AST: New Tool for Logical Analysis of Sentences based on Transparent Intensional Logic

Marek Medveď, Aleš Horák

Natural Language Processing Centre Faculty of Informatics, Masaryk University
Botanická 68a, 602 00 Brno, Czech Republic

December 5, 2015

Semantic Analysis

- semantic analysis
- based on selected logical system
  - first-order logic

Semantic Analysis

- semantic analysis
- based on selected logical system
    - ~~first-order logic~~
    - high-order logic $\rightarrow$ to capture higher order phenomena that occurs in NL

Automatic Semantic Analysis (AST)

- new standalone tool
- provides the semantic analysis in the form of Transparent Intensional Logic (TIL) constructions
- based on SYNT TIL logical analysis
- language and parser independent

AST Input

- syntactic tree
  - edges
  - morphological information
    - a lemma
    - a PoS tag

Example
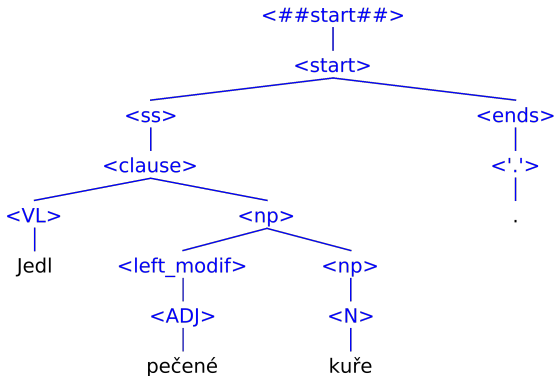
```
<tree>
{##start##
 {start
  {ss
   {clause
     {VL<leaf><idx>0</idx><w>Jedl</w><l>jíst</l>
      <c>k5eAaIgMnS</c></leaf>}
     {intr
      {adjp
       {ADJ<leaf><idx>1</idx><w>pečené</w><l>pečený</l>
        <c>k2eAgNnSc4</c></leaf>}
      }
      {np
       {N<leaf><idx>2</idx><w>kuře</w><l>kuře</l>
        <c>k1gNnSc4</c></leaf>}
      }
     }
   }
  }
  {ends
   {'.'<leaf><idx>3</idx><w>.</w><l>.</l><c>kX</c></leaf>}
  }
 }
}
</tree>
```

AST Input - graphical tree representation

Language Dependent Files

- lexical items
- verb valencies
- prepositional valencies
- a semantic grammar

Lexical Items

- o: representing the truth-values
- ι: class of individuals
- τ: class of time moments
- ω: class of possible worlds

Example (word "jíst")

```
jíst
/k5/otriv  (((o(oo_{τω})(oo_{τω}))ω)ι)
```
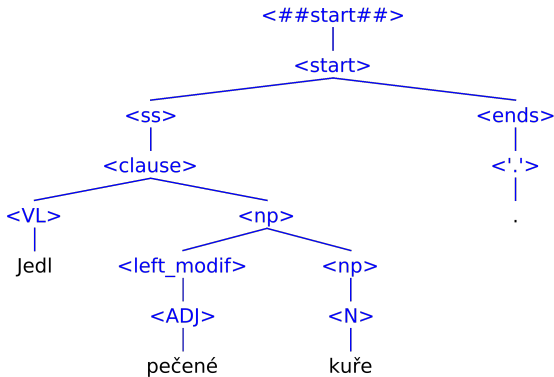
Semantic Grammar File

- resulting semantic construction
  - bottom-up analysis based on the input syntactic tree
  - semantic extension of the actual grammar used in the parsing process

Example (syntactic rule with TIL action for the <np> node)

```
np -> left_modif np
  rule_schema ( "[#1,#2]" )
```

AST Input - graphical tree representation

Semantic Grammar File

Example (semantic analysis for the <np> node)

```
rule_schema: 2 nterms, '[#1,#2]'
1, 3, +np -> . left_modif  np . @level 0
  nterm 1: 1, 2, +left_modif -> . left_modifnl  . @level 0,
           k2eAgNnSc4
       TIL: ⁰pečený...((oι)_τω(oι)_τω)
  nterm 2: 2, 3, +np -> . N  . @level 0, k1gNnSc4
       TIL: ⁰kuře...(oι)_τω

Processing schema with params:
    #1: ⁰pečený...((oι)_τω(oι)_τω)
    #2: ⁰kuře...(oι)_τω
Resulting constructions:

    [⁰pečený/((oι)_τω(oι)_τω), ⁰kuře/(oι)_τω]...(oι)_τω
```

Verb Valencies

- builds the resulting construction from the corresponding valency frame

Example (verb "jíst")

```
jíst
hPTc4 :exists:V(v):V(v):and:V(v)=[[#0,try(#1)],V(w)]
```

SA
○
AST
○○○○○○○○○○○●○
Error Analysis
○○○○○○○
Evaluation
○
Conclusions
○

Prepositional Valency Expressions

- semantic mappings of prepositional phrases to valency expressions based on the head preposition

Example (preposition "k")

```
                k
                3 hA hH
```

AST system parts

- the input parser
- the grammar parser
- the lexical item parser
- the schema parser
- the verb valency parser
- the prepositional valency expression parser

Error Analysis

- continuously evaluated in comparison with the original SYNT TIL logical analysis
- number of uncovered phenomena is described and implemented in AST

Error Analysis

- sentences with two items divided by "and"
- verb valency vs. clause valency
- verb valency schema update
- verb valency schema missing
- system errors

Sentences with Two Items Divided by "and"

Example ("and" sentence)

Vidíte zásadnější rozdíly mezi přístupy českých a západních informačních firem?

(Can you see the main difference between Czech and west information companies?)

Example (construction)

```
          [západní, [český, x8]] and [firma, x8]
                              ↓
[západní, x8] and [firma,x8] and [český,x9] and [firma,x9]
```

Verb Valency vs. Clause Valency

Example

Možná, že se tito lidé ani nesetkali.

(Maybe, these people never met each other.)

- the clause schema does not contain the reflexive pronoun "*se*"

SA
○

AST
○○○○○○○○○○○○○

Error Analysis
○○○○●○○

Evaluation
○

Conclusions
○

Verb Valency Schema Update

- verb valency file does not contain an option that can match with the created cause valency

Verb Valency Schema Missing

- verb valency list is created from the Czech VerbaLex lexicon
  - 10,000 Czech verb lemmata and their verb frames

System Errors

- construction checker that does not allow the dash character "-" in the name of an object construction

Example

$$[0(\text{Si-an}/\iota)]$$

Evaluation

Table: 200 sentences evaluated by the SYNT TIL system and the AST system

system	correct	correct in %	incorrect	incorrect in %
SYNT TIL	131	65.5 %	69	35.5 %
AST	158	**79.0 %**	42	**21.0 %**

Conclusions

- new language and parser independent tool for semantic analysis
- corrects several frequent errors of its predecessor → **14% increase**

Acknowledgements

Thank you for your attention.