Semantic Processing And Semantic Representations

Lu Xiao
lxiao04@syr.edu
213 Hinds Hall

adopted some materials developed in previous courses by Nancy McCracken, Liz Liddy and others; and some instructor resources for the book “Speech and Language Processing” by Daniel Jurafsky and James H. Martin

Lexical Semantics:

WordNet And Word Senses, Semantic Lexical Resources, and Ontologies

Word Senses

- We say that a word has more than one word sense if there is more than one definition.

Online dictionary definitions (partial) for the noun *plant*

1. a living organism of the kind exemplified by trees, shrubs, herbs, grasses, ferns, and mosses, typically growing in a permanent site, absorbing water and inorganic substances through its roots, and synthesizing nutrients in its leaves by photosynthesis using the green pigment chlorophyll.
2. a place where an industrial or manufacturing process takes place

Lexical Semantics

- Lexicons – words (or lexemes or stems) together with some information
- Dictionaries – a lexicon with definitions for each word sense
  - Most are now available online
- Thesauruses – add synonyms for each word sense
  - Roget Thesaurus
  - WordNet
- Semantic networks – add more semantic relations
  - WordNet
  - EuroWordNet
- Ontologies – add semantic relations and rules about entities, concepts and relations

Knowledge Resources - Semantic Lexicons

- Lexicon where each word is assigned to a semantic class
- Lexical resources have been developed to assign words to semantic classes in support of applications that need to detect opinion, sentiment, or other more subjective meanings
- Examples:
  - Subjectivity Lexicon
  - LIWC
  - ANEW

Semantic Classes: Subjectivity Lexicon

- Subjectivity Lexicon from the MPQA project with Jan Wiebe
  - Gives a list of 8,000+ words that have been judged to be weakly or strongly positive, negative or neutral in subjectivity
  - Examples:

type=weaksubj len=1 word1=abandoned pos1=adj stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=abandonment pos1=noun stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=abandon pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abase pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abasement pos1=anypos stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abash pos1=verb stemmed1=y priorpolarity=negative
type=weaksubj len=1 word1=abate pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=absolve pos1=verb stemmed1=y priorpolarity=positive
type=strongsubj len=1 word1=absolute pos1=adj stemmed1=n priorpolarity=neutral

Semantic Classes: LIWC

- Linguistic Inquiry and Word Count (http://www.liwc.net/)
  - Text analysis software based on dictionaries of word dimensions
  - Dimensions can be syntactic
    - Pronouns, past-tense verbs
  - Dimensions can be semantic
    - Social words, affect, cognitive mechanisms
  - Other categories
    - See http://liwc.wpengine.com/compare-dictionaries/
    - James Pennebaker, Univ. of Texas at Austin
- Often used for positive and negative emotion words in opinion mining

Semantic Classes: LIWC

	wc	function	pronoun	ppron	I	we	you
. it is properly set up to serve readers who are looking for one or another entity or building named verizon . it is hard to find `` verizon fios '' , for example , by scanning only within the current main verizon article . are the buildings sometimes known as `` verizon '' explicitly covered there ? if so , that would be odd , as probably not all of them are important in the context of a huge business . there is no requirement that the main verizon article should keep mentioning these at all , much less in a convenient list ; in fact it only mentions some of them in scattered places within long article . keeping the disambiguation page ensures that all usages are kept together to serve disambiguation need (for readers and for editors , too) . --	125	49.6	8.8	1.6	0	0	0
. as nominator seems to acknowledge , the organization is wikipedia-notable (passive-aggressively (?) stating it is `` possibly notable '' , in the nominator 's words) , with a complaining tone about not being successful in changing the article . noam chomsky 's participation is important and so is the participation of other listed notables . these endorsements are relevant and important , much like the endorsements of other scientists along with albert einstein were important in establishing the importance of the russellã¢â‚¬â„¢einstein manifesto , for just one example . the al-jazeera reference alone is plenty to establish notability . other complaints about `` promotionalism '' are for tagging or talk page discussion , not reasons to delete . i dunno , maybe i want to join this worthy-sounding organization . it is clearly wikipedia-notable . i do n't dismiss dgg 's concerns and i will watch the article and try to participate constructively , but the concerns seem not about notability . --	149	45.64	7.38	2.68	2.68	0	0

Semantic Classes For Words: ANEW

- Affective Norms for English Words
  - Provides a set of emotional ratings for a large number of words in the English language
- Participants gave graded reactions from 1-9 on three dimensions
  - Good/bad, psychological valence
  - Active/passive, arousal valence
  - Strong/weak, dominance valence
- From the NIMH Center for the Study of Emotion and Attention at the University of Florida
  - http://csea.phhp.ufl.edu/Media.html
  - See also the paper by Dodds and Danforth on Happiness of Large-Scale Written Expressions

Knowledge Resources - Dictionary

- For each word in the language vocabulary, a dictionary provides:
  - A list of meanings
  - Definitions (for all word meanings)
  - Typical usage examples (for most word meanings)

WordNet definitions(called glosses)/examples for synsets of the noun *plant*

1. buildings for carrying on industrial labor; "they built a large plant to manufacture automobiles“
2. a living organism lacking the power of locomotion
3. something planted secretly for discovery by another; "the police used a plant to trick the thieves"; "he claimed that the evidence against him was a plant"
4. an actor situated in the audience whose acting is rehearsed but seems spontaneous to the audience

Knowledge Resources - Thesaurus

- A thesaurus adds:
  - An explicit synonymy relation between word meanings

WordNet synsets for the noun "plant"
1. plant, works, industrial plant
2. plant, flora, plant life

Knowledge Resources – Semantic Network

- A semantic network adds relations for each word sense:
  - hypernymy/hyponymy (IS-A),
    - hypernyms are more general, hyponyms are more specific
  - meronymy/holonymy (PART-OF),
  - antonymy, entailment, etc.

WordNet related concepts for the meaning "plant life"

{plant, flora, plant life}

hypernym:	{organism, being}
hypomym:	{house plant}, {fungus}, …
meronym:	{plant tissue}, {plant part}
holonym:	{Plantae, kingdom Plantae, plant kingdom}

WordNet

- WordNet is a database of facts about words
  - Meanings and the relations among them
- Words are organized into clusters of synonyms
  - Synsets
- http://wordnet.princeton.edu/
- Organized into nouns, verbs, adjectives, and adverbs
  - Currently 170,000 synsets
  - Available for download, arranged in separate files (DBs)

WordNet Relations

- A more detailed list from Jurafsky and Martin

Relation	Also Called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	*breakfast*^{1} → *meal*^{1}
Hyponym	Subordinate	From concepts to subtypes	*meal*^{1} → *lunch*^{1}
Instance Hypernym	Instance	From instances to their concepts	*Austen*^{1} → *author*^{1}
Instance Hyponym	Has-Instance	From concepts to concept instances	*composer*^{1} → *Bach*^{1}
Member Meronym	Has-Member	From groups to their members	*faculty*^{2} → *professor*^{1}
Member Holonym	Member-Of	From members to their groups	*copilot*^{1} → *crew*^{1}
Part Meronym	Has-Part	From wholes to parts	*table*^{2} → *leg*^{3}
Part Holonym	Part-Of	From parts to wholes	*course*^{7} → *meal*^{1}
Substance Meronym		From substances to their subparts	*water*^{1} → *oxygen*^{1}
Substance Holonym		From parts of substances to wholes	*gin*^{1} → *martini*^{1}
Antonym		Semantic opposition between lemmas	*leader*^{1} ⟺ *follower*^{1}
Derivationally Related Form		Lemmas w/same morphological root	*destruction*^{1} ⟺ *destroy*^{1}

WordNet Hierarchies

```
Sense 3
bass, basso --
(an adult male singer with the lowest voice)
=> singer, vocalist, vocalizer, vocaliser
    => musician, instrumentalist, player
        => performer, performing artist
            => entertainer
                => person, individual, someone...
                    => organism, being
                        => living thing, animate thing,
                            => whole, unit
                                => object, physical object
                                    => physical entity
                                        => entity
                    => causal agent, cause, causal agency
                        => physical entity
                            => entity
```

Representations Of Semantics: Ontology

- Ontology is an approach of knowledge organization
  - We can think of ontology as categorizing everything in the world.
- Representations:
  - Concepts, e.g. *person, animal, food, table, movie, etc*.
  - Instances (or entities), e.g. Barack Obama is an *instance* of the concept "person".
  - Properties, e.g. a person has properties of *gender, height, weight, father, mother, etc.*
  - Relations, e.g. Syracuse University is *located in* Syracuse.
  - Rules between concepts, properties, and relations, e.g., if someone is *married*, then he/she should have a *spouse*.

Semantics: Word Sense Disambiguation

Word Sense Disambiguation

- Definition
  - Correct selection of the appropriate sense / meaning of a polysensous word in context
- In English, the most frequently occurring nouns have 7 senses and the most frequently occurring verbs have 11 senses
- How can we define different word senses? (Two ways)
  - Give a list of synonyms
  - Give a definition
- Coarse-grained senses distinguish core aspects of meaning; Fine-grained senses also distinguish peripheral aspects of meaning
  - Bass: 5 senses related to music and 3 related to fish

Difficulties With Synonyms

- True synonyms non-existent, or very rare
- Near-synonyms (Edmonds and Hirst)
  - Examples:
    - Error, blunder, mistake
    - Order, command, bid, enjoin, direct
  - Dimensions of synonym differentiation
    - Stylistic variation (e.g., formal/informal)
      - Pissed, drunk, inebriated
    - Expressive variation
      - Attitude: skinny, thin, slim
      - Emotion: father, dad, daddy
    - . . .

How Humans Sense Disambiguation

- Sources of influence known from psycholinguistics research:
  - local context (e.g., *book* in a sentence that has *flight, travel, etc.)*
    - the sentence or other surrounding text containing the ambiguous word restricts the interpretation of the ambiguous word
  - domain knowledge (e.g., *plant* in a biology article)
    - the fact that a text is concerned with a particular domain activates only the sense appropriate to that domain
  - frequency data
    - the frequency of each sense in general usage affects its accessibility to the mind

Lesk Algorithm

- Original Lesk definition: measure overlap between sense definitions for all words in context. (Michael Lesk 1986)
  - Identify simultaneously the correct senses for all words in context
- Simplified Lesk (Kilgarriff & Rosensweig 2000): measure overlap between sense definitions of a word and current context
  - Identify the correct sense for one word at a time
  - Current context is the set of words in the surrounding sentence/paragraph/document.

Lesk Algorithm: A Simplified Version

- Algorithm for simplified Lesk:

  1. Retrieve from machine readable dictionary all sense definitions of the word to be disambiguated

  2. Determine the overlap between each sense definition and the current context

  3. Choose the sense that leads to highest overlap

Example: disambiguate PINE in

"Pine cones hanging in a tree"

- PINE

  1. kinds of evergreen tree with needle-shaped leaves

  2. waste away through sorrow or illness

Pine#1 ∩ Sentence = 1
Pine#2 ∩ Sentence = 0

Sense Tagged Corpus

- Examples of text where words are annotated with their sense from WordNet

Bonnie and Clyde are two really famous criminals, I think they were **bank/1** robbers
My **bank/1** charges too much for an overdraft.
I went to the **bank/1** to deposit my check and get a new ATM card.
The University of Minnesota has an East and a West **Bank/2** campus right on the Mississippi River.
My grandfather planted his pole in the **bank/2** and got a great big catfish!
The **bank/2** is pretty muddy, I can't walk there.

Classification Approach To WSD

- Train a classification algorithm that can label each (open-class) word with the correct sense, given the context of the word
- Training set is the hand-labeled corpus of senses
- Result of training is a model that is used by the classification algorithm to label words in the test set, and ultimately, in new text examples
  - In the SensEval conferences, a number of systems in range of 70-80% accuracy for English Lexical Sample task

Features

- The context is represented as a set of "features" of the word and includes information about the surrounding words
- Word Similarity Features:
  - For each word in the context, compute a similarity measure between that word and the words in the definitions to be disambiguated
  - Similarity measures can be defined from a semantic relation lexicon, such as WordNet (hypernym, hyponym)
    - One example is path similarity: for any two words, gives a number between 0 and 1 based on the shortest path between the two words in the WordNet hypernym/hyponym hierarchy

Features

- Collocational features: Information about words in specific positions (i.e. previous word), and typical features include the word itself, its stem and its POS tag
  - Example feature set: 2 words to the left and right of the target word and their POS tags

    *An electric guitar and **bass** player stand off to one side, not really part of the scene, just as a sort of nod to gringo expectations perhaps.*

    [guitar, NN, and, CC, player, NN, stand, VB]
- Syntactic features (relationship between the word and the other parts of the sentence)
  - Predicate-argument relations
    - Verb-object, subject-verb
  - Heads of Noun and Verb Phrases

Features

- Associated words features:
  - for each word to be disambiguated, collect a small number of frequently-used context words.
    - Example: for each word, collect the 12 most frequent words from a collection of sentences drawn from the corpus as the limited set.

      For bass, the 12 most frequent context words from the WSJ are: [fishing, big, sound, player, fly, rod, pound, double, runs, playing, guitar, band]
  - Represent these words as a set of words feature:
    - The features of bass in the previous sentence (represented as 1 or 0 indicating the presence or not of the word in a window of size 10 in the input text/context):
      [0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0]

Semantic Processing at Sentence Level and Beyond

Related Tasks For Semantic Processing:

- Detect non-syntactic ambiguities. If a sentence is two ways ambiguous, characterize the meaning of each reading.
  The bill is large.
  - Related Topics:
    - Word Sense Disambiguation: if a word has more than one sense, decide the sense of the word as it occurs in a sentence
      The bill is large but I have enough money to cover it.
    - Semantic similarity of words
- Decide if one sentence is a paraphrase of another (two way).
  Your marks on the tests were excellent.

  You scored very high on the exams.
- Entailment: decide if the truth of one sentence implies the truth of another (one way).
  John lives in Toronto.
  implies *John's residence is in Canada.*

Relation Between Syntax And Semantics In NLP

- Syntactic analysis:
  - determines the syntactic category of the words
  - decides phrase structure – how words are grouped
  - assigns structural analysis to a sentence
- Semantic analysis:
  - creates a representation of the meaning of a sentence
- Clearly syntactic structure affects meaning (e.g. word order, phrase attachment).
  - *"The man with the telescope watched Mary."*
  - *"Mary watched the man with the telescope."*
- But meaning can determine syntactic structure
  - Recall that lexicalized statistical parsing used head word affinities (probabilities) to help determine parsing.

Semantic Systems and Semantic Representations

- A semantic system consists of different types of building blocks: entities, concepts, relations, and predicates.
- A semantic representation shows how to put together blocks of a semantic system to describe a situation or "semantic world"
  - Enables reasoning about that semantic world

Why Do We Need Semantic Systems and Semantic Representations?

Use of Semantic Systems and Semantic Representations

- To link the surface, linguistic elements to the non-linguistic knowledge of the world
  - Many words, few concepts
- To represent the variety at the lexical level at a unified conceptual level
  - Unambiguous representations; canonical forms
- Structures composed from a set of symbols
  - All languages have a predicate-argument structure
  - Correspond to relationships that hold among concepts underlying constituent words and phrases of a sentence, and then across sentences
- Can be used to reason, both to verify what is true in the world and to infer knowledge from the semantic representation

Building Blocks of Semantic Systems

- Semantics that words (or base noun phrases) represent – the objects
  - Entities – individuals such as a particular person, location or product
    - John F. Kennedy, Washington, D.C., Cocoa Puffs
  - Concepts – the general category of individuals such as
    - person, city, breakfast cereal

Building Blocks Of Semantic Systems

- Semantics indicated by verbs, prepositional phrases and other structures
  - Relations between entities and concepts
    - John F. Kennedy "is-a" person
  - Relations between entities or between concepts
    - Hierarchy of specific to more general concepts
    - Wide variety of other relations (e.g., people are related to organizations, locations are related to people, etc)
  - Predicates representing verb structures, sometimes called events
    - Semantic roles, case grammar
    - Can also be used for relations between objects

Semantic Representation Approaches

- Some possible representation approaches:
  - First Order Logic
  - Semantic Nets
  - Conceptual Dependency
  - Frames
  - Rule-Based
  - Conceptual Graphs
  - Case Grammar

First Order Logic For Semantic Representation

- Also known as Predicate Calculus
- A symbolic language whose symbols have precisely stated meanings and uses
  - The symbols can be used as meanings in the real world
  - Typically express properties of entities in the world
  - Example – *if Socrates is a man, then Socrates is a mortal*
    Man (Socrates) -> Mortal (Socrates)
- First Order Logic (FOL) often used in AI systems found in such applications as robotics and computational control systems
  - Systems have automatic reasoning to make decisions or supply information
  - FOL allows a natural language interface to such systems

FOL Language

- FOL uses terms to represent objects in the real world
  - Constants are specific objects in the world - entities
    - Socrates, Pastabilities
  - Functions represent concepts about objects
    - LocationOf (Pastabilities)
      - Note the value of a function is a concept or entity
  - Variables are used to stand for any object
    - X
- FOL uses predicates to state relations between objects
  - Note the value of a predicate is True or False representing facts in the world
  - "IsRestaurant" could be a predicate that when applied to an object returns True if it is a restaurant
    - IsRestaurant (Pastabilities)
  - If "Serves" is a predicate taking a restaurant and a type of food as arguments, we can state that a restaurant serves a type of food
    - Serves (Pastabilities, VegetarianFood)

FOL Language, Operations And Quantifiers

- FOL uses connectives *and* and *or* to combine statements
  - Serves (Pastabilites, VegetarianFood) ^ IsExpensive(Pastabilities)
- FOL uses the implication connection to mean if the first statement is true, then the second one is also true
  - Serves(Pastabilities, VegetarianFood) => IsRestaurant(Pastabilities)
    - Is this true?
- FOL uses the existential quantifier to assert that an object with particular properties exists

  ∃ X IsRestaurant(X) ^ Serves(X, VegetarianFood)
- FOL uses the universal quantifier to assert that particular properties are true for all objects (using ∀ for the "forall" symbol)
  - ∀X IsRestaurant(X) => Serves(X, VegetarianFood)

  (this is definitely false because not all restaurants serve vegetarian food)

Example - Syracuse Restaurant Semantic World

- Objects:

Pastabilities, ElCanjelo, FunkNWaffles, *# restaurants in Syracuse*

VegetarianFood, MexicanFood, IndicanFood *# types of food*
SyracuseUniversity, TheWarehouse *# locations in Syracuse*

- Functions: LocationOf( ) # gives the location of the argument

- Predicates:

IsRestaurant( ) # true if the argument is a restaurant
Serves (,) # true if the 1st argument serves the type of food in the 2nd arg
Near (,) # true if the 1st arg location is near the 2nd arg location

Example - Syracuse Restaurant Semantic World

- FOL representation of example sentences:
  - More than one sentence could map to the same representation (paraphrase check)

The restaurant El Canjelo serves Mexican food.

El Canjelo is a restaurant specializing in Mexican food.

IsRestaurant(ElCanjelo) ^ Serves(ElCanjelo, MexicanFood)

  - Note: some nuances are lost ("*specializing*")

  - Some information (restaurant is new) may not be in the semantic world

Funk N Waffles is a new Indian restaurant near Syracuse University.

IsRestaurant(FunkNWaffles) ^ Serves(FunkNWaffles, IndianFood)
^ Near (LocationOf(FunkNWaffles), SyracuseUniversity)

Reasoning With FOL

- FOL allows inference to make conclusions of new information
  - Inference rule is called "modus ponens", informally is if-then reasoning
    if we know that A is true and we know that A => B is true,
    we can conclude that B is true

- This type of inference has efficient implementations to allow systems to reason from facts given in the semantic world or in text.
  - For example, reasoning could find answers for a question answering system
    Find me a restaurant serving Mexican food near the Warehouse
  - Find the X such that

IsRestaurant(X) ^ Serves(X, MexicanFood) ^ Near(LocationOf(X)), TheWarehouse)

Events In First Order Logic

- So far the predicates have captured state, properties that remain unchanged over some period of time
- Events denote changes in some state and can have a host of participants, props, times and locations.
- One way to give events in FOL is to state the existence of an event that has all the participants, etc.

I ate a turkey sandwich for lunch at my desk on Tuesday.

$\exists$ e Eating(e) ^ Eater(e, Speaker) ^ Eaten(e, TurkeySandwich)
^ Meal(e, Lunch) ^ LocationOf(e, Desk) ^ Time(e, Tuesday)

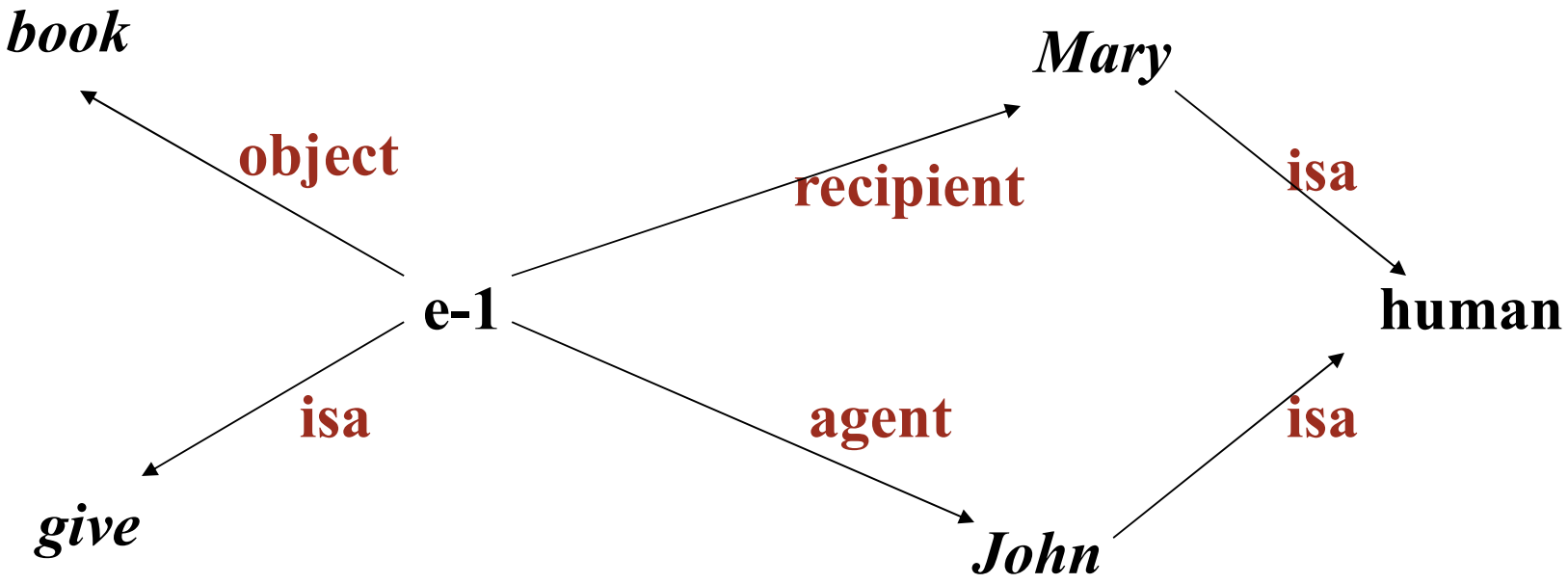
Difficulties With First Order Logic

- Problem for NLP:
  - 'semantics' of logic does not necessarily equate to 'meaning' in the real world
  - Not everything is as clear cut as required by a formal logic
  - May not be enough "real world" predicates in the FOL system to capture semantics of text
    - This is a problem for all the semantic representations
    - Semantic systems better developed for objects and actions
    - Not as well developed to represent ideas and beliefs
  - See Cyc Corp efforts to embody all world knowledge in (essentially) First Order Logic in their "Knowledge Base"
    - http://www.cyc.com/kb/

Semantic Networks

- A network or graph of nodes joined by links where:
  - nodes represent concepts (book, human) and entities (John, Mary)
  - links (labelled, directed arcs) represent relations (e.g. ISA)

John gives a book to Mary.



Frames

- A type of structured representation or *schema*
- Introduced by Marvin Minsky in 1975
  - "A Framework for Representing Knowledge"
  - Most widely referenced paper on knowledge representation
  - Explicitly attempts to represent human processing
- A way of grouping information about an entity or an event in terms of a record of 'slots' and 'fillers'
  - Each object has a frame with slots
  - One slot filled by the name of the object that the node stands for
  - Other slots filled with a property or relation and the value of the property or the entity that is related

Example Of Frames

- Wikipedia Info Box is an example of a frame structure
  - Slot names are properties or relations
  - An property value is information such as a date or height
  - A relation value is another entity, which may have its own frame

Norwegian Lundehund

A Norwegian Lundehund

Other names	Norsk Lundehund Norwegian Puffin Dog
Nicknames	Lundehund
Country of origin	Norway
Traits	[hide]
Weight	6–7 kilograms (13–15 lb)
Height	30–40 centimetres (12–16 in)
Classification / standards	[hide]
FCI	Group 5 Section 2 #265 standard
AKC	Non-Sporting standard
UKC	Northern Breed standard

Dog (*Canis lupus familiaris*)

Example Of Frames

- More formal frame systems (such as those for information extraction) require uniformity of slot names and value syntax

- Reasoning with Frames can use FOL:
  (∃ X) (Name(X) = Barack Obama) ^ Birthplace(X) = Honolulu) etc.

Name	Barack Obama
Birthdate	August 4, 1961
Birthplace	Honolu, Hawaii
Height	6’ 1” (1.85 m)
Parents	Barack Obama Sr., Ann Dunham
Children	Natasha Obama, Malia Ann Obama

Applications Of Semantic Representations

- Paraphase task: two sentences map to the same semantic representation (Microsoft Research Paraphrase Corpus)
- Entailment task: the semantics of the first sentence implies the semantics of the second under reasoning (EDITS: http://edits.fbk.eu/)
- Semantic representations are used to represent entities with their properties and relations in information extraction and question answering systems
- Semantic representations are used in AI systems such as robot manipulations through reasoning
  - Could also be used in dialog systems
  - Works best in a small environment where the amount of world knowledge needed is small

Getting Semantic Representation From Text

- Use a syntactic parse tree to identify predicates and possible relations structures
- Algorithms map syntactic structure to relations, given the words in the text
  - Semantic role labeling is one important algorithm (next week)
  - Some systems employ a First Order Logic mapper
  - Watson (IBM question answering system) mapped dependency parses to frames