

Error analysis for anaphora resolution in Russian: new challenging issues for anaphora resolution task in a morphologically rich language

Svetlana Toldova¹, Alina Ladygina², Maria Vasilyeva³, Ilya Azerkovich¹, Anna Roytberg¹

¹National Research University HSE, ²University of Tübingen, ³Lomonosov MSU

Introduction

This project outlines a quantitative and qualitative error analysis of Russian anaphora resolvers which participated in the RU-EVAL evaluation campaign [Toldova et al. 2014].

The analysis explores language-specific issues having impact on the systems performance, such as a high level of grammatical ambiguity, specific binding conditions, free word order and cases of non-referential pronouns.

Evaluation

- ▶ Evaluation set - Russian Coreference Corpus (RuCor) [Toldova et al. 2016]
- ▶ 85 texts, manually annotated with coreferential and anaphorical relations (1600 chains, 2300 pairs)



Figure 1: RuCor web annotation tool

Run	Algorithm type	P	R	F-score
sys1	rule-based+onto	0.82	0.70	0.76
sys2	rule-based	0.71	0.58	0.64
sys3	rule-based	0.63	0.50	0.55
sys4	logreg+onto	0.54	0.51	0.53
sys5	svm+sem	0.58	0.42	0.49
sys6	decision tree	0.36	0.15	0.21

Table 1: Evaluation results of RU-EVAL-2014

Error analysis

- ▶ The performance is evaluated on the following set of pronouns:

pronoun type	label	frequency
3rd person pronouns (subject)	ana_nom	640
3rd person pronouns (direct object)	ana_acc	217
3rd person pronouns (in PPs)	ana_pp	195
3rd person pronouns (other positions)	ana_other	174
3rd person possessives reflexive	ana_poss	298
reflexive possessive	refl	126
relative	refl_poss	294
	rel	357
total		2301

Table 2: Statistics on pronoun types

Results

- ▶ The least problematic cases are possessive reflexives, relatives and 3rd person pronouns in nominative case.
- ▶ The most difficult is the resolution of reflexives and 3rd person pronouns in accusative case

	sys1	sys2	sys3	sys4	sys5	sys6
ana_nom	0.20	0.33	0.43	0.46	0.44	0.72
ana_acc	0.27	0.36	0.43	0.58	0.5	0.75
ana_pp	0.21	0.39	0.46	0.53	0.45	0.77
ana_other	0.23	0.35	0.44	0.45	0.41	0.69
ana_poss	0.20	0.35	0.46	0.42	0.47	0.68
refl	0.20	0.34	0.52	0.83	0.86	0.65
refl_poss	0.17	0.29	0.41	0.55	0.44	0.60
rel	0.19	0.29	0.43	0.55	0.57	0.71
mean	0.21	0.34	0.45	0.55	0.53	0.69

Table 3: Precision error rate for different pronoun types

Normalized error rate

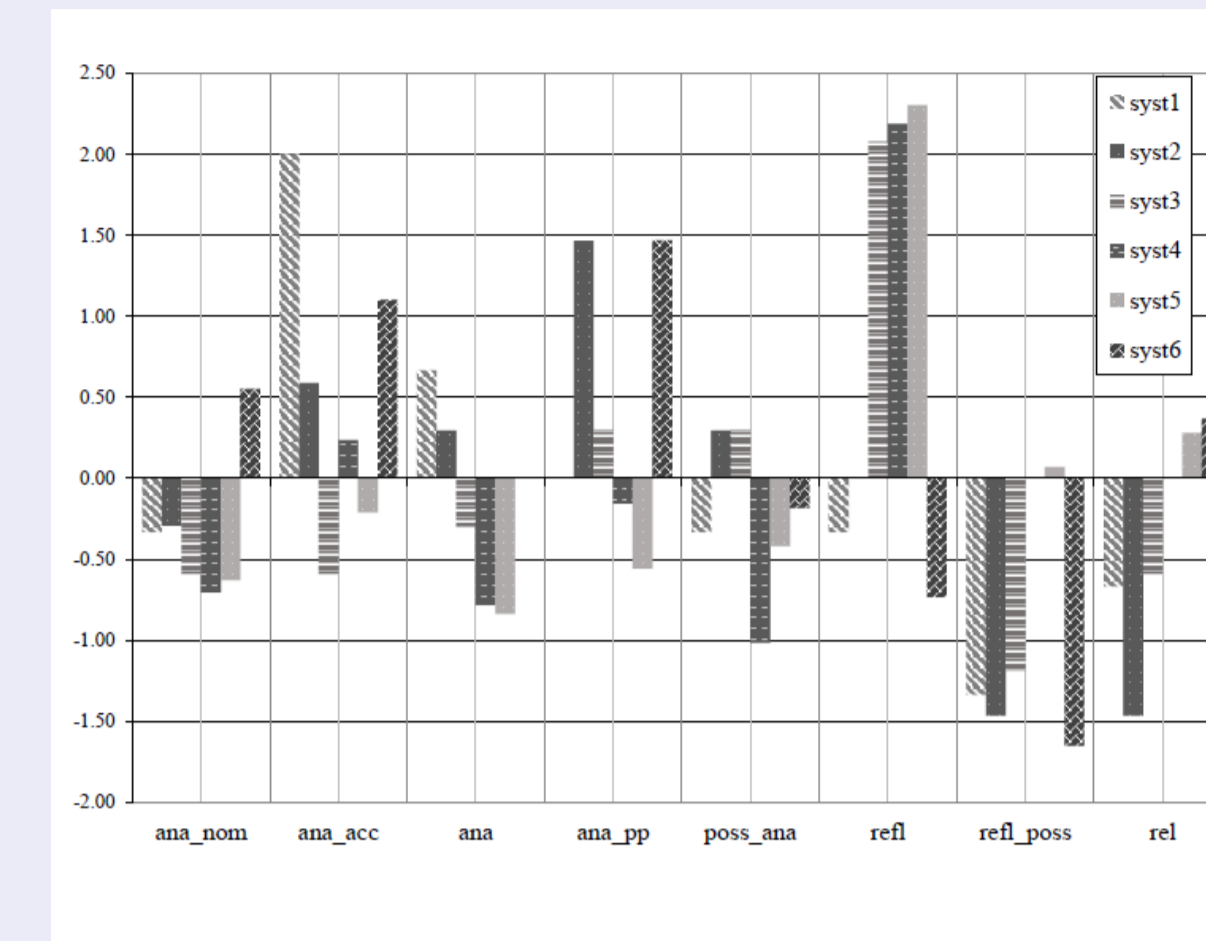


Figure 2: Deviation in error rates for pronoun types across systems

Errors due to morphological ambiguity

- ▶ Pronoun animacy deficiency: *project* (inanim.) vs. *Professor* (anim.) - both are candidates for *on* 'he/it'
- ▶ Lack of gender contrasts in pronouns
- ▶ Nominal case-number syncretism

Incorrect binding

- ▶ Ignoring long distance binding conditions (in infinitive clauses and NPs):
(1) *Ona_i prishla zabrat' svoego_i syna* 'She came to pick up her son'
- ▶ Resolving cataphora by finding an antecedent in the preceding clause:
(2) *Zapretit' etu partiju predlozhili deputaty_s. V svoju_{is} ochered', mestnyj parlament_i ...* '(The) deputy_s suggested to ban the party. In its_{is} turn, (the) local parliament_i ...'
- ▶ Binding reflexives outside the local domain
- ▶ Binding 3rd person pronouns within the same clause

Non-referential pronouns

- ▶ Idiomatic expressions: *svoja rubashka blizche k telu* - 'self before all', *Ego prevoshoditelystvo* - 'His excellency'
- ▶ Associative plural: (3) *Masha obizhaetsya chto my ih ne zovem* 'Mary takes offence that we don't invite them (Mary and her friends)'

Other issues

- ▶ Pre-processing issues: wrong matching of NP boundaries and multi-word expressions
- ▶ NP embedding - ambiguity in antecedent selection: *zdanije ministrestva* 'building of the Ministry'...*ego* 'its/his'

Conclusion

- ▶ The most common issues are relevant for all systems which participated in RU-EVAL-2014 evaluation campaign, despite the difference in their approaches and models.
- ▶ Language-specific properties require a joint fine-grained analysis of morphology, syntax and semantics, as well as particular rules for binding

Acknowledgments

The reported study was supported by the Russian Foundation of Basic Research, research project No. 15-07-09306 "Evaluation benchmark for information retrieval".