

---

# Ontology, Verb Meaning and Semantic Roles

---

Fernando Gomez

Department of EECS

University of Central Florida

---

# Verb Argument Structure

- However, the African sculptors often incorporate traditional images or themes into their works.(World Book Encyclopedia)

AGENT = “African sculptors”

THEME = “traditional images or themes”

DESTINATION = “into their works”

---

---

## Example 2

- **Grand Teton National Park was created in 1929, after efforts to incorporate the area in Yellowstone National Park had failed.(World Book Encyclopedia)**

**INANIMATE-CAUSE = efforts THEME = the area  
DESTINATION = Yellowstone National Park**

- **Current design incorporates desiccants to maintain the internal humidity to insure proper JM performance.( NASA)**

**INANIMATE-CAUSE = current design THEME =  
desiccants**

---

# Example 3

**The dog recognized his master when Ulysses returned home after an absence of 20 years. (World Book Encyclopedia)**

**Agent = Ulysses T0-LOC = home TEMP-ADJUNCT = "after an absence of 20 years"**

**This food, or cud, is later returned to the mouth in a ball-like glob, and the camel chews it. (World Book Encyclopedia)**

**THEME="food" TO-LOC = "the mouth" MANNER-ADJUNCT = "in a ball-like glob"**

**The ISS safer system, when returning on a Soyuz vehicle, shall withstand post-landing temperatures as low as -58 f. (NASA)**

**THEME = "ISS safer system" INSTRUMENT = "a Soyuz vehicle"**

**The Abort Motor igniter provides initiation energy to the Abort Motor (NASA)**

**INANIMATE-CAUSE = "The Abort Motor igniter" Theme = "initiation energy" TO-POSS = "the Abort Motor"**

---

# Statistical Parsers

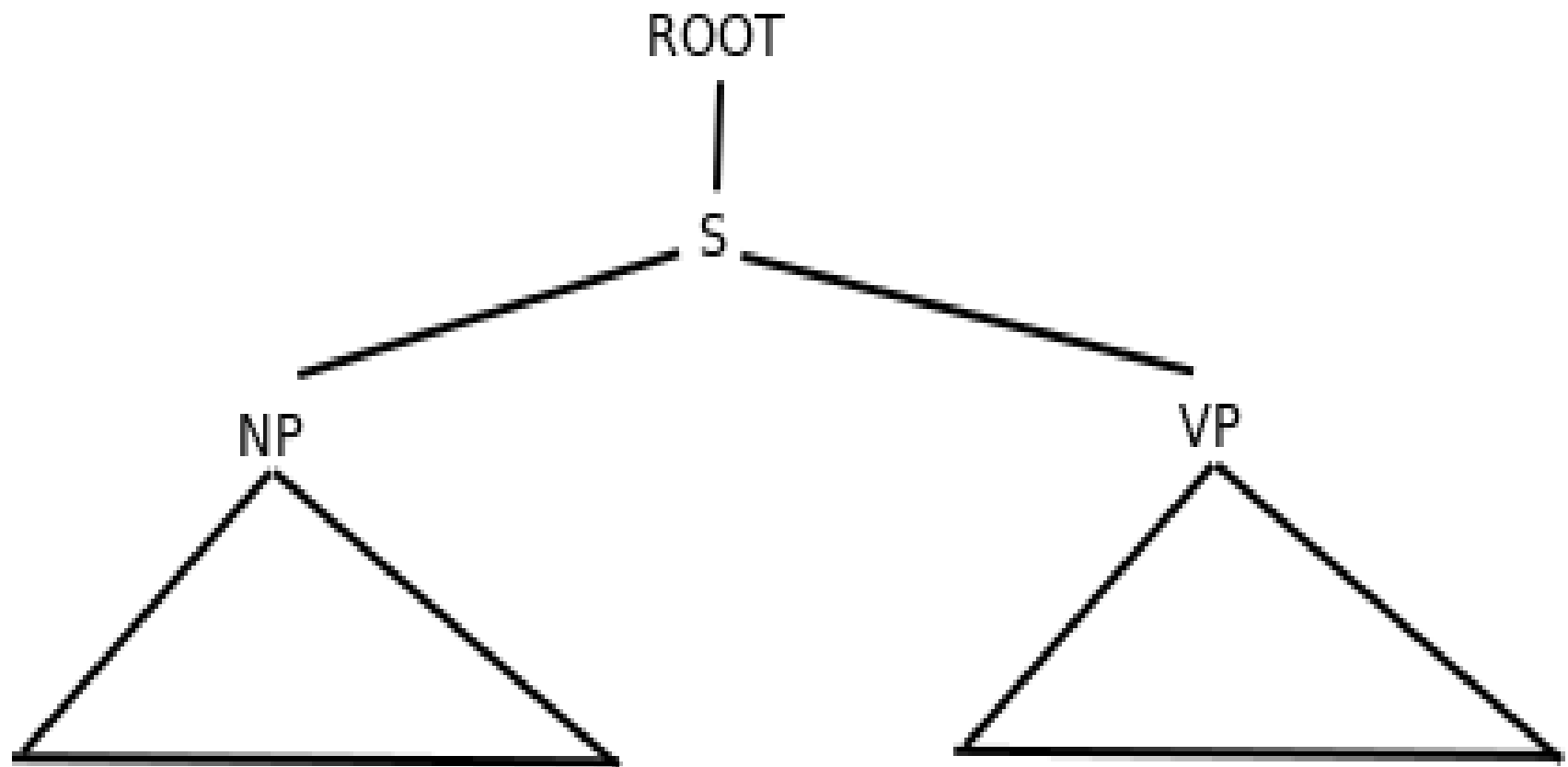
- **Statistical Parsers provide an impoverished parse tree in which long distance dependency and missing elements on the tree are not indicated, let alone filled with their referents.**
  - **Linguists refer to these missing elements, or null elements, as empty categories.**
  - **The recovery of these missing elements is a necessary condition for extracting semantic information from a parse tree.**
-

---

# Two Statistical Parsers

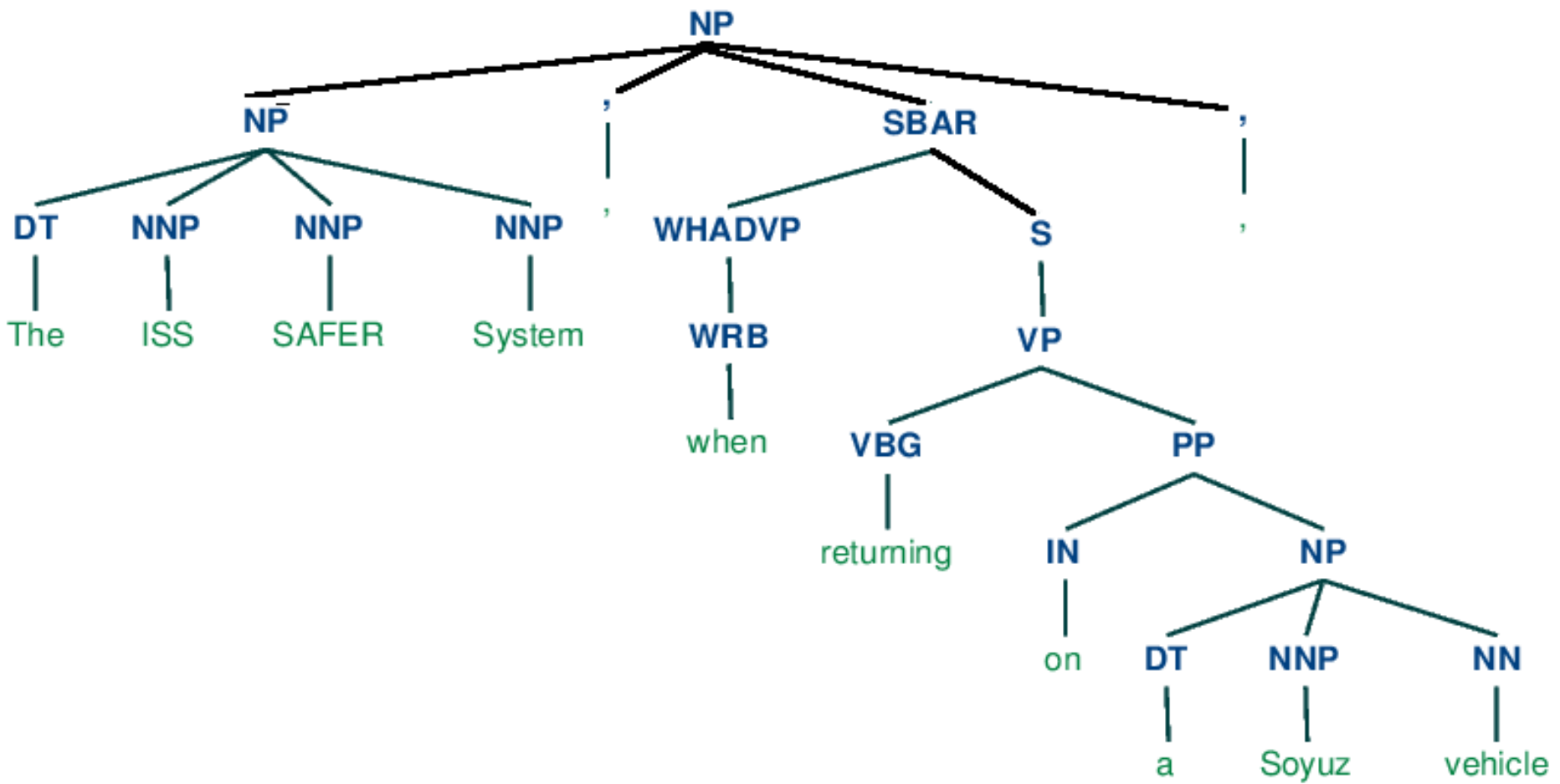
- Charniak Parser (E. Charniak “A maximum-entropy inspired parser, NAACL 2000)
- Stanford Parser (D. Klein and Chris Manning “Accurate Unlexicalized Parser” ACL 2003)



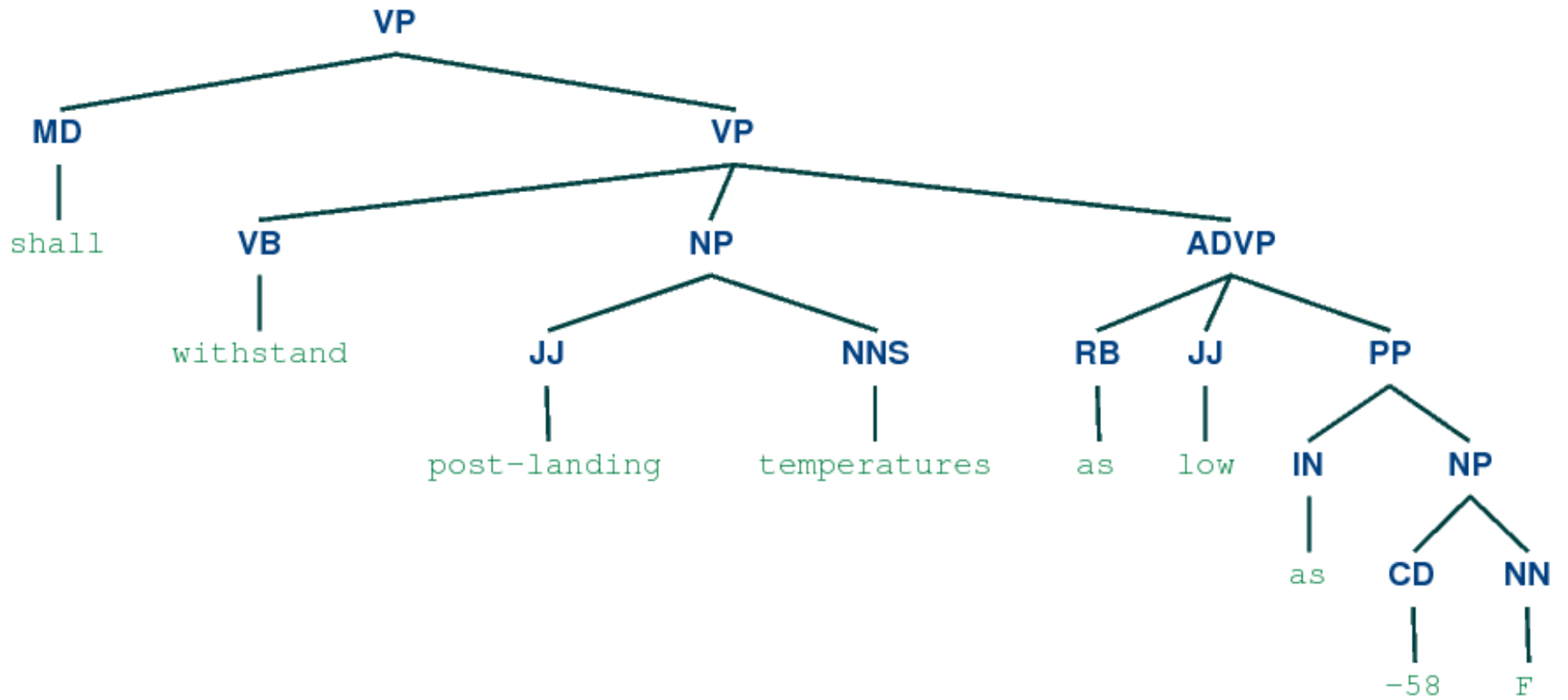


The ISS SAFER System, when returning on a Soyuz vehicle,

shall withstand post-landing temperatures as low as -58 F.







# Examples of Parses by Charniak's Parser (Charniak 2000)

- "The ISS safer system, when returning on a Soyuz vehicle, shall withstand post-landing temperatures as low as -58 f."
  - "(S1 (S (NP (DT The) (JJ ISS) (JJR safer) (NN system)) (, ,) (SBAR (WHADVP (WRB when)) (S (VP (VBG returning) (PP (IN on) (NP (DT a) (NNP Soyuz) (NN vehicle)))))) (, ,) (VP (MD shall) (VP (VB withstand) (NP (JJ post-landing) (NNS temperatures)) (ADVP (ADVP (RB as) (JJ low)) (PP (IN as) (NP (CD -58) (NN f)))))) (. .)))"
  - He often accused some newspapers and TV networks of presenting news in a way that was prejudiced against the Administration. (World Book Encyclopedia)
  - "(S1 (S (NP (PRP He)) (ADVP (RB often)) (VP (VBD accused) (NP (NP (DT some) (NNS newspapers)) (CC and) (NP (NN TV) (NNS networks))) (PP (IN of) (S (VP (VBG presenting) (NP (NN news)) (PP (IN in) (NP (NP (DT a) (NN way)) (SBAR (WHNP (WDT that)) (S (VP (AUX was) (ADJP (VBN prejudiced) (PP (IN against) (NP (DT the) (NNP Administration)))))))))) (. .)))"
  - "She visited Spain to give a concert."
- "(S1 (S (NP (PRP she)) (VP (VBD visited) (NP (NNP Spain)) (S (VP (TO to) (VP (VB give) (NP (DT a) (NN concert)))))) (. .)))"

---

# Looking Past the Verb

- **"Widely expected to open 10 to 15 cents a barrel higher on the strength of statistics from the American Petroleum Institute , the December contract managed to start the session only eight cents higher ." (Ontonotes)**

"(S1 (S (S (ADVP (RB Widely)) (VP (VBN expected) (S (VP (TO to) (VP (VB open) (ADVP (NP (NP (QP (CD 10) (TO to) (CD 15)) (NNS cents)) (NP (DT a) (NN barrel))) (RBR higher)) (PP (IN on) (NP (NP (DT the) (NN strength)) (PP (IN of) (NP (NP (NNS statistics)) (PP (IN from) (NP (DT the) (NNP American) (NNP Petroleum) (NNP Institute)))))))))) (, ,) (NP (DT the) (NNP December) (NN contract)) (VP (VBD managed) (S (VP (TO to) (VP (VB start) (NP (DT the) (NN session)) (ADVP (NP (QP (RB only) (CD eight)) (NNS cents)) (RBR higher)))))) (. .)))")"

---

# Object Gap

- ***It is also unclear whether they made their own raw glass or imported the pre-made ingots, which they melted and finished.*** (World Book Encyclopedia)

***"(S1 (S (NP (PRP It)) (VP (AUX is) (ADVP (RB also)) (ADJP (JJ unclear)) (SBAR (IN whether) (S (NP (PRP they)) (VP (VP (VBD made) (NP (PRP\$ their) (JJ own) (JJ raw) (NN glass)))) (CC or) (VP (VBD imported) (NP (NP (DT the) (JJ pre-made) (NNS ingots)) (, ,) (SBAR (WHNP (WDT which)) (S (NP (PRP they)) (VP (VBD melted) (CC and) (VBD finished)))))))))) ( . .)))"***

---

# Machine Learning Approaches

- In the Penn Treebank corpus (Marcus et al., 1993) these null elements, or empty categories and their antecedents are marked on the tree by the human annotators.
  - The dominant approach to this problem has been machine learning, ranging from pattern matching algorithms (Johnson, 2002) to training the PCFG parser into the enhanced parse trees (Dienes and Dubey, 2003a, b).
-

---

# Linguistic Approaches

## Using Linguistic Principles to Recover Empty Categories (R. Campbell, 2004)

- Insert empty categories in the parse tree
  - Index them with their antecedents (if any)
  - Assign function tags
  - Outperform machine learning approaches
  - Does not use valency/subcategorization of the verb (A problem as Campbell himself recognizes)
-

# Argument Structure Revisited

- Even if all empty categories will be recovered, we are still left with the problem of determining arguments vs. adjuncts in the parse tree.
- “She knew when Peter left for the library.”
- "(S1 (S (NP (PRP she)) (VP (VBD knew) (SBAR (WHADVP (WRB when)) (S (NP (NNP Peter)) (VP (VBD left) (PP (IN for) (NP (DT the) (NN library)))))))) (. .)))"
- (g “She ate when Peter left for the library.”)
- "(S1 (S (NP (PRP she)) (VP (VBD ate) (SBAR (WHADVP (WRB when)) (S (NP (NNP Peter)) (VP (VBD left) (PP (IN for) (NP (DT the) (NN library)))))))) (. .)))"

---

# MCR (Minimal Clause Reconstruction (C. Millward and F. Gomez, 2010))

- Some similarity to (Winograd, 1983 ) analysis of systemic grammar (Halliday, 1973) for parsing.
  - Uses linguistic principles to transform a parse tree into a set of clauses.
  - Builds a clause for each non-auxiliary verb.
  - Recovers the verb sequence
  - Recovers subjects (in a minimal way).
  - Resolve gaps resulting from relative clauses.
-



---

# MCR (continuation)

- Fills all post-verbal constituents of the VP (all those constituents dominated by the VP), and names them as obj1 , obj2 ... obji
  - It does not decide which constituents of the clause are adjuncts and which ones are arguments.
  - May break up attachment of post-verbal prepositional phrases.
-

---

# Example

" Instead, Alcestis took his place, but Heracles managed to persuade Thanatos, the god of death, to return her to the world of the living." (Wikipedia)

- "(S1 (S (S (ADVP (RB Instead)) (, ,) (NP (NNP Alcestis)) (VP (VBD took) (NP (PRP\$ his) (NN place)))) (, ,) (CC but) (S (NP (NNP Heracles)) (VP (VBD managed) (S (VP (TO to) (VP (VB persuade) (NP (NP (NNP Thanatos)) (, ,) (NP (NP (DT the) (NNP god)) (PP (IN of) (NP (NN death)))) (, ,)) (S (VP (TO to) (VP (VB return) (NP (PRP her)) (PP (TO to) (NP (NP (DT the) (NN world)) (PP (IN of) (NP (DT the) (NN living))))))))))))))))) (. .)))"
-

# MCR OUTPUT

- **(G1189**  
**(ADVERB ((RB INSTEAD)) SUBJ ((PN ALCESTIS)) VERB**  
**((MAIN-VERB TAKE TOOK) (TENSE VBD)) SS (T) COORD (S G1190) OBJ**  
**((PRP\$ HIS) (NOUN PLACE)))**
- **G1190**  
**(SUBJ ((PN HERACLES)) VERB ((MAIN-VERB MANAGE MANAGED) (TENSE VBD)) SS**  
**(T) COORD (S G1189) OBJ ((GENSYM G1191 INFINITIVE)))**
- **G1191**  
**(SUBJ ((PN HERACLES)) VERB ((MAIN-VERB PERSUADE PERSUADE) (TENSE VB))**  
**TYPE (INFINITIVE) SS (NIL) PARENT-VERB (MANAGE G1190) OBJ**  
**((PN THANATOS)) APOS**  
**((DT THE) (PN GOD)) PREP (OF ((NOUN DEATH)))) OBJ2**  
**((GENSYM G1192 INFINITIVE)))**
- **G1192**  
**(SUBJ (((PN HERACLES)) ((PN THANATOS))) VERB**  
**((MAIN-VERB RETURN RETURN) (TENSE VB)) TYPE (INFINITIVE) SS**  
**(NIL) PARENT-VERB (PERSUADE G1191) OBJ ((PRON HER)) PREP**  
**(TO ((DT THE) (NOUN WORLD))) PREP (OF ((DT THE) (NOUN LIVING))))**

# Another Example

- " It is also unclear whether they made their own raw glass or imported the pre-made ingots, which they melted and finished."
- "(S1 (S (NP (PRP It)) (VP (AUX is) (ADVP (RB also)) (ADJP (JJ unclear)) (SBAR (IN whether) (S (NP (PRP they)) (VP (VP (VBD made) (NP (PRP\$ their) (JJ own) (JJ raw) (NN glass)))) (CC or) (VP (VBD imported) (NP (NP (DT the) (JJ pre-made) (NNS ingots)) (, ,) (SBAR (WHNP (WDT which)) (S (NP (PRP they)) (VP (VBD melted) (CC and) (VBD finished)))))))))) ( . .)))"

# MCR Output

- **G876**  
(SUBJ ((PRON It)) VERB ((MAIN-VERB be is) (TENSE AUXVB)) SS (T) ADVERB ((RB also)) PRED ((ADJ unclear)) PRED2 (AND-S (((GENSYM G878 SBAR whether))) (((GENSYM G877 SBAR whether))))
- **G877**  
(SUBJ ((PRON they)) VERB ((MAIN-VERB make made) (TENSE VBD)) TYPE (SBAR whether) SS (T) COORD (VP G878) PARENT-VERB (be G876) OBJ ((PRP\$ their) (ADJ own) (ADJ raw) (NOUN glass)))
- **G878**  
(SUBJ ((PRON they)) VERB ((MAIN-VERB import imported) (TENSE VBD)) TYPE (SBAR whether) SS (T) COORD (VP G877) PARENT-VERB (be G876) OBJ ((DT the) (ADJ pre-made) (NOUN ingots)) RELATIVE (AND-S (((GENSYM G880 REL WHNP-M which COMMA))) (((GENSYM G879 REL WHNP-M which COMMA))))
- **G879**  
(SUBJ ((PRON they)) VERB ((MAIN-VERB melt melted) (TENSE VBD)) TYPE (REL WHNP-M which COMMA) SS (T) COORD (VP G880) PARENT-VERB (import G878) MOVED-OBJ ((DT the) (ADJ pre-made) (NOUN ingots)))
- **G880**  
(SUBJ ((PRON they)) VERB ((MAIN-VERB finish finished) (TENSE VBD)) TYPE (REL WHNP-M which COMMA) SS (T) COORD (VP G879) PARENT-VERB (import G878) MOVED-OBJ ((DT the) (ADJ pre-made) (NOUN ingots))))

---

# Supervised Machine Learning Approaches to Argument/Role Labeling

- 1) Supervised approaches rely on human annotation of a given corpus.
  - 2) They are very much depending on the syntax of the parse tree.
  - 3) Performance drops off 15% or more when the algorithms are presented with new domains.
  - 4) They do not address the issue of verb meaning/verb predicate.
  - 5) These systems fail to generalize to the language of a domain they have not seen before.
-

# Sentence Simplification for Semantic Role Labeling (D. Vickrey and D. Koller, 2008)

- The classifier relies on the path of the tree to determine the roles.
- The authors discussed the sentence "He expected to receive a prize for winning."
- "(S1 (S (NP (PRP He)) (VP (VBD expected) (S (VP (TO to) (VP (VB receive) (NP (NP (DT a) (NN prize)) (PP (IN for) (S (VP (VBG winning)))))))))))))"
- If the sentence had said "did not expect" the path would be different, and the classifier will find the syntax of the two sentences totally unrelated.
- "(S1 (S (NP (PRP he)) (VP (AUX did) (RB not) (VP (VB expect) (S (VP (TO to) (VP (VB receive) (NP (NP (DT a) (NN prize)) (PP (IN for) (S (VP (VBG winning))))))))))))) (. .)))"
- The authors propose some rules to simplify the parse and methods to train the classifier in these simplified sentences.

---

# Determining Verb Meaning and Semantic Roles (Gomez, 2007, 2004, 2001)

- Arguments are decided using verb valency/verb subcategorization.
  - Verb Semantics
  - Noun Ontology (WordNet)
  - Strong relation between arguments and verb meaning
-



---

# Semantic Interpretation Algorithm

- **The input to the semantic interpretation algorithm is the output of the MCR.**
  - **The MCR breaks up attachment of post-verbal prepositional phrases.**
  - **The determination of verb meaning and semantic roles is interdependent.**
  - **A predicate explains a grammatical relation if it has a semantic role realized by that grammatical relation.**
  - **The predicate that has the most semantic roles realized is selected as the meaning of the verb.**
-

# Some examples for “drive”

- On the first EVA the crew drove the Rover south to sample the area at Elbow Crater.
- An ELP must be driven by relatively high voltage AC power, which is provided by an inverter circuit.
- At this point, the oxygen-poor blood from upper and lower body flows through the lungs without being pumped (driven only by the pressure that builds up in the veins).
- After Mike leaves, Joshua is bitten by an escaped tarantula and is driven into a frenzy. (hotel)
- In the decades after the Second World War, a massive restructuring plan drove Japan to become the world's second-largest economy

---

# Two Predicates for “leave”

[leave-a-place

(is-a(change-location)) (wn-map(leave1))

(agent(human animal1)(subj))

(from-loc( location) (obj (prep for)))

(instrument(vehicle)( subj (prep in on)))]

---

---

# Predicate for “leave”

[leave-give

(is-a(give)) (wn-map(leave12))

(agent(human-agent) (subj-if-obj))

(theme(possession thing) (obj ))

(to-poss (human-agent) (obj-if-obj2)

(human-agent animal1) ((prep for to)))

(thematic-rule (require(to-poss))))]

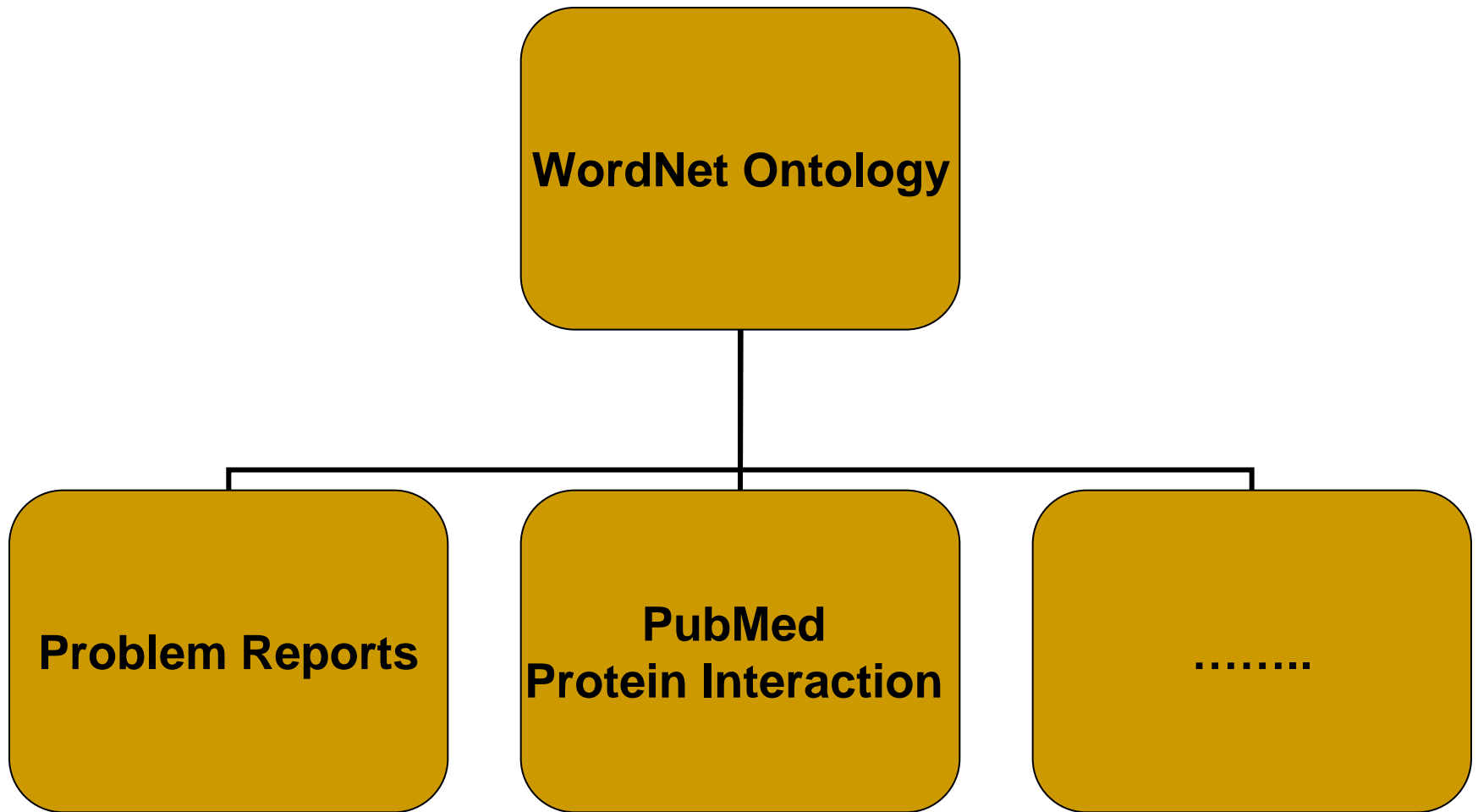
---

---

# Some Sentences

- The farmer left the land to her daughter.
  - The farmer left the land on his tractor.
  - The teacher left for school on her car
  - The colonial wars left Britain in debt.
  - They left much of the country in ruins.
-

# Domain Specific Ontologies



---

# Conclusions

- We have explained methods for the extraction of knowledge from texts
  - The MCR algorithm may take care about 80% or more of the effort needed to build an extraction application.
  - The MCR algorithm is domain independent.
  - The remainder 20% percent effort will be the construction of the domain-dependent ontology, including the verb predicates.
-