Discovery of Manner Relations and their Applicability to Question Answering

Roxana Girju[1,2], Manju Putcha[1] and Dan Moldovan[1]
Human Language Technology Research Institute[1]
University of Texas at Dallas
and
Department of Computer Science[2]
Baylor University
girju@ecs.baylor.edu, moldovan@utdallas.edu

Abstract

The discovery of semantic relations from text becomes increasingly important for applications such as Question Answering, Information Extraction, Summarization, Text Understanding and others. This paper presents a method for the automatic discovery of *manner* relations using a Naive Bayes learning algorithm. The method was tested on the UPenn Treebank2 corpus, and the targeted manner relations were detected with a precision of 64.44% and a recall of 68.67%.

1 Introduction

1.1 Problem description

An important semantic relation for several NLP applications is the *manner* relation. Consider the sentence (from the Democratic response to the President Bush' 2003 State of the Union Address):

We want to work together to build our new economy, creating jobs by investing in technology so America can continue to lead the world in growth and opportunity.

There are four manner relations in this text: (1) *together* is a manner adverb that modifies the verb *work*, (2) *creating jobs* is an adverbial phrase attached through a manner relation to the verb *work*, (3) *by investing in technology* is a prepositional phrase that expresses manner and attaches to the verb *create*, and (4) *in growth and opportunity* is a manner prepositional phrase that modifies the verb *lead*.

The discovery of manner relations in open text allows Question Answering systems to identify these relations and formulate answers to *manner* questions that otherwise are not possible even with state-of-the-art QA systems. For example, by identifying the manner relations in the example above, the following *how* questions may be answered:

Q: *How do Democrats want America to lead the world ?* A: *in growth and opportunity*
Q: *How do Democrats want to work?* A: *work together (with Republicans).*
Q: *How do Democrats want to build the economy ?* A: *by creating jobs*;
Q: *How do Democrats want to create jobs?* A: *by investing in technology*

This paper provides a method for discovering manner semantic relations in open text.

1.2 The semantics of manner relation

In WordNet, the manner relation is defined as *a way of acting or behaving*. Similar definitions are provided by psychology researchers (Graesser et al., 2000).

There are different ways of expressing manner and the difficulty arises that the same lexico-syntactic patterns that express manner also express other semantic relations in different contexts. A possible way to check whether or not a verb expression conveys manner is to answer correctly the question "*In what manner/how* $<$ *to_verb* $>$?" For example, for *run quickly*, we ask *how to run?* However, this test holds only when there are no other answers to questions like "*Where* $<$ *verb* $>$?", or "*When* $<$ *verb* $>$?" that make sense. For example, *jump over the fence* or *jump always* are not manner relations although they may answer correctly a *how* question.

1.3 Previous work

Although manner relations were studied by philosophers (Aristotle, 350BC), logicians, psychologists and linguists (Quirk et al., 1985), (Fellbaum, 2002), not much work has been done to automatically identify the manner relations in texts. Hearst (Hearst, 1998) developed a method for the automatic acquisition of hypernymy relations by identifying a set of frequently used and unambiguous lexico-syntactic patterns. Then, she tried applying the same method to other semantic relations, such as part-whole, but without much success, as the patterns detected were ambiguous.

2 Lexico-syntactic patterns expressing manner

2.1 Manner as semantic role

The most frequently occurring form of manner is as a semantic role (Quirk et al., 1985). In this case, manner is encoded as a relationship between a verb and one of its arguments which can be represented by various parts of speech, the most common ones being *adverb, adverbial phrase*, *prepositional phrase*, *noun phrase*, and *clause*.

Verb-adverb patterns
One of the most frequently used patterns expressing manner is *verb-adverb*. In English, there are different kinds of adverbs (Quirk et al., 1985): adverbs of time, manner, degree, location, direction, frequency, transition and hedges.

Based on the classification provided by Quirk et al. (Quirk et al., 1985) and our statistics of English texts, we present below the adverbial patterns in order of their frequency of occurrence:

a) Adverbs of manner that end in "-ly"
This manner adverbs are the most frequently used. Their position is not fixed, as they can be placed either before or after the verb they modify. These adverbs can be modified by other adverbs forming this way adverbial expressions. Examples: *slowly, heavily, angrily*, etc.

b) Adverbs of manner that do not end in "-ly"
These adverbs also called *Quality description adverbs* provide a description of a particular quality.

Example: *fast, good, well*, etc.

c) Adverbial expressions
These are expressions that modify the underlying verb and refer along with the verb to a manner relation. Examples of such patterns are: <as adv_manner as $NP/adv/S$>, < NP as adv_manner >, < as adv_manner S >.

Examples: *several times as fast, as much as 60% faster, louder than ever, all around*, etc.

d) Compound adverbs of manner
These adverbs are usually formed with words linked by hypens. Examples: *radio-style, tax-free, flat-out, first-hand*, etc

e) Foreign adverbial expressions
There are expressions boroughed from other languages that are in a manner relationship with the underlying verb. Examples: *in flagrante, a la Gorbachev, en masse*, etc.

2.2 Other forms of manner relations

In addition to the manner roles expressed as verb-adverb pairs, manner relations are also expressed as (1) complex nominals (*fast car*), (2) verbs of implicit manner (for example *whisper* is a manner of speaking), (3) verb-PP (*I took your coat by mistake*), (4) verb-NP (*He breathed a deep breath*), (5) verb clauses (*I cook vegetables as Chinese do*), and others.

All these lexico-syntactic patterns are ambiguous. Thus we need some syntactic and semantic constraints to differentiate the manner relations from the other possible meanings these patterns may have.

In this paper we focus only on the discovery of manner semantic roles expressed as verb- adverb pairs. The method, however, is extendable to many other manner forms and even to other semantic relations.

3 Approach

The learning procedure proposed here is *supervised*, for the learning algorithm is provided with a set of inputs along with the corresponding set of correct outputs. In this paper we use the Naive Bayes Classifier approach to determine whether or not a verb-adverb pair indicates a manner relation. This method is similar with the basic algorithm for Document Classification (Mitchell, 1997).

Nr.	Feature
1	Specific adverb statistics
2	Parent phrase type
3	Present or not in the Adverb Dictionary
4	Distance between verb and adverb
5	Component before adverb
6	Component after the adverb
7	Adverbs ends or not with 'ly

Table 1: Summary of Manner Features.

This approach requires a decision on how to represent an arbitrary text in terms of attribute (or features) values and how to estimate their probabilities as required by the Naive Bayes Classifier.

4 Selecting features

Many researchers ((Blaheta-Charniak, 2000), (Gildea-Jurafsky, 2000), (Gildea-Palmer, 2002)) showed that lexical and syntactic information is very useful for predicate-argument recognition tasks. Their systems are statistical-based and have been trained to automatically label semantic roles only from the output of syntactic parsers.

However, lexical and syntactic information alone is not sufficient for the detection of the manner semantic roles, semantic information is necessary as well.

To represent the text for the discovery of manner relations, seven features which contribute the most to the classification were chosen. These features capture the context of the adverb and help in deciding the presence of the manner (MNR) component.

We have developed an Adverb Dictionary that is a source for some of the features. The Adverb Dictionary is created with adverbs from WordNet and TreeBank. The adverbs that contain the pattern "in a —– manner" in their gloss were extracted from WordNet. The adverbs that are annotated in TreeBank as MNR adverb-verb pairs are also included in the Dictionary. A total of 2183 adverbs were included in the Dictionary.

The features are explained with the help of the following example:

(S1 (S (NP (DT The) (NN bank)) (VP (AUX is) (ADVP (RB now))(VP (ADVP (RB aggressively)) (VBG marketing) (NP (JJ retail)(NNS services)) (PP (IN at) (NP (PRP$ its) (JJ domestic) (NNS branches))))) (. .)))

(1) Specific adverb statistics

Feature 1 checks if a specific adverb is present in the Dictionary or not. For example, *aggressively* is part of the Dictionary, where as *now* is not. The positive frequency calculated from this feature is the total number of times that adverb was encountered in the training corpus. In the case the adverb of a sentence in the testing corpus is part of the Dictionary, this feature helps in deciding what are its chances of being a Positive/Negative Indicator of Manner. This is a good feature as long as the training corpus is very rich (i.e it covers all adverbs).

(2) Parent phrase type

The second feature is the phrase type to which the adverb attaches. Here both *now* and *aggressively* attach to "VP". Most of the MNR indicating adverbs attach to verbs. This feature helps eliminate adverbs, which modify nouns or adjectives.

(3) Whether or not Adverb is present in the Dictionary

Feature 3, like feature 1 checks whether or not an adverb is present in the Adverb Dictionary. The difference is that its statistics are not calculated on the training corpus like in feature 1, but instead it takes the probability of being a manner adverb in the Adverb Dictionary.

The usefulness of feature 3 is realized when the test corpus has an adverb which was not encountered in the training corpus. The estimates from feature 1 fail to be of any use at such a point because it is a missing value and both positive and negative frequencies are the same. However, feature 3 assigns the probabilities of that adverb being a manner adverb in the Adverb Dictionary. So, we still have a good estimate from this feature to decide if it is a potential MNR indicator or not (which would have been nullified, had we relied only on feature 1).

For example, let's say we encounter the adverb *excitedly* in the test corpus and it is present in the Adverb Dictionary but not in the training corpus. Feature 1 will not contribute to the decision while feature 3 will help. We can use the lookup table for feature 3 and it is evident that an adverb present in the Dictionary has a higher probability of indicating manner.

(4) Distance between verb and adverb

The fourth feature is the distance between verb and adverb. This doesn't take into consideration whether the adverb precedes or succeeds the verb. Distance refers to the number of English words that separate them. For example, there are no words between *aggressively* and *marketing*, thus the distance is 0. Similarly, the distance between *now* and *marketing* is 1. The rational of this feature is based on the observation that most frequently a MNR indicating adverb appears immediately next to a VB.

(5) Component before the adverb

The fifth feature concerns the POS of the word preceding the adverb. This captures the context of the adverb. This is based on the observation that an adverb that succeeds an AUX is usually not a MNR indicator. For example *now* is preceeded by "AUX" and *aggressively* is preceded by an "ADVP".

(6) Component after the adverb

The sixth feature concerns the POS of the word after the RB. For example *now* is succeeded by an "AUX" and *aggressively* by an "VBG".

(7) Adverb ends in "ly"

This feature is 1 when the adverb ends in "ly" and 0 otherwise. The rational for this feature is that many adverbs in manner roles end in "ly".

Estimating Probabilities

The next step is to calculate the probabilities required by the Naive Bayes Classifier.

a. Class prior probabilities. This is the ratio between the number of adverbs of each class over the total number of adverbs in the training examples. In our case the classes are positive (or Manner) and negative (not Manner). This is defined as:

$$P(V_j) = E_j/E$$

where E_j is the total number of examples for which the target value is V_j and E is the total number of examples.

b. Class conditional probability. This is the probability that any of the seven features drawn from the parsed text tagged positive or negative will belong to the domain of the corresponding features. We use the m-estimate to avoid the cases when pos_freq and neg_freq are very small.

$$Prob(+) = (pos\_freq + 1)/(VOCAB + TEXT)$$
$$Prob(-) = (neg\_freq + 1)/(VOCAB + TEXT)$$

where pos_freq is the number of times the feature occurred in the Positive class, neg_freq is the number of times the feature occurred in the Negative class, $VOCAB$ is the distinct number of positive and negative instances for a given feature, and $TEXT$ is the total number of all positive and negative instances in the examples.

4.1 Learning Algorithm

The algorithm learns the probability that a given adverb indicates manner (i.e. how many times the adverb occurred in the positive class and how many times in the negative class). Similarly, it learns the probability that it attaches to a VP/NP/... in each of the positive and negative classes. The same is true for all features.

At the end of the learning process, the algorithm creates look-up tables for all the features. These are used by the classifier. The learning step along with the output are explained in the next section.

$$V_{nb} = \underset{v_j \in V}{argmax} P(v_j) \prod_{f_i \in F} P(f_i/v_j)$$

where V_{nb} is the output of the Naive Bayes Classifier, v_j is the class in the target set V, and f_i are the individual features from the set F of the seven features.

5 Experimental Setting

5.1 Building the Training and Test Corpus

In order to learn the constraints, we used the Treebank2 (Marcus, 1994) text collection and LA Times Corpus. Treebank2 is a corpus featuring one million words of 1989 Wall Street Journal material annotated with several predicate-argument structures. It is annotated with the following semantic roles: *BNF* (beneficiary), *DIR* (direction), *EXT* (spatial extent), *LOC* (location), *MNR* (manner), *PRP* (purpose and reason), and *TMP* (temporal). Treebank2 contains different types of manner annotations: ADVP-MNR (1683), PP-MNR(952), SBAR-MNR (60), NP-MNR(54), S-MNR(48), UCP-MNR (8), ADJP-MNR(1). For the work in this paper we used the ADVP-MNR annotations from Treebank2.

The input to the program is a parsed text. For training and testing the Treebank2 corpus is split in the 3:1 ratio. The algorithm doesn't work on the parsed text directly. Instead, the parsed text is converted into the 7-feature format augmented with the value of the target function as shown in Table 2

Adverb	Parent	In-Dict	Distance	Before	After	ly	target
now	VP	0	1	AUX	VP	0	no
then	S	0	3	no_before	NP	0	no
long	VP	0	1	AUX	VP	0	no
back	VP	0	5	NP	SBAR	0	no
aggressively	VP	1	0	ADVP	VBG	1	yes
magisterially	VP	1	2	NP	.	1	yes
directly	VP	1	0	VBN	PP	1	yes
rapidly	VP	1	0	AUX	VP	1	yes

Table 2: A sample of training data

Creation of the Look-Up table
Given this format as input, the learning algorithm creates LookUp tables using the Class Conditional Probability and Reference files. These files contain the domain of the features. Each feature can take a set of legal values encountered during training. Table 3 exemplifies the lookup entries for some feature examples.

6 Results for discovering manner relations

Let us define the *precision* and *recall* performance metrics in this context.

$$precision = \frac{Number\ of\ correctly\ retrieved\ relations}{Number\ of\ relations\ retrieved}$$

$$recall = \frac{Number\ of\ correctly\ retrieved\ relations}{Number\ of\ correct\ relations}$$

The experiments were conducted with the annotations in UPenn's Treebank2. The results of the first experiment are shown in Tables 4.

First experiment
Training = (1176 Positive + 2546 Negative) = 3722 examples
Testing = (507 Positive + 1183 Negative) =1690 examples.
Output of the program:
Prior Positive Probability = 0.315959162
Prior Negative Probability = 0.684040838
Precision = 191/242 = 78.92%
Recall = 191/507 = 37.62%

Second experiment
Based on the results from the previous set of results it is observed that considering adverbs like *moreover, then, thus* which can never indicate MNR reduces both the precision and recall. Therefore they were removed from the set of negative examples. Similarly the intensifiers like *much, very, so* were also removed from the positive examples.
Training examples = 1103 Pos + 1352 Neg = 2355
Test Corpus = 508 Pos + 1183 Neg = 1690
Prior Positive Probability = 0.4492
Prior Negative Probability = 0.5740
The results are shown in Table 5.

Relations	No. of relations
Nr of MNR relations in corpus	507
Number MNR relations retrieved	242
Number of correctly retrieved rel	191
Precision	191/242 = 78.92%
Recall	191/507 = 37.62%

Table 4: The precision and recall for experiment 1

Relations	No. of relations
Nr of MNR relations in corpus	507
Number MNR relations retrieved	540
Number of correctly retrieved rel	348
Precision	348/540 = 64.44%
Recall	348/507 = 68.67%

Table 5: The precision and recall for experiment 2

7 Application to Question Answering

The manner semantic relation occurs with high frequency in open text. Its discovery is paramount for many applications, such as Information Extraction, Text Mining, Knowledge Base construction, etc. In this section we mentioned only Question Answering.

The concepts and manner relations acquired from a collection of documents can be useful in answering difficult questions that normally can not be handled based solely on keywords matching and proximity. As the level of difficulty increases, Question Answering systems need richer semantic resources, including the discovery of semantic relations in open texts. In the case of a manner question, the answer

Feature	Feature Example	Nr Pos	Nr Neg	Prob(+)	Prob(-)
adverb	aggressively	18	2	0.000087	0.000014
	magisterially	4	0	000023	000005
	directly	34	0	0.000159	0.000005
parent	VP	1510	329	0.011178	0.002441
	no_par	99	243	0.000740	0.001805
	S	42	117	0.000318	0.000873
Dictionary yes	1	1175	1107	0.005371	0.005061
no	0	2	1440	0.000014	0.006582
distance	0	881	1036	0.004028	0.004736
	2	68	275	0.000315	0.001260
	1	142	515	0.000653	0.002356
POS preceding	ADVP	32	55	0.000151	0.000256
	NP	273	661	0.001251	0.003023
	VBN	107	101	0.000493	0.000466
POS after	VBG	57	27	0.000265	0.000128
	-	93	70	0.000429	0.000324
	PP	211	221	0.000968	0.001014
ends with "ly"	1	990	740	0.004526	0.003385
	0	185	1805	0.000850	0.008249

Table 3: Example of features look-up table

type of that question may be tagged as MNR. To provide the correct answer, often it is sufficient to locate first the paragraph where the potential answer is and then identify the MNR tag in that paragraph. In case when several such MNR tags exist, more reasoning is necessary. Consider the following examples which show the MNR tag in the answer sentence.

Q: *How did Bob Marley die?*

A1: Bob Marley died [of Melanoma][MNR].

Q: *How was little Johnny dressed last night?*

A1: Dressed [in a cowboy style][MNR], Johnny walked proudly on the street.

Q: *How does Marry dance?*

A1: Marry danced [as well as Bill][MNR].

Q: *How does Lina Mayors charms her audience?*

A1: Countering every unfruitful description, her work communicates and [impresses through the rhythm of the colors][MNR].

8 Conclusions

The method presented in this paper for the detection and validation of manner relations is automatic and novel. We combined lexical, syntactic and semantic features for a more accurate learning.

Naive Bayes Classifier assumes feature independence. Here, features 1 and 4 are independent, the rest are dependent on each other. This is the reason for 65-70% precision and recall. By using some heuristics like removing unambiguous adverbs these were helped. The improvement made in the second experiment is significant because if an adverb like *now*, or *moreover* is included in the negative examples, then other features which contribute to a positive example are nullified and the decision becomes less precise. For example, *apparently* attaches to VP and VP usually occurs in a positive class, and the inclusion of this example in the negative example reduces the estimates of VP to contribute to positive examples.

The Naive Bayes Classifier, though oversimplified by the independence assumption, proved to be a good classifier in the document classification and also promises to be a useful method for the discovery of semantic relations.

References

Aristotle. *On Sophistical Refutations*. On Sofistical Refulations, section 3, Translated by W. A. Pickard-Cambridge.

Don Blaheta and Eugene Charniak, 2000. *Assigning Function Tags to Parsed Text*. Proceedings of the 1st Annual Meeting of the North American Chapter of the Association for Computational Linguistics, Seattle, May 2000, pp. 234–240

Martha Palmer, Joseph Rosenzweig, William Schuler 1998 *Capturing Motion Verb Generalizations with Synchronous TAGs* Predicative Forms in NLP,pp 250-277, ed by Patrick St. Dizier, Kluwer Press, December, 1998.

Beth Levin - *English Verb Classes and Alternations* The University of Chicago Press

Cornelia Maria Verspoor 1997 *Contextually Dependent Lexical Semantics* The University of Edinburgh, 1997

Julia B. St.John *On the Semantics of Manner Adverbs* Carolina Working Papers in Linguistics Vol 1, Issue 1

2002 *Oriented Adverbs* Issues in Lexical Semantics of Event Adverbs, Von Wilhem Gauder, 2002

Christiane Fellbaum 2002 *On the Semantics of Troponymy* Cognitive Science Laboratory, Princeton University, December 2002.

Tom Mitchell 1997 *Machine Learning* McGraw Hill, 1997

Daniel Gildea and Daniel Jurafsky. 2000. *Automatic Labeling of Semantic Roles*. In Proceedings of the 38th Annual Conference of the Association for Computational Linguistics (ACL-00), pages 512-520, Hong Kong, October 2000.

Daniel Gildea and Martha Palmer. 2002. *The Necessity of Syntactic Parsing for Predicate Argument Recognition*. In Proceedings of the 40th Annual Conference of the Association for Computational Linguistics (ACL-02), Philadelphia, PA, 2002.

Arthur C Grasser, Peter Weimer Hastings and Katiga Waimer Hastings. 2002. *Constructing Inferences and Relations during Text Comprehension*.

M. Hearst. 1998. *Automated Discovery of WordNet Relations, An Electronic Lexical Database and Some of its Applications*. MIT Press, Cambridge MA, 1998.

Judith Levi. 1978. *The Syntax and Semantics of Complex Nominals*. NY: Academic Press.

Beth Levin. 1993. *English Verb Classes and Alternations*. The University of Chicago Press

M. Marcus. 1994. *The Penn treebank: A revised corpus design for extracting predicate-argument structur*. In Proceedings of hte ARPA Human Language Technology Workshop, Princeton, NJ, 1994.

R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. 1995. *A comprehensive grammar of English language*. Longman, Harlow, 1985