



National Standards Authority of Ireland

IRISH STANDARD

I.S. CEN/TR 14381:2003

ICS 35.040

**INFORMATION TECHNOLOGY – CHARACTER
REPERTOIRE AND CODING
TRANSFORMATIONS – EUROPEAN
FALLBACK RULES**

National Standards
Authority of Ireland
Dublin 9
Ireland

Tel: (01) 807 3800
Tel: (01) 807 3838

*This Irish Standard was
published under the
authority of the National
Standards Authority of
Ireland
and comes into effect on:
August 15, 2003*

**NO COPYING WITHOUT NSAI
PERMISSION EXCEPT AS
PERMITTED BY COPYRIGHT
LAW**

© NSAI 2003

Price Code T

Údarás um Chaighdeán Náisiúnta na hÉireann

TECHNICAL REPORT
RAPPORT TECHNIQUE
TECHNISCHER BERICHT

CEN/TR 14381

April 2003

ICS 35.040

English version

Information technology – Character repertoire and coding
transformations – European fallback rules

This Technical Report was approved by CEN on 2 June 2002. It has been drawn up by the Technical Committee CEN/TC 304.

CEN members are the national standards bodies of Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Malta, Netherlands, Norway, Portugal, Slovakia, Spain, Sweden, Switzerland and United Kingdom.



EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

Management Centre: rue de Stassart, 36 B-1050 Brussels

CEN/TR 14381:2003 (E)**Contents****Page**

Foreword	3
0 Introduction	4
0.1 Rationale for the provision of fallback rules	4
0.2 Basic concepts	5
0.3 Requirements	6
0.4 Satisfying the requirements	6
1. Scope and field of application	6
1.1 Scope	6
1.2 Field of application.....	6
2. Normative references.....	7
3. Definitions and abbreviations.....	7
3.1 Basic definitions	7
3.2 Other definitions.....	7
3.3 Abbreviations	8
4 Specification of the general fallback rules	8
Annex 1 The list of fallback specification per character.....	9
Annex II Examples of fallback representation of text in different languages and scripts	61
II.1 Multilingual original text.....	61
II.2 Multilingual text with fallbacks	62
Annex III Notes on fallback for Latin, Greek and Cyrillic characters	63
III.1 Fallback from extended Latin characters.....	63
III.2 Fallback from Greek characters to Latin characters	63
III.3. Fallback with a One-to-many transliteration	64
III.4 Fallback with a one-to-many transcription	64
III.5. Restoring Greek text from Latin script fallback text.....	64
III.6. Fallback with a one-to-one transliteration	65
III.7 Fallback from Cyrillic characters to Latin characters.....	65
III.8. Fallback with a one-to-many transliteration	65
III.9 Restoring Cyrillic text from Latin script fallback text	66
III.10 Fallback with a one-to-one transliteration	66
Bibliography	68

Foreword

This document (CEN/TR 14381:2003) has been prepared by Technical Committee CEN/TC 304 "Information and communications technologies - European localization requirements", the secretariat of which is held by IST.

The text of this technical report was written with the intent of it being published as a European pre-Norm (ENV). In light of the various formal and informal comments received on the document (some of which were only received after the closing of the ballot) the TC has resolved to turn this document into a CEN report as a recorded example of an attempt to formulate European wide fallback rules. It is evident that any fallback scheme in order for it to become acceptable by the users and the industry will need to be very carefully laid out and explained.

Resolutions no 7 of the 16th Meeting and no 4 of the 18th Meeting of CEN/TC 304 refer to this technical report:

Res. 7/16th. TC304 acknowledges that the Fallback project team has completed its contracted work. Although the proposed draft has received the sufficient support to be forwarded to CEN/BT for final adoption as an ENV, the nature of the comments received is such that it is decided to publish it as a CEN Report with editorial comments added by the secretary and reviewed by TC members and observers before final publication. Unanimous.

Res. 4/18th. TC304 accepts the Fallback document in N978 to be presented to CEN BT for adoption as a CR with the following text on Greek letters added in the foreword: "The method of performing fallback from Greek letters into Latin letters is especially seen as posing problems to Greek users and its use is not advised". Unanimous.

This technical report is intended to facilitate cross border communications and data exchange and to ensure that European cultural requirements are safeguarded in the increasingly interconnected world of today. It provides rules for fallback for multilingual European texts into the invariant set of ISO/IEC 646. These rules come into effect if data from different languages must be represented by equipment and systems that do not support the presentation of all the characters in the different language repertoires.

This technical report does not intend to influence, let alone substitute itself for, national standards or customs in this field. Nevertheless, national standards have the opportunity to adapt this Technical Report by declaring a formalized set of deviation rules (»delta«) if they so wish.

This document does not cancel or replace any other technical report or standard.

There is no known identical national technical report or standard in Europe.

According to the CEN/CENELEC Common Rules the following countries are bound to announce the existence of this Technical Report: Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Malta, Netherlands, Norway, Portugal, Slovakia, Spain, Sweden, Switzerland and United Kingdom.

CEN/TR 14381:2003 (E)**0 Introduction****0.1 Rationale for the provision of fallback rules**

Users who are trying to write text in a language which is not their mother tongue (native language) often wish to write that text using a character repertoire which does not contain all the letters needed for that language, especially those with diacritic marks. A method of character substitution would be useful for such users.

In spite of the computers being able to process larger repertoires of graphic characters than ever before, there are cases where it is not possible to render all the characters of a processing repertoire on an output device. In these cases, not all the characters in a processing repertoire are available in an output repertoire. In order to cater for these situations, a widely applicable standard method of character substitution (fallback) is required which will allow an approximate rendition to be made of the unsupported characters of the processing repertoire for output and rendition.

Examples of key applications are:

- a) a multilingual information service offered across Europe where personal or business documents come from different countries and are presented in a standardised rendition using MES 2 characters and which cannot be properly represented by the information service; and
- b) search engines in the World Wide Web which make use of "fuzzy" search techniques based on the use of search terms which have diacritical marks removed and make use of common substitutions for less frequently used letters of the Latin alphabet. Examples of the latter are - eth (ð), thorn (þ), æ, œ and the German sharp s (ß)

The provision of single fallback rules with a collection of fallback representations for MES 2 will enable the services to be improved and the applications easier for use by the human end user. A standard set of substitutions would be useful for such applications in order to avoid confusion. The same applies for the other two scripts represented by MES-2, Cyrillic and Greek.

The justification for preparing a technical report for these purposes is that the concept of representing the characters without diacritical marks is not useful for scripts originating outside Europe. Furthermore, Standardisation bodies of Europe that may wish to specify national schemes for fallback may modify the scheme given in this technical report for a limited set of characters and promulgate national standards for fallbacks. Greek and Cyrillic fallback representation specified in this technical report should be used with caution since transliteration into various Latin script languages depends on the target language. Local standards or local best practice should be referenced where they exist.

This is a free preview. Purchase the entire publication at the link below:

[Product Page](#)

-
- Looking for additional Standards? Visit Intertek Inform Infostore
 - Learn about LexConnect, All Jurisdictions, Standards referenced in Australian legislation
-