

COBOLにおける日本語機能の現状と今後の動向

* 床分 眞一 * 今城 哲二 *

(株) 日立製作所 ソフトウェア工場

プログラミング言語における日本語機能は、日本語データ処理の発展に伴い、必要不可欠の条件になっている。COBOLの日本語機能は、他のプログラム言語に先駆けて昭和54年から順次各メーカーの処理系がサポートし、今日に至っている。

一方、日本語機能の言語仕様は、処理系ごとに異なっているため、プログラムの可搬性が重要な問題になっており、標準化が望まれている。また、2バイト文字を必要とする国々（韓国や中国など）の適用や、欧米で開発されたCOBOLプログラムを日本国内でも販売・利用するため、それらのプログラムの日本語への適用も必要となってきた。

COBOLにおける日本語機能の標準化の動きは、ISO/ANSI COBOL WGの要請により、CODASYL COBOL委員会が、1988.2より検討を開始した。また、それと対応し、情報処理学会 SC22/COBOL WG も日本語機能の日本案を作成し、1988.5のCODASYL COBOL委員会に提示することになっている。

本報告では、COBOLにおける日本語機能の現状を紹介し、現状での問題点を明確にし、今後の動向としてCODASYL COBOLに対する日本語機能（案）について説明する。

* 情報処理学会 SC22/COBOL WG 委員 ** 情報処理学会 SC22/COBOL WG 主査, 日本語専門委員会 委員

Multi-Byte Character Facility
for
COBOL in Japan

Shinichi TOKOWAKE and Tetsuji IMAJYO

Software Works, Hitachi, Ltd.

549-6 Shinano Machi Totsuka Ku Yokohama-shi
Kanagawa Ken, 244 JAPAN

Multi-Byte Character Facility of programming language is getting very important with development of multi-byte character data management in Japan. Multi-Byte Character Facility of COBOL takes the lead. Each implementor has been providing extensions to COBOL since 1979.

We find problem of it, which specification of facility is independent of implementor. It is no portability to each system. We should make it "STANDARD SPECIFICATION".

We should think about this facility for Korean and Chinese character which need two-byte for represent one letter, and the application for Multi-Byte Character Facility by using COBOL program made in America or Europe.

Standardization of Multi-Byte Character Facility of COBOL has been examined by CODASYL COBOL Committee since Feb.1988. Japanese SC22/COBOL WG is making a technical paper now and the WG will propose it to CODASYL COBOL Committee at May 1988.

This presentation is report of presenting the condition of Multi-Byte Character Facility, cleaning the problems, and explaining the future trend of the technical paper for CODASYL COBOL.

1. はじめに

プログラミング言語COBOLは、周知の通り事務処理用言語である。日本語機能は、読み易さや分かり易さというユーザ・ニーズに答えるかたちで、他のプログラミング言語に先駆けて、昭和54年から各メーカーの処理系が順次サポートし、今日に至っている。

その間、COBOLにおける日本語機能の要求も、本機能の使用頻度の増加に伴い次のように変遷している。

- (1) 出力帳票の読み易さや分かり易さ
- (2) プログラムの読み易さや分かり易さ
- (3) プログラムの可搬性

特に項(3)は、近年重要な問題となっており、標準化が望まれている。

また、2バイト文字を必要とする国々(韓国や中国など)の適用の拡大や、欧米で開発されたCOBOLプログラムを日本国内でも販売・利用するため、それらのプログラムの日本語への適用も必要となってきた。

本稿では、第2章で日本語機能サポートの現状を紹介し、第3章でその問題点を明確にし、第4章で標準化の動向について述べる。

なお、第2章以降、「日本語機能」を「多バイト文字機能」と記述する。

2. 多バイト文字機能サポートの現状

2.1 多バイト文字のサポート範囲

COBOLにおける多バイト文字のサポート範囲は、次の4つのレベルに分けることができる。

レベル1: データとして多バイト文字が扱える。

レベル2: レベル1に加えてデータ名や手続き名などの利用者語(識別子)に多バイト文字が使用できる。

レベル3: レベル2に加えてMOVE, ADDなどの予約語(構文)も多バイト文字とする。(プログラムテキストをすべて多バイト文字とする。)

レベル4: レベル3に加えてデータもすべて多バイト文字とする。

以上のレベルをコード系で見ると次のようになる。

- ① レベル1, 2及び3は、単バイト文字のコード系と多バイト文字のコード系との併用
- ② レベル4は、単バイト文字のコードを廃止し、多バイト文字だけのコード系の採用

現状の多バイト文字の範囲は、レベル2が主流を占めており、第4章で述べる標準化もレベル2の範囲までである。

2.2 多バイト文字の言語仕様の概要

現状の多バイト文字の言語仕様は、第3次COBOL規格(ANS/ISO-85, JIS-88)の仕様に多バイト文字項目を追加している。主な機能について表-1に示す。

表-1 多バイト文字の主な言語仕様

項番	項目	内容
1	多バイト文字データ項目のPICTURE文字とUSAGE句	PICTURE文字は、N, X, J, K及びGの5種類。 USAGEはDISPLAY, DISPLAY-1及びDISPLAY-2の3種類。
2	多バイト文字定数	NC"~", "~", K"~", CC"~" 及び G"~" がある。
3	WRITE文におけるシフトコード挿入指定	APPLY句やCHARACTER TYPE句などで指定 あるいは、何も指定せず。
4	READ文におけるシフトコード削除	削除可及び削除不可がある。大半の処理系は削除不可。
5	多バイト文字データ項目の大小比較	大小比較ができるものとできないものがある。
6	利用者語の多バイト文字の文字数	最大文字数は15, 14 及び 13がある。

3. 現状の問題点

現状の多バイト文字機能サポートの問題点を以下に述べる。

(1) 仕様の不統一

多バイト文字機能は、機能面から見ると大筋一致しているが、表-1に示したように、同じ機能を記述する場合、その言語仕様が処理系ごとに異なっている。このため、プログラムの可搬性が阻害されている。

(2) シフトコードの扱いが不明確

多くの処理系がシフトコードを持つコード系を採用しているため、シフトコードの明確化が必要である。明確にすべき点は、次の項目であろう。

- ① 内部データ及び外部データとシフトコード
- ② プログラムテキストとシフトコード

4. 標準化の動向

4.1 標準化の検討経緯

COBOLにおける多バイト文字機能検討の経緯を表-2に示す。

表-2 多バイト文字機能検討の経緯

時 期	内 容
1979 ~	各処理系がCOBOLの多バイト文字機能を順次サポート。
1984.4 ~ 1985.3	電子工業振興協会 標準化専門委員会 国産7社及び海外4社の多バイト文字機能の仕様を調査し、COBOL第2次規格(ANS-74, ISO-78, JIS-80)をベースに「日本語COBOL」仕様案を作成。
1985.4 ~ 1986.3	電子工業振興協会 標準化専門委員会 COBOL第3次規格(ANS-85, ISO-85, JIS-88)対応に、上記「日本語COBOL」仕様を拡張。
1987.1 ~ 1988.3	情報処理学会 日本語専門委員会 ソフトウェア全般にわたる多バイト文字機能のあり方について検討し、ガイドラインを作成。
1987.9	ISO/IEC JTC1/SC22 日本の提案である「標準言語では、多バイト文字機能のサポートが必要である。」件が決議された。
1987.10	CODASYL COBOL委員会 ISOとANSIのWGからの要請により、多バイト文字の検討を最優先課題とすること決定。
1987.11 ~	情報処理学会 SC22/COBOL WG 「COBOL 多バイト文字機能」仕様案を作成し、CODASYL及びISOに日本から提案することを決定し、検討を開始。
1988.2 ~	CODASYL COBOL委員会 多バイト文字の検討を開始。なお、COBOL委員会での仕様承認目標時期は、1989.12である。

4.2 CODASYL COBOL 多バイト文字機能

COBOLの規格仕様は、CODASYL COBOLの完全な部分集合とする習慣であるため、多バイト文字機能の標準化（ISO化）のためには、まず、CODASYL COBOLでの多バイト文字機能のサポートが必要である。日本のSC22 COBOL WGからCODASYL COBOL委員会への提案概要を述べる。

- | | |
|------------------|---|
| (1) 多バイト文字機能の範囲 | : 第2章で述べたレベル2。 |
| (2) 分離符 | : 単バイト文字だけとする。 |
| (3) 文字列 | : 多バイト文字と単バイト文字の混在する文字列は不可。 |
| (4) 利用者語の等価性 | : 多バイト文字と単バイト文字の字形的に同じ文字については、等価とする。 |
| (5) 利用者語の長さ | : 単バイト文字と同じ長さ。(CODASYLは60文字、ISO/ANSIは30文字。) |
| (6) 定数 | : 多バイト文字定数の追加。(例えば、N"日本") |
| (7) データの字類 | : 多バイト文字だけからなるデータタイプの追加。(例えば、PIC N(2) USAGE IS DISPLAY-N) |
| (8) 転記の規則 | : 多バイト文字と単バイト文字の相互の転記は、関数を用意し別だてとする。 |
| (9) 比較条件 | : 多バイト文字と多バイト文字との比較は可。ただし、大小順序は作成者定義。 |
| (10) 多バイト文字が扱える文 | : INITIALIZE, MOVE文などがある。 |
| (11) 入出力時のシフトコード | : READ文では、入力時にシフトコードを自動削除する。
WRITE文では、出力時にシフトコードを自動挿入する。
なお、シフトコードの自動削除・挿入の構文を追加する。 |

5. おわりに

多バイト文字の標準化は、「ISO制定後JIS制定」との方針が明確になった。COBOLにおいても、本家本元のCODASYL COBOL委員会が検討を開始したので、標準化は加速を増すであろう。

第4章で述べたCODASYL COBOLの多バイト文字機能の内容は、第1回目の検討内容であるため、未検討要素が多く、まだ、流動的である。

情報処理学会 SC22/COBOL WGでは、次回のCODASYL COBOL委員会(1988.5.3~4)に出席し日本案を提示する。今後ともCODASYL COBOL委員会やISO SC22 COBOL WGなどに積極的に参加し、日本だけでなく国際的に通用する多バイト文字機能として標準化に寄与していきたいと考えている。

なお、本稿の作成にあたり、情報処理学会 SC22/COBOL WGの場での各委員の御意見に対し、感謝の意を表す。

[参考文献]

- 1) プログラム用言語の標準化に関する調査-日本語COBOL- 電子協 60-C-506(1985)
- 2) プログラム用言語の標準化に関する調査-COBOL74,85,との相違点,日本語COBOL, C言語ANSI案- 電子協 61-C-533(1986)
- 3) 床分,今城,花田: COBOL言語における日本語機能,情報処理学会, 第32回全国大会,5F-2(1986)
- 4) 高橋: 5年後の日本語プログラミング環境,情報処理学会, 計算機システムのヒューマンインタフェース-モデル・評価・展望-,pp.81-92(1988)