‘Kliment Std’ – a free font for Slavic Medievalists

Sebastian Kempgen, University of Bamberg
sebastian.kempgen@split.uni-bamberg.de

During the past twenty years, the author has developed many fonts of special interest to Slavicists. These inlude, but are not limited to, fonts like the Old Church Slavonic *Method* font series, Glagolitic fonts like *Kirill* (round Glagolica) and *Glagol* (square Glagolica), *RomanCyrillic*, a Cyrillic font with historic characters included, phonetic fonts like *Trubetzkoy*, fonts with flying accents like *TAccents* and several others. While these fonts are commercially available and widely used by the Academic community, the purpose of the current paper is to introduce a new font, *Kliment Std*, that is available for free to any Slavic medievalist. Another free recent release, *RomanCyrillic Std*, will be introduced in a separate article (forthcoming).

1. Download URL, license

The official web page for the *Kliment Std* font is

`http://kodeks.uni-bamberg.de/AKSL/Schrift/KlimentStd.htm`

As can be seen from the address, it is part of the material offered on the ‘Kodeks’ server run by the author, online since 1996. At the same time, the font is being mirrored on the ‘Repertorium’ web-site maintained by David Birnbaum at

`http://clover.slavic.pitt.edu/~repertorium/resources/fonts/fonts.html`

The Kodeks server is a server dedicated to help with teaching and studying Slavic medieval languages, and the cultural history of the Slavic peoples and countries, and it contains a section on the Slavic scripts, too. The font page is part of this sub-section of the server.

The licensing conditions of the font allow it to be freely used for any scholarly research and publication; however, commercial use is not covered by the license. That the font is free for use does not mean that is in the public domain; rather, the author retains all copyright to the outlines used for the characters and to the final product. The main reason for this is that the author has licensed the basic outlines himself from another, commercial font vendor. For any user of the font this has two important consequences: 1) The font is fully licensed and legal to use which means it does not violate anyone else’s copyrights – in contrast to many other free fonts available on the internet! This is the most important consequence. 2) This also means that the font may not be altered, modified, changed, renamed etc. by the end-user. For most users this is completely without importance and practical relevance. If someone would like to see additions, or has some suggestions regarding the font and ist character set, the solution to this is really simple: contact the author and I will see what can be done with implementing new features, characters, shapes etc.

2. Font format, platforms, encoding, compatibility

The font is made available as a Unicode TrueType font in the Windows font file format (i.e. as a file with name ending with the suffix .tff). Because any computer running Mac OS X can also handle .tff files, this file can be used on Windows PCs as well as Macintoshes – there is only one font file for use on both platforms. This means that there is complete compatibility and interoperability between these two platforms for any documents that use this font. The same is true for web-sites that assume the presence of a specific font in their html code (see the Repertorium site, for example).

The encoding of the font is Unicode-based. This means that whereever possible, the correct Unicode number and slot is being used for any given character. However, especially in the area of Slavic medieval studies, the Unicode standard itself is evolving, with more characters being added and introduced from time to time. Consequently, the font itself will evolve over time to incorporate new characters, accents etc. The present version of the font is specified as 1.7 and it implements Unicode version 4.1 within the scope – and limitations – of scripts and characters supported by this standard. Because the font itself will evolve to contain more and more characters, it is a good idea to check the website given above from time to time to see if a newer version might be available.

The font contains certain characters (like Ꙝ Λ Ҥ ІÔ ꙇ) which are not part of the current Unicode standard. If the Unicode standard evolves to include such a character which previously has not been part of it, and for which a stop-gap measure had been offered, this means that in future versions of the font certain characters might be moved to new slots or locations with their final, official Unicode number. To avoid any negative consequences for existing documents, the current version (1.7) will remain available alongside newer versions.

Many of the problems and areas concerning ‘missing characters’ and features in Unicode with respect to Slavic philology have been discussed recently in two articles by the author. These articles present background information on many of the features implemented in *Kliment Std*. The acknowledgement that Slavic scripts are not yet fully implemented in Unicode also lead Ralph Cleminson to collect relevant material, arguments and suggestions to formulate and finalize a submission to Unicode, Inc. for inclusion of some 40 new Cyrillic characters and diacritics. At the time of this writing, his submission has been positively reviewed by the relevant bodies but the process of adopting these new characters has not yet been finalized. As soon as the new additions have been officially sanctioned and been made publicly available from the Unicode web-site, the *Kliment Std* font will be updated accordingly.

3. Design

Kliment Std is a serifed font that most closely matches the ubiquituous *Times* font, a standard in desktop publishing and a basic font whose presence is presupposed by any PostScript-compliant printer. *Times* is a font manufactured by Linotype GmbH and

introduced to desktop typesetting and printing by Adobe Inc. on the Macintosh. Microsoft, however, licensed another very similar design, *Times New Roman*, from Monotype, for use on Windows machines. Many users will have both fonts installed and will not even be aware of the background just outlined. Although there are some subtle and some not so subtle differences between these two fonts, *Times New Roman* was considered to be so similar to *Times* that in earlier versions of Windows it was simply replaced by it when printed on a PostScript printer: *Times New Roman* was used and appeared on-screen, while *Times* was what got printed, at least on a PostScript printer. Because of the differences regarding the character set of the fonts, and because of considerable progress in printing and acceptance of TrueType fonts, this is usually no longer the case today.

Anyway, the suggestion is clear: *Kliment Std* can be and should be used along with *Times New Roman* and/or *Times* and can be mixed with those two fonts. This recommendation, is, of course, also true for any other font based on the same design, Cyrillic or Latin. *Kliment Std* uses the same outlines as the author's commercially available fonts, so there is no difference in printing quality, and the font can be used without limitation as to its quality. The only restriction that one should be aware of is that this free version of the font comes in one weight only, as a regular or upright typeface. At present, there is no free bold or italic version. An italic version may be mimicked by slanting the font - selecting 'italic' as the style in a word processor will result in a pseudo-italic oblique font. Similarly, selecting 'bold' may somewhat fatten the typeface but again this will not be a real bold font.

4. Characters, special features

In this section, we will outline some of the main features of the *Kliment Std* font. Some of them may be obvious and trivial, others may be more or less hidden and may merit special attention. The table in the Appendix presents a structured overview of what is available in *Kliment Std* in comparison to fonts like *Times* or *Times New Roman.* All in all, the font contains more than 1000 characters.

4.1. Latin characters

Kliment Std has a large set of Latin characters – several hundreds of them. In terms of 8-bit encodings or so-called 'code pages', they include the standard *Western* encodings (Windows 1252 Latin-1; MacOS Roman), the standard *Central European* (CE) encodings on both platforms (Windows 1250 Eastern Europe; MacOS Central European), and also encodings such as *Turkish* (MacOS Turkish; Windows 1254 Turkish), *Romanian* (MacOS Romanian; Windows: part of the East European character set), the *Baltic* languages (MacOS: included in the CE character set; Windows 1257 Baltic), *Croatian* (MacOS Croatian or Extended CE; Windows: included in the Eastern European encoding), Irish-Welsh (MacOS Welsh), and *Icelandic* (MacOS Icelandic; Windows: included in Latin-1).

In less technical terms, this means that all Slavic languages that use the Latin alphabet can be written with *Kliment Std* even though that's not its primary purpose. Also, all West-European and all non-Slavic East European languages (Romanian, Hungarian,

Baltic) are supported by the font. But the Latin characters available in the font are not limited to these languages: many more characters are present, for languages like Esperanto, Maltese and others. By supporting all the above-mentioned languages, the font inherently also supports the transliteration of Cyrillic and Glagolitic into the Latin alphabet. In effect, this means that many papers and articles on Slavic medieval philology do not need any other font beside *Kliment Std* and it should be a valuable tool for students and staff alike.

The support for Latin primarily concerns the following Unicode tables: *Basic Latin* and *Latin-1 Supplement* both of which are fully supported, and *Latin Extended B* which is partly supported. From this last table, the font supports all characters of relevance to Slavists, namely the *Serbocroatian digraphs* (DZ Dz dz – DŽ Dž dž – LJ Lj lj – NJ Nj nj), *Macedonian transliteration* (Ѓ ѓ), *nasal o* (Ǫ ǫ), *Štokavian accents* (lowercase only: à â è ê ì î ô ř̂ î ù û) plus some additional characters. Support for even more Latin letters is in the *Latin Extended Additional* block, where characters for the transliteration of *historical Russian*, and *Macedonian* (Ḱ ḱ Ḟ ḟ Ẏ ẏ) are available.

4.2. IPA characters

Phonetic signs (Unicode block 'IPA Extensions') are partly implemented; again, only the most important characters (for Slavists) were put into the font, symbols like

ɐ ɑ ə ɘ ɛ ɨ ɪ ɯ ʁ ʁ ʃ ʌ ʒ ʔ ʤ ʤ ʦ ʧ

Even these few characters, however, offer much more than standard fonts like *Times New Roman* or *Times* do – these fonts do not have any real support for IPA characters.

A unique feature of the font is the next Unicode table (see Fig. 1). Called '*Spacing Modifier Letters*', these are characters ('letters') with a positive width ('spacing letters'). That is, they appear onscreen by themselves, and the cursor moves as soon as the character is inserted.

02B0	02B1	02B2	02B3	02B4	02B5	02B6	02B7	02B8	02B9	02BA	02BB	02BC	02BD	02BE	02BF
ʰ	ʱ	ʲ	ʳ	ʴ	ʵ	ʶ	ʷ	ʸ	ʹ	ʺ	ʻ	ʼ	ʽ	ʾ	ʿ
02C0	02C1	02C2	02C3	02C4	02C5	02C6	02C7	02C8	02C9	02CA	02CB	02CC	02CD	02CE	02CF
ˀ	ˁ	˂	˃	˄	˅	ˆ	ˇ	ˈ	ˉ	ˊ	ˋ	ˌ	ˍ	ˎ	ˏ
02D0	02D1	02D2	02D3	02D4	02D5	02D6	02D7	02D8	02D9	02DA	02DB	02DC	02DD	02DE	02DF
ː	ˑ	˒	˓	˔	˕	˖	˗	˘	˙	˚	˛	˜	˝	˞	˟
02E0	02E1	02E2	02E3	02E4	02E5	02E6	02E7	02E8	02E9	02EA	02EB	02EC	02ED	02EE	02EF
ˠ	ˡ	ˢ	ˣ	ˤ	˥	˦	˧	˨	˩	˪	˫	ˬ	˭	ˮ	˯
02F0	02F1	02F2	02F3	02F4	02F5	02F6	02F7	02F8	02F9	02FA	02FB	02FC	02FD	02FE	02FF
˰	˱	˲	˳	˴	˵	˶	˷	˸	˹	˺	˻	˼	˽	˾	˿

Fig. 1: Spacing Modifiers

This block features special signs which modify the meaning of preceding symbol – this is why they are called 'spacing modifier letters'. Among them, we find the official transliteration symbols for the hard and the soft sign (see 02B9 and 02BA), the sign for a long vowel (02D0) and many other symbols which will be of importance for anyone concerned with Slavic phonology and phonetics.

Also fully implemented is the next block of symbols which is closely related to the preceding one. Called '*Combining Diacritical Marks*' (see Fig. 2), the important difference to the group above is that these symbols are non-spacing. This means that they can be put on top (above) or below other, normal, lowercase characters. Thus, for example, with the *dieresis below* (0324) or the *ring below* (0325) one can enter the base character first (like a b c r l u), and then the diacritic: r + ̥ = r̥ . It's in the nature of such combining diacritical marks that the fit between the base character and the diacritic cannot be perfect in each and every instance as one can see from the slight offset to the left of the ring below the 'r'. But such small imperfections aside it is still better to be able to typeset the desired combination than not to have this capability at all.

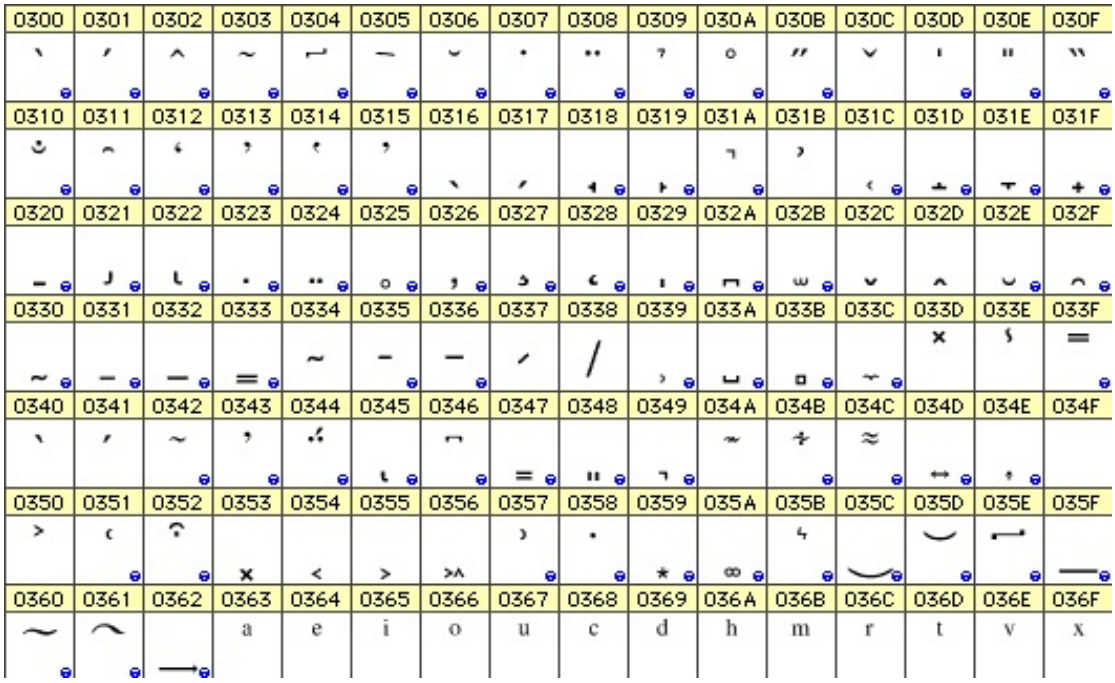

Fig. 2: Combining Diacritical Marks

To sum this up: the *Kliment Std* font fully supports *all* diacritics, spacing or non-spacing, which are defined in these sections of the Unicode standard v. 4.1.

4.3 Greek

The *Kliment Std* font also features a Greek alphabet (Unicode block 'Greek'), but it should be pointed out that this implementation is currently restricted to modern Greek and does not cover all the accents and breathing marks of Classical Greek (which is contained in another Unicode block). See Fig. 3 for an overview.

Fig. 3: Greek (incl. historical characters)

What may be worth mentioning is the fact the I have implemented all of the *historical characters* of the Greek script (see row 03D0 to 03DF) which are very useful for Slavicists, too, for example when writing about Old Church Slavonic number signs. Thus, the font features symbols like *archaic Koppa* (03D8, 03D9)*, Stigma* (03DA, 03DB)*, Digamma* (03DC, 03DD)*, Koppa* (03DE, 03DF), and *Sampi* (03E0, 03E1), all of them in lowercase and uppercase forms!

Several of these designs are a unique feature of this typeface, and have been developed by the author for the first time. If one takes a closer look at the preceding table, it can be seen that for all character pairs, true uppercase and true lowercase forms have been designed which are true to the design principles of the Greek script. For example, in a serifed printing type the uppercase digamma must not look like this: F. This is the design of a sans-serif font. The digamma is a character and not a symbol, and consequently, its shape should match the design of the rest of the alphabet. The design challenge, so to speak, is that certain historic characters are only attested in handwriting or in script form or sometimes not at all (for example, only the uppercase form – in today's terminology may be known). Starting from the uppercase form of the digamma and analyzing its construction, the author arrived at the lowercase form presented here (ϝ). In the Greek alphabet, all lowercase letters have a distinct script-like appearance, and any new lowercase design has to adhere to this basic principle. Another example is uppercase and lowercase archaic koppa (Ϙ ϙ) which must not look identical but must follow different design principle if they are to really reflect that they are characters not symbols which do not changes with fonts.

With the koppa (Ҁ ҁ), the problem was similar: again only the script lower-case form is attested, and for a serifed printing type the true uppercase form had to be developed for the first time. – Coptic characters have not been implemented in the font (see grey boxes of the last figure).

4.4. Cyrillic

Let us now turn our attention to Cyrillic. The font features a nearly full Cyrillic character set (see Fig. 4). Whereever possible, historic character form have been chosen instead of modern ones. Thus, for example, the first characters in the alphabet will look like АБВГДЕ абвгде where the 'b' has its older 'small uppercase' form, and the 'd' has its triangular form which is the older of the two and more closely related to its Greek origin. Other special character forms include 'Z' (Ꙁ ꙁ), 'Č' (Ҷ ҷ), ŠT (Щ щ), Y (Ъ ы), and YA (Ꙗ ꙗ). The OU character has been implemented in its vertical ligature form. The Omega with Veliki apostrof above is already implemented in its correct form here. Other fonts might have the older, incorrect outline present in the first versions of the Unicode standard.

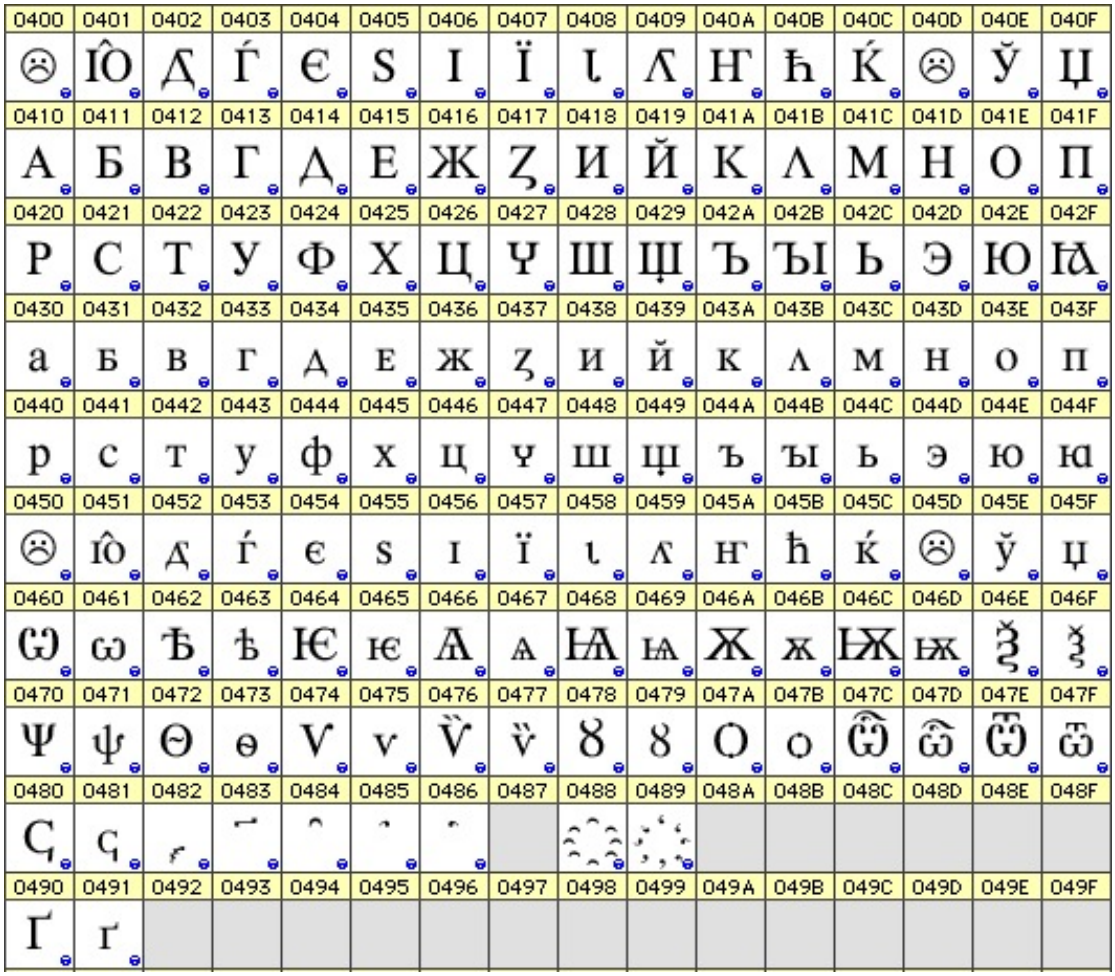

Fig. 4: Cyrillic (incl. historical characters)

The characters that have been left out in this section concern four accented modern Macedonian characters (lowercase and uppercase è ѝ). Because they are modern

characters, they have no place in a font whose aim is to help medievalists; in their places, one can see the 'smileys' in the table above. (If these characters need to be written using the *Kliment Std* font, one can always construct them from the base character with a combining accent added – that's the way they were produced here.)

The font also includes the full set of historical Slavic Cyrillic characters that are currently defined in Unicode 4.1 (see row 0460 to 047F and following). Here again, each character pair has its own carefully designed uppercase and lowercase variant. Several of the characters of Greek origin developed their own shapes in the context of the Cyrillic script, especially *ksi* and *psi*, and the uppercase *ksi* and the lowercase *psi* were given new, specially developed shapes that fit into a serifed Cyrillic font.

As can be seen from the figure, the font has the 'thousand' number sign (҂), the non-spacing, i.e. combining titlo (◌̾), the palatalization hook (◌҆) which is also a combining symbol, and the two breathing marks *dasi pneumata* and *psili pneumata* (◌̔ ◌̓) in their Cyrillic forms (not in the Glagolitic forms attested in the Kiev folia which were used for the official Unicode documents). Of course, all combining accents from the aforementioned IPA-section can also be used in conjunction with the Cyrillic characters which means the one can write accented Cyrillic with this font, too!

4.5. Glagolitic

Another unique feature of the *Kliment Std* font is its support for Glagolitic. The Glagolitic script had been adopted for version 4.1 of the Unicode standard, i.e. each character had been given its official 'slot' and unique number. *Kliment Std* does not include Glagolitic characters as such; instead, the author chose to implement automatic transliteration of Glagolitic into Cyrillic in this font. This means that where the first Glagolitic letter appears in the official Unicode table, *Kliment Std* has the matching Cyrillic character 'A' instead (see Fig. 5, slot 2C00). The next figure shows the transliteration used by the font. Again, old character shapes were used here if they exist. A special features is the introduction of a true *Cyrillic Iota* (uppercase and lowercase), the letters for 'I' and 'N' have their older shape which more closely match their respective Greek origin (horizontal arm for 'H' and right-downward pointing arm for 'N').

The 'U' has the vertical ligature here, and the Glagolitic nasals are correctly transliterated into Cyrillic with characters that are not available the Cyrillic section! A unique design is also the form of the so-called 'spidery X'. To indicate that it is a character that differs from the standard 'X', the author has designed a new character here for the first time. Its distinctive feature is the same trait that characterizes the Glagolitic character: the round 'belly' with its four legs. Implemented for a serifed font, this leads to the following shape: Ⱋ which, in my view, is the only logical way to mimic the Glagolitic character in a Cyrillic counterpart, and a much better solution than to write X_2.

The automatic transliteration of Glagolitic makes it possible, for instance, to take an already available Glagolitic text (as long as the encoding is Unicode-compliant), and to simply change the font to *Kliment Std* to have the text correctly transliterated. Of course, this also works the other way around: one can enter simple words in *Kliment*

Std using the characters from table 5 and change the font to a Glagolitic one to see what the word would look like in that script. Thus, the font is a useful educational tool and a tool for editors, librarians etc. alike.

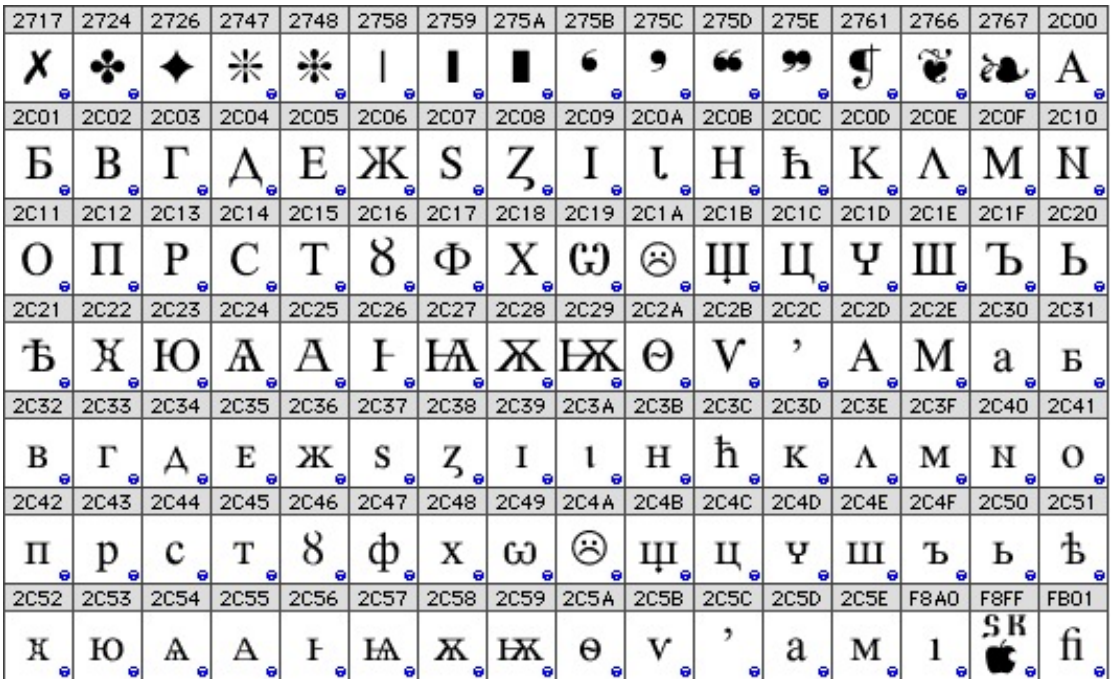

Fig. 5: Glagolitic table in transliterated implementation

4.6. Other symbols

Concluding this overview of the features of the *Kliment Std* font, I'd like to point out some additional areas that were given some sort of support in this typeface.

These areas include a full support for *Roman numerals* in uppercase and lowercase (MCDLIX), fractions (⅟ ⅓ ⅔ ⅕ ⅖ ⅙ ⅞), arrows (← ↑ → ↓ ↔ ↕), mathematical operators (∀ ∃ ∈ ∉), selected geometric shapes (■ □ ► ▷ ● ○), selected dingbats (★ ✆ ✓ ❊ ❦ ❧ † ✢ ✍) etc.

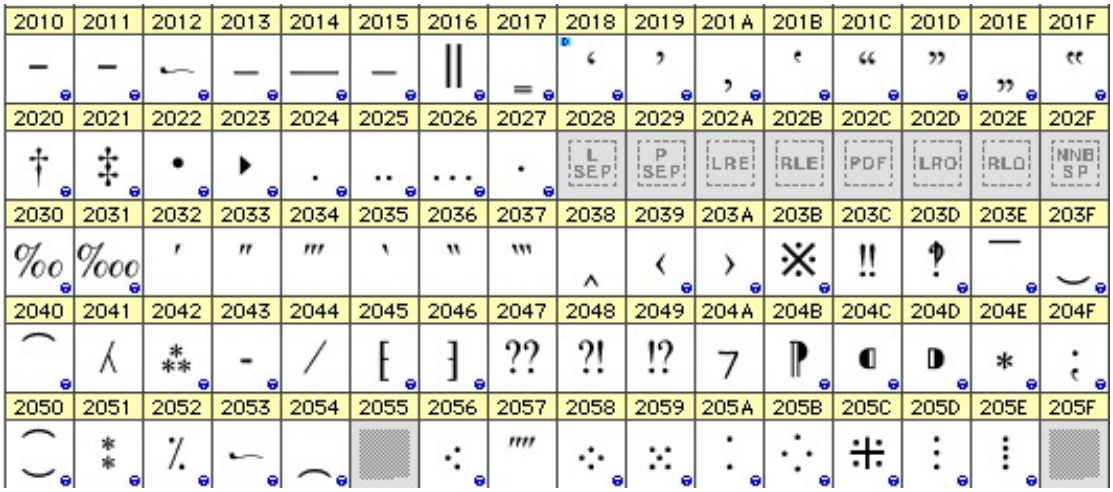

Fig. 6: General Punctuation

Punctuation has also been given extensive support (see Fig. 6), including symbols used for marking up text for correction, dashes, dots etc.

5. Conclusion

The *Kliment Std* font is meant as a tool for Slavicists working in the area of medieval studies. A free download from the author's web-server, it can be used for publishing either in print or electronically as well as on the web. It is a high-quality standard serifed font that matches the fonts usually found in scholarly publications. The font is fully legal to use and does not violate anyone else's copyrights. The font is free to use but not in the public domain. It features several solutions and designs that currently no other font can offer. As the Unicode standard evolves to include more historical Cyrillic characters, the font will also be updated to reflect these changes.

References

Kempgen, S.:

Unicode 4.1 and Slavic Philology – Problems and Perspectives (I). In: A. Miltenova (ed.), *Proceedings of the Conference "Azbuki net", Sofia, 24.–27. October 2005*, Sofia 2006, 25 pages (in print).

Unicode 4.1 and Slavic Philology – Problems and Perspectives (II). In: T. Berger, J. Raecke, T. Reuther (eds.), *Slavistische Linguistik 2004 & 2005*, München 2006, 26 pages (in print).

Unicode, Inc.: http://www.unicode.org/charts/

Orig. publ. in: Script & e-Scripta vols. 3-4, Sofia 2006, pp. 9-23.

Appendix: Features of 'Kliment Std' in comparison to other fonts

	Kliment Std v. 1.7 (Win/OS X)	Times / Helvetica v. 5.0d10e1 (OS X)	Times New Roman / Arial v. 3.05 (Win/OS X)
Basic **Latin**	✓	✓	✓
Latin-1 Supplement (= Western Europe)	✓	✓	✓
Latin Extended-A (= Eastern Europe & more)	✓	✓	✓
Latin Extended-B	some	most	some
Croatian Digraphs	✓	✓	– –
Maced. Translit. (ǵ)	✓	✓	– –
Štokavian Accents	✓	✓	– –
Nasal o	✓	✓	– –
Latin Extended Additional (256)	some	✓	ca. 1/3
Maced. Translit. (ḱ)	✓	✓	– –
Russ. Hist. Translit. (ḟ, ẏ)	✓	✓	– –
Sorbian (ḿ, ṕ)	✓	✓	– –
IPA – Phonetic	some	2/96	1/96
Spacing Modifiers	**80/80**	11/80	9/80
Translit. of Jers	✓	– –	– –
Combining Diacritics (= „Flying Accents“)	**112/112**	40/112	5/112
Greek			
Modern Greek	✓	✓	✓
Archaic Letters (Koppa, Stigma, Sampi…)	✓	– –	– –
Classical Greek	– –	✓	– –
Cyrillic			
Std. Russian & Slavic	✓	✓	✓
Macedonian Add. (è, ѝ)	(✓)	✓	– –
Hist. Add. (ѣ ѧ ѫ …)	✓	– –	– –
Ukrainian Ghe (Ґ ґ)	✓	✓	✓
Non-Slavic Cyrillic (ex GUS-Countries)	– –	ca 1/2	ca. 1/10
Glagolitic	– –	– –	– –
Transliteration into Cyrillic	✓	– –	– –

Orig. publ. in: Script & e-Scripta vols. 3-4, Sofia 2006, pp. 9-23.

	Kliment Std v. 1.7 (Win/OS X)	Times/ Helvetica v. 5.0d10e1 (OS X)	Times NR/ Arial v. 3.05 (Win/OS X)
Armenian, Georgian, Hebrew, Arabic, Ethiopian	– –	– – (supported by other fonts)	Hebrew, Arabic
General Punctuation	**70/112**	18/112	27/112
Superscripts/Subscripts (0…9)	**30/46**	– –	– –
Currency (Euro…)	1/48	3/48	6/48
Comb. Diacr. for Symbols (O)	✓	– –	– –
Number Forms	**49/64**		
Add. Fractions (2/3…)	✓	– –	6/13
Roman Numerals	✓	✓	– –
Arrows	6/112	– – (complete in Apple Symbols)	7/112 (complete in Wingdings)
Mathematical Operators (∏, ∫, ≠ …)	45/256	12/256 (complete in Apple Symbols font)	15/256 (complete in other fonts)

Orig. publ. in: Script & e-Scripta vols. 3-4, Sofia 2006, pp. 9-23.