

ISSN 2186-7437

NII Shonan Meeting Report

No. 2015-10

Simultaneous Machine Interpretation

Jordan Boyd-Graber
Hal Daumé III
Yusuke Miyao

August 03–06, 2015



National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

Simultaneous Machine Interpretation

Organizers:

Jordan Boyd-Graber (University of Colorado, USA)

Hal Daumé III (University of Maryland, USA)

Yusuke Miyao (National Institute of Informatics, Japan)

August 03–06, 2015

1 Meeting Overview

Synchronous machine interpretation is the task of automatically translating a foreign language while it is being spoken. When a sentence is being produced one word at a time in a foreign language, the goal is to produce a translation in a target language (e.g., English) simultaneously: with as little delay between the uttering of a foreign language word and the production of its English translation. A requirement for successful synchronous machine translation is to be able to predict what words are going to appear in the input stream, before they have been seen. This is particularly important when translating from verb-final languages like Japanese or German, into verb medial languages like English. Without the ability to predict the yet-to-be-seen verb, the translator has to wait until the entire sentence is uttered before anything beyond the subject can be translated. This is not just an issue of final-verb prediction; any significant reordering involves prediction in order to maintain simultaneity: for instance relative clauses or postpositions in Japanese.

This is an important task to study because effective human interaction with technology requires low latency. If translation systems have their built-in latency plus the latency of a sentence, there is no chance for translation systems to be used widely for everyday conversations. Creating systems that can provide low-latency translations will enable better communication across the divide of language and culture.

2 Meeting Schedule

- August 2, Sunday Evening
 - Welcome Banquet
- August 3, Monday Morning
 - Opening Remarks
 - Short Presentations on Existing Work
- August 3, Monday Afternoon

- Group Photo
- Short Presentations on Existing Work
- August 4, Tuesday Morning
 - Panel Discussions on Evaluation and Data
- August 4, Tuesday Afternoon
 - Panel Discussions on Evaluation and Data
- August 5, Wednesday Morning
 - Panel Discussions on Frameworks
- August 5, Wednesday Afternoon
 - Excursion
- August 6, Thursday Morning
 - Invited Talk
 - InproTK Tutorial
 - Closing Remarks

3 Abstracts of Talks and Panel Discussions

Don't Until the Final Verb Wait

Jordan Boyd-Graber

We introduce a reinforcement learningbased approach to simultaneous machine translationproducing a translation while receiving input words between languages with drastically different word orders: from verb-final languages (e.g., German) to verb-medial languages (English). In traditional machine translation, a translator must wait for source material to appear before translation begins. We remove this bottleneck by predicting the final verb in advance. We use reinforcement learning to learn when to trust predictions about unseen, future portions of the sentence. We also introduce an evaluation metric to measure expeditiousness and quality. We show that our new translation model outperforms batch and monotone translation strategies.

Interpretese vs. Translationese

Hal Daume III

Real-time simultaneous interpretation is a demanding task. An interpreter must translate partial sentences before knowing how the sentence will end. Computational approaches for this task are held back by our lack of a quantitative understanding of the tactics human simultaneous interpreters use. We produce a parallel, comparative corpus of translation and simultaneous interpretation. We study differences between interpreted text and translated text through an

automatic, computational approach. Our analysis reveals that human interpreters regularly apply several effective tactics to reduce translation latency, such as segmentation and passivization. Furthermore, we build a model that can predict tactics given only partial text input. This forms a necessary first step to integration with machine translation systems.

Paraphrase & Phrase Alignment

Yuki Arase

Paraphrases are a useful resource for machine translation. Translation quality can be notably improved by simply paraphrasing difficult-to-translate input sentences. Also paraphrases enable evaluation of machine translation outputs closer to what human does, by allowing other expressions to convey the same meaning. In such applications, we believe syntactic structures in paraphrases play an important role. We would like to introduce our latest work on this direction; monolingual phrase alignment to extract paraphrases with syntactic structures. Our method aligns parse-forests of an input sentential paraphrase pairs and identify the most plausible node pairs as phrasal paraphrases with their deep syntactic structures. We believe the same mechanism can be applied to bilingual parsing as well in order to obtain syntactic translation knowledge.

Simultaneous Translation for Hiero

Anoop Sarkar

Hierarchical phrase-based machine translation (Chiang, CL 2005) (Hiero) is a prominent approach for Statistical Machine Translation usually comparable to or better than conventional phrase-based systems. But Hiero typically uses the CKY decoding algorithm which requires the entire input sentence before decoding begins, as it produces the translation in a bottom-up fashion. Left-to-right (LR) decoding (Watanabe, ACL 2006) is a promising decoding algorithm for Hiero that produces the output translation in left to right order. In this paper we focus on simultaneous translation using the Hiero translation framework. In simultaneous translation, translations are generated incrementally as source language speech input is processed. We propose a novel approach for incremental translation by integrating segmentation and decoding in LR-Hiero. We compare two incremental decoding algorithms for LR-Hiero and present translation quality scores (BLEU) and the latency of generating translations for both decoders on audio lectures from the TED collection.

Promoting Science and Technology Exchange using Machine Translation

Toshiaki Nakazawa

There are plenty of useful scientific and technical documents which are written in languages other than English, and are referenced domestically. Accessing these domestic documents in other countries is very important in order to know what has been accomplished and what is needed next in the science

and technology fields. However, we need to surmount the language barrier to directly access these valuable documents. One obvious way to achieve this is using machine translation systems to translate foreign documents into the users' language. Even after the long history of developing machine translation systems among East Asian languages, there is still no practical system. We have launched a project to develop practical machine translation technology for promoting science and technology exchange. In this presentation, we introduce the background, goal and status of the project.

Incremental Speech Synthesis

Timo Baumann

Incremental speech synthesis deals with starting to deliver an utterance before the full utterance specification becomes available. This is challenging, as speech prosody may contain long-range dependencies that are usually computed from full sentences, rather than from partial structure. However, prosody is not just a nice topping for speech output, but also serves communicative functions. For translation, a speech synthesizer may have to communicate two aspects on this prosodic side-channel on short notice: (a) MT taking overly long so that a filled pause needs to be rendered to signal that more speech is to follow, and (b) MT rethinking already delivered partial utterance content. In the latter case, realistic corrections will need to be synthesized that integrate well with the existing material.

I will present an architecture for incremental speech synthesis (and other aspects of incremental processing) as well as the realization in a speech synthesis system. I will try to steer you towards the temporal aspect of speech, both for spoken input as well as spoken output, and present strategies for shaping how time is ideally allocated to result in translation experience that is both smooth and quick (or the best tradeoff of the two).

Strategies for Simultaneous Speech Translation using Short Units

Yusuke Oda

One important difference between usual speech translation and simultaneous speech translation is whether the speech to be translated is the sentence or more shorter units. If speakers continue speaking without explicit sentence boundaries (e.g. lectures, presentations), we must translate long sentences, which imposes long waiting times on listeners. Segmentation and prediction are simple and strong ideas to avoid long waiting times. "Segmentation" finds additional boundaries to separate the input sentence and start partial translation without waiting for the end of the sentence. "Prediction" makes up for the context dropped by segmentation, which helps prevent a drop of segmentation quality due to segmentation mistakes. This presentation shows an overview of segmentation and prediction methods with a focus on works of our research group.

The European Parliament Interpreting Corpus (EPIC): a trilingual resource for the study of simultaneous interpreting

Mariachiara Russo

The European Parliament Interpreting Corpus (EPIC) is one of the first machine-readable corpora available in the field of Interpreting Studies. It was created between 2004 and 2006 in the Department of Interpreting and Translation of the University of Bologna at Forl and consists of 9 sub-corpora in total: three sub-corpora of source language speeches (Italian, English and Spanish) and six sub-corpora of simultaneously interpreted speeches, thus comprising all possible directions and combinations of the three languages involved. The paper describes the methodology used to create the EPIC multimedia archive and the EPIC corpus, and provides an overview of the main teaching and research applications of the corpus to date. Finally, current developments are briefly described.

Human Elements in Automatic Simultaneous Speech Translation

Graham Neubig

In this presentation I will talk about two separate lines of work on considering human elements in the study of automatic simultaneous speech translation. In the first part, I will explain about using data from human simultaneous interpreters in engineering-oriented studies of speech translation, including 1) a description of available corpora including the one that we have created at NAIST, and 2) methods for learning speech translation systems from this data. In the second part of the talk, I will describe our recent work on creating evaluation measures for simultaneous speech translation systems that consider both the speed and accuracy of translation, and correlate well with human perceptions of the quality of the translation results.

Preordering for statistical machine translation based on the maximization of Kendall's tau

Yusuke Miyao

This talk presents a simple model for preordering input sentences for improving statistical machine translation. Given binary constituent trees, the model classifies whether children of each node should be reordered. We introduce a procedure to determine oracle reordering decisions from word alignments, by maximizing Kendall's τ . Experiments on Japanese-to-English translation revealed a significant boost in accuracy, which outperforms state-of-the-art methods in Japanese-to-English translation.

Revamping traditional Japanese parsing pipeline

Masaaki Nagata

Traditional Japanese parsing pipeline consists of morphological analysis, bunsetsu-based unlabeled dependency parsing, and predicate argument structure analysis including zero pronoun anaphora resolution. This is very different from the ones for other languages such as English and Chinese, and makes our SMT software asymmetric (translation from Japanese to Language-X and Language-X to Japanese are very different). We show that we can solve the problem if we reorganize Japanese analysis into multi-layer labeled segmentation, word-based labeled dependency parsing, and empty category detection.

Improving the vocabulary coverage of selected data

Amitai Axelrod

Domain & Discourse Effects in Machine Translation Lexical Choice

Marine Carpuat

In this talk, I will present two pieces of work on lexical choice in machine translation that might be useful for machine interpretation. First, I will argue that machine interpretation faces a domain adaptation problem since most training examples are drawn from translation rather than interpretation data. I will describe a method for automatically analyzing the impact of such a domain shift on lexical choice errors. Then, I will turn to discourse factors. I will describe an empirical analysis supporting the "one translation per discourse" hypothesis, which suggests that document level context could benefit systems that translate from partial input sentences in an interpretation setting.

Inducing tree substitution grammars from small corpora

Pascual Gomez

Most methods (such as the GHKM algorithm) extract rules of tree substitution grammars from bilingual corpora that is large enough to estimate reliable word-to-word alignments. However, in some cases, there might not be so much data to learn from. We propose a method to extract general rules from a small amount of bilingual trees, discuss the key algorithms and evidence the challenges associated to this problem.

Speech Translation

Stephan Vogel

Latency, Accuracy, and Evaluation for Simultaneous Translation

Anoop Sarkar, Yusuke Miyao, and Graham Neubig

What should grammar formalisms look like for synchronous translation

Yuki Arase, Masaaki Nagata, Marine Carpuat

How can we create select the best data for automatic synchronous translation systems

Amittai Axelrod, Mariachiara Russo, Toshiaki Nakazawa

Are Humans the Best Role Models for Simultaneous Interpretation

Hal Daume III and Graham Neubig

Frameworks for Synchronous Machine Translation

Stephan Vogel, Yusuke Oda

Fixing Bad Predictions

Graham Neubig, Jordan Boyd-Graber

Human simultaneous interpreters often predict content that has not yet been spoken in sentences to allow them to start translating more quickly. Recently, there have been a few works on incorporating this type of prediction into simultaneous translation. For example, Grissom et al. (2014) guess sentence-final verbs, and Oda et al. (2015) predict unseen syntactic constituents. However, in both human or automatic translation there is potential for bad predictions, where the predictions don't match the actual content of the following utterance. In this situation, human interpreters repair the utterance by rephrasing or correcting content. In this discussion, we'll talk about how we can implement similar functionality in simultaneous machine translation, and also about how we could perform evaluation for a translation system with this functionality.

Japanese to English Simultaneous Interpreting: Challenges, Enabling Factors and A Personal Perspective on Simultaneous Interpretation Machines

Noriaki Okumura

The speaker will provide a comprehensive overview, illustrated using examples in Japanese and English, of the practical techniques commonly used by conference interpreters when interpreting between Japanese and English in the simultaneous mode, including: anticipation, automation, simplification, fillers and word-for-word translation. The speaker will also offer an insiders perspective on the challenges typically faced by interpreters, common traits and trends among speakers, and the skills and non-linguistic expertise required to successfully work in this field.

InproTK Tutorial

Timo Baumann

4 List of Participants

- Amittai Axelrod, Johns Hopkins University
- Jordan Boyd-Graber, University of Colorado
- Marine Carpuat, University of Maryland
- Hal Daume III, University of Maryland
- Pascual Martnez-Gmez, Ochanomizu University
- Nagata Masaaki, NTT CSL
- Miyao Yusuke, National Institute of Informatics
- Graham Neubig, NAIST
- Oda Yusuke, NAIST
- Mariachiara Russo, University of Bologna
- Anoop Sarkar, Simon Fraser
- Stephan Vogel, Qatar Computing Research Institute
- Timo Baumann, Hamburg University
- Nakazawa Toshiaki, Japan Science and Technology Agency
- Arase Yuki, Osaka University