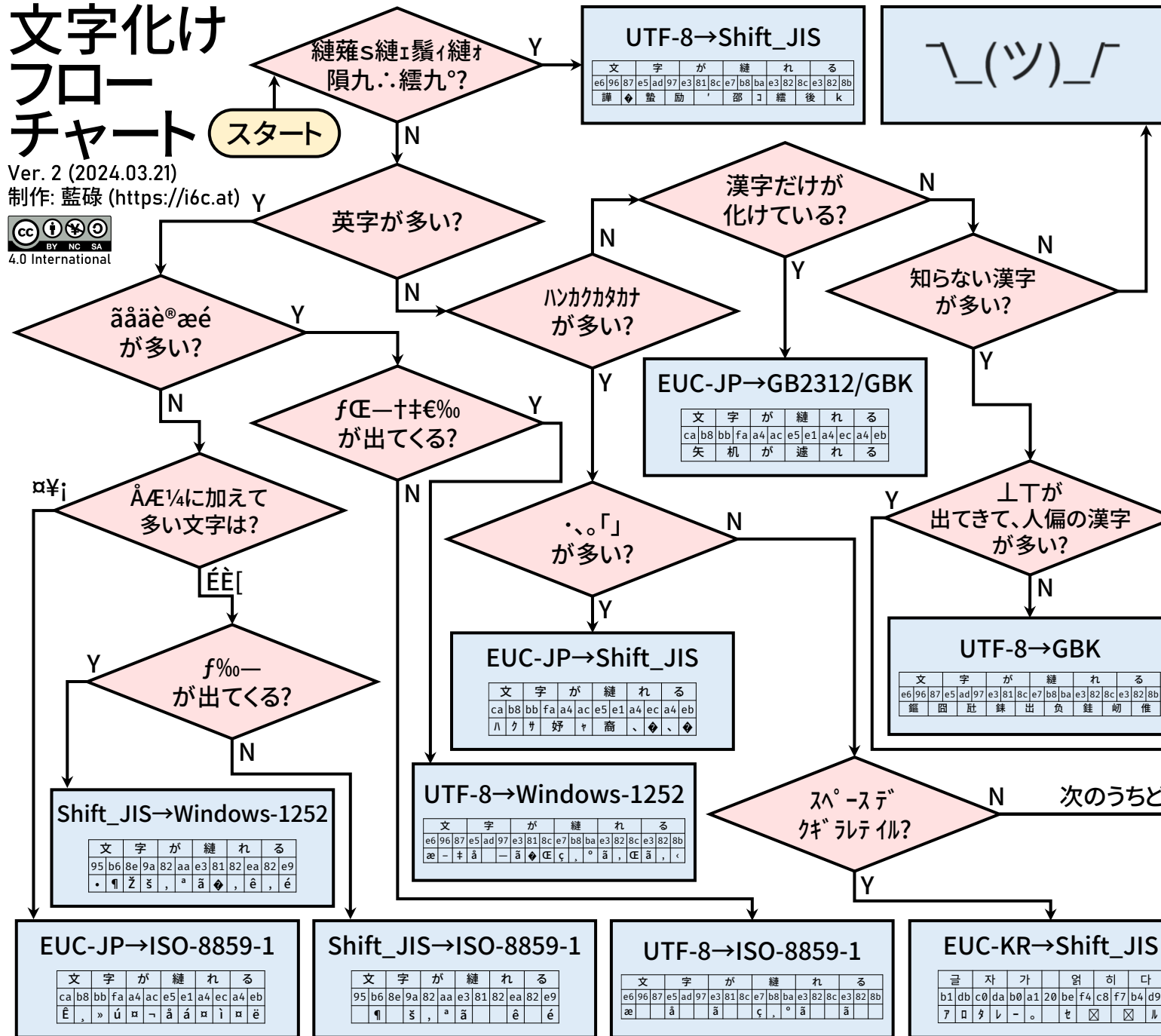


# 文字化け フロー チャート

Ver. 2 (2024.03.21)

制作: 藍碌 (<https://i6c.at>)



## 大まかな文字コードの解説

ASCII	7ビット, 米国用標準 (1963年) 国際版はISO/IEC 646 ('67)
JIS X 0201	8ビット, ISO/IEC 646に片仮名 などを追加 (1969年)
ISO/IEC 8859	8ビット, ASCIIの上位集合集で 欧州・中華・南亜用 (1987年)
ISO-8859-1	上の西欧用 (Latin-1)にHTTP 用に制御文字を追加した ('92)
Windows-1252	Latin-1の一部を文字・全角 約物で置換 別名CP1252 ('91)
JIS X 0208	日本語の漢字など利用可能 別名区点コード, JISコード ('78)
Shift_JIS	JIS X 0201・0208を共存させる (1982年発生 1997年制定)
Windows-31J	シフトJISのWindowsによる拡張 別名CP932, CP943 (1993年)
ISO/IEC 2022	ISO/IEC 646・8859等の併用法 逃避文字で符号表を切替 ('73)
ISO-2022-JP	7ビット, 区点コードを含むJUNET コードの規格化 ('86発生 '93制定)
EUC	ISO/IEC 2022の下位集合 東アジアでUNIXが使用
EUC-JP	シフトJIS・JIS X 0213補助漢字 をEUCで符号化 (1986年)
GB2312-80	中国版区点コード 主にEUCで符 号化して使用 (EUC-CN) ('80)
GBK (CP936)	2バイト, 上記+Unicode漢字の Windowsによる実装 (1995年)
GB18030-2000	可変長(1, 2, 4バイト) GBK及び Unicode 3.0と互換 (2000年)
Unicode	森羅万象文字コード (1992年) 代理組で17面に拡張 ('96)
UTF-16	2バイト, Unicodeの当初の 表現形 追加面はサロゲートペア
UTF-8	可変長(1~4バイト), ASCIIを そのまま扱える (1993年)