



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 — Symbols used in Dictionaries
Source: Asmus Freytag
Status: Expert contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Related: N2656. Responding to N2645, N2647, N2587, and other proposals

Several recent documents propose characters that are used in dictionaries, or relate to practices used in dictionaries. This document intends to extend the research on some of these issues, document additional instances, and propose additional characters.

Note: A significant amount of information about use of already encoded special characters in dictionaries can also be found in Unicode Standard Annex #14, “Line Breaking Properties”, which can be accessed at <http://www.unicode.org/reports/tr14/>.

Use of symbols for subject classification in dictionaries

Document N2587 proposes a set of characters for genealogical purposes. One of the characters (for ‘killed in action’) is part of a larger set of characters used in dictionaries. This document extends the research to complete coverage of that usage.

Dictionaries often need a shorthand notation to classify terms by subject matter or by other usage. A system of using iconic symbols for subject matter classification is fairly widespread, especially on the European continent. However, there are both differences in elaboration as well as some differences in particular symbols chosen.

Dictionaries represent a category of texts that are in common and active use and which are not infrequently cited in other works. Being able to represent the contents of a dictionary entry in plain text increases the robustness and portability of such usage.

In this context, it is less important that all dictionaries be ‘normalized’ to using the same symbol for the same semantic, or even just the same set of symbols. What matters for plain text is that the usage of a given source finds a recognizable representation in Uni-

code. The symbols used for this classification are dingbats, but their function is to serve directly as a shorter and more easily recognizable stand-in for one or more words or abbreviations.

Where glyphs are equivalent depictions of the same object, and can stand for each other in more than one context, unification to a set of generic dingbats would be appropriate. However, where the representative object for a category is different, for example a hypothetical use of both *ship* and *anchor* to indicate nautical terms, both symbols should remain distinct. In other words, what would be encoded in this hypothetical situation would be a symbol for a ship and a symbol for an anchor. The fact that they are both used to indicate ‘nautical term’ is an externally applied convention.

The set of proposed characters derived from the examples below are based on that underlying analysis of the issue.

Note: some of the symbols are more traditional and are both used, and readily recognized outside the context of the notation for a given publication. The symbols are arranged in that order in the table of proposed symbols.

Examples of subject classification symbols

The following examples show sets of symbols actually used for classification of subject matter in dictionaries.

Sample 1

1. Symbols — Bildliche Zeichen			
~}	siehe Seite 13: Fette und einfache Tilde;	⚡	Eisenbahn, <i>railway</i> .
o}	<i>see page 13: Bold-faced and Simple Tilde.</i>	✈	Flugwesen, <i>aviation</i> .
F	familiär, <i>familiar</i> ; Umgangssprache, <i>colloquial</i> .	✉	Postwesen, <i>postal affairs</i> .
V	vulgär, <i>vulgar</i> ; unanständig, <i>indecent</i> .	♪	Musik, <i>musical term</i> .
🌿	Botanik, <i>botany</i> .	🏠	Architektur, <i>architecture</i> .
⊕	Handwerk, <i>handicraft</i> ; Technik, <i>engineering</i> .	⚡	Elektrotechnik, <i>electrical engineering</i> .
⚒	Bergbau, <i>mining</i> .	⚖	Rechtswissenschaft, <i>legal term</i> .
✖	militärisch, <i>military term</i> .	ℵ	Mathematik, <i>mathematics</i> .
⚓	Schiffahrt, <i>nautical term</i> .	🌾	Landwirtschaft, <i>agriculture</i> .
†	Handel u. Wirtschaft, <i>commercial term</i> .	🧪	Chemie, <i>chemistry</i> .
		🩺	Medizin, <i>medicine</i> .
		→	siehe Seite 14: Verweiszeichen; <i>see page 14: Mark of Reference</i> .

The electric arrow can probably be unified with U+21AF ↘ even though the representative glyph in the standard is less canonical than the glyph here. The musical note can be unified with U+266A ♪.

A nearly identical set of symbols has been identified by Michael Everson in an English-Turkish dictionary.

Sample 2

Symbols — Semboller

F	<i>colloquial language</i>	<i>konuşma dili</i>
P	<i>provincialism</i>	<i>taşra dili</i>
V	<i>vulgar</i>	<i>kaba konuşma</i>
☘	<i>botany</i>	<i>botanik</i>
⊕	<i>mechanics</i>	<i>mekanik</i>
⌘	<i>mining</i>	<i>madencilik</i>
⌘	<i>military term</i>	<i>askeri terim</i>
⚓	<i>nautical term</i>	<i>denizcilik terimi</i>
🚂	<i>railway, railroad</i>	<i>demiryolu</i>
✈	<i>aviation</i>	<i>havacılık</i>
✉	<i>postal affairs</i>	<i>postacılık</i>
🎵	<i>musical term</i>	<i>müzik terimi</i>
⚡	<i>electrical engineering</i>	<i>elektrik mühendisliği</i>
⚖	<i>legal term</i>	<i>hukuk terimi</i>
△	<i>mathematics</i>	<i>matematik</i>
→	<i>agriculture</i>	<i>ziraat</i>
⚗	<i>chemistry</i>	<i>kimya</i>
⚕	<i>medicine</i>	<i>tıp</i>

A slightly more elaborate set of symbols was located by Michael Everson in an Icelandic dictionary). See Sample 3 below. All three samples show a different design for the symbol for aviation. That clearly indicates that the *aviation* symbol, as one would have expected, can be unified with a generic depiction of an airplane, so a unification with 2708 AIRPLANE seems allowable.

The three-leaf clover from the sample below can be unified with a *Shamrock*. Such a symbol is already proposed in N2586R, and has been reviewed favorably by UTC. Therefore it is not listed separately in the Table of Proposed Symbols. Should N2586R not be approved, then the shamrock character should be considered part of this proposal.

Sample 3

Ⓐ Ásatrú, norræn goðafræði	🏠 byggingarlist	
Ⓑ bókmenntafræði	* endurgerð orðmynd (oftast frumnorræn)	⊕ sjaldgæ.
Ⓓ dýrafræði	✈ flugmál	↓ sjómennska
Ⓔ eðlisfræði og efnafræði	† fornt eða úrelt mál, óbundið	* skáldamál, gamalt og nýtt
Ⓕ heimspeki	♣ grasfræði	○ slangur eða óstaðfest nýyrði
Ⓖ jarðfræði	✳ guðfræði og trúarbrögð	⊙ staðbundið málfar
Ⓜ málfræði, þar með bragfræði	£ hagfræði og verslun	☞ stjórnufræði
Ⓢ sálfræði	♏ iðnaður	∞ stærðfræði
Ⓤ uppeldisfræði, skólamál	⚡ íþróttir) tímatal
Ⓥ veðurfræði	🏡 landbúnaður	🎵 tónlist
Ⓟ þjóðtrú	💰 líffræði, lífeðlisfræði, læknisfræði og fleiri skyldar greinar	⚙ tækni
? vont mál, orð eða merking sem forðast ber í íslensku (yfirleitt aðeins sett þar sem betra orð er sýnt í skýringu)	⚖ lögfræði og önnur samfélagsfræði	→ verður að
o: skýring, það er	⚡ rafmagnsfræði	← komið af
# afleidd eða huglæg merking	‡ sagnfræði, sögulegt efni (við-	➡ sjá, samanber

Sample 4

A sample from a Danish-Islandic dictionary, showing a glyph variant on the anchor.

Ⓑ Bókmenntir, bragfræði, stílfraði	✂ Hernaður
Ⓓ Dýrafræði	♣ Iðnaður, handverk
Ⓔ Eðlis- og efnafræði	⚡ Íþróttir
Ⓕ Félagsfræði, stjórnsmál	💰 Líffræði, læknisfræði, lyfjafræði og fleiri skyldar greinar
Ⓕ Heimspeki	⚖ Lögfræði
Ⓖ Jarðfræði	⚡ Sagnfræði
Ⓜ Listir (nema tónlist og bókmenntir)	🎵 Sjómennska
Ⓜ Málfræði, hljóðfræði	🎮 Spil
Ⓢ Sálar- og uppeldisfræði, kennsla	☆ Stjórnufræði
Ⓤ Tölvufræði	π Stærðfræði
🏠 Byggingarlist	🎵 Tónlist
♣ Grasfræði, garðyrkja	
✳ Guðfræði	
£ Hagfræði, verslun, viðskipti	* fyrir framan orð tákna nýyrði

Sometimes only a subset of these symbols are used, but as can be seen, with very consistent representation. The following excerpt is from the list of symbols used in an English-Swedish dictionary. In addition to the symbols shown here, that dictionary also uses the sans-serif capital letters F and V with the same meanings as in the sample above. We may

consider this as pointing towards an overall pattern of usage that transcends individual editors' or publishers' style, but is fairly common.

Sample 5

- * sällsynt.
- ✕ militärterm.
- ⚓ sjöterm.

Here is a sample from a typical page of the English-German dictionary showing the symbols as they are used in an actual entry:

Sample 6

worries von Sorgen bedrückt; **ab-**
'sorb·ent [-bənt] **I.** *adj.* absorbie-
rend, aufsaugend: ~ cotton ⚓ Ver-
bandwatte; **II.** *s.* Absorpti'ons-
mittel *n*; **ab'sorb·ing** [-biŋ] *adj.*
□ 1. aufsaugend; *fig.* fesselnd, pak-
kend; 2. ⊕, *biol.* Absorptions...,
Aufnahme... (*a.* †); **ab·sorp·tion**
[əb'sɔ:pʃən] *s.* 1. *a.* ⚡, ⚓, ⊕, *biol.*,
phys. Auf-, Einsaugung *f*, Aufnah-
me *f*, Absorpti'on *f*; Vereinigung *f*;

Three other canonical subject classification symbols, not documented here, include the mask for *theater*, the film clip for *cinematic*, or *cinematography* and the palette for *art*.

Other symbols used in dictionaries

There are a number of symbols in widespread use in dictionaries that are missing from the Unicode standard. For pronunciation-related symbols see a separate document (N26xx).

The use of some form of tilde to repeat terms is common. Where the capitalization changes in the new context, a circle is added:

Sample 7

Wenn sich die Anfangsbuchstaben ändern (groß zu klein oder umgekehrt), steht statt der Tilde das Zeichen ̣ oder ̤.

Beispiele:

Brit|e *m*; ̣**in** *f*; ̤**isch** *adj.*; **falt|bar** *adj.*; ̤-**blatt** *n*; ̤**boot** *n*; **höchst...**: ̤**wert** *m*; ̤**zahl** *f*; ̣**zulässig** *adj.*; **Güterstand** *m*: ehelicher ̣ matrimonial regime; **heilig** *adj.* holy *etc.*; der ̤e **Geist** the Holy Ghost; **hängen** *v/i.* hang *etc.*; ̤ *n* hanging; **harsch** *adj.* harsh *etc.*; ̤ *m* crust.

As can be seen in this sample, the tilde is not the ordinary tilde, but a low tilde that rests on the baseline. The symbol that combines the tilde with the circle is only topologically similar to the combination of U+007E TILDE and U+030A COMBINING CIRCLE ABOVE, but in actual practice such a combination (̤) looks noticeably different (and the circle, which is a distinguishing feature appears de-emphasized). Note that while the text mentions the tilde by name, not describe it as a tilde with a circle added, rather treating it as a symbol of its own.

The following is a sample from Duden Rechtschreibung, a work usually considered the authority on German orthography, containing three further symbols not yet encoded in Unicode.

Sample 8

- ⋮ Die senkrechte punktierte Linie dient zur Angabe der Nottrennung, z. B. Nati:on (vgl. S. 50, R 176).
 - Ⓜ Als Warenzeichen geschützte Wörter sind durch das Zeichen Ⓜ kenntlich gemacht. Etwaiges Fehlen dieses Zeichens bietet keine Gewähr dafür, daß es sich hier um ein Freiwort handelt, das von jedermann benutzt werden darf.
 - Der waagerechte Strich vertritt das Stichwort buchstäblich, z. B. ab; - und zu; oder: Brett s; -[e]s, -er; oder: Allerlei s; -s, -s; Leipziger -.
 - ... Drei Punkte stehen bei Auslassung von Teilen eines Wortes, z. B. Buntdruck (*Mehrz. ...drucke*); oder: Streß *m*; ...s*ses*, ...s*se*. Über die Punkte bei Aussprachebezeichnungen vgl S. 13, VI.
 - ⌣ Der Bogen steht innerhalb einer Ableitung oder Zusammensetzung, um anzuzeigen, daß der vor ihm stehende Wortteil bei den folgenden Wörtern an Stelle der drei Punkte zu setzen ist, z. B. Augen**⌣**braue, ...diagnose.
- The first is a series of four dots, spanning the height of the entire line and used in the cited work to indicate places for emergency word breaks.
 - The second is based on the German word for trademark and is used where an English language work might use ® or ™. [Indirect evidence for the use of this symbol in online reference works for Microsoft Office 2000 comes from the fact that the enclosed ‘Verdana Ref’ font contains a glyph for this symbol – see table.]
 - The third is used to split a term into two parts, one of which can then be repeated by an ellipsis (...). Other dictionaries use a vertical bar and low tilde for the same purpose.

Use of geometrical shapes

Many dictionaries use geometrical shapes as symbols, cf. the use of □ in sample 6 above. Most of these symbols directly correspond to an existing symbol or geometric shape, already encoded in Unicode. For example the shapes ● and ► are widely used, except that the size of ● usually correspond to the ‘medium circle’ not yet encode in Unicode, but proposed in document N2687r. In some cases, the white square has a deliberately heavy outline, which gives the symbol a different character. The use of ■ can be attested, but so can the use of a 90 degree rotated version of this symbol, which is not encoded in Unicode, as in the following two samples. Those samples also show the use of a symbol that is superficially similar to U+25EB □ WHITE SQUARE WITH VERTICAL BI-SECTING LINE, but deliberately emphasizes the dividing line, rather than the two rectangular halves of the square.

Sample 9

<p>COUNT STROKES OF SHADED PART</p> <p>Count the strokes of the shaded part to get the second part of the SKIP number. ▶ 相 <input type="checkbox"/> 1-4-</p> <p>REFERENCE: Appendix 2. How to Count Strokes</p>	<p>DETERMINE TOTAL STROKE-COUNT</p> <p>Determine the total stroke-count of your character to get the second part of the SKIP number. ▶ 下 <input checked="" type="checkbox"/> 4-3-</p> <p>REFERENCE: Appendix 2. How to Count Strokes</p>
<p>COUNT STROKES OF BLANK PART</p> <p>Count the strokes of the blank part to get the third part of the SKIP number. ▶ 相 <input type="checkbox"/> 1-4-5</p>	<p>IDENTIFY SOLID SUBPATTERN</p> <p>Determine to which of the four solid subpatterns your character belongs to get the third part of the SKIP number. Select from: <input type="checkbox"/> 1, <input type="checkbox"/> 2, <input type="checkbox"/> 3, or <input type="checkbox"/> 4. ▶ 下 <input checked="" type="checkbox"/> 4-3-1</p>

Sample 10

Here is a second sample from the same source, completing the set of symbols used.

speed with minimum effort. With the help of simple rules, each character is unambiguously classified under one of four easy-to-identify geometrical patterns: 1 left-right, 2 up-down, 3 enclosure, and 4 solid. Within each group the characters are further subdivided into progressively smaller subgroups until each is assigned its own position.

Table of Proposed Symbols

The table is divided into three sections corresponding to the sections in the preceding discussion

Code	Glyph	Suggested Name
26A2		HAMMER AND CHISEL <i>mining, working day(in timetables)</i>
26A3		ANCHOR <i>nautical term</i>
26A4		CROSSED SWORDS <i>military term, battleground (on maps), killed in action</i>
26A5		ROD OF AESCLEPIUS <i>medical term</i> This is not U+2624 CADUCEUS (see below). An inclined rendering of this symbol is common.
26A6		SCALES <i>legal term, jurisprudence</i>
26A7		ALEMBIC <i>chemical term, chemistry</i>
26A8		FLOWER <i>botanical term</i> This is distinct from <i>shamrock</i> (proposed in N2586R)
26A9		GEAR <i>technology, tools</i> An alternate glyph appears to be  for which unification with this symbol is proposed
26AA		STAFF OF HERMES <i>commercial</i> A herald's staff with a winged wheel. See discussion on caduceus below.
26AB		LOKOMOTIVE <i>railway</i>
26AC		HORN <i>postal term</i>
26AD		TRIANGLE WITH PLUMB LINE <i>architectural term</i>
26AE		BUILDING
26AF		COMPASS <i>mathematical term</i>
26B0		SPADE <i>agricultural term, agricultur</i>
26B1		LIVESTOCK
26B2		SWIMMING <i>sports</i>

Proposal — Symbols used in Dictionaries

26B3		RUNNING <i>sports</i>
205A		VERTICAL FOUR DOTS Used a.o. to indicate legal, but undesirable mid-word line break opportunities in dictionaries, glyph extends the whole height of the line. Allows a line break after (properties: lb=BA; gc=Po; eaw=N).
2B18		WARENZEICHEN <i>trademark (German)</i> Certain German dictionary usage instead of either ® or ™
205B		BOGEN Unlike U+203F ◡ UNDERTIE this symbol rests on top of the baseline and is quite narrow. It is used a.o. to indicate part of a term which will be marked by ... later in the entry. It is normally used at a syllable or other intra-word boundary, and therefore, unlike U+203F allows a word-break after (properties: lb=BA; gc=Po; eaw=N)
<007E, FE00>		LOW TILDE <i>repeated term</i> This tilde sits low on the baseline to accommodate the large circle for the circled variant. As it is not used in contrast to a midline symbol, it might be better to consider it as a variation rather than a character.
205C		LOW TILDE WITH CIRCLE <i>repeated term with difference in capitalization</i> The shape looks noticeably different from the one achieved by a combination of ~ and combining °: ◌̃ – instead it looks like combination of a lowered ~ and a degree sign °.
2B12		SQUARE WITH TOP HALF BLACK used contrastively to 25E7 SQUARE WITH LEFT HALF BLACK and U+25E8 SQUARE WITH RIGHT HALF BLACK
2B13		SQUARE WITH BOTTOM HALF BLACK This is proposed as an obvious completion
2B14		WHITE SQUARE WITH BLACK OUTLINE used contrastively to WHITE SQUARE
2B15		WHITE SQUARE WITH TOP BLACK LINE
2B16		WHITE SQUARE WITH BOTTOM BLACK LINE
2B17		SQUARED VERTICAL LINE At first glance this appears to be a legitimate glyph alternate for U+25EB ◻ WHITE SQUARE WITH VERTICAL BISECTING LINE, but the centerline here is deliberately emphasized.

Note on character properties

Unless otherwise noted all characters are of general category “So”, EAW “N”, Line-Breaking “AL”, and all other properties assigned appropriately. BidiMirrow=N and CC=0 for all characters.

On Caduceus and Rod of Asclepius

The Wikipedia contains two interesting articles on Caduceus and Rod Of Asclepius, here in excerpt:



A **caduceus** ([kerykeion](#) in [Greek](#)) staff with two snakes wrapped around it.

The caduceus is a symbol of commerce and is associated with the Greek god [Hermes](#). It was originally a herald's staff, sometimes with wings, with two white ribbons attached. The ribbons eventually evolved into snakes in the figure-eight shape.

In modern times it is often mistakenly used interchangeably with the [Rod of Asclepius](#) (particularly in the [United States](#)), although the two symbols have distinct and unrelated meanings.



The **Rod of Asclepius** is an [ancient Greek](#) symbol associated with [medicine](#). It consists of a serpent entwined around a staff. [Asclepius](#) was the god of healing in the ancient Greek pantheon. He was instructed in medicine by the [centaur Chiron](#).

The Rod of Asclepius symbolizes the healing arts by combining the serpent, which in shedding its skin is a symbol of rebirth and fertility, with the staff, a simple tool.

A similar symbol, [Nehushtan](#), is mentioned in the [Bible](#) as having been used for healing of snakebite.

The Rod of Asclepius is frequently confused with the [caduceus](#), which is a symbol of commerce associated with the god [Hermes](#).

For more information on the distinction between these two symbols see: Blayney, Keith. *The Caduceus vs the Staff of Asclepius*, found at <http://drblayney.com/Asclepius.html>, from which the picture of the rod is taken.

About the use of oblique double hyphen in dictionaries

A lot has been written recently about the use of Fraktur hyphen (“oblique double hyphen”). See for example document N2639. For completeness, here is a citation of the use of this symbol in a dictionary context.

Webster's Ninth New Collegiate Dictionary, ISBN 0-87779-508-8
Copyright 1989 by Merriam-Webster Inc.

Explanatory Notes, End-of-line division (p11)

Proposal — Symbols used in Dictionaries

A double hyphen at the end of a line in this dictionary (as in the definition at indexation) stands for a hyphen that belongs at that point in a hyphenated word and that is retained when the word is written as a unit on one line.

And the definition of indexation reads (line breaks indicated by |):

a system of economic control |
in which certain variables (as wages and interest) are tied to a cost-of<double-> |
living index so that both rise or fall...

Graphically, the double hyphen is slanted upward.

References

Note that a large number of additional dictionaries were researched, but since they either use no symbols, or other symbols already encoded in Unicode, or simply duplicate the set of symbols proposed here, they have not been cited here.

Akdikmen, Resuhi. 1992. Langenscheidt's pocket Turkish dictionary: Turkish-English English-Turkish. Berlin & München: Langenscheidt. ISBN 0-245-60405-7

Árni Böðvarsson. 1992. Íslensk orðabók. Reykjavík: Mál og menning. ISBN 9979-3-0446-4

Engelsk-svensk ordbok, skoluplaga, Esselte Herzogs, Nacka 1976, ISBN 91-24-19070-2

Der Große Duden, Band 1, *Rechtschreibung*, Bibliographisches Institut, 1968, Mannheim, Zürich.

Langenscheidt's New College German Dictionary, German-English
© 1973 Langenscheidt KG, Berlin and Munich, ISBN:0-88729-018-3

NTC's New Japanese-English Character Dictionary, Jack Halpern ed., Tokyo 1993.

The Random House College Dictionary, revised edition, Random House, New York 1975, ISBN 0-394-436008-8

Webster's New World Dictionary, Second College Edition, Williams Collins, Cleveland 1979, ISBN 0-529-05234-1

Webster's Ninth New Collegiate Dictionary, ISBN 0-87779-508-8
Copyright 1989 by Merriam-Webster Inc.

Web sites referenced

<http://www.wikipedia.org/wiki/Caduceus>

http://www.wikipedia.org/wiki/Rod_of_Asclepius

**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 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.

A. Administrative

1. Title:	<u>Proposal — Symbols used in Dictionaries</u>
2. Requester's name:	_____ Asmus Freytag _____
3. Requester type (Member body/Liaison/Individual contribution):	_____ Individual Contribution _____
4. Submission date:	_____ 2003-10-15 _____
5. Requester's reference (if applicable):	_____
6. (Choose one of the following:)	
This is a complete proposal:	_____ X _____
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:	_____ X _____
Name of the existing block:	_____ Various _____
2. Number of characters in proposal:	_____ 29 _____
3. Proposed category (see section II, Character Categories):	_____
4. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000):	_____ 1 _____
Is a rationale provided for the choice?	_____ No _____
If Yes, reference:	_____
5. Is a repertoire including character names provided?	_____
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?	_____ proposer _____
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:	_____ _____ _____
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)?	_____ _____
9. Additional Information:	
Submitters are invited to provide any additional information about Properties of the proposed Character(s) ...	

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	_____ No _____
If YES explain _____	

¹ Form number: N2352-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09)

Proposal — Symbols used in Dictionaries

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? Irish NB, Michael Everson
 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? _____
 Reference: Publishers and users of common, widely available dictionaries of several languages

4. The context of use for the proposed characters (type of use; common or rare) _____ common _____
 Reference: In active use, see references section

5. Are the proposed characters in current use by the user community? _____ yes _____
 If YES, where? Reference: see references section

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?
 _____ 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 an existing character?
 _____ yes _____
 If YES, is a rationale for its inclusion provided?
 _____ yes _____
 If YES, reference: see attached

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: _____