

# Lemmatic Machine Translation

Stephen Soderland, Christopher Lim, Mausam, Bo Qin, Oren Etzioni, and Jonathan Pool

Turing Center, Department of Computer Science and Engineering

University of Washington, Box 352350

Seattle, WA 98195, USA

{soderlan,chrislim,mausam,boqin,etzioni,pool}@cs.washington.edu

## Abstract

Statistical MT is limited by reliance on large parallel corpora. We propose *Lemmatic MT*, a new paradigm that extends MT to a far broader set of languages, but requires substantial manual encoding effort.

We present PANLINGUAL TRANSLATOR, a prototype Lemmatic MT system with high translation adequacy on 59% to 99% of sentences (average 84%) on a sample of 6 language pairs that Google Translate (GT) handles. GT ranged from 34% to 93%, average 65%. PANLINGUAL TRANSLATOR also had high translation adequacy on 27% to 82% of sentences (average 62%) from a sample of 5 language pairs not handled by GT.

## 1 Introduction

Current Machine Translation (MT) systems can achieve high quality translations for a relatively small number of language pairs. Statistical MT techniques (Brown et al., 1993; Koehn et al., 2003) are limited by the need for large parallel corpora, and even then learn statistics specific to the domain and genre of the corpora (*e.g.*, parliament speeches). Rule-based MT systems (Bond et al., 2005) require manually engineered knowledge specific to the language pair and domain. The unavailability of aligned corpora or bilingual experts, for most language pairs, limits the applicability of current MT.

We propose *Lemmatic MT*—a new paradigm that scales MT to a far broader set of language pairs. The intuition behind Lemmatic MT is based on a situation that many have experienced. Someone is in a foreign country with a good translation dictionary,

Lemmatic encoding:

```
[ want two person room ]  
[ arrive June 1 ]  
[ 3 night stay ]  
[ price , with shower or bath ? ]
```

Intended meaning:

```
We want a room for two persons.  
We will arrive on June 1.  
We will stay for 3 nights.  
What is the price with shower or bath?
```

Figure 1: An example of lemmatic encoding of a request for hotel booking.

but no knowledge of the grammar of the language. Effective communication is still possible by stringing together short sequences of translated words selected from the dictionary. We refer to this style of encoding as *lemmatic encoding* (see Figure 1).

Of course, Lemmatic MT will not yield fluent, grammatical sentences in the target language, nor will it intelligibly translate idioms or other expressions that do not have compositional semantics. The burden is on the author to avoid such expressions. Nevertheless, Lemmatic MT is feasible for the many language pairs for which statistical MT is not, because Lemmatic MT merely requires a dictionary to translate between two languages, but no parallel corpora. Figure 1 gives an example of lemmatic encoding of a request for a hotel booking. We describe lemmatic encoding in more detail in Section 2.1.

This paper introduces PANLINGUAL TRANSLATOR, a prototype Lemmatic MT system. In PANLINGUAL TRANSLATOR, an author encodes each sentence to be translated by selecting only the words present in the system’s translation dictionary. The

system then translates the encoded word sequence into the target language, choosing an appropriate translation for each word. PANLINGUAL TRANSLATOR must still address the problem of word sense disambiguation (WSD) since the original sentence may contain polysemous words. To address this problem, we investigated two methods, a method that selects the dominant sense automatically (Section 2.3) and another that uses corpus-based statistics for cross-language WSD (Section 2.4).

We tested PANLINGUAL TRANSLATOR on a variety of genres: a request for hotel booking, a paraphrased news story, a wedding invitation, and a paraphrased encyclopedia article. It achieved high translation adequacy (conveying the intended meaning) for a broad range of language pairs.

PANLINGUAL TRANSLATOR had high translation adequacy for 59% to 99% of the sentences on a sample of 6 language pairs supported by Google Translate (GT), with an average of 84%. GT had high adequacy for 34% to 93% of the sentences, with an average of 65%.

Moreover, PANLINGUAL TRANSLATOR achieved high translation adequacy for 27% to 82% of sentences (average 62%) from a sample of 5 language pairs that GT does not handle. PANLINGUAL TRANSLATOR’s translation adequacy depends on the accuracy and coverage of its translation dictionary. We show high adequacy for 93% of sentences where the dictionary has high confidence translations for all words in the sentence, regardless of the source and target languages.

The key contributions of the paper are as follows:

- We introduce a new paradigm for MT – text translation based on lemmatic encoding.
- We present a novel cross-language WSD algorithm that scales to a large number of language pairs by using only monolingual corpora.
- We embody Lemmatic MT in the PANLINGUAL TRANSLATOR prototype, which can translate between any language pair for which a dictionary provides high precision translations.
- We report on a set of experiments in which PANLINGUAL TRANSLATOR has translation adequacy superior to that of Google Translate (GT) on several language pairs supported by

GT, and has high adequacy on language pairs that GT does not handle.

The remainder of this paper is organized as follows. Section 2 describes components of PANLINGUAL TRANSLATOR. We present empirical results in Section 3, related work in Section 4, and conclusions in Section 5.

## 2 Lemmatic Translation

We have investigated Lemmatic MT by creating a set of guidelines for lemmatic encoding (Section 2.1), building an interface to guide the user in encoding sentences lemmatically (Section 2.2), and designing translation algorithms (Sections 2.3 and 2.4).

### 2.1 Lemmatic Encoding

The purpose of lemmatic encoding is to enable lexical translation to convey the intended information clearly to the recipient of a message. We have found the following guidelines for lemmatic encoding to be effective.

1. Break the text into short, sentence-like sequences of words.
2. Select words from the translation vocabulary that will maintain the intended word sense when translated word-by-word.
3. Lexicalize information that would ordinarily be conveyed by word inflection (*e.g.*, verb tense).

As an example of these guidelines, consider an email sent to a hotel to book a room. The intended meaning is “We would like a double room on June 1st for 3 nights”. A possible encoding of this breaks it into the first three sentences in Figure 1.

Breaking the text into short sentences helps prevent confusion due to different word order in the source and target languages. The intended meaning of these short sequences of words is largely independent of word order, even if encoded with a verb-final word order and with modifier-noun attachment in the opposite order, *e.g.* [ room two person want ] or [ 1 June arrive ].

As future work, we intend to include rules for systematic word re-ordering, which would require a part-of-speech tagger for the source language and knowledge of word order for the target language.

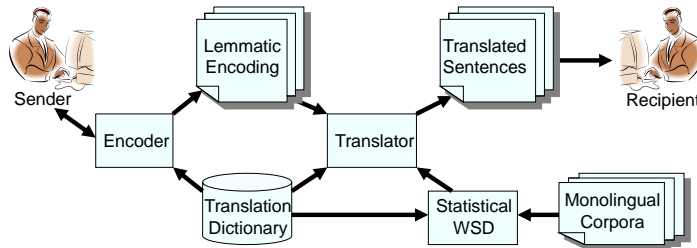


Figure 2: PANLINGUAL TRANSLATOR architecture for user-guided translation.

Another key to effective lemmatic encoding is to select words that will retain their sense when translated in isolation. (We later explore cross-language word sense disambiguation in Section 2.4.) The idiomatic use of “double” in “double room” is encoded as “two person room”.

Unfortunately, the translation dictionary may still select a word in the target language with the wrong sense. In an earlier version of PANLINGUAL TRANSLATOR, we found that such sense shifts happened for nearly 25% of the words. So we added *back translations* to the user interface to guide the encoder to avoid sense shifts. If PANLINGUAL TRANSLATOR translates a word  $w_s$  in source language  $l_s$  as word  $w_t$  in target language  $l_t$ , then the *back translation* is defined as PANLINGUAL TRANSLATOR’s translation of  $w_t$  back into  $l_s$ .

## 2.2 A User Interface for Lemmatic Encoding

We have built a graphical interface for PANLINGUAL TRANSLATOR that lets a user select a source and target language, and then encode sentences freely. PANLINGUAL TRANSLATOR looks up the words being typed and displays an auto-completion list with words that can be translated into the target language. This guides a user in selecting words that can be translated effectively.

For each possible completion of a word, PANLINGUAL TRANSLATOR shows the source word, the best translation (according to the algorithm described in Section 2.3) and a translation back into the source language.

Suppose that a user who is translating from English to Swahili encodes the word “want” with the intended meaning of wish or desire. The PANLINGUAL TRANSLATOR interface indicates that this will be translated as “uchehefu” with a back translation of “deficiency”. This is a valid translation of “want”, but clearly not in the intended word sense. If

the user substitutes the word “need”, PANLINGUAL TRANSLATOR gives the translation as “haja” with the back translation “need”. In the context of a request for hotel booking, this is an acceptable synonym.

When the language pair has high coverage in PANLINGUAL TRANSLATOR’s dictionary, we found that users can encode sentences nearly as fast as they are able to type. The most obvious word in the source language is usually found in the target language in the intended sense. For language pairs where the coverage is spotty, such as English to Swahili, encoding can become a search for source words that will be translated reliably, adding several minutes to the encoding of each text.

## 2.3 Context-Free Lexical Translation

Although the Lemmatic MT paradigm does not depend on a particular translation dictionary, we selected PANDICTIONARY (Mausam et al., 2009) as the basis for PANLINGUAL TRANSLATOR. PANDICTIONARY is a sense-distinguished translation dictionary compiled from Wiktionaries<sup>1</sup> and more than 600 machine readable bilingual dictionaries. PANDICTIONARY uses probabilistic inference to find translations that are not in any of the source dictionaries. The dictionary is organized as a set of word senses  $s$  that each have a list of translations  $w$  in multiple languages with an associated probability that  $w$  has sense  $s$  (denoted  $pr(w \in s)$ ).

PANDICTIONARY has 81,000 word senses with over 1.6 million translations with an estimated precision of 0.90 in over 700 languages. At a precision of 0.70, there are nearly 9 million translations in over 1,000 languages.

Its organization in terms of language-neutral senses means that if a language has high coverage in PANDICTIONARY, then there are translations

<sup>1</sup>www.wiktionary.org

into *all* other languages. This allows PANLINGUAL TRANSLATOR to scale to language pairs where parallel texts and bilingual dictionaries are lacking.

We now describe PANLINGUAL TRANSLATOR’s algorithm for selecting the best translation of each word into the target language. The basic translation algorithm begins with a source word  $w_s$  in language  $l_s$  and combines evidence from PANDICTIONARY to find the most probable context-free translation  $w_t$  in a target language  $l_t$ . A refinement of this that uses cross-language WSD is presented in the Section 2.4.

There are often multiple senses  $s$  in PANDICTIONARY in which a given  $w_t$  is a translation of  $w_s$ . This is because of synonymous or nearly synonymous senses from the various Wiktionaries and because of shared polysemy between  $w_s$  and  $w_t$ .

The *AllSenses* translation method we use in PANLINGUAL TRANSLATOR combines evidence from multiple PANDICTIONARY senses. For each possible translation  $w_t$  of  $w_s$  into language  $l_t$ , sum the probability that both  $w_s$  and  $w_t$  share sense  $s$  over all senses that contain both words. Return the  $w_t$  that has the maximum total probability.

$$\text{AllSenses}(w_s, l_t) = \underset{w_t}{\operatorname{argmax}} \left( \sum_{s \in \text{senses}} pr(w_s \in s) * pr(w_t \in s) \right) \quad (1)$$

The *AllSenses* algorithm is effective in finding translations of the dominant sense of a source word, especially for language pairs with high coverage in PANDICTIONARY. This is the translation algorithm used in the basic PANLINGUAL TRANSLATOR system described in our experiments. We next describe an enhancement to PANLINGUAL TRANSLATOR that uses corpus statistics to do automatic cross-language WSD.

## 2.4 Cross-language WSD

Since the goal of Lemmatic MT is to scale to large numbers of language pairs, we explored a WSD algorithm that does not need separate corpus statistics for each language pair.

There has been considerable research in statistical WSD (Navigli, 2009) since Yarowsky’s seminal paper (1992) that estimates the probability that a target word  $w$  will have a sense  $s$  given a context window of  $k$  words to the right and left of  $w$  based on corpus

statistics. Most cross-lingual WSD assumes a bilingual corpus. In contrast, our method cannot rely on the existence of even a monolingual corpus for each source language, since our techniques must work for hundreds of resource-poor languages.

Statistical WSD typically estimates the probabilities by compiling corpus co-occurrence statistics between senses and context words. If we represented the context of  $w$  in terms of actual *words*, we would need a separate corpus for each source language. Instead, we represent the context of  $w$  in terms of language-neutral evidence: the PANDICTIONARY senses. Thus, instead of collecting word co-occurrence statistics we collect sense co-occurrence statistics. Our approach is related to Dagan and Itai’s dictionary-based WSD (1994); however, their approach requires a different bilingual dictionary for every source language. In contrast, we can make use of PANDICTIONARY and perform WSD for many languages in one go.

We compile sense co-occurrence statistics from monolingual corpora in three diverse languages: English, Spanish, and Chinese. For each word  $w$  we look up its possible senses in PANDICTIONARY, looking up a stemmed version of  $w$  if the surface form is not found. As context for  $w$ , we look up the PANDICTIONARY senses for words  $c$  in a context window of  $k$  words to the right and left. We then tabulate the co-occurrence,  $N(s_w, s_c)$ , of the senses of  $w$  with the senses of each  $c$ , giving a fractional count wherever  $w$  or  $c$  has multiple possible senses. We also store the frequency of occurrence of each sense,  $N(s)$ , giving partial counts if a word belongs to many senses.

Based on these counts, given a message in any language,<sup>2</sup> we can estimate the probability of a sense  $s$  for the source word  $w_s$  and context vector  $\bar{c}$  by Bayes rule:

$$\begin{aligned} pr(s|\bar{c}) &\propto pr(s)pr(\bar{c}|s) \\ &= pr(s) \prod_{c \in \bar{c}} \{pr(c|s)\} \\ &= \frac{N(s)}{N} \prod_{c \in \bar{c}} \left\{ \sum_{s_c} \frac{N(s_c, s)}{|\text{senses}(c)|N(s)} \right\} \end{aligned} \quad (2)$$

Here  $|\text{senses}(c)|$  denotes the number of senses of  $c$ , and  $N$  is the total number of words processed in

<sup>2</sup>It does not have to be one of the three corpus languages.

the corpus. In our implementation we remove the context words that do not give us sufficient information, *e.g.*, appear very limited times in the corpus or do not co-occur with most senses under consideration. Moreover, we treat the statistics from different corpora independently and use the independence assumption to combine them together.

Given this context-sensitive estimate of probability we can use a method similar to the AllSenses algorithm of the previous section to compute the best target translation.

$$CSAllSenses(w_s, l_t) = \operatorname{argmax}_{w_t} \left( \sum_s pr(w_s \in s) pr(w_t \in s) pr(s|\bar{c}) \right) \quad (3)$$

### 3 Experimental Results

The principle question that we investigated is how well does lemmatic encoding using PANLINGUAL TRANSLATOR maintain translation adequacy, *i.e.*, to what degree does the recipient of the translation understand the intended meaning of the translated text?

We compared PANLINGUAL TRANSLATOR with a widely used translation tool, Google Translate<sup>3</sup> (GT), which is based on statistical MT. Because GT’s scope is smaller than that of PANLINGUAL TRANSLATOR, we restricted the comparison to a sample of language pairs that both systems can handle. In addition, we measured PANLINGUAL TRANSLATOR’s adequacy on a sample of language pairs from the broad set that is not covered by GT.

#### 3.1 Methodology and Test Suite

Our experiments consisted of three steps:

1. Encode a test set of texts lemmatically and record the intended meaning as full English sentences.
2. Show the encoded texts translated into the target language to bilingual informants, who write their interpretation of each sentence in English.
3. Have three evaluators compare the informants’ interpretations with the intended meaning, and score each sentence for translation adequacy on a scale of 1 to 5.

<sup>3</sup>[www.translate.google.com](http://www.translate.google.com)

The advantage of this procedure is that it does not require any informants who are fluent in both the source language and target language, which would be problematic for pairs such as Japanese to Slovenian or Russian to Afrikaans.

The source languages for our experiments were English, Japanese, and Russian, which are among the highest coverage languages in PANDICTIONARY, and which one of the authors of this paper understood. We selected six language pairs that can be handled by Google Translate (GT): English to Japanese and Russian; Japanese to Russian and Slovenian; Russian to Japanese and Slovenian<sup>4</sup>. Slovenian was chosen arbitrarily as one of the more resource-poor in the languages handled by GT.

We also included pairs that are not handled by GT: source language English, Japanese, or Russian to target languages Afrikaans, Faroese, and Swahili.

We composed four test texts, each consisting of eight encoded sentences, from a variety of genres: a request for hotel booking, a paraphrased news story about British troops in Iraq, a wedding invitation, and a paraphrased encyclopedia article on the evolution of horses. The texts were encoded by an experienced user (one of the authors), which gives an upper bound on performance by novice encoders.

We recruited bilingual informants using Amazon Mechanical Turk, giving each informant the texts in random order without revealing the translation method or source language. We needed to screen the responses to eliminate those whose gaps in understanding and errors indicated someone who used Google Translate or who relied on English cognates and wild guesses.

For our GT baseline, we broke the texts into short, simple sentences and entered them into the Google Translate Web interface<sup>5</sup>.

The scoring was done by three evaluators, who were shown the interpretations in random order by text, without knowing the source or target language, or the translation method. They scored each sen-

<sup>4</sup>We note that this is not a random sample, and that GT is likely to perform better on pairs where it has more parallel text such as English and French.

<sup>5</sup>We also did an experiment where we copied the GT output into the input box to check back translations of each sentence. Even spending several minutes per text searching for good back translations did not improve the translation adequacy.

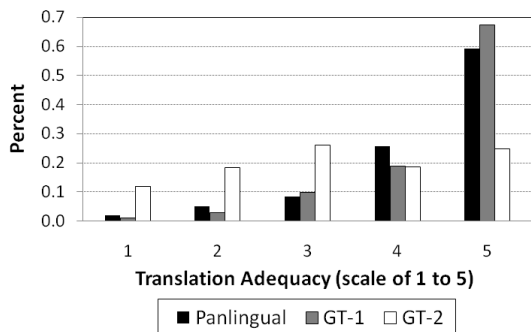


Figure 3: PANLINGUAL TRANSLATOR has high translation adequacy (score 4 or 5) for 84% of sentences in language pairs handled by Google Translate. For some language pairs GT has high adequacy for 86% of the sentences (GT-1), but only 43% for other pairs (GT-2).

tence from 1 to 5 where 5 means that the interpretation contains all the intended meaning; 4 means a small error<sup>6</sup>; 3 means half right; 2 means mostly wrong; and 1 means a totally incorrect interpretation. There was high inter-evaluator agreement – scores were within one point of all other evaluators’ scores 94% of the time.

### 3.2 Translation Adequacy

Figure 3 shows translation adequacy scores for PANLINGUAL TRANSLATOR on a sample of language pairs that GT also handles. PANLINGUAL TRANSLATOR had high translation adequacy (score 4 or 5) for 59% to 99% of the sentences, averaging 84%.

Google Translate’s adequacy varied even more widely, presumably due to the amount of training for different language pairs. GT-1 is the set of language pairs English-Russian, Russian-Slovenian, and English-Japanese, where GT had high scores on 93%, 85%, and 80% of the sentences, respectively. Performance fell dramatically for Japanese-Russian, Japanese-Slovenian, and Russian-Japanese (GT-2), with high adequacy on 34%, 44%, and 51% of the sentences. The overall average for GT-1 and GT-2 is 65%. PANLINGUAL TRANSLATOR outperformed GT on 4 of these 6 language pairs.

Overall, PANLINGUAL TRANSLATOR has an average translation adequacy of 4.33 vs. 3.83 for GT. This difference is statistically highly significant using a paired t-test ( $p < 10^{-5}$ ).

On language pairs that are not covered by GT,

<sup>6</sup>e.g. a word in almost the right sense or a wrong verb tense

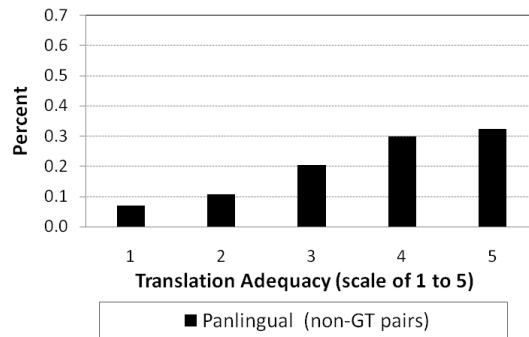


Figure 4: PANLINGUAL TRANSLATOR can translate language pairs that are not handled by Google Translate, with high translation adequacy for 62% of the sentences.

PANLINGUAL TRANSLATOR has high scores for 62% of the sentences (Figure 4). This falls short of the performance of GT on its best language pairs, but is better than GT on the remainder of the language pairs that GT handles.

It is remarkable that Lemmatic MT, with no language model and no attempt at fluent translations, can give better translation adequacy than statistical MT on so many language pairs. Of course, statistical MT is still the best choice where there is ample parallel text, such as English to Japanese or to Russian. Often, however, there is no readily available parallel text, even between major languages such as Japanese to Russian. In such cases Lemmatic MT can give higher translation adequacy than statistical MT, so long as a sense-distinguished translation dictionary has good coverage for the language pair.

### 3.3 The Effect of Dictionary Coverage

We analyzed our results to determine the effect of dictionary coverage on PANLINGUAL TRANSLATOR by examining PANDICTIONARY’s confidence in the translation of each encoded word. Figure 5 demonstrates the strong correlation between PANDICTIONARY confidence in word translation and errors that affected translation adequacy. PANLINGUAL TRANSLATOR achieves perfect scores for 68% of the sentences and scores of 4 or 5 for 93% of the sentences where all words are translated with high confidence by PANDICTIONARY.

About 80% of sentences consist entirely of words with high confidence translations where the target language is Russian or Japanese, but only 11% of the sentences where the target language is Slovenian

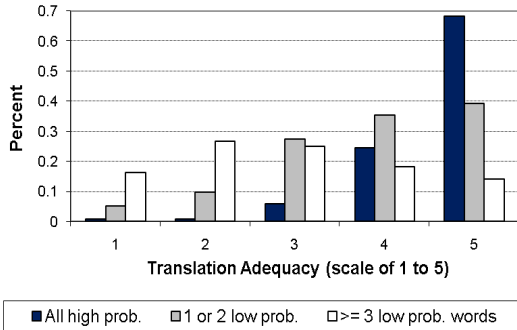


Figure 5: PANLINGUAL TRANSLATOR has high translation adequacy for 93% of sentences where the dictionary has high confidence in the translation for all the words. This falls off drastically for sentences with 3 or more low probability translations.

or Swahili. PANLINGUAL TRANSLATOR achieves surprisingly good performance for translation into languages with such low dictionary coverage.

### 3.4 The Effect of Cross-language WSD

The version of PANLINGUAL TRANSLATOR that we have presented thus far has implicit *manual* disambiguation – the system presents back translation feedback that guides the encoder to select only words that will be translated in the intended sense.

For our experiments with corpus-based WSD, we began with sentences that were encoded without this manual disambiguation. We took the words that were selected without back translation feedback as input to the WSD algorithm of Section 2.4.

This experiment is on an earlier set of language pairs that has source languages English, Japanese, and Russian, and has target languages Albanian, Catalan, Slovenian, and Swahili. In our later experiments we added target languages Japanese and Russian, dropped most of the other languages handled by GT, and added more languages not handled by GT.

Figure 6 shows that corpus-based WSD gives a significant improvement over the version of PANLINGUAL TRANSLATOR that lacks manual WSD. WSD boosts PANLINGUAL TRANSLATOR’s performance from high translation adequacy on 52% of the sentences to 66%.

## 4 Related Work

Attempts to carry out word-by-word translation based on dictionaries date back to the 1950s, if not

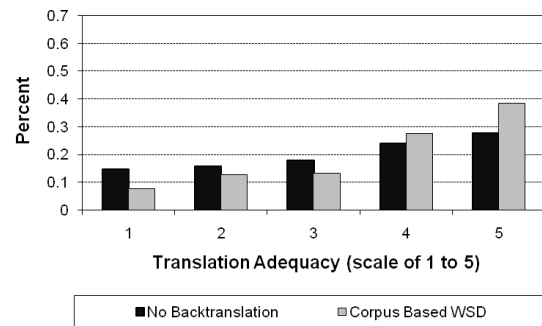


Figure 6: Adding corpus-based WSD gives a boost in translation adequacy over a version of PANLINGUAL TRANSLATOR that did not use back translation feedback.

earlier (Reifler, 1958; Bar-Hillel, 1960). MT based on an interlingual-thesaurus has also been explored in the past (*e.g.*, (King, 1958)), with limited success due to the difficulties in constructing such a thesaurus.

Recent research in dictionary-based MT includes work by Carbonell *et al.* (2006) that combines a bilingual dictionary with a language model derived from a monolingual corpus; and by Muegge (2006) for translation in the restricted domain of product descriptions. While PANLINGUAL TRANSLATOR cannot produce the fluent translations of a system like Carbonell’s, its massive, multilingual dictionary allows it to scale to large numbers of language pairs including resource-poor languages.

Broadly speaking, Lemmatic MT is a novel use of lexical translation, which has a long history (Helmreich *et al.*, 1993; Copestake *et al.*, 1994; Hull and Grefenstette, 1996) and more recently Etzioni *et al.* (2007).

Our work also falls in the realm of human-aided machine translation (Cole *et al.*, 1997). PANLINGUAL TRANSLATOR uses a novel form of human aid—lemmatic encoding of the original text. This is reminiscent of controlled language encoding (*e.g.*, (Fuchs *et al.*, 1998; Schwitter, 2002; Nyberg and Mitamura, 1996)), however, controlled languages tend to have a small, domain-dependent vocabulary and rigidly defined rules that allow parsing into an interlingua. Our lemmatic encoding is based on a large, *domain-independent* vocabulary and is intended for human decoding.

Our cross-language WSD algorithm, which uses multiple monolingual corpora and tabulates statis-

tics using a multilingual dictionary, is a novel approach to WSD for translation. It is most closely related to dictionary based WSD, which uses similar principles, but gathers statistics using a bilingual dictionary (Dagan and Itai, 1994).

## 5 Conclusions and Future Work

We have introduced a new paradigm, Lemmatic MT, in which an author encodes sentences as sequences of words from a translation dictionary. While this does not produce grammatically correct translations, we demonstrated that Lemmatic MT can convey the author's intended meaning effectively for a sample of language pairs.

Our prototype Lemmatic MT system, PANLINGUAL TRANSLATOR, proved to have high translation adequacy for 84% of sentences in a sample of language pairs supported by Google Translate (GT), ranging from 59% to 99%.

GT had high adequacy on 65% of the sentences, ranging from 34% to 93%. PANLINGUAL TRANSLATOR also performed well on language pairs that GT does not handle, with high translation adequacy on 62% of the sentences.

Additional experiments are necessary to test and extend Lemmatic MT, particularly for encoding in the hands of naive users. We are currently working on a PANLINGUAL TRANSLATOR Web service based on Lemmatic MT that enables practical translation between a thousand language pairs.

## Acknowledgements

This research was supported by a gift from Utilika Foundation to the Turing Center at University of Washington.

## References

- Yehoshua Bar-Hillel. 1960. The present status of automatic translation of languages. *Advances in Computers*, 1:91–163.
- F. Bond, S. Oepen, M. Siegel, A. Copestake, and D. D. Flickinger. 2005. Open source machine translation with DELPH-IN. In *MT Summit X*.
- P. Brown, S. Della Pietra, V. Della Pietra, and R. Mercer. 1993. The mathematics of machine translation: parameter estimation. *Computational Linguistics*, 19(2):263–311.
- J. Carbonell, S. Klein, D. Miller, M. Steinbaum, T. Grasziany, and J. Frey. 2006. Context-based machine translation. In *AMTA*.
- Ronald Cole, Joseph Mariani, Hans Uszkoreit, Giovanni Battista Varile, Annie Zaenen, and Antonio Zampolli, editors. 1997. *Survey of the State of the Art in Human Language Technology*. Morgan Kaufman.
- A. Copestake, T. Briscoe, P. Vossen, A. Ageno, I. Castellon, F. Ribas, G. Rigau, H. Rodriguez, and A. Samiotou. 1994. Acquisition of lexical translation relations from MRDs. *Machine Translation*, 3(3–4):183–219.
- Ido Dagan and Alon Itai. 1994. Word sense disambiguation using a second language monolingual corpus. *Computational Linguistics*, 20:563–596.
- O. Etzioni, K. Reiter, S. Soderland, and M. Sammer. 2007. Lexical translation with application to image search on the Web. In *Machine Translation Summit XI*.
- Norbert E. Fuchs, Uta Schwertel, and Rolf Schwitter. 1998. Attempto Controlled English (ACE), language manual, version 2.0. Institut fr Informatik, University of Zurich.
- S. Helmreich, L. Guthrie, and Y. Wilks. 1993. The use of machine readable dictionaries in the Pangloss project. In *AAAI Spring Symposium on Building Lexicons for Machine Translation*.
- D.A. Hull and G. Grefenstette. 1996. Querying across languages: a dictionary-based approach to multilingual information retrieval. In *Proceedings of ACM SIGIR 1996*, pages 49–57.
- G. W. King. 1958. A thesaurus lattice approach to the structure of language and communication with words. A Report to National Science Foundation.
- P. Koehn, F. J. Och, and D. Marcu. 2003. Statistical phrase-based translation. In *HLT/NAACL*.
- Mausam, S. Soderland, O. Etzioni, M. Skinner, D. S. Weld, and J. Bilmes. 2009. Compiling a massive, multilingual dictionary via probabilistic inference. In *ACL 2009*.
- Uwe Muegge. 2006. An excellent application for crummy machine translation: Automatic translation of a large database. In *Annual Conference of the German Society of Technical Communicators*, pages 18–21.
- R. Navigli. 2009. Word sense disambiguation: a survey. *ACM Computing Surveys*, 41(2):1–69.
- E. H. Nyberg and T. Mitamura. 1996. Controlled language and knowledge-based machine translation: Principles and practice. In *Workshop on Controlled Language Applications (CLAW-96)*.
- E. Reifler. 1958. The machine translation project at the University of Washington, Seattle. In *Intern. Congr. Linguists*, pages 514–518.
- Rolf Schwitter. 2002. Developing a black box specification using controlled English. *Computational Linguistics*.
- D. Yarowsky. 1992. Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. In *COLING*.