

4 章 漢字等の多オクテット文字機能

この章では、COBOL における漢字等の多オクテット文字の扱いについて説明します。国際符号化文字集合(UCS)等の多オクテット文字を標準的にサポートするため、データ定義などに拡張がされています。

4.1 文字集合

COBOL における文字集合には、COBOL 文字一覧、計算機コード化文字集合、そして符号系の3種類があります。

COBOL 文字一覧は、言語構文を定義するために使用される文字の一組で、従来の COBOL 言語と同様です。

計算機コード化文字集合は、実行時に内部処理用のデータを表現するために使用され、ソースコード中ではコメントや16進表現でない定数の内容が、計算機コード化文字集合で表現されます。この文字集合は、USAGE DISPLAY と記述された計算機英数字コード化文字集合と、USAGE NATIONAL と記述された計算機国別文字コード化文字集合とがあります。

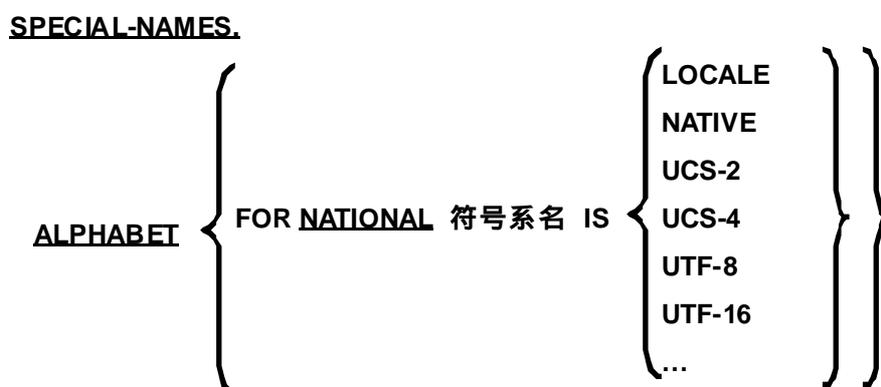
符号系は、名前のついた、コード化文字集合や照会順序の仕様です。特殊名段落で符号系を命名したり、利用者定義のコード化文字集合や照会順序を指定することができます。

4.1.1 特殊名段落 (SPECIAL-NAMES)

特殊名段落では符号系名をコード化文字集合や照会順序に関係付けることができます。このとき、国別文字コード化文字集合として、UCS-2、UCS-4、UTF-8、UTF-16 を指定することができます。

4.1.1.1 符号系名句 (ALPHABET)

ある名前を、コード化文字集合か照会順序、あるいはその両方に関連付けることができます。符号系名を実行用計算機段落の PROGRAM COLLATING SEQUENCE 句において、あるいは SORT 文や MERGE 文の COLLATING SEQUENCE 指定において参照した場合、符号系名は照会順序を表します。字類条件中や、データ部の CLASS 句、SYMBOLIC CHARACTER 句、あるいはファイル記述句の CODE-SET 句において参照した場合には、コード化文字集合を表します。



4.2 データ部

データをできるだけ機種によらないものにするために、データの特長や性質を記述するのに、「図形文字」を使用して表現される標準データ形式を用います。

実行時に計算機上で別なコード化文字集合が有効となる場合、定数は計算機コード化文字集合に変換されます。

ファイル入出力では、論理レコードと物理単位の間で転記が行われる際に、CODE-SET 句が指定されている場合、必要とされる変換、埋め文字の追加・削除が行われます。

4.2.1 国別文字定数

書き方 1

$$\left\{ \begin{array}{l} N'' \left\{ \text{文字 1} \dots \right\}'' \\ N' \left\{ \text{文字 1} \dots \right\}' \end{array} \right\}$$

書き方 2

$$\left\{ \begin{array}{l} NX'' \left\{ \begin{array}{l} 16 \text{ 進文字並び} \\ 1 \end{array} \right\}'' \\ NX' \left\{ \begin{array}{l} 16 \text{ 進文字並び} \\ 1 \end{array} \right\}' \end{array} \right\}$$

4.2.2 データ記述項

4.2.2.1 PICTURE 句

国別文字項目を定義する場合には PICTURE 文字列として記号「N」だけを使用します。国別文字編集項目を定義するには、PICTURE 文字列は記号「N」、「B」、「0」、「/」の組み合わせで記述し、「N」が少なくとも 1 つと「B」、「0」、「/」のどれかが共に含まれている必要があります。

4.2.2.2 USAGE 句

計算機の記憶域内でデータ項目を表すために国別文字コード化文字集合を使用することおよび、データ項目を文字境界に桁詰めすることを指定するために、USAGE NATIONAL を指定します。桁詰めにより、暗黙の FILLER ビット位置が生成されることがあります。国別文字は、計算機の記憶域中で表現される場合に、計算機の英数字文字集合の文字の大きさ以上の、一定の大きさを持つ文字表現でなければなりません。

4.3 組込み関数

国別文字コードを扱うために次の組込み関数が用意されます。

- CHAR-NATIONAL
- DISPLAY-OF
- NATIONAL-OF

また、次の組込み関数で引数として国別文字コードが使用できます。

- BYTE-LENGTH
- LENGTH
- LOCALE-COMPARE
- LOCALE-DATE
- LOCALE-TIME
- LOWER-CASE, UPPER-CASE
- MAX, MIN
- NUMVAL, NUMVAL-C, NUMVAL-F
- ORD, ORD-MAX, ORD-MIN
- REVERSE
- STANDARD-COMPARE
- TEST-NUMVAL, TEST-NUMVAL-C, TEST-NUMVAL-F

4.3.1 CHAR-NATIONAL 関数

FUNCTION CHAR-NATIONAL_(引数)

国別文字プログラム照会順序中において、引数の値に等しい順序位置にある国別文字の 1 文字を返します。

4.3.2 DISPLAY-OF 関数

FUNCTION DISPLAY-OF_(引数 1 [引数 2])

引数中の国別文字を混在英数字外部データ形式へと変換した値を含んでいる文字列を返します。

4.3.3 NATIONAL-OF 関数

FUNCTION NATIONAL-OF_(引数 1 [引数 2])

引数中の文字を国別文字の内部表現に変換した値を含んでいる文字列を返します。

