

BDS 1520:2011

Bangladesh Standard  
SPECIFICATION FOR  
BANGLA CODED CHARACTER SET FOR  
INFORMATION INTERCHANGE  
(Second Revision)



BANGLADESH STANDARDS AND TESTING INSTITUTION  
116-A, TEJGAON INDUSTRIAL AREA, DHAKA-1208  
BANGLADESH

**Bangladesh Standard  
SPECIFICATION FOR  
BANGLA CODED CHARACTER SET FOR  
INFORMATION INTERCHANGE**

Committee on Standardization of Bangla for use in ICT

**CHAIRMAN**

1. Executive Director

**REPRESENTING**

Bangladesh Computer Council

**MEMBER**

- |  |  |
|--|--|
| 2. Director General or<br>(An experienced representative from his<br>organization) | Bangla Academy   |
| 3. Joint Secretary (Development)   | Ministry of Science and ICT  |
| 4. Head of the Department of CSE   | Bangladesh University of Engineering and<br>Technology                                     |
| 5. Professor Monsur Musa   | Institute of Modern Language, University of Dhaka  |
| 6. Representative  | Center for Research on Bangla Language<br>Processing, BRAC University                      |
| 7. Representative  | Bangladesh Standards and Testing Institution   |
| 8. Representative  | Bangladesh Computer Samity (BCS)   |
| 9. Representative  | Bangladesh Association of Software & information<br>Services (BASIS)                       |
| 10. Representative   | Association of Mobile Telecom Operator in<br>Bangladesh (AMTOB)                            |
| 11. Representative   | Bangladesh Open Source Network (BDOSN)   |
| 12. Representative   | Internet Service Provider Association of<br>Bangladesh (ISPAB)                             |
| 13. Mr. S M Kamal  | Former Country Manager of IBM, Bangladesh and<br>Former Managing Director of BEXIMCO Group |
| 14. Deputy Director (Systems)  | Bangladesh Computer Council  |
| <b>MEMBER SECRETARY</b>  |  |
| 15. Senior Systems Analyst   | Bangladesh Computer Council  |

Bangladesh Standard  
**SPECIFICATION FOR  
BANGLA CODED CHARACTER SET  
FOR INFORMATION INTERCHANGE**  
(Second Revision)

0. **FOREWORD**

This Bangladesh Standard was adopted by the Bangladesh Standards and Testing Institution on 15-02-2011 after the standard was revised by the Committee on Standardization of Bangla for use in ICT had been approved by the Electrical and Electronics Divisional Committee.

- 0.1 This standard contains a set of Bangla Characters (graphic character such as letters, digits & symbols) with their coded representation. Most of these characters are mandatory and unchangeable.
- 0.2 This standard specifies Bangla Character Codes intended for the interchange of information among data processing systems including the recording of data in the form of codes on media.
- 0.3 This character set is applicable to all letters of Bangla Characters.
- 0.4 This standard is based on latest revision of Unicode, a 16-bit coded character set for information interchange, issued by Unicode Consortium and thus acknowledged with thanks.

---

BANGLADESH STANDARDS AND TESTING INSTITUTION  
(MAAN BHABAN)  
116-A, TEJGAON INDUSTRIAL AREA, DHAKA-1208

## 1. SCOPE

- 1.1 This standard specifies Bangla version of character code which are included in the Unicode version 6.0.0 and specification of the codes and with their coded representations.
- 1.2 This standard specifies Bangla Character Codes intended for the interchange of information among data processing systems including the recording of data in the form of codes on media
- 1.3 This character set is applicable to alphabets of the Bangla script.

## 2. CONFORMANCE

- 2.1 **Conformance of Information interchange-** A coded-character-data-element (CC-data-element) within coded information for interchange is in conformance with this Standard if all the coded representations of characters within that CC-data-element conform to this Standard.
- 2.2 **Conformance of device-** A device is in conformance with this Standard if it conforms to the requirement of 2.2.1 and either or both of 2.2.2 and 2.2.3 below. A claim of conformance shall identify the version adopted.
  - 2.2.1 **Device description-** A device that conforms to this standard shall be the subject of a description that identifies the means by which the user may supply characters to the device, or may recognize them when they are made available to him, as specified respectively in 2.2.2 and 2.2.3.
  - 2.2.2 **Originating devices-** An Originating device shall allow its user to supply any sequence of characters from the version adopted, and shall be capable of transmitting their coded representations within a CC-data-elements.
  - 2.2.3 **Receiving devices-** A receiving device shall be capable of receiving and interpreting any coded representations of characters that are within a CC-data-element and that conform to 2.1 and shall made the corresponding characters available to its user in such a way that the user can identify them from among those of the version adopted and can distinguish them from each other.

## 3. IMPLEMENTATION

This character set should be regarded as a basic alphabet in an abstract sense. Its practical use requires definitions of its implementation in various media. For example, this could include punched tapes, punched cards, magnetic and optical interchangeable media and transmission channels, thus permitting interchange of data to take place either indirectly by means of an intermediate recording on a physical medium, or by local electrical connection of various units (such as input and output devices and computers) or by means of data transmission equipment.

The implementation of this code character set in physical media and for transmission, taking into account the need for error checking, is not covered by this standard.

## 4. TERMINOLOGY

For the purpose of this standard the following definitions apply:

- 4.1 **Active position-** The character position which is to image the graphic symbol representing the next graphic character or relative to which the next control function is to be executed.  
**Note 1:** In general the active position is indicated in a display by a cursor.
- 4.2 **Bit combination-** An ordered set of bits used for the representation of characters
- 4.3 **Character-** A member of a set of elements used for the organization, control or representation of data.

- 4.4 **Character position-** The portion of a display that is imaging or is capable of imaging a graphic symbol.
- 4.5 **Coded character set code-** A set of unambiguous rules that establishes a character set and the one-to-one relationship between the characters of the set and their bit combinations.
- 4.6 **Coded-character-data-element (CC-data-element)-** An element of interchanged information that is specified to consist of a sequence of coded representations of characters in accordance with one or more identified standards for coded character sets.
- Note 2:** In a communication environment according to the Reference Model for Open Systems Interconnection (ISO 7498), a CC-data-element will form all or part of the information that corresponds to the Presentation-Protocol-Data-Unit (PPDU) defined in that International Standard.
- Note 3:** When information interchange is accomplished by means of interchangeable media, a CC-data-element will form all or part of the information that corresponds to the user data and not that recorded during formatting and initialization.
- 4.7 **Code extension-** The techniques for the encoding of characters that are not included in the character set of a given code.
- 4.8 **Code table-** A table showing the character allocated to each bit combination in a code.
- 4.9 **Control character-** A control function the coded representation of which consists of a single bit combination.
- 4.10 **Control function-** An action that affects the recording, processing, transmission, or interpretation of data, and that has a coded representation consisting of one or more bit combinations.
- 4.11 **Device-** A component of information processing equipment which can transmit, and or receive, coded information within CC-data-elements.
- Note 4:** It may be an input/output device in the conventional sense, or a process such as an application program or gateway function.
- 4.12 **Escape sequence-** A string of bit combinations that is used for control purposes in code extension procedures. The first of these bit combinations represents the control function ESCAPE.
- 4.13 **Final Byte-** The bit combination that terminates an escape sequence or a control sequence.
- 4.14 **Graphic character-** A character, other than a control function, that has a visual representation normally handwritten, printed or displayed, and that has a coded representation consisting of one or more bit combinations.
- 4.15 **Graphic symbol -**A visual representation of a graphic character or of a control function.
- 4.16 **Repertoire-** A specified set of characters that are represented by means of one or more bit combinations of a coded character set.
- 4.17 **User-** A person or other entity that invokes the services provided by a device.
- Note 5:** This entity may be a process such as an application program if the 'device' is code converter or a gateway function, for example.
- Note 6:** The characters as supplied by user or made available to him, may be in the form of codes local to the device or of non-conventional visible representation.
- 4.18 **Bangla Characters-** Bangla letters, digits and special symbols which can be grouped as follows:

4.18.1 **Bangla alphabets**

4.18.1.1 **Consonants**

ক খ গ ঘ ঙ চ ছ জ ঝ ঞ ট ঠ ড ঢ ণ ত থ  
দ ধ ন

প ফ ব ভ ম য র ল শ ষ স হ ড় ঢ় য় ং ঃ ঁ

4.18.1.2 Vowels

অ আ ই ঈ উ ঊ ঋ ঌ এ ঐ ও ঔ

4.18.1.3 Bangla Digits

০ ১ ২ ৩ ৪ ৫ ৬ ৭ ৮ ৯

4.18.1.4 Generic addition

ঋ ঌ ঍ ঎ এ ঐ ঑ ঒ ও ঔ ক খ (Ganda)

4.18.1.5 Bangla specific addition

ব ব̂ ট (Bangla Taka sign) | || (Dari) (Double Dari)

5. STANDARD CODE TABLE

The standard 8-bit code table (Table-1) is made up of 16 columns numbered 0 to 15 and 16 rows numbered 0 to 15 containing 256 code positions. Columns 8 to 15 contains 90 graphic characters.

6. EXPLANATORY NOTES

6.1 **Numbering of the position in Table 1-** Within any one character the bits are identified by  $b_8 b_7 \dots b_2 b_1$  where  $b_8$  is the highest order, or most significant bit, and  $b_1$  is the lowest order or least significant bit.

If desired, these may be given numerical significance in the binary system thus:

Bit identification :  $b_8 b_7 b_6 b_5 b_4 b_3 b_2 b_1$

Signification: 128 64 32 16 8 4 2 1

In the table the columns and rows are identified either by bit combination or by its column and row numbers. For instance the position containing the digit (one) may be identified.

- by its bit combination in order of decreasing significance i.e, 1110 0111
- by its column and row numbers, i.e., E/7

The column number is derived from its bits  $b_8 b_7 b_6$  &  $b_5$  giving them weights of 8, 4, 2, & 1 respectively. The row number is from bits  $b_4 b_3 b_2 b_1$  giving them weights 8, 4, 2 & 1 respectively.

6.2 **Code table-** A 8-bit table consists of 256 positions arranged in 16 columns and 16 row. The columns are numbered 0 to 15 and the rows 0 to 15. The code table positions are identified by notations of the form  $x/y$ , where  $x$  is the column number and  $y$  is the row number. The positions of the code table are in one-to-one correspondence with the bit combinations of the code. The notation of a code table position, of the form  $x/y$ , is the same as that of the corresponding bit combination.

6.3 **Names-** This standard assigns at least one name to denote each of the graphic characters displayed in Table 1.

The names chosen to denote graphic characters are intended to reflect their customary meanings. However, this standard does not define and does not restrict the meanings of graphic characters. In addition, it does not specify particular style or font design for the graphic characters.

6.4 **Uniqueness of character Allocation-** A character allocated to a position in Table 1, may not be placed elsewhere in the table. The graphic characters are specified in Table 2.

**Table 1 : Basic Code Table**

Code position for Bangla Characters

				Code position for Bangla Characters																	
				b8	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
				b7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	
				b6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	
				b5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	
					0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	
b4	b3	b2	b1		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0	0	0	0	0	0									ঐ	ঠ	র	ী		ঋ	ষ	
0	0	0	1	1	1								ং	ড		ু		্য	ঞ	ষ	
0	0	1	0	2	2								ং	ঢ	ণ	ু		্য	ঞ	ষ	
0	0	1	1	3	3								ং	ঙ	ণ	ু		্য	ঞ	ষ	
0	1	0	0	4	4									ভ	ত	ু		।	।	।	
0	1	0	1	5	5									অ	ক	খ			॥	২	
0	1	1	0	6	6									আ	খ	দ	শ		০	৩	
0	1	1	1	7	7									ই	গ	ধ	ষ	ে	ী	১	।
1	0	0	0	8	8									ঈ	ষ	ন	স	ে		২	৬
1	0	0	1	9	9									ঊ	ঙ		হ			৩	০
1	0	1	0	A	10									ঋ	ঢ	প				৪	৩
1	0	1	1	B	11									ঋ	জ	ফ		ো		৫	৮
1	1	0	0	C	12									খ	জ	ব	়	ৌ	ড	৬	
1	1	0	1	D	13										ঝ	ড	হ	্	ঢ	৭	
1	1	1	0	E	14										ঞ	ম	া	ৎ		৮	
1	1	1	1	F	15										এ	ট	য	ি		য়	৯

Table 2 : Graphic Character Allocations

Decimal and hexadecimal coded representation for Bangla characters

Graphic Symbol	Name	Code Representation	Hexa Code
◌̣	Bangla sign chandrabindu	8/1	81
◌̣̇	Bangla sign Anusvara	8/2	82
◌̣̈	Bangla sign Visarga	8/3	83
অ	Bangla letter A	8/5	85
আ	Bangla letter AA	8/6	86
ই	Bangla letter I	8/7	87
ঈ	Bangla letter II	8/8	88
উ	Bangla letter U	8/9	89
ঊ	Bangla letter UU	8/10	8A
ঋ	Bangla letter Vocalic R	8/11	8B
ঌ	Bangla letter Vocalic L	8/12	8C
এ	Bangla letter E	8/15	8F
ঐ	Bangla letter AI	9/0	90
ও	Bangla letter O	9/3	93
ঔ	Bangla letter AU	9/4	94
ক	Bangla letter KA	9/5	95
খ	Bangla letter KHA	9/6	96
গ	Bangla letter GA	9/7	97
ঘ	Bangla letter GHA	9/8	98
ঙ	Bangla letter UNGA	9/9	99
চ	Bangla letter CA	9/10	9A
ছ	Bangla letter CHA	9/11	9B
জ	Bangla letter JA	9/12	9C
ঝ	Bangla letter JHA	9/13	9D
ঞ	Bangla letter NYA	9/14	9E
ট	Bangla letter TTA	9/15	9F
ঠ	Bangla letter TTHA	10/0	A0
ড	Bangla letter DDA	10/1	A1
ঢ	Bangla letter DDHA	10/2	A2
ণ	Bangla letter NNA	10/3	A3
ত	Bangla letter TA	10/4	A4



থ	Bangla letter THA	10/5	A5
দ	Bangla letter DA	10/6	A6
ধ	Bangla letter DHA	10/7	A7
ন	Bangla letter NA	10/8	A8
প	Bangla letter PA	10/10	AA
ফ	Bangla letter PHA	10/11	AB
ব	Bangla letter BA	10/12	AC
ভ	Bangla letter BHA	10/13	AD
ম	Bangla letter MA	10/14	AE
য	Bangla letter YA	10/15	AF
র	Bangla letter RA	11/0	B0
ল	Bangla letter LA	11/2	B2
শ	Bangla letter SHA	11/6	B6
ষ	Bangla letter SSA	11/7	B7
স	Bangla letter SA	11/8	B8
হ	Bangla letter HA	11/9	B9
◌	Bangla sign NUKTA	11/12	BC
ং	Bangla sign AVAGRAHA	11/13	BD
া	Bangla Vowel Sign AA	11/14	BE
ি	Bangla Vowel Sign I	11/15	BF
ী	Bangla Vowel Sign II	12/0	C0
ু	Bangla Vowel Sign U	12/1	C1
ূ	Bangla Vowel Sign UU	12/2	C2
্	Bangla Vowel Sign Vocalic R	12/3	C3
্	Bangla Vowel Sign Vocalic RR	12/4	C4
ে	Bangla Vowel Sign E	12/7	C7
ৈ	Bangla Vowel Sign AI	12/8	C8
ো	Bangla Vowel Sign O	12/11	CB
ৌ	Bangla Vowel Sign AU	12/12	CC
্	Bangla Sign Virama	12/13	CD
ৎ	Bangla letter KHANDATA	12/14	CE
া	Bangla AU Length Mark	13/7	D7
ড়	Bangla letter RRA	13/12	DC
ঢ	Bangla letter RHA	13/13	DD

য়	Bangla letter YYA	13/15	DF
৷	Bangla letter Vocalic RR	14/0	E0
৷	Bangla letter Vocalic LL	14/1	E1
◌̆	Bangla Vowel Sign Vocalic L	14/2	E2
◌̆̆	Bangla Vowel Sign Vocalic LL	14/3	E3
	Bangla Daari	14/4	E4
	Bangla Double Daari	14/5	E5
০	Bangla Digit Zero	14/6	E6
১	Bangla Digit One	14/7	E7
২	Bangla Digit Two	14/8	E8
৩	Bangla Digit Three	14/9	E9
৪	Bangla Digit Four	14/10	EA
৫	Bangla Digit Five	14/11	EB
৬	Bangla Digit Six	14/12	EC
৭	Bangla Digit seven	14/13	ED
৮	Bangla Digit Eight	14/14	EE
৯	Bangla Digit Nine	14/15	EF
ঝ	Bangla letter RA with Middle Diagonal	15/0	F0
ঞ	Bangla letter RA with Lower Diagonal	15/1	F1
৳	Bangla Taka Mark	15/2	F2
ট	Bangla Taka sign	15/3	F3
৳	Bangla Currency Numerator One	15/4	F4
৳	Bangla Currency Numerator Two	15/5	F5
৳	Bangla Currency Numerator Three	15/6	F6
	Bangla Currency Numerator Four	15/7	F7
৳	Bangla Currency Numerator One less than the Denominator	15/8	F8
৳	Bangla Currency Denominator Sixteen	15/9	F9
ঈ	Bangla ISSHAR	15/10	FA
৳	Bangla Ganda Mark	15/11	FB