Demonstration of the German to English METIS-II MT System

Michael Carl, Sandrine Garnier and Paul Schmidt
Institut für Angewandte Informationsforschung,
66111 Saarbrücken, Germany
{carl,sandrine,paul}@iai.uni-sb.de

Starting in October 2004, METIS-II is the continuation of the successful project METIS I (IST-2001-32775). Like METIS I, METIS II aims at translation of free text input based on rule-based, statistical and pattern matching methods. The METIS-II project has four partners, translating from their 'home' languages Greek, Dutch German, and Spanish into English.

The following goals and premises were defined for the project:

1. use of a bilingual hand-made dictionary
2. use of 'basic' NLP tools and resources
3. different tag-sets for SL and TL possible
4. translation units below the sentence border
5. use a monolingual target language corpus
6. no bilingual corpus required

In particular, the availability of the monolingual target language corpus makes METIS-II a data-driven MT system. However, parallel corpora as in SMT/EBMT are not required. For our German-to-English METIS-II system we have designed and implemented an architecture which uses rule-based devices to generate sets of partial translation hypotheses and a statistical *Ranker* to evaluate and retrieve the best hypotheses in their context[1]. Similar architectures have already been suggested as EBMT systems (Sato and Nagao, 1990), for instance with their MBT2 system. Methods to integrate knowledge bases and statistics have also been explored in (Knight et al., 1994) and recently in the LOGON-project (Oepen et al., 2007) which uses statistical feature functions to select the best rule-induced structures at various stages during processing.

In the German-to-English METIS-II system, rule-based devices generate an acyclic AND/OR graph which allows for compact representation of many different translations while the *Ranker* is a beam search algorithm which tries to find most likely paths through the AND/OR graph. The architecture consists of the following five steps:

1. The *Analyser* lemmatises and morphologically analyses the SL sentence. It produces a (flat) grammatical analysis of the sentence, detecting phrases and clauses and potential subject candidates. The *Analyser* uses the linguistic technology available at the IAI.
2. *Dictionary Lookup* matches analysed SL sentence on the transfer dictionary and retrieves TL equivalences. This procedure retrieves ambiguous and/or overlapping entries and stores them in a partial OR graph. Our German to English dictionary contains more than 629.000 single and multi-word entries. Since matching proceeds on morphemes and lemmatised forms, a sophisticated compilation of the dictionary into a database is required. As described in (Carl and Rascu, 2006), the matching procedure is also suited to retrieve discontinuous entries.
3. The *Expander* inserts, deletes, moves and permutes items or chunks in the graph generated by the *Dictionary Lookup* according to TL syntax. The *Expander* is a rule-based device and extends the AND/OR graph with further partial translation hypotheses. It is called *Expander* because it expands the search space with additional paths. The operations of the *Expander* and its modifications on the graph are such that each path through the graph consumes exactly once the translation(s) of each word of the source language sentence. For our German-to-English implementation we have currently ca. 50 rules.
4. The *Ranker* is a beam search algorithm that iteratively traverses the AND/OR graph and computes the most likely translations in a log-linear fashion (Och and Ney, 2002). Unlike a usual statistical decoder (Koehn, 2004) — but similar to the method suggested by (Knight et al., 1994) — our *Ranker* traverses the search graph to grade alternative paths and outputs a list of the n-best translations. The *Ranker* itself does not modify the graph. It does not permute chunks or items and it does not generate additional paths which are not already contained in

[1] A full description of the system is provided in (Carl, 2007).

the graph.

5. A *Token Generator* generates surface wordforms from the lemmas and PoS tags. The Token Generator has been described in (Carl et al., 2005).

The *Ranker* and the *Token Generator* are trained on the British National Corpus (BNC[2]). It is a collection of tagged texts making use of the CLAWS5 tag set which comprises roughly 70 different tags[3]. The heuristic functions of the *Ranker* are trained with the CMU-language modelling toolkit.

Evaluation In a first experiment we have tested the system on four languages (Dutch, German, Greek and Spanish) into English based on 50 sentences for each of the languages. The results are shown in table (1). A separate set of *Expander* rules was developed for each source language, consisting of five rules for Greek up to approx. 20 rules for German.

Language	BLEU	NIST
Dutch	0.4034	6.4489
Spanish	0.3701	5.7304
Greek	0.2138	5.1220
German	0.1671	3.9197

Table 1: Results of first Experiment

Another set of evaluations was conducted one a German test set of 200 sentences after enhacing the *Dictionary Lookup*, *Expander*, and *Ranker* modules. Our best results are shown in the first line in table (2). However, they (still) lag behind those produced by Systran (Babelfish) on the same test set, as shown last line in table (2).

NIST	BLEU	token model	tag model
5.3193	0.2231	5M-n3	5M-n7
6.3644	0.3133	—	—

Table 2: Results of 200 test translations

A full description of the system is provided in (Carl, 2007).

References

Michael Carl and Ecaterina Rascu. 2006. A dictionary lookup strategy for translating discontinuous phrases. In *Proceedings of the 11th EAMT Conference*, pages 49–58, Oslo, Norway.

Michael Carl, Paul Schmidt, and Jörg Schütz. 2005. Reversible Template-based Shake & Bake Generation. In *In Proceedings of the Example-Based Machine Translation Workshop held in conjunction with the 10th Machine Translation Summit*, pages 17–26, Phuket, Tailand.

Michael Carl. 2007. METIS-II: The German to English MT System. In *Proceedings of the 11th Machine Translation Summit*, Copenhagen, Denmark.

Kevin Knight, Ishwar Chandler, Mattew Haines, Vasileios Hatzivassiloglou, Edouard Hovy, Masayo Ida, K. Luk, Steve, Akitoshi Okumura, Richard Whitney, and Kenji Yamada. 1994. Integrating Knowledge Bases and Statistics in MT. In *Proceedings of the AMTA*.

Philipp Koehn. 2004. Pharaoh: A beam search decoder for phrase-based statistical machine translation models. In *AMTA*, pages 115–124.

Franz J. Och and Hermann Ney. 2002. Discriminative Training and Maximum Entropy Models for Statistical Machine Translation. In *Proceedings of the 40th annual ACL Conference*, pages 295–302, Philadelphia, PA.

Stephan Oepen, Erik Velldal, Jan Tore Lonning, Paul Meurer, and Victoria Rosen. 2007. Towards Hybrid Quality-Oriented Machine Translation – On Linguistics and Probabilities in MT. In *Proceedings of the 11th International Conference on Theoretical and Methodological Issues in Machine Translation (TMI)*, Skövde, Sweden.

Satoshi Sato and Makato Nagao. 1990. Towards memory-based translation. In *Proceedings of the 13th International Conference on Computational Linguistics (COLING)*, pages 247–252, Helsinki, Finnland.

[2]The BNC consists of more than 100 million words in more than 6 million sentences `http://www.natcorp.ox.ac.uk/`

[3]To reach reversibility of the lemmatiser, and discriminate between otherwise ambiguous analyses, we have added a few tags.