

# App Connect Enterprise コード変換カスタマイズ・ガイド

• 2024年3月29日 改訂

日本アイ・ビー・एम システムズ・エンジニアリング



# はじめに

- 本書は、IBM App Connect Enterprise(ACE) V12.0.10.0の調査結果・検証結果を基に作成しており、具体的な手順は便宜上、Windows版、Linux版での手順を示しています。
  - ◆ 一部、V12.0.12.0での修正点についても記載しています。
- 当資料の内容は、今後のFixpack適用やバージョンアップによって変更される可能性があります。
- 当資料はACEの基本的な知識と実装スキルを保有している方を対象とします。

当資料に記載している内容は可能な限り稼動確認を取っておりますが、日本アイ・ビー・エム株式会社、及び日本アイ・ビー・エムシステムズ・エンジニアリング株式会社の正式なレビューを受けておらず、当資料で提供される内容に関して日本アイ・ビー・エム株式会社、日本アイ・ビー・エムシステムズ・エンジニアリング株式会社は何ら保証するものではありません。従って、当資料の利用またはこれらの技法の実施はひとえに使用者の責任において為されるものであり、当資料によって受けたいかなる被害に関しても一切の補償をするものではありません。

# 内容

- ACEでのコード変換
- コード変換のカスタマイズ
  - ◆ コード変換テーブルのカスタマイズ例
  - ◆ マッピング・テーブルのカスタマイズ例

# ACEでのコード変換

# ACEでのコード変換

## ■ACEはメッセージ・ツリーの文字フィールドをUnicodeで扱う

- ◆メッセージのパーズ/組立て時に、入出力メッセージの文字コード⇔Unicode変換が実施される

## ■宛先システムに合わせて出力メッセージの文字コードを変更することが可能

- ◆入力メッセージのコード・ページ/エンコーディングはメッセージ・ツリーの下記エレメントにセットされる

- Root.Properties.CodedCharSetId : コード・ページ (CCSID)

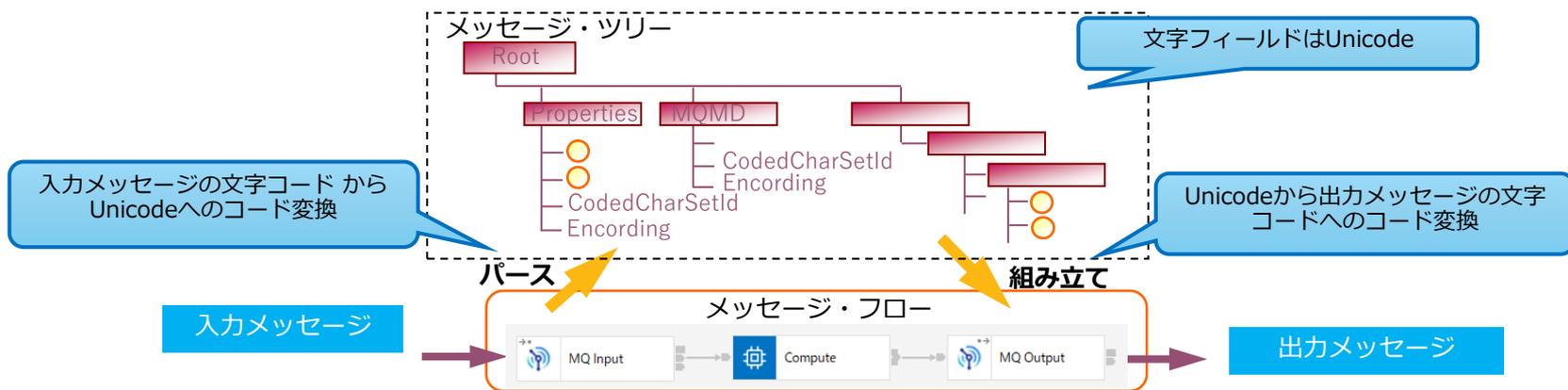
- Root.Properties.Encoding : エンコーディング

※MQメッセージの場合は下記エレメントにもセットされる

- Root.MQMD.CodedCharSetId、Root.MQMD.Encoding

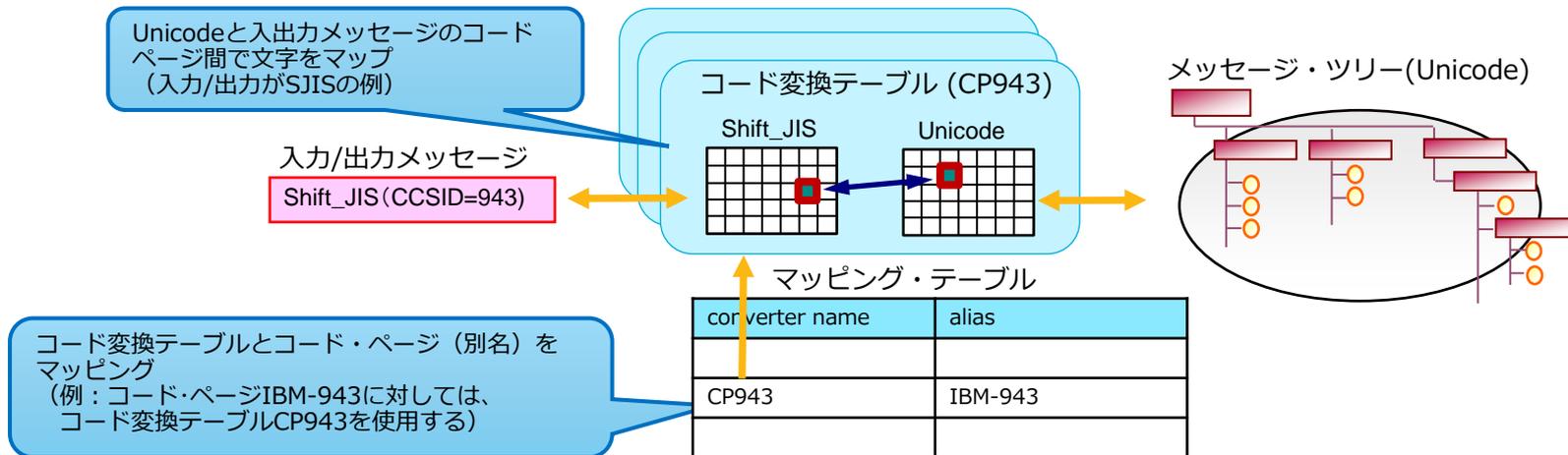
- ◆上記エレメント値を変更して出力することで、出力メッセージの文字コードを変更可能

- MQ連携の場合は基本的にMQMD下のエレメントを利用



# ACEでのコード変換

- ACEは内部的にICU(International Components for Unicode)をコード変換に使用
  - ◆ ICUは以下2つのテーブルを利用して、Unicode⇔他文字コードの変換を行う
    - コード変換テーブル : Unicode (メッセージ・ツリー) と入出力メッセージのコード・ページをマップ
      - ✓ 1文字ごとのコード・ポイントの対応表
      - ✓ サポートするコード・ページ毎にテーブルを持つ
    - マッピング・テーブル : コード変換テーブルとコード・ページをマップ
      - ✓ 入出力時に指定されたコードページに対し、どのコード変換テーブルを使用するかの対応表
  - ◆ サポートされるコード・ページの一覧については、以下URLを参照
    - <https://www.ibm.com/docs/en/app-connect/12.0?topic=flows-supported-code-pages>
    - ACEバージョンと使用されるICUバージョンの対応については、次ページを参照



# ACEバージョンとICUバージョンの対応

- ACEのバージョンやパーサーによって使用するICUのバージョンは異なる
  - ◆ DFDLパーサーとその他のパーサー（XMLNSCやMRMなど）は異なるバージョンのICUを使用
  - ◆ Fixpackの適用によってICUのバージョンが変わることもあるため、Fixpack適用時は要注意
    - 最新FixレベルではFixpackのページにDFDLパーサーとその他のパーサーで使用するICUバージョンの記載あり
      - ✓ <https://www.ibm.com/support/pages/node/7033840> (V12.0.10.0のサイト)

IIBバージョン	DFDL以外のパーサー	DFDLパーサー
ACE V11	ICU 72.1 (11.0.0.23時点)	ICU 70.1 (11.0.0.23時点)
ACE V12	ICU 72.1 (12.0.10.0時点)	ICU 70.1 (12.0.10.0時点)

- ◆ ICUをカスタマイズする際には、どのバージョンのICUが使用されるかを意識する必要がある
  - カスタマイズには使用するICUバージョンのツールの入手が必要

# (補足) ACEが使用するICUのバージョンの確認

- ◆ V12.0.12.0以降でのACEが使用するICUバージョンの確認方法
  - mqsiservice -t コマンド (DFDLパーサーと、DFDL以外のパーサーが使用するICUバージョンの両方を確認可能)

```
$ mqsiservice -t
BIPmsgs ja_JP
  Console OEM CP=932, ICU CCSID=943
  Default codepage=ibm-943_P15A-2003, in ascii=ibm-943_P15A-2003
  JAVA console codepage name=cp943c
```

ICU Version: 72.1

ICU data directory: C:¥z\_Test¥ACE10  
ICU TZ Data Version: 2022e

DFDLパーサー以外のICUバージョン

DFDL ICU Version: 70.1

DFDL ICU data directory: C:¥z\_Test¥ACE10  
DFDL ICU TZ Data Version: 2021a3

DFDLパーサーのICUバージョン

- mqsiservice -l コマンドでカスタマイズしたコード変換テーブル等が有効になっていることを確認可能

```
$ mqsiservice -l
:
ibm-941_P13A-2001: ibm-941 ibm-941_P13A-2001 (No JAVA aliases)
ibm-942_P12A-1999: cp932 ibm-932 ibm-932_VSUB_VPUA ibm-942 ibm-942_P12A-1999 ibm-942_VSUB_VPUA shift_jis78 sjis78 x-
IBM942 x-IBM942C (No JAVA aliases)
ibm-943_P130-1999: 943 Shift_JIS cp943 ibm-943 ibm-943_P130-1999 ibm-943_VASCII_VSUB_VPUA x-IBM943 (No JAVA aliases)
ibm-943_P15A-2003: IBM-943C MS_Kanji Shift_JIS cp932 cp943c csShiftJIS csWindows31J ibm-943 ibm-943_P15A-2003 ibm-
943_VSUB_VPUA ms932 pck sjis windows-31j windows-932 x-JISAutoDetect x-MS932_0213 x-ms-cp932 x-sjis (No JAVA aliases)
ibm-9447_P100-2002: cp1255 ibm-9447 ibm-9447_P100-2002 windows-1255 (No JAVA aliases)
:
```

# (補足) ACEが使用するICUのバージョンの確認

- ◆ V12.0.12.0より前のバージョンでのACEが使用するICUバージョンの確認方法

- DFDL以外のパーサー
  - ✓ mqsiservice -t コマンドを実行
    - コマンド実行例

```
$ mqsiservice -t
BIPmsgs ja_JP
Console CCSID=1208, ICU CCSID=1208
Default codepage=UTF-8, in ascii=UTF-8
JAVA console codepage name=UTF-8
```

```
ICU Version: 72.1
ICU TZ Data Version: 2022e
. . .
```

- DFDLパーサー
  - ✓ <install\_dir>/server/dfdl/bin/dfdljvarを実行
    - コマンド実行例

```
$ ./dfdljver
DFDL Java version information:

Build stream: dfdl4j
Build ID: DFDL1100-20230920-0925 (main/35e1410e1cee7a5c71127d4396cfb55247713282)
Build version: 1.1.3.1 (1.1.3.1)
Build platform: java8
Build type: production
ICU version: 70.1.0.0
```

# コード変換のカスタマイズ

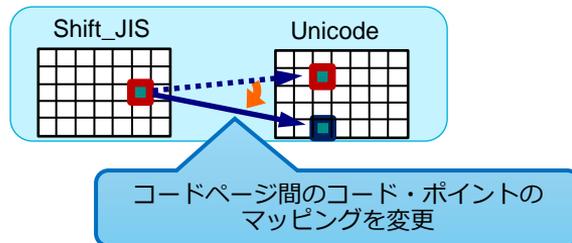
# コード変換のカスタマイズ

## ■ 要件に合わせてコード変換をカスタマイズ可能

- ◆ 特定の文字に対するコード・ポイントのマッピングを変えたい

⇒ コード変換テーブルをカスタマイズ

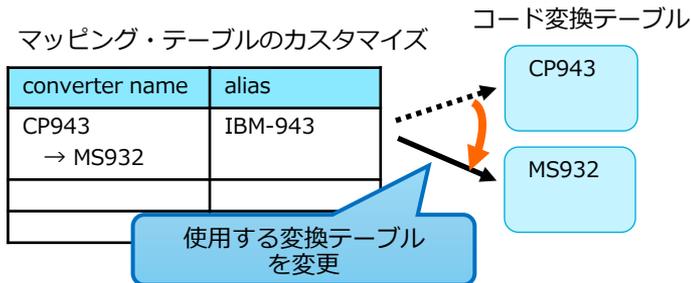
コード変換テーブルのカスタマイズ (Shift\_JIS → Unicodeの例)



- ◆ コード変換に使用される変換テーブルをデフォルトから変えたい

- Shift\_JIS⇔Unicodeのように複数のコード変換テーブルが選択可能なケースにて、  
連携システムに合わせて変換テーブルを変更する場合など

⇒ マッピング・テーブルをカスタマイズ



# コード変換のカスタマイズ

## ■ コード変換のカスタマイズ方法

- ◆ ICUのサイトより雛形となる変換テーブルやマッピング・テーブルをダウンロードして、カスタマイズ
- ◆ カスタマイズしたファイルをブローカー稼働環境でビルドして使用
  - ・ カスタマイズ対象のファイル名
    - ✓ カスタマイズ後、ICUツールを使用してバイナリ変換し、利用する

カスタマイズ対象	カスタマイズするファイル (テキスト)	ランタイムの使用ファイル (バイナリ)
コード変換テーブル	<コード・ページ名>.ucm (UniCode Mapping file)	<コード・ページ名>.cnv (binary converter file)
マッピング・テーブル	convrtrs.txt	cnvalias.icu (Converter Alias data file)

※DFDL以外のパーサー (MRM, XMLNSC等) とDFDL用のテーブルはバージョンが異なるため、それぞれダウンロードしてカスタマイズする必要があります

## ■ 参考

- ◆ ICUコード変換テーブル、マッピングテーブルはICUのサイトから入手  
<https://icu.unicode.org/charts/charset>
- ◆ ICUツール、およびマッピングテーブルについてはV12.0.12.0より製品同梱
  - ・ V12.0.12.0より前のバージョンではICUツールはIBM サポートよりAPAR IT38033を指定して入手が必要
- ◆ ICU ユーザーガイド  
<http://userguide.icu-project.org/>

# コード変換のカスタマイズ時の考慮点

- ACEのバージョン・アップ時にICUのバージョンが変更された場合は再ビルドが必要
  - ◆ ICUのバージョン間でコード変換テーブルのフォーマットが異なるため、互換性がない
- カスタマイズしたコード変換は、統合ノード実行環境でのみ有効
  - ◆ Integration Toolkit テストツールでの構文解析やシリアライズのテストでは、カスタマイズした変換を使用することはできない

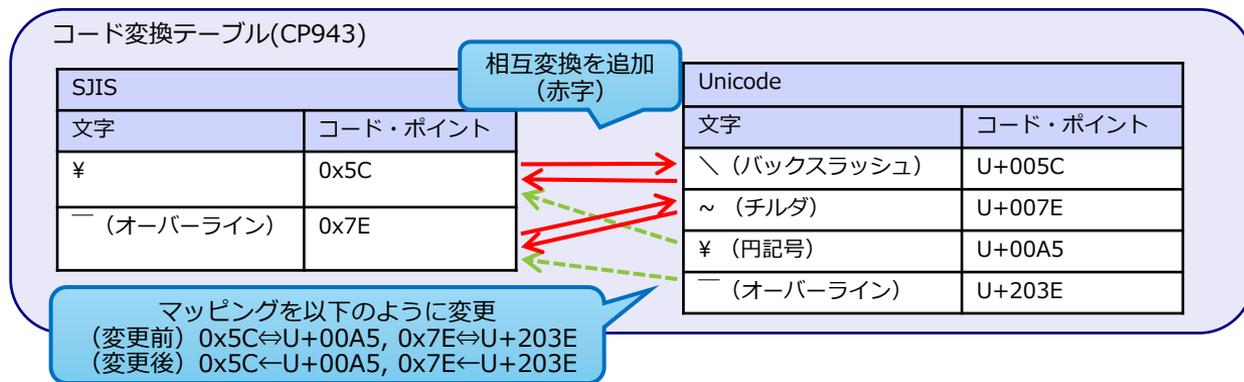
# コード変換のカスタマイズ例

## ■ コード変換のカスタマイズ例

- ◆ Shift\_JIS( CCSID943)⇔Unicodeの変換をカスタマイズ

### A) コード変換テーブルをカスタマイズするケース

- コード変換テーブル(CP943)のバックスラッシュ「\」、チルダ「~」のコードポイントを以下のように変更する
  - ✓ CP943C相当の変換テーブルとなるよう、CP943のテーブルをカスタマイズする



補足.

一般的にShift\_JIS(CCSID=943)⇔Unicode の変換テーブルとして、MS932, CP943C, CP943などがありますが、ICUではCP943 (デフォルト), MS932のみが選択可能です。(テスト環境での検証結果より)  
そのため、変換テーブルをCP943Cに変更する場合は、上記例のようにCP943をベースとしたコード・ポイントの変更で対応する必要があります。

# コード変換のカスタマイズ例

## B) マッピング・テーブルをカスタマイズするケース

- コード変換テーブルをCP943（デフォルト）からMS932に変更する
  - ✓ ICUではCP943はibm-943-P130-1999、MS932はibm-943-P15A-2003で表される



# コード変換のカスタマイズ例

## ■ コード変換テーブルのカスタマイズ、マッピング・テーブルのカスタマイズの手順一覧

### ◆ 手順1, 6は共通

手順	A) コード変換テーブルのカスタマイズ手順	B) マッピング・テーブルのカスタマイズ手順
1	ICU関連ツールの入手とインストール（※V12.0.12.0以降では製品同梱のため、対応不要）	
2	マッピング・データ・ファイル(.ucm)のダウンロードとファイル名の変更	マッピング・テーブル(.txt)のダウンロード
3	マッピング・データ・ファイル(.ucm)の修正	マッピング・テーブル(.txt)の編集
4	マッピング・データ・ファイル(.ucm)をバイナリ(.cnv)に変換	作成したマッピング・テーブル・ファイル(.txt)をバイナリ(.icu)に変換
5	マッピング・データ・ファイル(.cnv)を統合ノードがアクセス可能なディレクトリに配置	マッピング・テーブル・ファイル(cnvalias.icu)を統合ノードがアクセス可能なディレクトリに配置
6	（ICU_DATA環境変数を使用する場合）ファイルの配置ディレクトリのパスの設定	

# ICUツールの準備

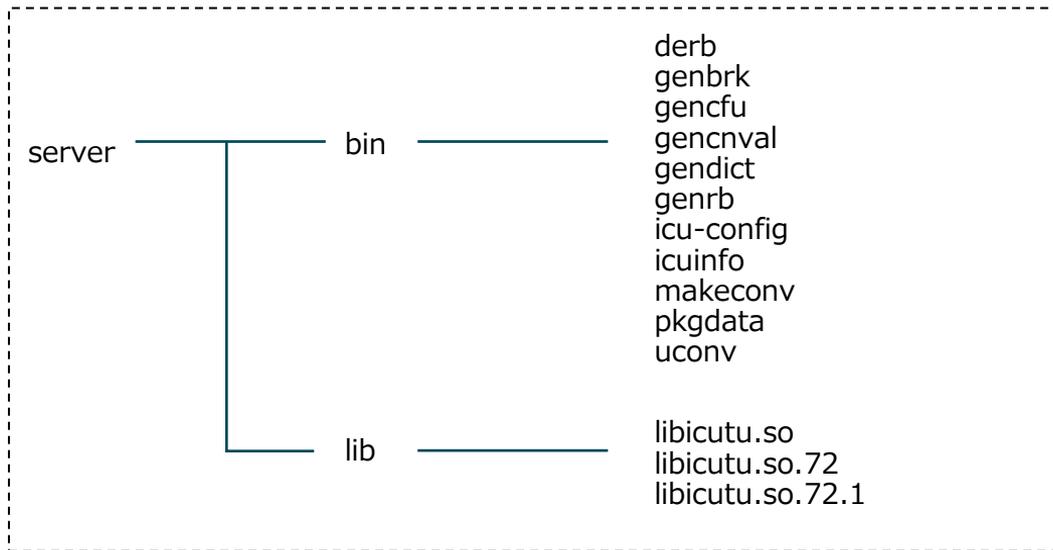
A) コード変換のカスタマイズ

B) マッピング・テーブルのカスタマイズ

## 1. V12.0.12.0より前のバージョンではICU関連ツールを入手

- ◆ IBM サポートからAPAR IT38033 を指定して、使用システムに応じたユーティリティを入手
  - APARに含まれるREADMEを参照し、ツールをインストールする

IT38033に含まれるツール群 (例)



- ◆ V12.0.12.0からICU関連ツールを製品同梱

# A) コード変換テーブルの準備

## 2. マッピング・データ・ファイル(.ucm)のダウンロードとファイル名の変更

- ◆ CCSID=943の雛形にするマッピング・データ・ファイルをダウンロード

<https://icu.unicode.org/charts/charset>

ibm-943\_p130-1999.ucm をダウンロード

- ◆ ダウンロードしたファイルのファイル名を「ibm-ccsid.ucm」のネーミングに変更

- 特定のCCSIDは `ibm-ccsid` で検索される
- サポートされるCCSIDは以下で確認

<https://www.ibm.com/docs/en/app-connect/12.0?topic=flows-supported-code-pages>

ibm-943\_p130-1999.ucm を `ibm-943.ucm` に名前を変更

# A) コード変換テーブルの編集

## 3. マッピング・データ・ファイルの修正

- ◆ マッピング・データ・ファイルのフォーマット、詳しくは以下のICUユーザー・ガイドを参照

<http://userguide.icu-project.org/conversion/data>

文字「\」を「¥」、「~」を「—」に置き換えるようマッピング・データ・ファイルを修正

	\	~	¥	—
UTF-16	005C	007E	00A5	203E
943	(なし)	(なし)	5C	7E

(以下、ibm-943.ucmより、該当する記述を抜粋)

修正前

```
<U005C> ¥x7F |2
:
<U007E> ¥x7F |2
:
<U00A5> ¥x5C |0
:
<U203E> ¥x7E |0
:
```

UNICODE SJIS(hex)

サポートする変換の向き  
 0: 相互  
 1: UNICODE → xxx  
 2: マップなし (ヘッダーに記載の subchar1で代替)  
 3: xxx → UNICODE

修正後

```
<U005C> ¥x5C |0 → ¥
:
<U007E> ¥x7E |0 → —
:
<U00A5> ¥x5C |1 → ¥
:
<U203E> ¥x7E |1 → —
:
```

変換の向きを数値で表す

マッピングを以下のように変更

(変更前) 0x5C ⇔ U+00A5, 0x7E ⇔ U+203E  
 (変更後) 0x5C ← U+00A5, 0x7E ← U+203E  
 0x5C ⇔ U+005C, 0x7E ⇔ U+007E

# A) コード変換テーブルの編集

## ■ (補足) マッピング・データ・ファイルのフォーマット

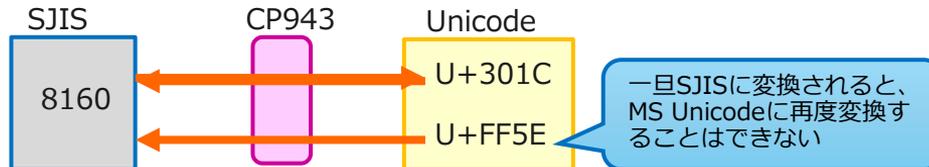
- ◆ マッピング・データ・ファイルでは、コード・ポイント間の変換の向きが数値で表される
  - 相互変換「0」、片方向変換「1」、「3」で異なる番号が使用される

番号	変換の向き	備考
0	相互	
1	Unicode → 他文字コード	DFDLではサポートされない (次ページ参照)
2	マップなし	ヘッダーに記載のsubchar1で代替される
3	他文字コード → Unicode	

- マッピングの向きがコード・ポイントにより異なる例
  - ✓ マッピングテーブルCP943における、SJIS(CCSID943)とUnicodeのマッピングでは、文字「~」は、Unicodeにおいて、コードポイントが2つ存在する

文字	SJIS コード	IBM Unicode	MS Unicode
~	8160	U+301C	U+FF5E

- ✓ IBM Unicodeである「U+301C」はSJISと相互に変換可能であるが、MS Unicode「U+FF5E」については、UnicodeからSJISへの片方向の変換のみサポートされる



# A) コード変換テーブルの編集

✓ (該当文字のマッピング・データ

```
<U301C> ¥x81¥x60 |0
```

..

```
<UFF5E> ¥x81¥x60 |1
```

U301Cは相互の変換なので"0"

UFF5EはUnicode→SJISの片方向  
変換なので"1"

## ◆ DFDLパーサー使用時の注意点

- DFDLでは、Unicode → 他文字コード変換において、マッピング方向の番号「1」がサポートされない
- ⇒ CP943における「U+FF5E」のように、Unicode → 他文字コード変換のみ行う文字すべてについて、番号を「4」に置き換える必要がある

```
<U301C> ¥x81¥x60 |0
```

..

```
<UFF5E> ¥x81¥x60 |4
```

DFDLパーサーを使用する場合、  
Unicode → 他文字コードの片方  
向の変換には番号「4」を指定

# A) コード変換テーブルのビルド

## 4. 作成したマッピング・データ・ファイル(.ucm)をバイナリ(.cnv)に変換

- ◆ DFDL以外のパーサー用とDFDLパーサー用で異なるコマンドを提供 (V12.0.12.0以降)

- makeconv: DFDL以外のパーサー用
- dfdl\_makeconv: DFDLパーサー用

- ◆ makeconv -v -d destination\_directory mapping\_file.ucm

```
$ makeconv -v -d /tmp/icucust ibm-943.ucm
```

⇒ ibm-943.cnv が生成される

- ◆ dfdl\_makeconv -v -d destination\_directory mapping\_file.ucm

```
$ dfdl_makeconv -v -d /tmp/icucust ibm-943.ucm
```

⇒ ibm-943.cnv が生成される

### makeconvコマンド出力例

```
the total number of offsets is 0x2c10=11280
compacting toUnicode data saves 856 bytes
fromUnicode number of uint32_t in stage 2:
0x799=1945
fromUnicode number of 2-byte stage 3 mapping entries:
0x6050=24656
Warning: ibm-943.cnv claims to be 'ibm-943_P130-
1999'
- Opened udata ibm-943.cnv
size of extension data: 3876
- Wrote 89892 bytes to the udata.
ibm-943.cnv
```

# B) マッピング・テーブルの準備、編集

B) マッピング・テーブルのカスタマイズ

## 2. マッピング・テーブルのダウンロード

- ◆ V12.0.12.0以降は製品同梱
- ◆ V12.0.12.0より前のバージョンでは雛形にするマッピング・データ・ファイルをダウンロード  
<https://github.com/unicode-org/icu/blob/main/icu4c/source/data/mappings/convrtrs.txt>

## 3. マッピング・テーブルの編集

- ◆ ダウンロードしたマッピング・テーブルを編集する

CCSID=943の時に使用するコード変換テーブルを“ibm-943\_P15A-2003”に設定

修正内容

ibm-xxxはCCSIDxxxに対応

```
ibm-943_P15A-2003 { UTR22* }  
  ibm-943 { IBM* JAVA }  
  # ibm-943  
  Shift_JIS { IANA* MIME* WINDOWS JAVA }  
  :
```

CCSID=943として使用するucmファイルのエントリーに {IBM\* JAVA}タグをつける  
ibm-943のエントリーはコメントアウト

```
ibm-943_P130-1999 { UTR22* }  
  # ibm-943 { IBM* JAVA }  
  Shift_JIS # Leave untagged because this isn't the default  
  cp943 { JAVA* } # This is slightly different, but the backslash mapping is the same.  
  :
```

ibm-943\_P130-1999のibm-943のエントリーはコメントアウト

# B) マッピング・テーブルの準備、編集

B) マッピング・テーブルのカスタマイズ

※DFDLパーサー使用時の注意点

DFDLパーサー用にカスタマイズする場合に以下のように変更する（ISE検証結果より）

修正内容

```
ibm-943_P15A-2003 { UTR22* }  
  ibm-943 { IBM* JAVA }  
  # ibm-943  
  943 {JAVA}  
  Shift_JIS { IANA* MIME* WINDOWS JAVA }  
  :  
  
ibm-943_P130-1999 { UTR22* }  
  # ibm-943 { IBM* JAVA }  
  Shift_JIS # Leave untagged because this isn't the default  
  cp943 { JAVA* } # This is slightly different, but the backslash mapping is the same.  
  # 943 {JAVA}  
  :
```

943 {JAVA} を追加

943 {JAVA} の行をコメントアウト

## B) マッピング・テーブルのビルド

B) マッピング・テーブルのカスタマイズ

### 4. 作成したマッピング・テーブル・ファイル(.txt)をバイナリに(.icu)に変換

- ◆ DFDL以外のパーサー用とDFDLパーサー用で異なるコマンドを提供（V12.0.12.0以降）
  - gencnval: DFDL以外のパーサー用
  - dfdl\_gencnval: DFDLパーサー用

- ◆ gencnval -v -s source\_directory -d destination\_directory

```
$ gencnval -v -s /tmp/icucust -d /tmp/icucust
```

⇒ cnvalias.icu が生成される

- ◆ dfdl\_gencnval -v -s source\_directory -d destination\_directory

```
$ dfdl_gencnval -v -s /tmp/icucust -d /tmp/icucust
```

⇒ cnvalias.icu が生成される

gencnvalコマンド出力例

```
/work/icutest/convtrs.txt:563: warning: duplicate aliases cp858 and CP00858 found for
standard JAVA and converter ibm-858_P100-1997
:
folding cp932 into cp932 -> ibm-943_P15A-2003 (alias conflict)
:
```

※編集したibm-943についてのエラーがなく、cnvalias.icuファイルが生成されればOKです。

# 所定のディレクトリへのファイルの配置

A) コード変換のカスタマイズ

B) マッピング・テーブルのカスタマイズ

## 5. 作成したビルド済みファイルを所定のディレクトリにコピー

### ◆「<導入ディレクトリ>/server/lib/data/icudt<version><platform-suffix>」に配置

•<version> バージョン指定

✓V72.1の場合は「72」

•<platform-suffix> l : リトル・エンディアン ASCII プラットフォーム (Windows, Linux/x86など)

b : ビッグ・エンディアン ASCIIプラットフォーム (AIX, Linux/zなど)

e : EBCDICプラットフォーム

※DFDL以外のパーサーとDFDLを両方使用する場合、icudt72l, icdt70lなど、それぞれのICUバージョンに応じたディレクトリを用意する必要があります。

### ◆マッピング・データ・ファイル(.cnv) を配置する場合

```
$ cp -p /tmp/icucust/ibm-943.cnv /opt/IBM/ace-12.0.10.0/server/lib/data/icudt72l
```

### ◆マッピング・テーブル・ファイル(cnvalias.icu) を配置する場合

```
$ cp -p /tmp/icucust/cnvalias.icu /opt/IBM/ace-12.0.10.0/server/lib/data/icudt72l
```

※V12.0.12.0現在、上記の所定ディレクトリに配置しても、うまく読み込まれないという製品障害が発生しています。現時点では以下のディレクトリに配置します。

Linux環境 : <導入ディレクトリ>/server/xml4c/data/icudt<version><platform-suffix>

Windows環境 : <導入ディレクトリ>/server/icudata/icudt<version><platform-suffix>

詳細はIBMサポート部門にお問い合わせください。また次ページのICU\_DATA環境変数を使用することで問題を回避することも可能です。

# (参考) 任意のディレクトリへの配置

A) コード変換のカスタマイズ

B) マッピング・テーブルのカスタマイズ

ICU\_DATA環境変数を使用し、任意のディレクトリに配置することも可能

- ◆ 既存バージョンからの移行で、ICU\_DATAを使用している場合など
- ◆ デフォルトでICU\_DATAは無効になっているため、ICU\_DATAの有効化が必要
  - MQSI\_DO\_NOT\_SET\_ICUDATA\_DIR 環境変数を“yes”に設定
- ◆ ICU\_DATA環境変数で配置ディレクトリを指定
  - 配置ディレクトリは「icudt<version><platform-suffix>」
    - ✓ ネーミングは前ページを参照
  - ICU\_DATA環境変数では「icudt <version>< platform-suffix>」フォルダの上位フォルダを指定

```
$ export MQSI_DO_NOT_SET_ICUDATA_DIR=yes  
$ export ICU_DATA=/var/mqsi/converters
```

・MQSI\_DO\_NOT\_SET\_ICUDATA\_DIR環境変数および、ICU\_DATA環境変数はブローカーが認識する必要があるため、ブローカーの起動シェル等で mqsiprofileの後に指定してください。