School of Engineering

Project Work
HS17 Degree: Computer Science

Techniques for ML-Assisted Language Translation

Authers	Nicolas Hoferer Daniel Einars
Main Referee	Marc Cieliebak
Assisting Referees	Kurt Stockinger Jan Milan Deriu
Date	30.10.2017

s

Abstract

This is just some normal text that goes here

Preface

thank-yous go here

DECLARATION OF ORIGINALITY
Project Work at the School of Engineering

By submitting this Project Work, the undersigned student confirms that this thesis is his/her own work and was written without the help of a third party. (Group works: the performance of the other group members are not considered as third party).

The student declares that all sources in the text (including Internet pages) and appendices have been correctly disclosed. This means that there has been no plagiarism, i.e. no sections of the Bachelor thesis have been partially or wholly taken from other texts and represented as the student's own work or included without being correctly referenced.

Any misconduct will be dealt with according to paragraphs 39 and 40 of the General Academic Regulations for Bachelor's and Master's Degree courses at the Zurich University of Applied Sciences (Rahmenprüfungsordnung ZHAW (RPO)) and subject to the provisions for disciplinary action stipulated in the University regulations.

City, Date:

. . . .

Signature:

. . . .

. . . .

. . . .

The original signed and dated document (no copies) must be included after the title sheet in the ZHAW version of all Bachelor thesis submitted.

Zurich University of Applied Sciences

Inhaltsverzeichnis

1. Introduction

- machines becoming better at processing human language (accuracy)

- conversation with machines are possible to a limited degree

- information retrieval through voice recognition still a challenge due to
    - attention to correct words
    - database structures
    - multiple ways to ask for identical information

- multiple solutions proposed
    - KBQA: Learning Question Answering over QA Corpora and Knowledge Bases
    - Eric, Manning - 2017 - Key-Value Retrieval Networks for Task-Oriented Dialogue - With Highlights

- Asking your Assistant (Google, Siri or S-Voice) weather you have an appointment tomorrow and ask follow-up questions about this appointment is currently not possible (due to above challenges but could be if these papers prove implementable

1.1. Initial Position

- No Response from KBQA for Code

- Refusal to share code from Manning

- Ultimate new goal: Implement Manning's solution without his code

1.2. Task

- Small Steps
    - implement seq2seq network for translation
        * implement char-based
        * implement word-based
        * try multiple different implementations (reversed-input, multiple LSTMs) and compare against each other
        * get decent results on both and move on
    - implement seq2seq with attention
        * attempt various attention mechanism

- One Large Step
    - map best working models and tools to KBQA and get better results than Stanford
    - Rub better results in Eric's face.
    - Profit.

2. Theoretical Principles

Test this one here too, please

```
1  int getRandomNumber()
2  {
3      return 4; // chosen by fair dice roll.
4                // guaranteed to be random.
5  }
```

2.1. Definitions

- Hypothesis
- Reference

2.2. Recurrent Neural Networks

- Standard Neural Networks
- Recurrent Neural Networks
  - Problems
  - Solutions

2.3. Seq2Seq

- encoder
- decoder

2.4. Attention

- Mechanisms

2.5. Performance Evaluation

- Translation
  - Bleu
  - others
- KB-Retrieval
  - Bleu
  - sent2vec

3. Experiments

- list experiments

- list results

4. Results

5. Discussion and Prospects

Wie in (**?**, Kapitel 2, Seite 215) nachzulesen, gibt es sogenannte Gleichungen.HRZélitismΩ

6. Index

6.1. Glossary

This document is incomplete. The external file associated with the glossary 'main' (which should be called zhawDocument.gls) hasn't been created.
Check the contents of the file zhawDocument.glo. If it's empty, that means you haven't indexed any of your entries in this glossary (using commands like \gls or \glsadd) so this list can't be generated. If the file isn't empty, the document build process hasn't been completed.
You may need to rerun LaTeX. If you already have, it may be that TeX's shell escape doesn't allow you to run makeindex. Check the transcript file zhawDocument.log. If the shell escape is disabled, try one of the following:

- Run the external (Lua) application:

  makeglossaries-lite.lua "zhawDocument"

- Run the external (Perl) application:

  makeglossaries "zhawDocument"

Then rerun LaTeX on this document.
This message will be removed once the problem has been fixed.

6.2. List of Figures

6.3. List of Tabels

6.4. Listingverzeichnis

6.5. Symbol Glossary

Ω unit of electrical resistance 10

6.6. Acronym Glossary

This document is incomplete. The external file associated with the glossary 'acronym' (which should be called zhawDocument.acr) hasn't been created.
Check the contents of the file zhawDocument.acn. If it's empty, that means you haven't indexed any of your entries in this glossary (using commands like \gls or \glsadd) so this list can't be generated. If the file isn't empty, the document build process hasn't been completed.
You may need to rerun LaTeX. If you already have, it may be that TeX's shell escape doesn't allow you to run makeindex. Check the transcript file zhawDocument.log. If the shell escape is disabled, try one of the following:

- Run the external (Lua) application:

  makeglossaries-lite.lua "zhawDocument"

- Run the external (Perl) application:

  makeglossaries "zhawDocument"

Then rerun LaTeX on this document.
This message will be removed once the problem has been fixed.

[title=Index]

A. Appendix

A.1. Projektmanagement

A.2. Final Words