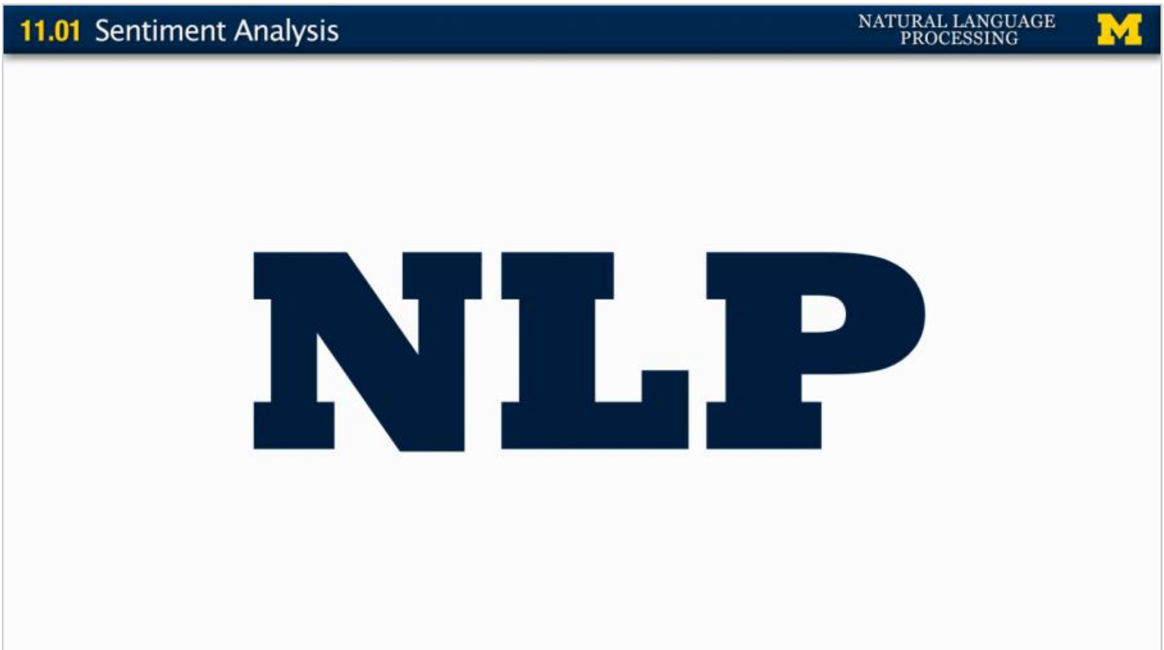
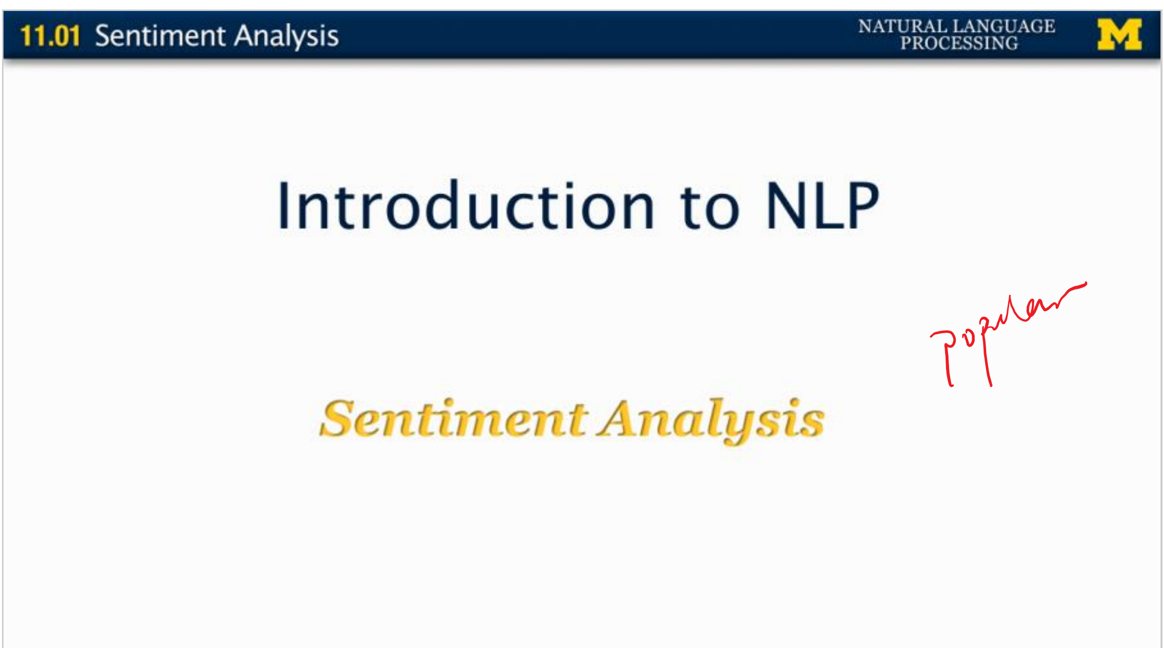
Week 11

Friday, August 12, 2016 14:12

11.1





Reviews of 1Q84 by Haruki Murakami

- "*1Q84* is a tremendous feat and a triumph . . . A must-read for anyone who wants to come to terms with contemporary Japanese culture."
  —Lindsay Howell, *Baltimore Examiner*
- "Perhaps one of the most important works of science fiction of the year . . . *1Q84* does not disappoint . . . [It] envelops the reader in a shifting world of strange cults and peculiar characters that is surreal and entrancing."
  —Matt Staggs, Suvudu.com
- Ambitious, sprawling and thoroughly stunning . . . Orwellian dystopia, sci-fi, the modern world (terrorism, drugs, apathy, pop novels)—all blend in this dreamlike, strange and wholly unforgettable epic."
  —*Kirkus Reviews* (starred review)

Reviews of 1Q84 by Haruki Murakami

- "*1Q84* is a tremendous feat and a triumph . . . A must-read for anyone who wants to come to terms with contemporary Japanese culture."
  —Lindsay Howell, *Baltimore Examiner*
- "Perhaps one of the most important works of science fiction of the year . . . *1Q84* does not disappoint . . . [It] envelops the reader in a shifting world of strange cults and peculiar characters that is surreal and entrancing."
  —Matt Staggs, Suvudu.com
- Ambitious, sprawling and thoroughly stunning . . . Orwellian dystopia, sci-fi, the modern world (terrorism, drugs, apathy, pop novels)—all blend in this dreamlike, strange and wholly unforgettable epic."
  —*Kirkus Reviews* (starred review)

Sentiment About Companies

Other Examples

- Movie reviews
- Product reviews
- Debates
  - www.createdebate.com

debates

Introduction

- Many posts, blogs
- Expressing personal opinions
- Research questions
  - Subjectivity analysis → decide whether something is subjective
  - Polarity analysis (positive/negative, number of stars)
  - Viewpoint analysis (Chelsea vs. Manchester United, republican vs. democrat)
- Sentiment target
  - entity
  - aspect

11.01 Sentiment Analysis
NATURAL LANGUAGE PROCESSING

Introduction

- Level of granularity
  - Document
  - Sentence
  - Attribute
- Opinion words
  - Base
  - Comparative (better, slower)

Introduction

- Just counting negative words is not enough
- Negation analysis

Some are negated.

Reviews of 1Q84 by Haruki Murakami

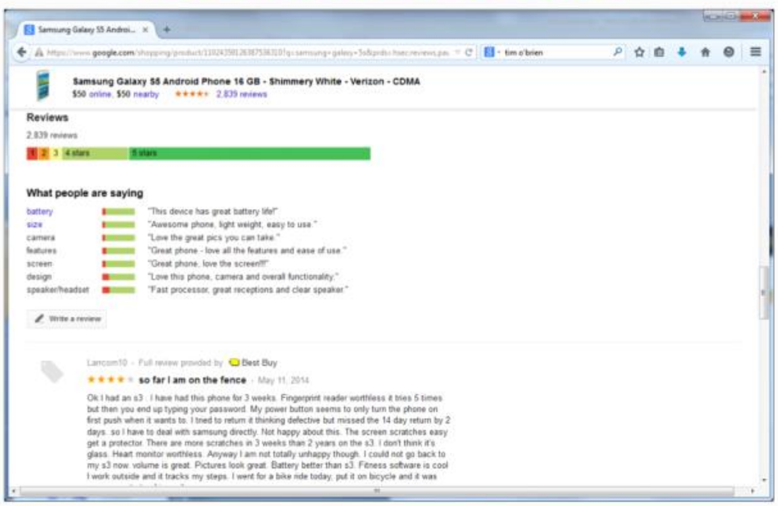
- "*1Q84* is a tremendous feat and a triumph . . . A must-read for anyone who wants to come to terms with contemporary Japanese culture."
  —Lindsay Howell, *Baltimore Examiner*
- "Perhaps one of the most important works of science fiction of the year . . . *1Q84* does not disappoint . . . [It] envelops the reader in a shifting world ~~of~~ strange cults and peculiar characters that is surreal and entrancing."
  —Matt Staggs, Suvudu.com
- Ambitious, sprawling and thoroughly stunning . . . Orwellian dystopia, sci-fi, the modern world (terrorism, drugs, apathy, pop novels)—all blend in this dreamlike, strange and wholly unforgettable epic."
  —*Kirkus Reviews* (starred review)

Reviews of 1Q84 by Haruki Murakami

- "*1Q84* is a tremendous feat and a triumph . . . A must-read for anyone who wants to come to terms with contemporary Japanese culture."
  —Lindsay Howell, *Baltimore Examiner*
- "Perhaps one of the most important works of science fiction of the year . . . *1Q84* does not disappoint . . . [It] envelops the reader in a shifting world of strange cults and peculiar characters that is surreal and entrancing."
  —Matt Staggs, Suvudu.com
- Ambitious, sprawling and thoroughly stunning . . . Orwellian dystopia, sci-fi, the modern world (terrorism, drugs, apathy, pop novels)—all blend in this dreamlike, strange and wholly unforgettable epic."
  —*Kirkus Reviews* (starred review)

11.01 Sentiment Analysis NATURAL LANGUAGE PROCESSING

Product Reviews



Twitter Sentiment

11.01 Sentiment Analysis
NATURAL LANGUAGE PROCESSING

Problems

- Subtlety
- Concession
- Manipulation → 水军.
- Sarcasm and irony

11.01 Sentiment Analysis

NATURAL LANGUAGE PROCESSING

SA as a Classification Problem

- Set of features
  - Words
    - Presence is more important than frequency
  - Punctuation
  - Phrases
  - Syntax
- A lot of training data is available -> use as training dataset
  - E.g., movie review sentences and stars
  - http://www.cs.cornell.edu/people/pabo/movie-review-data Bo Pang & Lillian Lee
- Techniques
  - MaxEnt
  - SVM
  - Naive Bayes

11.01 Sentiment Analysis

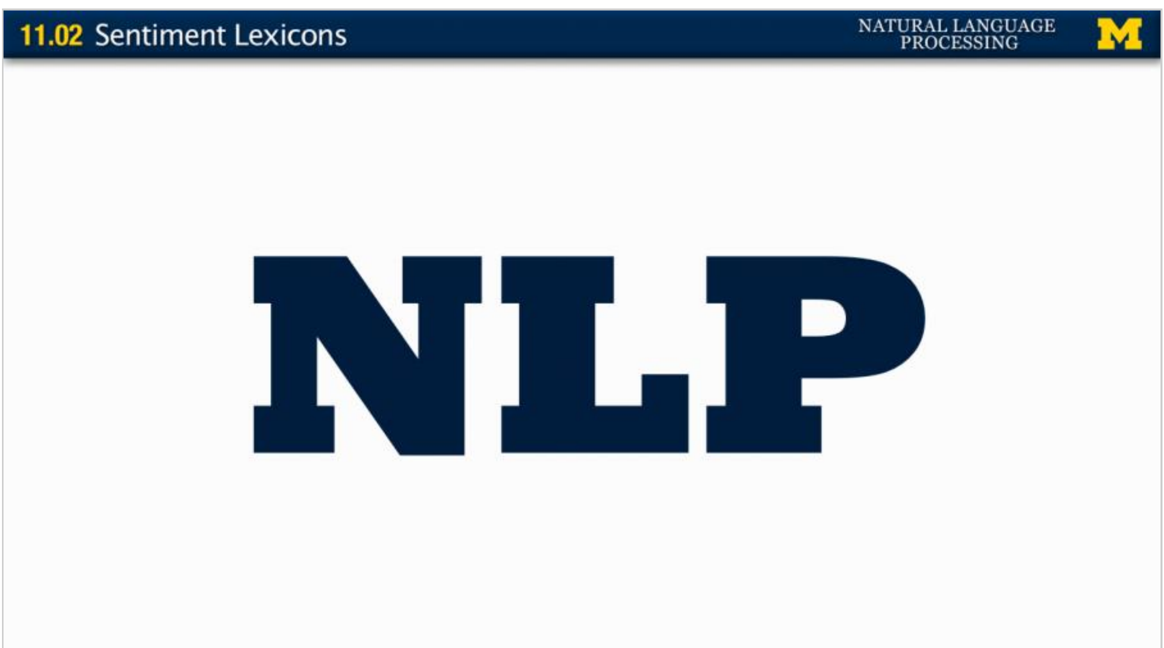
NATURAL LANGUAGE PROCESSING

Resources

- CMU Twitter parser
  - http://www.ark.cs.cmu.edu/TweetNLP/


11.01 Sentiment Analysis
NATURAL LANGUAGE
PROCESSING
NLP

11.2


11.02 Sentiment Lexicons
NATURAL LANGUAGE
PROCESSING
NLP

Introduction to NLP

Sentiment Lexicons

Sentiment Lexicons

manually built.

- SentiWordNet
  - http://sentiwordnet.isti.cnr.it/
- General Inquirer
  - 2,000 positive words and 2,000 negative words
  - http://www.wjh.harvard.edu/~inquirer/
- LIWC
  - http://www.liwc.net/
- MPQA subjectivity lexicon
  - http://www.cs.pitt.edu/mpqa/subj_lexicon.html

General Inquirer

- Annotations
  - Strong Power Weak Submit Active Passive Pleasur Pain Feel Arousal EMOT Virtue Vice Ovrst Undrst Academ Doctrin Econ@ Exch ECON Exprsv Legal Milit Polit@ POLIT Relig Role COLL Work Ritual SocRel Race Kin@ MALE Female Nonadlt HU ANI PLACE Social Region Route Aquatic Land Sky Object Tool Food Vehicle BldgPt ComnObj NatObj BodyPt ComForm COM Say Need Goal Try Means Persist Complet Fail NatrPro Begin Vary Increas Decreas Finish Stay Rise Exert Fetch Travel Fall Think Know Causal Ought Perceiv Compare Eval@ EVAL Solve Abs@ ABS Quality Quan NUMB ORD CARD FREQ DIST Time@ TIME Space POS DIM Rel COLOR Self Our You Name Yes No Negate Intrj IAV DAV SV IPadj IndAdj PowGain PowLoss PowEnds PowAren PowCon PowCoop PowAuPt PowPt PowDoct PowAuth PowOth PowTot RcEthic RcRelig RcGain RcLoss RcEnds RcTot RspGain RspLoss RspOth RspTot AffGain AffLoss AffPt AffOth AffTot WltPt WltTran WltOth WltTot WlbGain WlbLoss WlbPhys WlbPsyc WlbPt WlbTot EnlGain EnlLoss EnlEnds EnlPt EnlOth EnlTot SklAsth SklPt SklOth SklTot TrnGain TrnLoss TranLw MeansLw EndsLw ArenaLw PtLw Nation Anomie NegAff PosAff SureLw If NotLw TimeSpc
- http://www.webuse.umd.edu:9090/tags/
  - Positive: able, accolade, accuracy, adept, adequate…
  - Negative: addiction, adversity, adultery, affliction, aggressive…

Small section about positive/negative word

Dictionary-based Methods

- Start from known seeds
  - e.g., happy, angry
- Expand using WordNet
  - synonyms
  - hypernyms
- Random-walk based methods
  - words with known polarity as absorbing boundary

Automatic Extraction of Sentiment Words

- Semi-supervised methods

Vasileios Hatzivassiloglou and Kathleen R. McKeown. 1997. Predicting the Semantic Orientation of Adjectives. ACL, 174–181

Molistic

- NACLO problem (2007)

11.02 Sentiment Lexicons

NATURAL LANGUAGE PROCESSING

Imagine that you have heard these sentences:

Jane is molistic and slatty.
Jennifer is cluvious and brastic.
Molly and Kyle are slatty but danty.
The teacher is danty and cloovy.
Mary is blitty but cloovy.
Jeremiah is not only sloshful but also weasy.
Even though frumsy, Jim is sloshful.
Strungy and struffy, Diane was a pleasure to watch.
Even though weasy, John is strungy.
Carla is blitty but struffy.
The salespeople were cluvious and not slatty.

A1. Then which of the following would you be likely to hear?

a. Meredith is blitty and brastic.
b. The singer was not only molistic but also cluvious.
c. May found a dog that was danty but sloshful.

A2. What quality or qualities would you be looking for in a person?

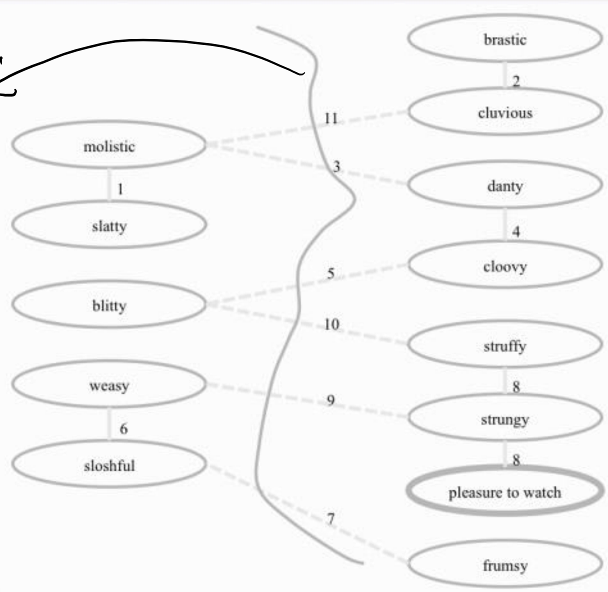
a. blitty
b. weasy
c. sloshful

clovy.
cluvious. brastic. danty.
slatty molistic blitty.

11.02 Sentiment Lexicons

NATURAL LANGUAGE PROCESSING

decision boundary that goes through all dash line.



+

PMI (Turney)

PMI again

- PMI=pointwise mutual information
- Check how often a given unlabeled word appears with a known positive word ("excellent")
- Same for a known negative word ("poor")

appear with a known positive word

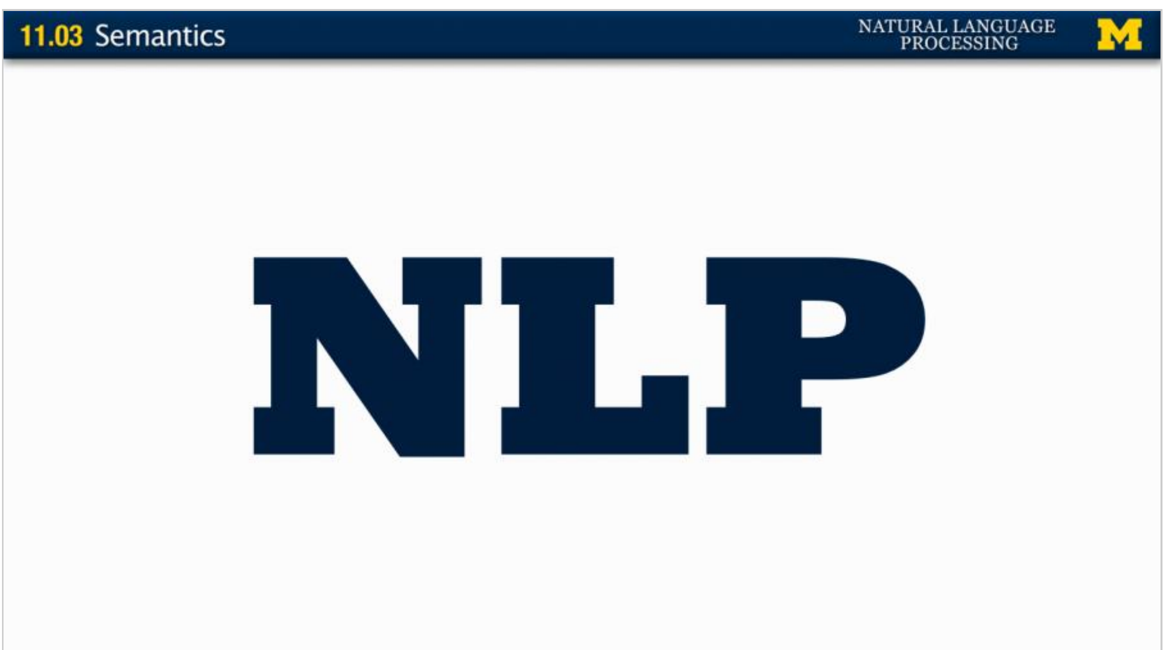
$$\text{PMI}(word_1, word_2) = \log_2 \frac{hits(word_1 \text{ NEAR } word_2)}{hits(word_1)hits(word_2)}$$

Dataset

- http://www.cs.jhu.edu/~mdredze/datasets/sentiment/

ties with logic, philosophy. knowledge representation

11.3



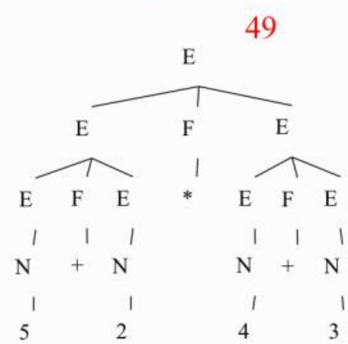
Introduction to NLP

Semantics

11.03 Semantics NATURAL LANGUAGE PROCESSING

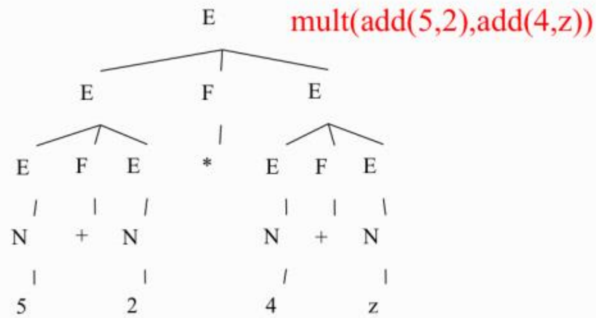
Semantics

- What is the meaning of: (5+2)*(4+3)?
- Parse tree



Semantics

- What if we had (5+2)*(4+z)?



What About (English) Sentences?

- Every human is mortal.
- ??

What's the meaning of this sentence

Representing Meaning

- Capturing the meaning of linguistic utterances using formal notation
- Linguistic meaning
  - "It is 8 pm"
- Pragmatic meaning
  - "It is time to leave"
- Semantic analysis:
  - Assign each word a meaning
  - Combine the meanings of words into sentences
- *I bought a book*:

  $\exists\ x,y$: Buying(x) ^ Buyer(*speaker*,x) ^ BoughtItem(y,x) ^ Book(y)

  Buying (Buyer=speaker, BoughtItem=book)

Entailment and Presupposition

- Entailment
  - One fact follows from another
  - "All cats have whiskers" and "Martin is a cat" entail the statement "Martin has whiskers"
  - "Martin has whiskers and a tail" entails "Martin has whiskers"
- Presupposition
  - "The Queen of Utopia is dead" presupposes that Utopia has a queen

Entailment and Presupposition

- NACLO problem from 2010
- Author: Aleka Blackwell
- http://www.nacloweb.org/resources/problems/2010/M.pdf
- http://www.nacloweb.org/resources/problems/2010/MS.pdf

Think about the meaning of the following sentence:

(1) The 2010 Winter Olympics were in Canada.

Assuming that we only know sentence 1 to be true, is sentence 2 necessarily true?

(2) The 2010 Winter Olympics were in Vancouver.

The answer is no. Assuming we only know sentence 1 to be true, the 2010 Winter Olympics could have taken place in any Canadian city, but not necessarily in Vancouver.

Now examine the relationship between sentences 3 and 4. Assuming sentence 3 is true, is sentence 4 now necessarily true?

(3) The 2010 Winter Olympics were in Vancouver.
(4) The 2010 Winter Olympics were in Canada.

Now the answer is yes. Since Vancouver is a Canadian city, any event which occurs in Vancouver necessarily occurs in Canada.

The logical relationship which holds between sentences 3 and 4 is called an *entailment*. In formal terms, sentence A entails sentence B if whenever A is true, B is necessarily true. The entailment relationship is typically represented graphically this way: A ||- B.

Here are some more examples of the entailment relationship between sentences:

(5) Shaun White is a Winter Olympian ||- Shaun White is an Olympian
(6) Shaun White is an Olympian ||- Shaun White is an athlete
(7) Shaun White won a gold medal ||- Someone won a gold medal

Notice that the entailment relationship must hold in the specified direction but will not necessarily hold in both directions. So, sentence 3 entails sentence 4 even though sentence 4 does not entail sentence 3.

Entailment and Presupposition

Now examine the relationship between sentences 8 and 9.

(8) I did not see Shaun White win the gold medal in the 2010 Winter Olympics.
(9) Shaun White won the gold medal in the 2010 Winter Olympics.

Sentences 8 and 9 illustrate a relationship called *presupposition*. In this pair of sentences, the information presented in sentence 9 is what the speaker assumes (or presupposes) to be the case when uttering sentence 8. That is, to say *"I did not see Shaun White win the gold medal"* assumes the belief that Shaun White won a gold medal. In formal terms, sentence A presupposes sentence B if A not only implies B but also implies that the truth of B is somehow taken for granted. A presupposition of a sentence is thus part of the background against which its truth or falsity is judged. The presupposition relationship is typically represented graphically this way: A >> B

Here are some more examples of presuppositions (where the first sentence in each pair presupposes the second):

(10) I regret not seeing Shaun White's gold medal run >> Shaun White had a gold medal run
(11) Shaun White continues to rule the halfpipe >> Shaun White had been ruling the halfpipe
(12) Snowboarding is now an Olympic sport >> Snowboarding was once not an Olympic sport

Entailment and Presupposition

For any given pair of sentences, the entailment and presupposition relationships may or may not hold, together or separately.

For each of the following possible combinations, your task is to provide one example of a pair of sentences with an explanation of your reasoning for proposing your pair of sentences as a valid and convincing example in each case.

a. A pair of sentences in which sentence A **neither entails nor presupposes** sentence B.

b. A pair of sentences in which sentence A **entails and presupposes** sentence B.

c. A pair of sentences in which sentence A **presupposes but does not entail** sentence B.

d. A pair of sentences in which sentence A **entails but does not presuppose** sentence B.

11.03 Semantics NATURAL LANGUAGE PROCESSING

Answers

For any given pair of sentences, the entailment and presupposition relationships may or may not hold, together or separately.

a. A pair of sentences in which sentence A **neither entails nor presupposes** sentence B.
   - A. Shaun White is a Winter Olympian.
   - B. The 2010 Winter Olympics were in Vancouver.

Explanation: Sentences A and B are unrelated.

Entailment: Given that sentence A is true, there is no way to know whether sentence B is true or false. If Shaun White is a Winter Olympian, the 2010 Winter Olympics may or may not have taken place in Vancouver. Thus, there is no entailment relationship between these two sentences.

Presupposition: When uttering sentence A, a speaker would not take sentence B for granted (or assume that sentence B is background information against which the truth or falsity of sentence A would be judged). A speaker would not utter "Shaun White is a Winter Olympian" and assume the belief/take for granted that the 2010 Winter Olympics were in Vancouver.

b. A pair of sentences in which sentence A **entails and presupposes** sentence B.
   - A. Shaun White continues to rule the halfpipe
   - B. Shaun White had been ruling the halfpipe.

Entailment: If sentence A is true, sentence B is necessarily true. The entailment relationship between these sentences relies on the meaning of the verb *continue* – to *continue to rule* the halfpipe, Shaun White had to be ruling the halfpipe already. Thus, sentence A entails sentence B.

Presupposition: When uttering sentence A, a speaker would take sentence B for granted (or assume that sentence B is background information against which the truth or falsity of sentence A would be judged). A speaker who utters "Shaun White continues to rule the halfpipe" assumes the belief/takes for granted that Shaun White had been ruling the halfpipe. Thus, sentence A presupposes sentence B.

11.03 Semantics NATURAL LANGUAGE PROCESSING

Entailment and Presupposition

c. A pair of sentences in which sentence A **presupposes but does not entail** sentence B.
   - A. I did not see Shaun White win the gold medal in the 2010 Winter Olympics.
   - B. Shaun White won the gold medal in the 2010 Winter Olympics.

Entailment: If sentence A is true, sentence B *may or may not* be true. The absence of an entailment relationship between these sentences relies on the words "did not see" – if it is true that I *did not see* Shaun White win the gold medal, then Shaun White may or may not have won the gold medal. Thus, sentence A does not entail sentence B.

Presupposition: When uttering sentence A, a speaker would take sentence B for granted (or assume that sentence B is background information against which the truth or falsity of sentence A would be judged). Specifically, a speaker who utters "I did not see Shaun White win the gold medal in the 2010 Winter Olympics" assumes the belief that Shaun White did actually win the gold medal in the 2010 Winter Olympics. Thus, sentence A presupposes sentence B.

Entailment and Presupposition

d. A pair of sentences in which sentence A **entails but does not presuppose** sentence B.

A. Shaun White did not win the gold medal in the 2010 Winter Olympics.

B. Shaun White did not both win the gold medal in the 2010 Winter Olympics and injure his ankle.

Entailment: If Shaun White did not win the gold medal in the 2010 Winter Olympics, then he necessarily did not *both* win that gold medal *and* injure his ankle, since he definitely did not win the gold medal. If one fact is not the case (the fact presented in sentence A), then both facts cannot be the case, either (the fact presented in sentence A + the new fact added to it in sentence B). Thus if sentence A is true, sentence B is *necessarily* true. Thus, sentence A entails sentence B.

Presupposition: When uttering sentence A, a speaker would not take sentence B for granted (or assume that sentence B is a background against which the truth or falsity of sentence A would be judged). Specifically, by uttering "Shaun White did not win the gold medal in the 2010 Winter Olympics" a speaker could not assume the belief that Shaun White did not both win the gold and injure his ankle, or that Shaun White either won a gold medal or injured his ankle. Whether Shaun White injured his ankle would not be information taken for granted when uttering "Shaun White did not win the gold medal in the 2010 Winter Olympics." Thus, sentence A does not presuppose sentence B.

11.03 Semantics

NATURAL LANGUAGE PROCESSING

NLP

PDF

11.4

NLP

11.04 Representing and Understanding Meaning
NATURAL LANGUAGE PROCESSING

Introduction to NLP

Representing and Understanding Meaning

Understanding Meaning

- If an agent hears a sentence and can act accordingly, the agent is said to understand it
- Example
  - Leave the book on the table
- Understanding may involve inference
  - Maybe the book is wrapped in paper?
- And pragmatics
  - Which book? Which table?
- So, understanding may involve a procedure

Properties

- Verifiability
  - Can a statement be verified against a knowledge base (KB)
  - Example: does my cat Martin have whiskers?
- Unambiguousness
  - Give me the book
  - Which book?
- Canonical form
- Expressiveness
  - Can the formalism express temporal relations, beliefs, ...?
  - Is it domain-independent?
- Inference

Representing Meaning

logic rep.

- One traditional approach is to use logic representations, e.g., FOL (first order logic)
- One can then use theorem proving (inference) to determine whether one statement entails another

Syntax of Propositional Logic

- The simplest type of logic
- The proposition symbols P_1, P_2, … are sentences
  - If S is a sentence, $\neg S$ is a sentence (negation)
  - If S_1 and S_2 are sentences, $S_1 \wedge S_2$ is a sentence (conjunction)
  - If S_1 and S_2 are sentences, $S_1 \vee S_2$ is a sentence (disjunction)
  - If S_1 and S_2 are sentences, $S_1 \Rightarrow S_2$ is a sentence (implication)
  - If S_1 and S_2 are sentences, $S_1 \Leftrightarrow S_2$ is a sentence (biconditional)

Propositional Logic in Backus Naur Form

- Sentence → AtomicSentence | ComplexSentence
- AtomicSentence → True | False | S | T | U ...
- ComplexSentence → (Sentence)

  | ¬Sentence

  | Sentence ∧ Sentence   conjunction

  | Sentence ∨ Sentence   disjunction

  | Sentence ⇒ Sentence

  | Sentence ⇔ Sentence

11.04 Representing and Understanding Meaning NATURAL LANGUAGE PROCESSING

Operator Precedence

¬ (highest)

∧

∨

⇒

⇔ (lowest)

11.04 Representing and Understanding Meaning

NATURAL LANGUAGE PROCESSING

Translating Propositions to English

logics ...

- A = Today is a holiday.
- B = We are going to the park.

- $A \Rightarrow B$ Q.
- $A \land \neg B$
- $\neg A \Rightarrow B$
- $\neg B \Rightarrow A$
- $B \Rightarrow A$

11.04 Representing and Understanding Meaning

NATURAL LANGUAGE PROCESSING

Translating Propositions to English

- A = Today is a holiday.
- B = We are going to the park.

- $A \Rightarrow B$
  If today is a holiday, we are going to the park.
- $A \land \neg B$
  Today is a holiday and we are not going to the park.
- $\neg A \Rightarrow \neg B$
  If today is not a holiday, then we are not going to the park.
- $\neg B \Rightarrow \neg A$
  If we are not going to the park, then today is not a holiday.
- $B \Rightarrow A$
  If we are going to the park, then today is a holiday.

Semantics of Propositional Logic

logic

- $\neg S$ is true iff S is false
- $S_1 \wedge S_2$ is true iff S_1 is true and S_2 is true
- $S_1 \vee S_2$ is true iff S_1 is true or S_2 is true
- $S_1 \Rightarrow S_2$ is true iff S_1 is false or S_2 is true
- i.e., is false iff S_1 is true and S_2 is false
- $S_1 \Leftrightarrow S_2$ is true iff $S_1 \Rightarrow S_2$ is true and $S_2 \Rightarrow S_1$ is true
- Recursively, one can compute the truth value of longer formulas

11.04 Representing and Understanding Meaning
NATURAL LANGUAGE PROCESSING

Connectives

P false Q true
第二一己
P ⇒ Q 结果false

P	Q	¬P	P∧Q	P∨Q	P⇒Q	P⇔Q
F	F	T	F	F	T	T
F	T	T	F	T	T	F
T	F	F	F	T	F	F
T	T	F	T	T	T	T

Logical Equivalence

$$
\begin{array}{rcll}
(\alpha \wedge \beta) & \equiv & (\beta \wedge \alpha) & \text{commutativity of } \wedge \\
(\alpha \vee \beta) & \equiv & (\beta \vee \alpha) & \text{commutativity of } \vee \\
((\alpha \wedge \beta) \wedge \gamma) & \equiv & (\alpha \wedge (\beta \wedge \gamma)) & \text{associativity of } \wedge \\
((\alpha \vee \beta) \vee \gamma) & \equiv & (\alpha \vee (\beta \vee \gamma)) & \text{associativity of } \vee \\
\neg(\neg\alpha) & \equiv & \alpha & \text{double-negation elimination} \\
(\alpha \Rightarrow \beta) & \equiv & (\neg\beta \Rightarrow \neg\alpha) & \text{contraposition} \\
(\alpha \Rightarrow \beta) & \equiv & (\neg\alpha \vee \beta) & \text{implication elimination} \\
(\alpha \Leftrightarrow \beta) & \equiv & ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha)) & \text{biconditional elimination} \\
\neg(\alpha \wedge \beta) & \equiv & (\neg\alpha \vee \neg\beta) & \text{de Morgan} \\
\neg(\alpha \vee \beta) & \equiv & (\neg\alpha \wedge \neg\beta) & \text{de Morgan} \\
(\alpha \wedge (\beta \vee \gamma)) & \equiv & ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma)) & \text{distributivity of } \wedge \text{ over } \vee \\
(\alpha \vee (\beta \wedge \gamma)) & \equiv & ((\alpha \vee \beta) \wedge (\alpha \vee \gamma)) & \text{distributivity of } \vee \text{ over } \wedge
\end{array}
$$

NLP

11.5

11.05 First Order Logic

NATURAL LANGUAGE PROCESSING

NLP

11.05 First Order Logic

NATURAL LANGUAGE PROCESSING

Introduction to NLP

First Order Logic

Properties of Propositional Logic

- Pros
  - Compositional
  - Declarative
- Cons
  - Limited expressive power
  - Represents facts

First Order Logic

- Used to represent
  - Objects – Martin the cat
  - Relations – Martin and Moses are brothers
  - Functions – Martin's age

First Order Logic

- *Formula* → AtomicFormula | Formula Connective Formula
  | Quantifier Variable Formula | ¬ Formula |
  (Formula)
- AtomicFormula → Predicate (Term…)
- Term → Function (Term…) | Constant | Variable
- Connective → ∧ | ∨ | ⇒
- Quantifier → ∀ | ∃
- Constant → M | Martin
- Variable → *x* | *y* | …
- Predicate → Likes | Eats | …
- Function → AgeOf | ColorOf | …

Common Mistake (1)

- ⇒ is the main connective with ∀
- Common mistake: using ∧ as the main connective with ∀:

  ∀x Cat(x) ∧ EatsFish(x)

  means "Everyone is a cat and everyone eats fish"

$\forall x \; Cat(x) \Rightarrow EatsFish(x)$

11.05 First Order Logic NATURAL LANGUAGE PROCESSING

Common Mistake (2)

- $\wedge$ is the main connective with $\exists$
- Common mistake: using $\Rightarrow$ as the main connective with $\exists$:

$$\exists x \; \mathrm{Cat}(x) \Rightarrow \mathrm{EatsFish}(x)$$

is true if there is anyone who is not a cat!

11.05 First Order Logic NATURAL LANGUAGE PROCESSING

First Order Logic

- NACLO problem from 2014
- Author: Ben King
- http://www.nacloweb.org/resources/problems/2014/N2014-H.pdf
- http://www.nacloweb.org/resources/problems/2014/N2014-HS.pdf

First Order Logic

(H) Bertrand and Russell (1/3) [10 points]

Teachers can be hard to understand sometimes. Case in point, the math teacher, Mr. Whitehead. Just this morning, he told the class, "It's not the case that if at least one student studied for the test, then every student failed the test." What does that even mean?

Well, the two new kids in the class, Bertrand and Russell, have come up with a plan to make sense of Mr. Whitehead's statements. They call it first-order logic (FOL), a way to map these confusing statements into an unambiguous representation. Bertrand says the whole system is built the idea of propositions, a statement that is either true or false. Propositions can be statements about people or things like *studied_for(John, test)* or *is_hard(test)*. Propositions can also be combined to make more complex statements with the following symbols:

Symbol	Example statement	Interpretation	Explanation
¬	¬ *studied_for(John, test)*	John did not study for the test.	The statement is true if and only if John did not study for the test.
∧	*is_hard(test)* ∧ *is_long(test)*	The test is long and hard.	This statement is true whenever the test is long and the test is hard.
∨	*is_hard(test)* ∨ *is_long(test)*	The test is long or hard.	This statement is true if the test is long, or if the test is hard, or both.
⟹	*studied_for(John, test)* ⟹ *aced(John, test)*	If John studied for the test, then he aced it.	This is true if the statement on the right side of the arrow is always true whenever the statement on the left side of the arrow is true. If the statement on the left is false, then the whole statement is true by default (if John didn't study, we don't know how he did on the test).

11.05 First Order Logic

NATURAL LANGUAGE PROCESSING

"But," says Russell, "the most important part of first-order logic is the quantifiers." Quantifiers allow you to make general statements like Mr. Whitehead loves to do.

Symbol	Example statement	Interpretation	Explanation
∀	[$\forall_x$: *student*(x) $\Longrightarrow$ *studied_for*(x, *test*)]	Every student studied for the test.	The ∀ symbol makes a statement about every possible object (whether a student or not). It temporarily gives it the name x to make such a statement. We use the $\Longrightarrow$ symbol because we don't want to make any claims about whether non-students studied.
∃	[$\exists_x$: *student*(x) ∧ *aced*(x, *test*)]	There exists at least one student who aced the test.	The ∃ symbol makes the claim that there is at least one (possibly more) object in the universe, temporarily called x, that satisfies the statement listed.

11.05 First Order Logic

NATURAL LANGUAGE PROCESSING

Bertrand and Russell also note that there are also a couple other things we can say about individuals (but not propositions or quantifiers). For example, if the names Jonathan and Jon both refer to the same person, we can say *Jon = Jonathan*. If we want to emphasize that John and Jon are different people, we can say *John ≠ Jon*.

H1. Translate Mr. Whitehead's statements into first-order logic by finding the proposition below that is equivalent to each statement and writing the letter of the proposition in the blank. Each statement has exactly one correct answer; not every proposition will be used.

D	Everyone either passed or failed the test.
J	Every student did not pass the test.
B	Exactly one student passed the test.
A	A student did not pass the test.
E	It is not the case that if at least one student studied for the test, then every student failed the test.

A.	[$\exists_x$: *student*(x) ∧ ¬ *passed*(x, *test*)]
B.	[$\exists_x$: *student*(x) ∧ *passed*(x, *test*) ∧ [$\forall_y$: *passed*(y, *test*) $\Longrightarrow$ $x = y$]]
C.	[$\exists_x$: *student*(x) ∧ *passed*(x, *test*) ∧ [$\exists_y$: *passed*(y, *test*) ∧ $x = y$]]
D.	[$\forall_x$: *passed*(x, *test*) ∨ *failed*(x, *test*)]
E.	¬ ([$\exists_x$: *student*(x) ∧ *studied_for*(x, *test*)] $\Longrightarrow$ [$\forall_x$: *student*(x) $\Longrightarrow$ *failed*(x, *test*)])
F.	[$\exists_x$: *passed*(x, *test*) ∧ *failed*(x, *test*)]
G.	[$\forall_x$: ¬ *student*(x) $\Longrightarrow$ *passed*(x, *test*)]
H.	[$\exists_x$: *student*(x) ∧ *studied_for*(x, *test*)] $\Longrightarrow$ ¬ [$\forall_x$: *student*(x) $\Longrightarrow$ *failed*(x, *test*)]
I.	¬ [$\exists_x$: *student*(x) ∧ ¬ *passed*(x, *test*)]
J.	[$\forall_x$: *student*(x) $\Longrightarrow$ ¬ *passed*(x, *test*)]

H2. Translate first-order logic propositions into their equivalent English sentences by finding the statement below that is equivalent to each proposition and writing the letter of the statement in the blank. Each proposition has exactly one correct answer; not every statement will be used.

	$[\forall_x : student(x) \implies studied\_for(x, test)] \lor [\forall_y : student(y) \implies passed(y, test)]$
	$[\forall_x : student(x) \implies [studied\_for(x, test) \lor passed(x, test)]]$
	$[\forall_x : (test(x) \land long(x)) \implies hard(x)]$
	$[\exists_x : test(x) \land (long(x) \lor hard(x))]$
	$[\forall_x : test(x) \land \neg (long(x) \land hard(x)) \implies \neg [\forall_y : student(y) \implies failed(y, x)]]$

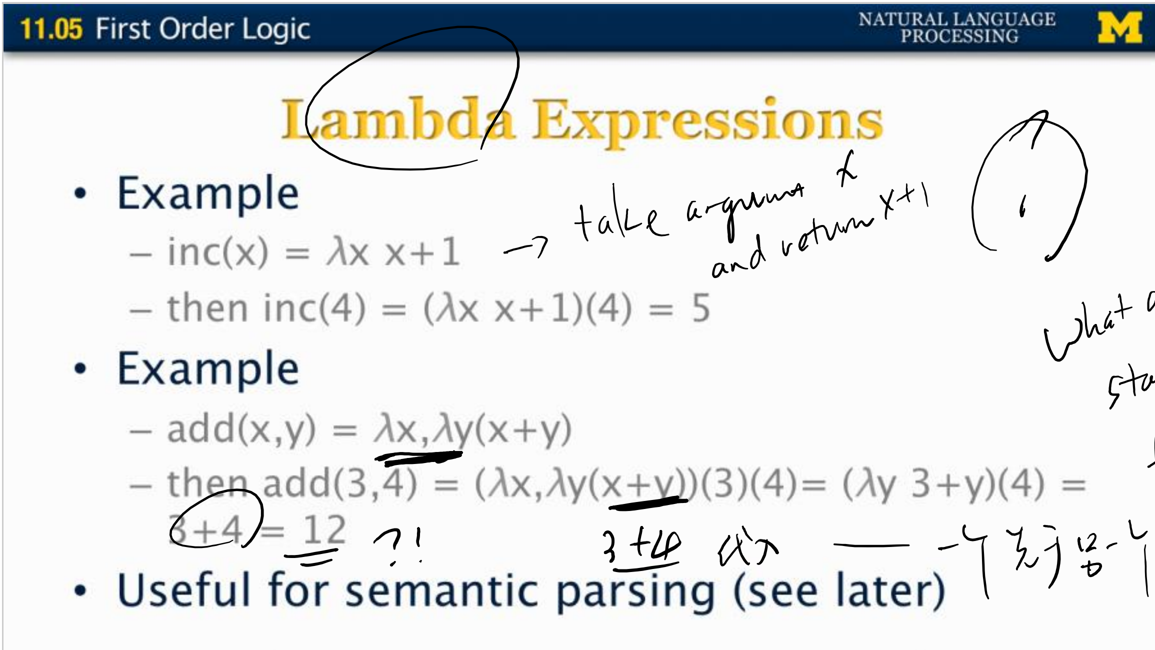
A.	There is a test that is long or hard.
B.	If a test is not long and not hard, then every student did not fail it.
C.	Every student studied for or passed the test.
D.	Every test that is long is also hard.
E.	Every student studied for the test or every student passed the test.
F.	If there is a test that is hard or not long, then at least one student failed it.
G.	Every test is long and hard.
H.	If a test is not both long and hard, then not every student failed it.

11.05 First Order Logic
NATURAL LANGUAGE PROCESSING

Solutions

1. D
   J
   B
   A
   E

2. E
   C
   D
   A
   H



What does λ stand for?
function?

11.6

NLP

11.06 Knowledge Representation
NATURAL LANGUAGE PROCESSING

Introduction to NLP

Knowledge Representation

Knowledge Representation

- Ontologies relationships between objects.
- Categories and objects
- Events
- Times
- Beliefs

11.06 Knowledge Representation
NATURAL LANGUAGE PROCESSING

Knowledge Representation

- Object
  - Martin the cat
- Categories
  - Cat
- Ontology
  - Mammal includes Cat, Dog, Whale
  - Cat includes PersianCat, ManxCat
- ISA relation
  - ISA (Martin,Cat) object in a category
- AKO relation
  - AKO (PersianCat,Cat) A kind of: two obj. Persian Cat is a kind of cat.
- HASA relation
  - HASA (Cat, tail) all cats have tails

Semantics of FOL

- FOL sentences can be assigned a value of *true* or *false.*

  ISA(Milo,Cat) = true

- *Milo is younger than Martin*

  <(AgeOf(Milo),AgeOf(Martin)) = true

  =(AgeOf(Milo),AgeOf(Martin)) = false

11.06 Knowledge Representation
NATURAL LANGUAGE PROCESSING

Examples with Quantifiers

- All cats eat fish

  ∀x:ISA(x,Cat)⇒EatFish(x)

Representing Events

- Martin ate
- Martin ate in the morning
- Martin ate fish
- Martin ate fish in the morning

One Possible Representation

- FOL representations
  - Eating1(Martin)
  - Eating2(Martin,Morning)
  - Eating3(Martin,Fish)
  - Eating4(Martin,Fish,Morning)
- Meaning postulates
  - Eating4(x,y,z) -> Eating3(x,y)
  - Eating4(x,y,z) -> Eating2(x,z)
  - Eating4(x,y,z) -> Eating1(x)

predicate. hard to get such predicates

many types of things. Can't exhaust all

Example from Jurafsky and Martin

Second Possible Representation

- Eating4(x,y,z)
  - With some arguments unspecified
- Problems
  - Too many commitments
  - Hard to combine Eating4(Martin,Fish,z) with Eating4(Martin,y,Morning)

Example from Jurafsky and Martin

Third Possible Representation

- Reification
  - $\exists$ e: ISA(e,Eating) $\wedge$ Eater(e,Martin) $\wedge$ Eaten(e,Fish)

Example from Jurafsky and Martin

Representing Time

- Example
  - Martin went from the kitchen to the yard
  - ISA(e,Going) ^ Goer(e,Martin) ^ Origin (e,kitchen) ^ Target (e,yard)
- Issue
  - no tense information: past? present? future?
- Fluents
  - A predicate that is true at a given time: T(f,t)

11.06 Knowledge Representation
NATURAL LANGUAGE PROCESSING

Representing Time

two events .

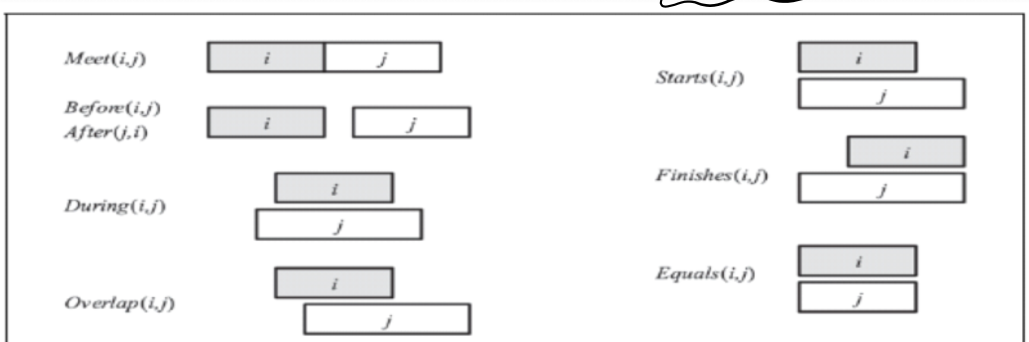


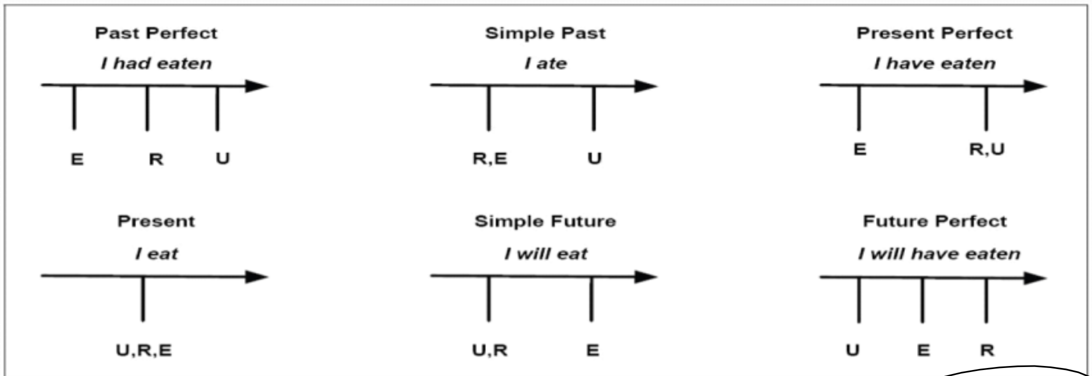
Figure 12.2 Predicates on time intervals.

Example from Russell and Norvig

11.06 Knowledge Representation

NATURAL LANGUAGE PROCESSING

Representing Time



Example from Jurafsky and Martin

11.06 Knowledge Representation

NATURAL LANGUAGE PROCESSING

Representing Beliefs

- Example
  - Milo believes that Martin ate fish
- One possible representation
  - $\exists$ e,b: ISA(e,Eating) $\wedge$ Eater(e,Martin) $\wedge$ Eaten(e,Fish) $\wedge$ ISA(b,Believing) $\wedge$ Believer(b,Milo) $\wedge$ Believed(b,e)
- However this implies (by dropping some of the terms) that "Martin ate fish" (without the Belief event)
- Modal logic
  - Possibility, Temporal Logic, Belief Logic

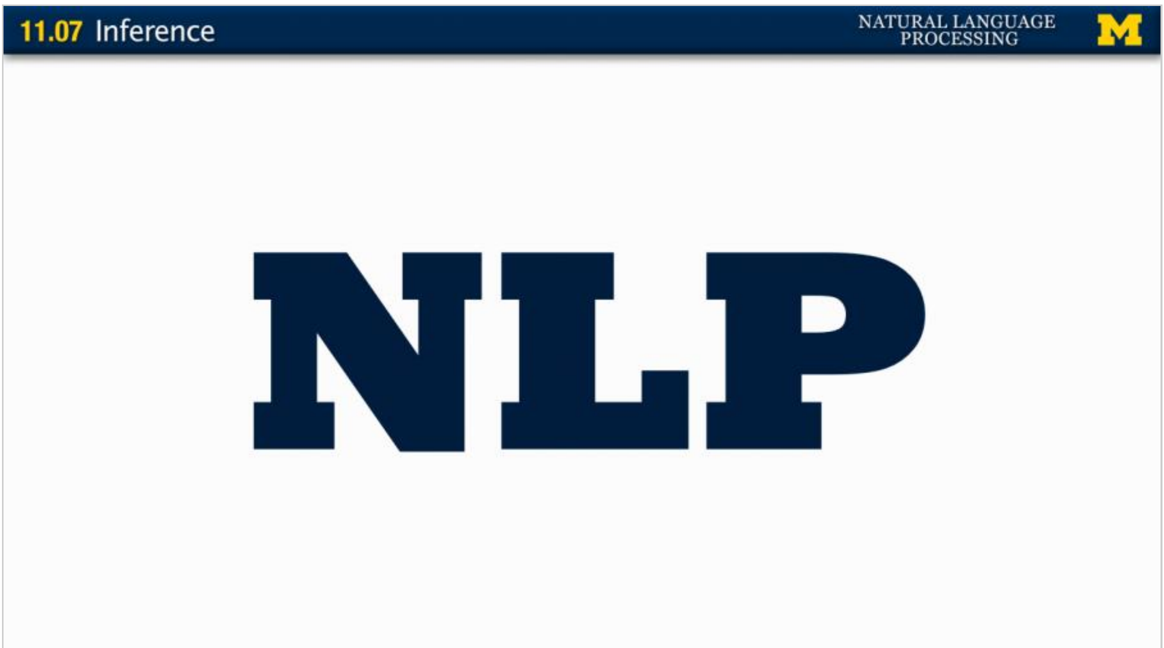
then this is not right --> should read more about modal logic to solve this.

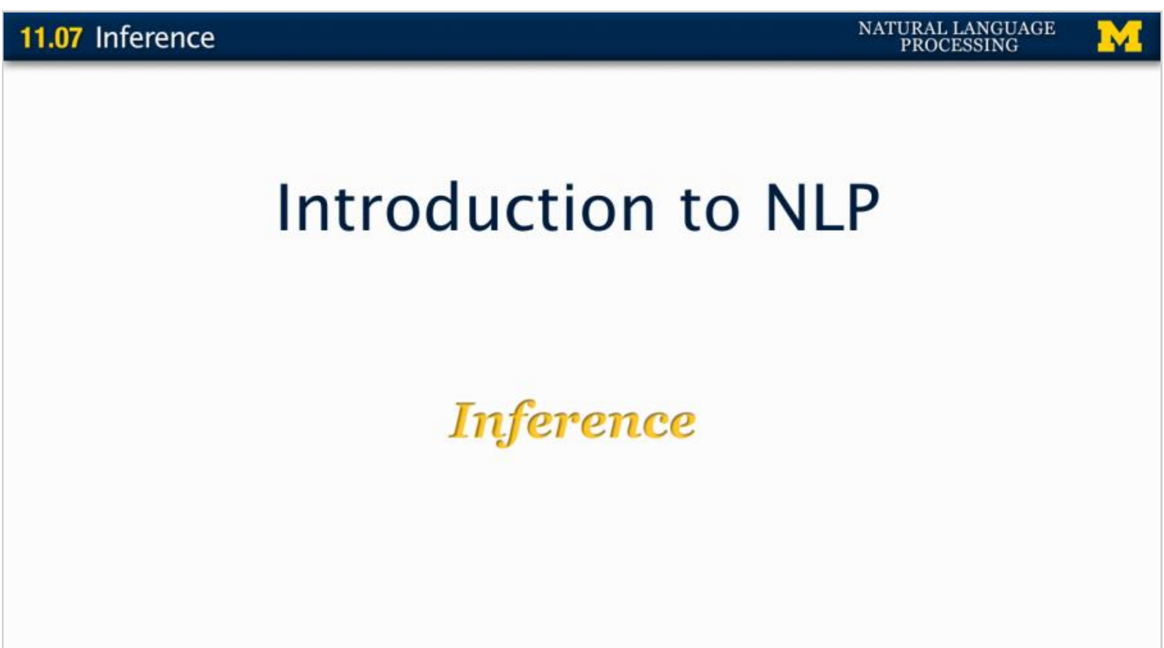
11.06 Knowledge Representation

NATURAL LANGUAGE PROCESSING

NLP

11.7


11.07 Inference
NATURAL LANGUAGE PROCESSING
NLP


11.07 Inference
NATURAL LANGUAGE PROCESSING
Introduction to NLP
Inference

Modus Ponens

- Modus ponens:

  α
  $\alpha \Rightarrow \beta$
  β

- Example:

  Cat(Martin)
  ∀ x: Cat(x) ⇒ EatsFish(x)
  EatsFish(Martin)

Inference

- Forward chaining
  - as individual facts are added to the database, all derived inferences are generated
- Backward chaining
  - starts from queries
  - Example: the Prolog programming language
- Prolog example
  - father(X, Y) :- parent(X, Y), male(X).
    parent(john, bill).
    parent(jane, bill).
    female(jane).
    male (john).
    ?- father(M, bill).

Prolog

Examples

The kinship domain:

- Brothers are siblings
  $\forall x,y\ Brother(x,y) \Rightarrow Sibling(x,y)$
- One's mother is one's female parent
  $\forall m,c\ Mother(c) = m \Leftrightarrow (Female(m) \wedge Parent(m,c))$
- "Sibling" is symmetric
  $\forall x,y\ Sibling(x,y) \Leftrightarrow Sibling(y,x)$

Universal Instantiation

- Every instantiation of a universally quantified sentence is entailed by it:

$$\frac{\forall v\ \alpha}{\mathrm{Subst}(\{v/g\}, \alpha)}$$

  for any variable v and ground term g

- E.g., $\forall x\ Cat(x) \wedge Fish(y) \Rightarrow Eats(x,y)$ yields:
  $Cat(Martin) \wedge \mathrm{Fish}(\mathrm{Blub})$

Existential Instantiation

- For any sentence α, variable *v*, and constant symbol *k* that does not appear elsewhere in the knowledge base:

$$\frac{\exists v\ \alpha}{\text{Subst}(\{v/k\}, \alpha)}$$

- E.g., $\exists x\ Cat(x) \wedge EatsFish(x)$ yields:

$$Cat(C_1) \wedge EatsFish(C_1)$$

provided C_1 is a new constant symbol, called a Skolem constant

Skolem constant.

Unification

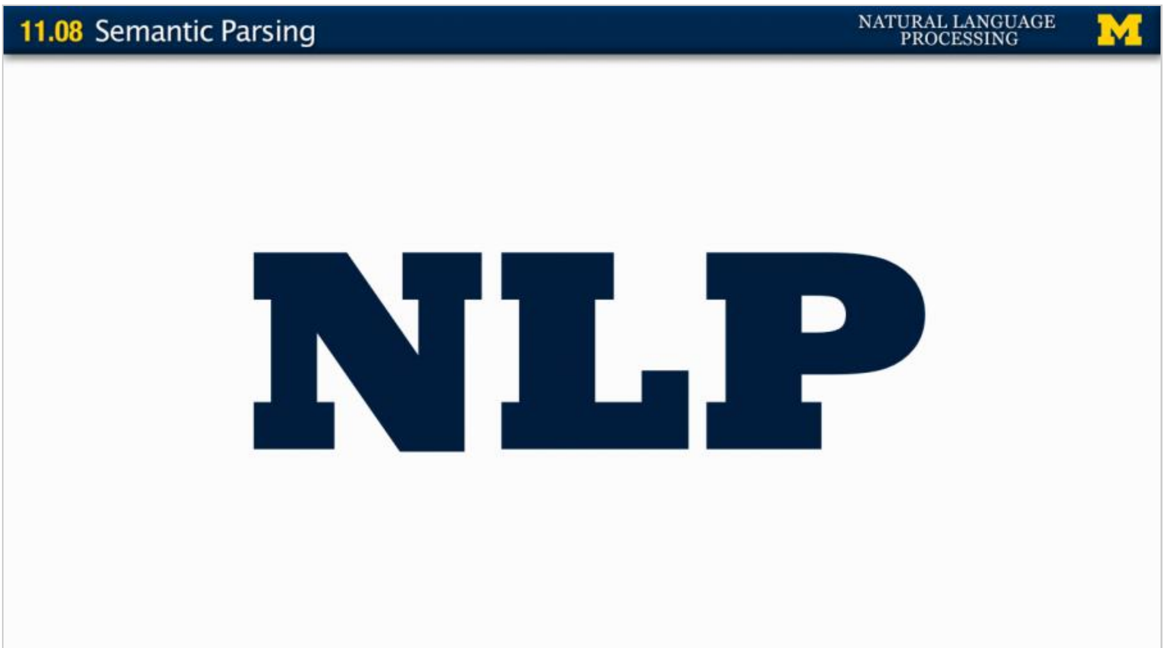
- If a substitution θ is available, unification is possible
- Examples:
  - p=Eats(x,y), q=Eats(x,Blub), possible if θ = {y/Blub}
  - p=Eats(Martin,y), q=Eats(x,Blub), possible if θ = {x,Martin,y/Blub}
  - p=Eats(Martin,y), q=Eats(y,Blub), fails because Martin≠Blub
- Subsumption
  - Unification works not only when two things are the same but also when one of them subsumes the other one
  - Example: All cats eat fish, Martin is a cat, Blub is a fish

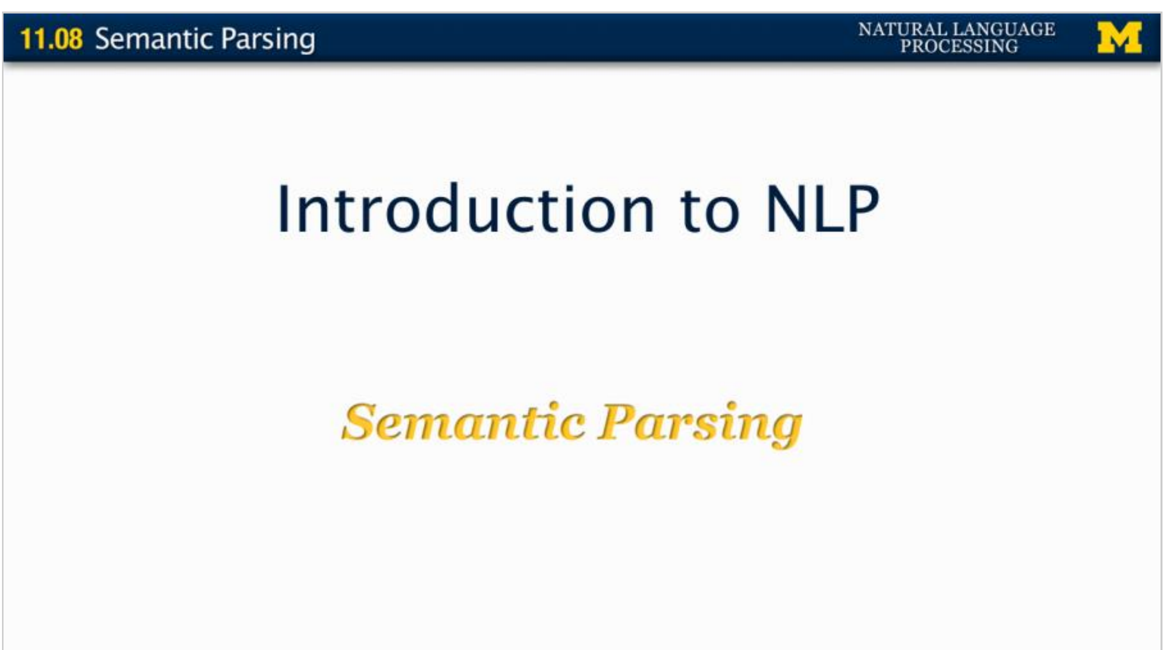
unification 就是放在同一级层面上.

Subsume 是指不同级的关系 一个 substitution 的条件

NLP

11.8


11.08 Semantic Parsing
NATURAL LANGUAGE PROCESSING
NLP


11.08 Semantic Parsing
NATURAL LANGUAGE PROCESSING
Introduction to NLP
Semantic Parsing

Semantic Parsing

- Converting natural language to a logical form
  - e.g., executable code for a specific application
- Example:
  - Airline reservations
  - Geographical query systems

11.08 Semantic Parsing NATURAL LANGUAGE PROCESSING

Stages of Semantic Parsing

- Input
  - Sentence
- Syntactic Analysis
  - Syntactic structure
- Semantic Analysis
  - Semantic representation

~~Syntactic~~

Semantic

Compositional Semantics

- Add semantic attachments to CFG rules
- Compositional semantics
  - Parse the sentence syntactically
  - Associate some semantics to each word
  - Combine the semantics of words and non-terminals recursively
  - Until the root of the sentence

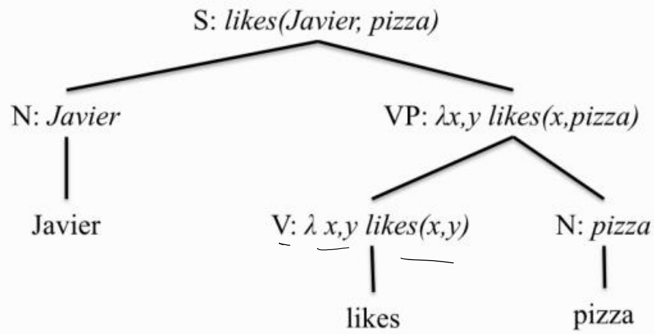
Example

- Input
  - Javier likes pizza
- Output
  - *like(Javier, pizza)*

Semantic Parsing

dependency parsing

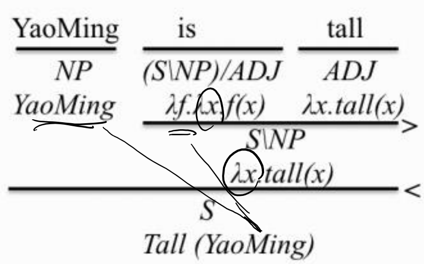
- Associate a semantic expression with each node



11.08 Semantic Parsing NATURAL LANGUAGE PROCESSING

Using CCG (Steedman 1996)

- CCG representations for semantics
  - ADJ: $\lambda x.tall(x)$
  - (S\NP)/ADJ : $\lambda f.\lambda x.f(x)$
  - NP: YaoMing



CCG

- NACLO problem from 2014
- Authors: Jonathan Kummerfeld, Aleka Blackwell, and Patrick Littell
- http://www.nacloweb.org/resources/problems/2014/N2014-O.pdf
- http://www.nacloweb.org/resources/problems/2014/N2014-OS.pdf
- http://www.nacloweb.org/resources/problems/2014/N2014-P.pdf
- http://www.nacloweb.org/resources/problems/2014/N2014-PS.pdf

Top Bar Reserved for U-M Branding and Course Information

CCG

One way for computers to understand language is by forming a structure that represents the relationships between words using a technique called Combinatorial Categorial Grammar (CCG). Computer scientists and linguists can use CCG to parse sentences (that is, try to figure out their structure) and then extract meaning from the structure.

As the name suggests, Combinatorial Categorial Grammar parses sentences by combining categories. Each word in a sentence is assigned a particular category; note that / and \ are two different symbols:

I	NP
books	NP
sleep	S \ NP
enjoy	(S \ NP) / NP

CCG

One way for computers to understand language is by forming a structure that represents the relationships between words using a technique called Combinatorial Categorial Grammar (CCG). Computer scientists and linguists can use CCG to parse sentences (that is, try to figure out their structure) and then extract meaning from the structure.

As the name suggests, Combinatorial Categorial Grammar parses sentences by combining categories. Each word in a sentence is assigned a particular category; note that / and \ are two different symbols:

I	NP
books	NP
sleep	S \ NP
enjoy	(S \ NP) / NP

These categories are then combined in systematic ways. We will not explain how, but we will give you two successful parses…

$$\dfrac{\overset{\text{I}}{NP} \quad \overset{\text{sleep}}{S \backslash NP}}{S}<$$

$$\dfrac{\overset{\text{I}}{NP} \quad \dfrac{\overset{\text{enjoy}}{(S \backslash NP) / NP} \quad \overset{\text{books}}{NP}}{S \backslash NP}>}{S}<$$

...and four unsuccessful parses…

$$\dfrac{\overset{\text{enjoy}}{(S \backslash NP) / NP} \quad \overset{\text{books}}{NP}}{S \backslash NP}>$$

$$\dfrac{\overset{\text{I}}{NP} \quad \overset{\text{sleep}}{S \backslash NP}}{S}< \quad \overset{\text{books}}{NP}$$

$$\overset{\text{I}}{NP} \quad \overset{\text{enjoy}}{(S \backslash NP) / NP}$$

$$\overset{\text{Books}}{NP} \quad \dfrac{\overset{\text{I}}{NP} \quad \overset{\text{sleep}}{S \backslash NP}}{S}<$$

If a parse is successful, the sentence is declared "grammatical"; if not, the sentence is declared "ungrammatical".

CCG

O1. Using the above examples as evidence, figure out how CCG parses sentences, and describe it briefly here:
O2. In the sentence "I enjoy long books", list all of the categories that, if assigned to "long", make the sentence have a successful parse.
O3. Not every grammatical sentence of English will be declared "grammatical" by the process above. Using only the words "I", "books", "sleep", and "enjoy", form a grammatically correct English sentence that will fail to parse given the categories above. You don't have to use all four of the words.

Answer

O1. CCG assigns a category to each word and constructs a parse by combining pairs of categories to form an S. Not all pairs of categories can combine. A pair is allowed to combine if one category (e.g. A) is contained within the category next to it (e.g. B / A) and lies on the side indicated by the slash (\ for left, / for right). When two categories combine, the result is a new category, taken from the left of the slash (B in this example).

O2. There are four categories that 'long' could have that would create a successful parse of 'I enjoy long books':

1. NP / NP
2. ((S \ NP) \ ((S \ NP) / NP)) / NP
3. ((S \ NP) / NP) \ ((S \ NP) / NP)
4. ((S / NP) \ NP) \ ((S \ NP) / NP)

The first of these is probably the most appropriate. Some possible reasons:

- It is by far the simplest. (After all, all our other categories are relatively simple.)
- It keeps the existing structure of the sentence (where "enjoy" combines with what follows it and then with what precedes it).
- "Long" describes "books" and not "enjoy", so it might make sense to keep them together.
- The first would be the only one to work if "long books" were in any other position.

O3. Possible answers: "I enjoy sleep", topicalized object sentences like "Books I enjoy" and "Sleep I enjoy".

Top Bar Reserved for U-M Branding and Course Information

This problem is a follow-up to problem O and has to be solved after that problem. Tok Pisin (also referred to as New Guinea Pidgin or Melanesian Pidgin) is a creole language spoken in the northern mainland of Papua New Guinea and surrounding islands. It is an official language and the mostly widely used language in the country, spoken by over 5 million people.

Many Tok Pisin words come originally from English – its name comes from "talk" and "pidgin"[1] -- but Tok Pisin isn't just English. It has a distinct grammar and uses these words in different (but systematic!) ways.

P1. Below are sentences in Tok Pisin with a scrambled list of English translations. Match each sentence to its English equivalent.

1.	Brata bilong em i stap rit.	
2.	Ol i stap dringim wara.	
3.	Ol i ken ritim buk bilong mi.	
4.	Em i ritim buk pinis.	
5.	Em i laik rit.	
6.	Susa bilong em i ken rait.	
7.	Susa bilong mi i boilim wara.	
8.	Wara i boil pinis.	

A.	He has read the book.
B.	My sister boils the water.
C.	They can read my book.
D.	His sister can write.
E.	His brother is reading.
F.	The water has boiled.
G.	He wants to read.
H.	They are drinking water.

Top Bar Reserved for U-M Branding and Course Information

CCG

P2. Translate the following Tok Pisin sentence into English:

Brata bilong mi i stap ritim buk bilong susa bilong mi.

____

P3. Translate the following English sentence into Tok Pisin:

Their sister wants to write a book.

____

Top Bar Reserved for U-M Branding and Course Information

P4. Describing these words in terms of their CCG categories (introduced in Problem O) highlights that these aren't English words combined according to English rules, but are Tok Pisin words combined according to Tok Pisin rules.

Match each Tok Pisin word to its CCG category. Some categories will be used more than once. The symbol S_b is short for 'Bare Clause'.

1.	bilong	
2.	brata	
3.	boil	
4.	boilim	
5.	buk	
6.	dringim	
7.	em	
8.	i	
9.	ken	
10.	laik	

11.	mi	
12.	ol	
13.	pinis	
14.	stap	
15.	raitim	
16.	rit	
17.	ritim	
18.	susa	
19.	wara	

A.	NP
B.	(NP \ NP) / NP
C.	(S \ NP) / (S_b \ NP)
D.	(S_b \ NP)
E.	(S_b \ NP) / NP
F.	(S_b \ NP) \ (S_b \ NP)
G.	(S_b \ NP) / (S_b \ NP)

P5. Explain your answer.

Top Bar Reserved for U-M Branding and Course Information

CCG

P1.

1.	*Brata bilong em i stap rit.*	E
2.	*Ol i stap dringim wara.*	H
3.	*Ol i ken ritim buk bilong mi.*	C
4.	*Em i ritim buk pinis.*	A
5.	*Em i laik rit.*	G
6.	*Susa bilong em i ken rait.*	D
7.	*Susa bilong mi i boilim wara.*	B
8.	*Wara i boil pinis.*	F

A.	He has read the book.
B.	My sister boils the water.
C.	They can read my book.
D.	His sister can write.
E.	His brother is reading.
F.	The water has boiled.
G.	He wants to read.
H.	They are drinking water.

P2. My brother is reading my sister's book.

P3. Susa bilong ol i laik raitim buk.

CCG

P4.

1.	bilong	B
2.	brata	A
3.	boil	D
4.	boilim	E
5.	buk	A
6.	dringim	E
7.	em	A
8.	i	C
9.	ken	G
10.	laik	G

11.	mi	A
12.	ol	A
13.	pinis	F
14.	stap	G
15.	raitim	E
16.	rit	D
17.	ritim	E
18.	susa	A
19.	wara	A

A.	NP
B.	(NP \ NP) / NP
C.	(S \ NP) / (S_b \ NP)
D.	(S_b \ NP)
E.	(S_b \ NP) / NP
F.	(S_b \ NP) \ (S_b \ NP)
G.	(S_b \ NP) / (S_b \ NP)

Top Bar Reserved for U-M Branding and Course Information

CCG

P5. A. Any noun or pronoun is category A (NP) because they can be used as a noun.

B. The word "bilong" shows possession of the preceding NP by the following NP; therefore, it is (NP\NP)/NP. Also, the phrase [NP bilong NP] yields a noun phrase (NP).

C. The word "i" is necessary for a grammatical sentence, so it is (S\NP)/(S_b\NP). It wants a following verb phrase (indicated by (S_b\NP)) and a preceding noun phrase (NP). NP+i+(S_b\NP) forms a sentence.

D. Each intransitive verb (boil and rit) can stand on its own as S_b\NP, forming the verb phrase.

E. Transitive verbs (boilim, dringim, raitim, ritim; the ones ending in -im), need a following NP, so they are categorized as (S_b\NP)/NP, a verb phrase followed by a noun phrase.

F. The verbs "stap," "ken," and "laik" precede the primary verb phrase and need another verb phrase to create an S_b\NP, so they are the category (S_b\NP)/(S_b\NP).

G. The verb "pinis" comes after the main verb, so it is of the category (S_b\NP)\(S_b\NP) which requires a (S_b\NP) to precede it.

GeoQuery (Zelle and Mooney 1996)

Earliest sys

What is the capital of the state with the largest population?

```
answer(C, (capital(S,C), largest(P, (state(S),
population(S,P))))).
```

What are the major cities in Kansas?

```
answer(C, (major(C), city(C), loc(C,S),
equal(S,stateid(kansas)))).
```

Form	Predicate
capital(C)	C is a capital (city).
city(C)	C is a city.
major(X)	X is major.
place(P)	P is a place.
river(R)	R is a river.
state(S)	S is a state.
capital(C)	C is a capital (city).
area(S,A)	The area of S is A.
capital(S,C)	The capital of S is C.
equal(V,C)	variable V is ground term C.
density(S,D)	The (population) density of S is P
elevation(P,E)	The elevation of P is E.
high_point(S,P)	The highest point of S is P.
higher(P1,P2)	P1's elevation is greater than P2's.
loc(X,Y)	X is located in Y.
low_point(S,P)	The lowest point of S is P.
len(R,L)	The length of R is L.
next_to(S1,S2)	S1 is next to S2.
size(X,Y)	The size of X is Y.
traverse(R,S)	R traverses S.

represented in this form

Type	Form	Example
country	countryid(Name)	countryid(usa)
city	cityid(Name, State)	cityid(austin,tx)
state	stateid(Name)	stateid(texas)
river	riverid(Name)	riverid(colorado)
place	placeid(Name)	placeid(pacific)

Zettlemoyer and Collins (2005)

Utah $:=$ NP
Idaho $:=$ NP
borders $:=$ $(S\backslash NP)/NP$

a) What states border Texas
$\lambda x.state(x) \wedge borders(x, texas)$

b) What is the largest state
$\arg\max(\lambda x.state(x), \lambda x.size(x))$

c) What states border the state that borders the most states
$\lambda x.state(x) \wedge borders(x, \arg\max(\lambda y.state(y), \lambda y.count(\lambda z.state(z) \wedge borders(y, z))))$

a) Utah borders Idaho

- Utah: NP : $utah$; borders: $(S\backslash NP)/NP$: $\lambda x.\lambda y.borders(y,x)$; Idaho: NP : $idaho$
- $>$: $(S\backslash NP)$: $\lambda y.borders(y, idaho)$
- $<$: S : $borders(utah, idaho)$

b) What states border Texas

- What: $(S/(S\backslash NP))/N$: $\lambda f.\lambda g.\lambda x.f(x) \wedge g(x)$; states: N : $\lambda x.state(x)$; border: $(S\backslash NP)/NP$: $\lambda x.\lambda y.borders(y,x)$; Texas: NP : $texas$
- $>$: $S/(S\backslash NP)$: $\lambda g.\lambda x.state(x) \wedge g(x)$; $>$: $(S\backslash NP)$: $\lambda y.borders(y, texas)$
- $>$: S : $\lambda x.state(x) \wedge borders(x, texas)$

Utah $:=$ $NP : utah$
Idaho $:=$ $NP : idaho$
borders $:=$ $(S\backslash NP)/NP : \lambda x.\lambda y.borders(y,x)$

11.08 Semantic Parsing

NATURAL LANGUAGE PROCESSING

Zettlemoyer and Collins (2005)

states	:=	$N : \lambda x.state(x)$
major	:=	$N/N : \lambda f \lambda x.major(x) \wedge f(x)$
population	:=	$N : \lambda x.population(x)$
cities	:=	$N : \lambda x.city(x)$
rivers	:=	$N : \lambda x.river(x)$
run through	:=	$(S \backslash NP)/NP : \lambda x.\lambda y.traverse(y, x)$
the largest	:=	$NP/N : \lambda f.\arg\max(f, \lambda x.size(x))$
river	:=	$N : \lambda x.river(x)$
the highest	:=	$NP/N : \lambda f.\arg\max(f, \lambda x.elev(x))$
the longest	:=	$NP/N : \lambda f.\arg\max(f, \lambda x.len(x))$

Figure 6: Ten learned lexical items that had highest associated parameter values from a randomly chosen development run in the Geo880 domain.

states	:=	$N : \lambda x.state(x)$
major	:=	$N/N : \lambda f \lambda x.major(x) \wedge f(x)$
population	:=	$N : \lambda x.population(x)$
cities	:=	$N : \lambda x.city(x)$
rivers	:=	$N : \lambda x.river(x)$
run through	:=	$(S \backslash NP)/NP : \lambda x.\lambda y.traverse(y, x)$
the largest	:=	$NP/N : \lambda f.\arg\max(f, \lambda x.size(x))$
river	:=	$N : \lambda x.river(x)$
the highest	:=	$NP/N : \lambda f.\arg\max(f, \lambda x.elev(x))$
the longest	:=	$NP/N : \lambda f.\arg\max(f, \lambda x.len(x))$

11.08 Semantic Parsing

NATURAL LANGUAGE PROCESSING

NLP