Korea JTC1/SC2, Committee on Character Codes

Author: KIM, Kyongsok, Chairperson of Korea JTC1/SC2
Date: 2013.05.17.
Status: NB position
Subject: How to reduce # candidate chars to 4000 for IRG review

1. In general, the total number of characters submitted for each project (e.g., ExtG) from all MBs/IOs can exceed 4000.
   - According to IRG PnP, IRG needs to reduce the number of candidate characters to 4000 for IRG review.
   - A method is suggested below on reducing # of candidate characters to 4000.

2. How to reduce # candidate chars to 4000 for IRG review
   - It is suggested that each MB/IO can submit at most 1,000 chars for each project (e.g., ExtG). Based on this assumption, a method to reduce the number of candidate chars to 4000 is explained below step by step.

2.1 Calculate the total number of chars submitted
   - Suppose that, as an example, the number of characters submitted by MBs/IOs is as follows (see column S). Assume that MBs are sorted in ascending order of number of chars submitted.

<table>
<thead>
<tr>
<th>subm</th>
<th>step 1</th>
<th>step 2</th>
<th>step 3</th>
<th>step4 [final step]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(200)</td>
<td>(600)</td>
<td>(700)</td>
<td>(900)  (900)</td>
</tr>
<tr>
<td>MB1:</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200    200</td>
</tr>
<tr>
<td>MB2:</td>
<td>600</td>
<td>200</td>
<td>600</td>
<td>600    600</td>
</tr>
<tr>
<td>MB3:</td>
<td>700</td>
<td>200</td>
<td>600</td>
<td>700    700</td>
</tr>
<tr>
<td>MB4:</td>
<td>900</td>
<td>200</td>
<td>600</td>
<td>700    900 833</td>
</tr>
<tr>
<td>MB5:</td>
<td>1000</td>
<td>200</td>
<td>600</td>
<td>700    900 833</td>
</tr>
<tr>
<td>MB6:</td>
<td>1000</td>
<td>200</td>
<td>600</td>
<td>700    900 833</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>-------------------</td>
</tr>
<tr>
<td>total</td>
<td>4400</td>
<td>1200</td>
<td>3200</td>
<td>3600   4200 3999</td>
</tr>
</tbody>
</table>
2.2 determining the max. number of chars for each MB

- Principle: Assign one char to each of six MBs at each round while the total number of chars assigned to 6 MBs is less than or equal to 4,000.

- **step 1**: Try to assign 200 (the number of chars submitted by MB1) chars for all six MBs. It is shown in “step 1” column. The total number of chars assigned to 6 MBs is 1,200. It is less than 4,000. We are done with MB1.

- **step 2**: Try to assign 600 (the number of chars submitted by MB2) chars for the remaining five MBs (i.e., MB2 ~ MB6). It is shown in “step 2” column. The total number of chars assigned to 6 MBs is now 3,200. It is less than 4,000. We are done with MB2 (and MB1 in step 1).

- **step 3**: Try to assign 700 (the number of chars submitted by MB3) chars for the remaining four MBs (i.e., MB3 ~ MB6). It is shown in “step 3” column. The total number of chars assigned to 6 MBs is now 3,200. It is less than 4,000. We are done with MB3 (and MB1 and MB2 in previous steps).

- **step 4**: Try to assign 900 (the number of chars submitted by MB4) chars for the remaining three MBs (i.e., MB4 ~ MB6). It is shown in “step 4” column. The total number of chars assigned to 6 MBs is now 4,200. It is greater than 4,000. We need to step back and take a final step.

- **final step**: Assign 400 (= 4,000 - 3,600, which is the total number of chars assigned to 6 MBs in step 3) chars equally to the remaining three MBs (MB4 ~ MB6) in addition to 700 chars already assigned in step 3. In other words, assign 833 (= 700 + 400 / 3) chars to MB4, MB5, and MB6.

  It is shown in “final step” column. The total number of chars assigned to 6 MBs is now 3,999. We are finally done.
3. Parameters

3.1 There is ONE parameter that can be adjusted/determined by IRG in the above steps.
- The number of chars that each MB/IO can submit (1000 in the above example). This number seems reasonable; however, IRG can decide the value.

3.2 To implement the above method, each MB will assign a priority number for each char in the submission form (we need to modify IRG PnP by adding one more column to submission form). A priority number will start with 1 and the max. value is the same as the number of chars from each MB.

- In the above example, MB4 submitted 900 chars. Each char is allocated a priority number whose value is between 1 and 900 inclusive. Priority numbers will be assigned by MB/IO based on the importance of each char.

- According to the above method, only 833 chars from MB4 can become candidate chars for IRG review.

- Then 833 chars whose priority numbers are between 1 and 833 inclusive will become candidate chars for IRG review and the remaining 67 chars whose priority numbers are between 834 and 900 inclusive will be excluded from IRG review for this project (e.g. ExtG). Those 67 chars could be possibly submitted for ExtH later.

- Comments are welcome. Thanks.

***