

JIS X 0213 は、日本の文字コード規格です。名称を「7ビット及び8ビットの2バイト情報交換用符号化拡張漢字集合」といいます。日本の漢字を含む2バイトコード規格のうち、日本のテキストの符号化に最も適したものです。

漢字、平仮名、片仮名、アルファベット、記号類など、11,233文字のコードを定義しています。制定年や改正年をとって、JIS2000 や JIS2004 と呼ばれます。

符号化文字集合の概要

JIS X 0208 に不足していた現代日本の文字を符号化することをねらいとし、JIS X 0208 の上位互換として設計されています。ISO/IEC 2022 に整合的な形の、漢字集合1面・漢字集合2面という2つの2バイト符号化文字集合が定義されています。

このうち漢字集合1面が JIS X 0208 の上位互換です。平仮名・片仮名やダイアクリティカルマーク付きラテン文字、丸付き数字や記号類といった非漢字、それに比較的使用頻度の高い第3水準漢字が追加されています。漢字集合2面にはより使用頻度の少ない第4水準漢字が含まれていません。

追加すべき文字の選定にあたっては学校の教科書も全数調査され、学術用語として学校教科書に載っていないながら JIS X 0208 には無かった漢字も、JIS X 0213 でカバーされました。小学3年生で学習する日本語のローマ字綴りに用いられる長音符号付きの母音字は、JIS X 0208 にはありませんでしたが JIS X 0213 で対応しました。

その他、文字レパートリーの概要については、漢字集合1面、漢字集合2面の項をそれぞれ参照してください。

符号化方式の概要

ASCII など他の符号化文字集合と合わせた運用方式、いわゆる文字符号化方式について、JIS X 0213 は、シフト JIS や EUC といった、JIS X 0208 の全ての符号化方式に対応しています。

規格本体に6種類を規定しているほか、附属書にて、下記の符号化方式を定義しています。

- Shift JIS-2004 (附属書1)
- ISO-2022-JP-2004 (附属書2)
- EUC-JIS-2004 (附属書3)

それぞれ、JIS X 0208 における Shift JIS、ISO-2022-JP、EUC-JP に対応する符号化方式(おおまかに言ってこれらの拡張版)です。

規格本体で規定される6種類の符号化方式は下記の通りです。これらは JIS X 0208:1997 から踏襲されています。

- ・漢字用 7 ビット符号 - 漢字集合 1 面・2 面を SI/SO で切り替える 7 ビットの符号
- ・漢字用 8 ビット符号 - 漢字集合 1 面を GL 領域に、漢字集合 2 面を GR 領域に割り当てる 8 ビット符号
- ・国際基準版・漢字用 7 ビット符号 - ISO/IEC 646 国際基準版と漢字集合 1 面を SI/SO で切り替え。SS3 で漢字集合 2 面
- ・国際基準版・漢字用 8 ビット符号 - ISO/IEC 646 国際基準版を GL に、漢字集合 1 面を GR に。SS3 で漢字集合 2 面 (EUC-JIS-2004 とほぼ同じ)
- ・ラテン文字・漢字用 7 ビット符号 - JIS X 0201 ラテン文字集合と漢字集合 1 面を SI/SO で切り替え。SS3 で漢字集合 2 面
- ・ラテン文字・漢字用 8 ビット符号 - JIS X 0201 ラテン文字集合を GL に、漢字集合 1 面を GR に。SS3 で漢字集合 2 面

実装水準

上記の符号化方式には、それぞれ、実装水準 3、実装水準 4 というふたつの実装水準が設けられています。実装水準 3 は、漢字集合 2 面に対応しない実装であり、実装水準 4 は漢字集合 1 面・漢字集合 2 面の両方に対応した完全版の実装です。実装水準 3 の場合は漢字は第 1 水準から第 3 水準までの対応で、一方実装水準 4 では第 1 水準から第 4 水準漢字まで全ての対応となります。

このような実装水準が設けられているのは、漢字集合 2 面の対応には、従来の JIS X 0208 だけの実装には用いられなかった機構が要ることがあるためです。例えば EUC の場合ならシングルシフト 3 による対応が必要になりますし、SJIS なら漢字集合 2 面は区点番号との計算方法が漢字集合 1 面のものとは異なります。

従来の JIS X 0208 対応の実装に対して最小限の変更で済ませられるような簡易な適合性として実装水準 3 が用意されています。

Unicode ないし ISO/IEC 10646 との関係

JIS X 0213 はその各文字について、Unicode (ISO/IEC 10646 UCS) の対応する文字名・符号位置を定めています。これに基づいて、Unicode との間でコード変換ができます。

JIS X 0213 が 2000 年に制定された当初は Unicode がない文字もありましたが、その後 Unicode 3.2 で JIS X 0213 の全ての文字に対応しました。

ただし、Unicode では結合文字を使って複数の符号位置の並びとして表現する必要のある文字もあります。例えば鼻濁音表記用の平仮名・片仮名や、アイヌ語表記用の片仮名の一部がこれに当たります。

また、漢字の一部が CJK 互換漢字 による対応になっています。これらは Unicode 正規化の処理によって CJK 統合漢字 に置き換えられてしまいます。例えば「福」のへんが「ネ」でなく「示」の字体 (面区点 1-89-33、第 3 水準) がこれに当たります。

制定・改正履歴

- ・2000 年 制定。
- ・2004 年 改正。追補 1 を発行。例示字形を表外漢字字体表に合わせる変更。これに伴い第 3 水準漢字に「表外漢字 UCS 互換」10 文字を追加。
- ・2012 年 改正。追補 2 を発行。2010 年に改正された常用漢字表との対応を参考情報として

追加。

他の規格との関係

- [JIS X 0208](#) - [JIS X 0213](#) はこれの拡張版。
- [JIS X 0201](#), [ISO/IEC 646](#) - これらの 1 バイトコードと同時に運用する符号化方式が [JIS X 0213](#) で定義されている。また、これらのコードの文字を [JIS X 0213](#) は全て含む。
- [ISO/IEC 8859](#) - 第 1 部 (Latin-1)、第 2 部 (Latin-2) の文字を [JIS X 0213](#) は全て含む。
- [ISO/IEC 10646](#) ([JIS X 0221](#)) または [Unicode](#) - [JIS X 0213](#) は各文字についてこれに対応する符号位置を記述。(変換表として使用可能)
- [ISO/IEC 2022](#) - [JIS X 0213](#) はこれに整合的な構造であり、符号拡張の方式を用いて他の文字コードと組み合わせて使用可能。

参考

- [JIS X 0213 の代表的な符号化方式](#) - 各符号化方式の説明。
- [JIS X 0213 のあまり代表的でないかもしれない符号化方式](#) - 規格本体の 6 種類の符号化方式の説明。
- [JIS X 0213 のコード対応表](#) - [Unicode](#) との変換表があります。
- [漢字集合 1 面の文字コード表](#)
- [漢字集合 2 面の文字コード表](#)
- [JSA Webdesk](#) - 規格票を紙または PDF の形で購入できます。規格番号「X0213」で検索のこと。
- 当サイト全般