CJK Integration ==Procedure== (v0.==5==)

Precondition

- Each of CJK Generation Panels (GPs) generates LGR for each language TLD before integration.
  - ➢ CJK GPs pick up ideographic variants for CJK from domain name usage perspective.
  - ➢ CJK GPs don't elaborate ideographic variants for CJK from linguistic perspective.
- CJK GPs agree on the mechanism (steps) to integrate and extract each language (script) LGR.

Step 1: Each CJK GP generates its own LGR (hereinafter, LGR-alpha)

- LGR-alpha generation process is left to each CJK GP.
- LGR-alpha format must follow XML schema for LGR.
  <https://datatracker.ietf.org/doc/draft-ietf-lager-specification/>
- Each code point must have individual <char> element (i.e. don't use <range> element).
- Each <char> element in LGR-alpha must have reflexive mapping as <var> element (i.e. each code point must have explicit variant type/subtype).
- WLE of Each LGR-alpha must not conflict (Conflict of WLE must be coordinated and solved at this step).

Step 2: CJK GPs collectively generate a merged table of each LGR-alpha (hereinafter, LGR-M)

1. Extract every <char> element tagged "sc:Hani" from each LGR-alpha.
2. For each extracted <char> element, check the existence of another <char> element with the same code point ("cp" value), and if exitsts, merge them into one element. At this time, "type" attribute of <var> element must be removed.
3. After the check was finished, record every merged <char> elements to LGR-M.

- Repertoire of LGR-M is the union of all sc:Hani in each CJK LGR-alpha.

36   ●   Variants of each <char> elements in LGR-M is the union of all variants defined
37   for the code point in each CJK LGR-alpha.
38   ●   LGR-M does not have following information.
39   ➢   Language tag.
40   ➢   Variant type/subtype attribute of <var> elements.
41   ➢   WLE (Whole Label Evaluation rules).
42

Step 3: Each CJK GP extract its original repertoire with integrated variants from LGR-M.

46 1. For each <char> element in its LGR-alpha (hereinafter, ${char:: LGR-alpha}), extract <char> element of the same code point ("cp" value) from LGR-M (hereinafter, ${char::LGR-M}).

49   1.1. For each <var> element in ${char::LGR-M} (hereinafter, ${var::LGR-M}), compare with <var> elements in ${char::LGR-alpha} (hereinafter, ${var::LGR-alpha}).

52   1.1.1. If ${var::LGR-M} has the same code point ("cp" value) with one of ${var::LGR-alpha}, then copy "type" attribute of corresponding ${var::LGR-alpha} to ${var::LGR-M}.

55   1.1.2. Otherwise, if ${var::LGR-M} has the same code point ("cp" value) with any of <char> elements in LGR-alpha, set "type" attribute with value "review" to ${var::LGR-M}.

58   1.1.3. Otherwise, set "type" attribute with value "blocked" to ${var::LGR-M}, and record the code point of ${var::LGR-M} to Out-of-Repertoire list (hereinafter, ${OoR-list}).

61   1.2. Record ${char::LGR-M} to Integrated-Repertoire list (hereinafter, ${IR-list}).

62   1.3. If ${IR-list} includes "review" type, ==responsible GP== execute either (a) return to Step 1 and redefine appropriate type/subtype to corresponding <var> elements, or (b) replace all "review" by "blocked".

66   ●   ==Whether "review" type is included in ${IR-list} is differ by each GP.   Therefore,==
67   ==each GP should decide "review" type treatment independently.==
68

Step 4: Each CJK GP add "Out of Repertoire" code points for symmetry.

71 1. For each code point in ${OoR-list} (hereinafter, ${cp::OoR-list}), extract <char>

72 element of the same code point ("cp" value) from LGR-M (hereinafter,

73 ${char::LGR-M}).

    1.1. For each \<var> element in ${char::LGR-M} (hereinafter, ${var::LGR-M}),

        compare ${cp::OoR-list} and code point ("cp" value) of ${var::LGR-M}.

    1.2. If the two code points are the same, add "type" attribute with value "out-of-repertoire-var" to ${var::LGR-M}.

    1.3. Otherwise, add "type" attribute with value "blocked" to ${var::LGR-M}.

    1.4. Record ${char::LGR-M} to ${IR-list}.

Step 5: Each CJK GP merge WLE in LGR-alpha into one.

1. Each GP extract \<rules> element from each LGR-alpha and merge them into one WLE (generate integrated \<rules> element, hereinafter, ${rules::LGR-M}).

2. Each GP add following rule to ${rules::LGR-M} for handling "out-of-repertoire-var" variant type.

   \<action disp="invalid" any-variant="out-of-repertoire-var" />

Step 6: Each CJK GP generates integrated LGR (hereinafter, LGR-beta).

1. Each GP extract preambles from its LGR-alpha (hereinafter, ${preamble::LGR-alpha}).

2. Each GP extract all \<char> elements with "tag" value other than "sc:Hani" and record to IR-list.

3. Each GP merge ${preamble::LGR-alpha}, IR-list and ${rules::LGR-M} into LGR-beta.

● In other words, this step replaces body of \<data> element and \<rules> element of LGR-alpha to IR-list and ${rules::LGR-M} respectively.