Lexical Semantic + Students Presentations ICS 482 Natural Language Processing

Lecture 26: Lexical Semantic + Students Presentations

Husni Al-Muhtaseb

بسم الله الرحمن الرحيم

ICS 482 Natural Language Processing

Lecture 26: Lexical Semantic + Students Presentations

Husni Al-Muhtaseb

NLP Credits and Acknowledgment

These slides were adapted from presentations of the Authors of the book

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

and some modifications from presentations found in the WEB by several scholars including the following

NLP Credits and Acknowledgment

If your name is missing please contact me
muhtaseb
At
Kfupm.
Edu.
sa

NLP Credits and Acknowledgment

Husni Al-Muhtaseb
James Martin
Jim Martin
Dan Jurafsky
Sandiway Fong
Song young in
Paula Matuszek
Mary-Angela Papalaskari
Dick Crouch
Tracy Kin
L. Venkata Subramaniam
Martin Volk
Bruce R. Maxim
Jan Hajič
Srinath Srinivasa
Simeon Ntafos
Paolo Pirjanian
Ricardo Vilalta
Tom Lenaerts

Heshaam Feili
Björn Gambäck
Christian Korthals
Thomas G. Dietterich
Devika Subramanian
Duminda Wijesekera
Lee McCluskey
David J. Kriegman
Kathleen McKeown
Michael J. Ciaraldi
David Finkel
Min-Yen Kan
Andreas Geyer-Schulz
Franz J. Kurfess
Tim Finin
Nadjet Bouayad

Khurshid Ahmad
Staffan Larsson
Robert Wilensky
Feiyu Xu
Jakub Piskorski
Rohini Srihari
Mark Sanderson
Andrew Elks
Marc Davis
Ray Larson
Jimmy Lin
Marti Hearst
Andrew McCallum
Nick Kushmerick
Mark Craven
Chia-Hui Chang
Diana Maynard

Martha Palmer
julia hirschberg
Elaine Rich
Christof Monz
Bonnie J. Dorr
Nizar Habash
Massimo Poesio
David Goss-Grubbs
Thomas K Harris
John Hutchins
Alexandros Potamianos
Mike Rosner
Latifa Al-Sulaiti
Giorgio Satta
Jerry R. Hobbs
Christopher Manning
Hinrich Schütze
Alexander Gelbukh
Gina-Anne Levow
Guitao Gao

Previous Lectures

- NLP Applications - Chatting with Alice
- Introduction and Phases of an NLP system
- Finite State Automata & Regular Expressions & languages
- Morphology: Inflectional & Derivational
- Parsing and Finite State Transducers, Porter Stemmer
- Statistical NLP – Language Modeling
- N Grams, Smoothing
- Parts of Speech - Arabic Parts of Speech
- Syntax: Context Free Grammar (CFG) & Parsing
- Parsing: Earley's Algorithm
- Probabilistic Parsing
- Probabilistic CYK (Cocke-Younger-Kasami)
- Dependency Grammar
- Invited Speech: Lexicons and Morphology
- Semantics: Representing meaning
- Semantics: First Order Predicate Calculus
- Semantic Analysis: Syntactic-Driven Semantic Analysis
- Information Extraction

Today's Lecture

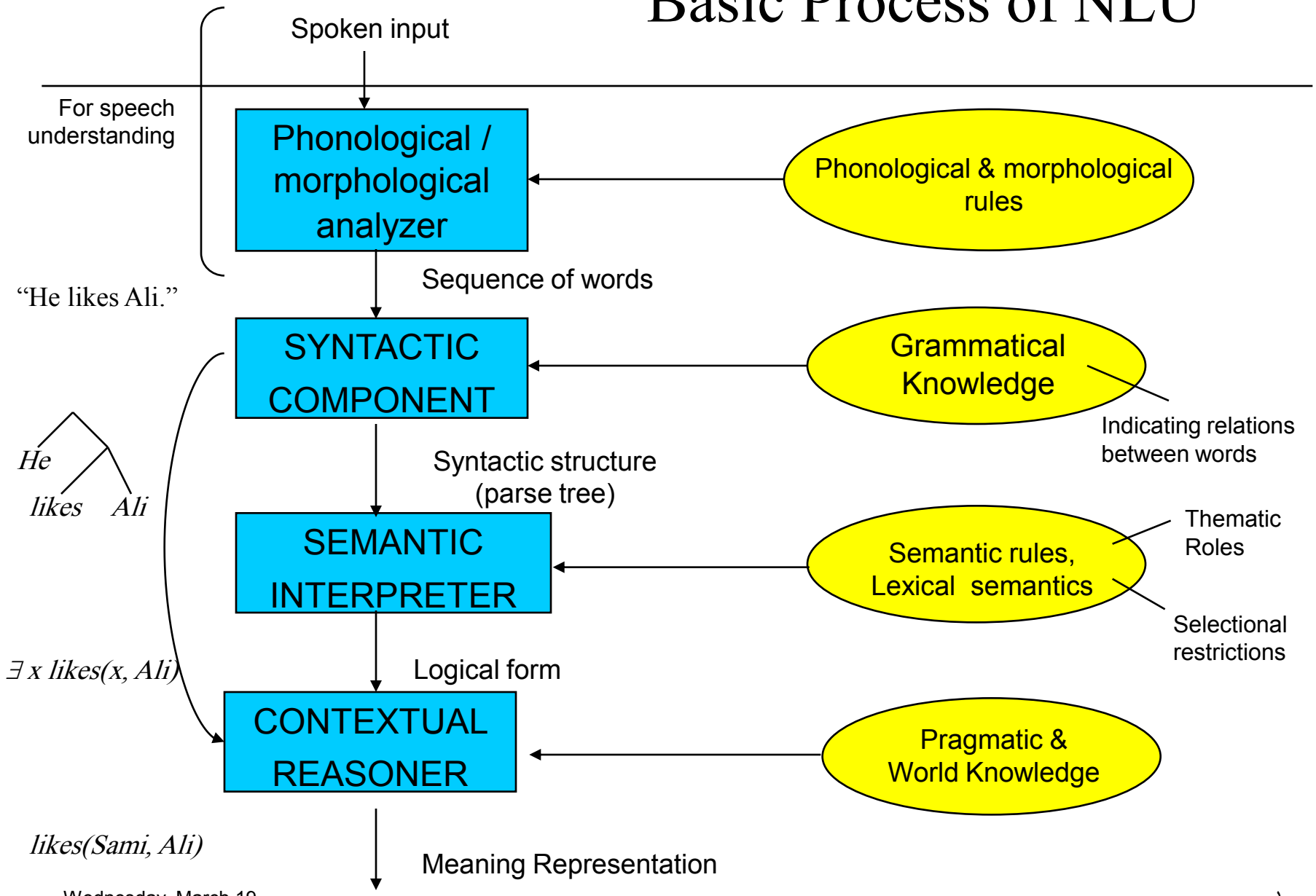
- Lexical Semantics (Ch 16)
- Students' Presentations

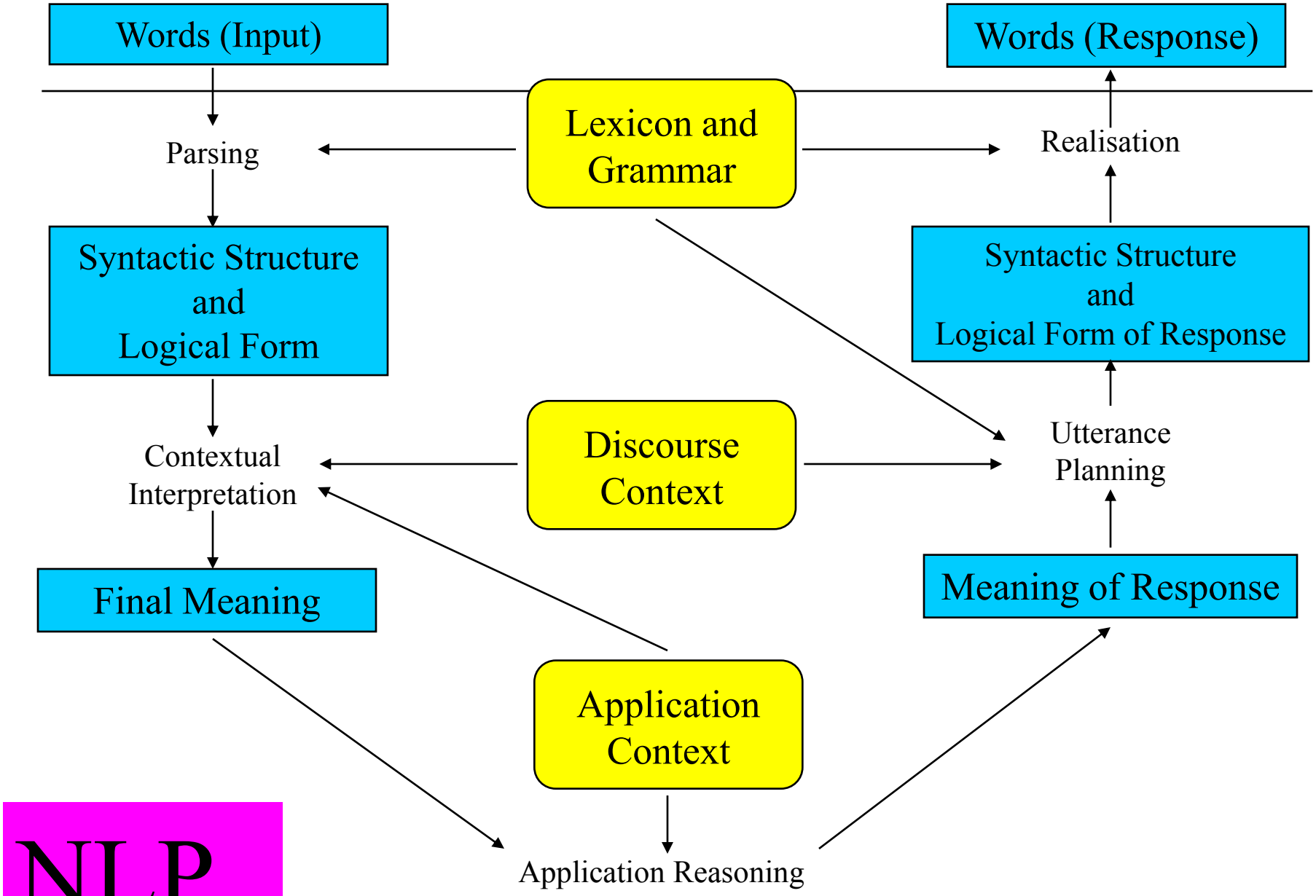
Students' Presentations

- Sunday May 20 (Today)
  - Abbas Al-Julaih - An Ambiguity-Controlled Morphological Analyzer for Modern Standard Arabic Modeling
  - AbdiRahman Daoud - Online Arabic Handwriting Recognition Using HMM (Did not attend)
  - Shaker Al-Anazi - How Do Search Engines Handle Arabic Queries?
- Tuesday, May 22 (Next Time)
  - Hussain AL-Ibrahem - Arabic Tokenization, Part-of-Speech Tagging
  - Ahmed Bukhamsin - Hybrid Method for Tagging Arabic Text
  - Al-Ansari, Naser - Light Stemming for Arabic Information Retrieval

Lexical Semantics (Chapter 16)

Basic Process of NLU




Words (Input)
Words (Response)
Lexicon and Grammar
Parsing
Realisation
Syntactic Structure and Logical Form
Syntactic Structure and Logical Form of Response
Discourse Context
Contextual Interpretation
Utterance Planning
Final Meaning
Meaning of Response
Application Context
Application Reasoning

NLP

Meaning

- Traditionally, meaning in language has been studied from three perspectives
  - The meaning of a text or discourse
  - The meanings of individual sentences or utterances
  - The meanings of individual words
- We started in the middle, now we'll move down to words and then we should move back up to discourse

Word Meaning

- We didn't assume much about the meaning of words when we talked about sentence meanings
  - Verbs provided a template-like predicate argument structure
  - Nouns were practically meaningless constants
- There has be more to it than that

Preliminaries

- What's a word?
  - Types, tokens, stems, roots, inflected forms, etc...
  - Lexeme: An entry in a lexicon consisting of a pairing of a form with a single meaning representation
  - Lexicon: A collection of lexemes
- Lexeme: an entry in the lexicon that includes
  - an orthographic representation
  - a phonological form
  - a symbolic meaning representation or sense
- Dictionary entries:
  - Red ('red) n: the color of blood or a ruby
  - Blood ('bluhd) n: the red liquid that circulates in the heart, arteries and veins of animals

Relation Among Lexemes & Their Senses

- Homonymy
- Synonymy
- Polysemy
- Metonymy
- Hyponymy/Hypernym
- Meronymy
- Antonymy

Relation Among Lexemes & Their Senses

- Homonymy:
  - Lexemes that share a form
    - Phonological, orthographic or both
  - example:
    - Bat مضرب (wooden stick-like thing) vs
    - Bat وطواط(flying scary mammal thing)

Synonymy

- Different ways of expressing related concepts
- Examples
  - cat, feline, Siamese cat
- Overlaps with basic and subordinate levels
- Synonyms are almost never truly substitutable:
  - Used in different contexts
  - Have different implications
    - This is a point of debate

Polysemy

- Most words have more than one sense
  - Homonym: same word, different meaning
    - bank (river)
    - bank (financial)
  - Polysemy: different senses of same word
    - That dog has floppy ears.
    - He has a good ear for jokes.
    - bank (financial) has several related senses
      - the building, the institution, the notion of where money is stored

Metonymy

- Use one aspect of something to stand for the whole
  - Newscast: “The White House released new figures today.”
  - Metaphor: Assuming the White house can release figures (like a person)

Hyponymy/Hypernym

- ISA relation
- Related to Superordinate and Subordinate level categories
  - hyponym(robin,bird)
  - hyponym(bird,animal)
  - hyponym(emus,bird)
- A is a hypernym of B if B is a type of A
- A is a hyponym of B if A is a type of B

Basic-Level Categories

- Folk biology:
  - {Unique beginner}: plant, animal
    - Life form: tree, bush, flower
      - *Generic name: pine, oak, maple, elm*
        - Specific name: Ponderosa pine, white pine
          - Varietals name: Western Ponderosa pine
- No overlap between levels
- Level 3 is basic
  - Corresponds to genus
  - Folk biological categories correspond accurately to scientific biological categories only at the basic level

Psychologically Primary Levels

SUPERORDINATE	animal	furniture
BASIC LEVEL	dog	chair
SUBORDINATE	terrier كلب صيد	rocker كرسي هزاز

- Children take longer to learn superordinate
- Superordinate not associated with mental images or motor actions !

Meronymy

- Parts-of relation
  - part of(beak$_{منقار}$, bird)
  - part of(bark$_{لحاء}$, tree)
- Transitive conceptually but not lexically:
  - The knob is a part of the door.
  - The door is a part of the house.
  - ? The knob is a part of the house ?

Antonymy

- Lexical opposites
  - antonym(large, small)
  - antonym(big, small)
  - antonym(big, little)
  - but *not* large, little

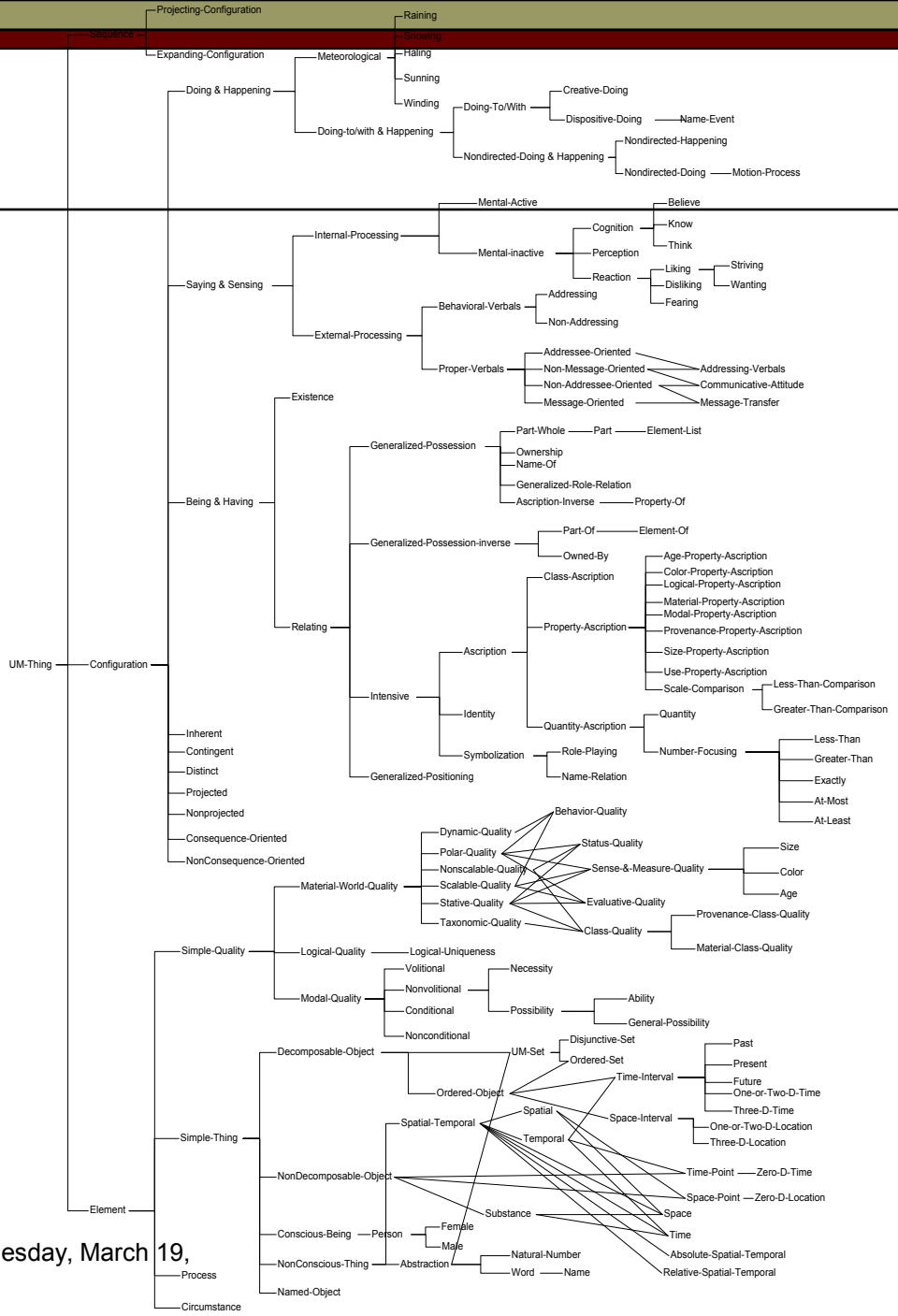
Thesauri and Lexical Relations

- Polysemy: Same word, different senses of meaning
  - Slightly different concepts expressed similarly
- Synonyms: Different words, related senses of meanings
  - Different ways to express similar concepts
- Thesauri help draw all these together
- Thesauri also commonly define a set of relations between terms that is similar to lexical relations

What is an Ontology?

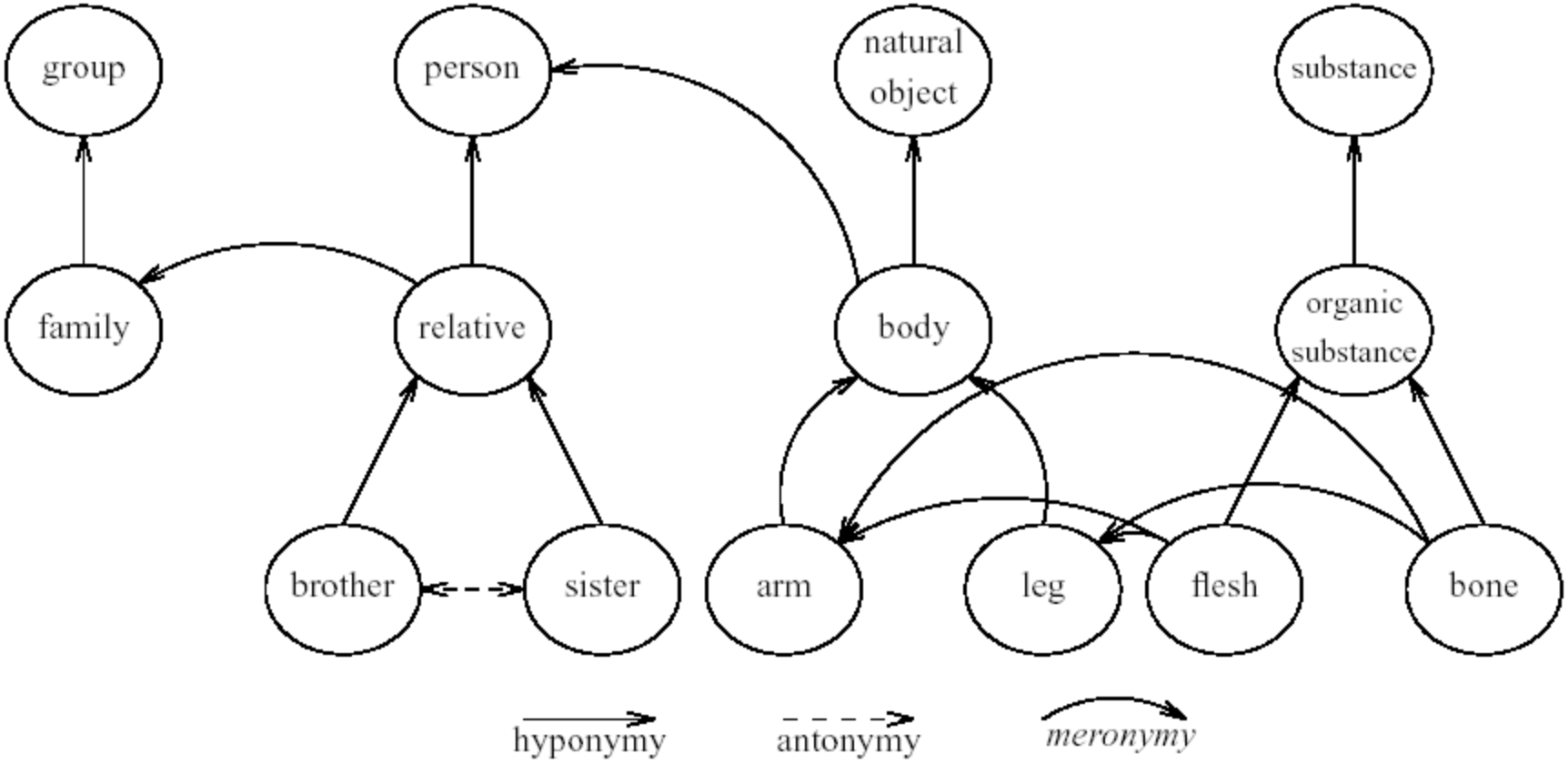
- From Merriam-Webster's Collegiate:
  - A branch of metaphysics concerned with the nature and relations of being
  - A particular theory about the nature of being or the kinds of existence
- Or:
  - A carving up of the world's meanings
  - Determine what things exist, but not how they inter-relate
- Related terms:
  - Taxonomy, dictionary, category structure
- Commonly used now in CS literature to describe structures that *function* as Thesauri

Example of Ontology



http://www.cogsci.princeton.edu/~wn/5papers.pdf

Figure 2. Network representation of three semantic relations among an illustrative variety of lexical concepts



http://www.cogsci.princeton.edu/~wn/5papers.pdf

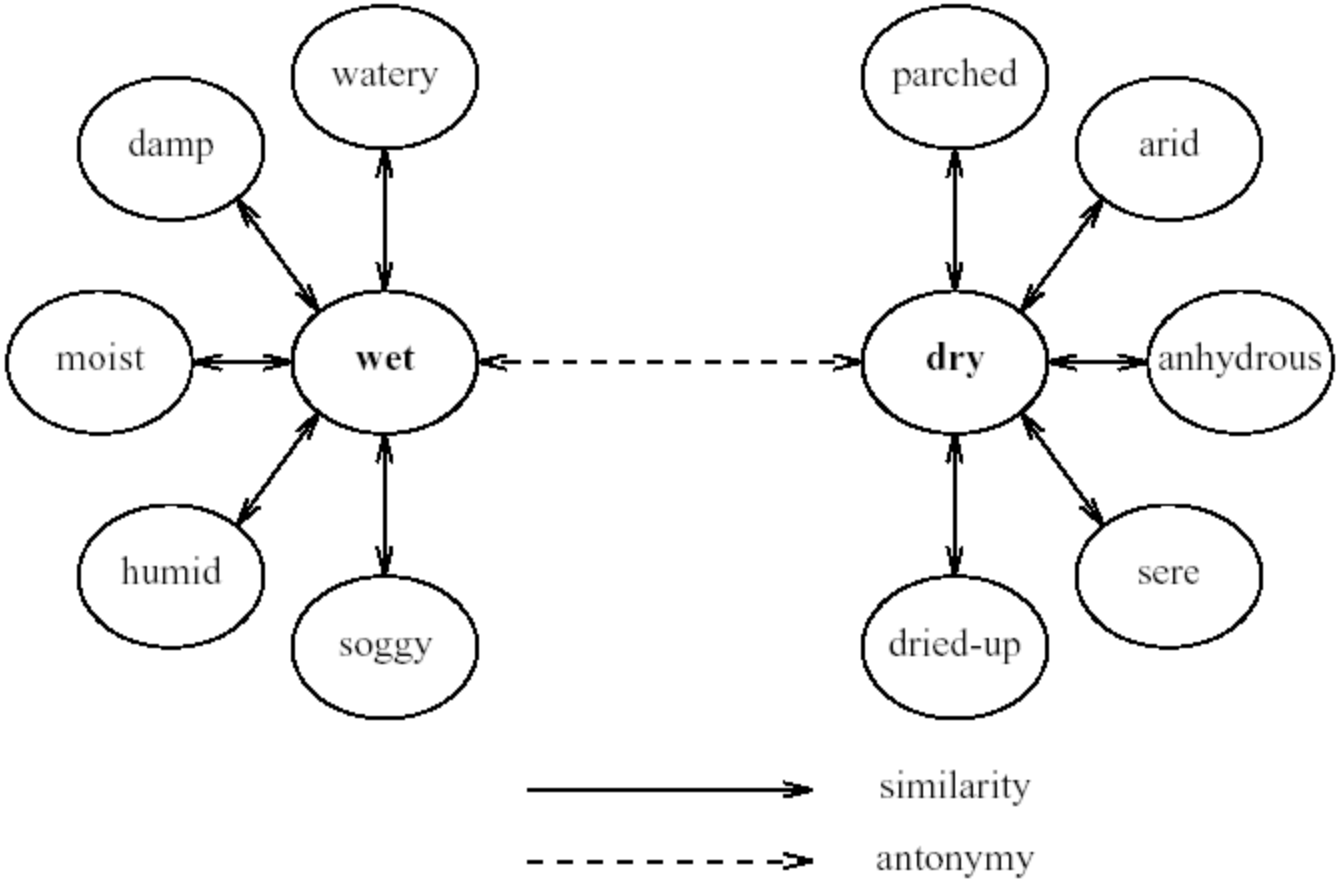


Figure 1. Bipolar Adjective Structure

Resources

- There are lots of lexical resources available these days…
  - Word lists
  - On-line dictionaries
  - Corpora
- The most ambitious one is WordNet
  - A database of lexical relations for English
    - Versions for other languages are under development

WordNet

- The critical thing to grasp about WordNet is the notion of a synset; its their version of a sense or a concept
  - Synset: set of synonyms, a dictionary-style definition (or gloss), and some examples of uses --> a concept
  - Databases for nouns, verbs, and modifiers
- Example: table as a verb to mean defer
  - > {postpone, hold over, table, shelve, set back, defer, remit, put off}
- For WordNet, the meaning of this sense of table is this list.

WordNet 2.1 newer than the one in the book

POS	Unique Strings	Synsets	Total Word-Sense Pairs
Noun	117097	81426	145104
Verb	11488	13650	24890
Adjective	22141	18877	31302
Adverb	4601	3644	5720
Totals	155327	117597	207016

Lexical Relations in WordNet

Relation	Definition	Example
Hypernym	From concepts to superordinates	*breakfast* → *meal*
Hyponym	From concepts to subtypes	*meal* → *lunch*
Has-Member	From groups to their members	*faculty* → *professor*
Member-Of	From members to their groups	*copilot* → *crew*
Has-Part	From wholes to parts	*table* → *leg*
Part-Of	From parts to wholes	*course* → *meal*
Antonym	Opposites	*leader* → *follower*

Relation	Definition	Example
Hypernym	From events to superordinate events	*fly* → *travel*
Troponym	From events to their subtypes	*walk* → *stroll*
Entails	From events to the events they entail	*snore* → *sleep*
Antonym	Opposites	*increase* ⟺ *decrease*

Relation	Definition	Example
Antonym	Opposite	*heavy* ⟺ *light*
Adverb	Opposite	*quickly* ⟺ *slowly*

Structure of WordNet

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
Synonym	A concept that means exactly or nearly the same as another. *WordNet* considers immediate hypernyms to be synonyms.	×	×	×	×	{ ***sofa, couch, lounge*** } are all synonyms of one another. { ***seat*** } is the immediate hypernym of the synset.
Antonym	A concept opposite in meaning to another.	×	×	×	×	{ ***love*** } is the antonym of { ***hate, detest*** }.
Hypernym	A concept whose meaning denotes a superordinate.	×	×			A { ***feline, felid*** } is a hypernym of { ***cat, true cat*** }.
Hyponym	A concept whose meaning denotes a subordinate.	×	×			A { ***wildcat*** } is a hyponym of { ***cat, true cat*** }.
Substance meronym	A concept that is a substance of another concept.	×				A { ***snowflake, flake*** } is substance of { ***snow*** }.

Structure of WordNet

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
Part meronym	A concept that is a part of another concept.	×				A { ***crystal, watch crystal, watch glass*** } is a part of a { ***watch, ticker*** }.
Member meronym	A concept that is a member of another concept.	×				An { ***associate*** } is a member of an { ***association*** }.
Substance of holonym	A concept that has another concept as a substance.	×				A { ***tear, teardrop*** } has { ***water, H20*** } as a substance.
Part of holonym	A concept that has another concept as a part.	×				A { ***school system*** } has a { ***school, schoolhouse*** } as a part.
Member of holonym	A concept that has another concept as a member.	×				{ ***organized crime, gangland, gangdom*** } has { ***gang, pack, ring, mob*** } as a member.
Attribute	An adjective that is the value of a noun.	×				{ ***fast (vs. slow)*** } is a value of { ***speed, swiftness, fastness***}

Structure of WordNet

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
Cause to	A verb that is the cause of a result.		×			{ ***give*** } is the cause of the result { ***have, have got, hold*** }
Entailment	A verb that involves unavoidably a result.		×			To { ***die, decease, perish, go, exit, pass away, expire*** } involves unavoidably to { ***leave, leave behind*** }.
Troponym	A verb that is a particular way to do another.		×			To { ***samba*** } is a particular way to { ***dance, trip the light fantastic*** }.
Pertainym	An adjective or adverb that relates to a noun.			×	×	{ ***criminal*** } relates to { ***crime*** }.
Attribute	An adjective that is the value of a noun.	×				{ ***fast (vs. slow)*** } is a value of { ***speed, swiftness, fastness***}
Value	A noun that has an adjective for a value.			×		{ ***weight*** } has { ***light (vs. heavy)*** } as a value.

WordNet Usage

- Available online if you wish to try it…

  http://wordnet.princeton.edu/

Arabic WordNet ?

Reminder: Choose your project

- Arabic POS Tagger
- Specific Information Picker
- An Arabic morphological analyzer
- An Arabic Spell checker w/ morphology analysis
- An Arabic Syntax analyzer
- Random syntactically-correct Arabic sentence generator
- An English to Arabic machine translation using word re-ordering
- Your Own: Let us discuss

Thank you

- السلام عليكم ورحمة الله