

Part 5

INSTALLATION DEPENDENT OPTIONS

Section 21

Adapting CDS/ISIS to local requirements

A. Creating a new language version

To create a new language version of CDS/ISIS you must translate all menus, system worksheets and messages into the language to be implemented. The following sections outline the procedures for each item.

1. Menu and system worksheet translation

Modify all xXLNG menus to include the new language and create a new menu xXLNG corresponding to the new language.

For each menu or system worksheet use the ISISUTL services to copy the english menu into the new language menu using option **D** of menu xXM2 (for worksheets) or xXM3 (for menus). Then, using the editor, translate the captions. The name of the new language worksheet or menu must be the same as the name of the english version except for the language code. For example the name of the french version of the menu EXGEN will be FXGEN.

In all menus which include the **L** option change the name of the xXLNG menu to be selected to the name of the xXLNG in the new language. For example, in the EXGEN menu, the **L** option points to menu EXLNG; in the FXGEN menu it should point to FXLNG, where the name of the languages will be in french.

If you change the option identifier in menus, be sure you are providing the correct *internal option identifier* (see under "Internal option identifier" on page 218). In this case you should have a printed copy of the original English menus supplied by Unesco.

When translating system worksheets, remember that *all* worksheets of the same type (e.g. all print worksheets) *must have the same number of fields (including title fields)* and *the order of the fields must be the same* as in the original version supplied by Unesco.

2. Message translation

To translate the system messages proceed as follow:

1. create the new language message data base by copying all the english message data base files; to do this use the following MSDOS commands:

```
COPY EMSG.*    xMSG.*
COPY ?EMSG.FMT ?xMSG.FMT
```

where x is the code of the new language;

2. select the data entry option for data base xMSG and translate all the messages. Be careful when translating messages which are used as table headings (6, 42, 248, 271, 272, 283) to respect the same alignment as in the english version. Also note that the following groups of messages must have the same length (pad the shorter one with spaces if necessary), in each language version:

75, 76, 77, 78
86, 87, 88, 89
101, 102

3. furthermore, submenu component lines should not be longer than 80 characters, otherwise they may cause undesirable screen scrolling. Submenus where this may occur are listed below:

49, 50
51, 52
54, 55
56
61, 62
71, 72
103, 104, 105

4. Message 343 (in the English message data base) contains the text [**unassigned**]. It is used to provide the text to be printed for unassigned messages in the various linguistic versions when printing the message file. Simply translate this message into the new language (e.g. [**no asignado**] in spanish). All unassigned messages will then print this message when listing the message data base using the **D** option of the ISISUTL services.

Note that message master files, unlike other master files, are stored in a special way in order to optimize the access time. In particular, records in a message master file cannot exceed 512 characters and there can be no more than 10 messages per record. Therefore when you translate the message file you should keep these restrictions in mind. "List of English messages" on page 235 of this manual contains the complete list of the English messages as supplied by Unesco. This list was produced using option **D** of the main menu of the ISISUTL services. After you have translated the message file you may produce a listing of it using this same option, by first selecting the new language code through the **L** option.

3. Customizing submenu options

You may customize submenu options to suit the particular requirements of a given language. As for menus, CDS/ISIS must be able to uniquely identify the option selected by the user from a submenu. The *internal* submenus identifiers must be the same as the ones provided by Unesco in the english message file. If you want to change the option identifiers in a given submenu, you must provide a correspondence table between the external and the internal option identifier. This table is stored at the end of each submenu line as follows:

\$e1i1e2i2e3i3 . . .

where **e1**, **e2**, etc. are the *external* option identifiers, and **i1**, **i2**, etc. are the corresponding *internal* option identifiers. For example, suppose you wanted to change submenu 51/52 to read:

```
Field attribute:  N[ormal]      R[everse video]  H[ighlighted]
                  U[nderlined]  B[linking]      I[nvisible]
```

then you must change messages 51 and 52 as follows:

```
Field attribute:  N[ormal]      R[everse video]  H[ighlighted]$NOR1H2
                  U[nderlined]  B[linking]      I[nvisible]$U3B4I5
```

B. Pseudo messages

Some of the messages stored in the message data base are not used as prompts, but define system parameters¹. These are described below.

- 273 the first character of this message defines the graphic to be used by the line editor for the FDT and FST table headings (default setting =)
- 274 this message provides the separator to be used by the ISISPRT services when printing an index (default setting ', ');
- 351 the first eight characters in this message provide the screen attributes to be used for screen background, message area, normal, reverse video, bold, underline, blink and invisible respectively. They may be changed by using option A of the ISISUTL services.
- 354 if defined, it must contain the segment address of the video refresh buffer of the display adapter. The address must be entered in hexadecimal and preceded by the # sign (e.g. #B000). This is only relevant for the IBM-PC version of CDS/ISIS, which uses memory-mapped techniques for video display (other versions such as VAX, WANG or the Arabic version do not use memory mapped display). In the message files distributed by Unesco, this message is not defined. CDS/ISIS will automatically determine the correct address of the video refresh buffer by checking the type of adapter installed on your computer (#B800 for the monochrome display adapter and #B000 for the color/graphics adapter). But if you are using a machine with a non-standard display adapter, you may provide here the relevant segment address. Note, however, that this message must first be modified on an IBM-compatible computer.

These messages should not be modified while creating a new language version.

¹ Other system parameters are defined in the file SYSPAR.PAR (see under "SYSPAR.PAR: Global parameters" on page 11).

C. Creating a stopword file

This is a very simple text file which you may create using a standard editor such as EDLIN. It contains one stopword per line and the stopwords must be in *upper case, left adjusted* (i.e. beginning in column 1) and in *ascending alphabetical order*. The maximum number of words contained in a stopword file is 799 and each word may not exceed 10 characters. A sample stopword file is given below:

```
A
AN
AND
AS
BY
FOR
FROM
IN
INTO
ITS
OF
ON
THE
TO
UPON
WITH
```

The stopword file used in conjunction with the FST for the inverted file must have the same name as the data base name with the file extension **STW**. Other stopword files used with the print program may have any name, but always the file extension **STW**.

D. Creating an ANY file

The ANY file is a text file which you may create with a standard editor such as EDLIN. Each line contains the ANY term starting in position 1 (left justified) and the associated search term starting in position 31 (also left justified). The Any term and/or the associated search terms may be entered in upper or lower case or a combination of the two.

The ANY file must have the same name as the data base name with the file extension **ANY**. A sample ANY file is given below:

```
ANY SCANDINAVIA      SCANDINAVIA
ANY SCANDINAVIA      DENMARK
ANY SCANDINAVIA      FAROE ISLANDS
ANY SCANDINAVIA      FINLAND
ANY SCANDINAVIA      GREENLAND
ANY SCANDINAVIA      ICELAND
ANY SCANDINAVIA      NORWAY
ANY SCANDINAVIA      SWEDEN
ANY CARIBBEAN        CARIBBEAN
ANY CARIBBEAN        BAHAMAS
ANY CARIBBEAN        BARBADOS
ANY CARIBBEAN        BRITISH VIRGIN ISLANDS
ANY CARIBBEAN        CAYMAN ISLANDS
```

21. Adapting CDS/ISIS to local requirements

ANY CARIBBEAN	CUBA
ANY CARIBBEAN	DOMINICAN REPUBLIC
ANY CARIBBEAN	GUADELOUPE
ANY CARIBBEAN	HAITI
ANY CARIBBEAN	JAMAICA
ANY CARIBBEAN	MARTINIQUE
ANY CARIBBEAN	MONTSERRAT
ANY CARIBBEAN	NETHERLANDS ANTILLES
ANY CARIBBEAN	PUERTO RICO
ANY CARIBBEAN	TRINIDAD AND TOBAGO
ANY CARIBBEAN	TURKS AND CAICOS ISLANDS
ANY CARIBBEAN	UNITED STATES VIRGIN ISLANDS
ANY CARIBBEAN	WEST INDIES ASSOCIATED STATES
ANY AGRICULTURE	AGRICULTURAL ECONOMICS
ANY AGRICULTURE	LAND ECONOMICS
ANY AGRICULTURE	AGRICULTURAL ENTERPRISES
ANY AGRICULTURE	AGRICULTURAL EQUIPMENT
ANY AGRICULTURE	AGRICULTURAL PRODUCTION
ANY AGRICULTURE	AGRICULTURAL RESEARCH
ANY AGRICULTURE	PLANT PRODUCTION
ANY AGRICULTURE	FORESTS
ANY AGRICULTURE	ANIMAL PRODUCTION
ANY AGRICULTURE	FISHERY

E. Uppercase conversion table (ISISUC.TAB)

This text file is used to convert text to upper case. It contains the decimal ASCII code of the uppercase equivalent of each of the 256 characters. Uppercase conversion is performed by replacing a given text character by the corresponding code in this table. Note that the table must contain 256 codes. Characters which are not alphabetic remain unchanged. The standard table supplied by Unesco for the IBM PC version is given below (different tables are provided for the WANG PC and the VAX versions):

```

000 001 002 003 004 005 006 007 008 009 010 011 012 013 014 015
016 017 018 019 020 021 022 023 024 025 026 027 028 028 030 031
032 033 034 035 036 037 038 039 040 041 042 043 044 045 046 047
048 049 050 051 052 053 054 055 056 057 058 059 060 061 062 063
064 065 066 067 068 069 070 071 072 073 074 075 076 077 078 079
080 081 082 083 084 085 086 087 088 089 090 091 092 093 094 095
096 065 066 067 068 069 070 071 072 073 074 075 076 077 078 079
080 081 082 083 084 085 086 087 088 089 090 123 124 125 126 127
067 085 069 065 065 065 065 067 069 069 069 073 073 073 065 065
069 069 069 079 079 079 085 085 089 079 085 155 156 157 158 159
065 073 079 085 078 078 166 167 168 169 170 171 172 173 174 175
176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191
192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207
208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223
224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239
240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255

```

Note that accented characters are converted to their non accented upper case form.

ISISUC.TAB may contain two 256-character tables. The second table, if missing, will, by default, be equal to the first one. The first table is used to translate search terms to upper case before storing them on the inverted file, and to translate text to upper case when the appropriate mode is selected in a display format (**mpu**, **mhu** or **mdu**). The second table is used to translate sort keys before storing them on the hit file. This is because certain languages require special sort sequences which are different from the normal ASCII collating sequence.

The sort key translate table provided in ISISUC.TAB is the system default table. In addition, you may provide a special sort key translate table for each data base, by creating a file `xxxxxx.SRT` (where `xxxxxx` is the data base name) in the data base path. If this file exists, CDS/ISIS will use it for generating the sort keys of the data base. If not, the system default table stored in ISISUC.TAB will be used.

Note that, as sort keys are translated *after* executing the FST data extraction format, if you provide a sort key translate table you should normally avoid using mode commands forcing upper case translation (**mdu**, **mpu** or **mhu**) in this format.

F. Alphabetic characters table (ISISAC.TAB)

This text file defines the decimal ASCII code of all alphabetic characters. It is used each time CDS/ISIS needs to know if a given character is alphabetic (e.g. when performing word indexing using indexing technique 4, or validating alphabetic fields). The standard table supplied by Unesco for the IBM PC version is given below (different tables are provided for the WANG PC and the VAX versions):

```
065 066 067 068 069 070 071 072 073 074 075 076 077 078 079 080
081 082 083 084 085 086 087 088 089 090 097 098 099 100 101 102
103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
119 120 121 122 128 129 130 131 132 133 134 135 136 137 138 139
140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 160
161 162 163 164 165
```

Note that accented characters are defined as alphabetic.