# Deep Understanding and Textual Entailment

James Allen
University of Rochester &
Institute for Human and Machine Cognition

# Entailment requires knowledge

Lots of it!

e.g.,

- when you fall asleep, you were awake before
- if a building burns down, it is destroyed
- if you buy something, you own it
- driving a car requires a key

### Much Recent Work on Learning by Reading

#### Levels of Learning

- simple facts given an ontology instantiate one or more relations (e.g.,
   Athens is in Greece, Chicago is in the US, ...)
- complex facts given an ontology, learn compositional knowledge (e.g., learn the route to get to a restaurant)
- new words, concepts and relations Extending our ontology as we learn new facts (e.g.,learn about plumbing)

### Much Recent Work on Learning by Reading

#### Levels of Learning

- simple facts given an ontology instantiate one or more relations (e.g., Athens is in Greece, Chicago is in the US, ...)
- complex facts given an ontology, learn compositional knowledge (e.g., learn the route to get to a restaurant)
  - new words, concepts and relations Extending our ontology as we learn new facts (e.g., learn about plumbing)

#### Our Approach

Learn definitions of concepts from online sources, bootstraping from a hand-built deep understanding system

# What is Deep understanding?

Students develop deep understanding when they grasp the relatively complex relationships between the central concepts of a topic or discipline. Instead of being able to recite only fragmented pieces of information, they understand the topic in a relatively systematic, integrated or holistic way. As a result of their deep understanding, they can produce new knowledge by discovering relationships, solving problems, constructing explanations and drawing conclusions.

Students have only shallow understanding when they do not or cannot use knowledge to make clear distinctions, present arguments, solve problems or develop more complex understanding of other related phenomena.

Dept. of Education, Queensland

# What is Deep understanding?

Students develop deep understanding when they grasp the relatively complex relationships between the central concepts of a topic or discipline. Instead of being able to recite only fragmented pieces of information, they understand the topic in a relatively systematic, integrated or holistic way. As a result of their deep understanding, they can produce new knowledge by discovering relationships, solving problems, constructing explanations and drawing conclusions.

Students have only shallow understanding when they do not or cannot use knowledge to make clear distinctions, present arguments, solve problems or develop more complex understanding of other related phenomena.

Dept. of Education, Queensland

in other words, connecting language to other cognitive abilities: Knowledge, reasoning, action, learning, ...

### Building a Knowledge Base from definitions ...

#### Bootstrap from a hand-built deep understanding system

- TRIPS 2500+ concept upper ontology, semantic roles, extensive grammar and core vocabulary, and wordnet mappings
- Why? a core competence in linking NLP & KR in order to learn new concepts

#### Build new concepts compositionally

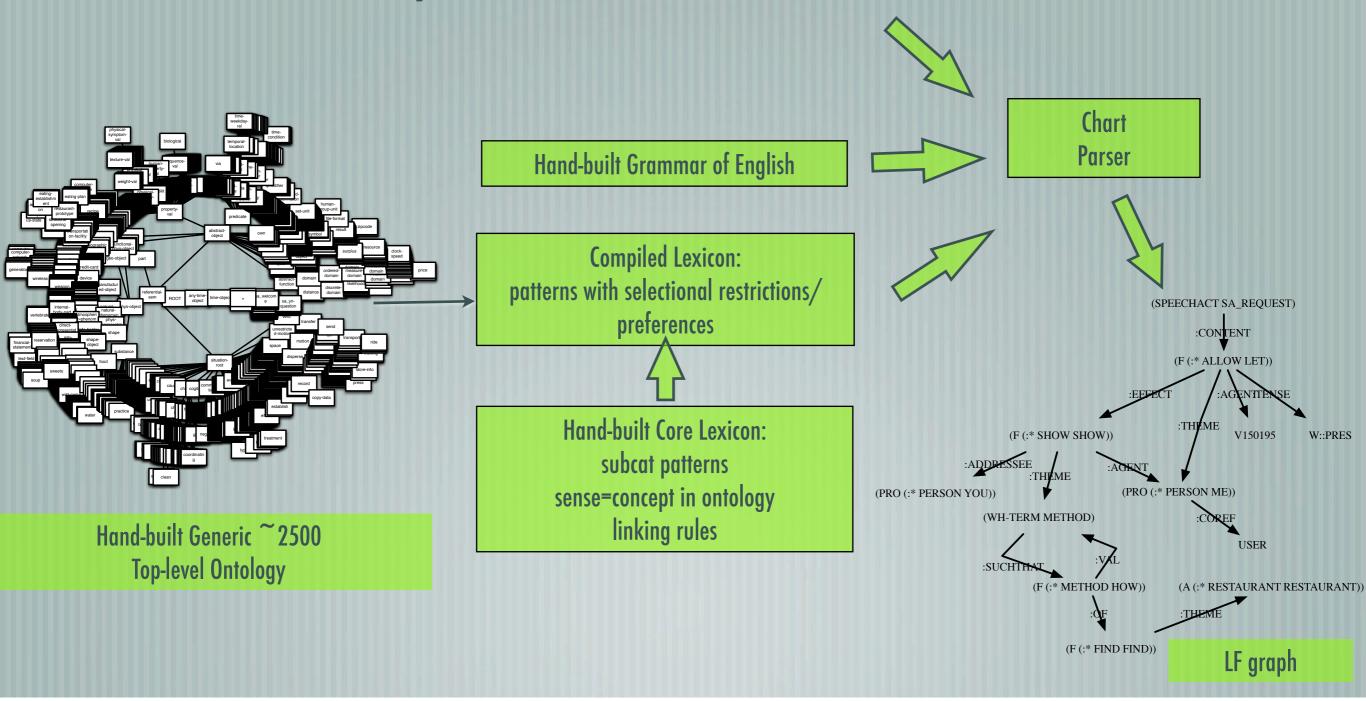
- new concepts expressed in OWL-DL (Description Logic)
  - Why? OWL designed to express complex compositional concepts; multiple inheritance; subsumption provide automatic classification of new concepts

#### Use existing sources for concept definitions

- Wordnet glosses
- Why? extensive source of definitions, and used in many applications

### TRIPS Language Processing

"Let me show you how to find a restaurant"



# But everyone knows deep understanding is not possible\*

Deep understanding only works in narrow domains using lots of hand-engineering!

Deep understanding is too brittle!

Deep understanding is not possible without solving the entire Al problem!

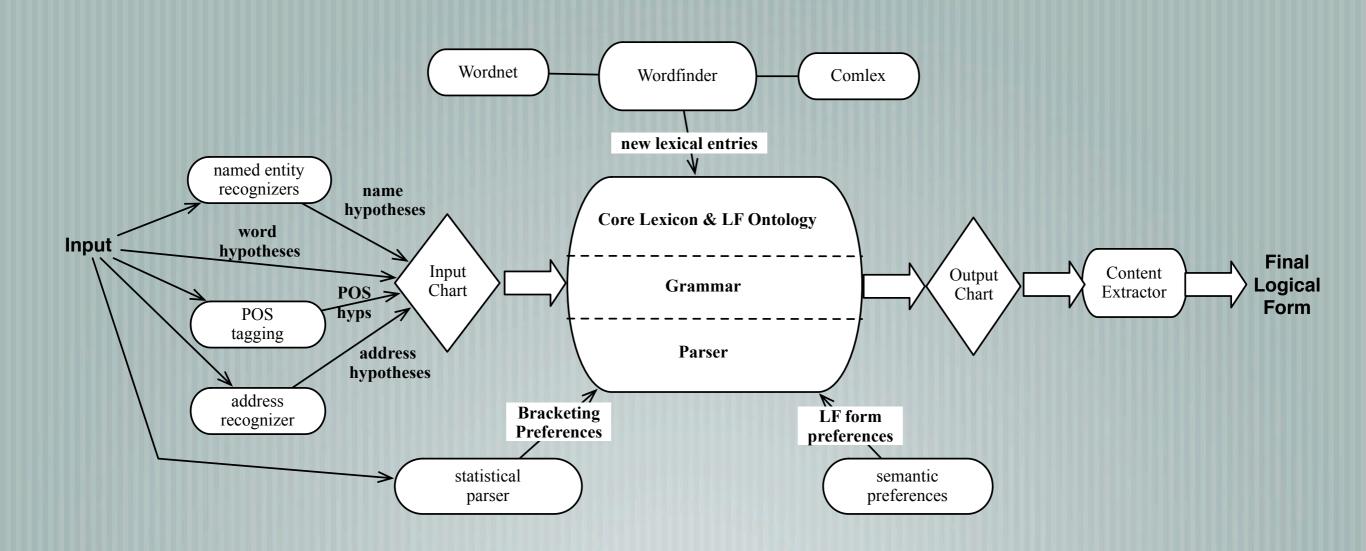
# But everyone knows deep understanding is not possible\*

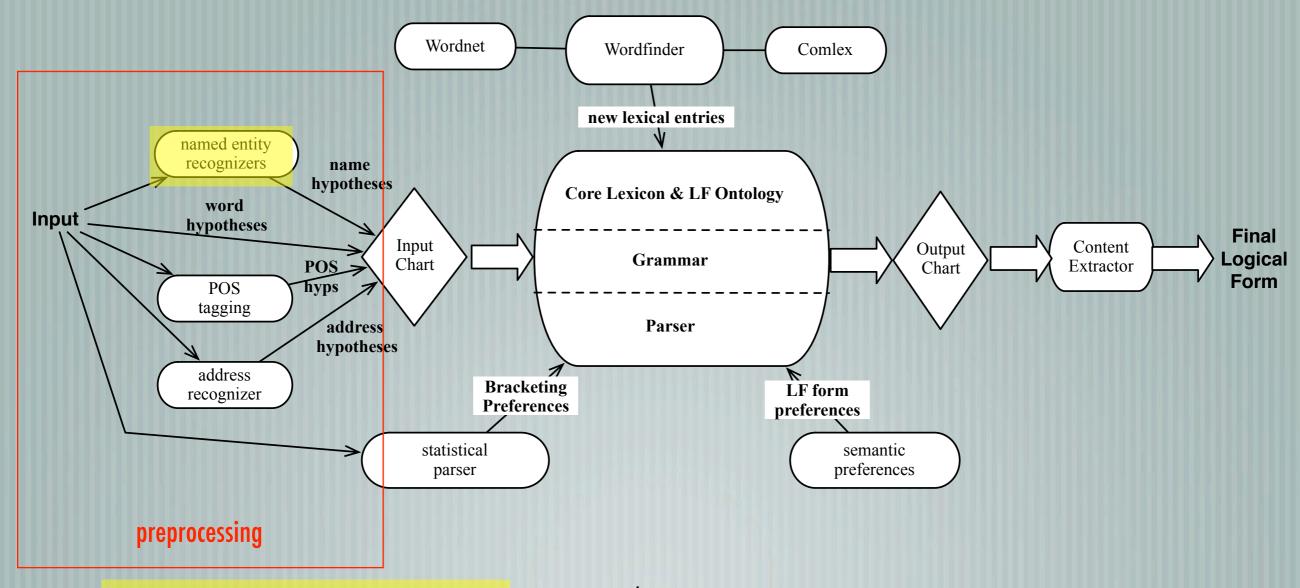
Deep understanding only works in narrow domains using lots of hand-engineering!

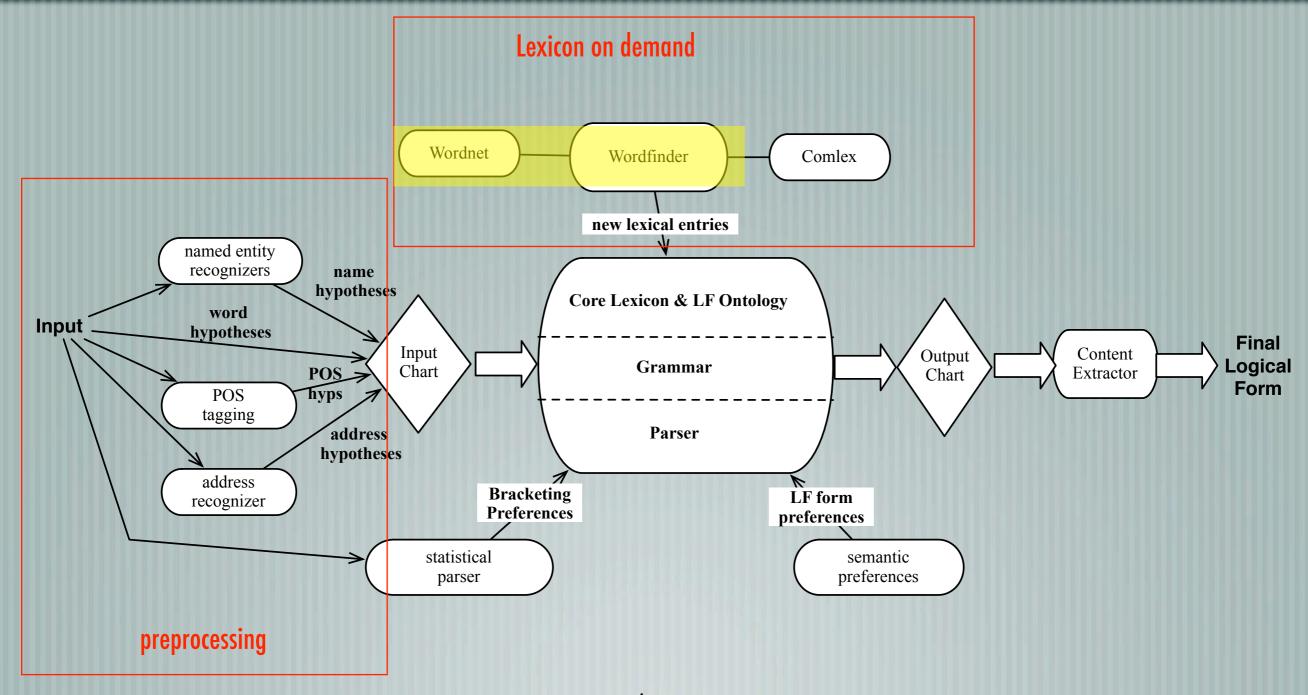
Deep understanding is too brittle!

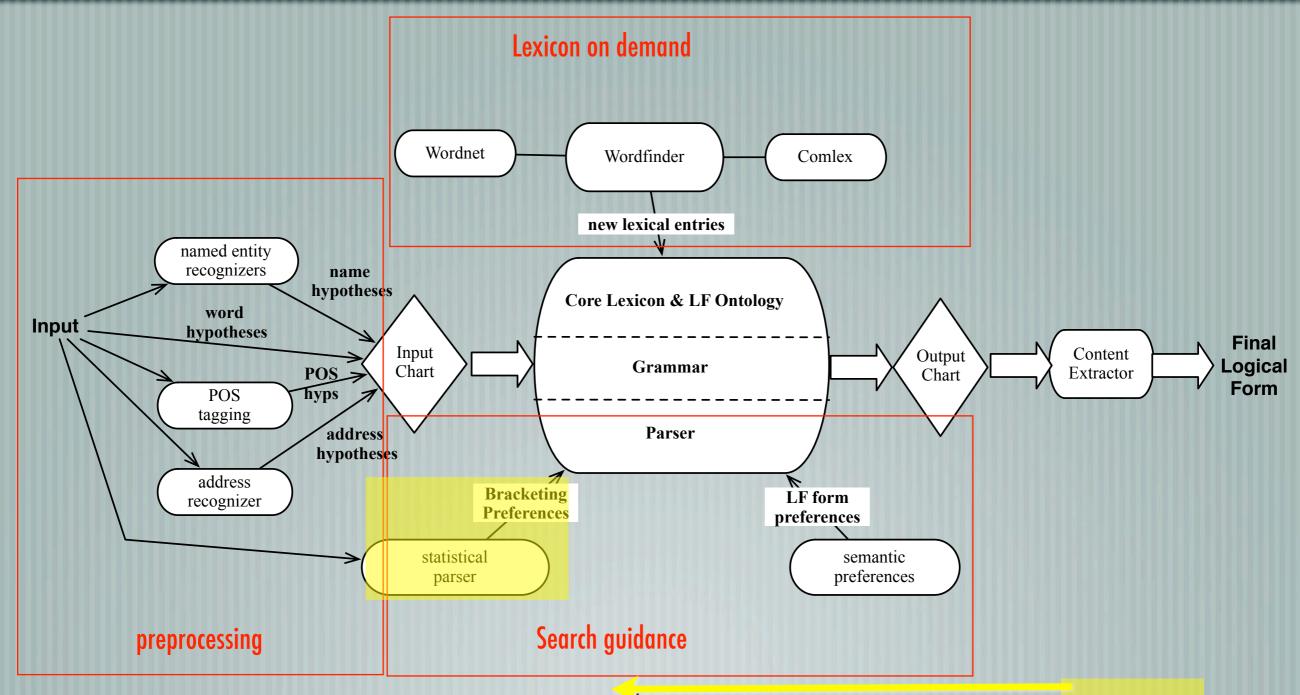
Deep understanding is not possible without solving the entire Al problem!

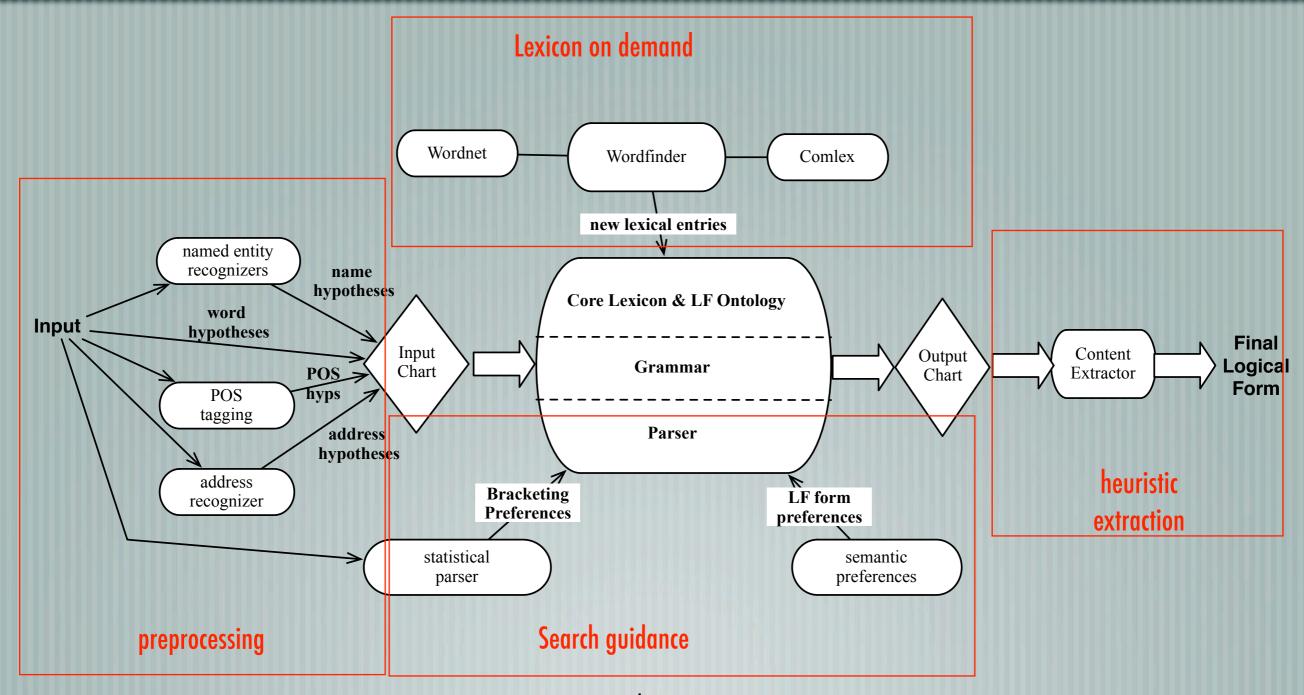












### Utilizing statistical preprocessing: TextTagger

#### TextTagger preprocesses text using a wide range of resources

#### Named entity recognitiOn

Stanford NER (-> ONT::ORGANIZATION, ONT::PERSON, ONT::GEOGRAPHIC-REGION)

Geographical Names (GNIS) -> ONT::GEOGRAPHIC-REGION

Street Address Recognizer

Alphanumerics, e.g., "AK-47"

#### Part of speech tagging

Stanford POS: emits both Penn treebank tags and TRIPS tags

#### **Structural Analysis**

Sentence and clause boundary detection and segmentation

Constituent bracketing (Stanford Parser)

#### Alternate spellings

e.g., behaviour, behavior

#### multi-word recognition based on Wordnet

... Ace Medical cooperating with us in ....

(phrase "Ace Medical" :frame (21 32) :tag (W::NAME) :ont-sense ONT::CORPORATION)

(word "cooperating" :frame (33 42) :tag (W::V) :penn-tag VBG :alt ("co-operating"))

(word "with" :frame (43 47) :tag (W::PREP W::ADV) :penn-tag IN)

(word "us" :frame (48 50) :tag (W::PRO) :penn-tag PRP)

(prefer (% W::NP) :frame (48 50) :score 8.799)

(prefer (% W::PP) :frame (43 50) :score 12.066)

### Building Lexical Entries Dynamically

```
"Cooperates" is not in core lexicon
      (word "cooperating" :frame (33 42) :tag (W::V) :penn-tag VBG :alt ("co-
      operating"))
          "co-operating" is also not in lexicon
      Look up in Wordnet
S: (v) collaborate, join forces, cooperate, get together (work together on a common enterprise of project) "The soprano and the pianist did not
get together very well"; "We joined forces with another research group"
   S: (v) work (exert oneself by doing mental or physical work for a purpose or out of necessity) "I will work hard to improve my grades"
      Map to TRIPS Ontology
```

{collaborate, join forces, ...} no mapping, move up wordnet hierarchy ...

{work} work % 2:41:02:: -> ONT::WORKING

which has essential role AGENT (+intentional object)

we adopt the linking rules for the words with a sense ONT::WORKING

#### **Derive Features**

Using penn tag VBG, we set the VFORM feature to "-ing"

From ONT::WORKING class and inheritance, we get other feature values

e.g., TIME-SPAN=EXTENDED, ASPECT=UNBOUNDED, ...

We now have a semantic lexical entry for "cooperates"

### Robustness: Heuristic Extraction

					VP
СР			ADVBL		
S			PP		
NP	VP		NP		
He	ran	and	to	the	store

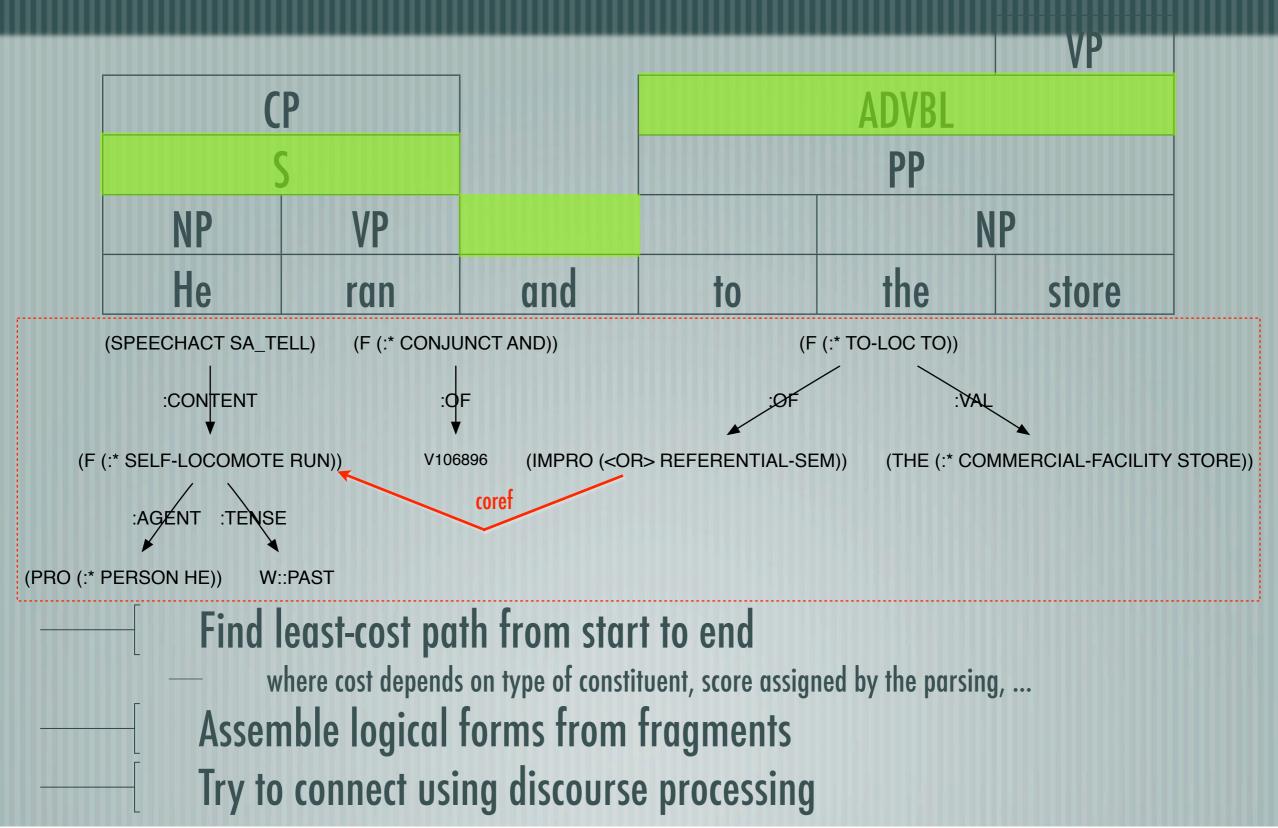
Find least-cost path from start to end

where cost depends on type of constituent, score assigned by the parsing, ...

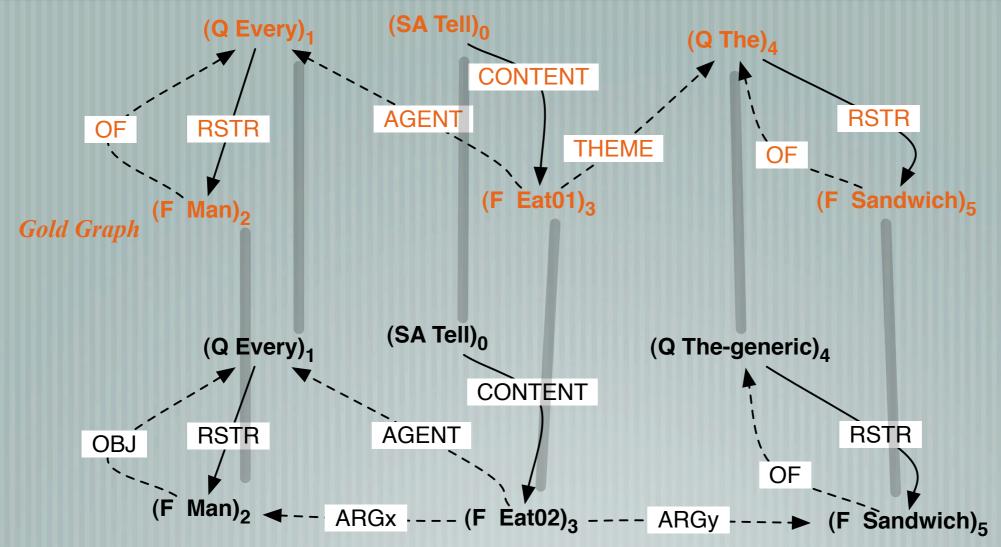
Assemble logical forms from fragments

Try to connect using discourse processing

### Robustness: Heuristic Extraction



### A Semantic Evaluation



Test Graph

 $Nscore_A(n) = 2$  if the indices match, and both the node type and sense in the label of node n matches the label of node A(n), 1 if one of them matches, and 0 otherwise.

 $Escore_A(e) = 1$  if e connect nodes n1 and n2, and there is an edge between A(n1) and A(n2) with same label, 0 otherwise.

```
Gscore(G,T) = max_A(Sum_{n,e}(Nscore_A(n)+Escore_A(e))

Precision(G,T) = Gscore(G,T)/Gscore(T,T)

Recall(G,T) = Gscore(G,T)/Gscore(G,G)
```

## Example 1: STEP '09

Six paragraphs submitted by six different research groups

Gold standard built by hand and verified, including word senses (all words including quantifiers, predicates, operators, ...), semantic roles

But not "deeper" content, including coreference, temporal relations, etc

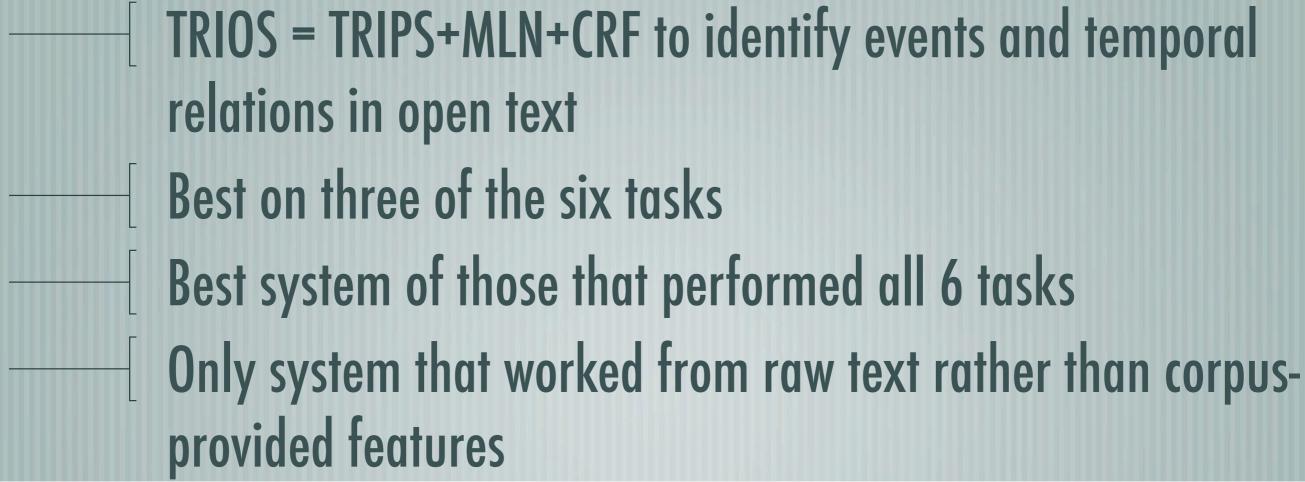
	Initial Baseline System	Baseline System II	w/ NER, POS, and UKW lookup	w/ constituent advice from Stanford Parser
Prec.	61.40%	74.4%	78.2%	79.0%
Recall	67.20%	74.4%	82.6%	82.8%
F-score	64.2%	74.4%	80.3%	80.9%

Amid the tightly packed row houses of North Philadelphia, a pioneering urban farm is providing fresh local food for a community that often lacks it, and making money in the process. Greensgrow, a one-acre plot of raised beds and greenhouses on the site of a former steel-galvanizing factory, is turning a profit by selling its own vegetables and herbs as well as a range of produce from local growers, and by running a nursery selling plants and seedlings. The farm earned about \$10,000 on revenue of \$450,000 in 2007, and hopes to make a profit of 5 percent on \$650,000 in revenue in this, its 10th year, so it can open another operation elsewhere in Philadelphia.

### Performance in TempEval-2 (2010)

Task	TRIOS Performance
Temp exp extraction	Second best
Event extraction	Third best
Temp relation event and temp exp	Best
Temp relation event and DCT	Second best
Temp relation main events consecutive sentence	Best
Temp relation events in same sentence	Second best

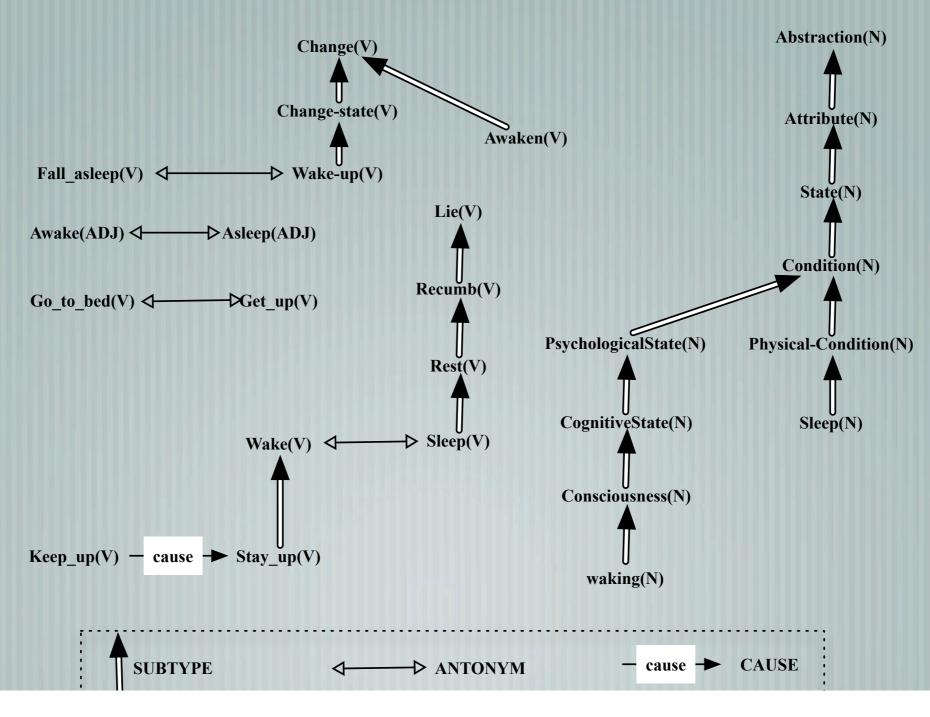
Table 1: Performance of TRIOS in TempEval-2

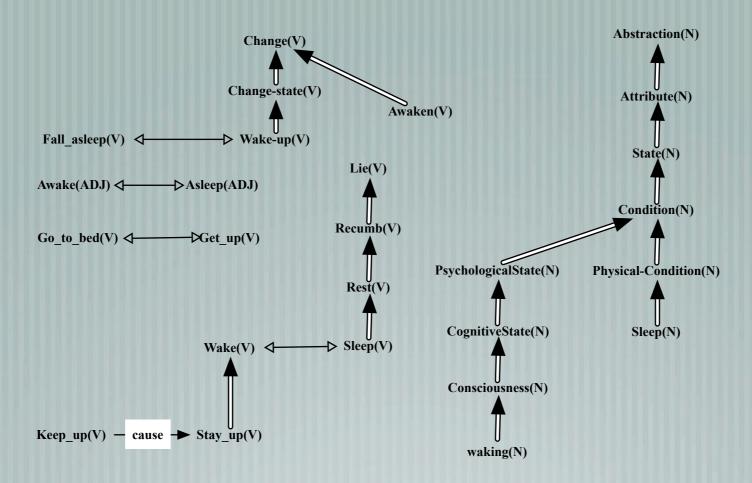


# Extracting Knowledge from WordNet

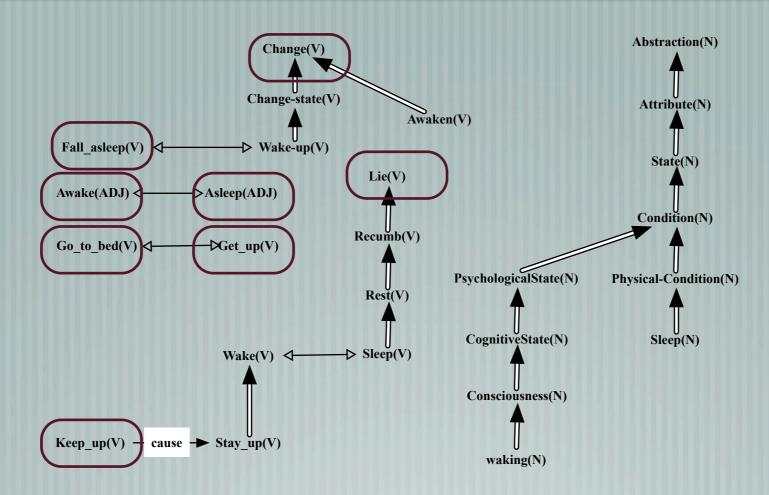
### Using Wordnet as a knowledge source

#### What WordNet says about Sleeping ...

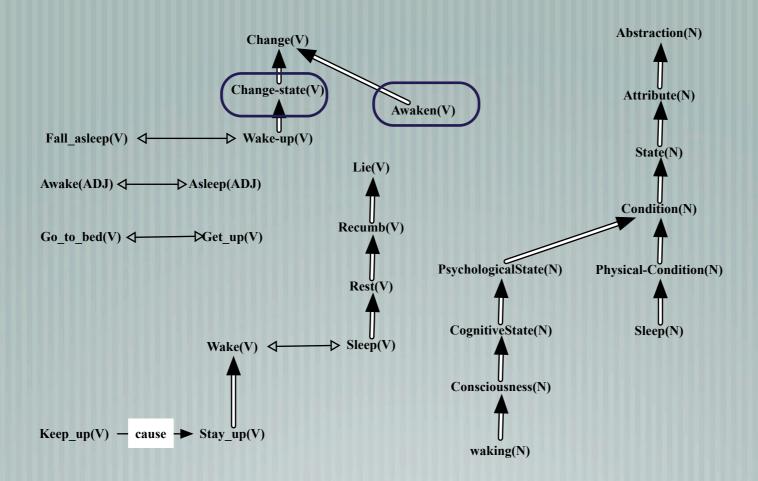




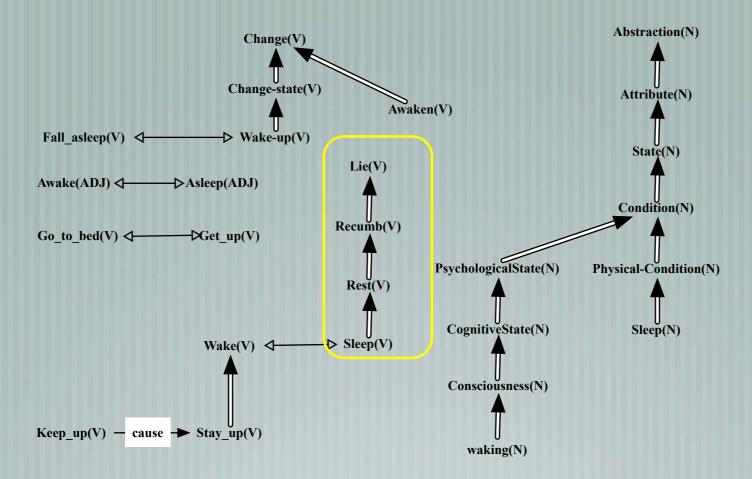
Shallow hierarchies except for nouns



Shallow hierarchies except for nouns Many top-level (uninterpreted) nodes



Shallow hierarchies except for nouns Many top-level (uninterpreted) nodes Fine-grained near synonyms



Shallow hierarchies except for nouns Many top-level (uninterpreted) nodes Fine-grained near synonyms Unclear what hierarchies mean

### Glosses in Wordnet have much richer information

Synset Name	Gloss
wake_up%2:29:00	to stop%2:42:00 sleeping%1:09:00
awaken%2:29:00	cause to become awake or conscious%3:00:00
sleeping%1:09:00	the state%1:03:00 of being asleep%4:02:00
asleep%4:02:00	into a sleeping state%1:03:00
fall_asleep%2:29:00	to change%2:30:00 from a waking%1:09:00 to a sleeping [sleep%2:29:00] state%1:03:00
waking%1:09:00	the state%1:03:00 of remaining awake
wake%2%29%0	Be awake, be alert, be there
sleep%2:29:00	to be asleep%4:02:00
awake%3:00:00	not in a state%1:03:00 of sleep%1:26:00
asleep%3:00:00	in a state%1:03:00 of sleep%1:26:00
sleep%1:26:00	a natural and periodic state%1:03:00 of rest during which consciousness%1:09:00 of the world %1:09:00 is suspended [suspend%2:30:02]
keep_up%2::29:00	prevent%2:41:00 from going to bed [go_to_bed%2:29:00]
stay_up%2:29:00	Not go to bed [go_to_bed%2:29:00]
go_to_bed%2:29:00	to prepare for sleep%2:29:00

#### Reading this, we (as people) learn facts such as

- before falling asleep, you are awake, and afterwards you are asleep
- if you awaken someone then they wake up

## Our target: OWL-DL

- OWL-DL is designed to be expressive as possible while still allowing provably efficient reasoning algorithms

  Concepts are arbitrary compositions of constraints, such as

  subtype constraints: e.g.,
  HUMAN is a subtype of ANIMAL
  - All LIVING-THINGS are either a PLANT or ANIMAL
  - role constraints, e.g.,
    - whose sex role is either MALE or FEMALE
    - who has exactly one head
    - whose every child is a HUMAN
  - and equality, e.g.,
    - BACHELOR = [PERSON :sex MALE :marital-status SINGLE]

"do away with, cause the destruction or undoing of"

"destroy completely; damage irreparably"

"defeat soundly"

"put (an animal) to death"

```
"do away with, cause the destruction or undoing of"

"destroy completely; damage irreparably"
```

defeat soundly"

"put (an animal) to death"

Multiple definitions in one gloss

"do away with, cause the destruction or undoing of"

"destroy completely; damage irreparably"

defeat soundly"

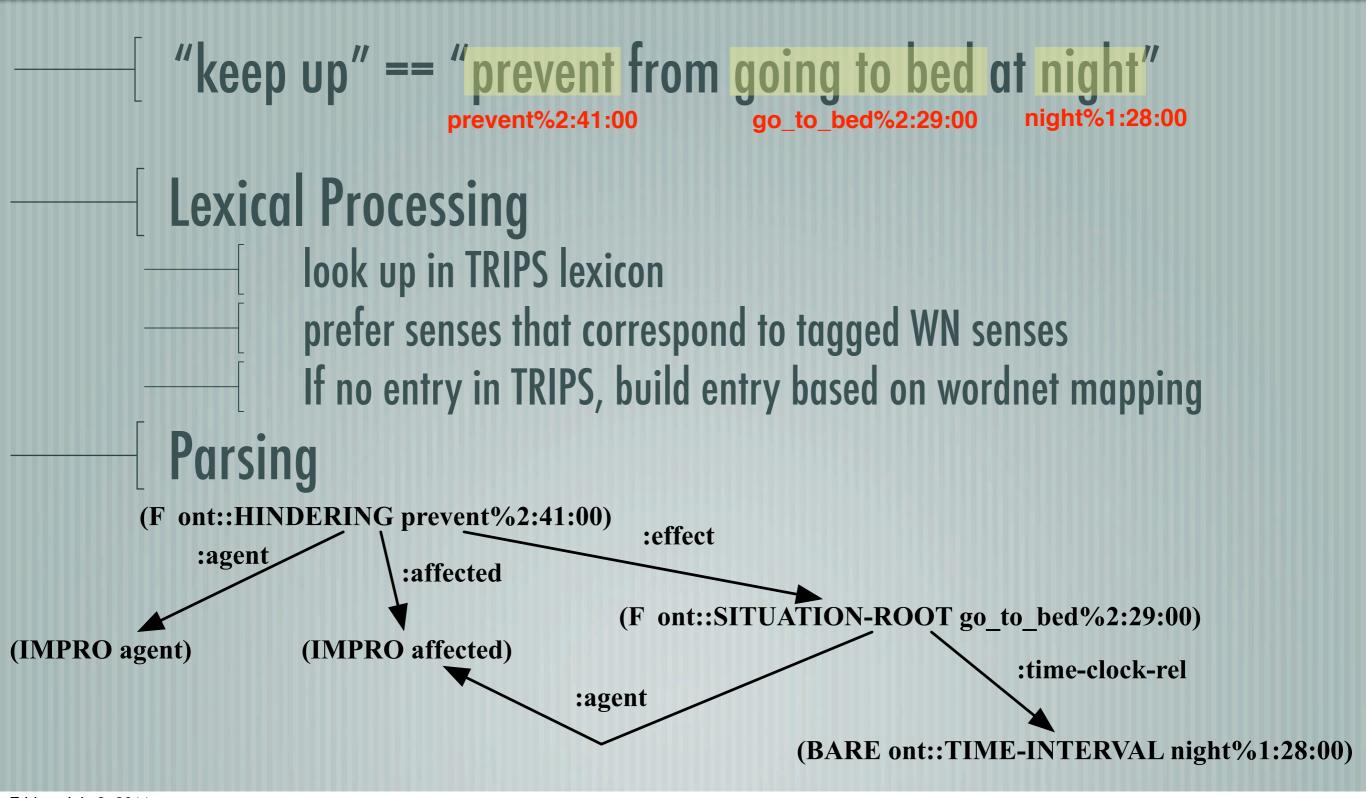
full put (an animal) to death"

Multiple definitions in one gloss Disjunctions and conjunctions

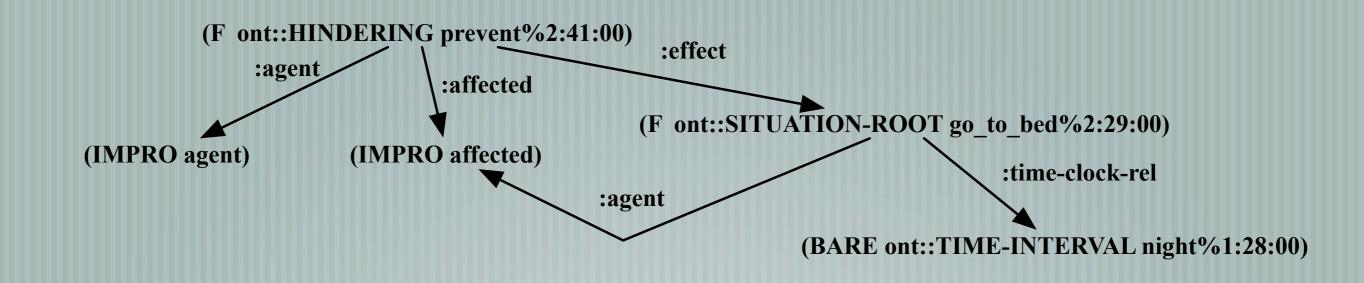
```
do away with, cause the destruction or undoing of"
"destroy completely; damage irreparably"
"defeat soundly"
"put (an animal) to death"
 Multiple definitions in one gloss
 Disjunctions and conjunctions
```

Gaps & ellipsis

### An Example:Defining a new concept

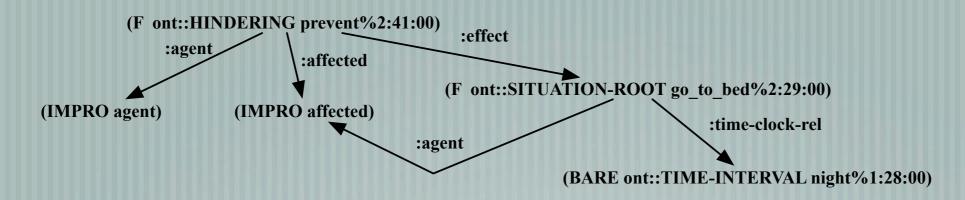


### Determining Roles



"Keep up" = "prevent from going to bed"
 acquires :agent and :affected role from ONT::HINDERING
 but no :effect role as that is explicitly filled in the definition
 We verify the hypothesized roles by parsing the examples that Wordnet provides

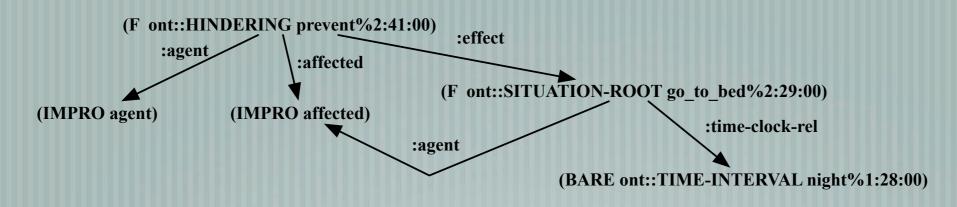
# Details: Constructing a Definition in OWL



#### **Constructing the Definition**

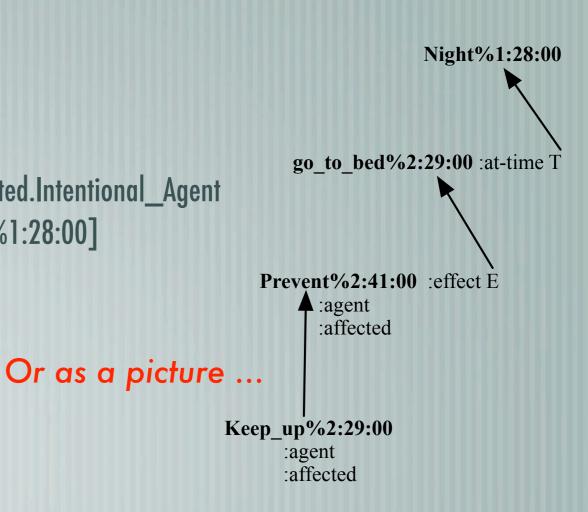
```
Keep_up%2:29:00 \subseteq Prevent%2:41:00 \cap \exists agent.Intentional_Agent \cap \exists affected.Intentional_Agent \cap \exists effect.[Go_to_bed%2:29:00 \cap \exists at-time.Night%1:28:00]
```

# Details: Constructing a Definition in OWL



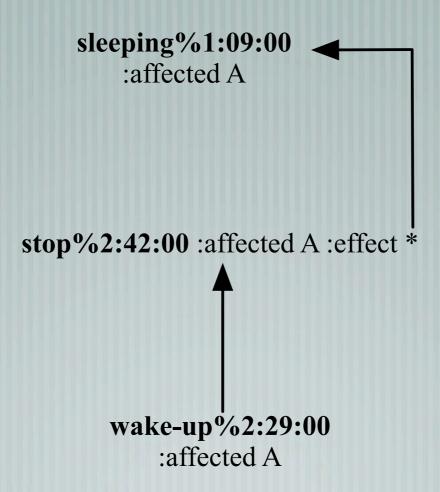
#### **Constructing the Definition**

Keep\_up%2:29:00  $\subseteq$  Prevent%2:41:00  $\cap$   $\exists$  agent.Intentional\_Agent  $\cap$   $\exists$  affected.Intentional\_Agent  $\cap$   $\exists$  effect.[Go\_to\_bed%2:29:00  $\cap$   $\exists$  at-time.Night%1:28:00]

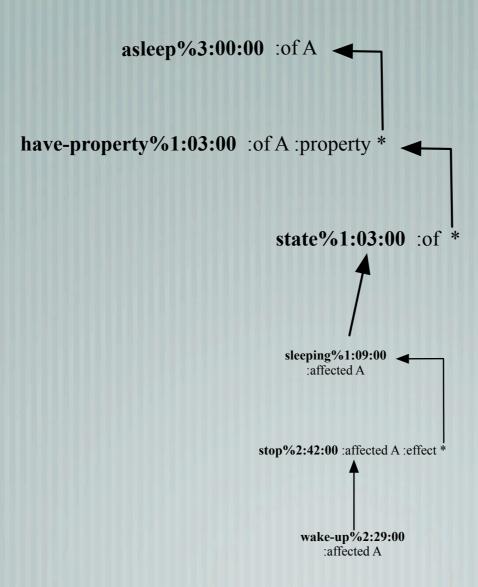


Synset Name	Gloss		
wake_up%2:29:00	to stop%2:42:00 sleeping%1:09:00		
awaken%2:29:00	cause to become awake or conscious%3:00:00		
sleeping%1:09:00	the state%1:03:00 of being asleep%4:02:00		
asleep%4:02:00	into a sleeping state%1:03:00		
fall_asleep%2:29:00	to change%2:30:00 from a waking%1:09:00 to a sleeping [sleep		
	%2:29:00] state%1:03:00		
waking%1:09:00	the state%1:03:00 of remaining awake		
wake%2%29%0	Be awake, be alert, be there		
sleep%2:29:00	to be asleep%4:02:00		
awake%3:00:00	not in a state%1:03:00 of sleep%1:26:00		
asleep%3:00:00	in a state%1:03:00 of sleep%1:26:00		
sleep%1:26:00	a natural and periodic state%1:03:00 of rest during which consciousness		
	%1:09:00 of the world%1:09:00 is suspended [suspend%2:30:02]		
keep_up%2::29:00	prevent%2:41:00 from going to bed [go_to_bed%2:29:00]		
stay_up%2:29:00	Not go to bed [go_to_bed%2:29:00]		
go_to_bed%2:29:00	to prepare for sleep%2:29:00		

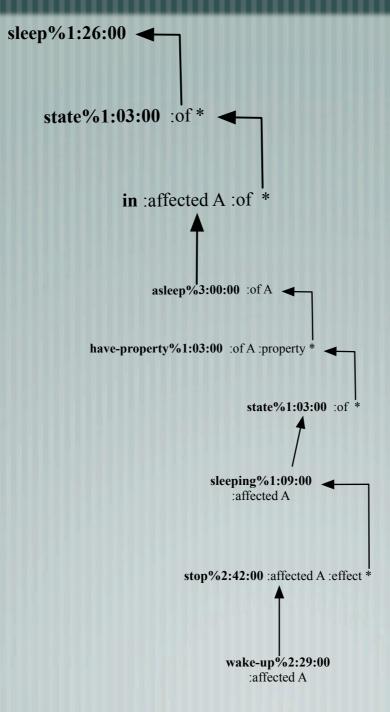
wake\_up%2:29:00 = to stop%2:42:00 sleeping%1:09:00



wake\_up%2:29:00 = to stop%2:42:00 sleeping%1:09:00

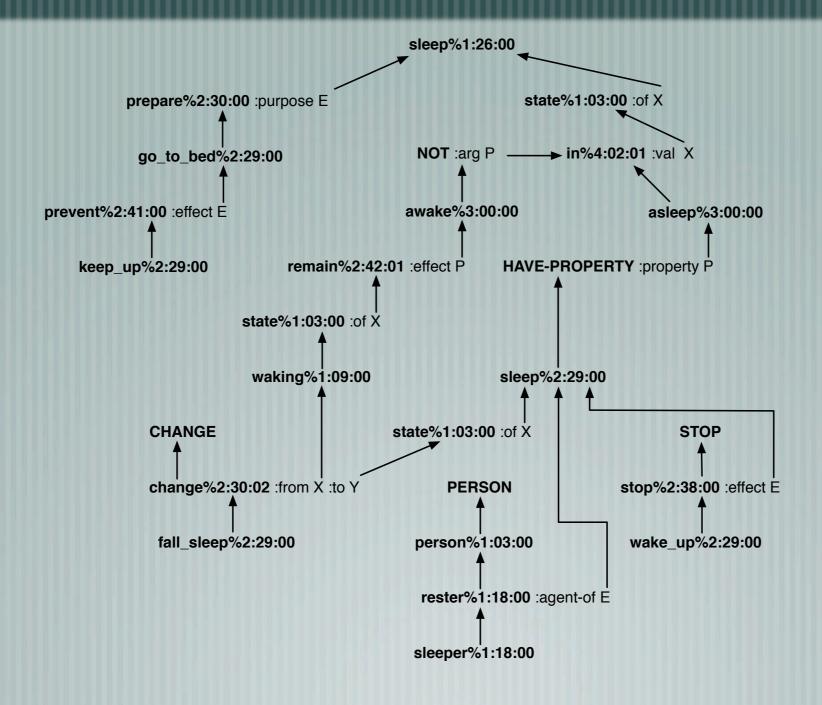


sleeping%1:09:00=the state%1:03:00 of being asleep%4:02:00

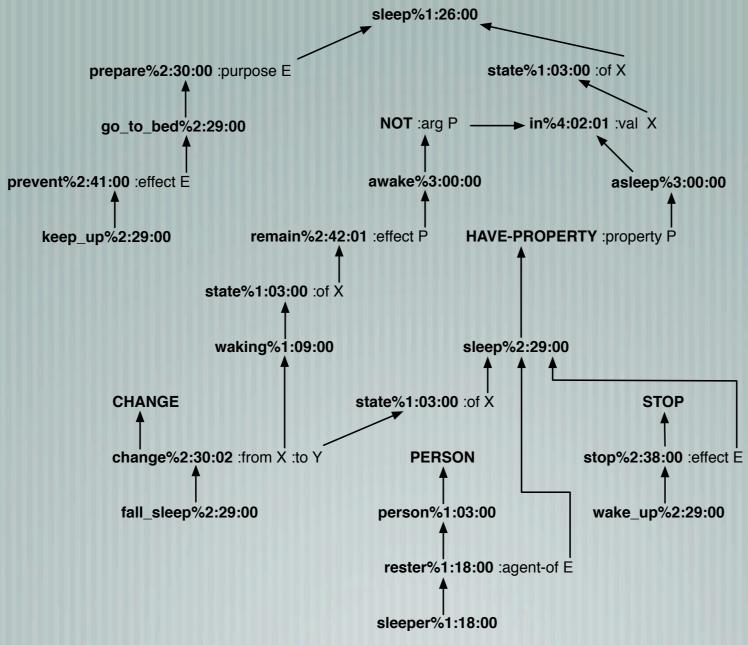


asleep%3:00:00=in a state%1:03:00 of sleep%1:26:00

## A Slightly Simplified KB from the definitions



# A Slightly Simplified KB from the definitions



once we have this, what can we do with it?

not much yet! Need a model of entailment

# Filling the gaps

Note that we didn't expand out definitions of terms like stop, change, cause, remain, ...

# Filling the gaps

Note that we didn't expand out definitions of terms like stop, change, cause, remain, ...

Why?

stop = come to a halt, stop moving halt = the event of something ending ending = be the end of end = the point in time when something ends

We don't learn what stopping means from reading!!

# Inference

Need an assertional layer to capture common-sense knowledge about basic causality, time, etc

Use variant Allen's temporal logic

T(<concept>, <time>)

— meaning <concept> HOLDS/OCCURS at time t

# Axiomatization

```
"when e stops it isn't occurring any more"

T([stop%2:38:00 :effect e], t) ⊃ ∃ t'. Meets(t, t') ^ ~T(e, t')

"If X is in a state of e, the e is occurring"

T([in%4:02:01 :val [state%1:03:00 :of e]], t) ⊃ T(e, t)

"If X changes to state Z, then X in in state Z afterwards"

T([change%2:30:02 :to z], t) ⊃ ∃ t' . t meets t' ^ T(z, t')
```

# Implementation

- Common-sense axioms encoded in SWRL
- Entailments arise as events are instantiated:
  - create event from parse
  - repeat as needed
    - instantiate new events entailed from definitions
    - apply SWRL rules to new events

#### Adding Logical Form

"John fell asleep"

Fall\_Asleep%2:29:00

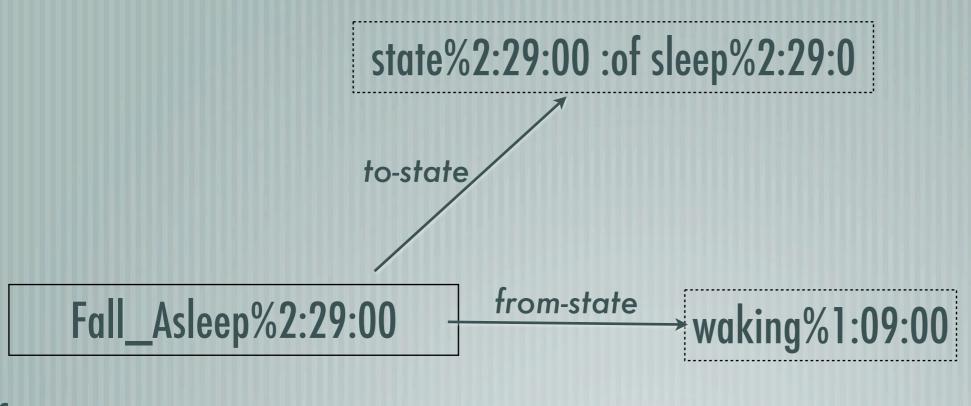
Concepts

**Timeline** 

Fall\_Asleep%2:29:00

# Simplified Example inherit properties from Defn

"John fell asleep"



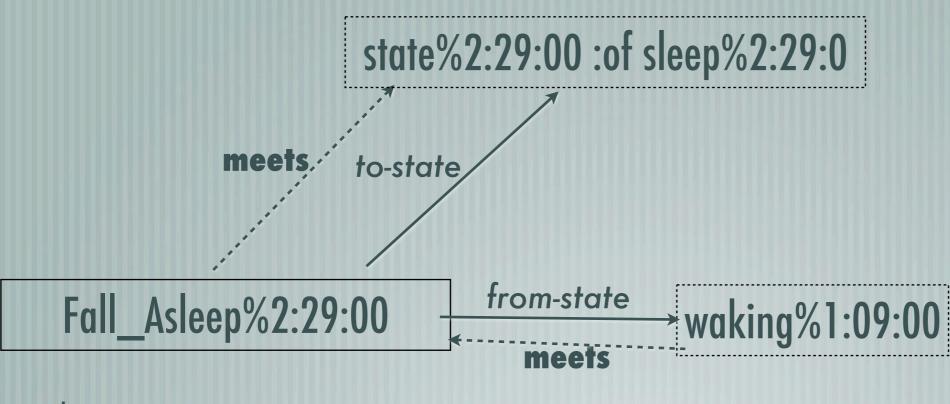
Concepts

**Timeline** 

Fall\_Asleep%2:29:00

#### **SWRL** inference

"John fell asleep"



Concepts

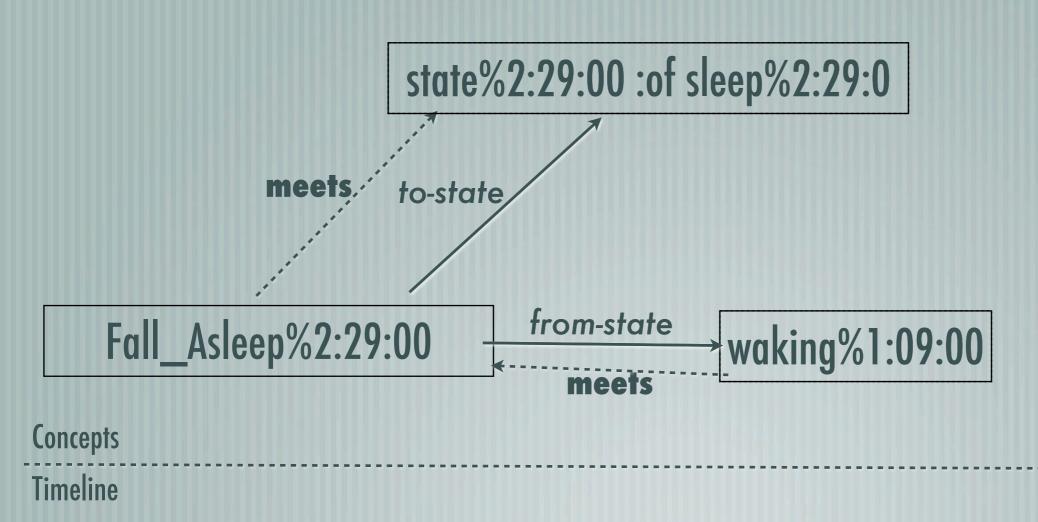
**Timeline** 

Fall\_Asleep%2:29:00

waking%1:09:00

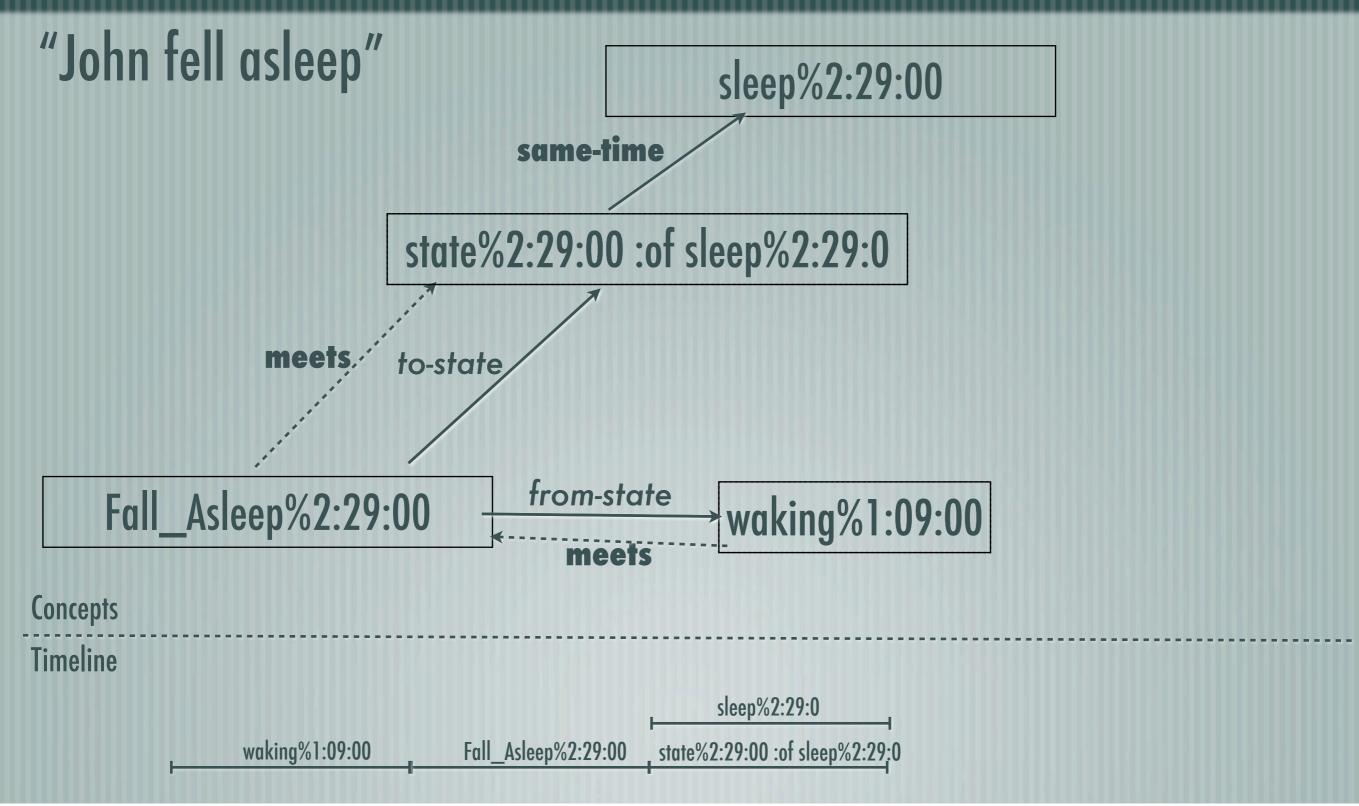
#### Instantiate events

"John fell asleep"

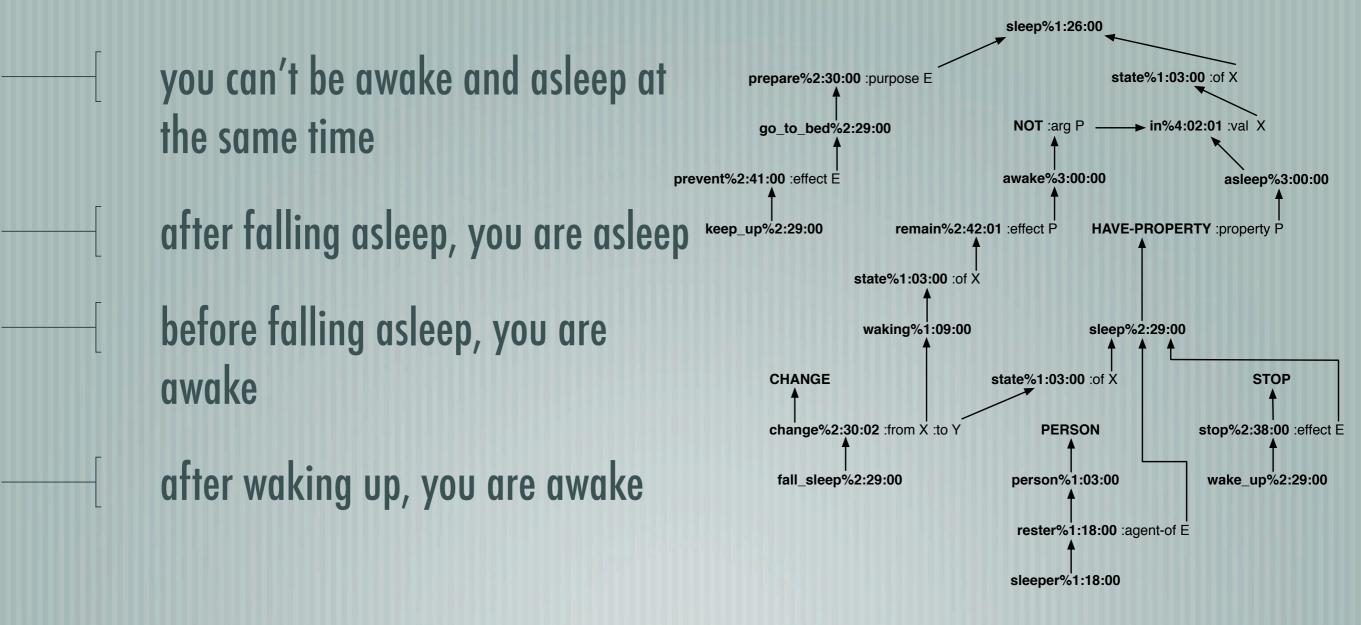


Fall\_Asleep%2:29:00 state%2:29:00 :of sleep%2:29:0

#### **SWRL** inference

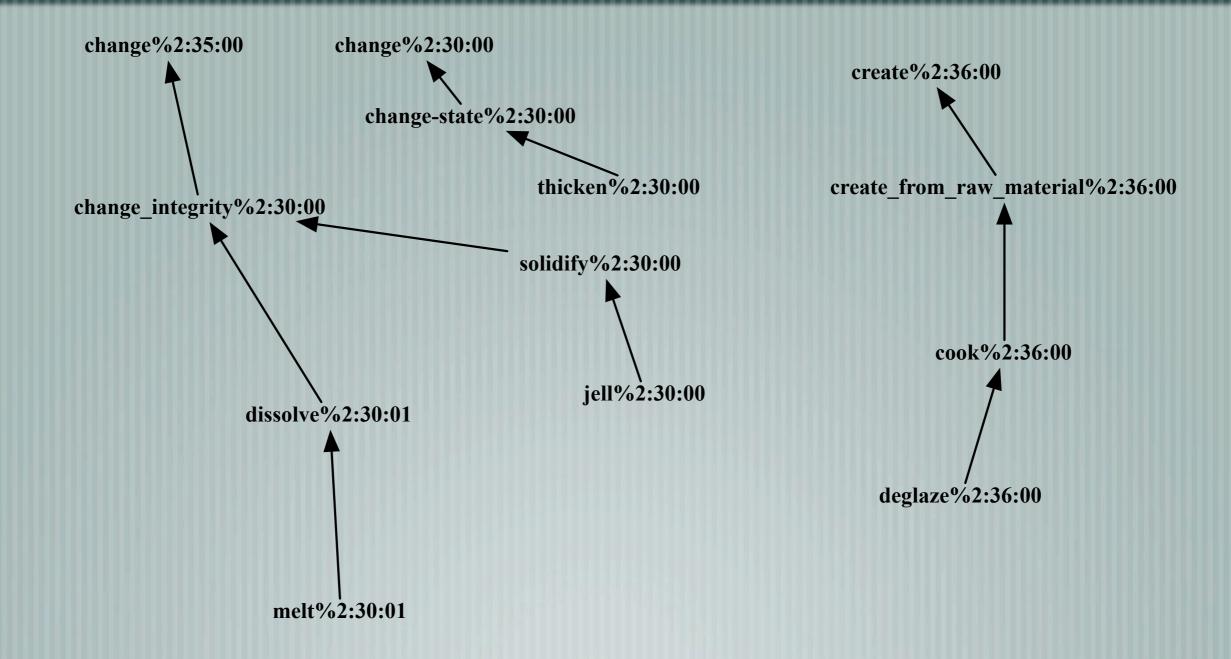


# Entailments from the KB



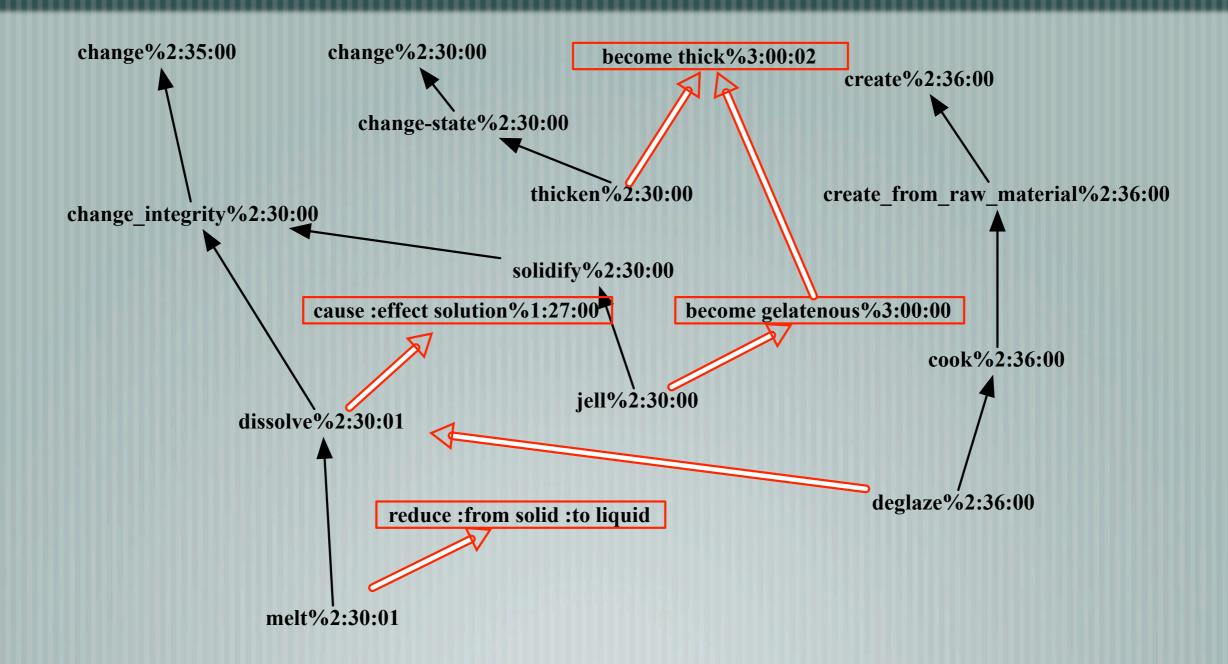
But not "if you are kept up, you are awake" keep up = prevent from going to bed

# Another Example: knowledge about cooking



The WordNet hierarchy for some selected cooking terms (out of 7318!)

# Another Example: knowledge about cooking



The WordNet hierarchy for some selected cooking terms (out of 7318!)

Augmented by understanding the glosses

# Evaluation

- Ultimate evaluation is a demonstration the knowledge we learn is useful
  - entailment tests, reading comprehension tests, ...
  - improved language handling: word sense disambiguation, discourse, ...
- Not there yet, so some preliminary explorations...
- do the wordnet hierarchies provide useful knowledge?
- can we build reasonable knowledge using glosses

# Example: Exploring the Noun Hierarchy

The noun hierarchy is the best developed in Wordnet, by far, so it a good place to explore capabilities

We evaluate our ability to understand glosses by building a new hierarchy based on glosses alone

#### Questions:

can parsing glosses recover the existing hierarchical information in WordNet?

## Learning Hierarchical Relations

#### Experiment

randomly select 100 nouns, with 20 known WN subtype pairs collect the glosses, and definitions of concepts used in the glosses

process and test subsumption between each pair of concepts

use WN hierarchy as the gold standard

## Learning Hierarchical Relations

#### Experiment

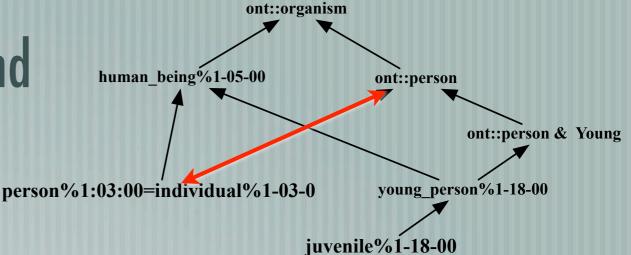
randomly select 100 nouns, with 20 known WN subtype pairs
 collect the glosses, and definitions of concepts used in the glosses
 process and test subsumption between each pair of concepts
 use WN hierarchy as the gold standard

Initial Results precision: 100% recall 44%

TRIPS	TRUE	FALSE
X subtype Y	13	0
no reln	9975	16

# Problems Revealed

Distinguishing being necessary and necessary & sufficient conditions



Complex definitions, requiring better parsing and/or more common sense

common-fraction%1-23-00 = "the quotient of two integers" one\_tenth%1:23:0 = "a tenth part; one part in ten equal parts"

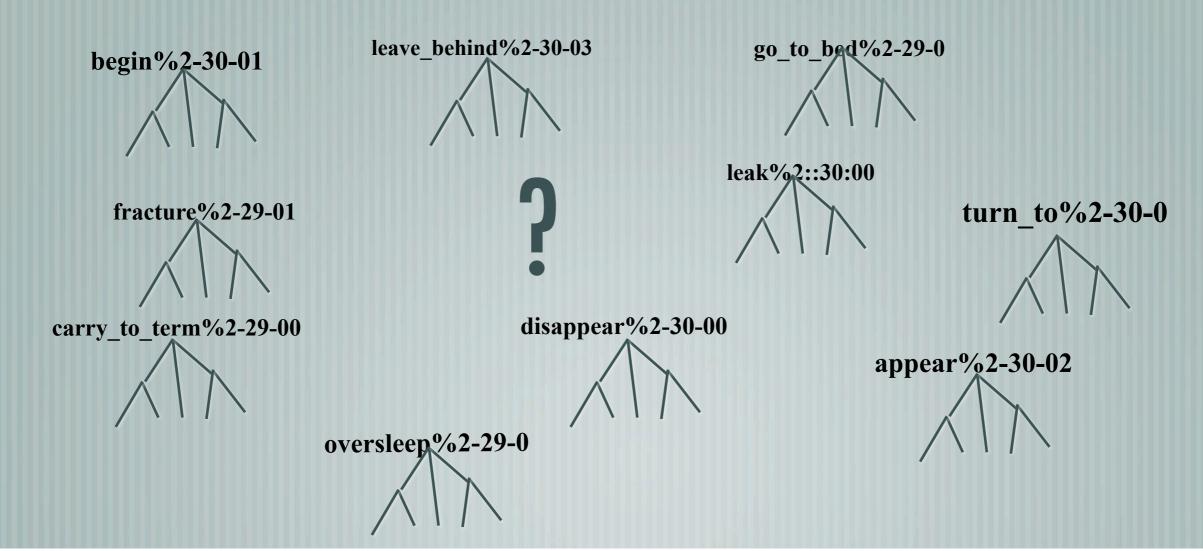
Questionable definitions/relations in WordNet

an emirate is a position (not true in glosses or real life)?? a shelf is a device??

# Example: Filling in the Wordnet Verb Hierarchy

Its not a hierarchy, but a forest of subtrees

550+ "top level" verbs in WN that have no supertypes



# Can we create a useful hierarchy from glosses?

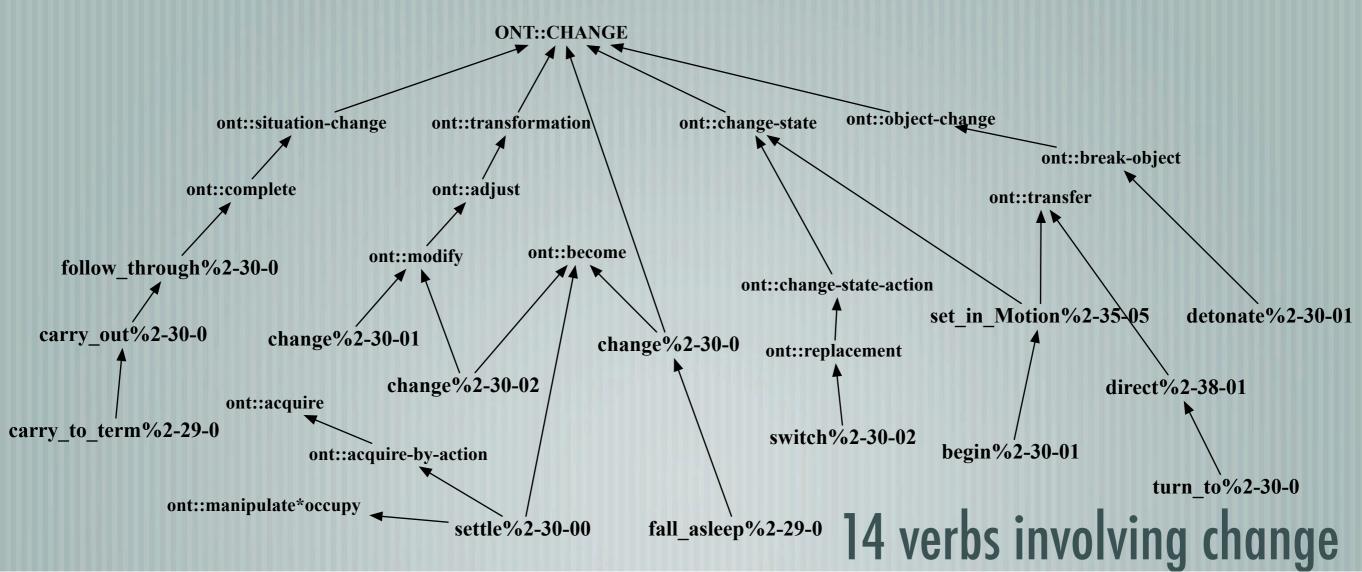
An Exploration

- select 50 "top-level" verb senses from Wordnet
- classify them into the TRIPS upper ontology based on the glosses

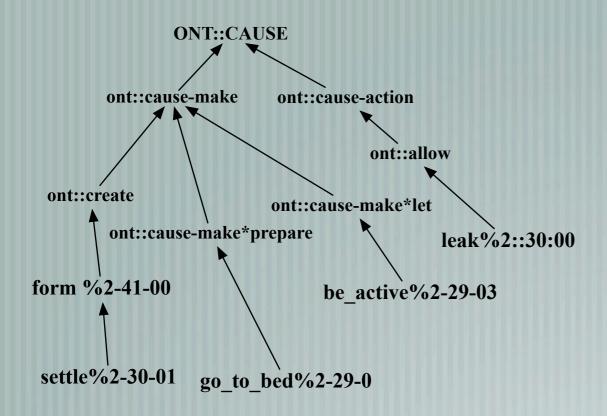
# Can we create a useful hierarchy from glosses?

#### An Exploration

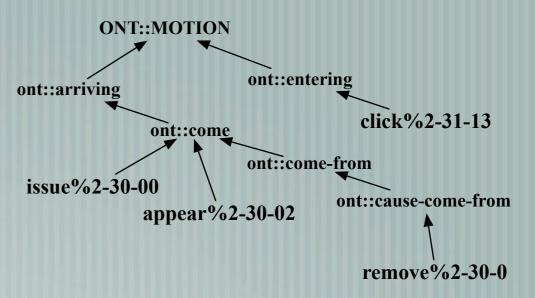
- select 50 "top-level" verb senses from Wordnet
- classify them into the TRIPS upper ontology based on the glosses



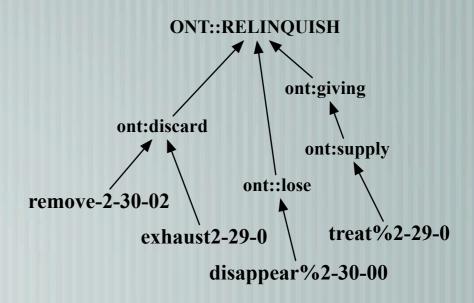
### Verb Hierarchies



5 verbs of causation



#### 4 verbs of motion



#### 4 verbs of relinquishing

# Can we make this work?

# 1. Better Deep Parsing

#### Better handling of conjunctions/disjunctions

- "highly seasoned Mediterranean soup or stew made of several kinds of fish and shellfish
- with tomatoes and onions or leeks and seasoned with saffron and garlic and herbs"
- current approach: 1) try full parse, 2) simplification edits, 3) robust extraction
- what's needed: better preprocessing to tag scope of conjuncts using similarity measures

#### Better word sense disambiguation

- current approach: 1) prefer explicit wordnet tag with backoff, 2) prefer most common senses in each syntactic class
- bootstrapping common semantic combinations that provide preferences

#### Better semantic role assignment/TRIPS ontology

reworking roles in context of reworking upper ontology (based on GUM)

# 2. Cleaning up the Knowledge Acquired

# Wordnet has many redundant concepts, e.g., — "young person" -> young\_person%1-18-00 "a young person (especially a young man or boy) — approach — check if meaning of the phrase subsumes the meaning of the gloss — if so, we can remove the defa ("young person" as a multiword)

- if so, we can remove the defn ("young person" as a multiword)
   "go to bed" -> go\_to\_bed%2-29-00 is retained since no meaning of
  - "go" and "bed" can subsume "prepare for sleep"
- Other senses may remain, but to a composite concept, not an atomic one
  - "spring chicken" (a synonym of "young person") -> PERSON ^
    YOUNG

# 3. Clustering and Abstraction of Senses

Once definitions are in OWL, we can define a semantic similarity metric between concepts

- combination of true subsumption (rare), class similarity ("size" of most specific subsuming class), and role similarity (requiring recursive class similarity)
- Standard clustering algorithms can then suggest sense merges, and we use KB to derive a new common subsuming class that captures the new sense

# 4. Wordnet as a Source of Knowledge?

```
Wordnet misses much commonsense knowledge, e.g.,
      keep_up%2:29:00 = "to be asleep%3:00:00"
      asleep%3:00:00 = "into a sleeping state"
      but what is sleeping???
Unhelpful definitions
      Circular Definitions
          stop%2:41:00 = "to stop from happening or developing"!
Still, Wordnet includes large amounts of knowledge if it
  can be extracted intelligently
      And there are many other sources of definitions (other dictionaries,
      Wikipedia, ...)
```

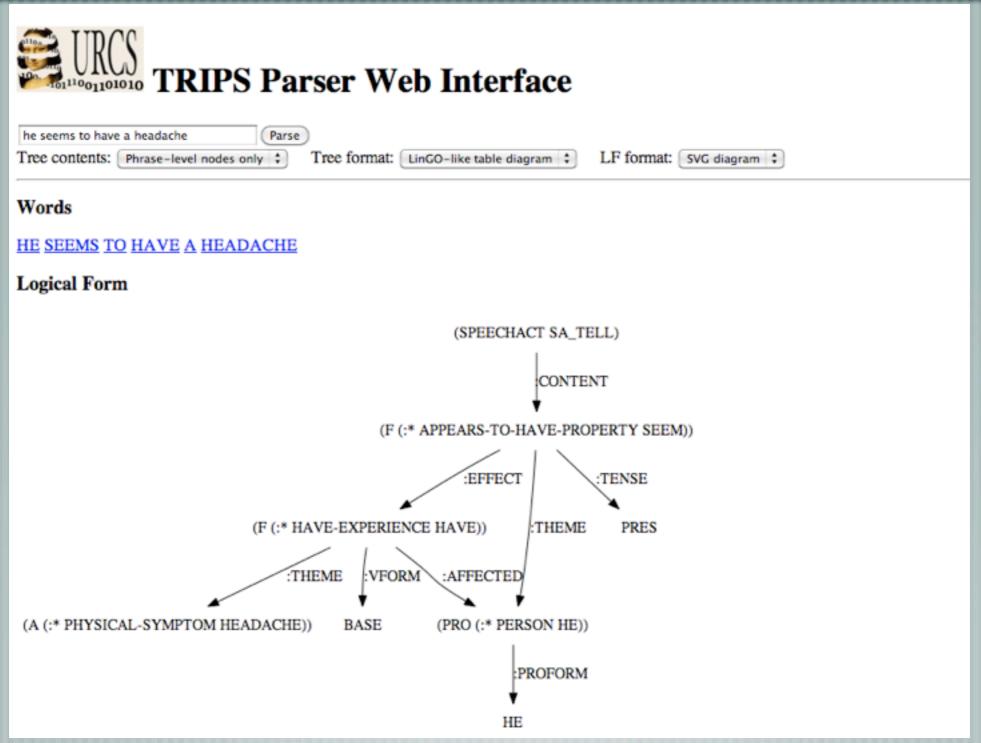
# Concluding Thoughts

We hope we have established the feasibility of building rich knowledge bases for entailment from online resources such as Wordnet and dictionaries

To obtain high quality knowledge we need

- better handling of complex phenomena in parsing especially conjunction and ellipsis
- reasoning-based cleaning techniques to detect and resolve inconsistencies, and generalize knowledge

# Try it out!



www.cs.rochester.edu/research/cisd/projects/trips/parser/cgi/web-parser-xml.cgi

# Try it out!



Look Up stop

#### stop

Noun Classes:

ONT::location (show synset) (show ancestors) Example: go to this stop

Verb Classes:

ONT::stop (show synset) (show ancestors)

Frames:

cause-effect-xp-templ (xp (% vp (vform ing)))

Example: he stopped the rioting

affected-templ

Example: he/the truck stopped

affected-effect-xp-templ (xp (% vp (vform ing)))

Example: the computers/managers stopped working

cause-affected-xp-templ

Example: he/the storm stopped the fair/the truck

effect-templ

Example: the rioting stopped

cause-effect-affected-objcontrol-templ (xp (% cp (ctype s-from-ing)))

Example: the hospital stops visitors from smoking

Data files last modified: Sun Jul 3 03:19:02 2011



### TRIPS Ontology Browser

Search for ONT type V root V any-sem > any-time-object V referential-sem (entity%1:03:00) (show words) > abstract-object (psychological feature%1:03:00, abstraction%1:03:00) > part (part%1:24:00) (show words) > phys-object (object%1:03:00) (show words) V situation-root accommodate (adapt%2:30:01) (show words) > acquire (take%2:33:08, take%2:35:14, take%2:40:08, subscribe%2:40:00, claim%2:32:01, accept%2:40:00, take%2:40:00, assume%2:30:00, get%2:40:00) (show words) V acting (act%2:29:00, act%1:03:00) (show words) > cause-effect (causal agent%1:03:00, cause%1:10:00, cause%1:11:00, cause%2:36:00, produce%2:36:03) (show words) - fail (show words) V inhibit-effect disable (show words) > hindering (restrain%2:41:01, prevent%2:41:00, restrict%2:30:00, prevent%2:41:01, handicap%2:33:00) (show words) prevent (prevent%2:41:00, prevent%2:41:01, prevention%1:04:00) (show words) prohibit (show words) refuse (deny%2:40:00) (show words) stop (discontinue%2:42:00, end%2:42:00, run out%2:42:00) (show words) > activity (activity%1:04:00) > activity-ongoing (show words)

offact avnariancer (chow words)

www.cs.rochester.edu/research/trips/lexicon/browse-ont-lex.html