Global Design Seminar

Text Normalization for Text-to-Speech Synthesis

Richard Sproat, Research Scientist, Google New York
(http://rws.xoba.com/newindex/cv.pdf)

A website for a Google-sponsored event back in April invited attendees to "Google's 1st European Doctoral Workshop on Speech Technology, taking place on 29 & 30 April 2014 at our Google office in London, UK." Any competent speaker of English reading this phrase aloud would read "29 & 30 April 2014" as something like "twenty ninth and thirtieth (of) April, twenty fourteen", "1st" as the ordinal number "first", and "UK" as "u k" (rather than "uck"). Yet, except perhaps for "UK", none of the written expressions are found in any dictionary; speakers must translate these "non-standard words" into ordinary words as part of their process of converting between text and speech. Text normalization is the problem of building computational algorithms that mimic this process --- for example as part of a text-to-speech synthesizer. Since there are many classes of "non-standard words" --- besides dates and times, there are currency amounts, measures, abbreviations, etc. --- building a wide-coverage text normalization system for a language is labor intensive. For some languages, such as Russian with its complex inflectional morphology, the process can be especially difficult.
In this talk I outline the problem, and give specific examples of why it is hard, and why large parts of text normalization systems are still constructed by hand, rather than trained using machine-learning algorithms. Nevertheless, there are some areas of the problem where machine-learning has made inroads, and I will discuss some of our own research along those lines.

7月24日（木）16:40～18:10 ※会場が変更になりました

工2号館 4階 245号講義室（部屋番号：401・402）

問合せ先：pim@gcl.i.u-tokyo.ac.jp

東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo