

Improving user friendliness of Singlish to Sinhala Unicode converters by word prediction

Kapugama K.D.C.G., Lorensuhewa S.A.S. and Kalyani M.A.L.

Department of Computer Science, University of Ruhuna, Matara, Sri Lanka.

Singlish is a method that uses Latin characters (English) for transcribing Sinhala words. Sri Lankan people are more familiar with this method of input than typing with the standard Sinhala keyboard layout (Wijesekara keyboard). There are some software applications to convert such Singlish text into Sinhala Unicode text. In such converters, there is an underlying Transliteration Scheme, where one or more English characters correspond to a single Sinhala character. So, the user interacting with the converter has to remember this transliteration schema when typing. Our main objective is to introduce a method to reduce the need of a user remembering a transliteration scheme and to save keystrokes in typing. In this paper, we propose a method to provide meaningful Sinhala word predictions for the Singlish words typed by the user. This method includes current word and next word prediction. For developing this, we used Ngram models and string based pattern matching. Also, we analyzed the impact of limiting those two ways of predictions to the user's intended domain. Our evaluation results show that our approach can reduce the number of keystrokes required in typing. Also, limiting the word predictions to the user's intended domain further reduces the number of keystrokes. From this kind of a keystroke reduction, the user-friendliness of the input system can be improved.

Keywords: singlish, sinhala unicode, unicode converters, word prediction

**Corresponding author: charakageethal@gmail.com*