

CODE COLLISIONS IN THE PROPOSAL OF MICHAEL EVERSON!

Working document with error samples from N3532

BAKONYI Gábor. (Hungary, Budapest, Csillaghegy)

February 23, 2009

1 Introduction

In the document N3532.pdf, Michael Everson tries to pack incompatible things to a common pack. A character has several equivalent glyph form, but this solution doesn't take it into consideration.

For example, we could see on the figure1, that during the centuries in different territories of the country, different character styles were used. We can see in the table, that disregard the small changes of the glyphs, and disregard the ortographical using, near the same base characters were present in all of the abc's. Sometimes forgotten characters are missing from the sets, but the existent characters are corresponding to the characters of the other sets. They seems to be different sytes of the same writing system. If we want to use the traditional native hungarian writing system, we have to sign one base character with one code.

Despite of these, Michael Everson want to sign one base character with two or more code, depend on the style. In this case, for example a text written in the Mezőkeresztesi style was not readable with the Telegdi ttf font on the internet. If we could read a text with only the same style of character set it was written, why to use the Unicode?

2 Damages made by diletant refomers.

In the XX. century, mainly in the last fifteen, twenty years, independent diletant groups made reforms, to deform the original writing system, using the well known latin abc as model. The several diletant group created a chaos and anarchy. Today, the different groups already has character sets incompatible with each others, and the original styles. The reformers never had curage to create new glyphs for the new characters, so they made the worst thing which was possible:

They selected glyphs from the old character sets, and they gave them new meanings. This way, using some old character sets, a larger character set was created. This harmful method was used widely among the deformers, but there is a very great number of possibilities to cobble a wide character set from the original sets and it caused an anarchy. The

	a	á	e	é	ë	i	í	o	ó	ö	ő	u	ú	ü	ű
<i>Petrovay János</i> 1903	4	4̄	ŷ	ŷ̄		†	†̄	◊	◊̄	K	K̄	⊗	⊗̄	z	z̄
<i>Magyar Adorján</i>	4	4	ŷ	ŷ		†	◊	ŷ	⊗	⊗	⊗	⊗	⊗	⊗	⊗
<i>Szakonyi István</i> 1963	4	4	ŷ	ŷ		†	◊	ŷ	⊗	⊗	⊗	⊗	⊗	⊗	⊗
<i>Forrai</i> 1996	4	4	ŷ	ŷ		†	†	◊	◊	K	⊗	⊗	⊗	⊗	⊗
<i>Forrai</i> 1994	4	4	ŷ	ŷ		†	†	◊	◊	K	⊗	⊗	⊗	⊗	⊗
<i>Libsch</i> 1998	4	4	ŷ	ŷ		†	†	◊	◊	z	K	⊗	⊗	⊗	⊗
<i>Nyíri Attila</i> <2008	4	4	ŷ	ŷ		†	†	◊	◊	◊	⊗	⊗	⊗	⊗	⊗
Badinyi	}	4	4	ŷ	ŷ	†	◊	ŷ	⊗	⊗	⊗	⊗	⊗	⊗	⊗
		P	P	ŷ	ŷ	†	◊	ŷ	⊗	⊗	⊗	⊗	⊗	⊗	⊗
Kiszely 2001	}	4	ŷ	†	◊	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗
		4	ŷ	†	◊	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗	⊗

Figure 2: Reform vowels.



Figure 3: Traditional 'U' glyph variants.

character set created by the deformer groups are incompatible not only with each others, but with the original sets too.

2.1 Detailed collisions.

The code collisions are mainly at he vowels.

2.1.1 Example: The character 'U'

We can see some characteristic glyph form of the letter 'U' on the figure3 above, from the wide glyph repertoire of the letter. If somebody wants to create a ttf font, the creator should choose freely a glyph among them, in the traditional use. In the reformers writing system, two glyph is selected: one of them as long 'U' and the other as short 'U' letter with

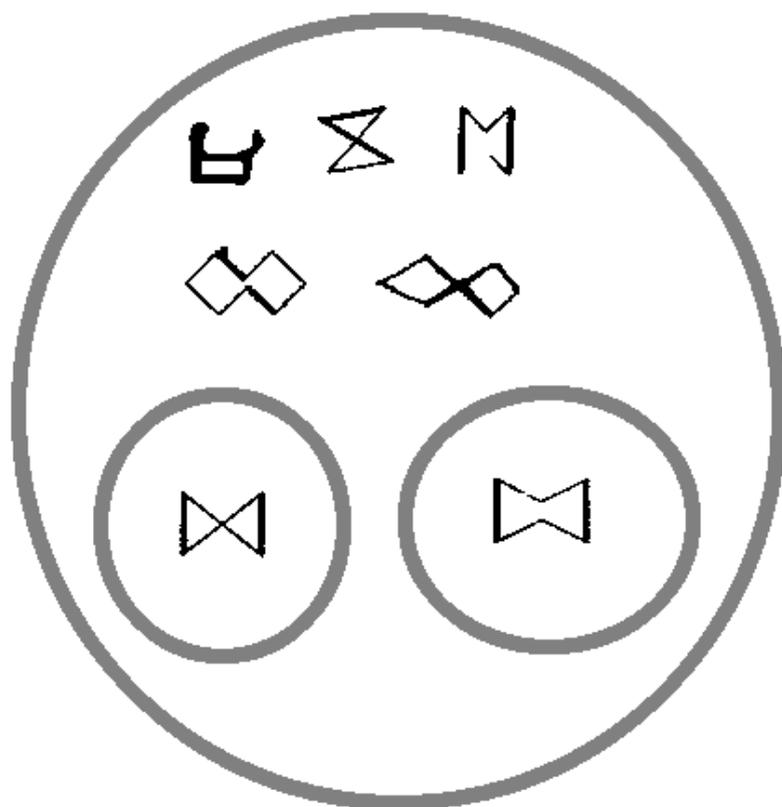


Figure 4: 'U' glyph heaps.

distinct meaning. The other glyph forms are undefined. The different diletant reformer groups created different solutions. We can see on Figure figure2, that for example the reformer Nyíri selected the opposite of the selection of the reformer Forrai.

This way, impossible to create a consistent ttf font. Michael Everson concentrating only the reform ambitions, decided to code the two glyph variants with two independent codes. Deciding from the loudness of the diletant reform groups was selected one of the solutions. There are no arguments, only violence.

On the other hand, this solution prohibits the other glyph variants of the traditional letter, a mandatory one is selected to use. In this case, impossible to use different styles!

2.1.2 Example: The character 'Ü' and 'Ö'

We have to talk about the 'Ü' and 'Ö' the same time, due to an ortographical anomaly of the so called Marsigli stick. On the Marsigli stick, this two letters was used transposed. The writer systemtically used 'Ü' instead of 'Ö' and 'Ö' instead of 'Ü'. The reformers wanted to increase the count of the characters, so they seized an opportunity.

I found too complicated to draw a diagramme to demonstate the heaps of all solutions to select long and short vowel pairs from these original heaps. There is a total anarchy. Deciding from the loudness of the diletant reformer groups, was selected one of the solutions by Mr. Everson. There are no arguments, only violence. Two glyph was selected to use as 'Ü' and two glyph variant to be used as 'Ö'. One member of the pairs is from the

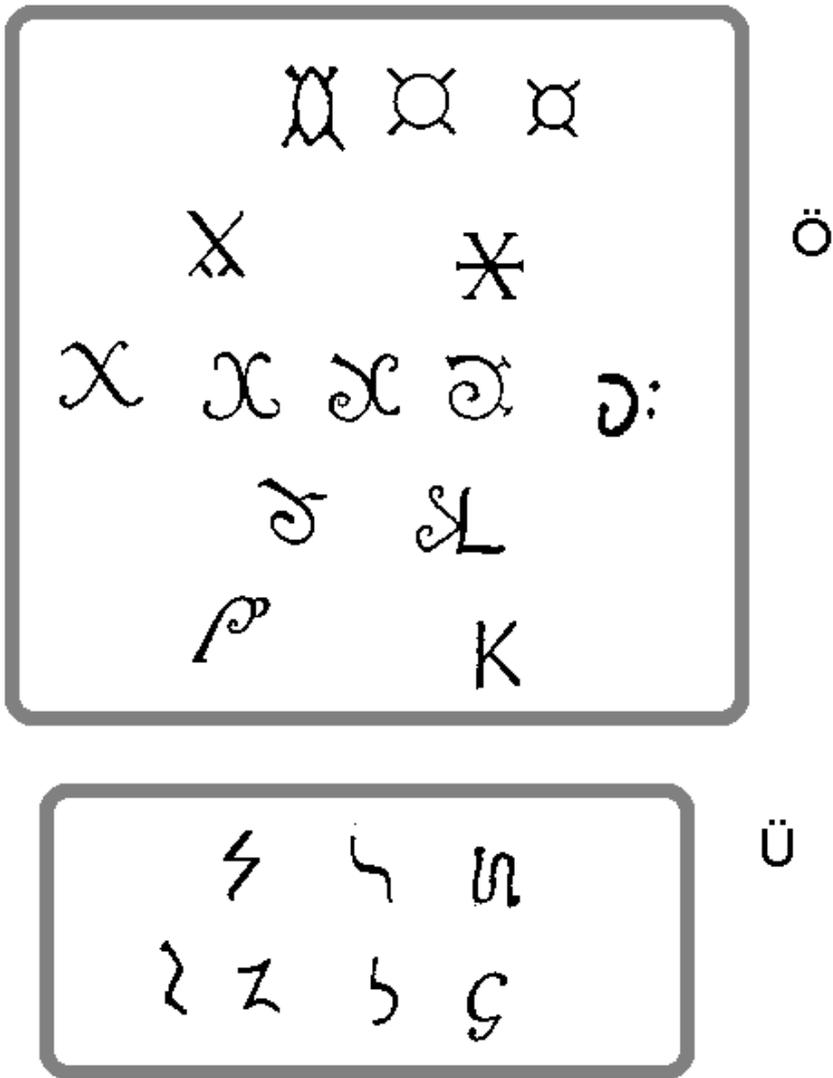


Figure 5: Traditional 'Ü' and 'Ö' glyph variants.

glyph variants of the traditional 'Ö' letter and the other from the the glyph variants of the traditional 'Ü' letter, without determined roles as long or short.

On the other hand, this solution prohibits the other glyph variants of the traditional letters, a mandatory pair is selected to use. In this case, impossible to use different styles!

2.1.3 The capital letters and small letters.

In the Hungarian Native writing system, the usage of the distinct capital letters and small letters are against the tradition. The reformers only wants to imitate the latin writing system. In this summer it was needn't, but a new wave of a brainstorm brought it. In the history, the small letters are the simplified forms of the capital letters. The natural evolution of the characters are the opposite of this brainstorm.

2.2 Other errors in the document.

Maybe my eyes are jumped a row, but at 10cb3 I saw a 'TPRUS' sign named 'ENT'. The names of the letters named as Nicolsburgian are disputables. Naming them as Bolognaisan would be correct.

3 Conclusion

Maybe Mr Everson likes to develop artifical writing systems, but the Hungarian native writing has about one thousand year history. I propose, to separate the new creations from the traditional characters. I dislike to see the damages which were made by his proposal to the traditional use, to make opportunities to brainstorms. I think, that his proposal is neither good for the reformers nor the traditional users, because it solves the incompatibility problems with violence.

I think, that the traditional characters and the newly created characters should be put to separated blocks, so an independent block served the brainstorm characters. If new glyphs were created for the newly created long vowel characters, it had had duplicated the count of the vowels, but this way, we shoud use at least three independent characters for a vowel. One for the original, one for the short and one for the long vowels. I know, that it is a very stupid thing, but the situation created by the reformers is very stupid too. In the brainstorm block the vowels had no determined glyph shape, but only a sound form. (They have really no traditional glyph form!) Therefore every reformer group would use their own glyphs, and it didn't disturbed anybody. However the traditional block got rid of this anarchy.

Michael Everson says, that my proposition is against the UNICODE rules. I don't know the UNICODE rules as well as him, but I dislike the anarchy, the dictatonic decissions and code collissions. As I know the hungarians, if a wrong standard would be accepted, everybody will change the glyph shapes to taste in the ttf fonts, and the anarchy would be greater than ever before! The users like to use diferent styles, and never will let to prohibit them.