Investigating the Relation between Automatic Parsing and Human Sentence Processing

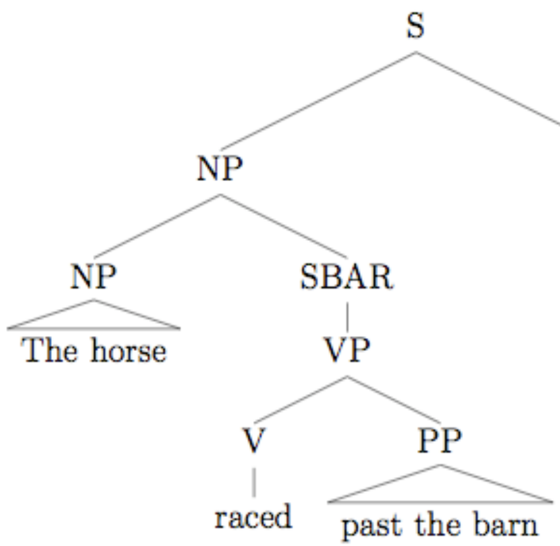
The Dependency Shift-Reduce Parser

Atanas Georgiev Chanev
PhD in Cognitive Science
Linguistic Modeling Lab
Bulgarian Academy of Sciences

Contents:

The Syntactic Parsing Task

- The task of building the syntactic structure of a sentence in a natural language
- It is also a cognitive function



Problems:

1. Sentences in natural languages can be structurally ambiguous
2. Humans have difficulties with the processing of certain syntactic structures

Automatic Parsing

- Natural Language Processing Technology

  Automatic parsing is a subtask used in areas such as Question Answering. Possibilities for its use in text processing tasks, such as Machine Translation and Information Extraction are being explored.

- Limitations of Automatic Parsers

  Current parsing techniques seem to have reached their upper performance limit. Novel techniques are needed to increase parsing performance.

Human Sentence Processing

- Understanding the way humans analyze sentences

  What is the architecture of the human parser? (Is it modular? Is it connectionist? Is it serial or parallel?) How is it related to other faculties of the human mind?

- Limitations of Models of the Human Sentence Parsing Mechanism

  Lack of wide-coverage. Inability to explain reaction times of human subjects in sentence processing experiments.

Lack of a Unified Framework

Parser/Quality	High Accuracy	Wide Coverage	Psychological Plausibility
Automatic Parsers (APs)	✓ (≈90% English)	✓Able to parse unseen input	X Unable to explain human difficulty
Models of the Human Parser	X Far from state-of-the art APs	X (only a limited number of models have wide-coverage)	✓Plausible for certain experimental materials

A unified framework: **Objectives**

- Develop a model of the human parser that combines the qualities of automatic parsers and those of architectures of the human sentence processing mechanism
- Use knowledge about the architecture of the human parser to increase the accuracy of automatic parsers
- Extend the psychological plausibility of models of the human parser

A Wide Coverage Model of The Human Sentence Processing Mechanism

A Wide-Coverage Model of the HSPM: Constraints

- General Constraints
  - Wide-Coverage
  - Accuracy
  - Robustness
  - Multilinguality
- Architectural Constraints
  - Incrementality
  - Non-projectivity
  - Number of stages/Parallel activation

A Wide-Coverage Model of the HSPM: Constraints (2)

- Informational Constraints
  - Lexical Preferences
  - Discourse Structure
  - Semantic Plausibility
  - Prosody
  - Syntactic Structure

The Shift-Reduce Dependency Parser Revisited

t1 t2 t3 t4 ...

Input

s1

...

Stack

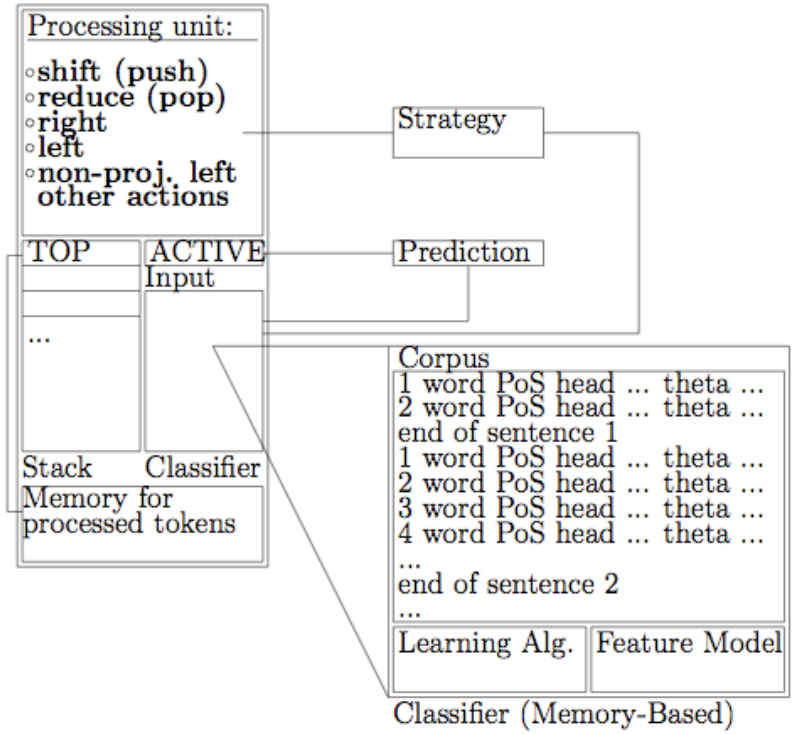
Actions:

- shift: push t1 on the top of the stack
- left: make s1 head of t1, (remove t1, pop s1 back to the input)
- right: make t1 head of s1, (pop s1)
- reduce: pop s1 from the stack

...

The Architecture

- Processing Unit
- Input
- Stack
- Memory for Processed Tokens
- Classifier
  - Linguistic Database
  - Learning Algorithm
  - Feature Model



Linking Hypotheses

- The Syntactic Prediction Locality Theory (SPLT), Gibson'98

  Human difficulty depends on the number of expected syntactic heads (the stack size) and the distance between them and their dependents.

- The Surprisal Model, Hale'01; Levy'05

  Human difficulty is proportional to the surprisal associated with the integration of the next word into the temporary syntactic structure

- Can locality effects be explained with surprisal?

Automatic Parsing and Emulations of Experiments with Human Subjects

Automatic Parsing

Lexical Models:

I have used various lexical models together with two shift-reduce parsers. The results are outlined in the table below.

Measure used for evaluation: Labeled Attachment Score (LAS) (the number of correctly assigned dependency relations divided by all the dependency relations in the evaluation set)

Language	Arabic	Bulgarian	English	Greek	Italian TUT	Italian ISST	Slovene
LAS	67.4%*	81.8%	81.2%†	76.6%†	83.7%	81.7%†	68.2%*

* A treebank transformation was applied to coordinated structures

† These results are from the parameter optimization session at the CoNLL 2007 shared task in dependency parsing

Automatic Parsing (2)

Integration of Semantics

I have integrated semantic theta roles and WordNet senses into parsing models.

Treebank	Baseline Model (LAS)	Semantic Model (LAS)
Turin University Treebank (it)	77.7%	81.3%
Boston Radio News Corpus *	77.1%	81.2%
Penn Treebank †	81.2%	82.7%

* I annotated the Boston Radio News Corpus with dependency syntax and semantic tags
† The Penn treebank has been annotated automatically with semantic tags.

Automatic Parsing (3)

- Models enriched with prosodic information

Treebank	Baseline Model (LAS)	Prosodic Model (LAS)
Boston Radio News Corpus (annotated with syntactic dependencies)	75.3%	75.2%

- Models enriched with constituency syntax

  An increase of accuracy has been achieved for Bulgarian (the BulTreeBank) using gold standard constituency labels in the learning model

Emulations of Experiments with Human Subjects

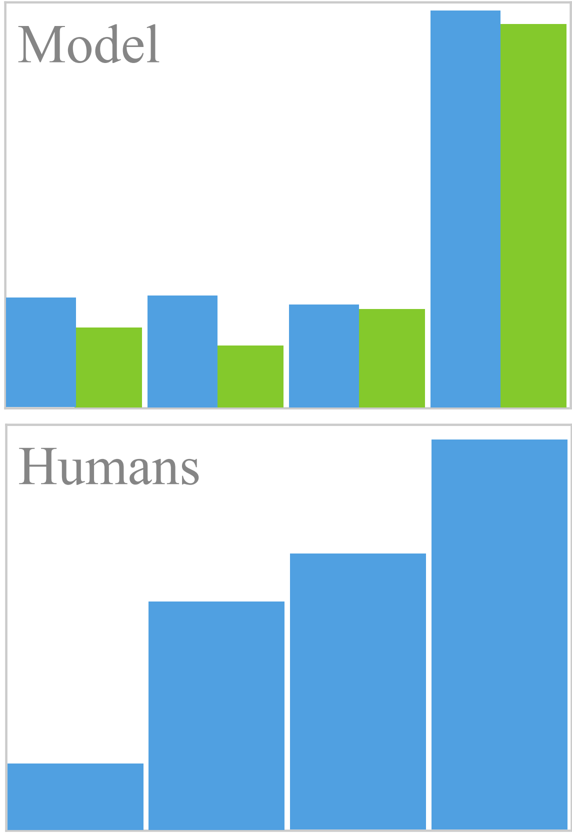
Parsing Accuracy on Text Segments and Surprisal:

For small sentences the surprisal would be proportional to 1 minus the probability of the sentence estimated by its parsing accuracy (LAS).

The concept can be generalized for larger text segments.

Chen et al., '05 (Exp. 1)

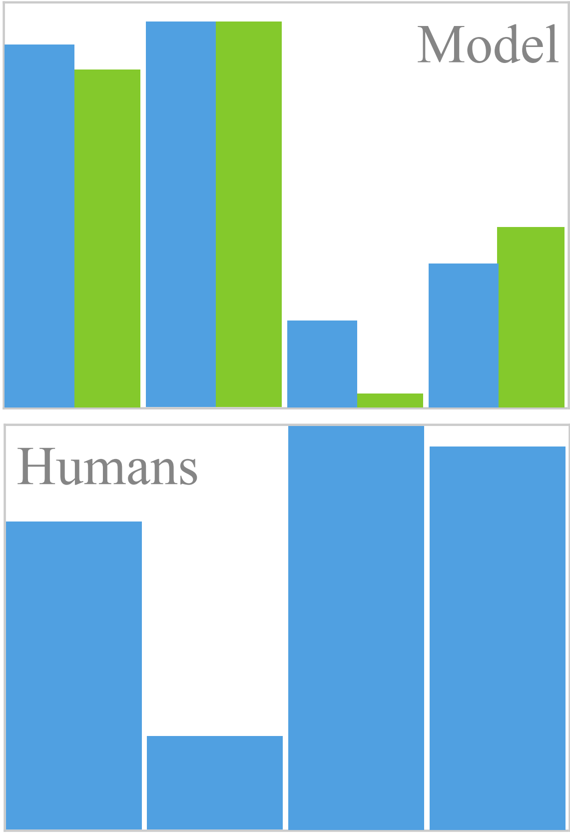
1. The detective suspected that the thief knew that **the guard protected the jewels** and so he reported immediately to the museum curator.
2. The detective suspected that the knowledge that **the guard protected the jewels** worried the museum curator.
3. The suspicion that the thief knew that **the guard protected the jewels** worried the museum curator.
4. The suspicion that the knowledge that **the guard protected the jewels** came from an insider worried the museum curator.



Notes: The moving window paradigm has been used. The critical region is in bold. The "surprisal" is proportional to 1 minus the parsing accuracy on the underlined text segment. Two parsing models: baseline and semantic (in green).

Ni et al., '98 (Exp. 1)

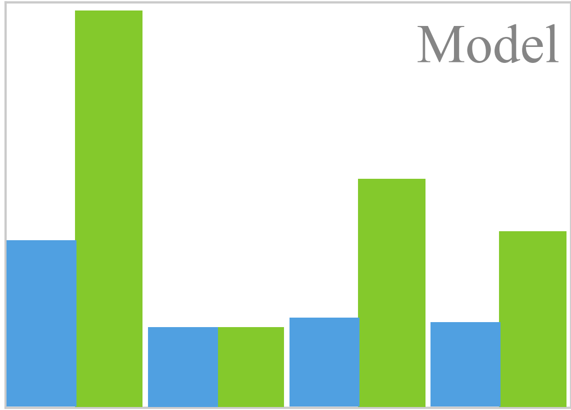
1. The businessmen loaned money at low interest **were told** to record their expenses.
2. Only businessmen loaned money at low interest **were told** to record their expenses.
3. The wealthy businessmen loaned money at low interest **were told** to record their expenses.
4. Only wealthy businessmen loaned money at low interest **were told** to record their expenses.

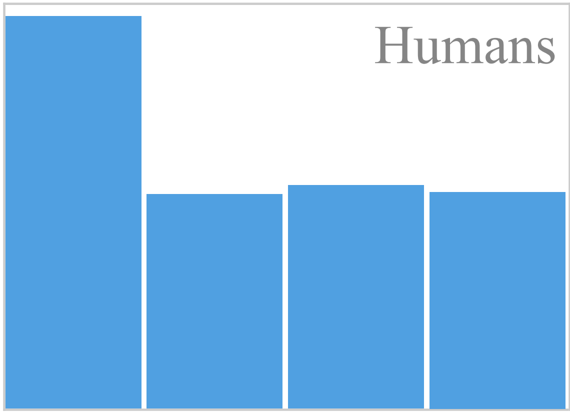


Notes: The moving window paradigm has been used. The critical region is in bold. The "surprisal" is proportional to 1 minus the parsing accuracy on the underlined text segment. Two parsing models: baseline and semantic (in green).

Trueswell & al. '94 (Exp. 1)

1. The defendant examined **by the lawyer** turned out to be unreliable.
2. The evidence examined **by the lawyer** turned out to be unreliable.
3. The defendant that was examined **by the lawyer** turned out to be unreliable.
4. The evidence that was examined **by the lawyer** turned out to be unreliable.





Notes: The eye-movement paradigm has been used. Only first pass reading times are considered. The critical region is in bold. The "surprisal" is proportional to 1 minus the parsing accuracy on the underlined text segment. Two parsing models: baseline and semantic (in green).

Gibson & Breen '05 (Exp. 2)

1. In the last scene, a character appeared **who was wounded** in the battle and the heroine wept when she saw him.
2. In the last scene, a character died **who was wounded** in the battle and the heroine wept when she saw him.
3. In the last scene, a character **who was wounded** in the battle appeared and the heroine wept when she saw him.
4. In the last scene, a character **who was wounded** in the battle died and the heroine wept when she saw him.

Notes: The moving window paradigm has been used. The critical region is in bold. The "surprisal" is proportional to 1 minus the parsing accuracy on the underlined text segment. Two parsing models: baseline and semantic (in green).

Discussion

- Comparison to other Automatic Parsers
  - Reasonable accuracy using lexicalized parsing models
  - Accuracy improvement using enriched parsing models
- Comparison to other Models of the Human Parser
  - Our model is serial and has wide coverage
  - It is constraint-based and can be trained using a neural network
  - It satisfies the constraints for a wide coverage model of the human parser and can be associated to multiple linking hypotheses

Conclusion

- I have prepared a list of constraints for wide coverage models of the human sentence processing mechanism
- I have proposed a wide-coverage model of the human parser that satisfies these constraints
- I have associated my architecture of the human parser to two linking hypotheses: the memory load component of SPLT and "surprisal"
- I have made a contribution to automatic parsing
- I have emulated human sentence processing

Thanks:

- Paolo Bouquet
- Alberto Lavelli
- Edward Gibson
- Kiril Simov
- Joakim Nivre
- Giuseppe Attardi
- Connexor, Finland

Thank you!