Chinese Character Synthesis Using Metapost

Candy L.K. Yiu, Wai Wong
Department of Computer Science, Hong Kong Baptist University
candyyiu@comp.hkbu.edu.hk wwong@comp.hkbu.edu.hk

March 11, 2003

Abstract The rapid advancement of the Internet and World Wide Web provides an effective means of information exchange. However, there is a very serious problem in exchanging Chinese documents. This is because the total number of Chinese characters which ever existed is unknownable. Furthermore, new characters are being created continually. Therefore, no character set can encode *all* Chinese characters.

Even if a character set could encode all Chinese characters, it is very expensive to create Chinese fonts using typical methods and a fairly large number of Chinese characters would be so rarely used that expense would be very difficult to justify.

One possible solution to this problem is to create an unencoded character according to its composition of strokes and radicals. Several experiments along this line were attempted in the past, but none were very successful. The key reason is that the composition of the strokes and radicals is very complex, and the previous attempts did not effectively divide and resolve the complexity.

Our approach to Chinese character synthesis resolves the complexity in two ways. First, we defined a high-level Chinese character description language, *HanGlyph*. It captures the abstract and topological relation of the strokes. Thus, the character description is compact and can be targeted to a variety of rendering styles. Secondly, we use metapost as our rendering engine to take its advantages of metaness and the ability of specifying paths and solving linear equations.

HanGlyph defines 41 basic strokes, 5 operators and a set of relations. A character is built by combining strokes using the operators recursively. *HanGlyph* allows the user to define macros to represent a stroke cluster which can then be re-used in building more complex characters.

We have developed a front-end to translate *HanGlyph* expressions into metapost programs. We have developed a library of metapost macros to implement the operators, relations and the basic strokes. By varying the parameters to these macros, or redefining the basic stroke macros, Chinese characters in different styles can be formed. Thus, it can create a variety of different fonts from the same *HanGlyph* description.