

# Tecto to AMR and translation

Ondřej Bojar, Silvie Cinková,  
Ondřej Dušek, Tim O’Gorman,  
Martin Popel, Roman Sudarikov,  
Zdeňka Urešová

August 1, 2014

# Introduction

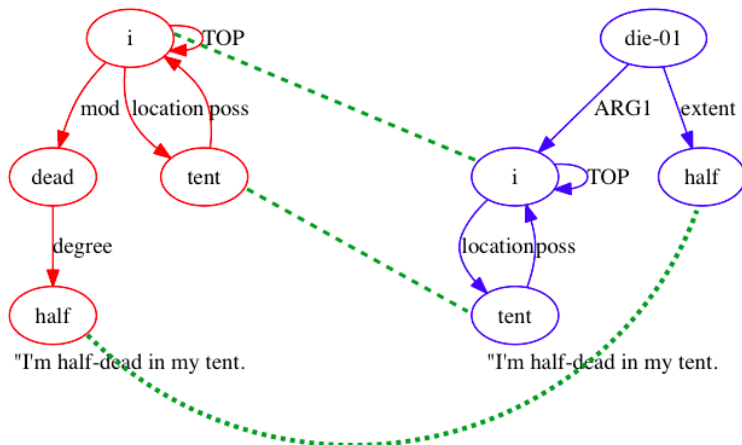
# Motivation

- ▶ We are investigating the value of parallel Abstract Meaning Representations (AMRs)
- ▶ Question 1: How similar are AMRs made in different languages? How do you compare them?
- ▶ Question 2: How could we get a large corpus of parallel AMRs?

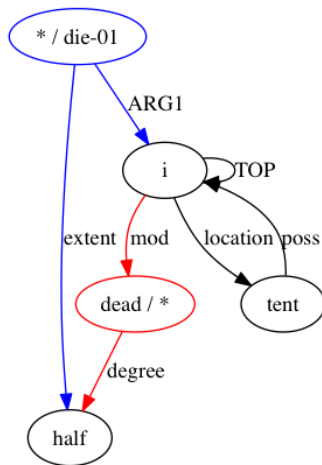
# AMRICA

- ▶ (AMR Inspector with Cross-language Alignment)
- ▶ Usual evaluation and alignment methods break across languages.
- ▶ Extension to Smatch (Cai & Knight 2012).

# Smatch Classic

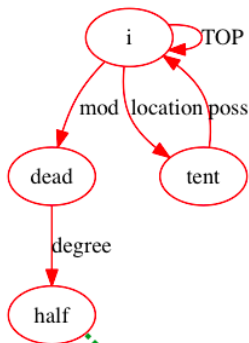


# Smatch Classic

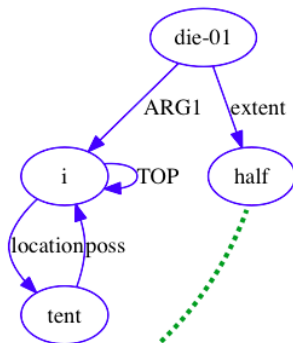


"I'm half-dead in my tent."

# Smatch Classic

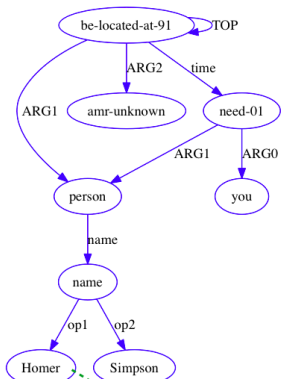


"I'm half-dead in my tent."

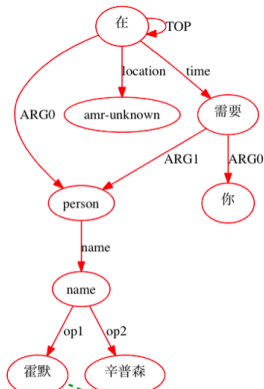


"I'm half-dead in my tent."

# AMRICA



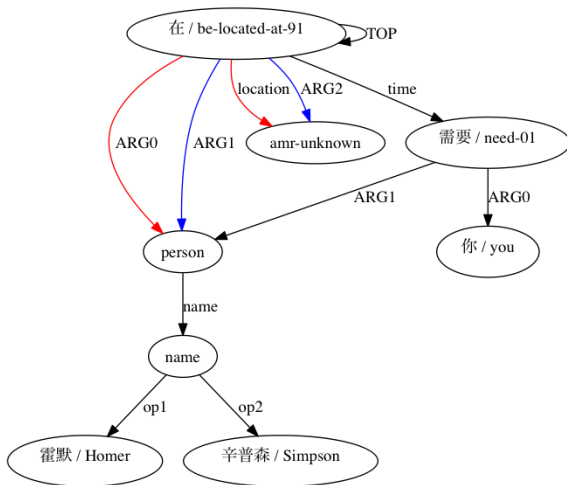
Where's Homer Simpson when you need him ?



当你需要他时，霍默辛普森在哪里？



# AMRICA



Where's Homer Simpson when you need him ?  
当你需要他时，霍默辛普森在哪里？

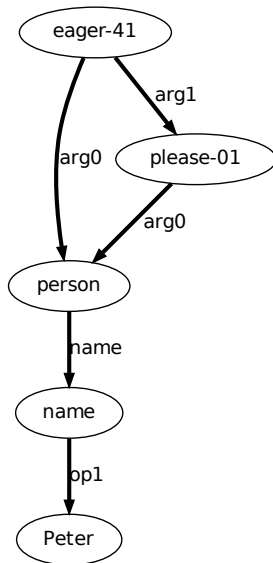
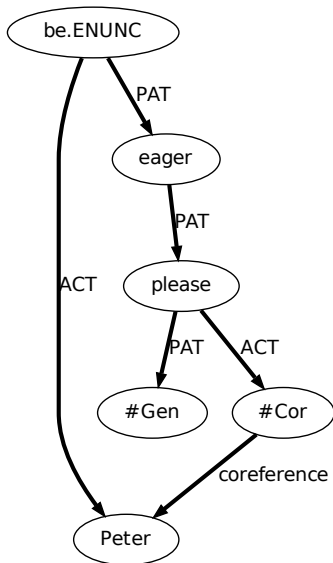
# T-layer to AMR conversion

- ▶ PCEDT: Large parallel corpus (WSJ) annotated with t-layer for English and Czech
- ▶ T-layer to AMR conversion would provide a large static parallel AMR corpus.
- ▶ Could be used dynamically to turn a "t-layer" parser into an AMR parser.

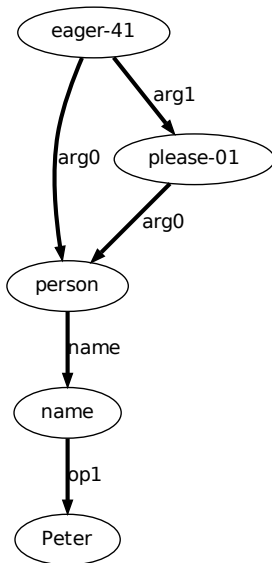
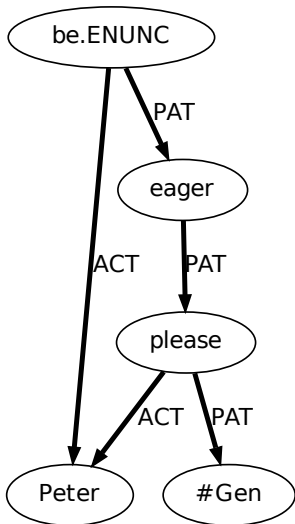
# Why this might work

- ▶ AMR and t-layer are very similar:
  - ▶ Both abstract away from syntax.
  - ▶ Both make all semantic links in a sentence in a graph format.
  - ▶ Both do coreference
- ▶ Various minor structural differences.
- ▶ AMR is more abstract, makes more inference.

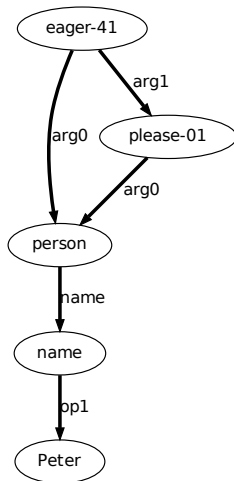
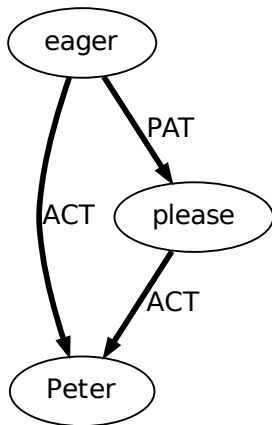
# “Peter is eager to please”



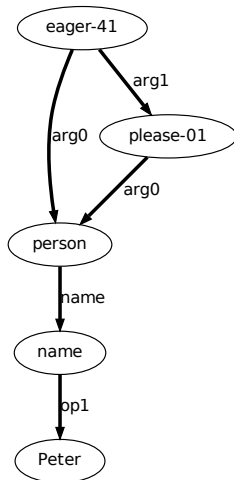
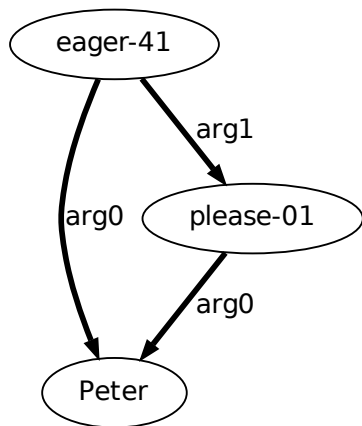
# Merging of Coreferent Nodes



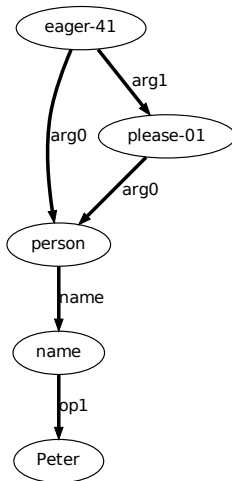
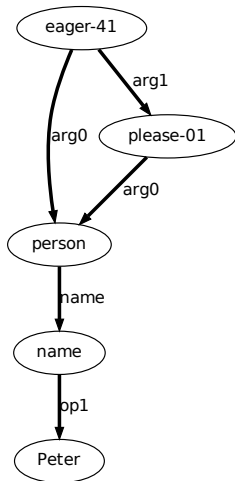
# Elimination of semantically light words



# Semantic Roles and Senses



# Add Named Entities





# Conversion Procedures

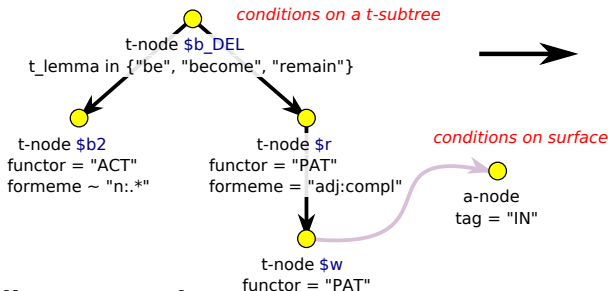
- ▶ Converted t-trees to AMR format
- ▶ Added named entities using NER systems (Stanford and NameTag)
- ▶ Tried two strategies for doing more complex changes to the graphs:
  - ▶ PML-TQ
  - ▶ Tsurgeon
- ▶ List-based verbalization and semantic role mapping

# PML-TQ rules

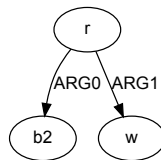
- ▶ Based on AMR guidelines (generalized)
- ▶ For copula, attributes, non-core roles ...

## A PML-TQ rule

### LHS (PML-TQ Query)



### RHS (AMR Subtree)

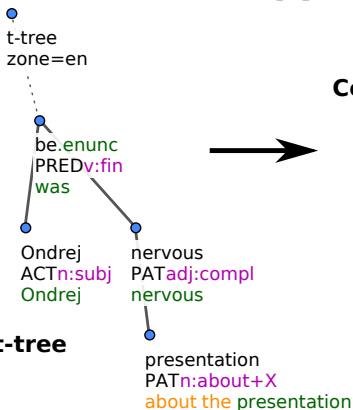


### Guidelines example:

*The boy is responsible for the work.*

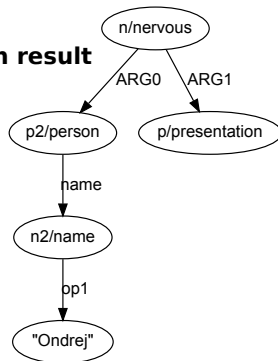
# PML-TQ rules

## Rule application



Matching t-tree

## Conversion result



## Matching sentence:

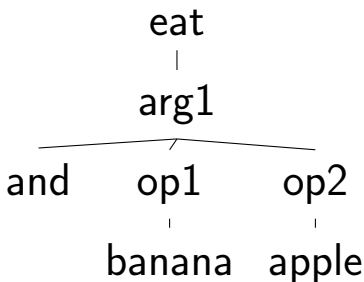
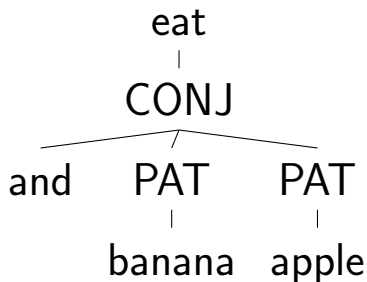
*Ondrej was nervous about the presentation.*

# Tsurgeon tree transformation rules

- ▶ We converted to constituency trees so as to use a tree transformation tool, **Tsurgeon** (Levy and Andrews 2006) to quickly implement hand-written rules.

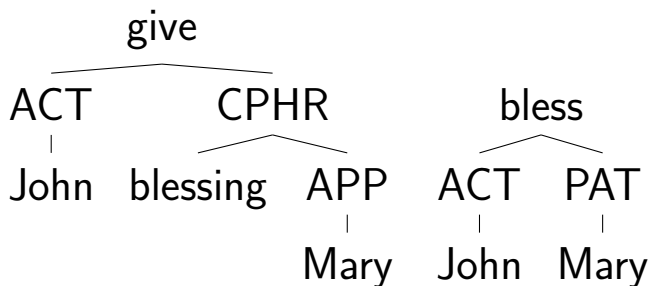
# Transformer tree transformation rules

- ▶ Many of the structural differences are just notational differences:



# List-based Methods

- ▶ Verbalizations are based on dictionary look-ups:
  - ▶ beekeeper → person :ARG0-of keep-01 :ARG1 bee
- ▶ As are complex predications:



# Using Existing Resources

|                                 | Vallex | Propbank Lexicon | Other WSJ annotation | Lexical Lists |
|---------------------------------|--------|------------------|----------------------|---------------|
| Map t-layer roles to AMR roles  | X      | X                | X                    |               |
| Verbalize nouns/adjectives      |        | X                |                      | X             |
| Introduce inferrable predicates |        |                  |                      | X             |
| Named Entities                  |        |                  | X                    | X             |

# Results of EN t-to-AMR Conv

|                              | Semantic Role Mapping |   |   |   | Named Entities | Verbalization Lists | Smatch | Smatch w/o senses |
|------------------------------|-----------------------|---|---|---|----------------|---------------------|--------|-------------------|
| Baseline (direct conversion) |                       |   |   |   |                |                     | 20     | 28                |
| Baseline (direct conversion) | X                     |   |   |   |                |                     | 33     | 41                |
| Baseline (direct conversion) | X                     | X |   |   |                |                     | 37     | 45                |
| Baseline (direct conversion) | X                     | X | X |   |                |                     | 40     | 48                |
| PML-TQ (guidelines-based)    | X                     |   |   | X |                |                     | 35     | 43                |
| PML-TQ (guidelines-based)    | X                     | X | X |   |                |                     | 38     | 47                |
| Tsurgeon (rule-based)        | X                     | X | X |   |                |                     | 44     | 52                |
| JAMR                         |                       |   |   |   |                |                     | 44     | 45                |