Universität Zürich[UZH]

Masterarbeit
zur Erlangung des akademischen Grades
Master of Arts
der Philosophischen Fakultät der Universität Zürich

(Titel)

Verfasserin/Verfasser: Jonathan Schaber
Matrikel-Nr: 11-771-359

Referentin/Referent: Dr. Simon Clematide

[Betreuerin/Betreuer: (Titel Vorname Name) [nur falls vom Ref. unterschiedlich]]

Institut für Computerlinguistik

Abgabedatum: (xx.xx.xxxx)

Abstract

This is the place to put the English version of the abstract.

Zusammenfassung

Und hier sollte die Zusammenfassung auf Deutsch erscheinen.

Acknowledgement

I want to thank X, Y and Z for their precious help. And many thanks to whoever for proofreading the present text.

Contents

List of Figures

List of Tables

List of Acronyms

POS	Part-Of-Speech
SRL	Semantic Role Labelling OR Semantic Role Labeller
STTS	Stuttgart-Tübingen-TagSet

1 Introduction

1.1 Motivation

Some words on your motivation would be nice.

1.2 Research Questions

The research questions that shall be answered in this thesis, are:

1. What do I do?
2. How do I do it?
3. And why?

1.3 Thesis Structure

In this first chapter ...
Chapter 2 introduces ...
Chapter 3 ...

2 Semantic Roles

2.1 Overview

"The main reason computational systems use semantic roles is to act as a shallow meaning representation that can let us make simple inferences that aren't possible from the pure surface string of words, or even from the parse tree." [Jurafsky and Martin, 2019, p. 375]

3 Data Sets

3.1 Why create an own corpus?

3.2 Corpora

3.2.1 deISEAR

As Troiano et al. [2019] write in their

3.2.2 MLQA_V1

Lewis et al. [2019] compiled

3.2.3 PAWS-X

Yang et al. [2019]

3.2.4 SCARE

Sänger et al. [2016]

3.2.5 XNLI

Conneau et al. [2018]

3.2.6 XQuAD

Artetxe et al. [2019]

4 Architecture

4.1 Overview

4.2 Semantic Role Labeller

State-of-the-art semantic role labellers (SRLs) are end-to-end models, implementing For my system, I implement the DAMESRL, a model presented by Do et al. [2018]. I use their pre-trained German Character-Attention model which, according to the authors, achieved an F1 score of 73.5 on the CoNLL'09 task [Hajič et al., 2009].

"A major advantage of dependency grammars is their ability to deal with languages that are morphologically rich and have a relatively free word order." [Jurafsky and Martin, 2019, p. 274] For extracting predicates, I rely on the dependency tree the ParZu parser Sennrich et al. [2013] generates for a given sentence. Since one sentence can have multiple predicate-argument structures, I need to device an algorithm to extract the relevent predicates in a sentence. This is not as straight forward as it seems on the first look. Consider the sentence `Die Pakete werden einzeln weitergeleitet, was manchmal zu unterschiedlichen Pfaden und einer fehlerhaften Übertragung führt.` whose dependency parse tree is shown in Figure 1: This sentence has three verbs in it, `werden`, texttweitergeleitet, and `führt`, but only the last two of them are relevant predicates, i.e. pradicates that carry "true" semantics.

I propose the following algorithm deciding whether a verb in a sentence is or isn't a predicate using the a heuristic, mostly relying on the STTS-tags [Schiller et al., 1999] of the parser's output:

- if the POS-tag of a token t is not "V" — i.e. it is a not verbal form —, it is immediately labelled `NOT_PRED`
- if the POS-tag of a token t is verbal and its STTS-tag is "VVFIN" — i.e. it is a inflected full-verb — it is immediately labelled `PRED`
- if the POS-tag of a token t is verbal and its STTS-tag is not "VVFIN", it is

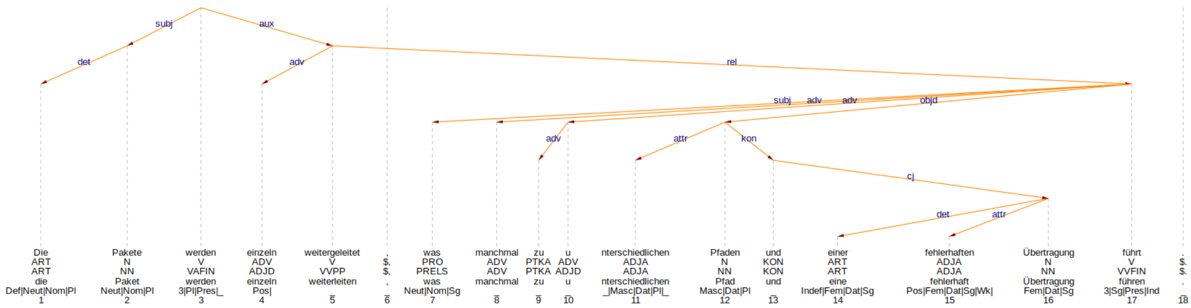

Figure 1: Example sentence with multiple predicates.

checked whether any of the other tokens in this subclause is POS-tagged as "V" and if it is dependent on t, i.e. if it points to the index of t. If this is the case for any token in the subclause, t gets labelled `NOT_PRED`, since it only modifies another verb. If there is no token that fulfills the above mentioned criteria, t is considered a predicate and gets labelled `PRED`.

4.3 German BERT

5 Results

5.1 BLEU Scores

Table 1 shows how to use the predefined tab command to have it listed.

language pair	ABC	YYY
EN→DE	20.56	32.53
DE→EN	43.35	52.53

Table 1: BLEU scores of different MT systems

And we can reference the large table in the appendix as Table 2

5.2 Evaluation

We saw in section 5.1

We will see in subsection 5.2.1 some more evaluations.

5.2.1 More evaluation

5.3 Citations

Although BLEU scores should be taken with caution (see Callison-Burch et al. [2006]) or if you prefer to cite like this: [Callison-Burch et al., 2006] ...

to cite: [Koehn, 2005, 30-31]
to cite within parentheses/brackets: [Koehn, 2005], [Koehn, 2005, 30-32]

to cite within the text: Koehn [2005], Koehn [2005, 37]
only the author(s): Callison-Burch et al.
only the year: 2006

5.4 Graphics

To include a graphic that appears in the list of figures, use the predefined fig command:

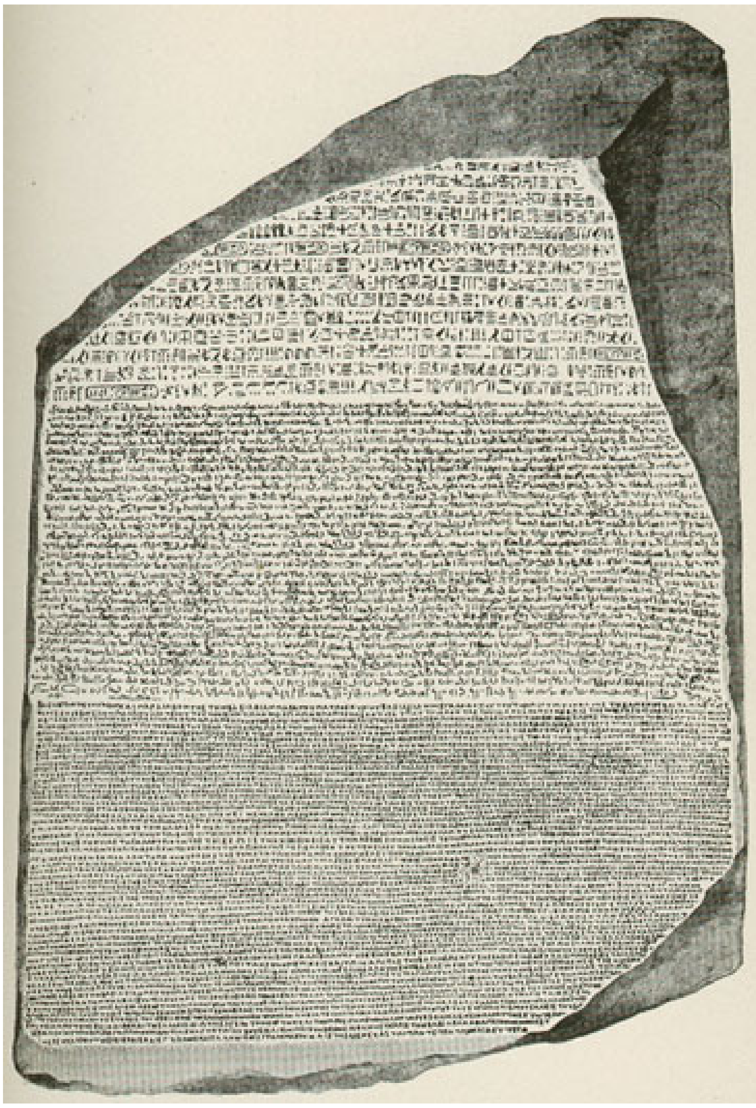

Figure 2: The Rosetta Stone

And then reference it as Figure 2 is easy.

5.5 Some Linguistics

(With the package 'covington')

Gloss:

(5.1)	*The*	*cat*	*sits*	*on*	*the*	*table.*
	die	Katze	sitzt	auf	dem	Tisch

'Die Katze sitzt auf dem Tisch.'

Gloss with morphology:

(5.2)	*La*	*gata*	*duerm*	*-e*	*en*	*la*	*cama.*
	Art.Fem.Sg	Katze	schlaf	-3.Sg	in	Art.Fem.Sg	Bett

'Die Katze schläft im Bett.'

6 Conclusion

In this project we have done so much.[1]

We could show that ...

Future research is needed.
The show must go on.

[1] Thanks to many people that helped me.

Glossary

Of course there are plenty of glossaries out there! One (not too serious) example is the online MT glossary of Kevin Knight [2] in which MT itself is defined as

> techniques for allowing construction workers and architects from all over the world to communicate better with each other so they can get back to work on that really tall tower.

accuracy A basic score for evaluating automatic **annotation tools** such as **parsers** or **part-of-speech taggers**. It is equal to the number of **tokens** correctly tagged, divided by the total number of tokens. [. . .]. (See **precision and recall**.)

clitic A morpheme that has the syntactic characteristics of a word, but is phonologically and lexically bound to another word, for example *n't* in the word *hasn't*. Possessive forms can also be clitics, e.g. The dog*'s* dinner. When **part-of-speech tagging** is carried out on a corpus, clitics are often separated from the word they are joined to.

[2] Machine Translation Glossary (Kevin Knight): `http://www.isi.edu/natural-language/people/dvl.html`

References

M. Artetxe, S. Ruder, and D. Yogatama. On the cross-lingual transferability of monolingual representations. *arXiv preprint arXiv:1910.11856*, 2019.

C. Callison-Burch, M. Osborne, and P. Koehn. Re-evaluating the Role of BLEU in Machine Translation Research. In *Proceedings of the European Association of Computational Linguistics (EACL)*, pages 249–256, 2006. URL `http://www.aclweb.org/anthology/E/E06/E06-1032.pdf`.

A. Conneau, G. Lample, R. Rinott, A. Williams, S. R. Bowman, H. Schwenk, and V. Stoyanov. Xnli: Evaluating cross-lingual sentence representations. *arXiv preprint arXiv:1809.05053*, 2018.

Q. N. T. Do, A. Leeuwenberg, G. Heyman, and M. F. Moens. A flexible and easy-to-use semantic role labeling framework for different languages. In *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, pages 161–165, 2018.

J. Hajič, M. Ciaramita, R. Johansson, D. Kawahara, M. A. Martí, L. Màrquez, A. Meyers, J. Nivre, S. Padó, J. Štepánek, et al. The conll-2009 shared task: Syntactic and semantic dependencies in multiple languages. 2009.

D. Jurafsky and J. H. Martin. Speech and language processing (draft). october 2019. *URL https://web. stanford. edu/˜ jurafsky/slp3*, 2019.

P. Koehn. Europarl: A Parallel Corpus for Statistical Machine Translation. In *Proceedings of the MT Summit 2005*, pages 79–86, 2005. URL `http://www.iccs.inf.ed.ac.uk/~pkoehn/publications/europarl-mtsummit05.pdf`.

P. Lewis, B. Oğuz, R. Rinott, S. Riedel, and H. Schwenk. Mlqa: Evaluating cross-lingual extractive question answering. *arXiv preprint arXiv:1910.07475*, 2019.

M. Sänger, U. Leser, S. Kemmerer, P. Adolphs, and R. Klinger. Scare—the sentiment corpus of app reviews with fine-grained annotations in german. In

Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), pages 1114–1121, 2016.

A. Schiller, S. Teufel, C. Stöckert, and C. Thielen. Guidelines für das tagging deutscher textcorpora. *University of Stuttgart/University of Tübingen*, 1999.

R. Sennrich, M. Volk, and G. Schneider. Exploiting synergies between open resources for german dependency parsing, pos-tagging, and morphological analysis. In *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, pages 601–609, 2013.

E. Troiano, S. Padó, and R. Klinger. Crowdsourcing and validating event-focused emotion corpora for german and english. *arXiv preprint arXiv:1905.13618*, 2019.

Y. Yang, Y. Zhang, C. Tar, and J. Baldridge. Paws-x: A cross-lingual adversarial dataset for paraphrase identification. *arXiv preprint arXiv:1908.11828*, 2019.

Lebenslauf

Persönliche Angaben

Ich Persönlich
Meinestrasse Nr
PLZ Wohnort
ichpersoenlich@uzh.ch

Schulbildung

2012-2014	Bachelor-Studium Computerlinguistik und Sprachtechnologie an der Universität Zürich
seit 2014	Master

Berufliche und nebenberufliche Tätigkeiten

2012–2013	Tutorate PCL I+II

A Tables

Part of speech	POS type	POS	number of labels in my corpus
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	DET	**35**	280
14	Total	**35**	280

Table 2: Some very large table in the appendix

B List of something

This appendix contains a list of things I used for my work.

- apples
  - `export2someformat`
- bananas
- oranges
  - `bleu4orange`
  - `rouge2orange`