CGP Update
Chinese Script Root Zone Label Generation Rules (LGR)

Kenny Huang, Ph.D. 黃勝雄博士

Co-chair, CGP

Executive Council Member, APNIC

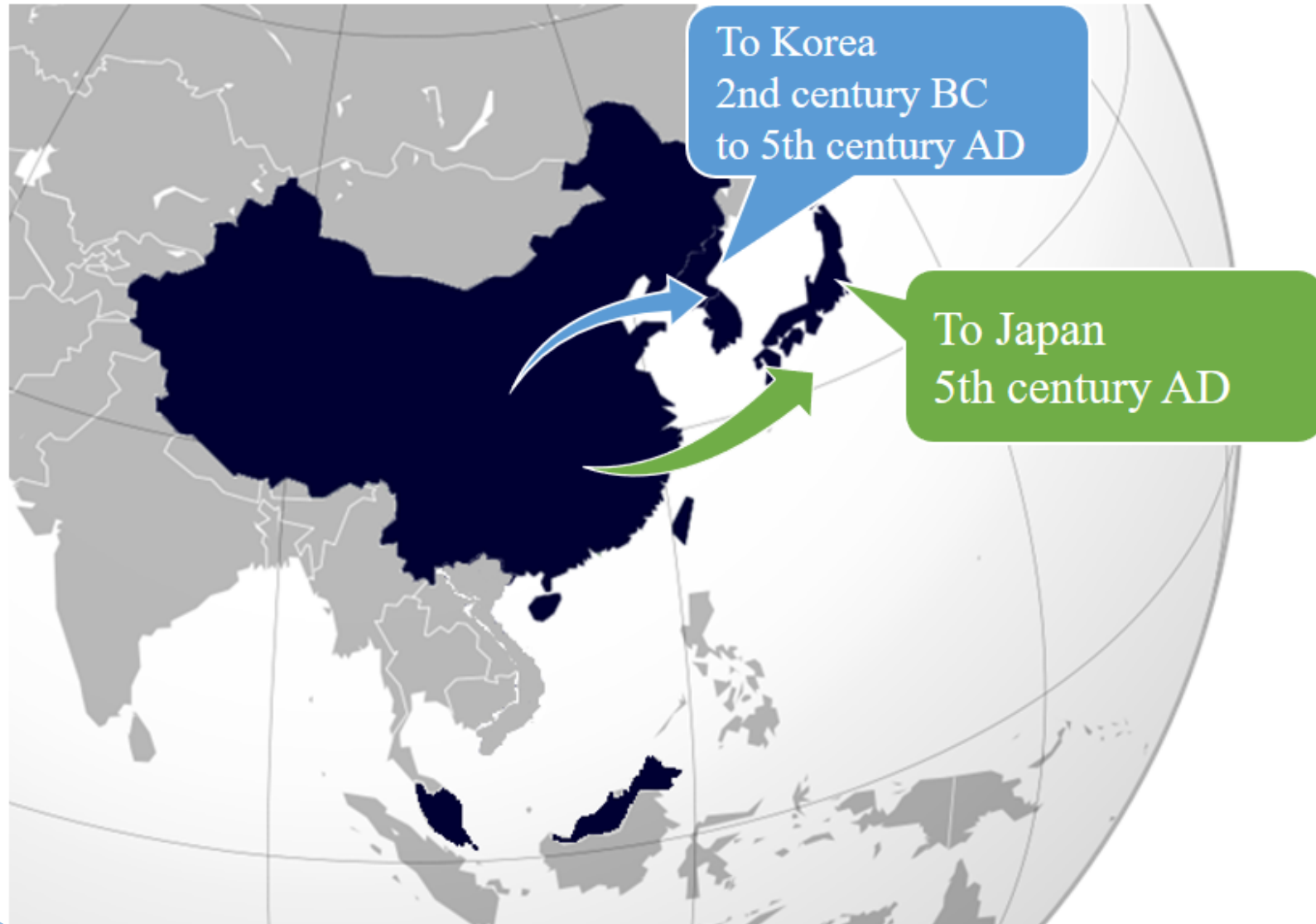huangksh@gmail.com

APrIGF Taipei

28 July 2016

Evolution of Chinese Character

source : CGP Proposal

Chinese Script Spread to Japan and Korea

To Korea
2nd century BC
to 5th century AD

To Japan
5th century AD

source : CGP Proposal

Countries and Regions Using Chinese Script

■ (dark green)	Traditional Chinese script used exclusively or almost exclusively (Taiwan, Macau and Hong Kong)
■ (green)	Simplified Chinese script used exclusively or almost exclusively (Mainland China and Singapore)
■ (light green)	Simplified Chinese script used formally but Traditional script still used widely (Malaysia)
■ (pale green)	Chinese script used with other systems of writing in the same language Kanji (Japan)
■ (grey-green)	Chinese script no longer officially used Hanja (Republic of Korea)

source : CGP Proposal
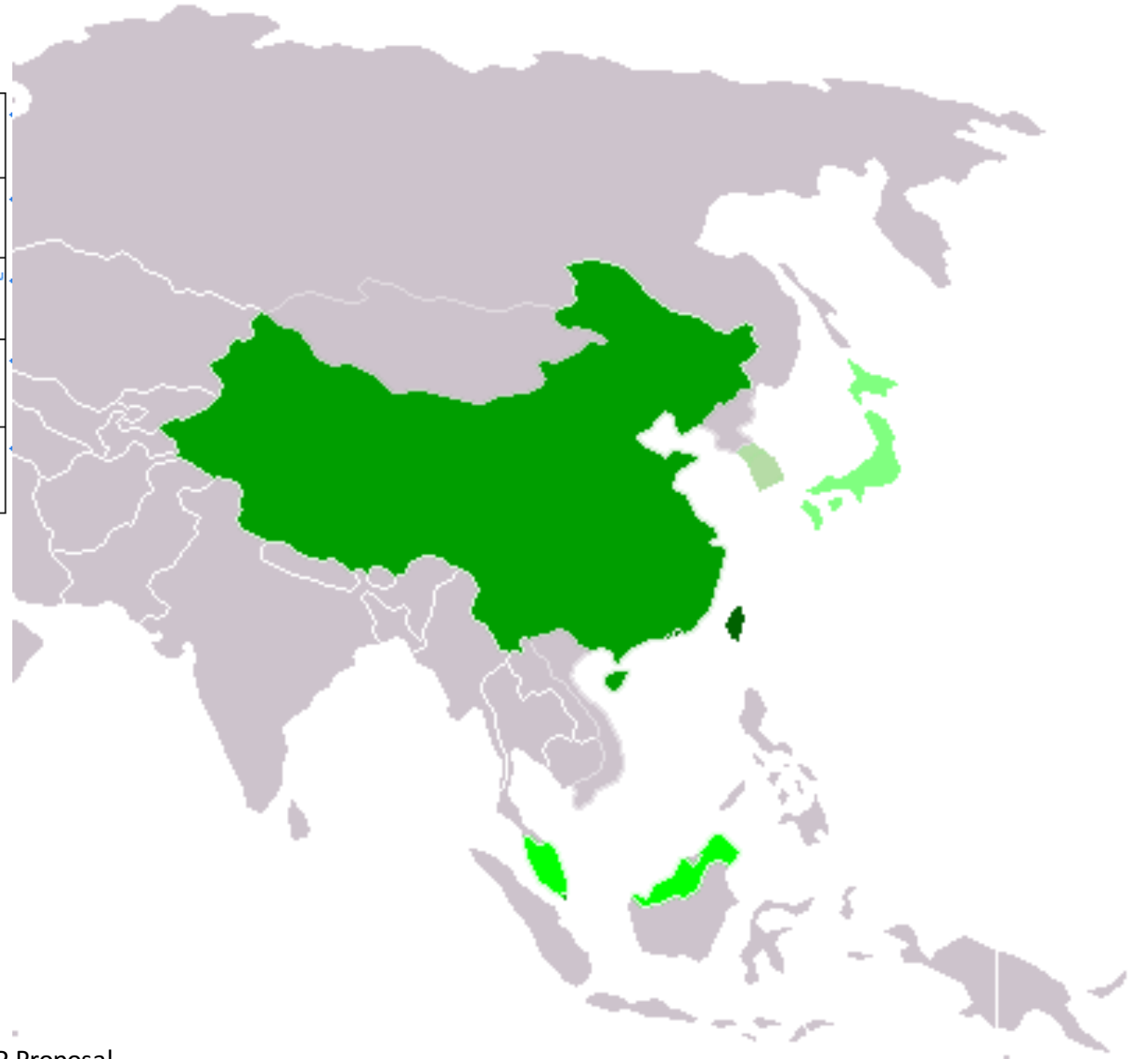
Hanzi, Kanji and Hanja
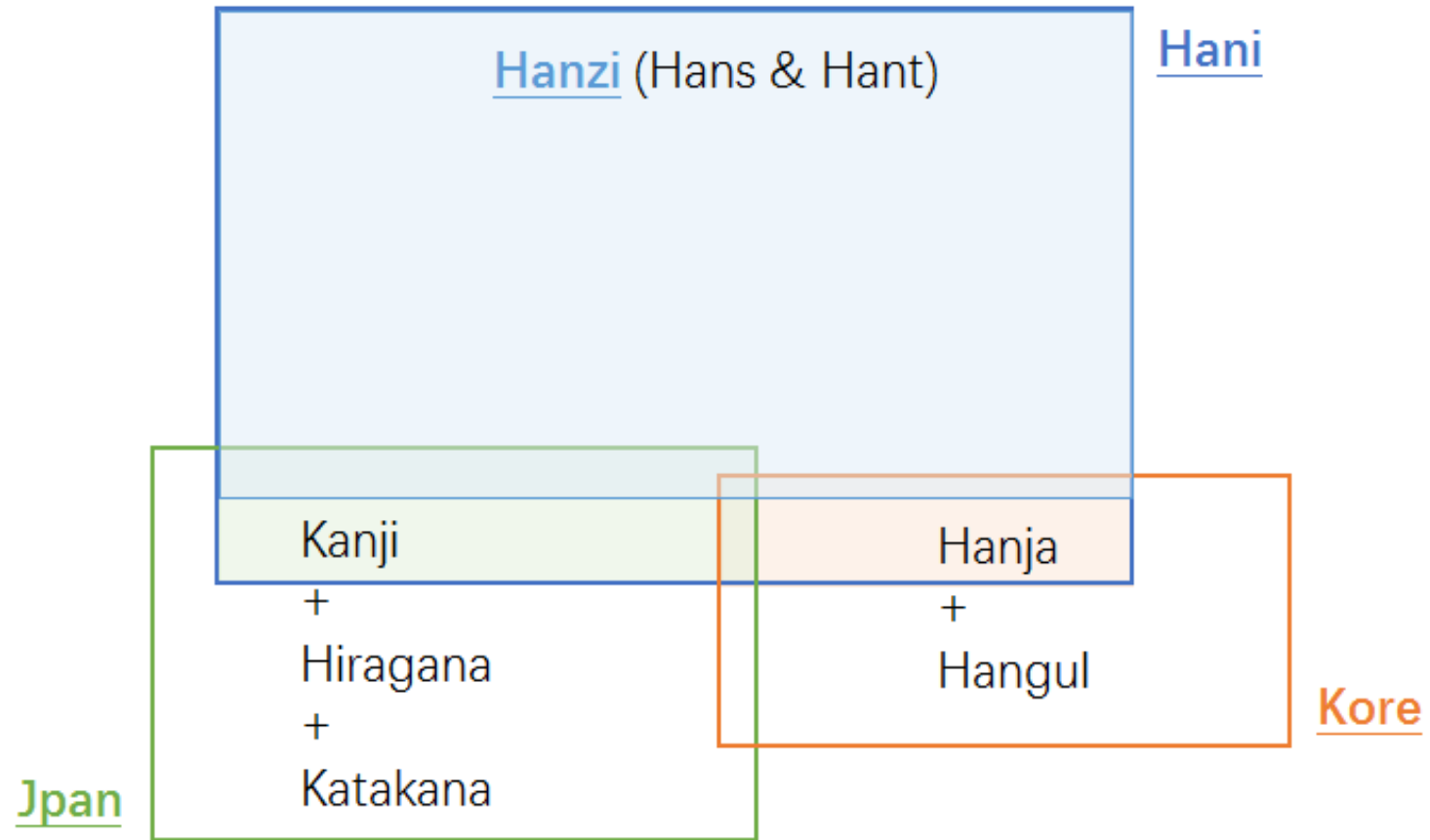
source : CGP Proposal

Variant solutions in different scripts

- **CDNC: RFC 3743 & 4713**
- Allocate all Applied-for IDL and Variant IDLs to the same registrant
- Delegate Applied-for IDL, Preferred SC IDL, Preferred TC IDL
- Reserve all the other Variant IDLs
- Delegate reserved variant IDLs when requested at a later date

- **JPRS: No Variant issue**
- Among Kanji characters, some are in a simplified form (called the "new character form"), derived from the traditional imported form (called the "old character form").
- It is appropriate to distinguish new and old forms as different and independent characters instead of pure variants. This understanding has been reflected in the IANA IDN table developed by the JPRS, in which no variants are identified for Kanji.

- **KISA: No Variant issue. Concern about overlapped code points with CDNC**
- Hanja is no longer widely used in the ROK. A law enacted in 2011 orders all ROK official government documents to be written ONLY in Hangul.
- KISA stated that its SLD IDN policy does not allow and nor does they have any intention of allowing the use of Hanja in their domestic market.

Coordination Principles

Each CJK panel creates an LGR and each LGR includes a repertoire and variants.

The variant mappings must agree for the same code point for all LGRs.

The variant types may be different (blocked or allocatable), the variant types do not have to agree across LGRs

The repertoires may be different

Allocatable

A potential allocation rule says that once the variant label is generated, that variant label may be allocated to the applicant for the original label.

Blocked

A blocking rule says that a particular label must not be allocated to anyone under any circumstances.

Coordination Example

For CGP

Code Point	Allocatable Variant	Blocked Variant	Tag
爱 U+7231	愛 U+611B	-	und-hani
愛 U+611B	爱 U+7231	-	und-hani

For JGP, probably

Code Point	Allocatable Variant	Blocked Variant	Tag
愛 U+611B	-	-	und-jpab

Code Point	Allocatable Variant	Blocked Variant	Tag
爱 U+7231	愛 U+611B	-	und-hani
愛 U+611B	爱 U+7231	-	und-hani
愛 U+611B	-	爱 U+7231	und-jpan

Iteration Model of CGP Work Process

source : CGP Proposal
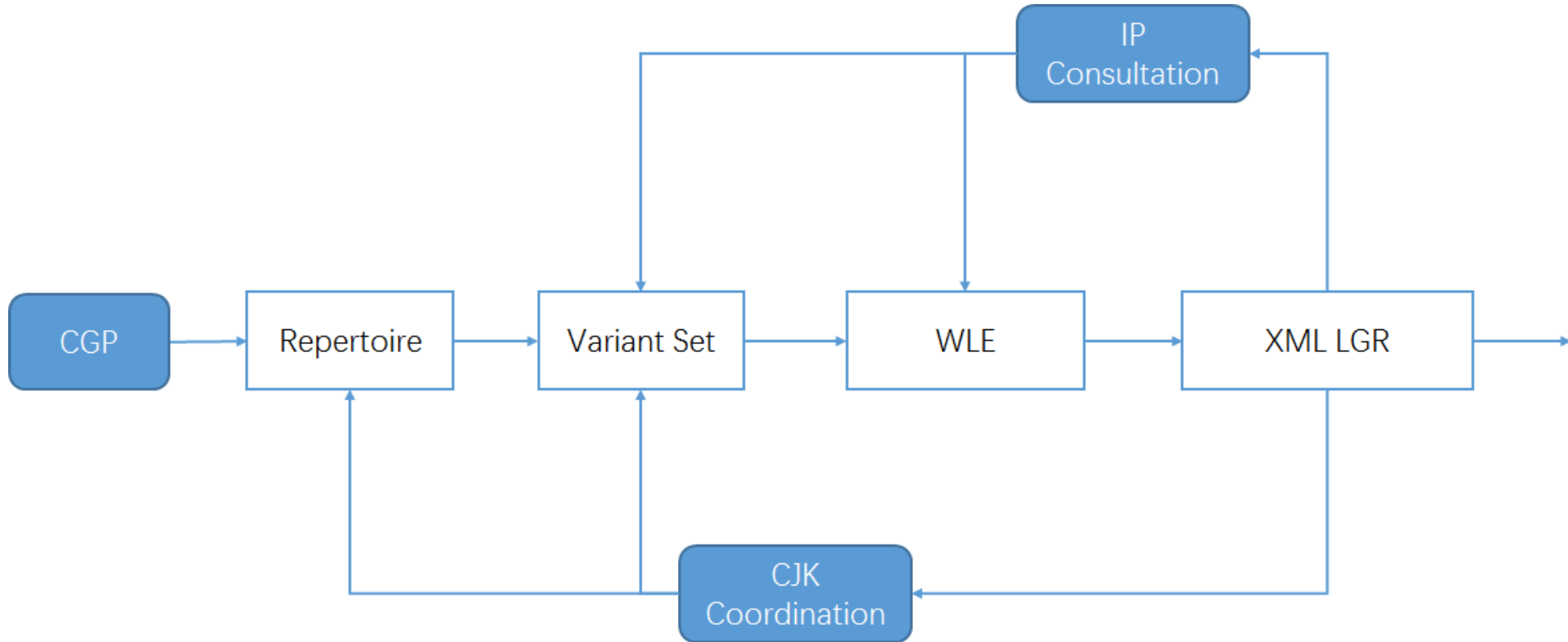
CGP Repertoire Components
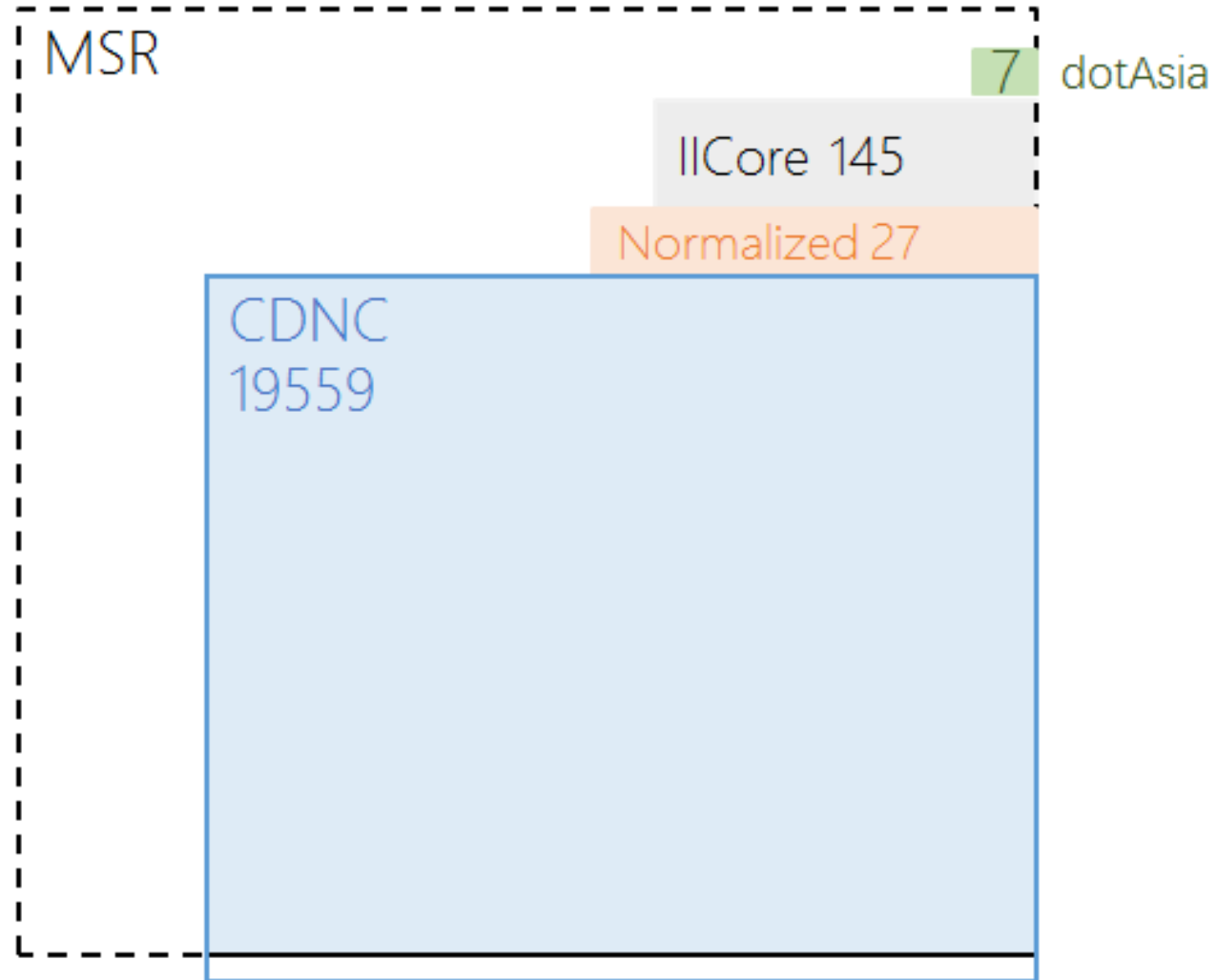
source : CGP Proposal

CGP Hanzi set, JGP Kanji set and KGP Hanja set

JGP: 6532

KGP: 4819

4147

6306

CGP: 19738

source : CGP Proposal
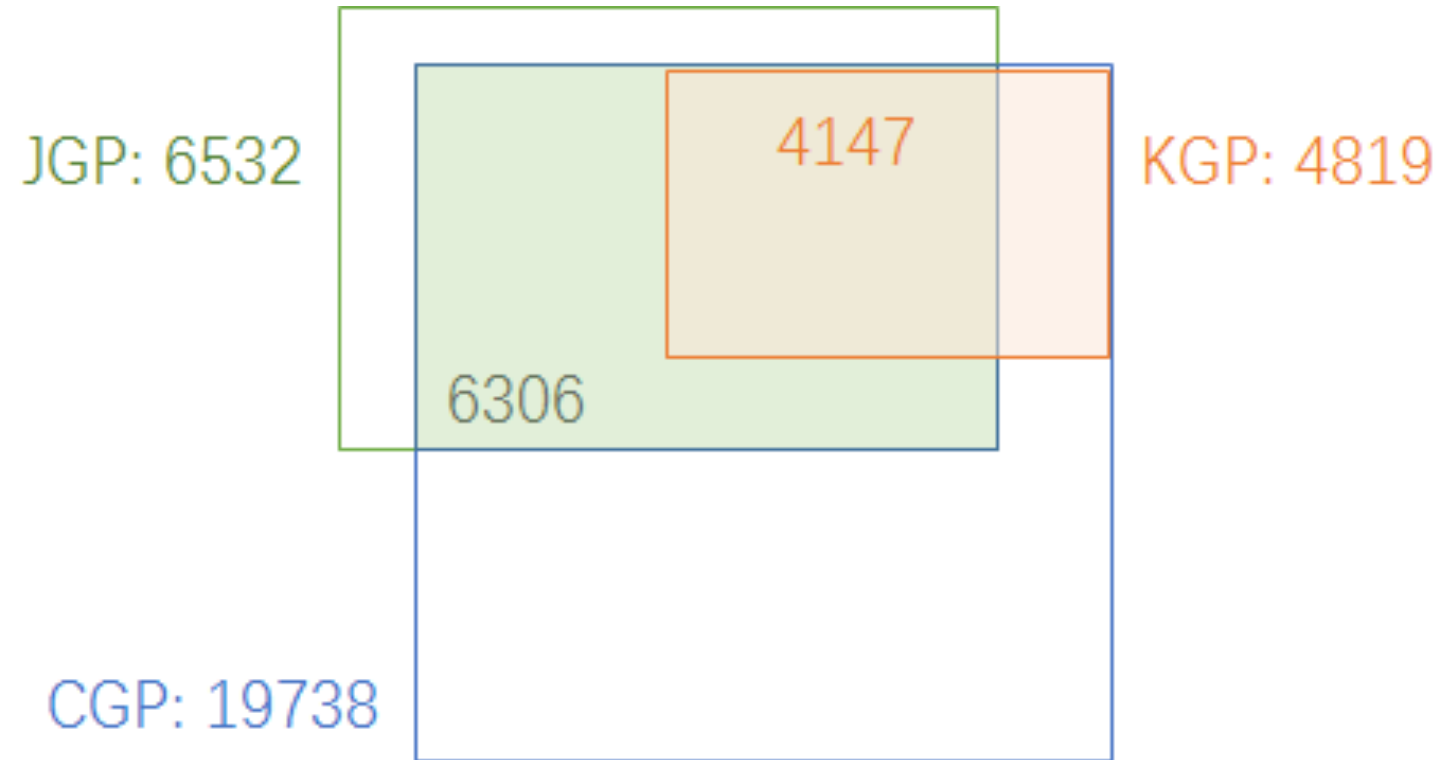
Variant Setting in CDNC IDN Table

86, China Country Code, preferred simplified variant

886, Taiwan Country Code, preferred traditional variant

All variants, including reflexive one.

```
U+4F4E(0);  U+4F4E(86),  U+4F4E(886);  U+4F4E(0);
U+4F4F(0);  U+4F4F(86),  U+4F4F(886);  U+4F4F(0);
U+4F50(0);  U+4F50(86),  U+4F50(886);  U+4F50(0);
U+4F51(0);  U+4F51(86),  U+4F51(886);  U+4F51(0);
U+4F52(0);  U+4EF0(86),  U+4F52(886);  U+4EF0(0),U+4F52(0);
U+4F53(0);  U+4F53(86),  U+9AD4(886);  U+4F53(0),U+8EB0(0),U+8EC6(0),U+9AB5(0),U+9AD4(0);
```

reserved variants

"simp"	preferred simplified;
"r-simp"	reflexive preferred simplified;
"trad"	preferred traditional
"r-trad"	reflexive preferred traditional
"both"	preferred simplified and traditional ones are the same
"r-both"	reflexive preferred simplified and traditional ones are the same

source : CGP Proposal

Framework of CJK LGR Integration for Han Characters

source : JGP
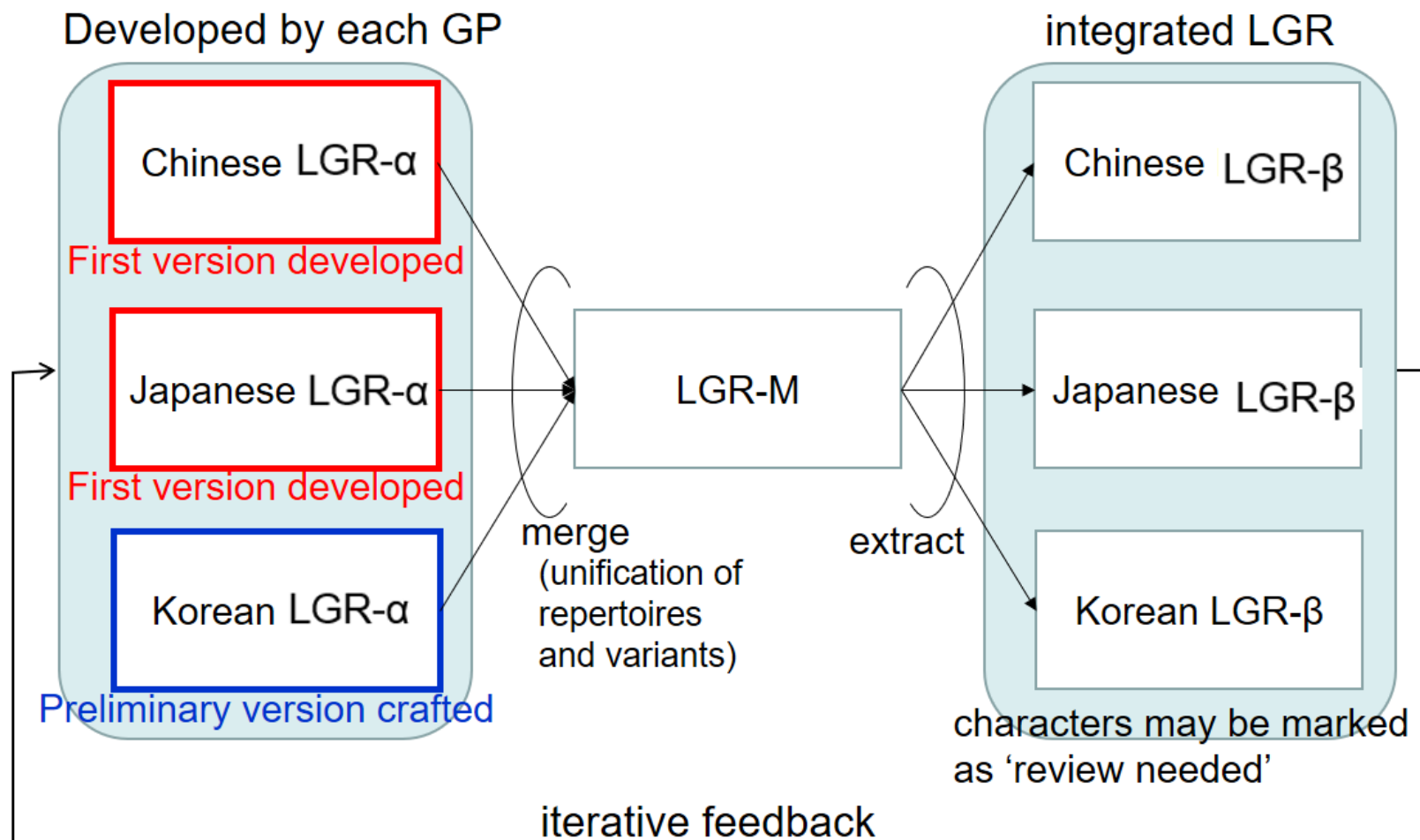
Next Step

- Coordination with CJK community
- Adopting guiding principles and seeking rough consensus
    - Technical feasibility
    - Cultural and social considerations
- A harmonized solution is anticipated before the end of 2016

Thank You
Question?