

DB65

新疆维吾尔自治区地方标准

DB 65/T 3689—2015

现行维吾尔文与维吾尔新文字 编码字符转换规则

Coded Character Conversion Rules between Uyghur and New Uyghur

2015 – 01 – 13 发布

2015 – 03 – 13 实施

新疆维吾尔自治区质量技术监督局 发布

前 言

本标准依照GB/T 1.1-2009《标准化工作导则 第1部分:标准的结构和编写》、GB 13000-2010《信息技术 通用多八位编码字符集(UCS)》、GB 21669-2008《信息技术 维吾尔文、哈萨克文、柯尔克孜文编码字符集》的规定编写。

本标准由新疆维吾尔自治区民族语言文字工作委员会提出。

本标准由新疆维吾尔自治区经济和信息化委员会归口。

本标准起草单位:新疆维吾尔自治区标准化研究院、新疆西北星信息技术有限责任公司、中科院新疆理化技术研究所、新疆维吾尔自治区民族语言文字工作委员会。

本标准主要起草人:董军、吾守尔·斯拉木、石常海、刘金龙、亚森·伊明、申文英、李晓、阿里木·哈萨尼、艾力·马合木提、王磊、吐尔根·伊布拉音、杨雅婷、塔依尔·阿布都外力。

引 言

现行维吾尔文与维吾尔新文字都是用于表达维吾尔语言的字母文字，但两种文字表达维吾尔语言字母的方式有许多显著区别，主要包括以下几个方面：

- 现行维吾尔文是从右向左书写的文字，维吾尔新文字是从左向右书写的文字。
- 现行维吾尔文字母不区分大小写，维吾尔新文字字母区分大小写。
- 现行维吾尔文用字符 ﻥ (该字符的编码为 0x0626) 与另外 8 个字符结合构成 8 个双字符字母，用这 8 个双字符字母表示 8 个维吾尔文元音字母，且根据上下文环境的不同，字符 ﻥ 有时需要出现，有时不需要出现；维吾尔新文字用 8 个字符分别表示这 8 个维吾尔文元音字母。
- 维吾尔新文字有 4 个双字符字母 Zh、Ch、Sh、Ng，分别表示 1 个维吾尔文字母，而在现行维吾尔文中这 4 个维吾尔文字母各用 1 个不同字符表示。
- 缩略语在维吾尔新文字中用每个单词的大写首字母表示，而在现行维吾尔文中则用每个单词的首字母加空格表示。

为了保证现行维吾尔文与维吾尔新文字编码字符转换过程中信息的完整性，应对这两种文字的书写方式进行必要的规范。因此，本标准提出了与编码字符转换有关的两种文字的书写规则，并制定两种文字编码字符的转换规则。

现行维吾尔文与维吾尔新文字编码字符转换规则

1 范围

本标准规定了现行维吾尔文与维吾尔新文字编码字符转换的术语和定义、书写规则及转换规则等。本标准适用于现行维吾尔文与维吾尔新文字编码字符的转换工具和各类信息系统。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB 13000 信息技术 通用多八位编码字符集(UCS)

3 术语和定义

下列术语和定义适用于本文件。

3.1

现行维吾尔文 uyghur

指阿拉伯字母为基础的维吾尔文。

3.2

维吾尔新文字 new uyghur

指拉丁字母为基础的维吾尔新文字。

3.3

字母 letter

注音及拼读的符号。

3.4

字符 character

供组织、控制或表示数据用的元素集合中的一个元素。

3.5

编码字符 coded character

字符及其编码表示。

3.6

单字符字母 mono letter

由一个字符表示的字母。

3.7

双字符字母 dual letter

由两个相邻字符按固定的先后顺序构成字符序列表示的字母。

3.8

阿拉伯字符 arabic character

用于表示阿拉伯字母和符号的编码字符，在GB 13000中，有5个阿拉伯字符的编码区域，分别是基本区（编码从0x0600到0x06FF）、补充区（编码从0x0750到0x077F）、扩展区（编码从0x08A0到0x08FF）、变形显现区A（编码从0xFB50到0xFDEE）、变形显现区B（编码从0xFE70到0xFEFF）。

3.9

拉丁字符 latin character

用于表示拉丁字母和符号的编码字符，在GB 13000中，有9个拉丁字符的编码区域，分别是基本区（编码从0x0000到0x007F）、补充区（编码从0x0080到0x00FF）、扩展区A（编码从0x0100到0x017F）、扩展区B（编码从0x0180到0x024F）、扩展区C（编码从0x2C60到0x2C7F）、扩展区D（编码从0xA720到0xA7FF）、附加扩展区（编码从0x1E00到0x1EFF）、连字区（编码从0xFB00到0xFB06）、全宽字母区（编码从0xFF00到0xFF5E）。

3.10

分隔符 delimiter

字符 '，主要用于音节分隔或区分双字符字母和单字符字母，该字符的编码为0x0027。

3.11

连接符 tatweel

字符 ـ，主要用于调整现行维吾尔文的编辑排版效果，该字符的编码为0x0640。

3.12

镜像字符 mirrored character

在GB 13000中，具有Bidi Mirrored属性的字符，例如各种括号。

3.13

布局和格式控制字符 layout and format control character

用于控制字符的连接行为、双向顺序行为和选择显示格式行为的字符，例如：编码为0x200C的零宽不连接字符，编码为0x200F的从右向左标记字符等。

3.14

零宽空格 zero width space

编码为0x200B的字符，宽度为零，主要用于指示单词的边界。

3.15

零宽不中断空格 zero width no-break space

编码为0xFEFF的字符，宽度为零，主要用于指示字符的前后即不分词也不断行。

3.16

一对一的编码字符转换 conversion of one coded character to one coded character

一个编码字符转换为一个编码字符。

3.17

一对多的编码字符转换 conversion of one coded character to several coded character

一个编码字符转换为多个编码字符。

3.18

多对多的编码字符转换 conversion of several coded character to several coded character

多个编码字符转换为多个编码字符。

3.19

多对一的编码字符转换 conversion of several coded character to one coded character

多个编码字符转换为一个编码字符。

4 与编码字符转换相关的现行维吾尔文书写规则

4.1 缩略语的每个字母之间应有一个空格（见示例 1）。

示例1: ش ئۇ ئار (شىنجاڭ ئۇيغۇر ئاپتونوم رايونى، 新疆维吾尔自治区。)

5 与编码字符转换有关的维吾尔新文字书写规则

5.1 句首字母应大写（见示例 2）。

5.2 双引号内的首字母应大写（见示例 3）。

5.3 专有名词(人名、地名、机构名称)的首字母应大写（见示例 4）。

5.4 双字符字母的大写形式只大写首字符（见示例 5）。

5.5 一个单词中，前一个音节是闭音节（以辅音字母结束的音节），并且后一个音节是元音字母开头时，前音节与后音节中间应标志分隔符 '（见示例 6）。

5.6 为了准确区分单字符字母和双字符字母，在一个单词中，如果相邻的两个字符 N 和 G，Z 和 H，S 和 H 分别代表一个维吾尔文字母，则应在他们之间标志分隔符 '（见示例 7）。

5.7 缩略语中的字母应大写（见示例 8）。

示例2: Mən məktəptə okuyman. (我在学校读书)

示例3: Alim mundak dedi: "Mən burun bundak igiz adəmni körüp bakmigan." (阿里木说: “我以前没有见过这么高的人。”)

示例4: Alim (阿里木), Turpan (吐鲁番)

示例5: Ng

示例6: qət'əl (外国)

示例7: nəmən'gan (纳曼干)

示例8: XUAR (新疆维吾尔自治区)

6 现行维吾尔文编码字符转换为维吾尔新文字编码字符的规则

6.1 总体要求

6.1.1 扩展区编码字符的转换

阿拉伯字符变形显现区A和变形显现区B的编码字符表达的现行维吾尔文，应首先转换为阿拉伯字符基本区编码字符表达的现行维吾尔文，然后再转换为对应的维吾尔新文字编码字符。

6.1.2 现行维吾尔文字母应先转换为维吾尔新文字字母的小写形式

如无特别说明，现行维吾尔文编码字符应先转换为对应的维吾尔新文字编码字符的小写形式，然后再根据维吾尔新文字的书写规则进行必要的大小写转换。

6.1.3 镜像字符的处理

如无特别说明，镜像字符保留不变。信息系统不能依据上下文环境选择镜像字符的正确字形用于显示的，需将一个镜像字符的编码字符转换为与之镜像的另一个编码字符。表1列出了部分镜像字符的镜像关系，完整的镜像字符列表可以参照GB 13000。

表1 部分镜像字符

文字	字符									
现行维吾尔文	(0028) 0029	[005B] 005D	{ 007B	} 007D	« 00AB	» 00BB	‹ 2039	› 203A
维吾尔新文字) 0029	(0028] 005D	[005B	} 007D	{ 007B	» 00BB	« 00AB	› 203A	‹ 2039

6.2 一对一的编码字符转换

6.2.1 现行维吾尔文 18 个辅音字母，其编码字符应直接转换为对应的维吾尔新文字字母的编码字符，具体见表 2（示例见 9 至 12）。

表2 一对一的编码字符转换

文字	字符								
现行维吾尔文	ب	پ	ت	خ	د	ر	ز	ژ	غ
维吾尔新文字	b	p	t	h	d	r	z	zh	q
现行维吾尔文	ف	ق	ك	گ	ل	م	ن	ھ	ي
维吾尔新文字	f	k	k	g	l	m	n	h	y

示例9: قول (手) 应转换为 kol

示例10: باش (头) 应转换为 bax

示例11: پۈت (足) 应转换为 put

示例12: كۆز (眼) 应转换为 kez

6.2.2 现行维吾尔文字母 ج 与两个维吾尔新文字字母 j 和 zh 对应，因此现行维吾尔文字母 ج 的编码字符需根据专用的词汇表判断转换为维吾尔新文字字母 j 或 zh 的编码字符。若专用的词汇表中没有包含需要转换的词条，转换现行维吾尔字母 ج 默认转换为维吾尔新文字字母 j 的编码字符，具体见表 3（示例见 13 至 14）。

表3 现行维吾尔文字母 ج 的转换

文字	字符	
现行维吾尔文	ج	ج
维吾尔新文字	j	zh

示例13: جەڭچى (战士) 应转换为 jǎngqi

示例14: جۇدې (朱德) 应转换为 Zhu De

6.2.3 现行维吾尔文字母 چ 与两个维吾尔新文字字母 q 和 ch 对应，现行维吾尔文字母 چ 的编码字符需根据专用的词汇表判断转换为维吾尔新文字字母 q 或 ch 的编码字符。若专用的词汇表中没有包含需要转换的词条，转换现行维吾尔字母 چ 默认转换为维吾尔文新文字字母 q 的编码字符，具体见表 5（示例见 15 至 16）。

表4 现行维吾尔文字母 چ 的转换

文字	字符	
现行维吾尔文	چ	چ
维吾尔新文字	q	ch

示例15: چەك (支票) 应转换为 qǎk

示例16: چاششا (长沙) 应转换为 Changsha

6.2.4 现行维吾尔文字母 س 与两个维吾尔新文字字母 s 和 c 对应，现行维吾尔文字母 س 的编码字符需根据专用的词汇表判断转换为维吾尔新文字字母 s 或 c 的编码字符。若专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔字母 س 默认转换为维吾尔文新文字字母 s 的编码字符，具体见表 5（示例见 17 至 18）。

表5 现行维吾尔文字母 س 的转换

文字	字符	
现行维吾尔文	س	س
维吾尔新文字	s	c

示例17: سان (数) 应转换为 San

示例18: سەيئى (蔡锷) 应转换为 Cai E

6.2.5 现行维吾尔文字母 ش 与两个维吾尔新文字字母 x 和 sh 对应，现行维吾尔文字母 ش 的编码字符需根据专用的词汇表判断转换为维吾尔新文字字母 x 或 sh 的编码字符。若专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔字母 ش 默认转换为维吾尔文新文字字母 x 的编码字符，具体见表 6（示例见 19 至 20）。

表6 现行维吾尔文字母 ش 的转换

文字	字符	
现行维吾尔文	ش	ش
维吾尔新文字	x	sh

示例19: شېر (狮子) 应转换为 xir

示例20: شاتىخەي (上海) 应转换为 Shanghai

6.2.6 现行维吾尔文 8 个元音字母是双字符字母，当出现在非词首位置时，根据上下文环境字符 ۋ 有时需要出现，有时不需要出现。如表 7 所示，当字符 ۋ 不出现时，则这 8 个元音字母直接转换为对应的维吾尔新文字元音字母的编码字符（示例见 21 至 25）。

表7 现行维吾尔文元音字母在非词首位置的转换

文字	字符							
现行维吾尔文	ا	ە	و	ۇ	ۆ	ۈ	ې	ى
维吾尔新文字	a	ə	o	u	e	ü	e	i

示例21: كىتاب (书) 应转换为 kitab

示例22: ۋەتەن (祖国) 应转换为 wətən

示例23: تومۇر (经脉) 应转换为 tomur

示例24: كۆمۈر (煤) 应转换为 kɵmür

示例25: ئېلېكتىر (电) 应转换为 elektir

6.2.7 现行维吾尔字母 ۋ, 编码字符与两个维吾尔新文字字母 W 和 V 的编码字符对应。需要根据专用的词汇表来确定现行维吾尔文字母 ۋ 的编码字符转换为维吾尔新文字字母 W 或 V 的编码字符。若专用词汇表中没有包含需要转换的词条, 转换现行维吾尔字母 ۋ 默认转换为维吾尔新文字字母 w 的编码字符(示例见 26 至 27)。

表8 现行维吾尔文字母ۋ的转换

文字	字符	
现行维吾尔文	ۋ	ۋ
维吾尔新文字	w	v

示例26: ۋەتەن (祖国) 应转换为 wətən

示例27: ۋېيتنام (越南) 应转换为 Vyetnam

6.3 一对多的编码字符转换

现行维吾尔文字母 ڭ, 其编码字符直接转换为对应的维吾尔新文字双字符字母 ng 的编码字符具体见表 9 (见示例 28)。

表9 现行维吾尔文字母ڭ的转换

文字	字符
现行维吾尔文	ڭ
维吾尔新文字	ng

示例28: شىنجاڭ (新疆) 应转换为 Xingjang

6.4 多对一的编码字符转换

现行维吾尔文 8 个元音字母是双字符字母, 当出现在词首位置时, 应忽略词首的字符 ۋ, 双字符字母中的另一个字符直接转换为对应的维吾尔新文字元音字母的编码字符, 具体见表 10 (示例见 29 至 36)。

表10 现行维吾尔文元音字母在词首位置的转换

文字	字符							
现行维吾尔文	ئا	ئە	ئو	ئۇ	ئۆ	ئۈ	ئي	ىي
维吾尔新文字	a	ə	o	u	ɵ	ü	e	i

示例29: ئانار (石榴) 应转换为 anar

示例30: ئەنجۈر (无花果) 应转换为 ənjür

示例31: ئوردا (皇宫) 应转换为 orda

示例32: ئۇرۇش (战争) 应转换为 urux

示例33: ئۆردەك (鸭子) 应转换为 ɵrdək

示例34: ئۈزۈم (葡萄) 应转换为 üzüm

示例35: ئېلان (广告) 应转换为 elan

示例36: ئىنكاس (反映) 应转换为 inkas

6.5 多对多的编码字符转换

6.5.1 现行维吾尔文8个元音字母是双字符字母，当出现在非词首位置时，根据上下文环境字符 ۋ 有时需要出现，有时不需要出现。当字符 ۋ 出现时，则字符 ۋ 转换为分隔符 ' 的编码字符，双字符字母的另一个字符转换为对应的维吾尔新文字元音字母的编码字符，具体见表11（示例见37至44）。

表11 现行维吾尔文元音字母在非词首位置的转换

文字	字符							
现行维吾尔文	ئا	ئە	ئو	ئۇ	ۆ	ۈ	ئي	ىي
维吾尔新文字	'a	'ə	'o	'u	'ə	'ü	'e	'i

- 示例37: ئىنىكئانا (奶妈) 应转换为 inik'ana
示例38: ئەسئەت (艾斯艾提) 应转换为 Əs'ət
示例39: رادىئو (收音机) 应转换为 radi'o
示例40: مەسئۇل (负责) 应转换为 məs'u'l
示例41: قارىئۆرۈك (乌梅) 应转换为 ari'er'ük
示例42: ئائۇمىد (失望) 应转换为 na'ümid
示例43: ئىتئېيىق (狗熊) 应转换为 it'eyi
示例44: جەمئىي (共计) 应转换为 jəm'iy

6.5.2 单词中有表12所示的现行维吾尔文字符序列，字符序列中的每个编码字符转换为对应的维吾尔新文字编码字符，并应在两个维吾尔新文字编码字符之间标志分隔符 '（示例见45至46）。

表12 现行维吾尔文字符序列نگ and زخ的转换

文字	字符	
现行维吾尔文	نگ	زخ
维吾尔新文字	n'g	z'h

- 示例45: نەمەنگان (纳曼干) 应转换为 Nəmən'gan
示例46: ئۆزخان (吾孜汗) 应转换为 Əz'han

6.5.3 现行维吾尔文字符序列 سخ，与维吾尔新文字字符序列 c'h 和 s'h 对应，其字符编码需根据专用词汇表确定转换为维吾尔新文字字符序列 c'h 或 s'h。若专用的词汇表中没有包含需要转换的词条，则转换时现行维吾尔文字符序列 سخ 默认转换为维吾尔新文字字母 s'h 的编码字符，具体见表13（见示例47）。

表13 现行维吾尔文字符序列سخ的转换

文字	字符	
现行维吾尔文	سخ	سخ
维吾尔新文字	c'h	s'h

- 示例47: پاسخا بايرىمى (复活节) 应转换为 Pas'ha Bayrimi

6.6 缩略语的编码字符转换

相邻的两个或两个以上现行维吾尔文字母左右及之间有1个空格，应忽略字母之间的空格，并将现行维吾尔文字母转换为对应的维吾尔新文字字母的大写形式（见示例48至50）。

- 示例48: ج ك پ (中国共产党) 应转换为 ZhKP
示例49: ج خ (公安) 应转换为 JH
示例50: ش ئۇ ئار (新疆维吾尔自治区) 应转换为 XUAR

6.7 标点符号的编码字符转换

现行维吾尔文标点符号的编码字符应转换为对应的维吾尔新文字标点符号的编码字符，具体见表9（示例见51至53）。

表14 现行维吾尔文标点符号的转换

文字	字符		
现行维吾尔文	‘ 060C	؟ 061F	؛ 061B
维吾尔新文字	’ 002C	? 003F	; 003B

示例51: ؟ نېمە؟ (这是什么?)应转换为 bu nemə?
示例52: مەن، سەن (我，你)应转换为 mən,sən
示例53: ئادەم؛ ھايۋان (人；动物)应转换为 adəm; aywan

6.8 非现行维吾尔文编码字符的转换

6.8.1 现行维吾尔文中拉丁字符的处理

现行维吾尔文中的拉丁字符应保留不变，同时在拉丁字符的前后分别增加零宽空格和零宽不中断空格。

6.8.2 零宽空格和零宽不中断空格的处理

忽略现行维吾尔文中的零宽空格和零宽不中断空格，同时零宽空格和零宽不中断空格之间的字符保留不变。

6.8.3 布局和格式控制字符的处理

现行维吾尔文中的布局和格式控制字符保留不变。

6.8.4 连接符的处理

当连接符出现在现行维吾尔文单词中，转换为维吾尔新文字时应忽略连接符。

6.8.5 单个不带字符ﻻ的现行维吾尔文元音字母的处理

现行维吾尔文的8个元音字母都是由包括字符 ﻻ 在内的两个字符构成的双字符字母，根据上下文环境，字符 ﻻ 有时需要出现，有时不需要出现。当不带字符 ﻻ 的现行维吾尔文元音字母单独出现的时候，在转换为维吾尔新文字时其编码字符保留不变。

6.8.6 其它非现行维吾尔文编码字符的处理

现行维吾尔文中的其它非现行维吾尔文编码字符保留不变。

6.9 现行维吾尔文与维吾尔新文字编码字符对照见附录 A。

7 维吾尔新文字编码字符转换为现行维吾尔文编码字符的规则

7.1 总体要求

7.1.1 扩展区编码字符的转换

如无特别说明，维吾尔新文字编码字符转换为对应现行维吾尔文字母名义形式的编码字符。信息系统不能依据现行维吾尔文字母名义形式的编码字符，选择正确的变形显现形式编码字符用于显示的，需

将现行维吾尔文字母名义形式的编码字符，依据现行维吾尔文的书写习惯，进一步转换为正确的变形显现形式编码字符。

7.1.2 维吾尔新文字字母大小写形式的处理

如无特别说明，在维吾尔新文字编码字符转换为现行维吾尔文编码字符的过程中，对同一个维吾尔新文字字母的大写形式和小写形式采用同样的方式处理。

7.1.3 镜像字符的处理

如无特别说明，镜像字符保留不变。信息系统不能依据上下文环境选择镜像字符的正确字形用于显示的，需将一个镜像字符的编码字符转换为与之镜像的另一个编码字符。表15列出了部分镜像字符的镜像关系，完整的镜像字符列表可以参照GB 13000。

表15 部分镜像字符的转换

文字	字符									
维吾尔新文字	(0028) 0029	[005B] 005D	{ 007B	} 007D	« 00AB	» 00BB	‹ 2039	› 203A
现行维吾尔文) 0029	(0028] 005D	[005B	} 007D	{ 007B	» 00BB	« 00AB	› 203A	‹ 2039

7.2 一对一的编码字符转换

7.2.1 维吾尔新文字 25 个辅音字母，其编码字符应直接转换为对应的现行维吾尔文字母的编码字符，具体见表 16（示例见 54 至 57）。

表16 维吾尔新文字辅音字母的转换

文字	字符												
维吾尔新文字	b	p	t	j	q	h	d	r	z	z	s	c	x
现行维吾尔文	ب	پ	ت	ج	چ	خ	د	ر	ز	ژ	س	س	ش
维吾尔新文字	q	f	k	k	g	l	m	n	h	y	w	v	
现行维吾尔文	غ	ف	ق	ك	گ	ل	م	ن	ھ	ي	ۇ	ۋ	

- 示例54: kol（手）应转换为 قول
- 示例55: bax（头）应转换为 باش
- 示例56: put（足）应转换为 پۈت
- 示例57: kəz（眼）应转换为 كۆز

7.2.2 维吾尔新文字 8 个元音字母在非词首时，转换为对应现行维吾尔文元音字母不带字符 ۋ 的形式 的编码字符，具体见表 17（示例见 58 至 62）。

表17 维吾尔新文字元音字母在非词首位置的转换

文字	字符							
维吾尔新文字	a	ə	o	u	ə	ü	e	i
现行维吾尔文	ا	ە	و	ۇ	ۆ	ۈ	ې	ى

- 示例58: kitab（书）应转换为 كىتاب
- 示例59: wətən（祖国）应转换为 ۋەتەن
- 示例60: tomur（经脉）应转换为 تومۇر
- 示例61: kmür（煤）应转换为 كۆمۈر
- 示例62: elektir（电）应转换为 ئېلېكتىر

7.3 一对多的编码字符转换

维吾尔新文字 8 个元音字母在词首位置时，转换为对应现行维吾尔文元音字母的编码字符，具体见表 18（示例见 63 至 70）。

表18 维吾尔新文字元音字母在词首位置的转换

文字	字符							
维吾尔新文字	a	ə	o	u	ə	ü	e	i
现行维吾尔文	ئا	ئە	ئو	ئۇ	ئۆ	ئۈ	ئې	ئى

示例63: anar（石榴）应转换为 ئانار

示例64: ənjür（无花果）应转换为 ئەنجۈر

示例65: orda（皇宫）应转换为 ئوردا

示例66: urux（战争）应转换为 ئۇرۇش

示例67: ərɗək（鸭子）应转换为 ئۆردەك

示例68: üzüm（葡萄）应转换为 ئۈزۈم

示例69: elan（广告）应转换为 ئېلان

示例70: inkas（反映）应转换为 ئىنكاس

7.4 多对一的编码字符转换

维吾尔新文字双字符字母的编码字符直接转换为对应的现行维吾尔文字母的编码字符，具体见表 19（示例见 71 至 74）。

表19 维吾尔新文字双字符字母的转换

文字	字符			
维吾尔新文字	ng	zh	ch	sh
现行维吾尔文	ڭ	چ	چ	ش

示例71: Xinjang（新疆）应转换为 شىنجاڭ

示例72: Zhu De（朱德）应转换为 جۇدې

示例73: Chencang（陈仓）应转换为 چېنساڭ

示例74: Shanghai（上海）应转换为 شائخەي

7.5 分隔符编码字符的转换

分隔符 '在维吾尔新文字字母 n 和 g, z 和 h, c 和 h, s 和 h 之间，转换时忽略分隔符；分隔符 '在维吾尔新文字元音字母之前，转换为字符 ۋ 的编码字符；分隔符 '在其它位置，则保留不变（示例见75至78）。

示例75: Nəmən'gan（纳曼干）应转换为 نەمەنگان

示例76: Əz'han（吾孜汗）应转换为 ئۆزخان

示例77: miras'hor（继承人）应转换为 مىراسخور

示例78: qət'al（外国）应转换为 قەتئەل

7.6 缩略语的编码字符转换

相邻的多个大写维吾尔新文字字母编码字符（双字符字母 Ng 作为大写维吾尔新文字字母）转换为现行维吾尔文编码字符时，直接转换为对应的现行维吾尔文编码字符，并在现行维吾尔文编码字符之间加1个空格编码字符（见示例79至81）。

- 示例79: JKP (中国共产党) 应转换为 ج ك پ
- 示例80: JH (公安) 应转换为 ج خ
- 示例81: XUAR (新疆维吾尔自治区) 应转换为 ش ئۇ ئا ر

7.7 标点符号的转换

表20所示的维吾尔新文字标点符号的编码字符应转换为对应的现行维吾尔文标点符号的编码字符(示例见82至84)。

表20 维吾尔新文字标点符号的转换

文字	字符		
维吾尔新文字	002C	?	;
现行维吾尔文	060C	؟	؛

- 示例82: bu nemə? (这是什么?) 应转换为 بۇ نېمە ؟
- 示例83: mən,sən (我, 你) 应转换为 مەن، سەن
- 示例84: adəm; aywan (人; 动物) 应转换为 ئادەم؛ ھايۋان

7.8 非维吾尔新文字编码字符的转换

7.8.1 维吾尔新文字中阿拉伯字符的处理

维吾尔新文字中的阿拉伯字符应保留不变,并在阿拉伯字符的前后分别增加零宽空格和零宽不中断空格。

7.8.2 零宽空格和零宽不中断空格的处理

忽略维吾尔新文字中的零宽空格和零宽不中断空格,零宽空格和零宽不中断空格之间的字符保留不变。

7.8.3 布局 and 格式控制字符的处理

维吾尔新文字中的布局 and 格式控制字符保留不变。

7.8.4 其它非维吾尔新文字编码字符的处理

维吾尔新文字中的其它非维吾尔新文字的编码字符保留不变。

7.8.5 有特殊要求需要保留不变的字符或字符串

有特殊要求需要保留的字符或字符串保留不变,例如罗马数字(I、II、III、IV、X)、字母名词(WTO、VCD、DVD)等。

7.9 维吾尔新文字与现行维吾尔文编码字符对照见附录 B。

附 录 A
(规范性附录)
现行维吾尔文与维吾尔新文字编码字符对照

A.1 现行维吾尔文与维吾尔新文字编码字符对照

现行维吾尔文与维吾尔新文字编码字符对照见表A.1。

表A.1 现行维吾尔文与维吾尔新文字编码字符对照表

序号	现行维吾尔文					维吾尔新文字	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
1	ئا 0626+0627	ئا ا FBEA FE8D			ئا ل FBEB FE8E	A 0041	a 0061
						'A 0027+0041	'a 0027+0061
2	ئە 0626+06D5	ئە ە FBEC FEE9			ئە ە FBED FEEA	Ə 018F	ə 0259
						'Ə 0027+018F	'ə 0027+0259
3	ب 0628	ب FE8F	ب FE91	ب FE92	ب FE90	B 0042	b 0062
4	پ 067E	پ FB56	پ FB58	پ FB59	پ FB57	P 0050	p 0070
5	ت 062A	ت FE95	ت FE97	ت FE98	ت FE96	T 0054	t 0074
6	ج 062C	ج FE9D	ج FE9F	ج FEA0	ج FE9E	J 004A	j 006A
						Zh 005A+0068	zh 007A+0068
7	چ 0686	چ FB7A	چ FB7C	چ FB7B	چ FB7D	Q 0051	q 0071
						Ch 0043+0068	ch 0063+0068
8	خ 062E	خ FEA5	خ FEA7	خ FEA8	خ FEA6	H 0048	h 0068
9	د 062F	د FEA9			د FEAA	D 0044	d 0064
10	ر 0631	ر FEAD			ر FEAE	R 0052	r 0072
11	ز 0632	ز FEAF			ز FEB0	Z 005A	z 007A
12	ژ 0698	ژ FB8A			ژ FB8B	Z 2C6B	z 2C6C

表 A. 1 (续)

序号	现行维吾尔文					维吾尔新文字	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
13	س 0633	س FEb1	س FEb3	س FEB4	س FEB2	S 0053	s 0073
						C 0043	c 0063
14	ش 0634	ش FEB5	ش FEB7	ش FEB8	ش FEB6	X 0058	x 0078
						Sh 0053+0068	sh 0073+0068
15	غ 063A	غ FECD	غ FECF	غ FED0	غ FECE	Q 01A2	q 01A3
16	ف 0641	ف FED1	ف FED3	ف FED4	ف FED2	F 0046	f 0066
17	ق 0642	ق FED5	ق FED7	ق FED8	ق FED6	K 2C69	k 2C6A
18	ك 0643	ك FED9	ك FEDB	ك FEDC	ك FEDA	K 004B	k 006B
19	گ 06AF	گ FB92	گ FB94	گ FB95	گ FB93	G 0047	g 0067
20	ڭ 06AD	ڭ FBD3	ڭ FBD5	ڭ FBD6	ڭ FBD4	Ng 004E+0067	ng 006E+0067
21	ل 0644	ل FEDD	ل FEDF	ل FEE0	ل FEDE	L 004C	l 006C
22	م 0645	م FEE1	م FEE3	م FEE4	م FEE2	M 004D	m 006D
23	ن 0646	ن FEE5	ن FEE7	ن FEE8	ن FEE6	N 004E	n 006E
24	ھ 06BE	ھ FBAA	ھ FBAC	ھ FBAD	ھ FBAB	H 2C67	h 2C68
25	ئو 0626+0648	و ئو FBEE FEED			و ئو FBEF FEEF	O 004F	o 006F
						'O 0027+004F	'o 0027+006F
26	ئۇ 0626+06C7	ۇ ئۇ FBF0 FBD7			ۇ ئۇ FBF1 FBD8	U 0055	u 0075
						'U 0027+0055	'u 0027+0075
27	ئۆ 0626+06C6	ۆ ئۆ FBF2 FBD9			ۆ ئۆ FBF3 FBDA	Θ 019F	θ 0275
						'Θ 0027+019F	'θ 0027+0275

表 A. 1 (续)

序号	现行维吾尔文					维吾尔新文字	
	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式	大写形式	小写形式
28	ئو 0626+06C8	ۈ ئو FBF4 FBDB			ۈ ئو FBF5 FBDC	Ü 00DC	ü 00FC
						'Ü 0027+00DC	'ü 0027+00FC
29	ۋ 06CB	ۋ FBDE			ۋ FBDF	W 0057	w 0077
						V 0056	v 0076
30	ئي 0626+06D0	ي ئي FBF6 FBE4	ي ئي FBF8 FBE6	ي FBE7	ي ئي FBF7 FBE5	E 0045	e 0065
						'E 0027+0045	'e 0027+0065
31	ئي 0626+0649	ي ئي FBF9 FEEF	ي ئي FBFB FBE8	ي FBE9	ي ئي FBFA FEF0	I 0049	i 0069
						'I 0027+0049	'i 0027+0069
32	ي 064A	ي FEF1	ي FEF3	ي FEF4	ي FEF2	Y 0059	y 0079
33	ئى 0626	ئى FE89	ئى FE8B	ئى FE8C	ئى FE8A	، 0027	، 0027
34	زخ 0632+062E	زخ FEAF+FEA5	زخ FEAF+FEA7	زخ FEB0+FEA7	زخ FEB0+FEA5	Z'H 005A+0027+0048	z'h 007A+0027+0068
35	سخ 0633+062E	سخ FEB1+FEA5		سخ FEB4+FEA8	سخ FEB4+FEA6	C'H 0043+0027+0048	c'h 0063+0027+0068
						S'H 0053+0027+0048	s'h 0073+0027+0068
36	نگ 0646+06AF	نگ FEE5+FB92	نگ FEE7+FB95	نگ FEE8+FB95	نگ FEE8+FB93	N'G 005A+0027+0047	n'g 006E+0027+0067
37	، 060C	، 060C	، 060C	، 060C	، 060C	، 002C	، 002C
38	؟ 061F	؟ 061F	؟ 061F	؟ 061F	؟ 061F	？ 003F	？ 003F
39	؛ 061B	؛ 061B	؛ 061B	؛ 061B	؛ 061B	； 003B	； 003B

附 录 B
(规范性附录)
维吾尔新文字与现行维吾尔文编码字符对照

B.1 维吾尔新文字与现行维吾尔文编码字符对照

维吾尔新文字与现行维吾尔文编码字符对照见表B.1。

表 B.1 维吾尔新文字与现行维吾尔文编码字符对照表

序号	维吾尔新文字		现行维吾尔文				
	大写形式	小写形式	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式
1	A 0041	a 0061	ئا 0626+0627	ا ئا FBEA FE8D			ئا ل FBEB FE8E
2	B 0042	b 0062	ب 0628	ب FE8F	ب FE91	ب FE92	ب FE90
3	C 0043	c 0063	س 0633	س FEb1	س FEb3	س FEB4	س FEB2
4	D 0044	d 0064	د 062F	د FEA9			د FEAA
5	E 0045	e 0065	ئې 0626+06D0	ي ئې FBF6 FBE4	ب ئې FBF8 FBE6	ب FBE7	ي ئې FBF7 FBE5
6	F 0046	f 0066	ف 0641	ف FED1	ف FED3	ف FED4	ف FED2
7	G 0047	g 0067	گ 06AF	گ FB92	گ FB94	گ FB95	گ FB93
8	H 0048	h 0068	خ 062E	خ FEA5	خ FEA7	خ FEA8	خ FEA6
9	I 0049	i 0069	ئى 0626+0649	ي ئى FBF9 FEEF	د ئى FBFB FBE8	د FBE9	ي ئى FBFA FEF0
10	J 004A	j 006A	چ 062C	چ FE9D	چ FE9F	چ FEA0	چ FE9E
11	K 004B	k 006B	ك 0643	ك FED9	ك FEDB	ك FEDC	ك FEDA
12	L 004C	l 006C	ل 0644	ل FEDD	ل FEDF	ل FEE0	ل FEDE
13	M 004D	m 006D	م 0645	م FEE1	م FEE3	م FEE4	م FEE2
14	N 004E	n 006E	ن 0646	ن FEE5	ن FEE7	ن FEE8	ن FEE6

表 B.1 (续)

序号	维吾尔新文字		现行维吾尔文				
	大写形式	小写形式	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式
15	O 004F	o 006F	ئو 0626+0648	و ئو FBEE FEED			و ئو FBEF FEEE
16	P 0050	p 0070	پ 067E	پ FB56	پ FB58	پ FB59	پ FB57
17	Q 0051	q 0071	چ 0686	چ FB7A	چ FB7C	چ FB7B	چ FB7D
18	R 0052	r 0072	ر 0631	ر FEAD			ر FEAE
19	S 0053	s 0073	س 0633	س FEb1	س FEb3	س FEB4	س FEB2
20	T 0054	t 0074	ت 062A	ت FE95	ت FE97	ت FE98	ت FE96
21	U 0055	u 0075	ئۇ 0626+06C7	ۇ ئۇ FBF0 FBD7			ۇ ئۇ FBF1 FBD8
22	V 0056	v 0076	ۋ 06CB	ۋ FBDE			ۋ FBDF
23	W 0057	w 0077	ۋ 06CB	ۋ FBDE			ۋ FBDF
24	X 0058	x 0078	ش 0634	ش FEB5	ش FEB7	ش FEB8	ش FEB6
25	Y 0059	y 0079	ي 064A	ي FEF1	ي FEF3	ي FEF4	ي FEF2
26	Z 005A	z 007A	ز 0632	ز FEAF			ز FEB0
27	Ü 00DC	ü 00FC	ئۇ 0626+06C8	ۇ ئۇ FBF4 FBDB			ۇ ئۇ FBF5 FBDC
28	Ə 018F	ə 0259	ئە 0626+06D5	ە ئە FBEC FEE9			ە ئە FBED FEEA
29	Ө 019F	ө 0275	ئۆ 0626+06C6	ۆ ئۆ FBF2 FBD9			ۆ ئۆ FBF3 FBDA
30	Q̣ 01A2	q̣ 01A3	غ 063A	غ FECF	غ FECF	غ FED0	غ FECE
31	Ḥ 2C67	ḥ 2C68	ھ 06BE	ھ FBAA	ھ FBAC	ھ FBAD	ھ FBAB
32	Ḳ 2C69	ḳ 2C6A	ق 0642	ق FED5	ق FED7	ق FED8	ق FED6
33	Ẓ 2C6B	ẓ 2C6C	ڭ 0698	ڭ FB8A			ڭ FB8B

表 B.1 (续)

序号	维吾尔新文字		现行维吾尔文				
	大写形式	小写形式	名义形式	独立字符形式	首写字符形式	中字符形式	尾字符形式
34	Ch 0043+0068	ch 0063+0068	چ 0686	چ FB7A	چ FB7C	چ FB7B	چ FB7D
35	Ng 004E+0067	ng 006E+0067	ڭ 06AD	ڭ FBD3	ڭ FBD5	ڭ FBD6	ڭ FBD4
36	Sh 0053+0068	sh 0073+0068	ش 0634	ش FEB5	ش FEB7	ش FEB8	ش FEB6
37	Zh 005A+0068	zh 007A+0068	ج 062C	ج FE9D	ج FE9F	ج FEA0	ج FE9E
38	C'H 0043+0027 + 0048	c'h 0063+0027 +0068	سخ 0633+062E	سخ FEb1+FEA5		سخ FEB4+FEA8	سخ FEB4+FEA6
39	N'G 005A+0027 +0047	n'g 006E+0027 +0067	نگ 0646+06AF	نگ FEE5+FB92	نگ FEE7+FB95	نگ FEE8+FB95	نگ FEE8+FB93
40	S'H 0053+0027 + 0048	s'h 0073+0027 +0068	سخ 0633+062E	سخ FEb1+FEA5		سخ FEB4+FEA8	سخ FEB4+FEA6
41	Z'H 005A+0027 +0048	z'h 007A+0027 +0068	زخ 0632+062E	زخ FEAF+FEA5	زخ FEAF+FEA7	زخ FEB0+FEA7	زخ FEB0+FEA5
42	، 0027	، 0027	، 0027	، 0027	، 0027	، 0027	، 0027
			ڧ 0626	ڧ FE89	ڧ FE8B	ڧ FE8C	ڧ FE8A
43	، 002C	، 002C	، 060C	، 060C	، 060C	، 060C	، 060C
44	؟ 003F	؟ 003F	؟ 061F	؟ 061F	؟ 061F	؟ 061F	؟ 061F
45	؛ 003B	؛ 003B	؛ 061B	؛ 061B	؛ 061B	؛ 061B	؛ 061B