

Corpus-based Mini Workshop

Motivation

2009.08.30.

Sungkeun Cho

skcho@ropas.snu.ac.kr

Cases using Corpus

- Language Translation
 - Automatic dictionary creation
 - Using specialized corpus to improve translation quality
- Spoken Language
 - Automatic language identification

- [1] S. Velupillai, M. Hassel, and H. Dalants. Automatic dictionary construction and Identification of parallel text pairs. In *UCCTS*, 2008.
- [2] Michael Wilkinson. Using specialized corpus to improve translation quality. In *Translation Journal*, 2005
- [3] Y. Muthusamy, R. Cole, and B. Oshika. The OGI multi-language telephone speech corpus. In *ICSLP*, 1992.

Cases using Corpus

- Robotics
 - Route instruction
- Media Arts
 - Finding happiness

- [4] G. Bugmann, E. Klein, S. Lauria, and T. Kyriacou. Corpus-based robotics: A route instruction example. In *Intelligent autonomous systems 8*, 2004.
- [5] R. Mihalcea and H. Liu. A corpus-based approach to finding happiness. In *AAAI*, 2006.

Corpus-based Shape Analysis

- Corpus in our project
 - The shapes of structures used in programs
- Encountered Problems
 - How to make corpus efficiently?
 - It is almost impossible to make it manually.
 - We are thinking about something that helps us to make the corpus.
 - How to use gathered corpus usefully?

4 Papers

1. M. Colon, S. Sankaranarayanan, and H. Sipma. Linear invariant generation using non-linear constraint solving. CAV'03.
2. S. Sankaranarayanan, H. Sipma, and Z. Manna. Scalable analysis of linear systems using mathematical programming. VMCAI'05.
3. S. Gulwani, S. Srivastava, and R. Venkatesan. Program analysis as constraint solving. PLDI'08.
4. A. Gupta, R. Majumdar, and A. Rybalchenko. From tests to proofs. TACAS'09.

This papers may give us some hints.

Let's start!