Afghan Transitional Islamic Administration
Ministry of Communications

Afghanistan

Computer Locale Requirements for Afghanistan

Background. In December 2002 the United Nations Development Programme Country Office for Afghanistan commissioned a report on the language support required for Pashto and Dari, Afghanistan's official languages. In addition, the Project Team (Michael Everson of Everson Typography, Dublin, and Roozbeh Pournader of the FarsiWeb Project, Tehran) was successful in collecting similar information for a number of other languages of Afghanistan, namely, Southern Uzbek and Brahui. (The survey sought information on other languages, such as Hazaragi, Aimaq, Southwest Pashai, Western Balochi, and Turkmen, but conclusive responses were not available.)

Need for support. The reason this report has been commissioned is that *none* of the major computer software providers currently supports *any* of these languages – adequately or at all – and this causes serious constraints and problems for all aspects of information technology for the entire country. Language support involves inputting, display and printing, processing, and provision of a user's locale format information.

Fonts: alphabets and glyph shapes. This document presents information on the letters required to support the languages of Afghanistan, and the minimum shaping behaviour required for correct rendering. We urge software manufacturers in the strongest possible terms to update their Arabic and Persian fonts to include the letters and combinations which are missing. A summary of these forms is found at the end of the document.

Input: keyboard layouts. This document also proposes keyboard layouts suitable for inputting the languages of Afghanistan. Responses to our questionnaire did not address keyboard layout issues; however, we did receive word-of-mouth comments during the interviews we held at the University of Kabul, the Afghanistan Academy of Sciences, and the Afghanistan Assistence Coordination Authority. Most people were using old Iranian-made mechanical typewriters based on the Iranian standard ISIRI 820:1973 (*Character arrangement on keyboards of Persian typewriters*), and because they were used to them, they said that they preferred these to anything else. Because, however, that layout is glyph-based and makes use of two to four keys to form each letter, it does not suit modern computer-based entry mechanisms which use one key per letter. Accordingly, we determined that it was necessary to develop practical keyboard layouts for Afghanistan. The ISIRI 2901:1994 (*Keyboard layout for Farsi: Characters in Computer*) layout, which we used in the Farsi sample in the questionnaires, resembles the general layout of ISIRI 820 quite closely, and is modernized for current usage. Therefore, we used the general layout of ISIRI 2901 to develop the Dari layout, and then modified that to derive the Pashto and Uzbek layouts to minimize the learning curve for users as much as possible. Each of the keyboard layouts enables the user to input characters used in each of the languages supported by this report. The basic letters of the language for which the keyboard layout is designed are always on the plain and shifted keys. In some software environments the AltGr keys may not be supported.

Computer Locale Requirements for Afghanistan
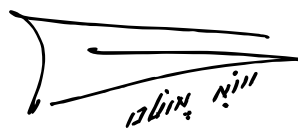
Locales: ICU data. Additional data we collected has been formatted in locale specifications via the ICU (International Components for Unicode) website hosted at "**http://oss.software.ibm.com/icu/**". In this way, the Afghan preferred formats for language, date format (including month and day names and calendar), time format, number format, and currency format will be available to vendors who can provide software which will be of use to the people of Afghanistan, at governmental, business, academic, and local level. Whenever ICU locale data format is adequate for describing a part of the locale information, the data in the ICU registry should be consulted for normative and accurate data. The specifications shown in this report are only informative examples in comparison with the ICU data files we have provided. However, in areas such as the Afghan calendar, which ICU locales cannot specify, the text in this report should be considered normative. (Note that the ICU files we have prepared are not standalone. They must be used in combination with other locale data, most importantly the "fa.txt" and "root.txt" files from ICU itself.)

Ordering. Sorting order specification is also provided with the ICU locales. This deserves some additional discussion here. In our analysis of the Pashto, Dari, and Uzbek alphabets provided to us, together with our analysis of authoritative Pashto-English and Pashto-Pashto dictionaries, as well as comparison of a Dari-English dictionary with a number of bilingual and monolingual Persian dictionaries, we discovered that, with the exception of the treatment of *hamza*, the alphabets were all mutually compatible with respect to ordering. Because the ordering of "singular" *hamza* (ء) is of little consequence – it is rare and typically comes toward the end of a word – we determined that placing it between *alef* and *beh* would have two advantages: first, it would allow all of the languages of Afghanistan to avail of a *single unified sort order*, and second, it would enable the closely-related Persian languages, Dari and Farsi, to be sorted in the same way. It should be mentioned, however, that the order shown in this report is only the general ordering for letters, and does not show the ordering of less significant elements like the ordering of vowel marks or the handling of punctuation or Unicode control characters. The complete and exact ordering is the one specified in the ICU locale data files.

We commend this report to the software vendors of the world, and trust that its specifications will be implemented as a matter of urgency, so that the people Afghanistan can avail of the essential tools which many of us take for granted.

Michael Everson
Everson Typography, Dublin
www.evertype.com

Roozbeh Pournader
FarsiWeb Project, Tehran
www.farsiweb.info

The latest version of this report, updates, and other resources (such as keyboard layout specifications in data files conformant to ISO/IEC 9995, *Information Technology – Keyboard layouts for text and office systems*, and ICU data files based on this report) are available at "**http://www.evertype.com/standards/af/**".

© Copyright 2003 United Nations Development Programme, Michael Everson, and Roozbeh Pournader. Everyone is permitted to copy and distribute verbatim copies of this document, but changing it is not allowed. The authors welcome feedback from people who download and use the report and data files.
Latest update: 2003-07-29

Computer Locale Requirements for Afghanistan
The Pashto Alphabet

The table below lists the letters used to write Pashto, including loanwords from Arabic. Our survey indicated a number of different sorting orders for Pashto. *All respondants* said that the alphabetic order was (reading from left to right) ج > څ > چ > ځ > ح > خ. This is the order agreed on at a 1991 meeting of Pashto experts in Peshawar. It should be noted that the most authoritative dictionary, the *Paṣtó-Paṣtó Descriptive Dictionary*, 4 volumes (Department of Linguistics, Institute of Languages and Literature, Academy of Sciences of Afghanistan, 1979–1987) uses the order ج > چ > څ > ځ > ح > خ, as does Rahimi and Rohi's *Pashto-English Dictionary* (1979). Materials for schoolchildren collected in Kabul, on the other hand, gave this same sequence as ج > چ > ح > خ > ځ > څ. Nevertheless, we suggest that the order agreed at the 1991 Peshawar meeting is the best to follow.

In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0622	ā	آ	ـآ	—	—	•	•	•	•	
0627	alif	ا	ـا	—	—	•	•	•	•	•
0623	alif-hamza	أ	ـأ	—	—	•	•	•	•	•
0671	alif-wasla	ٱ	ـٱ	—	—	•				
0625	al.-h.-below	إ	ـإ	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	be	ب	ـب	ـبـ	بـ	•	•	•	•	•
067E	pe	پ	ـپ	ـپـ	پـ		•	•	•	•
062A	te	ت	ـت	ـتـ	تـ	•	•	•	•	•
067C	ṭe	ټ	ـټ	ـټـ	ټـ			•		
0679	Urdu ṭe	ٹ	ـٹ	ـٹـ	ٹـ					•
062B	se	ث	ـث	ـثـ	ثـ	•	•	•	•	•
062C	jim	ج	ـج	ـجـ	جـ	•	•	•	•	•
0681	ǰe	ځ	ـځ	ـځـ	ځـ			•		
0686	če	چ	ـچ	ـچـ	چـ		•	•	•	•
0685	ce	څ	ـڅ	ـڅـ	څـ			•		
062D	he	ح	ـح	ـحـ	حـ	•	•	•	•	•

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
062E	xe	خ	ـخ	ـخـ	خـ	•	•	•	•	•
062F	dāl	د	ـد	—	—	•	•	•	•	•
0689	ḍāl	ڊ	ـڊ	—	—			•		
0688	Urdu ḍāl	ڈ	ـڈ	—	—					•
0630	zāl	ذ	ـذ	—	—	•	•	•	•	•
0631	re	ر	ـر	—	—	•	•	•	•	•
0693	ṛe	ړ	ـړ	—	—			•		
0691	Urdu ṛe	ڑ	ـڑ	—	—					•
0632	ze	ز	ـز	—	—	•	•	•	•	•
0698	že	ژ	ـژ	—	—		•	•	•	•
0696	ẓe	ږ	ـږ	—	—			•		
0633	sin	س	ـس	ـسـ	سـ	•	•	•	•	•
0634	šin	ش	ـش	ـشـ	شـ	•	•	•	•	•
069A	ṣin	�	ـښ	ـښـ	ښـ			•		
0635	swāt	ص	ـص	ـصـ	صـ	•	•	•	•	•
0636	zād	ض	ـض	ـضـ	ضـ	•	•	•	•	•
0637	twe	ط	ـط	ـطـ	طـ	•	•	•	•	•
0638	zwe	ظ	ـظ	ـظـ	ظـ	•	•	•	•	•
0639	ʿayn	ع	ـع	ـعـ	عـ	•	•	•	•	•
063A	ğayn	غ	ـغ	ـغـ	غـ	•	•	•	•	•
0641	fe	ف	ـف	ـفـ	فـ	•	•	•	•	•
0642	qāf	ق	ـق	ـقـ	قـ	•	•	•	•	•
06A9	kāf	ک	ـک	ـکـ	کـ		•	•	•	•
0643	Arabic kāf	ك	ـك	ـكـ	كـ	•				
06AB	gāf	ڪ	ـڪ	ـڪـ	ڪـ			•		
06AF	Persian gāf	گ	ـگ	ـگـ	گـ		•		•	•
0644	lām	ل	ـل	ـلـ	لـ	•	•	•	•	•
06B7	Brahui lhām	ڷ	ـڷ	ـڷـ	ڷـ					•
0645	mim	م	ـم	ـمـ	مـ	•	•	•	•	•
0646	nun	ن	ـن	ـنـ	نـ	•	•	•	•	•

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
06BA	nun ghunna	ں	ﮟ	—	—					•
06BC	ṇun	ﮞ	ﮞ	ﮧ	ﮨ			•		
0648	wāw	و	ﻮ	—	—	•	•	•	•	•
0624	wāw hamza	ؤ	ﺆ	—	—	•	•	•	•	•
06C7	Uzbek u	ۇ	ﯗ	—	—				•	
06C9	Uzbek o	ۉ	ﯙ	—	—				•	
0647	he	ه	ﻪ	ﻬ	ﻫ	•	•	•	•	•
0629	gərda te	ة	ﺔ	—	—	•	•	•	•	•
06CC	ye	ى	ﻰ	ﻴ	ﻳ		•	•	•	•
0649	alef maksura	—	—	ﻰ	ﯨ	•				
06D2	Urdu ye	ے	ﮯ	—	—					•
064A	saxta ye	ي	ﻲ	—	—	•		•	•	
06D0	pasta ye	ې	ﯤ	ﯧ	ﯦ			•	•	
06CD	ṣəĭna ye	ۍ	ﮯ	—	—			•		
0626	feʿli ye	ئ	ﺉ	ﺌ	ﺋ	•	•	•	•	

Note. Since the shapes of the ﻳ initial and ﻴ medial forms of the Pashto letters ى *ye* (U+06CC) and ي *saxta ye* (U+064A) are exactly the same, to avoid encoding ambiguities in Pashto data – and its known implications in security-related issues including possible Pashto domain names – we recommend that the Unicode character for *saxta ye*, namely ي U+064A, *never* be used in initial and medial forms in Pashto data. Where input data is explicitly known to be in Pashto, applications *should* automatically replace such usage in the input data (for example, in keyboard input) with the normal *ye* character, namely ى U+06CC. Applications *may not* automatically change U+064A to U+06CC in Pashto input if it is used in final or isolated forms.

Computer Locale Requirements for Afghanistan
The Dari Alphabet

The table below lists the letters used to write Dari, including loanwords from Arabic. In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0622	ā	آ	ﺎ	—	—	•	•	•	•	
0627	alef	ا	ﺎ	—	—	•	•	•	•	•
0623	alef-hamza	أ	ﺄ	—	—	•	•	•	•	•
0671	alif-wasla	ٱ	ﭑ	—	—	•				
0625	al.-h.-below	إ	ﺈ	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	be	ب	ﺐ	ﺒ	ﺑ	•	•	•	•	•
067E	pe	پ	ﭗ	ﭙ	ﭘ		•	•	•	•
062A	te	ت	ﺖ	ﺘ	ﺗ	•	•	•	•	•
067C	Pashto ṭe	ټ	ﭧ	ﭩ	ﭨ			•		
0679	Urdu ṭe	ٹ	ﭓ	ﭕ	ﭔ					•
062B	se	ث	ﺚ	ﺜ	ﺛ	•	•	•	•	•
062C	jim	ج	ﺞ	ﺠ	ﺟ	•	•	•	•	•
0681	Pashto je	ځ	ﭻ	ﭽ	ﭼ			•		
0686	če	چ	ﭻ	ﭽ	ﭼ		•	•	•	•
0685	Pashto ce	څ	ﭾ	ﮀ	ﭿ			•		
062D	he	ح	ﺢ	ﺤ	ﺣ	•	•	•	•	•
062E	xe	خ	ﺦ	ﺨ	ﺧ	•	•	•	•	•
062F	dāl	د	ﺪ	—	—	•	•	•	•	•
0689	Pashto ḍāl	ډ	ﮉ	—	—			•		
0688	Urdu ḍāl	ڈ	ﮈ	—	—					•
0630	zāl	ذ	ﺬ	—	—	•	•	•	•	•
0631	re	ر	ﺮ	—	—	•	•	•	•	•

Computer Locale Requirements for Afghanistan

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0693	Pashto ṛe	پڑ	پڑ	—	—			•		
0691	Urdu ṛe	ڑ	ڑ	—	—					•
0632	ze	ز	ز	—	—	•	•	•	•	•
0698	že	ژ	ژ	—	—		•	•	•	•
0696	Pashto ẓe	ږ	ږ	—	—			•		
0633	sin	س	س	ـسـ	سـ	•	•	•	•	•
0634	šin	ش	ش	ـشـ	شـ	•	•	•	•	•
069A	Pashto ṣin	ښ	ښ	ـښـ	ښـ			•		
0635	sād	ص	ص	ـصـ	صـ	•	•	•	•	•
0636	zād	ض	ض	ـضـ	ضـ	•	•	•	•	•
0637	tā	ط	ط	ـط	ط	•	•	•	•	•
0638	zā	ظ	ظ	ـظ	ظ	•	•	•	•	•
0639	ayn	ع	ع	ـعـ	عـ	•	•	•	•	•
063A	qayn	غ	غ	ـغـ	غـ	•	•	•	•	•
0641	fe	ف	ف	ـفـ	فـ	•	•	•	•	•
0642	qāf	ق	ق	ـقـ	قـ	•	•	•	•	•
06A9	kāf	ک	ک	ـکـ	کـ		•	•	•	•
0643	Arabic kāf	ك	ك	ـكـ	كـ	•				
06AB	Pashto gāf	ګ	ګ	ـګـ	ګـ			•		
06AF	gāf	گ	گ	ـگـ	گـ		•		•	•
0644	lām	ل	ل	ـلـ	لـ	•	•	•	•	•
06B7	Brahui lhām	ڷ	ڷ	ـڷـ	ڷـ					•
0645	mīm	م	م	ـمـ	مـ	•	•	•	•	•
0646	nūn	ن	ن	ـنـ	نـ	•	•	•	•	•
06BA	nūn ghunna	ں	ں	—	—					•
06BC	Pashto ṇūn	ڼ	ڼ	ـڼـ	ڼـ			•		
0648	vāv	و	و	—	—	•	•	•	•	•
0624	vāv hamza	ؤ	ؤ	—	—	•	•	•	•	•
06C7	Uzbek u	ۇ	ۇ	—	—				•	
06C9	Uzbek o	ۉ	ۉ	—	—				•	

Computer Locale Requirements for Afghanistan

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0647	he	ه	ـه	ـهـ	هـ	•	•	•	•	•
0629	te-ye gerd	ة	ـة	—	—	•	•	•	•	•
06CC	ye	ى	ـى	ـيـ	يـ		•	•	•	•
0649	alef-e maqsura	—	—	ـلـ	ـل	•				
06D2	Urdu ye	ے	ـے	—	—					•
064A	saxta ye	ي	ـي	—	—	•		•	•	
06D0	pasta ye	ې	ـې	ـېـ	ـېـ			•	•	
06CD	ṣəǰina ye	ۍ	ـۍ	—	—			•		
0626	ye hamza	ئ	ـئ	ـئـ	ئـ	•	•	•	•	

Computer Locale Requirements for Afghanistan
The Southern Uzbek Alphabet

The table below lists the letters used to write Southern Uzbek, including loanwords from Arabic. This is a new official orthography the development of which was inspired by the UNDP project. In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0622	alef-madha	آ	ـآ	—	—	•	•	•	•	
0627	alef	ا	ـا	—	—	•	•	•	•	•
0623	alef-hamza	أ	ـأ	—	—	•	•	•	•	•
0671	alif-wasla	ٱ	ـٱ	—	—	•				
0625	al.-h.-below	إ	ـإ	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	beh	ب	ـب	ـبـ	بـ	•	•	•	•	•
067E	peh	پ	ـپ	ـپـ	پـ		•	•	•	•
062A	teh	ت	ـت	ـتـ	تـ	•	•	•	•	•
067C	Pashto ṭeh	ټ	ـټ	ـټـ	ټـ			•		
0679	Urdu ṭeh	ٹ	ـٹ	ـٹـ	ٹـ					•
062B	seh	ث	ـث	ـثـ	ثـ	•	•	•	•	•
062C	jeem	ج	ـج	ـجـ	جـ	•	•	•	•	•
0681	Pashto ǰeh	ځ	ـځ	ـځـ	ځـ			•		
0686	cheh	چ	ـچ	ـچـ	چـ		•	•	•	•
0685	Pashto ceh	څ	ـڅ	ـڅـ	څـ			•		
062D	heh	ح	ـح	ـحـ	حـ	•	•	•	•	•
062E	kheh	خ	ـخ	ـخـ	خـ	•	•	•	•	•
062F	dal	د	ـد	—	—	•	•	•	•	•
0689	Pashto ḍal	ḍ	ـḍ	—	—			•		
0688	Urdu ḍal	ڈ	ـڈ	—	—					•
0630	zal	ذ	ـذ	—	—	•	•	•	•	•
0631	reh	ر	ـر	—	—	•	•	•	•	•

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0693	Pashto ṛeh	پړ	ـپړ	—	—			•		
0691	Urdu ṛeh	ڑ	ـڑ	—	—					•
0632	zeh	ز	ـز	—	—	•	•	•	•	•
0698	zheh	ژ	ـژ	—	—		•	•	•	•
0696	Pashto ẓeh	ژ	ـژ	—	—			•		
0633	sin	س	ـس	ـسـ	سـ	•	•	•	•	•
0634	shin	ش	ـش	ـشـ	شـ	•	•	•	•	•
069A	Pashto ṣin	ښ	ـښ	ـښـ	ښـ			•		
0635	swad	ص	ـص	ـصـ	صـ	•	•	•	•	•
0636	zwad	ض	ـض	ـضـ	ضـ	•	•	•	•	•
0637	tuy	ط	ـط	ـطـ	طـ	•	•	•	•	•
0638	zuy	ظ	ـظ	ـظـ	ظـ	•	•	•	•	•
0639	'ain	ع	ـع	ـعـ	عـ	•	•	•	•	•
063A	ghain	غ	ـغ	ـغـ	غـ	•	•	•	•	•
0641	feh	ف	ـف	ـفـ	فـ	•	•	•	•	•
0642	qaf	ق	ـق	ـقـ	قـ	•	•	•	•	•
06A9	kaf	ک	ـک	ـکـ	کـ		•	•	•	•
0643	Arabic kaf	ك	ـك	ـكـ	كـ	•				
06AB	Pashto gaf	ګ	ـګ	ـګـ	ګـ			•		
06AF	gaf	گ	ـگ	ـگـ	گـ		•		•	•
0644	lam	ل	ـل	ـلـ	لـ	•	•	•	•	•
06B7	Brahui lham	ڷ	ـڷ	ـڷـ	ڷـ					•
0645	meem	م	ـم	ـمـ	مـ	•	•	•	•	•
0646	nuun	ن	ـن	ـنـ	نـ	•	•	•	•	•
06BA	nuun ghunna	ں	ـں	—	—					•
06BC	Pashto ṇuun	ڼ	ـڼ	ـڼـ	ڼـ			•		
0648	waw	و	ـو	—	—	•	•	•	•	•
0624	waw hamza	ؤ	ـؤ	—	—	•	•	•	•	•
06C7	u	ۇ	ـۇ	—	—				•	
06C9	o'	ۉ	ـۉ	—	—				•	

Computer Locale Requirements for Afghanistan

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0647	heh	ه	ـه	ـهـ	هـ	•	•	•	•	•
0629	teh marbuta	ة	ـة	—	—	•	•	•	•	•
06CC	yih	ى	ى	ـيـ	يـ		•	•	•	•
0649	alef maksura	—	—	ـىـ	ىـ	•				
06D2	Urdu yih	ے	ے	—	—					•
064A	yeeh	ي	ي	—	—	•		•	•	
06D0	yeh	ې	ې	ـېـ	ېـ			•	•	
06CD	şəjina yih	ۍ	ۍ	—	—			•		
0626	yih hamza	ئ	ـئ	ـئـ	ئـ	•	•	•	•	

Note. To avoid encoding ambiguities in Uzbek data, and its known implications in security-related issues including possible Uzbek Internet domain names, we recommend that the Uzbek letter *u always* be encoded as ۇ U+06C7 and *never* as a sequence of *waw* and a *damma*. Where input data is explicitly known to be in Uzbek, applications *should* automatically replace a *waw* followed by a *damma* in the input data (for example, in keyboard input) with the U+06C7 character.

Computer Locale Requirements for Afghanistan

The Brahui Alphabet

The table below lists the letters used to write Brahui, including loanwords from Arabic. Brahui is not widely written either in Pakistan, where most Brahui speakers live, or in Afghanistan. An expert respondant informed us that in Pakistan, Brahui speakers use Urdu orthography with the addition of one letter, ‍ٱ *lhām*. If Brahui is written in Afghanistan, it may prefer the Pashto letters such as ټ *ṭe* to the corresponding Urdu ٹ *ṭe*. We have no evidence of this, but have given the Brahui alphabet here in order to ensure the support of its own unique letter.

In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0622	ā	آ	ـآ	—	—	•	•	•	•	•
0627	alif	ا	ـا	—	—	•	•	•	•	•
0623	alif-hamza	أ	ـأ	—	—	•	•	•	•	•
0671	alif-wasla	ٱ	ـٱ	—	—	•				
0625	al.-h.-below	إ	ـإ	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	be	ب	ـب	ـبـ	بـ	•	•	•	•	•
067E	pe	پ	ـپ	ـپـ	پـ		•	•	•	•
062A	te	ت	ـت	ـتـ	تـ	•	•	•	•	•
067C	ṭe	ټ	ـټ	ـټـ	ټـ			•		
0679	Urdu ṭe	ٹ	ـٹ	ـٹـ	ٹـ					•
062B	se	ث	ـث	ـثـ	ثـ	•	•	•	•	•
062C	jim	ج	ـج	ـجـ	جـ	•	•	•	•	•
0681	ǰe	ځ	ـځ	ـځـ	ځـ			•		
0686	če	چ	ـچ	ـچـ	چـ		•	•	•	•
0685	ce	څ	ـڅ	ـڅـ	څـ			•		
062D	he	ح	ـح	ـحـ	حـ	•	•	•	•	•
062E	xe	خ	ـخ	ـخـ	خـ	•	•	•	•	•
062F	dāl	د	ـد	—	—	•	•	•	•	•

Computer Locale Requirements for Afghanistan

Unicode	Name	*isolated*	*final*	*medial*	*initial*	Arabic	Dari	Pashto	Uzbek	Brahui
0689	ḍāl	ﮂ	ﮃ	—	—			•		
0688	Urdu ḍāl	ڈ	ـڈ	—	—					•
0630	zāl	ذ	ـذ	—	—	•	•	•	•	•
0631	re	ر	ـر	—	—	•	•	•	•	•
0693	ṛe	ړ	ـړ	—	—			•		
0691	Urdu ṛe	ڑ	ـڑ	—	—					•
0632	ze	ز	ـز	—	—	•	•	•	•	•
0698	že	ژ	ـژ	—	—		•	•	•	•
0696	ẓe	ږ	ـږ	—	—			•		
0633	sin	س	ـس	ـسـ	ـسـ	•	•	•	•	•
0634	šin	ش	ـش	ـشـ	شـ	•	•	•	•	•
069A	ṣin	ښ	ـښ	ـښـ	ښـ			•		
0635	swāt	ص	ـص	ـصـ	صـ	•	•	•	•	•
0636	zād	ض	ـض	ـضـ	ضـ	•	•	•	•	•
0637	twe	ط	ـط	ـطـ	طـ	•	•	•	•	•
0638	zwe	ظ	ـظ	ـظـ	ظـ	•	•	•	•	•
0639	ʿayn	ع	ـع	ـعـ	عـ	•	•	•	•	•
063A	ğayn	غ	ـغ	ـغـ	غـ	•	•	•	•	•
0641	fe	ف	ـف	ـفـ	فـ	•	•	•	•	•
0642	qāf	ق	ـق	ـقـ	قـ	•	•	•	•	•
06A9	kāf	ک	ـک	ـکـ	کـ		•	•	•	•
0643	Arabic kāf	ك	ـك	ـكـ	كـ	•				
06AB	Pashto gāf	ګ	ـګ	ـګـ	ګـ			•		
06AF	gāf	گ	ـگ	ـگـ	گـ		•		•	•
0644	lām	ل	ـل	ـلـ	لـ	•	•	•	•	•
06B7	lhām	ڷ	ـڷ	ـڷـ	ڷـ					•
0645	mim	م	ـم	ـمـ	مـ	•	•	•	•	•
0646	nun	ن	ـن	ـنـ	نـ	•	•	•	•	•
06BA	nun ghunna	ں	ـں	—	—					•
06BC	ṇun	ڼ	ـڼ	ـڼـ	ڼـ			•		

Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0648	wāw	و	ـو	—	—	•	•	•	•	•
0624	wāw hamza	ؤ	ـؤ	—	—	•	•	•	•	•
06C7	Uzbek u	ۇ	ـۇ	—	—				•	
06C9	Uzbek o	ۉ	ـۉ	—	—				•	
0647	he	ه	ـه	ـهـ	هـ	•	•	•	•	•
0629	te marbuta	ة	ـة	—	—	•	•	•	•	•
06CC	ye	ی	ـی	ـیـ	یـ		•	•	•	•
0649	alef maksura	—	—	ـىـ	ىـ	•				
06D2	Urdu ye	ے	ـے	—	—					•
064A	saxta ye	ي	ـي	—	—	•		•	•	
06D0	pasta ye	ې	ـې	ـېـ	ېـ			•	•	
06CD	ṣəǰina ye	ۍ	ـۍ	—	—			•		
0626	ye hamza	ئ	ـئ	ـئـ	ئـ	•	•	•	•	

Computer Locale Requirements for Afghanistan

Glyph-shape Requirements for Afghanistan

The table below lists the glyphs not usually found in most Arabic fonts, which it is essential to support in order to enable the people of Afghanistan to write and process their languages.

Unicode	Name	isolated	final	medial	initial	Dari	Pashto	Uzbek	Brahui
067E	pe	پ	‍پ	‍پ‍	پ‍	•	•	•	•
067C	Pashto ṭe	ټ	‍ټ	‍ټ‍	ټ‍		•		
0679	Urdu ṭe	ٹ	‍ٹ	‍ٹ‍	ٹ‍				•
0681	ǰe	ځ	‍ځ	‍ځ‍	ځ‍		•		
0686	če	چ	‍چ	‍چ‍	چ‍	•	•	•	•
0685	ce	څ	‍څ	‍څ‍	څ‍		•		
0689	ḍāl	ډ	‍ډ	—	—		•		
0688	Urdu ḍāl	ڈ	‍ڈ	—	—				•
0693	ṛe	ړ	‍ړ	—	—		•		
0691	Urdu ṛe	ڑ	‍ڑ	—	—				•
0698	že	ژ	‍ژ	—	—	•	•	•	•
0696	ẓe	ږ	‍ږ	—	—		•		
069A	ṣin	ښ	‍ښ	‍ښ‍	ښ‍		•		
06A9	Persian kāf	ک	‍ک	‍ک‍	ک‍	•	•	•	•
06AB	Pashto gāf	ګ	‍ګ	‍ګ‍	ګ‍		•		
06AF	gāf	گ	‍گ	‍گ‍	گ‍	•		•	•
06B7	lhām	ڷ	‍ڷ	‍ڷ‍	ڷ‍				•
06BA	nun ghunna	ں	‍ں	—	—				•
06BC	ṇun	ڼ	‍ڼ	‍ڼ‍	ڼ‍		•		
06C7	Uzbek u	ۇ	‍ۇ	—	—			•	
06C9	Uzbek o	ۉ	‍ۉ	—	—			•	
06CC	ye	ی	‍ی	‍ی‍	ی‍	•	•	•	•
06D2	Urdu ye	ے	‍ے	—	—				•
06D0	pasta ye	ۍ	‍ۍ	‍ۍ‍	ۍ‍		•	•	
06CD	ṣǝǰina ye	ڍ	‍ڍ	—	—		•		

Computer Locale Requirements for Afghanistan

Almost all of the languages used in Afghanistan require the use of diacritical marks which follow a base letter, to make pronunciation or grammar usage clear. These marks are known by different names in the different languages of Afghanistan, but are listed here with their Unicode character names. Font and application developers should ensure that they support these marks in their products. The *zwarakay* at the bottom of the list is used in some Pashto educational materials, and as a result of this survey is now being proposed for addition to the Unicode Standard. Until such time as it has been formally encoded, we recommend the Private Use Area code position U+E659 for the *zwarakay.*
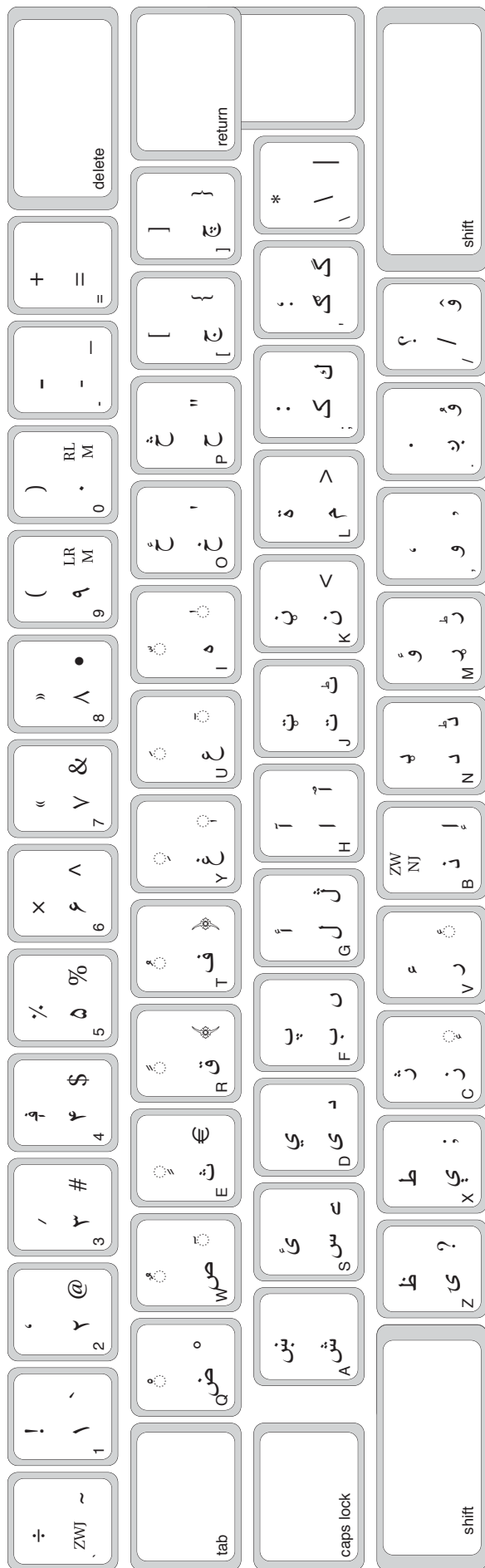
Unicode	Name		Unicode	Name	
064B	fathatan		0651	shadda	
064C	dammatan		0652	sukun	
064D	kasratan		0653	maddah above	
064E	fatha		0654	hamza above	
064F	damma		0670	superscript alef	
0650	kasra		(E659)	zwarakay	

One additional character appears to be in wide use by local banks and accountants in Afghanistan, namely the *afghani sign,* ﬗ. This character is also being proposed for addition to the Unicode Standard. Until such time as it has been formally encoded, we recommend the Private Use Area code position U+E0B4 for the *afghani sign.* It should be noted that glyph variants for this character exist, for example ﬗ and ﬗ, but in our discussions several experts indicated that the first one shown above is the most original and may therefore be more appropriate. Research on the best glyph shape is still ongoing. As of this writing, neither the Ministry of Finance nor Da Afghanistan Bank had formally adopted this symbol for use, although it is in use in both the public and private sectors.

It should also be mentioned that certain punctuation marks and control characters are used in Afghan computing. We have listed all of these characters on the collectively-equivalent keyboard layouts. Font developers are expected to include *all* of these additional characters in fonts which are intended to support Afghan languages.

Computer Locale Requirements for Afghanistan

Pashto keyboard layout for Afghanistan

by Michael Everson, Roozbeh Pournader, Said Marjan Zazai, & Mohammad Tamim Noori, 2003-07-05

Guide to the key assignments, reading top-down and left-right

Unshifted:

ZWJ, one, two, three, four, five, six, seven, eight, nine, zero, hyphen, equals
zād, swāt, se, qāf, fe, ğayn, ʿayn, he, xe, he, jim, če
šin, sin, ye, be, lām, alif, te, nun, mim, kāf, gāf, backslash
ṣəјina-ye, feʾli-ye, ze, re, zāl, dāl, ṛe, vāv, že, slash

Shifted:

division, exclamation, thousands, decimals, afghanẹy, percent, multiplication, guillemet, guillemet, parenthesis, parenthesis, kashida, plus
sukun, dammatan, kasratan, fathatan, damma, kasra, fatha, shadda, јe, ce, square-bracket, square-bracket
ṣin, pasta-ye, saxta-ye, pe, alef-hamza, ā, ṭe, ṇūn, gərda-te, colon, semicolon, asterisk
zwe, twe, že, hamza, ZWNJ, ḍāl, wāw-hamza, comma, full-stop, question

Alt-Gr:

tilde, grave, at, number, dollar, percent, circumflex, ampersand, bullet, LRM, RLM, underscore, —
degree, maddah, euro, ornament-parenthesis, ornament-parenthesis, alif-below, zwarakay, alif-above, apostrophe, quote, curly-bracket, curly-bracket
—, urdu-ye, alif-maksura, nun-ghunna, brahui-lhām, alif-wasla, urdu-ṭe, less-than, greater-than, arabic-kāf, persian-gāf, vertical-bar
question, semicolon, hamza-below, hamza-above, alif-hamza-below, urdu-ḍāl, urdu-ṛe, comma, uzbek-u, uzbek-o

shift	altgr	
	shift	
plain	altgr	

Unshifted spacebar → SPACE

Shifted spacebar → ZERO WIDTH NON-JOINER

Alt-Gr spacebar → NO-BREAK SPACE
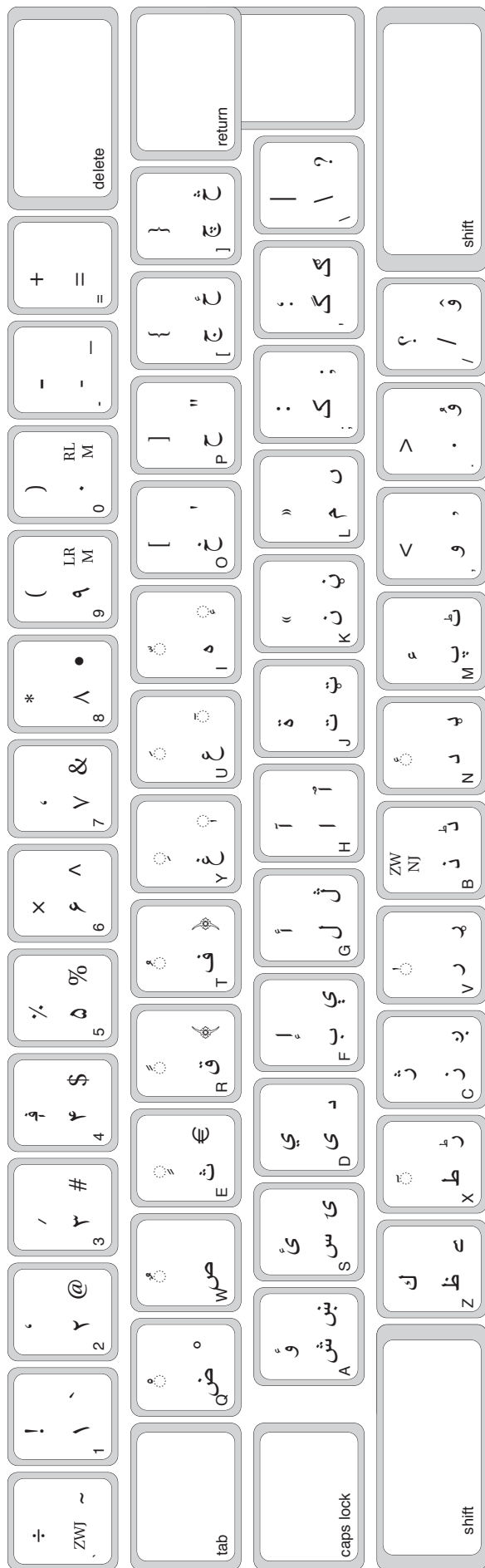
Keyboard layouts for Afghanistan

Key assignments for Pashto, using ISO/IEC 9995 notation:

	Unshifted	Shifted	AltGr
TLDE	200D	00F7	007E
E01	06F1	0021	0060
E02	06F2	066C	0040
E03	06F3	066B	0023
E04	06F4	AFGHANI	0024
E05	06F5	066A	0025
E06	06F6	00D7	005E
E07	06F7	00BB	0026
E08	06F8	00AB	2022
E09	06F9	0029	200E
E10	06F0	0028	200F
E11	002D	0640	005F
E12	003D	002B	<FREE>
D01	0636	0652	00B0
D02	0635	064C	0653
D03	062B	064D	20AC
D04	0642	064B	FD3E
D05	0641	064F	FD3F
D06	063A	0650	0656
D07	0639	064E	ZWARAKAY
D08	0647	0651	0670
D09	062E	0681	0027
D10	062D	0685	0022
D11	062C	005D	007D
D12	0686	005B	007B
C01	0634	069A	<FREE>
C02	0633	0626	06D2
C03	06CC	064A	0649
C04	0628	067E	06BA
C05	0644	0623	06B7
C06	0627	0622	0671
C07	062A	067C	0679
C08	0646	06BC	003E
C09	0645	0629	003C
C10	06A9	003A	0643
C11	06AB	061B	06AF
B01	06CD	0638	003F
B02	06D0	0637	003B
B03	0632	0698	0655
B04	0631	0621	0654
B05	0630	200C	0625
B06	062F	0689	0688
B07	0693	0624	0691
B08	0648	060C	002C
B09	0696	002E	06C7
B10	002F	061F	06C9
BKSL	005C	002A	007C
SPCE	0020	200C	00A0

Dari keyboard layout for Afghanistan
by Michael Everson & Roozbeh Pournader, 2003-07-05

Guide to the key assignments, reading top-down and left-right

Unshifted:
ZWJ, one, two, three, four, five, six, seven, eight, nine, zero, hyphen, equals
zād, sād, se, qāf, fe, qayn, ayn, he, xe, he, jim, če
šin, sin, ye, be, lām, alef, te, nūn, mīm, kāf, gāf, backslash
zā, tā, ze, re, zāl, dāl, pe, vāv, full-stop, slash

Shifted:
division, exclamation, thousands, decimals, afghani, percent, multiplication, comma, asterisk, parenthesis, parenthesis, kashida, plus
sukun, dammatan, kasratan, fathatan, damma, kasra, fatha, shadda, square-bracket, square-bracket, curly-bracket, curly-bracket
vāv-hamza, ye-hamza, saxta-ye, alef-hamza-below, alef-hamza, ā, te-ye-gerd, guillemet, guillemet, colon, semicolon, vertical-bar
arabic-kāf, maddah, že, alef-above, ZWNJ, hamza-above, hamza, less-than, greater-than, question

Alt-Gr:
tilde, grave, at, number, dollar, percent, circumflex, ampersand, bullet, LRM, RLM, underscore, —
degree, —, euro, ornament-parenthesis, ornament-parenthesis, alef-below, zwarakay, hamza-below, apostrophe, quote, pashto-ǰe, pashto-ce
pashto-šin, šəjina-ye, alef-e-maqsura, pasta-ye, brahui-lhām, alef-wasla, pashto-ṭe, pashto-ṛun, nūn-ghunna, semicolon, pashto-gāf, question
urdu-ye, urdu-ṛe, pashto-ẓe, pashto-ṛe, urdu-ḍāl, pashto-ḍāl, urdu-ṭe, comma, uzbek-u, uzbek-o

shift	altgr shift
plain	altgr

Unshifted spacebar →
SPACE

Shifted spacebar →
ZERO WIDTH NON-JOINER

Alt-Gr spacebar →
NO-BREAK SPACE

Computer Locale Requirements for Afghanistan
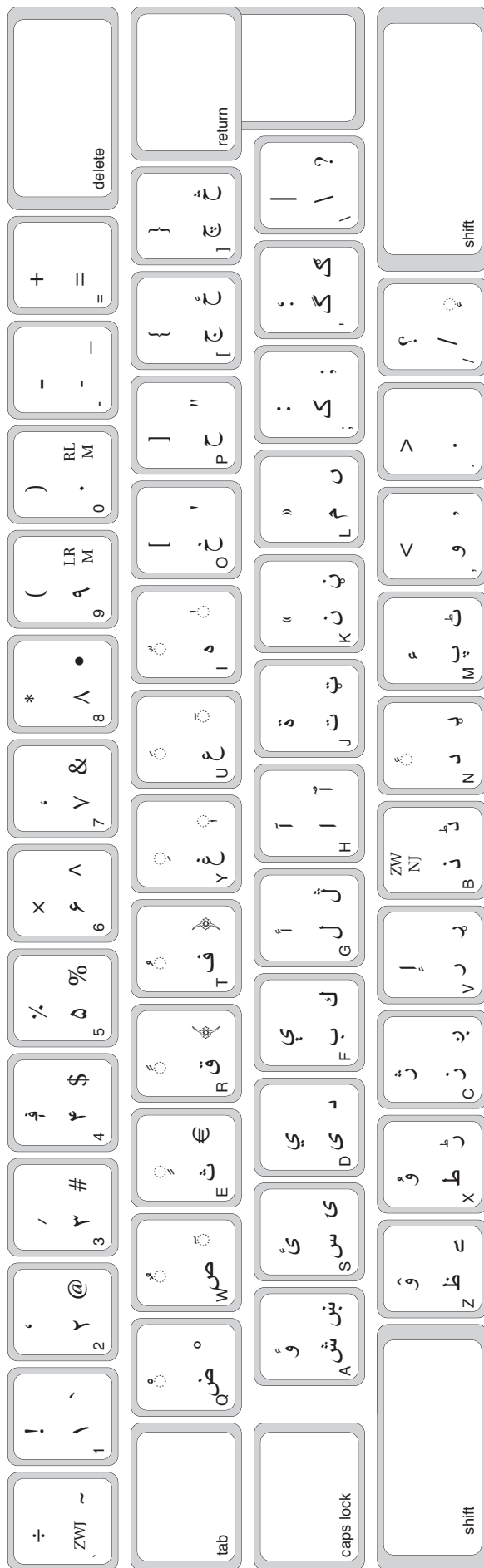
Keyboard layouts for Afghanistan

Key assignments for Dari, using ISO/IEC 9995 notation:

	Unshifted	**Shifted**	**AltGr**
TLDE	200D	00F7	007E
E01	06F1	0021	0060
E02	06F2	066C	0040
E03	06F3	066B	0023
E04	06F4	AFGHANI	0024
E05	06F5	066A	0025
E06	06F6	00D7	005E
E07	06F7	060C	0026
E08	06F8	002A	2022
E09	06F9	0029	200E
E10	06F0	0028	200F
E11	002D	0640	005F
E12	003D	002B	<FREE>
D01	0636	0652	00B0
D02	0635	064C	<FREE>
D03	062B	064D	20AC
D04	0642	064B	FD3E
D05	0641	064F	FD3F
D06	063A	0650	0656
D07	0639	064E	ZWARAKAY
D08	0647	0651	0655
D09	062E	005D	0027
D10	062D	005B	0022
D11	062C	007D	0681
D12	0686	007B	0685
C01	0634	0624	069A
C02	0633	0626	06CD
C03	06CC	064A	0649
C04	0628	0625	06D0
C05	0644	0623	06B7
C06	0627	0622	0671
C07	062A	0629	067C
C08	0646	00BB	06BC
C09	0645	00AB	06BA
C10	06A9	003A	003B
C11	06AF	061B	06AB
B01	0638	0643	06D2
B02	0637	0653	0691
B03	0632	0698	0696
B04	0631	0670	0693
B05	0630	200C	0688
B06	062F	0654	0689
B07	067E	0621	0679
B08	0648	003E	002C
B09	002E	003C	06C7
B10	002F	061F	06C9
BKSL	005C	007C	003F
SPCE	0020	200C	00A0

Southern Uzbek keyboard layout for Afghanistan

by Michael Everson & Roozbeh Pournader, 2003-07-07

Guide to the key assignments, reading top-down and left-right

Unshifted:

ZWJ, one, two, three, four, five, six, seven, eight, nine, zero, hyphen, equals

zwad, swad, seh, qaf, feh, ghain, ʼain, heh, kheh, heh, jeem, cheh

shin, sin, yih, beh, lam, alef, teh, nuun, meem, kaf, gaf, backslash

zuy, tuy, zeh, reh, zal, dal, peh, waw, full-stop, slash

Shifted:

division, exclamation, thousands, decimals, afghani, percent, multiplication, comma, asterisk, parenthesis, parenthesis, kashida, plus

sukun, dammatan, kasratan, fathatan, damma, kasra, fatha, shadda, square-bracket, square-bracket, curly-bracket, curly-bracket

waw-hamza, yih-hamza, yeeh, yeh, alef-hamza, alef-madha, teh marbuta, guillemet, guillemet, colon, semicolon, vertical-bar

oʼ, u, zheh, alef-hamza-below, ZWNJ, hamza-above, hamza, less-than, greater-than, question

Alt-Gr:

tilde, grave, at, number, dollar, percent, circumflex, ampersand, bullet, LRM, RLM, underscore, —

degree, maddah, euro, ornament-parenthesis, ornament-parenthesis, alef-below, zwarakay, alef-above, apostrophe, quote, pashto-ǰeh, pashto-ceh

pashto-şin, şejina-yih, alef-e-maqsura, arabic-kaf, brahui-lham, alef-wasla, pashto-ṭeh, pashto-ṇuun, nuun-ghunna, semicolon, pashto-gaf, question

urdu-yih, urdu-ṛeh, pashto-ẓeh, pashto-reh, urdu-ḍal, pashto-ḍal, urdu-ṭeh, comma, —, hamza-below

Unshifted spacebar → SPACE

Shifted spacebar → ZERO WIDTH NON-JOINER

Alt-Gr spacebar → NO-BREAK SPACE

shift	altgr
	shift
plain	altgr

Keyboard layouts for Afghanistan

Key assignments for Southern Uzbek, using ISO/IEC 9995 notation:

	Unshifted	**Shifted**	**AltGr**
TLDE	200D	00F7	007E
E01	06F1	0021	0060
E02	06F2	066C	0040
E03	06F3	066B	0023
E04	06F4	AFGHANI	0024
E05	06F5	066A	0025
E06	06F6	00D7	005E
E07	06F7	060C	0026
E08	06F8	002A	2022
E09	06F9	0029	200E
E10	06F0	0028	200F
E11	002D	0640	005F
E12	003D	002B	<FREE>
D01	0636	0652	00B0
D02	0635	064C	0653
D03	062B	064D	20AC
D04	0642	064B	FD3E
D05	0641	064F	FD3F
D06	063A	0650	0656
D07	0639	064E	ZWARAKAY
D08	0647	0651	0670
D09	062E	005D	0027
D10	062D	005B	0022
D11	062C	007D	0681
D12	0686	007B	0685
C01	0634	0624	069A
C02	0633	0626	06CD
C03	06CC	064A	0649
C04	0628	06D0	0643
C05	0644	0623	06B7
C06	0627	0622	0671
C07	062A	0629	067C
C08	0646	00BB	06BC
C09	0645	00AB	06BA
C10	06A9	003A	003B
C11	06AF	061B	06AB
B01	0638	06C9	06D2
B02	0637	06C7	0691
B03	0632	0698	0696
B04	0631	0625	0693
B05	0630	200C	0688
B06	062F	0654	0689
B07	067E	0621	0679
B08	0648	003E	002C
B09	002E	003C	<FREE>
B10	002F	061F	0655
BKSL	005C	007C	003F
SPCE	0020	200C	00A0

Computer Locale Requirements for Afghanistan

The official calendar of Afghanistan

The official calendar of Afghanistan is the solar Islamic calendar (also known as the Persian calendar or the *hejrī-e šamsī* calendar). This calendar counts up from the year of the Hegira of Muhammad. The Gregorian calendar is used in international activities or occasions, and the Lunar Islamic calendar (also known as *hejrī-e qamarī*) is used for religious ceremonies and some of the holidays.

The Persian calendar of Afghanistan, although very similiar to the Persian calendar of Iran, differs in the algorithm it uses to calculate leap years, which may lead to a one-day difference for some years. There are plans to synchronize these calendars, but the present report only describes the current calendar of Afghanistan. The Persian calendar has 12 months, consisting of 29 to 31 days. The names of the months in Dari and Pashto, together with the number of the days in each month, is given below. Other languages of Afghanistan usually use the Dari names of the months.

Dari		Pashto		Days
حمل	Hamal	وری	Wray	31
ثور	Sawr	غویی	Ğwayay	31
جوزا	Jawzā	غبرگولی	Ğbargolay	31
سرطان	Saratān	چنگاښ	Čungāṣ	31
اسد	Asad	زمری	Zmaray	31
سنبله	Sonbola	وږی	Waẓay	31
میزان	Mīzān	تله	Təla	30
عقرب	Aqrab	لړم	Laṛam	30
قوس	Qaws	لیندی	Lindəy	30
جدی	Jady	مرغومی	Marğūmay	30
دلو	Dalv	سلواغه	Salwāğə	30
حوت	Hūt	کب	Kab	29 (30 in leap years)

Iranian leap years are computed astronomically, but the Afghan leap years have an arithmetic formula. The formula is for synchronizing with the Gregorian calendar, and is rather simple in that regard: the Persian year x is a leap year if and only if the Gregorian year $x + 621$ is a leap year. For example, the Persian year 1383 will be a leap year in the Afghan Persian calendar, since the Gregorian year 2004 will be a leap year. For the sake of conversion between Gregorian dates and Persian years, one can use the following date for a reference point: **1 Hamal 1382 = 21 March 2003**.

The Lunar Islamic calendar used in Afghanistan is a civil Lunar Islamic calendar, based on pre-computation of the months. The number of the days of the months are 29 and 30 days alternatively. Leap years occur every two or three years, and make the month *Dhu l-Qa'da* a day longer. The exact algorithm used for leap years is reported to be complicated and was not made available to us. It is worth mentioning that because of the importance of Islamic months *Ramadan* and *Shawwal* in Muslim fasting and other ceremonies, the starting day of these two months is as observed by religious or governmental authorities rather than pre-computed. There have been (and will be) frequent cases of one day difference between the pre-computed calendar and the observed one. This difference is usually adjusted at the end of *Shawwal,* and the pre-computed calendar is used from the next month.

Acknowledgements