EINECS

European inventory of existing commercial chemical substances

June 1990

Office for Official Publications
of the European Communities

EINECS

European inventory of existing
commercial chemical substances

**Final publication in the
Official Journal of the European Communities**

June 1990

Contents

What is Einecs?

The introductory text as published in Official Journal
C 146 A of 15 June 1990 (90/C 146 A/101)

What is Einecs?

Einecs is the European inventory of existing commercial chemical substances. It has been drawn up by the European Commission in application of Article 13 of Directive 67/548/EEC, as amended by Directive 79/831/EEC, and in accordance with the detailed provisions of Commission Decision 81/437/EEC. It lists and defines those chemical substances which were on the European Community market between 1 January 1971 and 18 September 1981. In terms of Article 1(4) of amended Directive 67/548/EEC, these are substances to which the pre-marketing notification provisions of the Directive do not apply.

Einecs has full legal effect only six months after its publication, in all nine official languages, in the *Official Journal of the European Communities*. From that date it becomes the exclusive point of reference for the identification of substances which are exempt from the pre-marketing notification provisions of amended Directive 67/548/EEC. Meanwhile this Official Journal edition of Einecs has the greatest practical value as a means of identifying 'existing' chemical substances exempted from the pre-marketing notification requirements which have already been obligatory for 'new' substances since the coming into force of the provisions of Directive 79/831/EEC.

The advance edition of Einecs was published exclusively in English and in eight volumes — three of master inventory and five supplementary volumes of indexes (name index, molecular formula index, UVCB index, definition index, plant name index). The Official Journal edition, however, contains nine language versions of the master inventory, but no indexes. Those who possess the advance edition are therefore advised to retain it because of the additional value of its indexes in providing access to the Official Journal edition.

With some minor modifications and changes of presentation, e.g. use of IUPAC nomenclature for the chemical names, the English version of the master inventory in the Official Journal edition is the same as that in the advance edition. These modifications and changes are of course carried over to the eight language versions of the Official Journal edition, which are now published for the first time.

Einecs contains 100 106 entries, about 33 000 derived from the European core inventory (Ecoin) and 67 000 derived from supplementary declarations by industry. Of this number, about 82 000 substances are considered to be well

defined and about 18 000 substances are considered as poorly defined or 'substances of unknown or variable composition, complex reaction products and biological materials' (UVCB substances). However, of these 18 000 UVCBs, approximately 5 000 also have short descriptions in addition to the name.

Substances contained and defined in the master inventory are listed in order of Einecs and CAS (Chemical Abstracts Service) registry number.

The pre-marketing notification requirements of Directive 79/831/EEC

Chemicals have no respect for political boundaries. They are transported through international trade and by natural processes so that, as we now know, the sale and use of a chemical substance in one country may have serious impacts on the population of a neighbouring country, or even the entire globe.

Since the adoption of the 1967 Directive on dangerous substances,[1] the European Community has been developing a body of legislation to provide for the identification and control of risks from commercial chemicals. In the first stage, attention focused on the classification, labelling and control of certain dangerous chemical substances. In 1979, with the adoption of the sixth amendment[2] to the 1967 Directive, the Community moved in a new direction.

The overriding purpose of this new Directive was to establish a single testing and notification system for new chemical substances marketed in the European Community. A chemical substance marketed by the manufacturer or importer for the first time after 18 September 1981 must be notified to the national competent authority at least 45 days in advance.

The notification must contain a dossier of information about the characteristics of the chemical: identity, uses, quantities, disposal and results of physico-chemical, toxicological and ecotoxicological tests.

A summary of this notification dossier is transmitted to the Commission of the European Communities and from there to the other national competent authorities. Further information must be supplied at later dates when certain conditions are met.

---

[1] Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (OJ 196, 16.8.1967, p. 1).

[2] Council Directive 79/831/EEC amending for the sixth time Directive 67/54/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances (OJ L 259, 15.10.1979, p. 10).

By filing a notification for a 'new' chemical substance with the competent authority of a single Member State, a manufacturer or importer may market the chemical substance throughout the entire European Community.

What is a 'new' chemical?

The Directive drew a line between 'new' and 'existing' chemicals by calling for the Commission to provide an inventory of 'existing' chemicals, that is, those which would not be subject to notification because they had been marketed by 18 September 1981.

Hence, *prima facie*, a chemical substance is 'new' and must be notified if it cannot be found in Einecs. But a word of warning may be appropriate. As stated above, Einecs does not have full legal effect, nor can it conclusively identify a substance exempt from pre-marketing notification until six months after its publication in the *Official Journal of the European Communities*. Directive 79/831/EEC specifies in Article 1 (4) that exemption from pre-marketing notification applies until that date simply 'to substances placed on the market before 18 September 1981', and makes no specific reference to Einecs as an authority in the mean time.

Preparation of Einecs

In May 1981, the Commission adopted the procedure for the preparation of Einecs and criteria governing the supply of information for Einecs to the Commission by the Member States.[1]

Einecs is composed of a European core inventory (Ecoin), which was drawn up and published by the Commission in 1982, and a list of substances which are the subject of subsequent declarations communicated to the Commission by the Member States, based on declarations made to them by manufacturers or other persons resident in the European Communities who have marketed these substances.

Its development was divided into five phases, now complete:

1. The preparation of Ecoin and the guidance document 'Reporting for the Einecs inventory'.
2. The submission of declarations by industry to national contact points who forwarded them to the Commission.
3. The examination and processing of the declarations.
4. The publication of the advance edition of Einecs.

---

[1] Commission Decision 81/437/EEC laying down the criteria in accordance with which information relating to the inventory of chemical substances is supplied by the Member States to the Commission (OJ L 167, 24.6.1981, p. 31).

5. The publication of Einecs in the *Official Journal of the European Communities* in the nine official languages.

How to use Einecs

A chemical substance may be located in Einecs by identifying either the Einecs or the CAS (Chemical Abstract Service) number.

Einecs consists of a sequential list of chemicals in ascending order of their Einecs and CAS registry numbers. Each entry of a chemical substance also provides the chemical name, the molecular formula and, where appropriate, the definition of the chemical.

The corresponding CAS number appears on the right on the same line.

The chemical name is presented according to IUPAC nomenclature.

The molecular formula is given when the structure of the chemical is well defined.

A definition is provided in addition to the chemical name where the latter cannot be given precisely, such as for some UVCB substances. It usually contains a description of the process, physical or class terms, genus/species information, SDA (Soap and Detergent Association) or colour index numbers.

Further information

The data contained in Einecs are available, or may become available, in various electronic formats. For further information, contact:

Office for Official Publications of the European Communities,
2 rue Mercier,
L-2985 Luxembourg
Telex: 1324 PUBOF LU
Fax: (352) 49 00 03/49 57 19

Technical specifications of the Formex-like magnetic tapes

The structure of Einecs information

There is one magnetic tape for each of the nine language versions. Any combination of these can be purchased.

Important remark

Although known as 'Formex' the storage format for the data is not totally compatible with the Formex specifications described in the manual published by the Office for Official Publications of the European Communities.

For both technical and economic reasons, it has been decided that a physical block can contain more than one logical record, which is not foreseen in the abovementioned specifications.

A knowledge of the Formex standard is not required in order to process the tape, but an understanding of the SGML (Standard generalized mark-up language) principles will allow optimum use to be made of the data as and when necessary.

(a) Physical structure

The tape contains a group of labels arranged as follows:
VOL1, HDR1, HDR2, HDR3, UHL1, EOF1, EOF2, EOF3
Their physical organization being:
VOL1 HDR1 HDR2 HDR3 UHL1 (TM) .... DATA .... (TM) EOF1 EOF2 EOF3 (TM) (TM)
(TM) = TAPE MARK

Position	Field name	Length	Contents

VOL1: Volume header label

Position	Field name	Length	Contents
1-3	Label indentifier	3	'VOL'
4	Label number	1	'1'
5-10	Volume identifier	6	'EINECS'
11-37	Unused	27	spaces
38-51	Owner identifier	14	4 spaces, 'OPOCE-LX' 2 spaces
52-79	Unused	28	spaces
80	Label standard version	1	Indicates the version of the ISO 1001 standards to which the label and data formats in the volume conform. Currently equal to: '3'

HDR1: First header label

Position	Field name	Length	Contents
1-3	Label identifier	3	'HDR'
4	Label number	1	'1'
5-16	Catalogue number	12	'FXMC90146DAC' DANISH 'FXMC90146DEC' GERMAN 'FXMC90146ENC' ENGLISH 'FXMC90146ESC' SPANISH 'FXMC90146FRC' FRENCH 'FXMC90146GRC' GREEK 'FXMC90146ITC' ITALIAN 'FXMC90146NLC' DUTCH 'FXMC90146PTC' PORTUGUESE
17	Info level	1	'1'
18	Mark-up level	1	'2'
19-27	Unused	9	spaces
28-31	File section number	4	'0001'
32-35	File sequence number	4	'0001', '0002', ..., '0009'
36-39	Generation number	4	'0001'
40-41	Generation version	2	'01'
42-47	Creation date	6	'90166'= 15 June 1990
48-53	Expiration date	6	'99365'= 31 December 1999

9

Position	Field name	Length	Contents
54	Accessibility	1	space (= no restriction)
55-60	Block count	6	'000000'
61-80	Unused	20	spaces

HDR2: second file header label

Position	Field name	Length	Contents
1-3	Label identifier	3	'HDR'
4	Label number	1	'2'
5	Record format	1	'D' (variable length)
6-10	Block length	5	Specifies the maximum number of characters per block. Currently equal to '08192'
11-15	Record length	5	Specifies the maximum record length, including the count field. Currently equal to '08188'
16-80	Unused	65	spaces

HDR3: Third file header label

Position	Field name	Length	Contents
1-3	Label identifier	3	'HDR'
4	Label number	1	'3'
5-12	Owner identification	8	'OPOCE-LX'
13-56	File name	44	The first 17 characters as 5-21 in HDR1, equal spaces
57-80	Unused	24	spaces

UHL1: User file header label

Position	Field name	Length	Contents
1-3	Label identifier	3	'UHL'
4	Label number	1	'1'
5	CCF view indicator	1	'1' = CCF
6	SGML view indicator	1	'1', SGML processing is possible
7	Unused	1	space
8	Info level	1	'1' The file only contains strings of characters, which can be processed strictly sequentially for printing.
9	Mark-up level	1	'2' Text is marked up (basic SGML)
10-80	Unused	71	spaces

Position	Field name	Length	Contents
EOF1: First end of file			
1-3	Label identifier	3	'EOF'
4	Label number	1	'1'
5-54	Same as the corresponding fields in HDR1	50	Same as the corresponding fields in HDR1.
55-60	Block count	6	Numeric characters. Denotes the number of data blocks since the preceding 'beginning-of-file section label group'. This count excludes label blocks and tape-mark blocks.
61-80	Same as the corresponding fields in HDR1.	20	Same as the corresponding fields in HDR1.
EOF2: Second end of file			
1-3	Label identifier	3	'EOF'
4	Label number	1	'2'
5-80	Same as the corresponding fields in HDR2.	76	Same as the corresponding fields in HDR2.
EOF3: Third end of file			
1-3	Label identifier	3	'EOF'
4	Label number	1	'3'
5-80	Same as corresponding fields in HDR3.	76	Same as the corresponding fields in HDR3.

BLOCK STRUCTURE

Storage density is 6250 bpi. All the data, including labels VOL1, HDR1, HDR2, HDR3, UHL1, EOF1, EOF2, EOF3, are encoded in accordance with ISO 2022 with each character encoded on 8 bits (see character set used).

The data zone (located between the labels UHL1 and EOF1) is split into physical blocks of variable size and 8192 (bytes) in length (maximum).

A physical block can contain one or more physical records each of which corresponds to a logical record.

A logical record corresponds to a chemical entry:
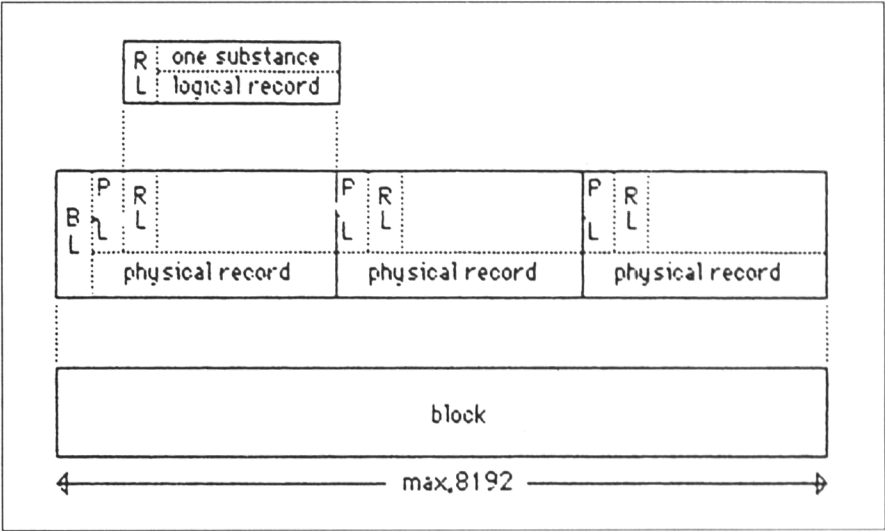
Definition	Representation
BL: Block length	4 numerical characters
PL: Physical record length	4 numerical characters
RL: Record length	5 numerical characters

Remark: RL (record length) = PL (physical record length) — 4 because PL represents the length of the physical record including RL.

A record cannot straddle two physical blocks. Each physical record starts with the RL (record length) identifier followed by a fixed-length label that contains the information necessary to access the data. This label is called 'record label'.

Each data zone must start by the start-of-data identifier, represented by the following character chain:

	0	0	IS_1	0
In hexadecimal	**30**	**30**	**1F**	**30**

Each physical record ends with the character IS_3 (ISO 646 code $1D_{hex}$).

Record label layout

Position	Field name	Length	Contents
1-5	Record length	5	Total length of the record, inclusive of the record label and the record separator. This number is right-justified and zero-padded.
6	Record status	1	'A': New record
7	Unused	1	space
8	Bibliographic level	1	'M': Monograph
9-10	Unused	2	space
11	Indicator length	1	'2'
12	Identifier length	1	'2'
13-17	Base address of data	5	'00024' (record label length)
18-20	Unused	3	spaces
21	Length of 'length-of-data field' in the directory	1	'4'
22	Length of 'length-of-starting character position' in the directory	1	'5'
23	Length of 'implementation-defined part' in the directory	1	'2'
24	Unused	1	space

BLOCK STRUCTURE
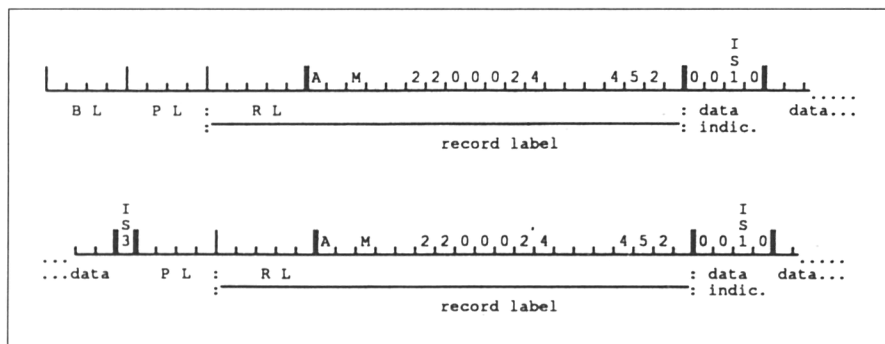
(b) Logical structure

The first Formex records do not contain chemical substance entries. Instead, they contain bibliographical data by SGML markers.

THE BIBLIOGRAPHICAL DATA

Remark: Each line indicated by ' ⇒ ' corresponds to a record on the magnetic tape and is preceded by a 'record label' and a 'data indicator', as previously defined.

\<TI\>   Title.

         EINECS – EUROPEAN INVENTORY OF EXISTING COMMERCIAL CHEMICAL SUBSTANCES

⇒   \<TI\> EINECS – EUROPEAN INVENTORY OF EXISTING COMMER-CIAL CHEMICAL SUBSTANCES.\</TI\>

Note: The LA, TYPE and RANK attributes are ignored.

\<BODY ROLE = ...\>

         Name of corporate body.

         ROLE = 020 Author
         COMMISSION OF THE EUROPEAN COMMUNITIES

         ROLE = 630 Publisher
         OFFICE FOR OFFICIAL PUBLICATIONS OF THE EUROPEAN COMMUNITIES

         ROLE = 260 Copyright holder
         © **ECSC-EEC-EAEC, Brussels — Luxembourg, 1990**

\<NA\>   Name

\<AD\>   Address in use with \<BODY...\>
         L-2985 Luxembourg

⇒   \<BODY ROLE = 20\> \<NA\> COMMISSION OF THE EUROPEAN COMMUNITIES \</NA\> \</BODY\>

⇒   \< BODY ROLE = 630\> \<NA\> OFFICE FOR OFFICIAL PUBLICA-TIONS OF THE EUROPEAN COMMUNITIES \</NA\> \<AD\> L-2985 Luxembourg \</AD\> \</BODY\>

⇒   \<BODY ROLE = 260\> \<NA\> © ECSC-EEC-EAEC, Brussels ● Luxembourg, 1990 \</NA\> \</BODY\>

Note: The RESP attribute is ignored; the \<OT\>, \<CI\> and \<LOC\> markers are ignored.

14

< ED >   Edition statement
                Final edition

⇒   < ED > Final edition, OJ C 146A, 15.6.90 < /ED >

Note: The LA attribute is ignored.

< DATA TYPE = REC >
                Date of the creation of the record
                In this case, the creation date of the first Formex tape.
                15 June 1990.

⇒   < DATE TYPE = REC > 19900615 > /DATE >

Note: The LA and STD attributes are ignored.

< ID SCHEME = ISSN >

                ISSN number   Version

                0378-7001   Danish
                0376-9461   German
                0378-6986   English
                0257-7763   Spanish
                0378-7052   French
                0250-815X   Greek
                0378-701X   Italian
                0378-7079   Dutch
                0257-7771   Portuguese

⇒ < ID SCHEME = ISSN > xxxx-xxxx < /ID >

< TARIFF > Price of the item: printed version, per language

< NOTE >   Information about item (info about tariff)
                Price (excluding VAT) in Luxembourg

< PRICE CURRENCY = ECU >
                Price of the item in ECU
                228

< PRICE CURRENCY = BFR >
                Price of the item in Belgian francs (approximative)
                9640

15

⇒   < TARIFF >   < NOTE > Price (excluding VAT) in Luxembourg
         < /NOTE > < PRICE   CURRENCY = ECU > 228 < /PRICE > < PRICE
         CURRENCY = BFR > 9640 < /PRICE > < /TARIFF

Note: The LA, STD, TYPE (NOTE) attributes are ignored.

THE CHEMICAL SUBSTANCE ENTRIES

An 'Einecs' chemical substance entry is composed of various items of informa-
tion describing the substance. This information may be any of the following:

(i)   the Einecs number: 9 positions with unique identifier (xxx-xxx-x);

(ii)   the CAS number: 11 positions with unique identifier (xxxxxx-xx-x);

(iii) the 'Chemical substance' name;

(iv)   the molecular formula, if any;

(v)   possibly a definition description generally used where no molecular formula
       exists (a descriptive text on the substance) called 'substance definition
       description';

(vi)   possibly one or more items of 'column one' information where a descrip-
       tion is available;

(vii) possibly one or more items of 'column two' information where a description
       is available.

Each piece of information can be identified, within a notice, by the following
SGML markers:

Einecs number	< EINECS >
CAS number	< CAS >
Chemical substance name	< CSN >
Molecular formula	< MF >
Substance definition description	< DD >
Substance definition column text one	< C1 >
Substance definition column text two	< C2 >

Note: The < /... > SGML markers do not appear. (Shorttag, Omittag)
These markers can only be present in the above order, except < C1 > and < C2 > that
normally alternate.
(< C1 > data < C2 > data < C1 > data < C2 > data ...)

As a rule, the absence of the < MF > marker necessitates the presense of the
< DD > , < C1 > and < C2 > markers. The markers < C1 > or < C2 > do not
always appear.

The standard cases are:

    < EINECS > ... < CAS > ... < CSN > ... < MF > ...

```
<EINECS> ... <CAS> ... <CSN> ... <DD> ...
<EINECS> ... <CAS> ... <CSN> ... <DD> ... <C1> ... <C2> ...
<C1> ...  <C2> ...
```

The following exceptions to the normal rules have been identified:

1. Einecs number 266-046-0 has the following structure:
   ```
   <EINECS> ... <CAS> ... <CSN> ... <DD> ... <C1> ... <C2 > ...
   <C1> ... <C2> ... <C2> ...
   ```

2. Einecs numbers 232-336-0 and 232-340-2 have the following structure:
   ```
   <EINECS> ... <CAS> ... <CSN> ... <DD> ...<C1> ... <C2>
   ... <C1>
   ```

EXTRACT FROM THE INVENTORY

(1) **266-064-9**   (2) **66009-14-1**
(3)   22-34-Gastrin I (pig), 22-ı-leucine-32-ı-leucine-
(4) $C_{75}H_{101}N_{16}O_{26}$

        266-068-0   66016-77-1
           2-chloro-a-phenyl[1,1'-biphenyl]-4-methanol $C_{19}H_{16}ClO$

        266-072-2   **66027-93-8**
           indium(3+) trisulphamate $H_3NO_3S.^1/_3In$

        266-043-4   65997-15-1
           Cement, portland, chemicals
(5)   Portland cement is a mixture of chemical substances produced by
           burning or sintering at high temperatures (greater than 1200°C
           (2192°F)) raw materials which are predominantly calcium
           carbonate, aluminium oxide, silica, and iron oxide. The
           chemical substances which are manufactured are confined in a
           crystalline mass. This category includes all of the chemical
           substances specified below when they are intentionally
           manufactured in the production of Portland cement. The
           primary members of the category are Ca_2SiO_4 and Ca_3SiO_5.
           Other compounds listed below may also be included in
           combination with these primary substances.

           (6) $CaAl_2O_4$ (7) $Ca_2Al_2SiO_7$
                 $CaAl_4O_7$ $Ca_4Al_6SO_{16}$
                 $CaAl_{12}O_{19}$ $Ca_{12}Al_{14}Cl_2O_{32}$
                 $Ca_3Al_2O_6$ $Ca_{12}Al_{14}F_2O_{32}$
                 $Ca_{12}Al_{14}O_{33}$ $Ca_4Al_2Fe_2O_{10}$
                 CaO $Ca_6Al_4Fe_2O_{15}$
                 $Ca_2Fe_2O_5$

(1) Einecs number   <EINECS>
(2) CAS number   <CAS>
(3) Chemical substance name   <CSN>
(4) Molecular formula   <MF>
(5) Substance definition description   <DD>
(6) Substance definition column text one <C1>
(7) Substance definition column text two <C2>

The 'Micro-Formex' character set

The character set uses the ISO 2022 structure, including the loading of a number of character tables.

When the magnetic tape was created a simplification facility was incorporated and some tables can be considered as loaded in memory.

The escape sequences for the calls to the tables G0, G1, C0 and C1 (e.g. ESC 2/8 4/0, ESC 2/10 6/12, etc.), are implicit and are not included on the tape.

Tables G0, G1, G2, C0 and C1 are default tables.

The access to the G2 table is done by temporary load (character by character), using the escape sequence SS_2 (single shift) of the C1 table, followed by the desired character of the G2 table.

Example: The $-$ (minus) sign of table G2 is encoded, on this tape, as follows:

$8E_{hex}$ 50_{hex}, same as SS_2 50_{hex}

The 'amno-6' character chain is encoded (in hexadecimal):

a	m	n	o	SS_2	$-$	6
61	**6D**	**6E**	**6F**	**8E**	**50**	**36**

(see the Formex documentation)

Certain special characters, not specified in the Formex tables, are represented by logical entities. These entities are:

Graphical presentation	Description	Hexadecimal representation	Entity
$\geqslant$	Equal or greater-than sign	26 45 47 54 3B	&EGT;
$\leqslant$	Equal or less-than sign	26 45 4C 54 3B	&ELT;

Remarks

(i)   The - (hyphen) sign found in Label VOL1 and as a separator in both Einecs and CAS numbers (66009-14-1) is encoded $2D_{hex}$

(ii) The $-$ (minus) sign found in chemical names (22-L-Leucine-32...) is represented by the character of Table G2 (e.g. $8E_{hex}$, 50_{hex})

(iii) The Greek 'beta' and the German '(SZ)' characters are both represented by the same code (e.g. $E2_{hex}$)

(iv) Although the labels are standard (ISO 1001), certain difficulties may arise when reading this tape given the fact that many mainframe systems use EBCDIC as a standard for character representation and not ISO 2022.

If that is the case, the user should go around the problem by using BTAM principles (Basic tape access method). BTAM's macro-instructions are normally available in programming module libraries.

(a) Printing markers

The information on the magnetic tape is expanded for printing by the addition of markers specially created for the purpose.

Note: The SGML does not use physical attributes but only logical ones. The transcoding carried out in order to store the information on the tape was based on fully-accentuated characters which prevented the SGML logical markers, such as valency, for example, from being included.

The printing markers are as follows:

— roman   no marking
— return to roman $<@\backslash>$
— roman subscript $<@L>$
— roman subscript $<@S>$
— roman subscript $<@H>$
— italic $<@I>$
— italic superscript $<@J>$
— roman small capital $<@C>$

Roman type is the default type.

Every marker applies to the characters following it until another marker is encountered or the cancellation marker $<@\backslash>$ appears.

Any marker cancels any previous markers' effects.

Example:

$C_6 H_{22} O_{11}$ 22-L Leucine 32

and

C $<@L>$ 6 $<@\backslash>$ H $<@L>$ 22 $<@\backslash>$ 0 $<@L>$ 11 $<@\backslash>$ 22- $<@C>$ L $<@\backslash>$ - Leucine 32

are equivalent.

(b) Formex documentation

(i) Introduction

As the data processed in the frame of Micro-Formex belong to a multilingual environment, the usual 8-bit character sets are insufficient to cover the various required combinations. As such sets only provide 256 positions (2^8), another technique must be used to cover the actual needs. The International Standards Organization (ISO) has been busy with these problems for a long time and therefore, in order to provide the maximum flexibility and compatibility, it has been decided to follow the standards defined by this organization.

Formex publishes documents supported by nearly all computer manufacturers.

Only standards in the area of character sets and information coding will be given here, for the purpose of these specifications.

These standards are, obviously, modified from time to time in order to keep them as up to date as possible and to take into account the varying needs of all member countries (or associations, such as ECMA). Therefore the present specifications are also subject to slight modifications, but they will only be notified and applied once the modifications are stable enough to be taken over to Formex.

The following sections describe the items of ISO as they apply to Micro-Formex.

The main items used for that purpose can be found in the following list:

ISO 646,
ISO 2022,
ISO 6937/2,
ISO 6937/3,
ISO 6937/6,
ISO 6937/7.

(ii) The principles

As explained in the previous section, standards were defined to give a precise definition of how data should best be coded and interchanged between computers using 7- or 8-bit character sets.

ISO 646 defines an international reference version of a widely used character set; even by other normalization organizations.

20

The ISO 2022 norm gives the rules to be followed in data interchange that cover sets holding more than 128 or 256 positions (respectively in 7- and 8-bit environments); it is the cornerstone of all interchanges covered by Formex.

The idea of ISO 2022 can be explained as follows:

☐ The basic character set (256 positions will be assumed, i.e. 8-bit combinations) can be divided in two areas each holding 128 positions (this allows compatibility with the 7-bit combinations). This subdivision can be seen in Figure 1.

☐ Each of these areas is divided like a matrix in 8 columns and 16 rows giving the actual entries of the area.

☐ In addition, each area is subdivided into two parts, concerning respectively: control functions, on the one hand, and actual character combinations on the other (Figure 1).

☐ The control function part is defined by the first two columns of each part and the character combinations are located in the next six columns; each part including the 16 defined rows.

On that basis, ISO 2022 allows an almost unlimited number of sets to be used by defining the mechanisms for replacing any of the above described parts by another. For doing this, the user needs two tools to define what he wants to replace and how he would like it to happen.

☐ A lot of control function parts and character combination parts were pre-defined by several norms allowing the invocation of these components by using a part identifier uniquely defined and registered.

☐ The way a new control function or character combination part can be invoked has also been defined: by using special pre-defined sequences according to where the user needs the set.

An example may be useful, at this stage, to show these principles:

☐ A text respecting ISO 2022 could, for instance, include several characters not present in the table (or part) covering the rest of the text. In such a case, the norm states that it is possible to define, in the flow of data, a sequence recognized as a call to another part (or table) including, this time, the characters missing in the first one used. All that is needed, therefore, is the identification of the table containing the missing characters and the pre-defined sequence allowing interruption of the use of the current table and its replacement by the new one.

The norm defines two possible substitutions for a part by another:

● first, a *long-term* substitution that makes a new part active until it is explicitly suppressed by another replacement, and

● second, a *short-term* substitution that applies only to the character directly following, including a complete return to the original part for the second-next character.

The abovementioned principles are general to cover all situations one can meet in the concerned area. The application of these rules in the frame of Micro-Formex represents, in fact, a subset of the norm. This subset completely satisfies the needs of communicating and interchanging data concerning official publications of the European Communities. It has been chosen to limit somewhat the complexity a user may feel when viewing the complete norm. Nevertheless, the complete procedure is described in Figure 1, showing:

- the 8-bit set used by the process,
- the splitting of the set in two areas,
- the splitting of each area in two parts,
- the possibility of replacing any part of the set by another,
- the existence of several repertoires (control function and graphic character repertoires),
- the invocation of a given graphic set (the actual substitution call),
- the designation or identification of, possibly, several repertoires.

The ISO 2022 norm defining table substitutions is completed by ISO 6937 describing the use, in the frame of ISO 2022, of several specific set: ISO 6937/2 concerns Latin and Latin-supplementary sets, ISO 6937/3 is related to the two control function sets: ISO 6937/6 describes the use of some special graphic characters, and ISO 6937/7 indicates how to process a Greek set.

The following restrictions or limitations to the norm ISO 2022 should be noted:

- only repertoires shown in white in Figure 1 will be used in Micro-Formex,
- the graphic sets G0, G1, G2 and G3 are pre-loaded and no other set should be used,
- the sets G0 and G1 are even pre-installed in the 8-bit set in use: NO invocation,
- the sets G2 and G3 may only be invoked by short-term substitutions: SS_2 and SS_3,
- no additional control function repertoire is used.

Taking these restrictions into account, the sets used, in the frame of Micro-Formex, may be called C0, C1, G0, G1, G2 and G3 if the user keeps in mind that this designation of control function and graphic sets has a locally defined interpretation.

The various sets will be completely described in the next section.
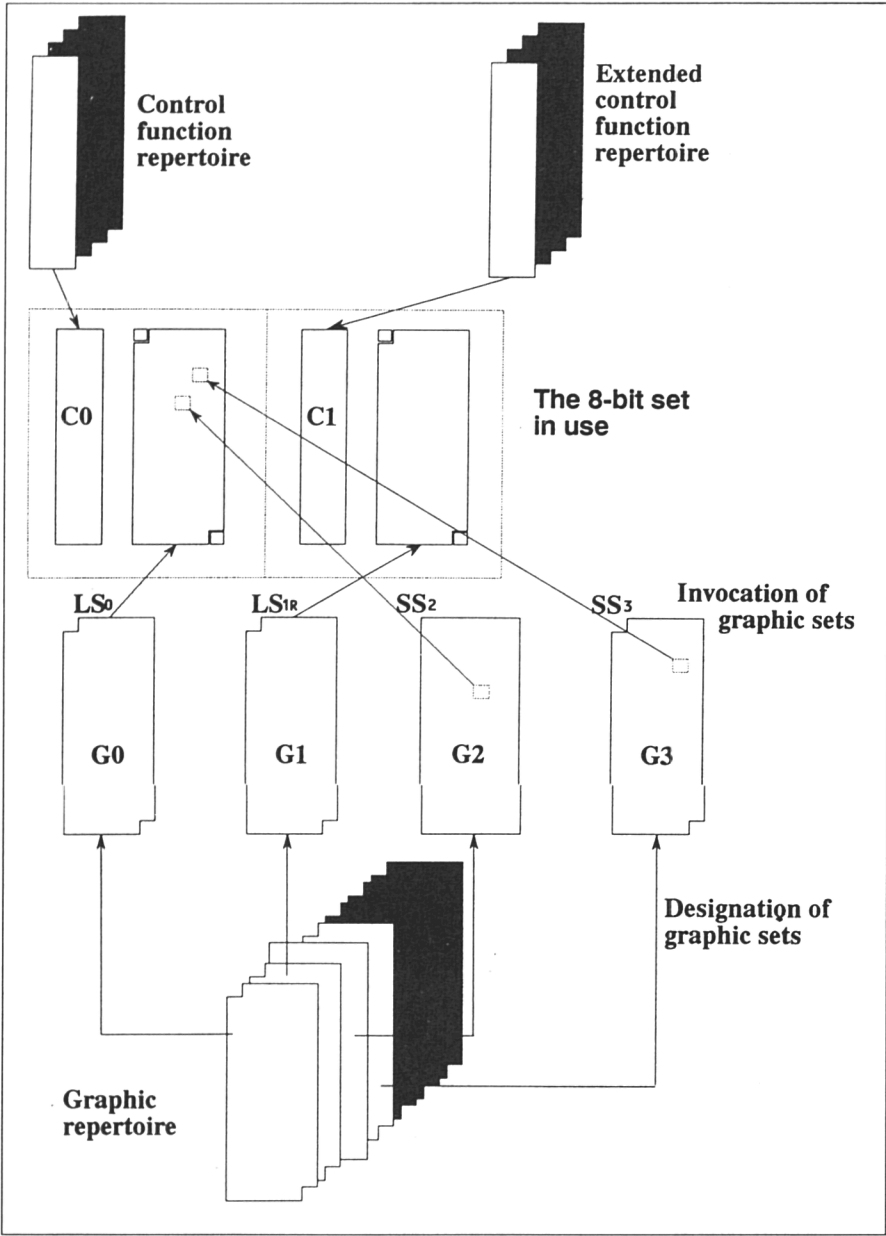
Figure 1: Set substitution principles

(iii) The sets in use under Micro-Formex

As stated in the previous section, all permitted bit-combinations belong to one of the six sets called C0, C1, G0, G1, G2 and G3 (see previous section for the meaning of this convention). The sets are described in this section (points (a) to (f)).

For each set, the following information will be provided:

- a list, sorted on hexadecimal representation, of all elements appearing in the set showing, for each character:
  - the identifier of the character,
  - the hexadecimal representation of it,
  - the literal definition of the entry;
- a table, given for illustration purposes only, grouping all elements in a matrix.

Notes:

1. The identifiers given here are used for reference only as the trend, in the ISO norms, seems to be their slow withdrawal. Nevertheless, the following rules apply for the present specifications:

— each identifier has six positions beginning with < and ending with >, leaving four actual identification positions;

— the first identification position shows the type of character according to the list:

      C   for control characters,
      G   for Greek characters,
      L   for Latin characters,
      N   for numbers,
      S   for signs;

— the second identification position defines the sub-type of characters for each type:

   within type **C**:
      E   for *extension*,
      F   for *format*,
      M   for *miscellaneous*,
      S   for *separator;*
   within types **G** and **L**:
      the actual presented letter or the transliteration;
   within type **N**:
      D   for *digit*,
      F   for *fraction*,
      S   for *superscript;*
   within type **S**:
      A   for *arithmetic*,
      B   for *block design*,
      C   for *currency*,

D for *diacritical,*
G for *graphic,*
M for *miscellaneous,*
P for *punctuation;*

— the positions three and four of the identification positions are used to distinguish several variations on a given character (accented versions, fractions, etc.).

2. Some table positions are not available in several sets (limited set or reserved for future use positions): they appear with a grey background in the tables.

(a) The control function set C0

This set is described in ISO 646 and 6937/3 and can be viewed, in the frame of Micro-Formex, as showing the following elements:

< CF10 >	00/08	backspace
< CF11 >	00/09	horizontal tabulation
< CF12 >	00/10	line feed
< CF13 >	00/11	vertical tabulation
< CF14 >	00/12	form feed
< CF15 >	00/13	carriage return
< CE08 >	00/14	locking shift 1
< CE07 >	00/15	locking shift 0
< CM01 >	01/00	data link escape
< CM02 >	01/10	substitute character
< CE03 >	01/11	escape
< CS03 >	01/13	information separator 3
< CS02 >	01/14	information separator 2
< CS01 >	01/15	information separator 1

Its table representation is, according to this list, as shown in Figure 2. Several characters were defined in the C0 set that are not used at present in the frame of Micro-Formex; therefore they are shown on a hachured background.

(b) The control function set C1

This set is described in ISO 6937/3 and shows, in the frame of Micro-Formex, the following elements:

< CF16 >	08/11	partial line down
< CF17 >	08/12	partial line up
< CE04 >	08/14	single shift 2
< CE05 >	08/15	single shift 3
< CE06 >	09/11	control sequence introducer

Its table representation is, according to this list, as shown in Figure 3,

	0	1
0	NUL	DLE
1	SOH	DC_1
2	STX	DC_2
3	ETX	DC_3
4	EOT	DC_4
5	ENQ	NAK
6	ACK	SYN
7	BEL	ETB
8	BS	CAN
9	HT	EM
A	LF	SUB
B	VT	ESC
C	FF	IS_4
D	CR	IS_3
E	LS_1	IS_2
F	LS_0	IS_1

Figure 2: The C0 set

	8	9
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
A		
B	PLD	CSI
C	PLU	
D		
E	SS_2	
F	SS_3	

Figure 3: The C1 set

(c) The graphic set G0

This set is described in ISO 646 and 6937/2 and can be viewed, in the frame of Micro-Formex, as showing the following elements:

< SP01 >	02/00	space
< SP02 >	02/01	exclamation mark
< SP04 >	02/02	quotation mark
< SM01 >	02/03	number sign
< SC01 >	02/04	general currency sign
< SM02 >	02/05	per cent sign
< SM03 >	02/06	ampersand
< SP05 >	02/07	apostrophe
< SP06 >	02/08	left parenthesis
< SP07 >	02/09	right parenthesis
< SM04 >	02/10	asterisk
< SA01 >	02/11	plus sign
< SP08 >	02/12	comma
< SM12 >	02/13	horizontal bar
< SP11 >	02/14	full stop, period
< SP12 >	02/15	solidus (slash, stroke)
< ND10 >	03/00	digit zero
< ND01 >	03/01	digit one
< ND02 >	03/02	digit two
< ND03 >	03/03	digit three
< ND04 >	03/04	digit four
< ND05 >	03/05	digit five
< ND06 >	03/06	digit six
< ND07 >	03/07	digit seven
< ND08 >	03/08	digit eight
< ND09 >	03/09	digit nine
< SP13 >	03/10	colon
< SP14 >	03/11	semicolon
< SA03 >	03/12	less-than sign
< SA04 >	03/13	equals sign
< SA05 >	03/14	greater-than sign
< SP15 >	03/15	question mark
< SM05 >	04/00	commercial 'a'
< LA02 >	04/01	capital A
< LB02 >	04/02	capital B
< LC02 >	04/03	capital C
< LD02 >	04/04	capital D
< LE02 >	04/05	capital E
< LF02 >	04/06	capital F
< LG02 >	04/07	capital G
< LH02 >	04/08	capital H
< LI02 >	04/09	capital I

< LJ02 >	04/10	capital J
< LK02 >	04/11	capital K
< LL02 >	04/12	capital L
< LM02 >	04/13	capital M
< LN02 >	04/14	capital N
< LO02 >	04/15	capital O
< LP02 >	05/00	capital P
< LQ02 >	05/01	capital Q
< LR02 >	05/02	capital R
< LS02 >	05/03	capital S
< LT02 >	05/04	capital T
< LU02 >	05/05	capital U
< LV02 >	05/06	capital V
< LW02 >	05/07	capital W
< LX02 >	05/08	capital X
< LY02 >	05/09	capital Y
< LZ02 >	05/10	capital Z
< SM06 >	05/11	left square bracket
< SM07 >	05/12	reverse solidus (slash, stroke)
< SM08 >	05/13	right square bracket
< SD15 >	05/14	circumflex accent
< SP09 >	05/15	low line
< SD13 >	06/00	grave accent
< LA01 >	06/01	small a
< LB01 >	06/02	small b
< LC01 >	06/03	small c
< LD01 >	06/04	small d
< LE01 >	06/05	small e
< LF01 >	06/06	small f
< LG01 >	06/07	small g
< LH01 >	06/08	small h
< LI01 >	06/09	small i
< LJ01 >	06/10	small j
< LK01 >	06/11	small k
< LL01 >	06/12	small l
< LM01 >	06/13	small m
< LN01 >	06/14	small n
< LO01 >	06/15	small o
< LP01 >	07/00	small p
< LQ01 >	07/01	small q
< LR01 >	07/02	small r
< LS01 >	07/03	small s
< LT01 >	07/04	small t
< LU01 >	07/05	small u
< LV01 >	07/06	small v
< LW01 >	07/07	small w
< LX01 >	07/08	small x

< LY01 >	07/09	small y
< LZ01 >	07/10	small z
< SM11 >	07/11	left curly bracket (brace)
< SM13 >	07/12	vertical line
< SM14 >	07/13	right curly bracket (brace)
< SD19 >	07/14	tilde accent
< CM03 >	07/15	delete

Its table representation is, according to this list, as shown in Figure 4.

The reader should note that the G0 set is a 96-position set, including both hexadecimal positions 02/00 and 07/15.

This set may also be used to mention the existence of so-called diacritical characters that enable the creation of any accented character by using the sequence:

diacritical sign   actual character

avoiding the use of backspace, or any other trick.

(d) The graphic set G1

This set is described in ISO 6937/7 and can be viewed, in the frame of Micro-Formex, as showing the following elements:

< SD51 >	11/04	accent ('tonos')
< SD53 >	11/05	accent plus diaeresis
< GA52 >	11/06	capital alpha with accent
< GE52 >	11/08	capital epsilon with accent
< GG52 >	11/09	capital eta with accent
< GI52 >	11/10	capital iota with accent
< GP52 >	11/12	capital omicron with accent
< GU52 >	11/14	capital upsilon with accent
< GZ52 >	11/15	capital omega with accent
< GI53 >	12/00	small iota with accent plus diaeresis
< GA02 >	12/01	capital alpha
< GB02 >	12/02	capital beta
< GC02 >	12/03	capital gamma
< GD02 >	12/04	capital delta
< GE02 >	12/05	capital epsilon
< GF02 >	12/06	capital zeta
< GG02 >	12/07	capital eta
< GH02 >	12/08	capital theta
< GI02 >	12/09	capital iota
< GK02 >	12/10	capital kappa
< GL02 >	12/11	capital lambda
< GM02 >	12/12	capital mu
< GN02 >	12/13	capital nu

	0	1	2	3	4	5	6	7
0			SPACE	0	@	P	`	p
1			!	1	A	Q	a	q
2			"	2	B	R	b	r
3			#	3	C	S	c	s
4			¤	4	D	T	d	t
5			%	5	E	U	e	u
6			&	6	F	V	f	v
7			'	7	G	W	g	w
8			(	8	H	X	h	x
9			)	9	I	Y	i	y
A			*	:	J	Z	j	z
B			+	;	K	[	k	{
C			,	<	L	\	l	|
D			—	=	M	]	m	}
E			.	>	N	^	n	~
F			/	?	O	_	o	DEL

Figure 4: The graphic set G0

<GO02>	12/14	capital xi
<GP02>	12/15	capital omicron
<GQ02>	13/00	capital pi
<GR02>	13/01	capital rho
<GS02>	13/03	capital sigma
<GT02>	13/04	capital tau
<GU02>	13/05	capital upsilon
<GV02>	13/06	capital phi
<GX02>	13/07	capital chi
<GY02>	13/08	capital psi
<GZ02>	13/09	capital omega
<GI18>	13/10	capital iota with diaeresis
<GU18>	13/11	capital upsilon with diaeresis
<GA51>	13/12	small alpha with accent
<GE51>	13/13	small epsilon with accent
<GG51>	13/14	small eta with accent
<GI51>	13/15	small iota with accent
<GU53>	14/00	small upsilon with accent plus diaeresis
<GA01>	14/01	small alpha
<GB01>	14/02	small beta
<GC01>	14/03	small gamma
<GD01>	14/04	small delta
<GE01>	14/05	small epsilon
<GF01>	14/06	small zeta
<GG01>	14/07	small eta
<GH01>	14/08	small theta
<GI01>	14/09	small iota
<GK01>	14/10	small kappa
<GL01>	14/11	small lambda
<GM01>	14/12	small mu
<GN01>	14/13	small nu
<GO01>	14/14	small xi
<GP01>	14/15	small omicron
<GQ01>	15/00	small pi
<GR01>	15/01	small rho
<GS03>	15/02	small final sigma
<GS01>	15/03	small sigma
<GT01>	15/04	small tau
<GU01>	15/05	small upsilon
<GV01>	15/06	small phi
<GX01>	15/07	small chi
<GY01>	15/08	small psi
<GZ01>	15/09	small omega
<GI17>	15/10	small iota with diaeresis
<GU17>	15/11	small upsilon with diaeresis
<GP51>	15/12	small omicron with accent
<GU51>	15/13	small upsilon with accent

< GZ51 >	15/14	small omega with accent

Its table representation is, according to this list, as shown in Figure 5.

It should be noted that the G1 set is a 94-position set, excluding both hexadecimal positions 02/00 and 07/15.

This set shows a variation on the accented characters subject: as few letters are accented in Greek and as an own set has been created, the most frequently accented letters were introduced into this set allowing direct use if the required character exists already accented in the set.

(e) The graphic set G2

This set is described in ISO 6937/2 and can be viewed, in the frame of Micro-Formex, as showing the following elements:

< SP31 >	02/00	no-break space
< SP03 >	02/01	inverted exclamation mark
< SC04 >	02/02	cent sign
< SC02 >	02/03	pound sign
< SC03 >	02/04	dollar sign
< SC05 >	02/05	yen sign
< SM24 >	02/07	section sign
< SP19 >	02/09	single quotation mark left
< SP21 >	02/10	double quotation mark left
< SP17 >	02/11	angle quotation mark left (guillemet)
< SM30 >	02/12	leftward arrow
< SM32 >	02/13	upward arrow
< SM31 >	02/14	rightward arrow
< SM33 >	02/15	downward arrow
< SM19 >	03/00	degree sign
< SA02 >	03/01	plus or minus sign
< NS02 >	03/02	superscript two
< NS03 >	03/03	superscript three
< SA07 >	03/04	multiplication sign
< SM17 >	03/05	micro sign
< SM25 >	03/06	pilcrow (paragraph sign)
< SM26 >	03/07	middle dot
< SA06 >	03/08	division sign
< SP20 >	03/09	single quotation mark right
< SP22 >	03/10	double quotation mark right
< SP18 >	03/11	angle quotation mark right (guillemet)
< NF14 >	03/12	fraction: one-quarter
< NF12 >	03/13	fraction: one-half
< NF34 >	03/14	fraction: three-quarters
< SP16 >	03/15	inverted question mark

	8	9	A	B	C	D	E	F
0					ï·	Π	ü·	π
1					A	Ρ	α	ρ
2					B		β	ς
3					Γ	Σ	γ	σ
4				'	Δ	Τ	δ	τ
5				·'·	E	Ͳ	ε	υ
6				'Α	Z	Φ	ζ	φ
7				H	X	η	χ
8				'Ε	Θ	Ψ	θ	ψ
9				'Η	I	Ω	ι	ω
A				'Ι	K	Ϊ·	κ	ï
B				Λ	Ͳ	λ	ü
C				'Ο	M	ά	μ	ό
D				N	έ	ν	ύ
E				'Ͳ	Ξ	ή	ξ	ώ
F				'Ω	O	ί	o	

Figure 5: The graphic set G1

<SD13>	04/01	grave accent
<SD11>	04/02	acute accent
<SD15>	04/03	circumflex accent
<SD19>	04/04	tilde accent
<SD31>	04/05	macron
<SD23>	04/06	breve
<SD29>	04/07	dot above
<SD17>	04/08	diaeresis or umlaut accent
<SD27>	04/10	ring
<SD41>	04/11	cedilla
<SD33>	04/12	underline
<SD25>	04/13	double acute accent
<SD43>	04/14	ogonek
<SD21>	04/15	caron
<SA08>	05/00	minus sign
<NS01>	05/01	superscript one
<SM53>	05/02	registered sign
<SM52>	05/03	copyright sign
<SM54>	05/04	trade mark sign
<SM39>	05/05	music note
<SM66>	05/06	not sign
<SM65>	05/07	broken bar
<NF18>	05/12	fraction: one-eighth
<NF38>	05/13	fraction: three-eighths
<NF58>	05/14	fraction: five-eighths
<NF78>	05/15	fraction: seven-eighths
<SM18>	06/00	ohm sign
<LA52>	06/01	capital Æ diphthong
<LD62>	06/02	capital D with stroke
<SM21>	06/03	ordinal indicator, feminine
<LH62>	06/04	capital H with stroke
<LI52>	06/06	capital IJ ligature
<LL64>	06/07	capital L with middle dot
<LL62>	06/08	capital L with stroke
<LO62>	06/09	capital O with slash
<LO52>	06/10	capital Œ ligature
<SM20>	06/11	ordinal indicator, masculine
<LT63>	06/12	small thorn, Icelandic
<LT62>	06/13	capital T with stroke
<LN62>	06/14	capital eng, Lapp
<LN63>	06/15	small n with apostrophe
<LK61>	07/00	small k, Greenlandic
<LA51>	07/01	small æ diphthong
<LD61>	07/02	small d with stroke
<LD63>	07/03	small eth, Icelandic
<LH61>	07/04	small h with stroke
<LI61>	07/05	small i without dot

< LI51 >	07/06	small ij ligature
< LL63 >	07/07	small l with middle dot
< LL61 >	07/08	small l with stroke
< LO61 >	07/09	small o with slash
< LO51 >	07/10	small œ ligature
< LS03 >	07/11	small sharp s, German
< LT64 >	07/12	capital thorn, Icelandic
< LT61 >	07/13	small t with stroke
< LN61 >	07/14	small eng, Lapp
< SP32 >	07/15	soft hyphen

Its table representation is, according to this list, as shown in Figure 6.

It can be seen that the G2 set is a 96-position set, including both hexadecimal positions 02/00 and 07/15.
Moreover, the diacritical signs were grouped into column 04/xx with two positions reserved for future use (04/00 and 04/09).

(f) The graphic set G3

This set is described in ISO 6937/6 and can be viewed, in the frame of Micro-Formex, as showing the following elements:

< SP33 >	02/01	em space
< SP34 >	02/02	en space
< SP35 >	02/03	3/em space
< SP36 >	02/04	4/em space
< SP37 >	02/05	digit space
< SP38 >	02/06	punctuation space
< SP39 >	02/07	thin space
< SP40 >	02/08	hair space
< SP52 >	02/09	em dash
< SP53 >	02/10	en dash
< SP10 >	02/11	hyphen
< SP54 >	02/12	significant blank symbol
< SP50 >	02/14	ellipsis
< SP51 >	02/15	double baseline dot
< NF13 >	03/00	fraction: one-third
< NF23 >	03/01	fraction: two-thirds
< NF15 >	03/02	fraction: one-fifth
< NF25 >	03/03	fraction: two-fifths
< NF35 >	03/04	fraction: three-fifths
< NF45 >	03/05	fraction: four-fifths
< NF16 >	03/06	fraction: one-sixth
< NF56 >	03/07	fraction: five-sixths
< SM55 >	03/08	care-of sign
< SB40 >	03/09	full block

	0	1	2	3	4	5	6	7
0			NBSP	°		–	Ω	ĸ
1			¡	±	`	¹	Æ	æ
2			¢	²	´	®	Đ	đ
3			£	³	˜	©	ª	ð
4			$	×	~	™	Ħ	ħ
5			¥	µ	‾	·		ı
6				¶	˘	¬	IJ	ij
7			§	■	■		Ŀ	ŀ
8				÷	¨		Ł	ł
9			'	'			Ø	ø
A			"	"	°		Œ	œ
B			«	»	¸		º	ß
C			←	¼	_	⅛	Þ	þ
D			↑	½	″	⅜	Ŧ	ŧ
E			→	¾	˛	⅝	Ŋ	ŋ
F			↓	¿	ˇ	⅞	'n	SHY

Figure 6: The graphic set G2

< SB41 >	03/10	upper half block
< SB42 >	03/11	lower half block
< SB50 >	03/12	25% shading
< SB51 >	03/13	50% shading
< SB53 >	03/14	75% shading
< SM67 >	03/15	marker
< SB01 >	04/00	light horizontal
< SB02 >	04/01	light vertical
< SB13 >	04/02	light top right corner
< SB15 >	04/03	light top left corner
< SB09 >	04/04	light bottom left corner
< SB07 >	04/05	light bottom right corner
< SB10 >	04/06	light vertical with light join right
< SB14 >	04/07	light horizontal with light join above
< SB12 >	04/08	light vertical with light join left
< SB08 >	04/09	light horizontal with light join below
< SB11 >	04/10	light vertical with light horizontal
< SB26 >	04/11	light vertical with heavy join right
< SB28 >	04/12	light horizontal with heavy jcin above
< SB27 >	04/13	light vertical with heavy join left
< SB25 >	04/14	light horizontal with heavy jcin below
< SB03 >	04/15	light vertical with heavy horizontal
< SB04 >	05/00	heavy horizontal
< SB05 >	05/01	heavy vertical
< SB22 >	05/02	heavy top right corner
< SB24 >	05/03	heavy top left corner
< SB18 >	05/04	heavy bottom left corner
< SB16 >	05/05	heavy bottom right corner
< SB19 >	05/06	heavy vertical with heavy join right
< SB23 >	05/07	heavy horizontal with heavy join above
< SB21 >	05/08	heavy vertical with heavy join left
< SB17 >	05/09	heavy horizontal with heavy join below
< SB20 >	05/10	heavy vertical with heavy horizontal
< SB30 >	05/11	heavy vertical with light join right
< SB32 >	05/12	heavy horizontal with light join above
< SB31 >	05/13	heavy vertical with light join left
< Sb29 >	05/14	heavy horizontal with light join below
< SB06 >	05/15	heavy vertical with light horizontal
< SG11 >	06/00	circle
< SG12 >	06/01	square
< SG13 >	06/02	rectangle
< SG14 >	06/03	triangle up
< SG15 >	06/04	triangle down
< SG16 >	06/05	star
< SG21 >	06/06	round bullet
< SG22 >	06/07	square bullet
< SG24 >	06/08	triangle up bullet

<SG25>	06/09	triangle down bullet
<SM77>	06/10	left pointer
<SM78>	06/11	right pointer
<SG01>	06/12	club
<SG02>	06/13	diamond
<SG03>	06/14	heart
<SG04>	06/15	spade
<SG51>	07/00	Maltese cross
<SG52>	07/01	dagger
<SG53>	07/02	double dagger
<SM84>	07/03	tick, check mark
<SM85>	07/04	ballot cross
<SM91>	07/05	musical sharp
<SM92>	07/06	musical flat
<SG31>	07/07	male symbol
<SG32>	07/08	female symbol
<SG41>	07/09	telephone symbol
<SG42>	07/10	telephone recorder symbol
<SG43>	07/11	phonograph copyright sign
<SM87<	07/12	caret
<SM88>	07/13	single low quotation mark
<SM89>	07/14	double low quotation mark

Its table representation is, according to this list, as shown in Figure 7.
It can be seen that the G3 set is a 94-position set, excluding both hexadecimal
positions 02/00 and 07/15.

	0	1	2	3	4	5	6	7
0				$^1/_3$	—	—	○	✠
1			EM SPACE	$^2/_3$	│	│	□	†
2			EN SPACE	$^1/_5$	└	└	▭	‡
3			3/EM SPACE	$^2/_5$	┘	┘	△	√
4			4/EM SPACE	$^3/_5$	┐	┐	▽	×
5			DIGIT SPACE	$^4/_5$	┌	┌	☆	#
6			PUNCT SPACE	$^1/_6$	├	├	●	♭
7			THIN SPACE	$^5/_6$	⊥	⊥	■	♂
8			HAIR SPACE	$^c/_o$	┤	┤	▲	♀
9			EM DASH	■	┬	┬	▼	☎
A			EN DASH	▬	+	+	◄	⌀
B			HYPHEN -	▬	├	├	►	℗
C			⊔		⊥	⊥	♣	✁
D					┤	┤	◇	'
E			...	▨	┬	┬	♡	,,
F			..	■	+	+	♠	

Figure 7: The graphic set G3

Some possible uses of Einecs tapes

1. Transfer from magnetic tape to hard disk

2. Structuring of chemical substances

Depending upon the objectives of the user of the tape information, one can discard the bibliographical data at the beginning of the tape and only treat the chemical substances data.

According to the different usage contexts, the user should define the proper support:

(i) a database if the objective is to allow various accesses to this enormous amount of data;

(ii) simple files if the objective is to print the information using another format and/or to print only subsets of this data or even to cross-relate this inventory with others.

3. How to explore the contents

There are a number of different ways to use this information, each client having his own reasons to purchase this tape or set of tapes.

Here are two examples:

(i) *Loading a database*

After carefully reading these specifications it is clear that each chemical substance has a certain number of fields (some of which are optional) that can be used to load a chemistry database after defining and selecting access keys.

The problems with the impoverishment of the rich character set as well as the representation of special characters can be solved (when reading or using the tape) according to the several output qualities foreseen in the application.

(ii) *Reprinting according to a personalized structure*

The majority of modern desktop publishing or word-processing tools have 'stylesheet' formatting possibilities that allow the presentation of the data in a very flexible way according to the user's needs.

This possibility, as the previous one, requires the user to define clearly the level of printing capacities and the corresponding impoverishment to take place.

European Communities—Commission

EINECS: European inventory of existing commercial chemical substances

Luxembourg: Office for Official Publications of the European Communities

1990 — 40 pp. — 14.8 × 21 cm

Venta y suscripciones • Salg og abonnement • Verkauf und Abonnement • Πωλήσεις και συνδρομές
Sales and subscriptions • Vente et abonnements • Vendita e abbonamenti
Verkoop en abonnementen • Venda e assinaturas

BELGIQUE / BELGIË

Moniteur belge /
Belgisch Staatsblad
Rue de Louvain 42 / Leuvenseweg 42
1000 Bruxelles / 1000 Brussel
Tél. (02) 512 00 26
Fax 511 01 84
CCP / Postrekening 000-2005502-27

Autres distributeurs /
Overige verkooppunten

Librairie européenne/
Europese Boekhandel
Avenue Albert Jonnart 50 /
Albert Jonnartlaan 50
1200 Bruxelles / 1200 Brussel
Tél. (02) 734 02 81
Fax 735 08 60

Jean De Lannoy
Avenue du Roi 202 /Koningslaan 202
1060 Bruxelles / 1060 Brussel
Tél. (02) 538 51 69
Télex 63220 UNBOOK B

CREDOC
Rue de la Montagne 34 / Bergstraat 34
Bte 11 / Bus 11
1000 Bruxelles / 1000 Brussel

DANMARK

J. H. Schultz Information A/S
EF-Publikationer
Ottiliavej 18
2500 Valby
Tlf. 36 44 22 66
Fax 36 44 01 41
Girokonto 6 00 08 86

BR DEUTSCHLAND

Bundesanzeiger Verlag
Breite Straße
Postfach 10 80 06
5000 Köln 1
Tel. (0221) 20 29-0
Fernschreiber:
ANZEIGER BONN 8 882 595
Fax 20 29 278

GREECE

G.C. Eleftheroudakis SA
International Bookstore
Nikis Street 4
10563 Athens
Tel. (01) 322 63 23
Telex 2194°0 ELEF
Fax 323 98 21

ESPAÑA

Boletín Oficial del Estado
Trafalgar, 27
28010 Madrid
Tel. (91) 446 60 00

Mundi-Prensa Libros, S.A.
Castelló, 37
28001 Madrid
Tel. (91) 431 33 99 (Libros)
431 32 22 (Suscripciones)
435 36 37 (Dirección)
Télex 49370-MPLI-E
Fax (91) 275 39 98

Sucursal

Librería Internacional AEDOS
Consejo de Ciento, 391
08009 Barcelona
Tel. (93) 301 86 15
Fax (93) 317 01 41

Generalitat de Catalunya:

Llibreria Rambla dels estudis
Rambla, 118 (Palau Moja)
08002 Barcelona
Tel. (93) 302 68 35
302 64 62

FRANCE

Journal officiel
Service des publications
des Communautés européennes
26, rue Desaix
75727 Paris Cedex 15
Tél. (1) 40 58 75 00
Fax (1) 40 58 75 74

IRELAND

Government Publications
Sales Office
Sun Alliance House
Molesworth Street
Dublin 2
Tel. 71 03 09

or by post

Government Stationery Office
EEC Section
6th floor
Bishop Street
Dublin 8
Tel. 78 16 66
Fax 78 06 45

ITALIA

Licosa Spa
Via Benedetto Fortini, 120/10
Casella postale 552
50125 Firenze
Tel. (055) 64 54 15
Fax 64 12 57
Telex 570466 LICOSA I
CCP 343 509

Subagenti:

Libreria scientifica
Lucio de Biasio - AEIOU
Via Meravigli, 16
20123 Milano
Tel. (02) 80 76 79

Herder Editrice e Libreria
Piazza Montecitorio, 117-120
00186 Roma
Tel. (06) 679 46 28/679 53 04

Libreria giuridica
Via 12 Ottobre, 172/R
16121 Genova
Tel. (010) 59 56 93

GRAND-DUCHÉ DE LUXEMBOURG

Abonnements seulement
Subscriptions only
Nur für Abonnements

Messageries Paul Kraus
11, rue Christophe Plantin
2339 Luxembourg
Tél. 499 88 88
Télex 2515
CCP 49242-63

NEDERLAND

SDU uitgeverij
Christoffel Plantijnstraat 2
Postbus 20014
2500 EA 's-Gravenhage
Tel. (070) 78 98 80 (bestellingen)
Fax (070) 47 63 51

PORTUGAL

Imprensa Nacional
Casa da Moeda, EP
Rua D. Francisco Manuel de Melo, 5
1092 Lisboa Codex
Tel. (01) 69 34 14

Distribuidora de Livros
Bertrand, Ld.ª
Grupo Bertrand, SA
Rua das Terras dos Vales, 4-A
Apartado 37
2700 Amadora Codex
Tel. (01) 493 90 50 - 494 87 88
Telex 15798 BERDIS
Fax 491 02 55

UNITED KINGDOM

HMSO Books (PC 16)
HMSO Publications Centre
51 Nine Elms Lane
London SW8 5DR
Tel. (071) 873 9090
Fax GP3 873 8463

Sub-agent:

Alan Armstrong Ltd
2 Arkwright Road
Reading, Berks RG2 0SQ
Tel. (0734) 75 18 55
Telex 849937 AAALTD G
Fax (0734) 75 51 64

CANADA

Renouf Publishing Co. Ltd
Mail orders — Head Office:
1294 Algoma Road
Ottawa, Ontario K1B 3W8
Tel. (613) 741 43 33
Fax (613) 741 54 39
Telex 0534783

Ottawa Store:
61 Sparks Street
Tel. (613) 238 89 85

Toronto Store:
211 Yonge Street
Tel. (416) 363 31 71

JAPAN

Kinokuniya Company Ltd
17-7 Shinjuku 3-Chome
Shinjuku-ku
Tokyo 160-91
Tel. (03) 354 01 31

Journal Department
PO Box 55 Chitose
Tokyo 156
Tel. (03) 439 01 24

MAGYAR

Agroinform
Központ
Budapest I., Attila út 93. H-1012

Levélcím:
Budapest, Pf. 15 H-1253
Tel. 36 (1) 56 82 11
Telex (22) 4717 AGINF H-61

ÖSTERREICH

Manz'sche Verlags-
und Universitätsbuchhandlung
Kohlmarkt 16
1014 Wien
Tel. (0222) 531 61-0
Telex 11 25 00 BOX A
Fax (0222) 531 61-81

SCHWEIZ / SUISSE / SVIZZERA

OSEC
Stampfenbachstraße 85
8035 Zürich
Tel. (01) 365 51 51
Fax (01) 365 54 11

SVERIGE

BTJ
Box 200
22100 Lund
Tel. (046) 18 00 00
Fax (046) 18 01 25

TÜRKIYE

Dünya süper veb ofset A.Ş.
Narlibahçe Sokak No. 15
Cağaloğlu
Istanbul
Tel. 512 01 90
Telex 23822 DSVO-TR

UNITED STATES OF AMERICA

UNIPUB
4611-F Assembly Drive
Lanham, MD 20706-4391
Tel. Toll Free (800) 274 4888
Fax (301) 459 0056
Telex 7108260418

YUGOSLAVIA

Privrednivjesnik
Rooseveltov Trg 2
41000 Zagreb
Tel. 44 64 28
44 98 35
43 32 80
44 34 22
Teleks 21524 YU

AUTRES PAYS
OTHER COUNTRIES
ANDERE LÄNDER

Office des publications officielles
des Communautés européennes
2, rue Mercier
L-2985 Luxembourg
Tél. 49 92 81
Télex PUBOF LU 1324 b
Fax 48 85 73
CC bancaire BIL 8-109/6003/700

2/90