LSA Colloquium, SDSU April 16, 2005

Semantic Role Labeling

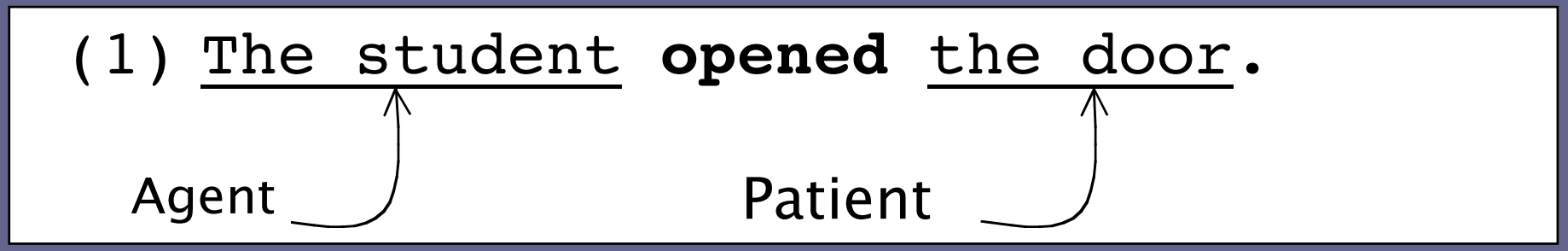
collaborative work by: Lucien Carroll (SDSU), Rebecca Colavin (SDSU), Hannah Rohde (UCSD), and Lara Taylor (UCSD)

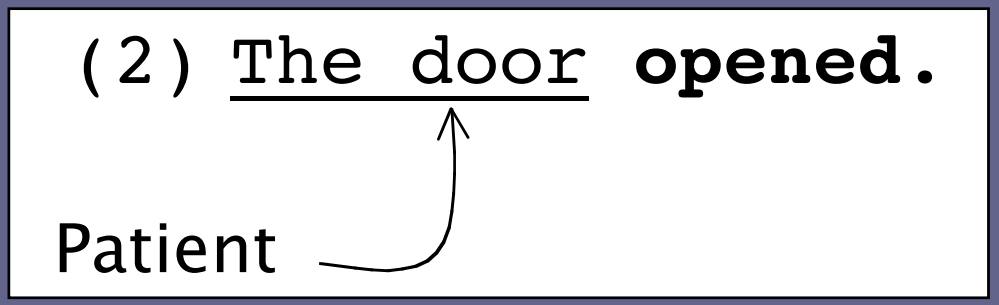
supervised by: Rob Malouf

The Problem

Computational task of semantic role labeling

- Identify verbal arguments
- Label with semantic role: Agent/Patient



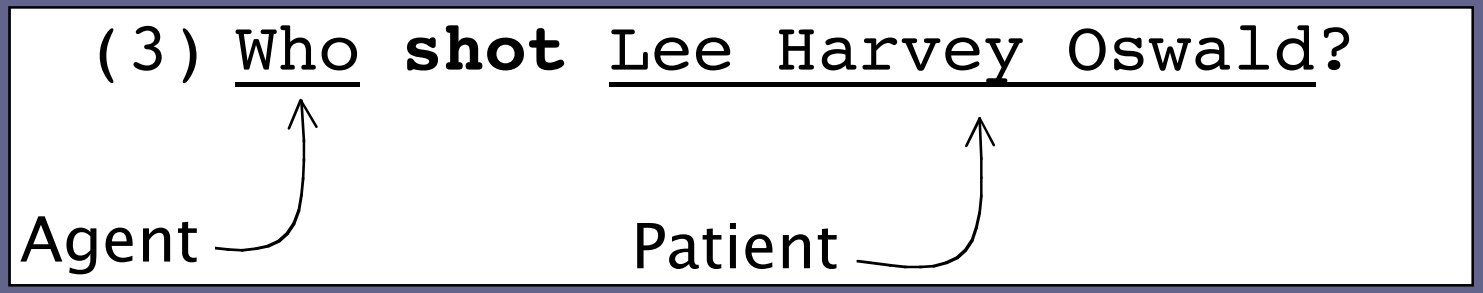


→ not all subjects are Agents

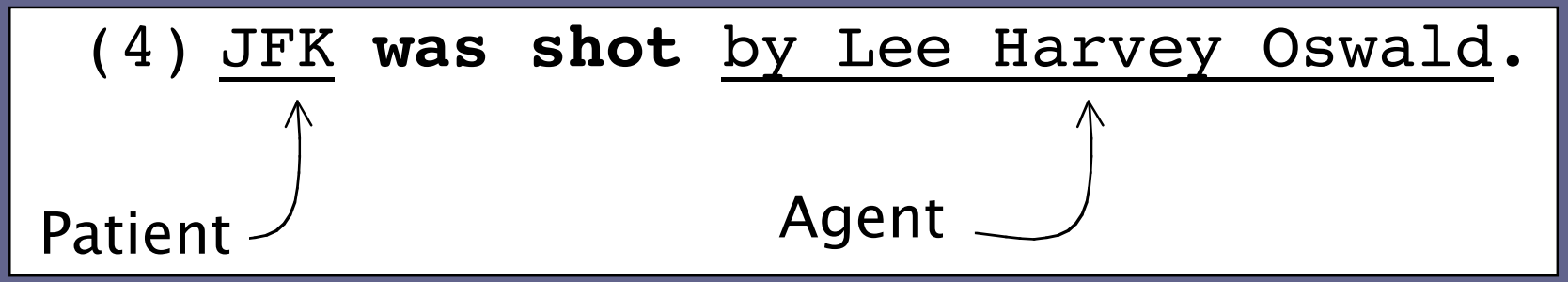
Relevance of General Task

- Why is semantic role labeling useful?

  Automatic question-answering systems have problems with:



  - Typical answer : JFK



  - Only use statistical co-occurrence
  - No information about semantic roles

Previous Approaches

COMPUTATIONAL:

- Syntax & word-order (Gildea & Palmer 2002, and others)
- Lexical information about verb (Gildea & Jurafsky 2002)
- Named entity recognition (Pradhan et al. 2004)

THEORETICAL: Dowty (1991)

- Roles as clusters of proto-properties

  (ex. prototypical agent has more 'agent-like' properties)

- Verbal Argument Selection Principle (1991: 576)

  (ex. given sentence with 2 arguments, argument with more 'agent-like' properties will be labeled the 'agent')

Our Question

BIG PICTURE: How can we use Dowty's (1991) theory of proto-properties for the computational task of semantic role labeling?

FIRST STEP: Use inherent semantic properties of head word to label arguments as Agent/Patient

Our Proposal

TRAIN:

- Use training data to learn which properties are predictive of which role labels

TEST:

- Use these properties to predict role labels for unseen test data

RESULTS:

- Evaluate our predictions against a hand-labeled answer key
- Compare our performance with that of a syntactic baseline on the same data

Training Phase

Given: 900 sentences of newspaper text

- Parser-selected arguments/heads
  Parsers : `Charniak, Collins, RASP`
  (finds arguments 93%)
- Dictionary of words
  Dictionary : `General Inquirer`
  with ~8000 words, 166 properties
  (coverage 78% of heads)

- Train
  - learn best predictors of each label by calculating conditional probabilities

Predictive Properties: Agent

Property	Example Word	Dowty proto-agent
collective	‘government’	volition
human	‘student’	sentience
kin	‘uncle’	perception
male	‘father’	causer
name	‘Italy’	movement
non-adult	‘child’	
perception	‘mistrust’	
political	‘queen’	
power	‘president’	
pronoun	‘it’	
role	‘expert’	

Predictive Properties: Patient

Property	Example Word	Dowty proto-patient
abstract	‘accuracy’	stationary
animal	‘fish’	causally affected
building part	‘kitchen’	incremental theme
activity	‘task’	undergoes change of state
communication	‘media’	
completion	‘defeat’	
common object	‘television’	
means to goal	‘payment’	
object	‘pistol’	
ordinal	‘eighth’	
solve	‘project’	
tool	‘fork’	

Testing Phase

Given: 100 test sentences and same tools

Further elimination: arguments Baseline doesn't label

- Test
  - given a verb and its arguments:
    1. find properties for each argument head
    2. for each arg, calculate how likely each label is
    3. pick argument with "strongest" preference for a label and assign it that label
    4. reduce set of remaining labels & remove labeled arg from consideration
    5. repeat 1-4 until no arguments remain

Results

100 arguments attempted:

- Random: precision 33%
- Baseline: precision 88%
- Semantic Labeler: precision 81%

BASELINE	SEMANTICS: Right	SEMANTICS: Wrong
Right	72	16
Wrong	9	3

Error Analysis I

Errors from learned semantic properties:

- Non-prototypical agents
- Lack of pronoun resolution

Error Analysis II

Non-prototypical agents:

- Personification

(5) …another machine is **doing** the work.

- Multiple Roles

(6) Applications won't have to be **rewritten** to **work** with Task Broker…

Error Analysis III

Lack of pronoun resolution:

(7) The program gets the $task_i$ and **splits** $\underline{it}_i$ up into parts…

Things we get right!

Where semantics wins over word order

☺ semantic properties ⇒ patient

(where Baseline depends on word order and mistakenly chooses agent)

- 'get' passives

(8) When elephants start fighting, ants get **killed**…

- unaccusative verbs

(9) If the battle **continues** much longer…

Conclusions & Future Research

- Predicted semantic roles (81% accuracy) using properties that overlap with Dowty's proto-properties
- Tension between syntax and semantics
- Inherent properties –
  now need event-level information (e.g. movement with respect to other participant)

(10) a. The batter hit the baseball.
     b. The baseball hit the batter.

- Larger project: implementation using conditional random fields, allows us to label all arguments simultaneously

References

Briscoe, E. and J. Carroll. 2002. Robust accurate statistical annotation of general text. In Proceedings of the 3rd Intl Conference on Language Resources and Evaluation: pp. 1499–1504.

Carreras, X. and Màrquez, L. 2004. Introduction to the CoNLL-2004 Shared Task: Semantic Role Labeling. In Proceedings of the 8th Conference on Natural Language Learning.

Charniak, Eugene. 2000. A Maximum-Entropy-Inspired Parser. In Proceedings of NAACL-2000.

Collins, Michael. 1996. A New Statistical Parser Based on Bigram Lexical Dependencies. In Proceedings of the 34th Annual Meeting of the ACL.

Dowty, David. 1991. Thematic proto-roles and argument selection. Language 67.3:547–619.

Gildea, Daniel & Daniel Jurafsky. 2002. Automatic Labeling of Semantic Roles. Computational Linguistics, 28(3):245-288,2002

Gildea, Daniel & Martha Palmer (2002). The Necessity of Parsing for Predicate Argument Recognition. In Proceedings of ACL 2002, Philadelphia, USA, 2002.

Pradhan, S., Ward, W., Hacioglu, K., Martin, J., Jurafsky, D., 2004. Shallow Semantic Parsing using Support Vector Machines. In Proceedings of HLT/NAACL.