Dialogue Modelling for Statistical Machine Translation

1   Relevance relative to the call for proposals

The proposed project sets out to improve the quality of *machine translation* (MT) technology. Machine translation, known by the general public through popular applications such as Google Translate, is defined as the automatic translation from one language to another through a computer algorithm – for instance, translating a text from Japanese to Norwegian or vice-versa. In a globalised world where the ability to access and produce multilingual content is becoming ever more important, machine translation has become a R&D area of strategic significance for both economic and cultural reasons. In addition to helping millions of people meet their daily translation needs, machine translation can also facilitate the dissemination of information across multiple linguistic communities, and thereby help preserve smaller languages and promote cultural diversity.

The project outlined in this proposal concentrates on a particular type of translation task, namely the automatic translation of (live or transcribed) *dialogues*. This type of translation comprises several important application domains such as the translation of subtitles for audiovisual content, the circulation of meeting transcripts across multiple languages, and (most difficult of all) speech-to-speech interpretation between individuals speaking different languages.

The project's key idea is to improve the translation quality in these domains by incorporating *contextual dialogue knowledge* into the translation pipeline. The project will develop new models, tools and algorithms that allow various contextual factors to be automatically extracted from the dialogue history and integrated into the translation models of a statistical machine translation system. The resulting "dialogue-aware" models will therefore be able to dynamically adapt their translation outputs according to the current dialogue context. The project aims to demonstrate that such context-rich approach is able to produce translations of a higher quality than standard methods. To this end, the project will include both theoretical investigations related to the formalisation and estimation of these rich statistical models and experimental work to implement these ideas in a full-fledged machine translation architecture. To demonstrate the practical significance of the project, the translation models and software tools developed through the project will be evaluated in a concrete application domain of high industrial relevance: the automatic translation of subtitles.

The project stands at the interface between two research areas: dialogue modelling on the one hand, and statistical machine translation on the other. These two fields have so far developed in relative isolation. The proposed project aspires to create a bridge between them and show how dialogue modelling methods can be fruitfully leveraged in statistical machine translation. To our knowledge, such an approach has never been thoroughly investigated before. The research realised through the course of the project is therefore expected to contribute to substantial scientific and technological advances in both dialogue modelling and machine translation.

To achieve this ambitious goal, the project will feature strong international collaborations with leading researchers in the field of machine translation. In particular, long research stays are planned both to the Swiss Federal Institute of Technology and the Department of Linguistics at Uppsala University to conduct joint research work on some of the key scientific challenges described in this proposal. In addition, the project is also proud to include two well-known industrial partners, NRK and Broadcast Text International, who will contribute to the project by providing us with a privileged access to their archives of professionally translated subtitles. These archives will be exploited to construct high-quality translation models for the project's application domain.

Owing to the innovative character of the envisioned research, its focus on foundational aspects of natural language processing technology and its clear international outlook, the FRINATEK program provides an excellent fit for the proposed work.

2 Aspects relating to the research project

2.1 Background and status of knowledge

Albeit considerable progress has been made in machine translation over the last decade due to the development of robust statistical techniques, many key problems are yet to be solved. One recurring question pertains to the use of contextual information in the translation process. When faced with the task of translating a sentence from one language to another, human translators routinely look at the larger context to select the words and grammatical constructions that are deemed most appropriate given the situation. Current machine translation technologies are unfortunately unable to exploit context in the same manner and typically view their source inputs as an unstructured collection of sentences that are assumed to be independent of one another. Although this assumption has obvious practical advantages, it also ignores the vast amount of linguistic information that are structured at the cross-sentential level. Interestingly, earlier rule-based architectures did often account for discourse or dialogue structure, as exemplified by the semantic transfer approach in VerbMobil (Emele et al., 2000), but many of these insights have been forgotten in contemporary data-driven approaches.

Fortunately, several research groups have recently started to re-address these issues in a statistical perspective, and the last five years have witnessed a renewed interest for the discourse aspects of machine translation.[1] In addition to the existing body of work on word–sense disambiguation (Vickrey et al., 2005; Carpuat and Wu, 2007), researchers have started to investigate new questions related to discourse connectives (Meyer et al., 2012), lexical cohesion and consistency (Tiedemann, 2010; Gong et al., 2011), verb tenses (Gong et al., 2012), and pronominal anaphora (Le Nagard and Koehn, 2010; Hardmeier and Federico, 2010). Hardmeier (2012) presents a survey of recent work on discourse-oriented statistical machine translation.

The aforementioned work is however strongly targeted towards discourse in text-based material such as news articles or legal documents. To date, few researchers have investigated how to improve the translation of *conversational* material by exploiting their internal structure. This lack of research on the dialogue aspects of machine translation is a surprising fact, since dialogue exhibits at least as much cohesiveness (if not more) than textual discourse. We describe below three key characteristics of dialogue that affect the translation process. We provide for each concrete examples of English-Norwegian translation pairs extracted from the *OpenSubtitles* parallel corpus.[2]

Dialogue structure

In a natural conversation, the contributions of the dialogue participants are not isolated sentences but are typically dependent on one another. As pointed out by a wealth of research in linguistics and psychology, dialogue is a fundamentally collaborative activity in which the communicative actions of the participants are tightly coupled (Clark and Schaefer, 1989). What a speaker is saying at a given point is therefore often only interpretable in relation to the preceding history. Consider these two examples:

A:	Which way goes into town?	*(Hvilken vei fører til byen?)*	
B:	Right.	*(Høyre)*	(1563595)

A:	So, those two don't work for Miletto.	*(Så de to arbeider ikke for Miletto)*	
	They work for Crenshaw.	*(De arbeider for Crenshaw)*	
B:	Right.	*(Riktig)*	(861643)

We see in these two examples that access to the local dialogue structure is a prerequisite to disambiguate the word "*right*" and translate it appropriately – that is, either as indicating a direction, or as a

---

[1]This year will for instance be organised the first workshop on Discourse in Machine Translation at ACL.

[2]The numbers on the bottom right corner of each example correspond to the line numbers in the corpus.

feedback on the information expressed by speaker A. The two responses are indeed distinct dialogue acts: the first is an answer to a question, whereas the second is a grounding act following a statement.

Fragmentary utterances

In spontaneous conversations, utterances rarely take the form of complete, well-formed sentences with a subject, a verb and subsequent arguments. Many utterances are fragmentary and contain only a few words without an overt predicate. These utterances (sometimes called elliptical or non-sentential utterances) must therefore be interpreted in context. Consider the two following excerpts where the fragment "*for me?*" builds upon the preceding utterance:

A:	Mother ... what was it like for you?	*(Mor... hvordan var det for deg?)*	
B:	For me?	*(For meg?)*	(27301)
A:	You made this?	*(Har du bygget den?)*	
	For me?	*(Til meg?)*	(1232764)

Although their surface forms are identical, the two "*for me?*" responses can be differentiated through their distinct semantic roles in relation to the previous utterance (experiencer for the first, beneficiary for the second), which respectively translate in Norwegian as "*for meg*" and "*til meg*".

Entrainment

The participants in a conversation are very often inclined to *align* their way of speaking with one another. This property in known in psychology as interactive alignment (Garrod and Pickering, 2009). Humans show a clear tendency to (unconsciously) imitate their conversational partners. In particular, they automatically align their choice of words, a phenomenon called lexical entrainment. But alignment also occurs on other levels such as grammatical constructions (Branigan et al., 2000), pronunciation (Pardo, 2006), speech rate, and even gestures and facial expressions. Below is another excerpt that illustrates this phenomenon:

A:	Please, don't make the mistake of not taking me seriously, Roschmann.	*(Ikke gjør den feilen å ikke ta meg på alvor, Roschmann.)*	
B:	I do take you seriously.	*(Jeg tar Dem på alvor)*	(28263)

In this example, speaker B reuses the expression "*take X seriously*", although many other expressions could be used to convey the same information. However, the recent dialogue history gives a much higher saliency to "*take X seriously*" compared to these alternative realisations.

To date, no statistical machine translation approach has attempted to systematically exploit these linguistic insights to improve the contextual relevance of the produced translations. The project outlined in this research proposal intends to remedy this important neglect.

2.2 Approaches, hypotheses and choice of method

The central hypothesis of the project is that the quality of machine translation outputs in conversational domains can be significantly improved through a better account of the broad dialogue context. To this end, the proposed project aims to develop a range of new methods for the integration and exploitation of dialogue-related features in a statistical machine translation architecture. The project will investigate the use of both *source-side* and *target-side* contexts to improve the translation results. We describe below these two strategies and then detail the implementation and evaluation scheme that will be followed to analyse their empirical performance.

Source-side dialogue context

The integration of richer linguistic features is one of the most active research topics in statistical machine translation (Gimpel and Smith, 2008). The recent introduction of factored translation models has facilitated such integration (Koehn and Hoang, 2007), allowing system developers to combine multiple knowledge sources in a flexible manner, without further exacerbating data sparsity issues. In a factored translation model, a word is not reduced to a simple token but can be augmented with various linguistic information such as part-of-speech, lemma, morphological markers, etc. The project will extend these models with the following source-side features:

- *Dialogue structure*: As mentioned in the previous section, the local dialogue structure (as expressed by the history of dialogue acts) can be exploited to provide more contextually relevant translations. Sridhar et al. (2011) recently demonstrated the potential of this approach in a speech-to-speech translation task. We will develop a dialogue act classifier trained through semi-supervised learning with both out-of-domain labelled data (such as the Switchboard corpus) and in-domain unlabelled data. The predicted dialogue acts will then be included as an additional factor in the translation model. We will conduct experiments to evaluate the accuracy of the classifier and how its use can influence the quality of the translation outputs.

- *Dependency relations*: It has been previously demonstrated that features expressing syntactic or semantic dependencies can improve machine translation outputs (Liu and Gildea, 2010; Wang et al., 2012). In order to deal with fragmentary utterances, we want to investigate how to extend these approaches with dependency relations extracted at a (local) cross-sentential level. This could be realised by detecting possible fragments, merging them with their antecedents and subsequently applying a dependency parser such as the Maltparser (Nivre et al., 2007) on the "reconstructed" utterances. While reconstructing fragments is a non-trivial operation in the general case given the variety of fragment types (corrections, completions, clarifications, etc.), it should be possible to identify sub-cases where the operation has a simple procedural form. This question will be investigated together with the Language Technology Group, which can draw on extensive experience on parsing non-canonical domains through the WeSearch project.[3] Once extracted, parts of the dependency structure can be used to augment the source-side factors with additional variables specifying the syntactic function of each word. The project will develop the techniques and tools to implement this approach, and will analyse the empirical effect of these linguistic features on the translation performance.

As is usually the case when dealing with statistical models, data sparsity is a key issue to address. The project will investigate various smoothing techniques such as back-off and interpolated back-off models. In addition, we will explore the use of partitioning techniques to reduce the size of the feature space. As we have shown in previous work on dialogue management (Lison, 2012), probabilistic models of dialogue can often profit from expressive representations that can capture key aspects of the domain structure and therefore bring improved generalisation performance.

Target-side dialogue context

The project will also take advantage of dialogue context on the target side to improve translation quality. As we have argued, dialogue is more than a sequence of unconnected utterances. Words and phrases employed by the different speakers tend to be frequently repeated, and their semantics is generally consistent through the dialogue. These repetitions are due to various factors:

- Dialogue is cohesive: a conversation usually revolves around one or two active topics that evolve over time, and these topics are associated with specific linguistic constructions;

---

[3] http://www.mn.uio.no/ifi/english/research/projects/wesearch/

- Dialogue is a collaborative activity where much of the interaction is spent giving positive and negative feedbacks to maintain common ground between the interlocutors. These grounding actions often reuse already uttered linguistic constructions (Clark and Schaefer, 1989);

- Finally, the widespread phenomena of lexical and syntactic entrainment also tend to increase the likelihood of specific lexical items both in the source and target languages.

Dynamic model adaptation constitutes one interesting strategy to strengthen the cohesiveness of the translations at the cross-sentential level. This adaptation can notably be realised via caching techniques (Bellegarda, 2004). Their key idea is to mix a classical static model (e.g N-grams) with a dynamic model estimated from recent items in the history. Caching techniques can be applied to both language and translation models (Tiedemann, 2010; Gong et al., 2011). The project will develop and evaluate similar caching techniques for dialogue. Due to their dynamic nature, cache-based models can capture the fact that linguistic constructions are frequently reused and repeated through the interaction. We hope to show that such adaptive models are able to deliver outputs that capitalise on these insights and thereby come closer to the reference translations.

Target-side context is however harder to leverage than its source-side equivalent. Contrary to the source-side context where the "true" inputs are directly accessible, the reference translations are not observable. The only observable target features are the actual translations produced by the system, which are potentially erroneous. There is therefore a substantial risk of error propagation, since an erroneous translation selected at the onset of the dialogue can potentially affect multiple subsequent translations further down in the translation process. To mitigate this risk, new decoding algorithms have recently been developed to optimise the translation at the document-level, through the use of multiple decoding passes (Gong et al., 2011) or document-wide local search triggered after a classical decoding step (Hardmeier et al., 2012). Alternatively, one can also use re-ranking techniques and let the decoder generate for each utterance an n-best list of possible translations that are subsequently reranked on the basis of more global features. We plan to investigate whether these global optimisation techniques can also be beneficial in the dialogue domain.

Implementation, resources and evaluation

The models and algorithms developed through the project will be built upon the well-known statistical machine translation toolkit Moses (Koehn et al., 2007).[4] Moses is an open source, widely used development platform in the field of statistical machine translation, supported by an active community of researchers and developers. It is also a success story for European ICT research and has received support from several EU-funded projects. In addition to its highly optimised decoder, Moses comes bundled with a wide array of development tools for training language and translation models from corpora and tuning the system parameters. The Moses toolkit provides support for phrase-based, syntax-based and factored translation models. The possibility to estimate factored translation models in the architecture will be crucial for the integration of source-side contextual features. The target-side part of the project will most likely require modifications to the decoder in order to incorporate a translation cache and perform dialogue-wide optimisation (Tiedemann, 2010; Hardmeier et al., 2012). A schema of the projected system architecture is provided in Figure 1.

The translation performance of the various approaches developed through the project will be evaluated in a specific application domain, namely the translation of subtitles for audiovisual content. Subtitles offer a good benchmark for our project, since they are (for the most part) made of conversational material,[5] and constitute to our knowledge the only conversational domain for which non-negligible amounts of parallel translations are available. The project will employ several databases

---

[4]http://www.statmt.org/moses/.

[5]Subtitles are of course scripted interactions that must satisfy specific timing and space constraints. This fact does not however change the conversational nature of these interactions, which is what matters most for the proposed project.
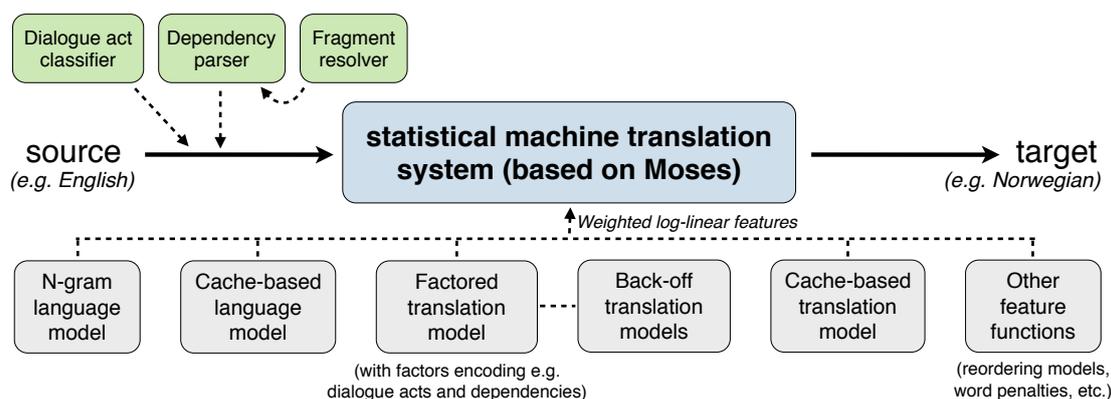
Figure 1: Simplified architectural schema of the translation pipeline.

of subtitles for training and evaluation. The first database is the *OpenSubtitles* corpus within the OPUS collection.[6] The corpus is a large repository of freely available fan-made subtitles that have been sentence-aligned. In addition, the project also boasts partnerships with two important players of the subtitling industry in Scandinavia: NRK's subtitling department and Broadcast Text International. These two partners will provide the project with a privileged access to their repositories of high-quality subtitles, which will complement the *OpenSubtitles* data set.

The project will attach special consideration to the inclusion of Norwegian (Bokmål) in the language pairs used for the experiments. The principal translation pair used for training and evaluation will therefore be English ↔ Norwegian.[7] In addition to the subtitle databases mentioned above, we will also consider out-of-domain resources such as the Oslo Multilingual Corpus[8] to increase the amount of data available for training the translation models. Besides English and Norwegian, the project will also carry out smaller scale experiments with other language pairs (provided they offer reasonable amounts of training material) to validate the approach on datasets of various sizes. It should however be stressed that the translation techniques devised during the course of the project are meant to be language-independent and could theoretically be applied to any language pair.

All the experiments conducted within the project will be rigorously evaluated using standard quality metrics and compared to their respective baselines. As is now common practice in the field, this evaluation will be performed with the help of reference-based metrics such as BLEU (Papineni et al., 2002) and METEOR (Lavie and Denkowski, 2009). Although these metrics have well-documented limitations, they have been shown to clearly correlate with user assessments of translation quality and remain the most practical methods to evaluate translation performance on a large scale basis (Koehn, 2010). In addition, the project will also consider more targeted evaluations to determine the performance of specific sub-components (such as the dialogue act classifier and semantic parser developed for the source-side modelling) or contrast how the translations of particular dialogue phenomena such as fragments are rendered with the different translation models.

To ensure the widest visibility of the research results and foster the involvement of all interested parties in the research effort, the tools, models and algorithms created through the project will be publicly released under an open source license and made available to the research community.

---

[6] http://opus.lingfil.uu.se/

[7] Both translation directions present their own challenges. Norwegian → English will have to discard some source-side features due to the more limited NLP resources available for Norwegian (e.g. absence of training data for dialogue acts). Some Norwegian resources are however in development and could be leveraged in the future, such as the dependency-based treebank currently collected at Språkbanken (cf. http://www.nb.no/Tilbud/Forske/Spraakbanken). English → Norwegian must face the lack of large n-gram models for Norwegian spoken language. New models will therefore need to be trained from in- and out-domain resources. To further improve the reordering, we can also include additional target factors such as POS tags in the translation model and train sequence models for them, as in Birch et al. (2007).

[8] www.hf.uio.no/iln/tjenester/kunnskap/sprak/korpus/flersprakligekorpus/omc/

2.3 The project plan, project management, organisation and cooperation

The project is to last for three years, from the beginning of February 2014 to the end of January 2017. The detailed project plan is provided in Figure 2.

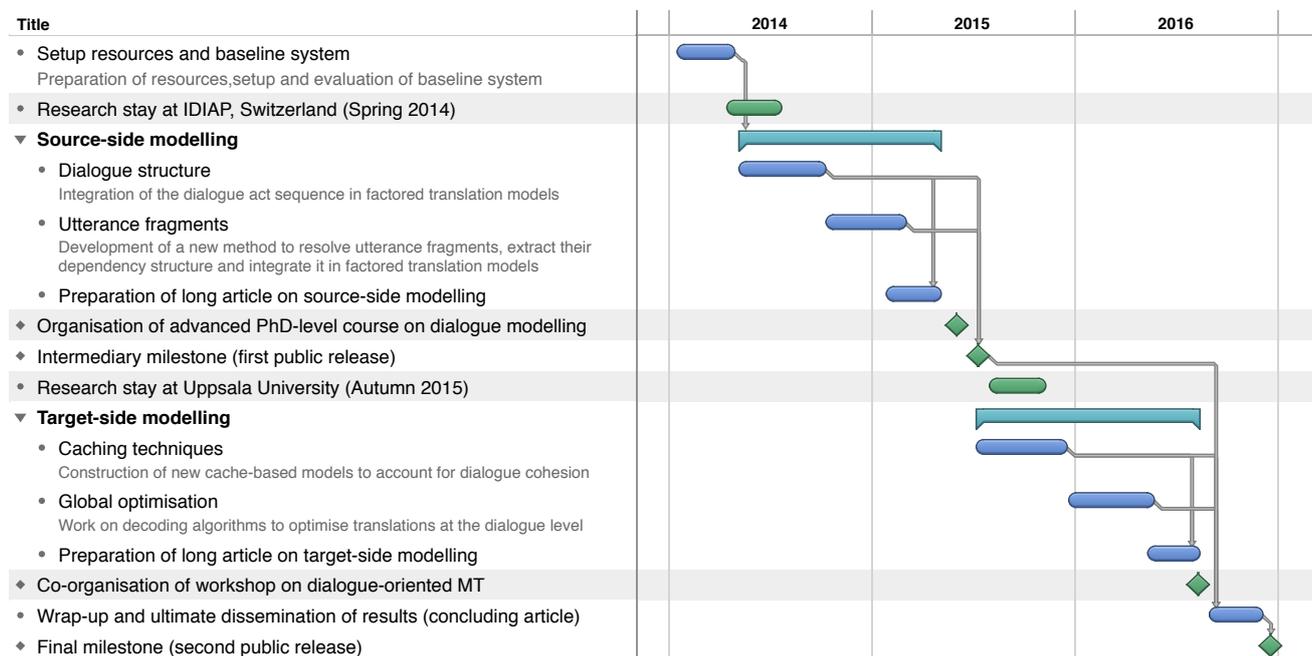| Title	2014	2015	2016
• Setup resources and baseline system			
Preparation of resources,setup and evaluation of baseline system			
• Research stay at IDIAP, Switzerland (Spring 2014)			
▼ **Source-side modelling**			
• Dialogue structure			
Integration of the dialogue act sequence in factored translation models			
• Utterance fragments			
Development of a new method to resolve utterance fragments, extract their dependency structure and integrate it in factored translation models			
• Preparation of long article on source-side modelling			
• Organisation of advanced PhD-level course on dialogue modelling			
• Intermediary milestone (first public release)			
• Research stay at Uppsala University (Autumn 2015)			
▼ **Target-side modelling**			
• Caching techniques			
Construction of new cache-based models to account for dialogue cohesion			
• Global optimisation			
Work on decoding algorithms to optimise translations at the dialogue level			
• Preparation of long article on target-side modelling			
• Co-organisation of workshop on dialogue-oriented MT			
• Wrap-up and ultimate dissemination of results (concluding article)			
• Final milestone (second public release)			

Figure 2: Gantt chart for the proposed project.

The project is divided in four main phases: (1) preparation of resources and baseline system, (2) modelling of the source-side dialogue context, (3) modelling of the target-side context, and (4) ultimate dissemination and wrap-up of the research results. Each phase comprises a full research cycle, from theoretical analyses to the development of translation models and experimental work to evaluate their performance. A minimum of three long journal articles are planned: one focusing on the source-side modelling, one on the target-side modelling, and a concluding article describing the complete framework, including its engineering aspects. Two milestones fix the expected date for the public release of the translation models and resources for the first and second half of the project.

I (Pierre Lison) will be in charge of managing this post-doctoral project, which will be carried through together with several local and international partners. I am currently finishing my Ph.D. in Computer Science at the University of Oslo on "*Structured Probabilistic Modelling for Dialogue Management*" (expected submission: October 2013). I can draw on substantial experience in project-driven research through my participation in three large EU-funded projects while working as a researcher at the German Research Centre for Artificial Intelligence (DFKI), and have published more than 20 peer-reviewed papers on topics related to dialogue modelling and management.

The greatest part of the research work will be carried out at the **Language Technology Group (LTG)** within the Department of Informatics of the University of Oslo. The LTG group specialises in hybrid approaches to NLP that combine rich linguistic knowledge with machine learning. The group coordinated the LOGON project on precision-oriented machine translation from 2002 to 2007, which was one of the flagship projects of the KUNSTI initiative. The group also participates in the META project and is involved in the QTLaunchpad initiative that prepares a large-scale research effort on quality translation for the upcoming Horizon 2020 programme. This blend of competence makes the group an ideal host for the planned research work. The group expertise will be key to the project success through multiple collaborations with group members (notably Stephan Oepen, Erik Velldal and Lilja Øvrelid) on topics such as dependency parsing and high-performance computing.

In addition, the following local and international partners will be collaborating to the project:

IDIAP Laboratory, Swiss Federal Institute of Technology: The IDIAP laboratory is an independent, internationally leading research institute conducting basic and applied research on multimedia management and human–machine multimodal interactions. A three-months visit during the spring 2014 is scheduled to IDIAP's Natural Language Processing Group led by Andrei Popescu-Belis. The group is involved in several Swiss and EU-funded projects, amongst which the COMTIS project on "Improving the Coherence of Machine Translation Output by Modelling Intersentential Relations". The purpose of the visit is to initiate a collaboration on the development of new methods to include rich source-side context into the translation models.

Department of Linguistics and Philology, Uppsala University: The Department of Linguistics and Philology at Uppsala University (Sweden) is one of the foremost centres in Computational Linguistics in Northern Europe, conducting research in areas such as dependency parsing (leading to the well-known MaltParser) and statistical machine translation. Of particular interest is the work of Jörg Tiedemann and colleagues on caching methods and document-wide decoding as part of the Swedish-funded project "Efficient Algorithms for Natural Language Processing Beyond Sentence Boundaries". The three-months visit to Uppsala will allow for collaboration on the target-side optimisation of translation outputs at the dialogue level.

Subtitling department, NRK: The Norwegian public broadcasting company has a long subtitling tradition and has its own in-house department dedicated to the production of high-quality Norwegian subtitles for its four TV channels. NRK will provide both their extensive translation experience and an access to their large repository of subtitles, which will allow us to develop precise language and translation models for Norwegian.

Broadcast Text International: Broadcast Text International (BTI) is one of the largest localisation companies in Europe (with headquarters in Stockholm and multiple regional offices), providing translations, subtitles and dubbing for both broadcast and DVD releases. In addition to offering their insider knowledge of the subtitling business, BTI also agreed to share part of their multilingual database of subtitles for the purpose of the project.

The project will also make extensive use of the high-performance facilities offered by the Abel computer cluster hosted by USIT and jointly funded by Forskningsrådet and UiO. These facilities will be instrumental to the technical success of the project, since statistical machine translation requires large amounts of computational resources for the estimation of its language and translation models. The project needs are anticipated to revolve at around 500 000 core hours per year, and will be applied for through the regular procedure for resource allocation put in place by the NOTUR project.

2.4   Budget

The project budget is described in the application form. It comprises a post-doctoral fellowship for a duration of three years, two outgoing researcher grants, and a small travel allowance.

3   Key perspectives and compliance with strategic documents

3.1   Compliance with strategic documents

The Language Technology Group was established from 2007 to 2012 as a prioritised focal research area within the Faculty of Mathematics and Natural Sciences of the University of Oslo. At the national level, machine translation is of high relevance for the R&D objectives outlined in the strategic roadmap "Veien videre for IKT-satsing", and more specifically its focus on "Teknologier for digitalt innhold". Machine translation is also a major research area within ICT in EU research initiatives, with

several large-scale integrated projects currently funded under FP7's ICT Work Programme (Objective 4.1) owing to its role as a key enabling technology for the "multilingual knowledge society".

3.2   Relevance and benefit to society

Machine translation is a technology of critical importance in today's society where the need to communicate across multiple languages is a daily experience for many, and a necessity for most businesses. Machine translation can notably serve as a vital tool for companies seeking to enter new, emerging markets and offer localised products and services.

The proposed research is likely to yield modest but non negligible improvements in the machine translation quality in conversational domains. As such, the research outcomes will have broad relevance for the ICT sector and various branches of the language industry. In particular, the developed translation models could potentially be exploited by localisation companies to boost the quality and efficiency of the subtitling process, as illustrated by the work done in the SUMAT project with the participation of the European subtitling industry.[9] The research results could also be applied to enhance the performance of speech-to-speech translation technology. While real-time, domain-independent speech-to-speech interpretation still remains technologically challenging, some applications have been successfully developed for limited domains. A particularly inspiring example is the use of speech translation software on mobile devices to break down communication barriers between doctors, volunteers and the local population in the field of humanitarian aid (Eck et al., 2010).

The project will also contribute to strengthening the machine translation expertise in Norway, and showcase how state-of-the-art machine translation technology can be applied to Norwegian, which is an important factor to ensure that Norwegian can continue to fulfil its societal role as "samfunnsbærende språk" (literally, "community-bearing language") in the future.[10] Finally, the technology could benefit multiple branches of the Norwegian public sector, in particular with regard to the public communication with minority groups where it could be applied to translate critical (e.g. health or safety related) information in their native language in addition to Norwegian.

3.3   Environmental impact

The proposed research has no immediate ramifications for the environment.

3.4   Ethical aspects

There are no (known) specific ethical issues relating to the proposed research or exploitation of results. The project will follow standard ethical guidelines for scientific research and publications as outlined by the National Committee for Research Ethics in Science and Technology.[11]

3.5   Gender issues

The questions of gender equality and gender perspectives are not applicable to the proposed project.

4   Dissemination and communication of results

The detailed dissemination plan is provided in the application form.

---

[9]http://www.sumat-project.eu
[10]cf. NOU 2013:4 Kulturutredningen 2014, Kapittel 11.7, Boks 11.6.
[11] http://www.etikkom.no/Forskningsetikk/Etiske-retningslinjer/Naturvitenskap-og-teknologi/

Bibliography

J. Bellegarda. Statistical language model adaptation: review and perspectives. *Speech Communication*, 42(1):93–108, 2004.

A. Birch, M. Osborne, and P. Koehn. CCG supertags in factored statistical machine translation. In *Proceedings of the Second Workshop on Statistical Machine Translation*, pages 9–16, 2007.

H. P. Branigan, M. J. Pickering, and A. A. Cleland. Syntactic co-ordination in dialogue. *Cognition*, 75(2):B13–B25, 2000.

M. Carpuat and D. Wu. Improving statistical machine translation using word sense disambiguation. In *Proceedings of the Joint 2007 EMNLP-CoNLL Conference*, pages 61–72, 2007.

H. H. Clark and E. F. Schaefer. Contributing to discourse. *Cognitive Science*, 13(2):259–294, 1989.

M. Eck, I. Lane, Y. Zhang, and A. Waibel. Jibbigo: Speech-to-speech translation on mobile devices. In *Proceedings of the IEEE Spoken Language Technology Workshop (SLT)*, pages 165–166, 2010.

M. Emele, M. Dorna, A. Lüdeling, H. Zinsmeister, and C. Rohrer. Semantic-based transfer. In W. Wahlster, editor, *Verbmobil: Foundations of Speech-to-Speech Translation*, Artificial Intelligence, pages 359–376. Springer, 2000.

S. Garrod and M. J. Pickering. Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1(2):292–304, 2009.

K. Gimpel and N. Smith. Rich source-side context for statistical machine translation. In *Proceedings of the Third Workshop on Statistical Machine Translation*, pages 9–17, 2008.

Z. Gong, M. Zhang, and G. Zhou. Cache-based document-level statistical machine translation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 909–919, 2011.

Z. Gong, M. Zhang, C. Tan, and G. Zhou. N-gram-based tense models for statistical machine translation. In *Proceedings of the Joint 2012 EMNLP-CoNLL Conference*, pages 276–285, 2012.

C. Hardmeier. Discourse in statistical machine translation. a survey and a case study. *Discours. Revue de linguistique, psycholinguistique et informatique*, (11), 2012.

C. Hardmeier and M. Federico. Modelling pronominal anaphora in statistical machine translation. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 283–289, 2010.

C. Hardmeier, J. Nivre, and J. Tiedemann. Document-Wide Decoding for Phrase-Based Statistical Machine Translation. In *Proceedings of the Joint 2012 EMNLP-CoNLL Conference*, July 2012.

P. Koehn. *Statistical Machine Translation*. Cambridge University Press, New York, NY, USA, 1st edition, 2010.

P. Koehn and H. Hoang. Factored translation models. In *Proceedings of the Joint 2007 EMNLP-CoNLL Conference*, pages 868–876, 2007.

P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. J. Dyer, O. Bojar, A. Constantin, and E. Herbst. Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics*, pages 177–180, 2007.

A. Lavie and M. J. Denkowski. The METEOR metric for automatic evaluation of machine translation. *Machine Translation*, 23(2-3):105–115, September 2009.

R. Le Nagard and P. Koehn. Aiding pronoun translation with co-reference resolution. In *Proceedings of the Joint Fifth Workshop on Statistical Machine Translation and Metrics*, pages 252–261, 2010.

P. Lison. Probabilistic dialogue models with prior domain knowledge. In *Proceedings of the 13th Annual Meeting on Discourse and Dialogue (SIGDIAL 2012)*, pages 179–188, 2012.

D. Liu and D. Gildea. Semantic role features for machine translation. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, pages 716–724, 2010.

T. Meyer, A. Popescu-Belis, N. Hajlaoui, and A. Gesmundo. Machine translation of labeled discourse connectives. In *Proceedings of the Tenth Conference of the Association for Machine Translation in the Americas (AMTA)*, 2012.

J. Nivre, J. Hall, J. Nilsson, A. Chanev, G. Eryigit, S. Kubler, S. Marinov, and E. Marsi. Maltparser: A language-independent system for data-driven dependency parsing. *Natural Language Engineering*, 13(2):95, 2007.

K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 311–318, 2002.

J. S. Pardo. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119:2382, 2006.

V. K. R. Sridhar, S. Bangalore, and S. Narayanan. Enriching machine-mediated speech-to-speech translation using contextual information. *Computer Speech & Language*, 2011.

J. Tiedemann. Context adaptation in statistical machine translation using models with exponentially decaying cache. In *Proceedings of the 2010 Workshop on Domain Adaptation for Natural Language Processing*, pages 8–15, 2010.

D. Vickrey, L. Biewald, M. Teyssier, and D. Koller. Word-sense disambiguation for machine translation. In *Proceedings of the Joint 2005 HLT-EMNLP Conference*, 2005.

R. Wang, P. Osenova, and K. Simov. Linguistically-enriched models for Bulgarian-to-English machine translation. In *Proceedings of the Sixth Workshop on Syntax, Semantics and Structure in Statistical Translation*, pages 10–19, 2012.