Proposal to Encode the Masaram Gondi Script in Unicode

Anshuman Pandey
Department of Linguistics
University of California, Berkeley
Berkeley, California, U.S.A.
anshuman.pandey@berkeley.edu

April 15, 2015

1   Introduction

This is a proposal to encode the Gondi script created by Mangal Singh Masaram in 1918 in Unicode. It replaces and supersedes the following documents:

- L2/10-207 "Preliminary Proposal to Encode the Gondi Script in the UCS"
- L2/12-235 "Revised Preliminary Proposal to Encode the Gondi Script"
- L2/15-005 "Proposal to Encode the Gondi Script"

This document provides a description of the writing system, a code chart and names list, character properties, and specimens that illustrate letterforms and usage. It is a revision of L2/15-005 and contains several changes to the encoding of Gondi proposed in that document. The major changes are as follows:

- The block name has been changed from 'Gondi' to 'Masaram Gondi'
- Separate encoding of REPHA and RA-KARA for cluster-initial and cluster-final forms of RA
- Redefinition of VIRAMA specifically as a control character used only for producing conjuncts
- Addition of HALANTA as a combining sign used solely for silencing the inherent vowel
- Addition of CANDRA for transcribing foreign vowel sounds

These changes were introduced as a result of discussions with experts such as Mukund Gokhale, members of the user community and with Unicode engineers at Google and Microsoft.

The 'Masaram Gondi' script is graphically and structurally distinct from another Gondi script known as 'Gunjala Gondi'. This latter script has been proposed for encoding in Unicode (see Pandey 2015b).

The symbol appears upon several sources containing Masaram's script and in various Gond religious contexts, in general. This symbol represents *persapen*, or the supreme spirit, in the indigenous Gond philosophical system known as *koya punem*. This symbol has been proposed for encoding in the 'Miscellaenous Symbols and Pictographs' block in Unicode, where several religious symbols are encoded (see Pandey 2015c).

2   Background

The script described here was invented by Munshi Mangal Singh Masaram of Kochewada, Balaghat District, Madhya Pradesh, India in 1918. It has no genetic relationship to other writing systems, but it is based upon the Brahmi model. The script was designed for writing Gondi (ISO 639-3: gon), a Dravidian language spoken by 2.6 million people, primarily in Madhya Pradesh and Maharashtra, with some speakers in Andhra Pradesh and Chhattisgarh. The language is generally written in both Devanagari and Telugu. Manuscripts containing yet another script, which appears to have been graphically inspired by these Modi writing system, was found in Gunjala in the Adilabad district of Telangana Masaram's Gondi script is actively used today for hand-written and printed materials. Fonts have been developed for the production of books. In 2011, the Akhil Gondvana Gondi Sahitya Parishad (Chandagadh, Maharashtra) passed a resolution adopting Masaram's script as the official script of the Gondi language.

Masaram's script has been slightly expanded and revised over the years in order to meet the needs and preferences of modern users. Innovations include the addition of new consonant letters, vowel signs, modifier signs, and the adoption of a Devanagari-style *halanta* for indicating the absence of the inherent vowel. Some of these new characters are included in the proposed repertoire, while others are not (see section 4.11).

3   Script Details

3.1   Name

Earlier versions of this proposal referred to name of the script block as 'Gondi'. While the script is certainly used for writing Gondi, it is one of many scripts used for the language. Moreover, the 'Gunjala Gondi' script is also associated with the language and culture of the Gonds. Given this, it is appropriate to assign an identifer for the script block that precisely defines which 'Gondi' script is contained within that block. A designation that includes the name of the script's creator seems appropriate. For this reason, the name of the script block in Unicode is 'Masaram Gondi' and the names of characters reflect the block name. Users may refer to the script as 'Gondi' or by whatever name they prefer.

3.2   Structure

Masaram's Gondi script is an alphasyllabic writing system that is written from left to right. Consonant letters possess the inherent vowel *a*, which is graphically represented by a horizontal stroke that extends rightward from the right edge of each consonant letter. A bare consonant is represented by removing this stroke, particularly in the rendering of consonant clusters. Some modern users represent a word-final bare consonant by writing the *halanta* beneath the stroke of the consonant letter. Consonant clusters are represented as conjuncts, which are rendered as a linear sequence using bare forms for all consonants except for the final, which occurs in its regular form. There are some exceptions to this rule, namely the forms of *ra* when in occurs in initial and final positions in a cluster. Independent and initial vowels are written using vowel letters, while medial and final vowels are expressed using dependent signs. There is no *mātrā* reordering.

3.3   Character Repertoire

A total of 75 characters are proposed for encoding in the 'Masaram Gondi' script block. A code chart and names list are attached. Names for characters follow the UCS convention for Brahmi-based scripts and align with the Latin transliteration of Devanagari analogues for Gondi letters given by B. S. Masaram (1951).

3.4 Glyphic Representations

The glyphic representations of some consonant letters have changed since the invention of the script. The majority of changes are observed with consonant letters. These differences result from the simplification of glyphs for ease of writing, ie. sets of independent circles being joined into a single-stroked loop. Representative glyphs are based upon forms shown in published script primers and reflect modern preferences.

4 Proposed Encoding

4.1 Vowel Letters

Ten vowel letters are proposed for encoding:

∩ˡ	MASARAM GONDI LETTER A	ᵾ	MASARAM GONDI LETTER UU
∩ᴱ	MASARAM GONDI LETTER AA	∟	MASARAM GONDI LETTER E
⊏	MASARAM GONDI LETTER I	╟	MASARAM GONDI LETTER AI
⌐ˡ	MASARAM GONDI LETTER II	∩ᴱ	MASARAM GONDI LETTER O
⊔ˡ	MASARAM GONDI LETTER U	∩ᴱ	MASARAM GONDI LETTER AU

Masaram's script does not have independent letters or dependent signs for the Dravidian long vowels /e:/ and /o:/, which correspond to Telugu ఏ *ē* and ఓ *ō*.

4.2 Vowel Signs

Ten dependent vowel signs are proposed for encoding:

○¯	MASARAM GONDI VOWEL SIGN AA	○ᵾ	MASARAM GONDI VOWEL SIGN VOCALIC R
○ᶠ	MASARAM GONDI VOWEL SIGN I	○ˡ	MASARAM GONDI VOWEL SIGN E
○ⁿ	MASARAM GONDI VOWEL SIGN II	○ⁱⁱ	MASARAM GONDI VOWEL SIGN AI
○ᵁ	MASARAM GONDI VOWEL SIGN U	○⊥	MASARAM GONDI VOWEL SIGN O
○ᵖ	MASARAM GONDI VOWEL SIGN UU	○ᵚ	MASARAM GONDI VOWEL SIGN AU

Vowel signs are written above and below the horizontal stroke of a consonant letter:

0⁻	0=	0ᶠ	0ⁿ	0ᵁ	0ᵖ	0ᵤ	0⊥	0ⁱⁱ	0±	0ᵚ
ka	*kā*	*ki*	*kī*	*ku*	*kū*	*kṛ*	*ke*	*kai*	*ko*	*kau*

They are represented in encoded text as follows:

kā	0=	<0⁻ KA, ○¯ SIGN AA>
ki	0ᶠ	<0⁻ KA, ○ᶠ SIGN I>

kī	O᷄	<O‑ KA, ○᷄ SIGN II>
ku	O᷄	<O‑ KA, ○᷄ SIGN U>
kū	O᷄	<O‑ KA, ○᷄ SIGN UU>
kṛ	O᷄	<O‑ KA, ○᷄ SIGN VOCALIC R>
ke	O᷄	<O‑ KA, ○᷄ SIGN E>
kai	O᷄	<O‑ KA, ○᷄ SIGN AI>
ko	O᷄	<O‑ KA, ○᷄ SIGN O>
kau	O᷄	<O‑ KA, ○᷄ SIGN AU>

There is no independent analogue for VOWEL SIGN VOCALIC R. The independent form of this vowel is represented using a consonant-vowel combination composed with the letter RA:

ṛ	()᷄	<()‑ RA, ○᷄ VOWEL SIGN VOCALIC R>

4.3 Consonants

There are 34 consonant letters:

O‑	MASARAM GONDI LETTER KA	ϑ‑	MASARAM GONDI LETTER DA
8‑	MASARAM GONDI LETTER KHA	ϙ‑	MASARAM GONDI LETTER DHA
Ⱶ‑	MASARAM GONDI LETTER GA	ω‑	MASARAM GONDI LETTER NA
Ɛ‑	MASARAM GONDI LETTER GHA	H‑	MASARAM GONDI LETTER PA
ϑ‑	MASARAM GONDI LETTER NGA	HH‑	MASARAM GONDI LETTER PHA
Ɔ‑	MASARAM GONDI LETTER CA	Ⱶ‑	MASARAM GONDI LETTER BA
3̄‑	MASARAM GONDI LETTER CHA	Ꙍ‑	MASARAM GONDI LETTER BHA
ϑ‑	MASARAM GONDI LETTER JA	Ⱶ‑	MASARAM GONDI LETTER MA
M‑	MASARAM GONDI LETTER JHA	Ϲ‑	MASARAM GONDI LETTER YA
Ƒ‑	MASARAM GONDI LETTER NYA	()‑	MASARAM GONDI LETTER RA
Ⴑ‑	MASARAM GONDI LETTER TTA	∩‑	MASARAM GONDI LETTER LA
+‑	MASARAM GONDI LETTER TTHA	8‑	MASARAM GONDI LETTER VA
ϑ‑	MASARAM GONDI LETTER DDA	Ⅴ‑	MASARAM GONDI LETTER SHA
∩‑	MASARAM GONDI LETTER DDHA	θ‑	MASARAM GONDI LETTER SSA
Ⴟ‑	MASARAM GONDI LETTER NNA	Ɛ‑	MASARAM GONDI LETTER SA
Ս‑	MASARAM GONDI LETTER TA	ω‑	MASARAM GONDI LETTER HA
✕‑	MASARAM GONDI LETTER THA	ꝏ‑	MASARAM GONDI LETTER LLA

The letter ꝏ‑ LLA is not part of Masaram's original script. It was introduced by modern users in order to properly represent Marathi ळ *ḷa* (see figure 19 for an example of LLA in usage).

4.4   Vowel modifiers

Three vowel modifiers are proposed for encoding:

    ◌˙   MASARAM GONDI SIGN ANUSVARA

    ◌¨   MASARAM GONDI SIGN VISARGA

    ◌˘   MASARAM GONDI SIGN CANDRA

Anusvara   The sign ◌˙ ANUSVARA is used for marking nasalization. It is placed above the horizontal stroke of a consonant. Its position differs slightly in different sources. In some documents the position is altered by the presence of an accompanying above-base vowel sign:

As shown above, the sign ANUSVARA occurs in its normal position when there is no vowel sign or the vowel sign is positioned below the stroke. Its position is raised when ◌¯ VOWEL SIGN AA is present. It is placed to the right of the following: ◌ᶠ VOWEL SIGN I, ◌ᴺ VOWEL SIGN II, ◌ᵞ VOWEL SIGN U, ◌ᵚ VOWEL SIGN UU, ◌ᴵ VOWEL SIGN E, ◌ᴵᴵ VOWEL SIGN AI. It is placed to the right and raised higher when it occurs with ◌ᴸ VOWEL SIGN O, ◌ᴸᴸ VOWEL SIGN AU. Some modern users prefer a more stationary position for ANUSVARA and place it above the body of the consonant:

These positional preferences are to be managed in the font. The sign ANUSVARA is used in encoded text as shown below. It is always placed after a vowel sign in the encoded sequence:

    kaṃ   ◌˙   <◌ KA, ◌˙ ANUSVARA>

    kāṃ   ◌˙   <◌ KA, ◌¯ VOWEL SIGN AA, ◌˙ ANUSVARA>

    kīṃ   ◌˙   <◌ KA, ◌ᴺ VOWEL SIGN II, ◌˙ ANUSVARA>

    kṛ̥ṃ   ◌˙   <◌ KA, ◌ᵤ VOWEL SIGN VOCALIC R, ◌˙ ANUSVARA>

Visarga   The ◌¨ VISARGA is used for the representation of Sanskrit words. It is written above the horizontal line of a consonant letter. When occurring with vowel signs its position is adjusted as follows:

Some modern sources show the ◌¨ VISARGA written as the glyphic variant ◌፥ . It is placed after the stroke:

The VISARGA is used in encoded text as follows:

    kaḥ   Oˑˑ   <O‑ KA, ◌̈ VISARGA>

    kāḥ   Oˑˑ   <O‑ KA, ◌⁻ VOWEL SIGN AA, ◌̈ VISARGA>

Candra   The sign ◌̆ CANDRA is used for transcribing vowel sounds that do not occur in Gondi. Examples are given in figure 25. It is derived from the sign ◌̆ used in Devanagari orthography for Marathi for representing the English vowel sounds /æ/ and /ɔ/. The CANDRA is written above the horizontal line of a consonant letter, and is used as follows:

    /æ/   Ľ   <L E, ◌̆ CANDRA>

    /ɔ/   ∩̆   <∩ AA, ◌̆ CANDRA>

    /kæ/   Ŏ   <O‑ KA, ◌̆ CANDRA>

    /kɔ/   Ŏ   <O‑ KA, ◌⁻ VOWEL SIGN AA, ◌̆ SIGN CANDRA>

In initial and independent contexts, CANDRA is used only with the vowel letters E and O for representing /æ/ and /ɔ/, respectively. The combination Ľ corresponds to ऍ U+090D DEVANAGARI LETTER CANDRA E, while ∩̆ corresponds to ऑ U+0911 DEVANAGARI LETTER CANDRA O. In dependent contexts, the CANDRA combines with a consonant letter for the sound /æ/, and with the VOWEL SIGN AA for the sound /ɔ/. The ◌̆ CANDRA corresponds to ◌ॅ U+0945 DEVANAGARI VOWEL SIGN CANDRA E, while the sequence <◌⁻ GONDI VOWEL SIGN AA, ◌̆ CANDRA> corresponds to ◌ॉ U+0949 DEVANAGARI VOWEL SIGN CANDRA O. Although the Gondi sequences correspond to atomic characters in Devanagari, there is no need to encode such precomposed letters and signs with CANDRA for Gondi.

Although the contexts in which CANDRA is used are limited, the sign technically may be used with any letter and may occur with any vowel sign. It would be positioned with other vowel signs as follows:

    Ŏ   Ŏ   Ŏ   Ŏ   Ŏ   Ŏ   Ŏ   Ŏ   Ŏ   Ŏ   Ŏ

4.5   Nukta

The ◌. MASARAM GONDI SIGN NUKTA is used for representing sounds that are not native to the Gondi language. It is written beneath the horizontal stroke of a consonant:

    O̤   O̤   O̤   O̤   O̤   O̤   O̤   O̤   O̤   O̤   O̤

Some users prefer to position the NUKTA below the body of the consonant letter:

    O̩   O̩   O̩   O̩   O̩   O̩   O̩   O̩   O̩   O̩   O̩

These positional preferences are to be managed in the font. The NUKTA is used in encoded text as follows:

    ṛa   ꙮ   <ꙮ DDA, ◌ NUKTA>

    ṛā   ꙮ   <ꙮ DDA, ◌ NUKTA, ◌ SIGN AA>

4.6   Halanta

The ◌ MASARAM GONDI SIGN HALANTA is proposed as a vowel silencer. It is used as follows:

    k   ꙮ   <ꙮ KA, ◌ HALANTA>

    kh   ꙮ   <ꙮ KHA, ◌ HALANTA>

The Gondi script as designed by Masaram does not have a native *halanta*, as the structure of the script does not require it. The embedding of the inherent vowel into the graphical structure of a consonant is a unique and innovative feature of the Gondi script. The horizontal stroke of each consonant letter represents the inherent vowel; removal of this stroke produces a bare consonant: ꙮ *ka* → ꙮ *k*, etc.

In most Indic scripts the inherent vowel is not part of the graphical structure of a consonant letter. As a result, these scripts require a mechanism for indicating the absence of the inherent vowel. In Devanagari this mechanism is a sign called ◌ U+094D DEVANAGARI SIGN VIRAMA: क *ka* + ◌ → क् *k*, etc.

However, modern users have adopted the Devanagari *halanta* (or VIRAMA) for marking a bare consonant at the end of a word (see figure 24). The GONDI SIGN HALANTA has been included in the repertoire in order to provide this functionality in the proposed encoding.

The proposed encoding for Gondi separates the two functions of the VIRAMA character as used in the models for most Indic scripts in Unicode. In Devanagari, for example, the VIRAMA functions both as a vowel silencer (*halanta*) and a control character for forming conjuncts. The default representation of VIRAMA in Devanagari is as a visible sign beneath the consonant with which it combines. If a consonant is placed after the VIRAMA it causes a conjunct to be formed from the two consonants around it. In order to display a visible VIRAMA between adjacent consonants, it is necessary to break conjunct formation. For this purpose, the control character ⌷ U+200C ZERO WIDTH NON-JOINER is placed after VIRAMA. This approach was proposed for Gondi in L2/15-005. However, in the interest of developing a simple encoding model for Gondi, and to eliminate the need for usage of ZERO WIDTH NON-JOINER or other invisible control characters, the encoding of the vowel silencing feature of the Indic VIRAMA has been encapsulated into the character ◌ GONDI SIGN HALANTA. The conjunct forming function is retained in the GONDI VIRAMA, described below.

4.7   Virama

The ◌ MASARAM GONDI VIRAMA is a control character that is used specifically for producing the bare form of a consonant letter. It is represented in the code chart as ⌷ in order to indicate that it is a special character. Conceptually, VIRAMA produces a half-form by removing the horizontal stroke from the glyph of the letter after which it is placed.

    k   ꙮ   <ꙮ KA, ◌ VIRAMA>

    kh   ꙮ   <ꙮ KHA, ◌ VIRAMA>

It is used for producing conjuncts, similar to the control function of ◌ U+094D DEVANAGARI SIGN VIRAMA. The Gondi VIRAMA, however, does not silence the inherent vowel; the ◌ GONDI SIGN HALANTA is to be used for that purpose.

4.8   Consonant Conjuncts

Consonant clusters are represented as conjuncts and are rendered in a linear sequence using half-forms of all but the final letter in a cluster, which appears in its full form, eg. OO‑ *kka*, OU‑ *kta*, OU8‑ *ktva*, etc. Consonants are placed sequentially in the conjunct in the order that they occur in the cluster.

Conjuncts are represented in encoded text by placing the control character ◌ MASARAM GONDI VIRAMA after each non-initial consonant in a cluster. The sequence <(*C*, ◌ VIRAMA)*, *C*$_f$> produces a half-form of all *C* and the regular full-form of *C*$_f$:

ka	O‑	<O‑ KA>
kta	OU‑	<O‑ KA, ◌ VIRAMA, U‑ TA>
ktva	OU8‑	<O‑ KA, ◌ VIRAMA, U‑ TA, ◌ VIRAMA, 8‑ VA>
ktvya	OU8C‑	<O‑ KA, ◌ VIRAMA, U‑ TA, ◌ VIRAMA, 8‑ VA, ◌ VIRAMA, C‑ YA>

There are some exceptions to the rule of conjunct formation. The following character are proposed in order to properly represent all Gondi conjuncts:

⌐n⌐	MASARAM GONDI REPHA
◌ı	MASARAM GONDI RA-KARA
M‑	MASARAM GONDI LETTER KSSA
ᎧⲎ‑	MASARAM GONDI LETTER JNYA
ᴈ‑	MASARAM GONDI LETTER TRA

The behavior of RA is described in section 4.8.1, and the use of atomic ligatures for three conjuncts is discussed in see section 4.8.2).

4.8.1   Forms of RA in conjuncts

Following the general rule of conjunct formation, the letter O‑ RA would occur as the half-form O when it is initial or medial in a cluster and in its regular full-form when in cluster-final position. The Gondi RA, however, does not behave entirely as expected in conjuncts and is rendered in several ways. There are three ways of writing O‑ RA in conjuncts. It occurs in its half-form O when cluster-initial or alternately as ◌ⁿ when cluster-initial and as ◌ı when cluster-final. These are described below:

- *Half-form*   The half-form of RA is used specifically for representing semantic distinctions of RA when the letter occurs at a morphological boundary. Such usage is influenced by Devanagari orthography for the Marathi language, in which र RA may occur as either the ◌̍ regular *repha* or the ◌ 'eyelash' *repha* when it is the initial consonant in a cluster; the 'eyelash' *repha* marks plural suffixes (दऱ्या *daryā* 'valleys' and दर्या *daryā* 'ocean') and inflectional suffixes (आचाऱ्यास *ācāryās* 'to the cook' and आचार्यास *ācāryās* 'to the teacher'). The half-form O corresponds to the Devanagari ◌ 'eyelash' *repha*. It is also used when RA occurs in cluster-medial position.

- *Repha*   When not used for marking morphological distinctions, cluster-initial RA is rendered as the ◌ⁿ *repha*. The logic of this character is based upon the Devanagari ◌̍ regular *repha*. The Gondi *repha*

attaches after the last letter in a conjunct, above the horizontal line or an extension of it, depending upon the presence of an above-stroke vowel sign, eg. 0ᴾ *rka*, 0═ᴾ *rkā*, etc.

0ᴾ   0═ᴾ   0ᴸᴾ   0ᴶᴾ   0ᴴᴾ   0ᴴᴾ   0ᴃ   0ᴸᴾ   0ᵐᴾ   0ᴸᴾ   0ᵐᴾ

Some modern sources show the ◌ᶯ *repha* represented using the form ◌ᴿ, eg. 0ᴿ *rka*. This form is simply the regular sign for ◌ᶯ *repha* with the left stroke drawn past the horizontal bar and curving to the right. It is a glyphic variant and is to be handled by the font.

- *Ra-kāra* In the current orthography, cluster-final ʀᴀ is rendered as ◌ı *ra-kāra* instead of as the full form 0⊢. The logic of the Gondi *ra-kāra* is based upon the Devanagari ◌ᵥ *ra-kāra* and ◌ *vattu*. The Gondi *ra-kāra* is positioned below the horizontal line of a consonant glyph or beneath an extension of the horizontal line:

0ₜ   0ₜ   0f   0ᴶ   0ᴴ   0ᴴ   0ᴛᴜ   0ᵻ   0ᵐ   0ᵻ   0ᵻ

Some modern sources show the ◌ı *ra-kāra* represented using the form ◌ı, eg. 0ᴦ *kra*. This form is a glyphic variant and is to be handled by the font.

- *Full-form* The full-form of ʀᴀ is rarely used at present when it is final in a cluster: 00⊢ *kra*. The preference is to use ◌ı *ra-kāra*.

The representation of ◌ᶯ *repha* and ◌ı *ra-kāra* requires an exception to the rule of conjunct formation in Gondi. The general rule states that the sequence <ʀᴀ, ᴠɪʀᴀᴍᴀ, *C*> is rendered using the half-form 0 of ʀᴀ and the full-form of *C*. It also states that <*C*, ᴠɪʀᴀᴍᴀ, ʀᴀ> would produce the half-form of *C* and the full-form 0⊢ of ʀᴀ. For this reason another method is required for the encoded representation of *repha* and *ra-kāra*, for which the expected encoded sequences according to the general model of Indic scripts would also be <ʀᴀ, ᴠɪʀᴀᴍᴀ, *C*> and <*C*, ᴠɪʀᴀᴍᴀ, ʀᴀ>, respectively.

There are four possible models for accommodating encoded representations of ʀᴀ in conjuncts. The first two are based upon the premise that the default behavior of ʀᴀ in conjuncts is similar to that of all other consonants, ie. it is rendered using the half-form 0 when cluster-initial and the full-form 0⊢ when cluster-final. These approaches treat *repha* and *ra-kāra* as exceptions. The third model deviates from this premise and establishes the *repha* and *ra-kāra* as default representations of ʀᴀ in conjuncts, and the half-form and full-form as exceptions.

1. *Use the Zero-Width Joiner*

   In L2/15-005, it was suggested that the generic control character ⬚ᶻʷ U+200D ᴢᴇʀᴏ ᴡɪᴅᴛʜ ᴊᴏɪɴᴇʀ (ᴢᴡᴊ) be used for representing *repha* and *ra-kāra* in encoded text. The usage of ᴢᴡᴊ was chosen because the character is used in various Indic scripts for controlling different forms of letters in conjuncts. The same principle was applied to Gondi, such that various forms of ʀᴀ in conjuncts would be produced as follows:

   half-form   0   <0⊢ ʀᴀ, ◌ ᴠɪʀᴀᴍᴀ, *C*>

   repha   ◌ᶯ   <0⊢ ʀᴀ, ⬚ᶻʷ ᴢᴇʀᴏ ᴡɪᴅᴛʜ ᴊᴏɪɴᴇʀ, ◌ ᴠɪʀᴀᴍᴀ, *C*>

   ra-kāra   ◌ı   <*C*, ◌ ᴠɪʀᴀᴍᴀ, ⬚ᶻʷ ᴢᴇʀᴏ ᴡɪᴅᴛʜ ᴊᴏɪɴᴇʀ, 0⊢ ʀᴀ>

Representation of approach #1 in encoded sequences would be as follows:

    ◌०–   *rka*   <◌– RA, ◌ VIRAMA, ०– KA>

    ०◌–   *kra*   <०– KA, ◌ VIRAMA, ◌– RA>

    ०ᵖ   *rka*   <◌– RA, ⬚ᶻ⪽ ZERO WIDTH JOINER, ◌ VIRAMA, ०– KA>

    ०ᴛ   *kra*   <०– KA, ◌ VIRAMA, ⬚ᶻ⪽ ZERO WIDTH JOINER, ◌– RA>

The Script Ad-hoc Committee recommended against the usage of ZWJ for such cases in Gondi. In L2/15-045, the Subcommittee suggested that an alternative model be developed and that "[f]or *repha* and *ra-kāra*, encoding separate characters may be useful, similar to the Malayalam *dot reph* and the *medial ra* in Myanmar and Tai Tham." The major concern is that usage of ZWJ introduces issues for both implementers and end users. For implementers, the use of ZWJ overloads the script-specific rules that must be tailored for each script. For end-users, ZWJ is problematic because it is an invisible control character and is neither readily available on keyboards or easily detectable in text sequence.

2. *Redefine the rule for rendering <RA, VIRAMA>*

The above model is based on the general rule of conjunct formation in Gondi: <*C*, VIRAMA> produces the half-form of *C* and therefore <RA, VIRAMA> should produce the half-form of RA. The *repha* and *ra-kāra* are treated as exceptions to the rule. A third approach for handling the different forms of RA in conjuncts is to redefine the default rendering for RA. With this approach the various forms of RA in conjuncts would be produced in a fashion similar to other Indic scripts:

    repha   ◌ᶰ   <◌– RA, ◌ VIRAMA, *C*>

    ra-kāra   ◌ı   <*C*, ◌ VIRAMA, ◌– RA>

Representation of approach #2 in encoded sequences would be as follows:

    ०ᵖ   *rka* <◌– RA, ◌ VIRAMA, ०– KA>

    ०ᴛ   *kra* <०– KA, ◌ VIRAMA, ◌– RA>

This approach, however, does not provide a means for producing the half-form ◌ of RA, or even the cluster-final full-form if ever needed. Some mechanism would be required to produce these forms. Inevitably, a control character such as ZWJ would be required:

    ◌०–   *rka* <◌– RA, ⬚ᶻ⪽ ZERO WIDTH JOINER, ◌ VIRAMA, ०– KA>

    ०◌–   *kra* <०– KA, ◌ VIRAMA, ⬚ᶻ⪽ ZERO WIDTH JOINER, ◌– RA>

This approach is essentially the converse of what was proposed in L2/15-005. It poses the same issues as it requires usage of ZWJ.

3. *Encode a ligating form of RA*

Another option is to encode a dummy letter whose cluster-initial form is ◌ᶰ *repha* and whose non-initial form is ◌ı *ra-kāra*. This letter might be called LIGATING RA and it would be defined for usage

only in conjuncts. Another possible name might be RRA, but this name is generally used in Indic scripts in Unicode for the Dravidian *ra* and as Gondi is a Dravidian language, this name is reserved for potentially encoding a such a character for the Gondi script. The representative glyph for this letter might be ⟨Ꙩ⟩, which is the letter Ꙩ RA placed within a dashed box that indicates that its function as a control character. The sequence ⟨LIGATING RA, VIRAMA, *C*⟩ would produce *repha*, while ⟨*C*, VIRAMA, LIGATING RA⟩ would produce *ra-kāra*. By extension, it does not have a half-form. The letter would be rendered as ⟨Ꙩ⟩ when used in any other context. Using this character the various forms of RA would be produced as follows:

ꙨꙨ	*rka*	⟨Ꙩ RA, ◌ VIRAMA, Ꙩ KA⟩
ꙨꙨ	*kra*	⟨Ꙩ KA, ◌ VIRAMA, Ꙩ RA⟩
Ꙩᴾ	*rka*	⟨Ꙩ LIGATING RA, ◌ VIRAMA, Ꙩ KA⟩
Ꙩⲧ	*kra*	⟨Ꙩ KA, ◌ VIRAMA, Ꙩ LIGATING RA⟩

Although LIGATING RA is an artificial letter and does not occur in the Gondi script, it facilitates the representation of *repha* and *ra-kāra* according to the general rules of conjunct formation, while also providing for the default rendering of RA as half-form and full-form in conjuncts:

Ꙩᴾ	*rkra*	⟨Ꙩ LIGATING RA, ◌ VIRAMA, Ꙩ KA, ◌ VIRAMA, Ꙩ LIGATING RA⟩
ꙨꙨꙨ	*rkra*	⟨Ꙩ RA, ◌ VIRAMA, Ꙩ KA, ◌ VIRAMA, Ꙩ RA⟩

This approach also aligns with the concept that *repha* and *ra-kāra* in Devanagari and other scripts are special forms of a RA character. The approach that utilizes LIGATING RA offers users with a clear method for producing normative and special forms of RA in conjuncts using VIRAMA, which follows the general rule for the encoded representation of conjuncts in Gondi.

4. *Encode 'repha' and 'ra-kāra' as separate characters*

In L2/15-045, the Script Ad-hoc Committee recommended that ○ᴺ *repha* and ○ᵢ *ra-kāra* be encoded as separate characters. While this approach provides the easiest means for encoding different forms of RA, it also raises several questions. Where should a *repha* character occur in an encoded sequence? Should it be handled logically and placed at the position where RA would normally occur, with the expectation that the rendering engine would reorder it to the end of the conjunct? Or should it be handled visually, and be placed manually at the end of the conjunct in an encoded sequence? For example, would the encoded representation for the hypothetical conjunct Ꙩᴾ *rkra* be ⟨○ᴺ *repha*, Ꙩ KA, ○ᵢ *ra-kāra*⟩ or ⟨Ꙩ KA, ○ᵢ *ra-kāra*, ○ᴺ *repha*⟩? Another issue concerns the classification of *repha*: is it to be considered a letter (ᴺ) or a sign (○ᴺ)? If *repha* is a letter, then is VIRAMA to be placed between it and the following consonant: ⟨ᴺ *repha*, ◌ VIRAMA, Ꙩ KA, ○ᵢ *ra-kāra*⟩?

This spproach requires the encoding of the following two characters:

⟨ᴺ⟩	REPHA
○ᵢ	RA-KARA

Furthermore, this approach requires the following considerations:

(a) The *repha* is defined a letter, not a combining sign. In the code chart the ᴖ *repha* is placed within a dashed box as ⸞ᴖ⸟ in order to indicate that it requires special rendering. The glyph ᴖ is used in the output.

(b) The logical model for *repha* is preferred over the visual model, ie. typing the character at the end of a cluster and any accompanying vowel signs. This approach provides the ability to type syllables according to the underlying phonology. The logical model requires that *repha* be placed at the same position in the encoded sequence that it occurs in phonetic expression. The rendering engine will re-order the *repha* to the end of the conjunct after any accompanying vowel signs.

(c) The *ra-kāra* is a regular combining sign. It is to be placed, as expected, in its logical position before any accompanying vowel sign.

(d) The VIRAMA is not to be used after either the *repha* or *ra-kāra*. Its presence in such contexts would be ignored, for example:

<ᴖ REPHA, ♀ VIRAMA, 0− KA>   →   ᴖ♀0−

<0− KA, ♀ VIRAMA, ◌ı RA-KARA>   →   0−♀◌ı

The REPHA and RA-KARA would be used in encoded sequences as follows:

0ᴬ   *rka*   <ᴖ REPHA, 0− KA>

0ᴛ   *kra*   <0− KA, ◌ı RA-KARA>

The usage of VIRAMA after and before RA would produce the expected output:

()0−   *rka*   <()− RA, ♀ VIRAMA, 0− KA>

0()−   *kra*   <0− KA, ♀ VIRAMA, ()− RA>

The most feasible of the above four is approaches #4. Representing cluster-specific forms of RA as separate characters offers a simpler implementation that does not require usage of control characters. For this reason, the suggestion made by the Script Ad-hoc Committee in In L2/15-045 has been adopted.

4.8.2   Conjunct letters

The clusters *kṣa, jña, tra* are represented not as regular conjuncts, but as distinctive letters. These are proposed for encoding as atomic letters:

ᛘ−   MASARAM GONDI LETTER KSSA

ᚾ−   MASARAM GONDI LETTER JNYA

ꝫ−   MASARAM GONDI LETTER TRA

Following the rules of conjunct formation, the expected representation of these three conjuncts would be:

kṣa   OƟ   <O⁻ KA, ♀ VIRAMA, Ɵ⁻ SSA>

jña   Ɵ꜀   <Ɵ⁻ JA, ♀ VIRAMA, ꜀ NYA>

tra   U⊙   <U⁻ TA, ♀ VIRAMA, ⊙⁻ RA>

In the Gondi script, each of these three letters phonetically represent a consonant cluster, but they all have the structure of an atomic letter. These forms are encoded as consonant letters because in all cases consonant conjuncts are written as linear sequences of half-forms, not as ligatures. While in most Indic scripts the written forms for *kṣa, jña, tra* have encoded representations as a character sequence, such an approach would not be consistent with this script.

It is evident that these ligatures were developed because distinctive forms exist in Devanagari. These three conjuncts are often shown at the end of Devanagari orthographies for various languages and are often interpreted by users as being distinctive letters that are fundamental elements of the script.

4.8.3   Rendering of conjuncts

As specified at the outset of this section, the sequence <$(C,$ ♀ VIRAMA$)^*, C_f$> produces a half-form of all C and the regular full-form of C_f. In order for the rendering of conjuncts to operate properly, the font must contain a complete set of consonants half-forms (glyphs without the horizontal stroke). The font should substitute each instance of <$(C,$ ♀ VIRAMA$)$> with the appropriate half-form for C. If this glyph is not available, then the output should show the full form of C followed by ♀ VIRAMA.

4.9   Digits

The script has a full set of digits:

O	MASARAM GONDI DIGIT ZERO	ß	MASARAM GONDI DIGIT FIVE
∪	MASARAM GONDI DIGIT ONE	ധ	MASARAM GONDI DIGIT SIX
U	MASARAM GONDI DIGIT TWO	⅃	MASARAM GONDI DIGIT SEVEN
∩	MASARAM GONDI DIGIT THREE	Ƽ	MASARAM GONDI DIGIT EIGHT
Ɵ	MASARAM GONDI DIGIT FOUR	Ϭ	MASARAM GONDI DIGIT NINE

Variant forms of digits are shown in figure 18.

4.10   Punctuation

Script-specific punctuation is not attested in Masaram's Gondi script. The *daṇḍā* and double *daṇḍā* are commonly used. These signs of punctuation are not proposed for separate inclusion in the 'Masaram Gondi' block, but are to be unified with ǀ U+0964 DEVANAGARI DANDA and ǁ U+0965 DEVANAGARI DOUBLE DANDA. Latin marks of punctuation, such as periods, are also used.

4.11   Characters Not Proposed for Encoding

The following are newly-invented characters. Actual usage of these characters, apart from their inclusion in new charts of the script, is unknown. For this reason, they are not proposed for encoding at present.

Additional vocalic vowel signs   The original script provided for the writing of only one Sanskrit vocalic vowel sign: ◌ VOWEL SIGN VOCALIC R. In order to accommodate anticipated usage of the script for linguistics research, Mukund Gokhale designed signs for the other three vocalic sounds:

    ◌ᵤ   *GONDI VOWEL SIGN VOCALIC RR

    ◌ᵤ   *GONDI VOWEL SIGN VOCALIC L

    ◌ᵤ   *GONDI VOWEL SIGN VOCALIC LL

The sounds represented by these four vowel signs are not found natively in Gondi. They occur only in Sanskrit words, and only ◌ VOWEL SIGN VOCALIC R is attested in usage. There is no corresponding independent letter for these signs, and in such contexts they are represented using a consonant-vowel combination composed with the letters RA and LA and the dependent sign for the vocalic letter:

    r̥   ◌   <◌ RA, ◌ VOWEL SIGN VOCALIC R>

    *r̥̄   ◌   <◌ RA, ◌ *VOWEL SIGN VOCALIC RR>

    *l̥   ◌   <◌ LA, ◌ *VOWEL SIGN VOCALIC L>

    *l̥̄   ◌   <◌ LA, ◌ *VOWEL SIGN VOCALIC LL>

The VOWEL SIGN VOCALIC R is included in the proposed repertoire, but the other three signs are not proposed for inclusion at present because their actual usage is not attested. Space has been reserved for these characters in the block in the event that a justifiable case to encode them is made in the future.

5   Character Data

5.1   Character Properties

The properties for 'Masaram Gondi' in the Unicode Character Database format are:

```
11B90;MASARAM GONDI LETTER A;Lo;0;L;;;;;N;;;;;
11B91;MASARAM GONDI LETTER AA;Lo;0;L;;;;;N;;;;;
11B92;MASARAM GONDI LETTER I;Lo;0;L;;;;;N;;;;;
11B93;MASARAM GONDI LETTER II;Lo;0;L;;;;;N;;;;;
11B94;MASARAM GONDI LETTER U;Lo;0;L;;;;;N;;;;;
11B95;MASARAM GONDI LETTER UU;Lo;0;L;;;;;N;;;;;
11B96;MASARAM GONDI LETTER E;Lo;0;L;;;;;N;;;;;
11B98;MASARAM GONDI LETTER AI;Lo;0;L;;;;;N;;;;;
11B99;MASARAM GONDI LETTER O;Lo;0;L;;;;;N;;;;;
11B9B;MASARAM GONDI LETTER AU;Lo;0;L;;;;;N;;;;;
11B9C;MASARAM GONDI LETTER KA;Lo;0;L;;;;;N;;;;;
11B9D;MASARAM GONDI LETTER KHA;Lo;0;L;;;;;N;;;;;
11B9E;MASARAM GONDI LETTER GA;Lo;0;L;;;;;N;;;;;
11B9F;MASARAM GONDI LETTER GHA;Lo;0;L;;;;;N;;;;;
11BA0;MASARAM GONDI LETTER NGA;Lo;0;L;;;;;N;;;;;
11BA1;MASARAM GONDI LETTER CA;Lo;0;L;;;;;N;;;;;
11BA2;MASARAM GONDI LETTER CHA;Lo;0;L;;;;;N;;;;;
11BA3;MASARAM GONDI LETTER JA;Lo;0;L;;;;;N;;;;;
11BA4;MASARAM GONDI LETTER JHA;Lo;0;L;;;;;N;;;;;
11BA5;MASARAM GONDI LETTER NYA;Lo;0;L;;;;;N;;;;;
11BA6;MASARAM GONDI LETTER TTA;Lo;0;L;;;;;N;;;;;
11BA7;MASARAM GONDI LETTER TTHA;Lo;0;L;;;;;N;;;;;
```

```
11BA8;MASARAM GONDI LETTER DDA;Lo;0;L;;;;;;N;;;;;
11BA9;MASARAM GONDI LETTER DDHA;Lo;0;L;;;;;;N;;;;;
11BAA;MASARAM GONDI LETTER NNA;Lo;0;L;;;;;;N;;;;;
11BAB;MASARAM GONDI LETTER TA;Lo;0;L;;;;;;N;;;;;
11BAC;MASARAM GONDI LETTER THA;Lo;0;L;;;;;;N;;;;;
11BAD;MASARAM GONDI LETTER DA;Lo;0;L;;;;;;N;;;;;
11BAE;MASARAM GONDI LETTER DHA;Lo;0;L;;;;;;N;;;;;
11BAF;MASARAM GONDI LETTER NA;Lo;0;L;;;;;;N;;;;;
11BB0;MASARAM GONDI LETTER PA;Lo;0;L;;;;;;N;;;;;
11BB1;MASARAM GONDI LETTER PHA;Lo;0;L;;;;;;N;;;;;
11BB2;MASARAM GONDI LETTER BA;Lo;0;L;;;;;;N;;;;;
11BB3;MASARAM GONDI LETTER BHA;Lo;0;L;;;;;;N;;;;;
11BB4;MASARAM GONDI LETTER MA;Lo;0;L;;;;;;N;;;;;
11BB5;MASARAM GONDI LETTER YA;Lo;0;L;;;;;;N;;;;;
11BB6;MASARAM GONDI LETTER RA;Lo;0;L;;;;;;N;;;;;
11BB7;MASARAM GONDI LETTER LA;Lo;0;L;;;;;;N;;;;;
11BB8;MASARAM GONDI LETTER VA;Lo;0;L;;;;;;N;;;;;
11BB9;MASARAM GONDI LETTER SHA;Lo;0;L;;;;;;N;;;;;
11BBA;MASARAM GONDI LETTER SSA;Lo;0;L;;;;;;N;;;;;
11BBB;MASARAM GONDI LETTER SA;Lo;0;L;;;;;;N;;;;;
11BBC;MASARAM GONDI LETTER HA;Lo;0;L;;;;;;N;;;;;
11BBD;MASARAM GONDI LETTER LLA;Lo;0;L;;;;;;N;;;;;
11BBE;MASARAM GONDI LETTER KSSA;Lo;0;L;;;;;;N;;;;;
11BBF;MASARAM GONDI LETTER JNYA;Lo;0;L;;;;;;N;;;;;
11BC0;MASARAM GONDI LETTER TRA;Lo;0;L;;;;;;N;;;;;
11BC1;MASARAM GONDI VOWEL SIGN AA;Mn;0;NSM;;;;;N;;;;;
11BC2;MASARAM GONDI VOWEL SIGN I;Mn;0;NSM;;;;;N;;;;;
11BC3;MASARAM GONDI VOWEL SIGN II;Mn;0;NSM;;;;;N;;;;;
11BC4;MASARAM GONDI VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
11BC5;MASARAM GONDI VOWEL SIGN UU;Mn;0;NSM;;;;;N;;;;;
11BC6;MASARAM GONDI VOWEL SIGN VOCALIC R;Mn;0;NSM;;;;;N;;;;;
11BCA;MASARAM GONDI VOWEL SIGN E;Mn;0;NSM;;;;;N;;;;;
11BCC;MASARAM GONDI VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
11BCD;MASARAM GONDI VOWEL SIGN O;Mn;0;NSM;;;;;N;;;;;
11BCF;MASARAM GONDI VOWEL SIGN AU;Mn;0;NSM;;;;;N;;;;;
11BD0;MASARAM GONDI SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11BD1;MASARAM GONDI SIGN VISARGA;Mn;0;NSM;;;;;N;;;;;
11BD2;MASARAM GONDI SIGN NUKTA;Mn;7;NSM;;;;;N;;;;;
11BD3;MASARAM GONDI SIGN CANDRA;Mn;0;NSM;;;;;N;;;;;
11BD4;MASARAM GONDI SIGN HALANTA;Mn;0;NSM;;;;;N;;;;;
11BD5;MASARAM GONDI VIRAMA;Mn;9;NSM;;;;;N;;;;;
11BD6;MASARAM GONDI REPHA;Lo;0;L;;;;;;N;;;;;
11BD7;MASARAM GONDI RA-KARA;Mn;0;NSM;;;;;N;;;;;
11BE0;MASARAM GONDI DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
11BE1;MASARAM GONDI DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
11BE2;MASARAM GONDI DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
11BE3;MASARAM GONDI DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
11BE4;MASARAM GONDI DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
11BE5;MASARAM GONDI DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
11BE6;MASARAM GONDI DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
11BE7;MASARAM GONDI DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
11BE8;MASARAM GONDI DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
11BE9;MASARAM GONDI DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
```

5.2 Linebreaking

Linebreaking properties given in the data format of `LineBreak.txt`:

```
11B90..11BC0; AL   # MASARAM GONDI LETTER A .. MASARAM GONDI LETTER TRA
11BC1..11BD4; CM   # MASARAM GONDI SIGN AA .. MASARAM GONDI SIGN HALANTA
```

```
11BD5; CM          # MASARAM GONDI VIRAMA
11BC6; AL          # MASARAM GONDI REPHA
11BD7; CM          # MASARAM GONDI RA-KARA
11BE0..11BE9; NU   # MASARAM GONDI DIGIT ZERO .. MASARAM GONDI DIGIT NINE
```

5.3   Syllabic Categories

Syllabic categories given in the format of `IndicSyllabicCategory.txt`:

```
# Indic_Syllabic_Category=Bindu
11BD0          ; Bindu                 # Mn        MASARAM GONDI SIGN ANUSVARA


# Indic_Syllabic_Category=Visarga
11BD1          ; Visarga               # Mc        MASARAM GONDI SIGN VISARGA


# Indic_Syllabic_Category=Nukta
11BD2          ; Nukta                 # Mn        MASARAM GONDI SIGN NUKTA


# Indic_Syllabic_Category=Virama
11BD5          ; Virama                # Mn        MASARAM GONDI VIRAMA


# Indic_Syllabic_Category=Pure_Killer
11BD4          ; Pure_Killer           # Mn        MASARAM GONDI SIGN HALANTA


# Indic_Syllabic_Category=Vowel_Independent
11B90..11B9B ; Vowel_Independent    # Lo   [10] MASARAM GONDI LETTER A .. AU


# Indic_Syllabic_Category=Vowel_Dependent
11BC1..11BC6 ; Vowel_Dependent      # Mn    [6] MASARAM GONDI VOWEL SIGN AA .. VOCALIC R
11BCA        ; Vowel_Dependent      # Mn        MASARAM GONDI VOWEL SIGN E
11BCC..11BCD ; Vowel_Dependent      # Mn    [2] MASARAM GONDI VOWEL SIGN AI .. O
11BCF        ; Vowel_Dependent      # Mn        MASARAM GONDI VOWEL SIGN AU
11BD3        ; Vowel_Dependent      # Mn        MASARAM GONDI SIGN CANDRA


# Indic_Syllabic_Category=Consonant
11B9C..11BC0 ; Consonant             # Lo   [40] MASARAM GONDI LETTER KA .. TRA


# Indic_Syllabic_Category=Consonant_Succeeding_Repha
11BD6          ; Consonant_Succeeding_Repha  # Lo  MASARAM GONDI REPHA


# Indic_Syllabic_Category=Consonant_Subjoined
11BD7          ; Consonant_Subjoined # Mn        MASARAM GONDI RA-KARA
```

5.4   Positional Categories

Positional data for Gondi combining signs in the format of `IndicPositionalCategory.txt`:

```
# Indic_Positional_Category=Top
11BC1..11BC5  ; Top     # Mn    [5] MASARAM GONDI VOWEL SIGN AA .. UU
11BCA         ; Top     # Mn        MASARAM GONDI VOWEL SIGN E
11BCC..11BCD  ; Top     # Mn    [2] MASARAM GONDI VOWEL SIGN AI .. O
11BCF         ; Top     # Mn        MASARAM GONDI VOWEL SIGN AU
11BD0         ; Top     # Mn        MASARAM GONDI SIGN ANUSVARA
11BD1         ; Top     # Mn        MASARAM GONDI SIGN VISARGA
11BD3         ; Top     # Mn        MASARAM GONDI SIGN CANDRA


# Indic_Positional_Category=Bottom
11BC6         ; Bottom  # Mn        MASARAM GONDI VOWEL SIGN VOCALIC R
```

```
11BD2          ; Bottom  # Mn       MASARAM GONDI SIGN NUKTA
11BD4          ; Bottom  # Mn       MASARAM GONDI SIGN HALANTA
11BD7          ; Bottom  # Mn       MASARAM GONDI RA-KARA
```

5.5 'Confusable' Characters

Gondi characters that bear resemblances to those of other scripts are listed below:

```
11BC1 MASARAM GONDI VOWEL SIGN AA  ;  0304 COMBINING MACRON
11BB1 MASARAM GONDI LETTER PHA     ;  1109D KAITHI LETTER NNA
11BBA MASARAM GONDI LETTER SSA     ;  0398 GREEK CAPITAL LETTER THETA
11BE2 MASARAM GONDI DIGIT TWO      ;  0055 LATIN CAPITAL LETTER U
```

6 References

Anderson, Deborah; et al. 2015. "Recommendations to UTC #142 February 2015 on Script Proposals" (L2/15-045). http://www.unicode.org/L2/L2015/15045-script-rec.pdf

गुरूजी, मनीराम दुर्गा [Gurūjī, Manīrāma Durgā]. गोंदी लम्क पुंदान [*Goṃḍī lamka puṃdāna*].

मण्डाले, सीताराम [Maṇḍāle, Sītārām]. कोयाबोली [*Koyābolī*]. गोंडी शब्द संग्रह - गोंडी, मराठी, हिन्दी [Goṃḍī Śabda Saṃgraha - Goṃḍī, Marāṭhī, Hindī].

Masaram, Bhava Singh. 1951. "गोंडी लिपि" [Goṃḍī lipi]. Central Institute of Indian Languages, Multimedia library, photograph no. 64.

Pandey, Anshuman. 2010. "Preliminary Proposal to Encode the Gondi Script in the UCS" (L2/10-207). http://www.unicode.org/L2/L2010/10207-gondi.pdf

———. 2012. "Revised Preliminary Proposal to Encode the Gondi Script" (L2/12-235). http://www.unicode.org/L2/L2012/12235-n4291-gondi.pdf

———. 2015a. "Proposal to Encode the Gondi Script" (L2/15-005). http://www.unicode.org/L2/L2015/15005-gondi.pdf

———. 2015b. "Preliminary Proposal to Encode the Gunjala Gondi Script in Unicode" (L2/15-086). http://www.unicode.org/L2/L2015/15086-gunjala-gondi.pdf

———. 2015c. "Proposal to Encode the 'Parsapen' Symbol in Unicode" (L2/15-111) http://www.unicode.org/L2/L2015/15111-parsapen-symbol.pdf

Ramakrishna, G., N. Gayathri, Debiprasad Chattopadhyaya. 1983. *An Encyclopaedia of South Indian Culture*. Calcutta: K. P. Bagchi & Co.

रामानन्द [Rāmānanda]. गोंडी अक्षर ज्ञान [Goṃḍī Akṣara Jñāna].

Vahia, M. N.; Ganesh Halkare. 2013. "Aspects of Gondi Astronomy". *Journal of Astronomical History and Heritage*, vol. 16, no. 1, pp. 29–44. http://www.tifr.res.in/~archaeo/VahiaGondTrip/JAHHGondsfinal.pdf

7 Acknowledgments

I am deeply indebted to Dr. Mukund Gokhale (Script Research Institute, Pune) for providing me with numerous materials in Masaram's Gondi script and for patiently answering my numerous questions about the script over the past few years. Dr. Gokhale's efforts in developing typefaces for this Gondi script have helped to address issues regarding the interaction of multiple glyphs with a base letter. I also appreciate him for contacting Dr. Motiravan Kangle (Akhil-Gondwana Gondi Sahitya Parishad, Nagpur) with my questions. I thank Dr. Kangle for indulging my inquiries and for providing examples of current orthography of *ra* in conjuncts, and the usage of *halanta* and *candra*. The Gondi chart shown in figure **??** was provided by B. A. Sharada and Suman Kumari of the Central Institute of Indian Languages (Mysore). Mark Penny provided the chart shown in figure 6. I am also thankful to Roozbeh Pournader (Google) and Andrew Glass (Microsoft) for sharing their insights regarding effective models for representing the Gondi RA in conjuncts.

	11B9	11BA	11BB	11BC	11BD	11BE
0	∩ 11B90	ઙ 11BA0	⊢ 11BB0	⅔ 11BC0	◌˙ 11BD0	0 11BE0
1	∩ 11B91	⊃ 11BA1	⊞ 11BB1	◌⁻ 11BC1	◌˙˙ 11BD1	∪ 11BE1
2	∟ 11B92	౩ 11BA2	⊢ 11BB2	◌ᶠ 11BC2	◌. 11BD2	∪ 11BE2
3	∼ 11B93	⊖ 11BA3	⊱ 11BB3	◌ᶯ 11BC3	◌˘ 11BD3	∩ 11BE3
4	Ч 11B94	ற 11BA4	⅂ 11BB4	◌ᴴ 11BC4	◌ˎ 11BD4	θ 11BE4
5	⊌ 11B95	⊬ 11BA5	⌓ 11BB5	◌ᴾ 11BC5	[◌♀] 11BD5	Ɓ 11BE5
6	∟ 11B96	∪ 11BA6	(·) 11BB6	◌ᵤ 11BC6	[∩] 11BD6	ⱳ 11BE6
7	▨	✝ 11BA7	∩ 11BB7	▨	◌ᵢ 11BD7	⌡ 11BE7
8	⊥ 11B98	ઙ 11BA8	⅄ 11BB8	▨	▨	9 11BE8
9	∩ 11B99	⋒ 11BA9	⌵ 11BB9	▨	▨	6 11BE9
A	▨	௯ 11BAA	θ 11BBA	◌ˡ 11BCA	▨	▨
B	⋒ 11B9B	∪ 11BAB	Ɛ 11BBB	▨	▨	▨
C	௦ 11B9C	✕ 11BAC	ⱳ 11BBC	◌ˡˡ 11BCC	▨	▨
D	8 11B9D	౨ 11BAD	∞ 11BBD	◌ᴸ 11BCD	▨	▨
E	⊦ 11B9E	౨ 11BAE	⋔ 11BBE	▨	▨	▨
F	Ɛ 11B9F	ⱳ 11BAF	ᴕ 11BBF	◌ᴵᴵ 11BCF	▨	▨

Vowels

11B90	∩	MASARAM GONDI LETTER A
11B91	∩	MASARAM GONDI LETTER AA
11B92	⊂	MASARAM GONDI LETTER I
11B93	⊃	MASARAM GONDI LETTER II
11B94	⊔	MASARAM GONDI LETTER U
11B95	⊌	MASARAM GONDI LETTER UU
11B96	∟	MASARAM GONDI LETTER E
11B97	▨	\<reserved\>
11B98	⊔	MASARAM GONDI LETTER AI
11B99	∩	MASARAM GONDI LETTER O
11B9A	▨	\<reserved\>
11B9B	∩	MASARAM GONDI LETTER AU

Consonants

11B9C	O⊢	MASARAM GONDI LETTER KA
11B9D	⊖	MASARAM GONDI LETTER KHA
11B9E	⊢	MASARAM GONDI LETTER GA
11B9F	⊖	MASARAM GONDI LETTER GHA
11BA0	⊖	MASARAM GONDI LETTER NGA
11BA1	⊃⊢	MASARAM GONDI LETTER CA
11BA2	∃⊢	MASARAM GONDI LETTER CHA
11BA3	⊕	MASARAM GONDI LETTER JA
11BA4	∽	MASARAM GONDI LETTER JHA
11BA5	⊬	MASARAM GONDI LETTER NYA
11BA6	∪	MASARAM GONDI LETTER TTA
11BA7	+	MASARAM GONDI LETTER TTHA
11BA8	⊱	MASARAM GONDI LETTER DDA
11BA9	∩⊢	MASARAM GONDI LETTER DDHA
11BAA	⊖	MASARAM GONDI LETTER NNA
11BAB	∪⊢	MASARAM GONDI LETTER TA
11BAC	✕	MASARAM GONDI LETTER THA
11BAD	⊃⊢	MASARAM GONDI LETTER DA
11BAE	⊋	MASARAM GONDI LETTER DHA
11BAF	⊍	MASARAM GONDI LETTER NA
11BB0	⊬	MASARAM GONDI LETTER PA
11BB1	⊬	MASARAM GONDI LETTER PHA
11BB2	∪⊢	MASARAM GONDI LETTER BA
11BB3	∪	MASARAM GONDI LETTER BHA
11BB4	⊬	MASARAM GONDI LETTER MA
11BB5	⊂⊢	MASARAM GONDI LETTER YA
11BB6	⊕⊢	MASARAM GONDI LETTER RA
11BB7	∩⊢	MASARAM GONDI LETTER LA
11BB8	⊛⊢	MASARAM GONDI LETTER VA
11BB9	⊬	MASARAM GONDI LETTER SHA
11BBA	⊖	MASARAM GONDI LETTER SSA
11BBB	⊖	MASARAM GONDI LETTER SA
11BBC	⊌	MASARAM GONDI LETTER HA
11BBD	∞⊢	MASARAM GONDI LETTER LLA

Conjunct letters

11BBE	⋔⊢	MASARAM GONDI LETTER KSSA
11BBF	∩⊢	MASARAM GONDI LETTER JNYA
11BC0	∋⊢	MASARAM GONDI LETTER TRA

Dependent vowel signs

11BC1	○⁻	MASARAM GONDI VOWEL SIGN AA
11BC2	○ᶠ	MASARAM GONDI VOWEL SIGN I
11BC3	○ᴺ	MASARAM GONDI VOWEL SIGN II
11BC4	○ᵁ	MASARAM GONDI VOWEL SIGN U
11BC5	○ᵁ	MASARAM GONDI VOWEL SIGN UU
11BC6	○ᵤ	MASARAM GONDI VOWEL SIGN VOCALIC R
11BC7	▨	\<reserved\>
11BC8	▨	\<reserved\>
11BC9	▨	\<reserved\>
11BCA	○ᴵ	MASARAM GONDI VOWEL SIGN E
11BCB	▨	\<reserved\>

11BCC	○"	MASARAM GONDI VOWEL SIGN AI
11BCD	○⊥	MASARAM GONDI VOWEL SIGN O
11BCE	▨	\<reserved\>
11BCF	○ᴵᴵ	MASARAM GONDI VOWEL SIGN AU

Various signs

11BD0	○˙	MASARAM GONDI SIGN ANUSVARA
11BD1	○¨	MASARAM GONDI SIGN VISARGA
11BD2	○.	MASARAM GONDI SIGN NUKTA
11BD3	○ˇ	MASARAM GONDI SIGN CANDRA
		• used for transcribing foreign vowels
11BD4	○ˎ	MASARAM GONDI SIGN HALANTA
		• used for silencing the inherent vowel

Virama

11BD5	⊡	MASARAM GONDI VIRAMA
		• used for producing conjuncts

Cluster-specific consonant forms

11BD6	⧈	MASARAM GONDI REPHA
		• cluster-initial form of RA
11BD7	○ˌ	MASARAM GONDI RA-KARA
		• cluster-final form of RA

Digits

11BE0	O	MASARAM GONDI DIGIT ZERO
11BE1	∪	MASARAM GONDI DIGIT ONE
11BE2	∪	MASARAM GONDI DIGIT TWO
11BE3	∩	MASARAM GONDI DIGIT THREE
11BE4	θ	MASARAM GONDI DIGIT FOUR
11BE5	B	MASARAM GONDI DIGIT FIVE
11BE6	ʊ	MASARAM GONDI DIGIT SIX
11BE7	J	MASARAM GONDI DIGIT SEVEN
11BE8	϶	MASARAM GONDI DIGIT EIGHT
11BE9	ϭ	MASARAM GONDI DIGIT NINE

Figure 1: A document illustrating the characters and basic principles of Mangal Singh Masaram's Gondi script. It was created by Mangal Singh's son Bhava Singh in 1951. Image courtesy of the Central Institute of Indian Languages, Mysore.
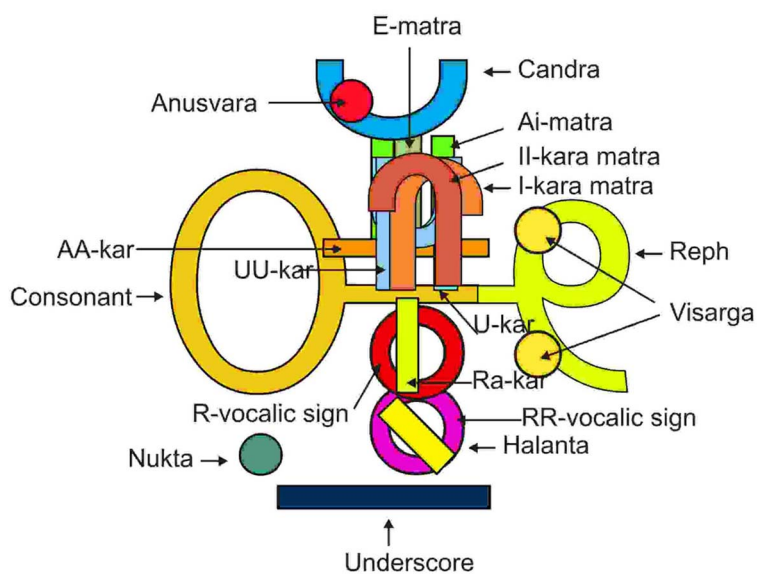
Figure 2: A schematic designed by Mukund Gokhale showing the position of all combining signs on a base letter. The diagram shows a sign labelled "RR-vocalic sign", which is a length variant of ◌ु VOWEL SIGN VOCALIC R introduced by Gokhale (see section 4.11).
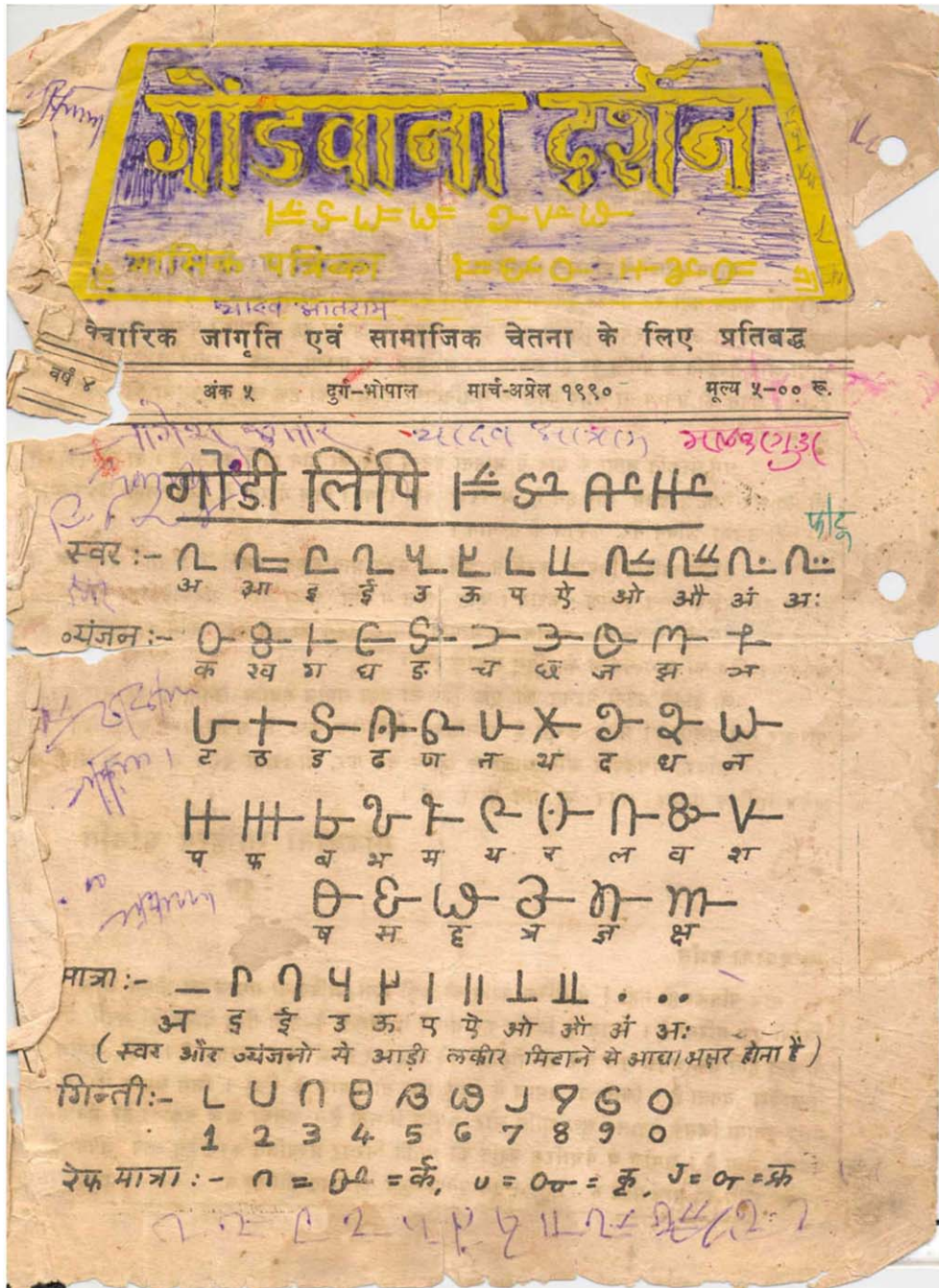
Figure 3: Cover of *Gondwana Darshan* (March-April 1990, vol. 5).

Figure 4: Page from *Goṇḍī Akṣara Jñāna* showing consonants and vowel signs, and consonant-vowel combinations (from Rāmānanda: 8).

Figure 5: Page from *Goṃḍī Lamk Pundan* showing consonants and vowel signs, and consonant-vowel combinations (from Guruji: 11).

Consonants व्यञ्जन

	Cons. only	− [voice]		+ [voice]		Nasals	Cons. with word
Aspiration		− [asp]	+ [asp]	− [asp]	+ [asp]		
Velar		౮	৪	⊦	Ɛ	⊃⋅	
Palatal		⊃	Ɔ	ᴀ	ᴦ	⨍	
Retroflex		৮	⊦	ৎ	ᴦ	ৎ	
Dental		Ʊ	⨯	৩	৯	ω	
Bilabial		Ⱨ	Ⱨⱨ	৳	৪	৭	
Liquids and Semi-Vowels		ৎ	(·)	⌃	৪		
Fricatives		Ʋ	θ	Ɛ	ω		
Affricates		ᴦ	Ƹ	ᴦ			

Vowels स्वर

ᴧ	ᴧᴴ	Ɛ	ᴧ	Ⴑ	ᴦ
∟	⅃∟	ᴧᴴ	ᴧᴴ	ᴧ∙	ᴧ∙∙

Figure 6: A handwritten chart of the Gondi script. Source: Ramesh Gedam and Mark Penny (2001).
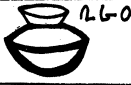
Figure 7: A handwritten chart of the Gondi script (Maṇḍāle 2008: 8).
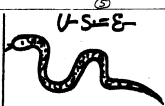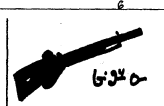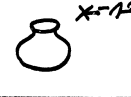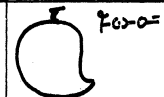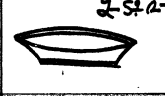
Figure 8: Page from *Goṃḍī Akṣara Jñāna* showing vowel letters (from Rāmānanda: 1–4).
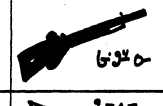
Figure 9: Page from *Goṃḍī Akṣara Jñāna* showing the letters ᴛᴀ .. ᴠᴀ (from Rāmānanda: 5–7).
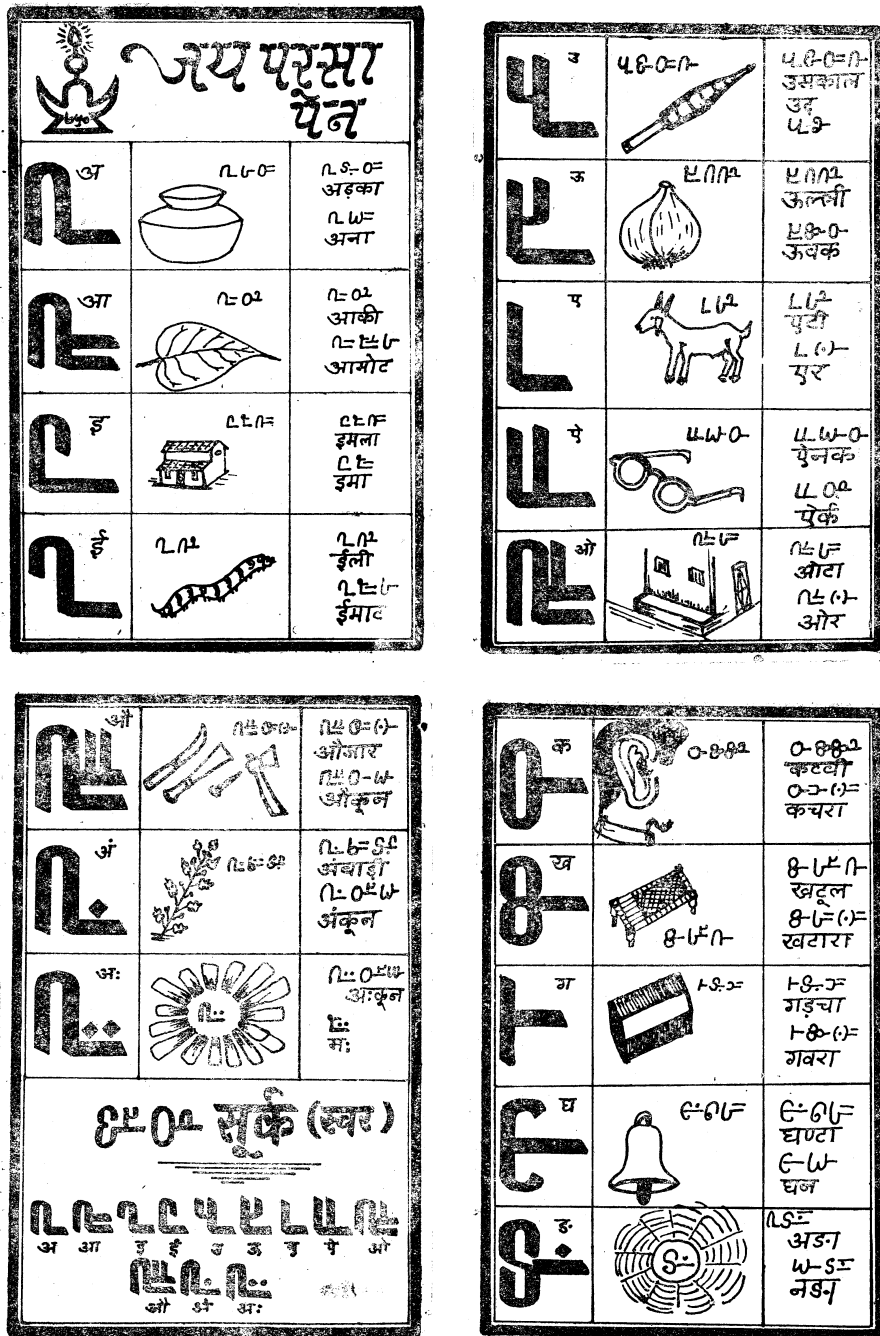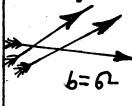
Figure 10: Page from *Goṃḍī Lamk Pundan* showing vowel letters (from Guruji: 1–4). The first page (top right corner) shows the *persapen* symbol.
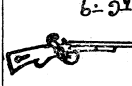
Figure 11: Page from *Goṃḍī Lamk Pundan* showing vowel letters (from Guruji: 5–8).
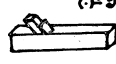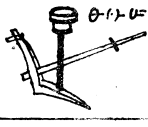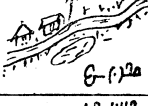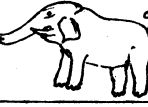
Figure 12: Page from *Goṇḍī Lamk Pundan* showing vowel letters (from Guruji: 9–10).

Figure 13: Page from *Goṃḍī Akṣara Jñāna* showing consonant-vowel combinations for ᴋᴀ .. ᴅᴀ (from Rāmānanda: 9, 10).

Figure 14: Page from *Goṃḍī Akṣara Jñāna* showing consonant-vowel combinations for DHA .. TRA (from Rāmānanda: 11, 12).

(जोड़ी लम्क) जोड़ अक्षर

(सूक्ति उण्ड वंजेगना आड़ो रबीरा इसूमतेन सूर्या लम्क आयतांग)

स्वर और व्यंजनों से आड़ी लकीर मिटाने से आधा अक्षर होता है। जैसे.

क्य = क + य = क्या
श्य = श + म = शमा
च्य = च + य = च्या

(हिन्दान) चिन्ह

(सयगो रोमी) . पूर्ण विराम
(मिन्द रोमी) , स्वल्प विराम
(सुर्या रोमी) ; अर्ध विराम
(अंगा रोमी) : विसर्ग
(पुसना रोमी) ? प्रश्न चिन्ह
(पर्चुक रोमी) ! उद्गार

Figure 15: Page from *Goṃḍī Akṣara Jñāna* explaining conjunct formation (from Rāmānanda: 13). Bottom half describes the usage of Latin marks of punctuation.

Figure 16:

ꓮ S F (पाड़ांग) गिनती		
उन्दी एक		सारंग छह
रंड दो		येरंग सात
मून्द तीन		अरंग आठ
नालुंग चार		नरंग नौ
सयुंग पांच		पद दस

नुर वेरी पाड़ांग (सौ तक गिनती)		
उन्दी	पवुदी	रण्डुन्दी
रण्ड	पदंड	रण्डन्ड
मून्द	पदूंद	रण्डुन्द
नालुंग	पनालू	रण्नालू
सयुंग	पसयुं	रन्सयुं
सारंग	पसारू	रन्सारू
येरंग	पदेरू	रन्डेरू
अरंग	पदारू	रण्डारू
नरंग	पनरू	रन्नरू
पद	रण्ड	मण्ड

Figure 17:

ꓮ S F (पाड़ांग) गिनती		
उंदी एक		सारंग छह
रंड दो		येरंग सात
मूंद तीन		अरंग आठ
नालुंग चार		नरंग नब
सयुंग पांच		पद दस

नुर वेरी पाड़ांग (सौ तक गिनती)		
उंदी	पदूंदी	रण्डुन्दी
रंड	पदंड	रण्डन्ड
मूंद	पदूंद	रण्डुन्द
नालुंग	पनालू	रन्नालू
सयुंग	पसयुं	रन्सयुं
सारंग	पसारू	रन्सारू
येरंग	पदेरू	रन्डेरू
अरंग	पदारू	रन्डारू
नरंग	पनरू	रन्नरू
पद	रंड	मण्ड

Figure 16: Page from *Goṇḍī Akṣara Jñāna* showing Gondi digits (from Rāmānanda: 14).

Figure 17: Page from *Goṇḍī Lamk Pundan* showing Gondi digits (from Guruji: 14).

36

Number	Name	Style in Maharashtra	Style in Madhya Pradesh
1	Undi		
2	Rand		
3	Munda		
4	Nalung		
5	Sayyung		
6	Sarung		
7	Yerung		
8	Arung		
9	Narung		
10	Pad		

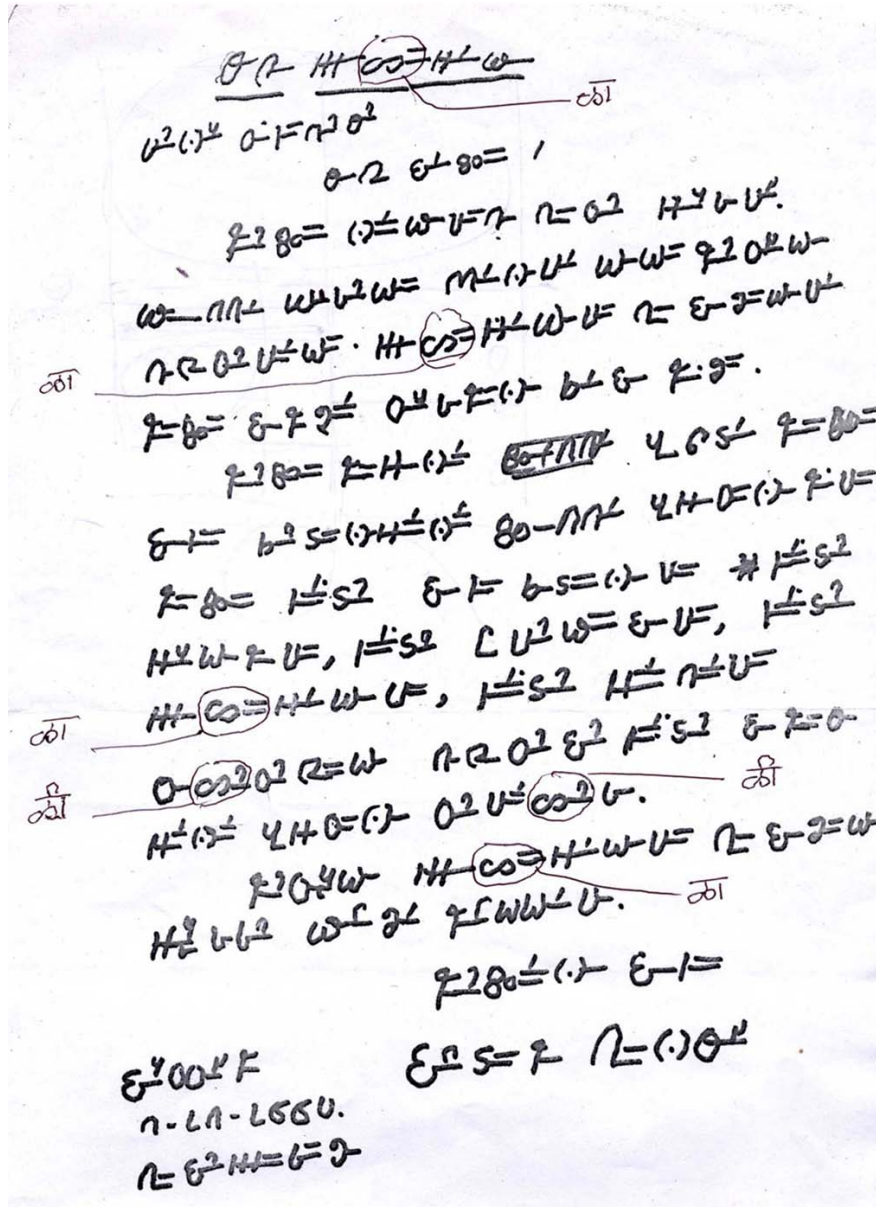Figure 18: Comparison of regional variants of Gondi digits (from Vahia and Halkare 2013: 33).

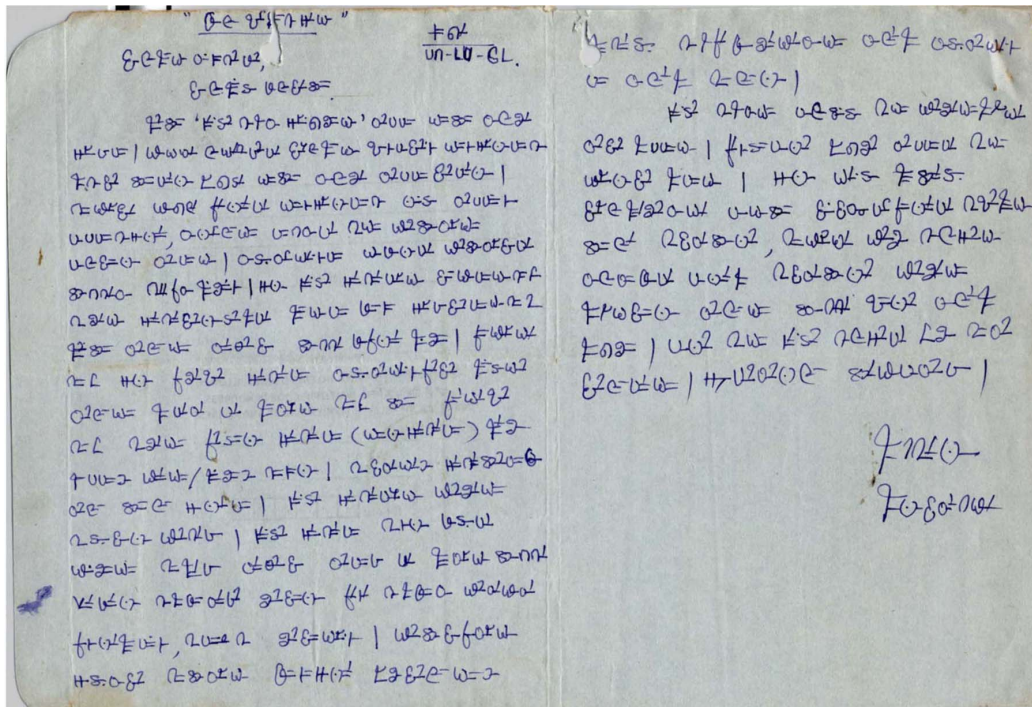Figure 19: A document showing usage of ꣲ LLA.

Figure 20: A letter written in Gondi.
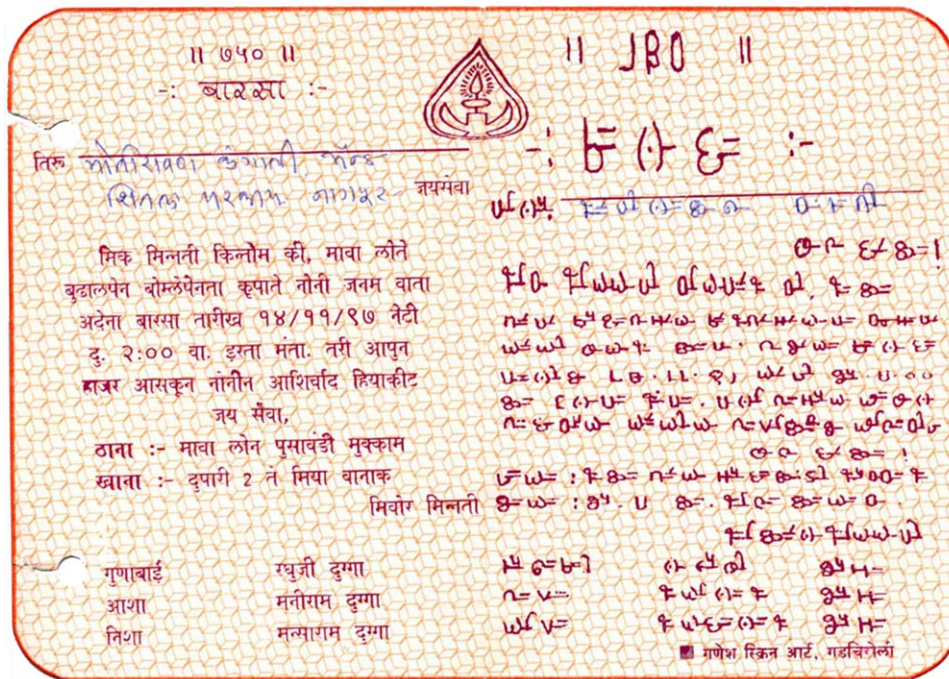
Figure 21: An invitation card written in the Devanagari and Gondi scripts. The *persapen* symbol appears in the top center of the card.

पेरसापेन पुयंनीना पाटा

ओ सल्ला शक्ति इमा, गांगरा शक्ति इमा ।
परसापेन मावा ओ परसापेन मावा ॥
जीवा शक्ति इमा, जायी शक्ति इमा ।
सजोर शक्ति मावा, ओ परसापेन मावा ॥
कयमो शक्ति मावा, करूमो शक्ति मावा ।
कामो शक्ति मावा, ओ परसापेन मावा ॥
मेंदोल शक्ति इमा, मती शक्ति इमा ।
मोह शक्ति मावा, ओ परसापेन मावा ॥
पुनो शक्ति इता, उनो शक्ति इमा ।
गुणो शक्ति मावा, ओ परसापेन मावा ॥
मान शक्ति इमा, माई शक्ति इमा ।
पुटसे शक्ति मावा, ओ परसापेन मावा ॥
कासे शक्ति इमा, मासे शक्ति इमा ।
गार शक्ति मावा, ओ परसापेन मावा ॥
तयमो शक्ति इमा, रयमो शक्ति इमा ।
सयमो शक्ति मावा, ओ परसापेन मावा ॥
गुर्रे शक्ति इमा, बुर्रे शक्ति इमा ।
यायाल शक्ति मावा, ओ परसापेन मावा ॥
मंत्रे शक्ति इमा, जंत्रे शक्ति इमा ।
बाबाल शक्ति मावा, ओ परसापेन मावा ॥
पुंगार शक्ति इमा, सिंगार शक्ति इमा ।
नींगार शक्ति मावा, ओ परसापेन मावा ॥
केंसार शक्ति इमा, वेंजार शक्ति इमा ।
लेंगार शक्ति मावा, ओ परसापेन मावा ॥
अगास शक्ति मावा, पताल शक्ति इमा ।
पोकराल शक्ति मावा, ओ परसापेन मावा ॥
विब्वासी शक्ति इमा, इवौसि शक्ति इमा ।
सल्ला शक्ति मावा, ओ परसापेन मावा ॥
ओ बुढालपेन मावा ॥ ओ परसापेन मावा ॥
॥ सेवा ॥ सेवा ॥ सेवा ॥

सीताराम मंडाले (सचीव)
आदिवासी गोंड समाज मंडळ, पुणे.

Figure 22: A Christian prayer typeset using digitized Gondi and Devanagari fonts. The Gondi font used in this specimen was designed by Mukund Gokhale.

|| ꙮ-ꙮ ꙮ-U-꙰-ꙮ ||

[Gondi script text]

Text is complete "Vande Mataram"
This 'Gondi' Script font is designed by : Prof Mukund Gokhale
Script Research Institute
mwgokhale_scriptresearch@yahoo.co.in
gokhale.mukund1@gmail.com

Dedicated to the inventor of the script 'Munshi Mangal Singh Masaram'

Figure 23: The lyrics to "Vande Mataram" transliterated into Gondi.

Figure 24: Usage of the sign HALANTA with regular forms of consonants for representation of bare consonants. Contents of letter described in section 4.6.
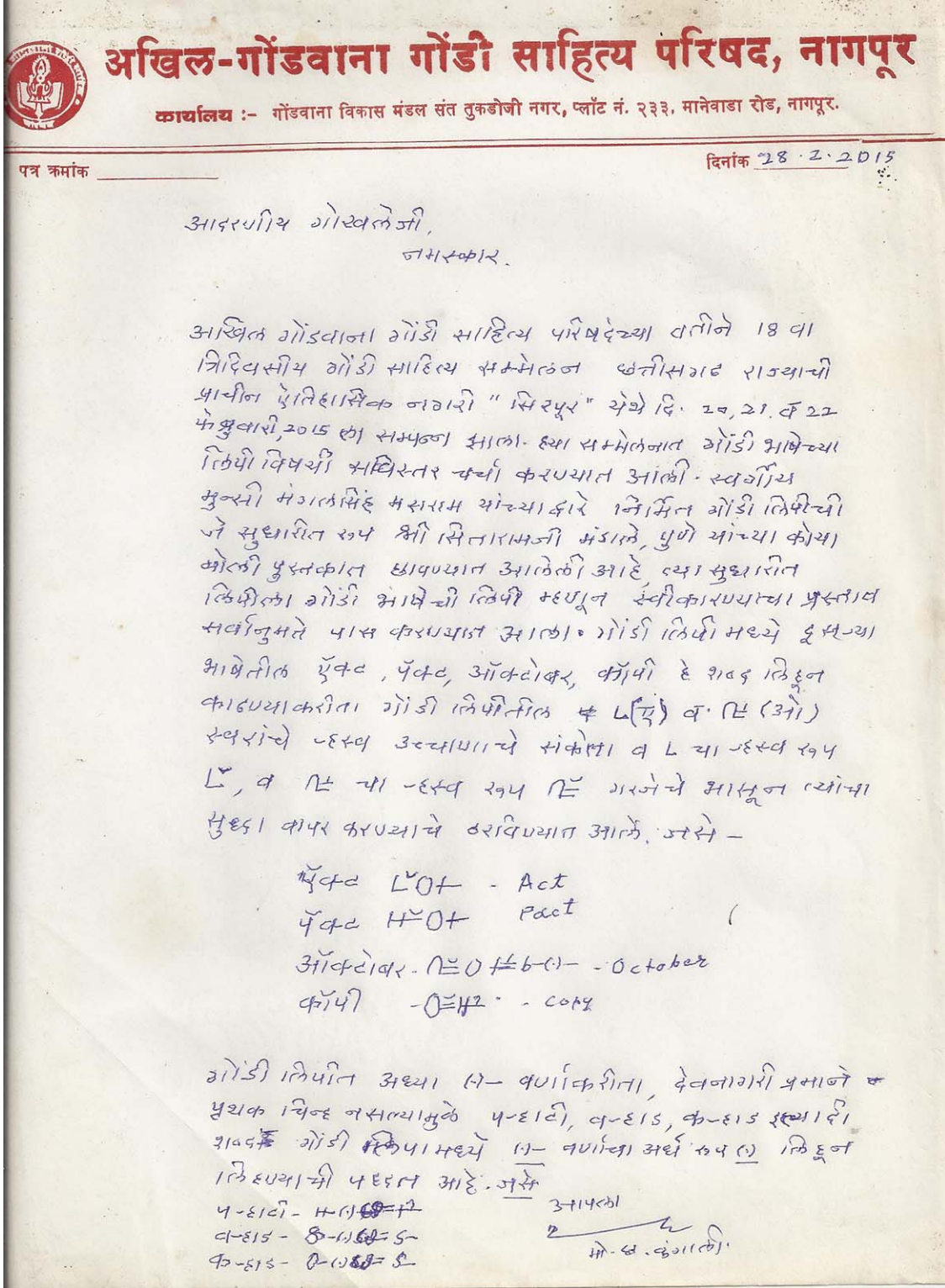
Figure 25: Usage of the sign CANDRA. Contents of letter described in section 4.4.

Figure 26: A calendar showing names for days of the week in Gondi and dates in Gondi digits.
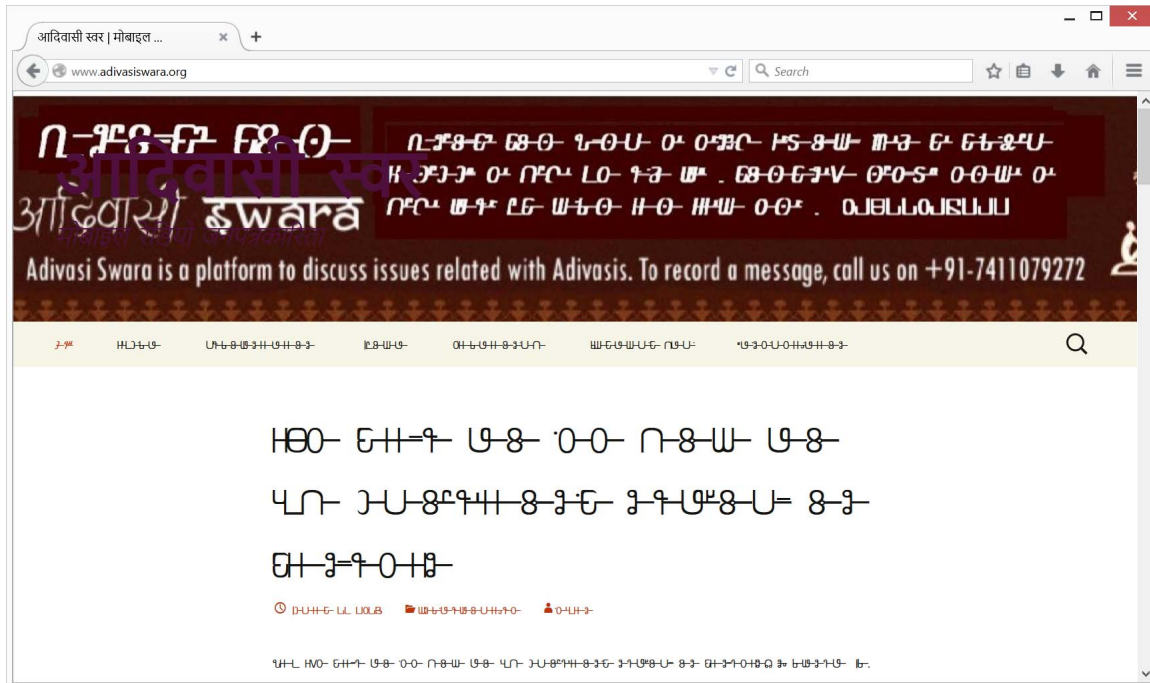
Figure 27: Screen-shot of adivasiswara.org showing content in Gondi script. The text in the banner is valid Gondi content. The text in the main frame of the site, however, is invalid: the script representing meaningless sequences of Gondi letters as the underlying text is Latin-script content in English. The Gondi text is represented using a server-side font based upon the Latin encoding that contains Gondi glyphs. The site appears to be a work-in-progress and is awaiting proper support for the Gondi script in Unicode in order to deliver proper content.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal to Encode the Masaram Gondi Script in Unicode*
2. Requester's name: *Script Encoding Initiative (SEI) / Anshuman Pandey (anshuman.pandey@berkeley.edu)*
3. Requester type (Member body/Liaison/Individual contribution): *Liaison contribution*
4. Submission date: *2015-04-15*
5. Requester's reference (if applicable):
6. Choose one of the following:
     This is a complete proposal: *Yes*
     (or) More information will be provided later:

B. Technical – General

1. Choose one of the following:
     a. This proposal is for a new script (set of characters): *Yes*
          Proposed name of script: *Masaram Gondi*
     b. The proposal is for addition of character(s) to an existing block:
          Name of the existing block:
2. Number of characters in proposal: *75*
3. Proposed category (select one from below - see section 2.2 of P&P document):
     A-Contemporary   **X**   B.1-Specialized (small collection)   B.2-Specialized (large collection)
     C-Major extinct   D-Attested extinct   E-Minor extinct
     F-Archaic Hieroglyphic or Ideographic   G-Obscure or questionable usage symbols
4. Is a repertoire including character names provided? *Yes*
     a. If YES, are the names in accordance with the "character naming guidelines"
          in Annex L of P&P document? *Yes*
     b. Are the character shapes attached in a legible form suitable for review? *Yes*
5. Fonts related:
     a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?
     Anshuman Pandey
     b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
     Anshuman Pandey (anshuman.pandey@berkeley.edu)
6. References:
     a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? *Yes*
     b. Are published examples of use (such as samples from newspapers, magazines, or other sources)
     of proposed characters attached? *Yes*
7. Special encoding issues:
     Does the proposal address other aspects of character data processing (if applicable) such as input,
     presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)? *Yes*

8. Additional Information:
Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

---

[1] Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? *Yes*
   If YES explain *Replaces L2/10-207, L2/12-235, L2/15-005*
2. Has contact been made to members of the user community (for example: National Body,
   user groups of the script or characters, other experts, etc.)? *Yes*
   If YES, with whom? *Mukund Gokhale (Script Research Institute, Pune)*
   Motiravan Kangle (Akhil Gondwana Gondi Sahitya Parishad, Nagpur)
   If YES, available relevant documents:
3. Information on the user community for the proposed characters (for example:
   size, demographics, information technology use, or publishing use) is included? *Yes*
   Reference:
4. The context of use for the proposed characters (type of use; common or rare) *Common*
   Reference: *Used for writing the Gondi language in India*
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
   in the BMP? *N/A*
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *Yes*
8. Can any of the proposed characters be considered a presentation form of an existing
   character or character sequence? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
9. Can any of the proposed characters be encoded using a composed character sequence of either
   existing characters or other proposed characters? *No*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
10. Can any of the proposed character(s) be considered to be similar (in appearance or function)
    to, or could be confused with, an existing character? *No*
    If YES, is a rationale for its inclusion provided?
    If YES, reference:
11. Does the proposal include use of combining characters and/or use of composite sequences? *Yes*
    If YES, is a rationale for such use provided? *Yes*
    If YES, reference: *Combining signs*
    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
    If YES, reference:
12. Does the proposal contain characters with any special properties such as
    control function or similar semantics? *Yes*
    If YES, describe in detail (include attachment if necessary) *Virama;*
    see text of the proposal

13. Does the proposal contain any Ideographic compatibility characters? *No*
    If YES, are the equivalent corresponding unified ideographic characters identified?
    If YES, reference: