

# Identification of Predicate-Argument Structure in Polish

Konrad Gołuchowski

Institute of Computer Science, Polish Academy of Science, Warsaw

## Predicate-Argument Structure

**Arguments:** who did what and how? etc...

are predicate-specific

**Adjuncts:** when? where?

ignored in my experiments

Mark ate a sandwich yesterday.

np(nom)

np(bier)

adjunct



## Experiment I: Finding a verb valence frame

**Problem:** What types of arguments are realized for each predicate?

### Argument types

Based on the syntactic realization of the argument:

- ★ Nominal argument: *np(nom)*
- ★ Prepositional argument: *prepn(w,loc)*
- ★ etc...

### Finding the valence frame

**Classifier:** Multi-output Random Forest

**Features:** extracted from syntactic heads of neighboring syntactic groups and predicate dependants

**Features types:**

- ★ Predicate PoS
- ★ Morphosyntactic
- ★ Valence lexicon (*Walenty*)
- ★ Dependency parser

**Corpus:** manually annotated tree bank *Składnica*

### Results in a best setting

- ★ **Frame-level accuracy:** 65.73%
- ★ **Precision:** 96.03%
- ★ **Recall:** 77.03%
- ★ **F-measure:** 85.47%

## Experiment II: Finding arguments head

**Problem:** Where are argument syntactic heads realized in the sentence? What argument types do they introduce?

### Labeling argument heads position

They are **going** to a pub and **drinking** a beer  
np \_ \_ prepnp \_ \_ \_ \_ np  
(nom) (do,gen) (bier)

### Automatic labeling

**Classifier:** CRF++

**Features:** (for each word)

- ★ Word PoS
- ★ Case (whenever applicable)
- ★ Lemma of prepositions, negation, reflexive *się*
- ★ Is the head of syntactic group?
- ★ Matching argument from valence lexicon
- ★ Dependency relation (if any)

### Results in a best setting

- ★ **Sentence-level accuracy:** 72.93%
- ★ **Precision:** 91.25%
- ★ **Recall:** 90.01%
- ★ **F-measure:** 91.25%

## Conclusions

- ★ The main problem in the experiment I is to achieve high recall. Better results were achieved in the experiment II.
- ★ Detailed error analysis has shown that the main source of errors are arguments that have the same syntactic type as adjuncts (mostly prepositional phrases).
- ★ To obtain full predicate-argument structure both experiments have to be merged.

**Acknowledgements:** The paper is co-funded by the European Union from financial resources of the European Social Fund. Project PO KL "Information technologies: Research and their interdisciplinary applications"