

文字コードにおいて包摂（ほうせつ）とは、ひとつの符号位置に複数の字体が対応することです。英語では unification といいます。

例えば、「逢」という字のしんによる点の数は 1 つでも 2 つでもあり得ますが、JIS X 0213 や Unicode といった文字コードではこうした 1 点しんによる・2 点しんによる字体を区別せずに同じ符号位置に対応させます。JIS X 0213 なら 1-16-9、Unicode なら U+9022 です。

漢字には字体の様々な揺れがありますが、包摂規準を参照することによって、その文字集合がどのような差を区別するか・しないかを確かめることができます。包摂という概念は漢字に限定されません。例えば算用数字の「4」や「7」にはストロークの接触や有無による字体差がありますが、常識的に誰でも分かるためわざわざ包摂規準を明示する必要がありません。それに比べて漢字は数が多く複雑なため、包摂規準を明示することが文字集合の体系としての理解のために有益です。

なお、誤解した説明がたまに見受けられますが、文字コードの包摂とは「ある符号位置が複数の字体を包摂する」のであって、「ある字体 A が別の字体 B を包摂する」ではありません。一般的に、ある符号位置に包摂される複数の字体間にいずれかの字体が優先されるといった関係はありません。

JIS X 0208 の包摂規準

JIS X 0208 は 1997 年改正で初めて規格内に明示的に包摂規準を規定しました。186 の規準が定義されています。これは全く何もないところからいきなり作られたわけではなく、過去の版の規格の「解説」に記載されていた包摂の例や、文字集合自体に内在する字体認識つまりどのような違いに対して別区点を割り当てているかといったことを考慮して定義されています。

ただし通常の包摂規準とは別に、83JIS において字体の大きく変えられた 29 の符号位置について、元の 78JIS 字体でも規格適合とするための救済用として特殊な包摂規準が設けられています (JIS X 0208:1997 6.6.4 節)。例えば、18 区 10 点の「鷗」は、78JIS ではへんが「區」であったのが 83JIS で変えられたものです。これは他の区点位置の例に照らせば本来 JIS X 0208 では包摂され得ない違いであることが分かりますが、83JIS が不適当な字体変更をしてしまったため、特例として、旧規格の実装を規格適合とするためにこの 2 つの字体を包摂しています。

JIS X 0213 の包摂規準

JIS X 0213 は JIS X 0208 の包摂規準を受け継ぎ、さらに 13 の規準を追加し 199 規準を設けています。

上記 83JIS の字体変更による特例の包摂規準は、JIS X 0213 では削除されています。該当する 29 区点位置に対応する 78JIS 字体は、JIS X 0213 において独立した区点位置が与えられたためです。例えば、上記「鷗」(1-18-10) の 78JIS 字体、つまりへんを「區」にしたものは、第 3 水準漢字として面区点位置 1-94-69 に配置されました。

ISO/IEC 10646 の包摂規準

ISO/IEC 10646 は JIS X 0213 ほど詳細な包摂規準を規定してはませんが、CJK 統合漢字においてどのような差異を統合（包摂）したかを説明している箇所があります (Annex S)。例えば上記のしんによりの 1 点・2 点が包摂対象となることがそこから分かります。

参考

- ・ JIS X 0208 と 0213 規格票の包摂関連項目（青空文庫）