==
Language in India www.languageinindia.com ISSN 1930-2940 Vol. 17:1 January 2017
==

Resolving Polysemy in Malayalam Verbs

Mohan Raj, S.N. M.A., M.Phil., Sachin Kumar, S. M.Tech., and S. Rajendran, Ph.D.
==

Abstract

Polysemy in verbs is a challenging problem in linguistics as well as in natural language processing (NLP). Verbs are the most polysemous words among all the grammatical categories. The polysemy leads to word sense ambiguity. Resolving polysemy in verbs requires certain steps leading to word sense disambiguation (WSD). The paper makes use of the methodology proposed by Rumshisky Anna (2008). The result of the method is encouraging. Further improvement can be done by making use of other knowledge sources like wordNet, dictionary and onto-thesaurus.

Key words: ambiguity, argument, context, contextualized similarity, corpus pattern analysis, disambiguation, distributional similarity, grammatical categories, homonymy, lexical ambiguity, polsemy, selectional equivalence structural ambiguity, wordNet,

1. Introduction

High degree of polysemy prevails in natural language and so whatever utterance we come across is liable to be interpreted in multiple ways. But the high degree of ambiguity does not hamper our understanding of the concerned utterance. Mostly the context nullifies the multiple interpretations and assigns a single interpretation to the given expression. It is the context which helps a native speaker to interpret an utterance or sentence correctly. Any automatic way of interpreting the sense of a lexical item expects the contexts to select or activate the correct sense out of the competitive senses of the concerned lexical items. All major word classes exhibit lexical ambiguity and the contextual factors relevant for the concerned word resolves the meaning of the targeted word. For example, the meaning assigned to an adjective may be determined by the semantics of the head noun; the meaning of a polysemous noun may be determined by the governing verb or a modifier; and the meaning of a verb may be determined by their argument structure or by the dependents and other elements of their syntactic frame (Rumshisky, 2008: 215). This is illustrated below and the relevant senses are given in parentheses:

1a. *veegatayuLLa kaar* 'fast car' (the car that is or can be driven fast)

1b. *veegatayuLLa joolikkaar* 'fast worker (one who works fast)'

2a. *patratte curuTTu* 'role the newspaper' (physical object)

2b. *patratte paTikku* ' read the newspaper' (content)

3a. *avar atinRe naSTatte uLkoNTu* 'They accepted the loss' (pay)

3b. *avar aa varttamaanam uLkoNTu* 'They understood the information' (learn)

The aim of this paper is to resolve the lexical ambiguity of the verbs by automatic means of finding sense distinctions using the semantics of the arguments of the targeted verbs. We can exploit the argument structure of a verb to interpret the correct sense of a verb. The corpus can give us the distribution of the verb with reference to its context. By means of distributional similarity we will be able to select the correct sense of a verb. The idea that semantic similarity between words must be revealed in the similarity of characteristic contexts in which words occur is fairly obvious. It has been expressed in many ways as found in the "strong contextual hypothesis" of Miller and Charles (1991), and in the well-known remark of Firth, "You shall know a word by the company it keeps" (Firth 1957:11). The possibility that similar senses of the same word will occur in similar contexts can be made use of for resolving lexical ambiguity. However, applying the idea of distributional similarity in computational tasks faces problems. One of the main problems is that one must be able to identify the sense in which a polysemous word is used in each case in order to use any kind of generalization based on distributional information,

The aim of this paper is to resolve the lexical ambiguity of the verbs by automatic means of finding sense distinctions that can be identified by looking at the semantics of the arguments of the targeted verbs. Any polysemous target word and it's 'selectors' i.e. the words with which it forms syntactic dependencies can be dealt using the same methodology. The insight about the data can help us to design a targeted distributional approach to the problem of polysemy. In this paper, we focus on identifying verbal ambiguities linked directly to the semantics of the words that occur in a particular argument position. As we will see, such words may activate the same sense of the target verb, and yet be quite distinct semantically. In other words, they need to be similar only with respect to the context provided by that verb. In line with Rumshisk (2008:216), we too develop a clustering method that relies on contextualized similarity to group such elements together.

2. Relevance of context for Resolving polysemy

The meaning is assigned to a word by the combination of two factors: (1) syntactic frame into which the word is embedded, and (2) semantics of the words with which it forms syntactic dependencies (Rumshisky, 2008: 217). Rumshisky (2008: 217) use the term 'selector' to refer to such words. This is done regardless of whether the target word is the headword or the dependent in the syntactic relation. Syntactic frame comprises of the minor categories such as adverbials, locatives, temporal adjuncts, etc. and the subphrasal cues such as genitives, partitives, negatives, bare plural/determiner distinction, infinitives, etc. The set of all 'usage contexts' in which a polysemous word occurs can usually be divided into groups. Each group roughly corresponds to a distinct 'sense'.

Consider following sentences with the verbs *niSeedhikkuka* 'deny' made use of in 4 and *paRayuka* 'say' made use of in 5 to illustrate the contribution of different context parameters to disambiguation. The difference in the syntactic patterns for the verb *niSeedhiccu 'deny'* as shown in 4a and 4b disambiguate between the two dominant senses: 'refuse to grant' and 'proclaim false' and similarly the difference in the syntactic patterns of *paRayuka* 'say' as shown in 5a and 5b disambiguate between 'complain' and 'blame'.

Syntactic frame

4a. *adhikaari atinu kaaraNam uNTu enna kaariyatte niSeedhiccu* {*enna*-clause)

'The authority denied that there is a reason for that' (proclaim false)

4b. *adhikaari kaaraNatte niSeedhiccu* [NP]

'The authority denied the reason' (proclaim false)

4c. *adhikaari raadhaykku visa niSeedhiccu* [NP]

'The authority dined visa to Radha' (refuse to grant)

5a. *avaL koccine aTiccu ennU avan kuRRam paRanjnu* [*ennu*-clause]

'He complained that she bet the child' (complain)

5b. *avan avaLe kuRRam paRanjnju* (blame) [NP]

'He blamed her'

Consider the following sentences with *ooTuka* 'run', *valicceTukkuka* ‘absorb’, and *perumaaRuka* ‘behave’. The contrasting argument and/or adjunct semantics shown in 6, 7 and 8 evoke the different the senses of *ooTuka* 'run', *valicceTukkuka* ‘absorb’, and *perumaaRuka* ‘behave’ respectively. The relevant argument type is shown in brackets and the corresponding sense in parentheses:

Semantics of the arguments and adjuncts/adverbials:

6a. *pooliis kaLLanRe puRakil ooTi*

'The police ran after the thief' [chased person] (run)

6b. *avaL avanRe kuuTe ooTippooyi*

'She ran away with him' [accompanying person] (elope)

7a.*panjnji veLLatte valicceTuttu*

‘The cotton absorbed the water’ [water] (absorb}

7b.*avaL avnRe kaiyil ninnum paisa valicceTuttu* [money] (take away)

‘She extracted money from him’

8a. *avan avaLooTU mariyadayaayi perumaaRi*

'He behaved with her with respect' [with respect] (behave)

8b. *pooliis avane nallavaNNam perumaaRi*

'Police beat him severely’ [severely] (beat)

A lexicographer establishes a set of senses available to a particular lexical item and (to some extent) specifies the context elements which typically activate each sense. This procedure is formalized in several current resource-oriented projects such as FrameNet and Corpus Pattern Analysis (CPA). FrameNet makes use of Fillmore’s case roles to represent semantics of the arguments. Corpus Pattern Analysis (CPA) (Hanks & Pustejovsky 2005) attempts to catalog prototypical norms of usage for individual words, specifying them in terms of context patterns. Each pattern gives a combination of surface textual clues and argument specifications. CPA uses the extended notion of syntactic frame, as outlined above. Semantics of the arguments is represented either through a set of shallow semantic types representing basic semantic features such as person, location, physical object, abstract, event, etc. or extensionally through ‘lexical sets’, which are effectively collections of lexical items.

3. Polysemy and Distributional Similarity

A number of tasks in NLP make use of the notion of distributional similarity. Distributional similarity is exploited in the areas such as word sense disambiguation (WSD), sense induction, automatic thesaurus construction, selectional preference acquisition, and semantic role labeling (Rumshisky, 2008: 220). Semantically similar words (as in thesaurus construction) or similar uses of the same word (as in WSD and sense induction are identified by making use of distributional similarity. Distributional similarity results in clusters of distributionally similar words. These clusters are often seen as means to address the problem of data sparsity faced by many NLP tasks. The generalization based on the distibutional similarity of a polysemous word must apply to different 'senses' rather than to its entire occurrence uniformly. Semantics of the arguments is often represented using information derived from external knowledge sources, such as FrameNet, machine-readable dictionaries, WordNet, etc (Rumshisky, 2008: 220).

Context is typically represented as a feature vector. Here each feature corresponds to some context element. The value of each feature is the frequency with which that element is encountered together with the target word. A word may be represented as a feature vector combining all the context features or as a probability distribution on the joint events of occurrence of the target word with each context element. Some approaches use distributional features based on bag-of words style co-occurrence statistics, others use context representations that incorporate syntactic information, and sometimes semantic information from external sources. In the latter case, each distributional feature may correspond to a grammatical relation populated with a particular word or an entity type.

Resolving polysemy implies separating out the occurrences corresponding to each sense from the distributional representation of the target word. Typically, this problem is resolved by either clustering similar occurrence contexts for each word, or clustering the actual words whose overall distributional profiles are similar (Rumshisky, 2008: 221).

4. Sense Assignment for a Polysemous Predicate

Sense inventories for polysemous predicates are often comprised by a number of related senses. Computational approaches to word sense disambiguation assign a sense to each word in an utterance from an inventory of senses. This simplified statement may not be true when the meaning of a complex expression is computed. Consider a polysemous target predicate with certain semantic preferences. In a given argument position, different senses of that predicate will be selected for different semantic features. Thus, in 6b, the ‘elope’ sense of *ooTuka* is selected by the argument indicating the accompanying person while 'run' sense in 6a is selected by the argument indicating the 'chased person'; similarly, in7b, the 'take away' sense is selected by the argument denoting 'money' in the direct object position whereas the 'absorb' sense in 7a is selected by the argument denoting the direct object 'water'. In the same way in 8b, the 'beat' sense is selected by the adverbial adjunct *'nallavaNNam* 'severely' and the 'behave' sense in 8a is selected by the adverbial adjunct *maryaadayaayi* 'with respect'.

4.1. Sense-Activating Argument Sets for a polysemous predicate

A number of semantically diverse arguments may activate the same sense of a predicate. Certain pertinent semantic feature will be central to the interpretation of the meaning for some of the predicate. It will be merely a contextual interpretation that they permit for other predicates. Effectively, each sense of the target predicate may be seen to induce an ad-hoc semantic category in the relevant argument position. For example, consider the senses of the verb *eTukkuka*: ‘take (something)’, ‘raise (hood like a snake’), 'take photo', 'score', 'copy', etc. Some of the lexical items that occur in direct object position for these senses are given below.

9a. *paampU paTam eTuttu*

'The snake raised its hood'

9b. *avan avaLuTe paTam eTuttu*

'He photographed her'

9c. *avan atinRe pakarppU eTuttu*

'He took the copy of it'

9d. *avan kaNakkinU nuuRu Satamaanam maarkU eTuttu*

'He scored hundred percent mark in mathematics'

The predicates need to be activated or evoked for a relevant sense by its arguments. The nouns in each argument set which are semantically quite distinct activate the relevant sense of the predicate. A particular aspect of the sense is selected by the context provided by the predicate. Mostly an argument set carry a central component of their meaning as well as certain other peripheral components. The core members of the argument set may be polysemous. In order to activate the appropriate sense of the predicate a 'bidirectional selection' process needs to be implemented. But notice that the interpretation of 10a and 10b, for example, is quite different.

10a. *pooliis avane kasTaTiyil eTuttu*

'The police arrested him'

10b. *avan aa jooli ceyyaan kuRaccu samayam eTuttu*

'He took some time to do the work'

In the above sentences both the words in the argument position activate the same sense of *eTukkuka*. But the disambiguation is between the EVENT reading and the TIME reading. It has to be noted that the different dependencies the word enters into rely on the different aspects of the meaning. For example, consider the use of the noun arguments with the verb *eTukkuka* in the sentences 11a and 11b given below.

11a. *kamsanRe kaaraagrhattil krishNan janmam eTuttu*

'Krishnan was born in Kamsan's jail'

11b. *innale kooTatiyil ninnum avane jaamyattil eTuttu*

'He has been bailed out from the court'

In the above examples, the words *janmam* 'birth' and *jaamyattil* 'bail-in' activate different senses for *eTukkuka* 'take'. The context provided by the verb effectively changes the relevant semantic components in the interpretation of the senses.

4.2. Sense Separation Based on Selector

The semantic components selected for different senses may be sufficiently distinct in the case of homonymy (Rumshisky, 2008: 224). The relevant lexical items can be grouped together by making us of the overall distributional similarity between arguments. For example, movement sense of *maTangkuka* 'return' is easily distinguished from the cluster of senses related to 'fold as paper'.

12a.*avaL oophiisil ninnum viiTTileekkU maTangi*

'She returned from office'

12b. *pustakattinRe peejukaLokke maTngippooyi*

'The pages of the book got folded up'

Similarly the movement sense of *naTakkuka* 'walk' is easily distinguished from the cluster of senses related to 'go on'.

13a. *koccU patukkeppatukke naTakkaan tuTangi*

'the child started walking slowly'

13b. *aviTatte tiyeeTTaRil nalla cinima naTakkunnu*

'A cinema is going on in the theatre'

But separating different senses of the verb is notoriously hard even for a trained eye in the case of polysemy. This problem has been the subject of extensive study in lexical semantics. It aims at addressing the question of selecting distinct senses based on context. There is no clear cut method to say when a context selects a distinct sense or when a context simply modulates the meaning (Pustejovsky 1995, Cruse 1995, Apresjan 1973). This is crucial for the computational method of word sense disambiguation. Lexicographers often face problems in deciding when to describe a set of usages as a separate sense while "lumping and splitting" senses during dictionary construction. The sense separation is often resolved on ad-hoc basis. It results in instances where the same occurrence may fall under more than one sense category simultaneously resulting in numerous cases of "overlapping senses". Resolving verbal polysemy often runs into this problem of indecision.

We need to determine which selectors are likely to activate what sense. While resorting to such attempt we should keeping in mind that at least some of the verb's senses are interrelated. The occurrences of a polysemous verb in a corpus verb cluster into 2-10 groups, each roughly corresponding to a sense (Rumshisky 2008: 225). There are a lot of cases in which the sense distinctions are clear-cut and easily discernible. But there are some boundary cases where the sense diction of the predicate is not clear cut. Thus, in a given argument position, three kinds of selectors are possible (Rumshisky 2008: 225): (i) Good disambiguators where the selectors immediately pick one sense of the target word. (ii) Poor disambiguators where the selectors that may be used with either sense and require more contexts to be disambiguated themselves (bidirectional selection doesn't work). (iii) Boundary Cases where the choice between two senses of the target is in fact impossible to make (i.e. selector activates both senses at once). For example, for the subject position with the verb *kaaNikkuka* 'show' in 14, *sarve* 'survey' and *paTam* 'photo' are good disambiguators, while g*raaph* 'graph' is a clear example of a boundary case.

14a. *aa paTam avaLuTe mughatte nalla vaNNam kaaNikkunnu* ('pictorially represent')

'That picture shows her face well'

14b. *aa sarve vyvasaaya meekhalayil uLLa sarkkaarinRe puroogatiye kaaNikkunnu*

'The survey shows the improvement of government in industrial sector'. ('demonstrated by evidence or argument')

14c. *ii* g*raaph ii maasattil uLLa maZyuTe Saraasari aLavine kaaNikiunnu*

'The graph shows an average rain fall in this month'. (both senses?)

Each individual sense needs to be clearly defined for the identification of boundary cases. Such cases are better construed as instances of 'multiple selection' (i.e. simultaneous activation of both senses), and not merely as evidence for overlapping sense definitions. Even syntactic pattern cannot always overrule the interpretation intrinsic to some selectors. For example, in 15, it is virtually impossible to resolve *niSeedhikkukaa* 'deny' between 'refuse to grant' and 'proclaim false'.

15 a. *vayoodhikarkkU avaruTe staanam nisheedhikkappeTunnu*

'Elders are often denied the status of adulthood'

15b. *cila jaadikkaar striikaLkkU svaatandryam nisheedikkunnu*

'People of certain caste denies autonomy to women'

On the other hand, in 16, the selector itself is polysemous, with two interpretations available for it, and it needs to be disambiguated by context before it can activate the appropriate sense of the predicate.

16a. *paNTatte aphipraaytte nisheedikku*

'Deny the traditional view ('proclaim false')

16b. *avanRe anuvaadam nisheedikku*

'Deny him permision ('refuse to grant')'

In the following sections, we discuss how these considerations can be taken into account while designing a computational strategy for automatic sense detection.

5. Similarity measure

The goal of a similarity measure is to allow us to tell automatically whether one word is "like" the other. But whether one word is like the other may vary, depending on the particular task. If our task is to determine the meaning of a predicate by looking at its arguments, two words in the same argument position will be "like" each other only if they pick the same sense of the predicate. We can capture this intuition by defining a measure aimed to assess 'contextualized similarity', i.e. similarity between two lexical items with respect to a particular context.

Rumshisky uses the term 'context' to refer to a singleton, i.e. a single populated syntactic relation in the following discussion (Rumshisky 2008: 226), For example, the verb *eTukkka* 'take' and the relation of the direct object above define a particular context of occurrence for the noun meaning. At its most basic, distributional similarity between frequency profiles of two words should reflect to what extent the contexts in which the two words occur overlap. Similarity between two words may be expressed as the frequency of their occurrence in identical contexts, relative to the average of their overall frequencies. Some normalization is also typically used since the two words may have very different corpus frequencies,. The result is a function of relation tuple frequency, typically referred to as the 'weighting' or the 'association score' between the word and the context attribute. Defined in this manner, distributional similarity will be high for lexical items whose overall distributional profiles are similar. This will be the case for words which are semantically very close in their dominant, most frequent sense. Or, in a less likely case, it may be that most of their senses are similar, and have similar relative frequencies. When several nouns from a given argument set activate the same sense of a polysemous verb, high similarity values may be obtained for the elements of the semantically uniform core of this argument set (if such a core is present). On the other hand, polysemous core elements for which the relevant semantic component is not dominant, as well as peripheral elements of this argument set, will slip through the cracks. Hindle (1990) remarks that while one can have and sell both beer and wine, it's the fact that you can drink both of them that makes them semantically close. In other words, when computing semantic similarity based on distributional behavior, some contexts are, to quote Orwell, "more equal than others". The reason we know that two words are used similarly in a given context is that there is a number of other contexts in language where they are used in the same way. Such 'licensing context' licenses the use of these lexical items with the same sense of the target word.

Computing similarity between contexts thus poses a separate problem. It is clearly incorrect to use overall distributional similarity between context-defining words to determine how close two contexts are. In order to be considered similar, two contexts must be similar with respect to their selectional properties, i.e. select for the same semantic component in the specified argument position. This problem is addressed by Rumshisky by introducing the notion of 'selectional equivalence' (Rumshisky 2008: 229).

6. Selectional Equivalence for Verbs

Rumishky defines selectional equivalence for two verbs with respect to a particular argument position and a particular sense for each verb (Rumshisky 2008: 229). We organize nouns can into lexical sets sharing a semantic feature. Similarly verbs can be organized into selectional equivalence sets, with arguments sharing a semantic feature. A lexical item W_1 is a 'selectional equivalent' ('contextual synonym') of lexical item W_2 with respect to a certain grammatical relation *R* if one of its senses selects for the same aspect of meaning as one of the senses of W_2 in the argument position defined by *R* (Rumshisky 2008: 229). It is not necessary that the selectional equivalents are synonyms or antonyms of each other. They are equivalents only in terms of the aspect of meaning they select. Verbs that are selectionally equivalent to one of the senses of the target verb effectively form a subset of all licensing contexts for that sense. Selectional equivalents can be grouped into clusters representing different senses of the target verb, if we can measure how close two contexts are with respect to the target context. The likelihood of associating each selector with the concerned sense can be determined by the resulting clusters. The clusters of selectional equivalents obtained for selected senses of *eTukkuka* 'take' and *nishedhikkuka* 'deny' are shown in 14 given below.

17. *eTukkuka* 'take'

sviikarikkuka 'accept', *grahikkuka* 'absorb', *seegharikkuka* collect, *koTukkuka* 'give' *vaanguka* 'buy'

18. *nishedhikkuka* 'deny',

niraakarikkuka 'deny', *viSadhamaakkuka* 'disclose' and *veLippeTuttuka* 'reveal' *nirasikkuka*, *maRuttupaRayuka* 'say against', *Staapikkuka* 'confirm', *teLiyikkuka* 'clarify', uRappikkuka 'endorse', uRappaakkuka 'ratify', *sammtikkuka* 'agree'

A specific kind of semantic similarity is suggested by 'selectional equivalence'. It overlaps only partially with what manually constructed resources typically aim to capture. Selectionally equivalent verbs may belong to the same frame, or to the frames related through some frame-to-frame relation in FrameNet. As one would expect, semantically uniform core elements to be similar when the verbs that operate on them are from the same situational frame. For example, *maRukuka* 'deny' and *staapikkuka* 'confirm' in 14 both evoke the same 'statement frame'; *viSadhamaakkuka* 'disclose' and *veLippeTuttuka* 'reveal' evoke the frame which inherits from statement. On the other hand, certain other pairs are not likely to evoke related frames. The same partial overlap can be observed with Levin classes and WordNet categories. In order to obtain clusters of selectional equivalents for each sense of the target verb, we need to be able to measure to what extent two verb senses share selectional properties. This measure of selectional equivalence effectively mirrors contextualized similarity as defined for selectors. The idea is to take all selectors that occur in the specified argument position with the target verb, identify the verbs that occur with these selectors, and cluster them according to the sense of the target with which they share selectional properties. Our model involves the assumption that two verbs tend to be selectionally close with respect to just one of their senses. Similarity between two verbs is estimated based on selectors that, for each of them, consistently activate the sense which is selectionally equivalent to one of the target's senses. In the next section, we outline the overall architecture of the algorithm and discuss in more detail the choice of reliable selectors. We then look at some results of the similarity computation based on the obtained selector lists.

7. Analysis of Annotation Decisions

In this paper, a simple recurrent neural network based learning approach is applied to identify the senses of *ooTu* 'run'. Recurrent Neural Networks (RNN) takes its idea from the human understanding of word from its previous word in the sequence. Traditional neural network did not have the mechanism to understand more about the present events based on the previous events. RNN's are neural networks with loops which will allow the information to exist in it. A simple RNN model is shown in figure

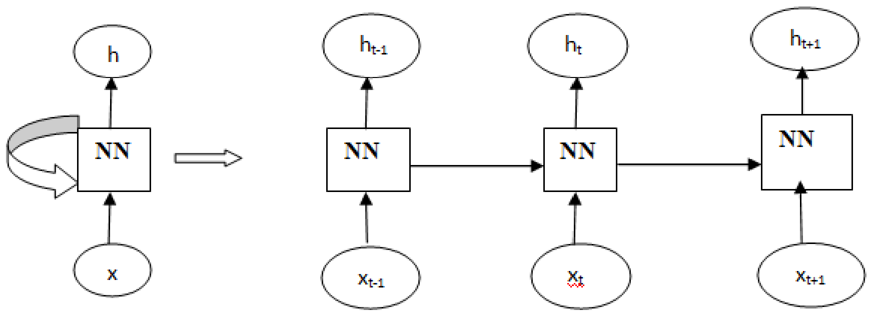

Figure 1: Simple RNN

In figure 1, 'NN' denotes any neural network architecture, ' x_t' denotes an input and ' h_t' denotes the output. Loop passes the output value to the next stage. RNNs essentially helps in captures the present information with the immediate previous value. However, it will omit if the immediate previous value or values will not contribute to understand the sequence of words. In such cases, long short-term memory, an extended version of RNN is used which can even capture long dependencies. In this paper, a simple RNN is used to distinguish the sense. 9 sense classes of *ooTu.*

Sense No.	**No. of data**
1	100
2	99
3	104
4	105
5	100
6	68
7	54
8	26
9	104

Table 1: No. of data in each class

Accuracy	**Precision**	**Recall**	**F-measure**
0.34328	0.368	0.343	0.332
0.69403	0.765	0.694	0.701
0.93284	0.938	0.933	0.933
0.97015	0.972	0.97	0.97

0.97761	0.981	0.978	0.978
0.85075	0.876	0.851	0.856
0.90299	0.905	0.903	0.902
0.98507	0.985	0.985	0.985
0.99254	0.993	0.993	0.993

Table 2: Result of 9 epochs. Each epoch is ran for 1000 steps

Table 2 shows the evaluation measures obtained for 9 epochs. When more words and its corresponding senses are created, the network can be tuned to get good models.

8. Conclusion

The result of the proposed method is encouraging as shown in the previous tables. It associates each of the target's senses with a cluster of selectional equivalents for that sense, with selectional equivalents represented as short contextualized vectors of reliable selectors. The resulting clusters serve to identify selectors that activate each sense, with association scores obtained for each selector indicating which sense it tends to activate. Even with certain assumptions about parallel sense distinctions and selector polysemy, we seem to be able to overcome some of the difficulties encountered by the previous attempts to address polysemy resolution. The evaluations discussed in the result section shows that simple RNN provides a competing result for identifying the senses of *ooTu* 'run'.

==

References

Apresjan, Ju. 1973. Regular polysemy. Linguistics, 142(5):5–32.

Briscoe, T. and J. Carroll. 2002. Robust accurate statistical annotation of general text. Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002), Las Palmas, Canary Islands, May 2002, pages 1499–1504.

Carpuat, M. and D. Wu. 2007. Improving statistical machine translation using word sense disambiguation. In Proc. of EMNLP-CoNLL, pages 61–72.

Chan, Y. S., H. T. Ng, and D. Chiang. 2007. Word sense disambiguation improves statistical machine translation. In Proc. of ACL, pages 33–40, Prague, Czech Republic, June.

Cruse, D. A. 1995. Polysemy and related phenomena from a cognitive linguistic viewpoint. In Dizier, Patrick St. and Evelyne Viegas, editors, Computational Lexical Semantics, pages 33–49. Cambridge University Press, Cambridge, England.

Cruse, D. A. 2000. Meaning in Language, an Introduction to Semantics and Pragmatics. Oxford University Press, Oxford, United Kingdom.

Firth, J.R. 1957. Papers in Linguistics 1934-1951. Oxford University Press, Oxford.

Hanks, P. and Pustejovsky, J. 2005. A pattern dictionary for natural language processing. Revue Franc¸aise de Linguistique Appliquee, 10:2

Hindle Donald 1990. Noun classification from predicate-argument structures. 28th Annual Meeting of the Association for Computational Linguistics Pittsburgh, PA. 268-275.

Hiroaki, S. 2003. FrameSQL: A software tool for FrameNet. In Proceedigns of ASIALEX '03, pages 251–258, Tokyo, Japan. Asian Association of Lexicography

http://colah.github.io/posts/2015-08-Understanding-LSTMs/ (Accessed on 11/09/2016)

http://karpathy.github.io/2015/05/21/rnn-effectiveness/ (Accessed on 11/09/2016)

Miller, G.A. and Charles, W.G. 1991. Contextual correlates of semantic similarity, Language, cognition and neuroscience 6(1): 1-28.

Pustejovsky, J. 1995. Generative Lexicon. Cambridge, MA: MIT Press.

Pustejovsky, J and Rumshisky, A 2010. Mechanisms of Sense Extension in Verbs. In G-M de Schryver (ed.) *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks.* Kampala: Menha Publishers.

Rumshisky, A 2008. Resolving Polysemy in Verbs: Contextualized Distributional Approach to Argument Semantics. In A. Lenci (ed.) *From Context to Meaning: Distributional Models of the Lexicon in Linguistics and Cognitive Science. Special Issue of the Italian Journal of Linguistics. Rivista di Linguistica, 20(1)*.

Rumshisky, A and Batiukova, O. 2008. Polysemy in verbs: systematic relations between senses and their effect on annotation. In HumanJudge '08 Proceedings of the workshop on Human Judgments in Computational Linguistics pages 33-41, Manchester, United Kingdom

.

Mohan Raj, S.N. M.A. Mphil
snmohanraj@gmail.com

Sachin Kumar, S. M.Tech
sachinnme@gmail.com

S. Rajendran. PhD
rajushush@gmail.com

Centre of Excellence in Computational Engineering and Networking (CEN)
Amrita Vishwa Vidyapeetham
Coimbatore 641 112
TamilNadu
India