The Road to Ethiopic TeX

Abass Andulem

In the olden days, when an Ethiopian scribe wrote a manuscript, the scribe used *bĕrāna* (made up of a goat or a sheep skin) to write on, a scribe's pen or *bĕrĕ* (made up of a reed), and ink (made up of different substances)[1]. After he/she gathered of all necessary materials, it was the scribe's responsibility to write and typeset a whole manuscript. The number of characters in Ethiopic alphabet was not a critical problem for the scribe although the entire process of writing a manuscript was time consuming and cumbersome.

When a modern printing press was established in Ethiopia at the end of the nineteenth century, new methods of printing books became easy, reliable, and fast. Old manuscripts that had been used in churches and in some other institutions were gradually replaced by newly printed books. Despite numerous technological advances in typesetting and in the printing press, the large number of Ethiopic characters is still a fundamental problem. There are 231 letters in the Ethiopic alphabet excluding special characters, numbers, and punctuation marks. As a result of that, the Ethiopic typewriter is extremely difficult to use compared with the English typewriter. Furthermore, this problem is exacerbated in computer programming where hardware restrictions are common in many applications.

There are very few application systems known to the author that are designed with the Ethiopic alphabet. Some of the output of these word processors, as observed from different advertisements, leaves much to be desired. Since there hasn't been published literature on any of these application systems, not much can be said about them.

The idea of designing a word processor that can produce high quality Ethiopic documents led us to begin a project, called "E^{T}HTeX". This project was initiated in the Fall of 1987 during a discussion with Dr. Brian Bourgeois at the University of Houston-Downtown. At that particular time, there was not a clear plan how this project would be carried out. However, Dr. Victor Espino's introduction of the TeX typesetting system in our university and Dr. Bourgeois' suggestion that the TeX typesetting and METAFONT system could be helpful to develop the system solved the fundamental problem.

The purpose of this article is to introduce the Ethiopic font that is designed with the METAFONT system, to discuss the effort that has continued to expand TeX's application into Ethiopic languages, and to raise some of the critical problems that have been encountered along the way.

Ethiopic Alphabet

According to the inscription on the monolith of Aksum, at Tigre province in Northern Ethiopia, Ethiopic script has been known for at least 2000 years[2]. The Ethiopic script has undergone several changes through the years. In particular, the introduction of a modern printing press in the country transferred the hand-lettering to printing which added a new style to the Ethiopic script. The usage of the Ethiopic alphabet has much in common with that of English. In the Ethiopic language, letters are written from left to right, words are separated by two dots (sometimes square dots or a blank), and sentences are separated by four dots (sometimes square dots). Basically, there are seven vowel sounds (*ä, u, i, a, e, ee, o*), representations of which are appended to each consonant letter. Therefore, unlike in English, the vowels are not independent of the consonant letters in written Ethiopic. For example, '*ѡ*', represents the consonant 's' sound. The sounds sä, su, si, sa, se, see, so are represented respectively by *ѡ*, *ѡ*, *ʻl*, *ʻl*, *ʻl*, *ʳ*, and *ʾ*. If we want the word TeX in Ethiopic alphabet, we write it as "ተሕ" in which the vowels sounds are expressed by the two consonant letters. The font table shows the 33 consonant letters families in which each family consists of seven characters.

Today, the dominant language Amharic, the official language of Ethiopia, uses Ethiopic script. The Amharic alphabet is made up of 268 characters consisting of 33 consonants each with 7 vowel sounds and of 37 rarely used additional characters. This number excludes numerals and punctuation marks. Presently, both Ethiopic numbers and Arabic numerals are used in Amharic languages for various applications.

Designing Ethiopic Script

Designing the Ethiopic font with METAFONT has been the primary goal in this project. Learning the language, using it to design the font, and becoming accustomed to the METAFONT system, were the day-to-day experiences. One of the biggest problems in this process was that of figuring out the so called "standard Ethiopic script" that would be used as a model. There are various types which are used in typesetting, but the distinction can be classified neither as stylistic nor as traditional. Presumably, this problem can be overcome in the future.

Ethiopic alphabet and its phonetic representation

U	[hä]	U-	[hu]	¥	[hi]	¥	[ha]	¥	[he]	U	[hee]	U	[ho]
∧	[lä]	∧-	[lu]	∧	[li]	∧	[la]	∧	[le]	∆	[lee]	∧-	[lo]
ih	[ḥä]	ih-	[ḥu]	ih	[ḥi]	ih	[ḥa]	ih	[ḥe]	ih	[ḥee]	ih	[ḥo]
ധo	[mä]	ധo-	[mu]	ൗ	[mi]	ൗ	[ma]	ൗ	[me]	ൗ	[mee]	ൗ	[mo]
w	[ṣä]	w-	[ṣu]	ʽ	[ṣi]	ʽ	[ṣa]	ʽ	[ṣe]	ൗ	[ṣee]	ൗ	[ṣo]
∠	[rä]	4-	[ru]	∠	[ri]	∠	[ra]	∠	[re]	C	[ree]	C	[ro]
∩	[sä]	∩-	[su]	∩	[si]	∩	[sa]	∩	[se]	∩	[see]	∩	[so]
ñ	[shä]	ñ-	[shu]	ñ	[shi]	ñ	[sha]	ñ	[she]	ñ	[shee]	ñ	[sho]
♦	[ḳä]	♦-	[ḳu]	♦	[ḳi]	♦	[ḳa]	♦	[ḳe]	♦	[ḳee]	♦	[ḳo]
∩	[bä]	∩-	[bu]	∩	[bi]	∩	[ba]	∩	[be]	∩	[bee]	∩	[bo]
✝	[tä]	✝-	[tu]	✝	[ti]	✝	[ta]	✝	[te]	✝	[tee]	✝	[to]
Ŧ	[chä]	Ŧ-	[chu]	Ŧ	[chi]	Ŧ	[cha]	Ŧ	[che]	Ŧ	[chee]	Ŧ	[cho]
┐	[ḫä]	┐-	[ḫu]	┐	[ḫi]	┐	[ḫa]	┐	[ḫe]	┐	[ḫee]	┐	[ḫo]
♭	[nä]	♭-	[nu]	♭	[ni]	♭	[na]	♭	[ne]	♭	[nee]	♭	[no]
¥	[gnä]	¥-	[gnu]	¥	[gni]	¥	[gna]	¥	[gne]	¥	[gnee]	¥	[gno]
∧	[ä]	∧-	[u]	∧	[i]	∧	[a]	∧	[e]	∧	[ee]	∧	[o]
┠	[kä]	┠-	[ku]	┠	[ki]	┠	[ka]	┠	[ke]	┠	[kee]	┠	[ko]
┰	[khä]	┰-	[khu]	┰	[khi]	┰	[kha]	┰	[khe]	┰	[khee]	┰	[kho]
⊕	[wä]	⊕-	[wu]	⊕	[wi]	⊕	[wa]	⊕	[we]	⊕-	[wee]	⊕	[wo]
O	[ä]	O-	[ú]	O	[í]	O	[á]	O	[é]	O	[ée]	P	[ó]
┠	[zä]	┠-	[zu]	┠	[zi]	┠	[za]	┠	[ze]	┠	[zee]	┠	[zo]
Ϊ	[jä]	Ϊ-	[ju]	Ϊ	[ji]	Ϊ	[ja]	Ϊ	[je]	Ϊ	[jee]	Ϊ	[jo]
૬	[yä]	૬-	[yu]	૬	[yi]	૬	[ya]	૬	[ye]	૬	[yee]	૬	[yo]
૨	[dä]	૨-	[du]	૨	[di]	૨	[da]	૨	[de]	૨	[dee]	૨	[do]
ૠ	[djä]	ૠ-	[dju]	ૠ	[dji]	ૠ	[dja]	ૠ	[dje]	ૠ	[djee]	ૠ	[djo]
┐	[gä]	┐-	[gu]	┐	[gi]	┐	[ga]	┐	[ge]	┐	[gee]	┐	[go]
ൗ	[dchä]	ൗ-	[dchu]	ൗ	[dchi]	ൗ	[dcha]	ൗ	[dche]	ൗ	[dchee]	ൗ	[dcho]
m	[ṭä]	m-	[ṭu]	m	[ṭi]	m	[ṭa]	m	[ṭe]	m	[ṭee]	m	[ṭo]
Å	[pä]	Å-	[pu]	Å	[pi]	Å	[pa]	Å	[pe]	Å	[pee]	Å	[po]
θ	[tsä]	θ-	[tsu]	θ	[tsi]	θ	[tsa]	θ	[tse]	θ	[tsee]	θ	[tso]
Å	[ṭsä]	Å-	[ṭsu]	Å	[ṭsi]	Å	[ṭsa]	Å	[ṭse]	Å	[ṭsee]	Å	[ṭso]
∠	[fä]	4-	[fu]	∠	[fi]	4	[fa]	∠	[fe]	Ç	[fee]	C	[fo]
Τ	[pä]	Ŧ	[pu]	Τ	[pi]	Ţ	[pa]	Τ	[pe]	Τ	[pee]	Τ	[po]

The font table shows the Ethiopic alphabet, which was designed with **METRFONT**, and its phonetic representation. The Computer Modern family of fonts served as a raw model almost in every aspect. Fortunately, with slight modification, the CM parameter and driver files (of course under different names) filled a gap that would have required intensive work in defining and generating the font. The font has distinctive characteristics compared with the 'traditional' one. The height of all characters, except the families of "ሽ, ፕ, ሽ, ፕ", is the same. Full calligraphic effects are not added because the 'usual types' don't feature identical calligraphic patterns.

Typesetting Ethiopic Alphabet

One of the serious challenges remaining in this project is that of incorporating the Ethiopic alphabet with the TEX typesetting system. Now, the testing process is underway, but some problems remain unsolved.

  a. Overall, there are more than 256 characters in the Ethiopic alphabet. Both **METRFONT** and TEX can handle 256 characters without problem, but hardware restrictions and the methods of usage require that the number of characters must be adjusted with the keyboard[3].

  b. Let's assume that problem (a) is solved using "ligtable" and "control sequences". Now, for a person who cannot speak and write English, but who wants to use the Ethiopic alphabet to typeset documents, there has to be an "Ethiopic editor". Thus, an interface between the TEX system and the Ethiopian language must be devised for non-English speaking users.

These are the main problems that require extensive effort in order to make the E⟂HTEX project fruitful. Hopefully, in the near future, these problems will be solved.

In the Fall of 1987, we were talking 'how we should start the project', but now we are talking 'how we should solve the remaining task', so that the project would be fruitful. The E⟂HTEX project was done partially as senior project at University of Houston-Downtown. Above all, it wouldn't have been possible to come this far without a profound commitment and genuine participation of Dr. Brian Bourgeois and Dr. Victor Espino, the faculty members of Department of Applied Mathematical Sciences at University of Houston-Downtown.

Bibliography

[1] E. A. Wallis Budge, KT, A HISTORY OF ETHIOPIA, Anthropological, 1966, page 557-560.

[2] E. Ullendorff, THE ETHIOPIANS, Oxford University Press, 1973, page 126-130.

[3] Donald E. Knuth, The TEXbook, Addison-Wesley, 1986 page 43-49.

[4] Donald E. Knuth, The *METAFONT*book, Addison-Wesley, 1986.

Typesetting Modern Greek with 128 Character Codes

Yannis Haralambous and Klaus Thull

> *"Ved dette været,*
> *når det regner*
> *så snør det"*
>
> — *Frøydis Frosk*

In european scripts where diacritical marks are common, there are (at least) two reasons to avoid TEX's accent mechanism in favor of many accented characters.

One is the possible misplacement of accents by dvitype's rounding algorithm; the second is lack or invalidity of hyphenation. For example, large portions of german text may be unhyphenatable, and, given the german inclination to long words, may not be in shape to be typeset at all. Thus, in Europe, the obvious thing to do is: let METRFONT put the accents onto the letters, then access these characters via TEX's ligature mechanism.

Accordingly, the greek fonts created by Silvio Levy[1] have 256 characters each, and are a fine tool to typeset greek texts, ancient as well as modern, except those containing the most recent unique accent " ' " (see below). But alas, there is the commercial world, whose device drivers just cannot do 256-code fonts (even .pxl-fonts were seen on the "Big-Tech" sales exhibition in West Berlin last winter). The free drivers are in better shape generally, but often the commercial ones cannot be disposed of in a hurry. So we decided to reduce these fonts to 128 characters. We kept only the ones strictly necessary for writing modern greek without misusing the \accent primitive. At the same time, we constructed some new fonts, which we describe below.