TRIOS-TimeBank Corpus: Extended TimeBank corpus with help of Deep Understanding of Text

The seventh international conference on Language Resources and Evaluation (LREC), Malta, 2010

Naushad UzZaman and James F. Allen
Computer Science Department
University of Rochester
Rochester, NY, USA

UNIVERSITY *of* ROCHESTER

TimeML, TimeBank, TempEval-1 and TempEval-2

- TimeML is the scheme for temporal annotation by James Pustejovsky et al.

- TimeBank is the first annotated corpus

- TempEval 1 and 2 are annotated corpus and was shared task on temporal information extraction
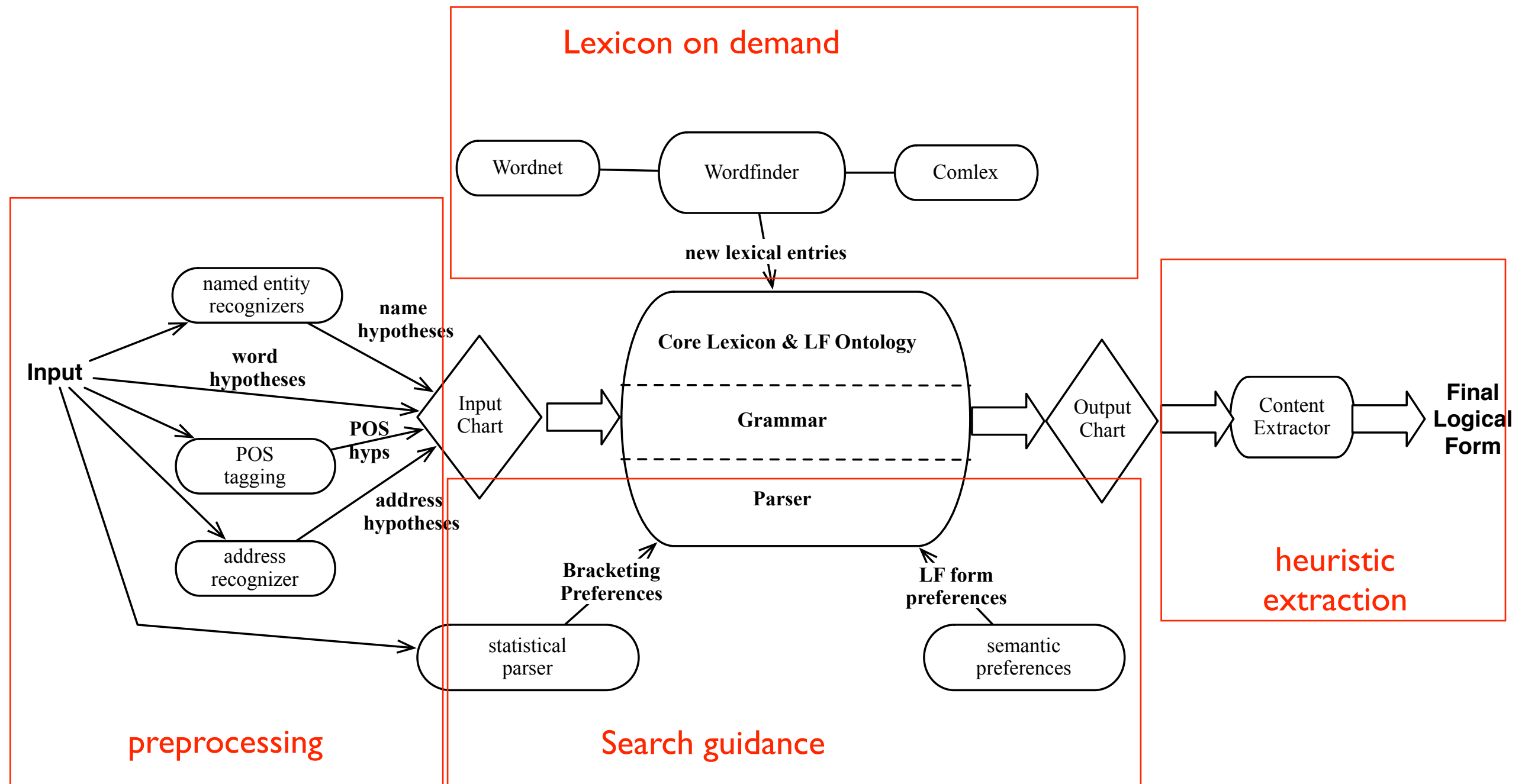
Outline

- **Our System**

- Extension to TimeBank and TimeML

  - Suggest new event

  - Suggest new temporal expressions

  - Suggest ontology type as new event features
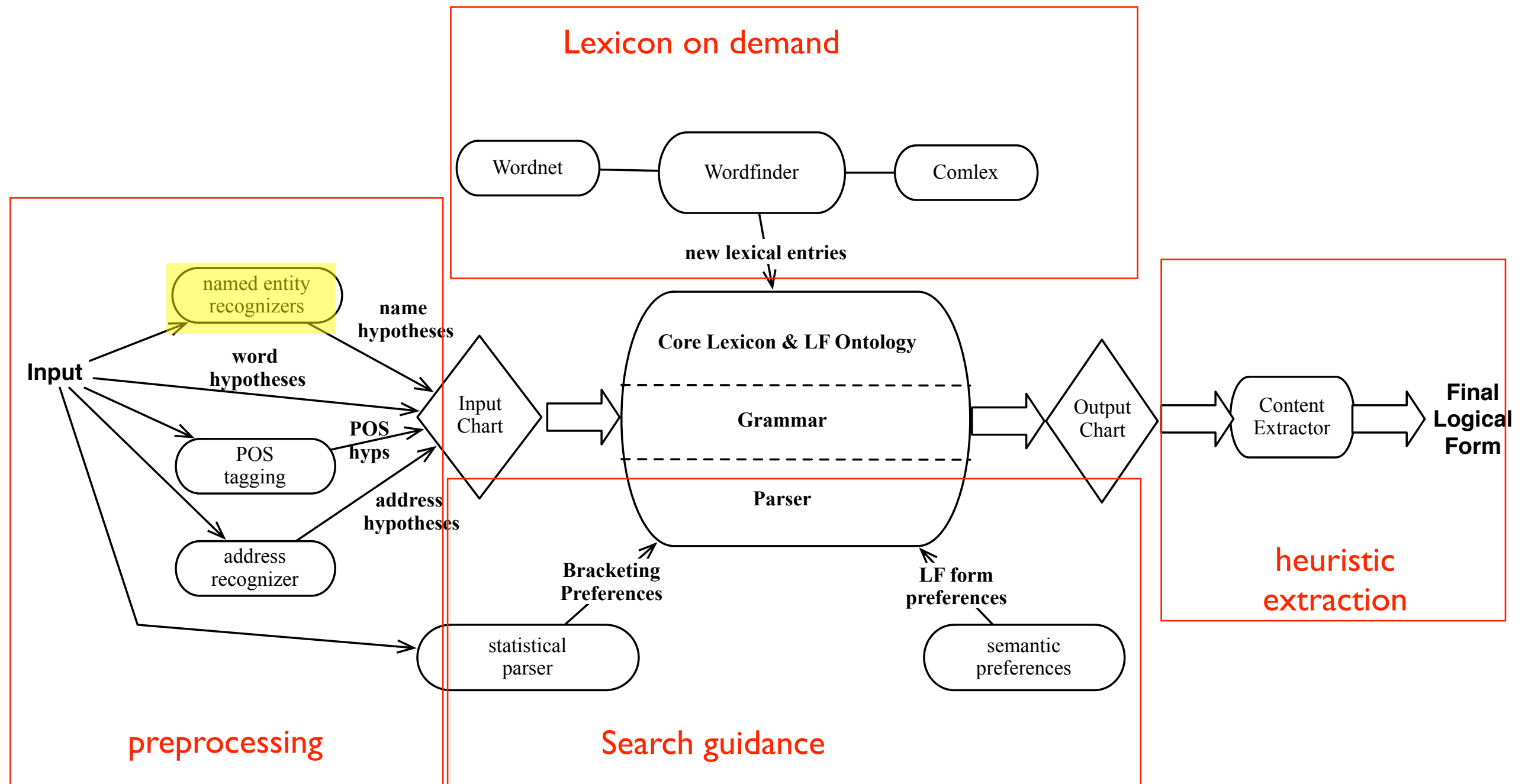
  - Suggest improved relations in TimeML

- Future Work

TRIPS Parser: Broad coverage deep parsing

"Advanced Medical paid $ 106 million in cash for its share in a unit of Henley 's Fisher Scientific subsidiary ."

UNIVERSITY *of* ROCHESTER

TRIPS Parser: Broad coverage deep parsing

"*Advanced Medical paid $ 106 million in cash for its share in a unit of Henley 's Fisher Scientific subsidiary .*"

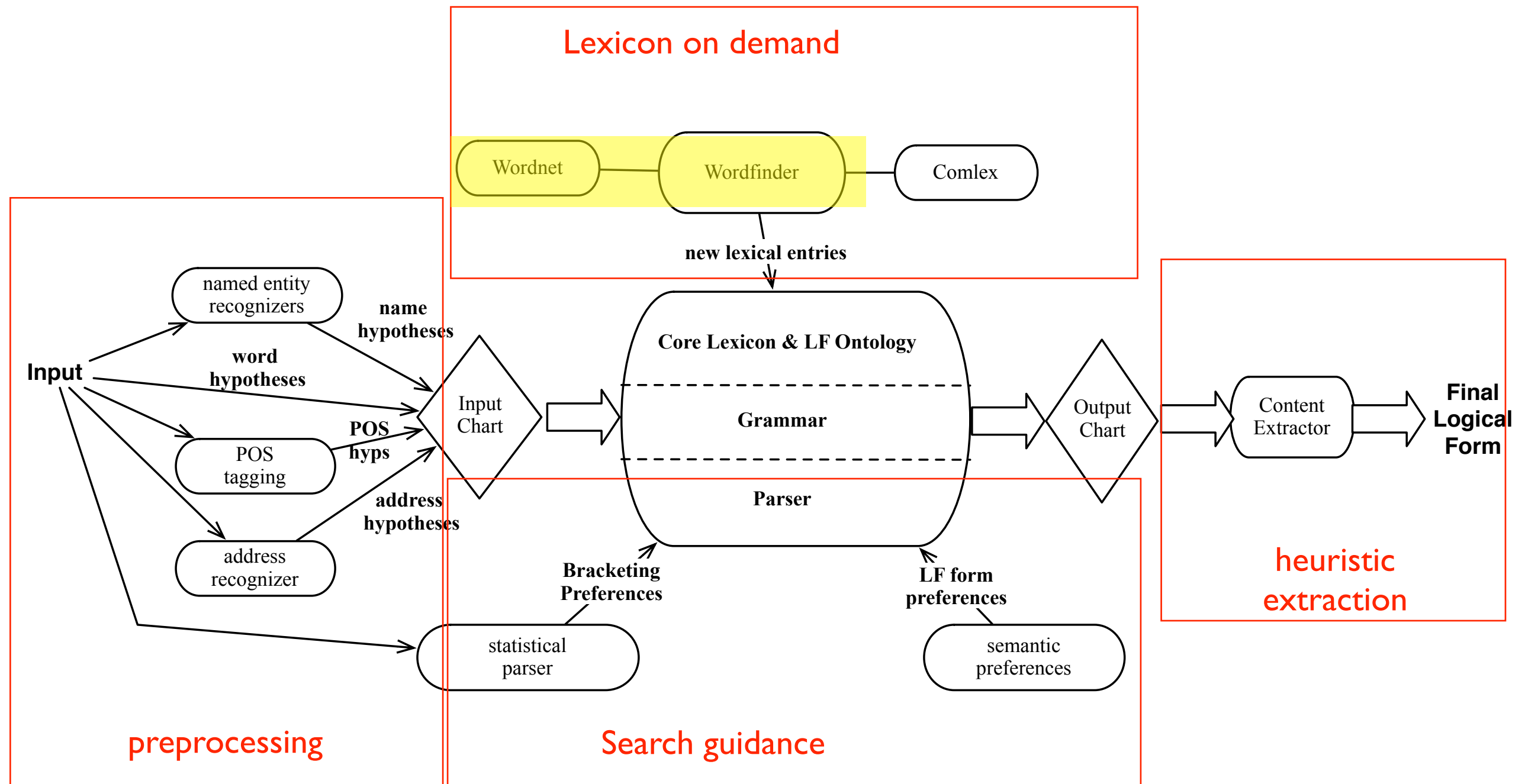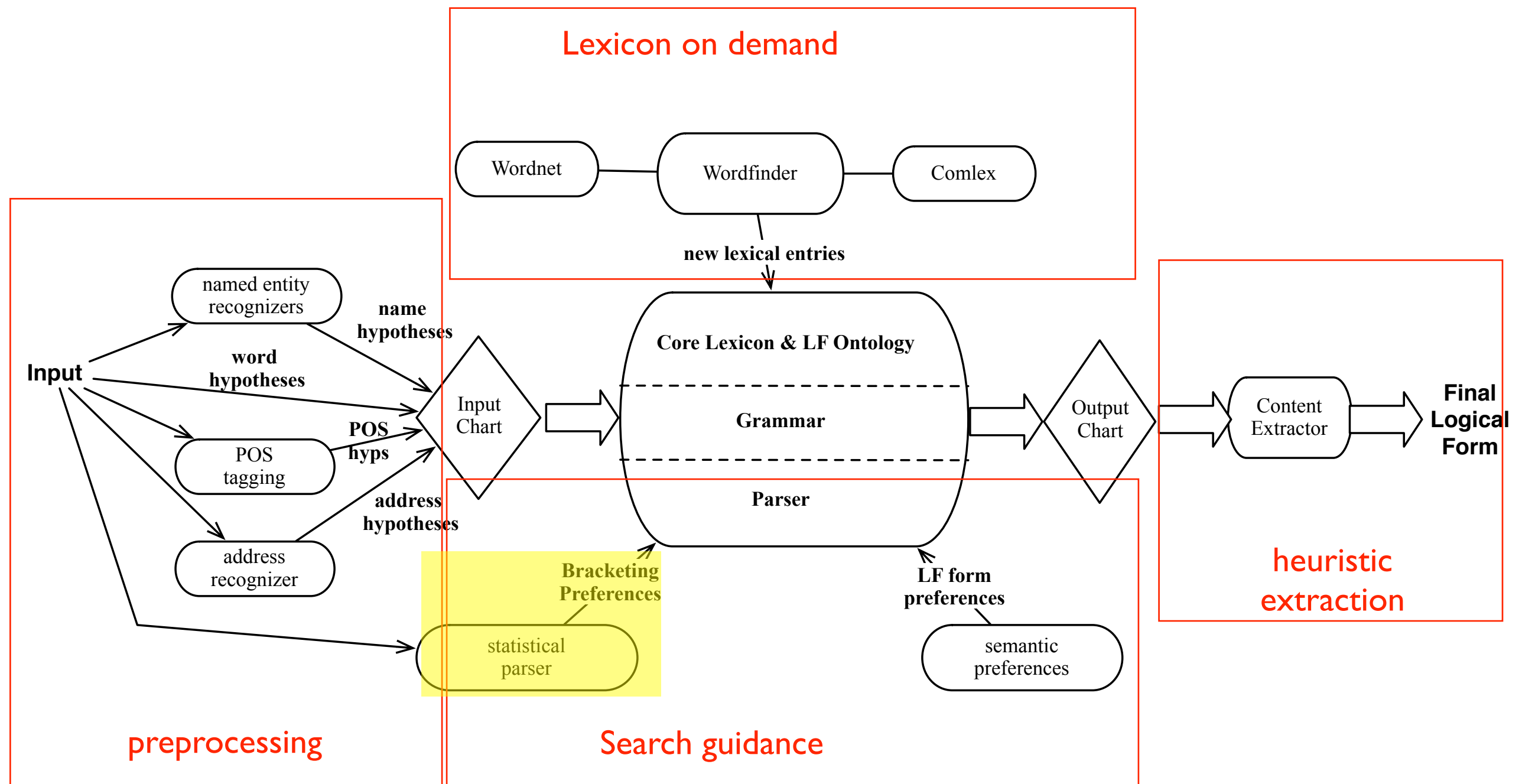# TRIPS Parser: Broad coverage deep parsing

"Advanced Medical paid $ 106 million in cash for its share in a unit of Henley 's Fisher Scientific subsidiary ."

UNIVERSITY *of* ROCHESTER

TRIPS Parser: Broad coverage deep parsing

Lexicon on demand

Wordnet — Wordfinder — Comlex

new lexical entries

Core Lexicon & LF Ontology

Grammar

Parser

named entity recognizers

name hypotheses

word hypotheses

POS hyps

POS tagging

address hypotheses

address recognizer

Input

Input Chart

Output Chart

Content Extractor

Final Logical Form

Bracketing Preferences

statistical parser

LF form preferences

semantic preferences

preprocessing

Search guidance

heuristic extraction

"Advanced Medical paid $ 106 million in cash for its share in a unit of Henley 's Fisher Scientific subsidiary .

UNIVERSITY *of* ROCHESTER

Events and event features extraction using TRIPS parser

Sentence

He fought in the war

TRIPS parser output

```
(SPEECHACT V1 SA-TELL :CONTENT V2)
(F V2 (:* FIGHTING FIGHT) :AGENT V3 :MODS (V4) :TMA ((TENSE PAST)))
(PRO V3 (:* PERSON HE) :CONTEXT-REL HE)
(F V4 (:* SITUATED-IN IN) :OF V2 :VAL V5) (THE V5 (:* ACTION WAR))
```

100+ Extraction rules

```
((THE ?x (? type SITUATION-ROOT))
  -extract-noms>
    (EVENT ?x (? type SITUATION-ROOT) :pos NOUN :class OCCURRENCE ))
```

Extracted with extraction rules

```
<EVENT eid=V2 word=FIGHT pos=VERBAL ont-type=FIGHTING tense=PAST
  class=OCCURRENCE voice=ACTIVE aspect=NONE polarity=POSITIVE nf-morph=NONE>
<RLINK eventInstanceID=V2 ref-word=HE ref-ont-type=PERSON relType=AGENT>
<SLINK signal=IN eventInstanceID=V2 subordinatedEventInstance=V5
  relType=SITUATED-IN>
<EVENT eid=V5 word=WAR pos=NOUN ont-type=ACTION class=OCCURRENCE
  voice=ACTIVE polarity=POSITIVE aspect=NONE tense=NONE>
```

Event Extraction Performance

TempEval-1 and TimeBank

System	Precision	Recall	Fscore	(P+R)/2
TRIOS avg	**0.8638**	0.7074	**0.7778**	**0.7856**
TRIPS avg	0.5801	**0.8513**	0.6900	0.7157
STEP	0.82	0.706	0.7587	0.763
Sim-Evita	0.812	0.657	0.727	0.7345
IAA	N/A	N/A	N/A	0.78

Table 1: Event Extraction Performance on Bethard and Martin's test data

TempEval-2

System	Precision	Recall	Fscore
TRIOS	0.80	0.74	0.77
TRIPS	0.55	**0.88**	0.68
Best (TIPSem)	**0.81**	0.86	**0.84**

Table 1: Performance of Event Extraction (Task B) in TempEval-2

UNIVERSITY *of* ROCHESTER

Performance on Temporal Expression Extraction

		TRIPS TRIOS	Best HeidelTime-1
Temp Exp extraction	Precision	0.85	0.90
	Recall	0.85	0.82
	Fscore	0.85	0.86
Normalization	*type*	0.94	0.96
	value	0.76	0.85

Table 1: Performance on Temporal Expression extraction (Task A)

UNIVERSITY *of* ROCHESTER

Outline

- Our System

- **Extension to TimeBank and TimeML**

  - **Suggest new event**

  - Suggest new temporal expressions

  - Suggest ontology type as new event features

  - Suggest improved relations in TimeML
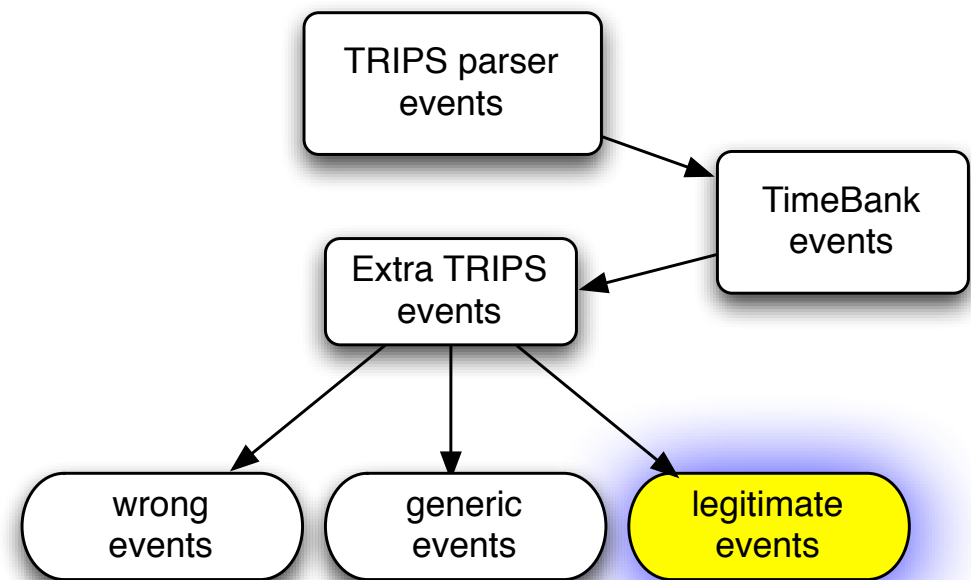
- Future Work

What not to tag?

- Not to tag generic interpretations

    - *Use of corporate jets for political travel is legal.*

- Complements are generics

    - *He said students are prohibited from fighting with each other.*

- Nominalization without extra information

    - *Newspaper reports have **said** ...*

TRIOS minus TimeBank

- Result of wrong parse

- Generic event

- Legitimate event but missed by annotators

Missed legitimate events

1. *At least one of the sensitive sites was a barracks of the elite Republican Guard, a well-placed source **told** The Associated Press.*

2. *Net interest income for the third quarter **declined** to $35.6 million from $70.1 million a year ago.*

3. *About $518 million of debt is **affected**.*

4. *If Iraq **chooses** a simple war of nerves and economic attrition, the Bush administration **knows** a long stalemate could try the patience of the American public and the West in general, and could **open** the possibility that moderate Arabs -- even including Saudi Arabia -- might **drop** out of the effort against Iraq and **accept** some deal from Saddam Hussein.*

5. *"It's the whole uncertainty about what's happening around us," **said** Valentin Von Korff, a trader at Credit Suisse First Boston in Frankfurt.*

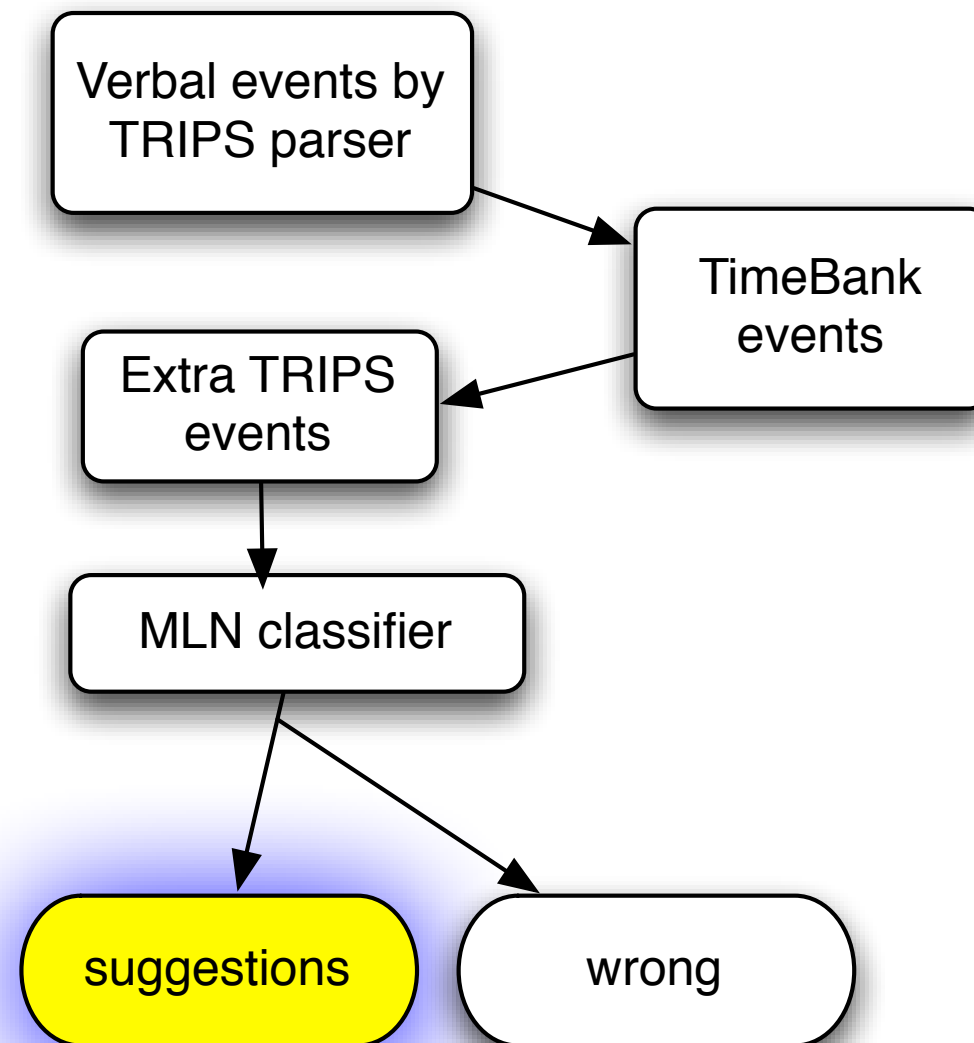# Suggest new verbal events to TimeBank

Table 1: Performance on suggested events

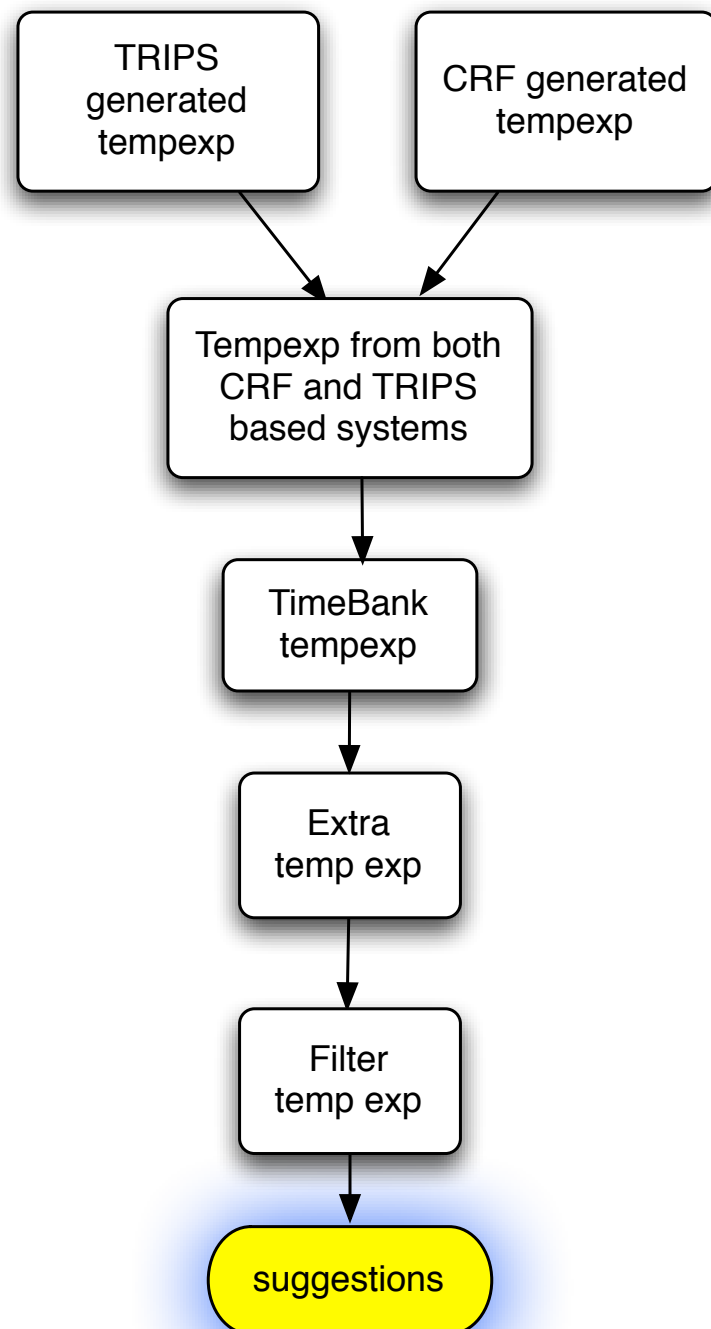| TimeML spec	Our Analysis	Number	Performance
Correct	Correct	90	60%
	Generic complement	5	3.3%
Generic	Generic	28	18.8%
	True throughout	8	5.3%
	Others	2	1.3%
Wrong	Wrong	13	8.7%
	Not appropriate	3	2%

Total event: 7935
suggestion: ~1%

true-throughout

If Iraq chooses a simple war of nerves and economic attrition, the Bush administration **knows** a long stalemate could try the patience of the American public and the West in general, …

UNIVERSITY *of* ROCHESTER

Suggest new temporal expressions

```
┌─────────────┐   ┌─────────────┐
│   TRIPS     │   │ CRF generated│
│  generated  │   │   tempexp    │
│  tempexp    │   │             │
└──────┬──────┘   └──────┬──────┘
       │                 │
       └────────┬────────┘
                ▼
       ┌─────────────────┐
       │ Tempexp from both│
       │  CRF and TRIPS   │
       │  based systems   │
       └────────┬────────┘
                ▼
         ┌────────────┐
         │ TimeBank   │
         │ tempexp    │
         └─────┬──────┘
               ▼
         ┌────────────┐
         │  Extra     │
         │ temp exp   │
         └─────┬──────┘
               ▼
         ┌────────────┐
         │  Filter    │
         │ temp exp   │
         └─────┬──────┘
               ▼
         ┌────────────┐
         │ suggestions│
         └────────────┘
```

Total timex: 1414
accuracy: 50/68 = 73.5%
suggestion: ~3.5%

- *At the end of the broadcast this evening, one more trip around Havana to see what it's been like since **the last time.***

- *Turks feel they have special ties to the whole region, which they ruled for **hundreds of years** during the Ottoman Empire.*

- *Weisfield's, based in Seattle, Wash., **currently** operates 87 specialty jewelry stores in nine states.*

- ***Previously**, watch imports were denied such duty-free treatment.*

UNIVERSITY *of* ROCHESTER

Add ontology-type as new event feature

- TimeML captures event information with coarse-grained *class* (7) or *pos*, or fine-grained *word*

- TRIPS Ontology type

  - more fine-grained than *class* or *pos*

  - coarse-grained than *word*

  - *Fighting* for *fight*, *Action* for *war*

  - Few other words with ont-type *Fighting*: contend, defend, struggle, etc.

- Mapping to WordNet

- Freely available

TimeML Class:
1) *Occurrence: die, crash, build;*
2) *State: on board, kidnapped;*
3) *Reporting: say, report;*
4) *I-Action: attempt, try, promise;*
5) *I-State: believe, intend, want;*
6) *Aspectual: begin, stop, continue;*
7) *Perception: see, hear, watch, feel.*

Outline

- Our System

- Extension to TimeBank and TimeML

  - Suggest new event

  - Suggest new temporal expressions

  - Suggest ontology type as new event features

  - **Suggest improved relations in TimeML**

- Future Work

UNIVERSITY *of* ROCHESTER

More SLINK instances

- SLINK or Subordinate Links: relation between two events

- TimeML includes: modal, factive, counter-factive, evidential, negative evidential and conditional

- Many other cases when one event is argument of other

SLINK examples

- "They have to **continue** to _tighten_ their belts," said Craig Kloner, an analyst at Goldman, Sachs amp Co. (Purpose)

- He **fought** in the _war_.

  ```
  <SLINK signal=IN eventInstanceID=V2
  subordinatedEventInstance=V5 relType=SITUATED-IN>
  ```

- Suggest ~900 SLINKs

New Relation Link, RLINK

- Dependency information improves temporal relation identification performance (Chambers et al. 2007 and Katsumasa et al. 2009)

- Chambers and Jurafsky (2008)'s narrative chain with AGENT

  ```
  <RLINK eventInstanceID=V2 ref-word=HE ref-ont-
  type=PERSON relType=AGENT>
  ```

- Suggest ~2000 RLINKs

Table 1: Most common relTypes used in SLINKs and RLINKs

Our Role	VerbNet equivalents	Lirics equivalents	SLINK Count	RLINK Count
Agent	Agent, Actor	Agent	19	709
Theme	Theme, Stimulus	Theme	336	1137
Affected	Patient	Patient	13	92
Cause	Cause	Cause		49
Goal-as-Loc	Destination	finalLocation	47	
To-Loc	Recipient	Goal	46	
At-Loc	Location	Location	42	
In-Loc	Location	Location	28	
On	Location	Location	20	
Situated-In	Location?	Location?	39	
Purpose	–	Purpose	226	

Approaching TempEval-2

Features	TE2 C	TE2 D	TE2 E	TE2 F
	TE1 A	TE1 B	TE1 C	
Event Class	YES	YES	e_1 x e_2	e_1 x e_2 [1]
Event Tense	YES	YES	e_1 x e_2	e_1 x e_2
Event Aspect	YES	YES	e_1 x e_2	e_1 x e_2
Event Polarity	YES	YES	e_1 x e_2	e_1 x e_2
Event Stem	YES	YES	e_1 x e_2	e_1 x e_2
Event Word	YES	YES	YES	YES
Event Constituent [2]		YES	e_1 x e_2	e_1 x e_2
Event Ont-type [3]	YES	YES	e_1 x e_2	e_1 x e_2
Event LexAspect [4] *x Tense*	YES	YES	e_1 x e_2	e_1 x e_2
Event Pos	YES	YES	e_1 x e_2	e_1 x e_2
Timex Word		YES		
Timex Type	YES	YES		
Timex Value	YES	YES		
Timex DCT relation	YES	YES		
Event's semantic role [5]	YES	YES		
Event's argument's ont-type	YES	YES		
TLINK event-time signal [6]	YES	YES		
SLINK event-event relation type [7]				e_1 x e_2

Table 1: Features used for TempEval-2 (TE2) Task C, D, E and F or TempEval-1 (TE1) Task A, B and C.

Performance in TempEval-2

Task	TRIPS		TRIOS		Best (with corpus features)
	Precision	Recall	Precision	Recall	Precision
Task C	0.63	**0.52**	**0.65**	**0.52**	0.63 (JU-CSE, UCFD, NCSU-indi)
Task D	0.76	**0.69**	0.79	0.67	**0.82** (TIPSem)
Task E	**0.58**	**0.50**	0.56	0.42	0.55 (TIPSem)
Task F	0.59	**0.54**	0.6	0.46	**0.66** (NCSU-individual)

Table 1: Performance of Temporal Relations on TempEval-2 (Task C-F)

Task	Description	Best
Task A	Temporal expression extraction	**TRIOS**
Task B	Event Extraction	TIPSem
Task C	Event-Timex relationship	**TRIOS**
Task D	Event-DCT relationship	TIPSem
Task E	Main event-event relationship	**TRIOS**
Task F	Subordinate event-event relationship	**TRIOS**

Table 1: Head-to-head comparison of TRIOS, TIPSem and JU-CSE-TEMP (teams that approached all tasks) in TempEval-2 challenge

UNIVERSITY *of* ROCHESTER

Outline

- Our System

- Extension to TimeBank and TimeML

  - Suggest new event

  - Suggest new temporal expressions

  - Suggest ontology type as new event features

  - Suggest improved relations in TimeML

- Future Work

Suggest Temporal Links

- Suggest new event-time temporal links that are missed by TimeBank annotators

- Suggest intra-sentence event-event temporal relations, which are ignored in TempEval-1

Automatically Building Larger Temporally Annotated Corpora

- Automatically build larger temporally annotated corpus for news domain, which can be reviewed by human annotators

- Automatically build temporally annotated corpus for other domains like medical domain

Summary

- Suggested New Events in TimeBank

- Added New Event Feature - *Ontology Type;* released TRIPS ontology

- Suggested New Temporal Expressions in TimeBank

- Added Improved Relations in Existing Annotation Scheme

    - SLINK or Subordinate links - relation between events

    - RLINK or Relation link - relation between event and its arguments

- Released TRIOS-TimeBank corpus, TRIPS ontology

Resources

- TRIOS-TimeBank Corpus & TRIPS Ontology:

  http://www.cs.rochester.edu/u/naushad/trios-timebank-corpus

- Other temporal related resources:

  http://www.cs.rochester.edu/u/naushad/temporal

Questions ?

- TRIOS-TimeBank Corpus & TRIPS Ontology:

  http://www.cs.rochester.edu/u/naushad/trios-timebank-corpus

- Other temporal related resources:

  http://www.cs.rochester.edu/u/naushad/temporal

Event Feature Extraction Performance

System	TRIPS	TRIOS	Best
Class	0.67	0.77	0.79 (TIPSem)
Tense	0.67	0.91	0.92 (Edinburgh-LTG)
Aspect	0.97	**0.98**	0.98
Pos	0.88	0.96	0.97 (TIPSem, Edinburgh-LTG)
Polarity	**0.99**	**0.99**	0.99
Modality	0.95	0.95	0.99 (Edinburgh-LTG)

Table 1: Performance of Event Features on TempEval-2 (Task B)

Overview

Benefits of TRIPS ontology

- Superior semantic ontology; better abstraction

- No problem with word sense disambiguation

- Considers semantic roles for disambiguation

- Helps to generate better links

Why RLINK

- RLINKs could be in a separate layer

- needed for complete temporally aware system

- hence included for a complete temporally annotated corpus

Markov Logic Network

- Problems with rule based system and machine learning techniques

- Markov logic = first order logic + markov network (probabilistic graphical model)

  - FOL with weights

  - weights determine how much penalty for a formula to be violated

Example: It is not going to _change_

```
tense(e1, INFINITIVE) & aspect(e1, NONE) => class(e1, OCCURRENCE)     weight = 0.319913
```