

PRECIS Framework

Peter Saint-Andre
PRECIS WG
IETF 87, Berlin

Order of Operations

- Modified order of operations in the PRECIS algorithm
- OldHangulJamo now **before** PrecisIgnorableProperties
- HasCompat now **before** LetterDigits and OtherLetterDigits

OldHangulJamo

- Some characters were properly categorized as `DISALLOWED` but because of the wrong rule (`PrecisIgnorableProperties` instead of `OldHangulJamo`), e.g.:
 - `HANGUL CHOSEON FILLER (U+115F)`
 - `HANGUL JUNGSEONG FILLER (U+1160)`

HasCompat

- Some characters seriously mis-categorized (PVALID instead of FREE_PVAL), e.g.:
- FEMININE ORDINAL INDICATOR
(U+00AA) = ^a (basically superscript 'a')
- VULGAR FRACTION ONE QUARTER
(U+00BC) = $\frac{1}{4}$

Old Order

```
If .cp. .in. Exceptions Then Exceptions(cp);
Else If .cp. .in. BackwardCompatible Then
BackwardCompatible(cp);
Else If .cp. .in. Unassigned Then UNASSIGNED;
Else If .cp. .in. ASCII7 Then PVALID;
Else If .cp. .in. JoinControl Then CONTEXTJ;
Else If .cp. .in. PrecisIgnorableProperties Then DISALLOWED;
Else If .cp. .in. Controls Then DISALLOWED;
Else If .cp. .in. OldHangulJamo Then DISALLOWED;
Else If .cp. .in. LetterDigits Then PVALID;
Else If .cp. .in. OtherLetterDigits Then SAFE_DIS or FREE_PVAL;
Else If .cp. .in. Spaces Then SAFE_DIS or FREE_PVAL;
Else If .cp. .in. Symbols Then SAFE_DIS or FREE_PVAL;
Else If .cp. .in. Punctuation Then SAFE_DIS or FREE_PVAL;
Else If .cp. .in. HasCompat Then SAFE_DIS or FREE_PVAL;
Else DISALLOWED;
```

Proposed New Order

```
If .cp. .in. Exceptions Then Exceptions(cp);
Else If .cp. .in. BackwardCompatible Then
BackwardCompatible(cp);
Else If .cp. .in. Unassigned Then UNASSIGNED;
Else If .cp. .in. ASCII7 Then PVALID;
Else If .cp. .in. JoinControl Then CONTEXTJ;
Else If .cp. .in. OldHangulJamo Then DISALLOWED;
Else If .cp. .in. PrecisIgnorableProperties Then DISALLOWED;
Else If .cp. .in. Controls Then DISALLOWED;
Else If .cp. .in. HasCompat Then ID_DIS or FREE_PVAL;
Else If .cp. .in. LetterDigits Then PVALID;
Else If .cp. .in. OtherLetterDigits Then ID_DIS or FREE_PVAL;
Else If .cp. .in. Spaces Then ID_DIS or FREE_PVAL;
Else If .cp. .in. Symbols Then ID_DIS or FREE_PVAL;
Else If .cp. .in. Punctuation Then ID_DIS or FREE_PVAL;
Else DISALLOWED;
```

Next Steps

- Validate algorithm changes with WG
- Compare output of createtables.rb and PrecisMaker.py (note: PrecisMaker needs some feature enhancements first)
- Perform comparative testing (output of stringprep + Unicode 3.2 vs. output of PRECIS + Unicode 6.2)