented by U+25A0 [BLACK SQUARE] (see Example 3-3), this block symbol does not have the correct character properties, and may be rendered at different sizes with different fonts, when the **moding** mark should always be rendered the same height and width as a CJK ideograph.

Example 3-1

This is a page from a Ming dynasty (1368-1644) edition of a collection of vernacular short stories entitled Qingpingshantang Huaben 清平山堂話本, in which there is a single **moding** mark where an ideograph is expected. Note that in the typeset edition of Qingpingshantang Huaben (Shanghai : Shanghai Guji Chubanshe 上海古籍出版社, 1987) this and other **moding** marks in the text are represented by a hollow square.


子從外而來將入來便望房裡去張待詔隨後■入來
被門檻一絆一交驚將覺來却是夢裡聽得鼓打三更
自思量道怪哉我道明白的事却是夢裡說由未了只
听得呀ㄈ地小兒哭响連忙看時已自妻子分娩了又
得快顧來的婦人伏侍張待詔見是箇女兒却和那没
服婆ㄈ一般相似當下張待詔甚是喜歡當日過了第
三做了三朝看ㄈ滿月不在話下光陰似箭日月如梭
漸ㄈ長成一周取名思量婆ㄈ的看經事取名蓮文又
早七年之期這女子件ㄈ聰明見經識經見書識書離
近又有一箇學堂教此女子入學讀書不過一年經史
皆通其實奇異父母惜如珠玉夫妻二人每日齋僧佈

Source : Qingpingshantang Huaben 清平山堂話本 (Beijing : Wenxue Guji Kanxingshe 文學古籍刊行社, 1987) page 308.

Example 3-2

This is a page from a Yuan dynasty (1206-1368) edition of a *zaju* play entitled Zhuye Zhou 竹葉舟 "The Bamboo Leaf Boat", in which there are two consecutive **moding** marks in place of two expected ideographs.

Example 3-3

On a Japanese web site the **moding** occuring in the text of the short story collection Liushijia Xiaoshuo 六十家小说 is encoded with U+25A0 ■ [BLACK SQUARE] :

> 且説張■進得房。［注20：原为墨丁，当是「狼」字］
>
> **Source :** http://www.rc.kyushu-u.ac.jp/~naka/r07a.html

With a Chinese font such as MingLiU or SimSun-18030 the black square may be drawn the same size as a Chinese ideograph, and looks correct, but with other fonts (e.g. Arial Unicode MS), the black square is not full-size, and it does not look like a **moding** at all.

Example 3-4

On a Chinese web site the **moding** occuring in the Chinese novel Sansui Pingyao Zhuan 三遂平妖传 is encoded as U+56D7 口 :

> 却说王则输了这一阵，正是刀添三个口[原文为墨钉]，人减七分威。
>
> **Source :** http://www.yifan.net/yihe/novels/classic/sansui/sansui18.html

U+56D7 口 is an uncommon Chinese ideograph (mostly used as a radical = U+2F1E) that is sometimes used in digital texts to replace the full-size hollow square that is used in Chinese typography to represent a missing or unclear ideograph.

---

4. IDEOGRAPHIC ITERATION MARKS

There are already two existing ideographic iteration marks within Unicode :

- U+3005 々 [IDEOGRAPHIC ITERATION MARK]
- U+303B 〻 [VERTICAL IDEOGRAPHIC ITERATION MARK]

These two characters cater for modern Japanese usage, but are not necessarily appropriate for Chinese texts (U+3005 is commonly used in modern hand-written Chinese, but to the best of my knowledge it does not occur in pre-twentieth century Chinese texts). In particular, several other symbols with the same semantics as these characters are frequently encountered in pre-modern Chinese printed and manuscript texts (good quality woodblock editions usually eschew the use of ideographic iteration marks, but they were common at the lower end of the commercial publishing market) :

1. Chinese Ideographic Iteration Mark-1 : a symbol which looks like the Chinese character *èr* 二 [U+4E8C] (see Examples 4-1-1 through 4-1-4).
2. Chinese Ideographic Iteration Mark-2 : a symbol which looks like the hiragana syllable *ku* く [U+304F] (see Examples 4-2-1 through 4-2-4).
3. Chinese Ideographic Iteration Mark-3 : a symbol which looks like the Chinese character *bǐ* 匕 [U+5315] (see Examples 4-3-1 through 4-3-2).

Whilst it is possible to substitute U+3005 or U+303B for these symbols in some instances, often it would not be appropriate to do so. For instance, during the process of textual transmission iteration marks sometimes become corrupted to graphically similar ideographs (e.g. Iteration Mark-1 may be corrupted to the ideograph *èr* 二 [U+4E8C]; wheras Iteration Mark-3 may be corrupted to the ideographs *qī* 七 [U+4E03] or *shàng* 上 [U+4E0A]). In documenting such textual corruption, it is necessary to be able to refer to the exact graphic form of the iteration mark. Furthermore, it is already established practise in scholarly transcriptions of early Chinese texts to transcribe Iteration Mark-1 as it appears (see Example 4-1-1), and no such character currently exists in Unicode.

Moreover the existence of the two already encoded ideographic iteration marks (U+3005 and U+303B) makes it more problematic to unify the three proposed Chinese ideographic iteration marks with them, as it is not evident which of U+3005 or U+303B each of the Chinese iteration marks could be considered a glyph variant of, with the result that in Text A Chinese Ideographic Iteration Mark-3 may be encoded using U+3005, but in Text B the same iteration mark may be encoded using U+303B, whereas Chinese Font A may map a glyph corresponding to Chinese Ideographic Iteration Mark-3 to U+3005, Chinese Font B may map the same glyph to U+303B, and Japanese Font C may map U+3005 and U+303B to the glyphs given in the Unicode code charts. Given that U+3005 and U+303B (which despite their different names may both be used in vertically laid out text) have been separately encoded, the simplest solution would be to follow this precedence and encode the remaining three Chinese ideographic iteration marks separately.

4.1 Ideographic Iteration Mark-1

This is the iteration mark with the longest history, occuring in manuscript texts dating from over two thousand years ago. Cursively, it may be written as a z-like squiggle similar to the VERTICAL IDEOGRAPHIC ITERATION MARK [U+303B]. However, there is an important difference between the two. From a Chinese perspective the squiggle is merely a cursive variant of the standard two-stroke form, whereas in the Japanese context the standard form of the iteration mark is a squiggle. Thus from a Chinese perspective U+303B would not be an appropriate way of representing the ideographic iteration mark in texts where the iteration mark consists of two distinct strokes (cf. the transcription of the Laozi manuscript given in Example 4-1-1).

Example 4-1-1

This example shows a facsimile of a small portion of a silk manuscript of the Book of Laozi [*Laozi Yiben* 老子乙本] that was copied circa 179-169 B.C., together with a scholarly transcription of part of the text. In this example an

iteration mark is used four times in the top section of the middle line.

Note that when a two-character phrase is to be repeated, the iteration mark is placed after each character of the phrase to be repeated (thus "靜是胃復二命二常也" stands for "靜，是胃復命。復命，常也。") This construct is common in very early texts, but is not found in woodblock printed texts.

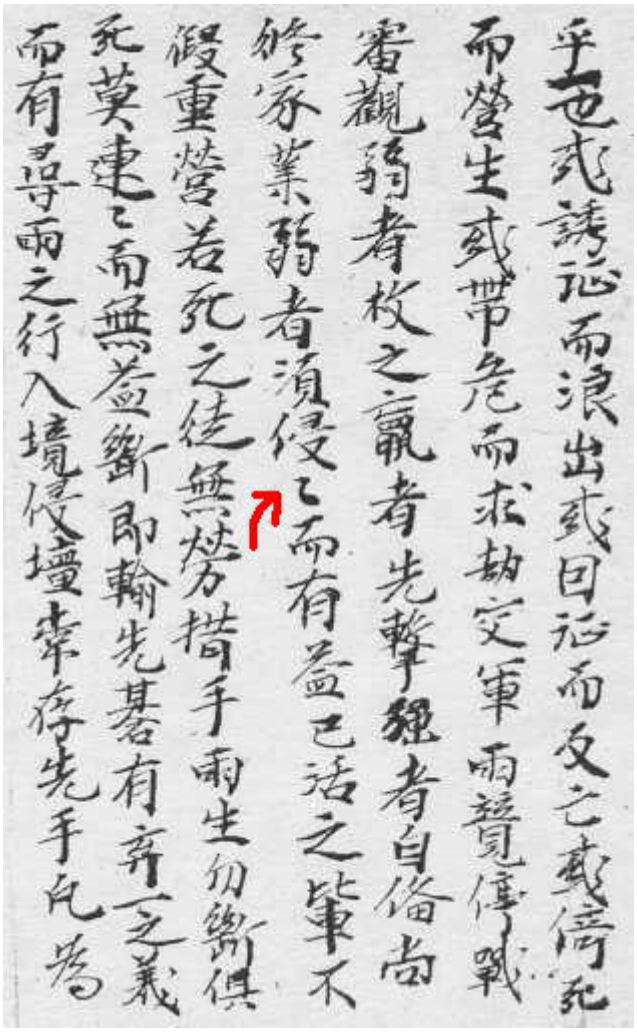
元復也天物云＝各復歸於元根曰靜＝是胃復＝命＝常也知常明也不知

Source : Mawangdui Hanmu Boshu 馬王堆漢墓帛書 (Beijing: Wenwu Chubanshe, 1980) vol.1.

Example 4-1-2

In cursive script, this iteration mark is often written as a z-like squiggle, as can be seen in this example from a manuscript copy of a manual on the game of Go ("弱者須侵，侵而有益" in the middle line). However, this z-like squiggle does not occur in printed Chinese texts, and thus when transcribing a text such as this the cursive

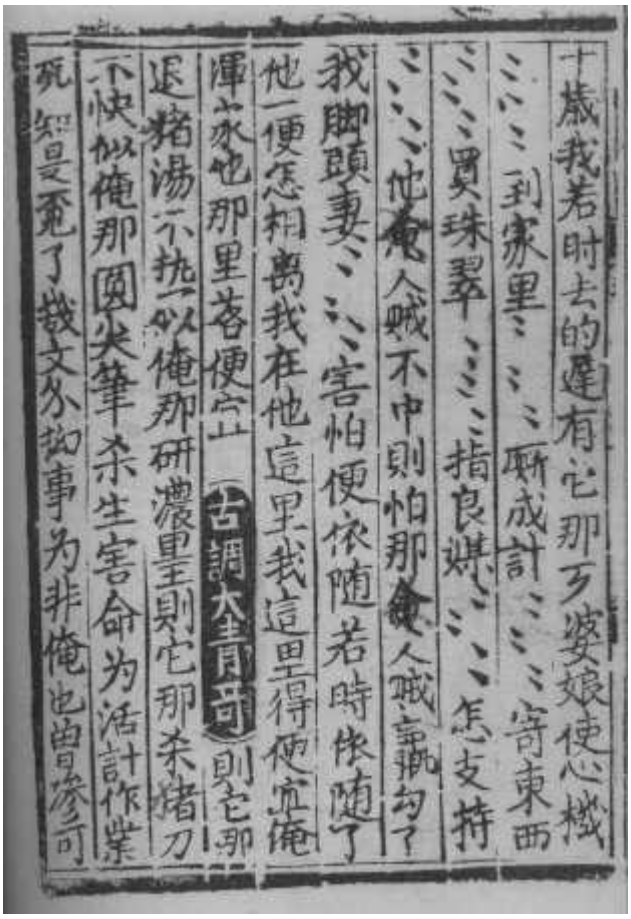
squiggle would be normalised as two unconnected horizontal lines (in the same way that cursive forms of Chinese ideographs would be normalised to their standard, dictionary form).



Source : Qi Jing 碁經. Tang dynasty (618-907) manuscript [S.5574].

Example 4-1-3

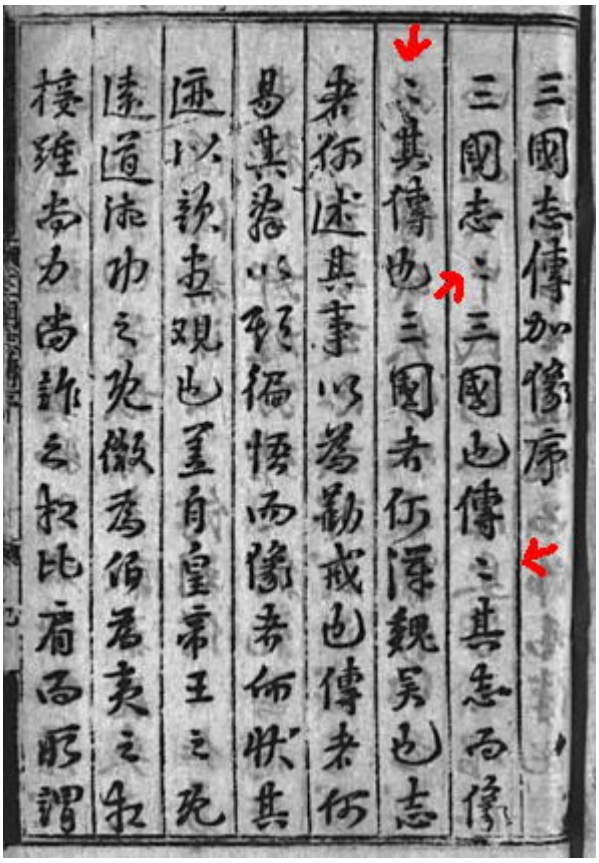
In this example from a Yuan dynasty (1206-1368) woodblock edition of a play, three consecutive iteration marks are used to indicate repetition of the preceding three-character phrase. Not how the iteration marks are alternately offset to the right and left.



Source : Xinbian Yue Kongmu Jie Tieguai Li Huanhun 新編岳孔目借鐵枴李還魂. In Yuankan Zaju Sanshizhong 元刊雜劇三十種 (Beijing: Shangwu Yinshuguan, 1958).

Example 4-1-4

In later Chinese texts, this iteration mark is more common in handwritten texts, as is seen in this woodblock print of a 1548 preface to the famous historical novel, Sanguo Yanyi [Romance of the Three Kingdoms] (the 4th and 9th characters of the 2nd line from the right, and the 1st character of the 3rd line : 三國志，志三國也。傳，傳其志，而像，像其傳也。).



Source : Preface to Xinkan Tongsu Yanyi Sanguozhi Shizhuan 新刊通俗演義三國志史傳. 1548 woodblock edition.

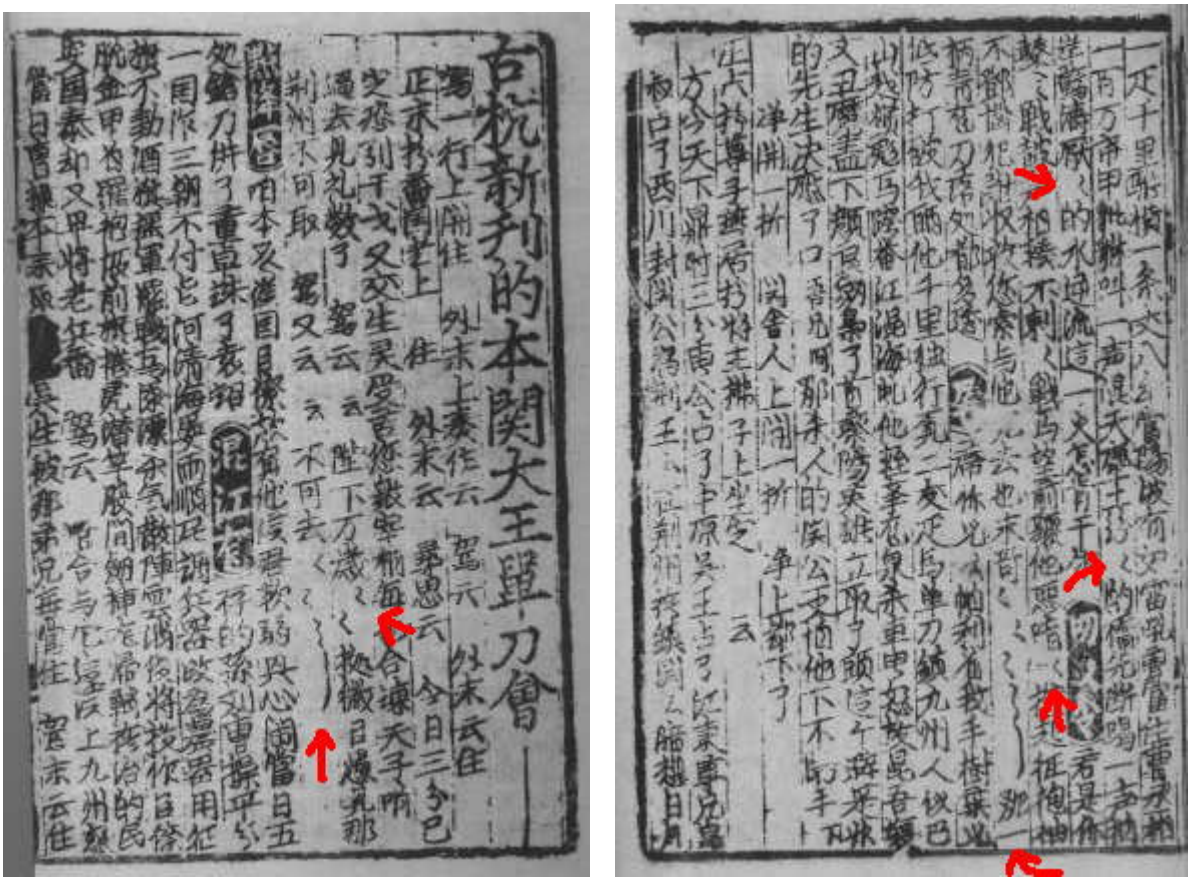
4.2 Ideographic Iteration Mark-2

This iteration mark is usually seen in texts published by downmarket commercial publishing houses in Jianyang (Fujian province) during the Yuan and Ming dynasties (as are all but the first example below). This symbol may sometimes be written with a slight gap between the two strokes, or sometimes, cursively, almost like the character *rén* 人 [U+4EBA].

This symbol is very similar in shape to the VERTICAL KANA REPEAT MARK ゞ[U+3031], but as U+3031 is used to iterate kana, and the Ideographic Iteration Mark-2 is used to iterate kanji, I believe that the latter mark needs to be encoded separately.

Example 4-2-1

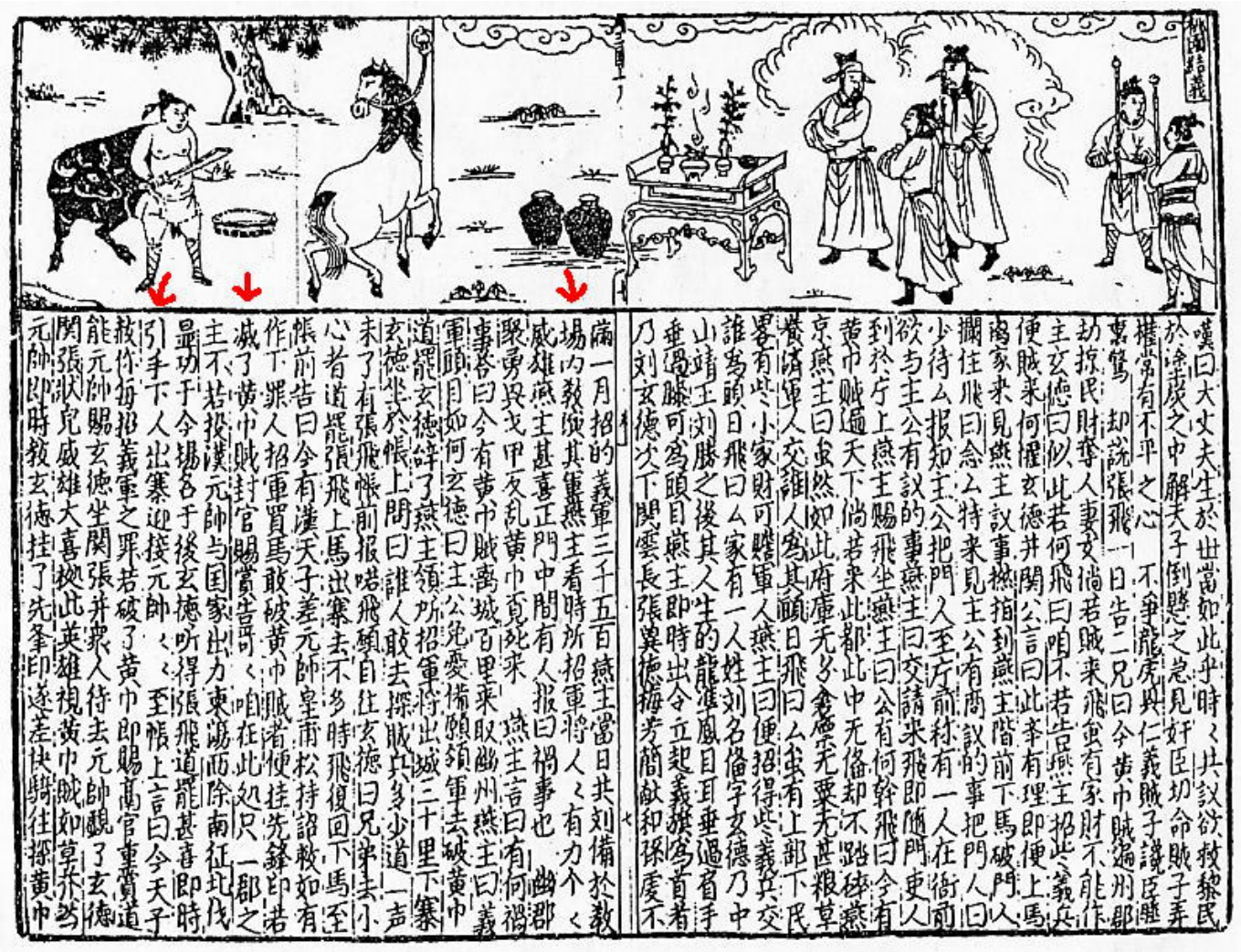
These are two pages from a Yuan dynasty (1206-1368) woodblock edition of a play. The page on the left shows two examples of double iteration marks, whilst the page on the right shows several examples of single iteration marks, and one example of triple iteration marks (after "也末哥"). Note that when the iteration marks are doubled or trebled, they are offset alternately to the right and left of centre.



Source : *Guhang Xinkan Diben Guan Dwang Dandao Hui* 古杭新刊的本關大王單刀會.In Yuankan Zaju Sanshizhong 元刊雜劇三十種 (Beijing: Shangwu Yinshuguan, 1958).

Example 4-2-2

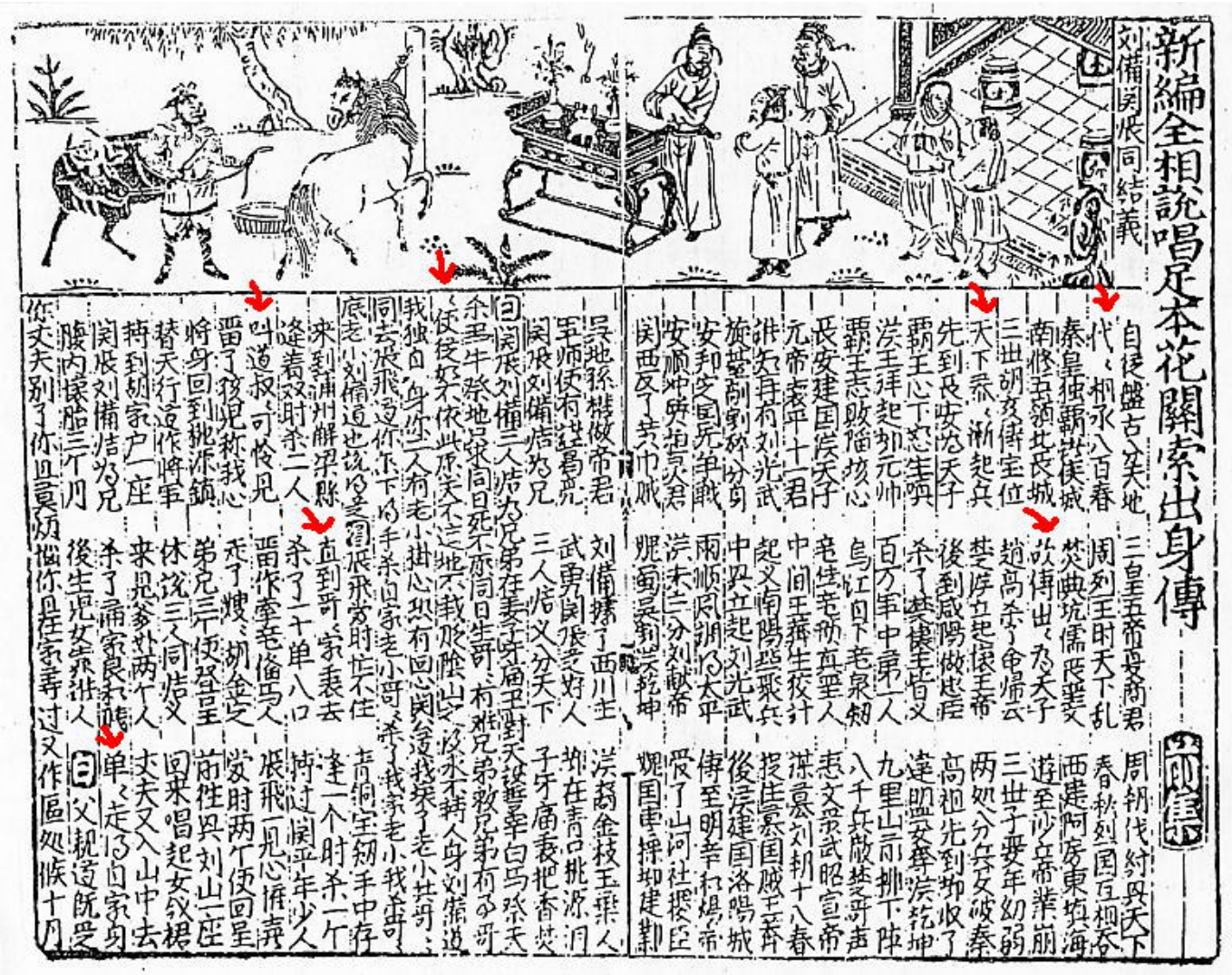
On the 2nd and 13th lines of the left-side page of this 1321-1323 woodblock edition of a popular historical narrative there are several examples of the iteration mark used to repeat the previous character, whilst on line 16th line is an example of double iteration marks to indicate the repetition of the two previous characters (元帥), here, atypically, not offset to the right and left.



Source : Sanguozhi Pinghua 三國志平話. 1321-1323 woodblock edition.

Example 4-2-3

In the opening verse of this 1478 woodblock edition of a verse narrative the iteration mark can be seen several times (e.g. "代く" in line 4, "世く" in line 11, and "莽く" in line 16).



Source : Xinbian Quanxiang Shuochang Zuben Hua Guansuo Chushen Zhuan 新編全相説唱足本花關索出身傳. 1478 woodblock edition.

Example 4-2-4

An example of the iteration mark can be seen in the first line from the right of this 1592 woodblock edition of the historical novel Sanguo Yanyi [Romance of the Three Kingdoms] ("陰雲漠く" for "陰雲漠漠").

Source : Yinshi Buyi Anjian Yanyi Quanxiang Piping Sanguozhi 音釋補遺按鑒演義全像批評三國志. 1592 woodblock edition.

4.3 Ideographic Iteration Mark-3

This is the most commonly encountered iteration mark in pre-modern Chinese woodblock texts, especially those dating to the Qing dynasty (1644-1911). This iteration mark is not obviously related to either of the other two proposed iteration marks.

Example 4-3-1

This page from a 1548 woodblock edition of the historical novel Sanguo Yanyi [Romance of the Three Kingdoms] shows two usages of the iteration mark.

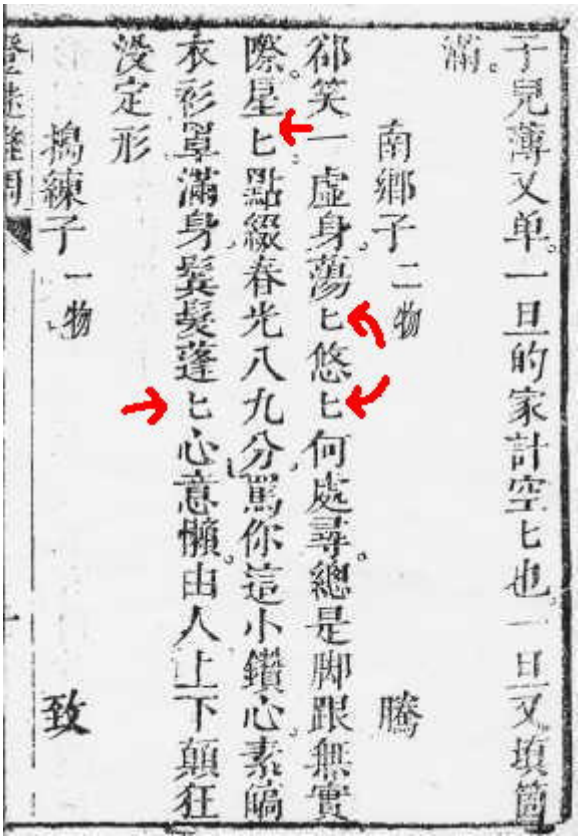
Lines 4-5 (from the right) show three examples of the ordinary use of the iteration mark to indicate repetition of the immediately preceding character ("爲術不足而化爲德ㄈ之不足而化爲仁ㄈ之不足而化爲儉ㄈ之不足而化爲仁義" for "爲術不足而化爲德，德之不足而化爲仁，仁之不足而化爲儉，儉之不足而化爲仁義").

Lines 6-10 show nine examples of duplicated iteration marks, which are used to indicate repetition of the previous two characters (i.e. "仁義", "三皇", "五帝", "三王", "四夷", "七雄", "楚漢", "新室", and "黃巾"). Where two Iteration Marks are used together, the two symbols are offset from each other longitudinally, the bottom symbol lying to the left of the top symbol.

Source : *Xinkan Tongsu Yanyi Sanguozhi Shizhuan* 新刊通俗演義三國志史傳. 1548 woodblock edition.

Example 4-3-2

In this example from an early 19th century woodblock edition of a collection of Lantern Riddles the iteration mark is used four times in the one riddle. Note how in this text, as is common with Qing dynasty texts, the iteration mark is a full-size character, unlike the half-sized forms in the previous example.



Source : *Dengmi Yadiao* 燈謎雅調. Woodblock edition.