Cross-lingual Transfer of Semantic Role Labeling Models

Mikhail Kozhevnikov and Ivan Titov
Semantic Role Labeling

Dependency-based, like in CoNLL 2009 ST

Aileen quickly showed Rob the book

Predicate Identification
Predicate Disambiguation
Argument Identification
Argument Classification
The Low-resource Setting

- Training requires large amounts of annotated data
- Even large corpora face coverage problems
- Very little or no data for many new languages
Unsupervised SRL

Aileen quickly showed Rob the book

The film was shown last night

Show me the map

[Grenager and Manning, 2006]
[Titov and Klementiev, 2012]
[Lang and Lapata, 2010]
Cross-lingual Approaches: Projection

- Run source-language model on the source side
- Propagate annotations through word-alignment links
- Train a target-language model on the output

Source-language model

Target-language model

The film was shown last night

La película se mostró anoche
Cross-lingual Approaches: Model Transfer

The film was shown last night.

La película fue muy aburrida.

Successfully applied to dependency parsing, NER, etc.

[Zeman and Resnik, 2008] [Durrett et al., 2012]
[Søgaard, 2011] [McDonald et al., 2011]
Overview

Purpose

- Create a simple model
- Compare against the alternatives in low-resource setting
- Figure out which features are useful

Model

- Independent linear classifiers for each argument
- No feature selection, no second-order features
Outline

- Motivation
- Cross-lingual Approaches
- Shared Feature Representation
- Evaluation and Baselines
- Results
- Conclusions and Outlook
Features of an argument instance include attributes of:

- Argument word
- Predicate word
- Parent
- Children
- Siblings
- Preceding and following words

Dog

Cross-lingual cluster ID: c123

Universal POS tag: NOUN

Distributed word representation: 0.84, -0.02, -0.11, ...

[Täckström et al., 2012]

[Petrov et al., 2012]

[Klementiev et al., 2012]
Dependency features

- (Unlabeled) dependency structure
  - Gold-standard dependencies stripped of dependency relations
- Direct transfer  [McDonald et al., 2011]

Dependency relations

- Currently only PCEDT
- Need more homogeneous treebanks
  
  - [Zeman et al., 2012]
  - [McDonald et al., 2013]

| Pair     | UAS |
|----------|-----|
| En-Zh    | 35% |
| Zh-En    | 42% |
| En-Cz    | 36% |
| Cz-En    | 39% |
| En-Fr*   | 67% |

En-Fr: evaluation against predicted, not gold syntax
Evaluation Data

Need: different languages, similar annotation

- **English-Czech**
  - Prague Czech-English Dependency Treebank 2.0
  - Similar dependency relations on the two sides
  - [Hajič et al., 2012]

- **English-French**
  - Annotation projection from English (PropBank)
  - 1000 sentences manually corrected
  - [van der Plas et al., 2011]

- **English-Chinese**
  - Chinese Treebank, guidelines similar to PropBank
  - Core roles only (no modifiers)
Baselines

- **Annotation projection baseline**
  - Apply annotation projection to parallel data (except for French)
  - Train a lexicalized model on the output
  - Compare in terms of F1 score

- **Unsupervised baseline**  \[\text{[Titov and Klementiev, 2012]}\]
  - Compare to an unsupervised SRL system using cluster measures
  - Classification only
Argument Identification

\[ F_1 \]

- **projection (t)**
- **transfer (t)**
- **projection (o)**
- **transfer (o)**

(t): transferred syntax
(o): original syntax

- **EN-ZH**
- **ZH-EN**
- **EN-CZ**
- **CZ-EN**
- **EN-FR**
Argument Classification: Top Ten PCEDT 2.0 Labels

$F_1$
Argument Classification: Supervised Evaluation

Proper projection model is somewhat better

Comparable performance

$F_1$
Argument Classification: Unsupervised Evaluation

Unsupervised model performs better with original dependencies

$F^C_i$
Feature Group Contribution: En-Fr, transferred syntax

- POS, Synt, Cls, Gloss
- POS, Cls, Gloss
- POS, Synt, Gloss
- POS, Synt, Cls
- POS, Gloss
- POS, Cls
- POS, Synt
- POS

$F_1$
Performance Drop Due to Transfer

![Bar chart showing performance drop due to transfer.

- **Source** and **Target** bars for each language pair (EN-ZH, ZH-EN, EN-CZ, CZ-EN, EN-FR).

- The chart compares $F_1$ scores for source and target languages.

- EN-ZH: Source - 90, Target - 70
- ZH-EN: Source - 90, Target - 70
- EN-CZ: Source - 80, Target - 60
- CZ-EN: Source - 80, Target - 60
- EN-FR: Source - 90, Target - 60

The chart highlights significant performance drops in translation quality when transferring between different language pairs.
Conclusions

- Transfer performs comparably to annotation projection
  - Easy to implement and trivial to apply
  - Does not require high-quality parallel data
- May work better than unsupervised SRL where no accurate dependency parser is available for the target language
Current and Future Work

- Better shared representation for dependency features
- Model inter-argument dependencies
- Multiple source languages
- Domain adaptation techniques
  - Refine using parallel data
Acknowledgments

- Alexandre Klementiev
- Ryan McDonald
- Oscar Täckström
- Slav Petrov