

Andy Way, Barbara Gawronska (Eds.)

# TMI 2007

Proceedings of the 11th International Conference on  
Theoretical and Methodological Issues in Machine  
Translation

Skövde, 7 – 9 September 2007



Skövde University Studies in Informatics  
2007:1

## 11th International Conference on Theoretical and Methodological Issues in Machine Translation

### Programme

#### Thursday 6<sup>th</sup> September

10.30 – 18.00 Conference excursion

18.00 – 20.00 Conference Registration

19.00 – 21.00 Welcome Reception

#### Friday 7<sup>th</sup> September

8.00 – 9.30 Conference Registration

9.15 – 9.45 Introductory Remarks

9.45 – 10.45 Invited Talk: *Rule-based and Statistical Machine Translation with a Focus on Swedish* -- Anna Sgvall-Hein

10.45 – 11.15 COFFEE

11.15 – 11.50 *How Phrase Sense Disambiguation outperforms Word Sense Disambiguation for Statistical Machine Translation* -- Marine Carpuat and Dekai Wu

11.50 – 12.25 *Exploiting Source Similarity for SMT using Context-Informed Features* -- Nicolas Stroppa, Antal van den Bosch and Andy Way

12.25 – 13.00 *Towards Hybrid Quality-Oriented Machine Translation. On Linguistics and Probabilities in MT* -- Stephan Oepen, Erik Velldal, Jan Tore Lnning, Paul Meurer and Victoria Rosen

13.00 – 14.15 LUNCH

14.15 – 14.50 *BLEUATRE: Flattening Syntactic Dependencies for MT Evaluation* -- Dennis Mehay and Chris Brew

14.50 – 15.25 *A Cluster-Based Representation for Multi-System MT Evaluation* -- Nicolas Stroppa and Karolina Owczarzak

15.25 – 16.00 *Combining translation models in statistical machine translation* -- Jesus Andrs-Ferrer, Ismael Garca-Varea and Francisco Casacuberta

16.00 – 16.30 COFFEE

16.30 – 17.05 *A Greedy Decoder for Phrase-Based Statistical Machine Translation* -- Philippe Langlais, Alexandre Patry and Fabrizio Gotti

17.05 – 17.40 *Word Reordering in Statistical Machine Translation with a POS-Based Distortion Model* -- Kay Rottmann and Stephan Vogel

Saturday 8<sup>th</sup> September

9.15 – 9.50 *Learning Bilingual Semantic Frames: Shallow Semantic Parsing vs Semantic Role Projection* -- Pascale Fung, Zhaojun Wu, Yongsheng Yang and Dekai Wu

9.50 – 10.15 *Reducing Human Assessment of Machine Translation Quality to Binary Classifiers* -- Michael Paul, Andrew Finch and Eiichiro Sumita

10.15 – 10.50 *A New Method for the Study of Correlations between MT Evaluation Metrics and Some Surprising Results* -- Paula Estrella, Andrei Popescu-Belis and Maghi King

10.50 – 11.20 COFFEE

11.20 – 11.55 *Alignment-Guided Chunking* -- Yanjun Ma, Nicolas Stroppa and Andy Way

11.55 – 12.30 *Sub-Phrasal Matching and Structural Templates in Example-Based MT* -- Aaron Phillips

12.30 – 13.05 *Extracting Phrasal Alignments from Comparable Corpora by Using Joint Probability SMT Model* -- Tadashi Kumano, Hideki Tanaka and Takenobu Tokunaga

13.05 – 14.20 LUNCH

14.20 – 15.50 Panel Session: *Is MT In Crisis?* (Moderator: Steven Krauwer)

15.50 – 16.20 COFFEE

16.20 – 17.50 Demonstration & Poster Session

- *Demonstration of the German to English METIS-II MT System* -- Michael Carl, Sandrine Garnier and Paul Schmidt
- *Automatic induction of shallow-transfer rules for open-source machine translation* -- Felipe Sánchez-Martínez and Mikel Forcada
- *Support Vector Machine Based Orthographic Disambiguation* -- Eiji Aramaki, Takeshi Imai, Kengo Miyo and Kazuhiko Ohe
- *EBMT Based on Finite State Automata Transfer Generation* -- Ren Feiliang, Zhang Li, Hu Minghan and Yao Tianshun
- *Demonstration of the Spanish to English METIS-II MT system* -- Maite Melero and Toni Badia
- *An Assessment of Language Elicitation without the Supervision of a Linguist* -- Alison Alvarez, Lori Levin, Robert Frederking and Jill Lehman
- *Phrase Alignment Based on Bilingual Parsing* -- Akira Ushioda
- *Demonstration of the Dutch-to-English METIS-II MT System* -- Peter Dirix, Vincent Vandeghinste and Ineke Schuurman
- *Hand in Hand: Automatic Sign Language to Speech Translation* -- Daniel Stein, Philippe Dreuw, Hermann Ney, Sara Morrissey and Andy Way
- *Combining Resources for Open Source Machine Translation* -- Eric Nichols, Francis Bond, Darren Scott Appling and Yuji Matsumoto
- *Demonstration of the Greek to English METIS-II MT System* -- Sokratis Sofianopoulos, Vassiliki Spilioti, Marina Vassiliou, Olga Yannoutsou and Stella Markantonatou

- *Reordering via N-Best Lists for Spanish-Basque Translation* -- Germán Sanchis and Francisco Casacuberta

Sunday 9<sup>th</sup> September

9.15 – 10.15 Invited Talk: *Statistical MT from TMI-1988 to TMI-2007: What Has Happened?*  
-- Hermann Ney

10.15 – 10.50 *Theoretical and Methodological Issues Regarding the Use of Language Technologies for Patients with Limited English Proficiency* -- Harold Somers

10.50 – 11.20 COFFEE

11.20 – 12.55 *Capturing Translational Divergences with a Statistical Tree-to-Tree Aligner* -- Mary Hearne, John Tinsley, Andy Way and Ventsislav Zhechev

12.55 – 12.30 *Breaking the barrier of context-freeness. Towards a linguistically adequate probabilistic dependency model of parallel texts* -- Matthias Buch-Kromann

12.30 – 13.05 Final Remarks

13.05 – 14.30 LUNCH

15.00 Bus to Copenhagen for *MT Summit XI* arriving at 20.30 (approx.)

# TMI-07

## 11th International Conference on Theoretical and Methodological Issues in Machine Translation

### Notes from the Programme Chair

The first International Conference on Theoretical and Methodological Issues in Machine Translation took place in upstate New York in 1985. More than 20 years later, we are pleased to be holding the 11<sup>th</sup> Conference in the series in Skövde, Sweden, from 7—9 September 2007.

As Programme Chair, I was absolutely thrilled that we received 63 submissions to the conference. From this, following a huge amount of hard work by the Programme Committee (listed on the next page), this has been whittled down to the programme that stands before you, consisting of 17 oral presentations, and 12 posters and demonstrations. As is traditionally the case at TMI, we have papers on a wide range of subjects, including statistical MT, example-based MT, rule-based MT, hybrid MT, MT evaluation, open source MT, alignment, inducing bilingual lexical information, parallel and comparable corpora, as well as multilingual applications.

In addition, we have two renowned keynote speakers in Anna Sågvald-Hein (Uppsala) and Hermann Ney (RWTH Aachen). Finally, we have a panel session chaired by Steven Krauwer (Utrecht) entitled *Is MT in Crisis?* which promises to be both fun and informative.

There are a number of people I would like to thank. Firstly, this conference would not be happening at all without the enormous effort provided by Barbara Gawronska and her team in Skövde. We would especially like to thank the School of Informatics for having the Proceedings printed for us. Secondly, my heartfelt thanks go to the members of the Programme Committee, who, as usual, did sterling work over and above what might have reasonably been expected from them. Thirdly, as you all know we decided a long time ago to co-locate the conference temporally with MT Summit XI, so I would like to thank Bente Maegaard and Viggo Hansen for their support in making this possible, and for supporting the discounted registration rates for both conferences. Finally, I would like to thank my students Karolina Owczarzak, Yanjun Ma, John Tinsley and Sara Morrissey for their help in preparing the website, proceedings and letters of support to enable attendees to travel.

We have an excellent programme assembled. I hope you enjoy what you hear.

Andy Way.