L2/07-093

ISO/IEC JTC 1/SC 35 N 113x

DATE: 2007-02-16

ISO/IEC JTC 1/SC 35

User Interfaces

Secretariat: AFNOR

DOC TYPE: WORKING DRAFT

TITLE: ISO/IEC 15897 – Information technology – Procedures for registration of cultural elements

SOURCE: Keld Simonsen

PROJECT: ISO/IEC 15897

DISTRIBUTION:P&O Members

MEDIUM:E

NO. OF PAGES: 54

Address Secretariat ISO/IEC JTC 1/SC 35 AFNOR - Philippe Magnabosco
11 rue Francis de Pressensé 93571 - La Plaine Cedex Saint-Denis - France
Telephone: +33 1 41 62 85 02; Facsimile: 33 1 49 17 93 00;
e-mail: philippe.magnabosco@afnor.org

Reference number of working document: **ISO/IEC JTC1/SC35 N113x**

Date: 2007-01-28

Reference number of document: **ISO/IEC 15897:200?(E)**

Committee identification: ISO/IEC JTC1/SC35

Secretariat: AFNOR

Information technology – Procedures for registration of cultural elements

Technologies de l'information - Procédures pour l'enregistement des éléments culturels

Contents Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

This International Standard was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee 22 *Programming languages, their environments and system software interfaces*. The first edition of this International Standard was the result of a fast-track ballot of the CEN specification ENV 12005:1996 named "Procedures for European registration of cultural elements".

This International Standard does not cancel or replace any other standard.

This International Standard has a number of changes to ISO/IEC 15897:1999 and CEN ENV 12005:1996. The changes are described in annex H.

This International Standard registers amongst other items Narrative Cultural Specifications and repertoiremaps, POSIX Locales and POSIX Charmaps as defined in ISO/IEC 9945 "POSIX", and other machine parsable cultural specifications such as ISO/IEC TR 14652 FDCC-sets, charmaps and repertoiremaps, and cultural specifications in SGML or XML.

The annexes A, C and E of this International Standard are normative, and the annexes B, D, F, G, and H are for information only.

Introduction

Cultural differences throughout the world make it necessary to adapt IT-equipment to each local culture. Standard methods, being developed by ISO/IEC JTC1/SC35, make such adaptation easier. Registering the local conventions in a globally available registry make it still easier. CEN/TC304 was the first committee to start work on such a registry and the European prestandard ENV 12005:1996 was the result. The first edition of this International Standard was the result of the fast-track of DS/ENV 12005, a Danish Standards Association standards publication equivalent to CEN ENV 12005. This edition of the International Standard adds support for registering specifications meant for machine processing such as ISO/IEC TR 14652 specifications, SGML and XML It enlarges the group of organizations that may be Sponsoring Authorities, and an effort has been made to align it with the registration procedures of ISO/IEC 2375.

The standard sets out the procedures for registering cultural elements, both as narrative text and in a more formal manner, using the techniques of ISO/IEC 9945 "POSIX", and other machine processable formats such as those specified in ISO/IEC TR 14652 "Specification method for cultural conventions", SGML or XML. The registration will be free-of-charge and the registered cultural elements will also be freely available on the Internet, see clause 6.3. This will make information on cultural conventions freely and easily available to producers in the IT market. Some of these specifications may even be applied without any change by downloading the formatted specifications and processing them, for example by POSIX compliant software.

Information technology - Procedures for registration of cultural elements

1 Scope

This International Standard specifies the information that may appear in a cultural specification and defines the procedures for registering such specifications. The cultural specifications may include freeform Narrative Cultural Specifications and Repertoiremaps as described in this International Standard, POSIX Locales and Charmaps conforming to ISO/IEC 9945, and other machine-parsable specifications such as FDCC-sets, repertoiremaps and charmaps following the recommendations of ISO/IEC TR 14652, and cultural specifications formatted using SGML or XML.The registry is in printed and electronic form.

Each cultural convention registration has unique identifiers in a particular standard format defined below in clause 15. A numeric identifier and a token identifier are assigned to each registered cultural convention specification, POSIX Locale, POSIX Charmap and Repertoiremap, FDCC-set and ISO/IEC TR 14652 Charmap, and other machine processable descriptions of cultural conventions . These identifiers are for unique identification of the cultural specification, and intended to be used with POSIX locale handling mechanisms and possibly other locale and charmap usage, such as in programming languages, database handling and communication protocols and for identification and specification by human operators.

The field of application of this International Standard is to provide reference for implementers, procurers, users, and standardization organizations, to determine cultural requirements in a given cultural environment. Registered items using certain POSIX formal specification methods can also be used by POSIX-conformant Operating Systems and other software capable of using such specifications.

> NOTE The title of this International Standard contains the term "cultural elements". The preferred term in this International Standard is "cultural conventions", but the term "cultural elements" is retained in the title for consistency with earlier versions of this International Standard.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 639-1:2002, *Code for the representation of names of languages- Part 1: Alpha-2 code.*

ISO 639-2:1998, *Code for the representation of names of languages - Part 2: Alpha-3 code*.

ISO/IEC 646:1991, *Information technology - ISO 7-bit coded character set for information interchange*.

ISO/IEC 2022:1994, *Information technology - Character code structure and extension techniques*.

ISO 3166 (all parts), *Codes for the representation of names of countries and their subdivisions*.

ISO 4217:2001, *Codes for the representation of currencies and funds*.

ISO 8601, *Data elements and interchange formats - Information interchange - Representation of dates and times*.

ISO/IEC 8824 (all parts), *Information technology -- Abstract Syntax Notation One (ASN.1)*

ISO/IEC 8825 (all parts), *Information technology - ASN.1 encoding rules.*

ISO/IEC 9945-1:2003, *Information technology - Portable Operating System Interface (POSIX) - Part 1: Base Definitions.*

ISO/IEC 9945-2:2003, *Information technology - Portable Operating System Interface (POSIX) - Part 2: System Interfaces*.

ISO/IEC 9945-3:2003, *Information technology - Portable Operating System Interface (POSIX) - Part 3: Shell and Utilities*.

ISO/IEC 9945-4:2003, *Information technology - Portable Operating System Interface (POSIX) - Part 4: Rationale*.

ISO/IEC 10646, *Information technology - Universal Multiple-Octet Coded Character Set (UCS).*

ISO 15924, *Information and documentation -- Codes for the representation of names of scripts.*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1
locale
The definition of the subset of a user's environment that depends on language and cultural conventions. See the clause 7 of the ISO/IEC 9945-1:2002 POSIX standard for a specification of the locale file format.

3.2
FDCC-set
A set of Formal Definitions of Cultural Conventions. The definition of the subset of a user's information technology environment that depends on language, territory, or other cultural customs. See clause 4 of ISO/IEC TR 14652 for a specification of the FDCC-set format.

3.3
charmap

A text file describing a coded character set. See the clause 6.4 of the ISO/IEC 9945-1:2002 POSIX standard for a description of the POSIX Charmap file format.

3.4
text file
A human-readable file that contains characters organized into one or more lines.

3.5
cultural convention
cultural element
A data item for computer use that may vary depending on language, territory, or other cultural circumstances.

3.6
cultural specification
Either a Narrative Cultural Specification, a POSIX Locale, an FDCC-set, a POSIX Charmap, a Repertoiremap, or other machine-processable description of cultural conventions such as ISO/IEC TR 14652 FDCC-sets, Charmaps or Repertoiremaps, or cultural specifications in SGML or XML.

3.7
narrative cultural specification
A narrative description of culturally dependent information pertaining to information technology. Such information may be useful when designing computer systems and software. See clauses 10, 11 and 12.

3.8
repertoiremap
A definition of a mapping between character names and characters for the repertoire of characters used in a Locale, further described in clause 13.

3.9
profile(of a standard)
a set of specifications of values of parameters in the standard, selections of optional items of the standard, or the recommendations concerning implementation-related matters of the standard.

3.10
token identifier
a string uniquely identifying the Cultural Specification, constructed from various key attributes of the specification, further described in clause 15.

4 International Register

4.1 Content

The International Register of Cultural Specifications (IR-15897) shall consist of two parts: a set of registrations of cultural specifications, and indices to the registrations.

4.2 Format

The International Register (IR) shall be available in electronic format through the Internet, and optionally on other electronic media. It may also be made available on paper.

Directions for access to the International Register on the Internet are given in clause 6.3.

4.3 Indices to the registrations

The International Register shall contain indices to the registrations by

- the registration number
- the token identifiers
- the type of cultural specification(s) in a registration
- the culture to which a registration applies
- the sponsoring authority of a registration
- other indices as deemed appropriate by the Registration Authority, or as requested by the subcommittee with administrative responsibility for this standard (see clause 5, ISO/IEC Supervisory Body).

4.4 Identification of an approved registration

Each approved registration must have a unique registration number and one or more unique token identifiers.

4.4.1 Structure of the identifiers

The structure of a registration number is one or more digits, without leading zeroes. The structure of a token identifier is given in clause 15.

4.4.2 Reference to an approved registration

A reference to an approved registration should be made by using one of its unique token identifiers, a part of its token identifier with for example part or all of the version number excluded as described in clause 15, or the prefix ISO-IR-15897 followed by a HYPHEN and the registration number, or one of the mechanisms listed in annex C. Examples of token identifiers are listed in clause 14. An example of using the ISO-IR-15897 prefix with the registration number 37 is ISO-IR-15897-37.

4.5 No modification nor deletion of registrations

The contents of an individual registration shall never be changed or deleted once the application for registration has been approved (except for name additions). Even the correction of editorial errors will make a new registration necessary. This is contrasted with a standard

which must be reviewed and/or revised periodically. If an existing entry in the registry is based on a standard that subsequently is revised, the existing registry entry is not changed. If a Sponsoring Authority desires recognition of such a revision, a new registration shall be made by following the normal procedures. Withdrawal of part of an existing registration is prohibited by clause 21.3.1.

A registration has a version number as part of its token identifier, defined in clause 15. By registering a Cultural Specification with a higher version number, a specification may be updated, and applications referring to the registration without the precise version will get the latest update, while applications referring the precise version, or the registration number, will get the older specification. The rules for the versioning mechanism is specified in clause 15.

5 ISO/IEC Supervisory body

The ISO/IEC JTC1 subcommittee concerned with internationalization has administrative responsibility for this standard and the content of the register.

> NOTE At the time of publication, subcommittee ISO/IEC JTC1/SC35, Programming languages, their environments and system software interfaces, has this responsibility.

6 Registration Authority

6.1 Appointment

6.1.1 The Registration Authority (RA) shall be an organization nominated by the ISO/IEC supervisory body for this International Standard and appointed by the ISO and IEC councils to act as the Registration Authority for this International Standard.

6.1.2 The Registration Authority shall be an organization actively participating in the work of the ISO/IEC supervisory body for this International Standard.

6.2 Responsibilities

6.2.1 The Registration Authority shall maintain the International Register of Cultural Specifications and assign their numeric and token identifiers.

6.2.2 The Registration Authority shall manage the execution of the registration procedure, including processing of:

- applications for registration (as specified in clauses 16 and 17);
- appeals (as specified in clause 18);
- corrections and revisions to existing registrations (as specified in clauses 19 and 20);
- withdrawal of existing registrations (as specified in clause 21)

The exact responsibilities of the Registration Authority are described in detail in clauses 16 and 17.

6.2.3 The contents of this register shall be available to ISO/IEC JTC 1 members and to the general public, both in printed and electronic form. In particular, the contents of the register shall be made available over the Internet.

6.2.4 One or more technical representatives of the Registration Authority may attend the meetings of the subcommittee with administrative responsibility for this standard and of appropriate working group(s) reporting to this subcommittee if required.

6.3 Identity

The identity of the designated Registration Authority and where the Registration Authority has published the 15897 Register on the Internet is available at:

http://www.iso.org/mara (In English)
http://www.iso.org/mara-fr (In French)

NOTE The initial Registration Authority was designated to be:

The Danish UNIX-systems User Group (DKUUG)
Fruebjergvej 3
DK-2100 København Ø
Danmark
Fax: +45 3920-8948
email: culturalregister@dkuug.dk

The data of the cultural register was initially available at:

http://www.dkuug.dk/cultreg/

7 Sponsoring Authority

7.1 Identity

A Sponsoring Authority is an organization that may submit applications for registration of cultural specifications to the Registration Authority. Applications for registration of Cultural Specifications may be made by the following bodies:

a) Any Member Body of ISO/IEC JTC1, for applications limited to the territory or territories for which they have authority;

b) Any National Member, Associate, or Affiliate of CEN, for applications limited to the territory or territories for which they have authority;

c) ISO/IEC JTC 1 and its Subcommittees and Working Groups, for any applications.

7.2 Responsibilities

The responsibilities of a Sponsoring Authority shall be:

a) to receive applications concerning Cultural Specifications from a Source of Information (see clause 8) operating in the area over which the Sponsoring Authority has jurisdiction;

b) to prepare applications for the registration of the Cultural Specifications according to the procedures and specifications in clause 15;

c) to assure that free distribution of the Cultural Specification is permitted without charge;

d) if any material in an application is under copyright, to obtain copyright clearance for the copyrighted material from the copyright holder, including permission to redistribute the copyrighted material as part of the Cultural Specification without restriction and without charge by the copyright holder;

e) Sponsoring Authorities may submit any applications for registration of the types Charmaps, Repertoiremaps, and Machine-parsable character set descriptions to support their other Cultural Specifications;

f) to effect such rationalization or coordination of all applications under consideration as the Sponsoring Authority may desire;

g) to submit applications for the registration of Cultural Specifications to the Registration Authority;

h) to announce the result of the registration procedure within its respective country, region, or organizations.

8 Source of Information

8.1 Identity

The Source of Information is the organization or individual responsible for the Cultural Specification.

8.2 Responsibilities

The responsibilities of a Source of Information shall be:

a) to prepare a Cultural Specification;

b) if any material in a Cultural Specification is under copyright, to assure that free

distribution of the Cultural Specification is permitted;

c) to submit the Cultural Specification to an appropriate Sponsoring Authority.

9 The Registration Authority's Joint Advisory Committee

9.1 Membership.

The Joint Advisory Committee (RA-JAC) shall consist of a representative of the Registration Authority and four other members who shall be representatives from the national member bodies on the subcommittee concerned with the maintenance of this standard or representatives from organizations with a liaison membership to this subcommittee.

The chair of the RA-JAC shall be the representative of the Registration Authority.

9.2 Appointment.

The subcommittee responsible for maintaining this standard shall appoint the members of the RA-JAC, except for the RA representative, which is appointed by the RA.

The subcommittee responsible for maintaining this standard shall appoint or confirm the members of the RA-JAC at its plenary meetings, except for the representative of the RA.

9.3 Responsibilities.

The responsibilities of the RA-JAC shall be as follows:

a) to determine whether an application for registration meets the technical requirements of clause 14;

b) to provide expert technical advice on comments if requested by the Registration Authority;

c) to consider and vote on appeals received by the Registration Authority;

d) to act as a mediator between the Registration Authority and the appealing party, or parties.

e) In addition, the RA-JAC may add comments to a registration.

10 Types and relationships of Cultural Specifications

10.1 Types of Cultural Specifications

A number of types of Cultural Specifications can be registered according to this International

Standard:

1. Narrative Cultural Specification
2. POSIX Locale
3. POSIX Charmap
4. Repertoiremap
5. Machine-parsable cultural specification
6. Machine-parsable coded character set specification

Type 1 is for Narrative Cultural Specifications, further specified in clause 11 and 12.

Types 2 and 3 are for POSIX specification of cultural conventions defined in ISO/IEC 9945.

Type 4 is for Repertoiremaps. Clause 13 defines the format of a repertoiremap included in an application for cultural registration.

> NOTE As far as Repertoiremaps according to ISO/IEC TR 14652 are also in accordance with clause 13, these can also be registered as Type 4.

Types 5 and 6 are for specification of cultural conventions in a machine-parsable format, such as specified in ISO/IEC TR 14652, XML or SGML table formats. Any format is allowed as long as it is machine parsable and adheres to the following rules: It is a TR 14652 FDCC-set, a TR 14652 charmap, or the first line of the file identifies the file format.

10.2 Relations between registration types

Registration types are related as follows:

10.2.1 The Narrative Cultural Specification

The Narrative Cultural Specification specifies cultural conventions in narrative form in any of the official ISO/IEC JTC 1 languages English, French and/or Russian, and it may give equivalent specifications in other languages. It may thus address issues which have not yet been codified by formal methods for specifications of cultural conventions. If parts of a Narrative Cultural Specification have been specified also in POSIX Locale or Charmap format, this Locale or Charmap should be referenced in the specification.

10.2.2 POSIX Locale

The POSIX Locale shall specify appropriate aspects of a Narrative Cultural Specification in formal POSIX syntax. The POSIX locale shall refer to either one or more POSIX charmaps it can use, or to a Repertoiremap.

When a POSIX locale is submitted, it should be accompanied by a corresponding Narrative Cultural Specification. The Narrative Cultural Narrative Cultural Specification and the POSIX Locale shall not be in contradiction.

10.2.3 POSIX Charmap

The POSIX charmap shall specify aspects of a Narrative Cultural Specification or a POSIX Locale that relate to coded character sets. A POSIX Charmap may refer to POSIX locales, Narrative Cultural specifications, or Repertoiremaps that it uses, but such references are not required.

10.2.4 Repertoiremap

A Repertoiremap is used as a tool to enable a POSIX Locale or a Narrative Cultural Specification to be independent of coded character sets, and to remove the requirement for POSIX Charmaps when registering a POSIX locale. It need not refer to other Cultural Specifications.

10.2.5 Other machine-parsable Cultural Specifications

Some aspects of a Narrative Cultural Specification may be specified in formal syntax as a TR 14652 FDCC-set, or other machine-parsable cultural specification, which may refer to the corresponding Narrative Cultural Specification. A TR 14652 FDCC-set may refer to the Repertoiremap being used, and should also list one or more Charmaps that can be used.

10.2.6 Other machine parsable character set specifications

In the case of a ISO/IEC TR 14652 Charmap, or other machine-parsable character set descriptions it shall specify aspects of a Narrative Cultural Specification or an FDCC-set that relate to coded character sets. A Charmap may refer to the Repertoiremap being used, and may refer to the FDCC-set or the Narrative Cultural Specifications using the Charmap.

> NOTE It is the intention to allow more formal specification methods in future revisions of this International Standard when they become standardized methods; for the time being these specifications can be registered as type 1, 5 or 6.

11 Format of a Narrative Cultural Specification

The format of a Narrative Cultural Specification shall contain clauses describing the following cultural conventions, which may also be described in a POSIX Locale, or an FDCC-set, or other machine-parsable cultural specification:

1 Alphanumeric deterministic ordering
2 Classification of characters
3 Numeric formatting
4 Monetary formatting
5 Date and time conventions
6 Affirmative and negative answers

The Narrative Cultural Specification may also include other culturally dependent

information, limited to the following list:

7 National or cultural Information Technology terminology
8 National or cultural profiles of standards
9 Character set considerations
10 Sorting and searching rules
11 Transformation of characters
12 Character properties
13 Use of special characters
14 Character rendition
15 Character inputting
16 Personal names rules
17 Inflection
18 Hyphenation
19 Spelling
20 Numbering, ordinals and measuring systems
21 Monetary amounts
22 Date and time
23 Coding of national entities
24 Telephone numbers
25 Mail addresses
26 Identification of persons and organizations
27 Electronic mail addresses
28 Payment account numbers
29 Keyboard layout
30 Man-machine dialogue
31 Paper formats
32 Typographical conventions

The list of clauses may be expanded in future revisions of this standard.

Clauses 10, 12, 20, 21, 22 and 30 are for description of cultural aspects in excess of what can be described in the corresponding POSIX clauses 1, 2, 3, 4, 5 and 6. In clause 9 it is possible to give further information on characters classified in clause 2.

NOTE Further information about the categories, along with specific examples illustrating their use may be found in clause 11, in annex D and in the Nordic Cultural Requirements on Information Technology (Summary Report).

Each clause shall begin on a new line after at least one blank line, and each clause shall be introduced by the string "Clause ", followed by the decimal clause number for the issue as listed above, then a colon and a space, and then the title of the clause, using the titles above (examples are given in annex D).

The body of the clause shall follow on the succeeding lines. A reference to a clause within the specification shall consist of the string "=> Clause " followed by the clause number. A reference to another specification shall consist of the string "=> Spec. " followed by the

registration number of the specification and, optionally, the string " Clause " and a clause number.

12 Contents of a Narrative Cultural Specification (informative)

The contents of a Narrative Cultural Specification are described in some detail in this clause. The specification builds on information from the POSIX Base Definitions standard (ISO/IEC 9945-1:2002) and the Nordic Cultural Requirements on Information Technology (Summary Report). Clauses 1 to 6 are related to POSIX. Clause 7 to 32 are to provide information, which is not presently expressible in POSIX notation. An example of a Narrative Cultural Specification is given in annex D.

> NOTE The numbering of the clauses are somewhat arbitrary, and if a Sponsoring Authority deems that some information in a subsequent clause is helpful for the understanding of a clause, then the clause with the additional information may be referenced via a ”see => Clause xx” specification.

12.1 Mandatory clauses

The format of a Narrative Cultural Specification shall contain the clauses (numbered 1-6) specified below. These clauses are POSIX categories. The Narrative Cultural Specification should be accompanied by a corresponding POSIX Locale specification. The information given in these clauses of the Narrative Cultural Specification may also be described in an FDCC-set, or other machine parsable cultural specification:

Clause 1: Alphanumeric deterministic ordering

In this clause the specification of a national standard for ordering should be listed. If there are more standards, or options for a standard, there should be one POSIX specification for each of the standards or options. An international standard such as ISO/IEC 14651 or a European Multilingual Ordering standard such as CEN CR 14400, could be referenced, and possible deviations, if any, could be described. Issues to cover may include whether there are letters that sort differently from common use in other languages, whether capital letters sort before small letters, or whether there is a specific ordering of diacritics. Further, this section may describe the ordering of scripts, and sorting levels -- that is, if there are cases when characters sort equally at first, but then may sort differently at other levels. Does the language require reordering of some characters before collation weighting (for example Thai)? Does the language sort on a syllabic basis, rather than merely letter-by-letter (for example Burmese)? Does the language make use of ideographs, and if so, how are they handled with respect to other characters? If aspects of the ordering for the language extend beyond what a POSIX specification can handle, then details can be described in Clause 10.

The specification is aimed at deterministic sorting, that is that if two strings of characters are compared, the result will always be the same, regardless for example of the encoding used for comparison, and the sequence of the two strings.

The clause is also intended for description of ordering of non-alphabetic scripts, such as ideographic scripts or syllable-oriented scripts, and for ordering of multiple scripts.

This is a POSIX category.

Clause 2: Classification of characters

The POSIX standard allows descriptions of what are alphabetic characters, capital and small letters, digits, hexadecimal digits, punctuation characters, spaces, graphical characters and control characters. Classification of

characters that cannot be handled in standard POSIX may be specified in Clause 12. This is a POSIX category.

Clause 3: Numeric formatting

This clause describes how numbers are formatted (for input and output), including the format of the decimal point and the thousands separator. Special considerations for how numbers should be formatted in narrative text and other numeric formatting that cannot be handled by the POSIX standard may be specified in Clause 20. This is a POSIX category.

Clause 4: Monetary formatting

This clause describes formatting rules for monetary amounts, as well as local and international currency symbols according to ISO 4217, as well as the relation between the amount, a sign and the currency symbol. Monetary formatting that cannot be handled by the POSIX standard may be specified in Clause 21. This is a POSIX category.

Clause 5: Date and time conventions

Various names for days and months are given, together with formats for writing date and time. Things to consider are: do day and month names start with a capital letter or a small letter? Are there well recognized abbreviations for the day and month names? Is ISO 8601 formatting widespread? As the date formats are for use in POSIX, for example when listing files, consideration should be given to possible POSIX conventions in the culture, and the abbreviated date formats should be of constant length for them to be used in lists. The long formats should be usable in narrative text such as letters. Handling of date formatting that cannot be handled by the POSIX standard may be specified in Clause 22. This is a POSIX category.

Clause 6: Affirmative and negative answers

In this clause the short notation for "yes" and "no" answers in the language can be specified. If the culture has strong relations to several languages, for example in a multilingual country, it should be permitted to answer in any of the languages. As English is widely used in many cultures, allowing responses in the English language should be considered. This is a POSIX category.

12.2 Optional clauses

The Narrative Cultural Specification may also include other culturally dependent information, specified in clauses 7-32. These clauses are not directly related to POSIX Locales:

Clause 7: National or cultural Information Technology terminology

In this clause terminology for a language or culture can be listed, for example a translation of ISO terminology for Information Technologies.

Clause 8: National or cultural profiles of standards

In this clause profiles of standards can be listed, for example, OSI national profiles, or profiles of the POSIX standards. See the POSIX ISO/IEC 9945-2:1993 standard Annex G for an example.

Clause 9: Character set considerations

This clause describes how characters are used in the culture, for example:
- which letters constitutes the alphabet, that is the necessary set of characters to write a particular language,
- which characters are used to give further precision in the language, for example allowed in the orthography of the language, like accented letters to denote stress,

- which characters are usually used in newspapers and books for writing of names and places, for example of foreign origin,
- which characters are used for historic writing of the language,
- and which characters are used for other purposes, such as characters ordinarily used in school for scientific or foreign language education.

This clause may also be used to specify which coded character sets are common in the culture and what coded character sets are recommended. Also further descriptions of coded character sets may be described; it is also possible to document these in the form of a POSIX Charmap registration.

Clause 10: Sorting and searching rules

This clause is for specifying sorting and searching rules that cannot be specified with POSIX specifications, such as ISO/IEC 14651 specifications, non-deterministic ordering, pre-handling and post-handling of records, such as how to split a record into sorting fields, and rules for common words like a, the which may be ignored when comparing or searching. Also sound based matching rules may be described here. What can be accomplished deterministically with POSIX should be described in Clause 1.

Clause 11: Transformation of characters

This clause describes transliterations and transformations of characters, for example transliteration rules between Latin, Greek and Cyrillic, or fallback notation for some frequent letters. Also this is the place to write about standards in the culture for character conversion. Examples of transliteration and fallback specifications in a syntax described in an earlier draft of ISO/IEC TR 14652 may be found in the bibliographic reference 4.

Clause 12: Character properties

Here additional considerations further than those given in clause 2 can be given, for example how small letters without a direct capital counterpart may be capitalized, or special capitalization rules.

Clause 13: Use of special characters

This clause describes the use of special characters, that is characters that are not letters, ideographs or syllables or other characters used to write words of a language, or that are not digits, nor control characters (control characters are a feature of information technology, and not special characters used in a writing system) . Examples of special characters are quotation marks, abbreviation marks, and punctuation marks. Information about culture-specific signs, and what should be avoided, may also be included in Clause 13; for example the preferred abbreviation for "number", the signs used to indicate a paragraph and a section in text, and the preferred sign used to indicate division. Spacing rules and the relation between different punctuation signs are also relevant here.

Clause 14: Character rendition

Special considerations about rendition such as what alternatives may be considered adequate, and acceptable glyphs, may be described in this clause.

Clause 15: Character inputting

A keyboard seldom has separate keys for all the characters needed. This clause is intended for description of keyboard inputting rules and other input methods.

Clause 16: Personal names rules

Personal naming differs from culture to culture, for example what is considered the family name, how titles are used, in which order the family name and given name come, and whether given names or initial are used. Also the rules for children inheriting their fathers' and mothers' family name, and what happens for married couples may be described here.

Clause 17: Inflection

Languages vary much with respect to inflection, that is different forms of a word depending of the context. In this clause the rules can be described or referenced.

Clause 18: Hyphenation

Hyphenation rules and lists may be described here, and also references to the hyphenation specifications or lists for a language may be done here.

Clause 19: Spelling

This clause is for specification of spelling rules and spelling lists, or reference to orthographic documentation.

Clause 20: Numbering, ordinals and measuring systems

This clause describes further aspects of number handling in excess of Clause 3, such as additional characters used for digits, how ordinals are written, and measurement systems (normally this is the ISO SI system). Use of decimal point and thousands grouping separator shall be described in clause 3.

Clause 21: Monetary amounts

Here further considerations to clause 4 can be described, such as old currencies, and when there be shifts of currencies either in the past or in the future, and possible conversion rates, and how to write numbers and do spelling of amounts on cheques.

Clause 22: Date and time

This clause is for date and time conventions which are not accommodated by clause 5, that is, other ways in which time may be expressed beyond POSIX date and time conventions. Such conventions should preferably be augmented by a description of the field or fields where the convention is used.

> NOTE Date and time conventions that are not specified by POSIX include time zone names, daylight savings rules, week numbering, national holidays, and colloquial expressions. Examples of colloquial expressions are the British "half seven," which means 07:30 or 19:30, and the German "halb sieben" (literally, "half seven") which means 06:30 or 18:30. Fields of use include narrative text, letters, and legal documents such as contracts.

Clause 23: Coding of national entities

This clause describes the coding for different entities, such as postal codes, administrative codes for local government, police districts, abbreviations for cities or provinces, and time zone names relating to different parts of the culture.

Also specifications should be given for identification of the whole culture, for example ISO country codes for a nation.

Clause 24: Telephone numbers

The formatting of telephone numbers, nationally and internationally.

Clause 25: Mail addresses

The formatting of postal addresses, where to put the title of the addressee, the street number and the postal code, what are the names of the floors, and other conventions used.

Clause 26: Identification of persons and organizations

A culture may have numbering schemes for persons and organizations, for example social security numbers, and general tax numbers for companies, together with registries for different organization forms such as limited companies and associations. This clause may be used to describe such numbering systems.

Clause 27: Electronic mail addresses

Cultural conventions for Internet and X.400 electronic addresses etc. may be described here.

Clause 28: Payment account numbers

Cultural conventions for bank account numbers can be described here.

Clause 29: Keyboard layout

This clause describes the conventions for keyboard layout.

Clause 30: Man-machine dialogue

Considerations for how to localize products may be described here.

Clause 31: Paper formats

This clause describes what the conventions are for paper size (normally ISO standards) and the use of window envelopes, etc. Also how punched holes are placed in paper may be relevant here.

Clause 32: Typographical conventions

This clause may be used to describe how layout is done, for example how to layout a business letter, an invoice, or a fax. Use of special characters, for example quotation marks, should be described in clause 13.

13 Format of a Repertoiremap

POSIX Locale, FDCC-set and Charmap sources shall be specified in a way that is independent of coded character sets, using character names. Relation between the character names and characters shall be specified via a Repertoiremap table, defined with a line for each character giving the character name and the ISO/IEC 10646 short character ID in the form of Uxxxx or Uxxxxxxxx, and optionally the long ISO/IEC 10646 character name. The character name and the ISO/IEC 10646 canonical encoding shall each be surrounded by angle brackets <>, and the fields shall be separated by one or more spaces or tabs on a line. If a right angle bracket or an escape character is used within a name, it shall be preceded by the escape character. An escape character can be used to escape the syntactical meaning of the following character, so that the following character is included literally. The default escape character for a Repertoiremap is the reverse solidus (\). The escape character for a Repertoiremap can be redefined from the default reverse solidus (\) with the first line of a Repertoiremap containing the string "escape_char" followed by one or more spaces or tabs and then the escape character.

> NOTE Character names as specified in ISO/IEC 9945-2:1993 annex G may be used, many of the charmap registrations do so.

14 Rules for Cultural Specifications

The following rules apply when registering a Cultural Specification:

14.1 An application for registration of a Cultural Specification shall be submitted as a Text File. A Narrative Cultural Specification may alternatively be submitted on paper, preferably A4, or one of the approved document formats of ISO/IEC JTC 1, as noted in the JTC 1 Directives, Annex H.4.

14.2 Requirements on formats of different types of Cultural Specifications

14.2.1 Format of a Narrative Cultural Specification

If a Narrative Cultural Specification is included in a registration, it shall follow the format described in clause 11.

14.2.2 Format of a POSIX Locale or Charmap

The format of the POSIX Locale and Charmap sources shall be conformant to ISO/IEC 9945. Locales which are otherwise conformant to ISO/IEC 9945, but which make use of the extension technique defined in annex E, are also permissible for registration.

14.2.3 Format of a Repertoiremap

If a Repertoiremap is included in a registration, it shall follow the format described in clause 13.

14.3 The POSIX Locale or FDCC-set shall define all standard categories. Individual categories may be copied from another Locale or FDCC-set. See Annex G for examples.

14.4 The coded character set of ISO/IEC 646 International Reference Version shall be used to represent text for the submitted files. For enhanced network portability it is recommended that only the invariant part of ISO/IEC 646, which contains 83 graphical characters (including space), be used. Comments shall be given in the English language, and equivalent comments may also be given in other languages. If characters outside ISO/IEC 646 International Reference Version are needed, character names defined in a repertoiremap formatted according to clause 13 shall be used.

14.5 The sources shall be delivered electronically, either via electronic mail or on physical storage media to the Registration Authority. Narrative Cultural Specifications may alternately be delivered on paper.

14.6 A written application shall accompany the Cultural Specification and be signed by authorized personnel on behalf of the contributing organization. It shall give permission to freely distribute the Cultural Specification.

14.7 The written Cultural Specification application shall contain information on the following items:

1. Cultural Specification type number (as in clause 10.1 above)
2. Organization name of Sponsoring Authority
3. Organization postal address
4. Name of contact person
5. Electronic mail address of the organization, or contact person
6. Telephone number for the organization, in international format.
7. Fax number for the organization, in international format.

For Types 1, 2, and 5, Narrative Cultural Specifications, POSIX Locales, and Machine-parsable cultural specifications:

8. Natural language, as specified in ISO 639-1, or ISO 639-2 terminology codes if an ISO 639-1 code does not exist, possibly also including script identifiers as specified in ISO 15924, preceded by a hyphen.

NOTE Guidance on the use of ISO 639, ISO 3166 and ISO 15924 codes for the creation of extended language codes can be found in RFC 3066.

9. Territory, as two-letter form of ISO 3166. The reserved ISO 3166 code EU will be used to designate the territory of Europe.

For Types 3, 4, and 6, POSIX Charmaps, Repertoiremaps, and Machine-parsable coded character set specifications:

10. Suggested Charmap or Repertoiremap or other name

All applications shall contain information on these items:

11. If not for general use, an indication of the intended user audience. The Registration Authority decides on a corresponding identifier element, to be used in the token identifier for the specification.
12. If for use of a special application, a description of the application. The Registration Authority decides on a corresponding identifier element, to be used in the token identifier for the specification.
13. Short name for Sponsoring Authority, for possible use in the token identifier. Blank if this is from a National Body.
14. Revision number consisting of digits and zero or more full stops (".").
15. Revision date in the format according to this example: "1995-02-05" meaning the 5th of February, 1995.

If any of the above information is non-existent, it must be stated in each case; the corresponding string is then the empty string. The default case in item 11 and 12 is also represented by an empty string. If required information is not present in any of the ISO

639 parts or ISO 3166, the relevant Maintenance Authority shall be approached by the Sponsoring Authority to get the needed item registered.

The information in items 8 to 14 is used in the token identifier for the Cultural Specifications. Items 8 to 13 may contain digits 0123456789 and the characters uppercase and lowercase forms of

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Item 10 may also contain the special characters:

/()*-.:_

NOTE All of these characters are included in ISO/IEC 10646 U0020..U007E.

Case of letters is not significant in token identifiers.

The form in Annex A shall be included as part of an application for registration of a Cultural Specification. Annex B gives an example of a completed form..

15 Specification of the token identifier

The information in clause 14.7 items 8 to 14 is used by the Registration Authority to construct a token identifier for the Cultural Specification according to the following rules. The token identifier may then be used to uniquely identify a Cultural Specification in a manner that may be more indicative of its contents than a mere numeric identifier. The maximum length of a token identifier is 200 characters.

For Narrative Cultural Specifications, POSIX Locales and FDCC-sets the token identifier will be:

8_9+11+12,13_14

For Charmaps and Repertoiremaps the token identifier will be:

10+11+12,13_14

where 11 and 12 and preceding pluses shall be omitted when not needed to specify position, and 13 may be omitted after request from the Sponsoring Authority, if this is a National Body.

The HYPHEN character "-" may be substituted for the UNDERLINE character "_", in order to align names with RFC 3066.

NOTE 1 A combination of a POSIX Locale or FDCC-set, and a Charmap may be designated by the Locale or FDCC-set identifier and the Charmap identifier separated by a solidus (/).

When referencing a Cultural Specification, the version number or parts thereof taken from the right may be omitted, to refer to the Cultural Specification with the highest digital version number available with the given version number prefix. If the item 13 is an empty string, referencing the token identifier without the preceding comma and items 13 and 14 shall give the Cultural Specification with the highest digital version number, thus giving preference to specifications from National Bodies.

NOTE 2 The version number may be used by the Sponsoring Authority to mark major releases, minor revisions and error corrections. It is recommended that major releases be reflected as the first number, minor revisions in the second number, and error corrections in the third number.

EXAMPLE 1: _EU,CEN_3.5 for the CEN European POSIX Locale
EXAMPLE 2: da_DK,_2.4 for the Danish Standards Danish POSIX Locale
EXAMPLE 3: ISO-8859-1:1987,DS_1.0 for the DS Charmap for ISO 8859-1

16 Initial registration procedures

16.1 The Sponsoring Authority shall prepare an application for registration according to clause 14.

16.2 The Sponsoring Authority shall submit an application for registration of a cultural specification to the Registration Authority.

16.3 The Registration Authority shall examine each application received. It shall ascertain that

The applicant is a Sponsoring Authority as identified in clause 7. The Registration Authority shall reject applications for registrations which come from sources other than the Sponsoring Authorities as defined in clause 7, or applications for registrations that the Sponsoring Authority does not have to authority to submit according to clause 7.1. The Registration Authority may refer the applicant to an appropriate Sponsoring Authority if one can be identified.

The proposed cultural specification is not identical to one already registered.

If the application fails to meet either of these requirements, the application shall be rejected.

When requested by the Registration Authority, the RA-JAC may provide an opinion on whether an application satisfies these requirements.

16.4 The Registration Authority shall also ascertain that

- The application for registration is legible and meets the presentation requirements of this international standard. See clause 14.1.

- The application for registration is in the proper syntax of the type of the Cultural

specification, See clauses 14.2, 14.3, and 14.4.

- The format of the application conforms to clause 14.5.

- The application includes the elements required from the Sponsoring Authority for the cover page. See clause 14.7.

- The application for registration includes any required copyright permissions and endorsements. See clause 14.6.

If the application for registration fails to meet any of these requirements, the Registration Authority shall inform the Sponsoring Authority of the changes needed to meet the requirements.

16.5 The Registration Authority shall submit the application to the RA-JAC for review for syntactical requirements required for the registration. The RA-JAC shall ascertain that

- The application is syntactically in accordance with this International Standard.

- The application for registration includes the required description of the cultural specification. See clause 14.7.

16.6 The RA-JAC shall report the results of its evaluation for registration requirements within 30 days to the Registration Authority and shall describe any formal concerns with the proposed registration.

16.7 The Registration Authority shall inform the Sponsoring Authority of any changes needed to satisfy the concerns of the RA-JAC regarding registration requirements.

16.8 After an application for registration has passed its review for syntactical and administrative errors by the Registration Authority and by the RA-JAC, the Registration Authority shall circulate the application and the proposed token identifiers to the members and liaison organizations of the subcommittee responsible for maintaining this standard and to the RA-JAC for an information and comment period of 90 days.

16.9 The Registration Authority shall forward any comments that prevent approval of the application for registration to the Sponsoring Authority for a response.

16.10 If the review comment is in error, the Sponsoring Authority shall submit an explanation of why it is in error to the Registration Authority. If the review comment is justified, the Sponsoring Authority shall either request withdrawal of the application (see clause 21.1) or submit a corrected application, A corrected application shall be subjected to the procedures specified in clauses 16 and 17 as though it was a new application.

16.11 The Registration Authority shall approve or reject the application for registration.

16.12 The Registration Authority shall process approved applications in accordance with clause 17.

16.13 The Registration Authority shall request the RA-JAC to provide expert advice on technical comments. The Registration Authority may attach comments from the RA-JAC that four fifths of the RA_JAC agree upon as described in clause 9.3 item e with the final registration.

16.14 When an application for registration is rejected, the Registration Authority shall inform the Sponsoring Authority and provide the reason for the rejection.

17 Processing of an approved application

Following completion of approval of an application for registration, the Registration Authority shall take the actions listed in this clause.

17.1 Assign a new Cultural Specification numeric identifier as follows:

- Numeric identifiers shall be allocated in ascending order. This allocation shall only be made immediately prior to publication of the registration, that is, after completion of all procedural steps.

- No numeric identifiers shall be reserved for future registration applications.

- A numeric identifier, once allocated to a registration, shall never be re-allocated for another registration.

17.2 The Registration Authority may also assign one or more token identifiers to the approved registration.

If the Cultural Specification is identical to one already registered, the new token identifiers shall be added to the existing registration, and the addition shall be noted in the version history of that registration;

17.3 The Registration Authority shall note the date of approval in the registration.

17.4 The Registration Authority shall publish the approved registration in the ISO/IEC 15897 register.

17.5 The Registration Authority shall notify the Sponsoring Authority of the publication of the registration.

17.6 The Registration Authority shall announce publication of the registration to the subcommittee responsible for maintaining this standard.

18 Appeal procedures

Appeal against the decision of the Registration Authority shall be made according to the procedures in this clause.

18.1 Appeals against rejection

If the Registration Authority rejects an application, the Sponsoring Authority may appeal that rejection based only on whether the application meets the syntactical or administrative requirements for a registration as described in clause 14.

18.2 Appeals against registration

18.2.1 The Registration Authority for shall accept an appeal from the subcommittee responsible for the maintenance of this International Standard when any Member Body objects to the forthcoming publication of a registration by the Registration Authority.

18.2.2 The Registration Authority shall accept appeals from the subcommittee responsible for the maintenance of this International Standard for the following reasons only:

1)disagreement with the Registration Authority on whether the application meets the syntactical or administrative requirements for a registration in clause 16.

2)disagreement with the Registration Authority on whether the application matches an existing registration.

3)disagreement on the correctness of some of the information in the cultural specification of the application.

18.3 Invalid reasons for an appeal

The following objections shall be considered invalid as grounds for an appeal:

- one or more registrations exist with identically the same field of application (Note that a national body is permitted to register different cultural specifications for its region, for example, alternate collation orders for different applications);

- an allegation is made that the technical content of the registration does not achieve its alleged purpose;

- editorial comments are rejected by the Registration Authority;

18.4 Procedure for filing an appeal

Appeals shall be sent to the Registration Authority by registered mail, facsimile, or electronic mail, either

within 30 days of receipt of the refusal of the Registration Authority; or

- within 30 days after the end of the circulation period to member bodies according to Clause 16.8.

17.5 Resolution of an appeal

18.5.1 Within 30 days after the receipt of an appeal, the Registration Authority shall submit the appeal to the members of the RA-JAC.

18.5.2 If four-fifths of the members of the RA-JAC consider an appeal against registration from a Sponsoring Authority to be administratively or syntactically justified, the Registration Authority shall approve the registration application.

18.5.3 If four-fifths of the members of the RA-JAC consider an appeal against registration from the subcommittee responsible for maintaining this standard to be administratively or syntactically justified, the Registration Authority shall disapprove the registration application.

18.5.4 If four-fifths of the members of the RA-JAC cannot agree on how to resolve an appeal, then the appeal shall be submitted to the P-members of the JTC 1 subcommittee responsible for the maintaining of this International Standard, according to the Procedures for the technical work of ISO/IEC JTC 1.

19 Revisions

19.1 In general, no changes to the content of a registration are permitted, as this would be contrary to the principles on which the registrations are based.

19.2 When a new registration application is based on an existing registration, either by the same Sponsoring Authority, or another Sponsoring Authority, then the Registration Authority shall create a new registration. The Registration Authority shall also add cross-reference notes to the two registrations.

20 Additions of token identifiers to an existing registration

When a Cultural Specification submitted for registration is identical to one already registered, the token identifier(s) for the new application shall be added to the existing registration.

21 Withdrawal

Withdrawal is a formal declaration by which the Sponsoring Authority informs the Registration Authority that it withdraws its support of a registration application or of all of an existing registration that it has sponsored.

Parts may be withdrawn only by withdrawing the whole registration, and then registering the parts in question according to the rules defined in this International Standard. Procedures on withdrawal of parts of a registration or application are specified in clause 21.3.

A declaration of withdrawal may, but need not, be accompanied by a statement of the reasons for the withdrawal.

21.1 Withdrawal of an application for registration

21.1.1 When the Registration Authority is notified, it shall take no further action to process the application.

21.1.2 If the application for registration is being circulated for comment according to clause 16.8, the Registration Authority shall notify the members of the subcommittee that the application has been withdrawn by the Sponsoring Authority.

21.2 Withdrawal of an entire existing registration

21.2.1 After withdrawal, the registration shall remain in the register and continue to be identified by the allocated numeric identifier and token identifiers.

21.2.2 After the date of withdrawal, the Registration Authority shall issue a new cover page for the registration and shall note on it that the registration was withdrawn by the Sponsoring Authority and give the date of withdrawal. When the Sponsoring Authority has given a reason for a withdrawal, the reason may be noted in the registration.

21.2.3 The Registration Authority shall inform the subcommittee responsible for maintaining this standard of the withdrawal of a registration.

21.3 Procedures for parts

21.3.1 Parts of an existing registration

Part of an existing registration cannot be withdrawn or modified. To change a part (either to add a part, modify an existing part, or withdraw a part), the whole registration must be withdrawn, and a revision submitted as a new application for registration.

21.3.2 Parts of an existing application

Part of an existing application cannot be withdrawn or modified. To change a part (either to add a part, modify a part, or withdraw a part), the whole application must be withdrawn, and resubmitted after revision.

Annex A

(normative)

Application form for a Cultural Specification

Please specify all data relevant for the Cultural Specification type, or enter "not applicable". Please fill out one form for each Cultural Specification submitted. When completed, please send it to the Registration Authority as listed in clause 6.3.

1. Cultural Specification type number: _____

2. Organization name of Sponsoring Authority: _____

3. Organization postal address: _____

_____

4. Name of contact person: _____

5. Electronic mail address of contact person: _____

6. Telephone number for contact person: + ___ _____

7. Fax number for contact person: + ___ _____

For Narrative Cultural Specifications, POSIX Locales or Machine-parsable cultural specifications (type 1, 2, and 5):

8. Natural language, as specified in ISO 639-1 (or ISO 639-2): _____

9. Territory, as two-letter form of ISO 3166: _____

For POSIX Charmaps, Repertoiremaps, or Machine-parsable coded character set specifications (type 3, 4 and 6):

10. Charmap, Repertoiremap or Machine-parsable coded character set name: _____

For all types:

11. If not for general use, an intended user audience, for example librarians: _____

12. If for use of a special application, the short application name: _____

13. Short name for Sponsoring Authority, used in token identifier: _____

14. Version number with zero or more dots: _____

15. Revision date in ISO 8601 format: _____

The Cultural Specification identified above, and of which we hold copyright, is allowed for free distribution.

Date: _____ Authorized signature: _____

Annex B

(informative)

Sample Application for a Cultural Specification

Please specify all data relevant for the Cultural Specification type, indicating non-available data by "not available". Please fill out one form for each Cultural Specification submitted. When completed, please send it to the Registration Authority as listed in clause 6.3.

1. Cultural Specification type number: 1, Narrative Cultural Specification

2. Organization name of Sponsoring Authority: National Standards Authority of Ireland
3. Organization postal address: Glasnevin, Dublin 9, Ireland

4. Name of contact person: Seán Citizen

5. Electronic mail address of contact person: nsai@nsai.ie

6. Telephone number for contact person: +353 1 807-3800

7. Fax number for contact person: +353 1 807-3838

For Narrative Cultural Specifications, POSIX Locales, or Machine-parsable cultural specifications (type 1, 2 and 5):

8. Natural language, as specified in ISO 639-1 (or ISO 639-2): ga (Irish Gaelic)

9. Territory, as two-letter form of ISO 3166: IE (Ireland)

For POSIX Charmaps, Repertoiremaps, or Machine-parsable coded character set specifications (type 3, 4 and 6):

10. Charmap, Repertoiremap or Machine-parsable coded character set name:

For all types:

11. If not for general use, an intended user audience, for example librarians:

12. If for use of a special application, short name of application:

13. Short name for Sponsoring Authority, used in token identifier:

14. Version number with zero or more dots: 0.5

15. Revision date in ISO 8601 format: 1996-01-28

The Cultural Specification identified above, and of which we hold copyright, is allowed for free distribution.

Date: 1996-03-16 Authorized signature: _____

Annex C

(normative)

External References to Cultural Specifications

C.1 Identification of Cultural Specifications

The Cultural Specifications registered according to this International Standard shall be referenced by object identifiers according to Abstract Syntax Notation 1 (ASN.1, ISO/IEC 8824 and ISO/IEC 8825).

NOTE Sponsoring authorities do not need to do anything to make this identification scheme happen.

C.2 Identification of Abstract Syntaxes

The definitions (abstract syntaxes) of the Cultural Specifications registered according to this International Standard shall be in form of ASN.1 defined arcs which follow the arc which defines this International Standard.

The ASN.1 arc defining this International Standard is:

iso(1) identified-organization(3) ewos(16) eg(2) tlg(0) cultural-register(2)

Separate arcs shall be defined for type 1, 2, 3, 4, 5 and 6 specifications:

abstract-syntaxes(1) <cultural-type> <registration-number>

where <cultural-type> is the number of the type defined in clause 10.1, and <registration-number> is the numeric identifier assigned as per clause 17.2.

In the case of a type 2 or 5 Cultural Specification the standard categories as defined by the POSIX standard may be specified in additional arcs. This arc shall be category(1) and shall follow the arc specifying the registration number:

abstract-syntaxes(1) <cultural-type> <registration-number> category(1) <locale-category>

The <locale-category> is a number as defined in clause 11 for the standard POSIX locale categories.

C.3 Object Descriptors

The object descriptors for the abstract syntax object identifiers defined in 2 above shall be the name of this International Standard followed by a corresponding <identifier>, either numeric or token identifier, as assigned per clause 17.2:

ISO/IEC 15897 Cultural Specification Type <cultural-type> <identifier>

C.4 Transfer Syntax

The transfer syntax as specified in ISO 8824 defines the encoding in which the contents of a registry entry might be transferred over a network. When transferring the contents of a registry entry over a network, the UTF-8 form of ISO/IEC 10646 shall be preferred for data encoding. Alternatively, transfer syntaxes as defined in ISO/IEC 2022 may be used.

Annex D

(informative)

Example of a Narrative Cultural Specification

A description of the clauses can be found in clause 12.The following is just an example, and not necessarily taken from the actual registry.

D.1. Danish language locale for Denmark, Narrative Cultural Specification

Users: general, applications: general
Source: Dansk Standard, date: 2002-10-08, version: 2.5
Token identifier: da_DK,_2.5

Clause 1: Alphanumeric deterministic ordering

Ordering in Danish is defined in Danish Standard DS 377, 3rd edition (1980) and the Danish Orthography Dictionary ("Retskrivningsordbogen, 3. udgave)", Alina A/S - Aschehoug Dansk Forlag A/S, København 2001. ISBN 87-23-01046-0).

Normal <a> to <z> ordering is used on the Latin script, except for the following letters: The letters <æ> <ø> <å> are ordered as 3 separate letters after <z>. <ü> is ordered as <y>, <ä> as <æ>, <ö> as <ø>, <ð> as <d>, <þ> as <t><h>, French <S> as <o><e>, Greenlandic <KRA> is ordered as <q>. Two <a>s are ordered as <å>, except when denoting two sounds (which is normally the case only in combined words). Non-accented letters come before accented letters, and capital letters come before small letters, when words otherwise compare equally. There is no explicit ordering of accents specified in "Retskrivningsordbogen", and whether case or accents are the most important is not specified.

Clause 2: Classification of characters

Danish uses normal classification of letters in uppercase and lowercase, this classification is also applicable to scripts like Greek and Cyrillic.

Clause 3: Numeric formatting

The decimal separator is COMMA <,>
The thousands separator is FULL STOP <.>
The grouping of large numbers is in groups of three digits.

Clause 4: Monetary formatting

International currency symbol: DKK 543,21
Domestic currency symbol: kr 543,21
Use of negative sign: kr -543,21
Thousands and decimal separators: kr 9.876.543,21

Clause 5: Date and time conventions

Both weekday and month names are written with an initial lower case letter in Danish (Normal capitalizing rules apply in the beginning of a sentence, etc.).

English name	Weekday names	Short weekday names

Sunday	søndag	søn
Monday	mandag	man
Tuesday	tirsdag	tir
Wednesday	onsdag	ons
Thursday	torsdag	tor
Friday	fredag	fre
Saturday	lørdag	lør

Short weekday names consisting of the two first letters are also commonly used.

English name	Month name	Short month name
January	januar	jan
February	februar	feb
March	marts	mar
April	april	apr
June	juni	jun
July	juli	jul
August	august	aug
September	september	sep
October	oktober	okt
November	november	nov
December	december	dec

Long date:	07 juni 1994
Abbreviated day and time:	tir 07 jun 1994 23:22:33 CET DST
long date with weekday:	onsdag den 21. december 1994
Abbreviated long date:	07 jun 1994
Numeric date:	1994-06-07
Time:	18:06:20

The 24 hour system is used in Denmark. There are no abbreviations commonly in use for before or after noon.

Clause 6: Affirmative and negative answers

Yes expressions	1JjYy	(= 1, Ja, Yes)
No expressions	0Nn	(= 0, Nej, No)

Clause 7: National or cultural Information Technology terminology

The official Information Technology terminology is "Edb-ordbog", DS 2049-1970, Gjellerup, København. A newer description can be found in Lars Frank: "edb-ordbogen", Kommunetryk, København 1984.

Clause 8: National or cultural profiles of standards

POSIX ISO/IEC 9945-1:1990 annex F and ISO/IEC 9945-2:1993 annex G contains example Danish POSIX profiles.

Clause 9: Character set considerations

The following is the Danish alphabet:

Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz Ææ Øø Åå

The combination of two <a>s are regarded as one <å>, originating from older orthography but still used in many person and place names.

For indicating stress, different pronunciation and long vowels, an accent can be used on all vowels:

Áá Éé ÍÍ Óó Úú Ýý

The following letters of foreign origin is commonly used in Danish newspapers and books, according to examples in "Retskrivningsordbogen":

Ââ Àà Çç Ðð Êê Ëë Èè Ôô Őő RS Þþ Üü Ää Öö

Other Latin letters are also used, for example in person and location names. Greek letters and some special letters are used in scientific environments, and also in general education.

The recommended character set is DS/ISO 8859-1; for a bigger repertoire DS/ISO/IEC 10646-1 is recommended.

Other character standards in use include ISO/IEC 8859-15, ISO/IEC 6937 and ISO/IEC 646 (a Danish version, DS 2089, of this has been withdrawn, but is still in use).

Vendor character sets in use include HP Roman 8, IBM CP 277, 278, 437, 850, 865, Macintosh, and MS CP 1252. The character sets have been described in the Internet RFC 1345, made in a Danish Standards Association and INSTA project, and they are also available in POSIX Charmap format.

Danish Internet Email exchange recommends MIME format and ISO 8859-1 encoding, if necessary in RFC1345 mnemonic format.

The Danish EDI council recommends ISO 8859-1 as the exchange coded character set, with possible RFC1345 mnemonic extensions.

Clause 10: Sorting and searching rules

The character oriented ordering is described in => Clause 1. More sophisticated ordering as described in "Retskrivningsordbogen" requires that numbers are expanded to their spelling, and also special characters be expanded to their spelling before sorting. Also a number of common words are to be discarded before sorting, such as "den", "det", "en", "et".

Clause 11: Transformation of characters

Transliteration of Cyrillic and Arabic is quite different from English conventions. Examples of transliterated Cyrillic names are Tjajkovskij, Gorbatjov, and Jeltsin; an example of a transliterated Arabic name is Khadaffi.

For a fallback notation of some letters, refer to the following table:

original letter	2-char	1-char
Æ	AE	E
Ø	OE	Y
Å	AA	O
Ü	Y	Y
Ä	Æ	Æ
Ö	Ø	Ø
Þ	TH	T

Clause 12: Character properties

For ordinary classification of characters, please refer to => Clause 2.

The Greenlandic letter KRA <ĸ> has the uppercase equivalent of K' ,which are seldom represented in a character set; if it is not represented, it is converted to a "Q" as also prescribed by modern Greenlandic orthography.

Clause 13: Use of special characters

For quoting, the character pairs <„><" >, <»><«> and <"><">are used,; the first character in each pair is used to start a quote, and the last to end the quote.

Various punctuation signs:

NUMBER SIGN <#> is seldom used, and should be avoided

AT SIGN <@> is not used for commercial purposes. It is used in Internet mail.

Double space after a FULL STOP <.> is seldom used.

DIVISION SIGN <÷> should not be used for division, as it is also used for subtraction, the sign is known as "minus" in Denmark. Use SOLIDUS </> instead.

SECTION SIGN <§> is often used in legal documents to refer to paragraphs.

In a sentence the FULL STOP <.> is placed as the last character, as in: Skibet hed "Titanic".

Clause 14: Character rendition

The Danish letters <Ø> and <ø> are often misprinted. The stroke in the letters is the problem. If you consider a rectangle box surrounding the letter, then the stroke should cross from the upper right corner to the opposite corner. The Danish letter <æ> is not to be rendered as the French <oe> ligature <œ>.

Clause 15: Character inputting

A proposed general input method is included in DS/ISO/IEC 9945-1 annex F.

Clause 16: Personal names rules

Children can get their father's or mother's last name, or any combination of these with or without a hyphen. Also in marriage the bride and the groom may take the other partner's name in any combination.

Personal names are commonly spelt with the full first given name, while use of initials only is seen also. People are mostly addressed by voice by their first name. The common address form is the informal "du", and the more formal "De" is becoming more common. The family name is never spelt in capital letters only, contrary to continental European habits. Titles are used in some circumstances.

Clause 17: Inflection

The Danish grammar is defined in "Retskrivningsordbogen". Danish has more inflections than English, for example nouns will have 8 forms based on indefinite/definite, singular/plural and nominative+others/genitive.

Danish tends to have longer words than English, as you can make combined words.

Clause 18: Hyphenation

Hyphenation rules are described in "Retskrivningsordbogen".

Clause 19: Spelling

Spelling of the Danish language is specified in "Retskrivningsordbogen". This spelling is approved by Danish Government, and used as authoritative in schools etc.

Clause 20: Numbering, ordinals and measuring systems

See => Clause 3 and => Clause 4 for a description of numeric and monetary formatting.

The measurement system is the SI system, DS/ISO 1000.

Temperatures are normally measured in degrees Celsius, the Kelvin scale is sometimes used in science.

Clause 21: Monetary amounts

See => Clause 4 for the POSIX specifications. Sometimes the local currency is written after the amount, with a space in between, like 3.500,00 kr. Sometimes a space is used for the thousands grouping.

Clause 22: Date and time

The time zone is UTC+0100 in the winter, UTC+0200 in the summer. The daylight savings period currently (2003) changes by one hour the last Sunday in March at 02:00, and back again by one hour the last Sunday in October at 03:00. This may change in the future. There is no official names for the time zones.

Use of week numbers are very common, and the week numbering is according to DS/ISO 8601.

The first day of the week is Monday, in accordance with DS/ISO 8601.

Date formatting according to DS/ISO 8601, for example 1995-04-13 for 13th of April 1995, is very common in technical business and in legal business, and other areas.

For POSIX date and time formatting, please see => Clause 5.

Clause 23: Coding of national entities

Denmark is a part of the Kingdom of Denmark, which also consists of Greenland and the Faroe Islands.

Denmark is situated about 54 - 58 degrees North, and 8 - 15 degrees East.
Denmark has an area of about 43.069 km2 and 5,2 mill inhabitants (1995).
The main language is Danish.

There are a number of standards giving a country code to Denmark:

ISO 3166 alpha-2	DK
ISO 3166 alpha-3	DNK
ISO 3166 numeric	208
CEPT-MAILCODE	DK
UN Genève 1949:68 Vehicle code	DK
CCITT E.163 international telephone prefix	45
CCITT X.121 X.25 numbering country code	238
ISO 2108 ISBN book numbering	87

The Alpha-2 code "DK" of ISO 3166 is for general use, and is use generally by the public as the abbreviation for Denmark.

The name of the country in Danish is "Danmark".

The language code according to ISO 639-1 for the Danish language is "da".

The name of the Danish language in Danish is "dansk".

The currency is Danish Kroner, in Danish, "danske kroner". The ISO 4217 code is DKK. The native abbreviation is "kr". 1 "krone" is equal to 100 "øre". See => Clause 4 for a POSIX description.

Postal codes ("postnumre") are 4 digits. See => Clause 25 for their use.

For public administration Denmark has 14 counties ("amter") and 275 communes ("kommuner"). The counties and communes have numbers, which can be found in Statistic Yearbook from Denmark's Statistics.

Clause 24: Telephone numbers

The international telephone prefix for Denmark is +45. There are no area codes; all numbers have 8 digits. The recommended format for telephone numbers is in groups of 2, for example 39 17 99 44.

Clause 25: Mail addresses

See => Clause 16 for how to write personal names.

The street number is placed after the street name.

The postal code is placed before the city name. The CEPT (Conference of European Postal and Telecommunications Association) country prefix should be placed in front of the postal code for international mail, this is even commonly done for mail within Denmark.
Postal codes are defined in "Post- og Telegrafhåndbogen - Postnummerdelen", obtainable at all postal offices, and may be found also in telephone directory books.

An example of a mail address is:

> Danish Standards Association
> Att: S142 u22
> Kollegievej 6
> DK-2920 Charlottenlund
> Danmark

According to CEPT recommendations, one should either use the French name of the country ("Danemark"), or the name in the local language "Danmark".

Storeyspecification is placed after the street number. The following conventions apply:

English	Danish	Danish abbreviation
Ground floor	stuen	st
1st floor	1. etage	1.
basement	kælderen	kld
right	til højre	th
left	til venstre	tv
middle	midt for	mf

An example of its use:

> Holger Danske
> Fremtidsvej 26, 2. tv

DK-2000 Frederiksberg

Clause 26: Identification of persons and organizations

In Denmark, persons are identified by a unique personal identity number ("personnummer" or "CPR-nummer"). This number incorporates the date of birth and the sex. The structure of the Danish personal identity number is:

DDMMYY-XNNB

where DD=day, MM=month, YY=year, X=running number, including century indication, NN=running number, and B=control digit with sex indication: odd=male; even=female.

Danish organizations are identified via the SE-number, which is also used as an identification for Value Added Tax (VAT) purposes. This is an 8-digit number, the VAT number may have a 2-digit area code attached in the end, after a hyphen.

There are a number of official registries for organizations, dependent of the organization form: "Aktieselskabsregistret", "Anpartselskabsregistret", "Fondsregistret".

Clause 27: Electronic mail addresses

The Danish X.400 email country code is DK, that is the ISO 3166 alpha-2 code.

The Danish Internet top domain is .DK (ISO 3166 alpha-2 code). Internet domain addresses have an organization name as the second level name. There are no economic sector (government, commercial, academic) indication.

The Danish X.500 service uses the character set T.61 with RFC 1345 mnemonic extensions for representing names and addresses.

Clause 28: Payment account numbers

The format of Danish bank account numbers have a 4-digit branch identification code, and then the numeric bank account number.

The format of the Danish Postal Giro accounts is 7 digits, an example is 123-4567.

Clause 29: Keyboard layout

A Danish keyboard has the layout of the alphabetic keys (first is lowercase, second is uppercase, third is alternate graphic):

```
½§  1!  2"@ 3#£ 4¤$ 5%  6&  7/{ 8([ 9)] 0=} +?  ´`|
     Q   W   E   R   T   Y   U   I   O   P   Å   ¨^~
      A   S   D   F   G   H   J   K   L   Æ   Ø   '*
       <>\ Z   X   C   V   B   N   M   ,;  .:  -_
```

´`¨^~ are normally dead keys.

Clause 30: Man-machine dialogue

Naturally, most Danish users require programs where all menus, names of icons, commands, information messages, help texts, manuals etc. are translated and adjusted to their language and culture.

Programmers and screen layout designers must bear in mind that when English text is translated into Danish - and most other languages - it will normally be longer, that is. require more space on the screen and occupy more computer memory.

Denmark has its own cultural symbols in some cases and use of non-Danish symbols as icons can create irritation and - if they are not easily recognized - confusion. Example: The typical suburban American mailbox with the raised flag is unusual in Denmark and hence not immediately associated with mail for most users.

Clause 31: Paper formats

ISO 216 paper sizes are used in Denmark. Two holes or 4 holes according to ISO 838 in A4 paper etc. is very common.

Clause 32: Typographical conventions

In Denmark the Didot point measure is used in typography, which is 7% larger than the point used in English and American typography.

When justifying text at both margins, extra space should be inserted between words, not between letters within a word.

Use of special characters are described in => Clause 13.

End of Narrative Cultural Specification.

Annex E

(normative)

"reorder-after" construct in LC_COLLATE

This is an extension technique for use with locale which are otherwise compliant with ISO/IEC 9945. This extension technique may be used in registration applications according to this International Standard.

For the LC-COLLATE definition in POSIX, additional keywords "reorder-after" and "reorder-end" are allowed when the "copy" keyword is used. This changes the specification for the "copy" keyword, as other keywords are then allowed. The "copy" shall refer a collating specification in source form not using the "copy" keyword.

The following keywords are recognized in a collating sequence definition, in addition to the keywords specified in ISO/IEC 9945:

reorder-after Redefine collating rules. Specify after which collating element the redefinition of collation order shall take order. This statement is followed by one or more collation order statements, reassigning character collation values and collation weights to collating elements.

reorder-end Specify the end of the "reorder-after" collating order statements.

E.1 "reorder-after" keyword

The "reorder-after" keyword shall be used to specify a modification to a copied collation specification of an existing locale. There can be more than one "reorder-after" statement in a collating specification. Using ISO/IEC 9945 specification methods, the syntax shall be:

"reorder-after %s\n",<collating-symbol>

The <collating-symbol> operand shall be a name, enclosed between angle brackets, and shall be present in the source locale copied via the "copy" keyword.

The "reorder-after" statement is followed by one or more collation identifier entries as described in the "Collating Order" section of ISO/IEC 9945, with the exception that the ellipsis symbol (...) shall not be used.

Each collation identifier entry reassigns character collation values and collation weights to collating elements existing in the copied collation specification, by removing the collating identifier entry from the copied specification, and inserting the collating element in the collating sequence with the new collation weights after the preceding collating element of the "reorder-after" specification, the first collating element in the collation sequence being the <collating-symbol> specified on the "reorder-after" statement.

A "reorder-after" specification is terminated by another "reorder-after" specification or the "reorder-end" statement.

E.2 "reorder-end" keyword

The "reorder-end" keyword shall specify the end of a list of collating identifier entries, initiated by the "reorder-after" keyword.

E.3 Example of "reorder-after"

```
reorder-after <y8>
<U:>        <Y>;<U:>;<CAPITAL>              % Ü
<u:>        <Y>;<U:>;<SMALL>                % ü
reorder-after <z8>
<AE>        <AE>;<NONE>;<CAPITAL>           % Æ
<ae>        <AE>;<NONE>;<SMALL>             % æ
<A:>        <AE>;<DIAERESIS>;<CAPITAL>      % Ä
<a:>        <AE>;<DIAERESIS>;<SMALL>        % ä
<O/>        <O/>;<NONE>;<CAPITAL>           % Ø
<o/>        <O/>;<NONE>;<SMALL>             % ø
<AA>        <AA>;<NONE>;<CAPITAL>  % Å U00C5 LATIN CAPITAL LETTER A WITH RING
<aa>        <AA>;<NONE>;<SMALL>    % å U00E5 LATIN SMALL LETTER A WITH RING
reorder-end
```

The example is interpreted as follows:

1. The collating element <U:> is removed from the copied collating sequence and inserted after <y8> in the collating sequence with the new weights. The collating element <u:> is removed from the copied collating sequence and inserted in the resulting collation sequence after <U:> with the new weights.

2. The second "reorder-after" statement terminates the first list of reordering collation identifier entries, and initiates a second list, rearranging the order and weights for the <AE>, <ae>, <A:>, <a:>, <O/>, <o/>, <AA> and <aa> collating elements after the <z8> collating symbol in the copied specification.

3. The "reorder-end" statement terminates the second list of reordering entries.

4. Thus for the original sequence

   ...(Aa Ää Åå) ... Ææ ... (Oo Øø) ... (Uu Üü) Vv Ww Xx Yy Zz

   this example reordering gives:

   Aa ... Oo ... Uu Vv Ww Xx (Yy Üü) Zz (Ææ Ää) Øø Åå

5. A complete example for Danish using characters from ISO/IEC 10646 is included in Annex G.1. For the sequence:

   ... (Aa Ää Åå Ǻǻ) ... (Ææ Ǽǽ Ǣǣ) ... (Oo Öö Øø)
   ...(Uu Ùù Úú Üü Űű) Vv Wv Xx (Yy Ýý) (Zz Žž)

the example reordering in Annex G.1 gives:

Aa ... Oo ... (Uu Ùù Úú) Vv Ww Xx (Yy Ýý Üü Űű) (Zz Žž)

(Ææ Ǽǽ Ǣǣ Ää) (Øø Ǿǿ Öö Őő) (Åå (AA Aa aA aa) Ǻǻ)

NOTE The characters on the last line are the following, with UCS short identifiers:

(00C6 00E6 01FC 01FD 01E2 01E3 00C4 00E4)
(00D8 00F8 01FE 01FF 00D6 00F6 0151 0150)
(00C5 00E5 (0041 0041 0041 0061 0061 0041 0061 0061) 01FA 01FB)

Annex F

(Informative)

Information on "reorder-after" construct in LC-COLLATE

F.1 "reorder-after" rationale

Much work has already been done on locales, and making them quite general. CEN/TC304 had on its programme of work the specification of a multilingual ordering for Europe, and also the collection of collating sequences of the different countries in Europe in a registry of cultural conventions; one of the formal specification techniques for this is using the POSIX standard (ISO/IEC 9945-2:1993). The POSIX standard introduces a copy command for all sections of the POSIX locale. This is good for many purposes and it ensures that two locales are equivalent for this category. A further step in building on previous locale work is defined in this International Standard.

Collating sequences often vary a bit from country to country, and from language to language, but generally much of the collating sequence is the same. For instance the Danish sequence is for the most part the same as the German, English or French collation, but for about a dozen letters it differs. The same can be said for Swedish or Hungarian: generally the Latin collating sequence is the same, but a few characters are different.

With the advent of the quite general, coded character set independent locales like the example Danish locales in the POSIX Shell and Utilities standard (ISO/IEC 9945-2:1993) annex G, and the European multilingual POSIX collating specification, it will prove to be convenient if the few differences could be specified just as changes to an existing one. Using the "reorder-after" construct will also help improve the overview of what the changes really are for implementers and other users.

An example of the use of the "reorder-after" construct is the following. A default European ordering for the Latin alphabet may be adequate for Danish, with the exception of the collation rules for the letters Ü, ü, Æ, æ, Ä, ä, Ø, ø, Å and å. By applying the "reorder-after" construct, the Danish specification can be made more easily by copying and reordering the existing European specification, rather than specifying collation parameters for all Latin letters (with or without diacritics). There is no obligation for Denmark to take this approach, but the normative annex E provides the mechanism for doing so if it is deemed desirable.

F.2 awk script for "reorder-after" construct

A script has been written in the "awk" language defined in the POSIX standard ISO/IEC 9945 to implement the "reorder-after" construct:

```
BEGIN { comment = "%"; back[0]= follow[0] = 0; }
/LC_COLLATE/ { coll=1 }
/END LC_COLLATE/ { coll=0; for (lnr= 1; lnr; lnr= follow[lnr]) print cont[lnr] }

{ if (coll == 0) print $0  ;
  else { if ($1 == "copy")   {
                        file = $2
                        while (getline < file )
                        if ( $1 == "LC_COLLATE" ) copy_lc = 1
                        else if ( $1 == "END" && $2 == "LC_COLLATE" ) copy_lc =0
                        else if (copy_lc) {
                                        lnr++
                                        follow[lnr_1] = lnr; back [ lnr ] = lnr_1
                                        cont[lnr] = $0; symb[ $1 ] = lnr
                        }
                        close (file )
        }
        else if ($1 == "reorder-after") { ra=1 ; after = symb [ $2 ] }
        else if ($1 == "reorder-end") ra = 0
        else {
                        lnr++
                        if (ra) follow [ lnr ] = follow [ after ]
                        if (ra) back [ follow [ after ] ] = lnr
                        follow[after] = lnr; back [ lnr ] = after
                        cont[lnr] = $0
                        if ( ra && $1 != comment && $1 != "" )  {
                                        old = symb [ $1 ];
                                        follow [ back [ old ] ] = follow [ old ];
                                        back [ follow [ old ] ] = back [ old ];
                                        symb[ $1 ] = lnr;
                        }
                        after = lnr
        }
  }
}
```

Annex G

(informative)

Sample POSIX Locale Specification for Danish

G.1 Danish example

This example uses characters from the repertoire of ISO/IEC 10646.

```
escape_char /
comment_char %

% Danish language locale for Denmark
% Source: Danish Standards Association
% Address: Kollegievej 6,
%   DK-2920 Charlottenlund, Danmark
% Contact: Keld Simonsen
% Email: Keld.Simonsen@dkuug.dk
% Tel: +45 - 39966101
% Fax: +45 - 39966102
% Language: da
% Territory: DK
% Revision: 4.1
% Date: 2002-10-21
% Application: general
% Users: general
% Repertoiremap: mnemonic,DS
% Charset: ISO-8859_1:1987
% Distribution and use is free, also
% for commercial purposes.

% The ordering algorithm is in accordance
% with Danish Standard DS 377 (1980)
% and the Danish Orthography Dictionary
% (Retskrivningsordbogen, 2001).
% It is also in accordance with
% Greenlandic orthography.

LC_COLLATE
collating-element <A-A> from "<A><A>"
collating-element <A-a> from "<A><a>"
collating-element <a-A> from "<a><A>"
collating-element <a-a> from "<a><a>"
copy en_DK
reorder-after <CAPITAL>
<CAPITAL>
<CAPITAL-SMALL>
<SMALL-CAPITAL>
<SMALL>
reorder-after <q8>
<kk> <Q>;<SPECIAL>;<SMALL>;IGNORE
reorder-after <t8>
<TH> "<T><H>";"<TH><TH>";"<CAPITAL><CAPITAL>";IGNORE
<th> "<T><H>";"<TH><TH>";"<SMALL><SMALL>";IGNORE
reorder-after <y8>
% <U:> and <U"> are treated as <Y> in Danish
<U:> <Y>;<U:>;<CAPITAL>;IGNORE
<u:> <Y>;<U:>;<SMALL>;IGNORE
<U"> <Y>;<U">;<CAPITAL>;IGNORE
<u"> <Y>;<U">;<SMALL>;IGNORE
reorder-after <z8>
% <AE> is a separate letter in Danish
```

```
<AE> <AE>;<NONE>;<CAPITAL>;IGNORE
<ae> <AE>;<NONE>;<SMALL>;IGNORE
<AE'>           <AE>;<ACUTE>;<CAPITAL>;IGNORE
<ae'>           <AE>;<ACUTE>;<SMALL>;IGNORE
<A3> <AE>;<MACRON>;<CAPITAL>;IGNORE
<a3> <AE>;<MACRON>;<SMALL>;IGNORE
<A:> <AE>;<SPECIAL>;<CAPITAL>;IGNORE
<a:> <AE>;<SPECIAL>;<SMALL>;IGNORE
% <O//> is a separate letter in Danish
<O//>           <O//>;<NONE>;<CAPITAL>;IGNORE
<o//>           <O//>;<NONE>;<SMALL>;IGNORE
<O//'>          <O//>;<ACUTE>;<CAPITAL>;IGNORE
<o//'>          <O//>;<ACUTE>;<SMALL>;IGNORE
<O:> <O//>;<DIAERESIS>;<CAPITAL>;IGNORE
<o:> <O//>;<DIAERESIS>;<SMALL>;IGNORE
<O"> <O//>;<DOUBLE-ACUTE>;<CAPITAL>;IGNORE
<o"> <O//>;<DOUBLE-ACUTE>;<SMALL>;IGNORE
% <AA> is a separate letter in Danish
<AA> <AA>;<NONE>;<CAPITAL>;IGNORE
<aa> <AA>;<NONE>;<SMALL>;IGNORE
<A-A>           <AA>;<A-A>;<CAPITAL>;IGNORE
<A-a>           <AA>;<A-A>;<CAPITAL-SMALL>;IGNORE
<a-A>           <AA>;<A-A>;<SMALL-CAPITAL>;IGNORE
<a-a>           <AA>;<A-A>;<SMALL>;IGNORE
<AA'>           <AA>;<AA'>;<CAPITAL>;IGNORE
<aa'>           <AA>;<AA'>;<SMALL>;IGNORE
reorder-end
END LC_COLLATE

LC_CTYPE
copy "en_DK"
END LC_CTYPE

LC_MONETARY
int_curr_symbol         "<D><K><K><SP>"
currency_symbol         "<k><r>"
mon_decimal_point       "<,>"
mon_thousands_sep       "<.>"
mon_grouping            3;3
positive_sign           ""
negative_sign           "<->"
int_frac_digits         2
frac_digits             2
p_cs_precedes           1
p_sep_by_space          2
n_cs_precedes           1
n_sep_by_space          2
p_sign_posn             4
n_sign_posn             4
END LC_MONETARY

LC_NUMERIC
decimal_point           "<,>"
thousands_sep           "<.>"
grouping                3;3
END LC_NUMERIC

LC_TIME
abday       "<s><o//><n>";"<m><a><n>";/
            "<t><i><r>";"<o><n><s>";/
            "<t><o><r>";"<f><r><e>";/
            "<l><o//><r>"
day         "<s><o//><n><d><a><g>";/
            "<m><a><n><d><a><g>";/
            "<t><i><r><s><d><a><g>";/
```

```
            "<o><n><s><d><a><g>";/
            "<t><o><r><s><d><a><g>";/
            "<f><r><e><d><a><g>";/
            "<l><o//><r><d><a><g>"
abmon       "<j><a><n>";"<f><e><b>";/
            "<m><a><r>";"<a><p><r>";/
            "<m><a><j>";"<j><u><n>";/
            "<j><u><l>";"<a><u><g>";/
            "<s><e><p>";"<o><k><t>";/
            "<n><o><v>";"<d><e><c>"
mon         "<j><a><n><u><a><r>";/
            "<f><e><b><r><u><a><r>";/
            "<m><a><r><t><s>";/
            "<a><p><r><i><l>";/
            "<m><a><j>";/
            "<j><u><n><i>";/
            "<j><u><l><i>";/
            "<a><u><g><u><s><t>";/
            "<s><e><p><t><e><m><b><e><r>";/
            "<o><k><t><o><b><e><r>";/
            "<n><o><v><e><m><b><e><r>";/
            "<d><e><c><e><m><b><e><r>"
d_t_fmt     "<%><a><SP><%><d><SP><%><b><SP><%><Y><SP><%><T><SP><%><Z>"
d_fmt       "<%><d><-><%><m><-><%><Y>"
t_fmt       "<%><T>"
am_pm       "";""
t_fmt_ampm  ""
END LC_TIME

LC_MESSAGES
yesexpr     "<<(><1><J><j><Y><y><)/>><.><*>"
noexpr      "<<(><0><N><n><)/>><.><*>"
END LC_MESSAGES
```

Annex H

(informative)

Differences from ISO/IEC 15897:1999 and CEN ENV 12005:1996

H.1 Changes from ISO/IEC 15897:1999

The edition was done as a normal revision of the standard, and thus a number of changes were done during the development phase and due to ballot comments. This list the major changes from ISO/IEC 15897:1999 to this revision of the International Standard.

1. Specifications from ISO/IEC TR 14652 and other machine parsable formats were added to the list of possible items for registration.

2. CEN/TC304 and other JTC1 Subcommittees and Working Groups were added to the list of Sponsoring Authorities.

3. The text was revised to align with International Standard ISO/IEC 2375:2003.

4. ISO 639-2 codes were added as an alternative to ISO 639 codes, when the latter did not exist., and the character "-" was added as a synonym to "_" for compatibility with RFC 3066.

5. A Registration Authority's Joint Advisory Committee (RA-JAC) was added to comment on registrations and in the appeals procedure.

6. French and Russian were added as languages for narrative cultural specifications.

7. Some parts of the text were moved around. For example, the former annex G which is now clause 11.

H.2 Changes from CEN ENV 12005:1996

To make the European Prestandard into International Standard ISO/IEC 15897:1999 the following changes were made: The word "European" was removed throughout the standard where appropriate and references to CEN were changed to ISO/IEC. CEN terminology such as "Prestandard" was changed to "International Standard" throughout the standard. The foreword and introduction were replaced by text from the ISO Central Secretariat, and editorial comments from ISO/CS were accommodated. This annex describing the changes from the CEN ENV was added to the International Standard.

The references to clauses in the following refer to the 1999 edition of ISO/IEC 15897. The detailed changes were:

The word "Prestandard" was changed to "International Standard" throughout the document where appropriate. The word "European" was removed from the title.

In clause 1 the erroneous reference to 6.11 was changed to 6.8.

In clause 4 the text "The Technical Board of CEN has designated as Registration Authority of this Prestandard:" was changed to: "The ISO and IEC councils shall designate a Registration Authority for this International Standard. The initial Registration Authority is designated to be:".

In clause 4 the contact information for the Registration Authority has been updated.

The audience in clause 4.d was changed from "CEN member bodies, Associated member bodies and Liaisons of the committee" to "ISO/IEC JTC1/SC22 members and liaisons".

The audience in clause 4.h was changed from "CEN member bodies, Associated member bodies and Liaisons" to

"ISO/IEC JTC1 members and liaisons".

Clause 5.b was changed to "ISO/IEC JTC1/SC22 for applications related to wider regions.".

In clause 7.2, "CEN Technical Committee" was changed to "JTC 1 subcommittee".

In clause 7.4, the first "CEN Technical Committee" was changed to " JTC 1 subcommittee", the second "CEN" was changed to "its", and remove the last subsentence after the last comma.

In annex B the contact data for the Irish member body was updated.

In clause C.3 "CEN ENV" was changed to "ISO/IEC 15897".

Bibliography

1.*Nordic Cultural Requirements on Information Technology*(Summary Report), INSTA technical report STRÍ TS3, Staðlaráð Íslands, Reykjavík 1992, ISBN 9979-9004-3-1.

2.ISO/IEC TR 14652:2004: *Information technology - Specification method for cultural conventions*.

3.RFC 3066. H. Alvestrand: *Tags for the Identification of Languages.* IETF, January 2001.

4.CEN CR 14400:2001: *European ordering rules - Ordering for Latin, Greek, Cyrillic, Georgian and Armenian scripts*

5.A number of translitteration and fallback specifications in the syntax of an earlier draft of ISO/IEC TR 14652 may be found at:
http://sources.redhat.com/cgi-bin/cvsweb.cgi/libc/localedata/locales/?cvsroot=glibc.
Examples are sr_YU translit_hangul translit_combining translit_cjk_variants translit_fractions.

End of International Standard