Poetic Machine Translation

Brad Girardeau Pranav Rajpurkar

December 13, 2013

1 Introduction

Translation is a complex, multifaceted challenge. It aims to produce an equivalent text in another language, but this equivalence is difficult to define. The translated text should communicate the same meaning, the same picture on a mental stage. It should match the style and structure of the original. Translation demands nuance and detail, a subtle, artistic pursuit.

Poetry is particularly difficult to translate, as meaning and form are intricately interwoven. Rhythm and sound play a key role in many poems, and an effective translation will capture these effects. It seems impossible for machines to capture the artistry of langauge in even mundane documents. Yet machine translation systems continue to improve at translating a variety texts. However, these systems focus on preserving meaning, not capturing the rhythm and sound qualities important in poetry.

We investigate how machine translation systems can be extended to take poetic qualities into account, aiming to create a system that more effectively translate lyrics and poems. To do so, we extend a contemporary machine translation system to incorporate rhythmic features and evaluate its performance.

2 Poetic Translation

To create a poetic machine translation system, we must first understand how language is used in poetry and its effect on the translation process. We examine a poem like Tarantella by Hilaire Belloc, which begins:

> Do you remember an Inn,
> Miranda?
> Do you remember an Inn?
> And the tedding and the spreading
> Of the straw for a bedding,
> And the fleas that tease in the High Pyrenees,
> And the wine that tasted of tar?

And the cheers and the jeers of the young muleteers
(Under the vine of the dark veranda)?

The meaning of the words has less of an impact on the reader or listener than the rhyme, syllable, and stress patterns that create the poem's driving energy. It is difficult to determine exactly how these factors intertwine to create a great poem in English, much less understand how to create a Tarantella in French or Chinese, if even possible. However, many poetic texts observe more structured patterns that can be approximated in translation.

Poetic forms like the haiku have a defining syllable count pattern, and in song, the number of syllables determines how the lyrics fit to the music. English poetry uses several common patterns of stressed and unstressed syllables, like the iambic unstress-stress pattern. Employing these patterns then creates English sentences that sound poetic or lyrical. While French does not use stress patterns, a poetic translation of a French poem into English would use these patterns to create a result that sounds poetic in English. On the other hand, both French and English employ rhyme for lyrical effect, indicating an effective translation will preserve the rhyme scheme across languages. This suggests two strategies to simplify the complex, artistic process of poetic translation: a poetic structure in the target language should be used, regardless of the source language, and poetic structures common to both languages should be preserved as much as possible.

3 Machine Translation

Developing computer systems to translate has been an area of active research for over half a century, beginning at the start of the Cold War. Early systems were rule-based, using morphological, syntactic, and semantic rules to analyze source sentences and generate equivalent outputs in the target language. While rule-based systems can capture linguistic knowledge about the source and target language, they require a time consuming development process and often struggle to adapt to the inherent ambiguity of language. Instead, a statistical machine translation approach is increasingly popular today, which uses machine learning algorithms to learn to translate from large parallel and monolingual corpora.

Statistical machine translation systems evaluate possible translations using patterns and observations derived from large datasets, searching for the best translations. As the search space of possible translations is so large, it seems likely that some of these translations are also poetic, suggesting that, with a measure of "poeticness," a statistical system could actually find these poetic translations [2]. Two core components of these systems are a language model and translation model, though other models are also frequently incorporated. The language model is created by computing frequencies of n-word phrases (n-grams) from a monolingual corpus, providing an estimate that the component parts of a sentence compose good phrases in the target language. The translation model learns probabilities of a phrase (or word) in the target language translating a phrase in the original language. The scores of a proposed translation on these models form a vector of features, which are then weighted and averaged to produce an overall score. This allows new features, or models, to easily be incorporated into the system.

4 Poetic Model

We propose to add a poetic model to a statistical machine translation system that evaluates the poetic or lyrical quality of a translation. This will cause the system to find more poetic translations by heavily rewarding those that satisfy a given poetic form. There are a variety of potential components in such a model, with the most relevant (in English), stress pattern, rhyme, and syllable count. The desired form can be encoded in a feature that tracks and evaluates the current state of translation with its potential to fit the form. For example, in the case of a haiku (with a translation unit of three lines), the haiku feature will penalize any candidate translations that do not contain 5 syllables before the first line break, 7 between the first and second, and 5 syllables on the last line, as well as any that do not have 3 lines by the end of translation. In addition, this shows that the most natural translation unit may be longer than a sentence or line, as is also the case with rhyme schemes. On the other hand, forms like iambic pentameter and preserving the number of syllables in lyrics can proceed in the traditional line by line manner.

We implement a poetic feature using Stanford Phrasal, an extensible machine translation system [1]. Phrasal allows for the addition of new features, described in feature classes written in Java. To experiment with poetic translation, we concentrate on maintaining the same syllable count when translating from French into English. This preserves an important characteristic of a text's rhythm, especially with lyrics set to music, as well as many poetic forms. Counting syllables accurately is itself a challenging problem in speech recognition, but a computer hyphenation algorithm described by Frank Liang in 1983 for use in the TEX typesetting system provides a good approximation, as hyphenation points typically occur between syllables [4]. We use the Hyfo library, which implements Liang's algorithm, to count syllables in both French and English. Speech corpora, like the CMU database, can also provide this information for some languages.

We create a Phrasal derivation feature that gives a score proportional to the difference in syllables between the original and target sentence, firing when the decoder proposes a full line translation. We also experiment with a rule feature that evaluates each individual phrase translation, but this is more restrictive to the translation. A future feature could use the Phrasal alignment data interface to generate intermediate syllable matching scores as the sentence is built, a middle ground between phrase and sentence level features. We train on a subset of the Europarl parallel and monolingual corpus provided by the CS224n class to test the syllable feature [3]. The feature weights are first tuned on a small news dataset, then the syllable feature weight is set manually to ensure it has a sufficient impact.

5 Results

We test the resulting system on several songs and poems to examine the effect of the syllable feature. Some of the lines translated were not changed by the additional feature, as the syllable count already exactly or nearly matched. Sometimes the mostly direct word to word or phrase to phrase mapping that is performed when less data is available corresponds in syllables between French and English. However, in the sentences that had different numbers of syllables originally, the system with the new feature chose synonyms that ensured the line matched the syllable count. For example, the two systems translated the French sentence "grande fête au château il y a bien longtemps," with 11 syllables, as:

> "great party to the castle there has long" (9 syllables, baseline system), and
> "large marks the castle there is beginning long" (11 syllables, new feature).

With this model of incorporating poetic features, they can shape a translation to fit a poetic form, in this case matching syllables.

6 Conclusion

As Robert Frost stated, "Poetry is that which is lost in translation." Poetic translation is a challenging task for human translators, much less computer systems. It lets us test the limits of machine translation. While machine translation systems may not be able to match high quality human translations today, they have made significant progress, and continue to improve. With more exploration of poetic features, machine translations of poetry and song can become increasingly artistic. We would like to continue explore the area of poetic translation, pushing the boundaries of machine translation. Maybe one day we will have a poet in our pockets.

References

[1] Cer, Daniel, Galley, Michel, and Jurafsky, Daniel (2010). Phrasal: A Toolkit for Statistical Machine Translation with Facilities for Extraction and Incorporation of Arbitrary Model Features. In *Proceedings of the North American Association of Computational Linguistics - Demo Session (NAACL-10).*

[2] Genzel, Dmitriy, Uszkoreit, Jakob, and Och, Franz (2010). "Poetic" Statistical Machine Translation: Rhyme and Meter. In *Proceedings of the 2010 conference on Empirical Methods in Natural Language Processing*, pages 158-166, MIT, Massachusetts, USA, October.

[3] Koehn, Philipp (2005). Europarl: A Parallel Corpus for Statistical Machine Translation. *MT Summit*.

[4] Liang, Franklin M (1983). Word Hy-phen-a-tion by Com-puter. *Stanford University*.