INTERNATIONAL STANDARD

ISO 10303-21

First edition
1994-12-15

Industrial automation systems and integration — Product data representation and exchange —

Part 21:
Implementation methods: Clear text encoding of the exchange structure

Systèmes d'automatisation industrielle et intégration — Représentation et échange de données de produits —

Partie 21: Méthodes de mise en application: Encodage en texte clair des fichiers d'échange

Reference number
ISO 10303-21:1994(E)

Contents

International Organization for Standardization
Case Postale 56 • CH-1211 Genève 20 • Switzerland
Printed in Switzerland

Annexes

Figures

Tables

Foreword

The International Organization for Standardization (ISO) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75% of the member bodies casting a vote.

International Standard ISO 10303-21 was prepared by Technical Committee ISO TC184, *Industrial automation systems and integration*, Subcommittee SC4, *Industrial data and global manufacturing languages*.

ISO 10303 consists of the following parts under the general title *Industrial automation systems and integration - Product data representation and exchange*:

— Part 1 Overview and fundamental principles;

— Part 11 Description methods: The EXPRESS language reference manual;

— Part 21 Implementation methods: Clear text encoding of the exchange structure;

— Part 22 Implementation methods: Standard data access interface;

— Part 31 Conformance testing methodology and framework: General concepts;

— Part 32 Conformance testing methodology and framework: requirements on testing laboratories and clients;

— Part 41 Integrated generic resources: Fundamentals of product description and support;

— Part 42 Integrated generic resources: Geometric and topological representation;

— Part 43 Integrated generic resources: Representation structures;

— Part 44 Integrated generic resources: Product structure configuration;

— Part 45 Integrated generic resources: Materials;

— Part 46 Integrated generic resources: Visual presentation;

— Part 47 Integrated generic resources: Shape variation tolerances;

— Part 49 Integrated generic resources: Process structure and properties;

— Part 101, Integrated application resources: Draughting;

— Part 104, Integrated application resources: Finite element analysis;

— Part 105, Integrated application resources: Kinematics;

— Part 201, Application protocol: Explicit draughting;

— Part 202, Application protocol: Associative draughting;

— Part 203, Application protocol: Configuration controlled design;

— Part 207, Application protocol: Sheet metal die planning and design;

— Part 210, Application protocol: Printed circuit assembly product design data;

— Part 213, Application protocol: Numerical control process plans for machined parts.

The structure of this International Standard is described in ISO 10303-1. The numbering of the parts of this International Standard reflects its structure:

— Part 11 specifies the description method;

— Parts 21 and 22 specify the implementation methods;

— Parts 31 and 32 specify the conformance testing methodology and framework;

— Parts 41 to 49 specify the integrated generic resources;

— Parts 101 to 105 specify the integrated application resources;

— Parts 201 to 213 specify the application protocols.

Should further parts be published, they will follow the same numbering pattern.

Annexes A, B and C form an integral part of this part of ISO 10303. Annexes D, E, and F are for information only.

Introduction

ISO 10303 is an International Standard for the computer-interpretable representation and exchange of product data. The objective is to provide a neutral mechanism capable of describing product data throughout the life cycle of a product, independent from any particular system. The nature of this description makes it suitable not only for neutral file exchange, but also as a basis for implementing and sharing product databases and archiving.

This International Standard is organized as a series of parts, each published separately. The parts of ISO 10303 fall into one of the following series: description methods, integrated resources, application protocols, abstract test suites, implementation methods, and conformance testing. The series are described in ISO 10303-1. This part of ISO 10303 is a member of the implementation methods series.

This part of ISO 10303 specifies a mechanism that allows product data represented using the EXPRESS language, specified in ISO 10303-11, to be transferred from one computer system to another.

Major subdivisions in this part of ISO 10303 are

— specification of the exchange structure syntax;

— mapping from an EXPRESS schema onto this syntax.

NOTE - The examples of EXPRESS usage in this part of ISO 10303 do not conform to any particular style rules. Indeed, the examples sometimes use poor style to conserve space or to concentrate on the important points. The examples are not intended to reflect the content of the information models defined in other parts of this International Standard. They are crafted to show particular features of EXPRESS or of the exchange structure. Many examples are annotated in a way that is not consistent with the syntax rules of this part of ISO 10303. These annotations are introduced by symbolic arrows, either horizontal '---->', or vertical. These annotations should be ignored when considering the parse rules. Any similarity between the examples and the normative models specified in other parts of this International Standard should be ignored. Several mapping examples have been provided throughout this document. Additional *spaces* and new lines have been inserted into some of these examples to aid readability. These *spaces* and new lines need not appear in an exchange structure.

Industrial automation systems and integration — Product data representation and exchange — Part 21: Implementation methods: Clear text encoding of the exchange structure

1 Scope

This part of ISO 10303 specifies an exchange structure format using a clear text encoding of product data for which the conceptual model is specified in the EXPRESS language (ISO 10303-11). The file format is suitable for the transfer of product data among computer systems.

The mapping from the EXPRESS language to the syntax of the exchange structure is specified. Any EXPRESS schema can be mapped onto the exchange structure syntax.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 10303. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 10303 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3788:1990, *Information processing — 9-Track, 12,7 mm (0.5 in) wide magnetic tape for information interchange using phase encoding at 126 ftpmm (3200 ftpi) — 63 cpmm (1600 cpi).*

ISO 8601:1988, *Data elements and interchange formats — Information interchange — Representation of dates and times.*

ISO 8859-1:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 1: Latin alphabet No. 1.*

ISO 8859-2:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 2: Latin alphabet No. 2.*

ISO 8859-3:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 3: Latin alphabet No. 3.*

ISO 8859-4:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 4: Latin alphabet No. 4.*

ISO 8859-5:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 5: Latin/Cyrillic alphabet.*

ISO 8859-6:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 6: Latin/Arabic alphabet.*

ISO 8859-7:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 7: Latin/Greek alphabet.*

ISO 8859-8:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 8: Latin/Hebrew alphabet.*

ISO 8859-9:1987, *Information processing — 8 bit single-byte coded graphic character sets — Part 9: Latin alphabet No. 5.*

ISO 10303-1:1994, *Industrial automation systems — Product data representation and exchange — Part 1: Overview and fundamental principles.*

ISO 10303-11:1994, *Industrial automation systems — Product data representation and exchange — Part 11: The EXPRESS language reference manual.*

ISO 10303-44:1994, *Industrial automation systems — Product data representation and exchange — Part 44: Product structure configuration.*

ISO/IEC 10646-1:1993-1, *Information Processing — Multiple octet coded character set — Part 1: Architecture and basic multilingual plane.*

3 Definitions

3.1 Terms defined in ISO 8859-1

This part of ISO 10303 makes use of the following terms defined in ISO 8859-1.

— byte;

— character;

— graphic character.

3.2 Terms defined in ISO 10646

This part of ISO 10303 makes use of the following term defined in ISO 10646.

— basic multilingual plane.

3.3 Terms defined in ISO 10303-1

This part of ISO 10303 makes use of the following term defined in ISO 10303-1.

— application protocol;

— exchange structure.

3.4 Terms defined in ISO 10303-11

This part of ISO 10303 makes use of the following terms defined in ISO 10303-11.

— data type;

— entity;

— token.

3.5 Terms defined in ISO 10303-44

This part of ISO 10303 makes use of the following terms defined in ISO 10303-44.

— directed acyclic graph;

— leaf node;

— node;

— tree.

3.6 Other definitions

For the purposes of this part of ISO 10303, the following definitions apply.

3.6.1 basic alphabet: the set of characters G(02/00) through G(07/14) of ISO 8859-1.

3.6.2 clear text encoding: the encoding of information that only uses 8-bit byte values corresponding to the basic alphabet.

3.6.3 control directive: a sequence of characters in the basic alphabet.

3.6.4 keyword: a special sequence of characters identifying an entity or a defined type in the exchange structure.

3.6.5 literal: an item in a language definition that stands for itself.

3.6.6 section: a collection of data of the same functional category of information.

3.6.7 sequential file: a file that can only be accessed in a sequential manner.

3.6.8 token separator: a sequence of one or more 8-bit bytes that separate any two tokens.

4 Abbreviations

For the purposes of this part of ISO 10303, the following abbreviations apply:

BMP Basic multilingual plane;

WSN Wirth Syntax Notation.

5 Exchange structure fundamental concepts and assumptions

5.1 Introduction

The exchange structure is described by an unambiguous, context-free grammar to facilitate parsing by software. The grammar is expressed in Wirth Syntax Notation that is described in annex B. The form of product data in the exchange structure is specified using a mapping from the EXPRESS language.

5.2 Notational and typographical conventions

Any *quotation marks* used in this part of ISO 10303 are not part of the text that appears in the exchange structure but serve to delimit that text. This statement applies to all places in the text where *quotation marks* are used. Table 2, table 3, and table 4 form an exception to this rule as the *quotation marks* used in those tables form part of the WSN rules.

In ISO 8859, each character is assigned an identifying name. When that name is used in this part of ISO 10303, it is typeset in *italics* to distinguish if from ordinary text. Thus *comma* is used to refer to ",", *low line* refers to "_", and *capital letter A* refers to "A".

Within examples in this part of ISO 10303, annotation is introduced by the sequence ----> where clarification is required.

5.3 Conformance

Two levels of conformance are specified:

— syntactical conformance of the exchange structure: an exchange structure conforms to ISO 10303-21 if the requirements of this part of ISO 10303 are satisfied;

— schema conformance of the exchange structure: the product data represented in the exchange structure conforms to the schema listed in the header section of the exchange structure if all

requirements and constraints of this schema are satisfied and the mapping requirements defined in clauses 10 and 11 of this part of ISO 10303 are satisfied.

Syntactical conformance is a prerequisite for schema conformance.

Two classes of syntactical conformance are defined by this part of ISO 10303, depending on the method chosen for the encoding of entity instances whose entity types are subtypes/sypertypes (see 11.2.5). An implementation that claims syntactical conformance to this part of ISO 10303 shall read or write files or both that exhibit syntactical conformance in (at least) one of these two conformance classes.

An implementation that claims schema conformance to this part of ISO 10303 shall read or write files or both that exhibit schema as well as syntactical conformance.

6 Formal definitions

6.1 Formal notation

Wirth Syntax Notation (WSN) is used in this part of ISO 10303 to specify the syntax of the exchange structure in a formal notation. WSN is described in annex B. Literals appearing within this notation are delimited by *quotation marks*.

6.2 Basic alphabet definition

The alphabet of the exchange structure is defined as the characters from G(02/00) to G(07/14) of ISO 8859-1. This alphabet is represented in the exchange structure by the set of 8-bit bytes with decimal values 32 to 126. Table 1 - divides the basic alphabet into subsets. G(x/y) is a notation for the character in position (16 times x) + y in the code table in ISO 8859-1.

NOTE - Table D.1 gives the correspondence between the 8-bit bytes and their graphic representation in ISO 8859-1.

6.3 Exchange structure

The exchange structure shall be a sequential file using a clear text encoding. The exchange structure shall consist of two sections: the header section and the data section. The header section provides data relating to the exchange structure itself. The structure of the header section is specified in clause 9. The data section provides the data to be transferred. The structure of the data section is specified in clause 10. The exchange structure is defined by the WSN in table 3.

The exchange structure is a stream of 8-bit bytes that are encodings of the graphic characters of the basic alphabet. The graphic characters are collected into recognizable sequences called tokens. Tokens may be separated by token separators. The exchange structure can be considered as a sequence of tokens and token separators.

Table 1 - WSN defining subsets of the basic alphabet

```
SPACE     = " " .

DIGIT     = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7"
          | "8" | "9" .

LOWER     = "a" | "b" | "c" | "d" | "e" | "f" | "g" | "h"
          | "i" | "j" | "k" | "l" | "m" | "n" | "o" | "p"
          | "q" | "r" | "s" | "t" | "u" | "v" | "w" | "x"
          | "y" | "z" .

UPPER     = "A" | "B" | "C" | "D" | "E" | "F" | "G" | "H"
          | "I" | "J" | "K" | "L" | "M" | "N" | "O" | "P"
          | "Q" | "R" | "S" | "T" | "U" | "V" | "W" | "X"
          | "Y" | "Z" | "_" .

SPECIAL   = "!" | """" | "*" | "$" | "%" | "&" | "." | "#"
          | "+" | "," | "-" | "(" | ")" | "?" | "/" | ":"
          | ";" | "<" | "=" | ">" | "@" | "[" | "]" | "{"
          | "|" | "}" | "^" | "`" .

REVERSE_SOLIDUS  = "\" .

APOSTROPHE = "'" .

CHARACTER = SPACE | DIGIT | LOWER | UPPER | SPECIAL
          | REVERSE_SOLIDUS | APOSTROPHE
```


6.4 Definition of tokens

The tokens used in the exchange structure are defined by the WSN in table 2.

6.5 WSN of the exchange structure

The syntax of the exchange structure is specified in table 3. Table 3 - references the tokens defined in table 2. The relationship between the syntax and the EXPRESS schema is specified in clause 11.

6.6 Token separators

A token separator is a element that separates two tokens. Token separators are *space*, the explicit print control directives, and comments. A token separator may appear between the terminals or non-terminals of the productions of table 3. Any number of token separators may appear wherever one token separator may appear. A token separator shall not appear within tokens except that explicit print control directives may also appear within binaries and within strings. Print control directives are defined in clause 12.

A comment shall be encoded as a *solidus asterisk* "/*" followed by any number of characters, and terminated by an *asterisk solidus* "*/". Any occurrence of *solidus asterisk* following the first occurrence shall not be significant, i.e. comments cannot be nested. All graphic characters appearing inside a comment shall not be significant to the exchange structure and are only intended to be read by humans.

Table 2 - WSN of token definitions

```
KEYWORD             = USER_DEFINED_KEYWORD ¦ STANDARD_KEYWORD .

USER_DEFINED_KEYWORD = "!" UPPER { UPPER ¦ DIGIT } .

STANDARD_KEYWORD    = UPPER { UPPER ¦ DIGIT } .

SIGN                = "+" ¦ "-".

INTEGER             = [ SIGN ] DIGIT { DIGIT } .

REAL                = [ SIGN ] DIGIT { DIGIT } "." { DIGIT }
                      [ "E" [ SIGN ] DIGIT { DIGIT } ].

NON_Q_CHAR          = SPECIAL ¦ DIGIT ¦ SPACE ¦ LOWER ¦ UPPER .

STRING              = "'" { NON_Q_CHAR ¦
                    APOSTROPHE APOSTROPHE ¦
                    REVERSE_SOLIDUS REVERSE_SOLIDUS ¦
                    CONTROL_DIRECTIVE } "'".

ENTITY_INSTANCE_NAME = "#" DIGIT { DIGIT } .

ENUMERATION         = "." UPPER { UPPER ¦ DIGIT } "." .

HEX                 = "0" ¦ "1" ¦ "2" ¦ "3" ¦ "4" ¦ "5" ¦ "6" ¦ "7" ¦
                      "8" ¦ "9" ¦ "A" ¦ "B" ¦ "C" ¦ "D" ¦ "E" ¦ "F" .

BINARY              = """" ( "0" ¦ "1" ¦ "2" ¦ "3" ) { HEX } """" .
```


6.7 Parameter lists

The PARAMETER_LIST production specified in table 3 is a list of instances of a simple or structured data type.

7 Tokens

In the exchange structure, a token is a special token, a keyword, or a simple data type.

7.1 Special tokens

The special token "ISO-10303-21;" shall be used to open an exchange structure, and the special token "END-ISO-10303-21;" shall be used to close an exchange structure.

The special token "HEADER;" shall be used to open the Header section of an exchange structure, and the special token "ENDSEC;" shall be used to close the Header section of an exchange section.

The special token "DATA;" shall be used to open the Data section of an exchange structure, and the special token "ENDSEC;" shall be used to close the Data section of an exchange structure.