



Redvers Consulting Ltd

Redvers COBOL XML Interface

User Guide

**Superfast XML Generator
RCFSTXML Version 2.7**

Contents

Preface	4
Overview	4
Installation.....	6
RCFSTCMP.....	6
RCFSTXML.....	7
Coding the COBOL Record Definition	8
Field Names.....	8
Default Tag Names	8
Override Tag Names.....	9
XML Attributes	9
XML Namespaces and SOAP	10
PICTURE Clause	11
Binary / Packed Fields	12
Other Clauses	13
BLANK WHEN ZERO	13
JUSTIFIED RIGHT	13
OCCURS	13
Clauses Not Supported	13
Structure.....	14
Mixed Content Elements	15
Non-tagged group level fields	15
Advanced Techniques.....	16
Formatting.....	16
Required Attributes and Elements	17
Optional Elements.....	18
Excluded Elements.....	19
Repeating Groups.....	20
Using OCCURS	20
Using Repeated Calls	21
Controlling the Point of Change	23
Orphan Repeats	25
Calling RCFSTXML	27
Parameters.....	27
CRD-OBJECT-AREA (input).....	27
CRD-RECORD-COUNT (input)	27
COBOL-RECORD (input).....	27
COBOL-RECORD-LENGTH (input)	27

XML-DOCUMENT (output)	28
XML-DOCUMENT-LENGTH (input & output)	28
FEEDBACK-CODE (output)	28
FEEDBACK-TEXT (output)	28
Calls to RCFSTXML.....	29
The First Call.....	29
Subsequent Calls	29
The Last Call	29
Sample Program Calling RCFSTXML.....	31
Structure Break	35
Normal Operation	35
What is a Structure Break?	36
How to code a Structure Break	36
When to use a Structure Break.....	37
Data Integrity	41
Character Range	41
Character References	41
Entity References	41
Empty Fields.....	42
CDATA	42
Maximum Document Size	43
Processing Instructions.....	43
Comments.....	43
User Maintained Variables	44
Program-Id.....	44
SELECT Statements	44
File Definition Statements.....	44
Maximum-COBOL-record-length	45
XML-declaration and XML-headers	45
Number-of-XML-headers	45
DTD-headers	45
Number-of-DTD-headers	45
End-of-line-chtrs	46
Maximum-number-of-fields	46
RCFSTCMP Compile Errors	47
RCFSTXML Feedback Messages	50
Index	52

Preface

This document describes the installation and operation of the Redvers COBOL XML Interface programs RCFSTCMP and RCFSTXML. It is designed for use by Information Technology departments familiar with the COBOL and XML computer languages.

Program RCFSTCMP is a batch compile process that reads a COBOL Record Definition (CRD) in order to produce a CRD object file for use in Redvers COBOL XML Interface programs RCFSTXML and RCFSTCOB.

Program RCFSTXML is a COBOL subroutine that generates XML documents from COBOL format data passed from customer applications. Output from RCFSTCMP provides RCFSTXML with all the information it needs to produce application specific XML. The generated well-formed XML standalone documents conform to the World Wide Web Consortium (W3C) Extensible Markup Language (XML) 1.0 (Second Edition) definition.

RCFSTXML is the counterpart to RCFSTCOB which uses a similar process to parse XML documents, returning the data in the form of a COBOL record.

This User Guide can be found on the internet at:

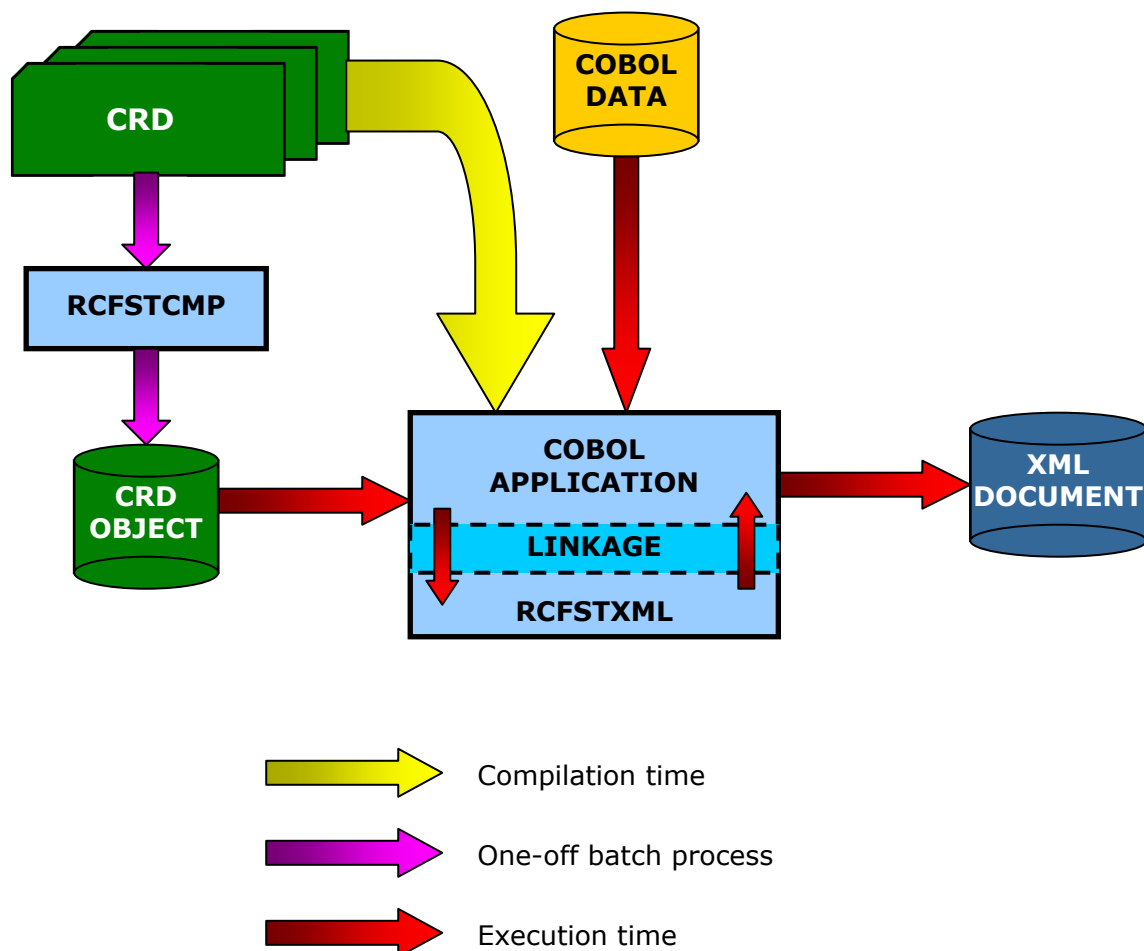
http://www.redversconsulting.com/downloads/user_guides/RCFSTXML_2.7_user.pdf.



Overview

RCFSTXML is a COBOL subroutine that is compiled and linked into application programs in order to generate well-formed XML documents. This particular interface level was built for large scale applications requiring fast and efficient XML generation. In order to achieve maximum speed and efficiency, the interpretation of the COBOL Record Definition (CRD) has been passed to an off-line batch compile process (RCFSTCMP).

Central to the operation of the interface is the COBOL Record Definition (CRD). This definition (usually a COBOL copybook member) is included in the application program source code where it describes the working storage fields that are to be passed to RCFSTXML. In addition, the CRD source is read by RCFSTCMP in a one-off batch process which creates a CRD object file containing the raw field positions, lengths, types etc. At execution time, application programs load this CRD object file into working storage and then pass it to RCFSTXML in a CALL statement.



Installation

RCFSTXML and RCFSTCMP are self-contained source programs that require no copybooks or objects. Before the product was despatched, the source code was passed through the [Redvers COBOL Cloaking Device](#) which removes the intellectual property within the source code without changing the logic.

To install RCFSTXML and RCFSTCMP, copy the “cloaked” source code into your source code library and paste the activation key (sent in an email when ordering/downloading) over the question marks in the last **VALUE** clause in working storage. The programs can then be compiled and linked using your standard COBOL compiler.

Various parameters including the program-id are defined as User Maintained Variables in the source code so that they may be set to alternative values if desired. See [User Maintained Variables](#) for details.

If your site uses single quotes (apostrophes) rather than double quotes (speech marks) to delimit literals, a global change of all double quotes to single quotes can be made. However, following this change in RCFSTXML, any single quotes within the XML declaration must be changed to double quotes before compilation. The XML declaration can be found at the start of working storage just above a comment line containing:

```
"<UMV> XML-declaration".
```

If you should encounter any problems during installation, please contact your account manager or use our “Contact” facility at: <http://www.redversconsulting.com/contact.php>.

RCFSTCMP

RCFSTCMP is the CRD compiler program required to pre-process CRD's used by the Redvers COBOL XML Interface (Superfast level).

To install, copy the RCFSTCMP source into your source code library and compile/link using your standard batch COBOL compiler.

To run, a batch compilation type job will be needed with the following file attributes:

- **RCCRDIN** – (input) fixed length 80 byte sequential file containing COBOL source code records of working storage field definitions (the CRD).
- **RCCRDOBJ** – (output) fixed length 132 byte sequential file with one record equating to each field on the input source code.
- File handling for output from **DISPLAY** statements, so that record counts or compilation errors can be viewed.

RCFSTXML

RCFSTXML is the XML generator subroutine component of the Redvers COBOL XML Interface (Superfast level).

To install, copy the source into your source code library and compile using your standard subroutine COBOL compiler.

To run, link RCFSTXML into your application program and issue a CALL to RCFSTXML passing the appropriate CRD object file created by RCFSTCMP. See [Calling RCFSTXML](#) for CALL parameter details.

Coding the COBOL Record Definition

The COBOL Record Definition (CRD) defines the layout of the COBOL record passed to the interface as well as the tags, format and structure of the XML document to be generated. Redvers Consulting offers two free tools to assist in drafting a CRD. These tools require either an XML W3C schema or Document Type Definition (DTD) as input and can be provided by your account manager or downloaded from the [Partners](#) area of our web site.

Field Names

Default Tag Names

The COBOL field names in the CRD are used as the default XML start/end tags and the data in the field becomes the XML element content.

Eg:

Field in CRD	Content
03 TV-program PIC X(20) .	"Sunday Night Live "

Generates:

XML Document
<TV-program>Sunday Night Live</TV-program>

If there is no data in the field, an empty element is generated (see [Empty Fields](#) for details on how empty fields are recognised).

Eg:

Field in CRD	Content
03 TV-program PIC X(20) .	spaces

Generates:

XML Document
<TV-program/>

Note: Upper/lower case settings in the tag/field name are preserved.

Override Tag Names

Because COBOL field names are subject to COBOL compiler rules such as character sets and the maximum length of a field name, the default tag may be overridden by coding an alternate name within "<" and ">" in the comment line or lines immediately following the field definition.

Eg:

Field in CRD	Content
03 TV-PROGRAM PIC X(20) . * <The_next_program_on_ * Channel-5_>	"Sunday Night Live "

Generates:

XML Document
<The_next_program_on_Channel-5_>Sunday Night Live</The_next_program_on_Channel-5_>

XML Attributes

If an XML attribute is required, this can be achieved by coding an "=" at the end of an override tag name in a subordinate field immediately after the group definition.

Eg:

Fields in CRD	Content
03 TV-program. 05 prog-time PIC X(5) . * <time=> 05 prog-name PIC X(20) .	"10:00" "Sunday Night Live "

Generates:

XML Document
<TV-program time="10:00"> <prog-name>Sunday Night Live</prog-name> </TV-program>

XML Namespaces and SOAP

The interface supports the generation of XML namespace declarations and Simple Object Access Protocol (SOAP) Envelopes using override tag names. Each namespace prefix is defined in the override tag name and the HTTP references are placed in each field's content.

Eg:

Fields in CRD	Content
03 TV-ENVELOPE. * <SOAP-ENV:Envelope>	
05 SOAP-NAMESP PIC X(34). * <xmlns:SOAP-ENV=>	"http://schemas.xmlsoap.org/soap "
05 UK-NAMESPACE PIC X(34). * <xmlns:uk=>	"http://www.greenwichmeantime.co.uk"
05 US-NAMESPACE PIC X(34). * <xmlns:us=>	"http://www.easternstandardtime.com"
05 TV-BODY. * <SOAP-ENV:Body>	
07 prog-name PIC X(20). 07 PR-UK-TIME PIC X(05). * <uk:time>	"Sunday Night Live " "10:00"
07 PR-US-TIME PIC X(05). * <us:time>	"05:00"

Generates:

XML Document
<pre>< SOAP-ENV:Envelope xmlns:SOAP-ENV=" http://schemas.xmlsoap.org/soap" xmlns:uk="http://www.greenwichmeantime.co.uk" xmlns:us="http://www.easternstandardtime.com"> <SOAP-ENV:Body> <prog-name>Sunday Night Live</prog-name> <uk:time>10:00</uk:time> <us:time>05:00</us:time> </SOAP-ENV:Body> </SOAP-ENV:Envelope></pre>

PICTURE Clause

The COBOL picture clause dictates how the data content of an XML element is written. Therefore COBOL data editing features can be utilised in order to produce the desired data format.

The interface uses the picture clause to determine if the field is to be left or right justified and will remove the appropriate leading/trailing spaces or zeroes but it will not remove blank characters if they are defined in the picture clause.

Eg:

Fields in CRD			Content
03	TV-program.		
05	prog-date	PIC 99/99/9999.	31122002
05	prog-name	PIC BBX(20)BB.	"Sunday Night Live "
05	prog-time	PIC X(5).	"10:00"
05	prog-duration	PIC ZZZ.999.	0.333
05	prog-rating	PIC S99.	5
05	prog-cost	PIC \$\$\$,\$\$\$,\$\$9.99DB.	-2500000

Generates:

XML Document
<pre> <TV-program> <prog-date>31/12/2002</prog-date> <prog-name> Sunday Night Live </prog-name> <prog-time>10:00</prog-time> <prog-duration>.333</prog-duration> <prog-rating>0E</prog-rating> <prog-cost>\$2,500,000.00DB</prog-cost> </TV-program> </pre>

Note: The imbedded sign in the prog-rating picture clause remains imbedded in the XML element content – not normally the desired result. A picture clause containing "-" or "+" (or the **SIGN TRAILING/LEADING SEPARATE** clause) would produce more readable XML.

Binary / Packed Fields

Due to the fact that XML is a character based language, binary and packed decimal numeric data cannot be safely represented within element tags. For this reason binary and packed decimal data is converted to a format known as base64. This format produces a character range of "A" through "Z", "a" through "z", "0" through "9", "+", "/" and "=". This subset of ISO 646 characters is represented identically in all versions of ASCII and EBCDIC which therefore ensures a smooth translation from ASCII to EBCDIC and vice-versa.

Eg:

Fields in CRD			Content
03	prog-rating	PIC 99 PACKED-DECIMAL.	8
03	prog-revenue	PIC S9(7)V99 BINARY.	1234567.89

Generates:

XML Document
<prog-rating>A18=</prog-rating> <prog-revenue>B1vNFQ==</prog-revenue >

Other Clauses

Data definition clauses can be used to edit the appearance of XML data within an element in just the same way they are used in COBOL. Some useful clauses are listed below:

BLANK WHEN ZERO

If the content of the field is zero this clause will result in the generation of an empty XML element.

JUSTIFIED RIGHT

This clause will cause the suppression of spaces to the left of text rather than to the right in the XML element data area.

OCCURS

This clause can be used to create a single dimension array of repeating data to be written to the XML document. It can be used at group or elementary level. Unpopulated occurrences within a populated array are generated as empty elements but trailing occurrences with no data are entirely suppressed from the document.

Arrays of more than one dimension are generated by issuing multiple calls to the interface subroutine – see [Repeating Groups](#) for details. This approach allows the interface to generate an unlimited number of occurrences in unlimited dimensions (which is the case for XML documents) using only a small amount of storage.

An **OCCURS 1** clause can be used to suppress optional elements entirely from the XML document, if there is no data to report – see [Optional Elements](#) for details. The use of **OCCURS 1** does not constitute an additional dimension in an array.

Clauses Not Supported

The following data definition clauses are not currently supported in the CRD:

OCCURS DEPENDING ON

REDEFINES

SYNCHRONIZED/SYNC

Symbol “P” in the picture clause

Structure

Just as COBOL has a hierarchical structure in the relationship between the various fields in a data record, XML has hierarchy between parent and child elements in a document. The interface uses the COBOL hierarchy to nest XML tags and data so that field relationships can be defined. All XML documents must start with a root element and this root element corresponds to the top level COBOL field in the CRD. Similarly, all lower COBOL levels are used to identify subordinate XML elements.

Eg:

Fields in CRD	Content
01 TV-listings.	
03 broadcast-date PIC 99/99/9999.	11122002
03 channel.	
05 channel-number PIC 99.	5
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-time PIC X(5) .	"10:00"
07 prog-rating PIC Z9.	8

Generates:

XML Document
<pre> <TV-listings> <broadcast-date>11/12/2002</broadcast-date> <channel> <channel-number>05</channel-number> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-time>10:00</prog-time> <prog-rating>8</prog-rating> </TV-program> </channel> </TV-listings> </pre>

Mixed Content Elements

When an XML element is a mixed content element (i.e. it contains data *and* other subordinate elements), it needs to be logically partitioned across a COBOL group definition. This type of generation is achieved using a group level data name with subordinate non-tagged fields. Non-tagged fields are either defined using **FILLER** or a null override tag name in the CRD.

Eg:

Fields in CRD		Content
03	TV-program.	
05	FILLER PIC X(20) .	"Sunday Night Live "
05	prog-time PIC X(5) .	"10:00"
05	END-TEXT PIC X(10) .	" tomorrow"
*	<>	

Generates:

XML Document
<pre><TV-program>Sunday Night Live <prog-time>10:00</prog-time> tomorrow </TV-program></pre>

Non-tagged group level fields

Non-tagged group level fields can be used in situations when fields need to be logically linked under a COBOL group level but the group level element itself is not required in the XML structure.

Eg:

Fields in CRD		Content
01	TV-listings.	
03	TV-program.	
*	<>	
05	prog-name PIC X(20) .	"Sunday Night Live "
05	prog-time PIC X(5) .	"10:00"

Generates:

XML Document
<pre><TV-listings> <prog-name>Sunday Night Live</prog-name> <prog-time>10:00</prog-time> </TV-listings></pre>

Advanced Techniques

Formatting

By using override tag names, non-tagged fields, attributes and field editing together, a variety of formats can be generated:

Eg:

Fields in CRD	Content
01 TV-listings.	
03 broadcast-date PIC 99/99/9999.	11122002
03 channel.	
05 channel-number PIC 99.	5
* <number=>	
05 TV-program.	
07 prog-name PIC BBX(20)BB.	"Sunday Night Live "
* <>	
07 prog-time PIC X(5) .	"10:00"
07 prog-rating.	
09 FILLER PIC X(20)B.	"I gave it "
09 mark PIC Z9.	8
09 out-of-text PIC BX(20) .	"out of ten "
* <>	

Generates:

XML Document
<pre> <TV-listings> <broadcast-date>11/12/2002</broadcast-date> <channel number="05"> <TV-program> Sunday Night Live <prog-time>10:00</prog-time> <prog-rating>I gave it <mark>8</mark> out of ten </prog-rating> </TV-program> </channel> </TV-listings> </pre>

Required Attributes and Elements

Under normal circumstances, an attribute will only be generated if the corresponding CRD field contains data (see [Empty Fields](#) for details of how the interface identifies a field as containing data or as being empty). However, DTD or schema validation rules may require the presence of an attribute even though there is no data to report. Required attributes would be defined in a DTD with the **#REQUIRED** parameter or in an XML schema with **use="required"**.

The introduction of release 2.5 enables the forced generation of *required* XML attributes by the coding of double "<" and ">" marks surrounding the override tag name in the CRD. Whenever the parent of a *required* attribute is generated, the interface will also generate the *required* attribute regardless of content.

Elements can also be marked as *required* using the double "<" and ">" marks, even though elements are generated by default anyway. The effect of defining empty elements as *required* causes the generation of separate start and end tags (**<tagname></tagname>**) rather than the usual empty tag (**<tagname/>**). For numeric elements, this may place zeroes in the element content depending on the picture clause.

Eg:

Fields in CRD	Content
03 channel.	
05 channel-number PIC 99.	00
* <<number=>>	
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-time PIC X(5) .	spaces
07 prog-rating.	
09 PROG-MARK PIC Z9.	00
* <<mark>>	

Generates:

XML Document
<pre> <channel number="00"> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-time/> <prog-rating> <mark>0</mark> </prog-rating> </TV-program> </channel> </pre>

Optional Elements

As mentioned in the Occurs section of [Other Clauses](#), trailing occurrences in a CRD array that are empty, are suppressed from the XML document. Therefore, using the same logic, an **OCCURS 1** clause can be used to suppress any element entirely from the XML document if it is deemed to be empty (see [Empty Fields](#) for details of how the interface identifies empty fields). The interface treats fields defined with **OCCURS 1** as it would the final occurrence of any multiply occurring field and will therefore not generate, if it is empty.

The **OCCURS 1** clause can therefore be coded on all optional elements, at group or elementary level, in order to remove them from the XML document when there is no data to report. Optional elements would be defined with an occurrence indicator of "?" or "*" in a DTD or with **minOccurs="0"** in an XML schema.

Eg:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-time OCCURS 1 PIC X(5) .	spaces
07 prog-rating OCCURS 1.	
09 PROG-MARK PIC Z9.	00
* <mark>	

Generates:

XML Document
<pre><channel> <TV-program> <prog-name>Sunday Night Live</prog-name> </TV-program> </channel></pre>

Excluded Elements

A COBOL Record Definition (CRD), used by the Redvers COBOL XML Interface is frequently in the form of a COBOL “copybook” or “include” and therefore may be used by an application for a variety of purposes, outside the calling of the Redvers XML generator routine. As a result, there may be fields in the CRD which relate to application processes outside the Redvers COBOL XML Interface and therefore should not result in the generation of an XML element for such a field.

From release 2.7, this issue has been addressed by providing a new type of CRD override tag name, within parentheses, in the form: `<(tagname)>`. Under these circumstances, the entry is still defined as a COBOL field and can still be used by the application for other purposes but no XML will be produced for this field, regardless of its content. In effect, the tag name specified becomes useful for documentation purposes only.

An example of an excluded field in a CRD can be seen below:

Field in the CRD		
*	03 TV-PROGRAM	PIC X(20) . <(Exclude_Me)>

Note: *Excluded fields must still be included in the total length of the CRD when populating the **COBOL-RECORD-LENGTH** parameter, otherwise a **FEEDBACK-CODE** of +0110 will be returned.*

Repeating Groups

In business applications, it's rare for an XML document to contain only a single set of information details. Elements and element groups are often repeated to reflect multiple sets and subsets of information. In order to generate this repeating data in XML, single dimension arrays can be defined in the CRD using the COBOL **OCCURS** clause.

Alternatively, when the number of occurrences is unknown or more than one dimension of repeating data is present, multiple calls can be made to RCFSTXML and only the changed information will be generated in XML. Using the latter technique, XML documents up to 99MB in length can be generated.

Using OCCURS

A single dimension array of repeating information can be passed to RCFSTXML in a CRD table which results in the generation of multiple sets of XML elements until all remaining occurrences in the CRD are unpopulated.

Eg:

Fields in CRD	Occurrence	Content
03 channel.		
05 channel-number PIC 99.		5
05 TV-program OCCURS 5.	1	
07 prog-name PIC X(17).	1	"Sunday Night Live"
07 prog-time PIC X(5).	1	"10:00"
07 prog-rating PIC Z9 OCCURS 1.	1	8
05 TV-program OCCURS 5.	2	
07 prog-name PIC X(17).	2	"News"
07 prog-time PIC X(5).	2	"11:30"
07 prog-rating PIC Z9 OCCURS 1.	2	0
05 TV-program OCCURS 5.	3	
07 prog-name PIC X(17).	3	spaces
07 prog-time PIC X(5).	3	spaces
07 prog-rating PIC Z9 OCCURS 1.	3	0
05 TV-program OCCURS 5.	4	
07 prog-name PIC X(17).	4	"Weather"
07 prog-time PIC X(5).	4	"11:55"
07 prog-rating PIC Z9 OCCURS 1.	4	0
05 TV-program OCCURS 5.	5	
07 prog-name PIC X(17).	5	spaces
07 prog-time PIC X(5).	5	spaces
07 prog-rating PIC Z9 OCCURS 1.	5	0

Generates:

XML Document
<pre> <channel> <channel-number>05</channel-number> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-time>10:00</prog-time> <prog-rating>8</prog-rating> </TV-program> <TV-program> <prog-name>News</prog-name> <prog-time>11:30</prog-time> </TV-program> <TV-program/> <TV-program> <prog-name>Weather</prog-name> <prog-time>11:55</prog-time> </TV-program> </channel> </pre>

Note: The **<prog-rating>** field and the fifth occurrence of **<TV-program>** are suppressed from the document when they are not populated but the third occurrence of **<TV-program>** is generated because subsequent occurrences exist.

Using Repeated Calls

An unlimited number of occurrences and dimensions (which is the case for XML documents) can be generated most efficiently by the use of repeated calls to the interface module. For each successive call, elements are generated for the lowest group level covering all changed fields in the CRD along with the minimum of related parent and child elements necessary to maintain the data relationships before and after the change. These repeated sets of related data items are collectively known to the program as the Point of Change (POC) for that call.

The example below generates a two dimensional array in XML for multiple **<TV-program>** elements within multiple **<channel>** elements using repeated calls and a smaller CRD.

Eg: The first call:

Fields in CRD	Content
03 channel.	
05 channel-number PIC 99.	5
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-time PIC X(5) .	"10:00"

The second call:

Fields in CRD	Content
03 channel.	
05 channel-number PIC 99.	5
05 TV-program.	
07 prog-name PIC X(20) .	"News "
07 prog-time PIC X(5) .	"11:30"

The third call:

Fields in CRD	Content
03 channel.	
05 channel-number PIC 99.	6
05 TV-program.	
07 prog-name PIC X(20) .	"Westenders "
07 prog-time PIC X(5) .	"08:00"

Generates:

XML Document
<pre> <channel> <channel-number>05</channel-number> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-time>10:00</prog-time> </TV-program> <TV-program> <prog-name>News</prog-name> <prog-time>11:30</prog-time> </TV-program> </channel> <channel> <channel-number>06</channel-number> <TV-program> <prog-name>Westenders</prog-name> <prog-time>08:00</prog-time> </TV-program> </channel> </pre>

Controlling the Point of Change

In order to provide a facility for the calling application to control the POC and force the regeneration of unchanged XML elements higher in the record hierarchy, a technique using dummy fields can be applied. When RCFSTXML reads a CRD, it converts any low-value characters in non-binary fields to spaces as it stores the information internally. Low-values remain in the application's CRD but the internal image of the data passed now contains spaces instead. The result of this process causes RCFSTXML to regard the dummy field as a changed field and therefore it generates additional XML to reflect this implied change. Of course, no XML is generated for the dummy field itself (having no tag and no data) but XML is generated for the parent of the dummy field. This implied change will continue for all subsequent calls, until low-values are removed from the dummy field by the application program.

In the example below, the dummy field in **TV-today** forces the second call to consider **TV-today** as the POC. Consequentially XML is generated for the entire **TV-today** group even though only **prog-name** and **prog-time** have changed.

Eg:

The first call:

Fields in CRD	Content
01 TV-listings.	
03 TV-today.	
05 Dummy-field PIC X.	Low-values
* <>	
05 broadcast-date PIC 99/99/9999.	11122002
05 channel.	
07 channel-number PIC 99.	5
07 TV-program.	
09 prog-name PIC X(20) .	"Sunday Night Live "
09 prog-time PIC X(5) .	"10:00"

The second call:

Fields in CRD	Content
01 TV-listings.	
03 TV-today.	
05 Dummy-field PIC X.	Low-values
* <>	
05 broadcast-date PIC 99/99/9999.	11122002
05 channel.	

07	channel-number	PIC 99.	5
07	TV-program.		
09	prog-name	PIC X(20) .	"News"
09	prog-time	PIC X(5) .	"11:30"

Generates:

XML Document
<pre> <TV-listings> <TV-today> <broadcast-date>11/12/2002</broadcast-date> <channel> <channel-number>05</channel-number> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-time>10:00</prog-time> </TV-program> </channel> </TV-today> <TV-today> <broadcast-date>11/12/2002</broadcast-date> <channel> <channel-number>05</channel-number> <TV-program> <prog-name>News</prog-name> <prog-time>11:30</prog-time> </TV-program> </channel> </TV-today> </TV-listings> </pre>

If there is a possibility that the CRD could be populated with exactly the same data in consecutive calls (duplicate records on the input), a similar technique can be used to generate a minimum of XML regardless of the COBOL input. This is done by placing a non-tagged dummy field, populated with low-values, just below the group level to be used as the minimum POC.

Orphan Repeats

Some XML documents require elements or groups of elements to repeat without the presence of a parent element for each occurrence. This type of generation is achieved using a non-tagged group level. Non-tagged group level fields are either defined using **FILLER** or a null override name in the CRD.

Eg:

The first call:

Fields in CRD			Content
03	channel.		
05	channel-number	PIC 99.	05
05	TV-program		
*	<>		
07	prog-name	PIC X(20).	"Sunday Night Live"
07	prog-time	PIC X(5).	"10:00"

The second call:

Fields in CRD			Content
03	channel.		
05	channel-number	PIC 99.	05
05	TV-program		
*	<>		
07	prog-name	PIC X(20).	"News"
07	prog-time	PIC X(5).	spaces

The third call:

Fields in CRD			Content
03	channel.		
05	channel-number	PIC 99.	05
05	TV-program		
*	<>		
07	prog-name	PIC X(20).	"Weather"
07	prog-time	PIC X(5).	"11:55"

Generates:

XML Document
<pre> <channel> <channel-number>05</channel-number> <prog-name>Sunday Night Live</prog-name> <prog-time>10:00</prog-time> <prog-name>News</prog-name> <prog-time/> </pre>

```
<prog-name>Weather</prog-name>  
<prog-time>11:55</prog-time>  
</channel>
```

Note: *The non-tagged **TV-program** group is the POC even though it doesn't appear on the XML document.*

Calling RCFSTXML

Parameters

A call to RCFSTXML requires eight parameters in the following sequence:

CRD-OBJECT-AREA (input)

This is the area of storage built by the application which holds the CRD object output from RCFSTCMP. The CRD defined for the XML document will have been passed through RCFSTCMP in a one-off batch process and the output made accessible to the application.

CRD-RECORD-COUNT (input)

This S9(8) binary field contains a count of the number of 132 byte records that make up **CRD-OBJECT-AREA**.

COBOL-RECORD (input)

This is the top level field name in the COBOL Record Definition (CRD) defined in the calling program. This area holds the COBOL format data to be converted to XML.

COBOL-RECORD-LENGTH (input)

This S9(8) binary field must be set to the logical length of **COBOL-RECORD**. It is used in validation only, to ensure **COBOL-RECORD** reflects the CRD in **CRD-OBJECT-AREA**.

If the logical length of **COBOL-RECORD** is difficult to determine (perhaps the same **COBOL-RECORD** storage area is being used to process several CRD's of different lengths), this value can be found in the appropriate **CRD-OBJECT-AREA**. The value is defined as an unsigned 7 digit packed decimal field (**PIC 9(7) COMP-3**) in positions 111 through 114 of **CRD-OBJECT-AREA**. These 4 bytes can be moved to the application program's working storage and a **REDEFINES** can be used to access the actual value.

XML-DOCUMENT (output)

This is the field name of the XML document area capable of holding the entire XML document to be generated.

XML-DOCUMENT-LENGTH (input & output)

This S9(8) binary field must be set to the length of the **XML-DOCUMENT** field and not changed during the generation process. When RCFSTXML has completed building the document, it will change this value to reflect the actual XML text length of the newly built document. The maximum XML document length supported by the interface is 99,999,999 characters.

FEEDBACK-CODE (output)

This S9(4) binary field is set by RCFSTXML to return the status of a call to the interface. This field should be checked for non-zero values after each call. If the call was successful this field will be set to zero otherwise it will contain an error code. See [Feedback Messages](#) at the end of this document for further information.

FEEDBACK-TEXT (output)

This eighty byte text field is set by RCFSTXML with diagnostic information on the results of each call. For the first successful call this area contains CRD information, for subsequent calls it contains the number of XML lines generated so far and the call count. For unsuccessful calls it contains an error message. See [Feedback Messages](#) at the end of this document for further information.

Note: *The parameter names used in this manual are suggestions only and may be changed to names more suitable to the application making the call.*

Calls to RCFSTXML

The First Call

If a Document Type Definition (DTD) is required at the start of the XML document, the first call to RCFSTXML must be made with the characters "DTD" in position one of **COBOL-RECORD** followed by spaces. **COBOL-RECORD-LENGTH** must be set to the full logical length of the CRD. If a Document Type Definition (DTD) is not required, the first call to RCFSTXML is made in the normal way - with a fully populated **COBOL-RECORD** as defined in the CRD.

Once processing is complete, **FEEDBACK-CODE** is set to zero and **FEEDBACK-TEXT** is populated with details of the CRD.

Subsequent Calls

Calls should be made to RCFSTXML (while **FEEDBACK-CODE** remains zero) with a fully populated **COBOL-RECORD** until all the application data is exhausted. In each case, **COBOL-RECORD** will represent a set of related data, similar to the way records are held in a sequential file.

After successful completion of each call, **FEEDBACK-CODE** is set to zero and **FEEDBACK-TEXT** is populated with the number of calls made and number of XML lines generated so far.

The Last Call

After all the application data has been passed to RCFSTXML, a final call must be made to allow the interface to complete the XML document and close its files. This final call is done, either by moving **HIGH-VALUES** to the whole CRD or by moving **ZERO** to the **COBOL-RECORD-LENGTH** parameter – either option will produce the same result.

After successful completion of this call, **XML-DOCUMENT-LENGTH** is set to the actual length of the generated XML document, **FEEDBACK-CODE** is set to zero and **FEEDBACK-TEXT** is populated with the total number of calls made and the total number of XML lines generated.

Note: *If RCFSTXML remains in memory after the last call, it can be reused to generate another XML document for the same, or a different, CRD by repopulating the calling parameters and restarting the call sequence from the first call.*

Note2: *If there is a requirement to generate multiple XML documents simultaneously, this can be done by calling a copy of RCFSTXML, compiled under a different Program-Id using different calling parameters ([Program-id](#) is a [User Maintained Variable](#)).*

Sample Program Calling RCFSTXML

```

...
000040*****
000050*   This sample program shows how RCFSTXML is called to generate *
000060*   an XML document in storage from COBOL record input.      *
000070*                                                           *
000080*   The CRD associated with this application will need to be   *
000090*   passed through the Redvers CRD compiler (RCFSTCMP) before  *
000100*   being loaded by the application. The CRD object file can   *
000110*   exist in storage, on a database or on a flat file. This    *
000120*   sample program uses a flat file (RCCRDOBJ).                *
000130*****
000132
000140 ENVIRONMENT DIVISION.
000150 INPUT-OUTPUT SECTION.
000160 FILE-CONTROL.
000170
000180         SELECT CRD-FILE                                ASSIGN RCCRDOBJ.
000190
000200 DATA DIVISION.
000210 FILE SECTION.
000220
000230 FD   CRD-FILE
000240         BLOCK CONTAINS 0 CHARACTERS
000250         LABEL RECORDS STANDARD.
000260 01   CRD-RECORD                                PIC X(132) .
000270
000280 WORKING-STORAGE SECTION.
000290
000300***   Storage area for the CRD object file:
000310 01   CRD-OBJECT-AREA.
000320         03   CRD-OBJECT-RECORD                PIC X(132) OCCURS 400.
000330
000340***   Start of COBOL Record Definition (CRD)
000350 01   TV-listings.
000360         03   channel.
000370                 05   channel-number                PIC 99.
000380*                 <number=>
000390                 05   channel-name                PIC X(20) .
000400*                 <>
000410                 05   TV-program                    OCCURS 10.
000420                         07   prog-name                PIC BBX(20)BB.
000430                         07   prog-time                PIC X(05) .
000440                         07   prog-rating              OCCURS 1.
000450*                         <rating>
000460                                 09   rating-txt1 PIC X(20)B.
000470*                                 <>
000480                                 09   mark                PIC Z9.
000490                                 09   rating-txt2 PIC BX(20) .
000500*                                 <>
000510                 05   channel-owner                PIC X(20) .
000520*                 <>
000530***   End of COBOL Record Definition (CRD)
000540
000550***   Storage area for the largest possible XML message:
000560 01   XML-DOCUMENT                                PIC X(16000)        VALUE SPACE.
000570
000572

```

```

000580 01  OTHER-PARAMETER-FIELDS.
000590      03  CRD-RECORD-COUNT          PIC S9(8)  BINARY VALUE ZERO.
000600      03  COBOL-RECORD-LENGTH       PIC S9(8)  BINARY VALUE ZERO.
000610      03  XML-DOCUMENT-LENGTH       PIC S9(8)  BINARY VALUE ZERO.
000620      03  FEEDBACK-CODE             PIC S9(4)  BINARY VALUE ZERO.
000630      03  FEEDBACK-TEXT             PIC X(80)          VALUE SPACE.
000640
000650 01  MISCELLANEOUS-FIELDS.
000660      03  CRD-TABLE-SIZE             PIC 9(4)          VALUE ZERO.
000670      03  CRD-FLAG                   PIC X              VALUE SPACE.
000680          88  START-OF-CRD              VALUE SPACE.
000690          88  END-OF-CRD                VALUE "E".
000700
000710 PROCEDURE DIVISION.
000720
000730 TOP-LEVEL SECTION.
000740*****
000750*  This section loads the CRD object file into CRD-OBJECT-AREA  *
000760*  and populates the COBOL record (CRD). It then calls RCFSTXML  *
000770*  once for each data group to generate the XML document.      *
000780*****
000790 TOP-ENTER.
000800
000810      COMPUTE CRD-TABLE-SIZE = LENGTH OF CRD-OBJECT-AREA
000820                          / LENGTH OF CRD-OBJECT-RECORD (1).
000830      MOVE LENGTH OF TV-LISTINGS TO COBOL-RECORD-LENGTH.
000840      MOVE LENGTH OF XML-DOCUMENT TO XML-DOCUMENT-LENGTH.
000850
000860      PERFORM A-LOAD-CRD.
000870
000930      INITIALIZE              TV-LISTINGS.
000940
000950      MOVE 3                  TO CHANNEL-NUMBER.
000960      MOVE "Popular TV"       TO CHANNEL-NAME.
000970      MOVE "Mr Nice Guy "     TO CHANNEL-OWNER.
000980
000990      MOVE "Sunday Night Live" TO PROG-NAME (1).
001000      MOVE "10:00"            TO PROG-TIME (1).
001010      MOVE "I gave it  "      TO RATING-TXT1 (1 1).
001020      MOVE 8                  TO MARK (1 1).
001030      MOVE "out of ten  "    TO RATING-TXT2 (1 1).
001040
001050      MOVE "News"              TO PROG-NAME (2).
001060      MOVE "11:30"            TO PROG-TIME (2).
001070
001080      MOVE "Weather"          TO PROG-NAME (3).
001090      MOVE "11:55"            TO PROG-TIME (3).
001100
001110*** The following call will generate XML for three programs on
001120*** channel 3 using the TV-program array.
001130      PERFORM C-CALL-RCFSTXML.
001140
001150      INITIALIZE              CHANNEL.
001160
001170      MOVE 5                  TO CHANNEL-NUMBER.
001180      MOVE "Soaps & Soaps"     TO CHANNEL-NAME.
001190      MOVE "Steady Eddie"      TO CHANNEL-OWNER.
001200
001210      MOVE "Westenders"        TO PROG-NAME (1).
001220      MOVE "09:00"            TO PROG-TIME (1).
001230

```



```

001240*** The following (repeated) call will generate XML for all
001250*** changed and populated elements. ie: channel 5.
001260     PERFORM C-CALL-RCFSTXML.
001270
001280     MOVE HIGH-VALUES                TO TV-LISTINGS.
001290
001300*** This final call completes the document.
001310     PERFORM C-CALL-RCFSTXML.
001320
001330*** Process XML-DOCUMENT (1:XML-DOCUMENT-LENGTH)
001400
001410     STOP RUN.
001420
001430 TOP-EXIT.
001440     EXIT.
001450
001460
001470 A-LOAD-CRD SECTION.
001480*****
001490*   This section reads the CRD object file into CRD-OBJECT-AREA.  *
001500*****
001510 A-ENTER.
001520
001530     OPEN INPUT CRD-FILE.
001540
001550     PERFORM
001560         UNTIL    END-OF-CRD
001570         OR      CRD-RECORD-COUNT = CRD-TABLE-SIZE
001580         READ CRD-FILE
001590         AT END
001600             SET END-OF-CRD        TO TRUE
001610             NOT AT END
001620             ADD 1                TO CRD-RECORD-COUNT
001630             MOVE CRD-RECORD      TO CRD-OBJECT-RECORD
001640                                 (CRD-RECORD-COUNT)
001650         END-READ
001660     END-PERFORM.
001670
001680     CLOSE CRD-FILE.
001690
001700     IF NOT END-OF-CRD
001710         DISPLAY "CRD-OBJECT-AREA TABLE IS FULL!"
001720         DISPLAY "CURRENT SIZE IS: " CRD-TABLE-SIZE
001730     STOP RUN
001740     END-IF.
001750
001760 A-EXIT.
001770     EXIT.
001780
001790
001800 C-CALL-RCFSTXML SECTION.
001810*****
001820*   This section executes the CALL to the interface and checks  *
001830*   the feedback code.                                           *
001840*****
001850 C-ENTER.
001860
001870     CALL "RCFSTXML"                USING CRD-OBJECT-AREA
001880                                     CRD-RECORD-COUNT
001890                                     TV-LISTINGS
001900                                     COBOL-RECORD-LENGTH

```

```
001910                                XML-DOCUMENT
001920                                XML-DOCUMENT-LENGTH
001930                                FEEDBACK-CODE
001940                                FEEDBACK-TEXT.
001950
001960      IF  FEEDBACK-CODE > ZERO
001970          DISPLAY "BAD RETURN FROM RCFSTXML - FEEDBACK CODE IS "
001980                                FEEDBACK-CODE
001990          DISPLAY "MESSAGE READS: "  FEEDBACK-TEXT
002000          STOP RUN
002010      END-IF.
002020
002030 C-EXIT.
002040      EXIT.
```

Structure Break

Structure Break processing was introduced with version 2.4 of the interface. It is used when an XML document is required to hold more than one hierarchy of information, giving rise to many-to-many relationships within the document. Alternatively, a structure break can be used when totals are to be reported at the end of the document which are not known when the initial population of the CRD is performed.

Normal Operation

The interface generator modules create XML elements from the series of data images passed to them in the COBOL Record Definition (CRD). For each call, the prior image of any changed fields are translated into XML elements along with any start/end tags necessary to ensure a well formed XML document. In addition to this, elements are generated in order to preserve the field relationships that exist for each image of the CRD passed to the subroutine.

Eg: The first call:

Fields in CRD	Content
03 TV-program.	
05 prog-name PIC X(20) .	"Sunday Night Live "
05 prog-cost PIC \$\$\$,\$\$9.99.	2500000

The second call:

Fields in CRD	Content
03 TV-program.	
05 prog-name PIC X(20) .	"Sunday Night Live "
05 prog-cost PIC \$\$\$,\$\$9.99.	1000000

The third call:

Fields in CRD	Content
03 TV-program.	
05 prog-name PIC X(20) .	"News "
05 prog-cost PIC \$\$\$,\$\$9.99.	1000000

Generates:

XML Document
<pre> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-cost>\$25,000.00</prog-cost> <prog-cost>\$10,000.00</prog-cost> </TV-program> <TV-program> <prog-name>News</prog-name> <prog-cost>\$10,000.00</prog-cost> </TV-program> </pre>

In the example above, "Sunday Night Live" has two costs associated with it and therefore two `<prog-cost>` elements are generated. This reflects the one-to-many relationship between `<prog-name>` and `<prog-cost>`. However, when "News" is passed to the interface with a `prog-cost` equal to the previous `prog-cost` (by pure coincidence) the interface does not just generate the `<prog-name>` element, instead it rebuilds the XML from the parent of `<prog-name>` and `<prog-cost>`. The interface does this because it records the fact that while "Sunday Night Live" existed in the CRD for both \$25,000.00 and \$10,000.00, "News" only existed for a cost of \$10,000.00. This is the desired outcome 99% of the time.

What is a Structure Break?

A Structure Break causes the generator subroutine to complete the creation of XML for data passed in the prior call and then wipe clean its record of what values previously existed on the CRD. Generation is then restarted by the next normal call.

How to code a Structure Break

A Structure Break is triggered by moving **LOW-VALUES** (binary zeroes) to the entire CRD area and calling the interface subroutine. On returning from the call, the generator will repopulate the CRD with all the data values that existed prior to the Structure Break call.

When to use a Structure Break

A Structure Break can be used to create "many-to-many" relationships between elements in an XML document. Usually, if a single data file is being used to create XML, only "one-to-many" relationships exist between the fields. For example, on one day there would be many channels, and one channel would have many programs, and one program would have many costs, etc. However, if the XML document is the result of more than one source file, "many-to-many" relationships may exist.

The example below shows how a Structure Break can be used to list television program information as well as all staff associated with a day's production, resulting in a "many to many" relationship between <TV-program> and <staff-name>.

Eg: The first call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-cost PIC \$\$\$,\$\$9.99.	2500000
05 channel-staff.	
07 staff-name PIC X(20) .	spaces

The second call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 channel-staff.	
07 staff-name PIC X(20) .	spaces

The third call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"News "
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 channel-staff.	
07 staff-name PIC X(20) .	spaces

The fourth call (Structure Break):

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	low values
07 prog-cost PIC \$\$\$,\$\$9.99.	low values
05 channel-staff.	
07 staff-name PIC X(20) .	low values

The fifth call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"News"
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 channel-staff.	
07 staff-name PIC X(20) .	John Smith

The sixth call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"News"
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 channel-staff.	
07 staff-name PIC X(20) .	Jane Jones

Generates:

XML Document
<pre> <channel> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-cost>\$25,000.00</prog-cost> <prog-cost>\$10,000.00</prog-cost> </TV-program> <TV-program> <prog-name>News</prog-name> <prog-cost>\$10,000.00</prog-cost> </TV-program> <channel-staff> <staff-name>John Smith</staff-name> <staff-name>Jane Jones</staff-name> </channel-staff> </channel> </pre>

In the previous example, the content of **staff-name** was spaces for all calls prior to the Structure Break, yet no empty element (`<staff-name/>`) was generated. This is because all record of the spaces in **staff-name** was erased by the Structure Break call. This leads to the other application for Structure Breaks: trailer totals. If an XML element is required at the end of the document, containing a control total of a numeric field, or even a hash total of several fields, this can be generated after a Structure Break, thereby delaying the need to populate the field in the CRD until the end of generation – when the value of the field is known.

The example below shows how a Structure Break can be used to generate the `<total-cost>` element for a day's programs only when the value of the total is known to the application.

Eg: The first call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-cost PIC \$\$\$,\$\$9.99.	2500000
05 total-cost PIC \$\$\$,\$\$9.99.	zero

The second call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"Sunday Night Live "
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 total-cost PIC \$\$\$,\$\$9.99.	zero

The third call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"News "
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 total-cost PIC \$\$\$,\$\$9.99.	zero

The fourth call (Structure Break):

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	low values
07 prog-cost PIC \$\$\$,\$\$9.99.	low values
05 total-cost PIC \$\$\$,\$\$9.99.	Low values

The fifth call:

Fields in CRD	Content
03 channel.	
05 TV-program.	
07 prog-name PIC X(20) .	"News"
07 prog-cost PIC \$\$\$,\$\$9.99.	1000000
05 total-cost PIC \$\$\$,\$\$9.99.	4500000

Generates:

XML Document
<pre> <channel> <TV-program> <prog-name>Sunday Night Live</prog-name> <prog-cost>\$25,000.00</prog-cost> <prog-cost>\$10,000.00</prog-cost> </TV-program> <TV-program> <prog-name>News</prog-name> <prog-cost>\$10,000.00</prog-cost> </TV-program> <total-cost>\$45,000.00</total-cost> </channel> </pre>

Data Integrity

Character Range

RCFSTXML accepts single byte characters in the hexadecimal range "00" through "FF". However, the use of hex "00" (null/low-values) has special meaning to the interface and these characters are converted to spaces before being passed to the XML document (see [Controlling the Point of Change](#) section). The low-values character is not within the XML character range defined by the W3C Extensible Markup Language (XML) 1.0 (Second Edition) definition. It is the application's responsibility to ensure that only characters within the permitted range of the XML protocol are generated.

Character References

Unicode character references (eg: î = î) may appear in entity declarations or as part of the data passed in the CRD. No attempt is made to interpret their character form.

Entity References

Data is usually transferred from CRD fields to XML elements without alteration. However, because certain characters are interpreted as instructions by XML parsers, these are automatically translated to their predefined entity references. The characters affected and their translations are listed below:

Character	Description	Entity Reference
>	greater than	>
<	less than	<
&	ampersand	&
'	apostrophe	'
"	double quote	"

RCFSTXML does not attempt to convert text in CRD fields to application specific entity references. If application specific entity references are required, they must be declared as general, internal, parsed entities in the DTD header table (see [DTD-headers](#) in the [User Maintained Variables](#) section) and passed to the interface in entity form (&xyz;). Under these circumstances the "&" character is passed unchanged to the XML document.

Empty Fields

An elementary alpha or alphanumeric field in the CRD is deemed empty if it contains only spaces and/or low values. An elementary numeric field in the CRD is deemed empty if it contains only zeroes and/or numeric edit characters ("£", "\$", "/", "*", ".", "-", "+", "CR", "DB" or ","). A group level CRD field is deemed empty if all of its subordinate fields are deemed empty and it contains no mandatory (required) attributes and it contains no mandatory (minOccurs>"0") child elements.

When an empty field has an override tag name ending with an "=" the attribute is not generated unless it is defined in the CRD as *required* (<<attrName=>>) – see [Required Attributes and Elements](#) for more details. When an empty field doesn't have an override tag name ending with an "=" an element is generated in the form of an empty tag (<tagName/>) unless it is a trailing occurrence in an array when it is not generated.

CDATA

CDATA sections can be generated in one of two ways. The first option is to use the COBOL STRING command to string the CDATA literals directly into the CRD field around a working storage variable:

Eg: STRING "<![CDATA[" WS-FIELD "]">" DELIMITED BY SIZE INTO CRD-FIELD.

Alternatively, the element can be defined to the CRD in a similar way to a mixed content element and the CDATA literals moved into the CRD or predefined using the **VALUE** clause:

Eg:

Fields in CRD			Content
03	TV-program.		
05	FILLER	PIC X(9) VALUE "<![CDATA[".	"<![CDATA[".
05	prog-name	PIC X(16) .	"News & Weather "
*	<>		
05	FILLER	PIC X(3) VALUE "]">".	"]]>"

Generates:

XML Document
<TV-program><![CDATA[News & Weather]]></TV-program>

Maximum Document Size

The Redvers COBOL XML Interface is designed to process XML documents up to 99,999,999 bytes in length. As this limit exceeds the maximum field size for most COBOL compilers, the picture clause for the the **XML-DOCUMENT** parameter in linkage is set to: `PIC X(9999999)` .

If a document length greater than 9,999,999 bytes is required, and if the platform can support a greater field length, the picture clause for the **XML-DOCUMENT** parameter in linkage may need to be changed from: `PIC X(9999999)` to a longer picture definition (up to 99,999,999 bytes).

Processing Instructions

It is not currently possible to generate processing instructions with this interface.

Comments

Comments can be coded into the start of the XML document using the **XML-headers** User Maintained Variable or the **DTD-headers** User Maintained Variable – See [User Maintained Variables](#) for details.

User Maintained Variables

One of the features of the Redvers COBOL XML Interface is that it is sold as COBOL source code. This means that certain parameters can be adjusted to suit the requirements of individual applications. These parameters are called User Maintained Variables and can be found in the "preserved" parts of the identification, environment and data divisions, marked by a following comment line beginning **<UMV>** with "*"s underlining the variable value.

NO PROCEDURE DIVISION CHANGES ARE NECESSARY.

These variables are defaulted to values that should be adequate in most circumstances while keeping storage requirements to a minimum.

Changes to User Maintained Variables in accordance with these instructions will not invalidate the warranty.

Program-Id

The program-id may be changed to suit site standards or to allow for multiple versions of RCFSTXML and RCFSTCMP with different User Maintained Variables.

SELECT Statements

In RCFSTCMP, external file names and other information specified in the SELECT statements can be changed to suit site standards and/or to satisfy platform compatibility requirements.

File Definition Statements

In RCFSTCMP, the input and output **FD** statements may be changed to suit site standards and/or to satisfy platform compatibility requirements. For example: "**BLOCK CONTAINS 0 CHARACTERS**" is frequently used on IBM platforms but not on HP platforms.

Maximum-COBOL-record-length

The length of this field dictates the maximum **COBOL-RECORD-LENGTH** (the logical length of the CRD) that can be passed to RCFSTXML in one call. This length must be increased if more than 4096 bytes of data are to be passed in the CRD or it may be decreased if storage is limited.

XML-declaration and XML-headers

These RCFSTXML text values provide the facility to control XML lines written at the start of the XML document (including the declaration). For example they can be used to include external DTDs or schemas in the document. Any single quote marks (apostrophes) within the texts are converted to double quote marks before being written. If populated, each text string generates a line on the XML document.

Additional header lines may be added by coding further **FILLER** fields of 100 bytes with text values, if the Number-of-XML-headers UMV is correspondingly increased.

Number-of-XML-headers

This value must be set to the number of XML declaration and header **FILLER** fields as described above.

DTD-headers

These RCFSTXML text values provide the facility to control XML lines written at the start of the DTD. For example they can be used to declare entity references. Any single quote marks (apostrophes) within the texts are converted to double quote marks before being written. If populated, each text string generates a line on the XML document.

Additional DTD header lines may be added by coding further **FILLER** fields of 100 bytes with text values, if the Number-of-DTD-headers UMV is correspondingly increased.

Number-of-DTD-headers

This value must be set to the number of DTD header **FILLER** fields as described above.

End-of-line-chtrs

In order to produce a more readable XML document, some applications may wish to include carriage return, line feed or new line characters at the end of each line in the XML document. The characters in this field provide this facility when they are populated with any hexadecimal values other than space. A single occurrence of any non-space character in this area will be appended to every logical line in the XML document.

Maximum-number-of-fields

The maximum number of discrete fields in the COBOL Record Definition (CRD) is defaulted to 400 in both RCFSTCMP and RCFSTXML (fields using the `occurs` clause are counted as one discrete field). If an application requires more than 400 fields in a single CRD, the number in the `occurs` clause for this UMV can be increased. Similarly, if storage is limited, this value can be decreased to save on storage requirements.

RCFSTCMP Compile Errors

The following error messages may be displayed by RCFSTCMP in the event that the input CRD could not be interpreted. If an error message is issued, no CRD object file will be produced.

Error Code	Error Text	Reason
+0101	NO FIELDS IDENTIFIED ON RECORD DEFINITION	No fields were identified on the input CRD .
+0102	TOO MANY FIELDS ON RECORD DEFINITION	The number of COBOL Record Definition fields on the input file exceeded the number allowed in the compiler's table. Increase this value. (See Maximum-number-of-fields in the User Maintained Variables section.)
+0103	PROCESSING EXCEPTION. PLEASE CONTACT REDVERS CONSULTING	There has been an internal logic error within the program. Please contact your Redvers Consulting account manager.
+0180	INVALID ACTIVATION KEY. PLEASE PLACE YOUR ACTIVATION KEY IN THE LAST W.S. FIELD	The Redvers COBOL XML Interface is supplied with a 32 character activation key. Please edit the RCFSTCMP source code and place this activation key in the VALUE clause literal for the last field definition in working storage. Then recompile.
+0190	30 DAY TRIAL PERIOD EXPIRED OR CALL LIMIT REACHED	The Redvers COBOL XML Interface can be downloaded free of charge for a thirty day trial period. The thirty days have now elapsed. Please contact Redvers Consulting for an additional thirty day trial or to arrange payment.
+0201	TOO MANY CHARACTERS FOR LEVEL NUMBER	RCFSTCMP was expecting a COBOL level number in the CRD file but found a string of more than two characters. <i>This message also points to the offending line number within the CRD file.</i>
+0202	INVALID LEVEL NUMBER	A COBOL level number that was not numeric or a level number greater than 49 but not 88 was found in the COBOL Record Definition. <i>This message also points to the offending line number within the CRD file.</i>
+0203	ILLEGAL DATA CLAUSE FOUND	An unsupported data clause was found in the field definition. See Clauses not Supported section. <i>This message also points to the offending line number within the CRD file.</i>

+0204	INVALID NUMBER OF OCCURRENCES	<p>An OCCURS clause was identified but it was not followed by a valid integer of more than zero and less than 9999.</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0205	INVALID PICTURE DEFINITION	<p>A PIC clause was identified but it was not followed by a valid string of characters less than 30 in length.</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0207	EXCEEDED MAXIMUM TAG SIZE	<p>An override tag name of more than 100 characters was encountered. The Superfast level interface is not designed to handle tag names of more than 100 characters. Please use one of the other Redvers COBOL XML Interface levels.</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0208	INVALID FILL CONSTANT	<p>An invalid override fill constant was found in an override tag name in the CRD. The value found didn't match one of the allowable fill constants: "L" for low-values, "H" for high-values, "S" for spaces, "Z" for zeroes and "Q" for quotes. (See "Override Fill Constants" section in the RCFSTCOB parser manual for details.)</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0209	INCOMPLETE FILL CONSTANT TAG	<p>An override fill constant was found in an override tag name in the CRD but the override was not completed with the ">" character. (See "Override Fill Constants" section in the RCFSTCOB parser manual for details.)</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0210	TAG DOES NOT START WITH AN ALPHABETIC CHARACTER	<p>All XML tags must start with an alphabetic character.</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0211	TAG NAME CONTAINS INVALID XML CHARACTERS	<p>XML tag names are confined to using only alphabetic, numeric, "-", "_", ":" or "." characters.</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>
+0214	ATTRIBUTES MAY NOT BE GROUP LEVEL ITEMS	<p>Attributes (tag names ending with "=") must be elementary data items.</p> <p><i>This message also points to the offending line number within the CRD file.</i></p>

+0215	ATTRIBUTE TAGS MAY NOT OCCUR MORE THAN ONCE	Attribute tags (those ending with "=") may occur only once in the parent/group level tag. <i>This message also points to the offending line number within the CRD file.</i>
+0220	GROUP ITEM MUST NOT HAVE A PICTURE	The field named in the message is at a higher level than the next field in the COBOL Record Definition and is therefore a group item. However, a picture clause was encountered for this field. <i>This error would normally have been caught in the compile stage.</i>
+0221	ELEMENTARY FIELD MUST HAVE A PICTURE	The field named is at an equal or lower level than the next field in the COBOL Record Definition and is therefore an elementary item. However, a picture clause was not given for this field. <i>This error would normally have been caught in the compile stage.</i>
+0222	MORE THAN ONE ROOT ELEMENT FOUND	The structure of the COBOL Record Definition may not have more than one root level. <i>This message includes the offending field tag name.</i>
+0224	NO TAG NAME FOR THE ROOT ELEMENT	A COBOL name or override tag name is required for the root element of any XML document.
+0225	THE ROOT ELEMENT CANNOT OCCUR MORE THAN ONCE	The maximum OCCURS value for a root element is 1.
+0226	INVALID POSITION FOR ATTRIBUTE	Attributes (tag names ending with a "=") must be coded within the group they relate to and they must be the first fields in that group. <i>This message includes the offending attribute name.</i>
+0227	MISSING GROUP TAG FOR ATTRIBUTE ELEMENT	The group level data item must have a tag if it is to hold an attribute (tag names ending with a "=") field. <i>This message includes the offending attribute name.</i>
+0230	ENCOUNTERED MULTIPLE DIMENSION ARRAY	An OCCURS clause greater than 1 has been nested within another OCCURS clause greater than 1. Only single dimension arrays are currently supported. To overcome this problem define a single dimension array and make multiple calls to the interface for each occurrence of the data item. <i>This message includes the offending field tag name.</i>

The +300 series indicate fatal errors encountered after the first call.

FEEDBACK -CODE	FEEDBACK-TEXT	Reason
+0101	NO FIELDS IDENTIFIED ON RECORD DEFINITION	The CRD-RECORD-COUNT parameter was zero. Probable causes are that the previous recompilation by RCFSTCMP failed in some way or the calling application hasn't loaded the compiled CRD in CRD-OBJECT-AREA .
+0103	PROCESSING EXCEPTION. PLEASE CONTACT REDVERS CONSULTING	There has been an internal logic error within the program. Please contact your Redvers Consulting account manager.
+0104	LENGTH OF RECORD DEFINITION IS > MAX COBOL RECORD	The total logical length of CRD-OBJECT-AREA , exceeds the size of the maximum COBOL-RECORD-LENGTH permitted. (See Maximum-COBOL-record-length in the User Maintained Variables section.)
+0110	RECORD DEFINITION / LINKAGE MISMATCH	The logical length of CRD-OBJECT-AREA was not the same as COBOL-RECORD-LENGTH passed in linkage. Probable causes are that the CRD needs to be recompiled by RCFSTCMP, the CRD has been changed but the calling program was not recompiled or that the layout used in the calling program is not the one in CRD-OBJECT-AREA .

+0180	INVALID ACTIVATION KEY. PLEASE PLACE YOUR ACTIVATION KEY IN THE LAST W.S. FIELD	The Redvers COBOL XML Interface is supplied with a 32 character activation key. Please edit the RCFSTXML source code and place this activation key in the VALUE clause literal for the last field definition in working storage. Then recompile.
+0190	30 DAY TRIAL PERIOD EXPIRED OR CALL LIMIT REACHED	The Redvers COBOL XML Interface can be downloaded free of charge for a thirty day trial period. This free version may be called up to 100 times in a single application execution. Either the thirty days have now elapsed or your application is trying to call RCFSTXML more than 100 times. Please contact Redvers Consulting for an additional thirty day trial or to arrange payment.
+0301	CALL SEQUENCE ERROR	A call has been made to RCFSTXML after processing has completed. Likely explanations are that FEEDBACK-CODE was not checked after a previous unsuccessful call or that the calling program has previously issued an end-of-run call (high-values in the CRD).
+0302	PARAMETERS MUST NOT CHANGE AFTER INITIAL CALL	One or more of the input calling parameters was found to have changed after the first call to the interface. Likely explanations are that the application has inadvertently overwritten one or more of the working storage fields used in the call to the interface or the application is attempting to generate more than one XML document without first issuing the final high-values call.
+0310	ATTEMPTED TO GENERATE MULTIPLE ROOT DOCUMENT	When generating repeating information the program found that the lowest group level common to the repeating information (the Point Of Change) was the root element and therefore cannot be repeated. To avoid this, code the root element as the only 01 level in the COBOL layout followed by a single 02 level so that repeating information can be generated under multiple 02 levels. If you don't want to pass this additional level to XML call it FILLER .
+0330	XML DOCUMENT LENGTH EXCEEDS DOCUMENT AREA	When building the XML document RCFSTXML has attempted to address beyond the length of the XML-DOCUMENT area in the calling program. Ensure that XML-DOCUMENT-LENGTH contains the full length of the XML-DOCUMENT area in the calling program. Otherwise the size of XML-DOCUMENT must be increased in the calling program.

Index

A

account manager, 6
activation key, 6
apostrophes, 6, 45
arrays, 13, 20, 21
ASCII, 12
attributes, 9, 16

B

base64, 12
Binary Fields, 12
binary zeroes, 36
BLANK WHEN ZERO, 13
BLOCK CONTAINS, 44

C

Calling, 27
Calls, 29
carriage return, 46
CDATA, 42
Character Range, 41
Character References, 41
Cloaking Device, 6
COBOL compiler, 6, 7
COBOL Record, 27
COBOL Record Definition, 8, 27
COBOL Record Length, 27, 29, 45
Comments, 43
Contact, 6
copybook, 19
CRD Object Area, 27
CRD Record Count, 27

D

Data Integrity, 41
dimensions, 21
Document Type Definition, 29
double quotes, 6

DTD, 8, 17, 18, 29, 45
DTD-headers, 43, 45
dummy fields, 23, 24

E

EBCDIC, 12
empty elements, 8, 13
empty fields, 8, 17, 18, 42
end tags, 8
End-of-line-chtrs, 46
entity declarations, 41
Entity References, 41
execution time, 5
Extensible Markup Language, 4, 41
external file names, 44

F

fatal, 50
FD, 44
Feedback Code, 28, 29, 50
Feedback Messages, 47, 50
Feedback Text, 28, 29, 50
field editing, 16
field names, 8
formatting, 16

H

hash total, 39
hierarchy, 14, 23
high-values, 29

I

imbedded sign, 11
include, 19
install, 6

J

JUSTIFIED RIGHT, 13

L

line feed, 46
low-values, 23, 36, 41

M

many-to-many, 37
Maximum Document Size, 43
Maximum-COBOL-record-length, 45
Maximum-number-of-fields, 46
minOccurs, 18, 42
mixed content, 15

N

namespaces, 10
new line, 46
non-tagged fields, 15, 16
non-tagged group, 15
non-tagged groups, 25
null, 41

O

OCCURS, 13, 18, 20
OCCURS DEPENDING ON, 13
one-to-many, 37
optional, 13
optional elements, 18
Orphan Repeats, 25
override tag names, 9, 10, 15, 16, 17

P

Packed Fields, 12
Parameters, 27
parent, 25
picture, 11, 17
Point of Change, 21, 23
Processing Instructions, 43
program-id, 6, 30, 44

Q

quote, 45

R

RCCMPRES, 31
RCFSTCMP, 4, 5, 6, 7, 27, 44, 47
RCFSTCOB, 4
REDEFINES, 13
repeating groups, 20
Repeating Groups, 20
required, 42
required attributes, 17
required elements, 17
root element, 14

S

Sample Program, 31
schema, 8, 17, 18
SELECT, 44
single quotes, 6
SOAP, 10
speech marks, 6
start tags, 8
STRING, 42
structure, 14
Structure Break, 35
SYNC, 13
SYNCHRONIZED, 13

T

tools, 8
totals, 35
trailer totals, 39

U

UMV, 44
Unicode, 41
User Maintained Variables, 6, 44

W

warranty, 44

World Wide Web Consortium (W3C), 4, 41

X

XML declaration, 6, 45

XML Document, 28

XML Document Length, 28, 29

XML-headers, 43, 45



Redvers Consulting Ltd

44 Broadway, London E15 1XH, UK
<http://www.redversconsulting.com/>