**Pedagogic application of regular expressions: Corpus-based online writing tool**

John Blake, University of Aizu

Undergraduate students of computer science are required to submit a short research article in English in order to graduate from the University of Aizu. This is particularly challenging for Japanese speakers who rarely function in English, have had little exposure to English in general and even less exposure to formal research writing.

A corpus-based online writing tool was created to feedback on surface-level errors in their final drafts. Regular expressions are harnessed to match particular sequences of characters or words. On matching, scripts for pre-determined tasks are executed to automatically provide feedback.

A learner corpus of over 300 draft research articles was compiled. Errors in the corpus were identified manually and automatically. Errors were classified into one of five categories (i.e. accuracy, brevity, clarity, objectivity and formality) mirroring the content of the in-house thesis writing course. Where possible, regular expressions and easy-to-understand actionable advice were created for each error. Client-side scripts were written that automatically colour, highlight, and display advice on the matched strings. A web interface was created to enable text submission around the clock. Further language feature detection functions (e.g. modality, tense and voice) were added to raise users awareness of expected generic conventions.

The writing tool reduces the need for teachers to provide feedback on commonly-occurring surface-level errors and gives students the opportunity to receive genre-specific advice on their drafts. The presentation concludes by arguing that presence of false positive results increases the pedagogic value of this tool.