CS114: Semantic Roles

March 17, 2015

James Pustejovsky

Brandeis University

Thanks for Jurafsky & Martin & Prof. Meteer for additional slides

What are semantic roles and what is their history?

- A lot of forms of traditional grammar (Sanskrit, Japanese, …) analyze in terms of a rich array of semantically potent case ending or particles
  - They're kind of like semantic roles
- The idea resurfaces in modern generative grammar in work of Charles ("Chuck") Fillmore, who calls them Case Roles (Fillmore, 1968, The Case for Case).
  - They're quickly renamed to other words, but various:
    - Semantic roles
    - Thematic roles
    - Theta roles
    - A predicate and its semantic roles are often taken together as an **argument structure**

Okay, but what *are* they?

- An event is expressed by a predicate and various other dependents
- The claim of a theory of semantic roles is that these other dependents can be usefully classified into a small set of semantically contentful classes
  - And that these classes are useful for explaining lots of things

Common semantic roles

- Agent: initiator or doer in the event
  - **Sue** killed **the rat**.
- Patient: affected entity in the event; undergoes the action
- Theme: object in the event undergoing a change of state or location, or of which location is predicated
  - **The ice** melted
- Experiencer: feels or perceive the event
  - **Bill** likes **pizza**.
- Stimulus: the thing that is felt or perceived

Slide from Chris Manning

Common semantic roles

- **Goal**:
  - Bill ran to Copley Square.
- **Recipient** (may or may not be distinguished from Goal):
  - Bill gave the book to Mary.
- **Benefactive** (may be grouped with Recipient):
  - Bill cooked dinner for Mary.
- **Source**:
  - Bill took a pencil from the pile.
- **Instrument**:
  - Bill ate the burrito with a plastic spork.
- **Location**:
  - Bill sits under the tree on Wednesdays

Common semantic roles

Try for yourself!

1. The submarine sank a troop ship.
2. Doris hid the money in the flowerpot.
3. Emma noticed the stain.
4. We crossed the street.
5. The boys climbed the wall.
6. The chef cooked a great meal.
7. The computer pinpointed the error.
8. A mad bull damaged the fence on Jack's farm.
9. The company wrote me a letter.
10. Jack opened the lock with a paper clip.

Linking of thematic roles to syntactic positions

- John opened the door
- AGENT THEME
- The door was opened by John
- THEME AGENT
- The door opened
- THEME
- John opened the door with the key
- AGENT THEME INSTRUMENT

Deeper Semantics

- From the WSJ…
  - He melted her reserve with a husky-voiced paean to her eyes.
  - If we label the constituents He and her reserve as the Melter and Melted, then those labels lose any meaning they might have had.
  - If we make them Agent and Theme then we can do more inference.

Problems

- What exactly is a role?
- What's the right set of roles?
- Are such roles universals?
- Are these roles atomic?
  - I.e. Agents
    - Animate, Volitional, Direct causers, etc
- Can we automatically label syntactic constituents with thematic roles?

Syntactic Variations

Yesterday, Kristina hit *Scott with a baseball*

Scott was hit by *Kristina yesterday with a baseball*

Yesterday, Scott was hit *with a baseball* by *Kristina*

With a baseball, Kristina hit *Scott yesterday*

Yesterday Scott was hit by *Kristina with a baseball*

Kristina hit *Scott with a baseball yesterday*

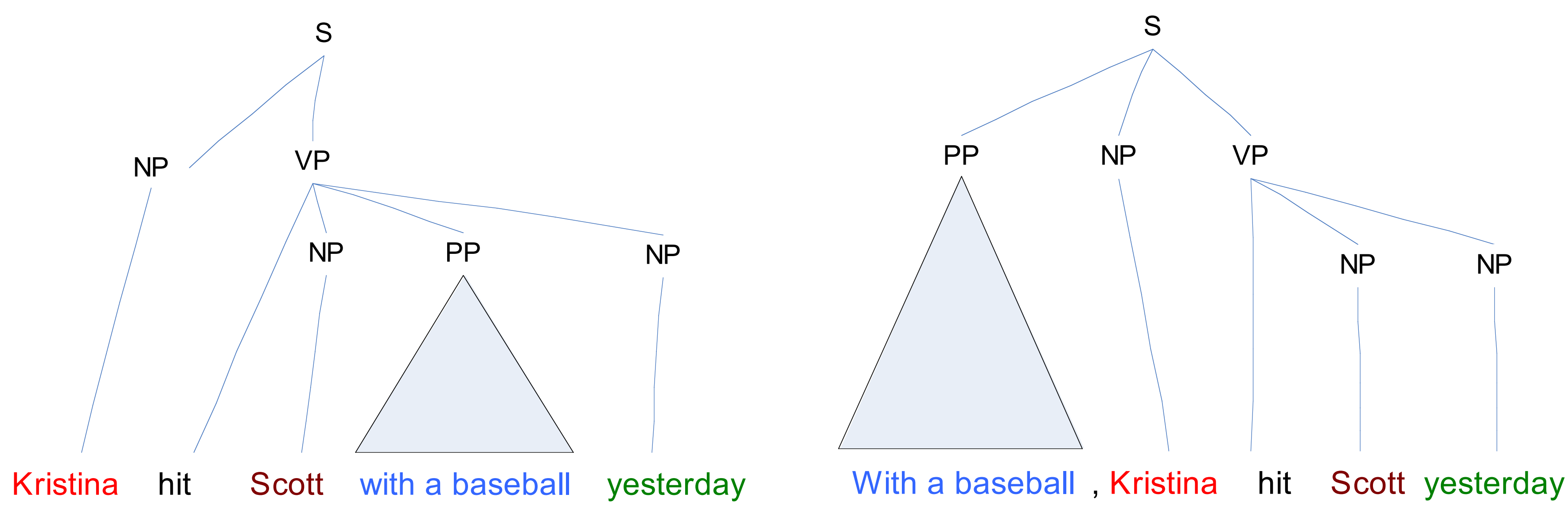
Agent, hitter

Thing hit

Instrument

Temporal adjunct

Syntactic Variations (as trees)



Semantic Role Labeling – *Giving Semantic Labels to Phrases*

- [$_{\text{AGENT}}$ John] **broke** [$_{\text{THEME}}$ the window]
- [$_{\text{THEME}}$ The window] **broke**
- [$_{\text{AGENT}}$Sotheby's] .. **offered** [$_{\text{RECIPIENT}}$ the Dorrance heirs] [$_{\text{THEME}}$ a money-back guarantee]
- [$_{\text{AGENT}}$ Sotheby's] **offered** [$_{\text{THEME}}$ a money-back guarantee] to [$_{\text{RECIPIENT}}$ the Dorrance heirs]
- [$_{\text{THEME}}$ a money-back guarantee] **offered** by [$_{\text{AGENT}}$ Sotheby's]
- [$_{\text{RECIPIENT}}$ the Dorrance heirs] will [$_{\text{ARM-NEG}}$ not] be **offered** [$_{\text{THEME}}$ a money-back guarantee]

Why is SRL Important – *Applications*

- Question Answering
  - Q: When was Napoleon defeated?
  - Look for: [PATIENT Napoleon] [PRED **defeat-synset**] [ARGM-TMP *ANS*]

- Machine Translation

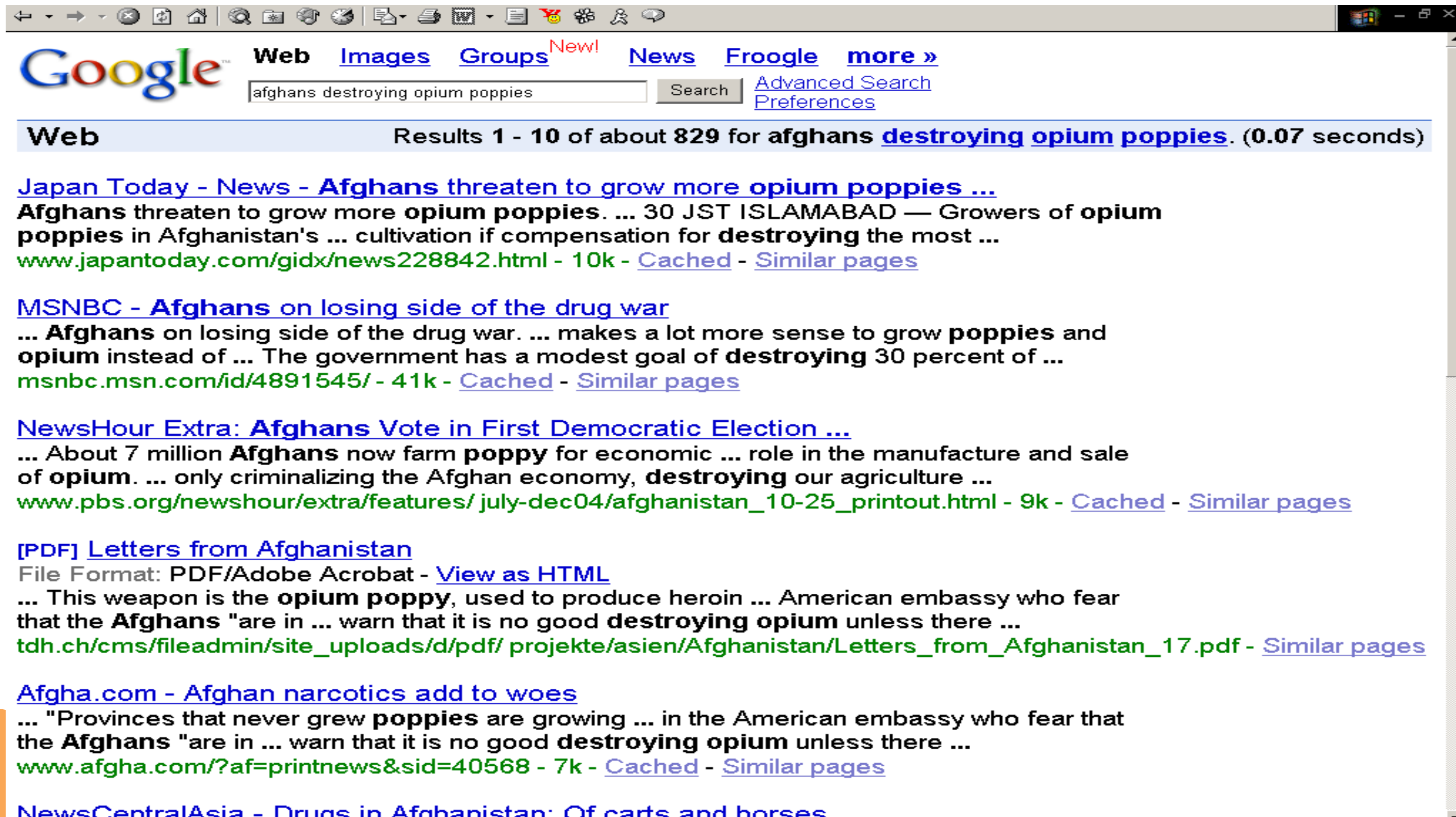
English (SVO)	Farsi (SOV)	
[AGENT The little boy]	[AGENT pesar koocholo]	boy-little
[PRED **kicked**]	[THEME toop germezi]	ball-red
[THEME the red ball]	[ARGM-MNR moqtam]	hard-adverb
[ARGM-MNR hard]	[PRED **zaad-e**]	hit-past

- Document Summarization
  - Predicates and Heads of Roles summarize content

- Information Extraction
  - SRL can be used to construct useful rules for IE

Application: Semantically precise search

Query: *afghans destroying opium poppies*



Some History

- Minsky 74, Fillmore 1976: *frames* describe events or situations
  - Multiple participants, “props”, and “conceptual roles”
- Levin 1993: verb class defined by sets of frames (meaning-preserving alternations) a verb appears in
  - {*break,shatter*,..}: *Glass X’s easily; John Xed the glass, …*
  - *Cut* is different: *The window broke; *The window cut.*
- FrameNet, late ’90s: based on Levin’s work: large corpus of sentences annotated with *frames*
- PropBank: addresses tragic flaw in FrameNet corpus

Alternations for verbs of contact:

conative:

Jean moved the table.
*Jean moved at the table.

body-part possessor ascension:

Janet broke Bill's finger.
*Janet broke Bill on the finger.

middle construction:

Bread cuts easily.
*Cats touch easily.

Verb Class

Alternation	Touch	Hit	Cut	Break
conative	N	Y	Y	N
body-part possessor ascension	Y	Y	Y	N
middle	N	N	Y	Y

Examples of verbs for each class:

Touch:	kiss, sting, tickle
Hit:	bash, hammer, tap
Cut:	chip, hack, scratch
Break:	hack, split, tear

Underlying hypothesis: *verbal meaning determines syntactic realizations*

Beth Levin analyzed thousands of verbs and defined hundreds of classes.

Frames in FrameNet

frame(TRANSPORTATION) frame_elements(MOVER(S), MEANS, PATH) scene(MOVER(S) move along PATH by MEANS)
frame(DRIVING) inherit(TRANSPORTATION) frame_elements(DRIVER (=MOVER), VEHICLE (=MEANS), RIDER(S) (=MOVER(S)), CARGO (=MOVER(S))) scenes(DRIVER starts VEHICLE, DRIVER controls VEHICLE, DRIVER stops VEHICLE)
frame(RIDING_1) inherit(TRANSPORTATION) frame_elements(RIDER(S) (=MOVER(S)), VEHICLE (=MEANS)) scenes(RIDER enters VEHICLE, VEHICLE carries RIDER along PATH, RIDER leaves VEHICLE)

Figure 1: A subframe can inherit elements an semantics from its parent

[Baker, Fillmore, Lowe, 1998]

FEG	Annotated Example from BNC
D	[$_D$ Kate] **drove** [$_P$ home] in a stupor.
V, D	A pregnant woman lost her baby after she fainted as she waited for a bus and fell into the path of [$_V$ a lorry] **driven** [$_D$ by her uncle].
D, P	And that was why [$_D$ I] **drove** [$_P$ eastwards along Lake Geneva].
D, R, P	Now [$_D$ Van Cheele] was **driving** [$_R$ his guest] [$_P$ back to the station].
D, V, P	[$_D$ Cumming] had a fascination with most forms of transport, **driving** [$_V$ his Rolls] at high speed [$_P$ around the streets of London].
D+R, P	[$_D$ We] **drive** [$_P$ home along miles of empty freeway].
V, P	Over the next 4 days, [$_V$ the Rolls Royces] will **drive** [$_P$ down to Plymouth], following the route of the railway.

Figure 2: Examples of Frame Element Groups and Annotated Sentences

FrameNet [Fillmore et al. 01]

Frame: Hit_target
(hit, pick off, shoot)

Agent	Means
Target	Place
Instrument	Purpose
Manner	Subregion
	Time

Core

Non-Core

Lexical units (LUs):
Words that evoke the frame
(usually verbs)

Frame elements (FEs):
The involved semantic roles

[$_{\text{Agent}}$ *Kristina*] ***hit*** [$_{\text{Target}}$ *Scott*] [$_{\text{Instrument}}$ *with a baseball*] [$_{\text{Time}}$ *yesterday*].

Methodology for FrameNet

1. Define a frame (eg DRIVING)
2. Find some sentences for that frame
3. Annotate them
4. If (remaining funding == 0) then exit;   else goto step 1.

- Corpora
  - FrameNet I – British National Corpus only
  - FrameNet II – LDC North American Newswire corpora
- Size
  - >8,900 lexical units, >625 frames, >135,000 sentences

http://framenet.icsi.berkeley.edu

Annotations in PropBank

- Based on Penn TreeBank
- Goal is to annotate *every tree* systematically
  - so statistics in the corpus are meaningful
- Like FrameNet, based on Levin's verb classes (via VerbNet)
- Generally more data-driven & bottom up
  - No level of abstraction beyond verb senses
  - Annotate every verb you see, whether or not it seems to be part of a frame

Some verb senses and "framesets" for propbank

Frameset: **decline.01** "go down incrementally"
Arg1: entity going down
Arg2: amount gone down by, EXT
Arg3: start point
Arg4: end point
Ex: ...[$_{\text{Arg1}}$ its net income] *declining* [$_{\text{Arg2-EXT}}$ 42%] [$_{\text{Arg4}}$ to $121 million] [$_{\text{ArgM-TMP}}$ in the first 9 months of 1989]. (wsj_0067)

Frameset: **decline.02** "demure, reject"
Arg0: agent
Arg1: rejected thing
Ex: [$_{\text{Arg0}}$ A spokesman$_i$] *declined* [$_{\text{Arg1}}$ *trace*$_i$ to elaborate] (wsj_0038)

FrameNet vs PropBank -1

FRAMENET ANNOTATION:

[$_{\text{Buyer}}$ Chuck] *bought* [$_{\text{Goods}}$ a car] [$_{\text{Seller}}$ from Jerry] [$_{\text{Payment}}$ for $1000].

[$_{\text{Seller}}$ Jerry] *sold* [$_{\text{Goods}}$ a car] [$_{\text{Buyer}}$ to Chuck] [$_{\text{Payment}}$ for $1000].

PROPBANK ANNOTATION:

[$_{\text{Arg0}}$ Chuck] *bought* [$_{\text{Arg1}}$ a car] [$_{\text{Arg2}}$ from Jerry] [$_{\text{Arg3}}$ for $1000].

[$_{\text{Arg0}}$ Jerry] *sold* [$_{\text{Arg1}}$ a car] [$_{\text{Arg2}}$ to Chuck] [$_{\text{Arg3}}$ for $1000].

FrameNet vs PropBank -2

FRAMENET ANNOTATION:

[$_{\text{Goods}}$ A car] was *bought* [$_{\text{Buyer}}$ by Chuck].

[$_{\text{Goods}}$ A car] was *sold* [$_{\text{Buyer}}$ to Chuck] [$_{\text{Seller}}$ by Jerry].

[$_{\text{Buyer}}$ Chuck] was *sold* [$_{\text{Goods}}$ a car] [$_{\text{Seller}}$ by Jerry].

PROPBANK ANNOTATION:

[$_{\text{Arg1}}$ A car] was *bought* [$_{\text{Arg0}}$ by Chuck].

[$_{\text{Arg1}}$ A car] was *sold* [$_{\text{Arg2}}$ to Chuck] [$_{\text{Arg0}}$ by Jerry].

[$_{\text{Arg2}}$ Chuck] was *sold* [$_{\text{Arg1}}$ a car] [$_{\text{Arg0}}$ by Jerry].

Proposition Bank (PropBank) [Palmer et al. 05]

- Transfer sentences to propositions
  - Kristina hit Scott → hit(Kristina,Scott)
- Penn TreeBank → PropBank
  - Add a semantic layer on Penn TreeBank
  - Define a set of semantic roles for each verb
  - Each verb's roles are numbered

…[**A0** the company] to … *offer* [**A1** a 15% to 20% stake] [**A2** to the public]
…[**A0** Sotheby's] … *offered* [**A2** the Dorrance heirs] [**A1** a money-back guarantee]
…[**A1** an amendment] *offered* [**A0** by Rep. Peter DeFazio] …
…[**A2** Subcontractors] will be *offered* [**A1** a settlement] …

Proposition Bank (PropBank)
Define the Set of Semantic Roles

- It's difficult to define a general set of semantic roles for all types of predicates (verbs).
- PropBank defines semantic roles for each verb and sense in the frame files.
- The (core) arguments are labeled by numbers.
  - A0 – Agent; A1 – Patient or Theme
  - Other arguments – no consistent generalizations
- Adjunct-like arguments – ***universal*** to all verbs
  - AM-LOC, TMP, EXT, CAU, DIR, PNC, ADV, MNR, NEG, MOD, DIS

Proposition Bank (PropBank) Frame Files

- hit.01 “strike”
  - A0: agent, hitter; A1: thing hit; A2: instrument, thing hit by or with

[$_{A0}$ *Kristina*] ***hit*** [$_{A1}$ *Scott*] [$_{A2}$ *with a baseball*] *yesterday*.

AM-TMP
Time

- look.02 “seeming”
  - A0: seemer; A1: seemed like; A2: seemed to

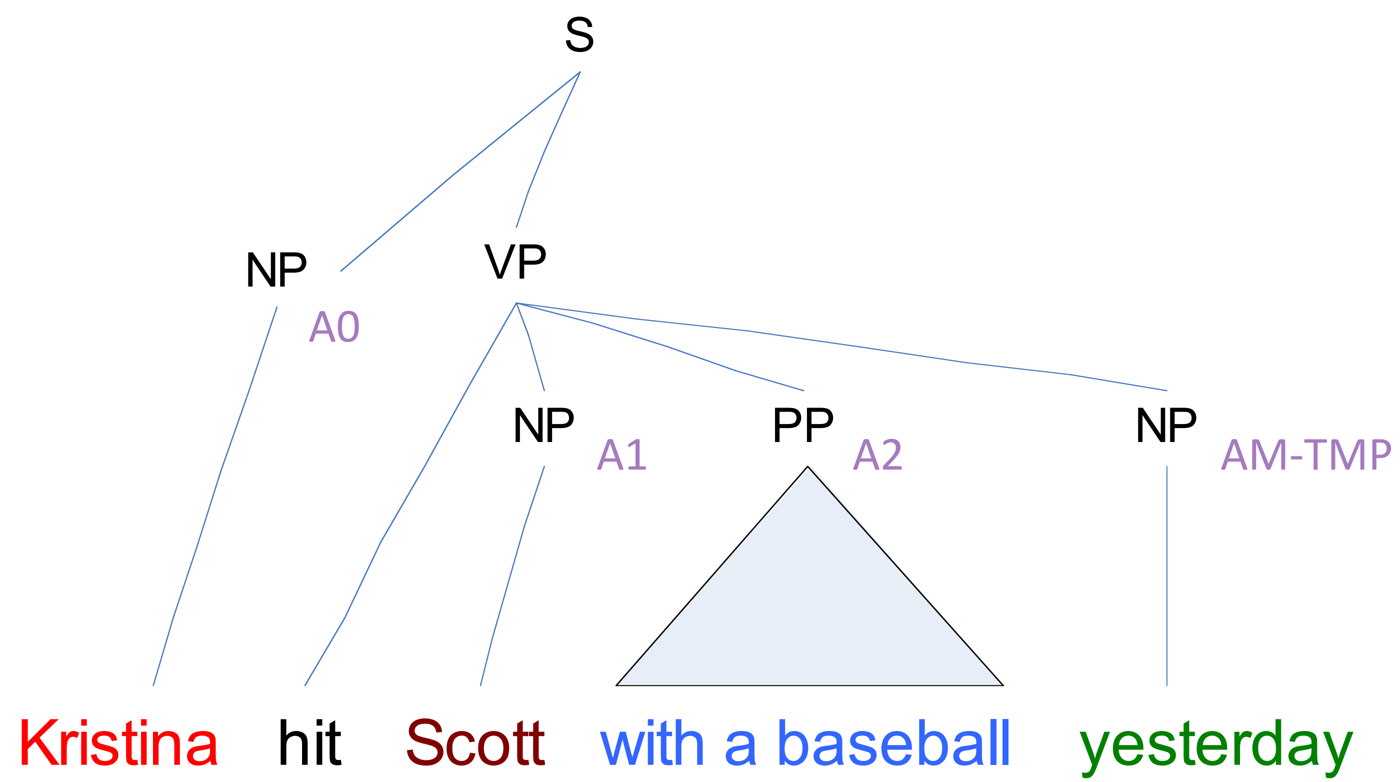
[$_{A0}$ *It*] ***looked*** [$_{A2}$ *to her] like* [$_{A1}$ *he deserved this*].

- deserve.01 “deserve”
  - A0: deserving entity; A1: thing deserved; A2: in-exchange-for

It looked to her like [$_{A0}$ *he*] ***deserved*** [$_{A1}$ *this*].

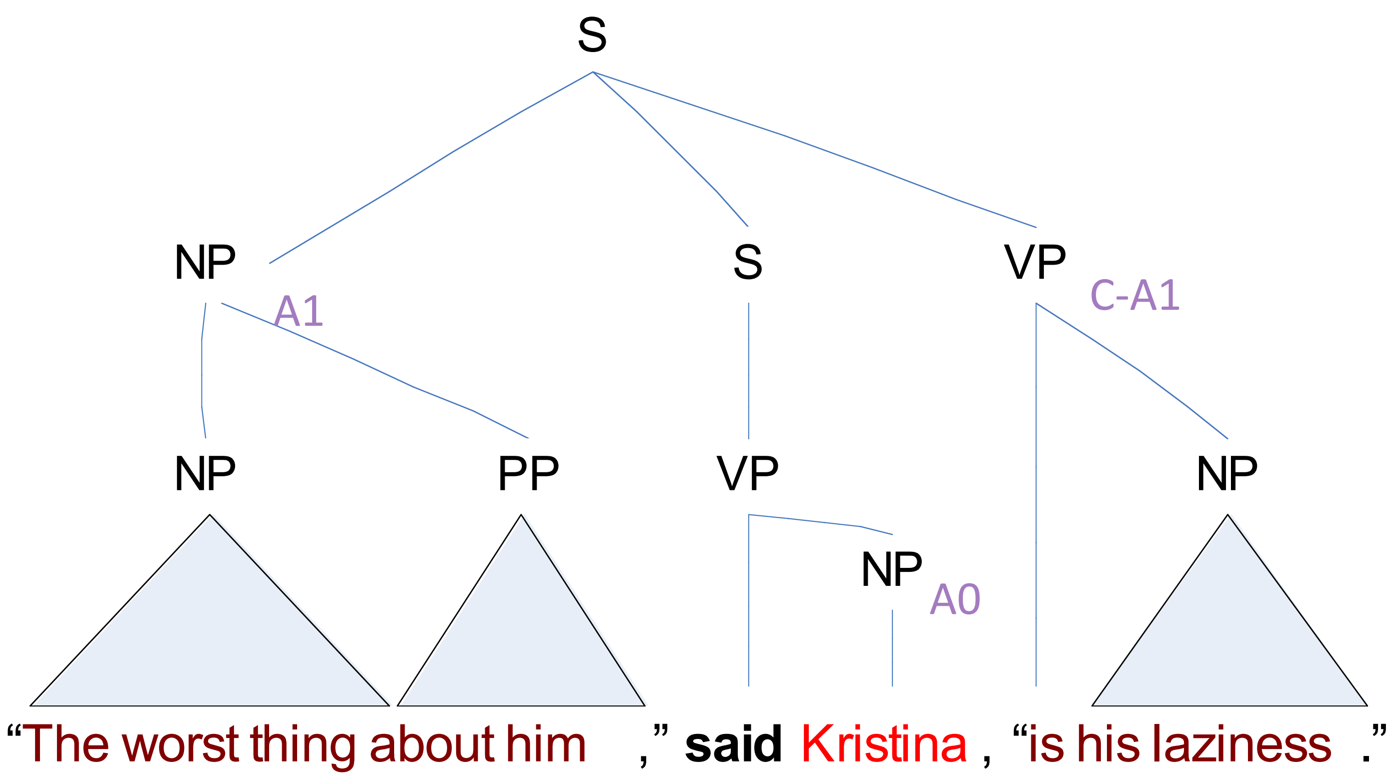
Proposition:
A sentence and
a target verb

Proposition Bank (PropBank) Add a Semantic Layer



[$_{\mathbf{A0}}$ *Kristina*] ***hit*** [$_{\mathbf{A1}}$ *Scott*] [$_{\mathbf{A2}}$ *with a baseball*] [$_{\mathbf{AM\text{-}TMP}}$ *yesterday*].

Proposition Bank (PropBank) Add a Semantic Layer – Continued



[$_{A1}$ *The worst thing about him*] ***said*** [$_{A0}$ *Kristina*] [$_{C\text{-}A1}$ *is his laziness*].

Proposition Bank (PropBank) Final Notes

- Current release (Mar 4, 2005): Proposition Bank I
  - Verb Lexicon: 3,324 frame files
  - Annotation: ~113,000 propositions

  http://verbs.colorado.edu/~mpalmer/projects/ace.html

- Alternative format: CoNLL-04,05 shared task
  - Represented in table format
  - Has been used as standard data set for the shared tasks on semantic role labeling

  http://www.lsi.upc.es/~srlconll/soft.html

1. faces(“the $1.4B robot spacecraft”, “a six-year journey to explore …moons”)
2. explore(“the $1.4B robot spacecraft”, “Jupiter and its 16 known moons”)

The	-	(A0*	(A0*
$	-	*	*
1.4	-	*	*
billion	-	*	*
robot	-	*	*
spacecraft	-	*)	*)
faces	face	(V*)	*
a	-	(A1*	*
six-year	-	*	*
journey	-	*	*
to	-	*	*
explore	explore	*	(V*)
Jupiter	-	*	(A1*
and	-	*	*
its	-	*	*
16	-	*	*
known	-	*	*
moons	-	*)	*)
.	-	*	*

1. lie("he",…)
2. leak("he", "information obtained from … he supervised")
3. obtain(X, "information", "from a wiretap he supervised")
4. supervise("he", "a wiretap")

He	-	(A0*)	(A0*)	*	*
is	-	*	*	*	*
also	-	*	*	*	*
accused	-	*	*	*	*
of	-	*	*	*	*
lying	lie	(V*)	*	*	*
under	-	(AM-LOC*	*	*	*
oath	-	*)	*	*	*
and	-	*	*	*	*
of	-	*	*	*	*
leaking	leak	*	(V*)	*	*
information	-	*	(A1*	(A1*)	*
obtained	obtain	*	*	(V*)	*
from	-	*	*	*	*
a	-	*	*	(A2*	(A1*
wiretap	-	*	*	*	*)
he	-	*	*	*	(A0*)
supervised	supervise	*	*)	*)	(V*)
.	-	*	*	*	*

Information Extraction versus Semantic Role Labeling

Characteristic	IE	SRL
Coverage	narrow	broad
Depth of semantics	shallow	shallow
Directly connected to application	sometimes	no

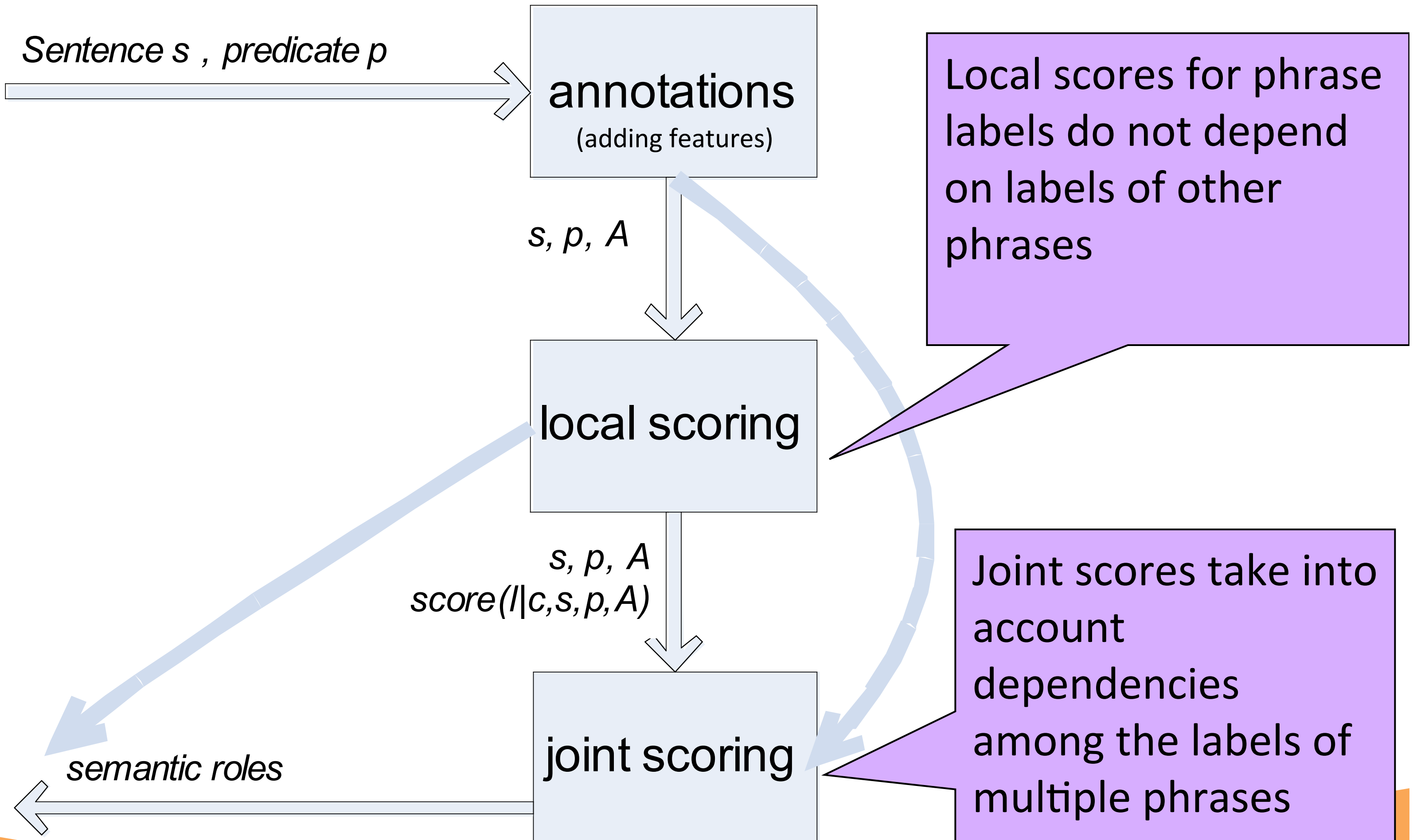
Evaluation Measures

Correct: [$_{A0}$ The queen] **broke** [$_{A1}$ the window] [$_{AM-TMP}$ yesterday]
Guess: [$_{A0}$ The queen] broke the [$_{A1}$ window] [$_{AM-LOC}$ yesterday]

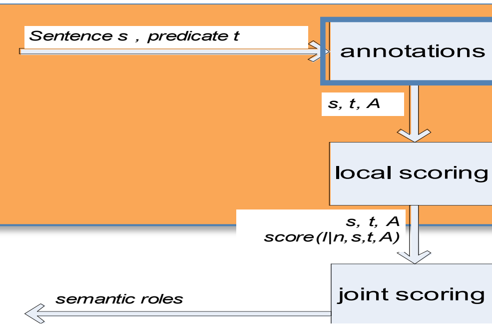
Correct	**Guess**
{The queen} →A0 {the window} →A1 {yesterday} ->AM-TMP all other → NONE	{The queen} →A0 {window} →A1 {yesterday} ->AM-LOC all other → NONE

- Precision ,Recall, F-Measure {*tp*=1,*fp*=2,*fn*=2} *p*=*r*=*f*=1/3
- Measures for subtasks
  - Identification (Precision, Recall, F-measure) {tp=2,fp=1,fn=1} *p*=*r*=*f*=2/3
  - Classification (Accuracy) acc = .5 (labeling of correctly identified phrases)
  - Core arguments (Precision, Recall, F-measure) {tp=1,fp=1,fn=1} *p*=*r*=*f*=1/2

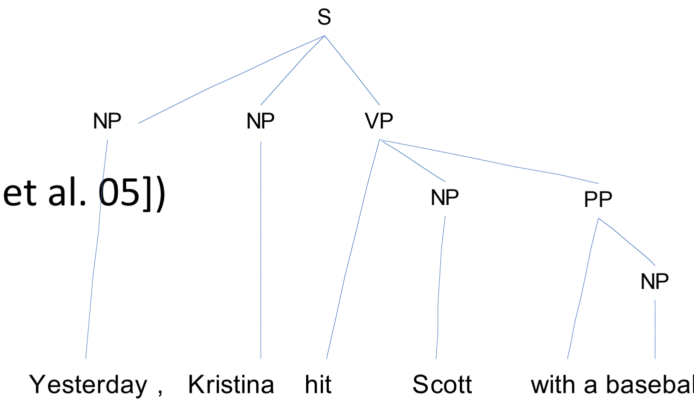
Basic Architecture of a Generic SRL System



Annotations Used



- Syntactic Parsers
  - Collins', Charniak's (most systems)
  - CCG parses ([Gildea & Hockenmaier 03],[Pradhan et al. 05])
  - TAG parses ([Chen & Rambow 03])



- Shallow parsers

  [$_{NP}$Yesterday] , [$_{NP}$Kristina] [$_{VP}$hit] [$_{NP}$Scott] [$_{PP}$with] [$_{NP}$a baseball].

- Semantic ontologies (WordNet, automatically derived), and named entity classes

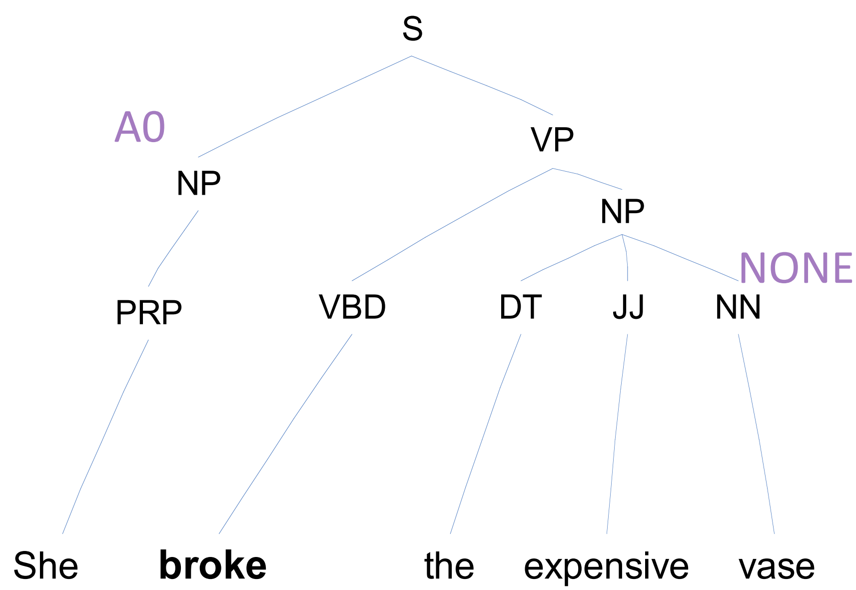
  (v) **hit** (cause to move by striking)

  WordNet hypernym

  propel, impel *(*cause to move forward with force*)*

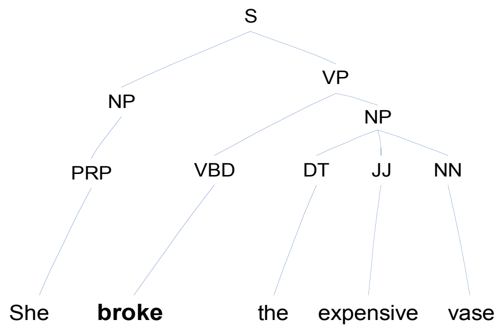
Labeling Parse Tree Nodes

- Given a parse tree *t*, label the nodes (phrases) in the tree with semantic labels
- To deal with discontiguous arguments
  - In a post-processing step, join some phrases using simple rules
  - Use a more powerful labeling scheme, i.e. C-A0 for continuation of A0

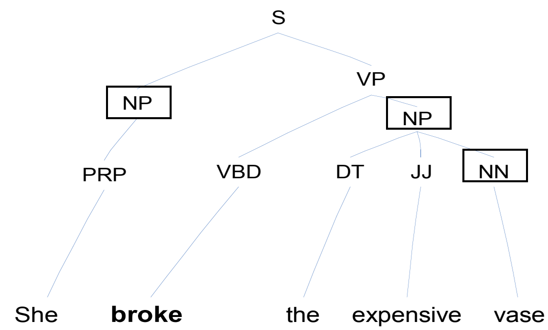


Another approach: labeling chunked sentences. Will not describe in this section.

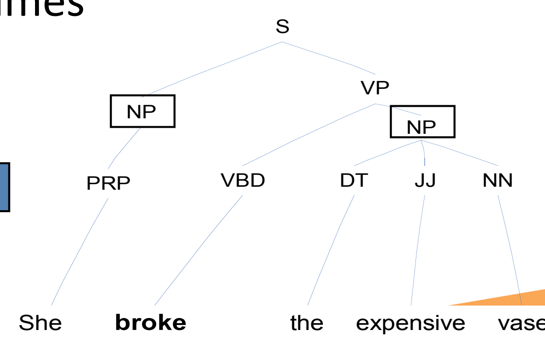
Combining Identification and Classification Models



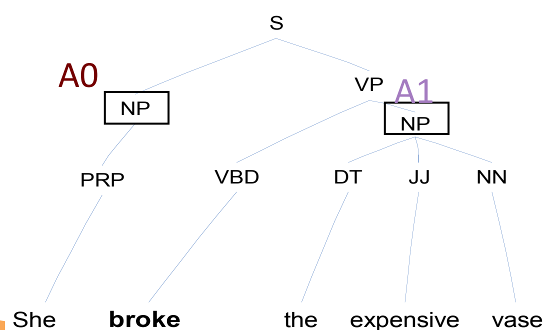
Step 1. *Pruning.* Using a hand-specified filter.



Step 2. *Identification.* Identification model (filters out candidates with high probability of NONE)



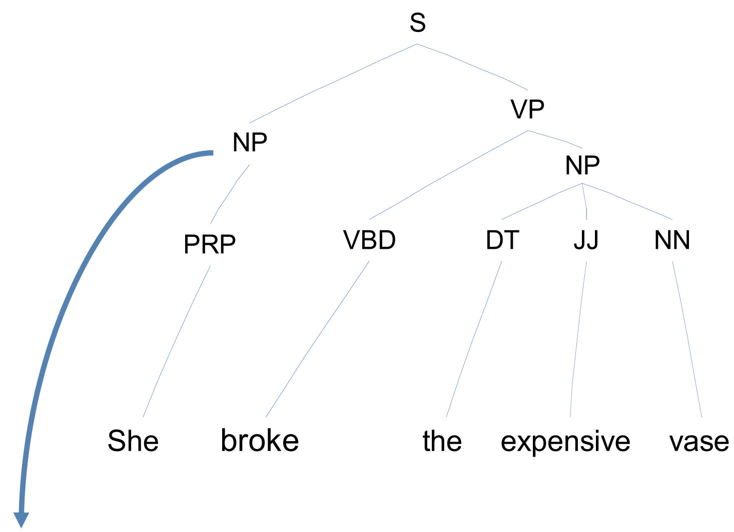
Step 3. *Classification.* Classification model assigns one of the argument labels to selected nodes (or sometimes possibly NONE)



Gildea & Jurafsky (2002) Features

- Key early work
  - Future systems use these features as a baseline

- Constituent Independent
  - Target predicate (lemma)
  - Voice
  - Subcategorization
- Constituent Specific
  - Path
  - Position (*left, right*)
  - Phrase Type
  - Governing Category (*S* or *VP*)
  - Head Word

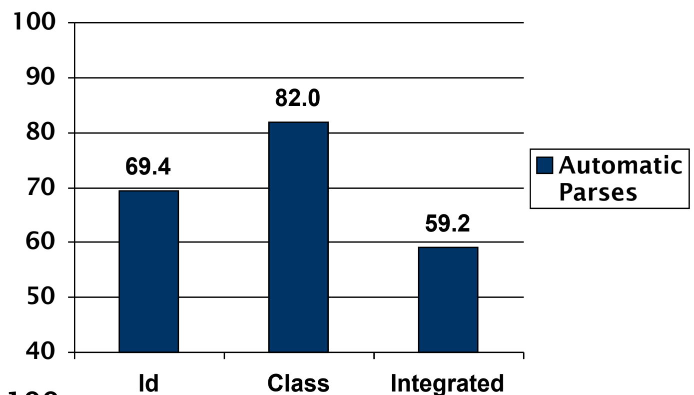


Target	*broke*
Voice	*active*
Subcategorization	*VP→VBD NP*
Path	*VBD↑VP↑S↓NP*
Position	*left*
Phrase Type	*NP*
Gov Cat	*S*
Head Word	*She*

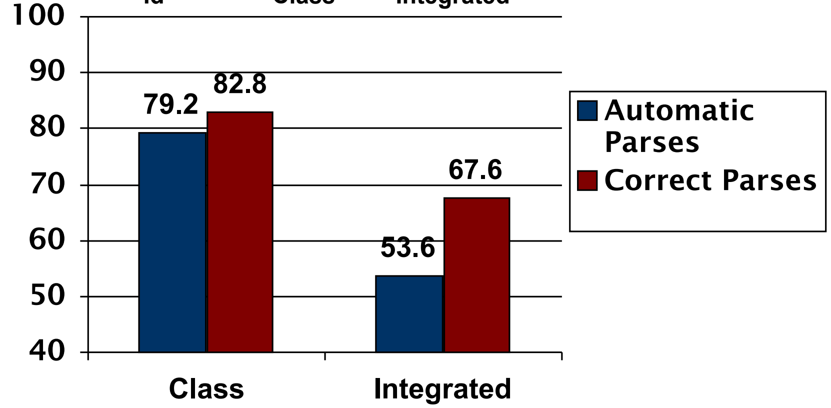
Performance with Baseline Features using the G&J Model

- **Machine learning algorithm:** interpolation of relative frequency estimates based on subsets of the 7 features introduced earlier

FrameNet Results



Propbank Results

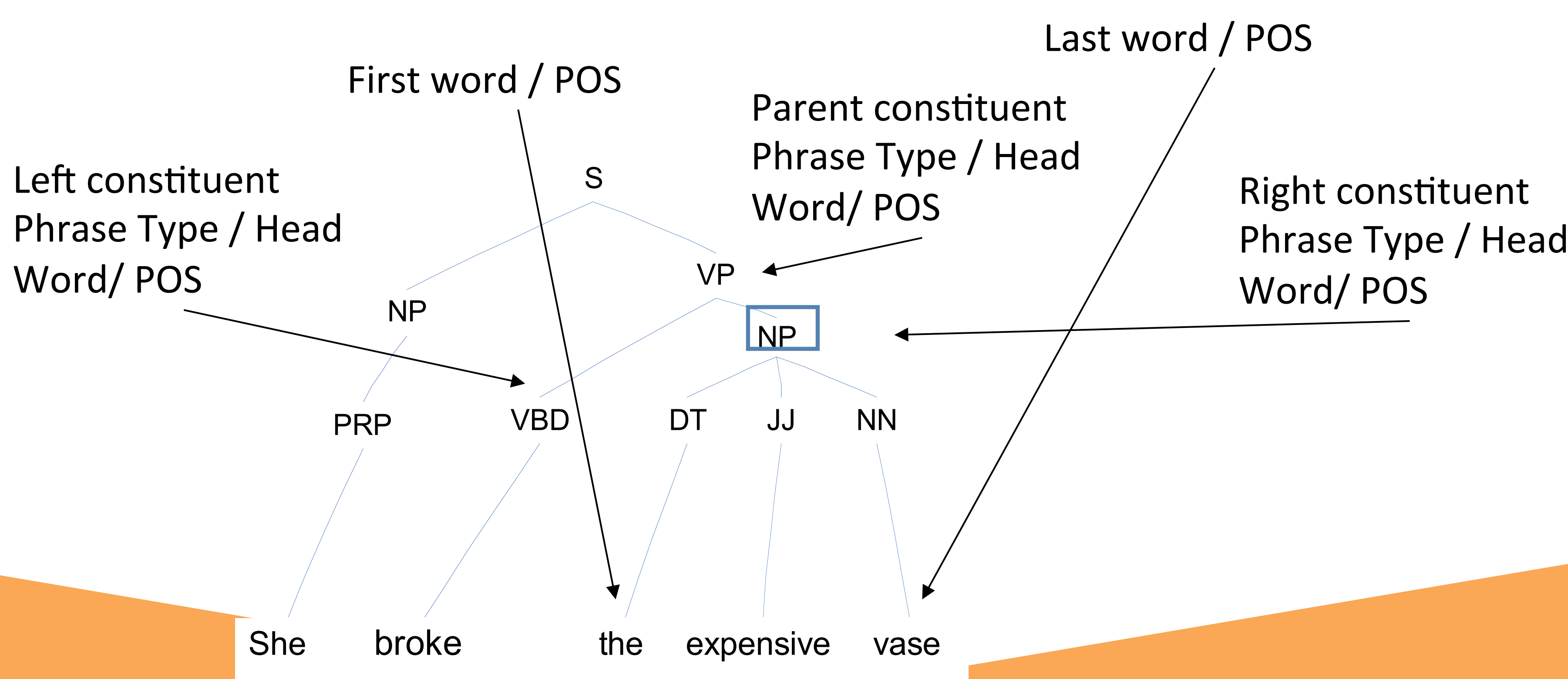


Performance with Baseline Features using the G&J Model

- Better ML: 67.6 → **80.8** using SVMs [Pradhan et al. 04]).
- Content Word (different from head word)
- Head Word and Content Word POS tags
- **NE labels (Organization, Location, etc.)**
- Structural/lexical context (phrase/words around parse tree)
- Head of PP Parent
  - If the parent of a constituent is a PP, the identity of the preposition

Pradhan et al. (2004) Features

- More (**31%** error reduction from baseline due to these + Surdeanu et al. features)



Joint Scoring: Enforcing Hard Constraints

- Constraint 1: Argument phrases do not overlap

  By [$_{A1}$ working [$_{A1}$ hard] , he] ***said*** *, you can achieve a lot.*

  - Pradhan et al. (04) – greedy search for a best set of non-overlapping arguments
  - Toutanova et al. (05) – exact search for the best set of non-overlapping arguments (dynamic programming, linear in the size of the tree)
  - Punyakanok et al. (05) – exact search for best non-overlapping arguments using integer linear programming
- Other constraints ([Punyakanok et al. 04, 05])
  - no repeated core arguments (good heuristic)
  - phrases do not overlap the predicate
  - (*more later*)

Joint Scoring: Integrating Soft Preferences

- Gildea and Jurafsky (02) – a smoothed relative frequency estimate of the probability of frame element multi-sets:

$$P(\{A0, AM_{TMP}, A1, AM_{TMP}\} | hit)$$

  – Gains relative to local model 59.2 → 62.9 FrameNet automatic parses

- Pradhan et al. (04) – a language model on argument label sequences (with the predicate included)

  – Small gains relative to local model for a baseline system 88.0 → 88.9 on core arguments PropBank correct parses

$$P(A0, AM_{TMP}, hit, A1, AM_{TMP})$$

- Toutanova et al. (05) – a joint model based on CRFs with a rich set of joint features of the sequence of labeled arguments (*more later)*
  – Gains relative to local model on PropBank correct parses 88.4 → 91.2 (24% error reduction); gains on automatic parses 78.2 → 80.0

- Also tree CRFs [Cohn & Brunson] have been used

Per Argument Performance

CoNLL-05 Results on WSJ-Test

- Core Arguments (Freq. ~70%)

	Best F_1	Freq.
A0	88.31	25.58%
A1	79.91	35.36%
A2	70.26	8.26%
A3	65.26	1.39%
A4	77.25	1.09%

- Adjuncts (Freq. ~30%)

	Best F_1	Freq.
TMP	78.21	6.86%
ADV	59.73	3.46%
DIS	80.45	2.05%
MNR	59.22	2.67%
LOC	60.99	2.48%
MOD	98.47	3.83%
CAU	64.62	0.50%
NEG	98.91	1.36%

Arguments that need to be improved

Data from Carreras&Màrquez's slides (CoNLL 2005)