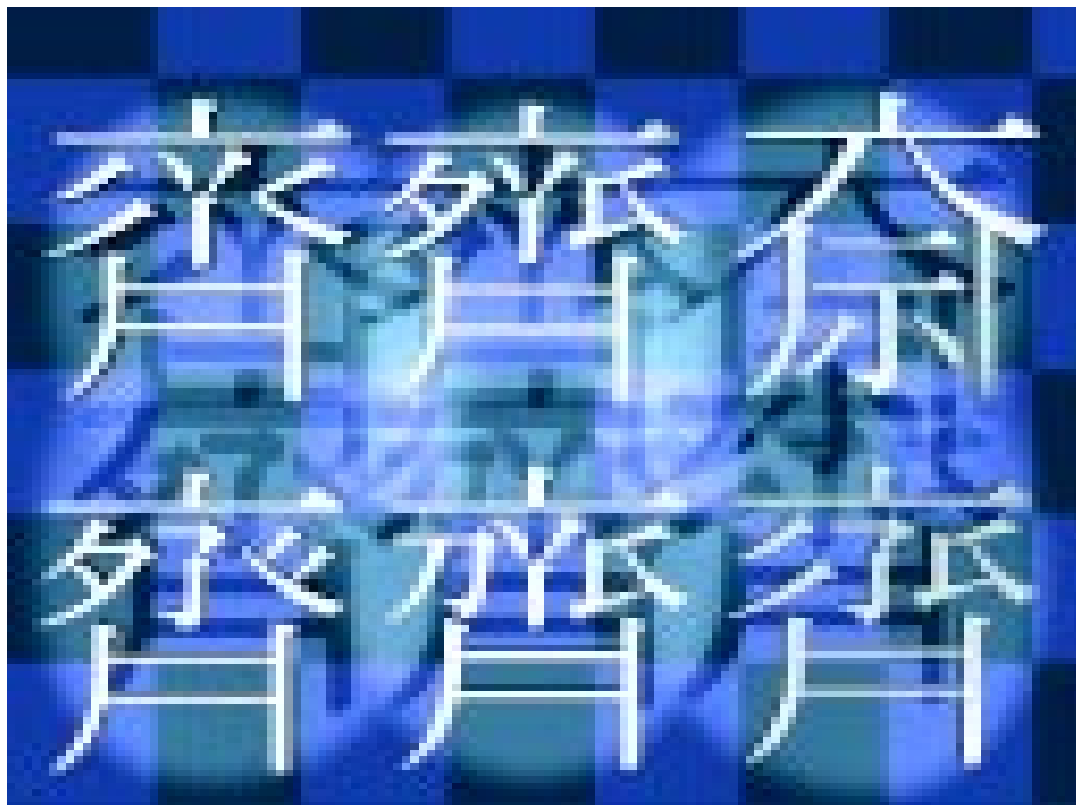


Windows NT[®] 日本語処理ガイドブック

著者 川俣 晶 + XKP 普及部会



Windows NT[®] 漢字処理技術協議会
Windows NT[®] Extended Kanji Processing Council

XKPはWindows NT[®]漢字処理技術協議会の商標です。

Windows の正式名称は Microsoft Windows Operating System です。

Microsoft、Windows、Windows NT は米国 Microsoft Corporation の米国およびその他の国における登録商標です。

その他記載されている会社名、製品名は各社の登録商標または商標です。

- 本書の内容は1998年10月現在のものです。本文中のWindows OSとはWindows3.1、9x、Windows NTの全般をさします。
 - 本書は情報提供のみを目的としています。Windows NT[®]漢字処理技術協議会は明示していると否とに関わらず本書においていかなる保証を行うものではありません。
-
-

はじめに

Windows NT漢字処理技術協議会(eXtended Kanji Processing、略称：XKP)は、1995年2月、地方自治体のパーソナルコンピュータを利用したシステム構築を推進する目的で作られた団体です。

XKP設立の経緯が、この団体の特徴をはっきりと表わしています。XKPは、コンピュータ・メーカーやソリューション・プロバイダーからマイクロソフト社への要請で始まりました。いままで、メインフレームやオフコンが担当していた戸籍、住民基本台帳などのコンピュータ処理をWindowsパソコンで行なおうとすると、JIS以外の文字が不足しており、また外字処理が標準化されていません。クライアント・サーバ型のアプリケーション・システムを構築する際に、マルチベンダー環境となるので、コンピュータ・メーカー独自の外字拡張方式が採用できません。これらの問題点を解決し、拡張漢字処理の標準化を行うために、マイクロソフト社とコンピュータ・メーカー、パッケージソフト・メーカー、ソリューション・プロバイダー、フォント・ベンダーなど、漢字処理に関係する企業が横断的に集まりXKPは設立されました。

「人名・地名外字のWindows NTでの対応」という大きなテーマであるため、地方自治体以外に公官庁や、金融・証券、文教などの「人名・地名」を処理する業界、そして印刷・出版などの業界からも大きな注目を集めています。

本書は、その活動の一環として作成しました。外字以前に、「Windows、Windows NTでの日本語処理」についての理解を深めていただきたいと思います、基礎的な内容に重点を置いています。

パーソナルコンピュータを使った日本語処理についての、正しい理解の一助となれば幸いです。

Windows NT漢字処理技術協議会
<http://www.xkp.or.jp/>

CONTENTS

Q パーソナルコンピュータとはなにか？	1
Q パソコンから漢字を出力する仕組みとは？	2
Q フォントってなんですか？	3
Q 文字コードってなんですか？	4
Q どんな仕組みで漢字を入力するの？	5
Q 文字コードにない文字は使えないのですか？	6
Q 外字エディタさえあれば外字は完璧？	7
Q 外字をIMEで入力するには？	8
Q 大型コンピュータの文字が使えますか？	9
Q JISコードってなんですか？	10
Q シフトJISってなんですか？	11
Q Unicodeってなんですか？	12
Q Unicodeは未来の文字コードですか？	13
Q XKPとは？	14
Q XKPで外字は通信できますか？	15
Q 文字コード数を越えた外字は使えない？	16
Q XKP外字は印刷できますか？	17

Q パーソナルコンピュータとはなにか？

A 1人が1台を占有できる多用途情報機器です。

パーソナルコンピュータは、安価で、小型な、コンピュータの一種です。

パーソナルコンピュータは、パソコン、あるいは、PCと呼ばれることもあります。（本書では、パソコンという表記で統一します）

パソコンは、安価で小型ではありますが、基本的には、汎用装置であり、ソフトウェアを変更することにより、様々な用途に使用できます。1台の機器が、ワープロとなり、表計算装置となり、データベース装置となり、大型コンピュータの端末機となり、たまには仕事の息抜きとしてゲームまで実行してしまうのが、パソコンです。

あらかじめ定められた機能しか使用できないワープロ専用機や電子手帳とは、この汎用性の点で、大きく異なります。

パソコンの性能は、従来から存在する大型コンピュータ（メインフレーム）に劣ります。しかし、大型コンピュータが、1台のコンピュータを多数のユーザーで利用する形態であるのに対し、パソコンは1人で1台を独占して利用することが前提であるため、一人あたりの処理能力としては、パソコンの方がはるかに優越します。

しかしながら、多数のユーザーより共有されるデータを扱う場合などには、1台で多数のユーザーから処理が集中するため、負荷によってはパソコンでは性能が不足する場合があります。そのような場合には、大型コンピュータが使用されます。このように、パソコンと大型コンピュータは、一概に、どちらが優れているとも言えず、用途に応じて適材適所で利用することが望ましいと言えます。

しかし、パソコンと大型コンピュータは、全く異なる文化の潮流より生まれ出た装置であるため、必ずしも仕様の面で整合性が取れていません。そのため、共存させて稼働させるには、いろいろな問題があります。文字について、この差異と解決方法を説明することが、本書のテーマの一つです。



Q パソコンから漢字を出力する仕組みとは？

A 文字コードとフォントがキーワード

パソコンで文字を扱う場合、主役として活躍するものが、文字コードと、フォントです。

通常、コンピュータは、文字を文字として理解するのではありません。文字ごとに、番号を割り当てて、この番号の違いで、文字の違いを理解します。

この数字のことを、「文字コード」と呼びます。

さて、文字コードは、ただの番号でしかありません。そのため、我々パソコンのユーザーは、文字コードだけでは、どんな文字がパソコンの中にあるか理解できません。そこで、文字コードに対応する文字の形も、パソコンに教え込まねばなりません。

そのために使用されるものが、「フォント」です。

フォントは、ユーザーの指定により容易に変更できます。これにより、同じ文字でも異なるデザインの表示に変更することが容易にできます。たとえば、明朝体からゴシック体への変更などが可能になるわけです。

文字コードとフォントは、どちらか一方だけあれば良い、というものではなく、両方そろって、はじめて、実際に文字が利用可能になるものです。



Q

フォントってなんですか？

A

文字の形を持つ一種の画像データです。

フォントとは、文字の形を記憶した一種の画像データです。

それぞれの文字コードに対応する画像が記録されています。

通常、フォントには、それ以上の情報は含まれていません。

フォントを複数用意するのは、明朝体とゴシック体のような書体を切り替えて利用するためです。

システムに標準で含まれない文字を持つフォントを組み込むだけでは、単に、特別な操作でその文字が表示可能になるだけで、検索などの処理は正常に機能しません。

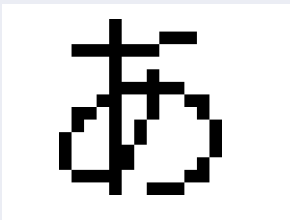
フォントには、大きく分けて、二つの種類があります。

一つは、ビットマップフォントと呼ばれ、文字の形を、点の集まりで表現します。

もう一つは、アウトラインフォントと呼ばれ、文字の形を、輪郭線の曲線で表現します。

ビットマップフォントは、簡単な仕組みですが、文字を拡大すると、四角い点があきりと見えてしまい、美しくありません。

それに対して、アウトラインフォントは、もともと曲線の輪郭線を記録してあるため、拡大しても、美しさが損なわれません。そのため、現在では、アウトラインフォントが主流となっています。



ビットマップフォント



アウトラインフォント

Q

文字コードってなんですか？

A

文字を代表して表す番号です。

パソコンで文字を扱う場合には、文字コードと呼ばれる番号に置き換えて使うのが普通です。

文字コードは、ただの番号でしかありませんから、違う文字に同じ番号が割り当てられてしまうと、どの文字を示しているか判別できなくなってしまいます。そのためどの文字に、どんな番号を割り当てるかのルールを厳密に決める必要があります。

通常は、JISなどの公的な機関が、ある範囲の文字を集め、重複しない番号を割り当てた「文字セット」というものを作成し、これを利用します。

さて、文字セットを作成する場合、ある範囲の微妙に異なる文字のグループに対して、一つの文字コードが割り当てられます。

これは、現実的な実用性を配慮してのことです。世の中には、同じ漢字でありながら、独自の書き方が複数存在する場合があります。たとえば、人名の漢字などで、微妙な差異(「吉」の上の部分は、下の棒が長い「土」とする等)、点の数や、はらいの有無などを重視して、「こう書かねば私の名前とは認めない」というような主張をする人がいます。

しかし、JISではこれらを同一の文字として、ひとつの文字コードしか割り当てていません。そのような主張を無制限に認めると、文字の種類が爆発的に増え、それに応じて、システムのコストも膨らむからです。そうすると、手頃な価格で活用できるパソコンのメリットが失われてしまいます。

そこで、文字セットを作る場合には、あるルールを定め、そのルールに従って、差異がある水準以上の文字に限り、文字コードを与えるという作業が行われています。

このように、微妙に異なる文字を一つにまとめることを、「包摂」といいます。

Q

どんな仕組みで漢字を入力するの？

A

IME が、読みを漢字に変換します。

パソコンで漢字を入力する仕組みは、Input Method と呼ばれています。

Windows OS の Input Method はいくつかの機能モジュールに分かれており、実際に入力やかな漢字変換を担当する機能を、IME=Input Method Editor と呼びます。

IME は、ユーザーが入力した、ローマ字またはひらがなによる「読み」の文字列を解釈して、漢字交じりの文章を候補として、表示します。ユーザーは、候補の中から、適切な単語を選択し、文字列を「確定」させます。

これにより、漢字交じりの日本語が入力できます。

さて、通常、IME は、文字列を入力して、文字列を確定させるソフトです。

文字列とは、文字コードが並んだものです。

ですから、そこには、文字コード以外の情報は含まれません。

そのため、フォント指定を含めた文字を登録しておき、それを確定させて、フォント指定ごと入力させることはできません。たとえば、MS 明朝の「亜」とMS ゴシックの「亜」を、別な読みで登録して使い分けるといったことはできないのです。

つまり、同じ文字コードに異なる文字を登録したフォントを複数用意しても、それだけでは、IME によって入力することはできません。



Q

文字コードにない文字は使えないのですか？

A

使えますが、いろいろな問題があります。

パソコンには、通常、外字機能と呼ばれるものがあります。

この機能は、ユーザーが、自分で使いたい文字を登録して、あたかも最初から登録済みの文字であるかのように、自由に利用できるものです。

この文字を外字と呼びます。

たいへん便利な機能であり、これがあれば、文字コードに無い文字を使いたいと思った場合、自分で登録するわけですから、どんな文字でも利用することができます。

しかしながら、現在では、外字機能には以下の問題が発生しています。

文字数が不足する場合がある

たとえば、Windows の標準的な利用可能外字数は 1880 文字です。しかし、この文字数では、大量の人名などの情報を蓄積する場合には、不足することがあります。

情報通信ができない

外字は、あくまで、1 台のパソコンに登録するもので、登録したパソコンでしか利用できません。しかし、現在では、LAN、インターネットなど、パソコンとパソコンを通信回線で接続することが当たり前になっています。ところが、外字を使って作成した文書を、それらのネットワークを経由して送ると、受け取り側のパソコンには目的の外字が登録されていないため、文字が再現されないという現象が起きます。そのため、電子メールなどの情報通信を行う場合には、マナーとして、外字を使わないということになっています。

なお、フロッピーディスクなどで文書ファイルを他のパソコンに持っていても、同じ現象が発生します。

これらの問題に対処するためには、二つの考え方があります。

一つは、より多数の文字を収録した文字セットを作り、それを使う方法。これには、第三水準・第四水準、JIS X 0213 と仮称される新しい JIS 規格の制定などの動きがありますが、現時点で実用的に利用可能な段階に達したものではありません。

もう一つは、より安全に外字を扱う仕組みを作る方法です。これには、XKP=Windows NT eXtended Kanji Processing などがあります。

Q

外字エディタさえあれば外字は完璧？

A

外字エディタはすべてではありません。

外字エディタには、2つの機能があります。

文字のデザイン機能（とフォントファイル作成機能）

作成されたフォントファイルを外字フォントとして登録する機能

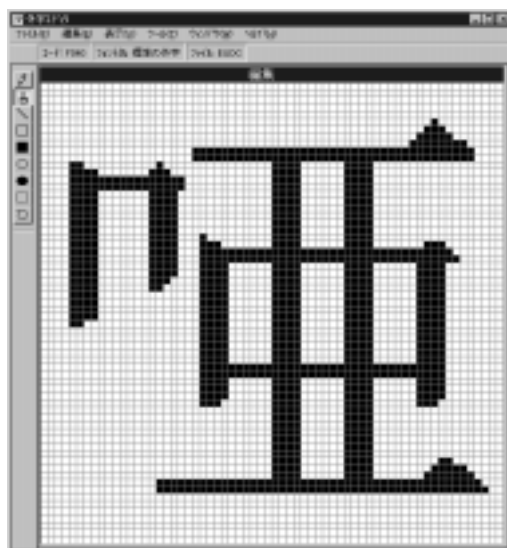
Windows OS標準付属の外字エディタでは、文字をデザインすることができますが、それと同時に、作成した外字フォントファイルを外字フォントとして登録する機能もあわせもっています。ですから、文字を作るとすぐにWindows OS上で使えるようになるのです。

しかし、外字フォントファイルには「読み」や「部首」などの情報は入っていませんから、そのままでは文字コードによってしか入力できません。読みを入力して変換させるためには、IMEの辞書に「読み」を登録する必要があるのです。

さて、本格的な外字活用を行う場合には、デザイン作成（フォントファイル作成）と、それを実際に使用するパソコンの外字として登録する作業は、まったく別個に発生します。そのため、プロ用の外字関連ツールでは、外字のデザインと、外字の登録は、別個のソフトウェアとなっています。

たとえば、XKPのような本格的な業務用の外字システムを使う場合には、単純に外字エディター一つで作業を行うのではなく、作業の流れに応じた複数のソフトウェアによって、作業を進めることとなります。

外字フォントにも、他のフォントと同じように、ビットマップフォントとアウトラインフォントがあります。点の数が少ないビットマップフォントの場合は、実際にフォントを必要とするユーザーが自らデザインする場面もありましたが、微妙な曲線を表現できるアウトラインフォントを適切にデザインするのは素人には難しく、専門業者に特注する場面が多くなっています。



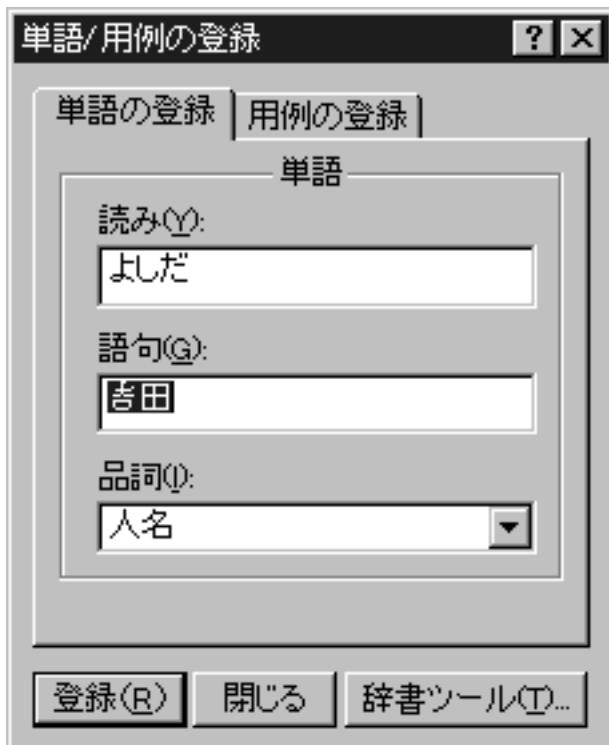
Windows NT 4.0付属の外字エディタ

Q**外字をIMEで入力するには？****A****外字の情報を登録しなければなりません。**

外字は、外字エディタで作成しただけでは、すぐにIMEで入力できません。

トラディショナルな方式の外字であれば、IMEの辞書に、外字とそれに対応する「読み」を登録する必要があります。

XKPのような、外字を集中管理する方式の場合は、管理データベースに、外字に関する各種情報を登録することで、入力可能となります。



MS-IME98で外字を含む単語を登録している例

Q

大型コンピュータの文字が使えますか？

A

解決すべき二つの課題があります。

大型コンピュータ（メインフレーム）は、各メーカーごとに、独自の文字セットを定めています。また、実際にそれを利用するユーザーの中には、それに加えて、ユーザー独自の文字を多数持っている場合もあります。

それらのシステムで利用していた文字は、そのままでは、パソコンでは使用できません。

利用可能にするための障害は二つあります。

文字コード

パソコンが通常使用するシフトJISと呼ばれる文字セットは、大型コンピュータに比べて、大幅に収録文字数が足りません。シフトJISにない文字を外字登録しようにも、利用可能な外字数が、足りないケースが多くなっています。

シフトJISにかわる、多数の文字を表現するための何らかのシステムが必要です。

フォント

大型コンピュータで使っていたドットフォントは、そのままではWindows上で使用できません。ですから、必要な文字を収録したフォントを入手しなければなりません。

この二つの問題を解決する方法のスタンダードを確立することを主テーマとしているものとして、XKPがあります。ただし、XKPは、文字コードやフォントを作成する団体ではありません。必要な文字はユーザーごとに異なっているので、それを標準化することも意味がないためです。あくまで、必要な文字を利用可能にする手段がテーマです。



Q

JIS コードってなんですか？

A

通常は JIS X 0208 を示します。

文字セットに関する規格は、JIS にも複数あります。

JIS X 0201-1997 7 ビット及び 8 ビットの情報交換用符号化文字集合
アルファベットとカタカナを規定しています。

JIS X 0202-1991 情報交換用符号の拡張法
1 バイトコードや 2 バイトコードのいくつかの文字コードを同一システムにおいて切り替えながら一緒に使う方法を決めています。

JIS X 0208:1997 7bit 及び 8bit の 2 バイト情報交換用符号化漢字集合
よく利用される漢字を中心とした文字セットを規定しています。6879 文字を収録しています。

JIS X 0212-1990 情報交換用漢字符号 - 補助漢字
JIS X 0208 に加えて使用するために、利用頻度の少ない漢字を集めた規格です。補助漢字と呼ばれます。6067 文字を収録しています。

JIS X 0221-1995 国際符号化文字集合(UCS) 第 1 部 体系及び基本多言語面
Unicode とほぼ同等の、全世界の文字を収録した国際文字セットです。日本語の文字としては、JIS X 0208 と JIS X 0212 相当のものが収録されています。

このうち、いわゆる JIS コードと呼ばれているのは、JIS X 0208 のことです。

ただし、この JIS X 0208 は、各社によって、独自の解釈の拡張がされています。それらの拡張された文字セットは、JIS X 0208 とは言えません。JIS X 0208 あるいは JIS コードと呼ぶ場合は、混乱を避けるため、外字や、メーカー独自文字(機種依存文字)などは、含まないものと考えべきでしょう。

また、JIS X 0208 そのものにも制定年度による内容の違いがあります。特に、1978 年制定の JIS X 0208-1978 から、1983 年制定の JIS X 0208-1983 へ移行するときに、いくつかの文字の位置が交換されました。これにより、JIS X 0208-1978 に準じる機器から、それ以降の規格に準じる機器にデータを送ると、文字が変わってしまう場合があります。

Q シフトJISってなんですか？

A パソコン用に加工された文字コードです。

シフトJISは、JIS規格であるJIS X 0201と、JIS X 0208をもとに、扱いやすい形になるよう、文字コードの数値を加工したものです。

そのため、利用できる文字の種類としては、上記の二つのJIS規格に従います。

ただし、利用できる文字の種類を、各メーカーが独自に拡張しているため、JIS規格に無い文字も使用できることが多くなっています。たとえば、丸付き数字や、ローマ数字は、独自拡張文字です。

しかしながら、それらの文字を使うことは、情報通信利用時にトラブルを起こすことが多いため、推奨はできません。文字種類が足りない場合は、より文字数の多い文字セットを利用するか、XKPのような通信上の整合性を配慮した仕組みを用いる必要があります。

基本的には、利用できる文字がJIS X 0208に制約されるので、大規模な利用には適しません。

また、メーカーや製品ごとに、準拠するJIS規格の年度が異なるため、内容に相違がある場合もあります。

更に、拡張性が無く、そのままでは、収録文字数を増やす方法がありません。

シフトJISには、半角文字と全角文字、あるいは、1バイト文字と2バイト文字という考え方があります。これらは、パソコンの性能が限定された1980年代には有効な表現手段として利用されていましたが、パソコンの性能が向上した今日では、半角、全角といった区別を廃して、フォントの選択によって文字のサイズなどを指定する方式が主流になりつつあります。



メーカーが独自に拡張した文字の例

Q

Unicodeってなんですか？

A

全世界の文字を収録した文字セットです。

Unicodeコンソーシアムという、主にアメリカの私企業が中心になった団体が作成した、全世界の文字を含む文字セットです。

公的な機関が定めた標準ではありませんが、ほぼ同等の内容のものが、国際規格 ISO/IEC-10646-1 として定められており、これに対応する国内の JIS 規格として、JIS X 0221 が制定されています。

Unicode は、世界中の主要な言語の文字をすべて 16bit 単位 (bit とは情報の単位で、0 または 1 の情報を持つ) で表現することを特徴とします (注)。そのため、7bit や 8bit を単位に文字を記述する従来の文字コードとの互換性はありません。

Unicode は、世界中の文字を収録した文字セットではありますが、Unicode 対応のシステムを利用すれば、世界中の文字が使用できるわけではありません。実際には、Unicode の全ての文字を収録したフォントというのは、事実上存在しません (フォントが無ければ、表示も印刷もできません)。また、各地域の各言語ごとに、独自の書き方のルール (日本語なら禁則処理など) があるため、それらを、きちんと扱えるアプリケーションソフトが必要になります。

現実には、Unicode は、複数の言語を同時に扱える文字セットとしてではなく、特定の言語をよりよく使うための文字セットとして利用されます。

たとえば、日本語に限定して、シフト JIS ではなく Unicode を使用した場合、以下のメリットが生じます。

JIS X 0212 補助漢字に相当する文字も使用できる

外字領域が 1880 文字から 6400 文字に拡大する

Unicode を使用すると、日本の書体と中国の書体が区別できなくなる、という批判がありますが、これは、必ずしも正しくはありません。微妙に外見の異なる文字を、あるルールに従って統合した結果、日本の文字と中国の文字が、同じ文字コードに割り当てられているケースがあります。しかし、地域・言語情報は文字コードとは別に適切に設定すべきものであり、それが設定されていれば、他言語の書体と混同されることはありません。

注： Unicode は、1 文字を 16bit で表現すると解説している資料がありますが、これは誤って流布した俗説です。JIS X 0208 が 7bit または 8bit データ 2 個で 1 文字を表現するのと同じように、Unicode でも、複数の 16bit データの組合せで 1 文字を表現する場合があります。

Q

Unicode は未来の文字コードですか？

A

既に実用的に使われています。

既に、Unicode は、様々な場所で活躍しています。

けて、未来の夢の文字コードではありません。

現在、Unicode が使用されている主要なソフトウェアを紹介します。

ワープロソフト(Wordなど)

Word などのワープロソフトでは、既に Unicode で文字を管理しています。そのため、フォントさえあれば、シフト JIS にはない、Unicode 独自の文字も使用することができます。

Windows NT

Windows NT は、Unicode を全面的に採用した OS です。シフト JIS で利用する場合は、OS に引き渡される際に、すべて Unicode に変換された上で処理されています。

Windows CE

小型システムでも稼働する Windows CE の場合、複数の文字コードをサポートするとシステムが重くなるため、Unicode のみをサポートして、それによって、世界中の言語を取り扱い可能としています。

IMAP4

電子メールを受信する新しいプロトコルとして注目されている IMAP4 では、文字に関する情報を Unicode で受け渡すことになっています。ですから、既に IMAP4 サーバ経由で電子メールを受け取っている人は、潜在的な Unicode の利用者と言えます。

Java 言語

Java 言語は、プログラム内部で扱う文字はすべて Unicode として扱うように設計されています。入出力を行うときには、従来の文字コードに変換しています。ですから、Java 言語で作成されたソフトを使用しているユーザーは、潜在的な Unicode ユーザーといえます。

XML

HTML の後継言語として、インターネットで注目されている XML ですが、どの文字コードで文書を記述する場合でも、名前を指定することで、Unicode のすべての文字が利用できます。(もちろん、表示するためには、フォントが必要です)

Q

XKPとは？

A

Windows NT[®]漢字処理技術協議会の略称です。

地方自治体の戸籍業務などをパソコンを用いてシステムを作成する場合、以下の問題が生じました。

利用可能な文字数がシフトJISでは足りない

特定の業者にすべてを依存するのではなく複数の業者から調達する方式が普及した

つまり、特定の業者のみが供給する特殊なシステムでは時代にニーズにそぐわず、かといって、標準的なパソコンでは、文字数が不足していたわけです。

そこで、漢字を扱う標準を定めるために設立されたのが、Windows NT[®]漢字処理技術協議会（略称=eXtended Kanji Processing=XKP）です。

XKPが作成した技術仕様は、「Windows NT漢字処理技術仕様」と呼ばれます。この仕様をさして、XKPと呼ぶ場合もあります。

XKP仕様には、以下の特徴があります。

戸籍業務などの、データベース+応用システムの構成を想定

Unicode とほぼ同等と言われる JIS X 0221 を採用

Unicode対応のOSであるWindows NTをプラットフォームとした

応用システムの開発を容易にするためにOLE Control（注）の標準仕様を作成

仕様書が公開され、それに対応する製品やサービスのビジネスへ誰でも参入できる

現在までに、XKP仕様準拠製品は複数のベンダーより供給され、地方自治体などへの導入実績があります。

また、Windows NTの開発と、XKPの方向性が一致しているので、Windows NTの次バージョンであるWindows NT 5.0では、よりXKP仕様との親和性が高くなることが期待されています。

注： OLE Control とは、マイクロソフト社のOLE技術を元に開発された、再利用可能なプログラム部品の標準です。これを活用することで、応用プログラムの開発コストと期間を低減することができます。

Q**XKPで外字は通信できますか？****A****できます。**

通常、外字を含む文書を通信で送る場合、相手に届くのは外字の文字コードだけなので、どんな文字であるかは分かりません。

これを解決する方法として、外字フォントと一緒に送る、というアイデアがあります。しかし、この方法は、送信者が自分で作った外字を送る場合は良いのですが、それ以外の外字フォントには、それぞれ著作権が存在し、利用契約を外れた範囲へ配布することを制限している場合も多いという問題が残ります。

そのため、外字フォントを送る、というアイデアは、100%の解決にはなりません。

XKP仕様では、フォントファイルを送信する方法も規定していますが、それだけでなく、ある文字が、どんな文字であるかを知るために必要な情報を送信する方法を規定しています。

これを利用すると、どこの組織が作成した何という文字セットの何番目の文字であるかが分かります。そして、外字フォントを入手するための連絡先が分かるようになっていきます。

更に、Windows NT 漢字処理技術協議会では、通信時の相互運用性を高めるために、コアモジュールのシングルバイナリー化、TCP/IPベースの外字情報交換の共通プロトコルの作成と、インターネット上の広域外字交換用サーバの構築、そして、XKP対応ソフトの互換性を確認するテストソフトウェアの作成、などを検討中です。



Q**文字コード数を越えた外字は使えない？****A****使えます。**

文字を指定するには、文字コードが割り当てられていなければならない、と考えるのは実は間違いです。

インターネットの文書を記述するHTML言語や、CALsなどで有名なSGML言語では、文字を文字コードとして記述する他に、文字の名前を記述することで、文字コードにない文字を扱う方法が用意されています。

もちろん、文字の名前を記述するためには、文字コードを使用します。

たとえば、HTML言語では、HTML言語自身が&記号に独自の機能を割り当てているため、そのままでは、文書に&記号を書き込めません。そこで、&記号にampという文字の名前を与え、&と書くことで、それを&記号そのものと見なす、というルールによって、文字コードで記述した場合は使用できない文字を、名前で書くことで利用可能にしています。

同じように、XKP仕様では、文字を名前として記述する方法を規定しています。コンピュータ内部では、外字専用の文字コードを使わず、下記のようなデータで記録されます。

吉田茂の吉は、&00012345;と記述します。

この方法を使用すると、外字領域を1文字も使用することなく、実用上無制限の数の外字を利用できます。

Q**XKP外字は印刷できますか？****A****できます。**

通常、XKP仕様で使用される外字は、TrueType形式で作成されます。

このフォントを切り替えながら、表示と印刷を行います。

そのため、特別な機器を用意しなくても、TrueTypeフォントを印刷できるプリンターがあれば、XKP外字を印刷できます。

TrueType形式のフォントは、アウトラインフォントのため、プリンターの解像度に合わせた美しい文字が印刷できます。プリンターの解像度が高ければ、それだけ、印字結果も美しくなります。



Windows NT[®]日本語処理ガイドブック

川俣晶 + XKP普及部会 著

1998年10月30日 初版

発行 Windows NT[®]漢字処理技術協議会

URL: <http://www.xkp.or.jp/>

E-Mail: webmaster@xkp.or.jp

**〒151-8583 東京都渋谷区代々木2-2-1 小田急サザンタワー
マイクロソフト株式会社 インダストリーマーケティング部内
(Windows NT[®]漢字処理技術協議会 事務局)**

印刷所 ムツミクリエテ株式会社



Windows NT[®] 漢字処理技術協議会

Windows NT[®] Extended Kanji Processing Council