Lexical and Compositional Semantics

www.CLE.org.pk

Sarmad Hussain
sarmad.hussain@kics.edu.pk
PROFESSOR AND HEAD
Center for Language Engineering
Al- Khwarizmi Institute of Computer Sciences
University of Engineering and Technology, Lahore, Pakistan

Important Note

Some of this material is from:

Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition

By Daniel Jurafsky, James H. Martin
Published by Pearson Prentice Hall, 2008
ISBN 0131873210, 9780131873216

These slides are for Computational Linguistics courses

For course use only and not for further circulation or reuse due to possible copyright violations; Please purchase original book

2

Semantics

- Need a mechanism to relate the phonological, morphological and syntactic structures to the knowledge of the world
- Allows to perform tasks
  - Writing an essay
  - Decide what to order in a restaurant
  - Learn to use a software

3

Lexical Semantics

- Study of meanings and relations of words
- Lexeme
  - Individual entry in a lexicon
    - Orthographic form
    - Phonological form
    - Sense

4

Relations in Lexical Semantics

- Homonymy
  - Least semantically interesting
  - Same orthographic and phonological form
  - Same part of speech
  - But unrelated meaning
    - Bank (of river)
    - Bank (with money)

5

Relations in Lexical Semantics

- Homophones
  - Same phonological form but different orthographic form
    - Wood
    - Would
- Homographs
  - Same orthographic form but different phonological form
    - Bass (type of fish)
    - Bass (musical instrument)

6

Polysemy

- All categories so far have unrelated meaning
  - some resemblance in form
- Polysemy is resemblance in meaning
  - Evidence through etymology
  - Serve
    - Food
    - In a company
    - Time in prison

7

Polysemy

- How many distinct meanings?
- How are these related?
- How can they be distinguished?

8

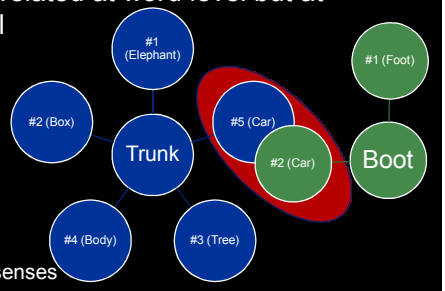
How many meanings?

- Identify different meanings
  - Trunk
    - Elephant
    - Tree
    - Car
    - Box
    - Body
- Label different meanings
  - Sense
    - Trunk#1, Trunk#2, …, Trunk#*n*
  - Explanation
  - Usage example in that sense

9

How is meaning related? Synonyms

- Words not related at word level but at sense level
  - Trunk
  - Boot



- Synoyms
  - synsets
    - Sets of senses
    - {trunk#5, boot#2, …}

10

How is meaning related? Hypernyms/Homonyms

- Meanings are not at the same level, but refer to categories at specific or general level in relation to each other
  - Hypernym
    - more general (parent) of a specific category
      - Vehicle is hypernym of car
  - Hyponym
    - More specific (child) of a general category
      - Dog is hyponym of animal

11

How is meaning related? Holonyms/Meronyms

- It is natural to see whole having parts or parts forming a whole
  - Holonym
    - Composite/entity with smaller parts/members
      - Tree is the holonym of trunk
  - Meronym
    - Part/member of a larger composite/entity
      - Bark is the meronym of trunk

12

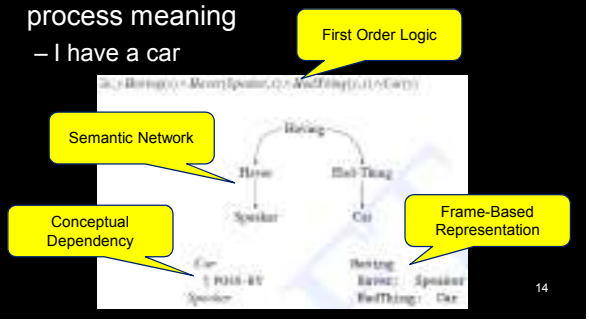
WordNet

- Lexical Database
  - Arranged by Senses/Synsets
    - Each synset has a unique number
  - Relationships
- Core – English WordNet
- Global WordNet (GWN)
  - Urdu WordNet – www.CLE.org.pk
  - Other languages of Pakistan?

13

Compositional Semantics

- Need a mechanism to represent and process meaning
  - I have a car



14

What is Needed in a Representation

- Verifiability
- Unambiguous Representation
- Canonical Form
- Inference and Variables
- Expressiveness

15

Verifiability

- To be able to determine the truth of the representation
- If ask a question: Does Maharani serve vegetarian food?
  - Need a knowledge base which contains facts, e.g. Serves(Maharani, VegetarianFood)
  - Need a computational system to match the representation of meaning in question with the knowledge base

16

Unambiguous Representation

- Multiple meanings of a sentence should be clearly represent-able
  - For example: I wanna eat someplace that's close to ICSI
- Need a mechanism to choose between multiple options
- Vagueness is not the same as ambiguous
  - Generic: I want to eat Italian food
  - Useful in various contexts

17

Canonical Form

- Different sentences with same meaning should be given the same form, called canonical form
  - E.g., these sentences mean the same
    - Does Maharani have vegetarian dishes?
    - Do they have vegetarian food at Maharani?
    - Are vegetarian dishes served at Maharani?
  - Difficult task as syntactic structure and lexical choices may be different
    - Need word senses and word sense disambiguation mechanism for lexical choices

18

Inference and Variables

- Inference: system's ability to derive conclusions based on input and stored facts
- Variables: ability to represent unknown entities; handle indefinite references
  - I would like to find a restaurant where I can get vegetarian food
  - Serves(x, VegetarianFood)

19

Expressiveness

- Can express wide range of subject matter, knowledge of the world and language
  - Hard to achieve

20

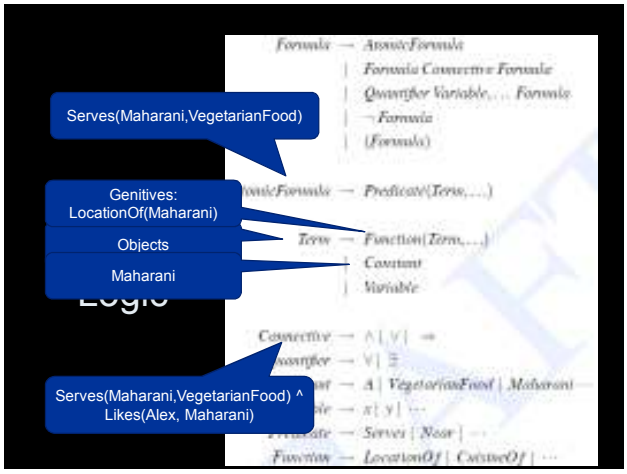
Model: Connecting Representation with the World

- Elements: Domain
  - Mathew, Fanco, Katie, Caroline
  - Frasca, Med, Rio
  - Italian, Mexican, Eclectic
- Properties: Sets of Elements
  - Noisy
    - Frasca, Med and Rio are noisy
- Relations: Sets of tuples of elements
  - Likes
    - Matthew likes the Med
    - Katie likes the Med and Rio
  - Serves
    - Med serves eclectic
    - Rio serves Mexican

First Order Logic

- Knowledge representation mechanism
- Provides computational basis for verifiability, inference, expressiveness
- Able to address the modeling requirements

22



FOL: Variables and Quantifiers

- Substitution semantics of quantifiers
  - A restaurant that serves Mexican food near ICSI
  - $\exists x\ Restaurant(x) \wedge Serves(x, MexicanFood) \wedge Near(LocationOf(x), LocationOf(ICSI))$
  - *All vegetarian restaurants serve vegetarian food*
  - $\forall x\ VegetarianRestaurant(x) \rightarrow Serves(x, VegetarianFood)$

24

FOL: Lambda Notation

- Provides the generic mechanism to define an expression to allow binding variables to specified terms
  - $\lambda x.P(x)$
  - This binding process is called lambda reduction
  - $\lambda x.P(x)\ (A) = P(A)$

25

FOL: Lambda Notation

- $\lambda x.\lambda y.\ Near(x,y)$
- $\lambda x.\lambda y.\ Near(x,y)\ (A) = \lambda y.\ Near(A,y)$
- $\lambda y.\ Near(A,y)\ (B) = Near(A,B)$

26

FOL: Inference Rules

- Modus Ponens
  - $\alpha \wedge \alpha \rightarrow \beta = \beta$

$$\frac{VegetarianRestaurant(Leaf) \qquad VegetarianRestaurant(x) \rightarrow Serves(x, VegetarianFood)}{Serves(Leaf, VegetarianFood)}$$

27

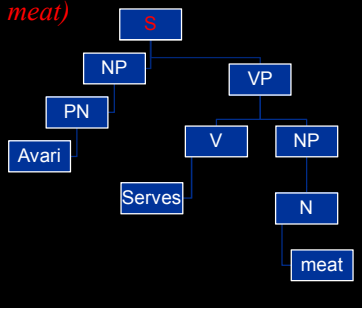
Semantic Analysis

- Avari serves meat
- $\exists e\ ISA(e,Serving) \wedge Server(e, Avari) \wedge Served(e, meat)$

28

Semantic Analysis

- $\exists e\ ISA(e,Serving) \wedge Server(e, Avari) \wedge Served(e, meat)$



29

Semantic Analysis

- Avari serves meat
- Lexicon
  - PN → Avari *{Avari}*
  - N → meat *{Meat}*
  - V → serves

  $\{\lambda x.\lambda y.\ \exists e\ ISA(e, Serving) \wedge Server(e, y) \wedge Served(e, x)\}$
    - Argument structure
    - roles

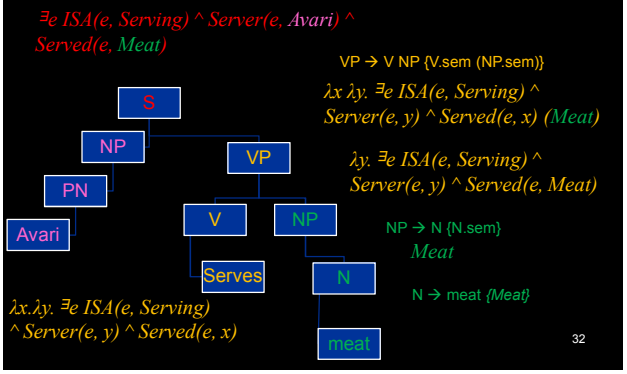
30

Semantic Analysis

- Avari serves meat
- Rules
  - NP → N {N.sem}
  - NP → PN {PN.sem}
  - VP → V NP {V.sem (NP.sem)}
  - S → NP VP {VP.sem (NP.sem)}

31

Semantic Analysis



32

Thank You

Questions

33