Incremental Semantic Role Labeling with Tree Adjoining Grammar

Ioannis Konstas

Joint work with Frank Keller, Vera Demberg and Mirella Lapata

Institute for Language, Cognition and Computation
University of Edinburgh

2 October 2014

Human Language Processing

Human language processing is *incremental:* we update our parse of the input for each new word that comes in.

Human Language Processing

Human language processing is *incremental:* we update our parse of the input for each new word that comes in.

Incrementality leads to local ambiguity, which we can observe in *garden path sentences:*

(1) a. The old man the boat.
    b. I convinced her children are noisy.

Human Language Processing

Many garden paths are not due to syntactic ambiguity alone, they also involve *semantic role ambiguity*

Human Language Processing

Many garden paths are not due to syntactic ambiguity alone, they also involve *semantic role ambiguity*

(2) The athlete realised her goals . . .

a. . . . at the competition.

b. . . . were out of reach.

This indicates that humans *incrementally* assign semantic roles.

Human Language Processing

Many garden paths are not due to syntactic ambiguity alone, they also involve *semantic role ambiguity*

(2) The athlete realised her goals ...

a. ... at the competition.

b. ... were out of reach.

This indicates that humans *incrementally* assign semantic roles.

Let's look at this example in more detail.

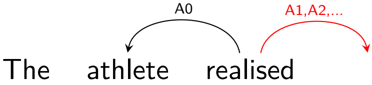
Human Language Processing - Example

A0

The athlete realised

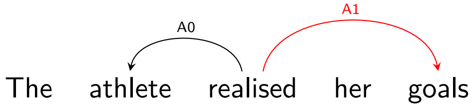
⟨A0,athlete,realised⟩

Human Language Processing - Example



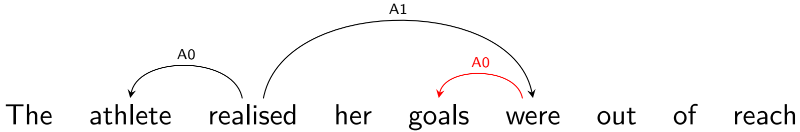
⟨A0,athlete,realised⟩
⟨[A1,A2],nil,realised⟩

Human Language Processing - Example



⟨A0,athlete,realised⟩
⟨A1,goals,realised⟩

Human Language Processing - Example



⟨A0,athlete,realised⟩
⟨A1,were,realised⟩
⟨A0,goals,were⟩

Incremental Semantic Role Labeling

- Determine Semantic Role Labels as the input unfolds

Incremental Semantic Role Labeling

- Determine Semantic Role Labels as the input unfolds
- Given a sentence prefix and its partial syntactic structure:

Incremental Semantic Role Labeling

- Determine Semantic Role Labels as the input unfolds
- Given a sentence prefix and its partial syntactic structure:
  1. Identify Arguments and Predicates

Incremental Semantic Role Labeling

- Determine Semantic Role Labels as the input unfolds
- Given a sentence prefix and its partial syntactic structure:
  1. Identify Arguments and Predicates
  2. Assign correct role labels

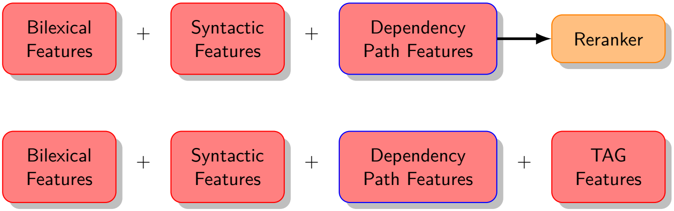
Incremental Semantic Role Labeling

- Determine Semantic Role Labels as the input unfolds
- Given a sentence prefix and its partial syntactic structure:
  1. Identify Arguments and Predicates
  2. Assign correct role labels
- Assign incomplete semantic roles

Non-incremental SRL

Pipeline approach

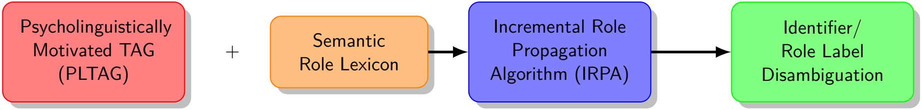
- Liu and Sarkar (2007)
- Màrquez et al. (2008)
- Björkelund et al. (2009) (MATE)



Màrquez et al. (2008), Björkelund et al. (2009)

Liu and Sarkar (2007)

Model


Psycholinguistically Motivated TAG (PLTAG)
+
Semantic Role Lexicon
Incremental Role Propagation Algorithm (IRPA)
Identifier/ Role Label Disambiguation

Psycholinguistically Motivated TAG (PLTAG)

Psycholinguistically Motivated TAG (PLTAG), is a variant of tree-adjoining grammar (Demberg et al., 2014):

Psycholinguistically Motivated TAG (PLTAG)

Psycholinguistically Motivated TAG (PLTAG), is a variant of tree-adjoining grammar (Demberg et al., 2014):

- in standard TAG, the lexicon consists of initial trees and auxiliary trees (both are lexicalized);

Psycholinguistically Motivated TAG (PLTAG)

Psycholinguistically Motivated TAG (PLTAG), is a variant of tree-adjoining grammar (Demberg et al., 2014):

- in standard TAG, the lexicon consists of initial trees and auxiliary trees (both are lexicalized);
- it adds unlexicalized predictive trees to achieve connectivity;

Psycholinguistically Motivated TAG (PLTAG)

Psycholinguistically Motivated TAG (PLTAG), is a variant of tree-adjoining grammar (Demberg et al., 2014):

- in standard TAG, the lexicon consists of initial trees and auxiliary trees (both are lexicalized);
- it adds unlexicalized predictive trees to achieve connectivity;
- the standard TAG operations are substitution and adjunction;

Psycholinguistically Motivated TAG (PLTAG)

Psycholinguistically Motivated TAG (PLTAG), is a variant of tree-adjoining grammar (Demberg et al., 2014):

- in standard TAG, the lexicon consists of initial trees and auxiliary trees (both are lexicalized);
- it adds unlexicalized predictive trees to achieve connectivity;
- the standard TAG operations are substitution and adjunction;
- it adds verification to verify predictive trees;

Psycholinguistically Motivated TAG (PLTAG)

Psycholinguistically Motivated TAG (PLTAG), is a variant of tree-adjoining grammar (Demberg et al., 2014):

- in standard TAG, the lexicon consists of initial trees and auxiliary trees (both are lexicalized);
- it adds unlexicalized predictive trees to achieve connectivity;
- the standard TAG operations are substitution and adjunction;
- it adds verification to verify predictive trees;

PLTAG supports parsing with incremental, fully connected structures.

PLTAG

Lexicon:

- Standard TAG lexicon
- Predictive lexicon (PLTAG)

Operations:

- Substitution
- Adjunction
- Verification (PLTAG)

PLTAG

Lexicon:

- Standard TAG lexicon
- Predictive lexicon (PLTAG)

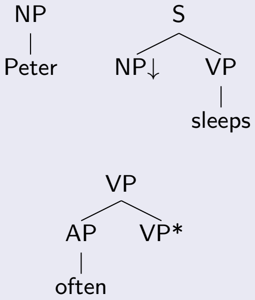
Operations:

- Substitution
- Adjunction
- Verification (PLTAG)

Example

Initial Tree:

Auxiliary Tree:



PLTAG

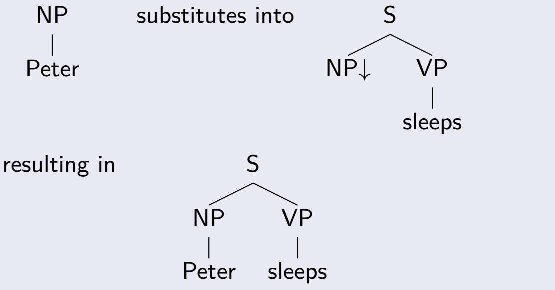
Lexicon:

- Standard TAG lexicon
- Predictive lexicon (PLTAG)

Operations:

- Substitution
- Adjunction
- Verification (PLTAG)

Example



PLTAG

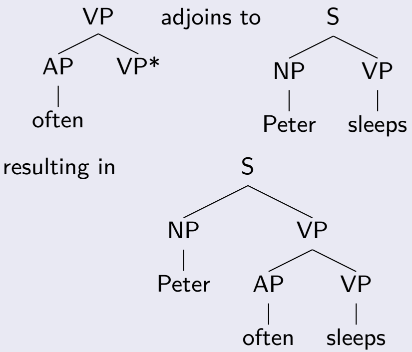
Lexicon:

- Standard TAG lexicon
- Predictive lexicon (PLTAG)

Operations:

- Substitution
- Adjunction
- Verification (PLTAG)

Example



PLTAG

Lexicon:

- Standard TAG lexicon
- Predictive lexicon (PLTAG)

Operations:

- Substitution
- Adjunction
- Verification (PLTAG)

Example

Prediction Tree:

$$\mathrm{S}_k \to \mathrm{NP}^k\!\downarrow \quad \mathrm{VP}^k_k$$

Index k marks predicted node.

PLTAG

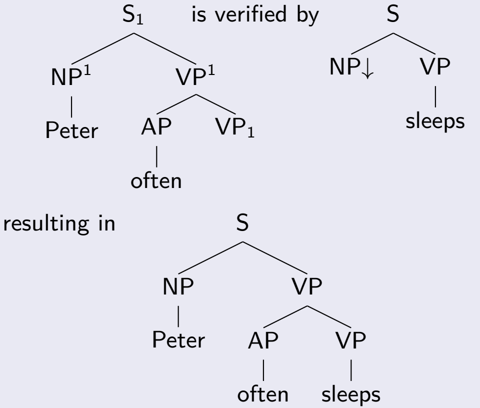
Lexicon:

- Standard TAG lexicon
- Predictive lexicon (PLTAG)

Operations:

- Substitution
- Adjunction
- Verification (PLTAG)

Example

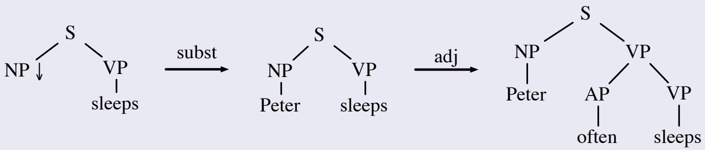


All nodes indexed with k have to be verified.

Comparison with TAG

TAG derivations are not always incremental.

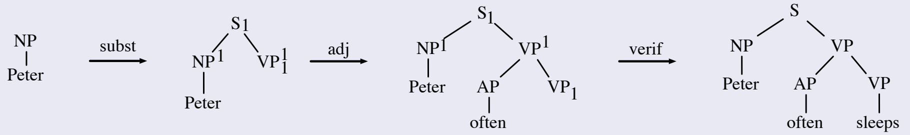
Example



Comparison with TAG

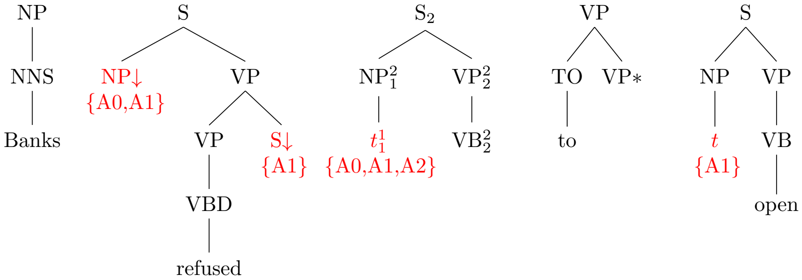
PLTAG derivation are always incremental and fully connected.

Example

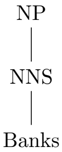


Semantic Roles in Lexicon

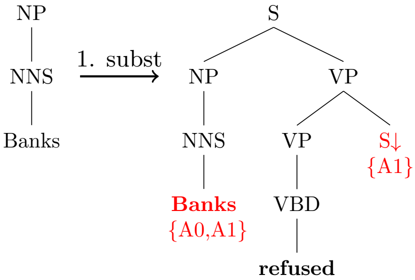
Used information for verb predicates *only*, derived from PropBank (Palmer, 2005)



Incremental Role Propagation Algorithm

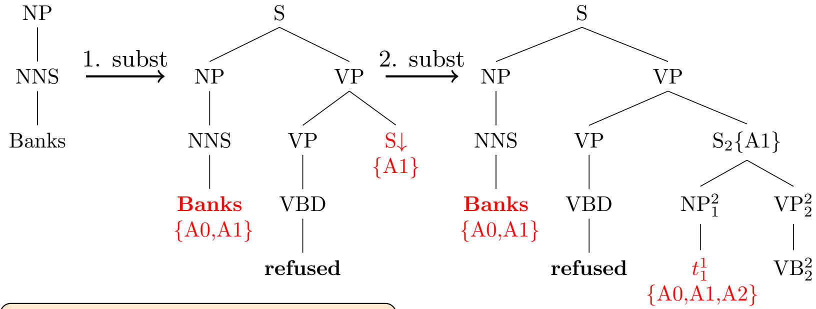


Incremental Role Propagation Algorithm



1. NP $\rightarrow$ ⟨{A0,A1},Banks,refused⟩
   S $\rightarrow$ ⟨A1,nil,refused⟩

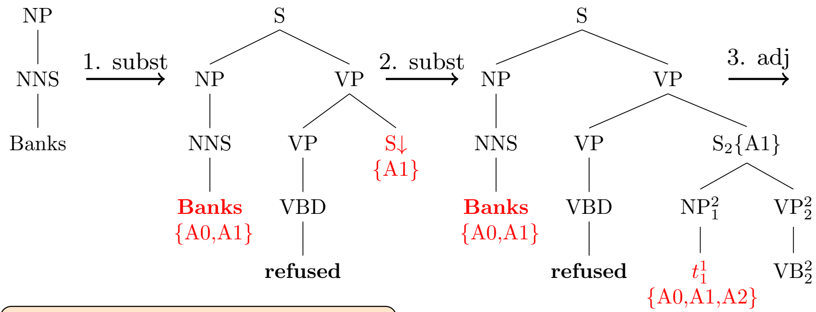
Incremental Role Propagation Algorithm



1. $\text{NP} \rightarrow \langle\{\text{A0,A1}\},\text{Banks},\text{refused}\rangle$
   $\text{S} \rightarrow \langle\text{A1},\text{nil},\text{refused}\rangle$

2. $\text{NP} \rightarrow \langle\{\text{A0,A1}\},\text{Banks},\text{refused}\rangle$
   $\text{S} \rightarrow \langle\text{A1},S_2,\text{refused}\rangle$
   $\text{NP} \rightarrow \langle\{\text{A0,A1,A2}\},t,\text{nil}\rangle$

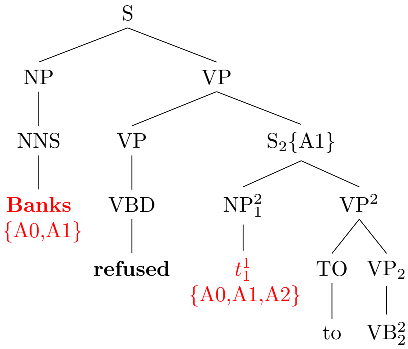
Incremental Role Propagation Algorithm



1. NP $\rightarrow \langle$\{A0,A1\},Banks,refused$\rangle$
   S $\rightarrow \langle$A1,nil,refused$\rangle$

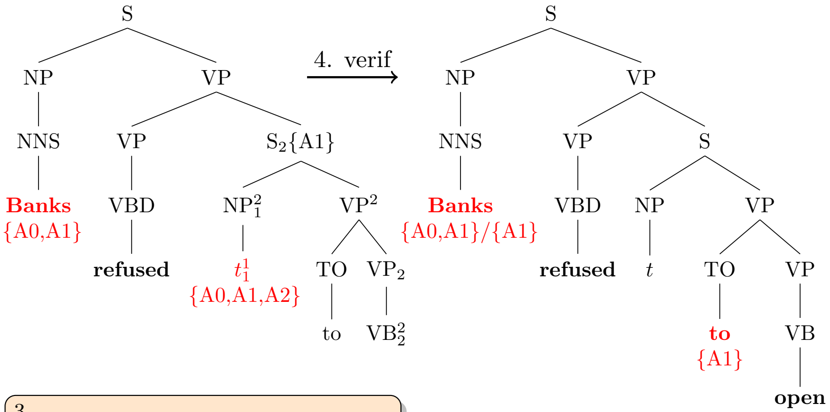
2. NP $\rightarrow \langle$\{A0,A1\},Banks,refused$\rangle$
   S $\rightarrow \langle$A1,S_2,refused$\rangle$
   NP $\rightarrow \langle$\{A0,A1,A2\},t,nil$\rangle$

Incremental Role Propagation Algorithm



3. —

Incremental Role Propagation Algorithm

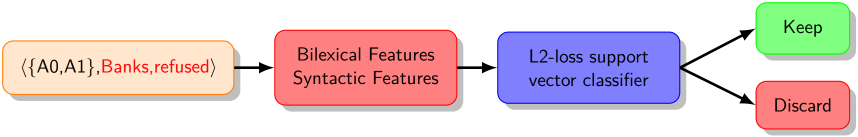


3. —

4. NP $\rightarrow$ ⟨{A0,A1},Banks,refused⟩
   S $\rightarrow$ ⟨A1,to,refused⟩
   NP $\rightarrow$ ⟨A1,Banks,open⟩

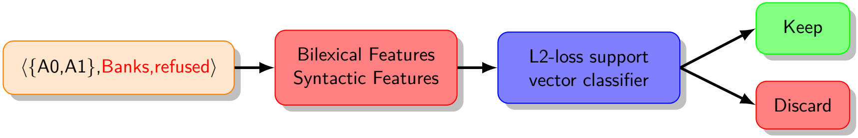
Argument Identification - Role Label Disambiguation

Argument Identification

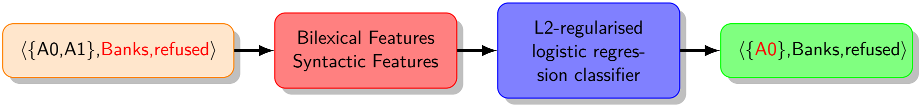


Argument Identification - Role Label Disambiguation

Argument Identification



Role Label Disambiguation



Experiments

- Train PLTAG on sections WSJ 02-21 (79.41% F_1)
- Train classifiers on CoNLL 2009 (Ident.: 92.18%, Lab.: 82.37%)
- Gold lexicon entries during parsing - CoNLL-SRL-only task

Experiments

- Train PLTAG on sections WSJ 02-21 (79.41% F_1)
- Train classifiers on CoNLL 2009 (Ident.: 92.18%, Lab.: 82.37%)
- Gold lexicon entries during parsing - CoNLL-SRL-only task

Evaluation

- Full sentence Accuracy (F_1)
- Unlabelled Prediction Score (UPS)
- Combined Incremental SRL Score (CISS)

Experiments

- Train PLTAG on sections WSJ 02-21 (79.41% F_1)
- Train classifiers on CoNLL 2009 (Ident.: 92.18%, Lab.: 82.37%)
- Gold lexicon entries during parsing - CoNLL-SRL-only task

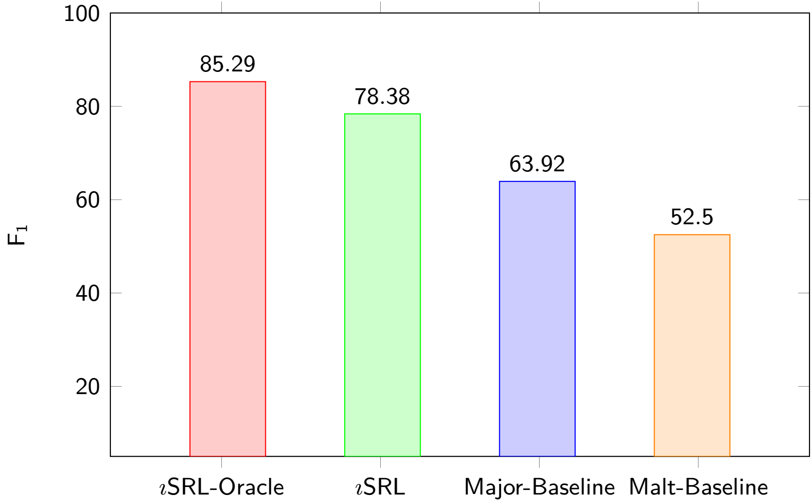
Evaluation

- Full sentence Accuracy (F_1)
- Unlabelled Prediction Score (UPS)
- Combined Incremental SRL Score (CISS)

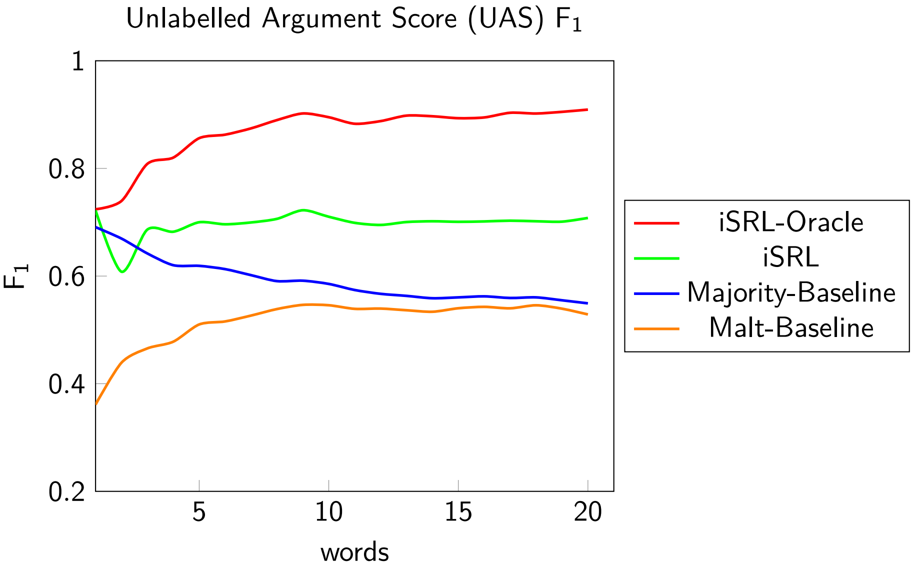
System Comparison

- $\imath$SRL -Oracle
- $\imath$SRL
- Majority-Baseline
- Malt-Baseline

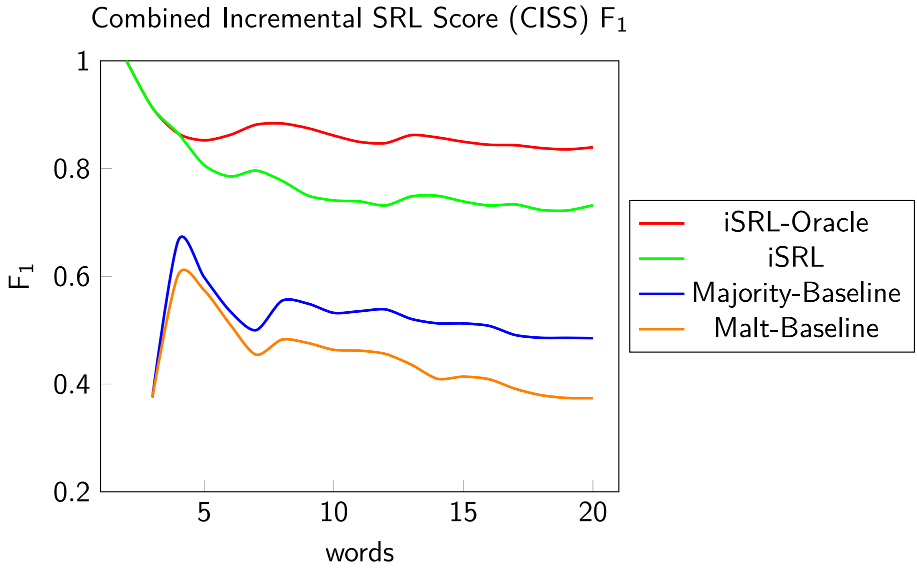
Results - Full sentence



Results - Incremental



Results - Incremental



Conclusions

- New task of Incremental Semantic Role Labeling
- Our system combines:
  - Psycholinguistically Motivated TAG (PLTAG)
  - Semantic Role Lexicon
  - Incremental Role Propagation Algorithm (IRPA)
  - Argument Identification, Role Disambiguation Classifiers
- Outperforms baselines
- Performs well incrementally: predicts (in)-complete triples early in the sentence

Fusing Syntax with Semantics

- Use $\imath$SRL labels as pivotal points and score with model of semantics
- PLTAG Parser Reranker

Fusing Syntax with Semantics

- Use $\imath$SRL labels as pivotal points and score with model of semantics
- PLTAG Parser Reranker

Banks

$$y^* \qquad f(d_1^*) \times \alpha$$

$$\hat{y} \qquad \begin{pmatrix} f(d_{11}) \times \alpha & & \\ f(d_{21}) \times \alpha & & \\ f(d_{31}) \times \alpha & & \\ f(d_{41}) \times \alpha & \rightarrow & \hat{y}_1 \\ f(d_{51}) \times \alpha & & \end{pmatrix}$$

$$\alpha \leftarrow \alpha + f(d_1^*) - f(d_{41})$$

Fusing Syntax with Semantics

- Use $\imath$SRL labels as pivotal points and score with model of semantics
- PLTAG Parser Reranker

$$
\begin{array}{lcc}
 & \text{Banks} & \text{refused} \\
y^* & f(d_1^*) \times \alpha & f(d_2^*) \times \alpha \\
\hat{y} & \begin{pmatrix} f(d_{11}) \times \alpha & & \\ f(d_{21}) \times \alpha & & \\ f(d_{31}) \times \alpha & & \\ f(d_{41}) \times \alpha & \rightarrow & \hat{y}_1 \\ f(d_{51}) \times \alpha & & \end{pmatrix} & \begin{pmatrix} f(d_{12}) \times \alpha & & \\ f(d_{22}) \times \alpha & \rightarrow & \hat{y}_2 \\ f(d_{32}) \times \alpha & & \\ f(d_{42}) \times \alpha & & \\ f(d_{52}) \times \alpha & & \end{pmatrix} \\
 & \alpha \leftarrow \alpha + f(d_1^*) - f(d_{41}) & \alpha \leftarrow \alpha + f(d_2^*) - f(d_{22})
\end{array}
$$

Fusing Syntax with Semantics

- Use $\imath$SRL labels as pivotal points and score with model of semantics
- PLTAG Parser Reranker

$$
\begin{array}{cccc}
 & \text{Banks} & \text{refused} & \text{to} \\
y^* & f(d_1^*)\times\alpha & f(d_2^*)\times\alpha & f(d_3^*)\times\alpha \\
\hat{y} &
\begin{pmatrix} f(d_{11})\times\alpha & \\ f(d_{21})\times\alpha & \\ f(d_{31})\times\alpha & \\ f(d_{41})\times\alpha & \rightarrow\ \hat{y}_1 \\ f(d_{51})\times\alpha & \end{pmatrix} &
\begin{pmatrix} f(d_{12})\times\alpha & \\ f(d_{22})\times\alpha & \rightarrow\ \hat{y}_2 \\ f(d_{32})\times\alpha & \\ f(d_{42})\times\alpha & \\ f(d_{52})\times\alpha & \end{pmatrix} &
\begin{pmatrix} f(d_{13})\times\alpha & \\ f(d_{23})\times\alpha & \rightarrow\ \hat{y}_3 \\ f(d_{33})\times\alpha & \\ f(d_{43})\times\alpha & \\ f(d_{53})\times\alpha & \end{pmatrix} \\
 & \alpha\leftarrow\alpha+f(d_1^*)-f(d_{41}) & \alpha\leftarrow\alpha+f(d_2^*)-f(d_{22}) & \alpha\leftarrow\alpha+f(d_3^*)-f(d_{23})
\end{array}
$$

Fusing Syntax with Semantics

- Use $\imath$SRL labels as pivotal points and score with model of semantics
- PLTAG Parser Reranker

$$
\begin{array}{ccccc}
 & \text{Banks} & \text{refused} & \text{to} & \text{open} \\
y^* & f(d_1^*)\times\alpha & f(d_2^*)\times\alpha & f(d_3^*)\times\alpha & f(d_4^*)\times\alpha \\
\hat{y} &
\begin{pmatrix} f(d_{11})\times\alpha & \\ f(d_{21})\times\alpha & \\ f(d_{31})\times\alpha & \\ f(d_{41})\times\alpha & \rightarrow \hat{y}_1 \\ f(d_{51})\times\alpha & \end{pmatrix} &
\begin{pmatrix} f(d_{12})\times\alpha & \\ f(d_{22})\times\alpha & \rightarrow \hat{y}_2 \\ f(d_{32})\times\alpha & \\ f(d_{42})\times\alpha & \\ f(d_{52})\times\alpha & \end{pmatrix} &
\begin{pmatrix} f(d_{13})\times\alpha & \\ f(d_{23})\times\alpha & \rightarrow \hat{y}_3 \\ f(d_{33})\times\alpha & \\ f(d_{43})\times\alpha & \\ f(d_{53})\times\alpha & \end{pmatrix} &
\begin{pmatrix} f(d_{14})\times\alpha & \\ f(d_{24})\times\alpha & \\ f(d_{34})\times\alpha & \rightarrow \hat{y}_4 \\ f(d_{44})\times\alpha & \\ f(d_{54})\times\alpha & \end{pmatrix} \\
 & \alpha \leftarrow \alpha + f(d_1^*) - f(d_{41}) & \alpha \leftarrow \alpha + f(d_2^*) - f(d_{22}) & \alpha \leftarrow \alpha + f(d_3^*) - f(d_{23}) & \alpha \leftarrow \alpha + f(d_4^*) - f(d_{34})
\end{array}
$$

Features

- Baseline PLTAG probability model score
- Syntactic Features
  - Current lexicon entry
  - Previous lexicon entry
  - Bigram lexicon entries
  - Unlexicalised features
- Current SRL triple(s)
- Semantic Score

Features

- Baseline PLTAG probability model score
- Syntactic Features
  - Current lexicon entry
  - Previous lexicon entry
  - Bigram lexicon entries
  - Unlexicalised features
- Current SRL triple(s)
- Semantic Score

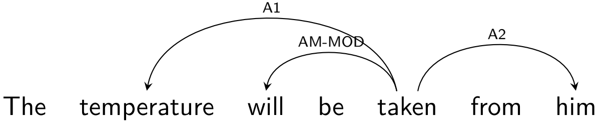
Semantic Score

- Blacoe and Lapata, 2013: CDT model trained using SRL instead of dependencies
- Sayeed and Demberg, ongoing: Baroni and Lenci, 2010 -inspired also trained using SRL instead of dependencies
- Baselines (No syntax)
  - Mikolov et al., 2013
  - Mitchell and Lapata, 2010

Semantic Score

- Blacoe and Lapata, 2013: CDT model trained using SRL instead of dependencies
- Sayeed and Demberg, ongoing: Baroni and Lenci, 2010 -inspired also trained using SRL instead of dependencies
- Baselines (No syntax)
  - Mikolov et al., 2013
  - Mitchell and Lapata, 2010

Multiple Triples (vary composition function)



⟨A1,temperature,taken⟩
⟨AM-MOD,will,taken⟩
⟨A2,him,taken⟩

Thank you