Semantic Role Labeling with Neural Networks for Texts in Russian

A.O. Shelmanov, D.A. Devyatkin

shelmanov@isa.ru,
devyatkin@isa.ru

FRC CSC RAS

2017, Moscow
Semantic role labeling

Semantic role labeling (SRL) (shallow semantic parsing):
1. Determines situations in sentence
2. Identifies arguments of situations
3. Classifies arguments of situations and assigns them thematic roles

Example:
“Month ago from Moscow to St. Petersburg, metropolitan restorers sent an unusual cargo.”

Imagine a tree diagram illustrating the semantic roles and their corresponding values:
- **Subject (Who?):** restorers
- **Object (What?):** cargo
- **Directive (Where to?):** to St. Petersburg
- **Ablative (From where?):** from Moscow
- **Temporarive (When?):** Month ago

Example sentence diagram:
- **sent отправили**
- **restorers реставраторы**
- **cargo груз**
- **Month ago Месяц назад**
- **to St. Petersburg в Петербург**
- **from Moscow из Москвы**
Applications of semantic role labeling

• Question-answering search
  • Shen D. and Lapata M., 2007

• Information extraction
  • Christensen J. et al., 2010

• Information search
  • Osipov G. et al., 2016

• Summarization
  • Khan A. et al., 2015

• Machine translation
  • Xiong D. et al., 2012
  • Bazrafshan M. and Gildea D., 2013

• Event extraction ~ Semantic role labeling
  • Wang X. et al., 2012
Related work

• Seminal work for statistical and machine learning methods for SRL:
  • Gildea D. and Jurafsky D., 2000

• Main corpora:
  • FrameNet (Baker C. F., Fillmore C. J., Lowe J. B., 1998)
  • PropBank (Kingsbury P. and Palmer M., 2002)

  • Hajič J. et al., 2009 and other publications

• New methods based on neural networks:
  • Collobert et al., 2011
  • FitzGerald et al., 2015
  • Roth and Lapata, 2016
  • Foland W. and Martin J., 2015
  • Marcheggiani et al., 2017
  • Zhou and Xu, 2015 (end-to end)
  • Swayamdipta et al., 2016
Semantic role labeling for Russian

• Rule-based semantic parsers:
  • AOT.ru (Sokirko, A. 2001)
  • The parser of ISA FRC CSC RAS (Shelmanov and Smirnov, 2014)
  • etc.

• Known corpora annotated with semantic roles:
  • The corpus from ISA FRC CSC RAS
    • Shelmanov and Smirnov, 2014
  • FrameBank
    • Lyashevskaya, 2012
    • Lyashevskaya and Kashkin, 2015

• Data-driven semantic role labelers:
  • SVM-based parser + feature engineering (Kuznetsov I., 2015) trained on pre-release version of FrameBank
  • The parser of ISA FRC CSC RAS (Shelmanov and Smirnov, 2014) – bootstrapping based on automatic annotation of SynTagRus using rule-based semantic semantic parser
FrameBank
https://github.com/olesar/framebank

- Semantic role hierarchy:
  - The FrameBank provides fine and coarse grained roles
  - It also provides generality relations between roles

- Lexicon of predicate frames:
  - Describes predicates frames and their roles in terms of morphological, syntactic, semantic, and other features

- Annotated text samples in Russian
  - Partially annotated text samples with predicates, arguments, and semantic roles (core and non-core)
  - In addition: morphology features, lemmas, sentences, etc.

- Statistics:
  - ~ 800 unique predicates (verbs)
  - ~ 70 unique roles
  - ~ 60 000 extractable arguments
Features for semantic role labeling

• Neural networks allow to use atomic features

• Categorical features:
  • Various types of morphological features of both an argument and a predicate: part of speech, grammar case, animacy, verb form, time, passiveness, and others ("morph")
  • Relative position of an argument in a sentence with respect to a predicate ("rel_pos")
  • Predicate lemma ("pred_lemma")
  • Preposition of an argument extracted from a syntax tree ("arg_prep")
  • Name of a syntax link from an argument to its parent in a syntax tree ("synt_link")

• Embeddings:
  • Embedding of an argument lemma ("arg_emdeddings")
  • Embedding of a predicate lemma ("pred_emdeddings")
A) Simple model

B) Complex model
Training model for “unknown” predicates

• Cannot use predicate lemma since it is not available in the training corpus for “unknown” predicates

• Instead of predicate lemma we are using word embeddings built using word2vec
  • Embeddings encode semantic similarities of words in a low dimensional vector space
    • Embeddings can encode similarities between predicate frames

• We used RusVectores 2.0 models (Kutuzov and Andreev, 2015):
  • Trained on Russian national corpus
  • 300 dimensions

• Used early stopping during training of neural networks
Data preparation

• Original FrameBank does not provide explicit correspondence between text offsets and SRL annotations
  • We created the automatic tool for mapping predicates and arguments with core roles to text tokens
    http://nlp.isa.ru/framebank_parser

• Parsed with Google’s SyntaxNet
  • For parsing we used SyntaxNet (McParsey model for Russian) (Andor D., 2016)
  • We prepared dockerized version of SyntaxNet for Russian (and other languages) + Python wrapper
    • https://github.com/IINemo/docker-syntaxnet_rus
    bash$> echo "мама мыла раму" | docker run --rm -i
    inemo/syntaxnet_rus
    • https://github.com/IINemo/syntaxnet_wrapper
  • The syntax structure corresponds to well-known Universal dependencies format
Example of FrameBank parsing via SyntaxNet


Красные сгустки и комки марли.

А Пелагея Ивановна уже встряхивает младенца и похлопывает его.

Аксинья гремит вёдрами, наливая в тазы воду.

Младенца погружают то в холодную, то в горячую воду.
Experiment setup for training of a model for “known predicates”

• Selected the subcorpus by keeping only predicates with > 10 examples
  • 572 predicates left
• Filtered infrequent semantic roles and fixed erroneous role labels
• Final version of experimental dataset:
  • 53,151 examples
  • 44 different semantic roles
• RusVectores problem:
  • Large portion of predicates (verbs) are not covered
  • 17,000 examples in our dataset have zero predicate embeddings
• Baseline: most frequent role in the corpus
• Five-fold cross-validation
## Results of neural-network models for “known” predicates

<table>
<thead>
<tr>
<th>Model + feature set</th>
<th>Macro $F_1$-score,%</th>
<th>Micro $F_1$-score,%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.5 ± 0.0</td>
<td>11.6 ± 0.2</td>
</tr>
<tr>
<td>Simple + morph</td>
<td>22.8 ± 0.6</td>
<td>35.4 ± 0.3</td>
</tr>
<tr>
<td>Simple + morph + pred_lemma</td>
<td>71.2 ± 0.6</td>
<td>76.1 ± 0.5</td>
</tr>
<tr>
<td>Simple + morph + pred_embeddings</td>
<td>62.0 ± 0.4</td>
<td>65.2 ± 0.3</td>
</tr>
<tr>
<td>Simple + morph + pred_lemma + arg_prep</td>
<td>75.9 ± 0.4</td>
<td>79.2 ± 0.2</td>
</tr>
<tr>
<td>Simple + morph + pred_lemma + arg_prep + synt_link</td>
<td>76.8 ± 0.5</td>
<td>80.3 ± 0.3</td>
</tr>
<tr>
<td>Simple + morph + pred_lemma + arg_prep + synt_link + arg_embeddings + pred_embeddings</td>
<td>78.6 ± 0.4</td>
<td>81.8 ± 0.2</td>
</tr>
<tr>
<td>Complex + morph + synt + pred_lemma + arg_embeddings + pred_embeddings</td>
<td>79.2 ± 0.3</td>
<td>82.3 ± 0.2</td>
</tr>
</tbody>
</table>
Results of non-neural-network models for “known” predicates

- Used the most complete feature set
- Tuned parameters using greedy strategy

<table>
<thead>
<tr>
<th>Model</th>
<th>Macro F₁-score,%</th>
<th>Micro F₁-score,%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LinearSVC</td>
<td>74.3 ± 0.2</td>
<td>77.6 ± 0.1</td>
</tr>
<tr>
<td>LogReg</td>
<td>75.1 ± 0.1</td>
<td>78.2 ± 0.3</td>
</tr>
<tr>
<td>LightGBM</td>
<td>71.3 ± 0.4</td>
<td>76.0 ± 0.1</td>
</tr>
<tr>
<td>Random Forest</td>
<td>69.7 ± 0.4</td>
<td>71.9 ± 0.1</td>
</tr>
<tr>
<td>+Top neural network</td>
<td>79.2 ± 0.3</td>
<td>82.3 ± 0.2</td>
</tr>
</tbody>
</table>
Example of SRL parsing

- [http://nlp.isa.ru/brat_framebank](http://nlp.isa.ru/brat_framebank)

В 1992 году «Фонд Караваева» заключил договор с долгопрудненским Заводом тонкого органического синтеза (ТОС) на производство препаратов и арендовал помещение под офис — возможность не приобретать, а арендовать как аппаратную платформу для клинической ИС, так и саму систему.

-- Федеральный закон запрещает продажу средств производства и торговых площадей тем, кто в течение многих лет честно арендовал их у государства.

Сейф он не завел, но почтовый ящик арендовал.
Experiments for “unknown” predicates

• The sets of **predicates** for training and testing do not intersect

• Performed evaluations for two different split methods:
  • The good split:
    • The test set contains highly similar predicates to the ones in the training set (by cosine similarity)
    • Easy for the models to restore semantic frame for “unknown” predicate
    • Training set: 49,709 examples
    • Testing set: 27 predicates; 3,442 examples
  • The bad split:
    • Test set contains predicates that are least similar to any of the “known” predicates
    • Hard for the models to restore semantic frame for “unknown” predicate
    • Training set: 50,093 examples
    • Testing set: 21 predicates; 3,058 examples

• Averaged across several fits using different random states
Results of models for “unknown” predicates

- Results for the “good” split

<table>
<thead>
<tr>
<th>Model + feature set</th>
<th>Macro $F_1$-score,%</th>
<th>Micro $F_1$-score,%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.4</td>
<td>9.6</td>
</tr>
<tr>
<td>Simple (only categ. Feats)</td>
<td>13.7 ± 0.4</td>
<td>24.6 ± 0.3</td>
</tr>
<tr>
<td>Complex + arg_embeddings</td>
<td>19.4 ± 0.3</td>
<td>31.9 ± 0.5</td>
</tr>
<tr>
<td>Complex + arg_pred_embeddings</td>
<td>41.4 ± 0.7</td>
<td>66.7 ± 1.1</td>
</tr>
</tbody>
</table>

- Results for the “bad” split

<table>
<thead>
<tr>
<th>Model + feature set</th>
<th>Macro $F_1$-score,%</th>
<th>Micro $F_1$-score,%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.7</td>
<td>13.2</td>
</tr>
<tr>
<td>Simple (only categ. Feats)</td>
<td>9.1 ± 0.2</td>
<td>24.8 ± 0.5</td>
</tr>
<tr>
<td>Complex + arg_embeddings</td>
<td>14.5 ± 0.7</td>
<td>27.2 ± 0.1</td>
</tr>
<tr>
<td>Complex + arg_pred_embeddings</td>
<td>24.1 ± 1.5</td>
<td>41.4 ± 2.2</td>
</tr>
</tbody>
</table>
Created electronical resources

• Preprocessed FrameBank + scripts + models:
  • [http://nlp.isa.ru/framebank_parser](http://nlp.isa.ru/framebank_parser)

• Original FrameBank + results of SRL and syntax parsing visualized via Brat tool (Stenetorp P. et al., 2012):
  • [http://nlp.isa.ru/brat_framebank](http://nlp.isa.ru/brat_framebank)

• Dockerized version of SyntaxNet for Russian:
  • `bash$> echo "мама мыла раму" | docker run --rm -i inemo/syntaxnet_rus`
  • [https://github.com/IINemo/docker-syntaxnet_rus](https://github.com/IINemo/docker-syntaxnet_rus)

• Python wrapper for SyntaxNet:
  • [https://github.com/IINemo/syntaxnet_wrapper](https://github.com/IINemo/syntaxnet_wrapper)

• Dockerized version of SyntaxNet for English:
  • `bash$> echo "Beware of a silent dog and still water" | docker run --rm -i inemo/syntaxnet_eng`

• Docker containers with SyntaxNet for other languages:
  • [https://github.com/IINemo/docker-syntaxnet](https://github.com/IINemo/docker-syntaxnet)
Conclusion and future work

• Results:
  • Presented the neural network models for semantic role labeling of Russian texts
  • Investigated the method for training a labeler for arguments of “unknown” predicates using word embeddings
  • Created benchmark based on FrameBank corpus for evaluation of parsers for SRL
  • The models and the benchmark are openly available

• Future work:
  • Adding global inference step for the SRL parser
  • Methods of merging several semantic parsers and their annotations for creating a corpus with a higher annotation recall
  • New methods base on semi-supervised learning

• Acknowledgments
  • The project is supported by the Russian Foundation for Basic Research, project number: 16-37-00425 “mol_a”.
References (1)


References (2)


References (3)


