Lexical Knowledge Networks

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay
(with contributions from Ramanand, Debasri, Manish, Akshay, Raghuvar, Mitesh, Aditya, Rajat and many others)

Introduction

- Lexico-semantic networks such as the Princeton WordNet are now considered vital resources
- Wordnets are being constructed in different languages as seen in the EuroWordNet project and the Hindi WordNet
- Competing lexical networks, such as ConceptNet, Hownet, MindNet, VerbNet, and FrameNet are also emerging as alternatives to wordnets
- Users are interested in knowing not only the relative merits from among a selection of choices, but also the intrinsic value of such resources

Outline

- What are *Lexical Knowledge Networks*?
- Brief look at different LKNs
- Elaborate discussion on wordnets
- A few universal combinatorial and graph theoretic properties of multilingual wordnets
- Towards evaluation of wordnets
- Conclusions and future work

Motivation (1/4)

- How do you disambiguate 'web' in "*the spider spun a web"* from "*go surf the web"*?
- How do you summarise a long paragraph?
- How do you automatically construct language phrasebooks for tourists?
- Can a search query such as *"a game played with bat and ball"* be answered as *"cricket"*?
- Can the emotional state of a person who blogs "*I didn't expect to win the prize!*" be determined?

Motivation 2/4

- Many of these issues can be (partially) resolved just by knowing more about the meaning of words - *lexical semantics theory*
- Need a lexicon that provides:
  - dictionary or thesaurus-like information
  - more rich associations among words
- *Key elements*:
  - collection of words
  - useful relations among them
  - ability to query the network

Motivation (3/4)

- **No major evaluations proposed or tried for lexical n/ws**
- Increasingly automated methods to grow networks (from web, wikipedia)
- Quality of such networks is unclear
- Other NLP fields have reasonably useful evaluation schemes – *BLEU* for MT

Motivation (4/4)

- LKNs embed a Conceptualisation of the world
- **Share universal properties across different languages**
- The properties are combinatorial and graph theoretic in nature and pertain to the *path length, degree, density etc.*
- They indicate the *level of maturity* of the LKN

Some Lexical Networks

WordNet

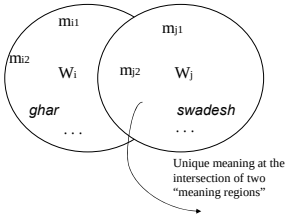
- Constituent elements are *senses*
- *Relational Semantics* as opposed to componential semantics
- *Principles*: Differentiation, Minimality, Coverage, Replaceability
- *Basic entities*: "Synsets", *i.e.,* Sets of Synonymous words
- *Relations*:
  - Lexical: Synonymy, Antonymy
  - Semantic: Hypernymy/Hyponymy, Meronymy/Holonymy and more

Componential semantics

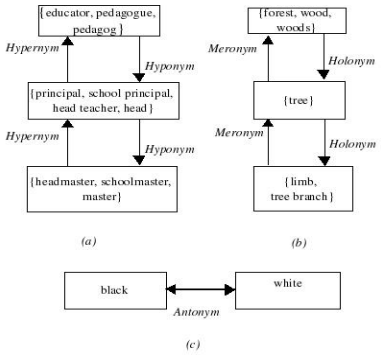
- Design a set of features, set/reset for each word, eg. +Human, +Animate, -Solid, -Moving etc.
- Create a feature matrix.
- Componential semantics depends on the feature space. Each sense has a unique feature vector.
- Disadvantages –
  - How to find the exhaustive feature space?
  - How to make features word-independent?
  - How to make it fine-grained enough?

Relational Symantics

- Relational – Disambiguate words by other words
- Put words denoting the same unique sense into a set called SYNSET



WordNet – illustration



Synsets and their Governing Principles

- **Synsets are sets of synonymous words: basic elements of WNs**
- **Minimality**: Only the minimal set that uniquely identifies the concept is used to create the sysnet, *e.g.*,{ghar, kamaraa} (*room*)
- **Coverage***:* The synset should contain all the words denoting a concept. The words are listed in order of (decreasing) frequency of their occurrence in the corpus {ghar, kamaraa, kaksh} (*room*)
- **Replaceability***:* The words forming the synset should be mutually replaceable in a specific context. Two synonyms may mutually replace each other in a context C, if the substitution of the one for the other in C does not alter the meaning of the sentence:
  - {svadesh, ghar} (*motherland*)– {apanaa desh} (*the country where one is born*)
  - amerikaa meN do saal bitaane ke baad shyaam svadesh/ ghar lauTaa
  - (America in two years stay after Shyam motherland returned)
  - (*'Shyam returned to his motherland after spending two years in America'*)

Elements of a synset

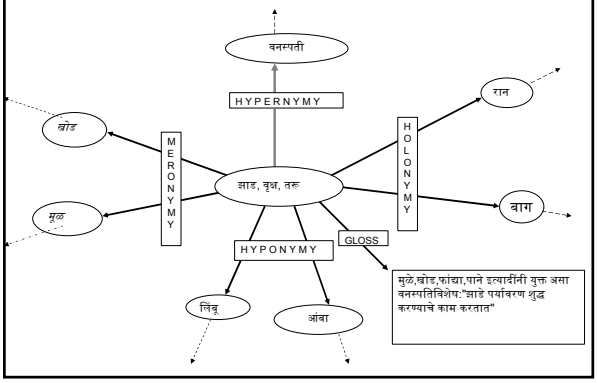
- A synset consists of the following elements.
  - **Synset:** {vidyaalay, paaThshaalaa, skuul} (*school*)
  - **Gloss** which consists of two parts.
    - **The text definition** that explains the concept denoted by the synset.
      - *vah sthaan jahaaM praathamik yaa maadhyamik star kii*
      - *aupachaarik shikshaa dii jaatii hai*
      - *'The place where formal education for primary or secondary level is given'*
    - **A sample sentence** that uses the word in a sentence
      - *is vidyaalay meM pahalii se paanchavii tak kii shikshaa dii jaatii hai*
      - *'Education from first to fifth class is given in this school'*

Lexico-semantic relations in wordnet

Semantic relations in WordNet:

- Synonymy
- Hypernymy / Hyponymy
- Antonymy
- Meronymy / Holonymy
- Gradation
- Entailment
- Troponymy

Semantic Relation subgraph (Noun)



Cross Part of Speech Linkages (important for word sense disambiguation)

- **Links between nouns and verbs:**
  - **Ability link** specifies the features inherited by a nominal concept
    - {machlii, macchii, matsya, miin, maahii} '*fish*'→ {tairnaa, pairnaa, paunrnaa} '*swim*'
  - **Capability link** specifies features acquired by a nominal concept
    - {vyakti, maanas}'*person*' → {tairnaa, pairnaa, paunrnaa} '*swim*'
  - **Function link** specifies function(s) associated with a nominal concept
    - {adhyaapak, shikshak}'*teacher*' → {paRhanaa, shikshaa denaa} '*teach*'
- **Links between nouns and adjectives**: indicate typical properties of a noun
  - {sher} '*tiger*' → {maansaahaarii} '*carnivorous*'.
- **Links between morphologically derived forms**
  - {bhaaratiiyataa} '*indianness*' is derived from {bhaaratiiya} '*Indian*' and is linked to it.

ConceptNet

- *'Common Sense'* semantic network
- Graph of simple concepts and rich relations about everyday knowledge
- 20 Relations, *e.g. "Causes", "Located At", "is-a", "is for"*
- Common Sense data collected from volunteers via the Web
- Data is processed to automatically yield the networks

ConceptNet – illustration

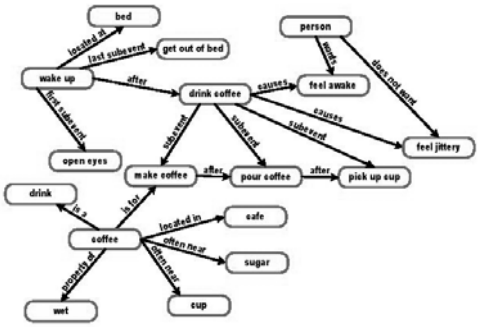


Fig: Singh, Liu: "Commonsense Reasoning in and over Natural Language")

Multilingual Wordnets for Indian Languages

Wordnet work at IIT Bombay

- http://www.cfilt.iitb.ac.in
- Follow the design principle(s) of the Princeton Wordnet for English paying particular attention to language specific phenomena (such as *complex predicates*)
- **Hindi Wordnet**
  - Total Number of Synsets: 23,067
  - Total Number of Unique Words: 48,725
- **Marathi Wordnet**
  - Total Number of Synsets: 11,908
  - Total Number of Unique Words: **18,093**

Status of other WNs

	Total Number of Synsets	**Total Unique Words**
WordNet (2.1)	117597	155327
GermaNet (2004)	53312	76563
Multi Word Net (1.39)	32,700	58,000

HWN and MWN Created Using Different Principles

(Tatsam, *i.e.*, Sanskrit words borrowed as

HWN entry:
{peR, vriksh, paadap, drum, taru, viTap, ruuksh, ruukh, adhrip, taruvar} 'tree'
jaR,tanaa, shaakhaa, tathaa pattiyo se yukt bahuvarshiya vanaspati
'perennial woody plant having root, stem, branches and leaves'
peR manushya ke lie bahut hi upayogii hai 'trees are useful to men'

MWN entry:
{jhaaR, vriksh, taruvar, drum, taruu, paadap} 'tree'
mule, khoR, phaanghaa, pane ityaadiinii yokt asaa vanaspativishesh
'perennial woody plant having root, stem, branches and leaves'
tii damuun jhaadacyaa saavlit baslii 'Being tired/exhausted she sat under the shadow of the tree'

Hindi WN: just recently **made free**

Towards Multilingual Indo-WN

- Through *Relation Borrowing* (illustrated through HWN and MWN)
- **When the meaning is found in both Hindi and Marathi**: This is the most common case, since Hindi and Marathi are sister languages
- **When the meaning is found in Hindi but not in Marathi**: Relation borrowing is not possible
  - For instance, {दादा [daadaa, grandfather], बाबा [baabaa, grandfather], आजा [aajaa, grandfather], दद्दा [daddaa, grandfather], पितामह [pitaamaha, grandfather], प्रपिता [prapitaa, grandfather]} are words in Hindi for paternal grandfather. There are no equivalents in Marathi.
- **When the meaning is not found in Hindi but is found in Marathi**: The relations must be set up manually
  - For example, {गुढीपाडवा [gudhipaadvaa, newyear], वर्षप्रतिपदा [varshpratipadaa, new year]} are words in Marathi which do not have any equivalents in Hindi.

Hindi Verb Knowledge Base (HVKB)

calanaa 'move'
(icl>act(agt>person))
***ve loga dhiire dhiire chal rahe hai.* 'They are moving slowly'.**
(gaman karnaa) 'to move'
Frame:NP1; NP1_NOM
[VINT, VOA, VOA-BACT]
→ ***caRhanaa 'climb'***
(icl>move{>act}(agt>person)
ve loga dhiire dhiire chaRha rahe hai. 'They are climbing slowly.'
***upar ki or jaanaa* 'to move upwards'**
Frame:NP1; NP1_NOM
[VINT, VOA, VOA-BACT]

Chakrabarti, Sarma and Bhattacharyya, Lexical Resources Engineering Journal (accepted)

Hierarchy for Compound Verbs

karnaa (do)
(icl> do(agt>person, obj>person)
tum kyaa kar rahe ho? (What are you doing?)
kisi kaarya ko karnaa (do something)
Frame:NP1 NP2
Case: NP1_ERG; NP2_NOM
[VTRANS, VOA, VOA-ACT]

→ maar Daalnaa (kill)
(icl> do(agt>person, obj>person)
DaakuoM ne aadmioM ko maar Daalaa (The dacoits killed the men)
kisii ko khatm kar denaa. (to kill somebody)
Frame: NP1 NP2
Case: NP1_ERG; NP2_NOM/ACC

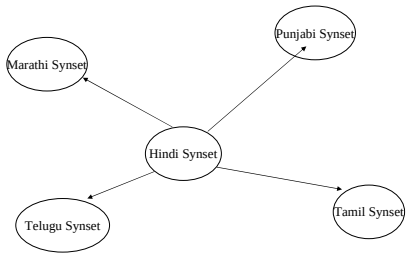
→ khaa jaanaa (eat)
(icl>do(agt>person,obj>thing))
wah saaraa khaanaa khaa gayaa. (He ate everything)
nindaniya Dhang se khaanaa(to eat disgracefully)
Frame: NP1 NP2
Case: NP1_NOM; NP2_NOM

Current statistics of HVKB

Total Verbs	Simple Verbs	Conjunct Verbs	Compound Verbs
1600	1100	200	300

Indian Language Dictionaries based on wordnets

- Every word in every language linked with each other through Hindi Synsets



Proposed Standard

Senses	Hindi	Marathi	Bangali	Oriya	Tamil
(W_1, W_2, W_3, W_4, W_5, W_6)	(W_1, W_2, W_3, W_4, W_5, W_6)	(W_1, W_2, W_3)	(W_1, W_2 , W_3)	(W_1, W_2, W_3, W_4)	(W_1, W_2, W_3)
(sun)	(सूर्य, सूरज, भानु, भास्कर, प्रभाकर, दिनकर, अंशुमान, अंशुमाली)	(सूर्य, भानू, दिवाकर, भास्कर, रवि, दिनेश, दिनमणी)	...	...	...
(cub, lad, laddie, sonny, sonny boy)	(लड़ का , बा लक, बच्चा , छो कड़ा , छो रा , लौ ड़ा)	(मु लगा , पो रगा , पो र, पो रगे)	...	...	...
(son, boy)	(पु त्र, बे टा , लड़ का , ला ल, सु त, बच्चा , नं दन, पू त, चिरं जी व, चिरं जी)	(मु लगा , पु त्र, ले क, चिरं जी व, तनय)	...	...	...

Wordnet Application: Word Sense Disambiguation

Different Approaches

- Knowledge-based
- Machine-learning-based
  - Supervised
  - Unsupervised

Knowledge-based WSD

- Based on argument structure and selectional preference
- *Equivalent concept in Indian Linguistics*
  - *Akanksha* (desire)
  - *Yogyata* (suitability)
  - *Sanniddhi (proximity)*

Examples

- She has to wash many dishes in the evening.
- She has to cook many dishes in the evening.
  - Agent of wash has to be animate.

Difficulties in knowledge-based approach

- Exhaustive enumeration in machine-readable form of:
  - Argument-structure of verbs
  - Selectional preferences of arguments
  - Description of properties of words such that meeting the selection preference criteria can be decided.

Resource Requirement

- Typically a hierarchical structure of words (Ontology) is called for:
- Entity
  - Animate
    - Human
    - Non-human
  - Inanimate

Probabilistic-formulation of the WSD problem

- Given a target word in a segment of text, we could like to obtain that sense which has the highest probability.
- Best sense, $\hat{s} = argmax_{s \varepsilon senses} Pr(s|w)$
- But, what is typically done is: *w* is represented by various information w_{info} about and around it.

w_{info} consists of:

- POS of *w*
- Features of w (semantic & syntactic)
- Collocation vector (set of words around it)
- Co-occurrence vector (number of times w occurs in bag of words around it)

Co-occurrence and Collocation vectors

- Co-occurrence vector:
  - Creates a vector of word frequencies in the environment of the target word
- Collocation vector:
  - Collocation vector typically consists of next word(+1), next-to-next word(+2), -2, -1 their POS's

Apply Bayes Rule

- $\hat{s} = \text{argmax}_{s \varepsilon senses} Pr(s|V_w)$; where V_w is the

  feature vector.
- Apply Bayes rule:
  - $Pr(s|V_w)=Pr(s).Pr(V_w|s)/Pr(V_w)$

Independence of Features

- $Pr(V_w|s)$ can be approximated by indepence assumption:

  $Pr(V_w|s)=Pr(V_w^1|s).Pr(V_w^2|s,V_w^1)...Pr(V_w^n|s,V_w^1,...,V_w^{n-1})$

  $=\Pi_{i=1}^n Pr(V_w^i|s)$
- $\hat{s} = \text{argmax}_{s \hat{\varepsilon} senses} Pr(s).\Pi_{i=1}^n Pr(V_w^i|s)$

Parameters

- Parameters in the probabilistic WSD are:
  - Pr(s)
  - $Pr(V_w^i|s)$
- Senses are marked with respect to sense repository=WORDNET
- Pr(s)= #c(s,w) / #c(w)

Probabilistic-formulation of WSD problem (contd..)

$Pr(V_w^i|s)=Pr(V_w^i,s)/Pr(s)$

$=(c(V_w^i,s,w)/c(w))/(c(s,w)/c(w))$

$=c(V_w^i,s,w)/c(s,w)$

- Make use of sense repositories like:
  - lookup SEMCOR
  - Lookup SENSEVAL

MRD-based approach (LESK Algorithm)

- MRD=machine-readable dictionary (e.g. WORDNET)
- "She prepares nice dishes which are delicious."
- {She, prepares, nice, delicious} -> **context bag**
- Words from the glosses and example sentences for each sense-> **Sense Bag** for that sense
- **Winner Sense:** the one with the maximum overlap with the context bag

The WSD Algorithm…

- **Let 'w' be the word whose disambiguation is to be done**
- **Construct the *context Bag***
- **Construct the *semantic Bag***
- **Filter bags through stemmer**
- **Using the *'Intersection Similarity',* find the *Overlap***
- **Output the sense 's' as the most probable sense which has the *maximum Overlap***

Dish in WN

- **Wordnet information about *dish***
  - **The noun dish has 6 senses (first 2 from tagged texts)**
  - **1. (9) dish -- (a piece of dishware normally used as a container for holding or serving food; "we gave them a set of dishes for a wedding present")**
  - **2. (3) dish -- (a particular item of prepared food; "she prepared a special dish for dinner")**
  - **3. dish, dishful -- (the quantity that a dish will hold; "they served me a dish of rice")**
  - **4. smasher, stunner, knockout, beauty, ravisher, sweetheart, peach, lulu, looker, mantrap, dish -- (a very attractive or seductive looking woman)**
  - **5. dish, dish aerial, dish antenna, saucer -- (directional antenna consisting of a parabolic reflector for microwave or radio frequency radiation)**
  - **6. cup of tea, bag, dish -- (an activity that you like or at which you are superior; "chemistry is not my cup of tea"; "his bag now is learning to play golf"; "marriage was scarcely his dish")**

Dish: Sense Bags

1. **{piece, dishware, normally, use, container, hold, serve, food, give, set, wedding, present}**
2. **{particular, item, prepare, food, prepare, special, dish, dinner}**
3. **{dishful, quantity, hold, serve, rice}**
4. **{smasher, stunner, knockout, beauty, ravisher, sweetheart, peach, lulu, looker, mantrap, dish, very, attractive, seductive, look, woman)**
5. **{dish aerial, dish antenna, saucer, directional, antenna, consist, parabolic, reflector, microwave, radio, frequency, radiation}**
6. **{cup of tea, bag, activity, like, superior, chemistry, cup, tea, bag, now, learn, play, golf, marriage, scarcely}**

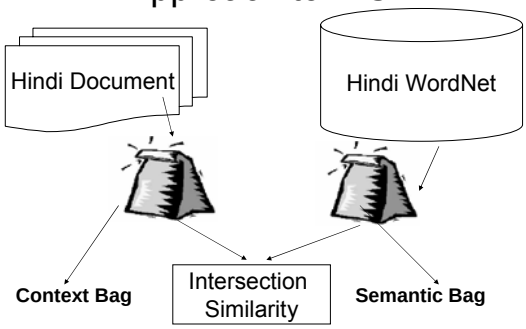
Compare Context and Sense Bags

- Only sense-2 has an overlap: *prepare*
- Winner sense is sense-2 and it is correct
- The Lesk algorithm suffers from *sparse match*: the possibility of word overlap is very slight.
- It can also be misled: *there were many delicious dishes during his marriage:* The last sense will be latched onto

Hindi Word Sense Disambiguation

Manish Sinha, Mahesh Reddy and Pushpak Bhattacharyya, *An Approach towards Construction and Application of Multilingual Indo-WordNet*, 3rd Global Wordnet Conference (GWC 06), Jeju Island, Korea, January, 2006.

Approach to WSD



The WSD Algorithm

Sense Bag

- *synonymy*, *hypernymy, hyponymy, meronymy* relations, their *Glosses* and *Example sentences* for semantic Bag

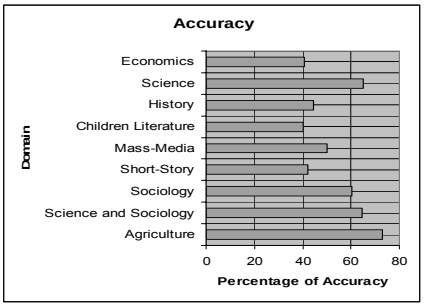
Context Bag

- *Current* , *previous* and *following* sentences in which word forms for context Bag

Experiment

- ***Nouns* only**
- **The test corpora has been taken from *CIIL, Mysore, India***
- **The system has been tested on corpus from 9 domains and each corpus containing around 2000 words on an average**

Results



Studies on wordnets: *small world properties*

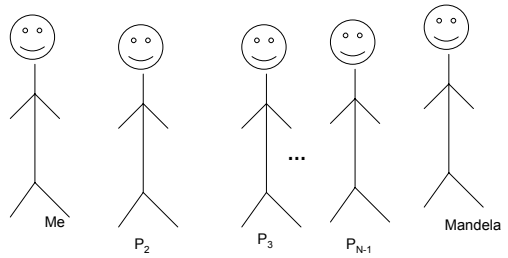
Ramanand, Ukey, Singh and Bhattacharyya, Mapping and Structural Analysis of Multilingual Wordnets, IEEE Data Engineering Bulletin, 30(1), March 2007

Human social graph

- How many links connect you to Adam Gilchrist/Nelson Mandela/Tim Berners-Lee?
- Sociological studies show that the diameter of human social graph is less than 10 (very, very small!)
- The concept of “Six Degrees of Separation”

Distance to Nelson Mandela

N<=10



Graphs and Measures

- Measures
  - Average Shortest Path Length
  - Clustering Coefficient
  - Degree Distribution
- Random Graphs: low Avg. Shortest Path
- Regular Graphs: high Avg. Shortest Path
- Small World Graphs: low Avg. Shortest Path

Cluster Coefficient

- Measures what fraction of neighbours of a node are related to each other
- *Cluster Coefficient C_i for a node i (with degree k_i)* of a directed graph:

$$C_i = \frac{|E(\Gamma_i)|}{2 \times \binom{k_i}{2}}$$

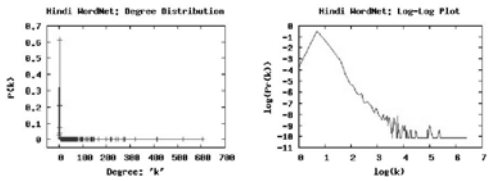
where Γ_i is the subgraph made of i and its neighbours, $|E(\Gamma_i)|$ is the number of edges of the subgraph, and $2 \times \binom{k_i}{2}$ is the total number of possible edges in Γ_i.

- Random Graphs: low CC (CC << 1)
- Regular Graphs: high CC (0.4 <= CC <= 0.7)
- Small World Graphs: high CC (0.4 <= CC <= 0.7)

But why study Small Worlds for NLP?

- Seen in language organization
- Seen in Wordnets!!

Degree Distribution in Wordnets



- Exponents observed:
  - English WordNet (Nouns): -2.063
  - Hindi WordNet: -2.592
  - Marathi WordNet: -2.841

High-degree nodes in Wordnets

- Eng. WordNet (Nouns):
  - (city,metropolis,urban center): 664, (law,jurisprudence): 611
  - (person,individual,someone,somebody,mortal,soul): 400
- Hindi WordNet:
  - (vyaktii, maanas, shaks, shakhs, ba.ndaa (person)): 607
  - (karm, karanii, kaam, kaarya, krtya, kaarvaaii, kaarvaahii (action)): 524
- Marathi WordNet:
  - (vyaktii, maaNus, isama, manushya, paTThaa, paThyaa (person)): 626
  - (karm, krtii, kriyaa, kaam kaarya, krtya (action)): 546

Cluster Coefficient in Wordnets

- Wordnet Avg. Cluster Coefficient:
  - English WN (Nouns): 0.526
  - Hindi WN: 0.268
  - Marathi WN: 0.358

Average Shortest Path Length in Wordnets

- Average Shortest Path Lengths observed:
  - English WordNet (Nouns): 8.878
  - Hindi: 4.378
  - Marathi: 4.255

Wordnet Evaluation

J. Ramanand and Pushpak Bhattacharyya, *Towards Automatic Evaluation of Wordnet Synsets,* Global Wordnet Conference, Szged, Hungary, Jan 2008

Need for Wordnet Evaluation

- Increasing emphasis on *evaluation* in all branches of NLP (e.g., *BLEU score in MT)*
- Varied applications of wordnets call for the creation of methods to evaluate their quality
- Validate synsets
- Of specific interest are synsets in which some members “do not belong”
- Our work, thus, is an attempt to flag human lexicographers’ errors by accumulating evidences from myriad lexical sources

Questions on Evaluation

- How to select one lexico-semantic network over another?
- Is a given wordnet sound and complete?
- Is this resource usable, scalable, and deployable?
- Is this wordnet suitable for a particular domain or application?

Key Questions

- What is the definition of a synonym?
- What are the necessary and sufficient conditions to determine that synonymy exists among a group of words?

Foundational observation

- if two words are synonyms, it is necessary that they must share one common meaning out of all the meanings they could possess.
- A sufficient condition could be showing that the words replace each other in a context without loss of meaning.

Key Points for Validation

- Are the words in a synset indeed synonyms of each other?
- Are there any words which have been omitted from the synset?
- Does the combination of words indicate the required sense?

Input and Output of the Validator

- Input
  - The *supposedly* synonymous words in the synset
  - The hypernym(s) of the synset
  - Other linked nodes, gloss, example usages
- Ouput
  - A verdict

Related Work on Automatic Synset Creation

- All these methods are based on web and corpora mining
- Existing Work on
  - Collection of synonyms in the medical domain from the Web by first building a taxonomy of words
  - Unsupervised learning method for extracting synonyms from the Web
  - Topic signature method to detect synonyms using document contexts and thus enrich large ontologies

Basic Idea of Our Approach

- ***If a word is present in a synset, there is a dictionary definition for it which refers to its hypernym or to its synonyms from the synset***

- Example*:*

(obtained from the website Dictionary.com):

snake: any of numerous limbless, scaly, elongate reptiles of the suborder Serpentes, comprising venomous and non-venomous species inhabiting tropical and temperate areas.

serpent: a snake

ophidian: A member of the suborder Ophidia or Serpentes; a snake.

This critical observation suggests that dictionary definitions may provide useful clues for verifying synonymy.

Three Groups of Rules

- The dictionary-based algorithm consists in applying three groups of rules in order
- The first group applies to each word individually, using its dictionary definitions
- The second group relies on a set of words collected for the entire synset during the application of the first group
- The final group consists of rules that do not use the dictionary definitions

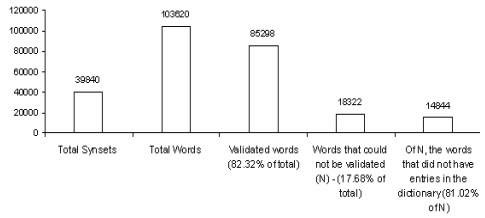
A Glimpse of the Rule Base

- *Group 1*
- *Rule 1 - Hypernyms in Definitions*
- Definitions of words for particular senses often make references to the hypernym of the concept. Finding such a definition means that the word's placement in the synset can be defended.
- e.g.
- Synset: *{brass, brass instrument}*
- Hypernym: *{wind instrument, wind}*
- Relevant Definitions:
- *brass instrument*: a musical *wind instrument* of brass or other metal with a cup-shaped mouthpiece, as the trombone, tuba, French horn, trumpet, or cornet.

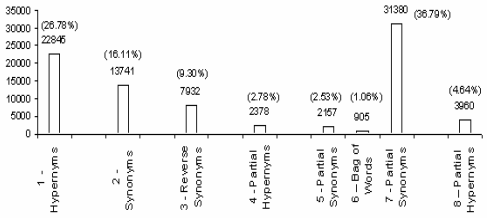
Experimentation

- The validation was tested on the Princeton WordNet (v2.1) noun synsets
- Out of the 81426 noun synsets, 39840 are synsets with more than one word – only these were given as input to the validator
- This set comprised of a total of 103620 words.

Results of Validation (1/2)



Rulewise Summary



An example case study

- Synset: *{visionary, illusionist, seer}*
- Hypernym: *{intellectual, intellect}*
- Gloss: *a person with unusual powers of foresight*
- The word *"illusionist"* was not matched in this context. This seems to be a highly unusual sense of this word (more commonly seen in the sense of *"conjuror"*). None of the dictionaries consulted provided this meaning for the word.

Summary

- Described
  - the motivation for studying lexical knowledge networks in particular wordnets
  - the work on Indian Language Wordnets
  - Word sense disambiguation and an instance of wsd for Hindi
  - common dictionary for all languages
  - invariances in wordnets: small world properties
  - Automatic validation of wordnets

Conclusions

- Multilinguality emerging as a norm rather than a fashion
  - Methods needed to tackle the challenges
- Invariances in multilingual computation and resources form an interesting study
- Wordnets are getting constructed by
  - Extension methods (based on another wordnet)
  - Web methods
  - Corpora methods
- They need to be validated: *flag possible lexicographer error*

URLs

- For resources

  www.cfilt.iitb.ac.in
- For publications

  www.cse.iitb.ac.in/~pb

Basic References

- 1. Miller, G., Beckwith, R., Felbaum C., Gross D., Miller, K.J.: Introduction to Wordnet: an on-line lexical database. The International Journal of Lexicography, vol. 3 (4), pp. 235-244 (1990)
- 2. Vossen, P.: EuroWordNet: a multilingual database with lexical semantic networks. Kluwer Academic Publishers (1998)
- 3. Narayan, D., Chakrabarty, D., Pande, P., and Bhattacharyya, P.: An Experience in building the Indo-Wordnet - A Wordnet for Hindi. In: First International Conference on Global Wordnet (GWC '02) (2002)
- 4. Liu, H., Singh, P.: Commonsense Reasoning in and over Natural Language. In: The proceedings of the 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems (2004)
- 5. Dong, Z., Dong, Q.:An Introduction to HowNet, available from http://www.keenage.com.
- 6. Richardson, S., Dolan, W., Vanderwende, L.: MindNet: acquiring and structuring semantic information from text. In: 36th Annual meeting of the Association for Computational Linguistics, vol. 2, pp. 1098-1102 (1998)
- 7. Kipper-Schuler, K.: VerbNet: A broad-coverage, comprehensive verb lexicon. Ph.D. dissertation, University of Pennsylvania (2005)
- 8. Baker, C. F. , Fillmore, C. J. , Lowe, J. B. : The Berkeley FrameNet project. In: Proceedings of the COLING-ACL (1998)