

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N 1165

Date: 2005-11-29

Doc. Type:	Member body contribution
Title:	Summary of Old Hanzi Forum Held in Taipei 2005.10.26-28
Source:	TCA
Status:	Input to IRG
Action:	To be discuss
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	2
Medium:	Electronic

On October 26-28, TCA hosted an old Hanzi forum. Experts from Taiwan and mainland China made some conclusions in the meeting. All experts agreed that a common consensus should be obtained between Taiwan and Mainland China before processing the realignment work separately.

Based on the conclusion of the form, TCA proposes the arrangement of Old Hanzi attributes plan as follow:

1. Start with building up (construction) an old Hanzi inscription database, including:
 - (1) Original image (figure) database,
 - (2) The database of glyph determination (隸定), and
 - (3) Attribute database of Old Hanzi inscription.
2. This Old Hanzi platform includes
 - (1) The image (figure) database: Store original images of the Old Hanzi and the explanation archives . Users can refer to the original image when needed. (refer to Fig. 1 & Fig. 2)
 - (2) The database of glyph determination: Store glyph determination and Corresponding Modern Character. (refer to Fig. 3)
 - (3) Attributes Database: Stores the attribute data of collected Old Hanzi inscriptions. (refer to Fig. 4)

Fig. 1 the image database frame

甲骨文

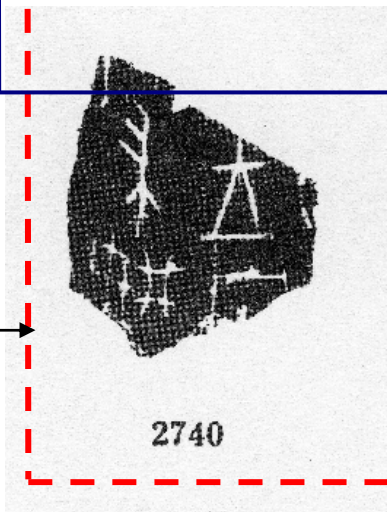
檔案(F) 其他(O)

指定釋文檔

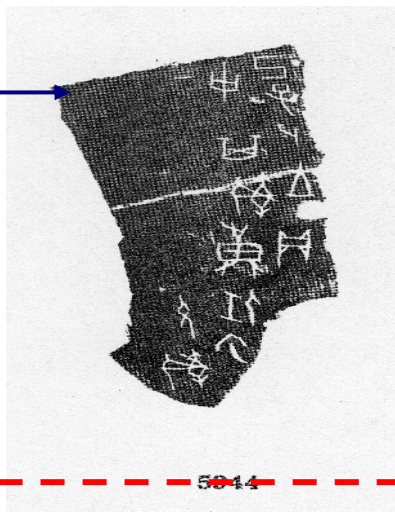
原稿管理 字體建構 甲骨檢索 古字檢索

編號	類別	出處	器名	時代	地域	材料
a-0-0001-00001	甲骨文	甲骨文合集		商	河南安陽	甲骨
c-0-0339-02169	甲骨文	小屯南地		商	河南安陽	甲骨
a-0-3713-30388	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0574-03337	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0091-00388	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-1005-06612	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2625-20305	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0510-02740	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0888-06075-1	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-3164-25077	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-3007-23354	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-4713-37952	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-4549-36534	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2852-22124	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2833-22072	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2116-14841	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2120-14887	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-4545-36514	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2380-17528	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0786-05447-0	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-3585-29243	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2122-14912	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-2770-21654	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0862-05944	甲骨文	甲骨文合集		商	河南安陽	甲骨
a-0-0888-06075-0	甲骨文	甲骨文合集		商	河南安陽	甲骨

瀏覽原稿 新增原稿 修改原稿 刪除原稿



甲骨文合集圖檔 2740



甲骨文合集圖檔 5944



甲骨文合集圖檔 23354

Fig. 2 the frame of using explanation archives search

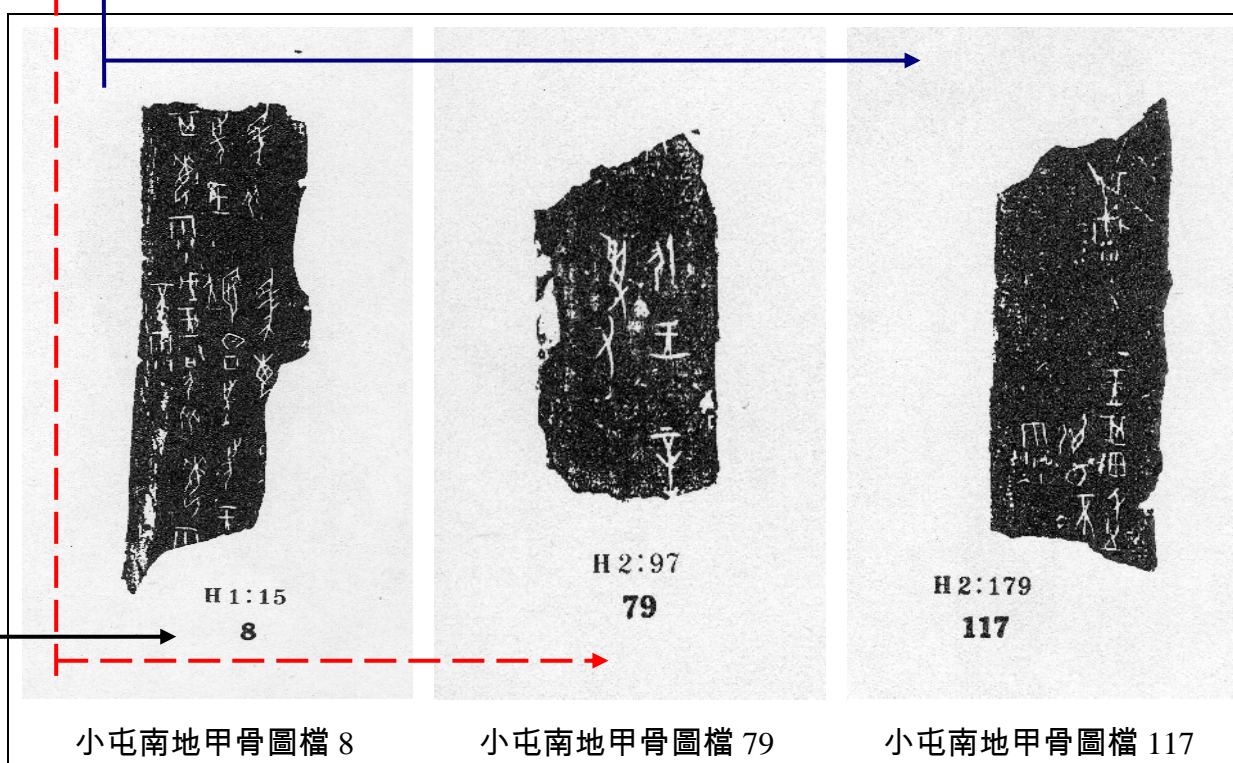
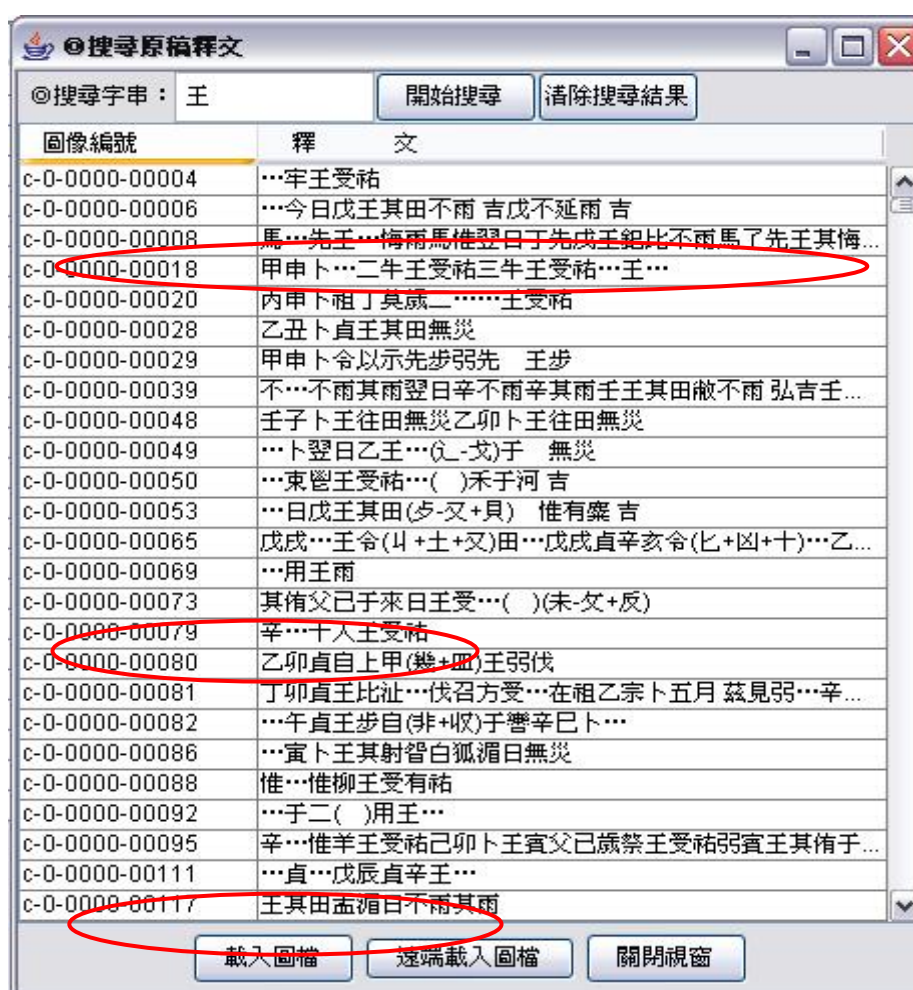


Fig. 3 the database of glyph determination

甲骨文

檔案(F) 其他(O)

指定釋文檔

原稿管理 字體建構 甲骨檢索 古字檢索

漢字屬性資料

◎	型別	今字	流水號	類別	出處	原形	隸定字	代表字
<input type="checkbox"/>	異構字	一	00001-A-05-...	甲骨文	甲骨文合集		一	原形
<input type="checkbox"/>	異構字	一	00002-A-05-...	甲骨文	小屯南地		一	隸定字
<input type="checkbox"/>	異構字	上	00003-A-05-...	甲骨文	甲骨文合集		上	上
<input type="checkbox"/>	異構字	上	00004-A-05-...	甲骨文	甲骨文合集		上	上
<input type="checkbox"/>	異構字	三	00005-A-05-...	甲骨文	甲骨文合集		三	三
<input type="checkbox"/>	異構字	三	00006-A-05-...	甲骨文	甲骨文合集		三	三
<input type="checkbox"/>	異構字	王	00007-A-05-...	甲骨文	甲骨文合集		王	王
<input type="checkbox"/>	異構字	王	00008-A-05-...	甲骨文	甲骨文合集		王	王
<input type="checkbox"/>	異構字	王	00009-A-05-...	甲骨文	甲骨文合集		王	王
<input type="checkbox"/>	異構字	王	000010-A-0-...	甲骨文	甲骨文合集		王	王
<input type="checkbox"/>	異寫字	王	000010-A-0-...	甲骨文	甲骨文合集		王	王

瀏覽屬性 新增異構字 新增異寫字 修改 刪除 設為異寫 設為異構

Fig. 4 attributes database

甲骨文

檔案(F) 其他(O)

指定釋文檔

原稿管理 字體建構 甲骨檢索 古字檢索

漢字屬性資料

◎	型別	今字	流水號	類別	出處	原形	隸定字	代表字
<input type="checkbox"/>	異構字	一	00001-A-05-...	甲骨文	甲骨文合集		一	
<input type="checkbox"/>	異構字	一	00002-A-05-...	甲骨文	小屯南地		一	
<input type="checkbox"/>	異構字	上	00003-A-05-...	甲骨文	甲骨文合集		上	
<input type="checkbox"/>	異構字	上	00004-A-05-...	甲骨文	甲骨文合集		上	
<input type="checkbox"/>	異構字	三	00005-A-05-...	甲骨文	甲骨文合集		三	
<input type="checkbox"/>	異構字	三	00006-A-05-...	甲骨文	甲骨文合集		三	
<input type="checkbox"/>	異構字	王	00007-A-05-...	甲骨文	甲骨文合集		王	
<input type="checkbox"/>	異構字	王	00008-A-05-...	甲骨文	甲骨文合集		王	
<input type="checkbox"/>	異構字	王	00009-A-05-...	甲骨文	甲骨文合集		王	
<input type="checkbox"/>	異構字	王	000010-A-0...	甲骨文	甲骨文合集		王	
<input type="checkbox"/>	異寫字	王	000010-A-0...	甲骨文	甲骨文合集			

瀏覽屬性 新增異構字 新增異寫字 修改 刪除 設為異寫 設為異構

3. The arrangement work procedure are:

- (1) Build up image database: scan all oracle inscription images, and input all the explanation archives (translations) as text files, thus forming an image database.
- (2) Choose glyphs: classify and choose words based on the constructed image database.
- (3) Glyph Determination: arrange attributes of the chosen glyphs and proceed glyph determination, data will be stored in database. The procedure of glyph determination are as follows:
 - a) Proceed glyph determination to the inscriptions of the Oracle Bones refer to the comments of experts
 - b) Compare and analysis the structure of glyphs with the same glyph determination result, judge between construct-variance and writing-variance according to the

conclusion made in the Old Hanzi Encoding and Modern Application Conference 2005.

- c) choose the most representing (complete construct, clear strokes) construct-variance and writing variance for the glyph, and store it into the database.
- (4) Counting representatives: there are many words for oracle bone inscription, every word has its own construct-variance and writing-variance, the representative is the word that represents the whole group of glyphs. Hence after the attributes arrangement, counting the number of representative is necessary.
- (5) Edit a wordlist: edit an oracle bone inscription wordlist, for WG2 as reference for encoding