Terminology for CJK coordination committee and IP (draft)

Background

We are collaborating to create Root zone LGR for CJK. CJK GPs' collaboration is not limited to among us, it is also including IP and ICANN. For the efficiency and the good progress, we should have common terminology definition for key concepts.

Terminology

(1) Script

Based on Unicode, script is a collection of letters and other written signs used to represent textual information in one or more writing systems. Ex. Han, Hiragana, Katakana and Hangul.

(2) Chinese script, Japanese script, Korean script

Script(s) used in the language. Chinese script is Han, Japanese script is a mixture of Han, Hiragana and Katakana, and Korean script is a mixture of Han and Hangul. In Root zone LGR context, Chinese script, Japanese script and Korean script are equivalent to und-hani, und-jpan and und-kore respectively. Here, hani, jpan and kore are from ISO 15924.

(3) Disposition

Result of whole label evaluation (WLE). Disposition is assigned to a label, not to a character.

(4) Variant type, variant subtype

Variant type is an attribute of a variant, which indicates the treatment of the variant in WLE. Variant type is one of (a) allocatable, (b) blocked and (c) out of repertoire var. The "allocatable" variant type can have subtype (variant subtype).   The variant subtype can be defined by each GP.   For example, CGP defines (i) simp, (ii) trad and (iii) both variant subtypes. Each variant subtype has to have one or more corresponding <action> element in WLE.

(5) LGR-1

LGR-1 is an LGR defined by each GP for CJK integration purpose. Each CJK GP can define LGR-1 independently from each other. LGR-1 must consist of language tag (one of und-hani, und-jpan, and und-kore), repertoire of allowed characters for applied-for label, variants of each character where variant type (and variant subtype if necessary) is associated with each variant characters, and WLE. The variant type must not have "out of repertoire var". LGR-1 is an intermediate product and it is not a final proposal of each GP.

(6) LGR-2

LGR-2 is a LGR generated by LGR-1 integration process. LGR-2 is generated for each Chinese/Japanese/Korean script. The differences between each LGR-2 are (A) language tag, (B) repertoire, (C) variant type (including variant subtype) of each variant. The set of variants and WLE are common between LGR-2s. LGR-2 is a final proposal of each CJK GP.

NOTE:
THESE TERMINOLOGY DEFINITIONS ARE VERY PRELIMINARY AND ARE NOT EXHAUSTIVE. THIS IS PREPARED FOR (Y)OUR REVIEW.