ISO/IEC JTC1/SC2/WG2 N3213

2007-02-23

Universal Multiple Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposed block allocation Cyrillic / Bamum**
Source: **Asmus Freytag – contributing editor**
Status: **Editorial contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Related: **N3211**

In preparing document N3211 I noticed that the proposed arrangements of the new Cyrillic Extended-B and Bamum blocks would result in a free single column space. Single Columns are too small to be allocated to any scripts and therefore cause fragmentation of the remaining encoding space. In conversation with some of the other editors the proposal emerged to combine the two new proposed blocks for Cyrillic Extended B and C into a new one starting at A630, by moving the proposed Bamum block up by two columms starting at A6A0. This extends the space proposed for Cyrillic extensions by one column.

The attached shows the proposed new arrangement for Cyrillic. This new arrangement implements these improvements:
- one fewer blocks needed overall
- room for additional 16 characters in the Cyrillic Extended-B block

It can be anticipated that, like Latin and Arabic, additional characters for the Cyrillic will be needed in the near future. Reserving a very modest amount of free space in the Cyrillic Extended-C block will prevent the need to add additional blocks in the near term.

WG2 has a policy to not extend blocks across a 0x80 (or 0x00) boundary in code space, if possible. This is particularly relevant for entire scripts. However, when the alternative is to create two, or even three blocks, and when the characters are an extension of an existing script — which means that the entire repertoire already extends across an 0x80 or 0x00 boundary — this policy does not prevent a placement like the one proposed here.

	A63	A64	A65	A66	A67	A68	A69
0	Ƶ A630	Ы A640		⁙ A660		Д A680	Т A690
1	ƶ A631	ы A641		▢ A661		д A681	т A691
2	Ƶ A632	Ѣ A642	Д A652	⁂ A662		Ƨ A682	Ч A692
3	ƶ A633	ѣ A643	д A653	✩ A663		ƨ A683	ч A693
4	Ƨ A634	Ю A644	Л A654			ЗЖ A684	ҕ A694
5	ƨ A635	ю A645	л A655			зж A685	ҕ A695
6	ι A636	Ꙇ A646	М A656			Ч A686	Щ A696
7	ι A637	ꙇ A647	м A657			ч A687	щ A697
8	А A638	А A648	Ѳ A658			Д A688	
9	а A639	а A649	ѳ A659			д A689	
A	Ꙩ A63A	Ꙋ A64A	Ꙛ A65A			Ћ A68A	
B	Ꙋ A63B	Ꙋ A64B	ꙛ A65B			ћ A68B	
C	Ꙍ A63C	Ꙍ A64C	Ꙝ A65C	◌̆ A66C		Т A68C	
D	ꙍ A63D	ꙍ A64D	ꙝ A65D	◌̌ A66D		т A68D	
E	Ъ A63E	Ꙏ A64E	Ꙟ A65E	◌̆ A66E		Ц A68E	
F	ъ A63F	ꙏ A64F	ꙟ A65F	ⷮ A66F		ц A68F	