Title: Allocating the Pau Cin Hau Scripts in the Unicode Roadmap
Author: Anshuman Pandey (pandey@umich.edu)
Status: Individual Contribution
Action: For consideration by UTC
Date: 2010-03-17

1 Introduction

The Pau Cin Hau writing systems have not yet been encoded in the Universal Character Set nor have they been allocated in the Unicode roadmaps. There are logographic and alphabetic forms of the Pau Cin Hau script. The alphabetic script is described in the preliminary proposal N3781 L2/10-080. Research on the logographic script is ongoing and a preliminary proposal is forthcoming.

2 Background

The Pau Cin Hau writing systems are the ecclesiastical scripts of the Laipian religious tradition. The scripts and the Laipian tradition were developed by Pau Cin Hau (1859–1948) at the turn of the 20th century in what is now Chin State, Myanmar. The script was originally designed to represent Tedim, a language belonging to the northern Kuki-Chin group of the Tibeto-Burman family, but contains characters used for writing other languages of this group.

There are logographic and alphabetic forms of Pau Cin Hau. The logographic form is believed to be the original script revealed to Pau Cin Hau in 1902. It is reported that this script consists of 1,050 characters. In 1931, the logographic script was modified into an alphabetic system consisting of 57 characters. Some characters of both forms resemble those found in Burmese, Latin, Indic scripts, and other writing systems, but these occurances are coincidental. Neither the logographic nor the alphabetic script has a genetic relationship with any other script.

Although practice of the Laipian tradition and usage of the Pau Cin Hau scripts have declined, they are still maintained by the Laipian community in Chin State. The scripts are used in recitation texts and dictionaries.

3 Recommendation

The logographic and alphabetic forms of Pau Cin Hau should be allocated in the Supplemental Multilingual Plane (SMP). Although research indicates that these scripts are currently used by a linguistic community, there is not enough space in the Basic Multilingual Plane (BMP) to accommodate them.

The logographic script will require at minimum 1,050 code points. This number is based upon the character repertoire of a Laipian recitation text that is attributed to the founder, Pau Cin Hau. Analysis of the recitation text indicates the use of diacritic tone marks, which reduces the number of unique logographic characters. Digits and punctuation are also attested. The actual number of characters likely to be proposed for encoding should be nearer to nine hundred.

The alphabetic script will require at minimum 57 code points. As specified in N3781 L2/10-080, this number consists of 30 consonant letters, 7 vowel letters, and 20 tone marks. Digits and punctuation are not attested.