

# 情報処理と標準化

和田 弘\*

1955~57年に電気試験所で ETL MK 3 と MK 4 を試作した際に入出力装置に使用したテレタイプの経験から、計算機で使える文字の種類 (character set) とそのコードについては標準化が必要であることを感じました。そのためには関係者が集って協力してまとめなくてはならないと思い、関係の方々と相談してコード会という会合を作り、1958~59年にわたって検討した末、一応の結論が得られました<sup>1)</sup>。

1960年に当学会が創立された際、この体験を基盤として工業技術院と交渉し、申し出を得て学会内に規格委員会が設置されることになったのは1961年8月であります。つまり、コード以外にもこのような協力が必要となる分野があり、それには専門の学識のある方方において、組織的に研究をするのが適当であると思ったからです。

このようなことについて、国際的な動きを捜し求めていたところ、果たせるかな、同じ趣旨の活動が ISO と IEC で始まったことがわかりました<sup>2)</sup>。そこで、早速たまたまヨーロッパに行かれた現会長の高橋先生

に、1962年10月パリで開かれた SC 2 の会議を傍聴していただき、私自身は1963年9月コペンハーゲンで開かれた IEC/53 D に出席し、さらに同年10月には元会長の山下先生に SC 2 に出席していただき、わが国としてコードについての考えを述べていただきました。そして、これらの会議を通して ISO の活動状況がわかってまいり、わが国としても正式に参加すべきであるとの意見にかたまりましたので、工業技術院にお願いをして P (参加) メンバになるとともに、1964年5月にニューヨークで開かれた諸会議には15名の代表を送り、TC 97 の全活動状況をは握いたしました。

それ以来 ISO の会議の実態について検討を行ない、会議には代表を派遣して世界の大勢に遅れないようにとつめてきております。そして規格委員会の活動状況の要旨は、毎年学会誌に報告してきております<sup>1), 2)</sup>。

ここでは、話しの順序として、ISO/TC 97 の構成をまず図1として示します。本日は、これらの中で諸君のご参考になりそうな項目についてお話しさせていた

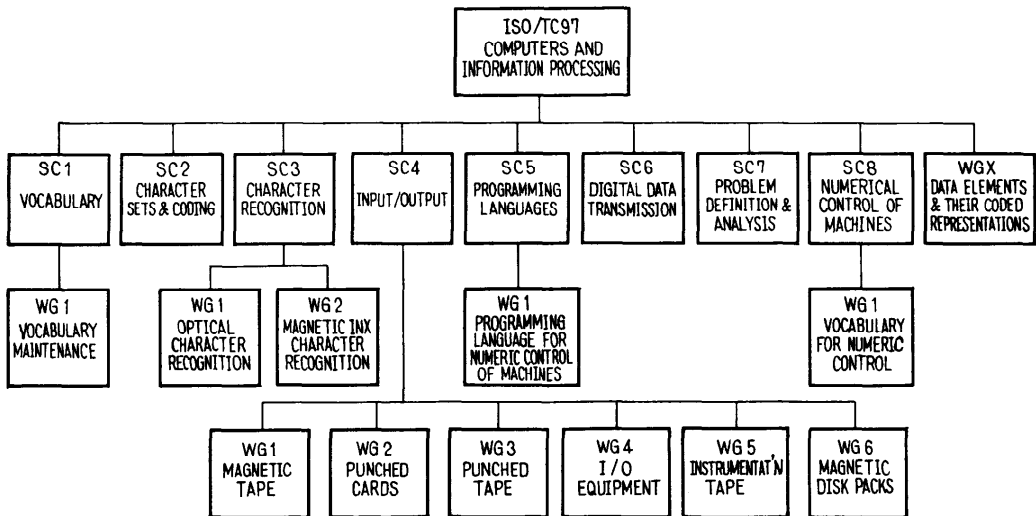


図 1 TC 97 の組織

\* 成蹊大学教授、情報処理学会規格委員会委員長  
 \*\* 第11回大会特別講演 (昭和 45 年 12 月 10 日)

できます。

### 1. 文字の種類とコード (SC 2)

冒頭に申したとおり、計算機に使える文字とそのコードを決定することは、最も重要な仕事として努力が続けられております。すでに R-646 として 7 bit コードが決定されました。JIS もこれに準じて制定され、実用化されております。国際的なこの文字の中に £ と \$ とが含まれていることは、いまなお議論を呼んでおります。いずれにしても、このコード表に含まれる文字は 128 字と限られていますので、それ以外の文字を使いたい場合にはどうすればよいか引き続き研究されてきております。これは 7 bit という hardware 的環境のもとで、文字の種類を拡張しようということですから extension と呼ばれております。図 2 に示すように、コード表で左の 2 列は機能文字に割り当てられていますので、文字だけの拡張を考えるなら 96 字 (正しくは 94 字) までならば、shift in/out で解決できることは、従来の電信の技術から簡単に理解できるのであります。各国が自由にそれらの文字を定義しますと、国際間ではどの国での shift であるかわからなくなります。そこで 3 図に示すように、extension の部分をそれぞれ 1 ページのようにみなして、機能文字 (escape) につづいて page の形式を示す記号

$b_7 b_6$	$b_5 0000$	00	01	10	11
		機能記号	数字および記号	文字	文字
		1111			

図 2 R-646 7 bit コード表

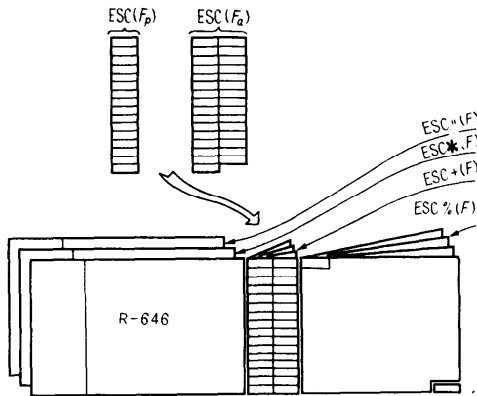


図 3 ページ式の extension 法

と page 番号に相当する文字、計 3 字を打つことによって、まず page を指定し、shift in/out でその page の使用と元にもどる操作をさせようというものであります (図 4)。

図 3 の右に page の形式を示しました。機能文字の取り扱いを慎重にしていることがわかっていただけると思います。

アメリカは最近になって、コード表自体を 8 bit に拡大することを提案しております。もちろん、現在の 7 bit でのコード表をその subset として含めることにしています (図 5)。そして右半分についても、まず 2 列は機能文字にあて、残る 6 列は文字にあてようとしており、その内容は詳細には決まっておりません。そして、これでも不足になる文字について、さきの 7 bit コードでの extension の手法をそのまま 8 bit を基礎として行ないたいと主張しております。

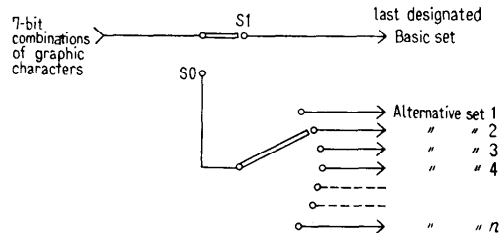


図 4 ESC によるページの切り替え

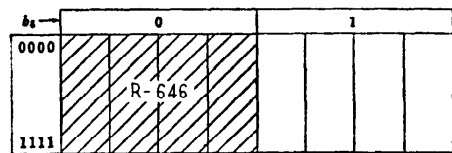


図 5 8 bit コード表

わが国ではカナ文字のために extension 法を必要とするので、早くからこの会議で発言しておりましたし、近くは漢字のために 1 page に代わって 1 冊とでもいうべき拡張法を提案して承認を得ております。

### 2. 入出力装置 (WG 4/4)

入出力装置は種類も多く進歩も顕著であり、その価格は高いものです。よい入出力装置をのぞんで接続したくとも、他社の製品ではできないのが現状です。これは接続法について標準がないためです、この互換性を確保するには logical, physical, electrical な特性の規準があればよいはずですが、このことは 1961 年の発足時からとり上げられていたのですが、欧米で具体的な意見の一致がみられなかったため、ほとんど進みま

せんでした。1967年にWG 4/4として再出発したものの、入出力装置と本体との境界(interface)をどこにするかさえ意見が分れて前進の徴候は見られませんでした。

すなわち、図6に示すようにinterfaceは3箇所考えられることができます。

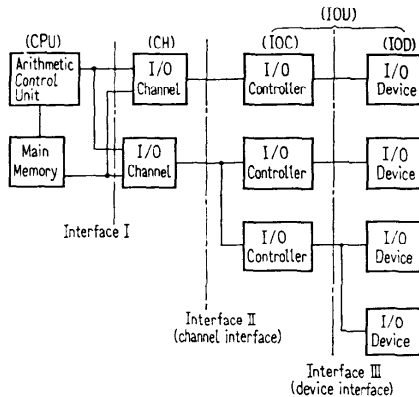


図6 入出力 interface

このころ、わが国では通産省の大型プロジェクトによる計算機で同じ困難に当面して、interfaceの確定につとめた結果、案の骨子がまとまりました。そこでこれをISOに原案として提出する準備が始められました。1969年9月に開かれた第2回の会議では、日本として初めて提出した案のほかに、イギリスからはdevice interfaceの案が出され、アメリカは国内の意見がまだまとまらないのでinterfaceの具備すべき条件を示す案を提出しただけでした。会議では原案が欧米以外の国から出たことはTC 97としては最初のことなので、かなり抵抗があったようであります。しかし、本年9月に開かれた第3回会議では、イギリスが提案を撤回したので、日本案を中心として議事が進むようになりました。

ところで、日本案の内容ですが、入出力装置の特性に依存しないchannel interfaceを基準として、転送速度は1 MB/s、ケーブルの長さは50 mまで、busは8 bitとparityを含めたものとし、芽づる式に接続してIOCは8台、IODでは64台まで動作させることを骨子としています。したがってinterfaceのところの単に1本に示した図6の内容は情報の往復のために計18本、制御用に6本、IOCの選択のために7本、時間など計測のために3本、合計34本が含まれることになっています。ケーブルには同軸24芯を採用し

て入力・出力の2本とし、残る14本は予備にあてることを考えています。

電氣的にはパルスの方式、出力電圧、波形など多くの特性とその試験条件も規定してありますが、本日はそこまでは触れないことにします。

### 3. プログラム言語 (SC 5)

Algol, FortranはISOの案が決まり、わが国でもこれに準拠したJISが刊行されています。Cobolは遅れてはいますが、最終案に各国が賛意を示しているため、近く正式に決まることになるでしょう。JISのCobolも目下原案が鋭意まとめられていますから、来年には脱稿されることになるかと待ちにしています。

CobolはCODASYLが中心となって常に改良・追加が行なわれています。そのつどこれに応じたcompilerを整備しようとする大変です。現に国内各社のcompilerにも互換性がないことは周知のとおりです。そこでISOでは、ある時点で区切りをつけてまとめることにしています。必ずしもCODASYLのものと同じではありません。たとえば、最近ではdataのstorage/retrieval, manipulation, communicationについても開発されており、しかし、これらは必ずしもCobolに限られた問題ではなく、Algol, Fortranにも同じことはいえるはずで、これらの問題にどう対応していくべきかは、課題として残されています。

さてPL/Iについて、ISOで審議にかかれとの意見は早くからありましたが、審議するにふさわしい権威ある母体が見つからないという理由で、現在まで一切触れられておりません。それに応じてヨーロッパにあるECMAは技術委員会を設けて、アメリカの規格協会、IFIP/TC 2と協力しながらBASIS-Iとして案をまとめにかかっています。すでに、その第3版ができております。この経過は来年5月のSC 5の会議にprogress reportとして報告される予定になっています。この計画によると、第9版としての最終案は、1972年12月にまとまることになっています。これは1973年5月に開催を予定されているSC 5に提出されることになりましょう。

さて、この動きをみますと、SC 5には最終案が形式的に提出されるだけでもみられますから、これを待ってわが国がこれを原案として審議を始めてもよほど有力な意見でない限り、その意見はISOで採択されるには無理があるように感じられます。関係各位の配慮をお願いしたいと存じます。

このような雰囲気ですから、会話型言語である BA-SIC, APL などを議題としてとりあげることは時期尚早でしょう。

つぎに operating system について述べます。SC 5 はプログラム言語のために設けられましたが、その後、その他の software でひろく用いられるものについても、その特性を審議するように改められています。OS は当然対象となるのですが、これとても最近では minicomputer がひろく使われるようになりましたから、OS を考える場合には、これも含めて考えるべきでしょう。各国はどうしたらよいか調査にさえまだ時間がかかりそうです。各位のご尽力でわが国から提案でもすれば、歓迎されることでしょう。

一方、利用の方をながめると data-management がますます増加してきています。そこで既存の言語から独立したもので、data-base management system がまとめられることが待望されます。プログラムとデータとをシステムの間で移動させるようにすることは、プログラム言語を標準化する以上困難なことです。それには、data として集積されたコード表示の内容は、どんな方針で記述されているかを明快に記述しなければなりません。この記述そのものを計算機に input することは、Cobol で Data-Division を input するのと同じ要領ですから、必須のことです。data を構成する data element についても記述しなければならないことは当然であります。そこで、つぎの話と関係してまいります。

#### 4. データ・コード (WG/K)

data といわれるものには、学会の論文、領収書、あるいは会社での帳簿など、文章の形式のものもあれば、領収書や帳簿のようにあらかじめ定められ Format にそって該当事項 (data-element) をいくつか記入すればよいものもあります。

ここでは後者に従って data とは data element の集合であるとみなして話を進めます。これと計算機との関係が問題なのです。

欧米の企業では、このような data を処理するための機械として Punched Card System がまず発達しました。sort/merge を始め計算を数値によって処理するので、たとえば社員の氏名には社員番号 (俗にコードといわれる) を付与して処理するのが当然となってきています。機械が PCS から計算機に変わっても、この PCS 時代の慣習は引き継がれて現在に至っています。けだし、data には時代を通して、一貫

性があるほうが便利なためでもありません。

一方、計算機の短い歴史をかえりみると、この間にさえいくらかの変遷がありました。machine language で書いていたプログラムやデータは、やがて symbolic な assembler language となり、さらに Fortran とか Cobol のような言語へと変わってきたではありませんか。計算機がプログラムを実行する際には machine language でやることは不変なので、文字や単語を使う現在のプログラム言語は compile に余計な時間を要しますけれども、人間に便利であるという理由からひろく実用されているのでしょう。計算機のことを近頃では情報処理装置と呼び、数値計算に代わって情報処理という言葉がわざわざ使われているのも、PCS は数字しか扱えなかったのに、計算機は文字も取り扱えるということの証査です。しかも事務への応用は給与計算のように限られたものから、次第に相互に関連の多いものへと変わって行って、MIS とが最近では Data Base と呼ばれたりする内容になりつつあります。この間にあって data element には数字でなく文字が使えるようになりました。Cobol の NAME PICTURE X(20) を指摘すれば充分でしょう。同様に element の樹枝構成も level で認められています。しかも、これらの文字は SC 2 の努力で alphabet 順に sort/merge できるように code が決められました。data element について man-machine communication ができるようになっています。

学術的にはこのような傾向にあるにもかかわらず、世間では data-code といわれて、data-element に code を付与することが情報処理の利用に必要な欠くべからざるものであるかのような議論が流布しています。この原因を考えてみましょう。

element がコード化されていれば、機内で能率がいいことは当然です。machine language が最も能率がよいのと同じ意味で。しかし、element がコードで機外にも使われることになると、番号であるから一般人にはその意味がわかりません。むしろ、暗号のような印象を与えます。

アメリカは data-element とそのコードについて国際的な互換性を設けたいと提唱して、1965年に ISO のなかに WG/K の設置を認めさせました。その際ヨーロッパ諸国は消極的でした。その本心では、労多くしてできそうもなく、手伝う意旨もない。できれば使ってもよいが、現在使っているコードと違うものならば採用しなければよいということのようです。

事実、日時、単位など基本的な element についてさえ各国の表示法はまちまちです。本日、昭和45年12月10日は Dec. 10, 1970 という表示法もあれば、10 Dec. 1970 というものもあります。これだけを統一しようとするさえ大変困難なことでありましょう。いわんや各種の element ということになれば、この辺がわかってか本年 Berlin での会議で、アメリカは対象を最小限に絞ろうと提案しました。もちろん承認されました。次の会議は面白いことでしょう。

わが国では現在、情報開発センターが中心になって各種の data-code の設定に努力が続けられています。仮に商品コードの設定を考えますと、その範囲を大きくすればするほどけた数がふえます。都内の電話番号は7けたですが、日本全体として見る場合には03-がつくし、世界の一部とみればさらに081-を追加しなくてはなりません。これは全世界の電話帳を1冊にまとめるようなものです。したがって、なるべく範囲を狭めて実用価値を重視すべきで、観念論から全部ということとは避けたほうがよいのではありますまいか。魚1匹でも大きさ、切身の位置など、いろいろの parameter がいます。これらをすべて共通するように考えることは無理でしょう。

いずれにしても、data-code を設定することが多くの人の要望であるのなら、学会もこれに協力すべきであります。しかし、それぞれの業界の内容については何も知らないのでありますから、学会の協力すべき scope としては、変動が多そうな data の記述法についてある洞察と定見とを確立し、設定されようとする data-code がこれと矛盾しないように指導できる guideline を開発することではありますまいか。

## 5. むすび

以上4つの例をひいて述べてきましたが、ISOとの10年に近い経験から

i) ISOに参加協力すると、外国で何が考えられているかがわかり、製品さえその発表以前に知ることができて、勉強になる利益があります。

ii) 一方、会議に代表を派遣することは旅費もかかるし、時間的にも大変です。

これらの長短にかかわらず、もっと積極的に参加するほうがよいように思われます。参加するというのは会議に出席することばかりでなく、国内で、すなわち学会内での研究に参加することも含まれます。

しかし TC 97 が無条件に礼賛すべきであるとも思われません。さきのような組織は、1961年に輪廓が定

められたのでありまして、この10年間に計算機の利用分野が大きく変わったのでありますから、それに応じて仕事の内容も見直されるべき時期に当面していると思われます。TC 97 自身もこれを認めていて、その改組案が待たれています。以下は私見ですが。

### 1) 応用分野の分離

SC 8 は工作機の数値制御を取り扱っています。いろいろな分野への応用という観点に立てば、化学工場とか事務分野とか documentation など沢山あります。

工作機だけに限るのはおかしい。これらについては各種分野に対して計算機を応用する際の guideline だけに限定して、総力を情報処理固有の問題に絞るべきではありますまいか。

### 2) system の重視

10年前の計算機は1台ずつの入出力装置をもって、batch 処理だけを対象とする孤立系でありました。最近の system はこれに対して異常に大きく複雑なものになっています。さらに、ますます複雑な system が論ぜられています。このように system が複雑になると、データ通信、入出力装置の interface、あるいは system の評価方法というようなものが重視されてしかるべきでしょう。現在では、まだそれらが断片的に考察されているにすぎないように思われます。

ひるがえって創立10周年を迎えた学会の活動を見渡すと本日の大会予稿集でもそうですがコーヤッタコーナーナッタというような範囲の狭い論文なり活動が多いように思われます。これらも結構ではあります。もっと全体に影響するような scope をもつ論文が欲しいと思います。そのためには会員各位が協力して、組織的な研究をしてその成果を発表するという心構えが必要ではないでしょうか。

規格を作るとか、標準化というと、ややもすると学術的価値にとぼしいというような考え方がみうけられますが、これらの仕事には学術的基礎がもっとも必要なのでありまして、及ぼす影響の大きいこれらの仕事は官庁とか業界とかメーカーのご都合主義で進まないようにすることも、学問に志すわれわれ会員としての大切な責任ではないでしょうか。

### 参考文献

- 1) 和田、高橋：コード会のコード、情報処理、1-2, p.107 (1960)
- 2) 和田：計算機についての国際的な標準化、情報処理、3-2, p.8.3 (1962)