

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

Doc Type: Working Group Document**Title: Towards an encoding of the Kaithi script in the SMP of the UCS****Source: Michael Everson****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2005-11-06**

Anshuman Pandey presented a proposal to encode the Kaithi script in the SMP of the UCS in L2/05-343. The proposal is good, but in my opinion needs some revisions in order to be acceptable for encoding.

1. The proposal uses a one-to-one mapping to the Devanagari block, which is wasteful of code positions, as Kaithi needs no extensions for Dravidian languages and the like. Holes in the encoding should be closed up, and the three nuktaed characters should be inserted into the sequence as Pandey shows in his Figure 5.
2. Evidence shows that, even while no longer used, the letters NGA and NYA existed and should be encoded as historical characters. See Pandey's Figures 6 7, 8, and 13 in L2/05-343.
3. The dandas should be encoded as script-specific punctuation. Recent practice has been consistent in encoding script-specific dandas for minority scripts, and that should apply here. The example in Pandey's Figure 18 shows clearly a script-specific variation in shape (1, ll) which is not normal for Devanagari.
4. The fractions should be decomposed into their six constituent parts. All of the fractions can be written by juxtaposing these six. Indeed, the use of these is simply additive. The base characters are $\surd = 1/16$, $\sphericalangle = 2/16$, $\sphericalright = 3/16$, $\lrcorner = 4/16$, $\llcorner = 5/16$, and $\lll = 6/16$. For the remaining fractions:

$$\begin{array}{lll} \sphericalright/16 = \lrcorner (\sphericalright/16 + 1/16); & \sphericalright/16 = \lrcorner (\sphericalright/16 + 2/16); & \sphericalright/16 = \sphericalright (\sphericalright/16 + 3/16); \\ \sphericalright/16 = \llcorner (\sphericalright/16 + 1/16); & \sphericalright/8 (\sphericalright/16) = \llcorner (\sphericalright/16 + 2/16); & 11/16 = \lll (\sphericalright/16 + 3/16); \\ 13/16 = \lll (\sphericalright/16 + 1/16); & 7/8 (\sphericalright/16) = \lll (\sphericalright/16 + 2/16); & 15/16 = \lll (\sphericalright/16 + 3/16). \end{array}$$

Following here are the Unicode properties, the code table, and the names list as I would recommend for encoding Kaithi. I have supplied character properties for the fractions, which were missing from Pandey's proposal. I have also changed a few of the character names in the last two columns.

```
11000;KAITHI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11001;KAITHI SIGN VISARGA;Mc;0;L;;;;;N;;;;;
11002;KAITHI LETTER A;Lo;0;L;;;;;N;;;;;
11003;KAITHI LETTER AA;Lo;0;L;;;;;N;;;;;
11004;KAITHI LETTER I;Lo;0;L;;;;;N;;;;;
11005;KAITHI LETTER II;Lo;0;L;;;;;N;;;;;
11006;KAITHI LETTER U;Lo;0;L;;;;;N;;;;;
11007;KAITHI LETTER UU;Lo;0;L;;;;;N;;;;;
11008;KAITHI LETTER E;Lo;0;L;;;;;N;;;;;
11009;KAITHI LETTER AI;Lo;0;L;;;;;N;;;;;
1100A;KAITHI LETTER O;Lo;0;L;;;;;N;;;;;
1100B;KAITHI LETTER AU;Lo;0;L;;;;;N;;;;;
1100C;KAITHI LETTER KA;Lo;0;L;;;;;N;;;;;
```

1100D;KAITHI LETTER KHA;Lo;0;L;;;;;N;;;;;
1100E;KAITHI LETTER GA;Lo;0;L;;;;;N;;;;;
1100F;KAITHI LETTER GHA;Lo;0;L;;;;;N;;;;;
11010;KAITHI LETTER NGA;Lo;0;L;;;;;N;;;;;
11011;KAITHI LETTER CA;Lo;0;L;;;;;N;;;;;
11012;KAITHI LETTER CHA;Lo;0;L;;;;;N;;;;;
11013;KAITHI LETTER JA;Lo;0;L;;;;;N;;;;;
11014;KAITHI LETTER JHA;Lo;0;L;;;;;N;;;;;
11015;KAITHI LETTER NYA;Lo;0;L;;;;;N;;;;;
11016;KAITHI LETTER TTA;Lo;0;L;;;;;N;;;;;
11017;KAITHI LETTER TTHA;Lo;0;L;;;;;N;;;;;
11018;KAITHI LETTER DDA;Lo;0;L;;;;;N;;;;;
11019;KAITHI LETTER DDDHA;Lo;0;L;;;;;N;;;;;
1101A;KAITHI LETTER DDHA;Lo;0;L;;;;;N;;;;;
1101B;KAITHI LETTER RHA;Lo;0;L;;;;;N;;;;;
1101C;KAITHI LETTER NNA;Lo;0;L;;;;;N;;;;;
1101D;KAITHI LETTER TA;Lo;0;L;;;;;N;;;;;
1101E;KAITHI LETTER THA;Lo;0;L;;;;;N;;;;;
1101F;KAITHI LETTER DA;Lo;0;L;;;;;N;;;;;
11020;KAITHI LETTER DHA;Lo;0;L;;;;;N;;;;;
11021;KAITHI LETTER NA;Lo;0;L;;;;;N;;;;;
11022;KAITHI LETTER PA;Lo;0;L;;;;;N;;;;;
11023;KAITHI LETTER PHA;Lo;0;L;;;;;N;;;;;
11024;KAITHI LETTER BA;Lo;0;L;;;;;N;;;;;
11025;KAITHI LETTER BHA;Lo;0;L;;;;;N;;;;;
11026;KAITHI LETTER MA;Lo;0;L;;;;;N;;;;;
11027;KAITHI LETTER YA;Lo;0;L;;;;;N;;;;;
11028;KAITHI LETTER YYA;Lo;0;L;;;;;N;;;;;
11029;KAITHI LETTER RA;Lo;0;L;;;;;N;;;;;
1102A;KAITHI LETTER LA;Lo;0;L;;;;;N;;;;;
1102B;KAITHI LETTER VA;Lo;0;L;;;;;N;;;;;
1102C;KAITHI LETTER SHA;Lo;0;L;;;;;N;;;;;
1102D;KAITHI LETTER SSA;Lo;0;L;;;;;N;;;;;
1102E;KAITHI LETTER SA;Lo;0;L;;;;;N;;;;;
1102F;KAITHI LETTER HA;Lo;0;L;;;;;N;;;;;
11030;KAITHI VOWEL SIGN AA;Mc;0;L;;;;;N;;;;;
11031;KAITHI VOWEL SIGN I;Mc;0;L;;;;;N;;;;;
11032;KAITHI VOWEL SIGN II;Mc;0;L;;;;;N;;;;;
11033;KAITHI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11034;KAITHI VOWEL SIGN UU;Mn;0;NSM;;;;;N;;;;;
11035;KAITHI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
11036;KAITHI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
11037;KAITHI VOWEL SIGN O;Mc;0;L;;;;;N;;;;;
11038;KAITHI VOWEL SIGN AU;Mc;0;L;;;;;N;;;;;
11039;KAITHI SIGN VIRAMA;Mn;9;NSM;;;;;N;;;;;
1103A;KAITHI RUPAYA SIGN;Sc;0;ET;;;;;N;;;;;
1103B;KAITHI WORD SEPARATOR;Po;0;L;;;;;N;;;;;
1103C;KAITHI SENTENCE SEPARATOR;Po;0;L;;;;;N;;;;;
1103D;KAITHI ABBREVIATION SIGN;Po;0;L;;;;;N;;;;;
1103E;KAITHI DANDA;Po;0;L;;;;;N;;;;;
1103F;KAITHI DOUBLE DANDA;Po;0;L;;;;;N;;;;;
11040;KAITHI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11041;KAITHI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11042;KAITHI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11043;KAITHI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11044;KAITHI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11045;KAITHI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11046;KAITHI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11047;KAITHI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11048;KAITHI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11049;KAITHI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
1104A;KAITHI FRACTION ONE SIXTEENTH;No;0;L;;;0.0625;N;;;;;
1104B;KAITHI FRACTION ONE EIGHTH;No;0;L;;;0.125;N;;;;;
1104C;KAITHI FRACTION THREE SIXTEENTHS;No;0;L;;;0.1875;N;;;;;
1104D;KAITHI FRACTION ONE QUARTER;No;0;L;;;0.25;N;;;;;
1104E;KAITHI FRACTION ONE HALF;No;0;L;;;0.5;N;;;;;
1104F;KAITHI FRACTION THREE QUARTERS;No;0;L;;;0.75;N;;;;;

TABLE XXX - Row 110: KAITHI

	1100	1101	1102	1103	1104
0	◌̇	᳚	᳛	᳜	᳝
1	◌̇:	᳞	᳟	᳠	᳡
2	᳢	᳣	᳤	᳥	᳦
3	᳧	᳨	ᳩ	ᳪ	ᳫ
4	ᳬ	᳭	ᳮ	ᳯ	ᳰ
5	ᳱ	ᳲ	ᳳ	᳴	ᳵ
6	ᳶ	᳷	᳸	᳹	ᳺ
7	᳻	᳼	᳽	᳾	᳿
8	᳾	᳿	ᳺ	᳻	᳼
9	᳾	᳿	ᳺ	᳻	᳼
A	᳾	᳿	ᳺ	᳻	᳼
B	᳾	᳿	ᳺ	᳻	᳼
C	᳾	᳿	ᳺ	᳻	᳼
D	᳾	᳿	ᳺ	᳻	᳼
E	᳾	᳿	ᳺ	᳻	᳼
F	᳾	᳿	ᳺ	᳻	᳼

G = 00
P = 01

TABLE XXX - Row 110: KAITHI

hex	Name
11000	KAITHI SIGN ANUSVARA
11001	KAITHI SIGN VISARGA
11002	KAITHI LETTER A
11003	KAITHI LETTER AA
11004	KAITHI LETTER I
11005	KAITHI LETTER II
11006	KAITHI LETTER U
11007	KAITHI LETTER UU
11008	KAITHI LETTER E
11009	KAITHI LETTER AI
1100A	KAITHI LETTER O
1100B	KAITHI LETTER AU
1100C	KAITHI LETTER KA
1100D	KAITHI LETTER KHA
1100E	KAITHI LETTER GA
1100F	KAITHI LETTER GHA
11010	KAITHI LETTER NGA
11011	KAITHI LETTER CA
11012	KAITHI LETTER CHA
11013	KAITHI LETTER JA
11014	KAITHI LETTER JHA
11015	KAITHI LETTER NYA
11016	KAITHI LETTER TTA
11017	KAITHI LETTER TTHA
11018	KAITHI LETTER DDA
11019	KAITHI LETTER DDDHA
1101A	KAITHI LETTER DDHA
1101B	KAITHI LETTER RHA
1101C	KAITHI LETTER NNA
1101D	KAITHI LETTER TA
1101E	KAITHI LETTER THA
1101F	KAITHI LETTER DA
11020	KAITHI LETTER DHA
11021	KAITHI LETTER NA
11022	KAITHI LETTER PA
11023	KAITHI LETTER PHA
11024	KAITHI LETTER BA
11025	KAITHI LETTER BHA
11026	KAITHI LETTER MA
11027	KAITHI LETTER YA
11028	KAITHI LETTER YYA
11029	KAITHI LETTER RA
1102A	KAITHI LETTER LA
1102B	KAITHI LETTER VA
1102C	KAITHI LETTER SHA
1102D	KAITHI LETTER SSA
1102E	KAITHI LETTER SA
1102F	KAITHI LETTER HA
11030	KAITHI VOWEL SIGN AA
11031	KAITHI VOWEL SIGN I
11032	KAITHI VOWEL SIGN II
11033	KAITHI VOWEL SIGN U
11034	KAITHI VOWEL SIGN UU
11035	KAITHI VOWEL SIGN E
11036	KAITHI VOWEL SIGN AI
11037	KAITHI VOWEL SIGN O
11038	KAITHI VOWEL SIGN AU
11039	KAITHI SIGN VIRAMA
1103A	KAITHI RUPAYA SIGN
1103B	KAITHI WORD SEPARATOR
1103C	KAITHI SENTENCE SEPARATOR
1103D	KAITHI ABBREVIATION SIGN
1103E	KAITHI DANDA
1103F	KAITHI DOUBLE DANDA
11040	KAITHI DIGIT ZERO
11041	KAITHI DIGIT ONE
11042	KAITHI DIGIT TWO
11043	KAITHI DIGIT THREE
11044	KAITHI DIGIT FOUR
11045	KAITHI DIGIT FIVE
11046	KAITHI DIGIT SIX
11047	KAITHI DIGIT SEVEN
11048	KAITHI DIGIT EIGHT
11049	KAITHI DIGIT NINE
1104A	KAITHI FRACTION ONE SIXTEENTH
1104B	KAITHI FRACTION ONE EIGHTH
1104C	KAITHI FRACTION THREE SIXTEENTHS
1104D	KAITHI FRACTION ONE QUARTER
1104E	KAITHI FRACTION ONE HALF
1104F	KAITHI FRACTION THREE QUARTERS

Group 00

hex	Name
-----	------

Plane 01

Row 110