AMHARIC CHARACTER GENERATOR

Gebre Amanuel Gessesse
Electrical Engineering Department
Addis Ababa University

INTRODUCTION

Computer based word processor systems are being used in applications requiring repetitive typing, text editing and document assembly. A word processing system in general requires: an input device, a keyboard; a means of display, usually a visual display terminal (VDT); file storage; a permanent output device, a printer; and the capability of searching, retrieving and editing stored information. In the office environment, computer word processors have been used for document creation, text editing and typewriter emulation [1]. Computer word processors are also being used as typesetters, [2].

Word processor systems could be designed to function in the different languages. One of the basic difference between word processors working in the different languages is the character generator. This unit converts the character code into a visual image of the character. Several techniques have been devised for character generation [3], [4], [5]. This paper describes the design of an Amharic character generator. Of the several possible techniques only one type, the dot raster generator, has been considered for the design.

THE AMHARIC SYLLABARY

The Geez syllabary with a set of diphthongs added to it is being used as the Amharic syllabary. As a result, in Amharic several symbols represent the same sound. The Amharic syllabary has been the subject of a seminar held in 1980 by the Ethiopian Languages Academy [6]. The syllabary recommended by the seminar is the one used in this paper. Two types of syllabary have been recommended. The difference between the two being one of font. One of the syllabaries has the same font as the one currently in use. This is the one considered for the character generator. To shift to the other syllabary one need only change the font.

The recommended syllabary consists of 27 basic characters with vocalic modifications of order 7 each (see Fig. 1). In addition, for 22 of the basic characters the 4th order diphthong has been included. On top of this there are 10 numerals (i.e. the Arabic numerals) and 18 punctuation marks in the recommendation. Of the latter two have been abandoned as they can be obtained by typing more than one key, and of the four quotation marks only two are realized. This brings the number down to 15. Five mathematical symbols are added to this raising the overall total to 241. A description of the design is contained in the following sections.

THE CHARACTER GENERATOR

Two possible alternatives are considered for the character generator design. The first alternative is based on the fact that the basic character and its vocalic modification, have high font resembalance. In addition the vocalic modifiers of the different basic characters resemble each other. Thus the basic characters and the vocalic modifications which cannot be easily obtained from them, punctuation marks, and numeral are stored in one erasable and programmable read-only memory (EPROM) called, in this paper, the basic EPROM and the vocalic modifiers in another EPROM, called the modifiers EPROM. Thus, some of the characters in the syllabary are displayed using the basic EPROM only while others are displayed using both the basic EPROM and the modifiers EPROM. The outputs of the two EPROMS are combined either through an OR gate or through an XOR gate. The latter alternative will serve to erase parts of the basic character unwanted in the vocalic modification.

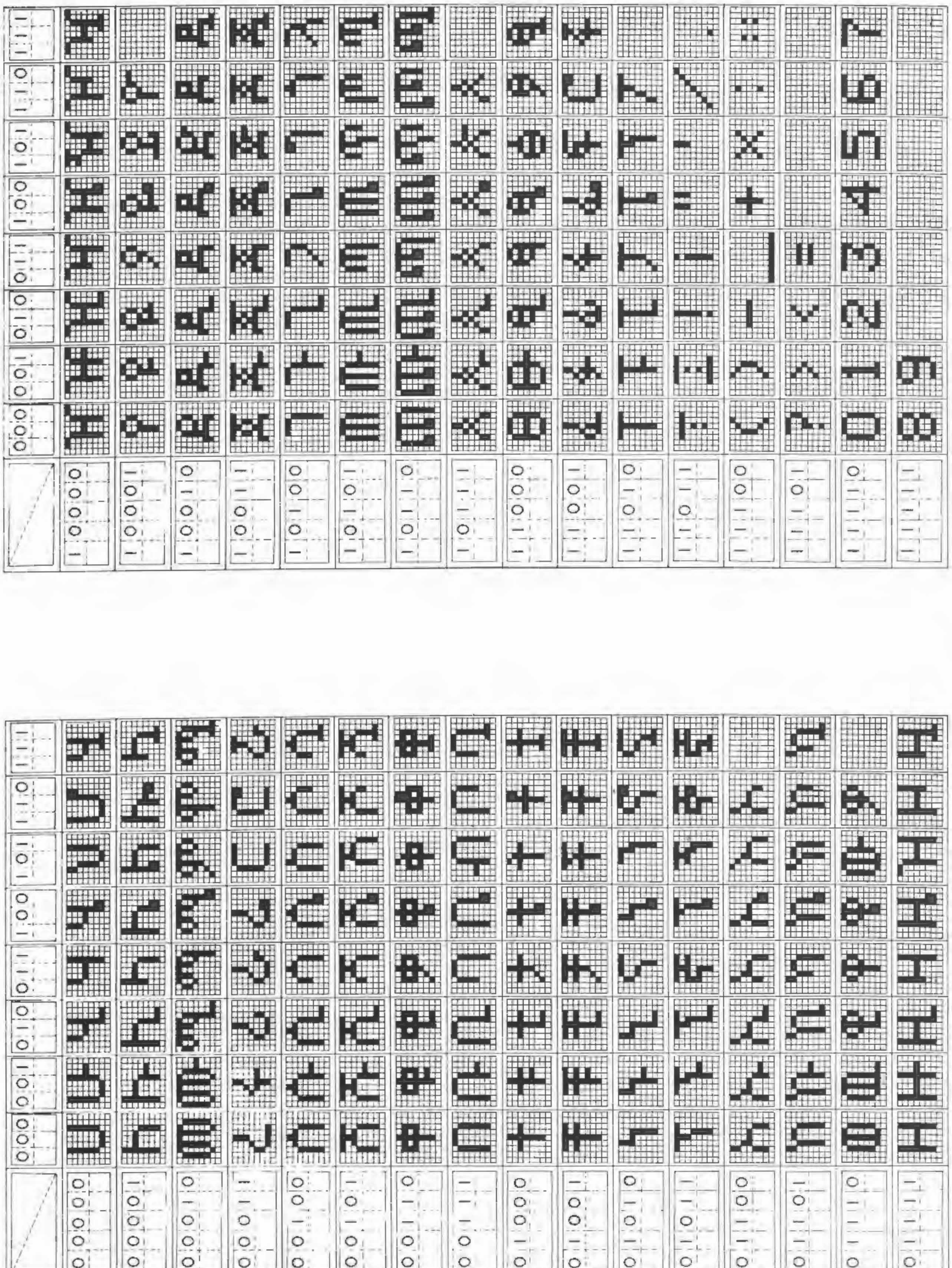

Fig. 1. Recommended Syllabary with Additional Mathematical Symbols

The block diagram of the character generator which works on this principle is shown in Fig. 2. Since in effect there are two character generators and since every position on the screen is a superposition of the outputs of the two generators, duplication of some components becomes a necessity. There are 13 modifiers in the modifier EPROM and this requires a 4 bit code. The basic EPROM has 82 symbols. This requires a 7 bit code. Thus, in all 11-bit code is necessary to access the character generator.

In the second alternative considered, the complete syllabary with punctuation marks and numerals is put within one character generator. The font selected is show in Fig. 1. The matrix size necessary for the selected font is 9X7. A 256X9X7 Eprom will be more than adequate for all the characters, numerals and punctuation marks. There are 15 additional spaces in which more symbols could be added. The character generator is accessed using an 8 bit code only instead of the 11 bits of the previous alternative. The block diagram for the generator is shown in Fig. 3.

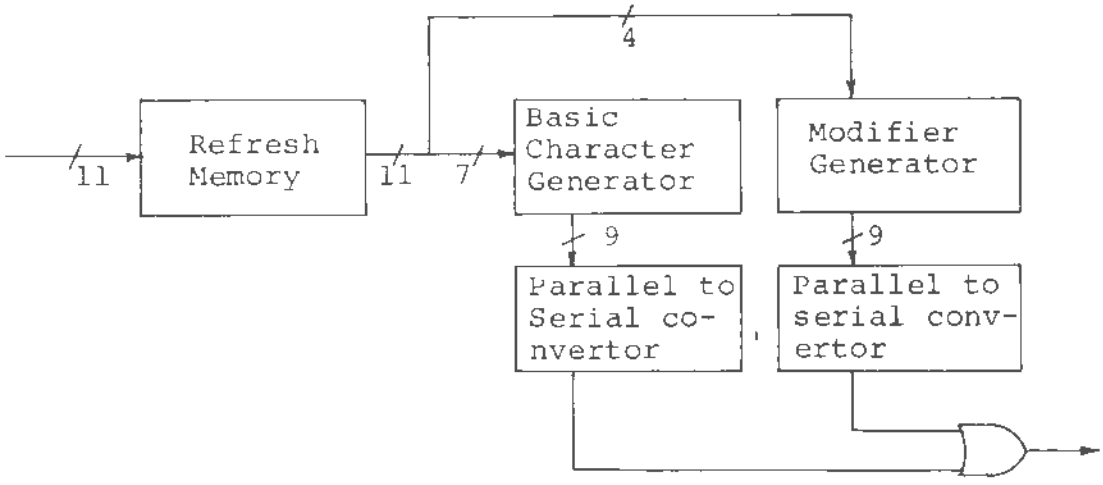

Fig. 2. Character generator using basic EPROM and modifier EPROM

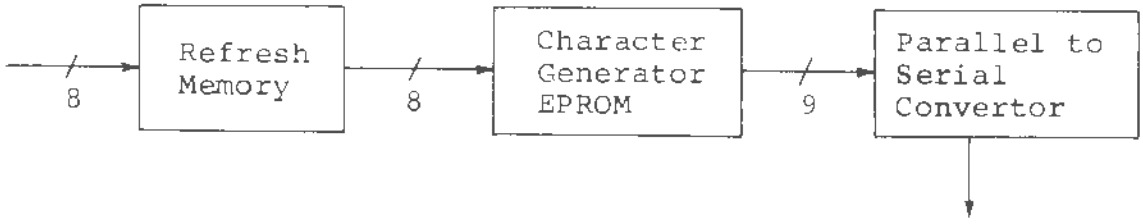

Fig. 3. Character Generator Using One EPROM Only

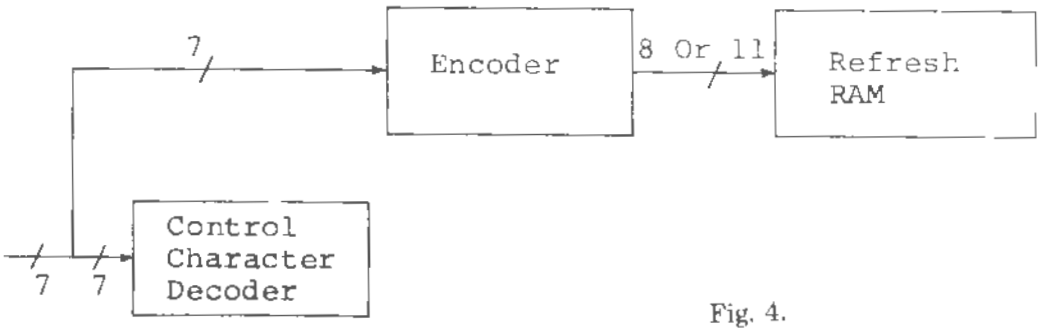

Fig. 4.

THE KEYBOARD INTERFACE

The Amharic typewriter keyboards that are popular at present have 92 symbols of which 9 are Arabic numerals, 8 or 10 are punctuation marks, 63 or 61 are characters, and 12 are vocalic modifiers. With the use of the recommended syllabary the keyboard could be modified so that all basic characters are typed with one key, all vocalic modifications with one or two keys, and all diphthongs with two or three keys. This will simplify encoding.

The keyboard is encoded using a 7-bit code. The non-printing characters in the code are chosen to be identical with the non-printing characters in the American Standard Code for Information Interchange (i.e. ASCII code). The rest of the code is used to encode the printing characters. With this code control characters are decoded using ASCII control character decoders.

The printing characters code could not be used directly to access the character generator. An encoder converting one, two or three seven bit codes into an eleven bit code for the first alternative or an eight bit code for the second alternative is required. The encoder position is shown in Fig. 4.

CONCLUSION

Due to lack of components the system has not been fully realized. The selected font has been printed on a dot matrix printer using graphics mode. An acceptable printout has been obtained but there is still room for further improvement. Certain liberties have been taken to modify the font of certain characters like ቈ . This modification enables to minimize the dot matrix size without loss of the distinctness of the characters involved.

REFERENCES

[1] Hartke, D., Sterling, W., and Shemer, J. Design of Raster Displair Processor for Office Applications. TEEE Transactios on Computers, Vol. C-27, pp. 337-348, April 1978.

[2] Bigelow, C., and Day, D. Digital Typography. Scientific American, Vol. 249, pp. 106-119, August 1983.

[3] Loewe, R.T., Sisson, R.L., and Horowitz, P. Computer Generated Displays. Proceedings of the IRE, Vol. 49, pp. 185-195, January 1961.

[4] Lewin, M.H. An Introduction to Computer Graphic Terminals. Proceedings of the IEEE, Vol. 55, pp. 1544-1552, September 1967.

[5] Sherr, S. Fundamentals of Display System Design. John Wiley and Sons, Inc., New York, 1970, Chapter 6.