# Proposal to Encode 9 Abbreviated Ideographs(Kanjis) in Japanese

Gen Kojitani(genkjtn@gmail.com)                                                  2023 June 22

I am requesting the addition of 9 Abbreviated Ideographs in Japanese. This document is revision of **L2/23-139**

## Background

This proposal requests the addition of 9 Ideograph(kanji) abbreviations used in Japan.

Among the kanji abbreviations in Japan that are not registered in Unicode, there are many abbreviations that have long been used in many handwritten documents and signs, although they are not officially recognized by the government. Although these kanji abbreviations are frequently seen by Japanese people in their daily lives, they have not yet been encoded in Unicode. Therefore, I have decided to submit a proposal requesting the addition of these kanji abbreviations.

Kanji abbreviation appears in many handwritten documents and media, mainly in creative works such as manga and anime, and public facility or building signboards. There are also cases where these abbreviations are used to prevent complicated characters from being indistinguishable due to being smashed and when the characters are illuminated, as in the Fig.1.

There are several patterns of kanji abbreviations in Japan. As shown in the example of the abbreviation for "摩(U+6469)" in Fig.1, there are patterns in which katakana is used to reduce the number of strokes in complex parts of the kanji, while others abbreviate by omitting some of the strokes or by converting cursive characters into printed form. In this article, I propose some of the most widely used of these abbreviations.
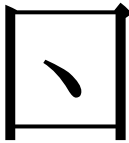


*Fig. 1:* Illuminated public event signage from *"Tama Center Illumination(多摩センターイルミネーション)"*
this picture is posted on Twitter by @himayasikatanow on 30 November 2021. URL:https://twitter.com/himayasikatanow/status/1465603206663213059

1

## Proposed Characters

First, a list of characters proposed for addition to Unicode is shown below. The list includes the letter form and information of each character.

Table 1. list of proposed characters

| No. | letter form | characters information |
|---|---|---|
| 1 | 庁 | • □广マ<br>• The abbreviation of U+9B54 魔 or U+6469 摩<br>• U+30DE KATAKANA LETTER MA "マ" is used to abbreviate the complex part. |
| 2 | 劷 | • □虍刂<br>• The abbreviation of U+5287 劇<br>• U+8C55 豕 is omitted. |
| 3 | 機 | • The abbreviation of U+6A5F 機<br>• The complex part is omitted as ⌄⌄-shape<br>• The middle stroke(dot) of ⌄⌄ is may be connected to under part(㭭) like Fig.14.<br>• A dot in the top right-hand corner may be omitted(㭭) like Fig.17. |
| 4 | 機 | • The abbreviation of U+6A5F 機<br>• The complex part is omitted as four dots.<br>• A dot in the top right-hand corner may be omitted(㭭) like Fig.18 and 21. |
| 5 | 迶 | • The abbreviation of U+9078 選<br>• The complex part is omitted as ⌄⌄-shape. |
| 6 | 屌 | • □尸ソ<br>• The abbreviation of U+5C64 層<br>• U+30BD KATAKANA LETTER SO "ソ" is used to abbreviate the complex part. |
| 7 | 庲 | • □广□タカ<br>• The abbreviation of U+9DF9 鷹<br>• U+30BF KATAKANA LETTER TA "タ" and U+30AB KATAKANA LETTER KA "カ" is used to abbreviate the complex part. |

2

| No. | letter form | characters information |
|---|---|---|
| 8 | 萉 | <ul><li>⬚⁺⁺⬚月卜</li><li>The abbreviation of U+85E4 藤</li><li>U+30C8 KATAKANA LETTER TO "卜" is used to abbreviate the complex part.</li></ul> |
| 9 | 囗 | <ul><li>⬚口丶</li><li>The abbreviation of U+56FD 国 or sometimes U+5712 園 like Fig.39-40</li><li>not to be confused with U+20B9A ⬚ (old letter of U+53E3 口), because the radical is different and the reason in the following item</li><li>The 丶 within the 口 can be written in a position straddling the right side of the 口(⬚), as in Fig.33 and 38, or in the lower left corner inside the 口(⬚), as in Fig.36-37.</li></ul> |

3

## Examples

The following images are examples for usage of abbreviated kanjis. Some of the original images in this document have been cropped for ease of viewing.

庁(囗广マ) for "魔"

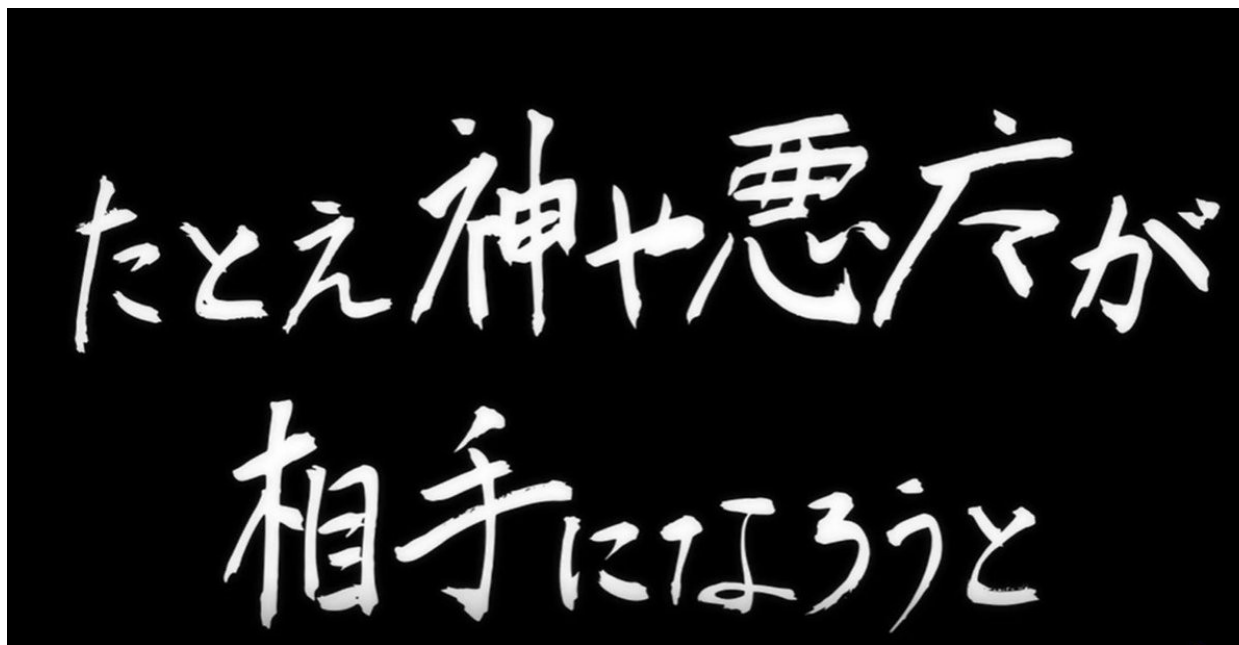This abbreviation appears mainly in handwritten characters in comic, anime, and other creative works.



*Fig. 2: "たとえ神や悪魔が相手になろうと"*

From the 11th episode of "Zombieland Saga Revenge" (broadcast June 17, 2021), a Japanese animated television series co-produced by MAPPA, Avex Pictures and Cygames.



*Fig. 3: "魔女とか"*

From "同居人に片思いしてる百合漫画" #9, p.6, comic by 灰島/藤乃

Fig. 4: *"雨ちゃんと魔人"*
From "悪魔のメムメムちゃん" #107, p.6, comic by Keitarō Yotsuya, publised 23 February 2021



Fig. 5: *"わっか邪魔では？"*
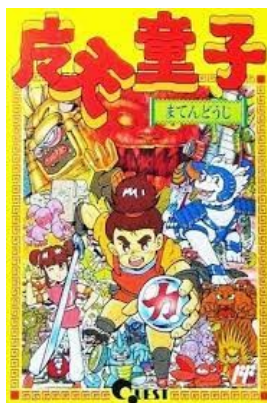From "ぐだぐだ太閤伝 ZIPANG" #11, p.3, comic by Keitarō Yotsuya, published in 2023



Fig. 6 *"魔天童子"*
From, "Maten Dōji," Family Computer software from Quest, released August 24, 1990.

<u>庁(囗广マ) for "摩"</u>

This abbreviation has the same shape as the abbreviation for "魔" and is mainly used to write "Tama(多摩)," the name of a place in Tokyo, on signs at public facilities and handwritten billboards. The letters in Fig. 1 are also an example of this usage.



*Fig. 7: "多摩ミシン商会"*
this picture is posted on Instagram by yosh003 at Oume, Tokyo on 31 March 2020 URL: https://www.instagram.com/p/B-YhHhhpunS
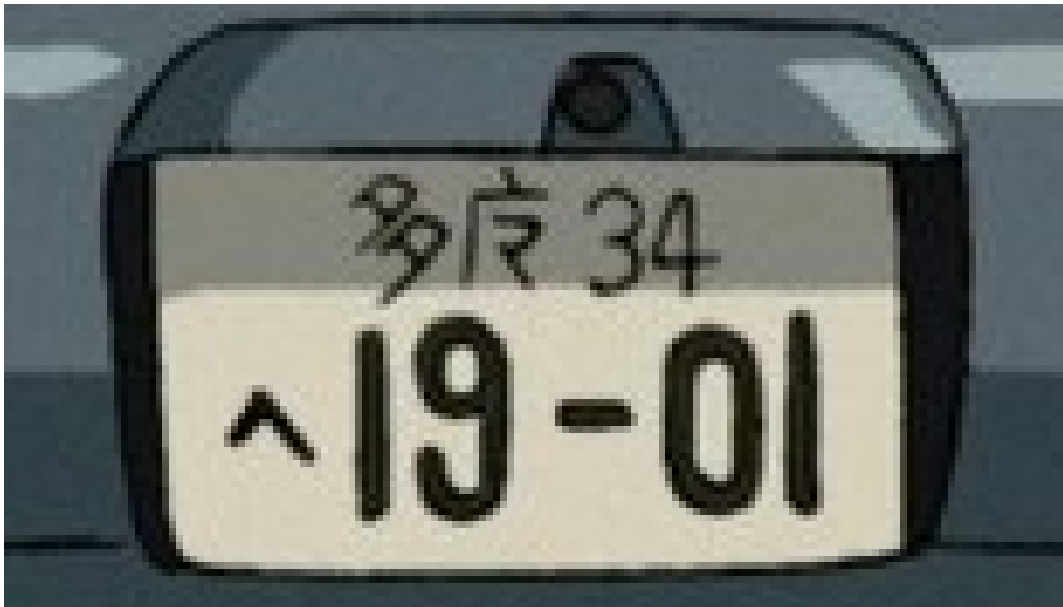


*Fig. 8: "多摩 34 へ 19-01"*
From "Spirited Away" (animated film released on July 20, 2001) by Hayao Miyazaki.

閉店

この度、一部の従業員達
（多摩大学学生）による 不衛
生な行為により営業を
停止させて頂く事になりま
した。皆様には深くおわび申し
上げると共に永きに渡り御
顧頂きまして本当にありが
とうございました
㈲泰尚
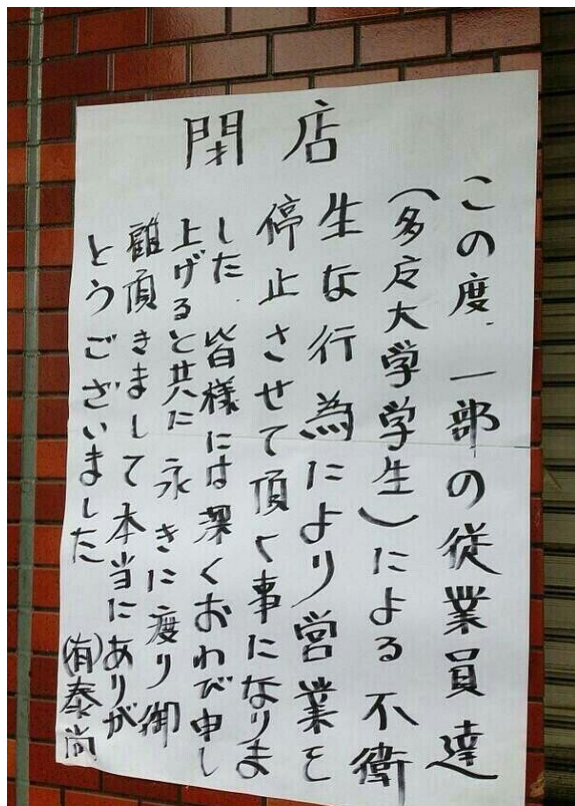
*Fig. 9: "多摩大学学生"*

this picture is posted on Twitter by @mosukoke on 29 April 2014 URL: https://twitter.com/mosukoke/status/461016869873262592

劇(<span>□</span>虍 刂) for "劇"

This abbreviation is mainly used on old Japanese theatre signs.



*Fig. 10:"*「劇」小劇場*"*
this picture is posted on Instagram by tm_walker1 at Shimokitazawa, Tokyo on 27 May 2023 URL: https://www.instagram.com/p/Csu4PrSvsSQ/



*Fig. 11:"オーエス劇場"*
this picture is posted on Instagram by tokyo5963 at Nishinari, Ōsaka on 19 October 2019 URL: https://www.instagram.com/p/BabCAASAW5Z/



*Fig. 12:"日劇"*
this picture is posted on Instagram by murasakiteikyotaro at Nishinari, Ōsaka on 27 May 2023 URL: https://www.instagram.com/p/CsutOGXSiTV/

These abbreviations are mainly used for signage at factories and facilities of companies that handle electrical goods. There are several variants, but this proposal introduces three of them that are widely used in Japan.

1. "枛" for "機"



*Fig. 13: "店建設機械鋼材"*

this picture is posted on Instagram by yakkll on 27 March 2016 URL: https://www.instagram.com/p/BDcmYwWNGBz/



*Fig. 14: "機械工具　昭和工機㈱"*

this picture is posted on Instagram by takeshi.unno.94 on 14 August 2022 URL: https://www.instagram.com/p/ChOeV5sPL04/

*Fig. 15: "事務機器センター"*
this picture is posted on Instagram by takeshi.unno.94 on 22 July 2022 URL: https://www.instagram.com/p/CgTnY3wPrn8/

*Fig. 16: "貸洗濯機" and "貸乾燥機"*
this picture is posted on Twitter by @watertowerUC on 4 September 2021 URL: https://twitter.com/watertowerUC/status/1433969408813142016



*Fig. 17: "消防機具庫"*
this picture is posted on Twitter by @V80Hh6 on 8 December 2022 URL: https://twitter.com/V80Hh6/status/1600802098098028545

2. "機" for "機"



*Fig. 18: "~力工機株式会社"*
this picture is posted on Instagram by yukiakari at Nakano Broadway, Tokyo on 17 April 2017 URL: https://www.instagram.com/p/BS-VCeUhJ-P/



*Fig. 19: "クボタ　農業機械　愛用者"*
this picture is posted on Twitter by @tenereisobe on 9 January 2023 URL:https://twitter.com/tenereisobe/status/1612125618812096513



*Fig. 20: "~う形散布機"*
this picture is posted on Twitter by @umesyua on 21 December 2022 URL:https://twitter.com/umesyua/status/1605377460983705601

*Fig. 21:* "*中有楽　機械室*"
this picture is posted on Twitter by @mashabow on 14 October 2017 URL:https://twitter.com/mashabow/status/919210505335578626



*Fig. 22:* "*住宅設備機器 展示場*"
this picture is posted on Twitter by @kirnura on 5 October 2022 URL:https://twitter.com/kirnura/status/1577599483906592769

"選" for "選"

This abbreviation is mainly used for old Japanese shop signs.



*Fig. 23: "特選靴"*
this picture is posted on Instagram by yaaaabuuuu on 16 November 2017 URL: https://www.instagram.com/p/BbhofjWDAFT/



*Fig. 24: "特選若鳥"*
this picture is posted on Instagram by _kacha_shimane on 27 November 2016 URL: https://www.instagram.com/p/BNSnXMtjbqQ/

"選" for "選"

尸 for 層



Fig. 25: "三層サクビ"
This picture is posted on Instagram by marusenokinawa on 11 December, 2017 URL : https://www.instagram.com/p/BcjBI9zlJzF/
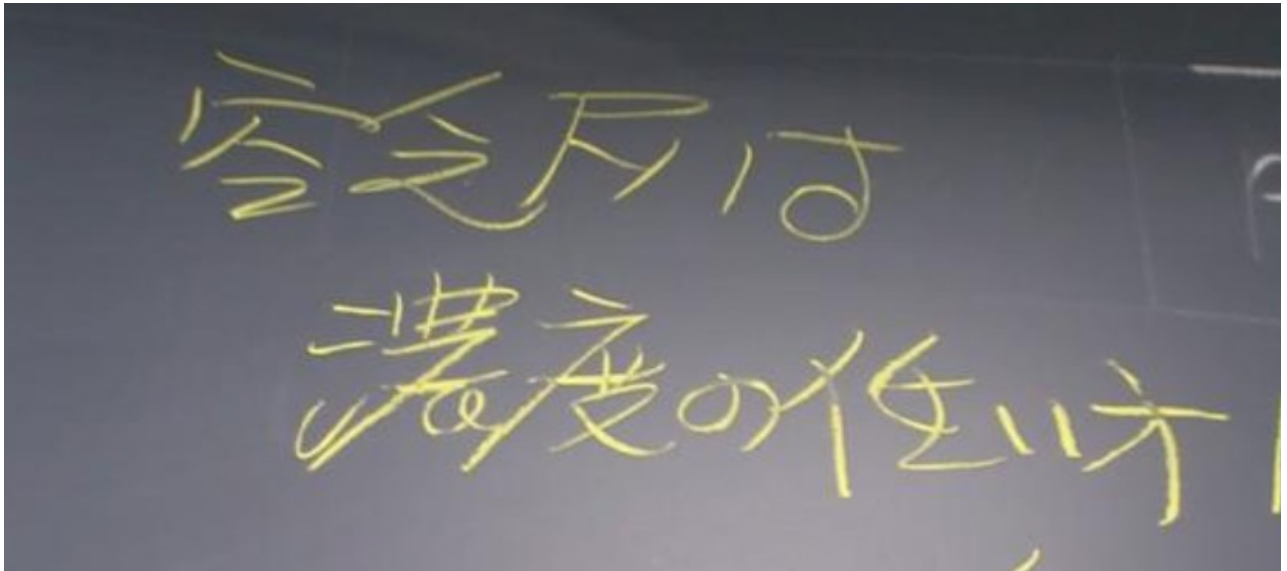


Fig. 26: "空乏層は濃度の低い方"
This picture is posted on Twitter by @mashabow on 21 January, 2013 URL : https://twitter.com/mashabow/status/293341660409102337
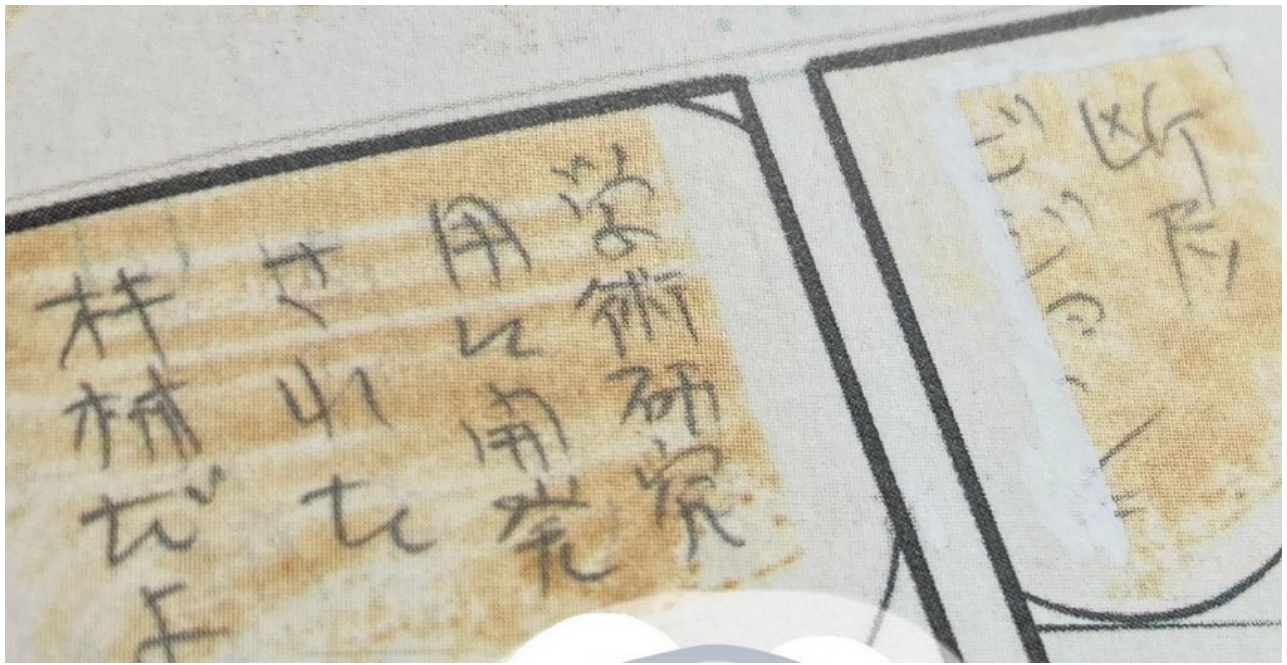
*Fig. 27: "断層ビジョン　学術研究用に開発された機械だよ"*

This picture is posted on Twitter by @coeos_ on 27 December, 2021 URL : https://twitter.com/coeos_/status/1475397378681503745
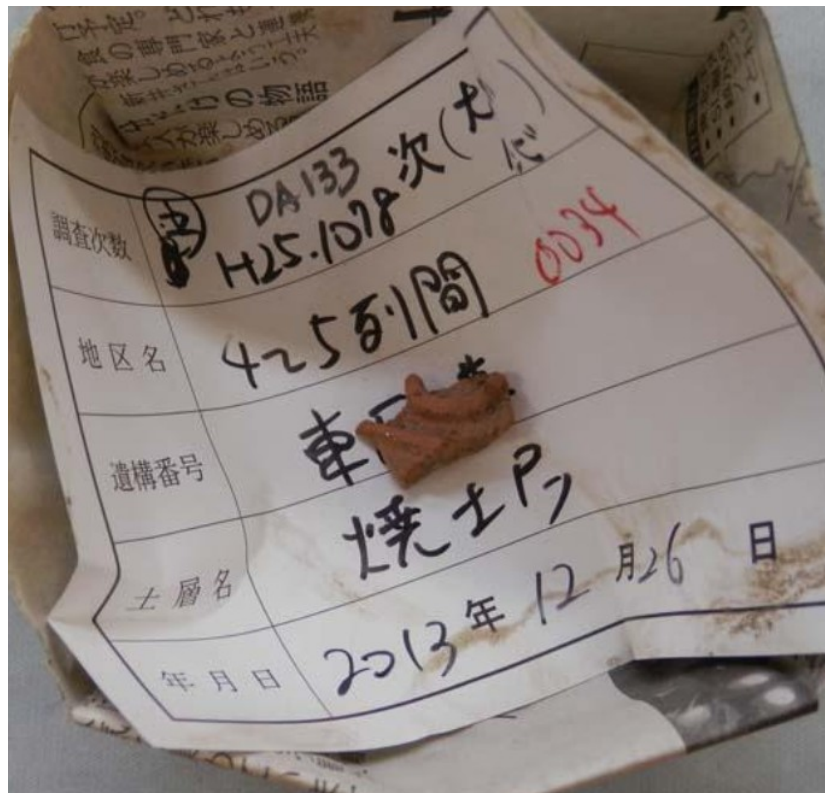


*Fig. 28: "焼土層"*

This picture is posted on Twitter by @Brise_Marine on 19 November, 2014 URL : https://twitter.com/Brise_Marine/status/535015994209472512

*Fig. 29: "綾鷹"*
*This picture is posted on Twitter by @wdicclub on 4 December, 2014 URL : https://twitter.com/wdicclub/status/540349731780575232*
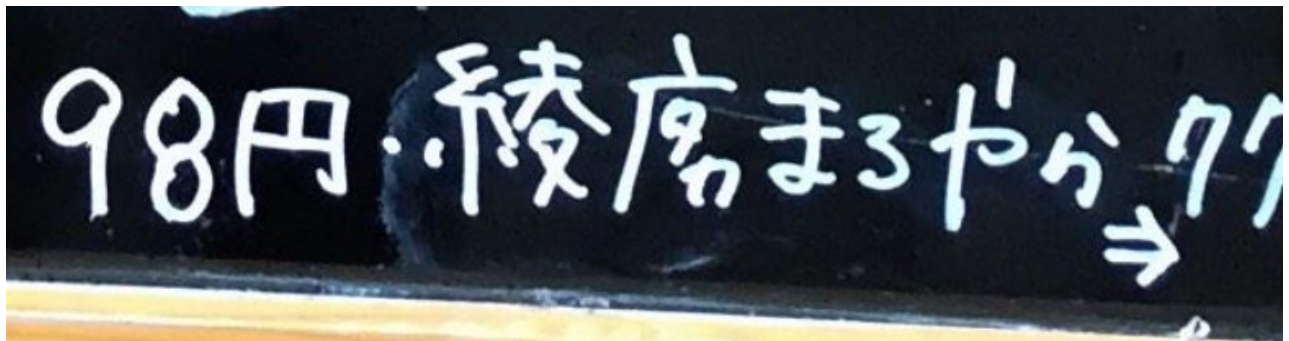


*Fig. 30: "綾鷹まろやか"*
*This picture is posted on Twitter by @wdicclub on 4 December, 2014 URL : https://twitter.com/wdicclub/status/540349731780575232*
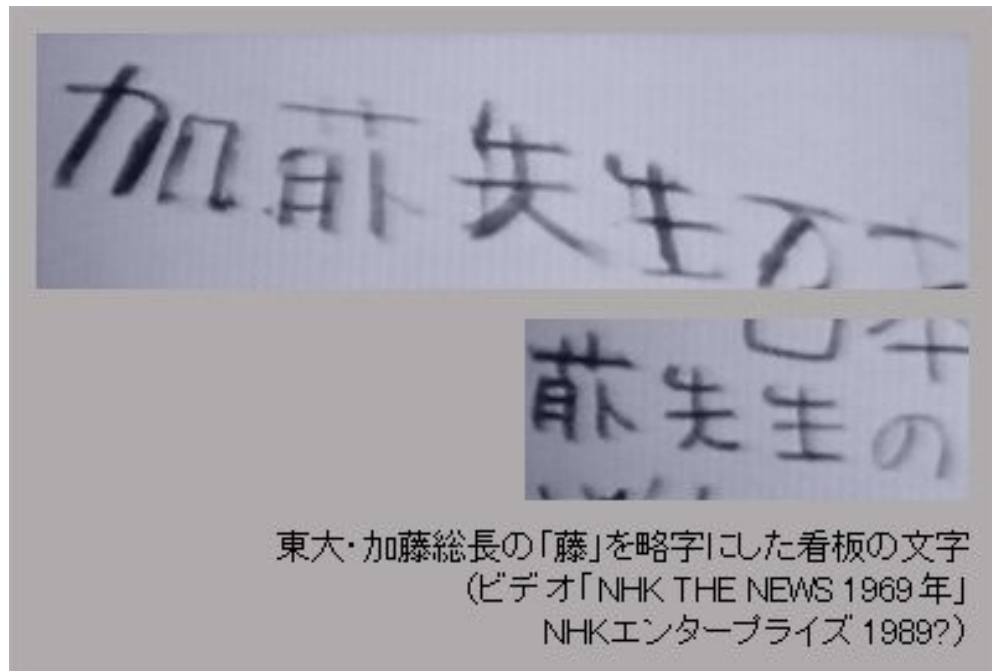
*Fig. 31: "加藤先生万才"*
from "NHK THE NEWS 1969", NHK エンタープライズ, 1969, URL : http://www.asahi-net.or.jp/~qm4h-iim/x_69_kato.jpg
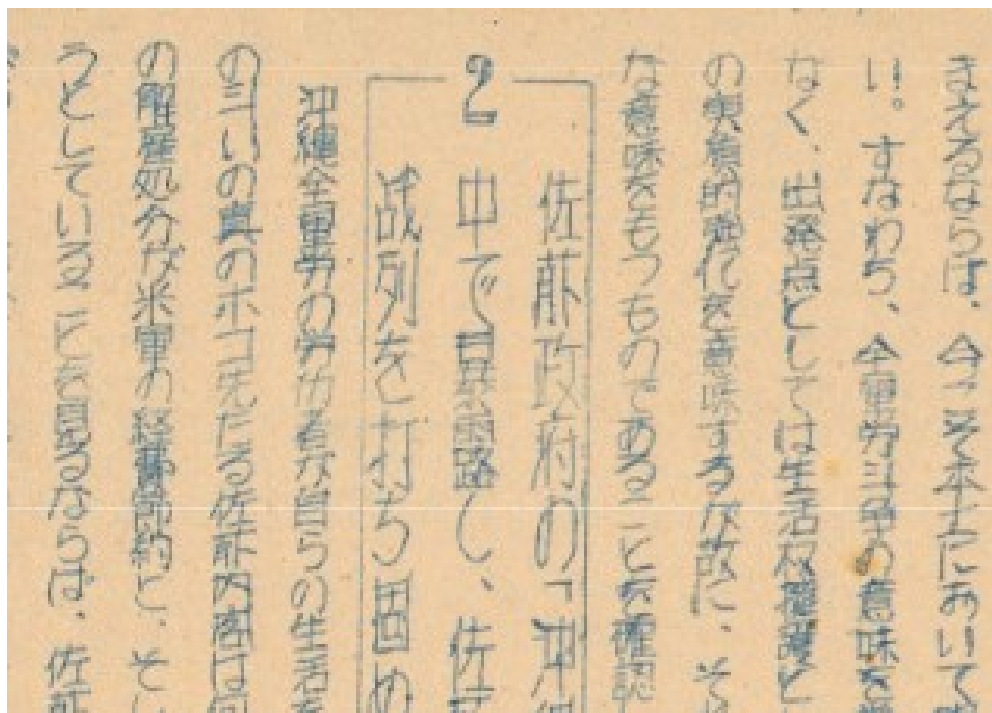


*Fig. 32: "佐藤政府の"*
from "学生共闘", 大阪市立大学, 17 January, 1970, URL : https://assert.jp/wordpress/data/bira/197001/MDS00282.pdf

☐ for 国
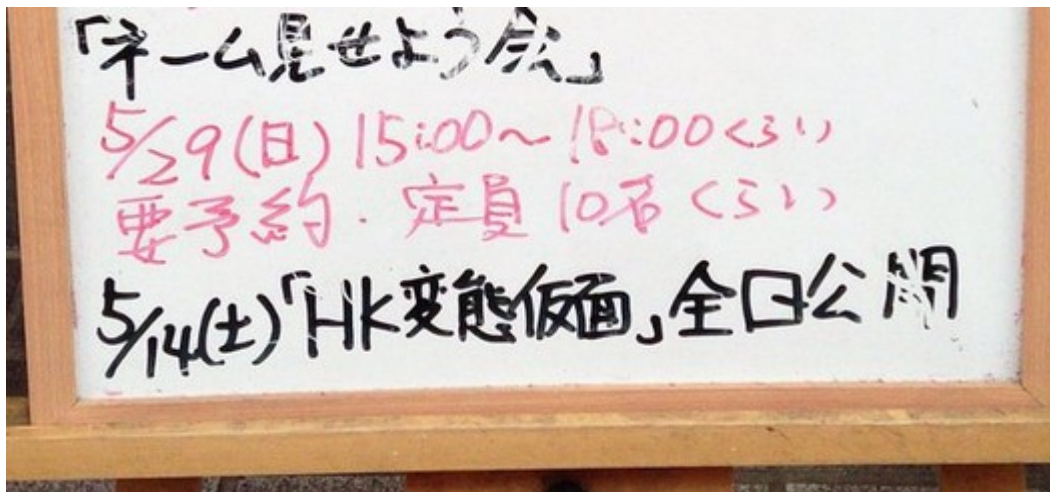


*Fig. 33: "「HK変態仮面」全国公開"*
This picture is posted on Twitter by @mashabow on 17 May, 2016 URL : https://twitter.com/mashabow/status/732545556132728833



*Fig. 34: "国葬"*
This picture is posted on Twitter by @farunifa on 28 September, 2022, at Kobe University URL :
https://twitter.com/farunifa/status/1574984240860901377

*Fig. 35: "斗争全国ゼネ..."*
*This picture is posted on Twitter by @mosukoke on 16 July, 2015 URL :*
*https://twitter.com/mosukoke/status/621598835317639168*
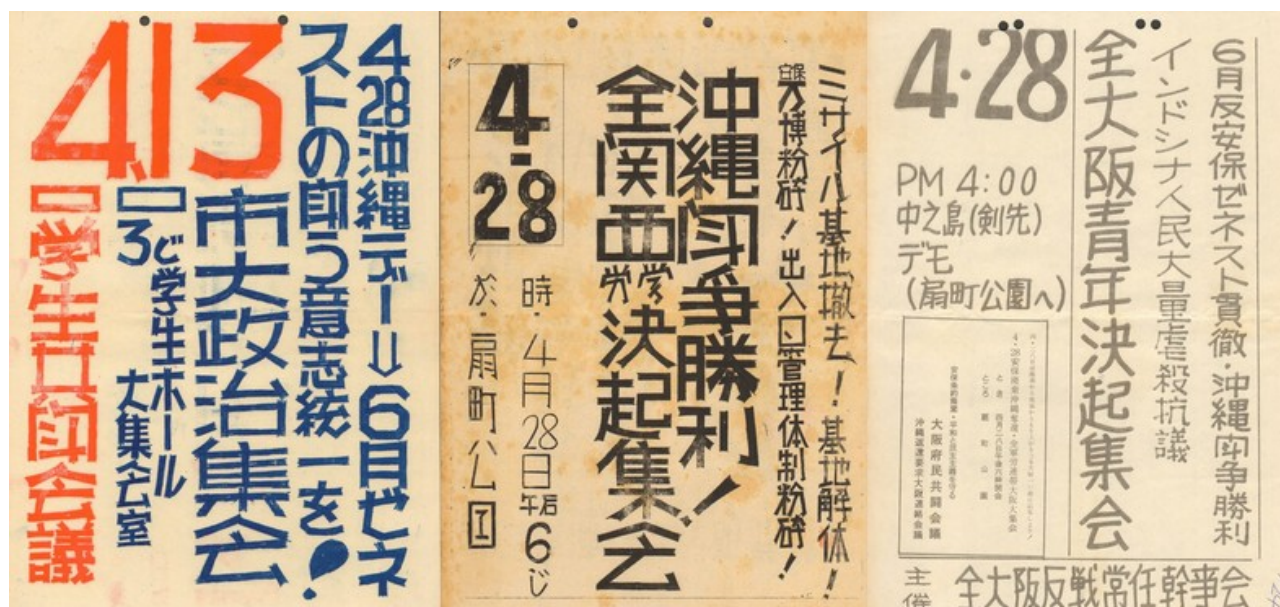


*Fig. 36: "出入国管理体制粉砕！"*
*from "大阪市立大学　ビラコレ　1970-04 ＜新歓期、4.28 沖縄闘争＞" URL : https://assert.jp/archives/5084*
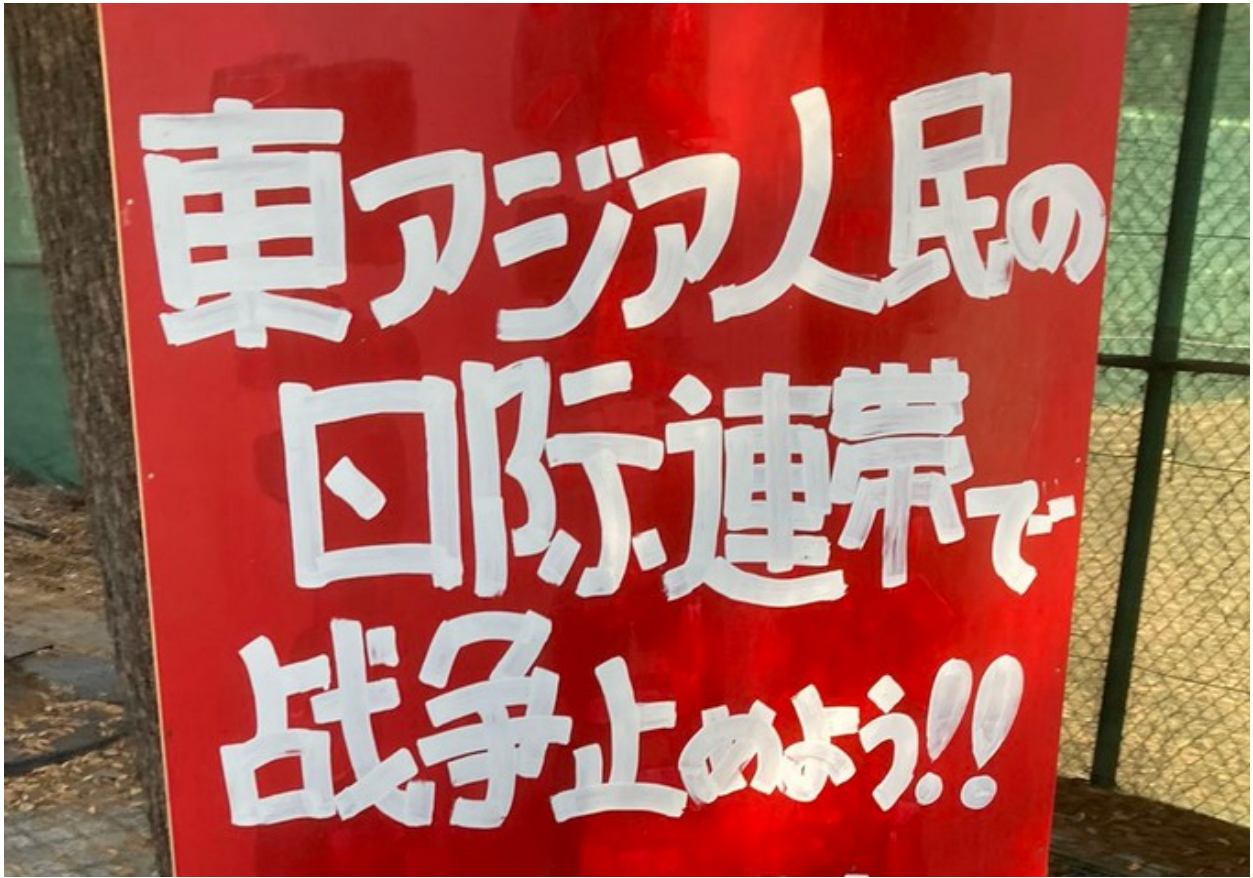
*Fig. 37:* "国際連帯"
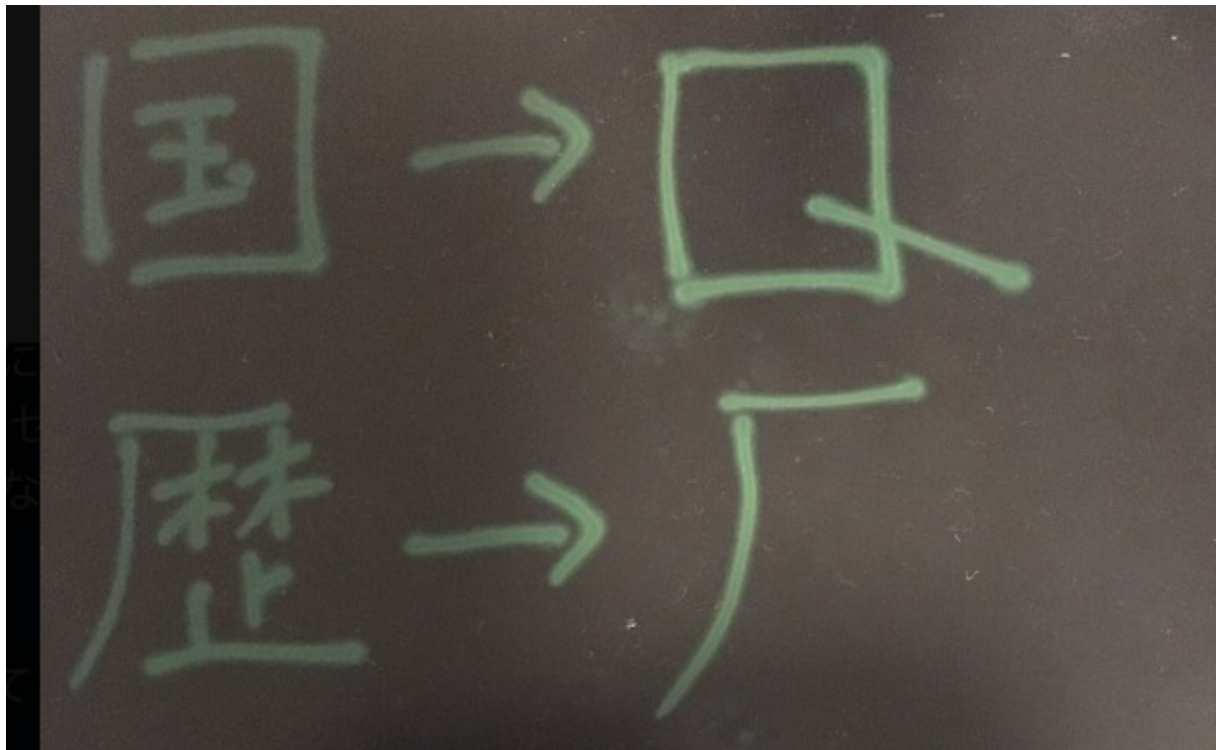
*Fig. 38:* The Sample of abbreviations
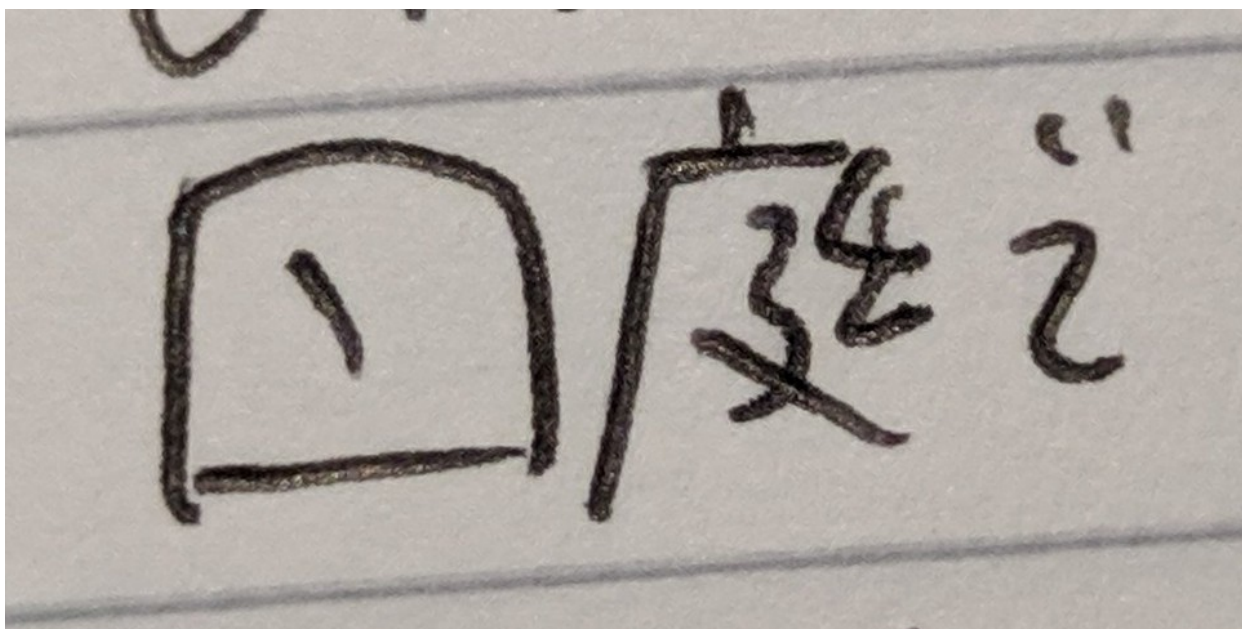
*Fig. 39: "園庭で"*
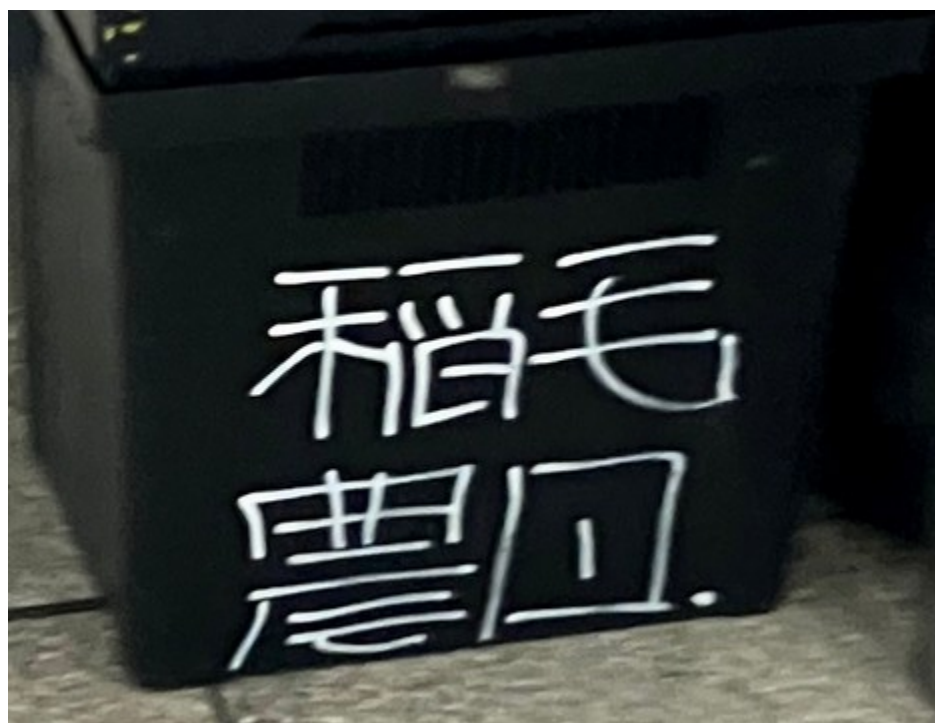
*Fig. 40: "稲毛農園"*

**Properties**

All characters proposed follow the CJK Ideographs' properties.

U+XXXXX; <CJK Ideograph-XXXXX>;Lo;0;L;;;;;N;;;;;

<table>
<tr><td colspan="2">

**ISO/IEC JTC 1/SC 2/WG 2**
**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS**
**FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**[1]
**Please fill all the sections A, B and C below.**
**Please read Principles and Procedures Document (P & P) from** http://std.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://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
**See also** http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

</td></tr>
</table>

**A. Administrative**

1. **Title:** *Proposal to Encode Abbreviated Ideographs(Kanjis) in Japanese*
2. Requester's name: *Gen Kojitani*
3. Requester type (Member body/Liaison/Individual contribution): *Individual contribution*
4. Submission date: *2023/06/22*
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):
        Proposed name of script:
    b. The proposal is for addition of character(s) to an existing block: *yes*
        Name of the existing block: *CJK Ideographs Extended-I?*
2. Number of characters in proposal: *9*
3. Proposed category (select one from below - see section 2.2 of P&P document):
    A-Contemporary    *X*    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? *no*
    a. If YES, are the names in accordance with the "character naming guidelines"
        in Annex L of P&P document?
    b. Are the character shapes attached in a legible form suitable for review?
5. Fonts related:
    a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
        *Gen Kojitani*
    b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
        *IPA Font License*
6. References:
    a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?
    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)? *no*

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.

---

## 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.)?  *no*
        If YES, with whom? _____
        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)  *common*
    Reference: _____
5. Are the proposed characters in current use by the user community?  no
    If YES, where?  Reference: _____
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
    in the BMP?  *no*
        If YES, is a rationale provided? _____
            If YES, reference: _____
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?  *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, or could be confused with, an existing character?  *no*
        If YES, is a rationale for its inclusion provided? _____
            If YES, reference: _____
11. Does the proposal include use of combining characters and/or use of composite sequences?  *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 characters?  *no*
    If YES, are the equivalent corresponding unified ideographic characters identified? _____
        If YES, reference: _____

25

**(Detach the following pages from the filled summary form before submitting).**

## Information accompanying submissions

The process of deciding which characters should be included in the repertoire of the standard by WG 2 depends on the availability of accurate and comprehensive information about any proposed additions. WG 2, at its San Francisco meeting 26, designed a form (template) that will assist the submitters in gathering and providing the relevant information, and will assist WG 2 in making more informed decisions.

This form is part of the WG 2 Principles and Procedures document and is available on line at http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html.

A duly completed proposal summary form must accompany each new submission. The latest version of this form must be used in submissions. Such a form will assist WG 2 to better evaluate the proposal, and progress the proposal towards a speedier acceptance and inclusion in the standard. Submitters are also requested to ensure that a proposed character does not already exist in the standard.

Submitters are encouraged to visit the "Where is my Character" page on the Unicode web site for more information on checking if their proposed character or script is already encoded in the standard, or a similar proposal has already been made by someone else. There are also several electronic discussion lists maintained by the Unicode consortium that one could use to discuss with other experts internationally on various subjects related to the standard. Submitters are also encouraged to familiarize themselves with ISO/IEC TR15285 – Character Glyph Model. The latest version of ISO/IEC 10646 as well as ISO/IEC TR15285 are available on line under 'Freely Available Standards' at: http://standards.iso.org/ittf/PubliclyAvailableStandards/ .

National bodies should take note of the guidelines in section *C.4 Some guidelines on proposing new material as ballot comments* in the P&P document.

In addition to text extracted from the standard in *Annex L: Character-naming guidelines* in the P&P document, the following definitions from the standard are also referenced in the proposal summary form:

> *Clause 4.12 Combining character*:
> A member of an identified subset of the coded character set of ISO/IEC 10646 intended for combination with the preceding non-combining graphic character, or with a sequence of combining characters preceded by a non-combining character (see also 4.14).
> NOTE – ISO/IEC 10646 specifies several subset collections, which include combining characters.

> *Clause 4.14 Composite sequence*:
> A sequence of graphic characters consisting of a non-combining character followed by one or more combining characters (see also 4.12).
> NOTE 1 – A graphic symbol for a composite sequence generally consists of the combination of the graphic symbols of each character in the sequence.
> NOTE 2 – A composite sequence is not a character and therefore is not a member of the repertoire of ISO/IEC 10646.

## Submitter's responsibilities

The national body or liaison organization (or any other organization or an individual) proposing new character(s) or a new script shall provide:

1. Proposed category for the script or character(s), character name(s), and description of usage.
2. Justification for the category and name(s).
3. A representative glyph(s) image on paper:
   If the proposed glyph image is similar to a glyph image of a previously encoded ISO/IEC 10646 character, then additional justification for encoding the new character shall be provided.
   *Note:* Any proposal that suggests that one or more of such variant forms is actually a <u>distinct</u> character requiring separate encoding should provide detailed, printed evidence that there is actual, contrastive use of the variant form(s). It is insufficient for a proposal to claim a requirement to encode <u>as characters</u> in the Standard, glyphic forms which happen to occur in another character encoding that did not follow the Character-Glyph Model that guides the choice of appropriate characters for encoding in ISO/IEC 10646.
   *Note:* WG 2 has resolved in Resolution M38.12 not to add any more Arabic presentation forms to the standard and suggests users to employ appropriate input methods, rendering and font technologies to meet the user requirements.
4. Mappings to accepted sources, for example, other standards, dictionaries, accessible published materials.
5. Font resource:
   A font must be provided to the Editor promptly after the characters have been approved by WG2 for use in printing the charts. If a font is not provided, the Editor cannot include the glyphs in the charts and, as a result, the repertoire of characters corresponding to these glyphs will not be included in draft amendments. Fonts must be in one of the following formats (in preferential order): OpenType, TrueType, Postscript Type 1.
   Fonts submitted must have no license restrictions that prevent embedding into PDF documents. Because of synchronization between ISO/IEC 10646 and The Unicode Standard, any grant of license must cover use in publishing both standards and related documents (see resolution M45.30). For technical reasons, the editors must be able under that license to freely modify or replace glyph outlines in their copies of the fonts. The submitter should be prepared to provide fonts of suitable quality and license conditions, unless equivalent fonts are already available to the editors.
6. List of all the parties consulted:
   Towards ensuring that all significant stakeholders have been consulted, WG 2 requests that submitters solicit feedback from other groups who use the proposed characters, and list these contacts. Submitters are encouraged to provide the email id-s of the submitters as well as other experts who have been consulted to facilitate any clarification queries.
7. Equivalent glyph images:
   If the submission intends using composite sequences of proposed or existing combining and non-combining characters, a list consisting of each composite sequence and its corresponding glyph image shall be provided to better understand the intended use.
8. Compatibility equivalents:
   If the submission includes compatibility ideographic characters, identify (per resolution M45.29):
   - the source, which contains two distinct code positions that correspond to a single unified CJK Ideograph character of ISO/IEC 10646
   - the ISO 10646 unified CJK Ideograph
   - the code position in the source for the unified CJK ideograph
   - the code position in the source for the proposed compatibility ideograph
9. Properties that may affect the BiDi processing
   Any BiDirectional algorithm related properties associated with the characters should be spelled out (see UAX#9 - http://www.unicode.org/reports/tr9/).
10. Any additional information that will assist in correct understanding of the different characteristics and linguistic processing of the proposed character(s) or script.
11. If any of the proposed characters are suitable as *syntax* characters please take note of guideline D.2.5 (on page 24 and of Unicode Standard Annex 31 at http://www.unicode.org/reports/tr31/.
12. If the proposal is for a new script, identify the standardized characters that are commonly used directly in the script, or proposed to be unified with the characters of the script, in particular standardized characters allocated in different blocks. Examples include punctuation marks and combining marks. Such information will assist in assigning properties for characters shared across multiple scripts or in identifying character repertoires needed to support particular languages.
13. If you are aware of already standardized characters that are visually close to any of the proposed characters, you are invited to list them in the proposal. This will assist in the analysis of the script for 'visually confusables', towards providing additional guidance on use of the standard from a security perspective (see UTR#36 - Unicode Security Considerations – at http://www.unicode.org/reports/tr36/).