A Comparison of Machine Translation Paradigms for Use in Black-Box Fuzzy-Match Repair

AMTA 2018, Boston, March 21st, 2018
Rebecca Knowles
John E. Ortega
Philipp Koehn
Overview

- Fuzzy-Match Repair
- Comparison of MT Paradigms
- Results & Analysis
- Future Work
Introduction to Fuzzy-Match Repair

01 The Source Sentence (s')
The cat **blinks** when the dog arrives

02 The TM Source (s)
The cat **runs** when the dog arrives

03 The TM Target (t)
El gato **corre** cuando llega el perro

Our Fuzzy-Match Repair algorithm will repair proposals from the TM and propose translation hypotheses closer to the source sentence.
Introduction to Fuzzy-Match Repair

The Translator
When working with fuzzy matches, the translator has to make changes to transform $t$ into an adequate translation of $s'$.

Translation Proposals
Our goal is to repair fuzzy matches and provide translation proposals so that the amount of post-editing by the translator is kept to a minimum.
FMR Algorithm

01  Align input source \((s')\) to TM source \((s)\)

02  Translate mismatches

03  Match translations to their TM target \((t)\)

04  Build pairs of repair operators \((\sigma',\sigma)(\tau',\tau)\)

05  Generate hypotheses \((t^*)\)
FMR Algorithm

The **blue** dog barks (**s’* source)**
The **red** dog barks (**s tm-source**)
El perro **rojo** ladra (**t tm-target**)

**σ’** - The **blue** dog, **blue** dog, **blue**
**σ** - The **red** dog, **red** dog, **red**
**τ’** - el perro **azul**, perro **azul**, **azul**
**τ** - el perro **rojo**, perro **rojo**, **rojo**

El perro **azul** ladra (**t* the best (oracle) of many hypotheses**
Oracle Evaluation (for FMR)

\[
\sum_{i=0}^{N} \frac{\text{ED}(t_i^*, r_i)}{\sum_{i=0}^{N} |r_i|}
\]

- Get TUs that meet fuzzy-match threshold
- If no TU meets threshold, use MT. Otherwise, get highest scoring TU and produce all possible hypotheses.
- Select repair with minimum edit distance.
Our approach to fuzzy-match repair allows the use of any external source of bilingual information (SBI) such as rule-based, statistical, or neural machine translation systems, dictionaries, and more ...
Introduction to Fuzzy-Match Repair

Previous Work
- FMR introduced
- Oracle evaluation on 3 language pairs

Current Work
- 3 MT Paradigms
- Oracle performance eval. & sub-segment analysis
Machine Translation Paradigms

**Rule-Based (RB)**
- Apertium

**Statistical (SMT)**
- Moses
  - *Training:*
    - Europarl, News Commentary, DGT-TM 2011-13
    + Large LM

**Neural (NMT)**
- Nematus
  - *Training:*
    - Europarl, News Commentary
    + DGT-TM 2011-13
Results & Analysis

- Compare System Performance: Translation & Oracle FMR
- Direct Comparison of Two Best Systems
- Analysis of Sub-Segment Translations
System Performance

SMT performs best for translation...

![Graph showing WER and BLEU scores for RB, SMT, and NMT systems. WER scores are 60.8, 35.2, and 36.8 respectively, while BLEU scores are 19.2, 57.2, and 52.6 respectively.]
SMT performs best for translation...but NMT performs best for FMR.

**System Performance**

**WER for FMR**

- TM
- RB
- SMT
- NMT

Fuzzy-Match Threshold: 60%, 70%, 80%
Direct Comparison: SMT vs. NMT

NMT is able to repair more segments
And it produces more repair options per segment

On a subset of segments that NMT & SMT both repair:
FMR performance is very similar between SMT and NMT

More repair options (NMT) gives better FMR performance
True under the oracle evaluation, but with a pessimal oracle, NMT suffers a greater performance drop than SMT
## Sub-Segment Translations

| Source | SMT | NMT |
|--------|-----|-----|
| annex 3 ; it cannot be furnished | el anexo 3 ; no podrán aportarse | anexo 3 |
| 's authorities shall within | dentro de las autoridades de | las autoridades de los estados miembros dispondrán de las autoridades nacionales competentes en el |
| ( place and date ) | ( lugar y fecha ) | ( lugar y fecha ) |
| ( 2 | ( 2 , | ( 2 |
Future Work

**Fuzzy-match Repair**
- Paper presented at AMTA with initial idea and concept in 2014

**Black-Box MT paradigms and sub-segment analysis**
- Presented at AMTA 2018

**2016**
- Idea formalized and algorithm released to the MT community at AMTA 2016

**2018+**
- Formalize features for Quality Estimation in FMR to rank hypotheses with unseen reference.
Thank you!

Rebecca Knowles (rknowles@jhu.edu)
John E. Ortega (jeo10@alu.ua.es)
Philipp Koehn (phi@jhu.edu)
This work was partially supported by a National Science Foundation Graduate Research Fellowship under Grant No. DGE-1232825 (to the first author) and by the Spanish government through the EFFORTUNE (TIN2015-69632-R) project (the second author).

Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors(s) and do not necessarily reflect the views of the National Science Foundation.