

EUC-JP は、JIS X 0208 を ASCII とともに運用する方式（いわゆる文字符号化方式）です。

符号の構造

8 ビット符号表に下記のように符号化文字集合を割り当てたものです。

- ・ GL 領域 (0x20-0x7F): ASCII (1 バイトコード)
- ・ GR 領域 (0xA0-0xFF): JIS X 0208 (2 バイトコード)
 - ・ ただし、0x8E の直後 1 文字は JIS X 0201 片仮名 (1 バイトコード)
 - ・ ただし、0x8F の直後 1 文字は JIS X 0212 (2 バイトコード)

0x8F, 0x8E は ISO/IEC 2022 で取り決められているシングルシフトの制御コードです。

なお、JIS X 0201 片仮名、JIS X 0212 の使用は元々オプションの扱いで、今日でも対応していない実装が見られます。

由来

EUC は Extended Unix Code という名のとおり Unix 系 OS での日本語用文字コードとして開発されたもので、元々は JIS で決められたものではありません。ただし、その中心部分、つまり ASCII と JIS X 0208 の部分は、JIS X 0208:1997 の「国際基準版・漢字用 8 ビット符号」と同一です。

JIS X 0213 との関係

JIS X 0213 では、EUC に対応する符号化方式として EUC-JIS-2004 が定義されています。EUC-JP の JIS X 0212 を用いない場合に対して上位互換となっています。

EUC-JP における JIS X 0212 は、EUC-JIS-2004 では JIS X 0213 の漢字集合 2 面に相当します。ここで、X0213 の漢字集合 2 面は JIS X 0212 で文字の割り当てられていない区点位置のみを使用しているため、JIS X 0213 か JIS X 0212 のどちらが使われているかを判断することができます。

関連項目

- ・ EUC-JIS-2004
- ・ 国際基準版・漢字用 8 ビット符号