The Necessity of Parsing for Predicate Argument Recognition

Daniel Gildea and Martha Palmer

Introduction

- Work has shifted from complex, rule-based systems to simpler finite-state and statistically based systems.
- Attention has been given to labeling corpora with semantic roles, such as in FrameNet and Propbank.
- Even for a single predicate, semantic arguments often have multiple syntactic realizations.
- In (Gildea and Jurfafski, 2002) and (Miller et al, 1998), both describe systems for which there are no gold-standard parses available.
- Because the Propbank labels data from the Penn Treebank, gold-standard parses are available for the Propbank.
- This paper compares performance of a system using gold-standard parses, with that of a system using automatic parser output.

(1) John will meet with Mary.
John will meet Mary.
John and Mary will meet.

(2) The door opened.
Mary opened the door.

Daniel Gildea and Daniel Jurfafski. 2002. Automatic Labeling of Semantic Roles. *Computational Linguistics.*

Scott Miller, Michael Crystal, Heidi Fox, Lance Ramshaw, Richard Schwartz, Rebecca Stone, Ralph Weischedel, and the Annotation Group. 1998. Algorithms that learn to extract information – BBN: Description of the SIFT system as used for MUC-7. In *Proceedings of the Seventh Message Understanding Conference (MUC-7)*, April.

Purpose

- Quantify the effect of parser accuracy on system performance in the task of semantic role identification.
- Examine whether a flatter, “chunked” representation can be as effective as a full parse.

Semantic Role Labeling

[$_{A0}$ He] [$_{AM-MOD}$ would] [$_{AM-NEG}$ n't] [$_{V}$ **accept**] [$_{A1}$ anything of value] from [$_{A2}$ those he was writing about] .

V: verb
A0: acceptor
A1: thing accepted
A2: accepted-from
A3: attribute
AM-MOD: modal
AM-NEG: negation

Source: “CoNLL-2005 Shared Task: Semantic Role Labeling” http://www.lsi.upc.edu/~srlconll/

The Data

FrameNet

- Project at International Computer Science Institute
- Annotation is performed on the British National Corpus.
- Labels verbs, nouns and adjectives
- Focuses on semantic “frames,” annotation is done by frame
- A frame is a schematic representation of a situation. Annotators define the frame, then its “frame elements.”
- Frame elements take on known conceptual roles.

Propbank

- Project at University of Pennsylvania
- Annotation is performed on the Penn Treebank.
- Only addresses verbs.
- Can be thought of as “FrameNet without the Frames,” annotation is done on a per-predicate basis
- Predicates are used. Arguments are labeled according to their position.
- Arguments are numbered. Similar verbs may share rolesets.

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)

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.

Source: Colin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of COLING/ACL,* pages 86-90, Montreal Canada

Propbank

11) *The company bought a wheel–loader from Dresser.*

Arg0:	The company
rel:	bought
Arg1:	a wheel–loader
Arg2–from:	Dresser

12) *TV stations bought "Cosby" reruns for record prices.*

Arg0:	TV stations
rel:	bought
Arg1:	"Cosby" reruns
Arg3–for:	record prices.

PURCHASE	BUY	SELL
Arg0: buyer	Arg0: buyer	Arg0: seller
Arg1: thing bought	Arg1: thing bought	Arg1: thing sold
Arg2: seller	Arg2: seller	Arg2: buyer
Arg3: price paid	Arg3: price paid	Arg3: price paid
Arg4: benefactive	Arg4: benefactive	Arg4: benefactive

Source: Paul Kingsbury and Martha Palmer. 2002. From Treebank to Propbank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002),* Las Palmas, Canary Islands, Spain.

The Model

P(r|pt, path, position, voice, hw, p)

- Phrase Type
- Parse Tree Path
- Position
- Voice
- Head Word

The Experiments

- The first experiment provided the system with arguments, the system merely had to label them. The following configurations were compared:
  - Propbank with:
    - Gold-standard parses
    - Automatic parses
    - Gold-standard parses, for which more than 10 examples were available
    - Automatic parses, for which more than 10 examples were available
  - FrameNet with automatic parses
- The second experiment was the same as the first, but the system also had to also find the arguments in this one.
- The first experiment was repeated with the path feature removed, using gold-standard Propbank parses.
- Two modifications of the path feature were tried
  - “Collapsed” paths
  - Two values: “NP under S” and “NP under VP”
- Experiments one and two were repeated using gold-standard chunks instead of parsing.

Chunking

- More coarse analyses than a full parse.
- It may be the case that systems using chunks are more robust to error than those using parsers.

$[_{NP}$ Big investment banks$]$ $[_{VP}$ refused to step$]$ $[_{ADVP}$ up$]$ $[_{PP}$ to$]$ $[_{NP}$ the plate$]$ $[_{VP}$ to support$]$ $[_{NP}$ the beleaguered floor traders$]$ $[_{PP}$ by$]$ $[_{VP}$ buying$]$ $[_{NP}$ big blocks$]$ $[_{PP}$ of$]$ $[_{NP}$ stock$]$, $[_{NP}$ traders$]$ $[_{VP}$ say$]$.

The Results

Arguments Provided

	Accuracy FrameNet	Propbank	Propbank > 10 *ex.*
Gold-standard parses		82.8	84.1
Automatic parses	82.0	79.2	80.5

Find Arguments and Roles

	FrameNet		Propbank		Propbank > 10	
	Precision	*Recall*	*Precision*	*Recall*	*Precision*	*Recall*
Gold-standard parses			71.1	64.4	73.5	71.7
Automatic parses	64.6	61.2	57.7	50.0	59.0	55.4

Arguments Provided, Chunking Results

Path	*Head*	*Accuracy*
gold parse	gold parse	82.3
auto parse	auto parse	79.2
not used	gold parse	81.7
not used	chunks	77.0

Find Arguments and Roles, Chunking Results

	Precision	*Recall*
gold parse	71.1	64.4
auto parse	57.7	50.0
chunk	27.6	22.0
chunk, relaxed scoring	49.5	35.1

Conclusions

- Other finite-state systems may do better than the chunking system in this experiment.
- By using a gold-standard chunking representation, better results have been achieved than could be expected from an automatic chunking system.
- Statistical parsers do a good job of providing information for this task. This information includes not only structure but also head words.
- Improvements in parsers will equate to improved performance on this task.