Second Year Review
Overview

Philipp Koehn, University of Edinburgh

November 25, 2013
Rationale

• Recently significant improvements to machine translation technology
  **but:** vast majority of this work is targeted towards bulk translation
  that is *good enough* or *fit for use*

• Translation for publication still almost exclusively provided by human translators

• Development of computer aided translation tools lags behind
  – translation memory standard in the translation industry
  – adoption of machine translation slowly becoming increasingly used practice
  – but: very simplistic, limited to post-editing, MT systems insufficiently adapted
    → widely resisted.

⇒ Clear need for better CAT technology
Scope and Objectives

• Cognitive studies of translators leading to insights into interface design
  → better understanding of translator needs

• Workbench with novel types of assistance to human translators
  – interactive translation prediction
  – interactive editing and reviewing
  – adaptive translation models
  → better tools for translators

• Demonstration of effectiveness in field tests with professional translators
  → increased translator productivity
Research and Development Pipeline

Basic scientific research

⇓

Applied scientific research

⇓

Development

⇓

Tests with intrinsic metrics

⇓

Lab tests

⇓

Field tests

⇓

Streamlined implementation and documentation

⇓

Public release of software

⇓

External usage
Project Organization

WP1
Cognitive Modelling

WP2
Interactive Translation Prediction

WP3
Interactive Editing

WP4
Adaptive Translation Models

WP5
Integration

WP6
Dissemination and Evaluation

inform
contribute
deploy
feedback
Partners

University of Edinburgh (UEDIN): expertise in statistical machine translation, interactive machine translation, cognitive studies

Copenhagen Business School (CBS): expertise in translation process studies

Polytechnic University of Valencia (UPVLC): expertise in statistical machine translation, interactive machine translation

Celer Solutions (CELER): expertise in employing novel computer aided technology in professional practice

All academic partners have developed a CAT tool before the project.
Modular design allows integration into existing CAT tools
Expected Results

• Better understanding of human translation process

• Open source workbench for
  – translation process studies
  – practical use by professional and volunteer translators

• Demonstrated effectiveness of novel types of assistance

• Dissemination of project outcomes
  – workshops and tutorials
  – user group of early adopters
  – integration into community translation platforms
Scientific Overview: Year 2

- Public release of beta version of workbench

- Second field trial (in month 20)

- Analysis and cognitive studies based on user activity data

- Advanced methods
  - new paradigms for interactive machine translation
    (multi-modality: e-pen and speech, active interaction, prediction based as machine learning problem, prediction for tree-based models)
  - adaptation of machine translation models
    (online learning, active learning, domain and user adaptation)
  - new types of assistance for human translators
    (advanced confidence measures, multiple translation options, authoring assistance)
WP5: Joint Development with MATECAT

- Development of basic workbench by MATECAT

- Integration of all CASMACAT functionality into MATECAT prototype

- Different focus
  - MATECAT: efficient tool for professional translators
  - CASMACAT: advanced types of assistance, logging, eye-tracking, replay

- Communication
  - shared git repository
  - technical meetings and phone calls
  - email
WP5: Workbench

- Modular design: allows use of selected CASMACAT components (e.g., multiple machine translation backends: Moses and Thot)

- Advanced CAT features
  - interactive translation prediction
  - display of multiple translation options
  - requesting alternative translations
  - word-level confidence measures
  - visualization of word alignment
  - search and replace function

- E-pen interaction

- Logging and replay

- Eye tracking
WP5: Interactive Machine Translation

- Alternative Implementation
- Inspired by "auto-complete" in text input
WP5: Gaze-to-Word Mapping

- Recorded gaze locations and fixations:

  Right eye gaze samples

  Left eye gaze samples

- Gaze-to-word mapping:
WP6: Field Trial

- English–Spanish, news domain
- User improved machine translation system (+2 BLEU)
- Improved baseline post-editing speed over last year
- Three modalities
  - post-editing
  - interactive translation prediction
  - interactive translation prediction with advanced visualization
- Results
  - no significant difference between modalities
  - learning curve: translators become much faster over time (3 days)
WP6: Data From Field Trial

- Many valuable data resources generated from field trial
  - Eye-tracking data enables cognitive studies
  - Word correction data enables better word confidence measures
  - Postediting time enables better sentence confidence measures
  - Test data for interactive MT research

- Data and models will be made publicly available
  (used in WMT 2013 shared task on confidence estimation)
Machine Translation Quality Matters

Experiment:
Post-editing with different machine translation systems
English–German, news stories

Note: study ongoing, very preliminary results
WP2: E-pen Interaction

deletion  \[\text{if} \ 1 \ \text{any} \ 2 \ \text{feature} \ 3 \ \text{not} \ 4 \ \text{is} \ 5 \ \text{available} \ 6 \ \text{on} \ 7 \ \text{your} \ 8 \ \text{network} \ 9\]

insertion  \[\text{if} \ 1 \ \text{any} \ 2 \ \text{feature} \ 3 \ \text{not} \ 4 \ \text{is} \ 5 \ \text{available} \ 6 \ \text{on} \ 7 \ \text{your} \ 8 \ \text{network} \ 9\]

substitution  \[\text{if} \ 1 \ \text{any} \ 2 \ \text{feature} \ 3 \ \text{not} \ 4 \ \text{is} \ 5 \ \text{available} \ 6 \ \text{on} \ 7 \ \text{your} \ 8 \ \text{network} \ 9\]

shift  \[\text{if} \ 1 \ \text{any} \ 2 \ \text{feature} \ 3 \ \text{not} \ 4 \ \text{is} \ 5 \ \text{available} \ 6 \ \text{on} \ 7 \ \text{your} \ 8 \ \text{network} \ 9\]

transposition  \[\text{if} \ 1 \ \text{any} \ 2 \ \text{feature} \ 3 \ \text{not} \ 4 \ \text{is} \ 5 \ \text{available} \ 6 \ \text{on} \ 7 \ \text{your} \ 8 \ \text{network} \ 9\]

Error rate with state-of-the-art gesture recognizer \(\approx 10\%\)
WP2: Confidence Measures

And on that the signs are mixed.  

Y en que los indicios son desiguales.

- Two highlight thresholds:
  - **Red**: Probable incorrect translations
  - **Orange**: Dubious translations
WP3: Translation Options in Context

Speaking in Latin to a small gathering of cardinals at the Vatican on Monday morning, Benedict

Treffen

Inform better about our 2nd 's pro-nuclear ladies and gentlemen, a 2011 event in that regard at a

Erfassung

systems - requirements and prevention and on used but comprehensive

Erhebung

An often elusive figure, he spent much of his papacy in the shadow of his beloved predecessor.
WP3: Display of Alternative Translations

Request
WP3: Display of Alternative Translations

Display

Norway’s rakfisk: Is this the world’s smelliest fish?

Norwegian rakfisk: Is this the world’s smelliest fish?

Norway’s five million people enjoy the highest standards of living in the world.

Could the secret to their well-being be the local appetite for guttapercha, a species of fish?

Take a selection of other dishes.

Norway’s rakfisk: Is this the world’s smelliest fish?

Пять миллионов людей Норвегии наслаждаются одним из
высочайших жизненных уровней не только в Европе, но и в мире

в мире

во всем мире

в остальном мире

на всем свете

в остальных частях мира

CANCEL ACCEPT
WP4: Adaptation

- Online learning
  - user translates a sentence
  - sentence pair is added to model immediately

- Active learning
  - system identifies most valuable sentence to translate
  - information gained from sentence translation is propagated
  - translator may only post-edit part of output

- Adaptation to domain and user
  - optimize performance for task at hand
  - novel adaptation method: Bayesian adaptation
Dissemination Strategy

- Year 1
  - raise awareness for the topic
  - initial contacts with external researchers

- Year 2
  - public release of beta version of workbench
  - promotion of workbench
  - collaboration with external researchers
  - initial contacts with end-users

- Year 3
  - public release of full version of workbench
  - integrate contributions from external researchers
  - collaboration with end-users
Dissemination

• Meetings
  – Machine Translation Summit: booth, demonstration of the workbench
  – 7th Workshop on Statistical Machine Translation: co-organized shared task on confidence measures
  – Third PhD Course in Translation Processes Research
  – Special symposium on Empirical approaches to gaze data analysis in reading, writing and translation
  – COLING workshop on Eye-tracking and Natural Language Processing
  – NLPCS Tutorial on Empirical Translation Process Research

• 12 invited talks

• 24 publications
Dissemination Highlights

• Public Release of Beta Version
  – installation instructions on web site
  – two publications promoting the tool (MT Summit, MT Marathon)
  – hands-on assistance with installation
  – first external use: Charles University for WMT evaluation

• Machine Translation Summit
  – great interest in workbench by researchers and industry
  – contacts that will lead to external field trials (Autodesk, EU DGT)

• SEECAT Workshop
  – research workshop at CBS
  – integration of speech recognition into workbench
User Group

- Initial solicitation of interest: project mailing list, 87 subscribers

- Goals for core user group: collaboration on research projects and field trials

- 9 Members
  - Autodesk, represented by Mirko Plitt
  - Charles University, represented by Ondrej Bojar
  - GXP Language Services, represented by Siegfried Armbruster
  - European Parliament, represented by Pedro Garcia-Dieuguez
  - Federal University of Minas Gerais, represented by Fabio Alves
  - Pactera Technology Spain, represented by Ana Guerberof
  - Pangeanic, represented by Manuel Herranz
  - Universitat Autònoma de Barcelona, represented by Anna Aguilar-Amat
  - Univesitat Mainz, represented by Čulo, Oliver
questions?