

Nanyang Technological University Update

Francis **Bond**
and many, many more

Computational Linguistics Lab (凡士研)
Linguistics and Multilingual Studies,
Nanyang Technological University

[<bond@ieee.org>](mailto:bond@ieee.org)

DELPH-IN 2017-08-07



- NTU

- ▶ Wordnets: Multi, Jpn, English, Chinese, Bahasa, *Myanmar, Kristang*
 - ★ the Collaborative Interlingual Index
- ▶ Grammars:
 - ★ Indonesian (INDRA)
 - ★ Chinese (Zhong [||]: Mandarin and Cantonese)
 - ★ Japanese (JACY) — bug fixes, start of enhancements
- ▶ Tools:
 - ★ OMW version II, LMF checking
 - ★ intsem.fx, visualkopasu
- ▶ Intelligent Tutoring and Feedback on Errors
- ▶ English Essay writing help

Open Multi-Lingual Wordnet

- Collaborative InterLingual Index (CILI)
 - ▶ Initial testing pretty much done
 - ▶ Wow, there are a lot of interesting things you can do
 - ★ And we check for/document many of them
- More languages and words
 - ▶ Sesotho, Tswana, Venda, Xhosa, Zulu, Gaelic, Abui, Moroccan Arabic, ...
 - ▶ National Cancer Institute Thesaurus, Colloquial Wordnet
 - ▶ 10,000+ new English synsets (enwordnet)
- More links
 - ▶ Interlanguage sense level links (Harmony Project)

NTU multilingual corpus

- Small, deeply analysed corpus
 - ▶ 6,000+ sentences x 3 languages (cmn, eng, jpn); 2,000 in Indonesian
 - ★ Mainichi Newspaper (NICT translations)
 - ★ Sherlock Holmes ([DANC](#), [SPEC](#), [REDH](#)[eng])
used in ACL workshop on Dialogue
 - ★ Cathedral and the Bazaar (many languages)
 - ★ Spider's thread 「蜘蛛の糸」
 - ★ Singapore Tourist data (plus Korean, Viet, Indo)
 - ▶ [Hand alignment](#), [WordNet tagging](#), Treebanking, Sentiment, ...
- Revised the DB structure; second round of wordnet tagging
 - ▶ Much behind schedule
 - ▶ Wordnet extension is being done in parallel
- Tagged Italian ([SPEC](#))
ready to start Bulgarian, Dutch, German, Polish, Abui (shared task)

New Projects

- A collaborative network for usage-based research on lesser-studied languages (JSPS: Tokyo University of Foreign Studies) — building large Indonesian corpus
- Digital Mapping the Literary Epigraph: Quantitative analysis of literary influence using network theory and thousands of epigraphs (NTU)
- Year four of research with Fuji-Xerox
 - ▶ WSD, Sentiment and transfer learning

Sabbatical

- UW kindly hosted me and Zhenzhen
 - ▶ Wordnet consolidation — 4.0 release, move to NTU
 - ▶ Machine translation/ISF/QMRS
 - ▶ Zhong
 - ▶ Jacy 維新 “the Jacy Restoration” (with Tsuneko)
- We had a mini meeting in January
Many thanks to Glenn for a lovely location
<http://moin.delph-in.net/CapitolHillTop>
useful discussions all round
- I joined the *Sound of the Baskervilles* and the *Hounds of the Internet*



Members

- Staff:
E umamaheswari Vasanthakumar, Takayuki Kuribayashi, Wenjie, Irfan, ...
- Visitors/Interns:
Hiroki Nomoto, Min Huan (just promoted to Assoc Prof in Southwest University)
- Students:
Zhenzhen, David, Tuan Anh, Mike, Jim, Giulia, Tommaso, Luis, Hannah, Martin

Welcome to GWC 2018



The 9th Global WordNet Conference

Nanyang Technological University (NTU), Singapore

8–12 January, 2018

3 (2.5) day Conference

Shared task: Annotating SPEC

Workshop on Distributional Semantics

Workshop on Technologically Enhanced Learning

We welcome you all!

