Minutes of the Adhoc meeting on submitted documents: N941, N942, N944, N945, N948, N949

Discussion on N941

We accept that in mainland Chinese there is a difference between stroke normalization, and simplification including derived simplification.

We recognizes that many processes need to distinguish characters in its simplified form and in the traditional form in display and other processes. However, other processes need to treat them the same. Therefore, the IRG should not reject the submissions from consideration. But a new character in the simplified form or in the traditional form still needs proof of use before you can submit it.

In the past, stroke normalization was considered for unification, but simplified characters were.

Discussion of N942
Different countries and regions have different ways of writing their characters. Since IRG editorial group requires the information on first stroke, editors are facing problems in checking the correctness of the first stroke as the editors are from different places. It is generally agreed that the IRG editorial group should provide some agreed first stroke list so that checking can be consistent.

We accept the contribution from Macao (N942) and we consider that it is a good starting point for this complied list.

The list should include two types of cases:

1.The same character/components in shape, but their first stroke is different, e.g. 女
2. The character/components that are considered variants, and they need to be listed

explicitly, e.g. a character with the following 3 glyphs, 户 , 尸 , 户 , as its components are all different, but since they need to be unified, the first stroke must also be agreed to facilitate the unification process.
It is agreed that we will try to produce a first stroke list and a review schedule during IRG 20. The total stroke count issue will also be discussed in this group.

Discussion on N944 and N945:

DPR Korea can provide the correct mapping and submit it to WG2 as an amendment or corrigendum. Such data will be accepted more readily by WG2 if there is confidence in the stability of the mapping. Although the quality of the mapping data is primarily the responsibility of the source submitting member, the IRG is willing to help out with the quality assurance. The IRG recommends that the DPRK do another

round of internal review, which can be followed by a review by the IRG at meeting 21.

After contact SC2, IRG was told that DPRK can proceed with the change of the mapping table as DPRK's mapping table was not yet published.

Discussion on N948:
IRG accepts the request from Vietnam to change their source names in C1.

Discussion on N949:
The purpose of the document is to write further explanatory notes for Annex S to so that IRG editors and anyone helping with unification can follow Annex S in a more consistent way. With this understanding, the group consider that most of the content from Section 1 to 3 of N949 is not appropriate for this purpose. Actually Section 1 to 3 are some discussion items from C1 Editorial Group. Therefore, we consider that they need to be discussed under a separate group meeting. The following is the Ad hoc group meeting on the discussion of Ext. C1 Unification related issues:

> There are unification rules under Annex S of ISO 10646. The IRG recognizes the fact that the examples given in the Annex S are typical examples, but not meant to be exhaustive. However, in the past IRG practice, editors who cannot find unification examples in Annex S, would have no methods to unify them, thus causing over-disunification. It is unanimously agreed that we should work out a procedure to avoid over-disunification. The suggestion is to produce an IRG Standing document that can collect more unification examples which are agreed by editors.
>
> Under circumstances where the Editorial Working Group cannot easily apply unification rules to doubtful characters under examination, IRG's recommendation to the editors is to propose unification whenever possible instead of proposing new characters. It is understood that under that current practice, a character already in a coded character set cannot be dis-unified/unified easily. However, for characters that are not yet coded in any national/regional standard it is easier to correct an improper unification than an improper disunification; even though, the dis-unification process is rather painless it is not without cost. The Editorial Group will keep the list of unification examples under the Annex S updated from time to time in its standing document, referred to as "IRG Unification Standing Document".

Continued discussion of N941:

Rules for determining when a simplified (or traditional) character with an encoded traditional (or simplified) form should be accepted:

1. Simplified forms found on government lists of official simplifications are always accepted.

2. Forms where the simplified (or traditional) counterpart is derived algorithmically need to have proof of use. There are thousands of theoretically possible simplifications, of which only a few are in use and necessary.

   a. General publications such as school books, or popular magazines are proof of use.
   b. Dictionaries are weak evidence of use. Many dictionaries include derived simplified or traditional forms for all characters which have them.
      i. A dictionary of simplified Chinese which includes a traditional form is not proof that the traditional form is in use.
      ii. A dictionary of traditional Chinese which includes a simplified form is not proof that the simplified form is in use.
      iii. A dictionary published originally in simplified Chinese is (weak) evidence of use for the simplified forms it contains. Large numbers of simplified forms from such a dictionary may be found only because they have been derived from a dictionary of traditional Chinese; as such, their inclusion in the dictionary is not proof of use. It would need to be proven that the dictionary is not merely taking its repertoire form a pre-existing traditional dictionary or other list of traditional characters.
      iv. The same is true for traditional Chinese, with the additional point that publication in a dictionary of traditional Chinese before the introduction of the simplified Chinese by the PRC government is (weak) evidence of use.

Discussion of 937:
We acknowledge the defect report from Vietnam. As soon as Vietnam provides a printed copy of the draft national standard, IRG is willing to review the mapping table.

The IRG urges its members to check future submissions against their own past submissions.

Discussion of N951:

IRG accepts the contribution N951. Every member is encouraged to give comments to Mr. James Seng within two months. James will produce a revised document by IRG#21.

Ad Hoc Group meeting closed at 2:30.