IndoWordNet and its Linking with Ontology

Brijesh Bhatt
Center For Indian Language Technology,
Department of Computer Science and Engineering,
Indian Institute of Technology, Mumbai
brijesh@cse.iitb.ac.in

Pushpak Bhattacharyya
Center For Indian Language Technology,
Department of Computer Science and Engineering,
Indian Institute of Technology, Mumbai
pb@cse.iitb.ac.in

Abstract

Reasoning about natural language requires combining semantically rich lexical resources with world knowledge, provided by ontologies. In this paper, we describe linking of WordNets of Indian languages with an upper ontology SUMO (Suggested Upper Merged Ontology). This creates multilingual resource for Indian languages which can be used in various natural language processing applications. This paper presents the architecture of IndoWordNet- Linking of WordNets of seventeen different Indian languages and provides a method to link it with upper ontology SUMO. Two different systems: IndoWordNet navigator and SIGMAKEE interface for Indian languages are developed to access this resource.

1 Introduction

WordNet (Fellbaum, 1998) has emerged as a great resource for the Natural Language Processing applications. Following English WordNet, WordNets are built for many languages of the world. Hindi WordNet (D. Narayan and Bhattacharyya, 2002) is the first WordNet built for an Indian language. WordNets for other 16 Indian languages are being built from Hindi WordNet using expansion approach (Vossen, 1998). Linking of all these WordNets provides rich knowledge base for the Indian languages, which can be useful for information extraction, retrieval and many other natural language processing applications.

1.1 WordNet and ontology

Ontology is defined as "Explicit specification of conceptualization"(Gruber, 1993). Ontology is a hierarchical structure of concepts related by subsumption relation, which can be shared between applications. Though WordNet is often considered as an ontology (the synset corresponds to concept and 'hypernymy-hyponymy' relation is similar to subsumption), it is a language specific resource which may vary from language to language. The ontological issues in WordNet discussed in (pease and Fellbaum, 2010; Gangemi et al., 2003; Niles and Pease, 2003) are as follows,

1. *Confusion between concept and individual*: WordNet synsets do not distinguish between concept and individual. For example, both *'Martial Art'* and *'Karate'* are considered as concept.

2. *Confusion between object level and meta level concept*: WordNet covers both object level and meta level concepts as hyponymy of same concept. For example, concept *'Abstraction'* includes both object level concept *'Time'* and meta level concept *'Attribute'*.

3. *Heterogeneous level of generality*: Two hyponyms of a concept may represent different level of generality. For example, as a hyponymy of concept *'Animal'*, there is a general concept *'chordate'* and a more specific concept *'Work Animal'*.

4. *Lexical gap*: A language may not have an indigenous lexeme to describe a concept. For example, vehicles can be divided into two classes, 1) Vehicles that run on the road and 2) Vehicles that run on the rail, but English language does not have specific words to describe these classes.

"Proceedings of ICON-2011: 9th International Conference on Natural Language Processing, Macmillan Publishers, India. Also accessible from http://ltrc.iiit.ac.in/proceedings/ICON-2011"

1.2 Benefits of linking WordNet with ontology

By keeping ontological relations in the formal ontology and linguistic relations in the lexicon, one can avoid merging two different levels of analysis and can capture the information needed about formal concepts and linguistic tokens (pease and Fellbaum, 2010). Linking of WordNet with ontology allows the language independent semantic relations of ontology to be used for inferencing on language specific words. The benefits of linking ontology and WordNet are as follows (Niles and Pease, 2003):

1. The formal specifications of the ontology can be used with the WordNet in the sense that the axioms corresponding to the words can be retrieved from the ontology.

2. The formal axioms of ontology can be used with natural language text.

3. Linking ontology concepts with WordNet can be used to check completeness of ontology.

4. Concepts of different languages can be compared and linked using ontology.

5. WordNet concepts can be refined and restructured using ontology.

6. Different domain ontologies can be linked using WordNet concepts.

1.3 Organization of the paper

The remaining of the paper is organized as follows: Section 2 describes the IndoWordNet. The selection criteria for upper ontology and a comparative study of upper ontology is given in Section 3 and 4. Section 5 describes the method to link IndoWordNet with ontology and interfaces designed to access the system. Observations and conclusions are discussed in section 6 and 7 respectively.

2 IndoWordNet

Seeing the enormous potential of WordNet, 17 out of 22 official languages of India have started developing WordNets. These languages are: (1) Hindi (2) Marathi (3) Konkani (4) Sanskrit (5) Nepali (6) Kashimiri (7) Assamese (8) Tamil (9) Malyalam (10) Telugu (11) Kannad (12) Manipuri (13) Bodo (14) Bangla (15) Punjabi (16) Gujarati and (17) Urdu. Together these languages represents three language families: Indo-Aryan, Dravidian and Tibeto-Barman. The comparative study of IndoWordNet with EuroWorNet is presented in (Bhattacharyya, 2010)

2.1 Synset Categorization

Synsets of the Hindi WordNet are used as basis to create synsets of WordNets of other languages. As a synset is represented by a gloss and a set of words of a particular language, in many cases, the synset representation of a concept may vary in sense across the languages. Also there exist some concepts for which there may not be words in all the languages. For example, Kashmiri language does not have words for the concepts like *ग्रह ('graaha', Planet), सोम ('som'), मंगल ('mangal')*. Kinship terms also vary across Indian languages. To handle this concept divergence across languages, synsets are divided into different categories. Table 1 describes these categories.

Such classification of synsets helps in linking concepts of different languages. For example, if a synset belongs to the universal synset then it is present in both Hindi and English. And if a synset belongs to the Pan-Indian category then it belongs to both Hindi and Gujarati. Thus, WordNet development using expansion approach will be faster by this method. This classification also helps in cross lingual information access. By identifying the category of a synset, its presence in another language can be easily predicted.

Till date, 7163 universal synsets and 1356 Pan-Indian synsets have been identified and are now linked across all languages. Language specific synsets are being developed and later they will be linked by translating them into Hindi and English.

3 Survey of upper ontology

Ontologies are categorized into three different types according to their level of generality (Guarino, 1998). Top level/Upper ontology, Domain specific ontology and application specific ontology. Upper ontology defines very general concepts independent of application or domain. Upper ontologies are useful in linking and development of more specific domain/application ontology. There are various upper ontologies like SUMO, DOLCE, CYC etc. The ontological choices for designing upper ontologies, discussed in (Oberle et al., 2007), are as follows,

Table 1: Synset categories in IndoWordNet

Category	Description
Universal	These concepts appear in all the languages. These concepts are essential and most frequently used. For example, *सूर्य ('soorya', sun)*
Pan-Indian	Concepts common to Indian languages and linkable across all Indian languages but does not have parallel concepts in English. For example, *तबला ('tabalaa', An Indian rhythm instrument)*
In-Family	These concepts are common in specific subset of Indian languages and linkable across all languages of the family. For example, *चाचा ('chacha' ,paternal uncle), भतिजा ('bhatija', brother's son)*
Language Specific	These are the concepts specific to a culture or a language. These may include local food, festivals, etc. For example, *बिहु ('bihu' , Name of a festival celebrated in Assam state of India)* word is very specific to the state and culture and does not appear in any other language.
Rare	This includes very specific concepts adopted in most of the languages. Specific scientific terms like *'ngram'* belongs to this category
Synthesized	Synset created in a language due to influence of another language. These synsets are not natural to the language but required to link WordNets of two different languages.

Table 2: Comparison of Upper ontologies

Case	SUMO	DOLCE	Open CYC
License	Open	Open	Free subset of Cyc
Modularity	Yes	No	Yes
Language	KIF, OWL	OWL	Cyc
Multiplicative(M)-Reductionist(R)	-	M	M
Descriptive(D)-prescriptive(Pr)	D	D	D
Endurant(E)-Perdurant(Pd)	No	Yes	Yes
Universal(U)-Particular(Pt)	Both	Pt	Both
Linking to WordNet	Full	Partial	Partial

1. *Endurant-Perdurant* : An *endurant* is an entity which exists in full in every instant at which it exists at all. A *perdurant* "unfolds itself over time in successive temporal parts or phases." Both endurants and perdurants are taken to be concrete particulars, *i.e.*, instances.

2. *Descriptive-prescriptive* : A *descriptive* ontology tries to capture more commonsense and social notions based on natural language usage and human cognition. Concepts are divided into things and events. A *prescriptive* ontology emphasizes upon the scientific and philosophical perspectives. All the constructs in revisionary ontology are space-time objects.

3. *multiplicative-reductionist* : In *multiplicative* upper ontology concepts can include anything that reality seems to require. Contrarily, *reductionist* ontology reduces the number of concepts to the fewest primitives sufficient to derive the rest of the complex reality.

4. *Universal-particular* : *Universals* are the entities that have instances, while *particulars* are entities that do not have instances.

A brief comparison of ontologies based on these choices and other parameters such as availability of system, support for domain specific ontology and linking with WordNet etc.is shown in table 2

Table 3: IndoWordNet Index

iwnindex(IWNID, POS, WNLINKTYPE, ENGID,SUMOTERM, SUMOLINKTYPE)
IWNID : IndoWordNet synset ID
POS : Part of speech of the synset
WNLINKTYPE : type of link through which Hindi and English synsets are connected. It can be either direct or Hyponymy
ENGID : English sense Id of the corresponding synset
SUMOTERM : SUMO term associated to concept
SUMOLINKTYPE : Type of link through which synset is connected to SUMO term. It can be any of the direct, subsumption or instance.

4 SUMO-Suggested Upper Merged Ontology

SUMO (Niles and Pease, 2001) is created at Teknowledge Corporation, by merging publicly available ontological content into a single, comprehensive, and cohesive structure. The knowledge representation language for the SUMO is SUO-KIF. SUMO is the largest freely available ontology which is linked to entire English WordNet. A mid-level ontology MILO (Niles and Terry, 2004) and many domain ontologies for the variety of domains like Communication, Countries and region, Distributed Computing, Finance, Military, Geography, Government, etc. are constructed using SUMO. Including all these SUMO provides around 23000 terms and 123000 axioms. SUMO is supported by resolution theorem prover: Vampire (Riazanov and Voronkov, 2002). SUMO is also one of the backbone ontologies for the Global WordNet Grid (Pease et al., 2008). Open source project SigmaKEE(Sigma Knowledge Engineering Environment) (Pease and Benzmüller, 2010) which provides environment for first order logic theory development is optimized for SUMO. All these factors make SUMO an obvious choice as backbone ontology for the IndoWordNet.

5 Implementation

WordNets of Indian languages are linked to SUMO by using English WordNet. A common index is designed to link all WordNets with SUMO. The common index- *iwnIndex* (shown in table 3) is defined as five tuple, which unambiguously classifies the concept.

The overall system architecture is as shown in the Fig 1.

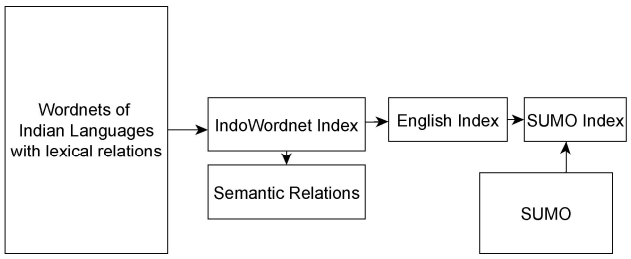

Figure 1: IndoWordNet architecture

As discussed in section 2, synsets are divided into six categories: Universal, Pan-Indian, In-Family, Language specific, Rare and Synthetic. Synsets of these categories are linked through IWNID. SynsetId of the synset represents this IWNID and gives unique identification to the synset. Universal, Pan-Indian, Rare and subset of In-family synsets for each language are developed from corresponding Hindi synsets, as these synsets are also present in Hindi WordNet. Therefore, IDs assigned to these synsets are synsetIDs of Hindi. For some of the in-family synsets and language specific synsets which are not present in Hindi, separate ID ranges are decided for each language group and these synsets are translated into Hindi and English language (Though there may not be a word specific to the synset).

5.1 Linking IndoWordNet and SUMO

The approach to link synsets and the method of using relations of Hindi WordNet to other Indian language WordNet are discussed in (Ramanand et al., 2007), (Sinha et al., 2006). The semantic relations defined in Hindi WordNet are inherited in all other languages. Lexical relations remain with the language specific WordNet, as they vary across languages.

Once synsets of all the Indian languages are linked through *IWNID*, linking of English WordNet and SUMO with the IndoWordNet is done by the process shown in table 4. This process takes a synset as an input and creates *iwnindex* by finding corresponding English synset ID and SUMO term.

readIwnIndex : This process returns the IWN ID for the given input word.

FindEnglishSynset : A semi automated approach (Saraswati et al., 2010) is used to link IWNID to EnglishID. It searches most appropriate English synset for the input Hindi Synset. If there

Table 4: SUMO linking process

```
readIwnIndex;
    findEnglishSynset ;
        findSumoConcept;
            findSumoMapType;
                linkSumo;
```

are multiple English synsets found for the given input Hindi Synset, the best synset is selected by manual verification.

Two types of semantic links are used to link Hindi synset with an English synset. 1. *Direct linkage*: There exists an identical concept in English WordNet for the given Hindi concept. 2. *Hypernymy linkage*: If there is no identical concept for the given Hindi concept, then Hindi concept is linked to the English concept which is in turn linked with nearest hypernymy of the Hindi concept.

FindSUMOConcept: The complete English WordNet is linked with SUMO. (Niles and Pease, 2003). Once the IWNID is mapped to English ID, the SUMO term is linked to form index.

linkSUMO: This process identifies the semantic type to link Hindi synset with SUMO term. There are three types of links defined, (Niles and Pease, 2003), to link a synset with the SUMO term.

1. *Direct mapping*: The Hindi synset is linked to the English synset through direct link and the English synset is linked to the SUMO term through equivalent link.

   For example, Hindi concept

   'आशा' , 'मन का यह भाव कि अमुक कार्य हो जाएगा'

   (*'aasha', mana ka vaha bhava ki aamuka karya ho jaayenga*)

   (*hope, mind's that feeling some work fulfills*)

   (*hope, feeling of mind that some work fulfills*)

   has equivalent synset in English *'Hope', 'General feeling that some desire will be fulfilled'*

   SUMO also has an equivalent concept *'Hope'*.

2. *Subsumption mapping*: The Hindi synset is linked through hypernymy to the English synset or English synset is linked to the SUMO term through subsumption link.

   For example, Hindi concept

   'शुभ', 'वह जो अच्छा हो'

   (*'shubha', vaha jo achchhaa ho'*)

   (*auspicious, something that good is*)

   (*auspicious, something that is good*)

   has direct equivalent relation with English concept *'Auspicious', 'auguring favorable circumstances and good luck'*

   As this is a subjective term, it does not have any direct concept mapped to SUMO, but SUMO defines more general concept *'SubjectiveAssessmentAttribute'* to link such terms with ontology.

3. *Instance mapping*: The Hindi synset is linked to the English synset either through direct or hypernymy link and the English synset is linked to the SUMO term through instance relation.

   For example, Hindi concept

   'मिथुन राशि', 'बाराह राशियो मे तीसरी राशि'

   (*'mithuna raashi' , 'baaraaha raashiyon mein tIsarI raashi'*)

   (*gemini , twelve zodiacs in third zodiac*)

   (*gemini , third of the twelve zodiacs*)

   has direct equivalent relation with English concept *'Gemini','The third sign of the zodiac'*

   It is mapped with the SUMO term *'AstronomicalBody'* using instance relation.

4. *Complement mapping*: WordNet defines concepts which are opposite to each other. Many times SUMO does not have concept for both the senses. In that case the synset is linked with SUMO by complement relation. For example, synset corresponding to the Hindi word '*घोषित*' ('ghoshit') is mapped to SUMO term *'stating'* using subsumption mapping and its opposite '*अघोषित*' ('aghoshit') is linked with same SUMO term *'stating'* using complemented subsumption mapping.

The following symbols are used to identify link type: Direct('='), Subsumption('+'), Instance('@'). Complements of these relations are identified by symbols '[', ']', ':', respectively.

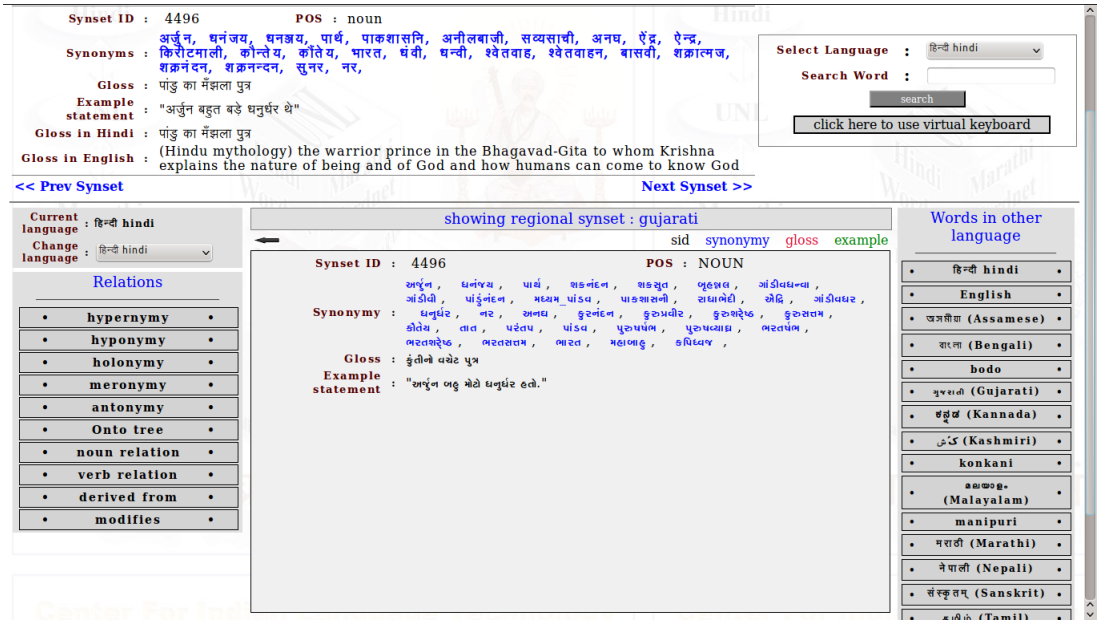

Figure 2: IndoWordNet browser

Table 5: iwnindex file: output of linking process

100, adj, Direct, 00935500, hasSkill, +
10980, adv, Direct, 00051848, ShapeAttribute, +
10035, noun, Direct, 06123363, FieldOfStudy, @
2534, verb, Direct, 02706046, Selecting, [

Some entries of the iwnindex file are as shown in table 5. Each row uniquely defines concept using- Indowordnet ID, Part-of-speech, Hindi-English link type, English sense Id, SUMO term and type of SUMO link.

5.2 Interface development

Two interfaces are developed to access the IndoWordNet system, IndoWordNet browser and SUMO browser. IndoWordNet browser is a web based interface to navigate WordNets of Indian languages. It generates xml file output which can be useful for different natural language processing applications. All the IndoWordNet synsets and relations are stored in MySQL database and the browser is designed using HTML, JSP, AJAX and MySQL technology. Fig. 2 shows snapshot of the interface.

The basic features of the system are as follows:

1. It provides easy access to all the WordNets of Indian languages
2. The interface can be browsed in 17 different Indian languages
3. It provides a comparative view of the synsets of different languages
4. Relations between synsets can be explored in all the languages

SIGMA knowledge Engineering Environment is available as an open source project. It is used to access SUMO ontology and its mapping to English WordNet.

An interface is created in SIGMA Knowledge Engineering Environment to access SUMO terms corresponding to the Hindi synset. Fig 3 shows snapshot of the system. The interface takes Hindi word as input and returns all the synsets for that word and its linking with SUMO. Fig 3 shows Hindi synsets and SUMO linking for the word *जीवन (‘Jeevana’, life)*.

6 Observation

Currently, 22148 Hindi synsets are linked with the English WordNet and SUMO. Table 6 shows the status of linking. Each row refers to synsets of specific part of speech and column refers to type of SUMO link.

Using the Hindi WordNet-SUMO link the formal semantic structure and axioms of SUMO can be used to process texts of different Indian languages.

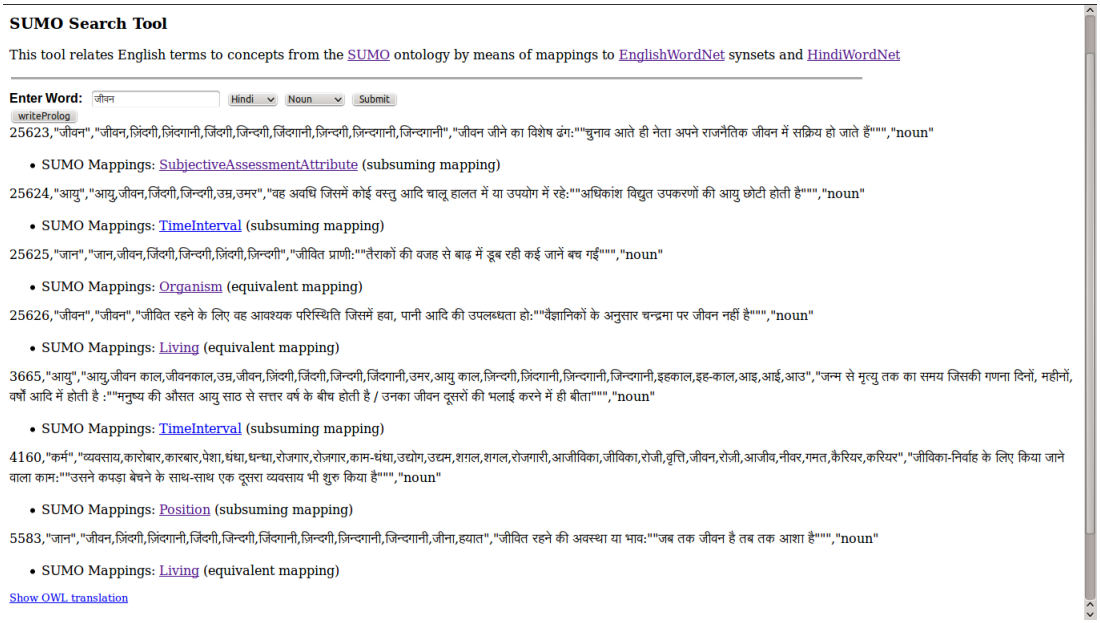

Figure 3: SUMO linking for the Hindi synsets

Table 6: Hindi synsets linked with SUMO

POS	Direct	Subsumption	Instance	Compl	Total
adjective	194	2487	310	102	3093
adverb	14	210	5	4	233
noun	2336	13690	1071	19	17116
verb	231	1473	2	0	1706

6.1 Inferencing using SUMO

Formal semantics of SUMO can be used for inferencing on natural language text using WordNet. For example, if there are two statements,

X: *उसे गाना पसंद है*
('use gaanaa pasanda hai')
(she singing likes)
(she likes singing)
Y: *उसे संगीत पसंद है*
('use sangeet pasanda hai')
(she music likes)
(she likes music)

then using WordNet there is no easy way to infer Y from X, as *गाना ('gaanaa', singing)* and *संगीत ('sangeet', music)* are not related through hypernymy-hyponymy relation. It requires processing gloss or to measure semantic distance to relate these two concepts. However in SUMO, गाना ('gaanaa', singing) is related to music by direct subsumption relation. So inferring Y from X becomes easier.

6.2 Restructuring WordNet concepts

Linking of IndoWordNet with SUMO can be used to restructure WordNet concept and extract taxonomy from the Indian language documents. Figure 4 shows the hyponymy of the Hindi synset containing words *जीव ('Jeev'), प्राणी ('PraNi')* linked with SUMO concept 'organism'. SUMO classifies the concepts 'organism' into three different categories 'animal', 'plant' and 'micro organism'. As shown in fig. 4, the WordNet synsets can be grouped into these three categories by SUMO link. The SUMO concept hierarchy can be extended by referring to WordNet synsets in all of these categories. This can lead to a large taxonomy of the 'organism'.

As shown in figure 4, hyponymy synsets represent the heterogeneous level of generality. It contains synsets corresponding to terms *प्राणी ('praaNI', Animal), मित्र ('mitra', Friend) , पैड ('peda' , Tree), पौराणीक पुरुष ('pauraNIka puRusha', Mythological being), etc*. SUMO link can be used to separate out concepts like *मित्र ('mitra', Friend)* from the structure of organism. Hindi WordNet classifies *जलचर ('jalachar')* as a separate concept. As per this classification the concept related to *कमल ('kamala' , Lotus)* is under the hyponymy of *जलचर ('jalachar')*. So it is not related to 'plant' or 'flower'. With SUMO linking it can be correctly categorized under the concept 'flowering plant'. This way, IndoWordNet SUMO linking helps in restructuring WordNet and domain

SYNSET: जीव, प्राणी, जीवधारी, जीवात्मा, अनीश, सजीव, प्राणधारी, तनुधारी,
Gloss : सजीव प्राणी या वह जिसमें प्राण हो
Example statement : "पृथ्वी पर विभिन्न प्रकार के जीव पाये जाते हैं"
Gloss in English : a living thing that has (or can develop) the ability to act or function independently

Direct Hyponymy in Hindi Wordnet	SUMO term
748 प्राणी , जीव , जीवधारी, वह जीवधारी जिसमें स्वैच्छिक गति होती है "पृथ्वी पर अनेकों प्रकार के जन्तु पाये जाते हैं"	organism **Animal** invertebrate arthropod mollusk worm vertebrate cold blooded vertebrate fish reptile warm blooded vertebrate bird mammal
715मित्र , दोस्त , साथी , वह जो सब बातों में सहायक और शुभचिन्तक हो "सच्चे मित्र की परीक्षा आपत्ति-काल में होती है"	normative attribute **subjective assessment attribute**
1249वनस्पति , पेड़-पौधा , वह सजीव जिसमें गति नहीं होती है और अधिकांशतः वह अपना भोजन स्वयं बनाता है "जंगलों में तरह-तरह की वनस्पतियाँ पायी जाती हैं"	organism **plant** flowering plant non flowering plant
1712पौराणिक जीव , वह जीव जिसका वर्णन पुराणों या धार्मिक ग्रंथों में मिलता है "हेमन्त पौराणिक जीव संबंधी कथाएँ सुनने में रुचि लेता है"	agent **sentient agent** cognitive agent
1554जलीय जीव , जलचर जल में पाया जानेवाला या रहनेवाला जीव (जंतु, वनस्पति आदि) "शैवाल, कमल, मोलस्क आदि जलीय जीव हैं"	Plant **flowering plant**
4490रोगाणु , वे सूक्ष्म जीव जो हवा या खाने-पीने की चीज़ों में मिले रहते हैं और अनेक प्रकार के रोगों के मूल कारण माने जाते हैं"रोगाणु मानव के लिए बहुत ही घातक होते हैं"	organism **microorganism** bacterium virus

Figure 4: relation of synsets in SUMO

specific taxonomies can be extracted by using this linking.

7 Conclusion

Linking WordNets of the Indian languages and SUMO creates a useful resource for natural language processing applications targeted at Indian languages. The language independent semantics and formalism of SUMO ontology can be used with WordNet for various text processing applications. The system is made available through two different interfaces: IndoWordNet browser that provides WordNet browsing in different Indian languages and SIGMAKEE open source package for first order theory development. This system can be useful for concept and relation extraction from Indian language documents. Future aim is to develop methods for automatic domain specific ontology extraction using this system.

Acknowledgements

The support of the Department of Information Technology (DIT), Ministry of Communication and Information Technology, Government of India and also of Ministry of Human Resource Development is gratefully acknowledged.

References

Pushpak Bhattacharyya. 2010. Indowordnet. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA).

D. Chakrabarty P. Pande D. Narayan and P. Bhattacharyya. 2002. An experience in building the indo wordnet - a wordnet for hindi. In *International Conference on Global WordNet (GWC02), Mysore, India*.

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. MIT Press.

Aldo Gangemi, Nicola Guarino, Claudio Masolo, and Alessandro Oltramari. 2003. Sweetening WordNet with DOLCE. *AI Magazine*, 24(3):13–24, sep.

T. R. Gruber. 1993. Towards principles for the design of ontologies used for knowledge sharing. In *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Deventer, The Netherlands. Kluwer Academic Publishers.

Nicola Guarino. 1998. Formal ontology in information systems : Proceedings of the 1st international conference june 6-8, 1998. pages 3–15. IOS Press.

Ian Niles and Adam Pease. 2001. Towards a standard upper ontology. In *FOIS '01 Proceedings of the international conference on Formal Ontology in Information Systems - Volume 2001*, pages 2–9. ACM Press.

Ian Niles and Adam Pease. 2003. Linking lexicons and ontologies: Mapping wordnet to the suggested upper merged ontology. In *Proceedings of the 2003 International Conference on Information and Knowledge Engineering (IKE 03), Las Vegas*, pages 23–26.

Ian Niles and A Terry. 2004. The milo: A general-purpose, mid-level ontology. In *Proceedings of the International Conference on Information and Knowledge Engineering, Las Vegas, Nevada, USA*, pages 5–19. CSREA Press.

Daniel Oberle, Anupriya Ankolekar, Pascal Hitzler, Philipp Cimiano, Michael Sintek, Malte Kiesel, B. Mougouie, S. Vembu, S. Baumann, Massimo Romanelli, Paul Buitelaar, R. Engel, D. Sonntag, N. Reithinger, Berenike Loos, R. Porzel, H.-P. Zorn, V. Micelli, C Schmidt, Moritz Weiten, F. Burkhardt, and J. Zhou. 2007. Dolce ergo sumo: On foundational and domain models in swinto (smartweb integrated ontology)). *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 5(3):156–174.

Adam Pease and Christoph Benzmüller. 2010. Sigma: An integrated development environment for logical theory development. In *The ECAI 2010 Workshop on Intelligent Engineering Techniques for Knowledge Bases (IKBET'2010), Lisbon, Portugal*.

Adam pease and Christiane Fellbaum. 2010. Formal ontology as interlingua: The sumo and wordnet linking project and global wordnet. In *Ontology and Lexicon, A Natural Language Processing perspective*, pages 25–35. Cambridge University Press.

Adam Pease, Christiane Fellbaum, and Piek Vossen. 2008. Building the global wordnet grid. In *Proceedings of the 18th International Congress of Linguists (CIL18), Seoul, Republic of Korea*.

J. Ramanand, Akshay Ukey, Brahm Kiran Singh, and Pushpak Bhattacharyya. 2007. Mapping and structural analysis of multi-lingual wordnets. *IEEE Data Eng. Bull.*, 30(1):30–43.

Alexandre Riazanov and Andrei Voronkov. 2002. The design and implementation of vampire. *AI Commun.*, 15:91–110, August.

Jaya Saraswati, Rajita Shukla, Ripple Goyal, and Pushpak Bhattacharyya. 2010. Hindi to english wordnet linkage: Challenges and solutions. In *IndoWordNet Workshop 2010, colocated with ICON 2010, Kharagpur, India*.

Manish Sinha, Mahesh Reddy, and Pushpak Bhattacharyya. 2006. An approach towards construction and application of multilingual indo-wordnet. In *3rd Global Wordnet Conference (GWC 06), Jeju Island, Korea*.

Piek Vossen, editor. 1998. *EuroWordNet: a multilingual database with lexical semantic networks*. Kluwer Academic Publishers, Norwell, MA, USA.