Coreference Resolution for Russian: The Impact of Semantic Features

Svetlana Toldova\textsuperscript{1}  
toldova@yandex.ru

Max Ionov\textsuperscript{2}  
max.ionov@gmail.com

\textsuperscript{1} National Research University Higher School of Economics  
\textsuperscript{2} Moscow State University, Goethe University Frankfurt

Dialogue 2017  
RSUH, 31.05.2017
Table of Contents

1. Coreference resolution
2. Setting the baseline
3. Semantic information
Coreference resolution:

Clustering noun phrases that refer to the same entity.

An important task for many high-level NLP tasks:
- Machine translation
- Discourse parsing
- Summarization

S. Toldova, M. Ionov
Coreference Resolution for Russian: Semantic Features
Coreference resolution:

- Clustering noun phrases that refer to the same entity
Coreference resolution:

- Clustering noun phrases that refer to the same entity
- An important task for many high-level NLP tasks:
Coreference resolution:

- Clustering noun phrases that refer to the same entity
- An important task for many high-level NLP tasks:
  - Machine translation
  - Discourse parsing
  - Summarization
  - ...
(1) Но дача была так расположена, что откуда бы я ни заходил, я мог видеть только небольшой угол двора. Он был так же пуст и невозделан, как и окружающая местность.

‘But the summer house was located in such a way that no matter from where I went, I could only see a small corner of a courtyard. It was as empty and uncultivated as the surroundings.’
But the summer house was located in such a way that no matter from where I went, I could only see a small corner of a courtyard. It was as empty and uncultivated as the surroundings.

‘But the summer house was located in such a way that no matter from where I went, I could only see a small corner of a courtyard. It was as empty and uncultivated as the surroundings.’
‘But the summer house was located in such a way that no matter from where I went, I could only see a small corner of a courtyard. It was as empty and uncultivated as the surroundings.’
‘But the summer house was located in such a way that no matter from where I went, I could only see a small corner of a courtyard. It was as empty and uncultivated as the surroundings.’

S. Toldova, M. Ionov
A lot of theoretical research on reference
A lot of theoretical research on reference

A lot of research on coreference resolution for English
A lot of theoretical research on reference

A lot of research on coreference resolution for English

Some research on coreference resolution for other languages
Coreference and coreference resolution

- A lot of theoretical research on reference
- A lot of research on coreference resolution for English
- Some research on coreference resolution for other languages — increasing as we speak
Coreference and coreference resolution

- A lot of theoretical research on reference
- A lot of research on coreference resolution for English
- *Some* research on coreference resolution for other languages — increasing as we speak
- A shared task on anaphora and coreference resolution for Russian in 2014:
A lot of theoretical research on reference
A lot of research on coreference resolution for English
*Some* research on coreference resolution for other languages — increasing as we speak
A shared task on anaphora and coreference resolution for Russian in 2014:
  - 3 teams participated in the coreference resolution track
Coreference and coreference resolution

- A lot of theoretical research on reference
- A lot of research on coreference resolution for English
- *Some* research on coreference resolution for other languages — increasing as we speak
- A shared task on anaphora and coreference resolution for Russian in 2014:
  - 3 teams participated in the coreference resolution track
  - None of them submitted a paper with a system description
A lot of theoretical research on reference

A lot of research on coreference resolution for English

Some research on coreference resolution for other languages — increasing as we speak

A shared task on anaphora and coreference resolution for Russian in 2014:

- 3 teams participated in the coreference resolution track
- None of them submitted a paper with a system description
- No open coreference resolution system trained for Russian available for research
Coreference resolution
Setting the baseline
Semantic information

Data: RuCor

- RuCor — Russian coreference corpus

S. Toldova, M. Ionov
Coreference Resolution for Russian: Semantic Features
Data: RuCor

- RuCor — Russian coreference corpus
- Introduced with a shared task in 2014
Data: RuCor

- RuCor — Russian coreference corpus
- Introduced with a shared task in 2014
- Published under an open license
Data: RuCor

- RuCor — Russian coreference corpus
- Introduced with a shared task in 2014
- Published under an open license
- Corpus size:
  - 180 texts
  - 3638 chains
  - 16557 noun phrases
- Automatically processed:
  - Sentence splitting, tokenization
  - Morphological annotation (partially fixed manually)
  - Dependency parsing
Data: RuCor

- RuCor — Russian coreference corpus
- Introduced with a shared task in 2014
- Published under an open license
- Corpus size:
  - 180 texts
  - 3,638 chains
  - 16,557 noun phrases
Data: RuCor

- RuCor — Russian coreference corpus
- Introduced with a shared task in 2014
- Published under an open license
- Corpus size:
  - 180 texts
  - 3,638 chains
  - 16,557 noun phrases
- Automatically processed:
  - Sentence splitting, tokenization
  - Morphological annotation (partially fixed manually)
  - Dependency parsing
Data: RuCor

- RuCor — Russian coreference corpus
- Introduced with a shared task in 2014
- Published under an open license
- Corpus size:
  - 180 texts
  - 3,638 chains
  - 16,557 noun phrases
- Automatically processed:
  - Sentence splitting, tokenization
  - Morphological annotation
Data: RuCor

- RuCor — Russian coreference corpus
-Introduced with a shared task in 2014
- Published under an open license

Corpus size:
- 180 texts
- 3,638 chains
- 16,557 noun phrases

Automatically processed:
- Sentence splitting, tokenization
- Morphological annotation (partially fixed manually)
- Dependency parsing
RuCor annotation guidelines

- Based on MUC-6 scheme
RuCor annotation guidelines

- Based on MUC-6 scheme
- Only real-world entities
RuCor annotation guidelines

- Based on MUC-6 scheme
- Only real-world entities (no abstract nouns or generic expressions)
RuCor annotation guidelines

- Based on MUC-6 scheme
- Only real-world entities (no abstract nouns or generic expressions)
- Only identity relations

S. Toldova, M. Ionov
Coreference Resolution for Russian: Semantic Features
RuCor annotation guidelines

- Based on MUC-6 scheme
- Only real-world entities (no abstract nouns or generic expressions)
- Only identity relations
- No singleton annotation
RuCor annotation guidelines

- Based on MUC-6 scheme
- Only real-world entities (no abstract nouns or generic expressions)
- Only identity relations
- No singleton annotation — NP is annotated only if it is a part of a coreference chain
Noun phrases are generated from syntactic annotations
Experiments design

- Noun phrases are generated from syntactic annotations
- Evaluation is performed using CoNLL reference coreference scorers
Experiments design

- Noun phrases are generated from syntactic annotations
- Evaluation is performed using CoNLL reference coreference scorers
- Based on exact matches of NPs
Experiments design

- Noun phrases are generated from syntactic annotations
- Evaluation is performed using CoNLL reference coreference scorers
- Based on exact matches of NPs
- Two evaluation strategies: gold mentions and gold boundaries:
Noun phrases are generated from syntactic annotations

Evaluation is performed using CoNLL reference coreference scorers

Based on exact matches of NPs

Two evaluation strategies: gold mentions and gold boundaries:

- Gold mentions: a set of NPs are taken from GS, coreference relations are predicted between them
- Gold boundaries: all NPs are considered, boundaries of the NPs are taken from GS
Experiments design

- Noun phrases are generated from syntactic annotations
- Evaluation is performed using CoNLL reference coreference scorers
- Based on exact matches of NPs
- Two evaluation strategies: *gold mentions* and *gold boundaries*:
  - **GOLD MENTIONS**: a set of NPs are taken from GS, coreference relations are predicted between them
  - **GOLD BOUNDARIES**: all NPs are considered, boundaries of the NPs are taken from GS
- Two coreference scores: MUC and B³
As it is often the case, some of the decisions about data modeling and annotation could be done differently:
Experiments design: Linguistics vs. NLP

- As it is often the case, some of the decisions about data modeling and annotation could be done differently:
  - What to annotate
  - How to annotate
As it is often the case, some of the decisions about data modeling and annotation could be done differently:

- What to annotate
- How to annotate

⇒ The task is not to resolve coreference *in general*...
As it is often the case, some of the decisions about data modeling and annotation could be done differently:

- What to annotate
- How to annotate

⇒ The task is not to resolve coreference *in general*... but to predict coreference relations according to the RuCor annotation guidelines.
Mention-pair model — the simplest model for coreference resolution
Mention-pair model

- Mention-pair model — the simplest model for coreference resolution
- For each NP there is a set of NPs — possible antecedents
Mention-pair model — the simplest model for coreference resolution

For each NP there is a set of NPs — possible antecedents

For each such pair we can predict if they are coreferent
Mention-pair model — the simplest model for coreference resolution
For each NP there is a set of NPs — possible antecedents
For each such pair we can predict if they are coreferent
After all the decisions are made, positive pairs are grouped together
A few simple rule-based baselines:

- **StrMatch**: two NPs corefer if their lemmas are the same (only for nouns and deictic pronouns).
- **StrMatchPro**: StrMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
- **HeadMatch**: two NPs corefer if their heads are the same (only for nouns and deictic pronouns).
- **HeadMatchPro**: HeadMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
Rule-based baselines

A few simple rule-based baselines:

- **StrMatch**: two NPs corefer if their lemmas are the same (only for nouns and deictic pronouns).

  **StrMatchPro**: StrMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.

- **HeadMatch**: two NPs corefer if their heads are the same (only for nouns and deictic pronouns).

  **HeadMatchPro**: HeadMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
Rule-based baselines:

- **StrMatch**: two NPs corefer if their lemmas are the same (only for nouns and deictic pronouns).
- **StrMatchPro**: StrMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
A few simple rule-based baselines:

- **StrMatch**: two NPs corefer if their lemmas are the same (only for nouns and deictic pronouns).
- **StrMatchPro**: StrMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
- **HeadMatch**: two NPs corefer if their heads are the same (only for nouns and deictic pronouns).
A few simple rule-based baselines:

- **StrMatch**: two NPs corefer if their lemmas are the same (only for nouns and deictic pronouns).
- **StrMatchPro**: StrMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
- **HeadMatch**: two NPs corefer if their heads are the same (only for nouns and deictic pronouns).
- **HeadMatchPro**: HeadMatch + non-deictic pronouns are paired with the nearest NP that agrees in gender and number.
### Rule-based baselines

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th>B³</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F₁</td>
<td>P</td>
<td>R</td>
<td>F₁</td>
</tr>
<tr>
<td>StrMatch</td>
<td>94.29</td>
<td>37.36</td>
<td>53.52</td>
<td>97.09</td>
<td>38.19</td>
<td>54.82</td>
</tr>
<tr>
<td>StrMatchPro</td>
<td>84.90</td>
<td>52.42</td>
<td>64.82</td>
<td>89.34</td>
<td>43.35</td>
<td>58.37</td>
</tr>
<tr>
<td>HeadMatch</td>
<td>87.78</td>
<td>47.06</td>
<td>61.27</td>
<td>92.11</td>
<td>43.64</td>
<td>59.22</td>
</tr>
<tr>
<td>HeadMatchPro</td>
<td>84.89</td>
<td>52.50</td>
<td>64.87</td>
<td>89.29</td>
<td>43.38</td>
<td>58.40</td>
</tr>
</tbody>
</table>

**Table 1:** Rule-based coreference systems, gold mentions
## Rule-based baselines

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>StrMatch</td>
<td>94.29</td>
<td>37.36</td>
</tr>
<tr>
<td>StrMatchPro</td>
<td>84.90</td>
<td>52.42</td>
</tr>
<tr>
<td>HeadMatch</td>
<td>87.78</td>
<td>47.06</td>
</tr>
<tr>
<td>HeadMatchPro</td>
<td>84.89</td>
<td>52.50</td>
</tr>
</tbody>
</table>

**Table 1:** Rule-based coreference systems, gold mentions

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>StrMatch</td>
<td>52.86</td>
<td>32.29</td>
</tr>
<tr>
<td>StrMatchPro</td>
<td>34.40</td>
<td>45.46</td>
</tr>
<tr>
<td>HeadMatch</td>
<td>35.26</td>
<td>41.38</td>
</tr>
<tr>
<td>HeadMatchPro</td>
<td>34.40</td>
<td>45.49</td>
</tr>
</tbody>
</table>

**Table 2:** Gold boundaries, mention detection f-score 51.38
Baseline ML models

- Basic set of features
Two ML models:

- Basic set of features
- Extended feature set:
Baseline ML models

- Basic set of features
- Extended feature set:
  - Distance features
  - Morphological features
  - Lexical features
  - Syntactical features
Basic feature set

- The distance between an anaphoric NP and a candidate antecedent is 1 sentence.
- Both NPs are not pronouns and after removing any demonstratives they match.
- NPs agree in animacy and if they are not pronouns their syntactic heads match.
- Anaphoric NP is a pronoun.
- Candidate antecedent is a pronoun.
- Both NPs are pronouns.
- NPs agree in gender.
- NPs agree in number.
- Both NPs are proper.
- An anaphoric NP is a demonstrative.
- NPs are in the appositive relation.

S. Toldova, M. Ionov
Coreference Resolution for Russian: Semantic Features
Extended features

- Distance features
  - Number of nouns between two NPs
  - Morphological features: NPs are pronouns of a specific type
  - Lexical features: Modifiers: one of the NPs equals to a noun modifier in another NP, Acronyms: one NP is an acronym of another
  - Syntactical features: NPs are subjects, NPs are objects, Syntactic parallelism: both NPs are in the beginning of sentences and they are both subjects
Extended features

- Distance features: number of nouns between two NPs
Extended features

- Distance features: number of nouns between two NPs
- Morphological features
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features:
  - Modifiers: one of the NPs equals to a noun modifier in another NP
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features:
  - Modifiers: one of the NPs equals to a noun modifier in another NP
  - Acronyms: one NP is an acronym of another
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features:
  - Modifiers: one of the NPs equals to a noun modifier in another NP
  - Acronyms: one NP is an acronym of another
- Syntactical features
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features:
  - Modifiers: one of the NPs equals to a noun modifier in another NP
  - Acronyms: one NP is an acronym of another
- Syntactical features:
  - NPs are subjects
  - NPs are objects
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features:
  - Modifiers: one of the NPs equals to a noun modifier in another NP
  - Acronyms: one NP is an acronym of another
- Syntactical features:
  - NPs are subjects
  - NPs are objects
  - Syntactic parallelism:
Extended features

- Distance features: number of nouns between two NPs
- Morphological features: NPs are pronouns of a specific type
- Lexical features:
  - Modifiers: one of the NPs equals to a noun modifier in another NP
  - Acronyms: one NP is an acronym of another
- Syntactical features:
  - NPs are subjects
  - NPs are objects
  - Syntactic parallelism: both NPs are in the beginning of sentences and they are both subjects
Extended features: examples

- Modifiers:

  (2) a. президент Обама ‘president Obama’ — президент ‘president’
Extended features: examples

- Modifiers:

(2) a. президент Обама ‘president Obama’ — президент ‘president’

NB классом этого друга ‘class of this friend’ — собственный класс мальчика ‘boy’s own class’
Extended features: examples

Modifiers:

(2) a. президент Обама ‘president Obama’ — президент ‘president’

NB классом этого друга ‘class of this friend’ — собственный класс мальчика ‘boy’s own class’

head comparison is not the same and leads to errors
Extended features: examples

**Modifiers:**

(2) а. президент Обама ‘president Obama’ — президент ‘president’

   NB классом этого друга ‘class of this friend’ — собственный класс мальчика ‘boy’s own class’

   Comparison is not the same and leads to errors

**Acronyms:**

(3) РФ ‘RF’ — Российская Федерация ‘Russian Federation’
Extended features: examples

- **Modifiers:**
  
  \[(2)\] а. президент Обама ‘president Obama’ — президент ‘president’

  NB классом этого друга ‘class of this friend’ — собственный класс мальчика ‘boy’s own class’

  comparison is not the same and leads to errors

- **Acronyms:**

  \[(3)\] РФ ‘RF’ — Российская Федерация ‘Russian Federation’ but not Россия ‘Russia’
## Baseline ML models: results

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>HeadMatchPro</td>
<td>84.89</td>
<td>52.50</td>
</tr>
<tr>
<td>MLMentionPair</td>
<td>73.98</td>
<td>62.24</td>
</tr>
<tr>
<td>MLUpdated</td>
<td>79.29</td>
<td>63.01</td>
</tr>
</tbody>
</table>

Table 3: ML-based coreference systems, gold mentions
### Baseline ML models: results

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th></th>
<th>B³</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F₁</td>
<td>P</td>
<td>R</td>
<td>F₁</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HeadMatchPro</strong></td>
<td>84.89</td>
<td>52.50</td>
<td>64.87</td>
<td>89.29</td>
<td>43.38</td>
<td>58.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLMentionPair</strong></td>
<td>73.98</td>
<td>62.24</td>
<td>67.61</td>
<td>71.40</td>
<td>49.34</td>
<td>58.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLUpdated</strong></td>
<td>79.29</td>
<td>63.01</td>
<td>70.22</td>
<td>79.42</td>
<td>48.39</td>
<td>60.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3:** ML-based coreference systems, gold mentions

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th></th>
<th>B³</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F₁</td>
<td>P</td>
<td>R</td>
<td>F₁</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HeadMatchPro</strong></td>
<td>34.40</td>
<td>45.49</td>
<td>39.18</td>
<td>26.89</td>
<td>39.58</td>
<td>32.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLMentionPair</strong></td>
<td>37.91</td>
<td>55.85</td>
<td>45.16</td>
<td>21.88</td>
<td>43.98</td>
<td>29.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MLUpdated</strong></td>
<td>37.94</td>
<td>53.87</td>
<td>44.52</td>
<td>25.00</td>
<td>42.61</td>
<td>31.51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4:** Gold boundaries, mention detection f-score 51.21
Table of Contents

1. Coreference resolution
2. Setting the baseline
3. Semantic information
A lot of cases is impossible to resolve without semantic information:

- Synonymy
- Hyponymy / hyperonymy

We test the impact of 3 ways to include semantic information:

- A list of named entities with their synonyms
- A word2vec model to check if two NPs are similar
- A thesaurus to check if two NPs are related
Coreference resolution
Setting the baseline
Semantic information

Semantic information: sources

- A lot of cases is impossible to resolve without semantic information:
  - Synonymy
  - Hyponymy / hyperonymy
Semantic information: sources

- A lot of cases is impossible to resolve without semantic information:
  - Synonymy
  - Hyponymy / hyperonymy

(4) профессор Вагнер ‘professor Vagner’ — необычайного человека ‘an extraordinary man’
A lot of cases is impossible to resolve without semantic information:

- Synonymy
- Hyponymy / hyperonymy

(4) профессор Вагнер ‘professor Vagner’ — необычайного человека ‘an extraordinary man’

We test the impact of 3 ways to include semantic information:
A lot of cases is impossible to resolve without semantic information:

- Synonymy
- Hyponymy / hyperonymy

(4) профессор Вагнер ‘professor Vagner’ — необычайного человека ‘an extraordinary man’

We test the impact of 3 ways to include semantic information:

- A list of named entities with their synonyms
A lot of cases is impossible to resolve without semantic information:

- Synonymy
- Hyponymy / hyperonymy

We test the impact of 3 ways to include semantic information:

- A list of named entities with their synonyms
- A word2vec model
A lot of cases is impossible to resolve without semantic information:
- Synonymy
- Hyponymy / hyperonymy

(4) профессор Вагнер ‘professor Vagner’ — необычайного человека ‘an extraordinary man’

We test the impact of 3 ways to include semantic information:
- A list of named entities with their synonyms
- A word2vec model to check if two NPs are similar
Semantic information: sources

- A lot of cases is impossible to resolve without semantic information:
  - Synonymy
  - Hyponymy / hyperonymy

(4) профессор Вагнер ‘professor Vagner’ — необычайного человека ‘an extraordinary man’

- We test the impact of 3 ways to include semantic information:
  - A list of named entities with their synonyms
  - A word2vec model to check if two NPs are similar
  - A thesaurus

S. Toldova, M. Ionov  Coreference Resolution for Russian: Semantic Features
A lot of cases is impossible to resolve without semantic information:

- Synonymy
- Hyponymy / hyperonymy

We test the impact of 3 ways to include semantic information:

- A list of named entities with their synonyms
- A word2vec model to check if two NPs are similar
- A thesaurus to check if two NPs are related
2 lists:

- A tiny list of frequent NEs from the corpus
- A large list of geographical names from GeoNames

The lists are used to check for synonyms and to check if the NE class is the same for both NPs. Both improve the recall (and F-measure as a result). Using a proper NER should improve further.
Named entities

- 2 lists:
  - A tiny list of frequent NEs from the corpus

S. Toldova, M. Ionov
Coreference Resolution for Russian: Semantic Features
2 lists:

- A tiny list of frequent NEs from the corpus
- A large list of geographical names
2 lists:
- A tiny list of frequent NEs from the corpus
- A large list of geographical names from GeoNames

Both lists improve the recall (and F-measure as a result) using a proper NER should improve further.
2 lists:

- A tiny list of frequent NEs from the corpus
- A large list of geographical names

The lists are used to check for synonyms and to check if the NE class is the same for both NPs
2 lists:
- A tiny list of frequent NEs from the corpus
- A large list of geographical names

The lists are used to check for synonyms and to check if the NE class is the same for both NPs

Both improve the recall
2 lists:
- A tiny list of frequent NEs from the corpus
- A large list of geographical names

The lists are used to check for synonyms and to check if the NE class is the same for both NPs

Both improve the recall (and F-measure as a result)
2 lists:
- A tiny list of frequent NEs from the corpus
- A large list of geographical names

The lists are used to check for synonyms and to check if the NE class is the same for both NPs

Both improve the recall (and F-measure as a result)

Using a proper NER should improve further
2 lists:
- A tiny list of frequent NEs from the corpus
- A large list of geographical names

The lists are used to check for synonyms and to check if the NE class is the same for both NPs

Both improve the recall (and F-measure as a result)

Using a proper NER should improve further — too few hits for items in the lists
Word2vec

“RusVectores” word2vec model

Used to look up the similarity of the heads of both NPs

Improved the results slightly increasing the recall

A threshold is used to determine if two words are similar

High threshold: very few cases

Low threshold: a lot of false positives

co-hyponyms

(5) a. муж 'husband' /emdash.cyr супруг 'spouse'
b. муж 'husband' /emdash.cyr жена 'wife'

lower threshold

Most of the results are also covered by RuThes

But: should help for the non-standard vocabulary
“RusVectores” word2vec model

Used to look up the similarity of the heads of both NPs

A threshold is used to determine if two words are similar

High threshold: very few cases

Low threshold: a lot of false positives

Co-hyponyms:

a. муж 'husband' /супруг 'spouse'

b. муж 'husband' /жена 'wife'

Most of the results are also covered by RuThes

But: should help for the non-standard vocabulary
Word2vec

- “RusVectores” word2vec model
- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
Word2vec

- “RusVectores” word2vec model
- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
Word2vec

- “RusVectores” word2vec model
- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
- High threshold: very few cases
Word2vec

- “RusVectores” word2vec model
- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
- High threshold: very few cases
- Low threshold: a lot of false positives
Word2vec

- “RusVectores” word2vec model
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
- High threshold: very few cases
- Low threshold: a lot of false positives

co-hyponyms

c. муж 'husband' /emdash.cyr супруг 'spouse'

b. муж 'husband' /emdash.cyr жена 'wife'

Most of the results are also covered by RuThes

But: should help for the non-standard vocabulary
“RusVectores” word2vec model

- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
- High threshold: very few cases
- Low threshold: a lot of false positives

(5) a. муж ‘husband’ — супруг ‘spouse’

co-hyponyms

S. Toldova, M. Ionov
“RusVectores” word2vec model

- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
  - High threshold: very few cases
  - Low threshold: a lot of false positives

(5) a. муж ‘husband’ — супруг ‘spouse’
    b. муж ‘husband’ — жена ‘wife’
“RusVectores” word2vec model
Used to look up the similarity of the heads of both NPs
Improved the results slightly increasing the recall
A threshold is used to determine if two words are similar
High threshold: very few cases
Low threshold: a lot of false positives

(5) a. муж ‘husband’ — супруг ‘spouse’
   b. муж ‘husband’ — жена ‘wife’
Word2vec

- “RusVectores” word2vec model
- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
  - High threshold: very few cases
  - Low threshold: a lot of false positives
- Most of the results are also covered by RuThes


- “RusVectores” word2vec model
- Used to look up the similarity of the heads of both NPs
- Improved the results slightly increasing the recall
- A threshold is used to determine if two words are similar
  - High threshold: very few cases
  - Low threshold: a lot of false positives

(5) a. муж ‘husband’ — супруг ‘spouse’
   b. муж ‘husband’ — жена ‘wife’

- Most of the results are also covered by RuThes
- But: should help for the non-standard vocabulary
Thesaurus RuThes-Lite

RuThes-Lite

- Used to look up two NPs or their heads:
  - If their domains are the same, they are from the same class
  - If there is a path from one to another using a parent relation, they are considered aliases

Improved the results slightly increasing the recall

6. a. работа 'work' / труд 'labor'
   b. лицо 'face / person' / человек 'man'

S. Toldova, M. Ionov

Coreference Resolution for Russian: Semantic Features
RuThes-Lite

- Thesaurus RuThes-Lite
- Used to look up two NPs or their heads:
  - If their domains are the same, they are from the same class
  - If there is a path from one to another using a parent relation, they are considered aliases

- Improved the results slightly increasing the recall

- a. работа 'work' /емdash.cyr труд 'labor'
- b. лицо 'face / person' /емdash.cyr человек 'man'
- homonymy

S. Toldova, M. Ionov
RuThes-Lite

- Thesaurus RuThes-Lite
- Used to look up two NPs or their heads:
  - If their domains are the same, they are from the same class
  - If there is a path from one to another using a parent relation, they are considered aliases
- Improved the results slightly increasing the recall
RuThes-Lite

- Thesaurus RuThes-Lite
- Used to look up two NPs or their heads:
  - If their domains are the same, they are from the same class
  - If there is a path from one to another using a parent relation, they are considered aliases
- Improved the results slightly increasing the recall

(6) a. работа ‘work’ — труд ‘labor’
RuThes-Lite

- Thesaurus RuThes-Lite
  - Used to look up two NPs or their heads:
    - If their domains are the same, they are from the same class
    - If there is a path from one to another using a parent relation, they are considered aliases
  - Improved the results slightly increasing the recall
    1. работа ‘work’ — труд ‘labor’
    2. лицо ‘face / person’ — человек ‘man’
RuThes-Lite

- Thesaurus RuThes-Lite
- Used to look up two NPs or their heads:
  - If their domains are the same, they are from the same class
  - If there is a path from one to another using a parent relation, they are considered aliases
- Improved the results slightly increasing the recall

(b)  a. работа ‘work’ — труд ‘labor’
    b. лицо ‘face / person’ — человек ‘man’
### Results

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th></th>
<th></th>
<th>B³</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
<td>F₁</td>
<td>P</td>
<td>R</td>
<td>F₁</td>
</tr>
<tr>
<td>MLMentionPair</td>
<td>73.98</td>
<td>62.24</td>
<td>67.61</td>
<td>71.40</td>
<td>49.34</td>
<td>58.36</td>
</tr>
<tr>
<td>MLUpdated</td>
<td>79.35</td>
<td>63.44</td>
<td>70.51</td>
<td>79.37</td>
<td>48.60</td>
<td>60.29</td>
</tr>
<tr>
<td>NamedEntities</td>
<td>79.43</td>
<td>63.72</td>
<td>70.71</td>
<td>79.37</td>
<td>48.86</td>
<td>60.48</td>
</tr>
<tr>
<td>word2vec</td>
<td>79.29</td>
<td>63.49</td>
<td>70.52</td>
<td>79.25</td>
<td>48.64</td>
<td>60.28</td>
</tr>
<tr>
<td>RuThes</td>
<td>79.19</td>
<td>63.79</td>
<td>70.66</td>
<td>78.92</td>
<td>48.78</td>
<td>60.29</td>
</tr>
<tr>
<td>All</td>
<td>79.19</td>
<td>63.97</td>
<td>70.77</td>
<td>78.85</td>
<td>48.94</td>
<td>60.39</td>
</tr>
</tbody>
</table>

**Table 5:** The impact of semantic information, gold mentions
### Results: gold boundaries

<table>
<thead>
<tr>
<th></th>
<th>MUC</th>
<th>B³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td><strong>MLMentionPair</strong></td>
<td>37.91</td>
<td>55.85</td>
</tr>
<tr>
<td><strong>MLUpdated</strong></td>
<td>37.94</td>
<td>53.87</td>
</tr>
<tr>
<td><strong>NamedEntities</strong></td>
<td>38.01</td>
<td>54.10</td>
</tr>
<tr>
<td><strong>word2vec</strong></td>
<td>37.69</td>
<td>53.92</td>
</tr>
<tr>
<td><strong>RuThes</strong></td>
<td>36.27</td>
<td>54.20</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td>36.08</td>
<td>54.32</td>
</tr>
</tbody>
</table>

**Table 6:** Gold boundaries, mention detection f-score 51.21
Distributional models are not able to resolve general hyperonyms:

- Distributional models are not able to resolve general hyperonyms:

  - Distributional models are not able to resolve general hyperonyms:
Distributional models are not able to resolve general hyperonyms:

(7) профессор ‘professor’ — человек ‘man’
Discussion

Distributional models are not able to resolve general hyperonyms:

(7) профессор ‘professor’ — человек ‘man’

RusVectores output for профессор:

- доцент 0.68
- проф 0.66
- преподаватель 0.66
- ректор 0.66
- ученый 0.63
- академик 0.62
- доктор 0.59
- декан 0.58
- преподавать 0.57
- адъюнкт-профессор 0.57
Ontologies and thesauri should help with this:
Ontologies and thesauri should help with this:

Профессор < научный работник < служащий(работник) < человек < живой организм
Ontologies and thesauri should help with this:

- профессор < научный работник < служащий(работник) < человек < живой организм

But some cases are problematic:

(8) дача — таинственное жилище
Ontologies and thesauri should help with this:

препрофессор < научный работник < служащий(работник) < человек < живой организм

But some cases are problematic:

(8) дача — таинственное жилище

Ruthes output:

• дача < загородный дом < жидкое здание < здание < недвижимое имущество
• жилище < место в пространстве
Even though there are some limitations, these approaches improve the quality.
Even though there are some limitations, these approaches improve the quality further:

- Using a NER system
- Using distributional models in a more complex way
- Handling ontologies more carefully to minimize the amount of generated homonymy
Even though there are some limitations, these approaches improve the quality.

Further elaboration of each of them could improve the overall quality further:

- Using a NER system
Even though there are some limitations, these approaches improve the quality.

Further elaboration of each of them could improve the overall quality further:

- Using a NER system
- Using distributional models in a more complex way
Even though there are some limitations, these approaches improve the quality further:

- Using a NER system
- Using distributional models in a more complex way
- Handling ontologies more carefully to minimize the amount of generated homonymy
Thank you!
Any questions?

RuCor corpus: http://rucoref.maimbava.net
Jupyter notebooks: https://github.com/max-ionov/rucoref/tree/master/notebooks/coreference-dialog-2017