

UNIMARC TO MARC 21 CONVERSION SPECIFICATIONS

(Version 3.0)

Prepared by the  
Network Development and MARC Standards Office  
Library of Congress

Washington, D.C.

August 2001

## TABLE OF CONTENTS

### INTRODUCTION

SECTION I. Conversion of UNIMARC Record Label, Directory, Fields 001, 100-135

SECTION II. Conversion of UNIMARC Fields 200 to 802

SECTION III. Procedures and Punctuation Checks

- Procedure 1: Personal Names
  - Procedure 2: Corporate Names
  - Procedure 3: Conference Names
  - Procedure 4: Titles
  - Procedure 5: Creation of MARC 21 Field 886
  - Procedure 6: Setting of Nonfiling Indicators
  - Procedure 7: Linking Entry Fields (4XX)
  - Procedure 8: Setting of the Second Indicator in Fields 600-607
  - Procedure 9: Construction of MARC 21 Leader, Directory, Field 008  
(Fixed-Length Data Elements)
  - Procedure 10: MARC 21 Field 007 for Visual Materials
  - Procedure 11: MARC 21 Field 007 for Music
  - Procedure 12: MARC 21 Field 007 for Microforms
  - Procedure 13: MARC 21 Field 007 for Maps
- Punctuation Check A  
Punctuation Check B

SECTION IV. Tables

- Table 1: ISO 3166 to MARC Country Code Conversion
- Table 2: Relator Code Conversion
- Table 3: Error Messages
- Table 4: Character Set Conversion
- Table 5: National Bibliography Number Source Conversion

## INTRODUCTION

### SCOPE AND PURPOSE

This document provides specifications for the machine conversion of bibliographic records in the UNIMARC format to the MARC 21 format. It covers various types of materials including language material (books), notated music (scores), cartographic materials (maps), projected and video material, sound recordings, computer media and three-dimensional artifacts (realia). The specifications cover the conversion of items at four bibliographic levels: analytic, collection, monograph, and serial. This document was originally written for a UNIMARC conversion program developed by the Library of Congress (LC) in 1984. Although updated in 2001 for UNIMARC users, resources were not available for exhaustive review. Some UNIMARC or MARC 21 elements may be missing from this specification.

### ARRANGEMENT OF THE DOCUMENT

After this general introduction, Detailed Conversion Specifications are presented in tabular form, arranged in the order of the UNIMARC format data elements from the Record Label through the last variable field (field 802). Following the mappings of UNIMARC to MARC 21 data elements are Procedures and Punctuation Checks which contain conversion routines used for several different fields. The procedures and punctuation checks are referenced in the specifications for applicable fields. At the end are Tables which give the correspondence of UNIMARC codes to equivalent MARC 21 codes, a listing of valid graphic and control characters, and a summary of recommended error and warning messages that might be generated during conversion.

### DETAILED CONVERSION SPECIFICATIONS

The valid UNIMARC data elements represent a composite of those found in the UNIMARC Manual, Bibliographic Format (2nd ed., 1994). The MARC 21 fields are based on the MARC 21 Format for Bibliographic Data (1999 ed., with Updates No. 1 and 2).

The column headings are self-explanatory. When conversion instructions are simple, they are given in the "Notes" column. When they are more complex, but field specific, they are given immediately following the subfield lists in a "Processing Note". Conversion instructions that are used repetitively are usually given separately as a procedure.

### KEY TO SYMBOLS USED IN THE CONVERSION TABLES

Symbols associated with UNIMARC data elements are:

- # Blank or space (hexadecimal character '20')
- \$ Subfield delimiter (hexadecimal control character '1F')
- + Following a data element indicates that it is provisional
- Following a data element indicates that it is now obsolete

- (D) In the notes column, following a coded value or indicator, indicates the default for that data element or indicator. Any value in that position in the incoming UNIMARC record for which no specific instruction appears should be converted to the default value.

## GENERAL SPECIFICATIONS

### 1. Designation of 008 Field Configuration in MARC 21

In the UNIMARC format general coded data for all forms of material is recorded in field 100. Coded data specific to certain forms of material is recorded in fields 105-135. The UNIMARC data elements from these LXX fields generally convert to one of seven (7) distinct configurations of MARC 21 field 008. In MARC 21 the field 008 configurations are Books (BK), Mixed Material (MX), Computer Files (CF), Maps (MP), Music (includes sound recordings) (MU), Visual Materials (includes films) (VM), and Serials (SE). In the conversion specifications, the term "If BOOKS", etc. may be used in the instructions to identify specific forms of material in UNIMARC. The following conditions determine the corresponding MARC 21 field 008 configuration and thus form of material:

	UNIMARC LDR/06	UNIMARC LDR/07	MARC21 LDR/06	MARC21 LDR/07	MARC21 LDR/08
BK	a, h	c, m	a	c, m	Ø
AM	b	c, m	-b, t	c, m	Ø
CF	l	c, m	m	c, m, s	Ø
MP	e, f	c, m, s	e, f	c, m, s	Ø
MU	c, d, i, j	c, m	c, d, i j	c, m, s	Ø
VM	g, k, m, n	c, m	g, k, o	c, m, s	Ø
SE	a, h	s	a	s	a

### 2. File Parameter (for LC conversion program only!)

In order to generate some MARC 21 data elements correctly, a parameter setting identifying certain characteristics of the UNIMARC file being processed should be made available during conversion. This parameter is referenced at several points in the program, e.g. Procedures 5.

As part of the error reports listing, the file parameter set for that run of the program should be printed out ahead of and on the same page as the list of error message texts in the form: "File parameter = XX", etc. as appropriate.

### 3. Validation

These specifications should result in the creation of MARC 21 records that are structurally sound and that conform in most respects to the MARC 21 Format for Bibliographic Data. The instruction in Unrecognized Content Designation Section (below) and at the field level attempt to ensure that tags, indicators, coded values, and characters that are invalid in MARC 21 are not present in the converted record. The program, however, includes only minimal "form-specific" validation (i.e., programmatic editing of fields according to the form of material identified by Record type and Bibliographic level values in the MARC 21 Record Leader).

#### 4. Unrecognized Content Designation

- a. Unrecognized Indicators and Coded Values. Instructions for handling unrecognized UNIMARC indicators and coded values are given at the field level in the Detailed Specifications. In general, these instructions specify that whenever the value of a MARC 21 code or indicator is dependent on the corresponding UNIMARC value, and the UNIMARC value is unrecognized, a default MARC 21 value is used to ensure the validity of the MARC 21 record.
- b. Unrecognized Field Tags. Any unrecognized field tag (i.e., an incoming UNIMARC field for which no conversion has been specified) should be transferred to a MARC 21 field 886 (Foreign MARC Information Field) according to Procedure 5 and an error message generated (Msg 07).
- c. Unrecognized Subfield Codes. All unrecognized subfield codes (i.e., incoming UNIMARC subfield identifiers for which no conversion has been specified) encountered in a field being converted from UNIMARC to MARC 21 should be handled as follows:
  - (1) if the unrecognized subfield code is the first subfield code in that field, the subfield code and data should be dropped and an error message generated (Msg 09)
  - (2) if the unrecognized subfield code is not the first subfield code in the field, it should be replaced by a single blank; the data in that subfield should be concatenated to the end of the immediately preceding subfield; and an error message should be generated (Msg 08)
- d. Unrecognized Characters. Any unrecognized character (i.e., an incoming UNIMARC character for which no conversion has been specified in Table 4, Character Set Conversion Specification) encountered in a data being converted from UNIMARC to MARC 21, should be converted to the fill character (hex '7C') and an error message generated (Msg 10).

#### 5. Error and Warning Messages

The format of the required error and warning messages is given in Table 3.

#### UPDATES

Version 1.0: Base document, 11-30-1984

Version 1.1: Update, 01-18-1985

Version 1.2: Update, 01-25-1985

Version 1.2 Consolidated: Supersedes previous base and updates, 09-01-1987

Version 2.0: Superseded Version 1.0, 11-15-1992

Version 3.0: Superseded Version 2.0, 08-28-2001