Learning WordNet-Based Classification Rules

Sebastian van Delden

Department of Computer Science
University of Central Florida
Orlando, Florida 32816
sdelden@cs.ucf.edu

Abstract

A new classification paradigm, which automatically acquires WordNet-Based rules from a corpus, is presented. The approach is applied to developing an autonomous software agent that can recognize emotions which are expressed in natural language during an interactive human-computer environment. Such an agent could adapt to a user's emotional state and dynamically adjust its interaction etiquette. Hierarchical concepts of WordNet's noun and verb hypernymy are the basic building blocks of the classification rules. A greedy learning algorithm automatically determines which hierarchical concepts are best suited for each rule. A corpus of 5000 emotional sentences has been compiled from 502 test subjects and serves as input to the system.

1 Introduction

The significance of WordNet(Miller, 1993) in the natural language community is apparent by the multiplicity of applications and papers which have included WordNet in their design. We present a new classification paradigm that automatically acquires WordNet-Based rules. Hierarchical concepts of the WordNet noun and verb hypernymy are the basic building blocks of the classification rules. A greedy algorithm learns which hierarchical concepts are best suited for each rule.

This approach is applied to learning rules that determine if an emotion is implicitly being expressed in a natural language sentence. For example: *I won the lottery*. We assume that *I* am *happy*, since I have *acquired wealth* or *won a game*. The WordNet hierarchical concepts of the syntactic relations in the sentence are used to form rules which categorize the emotion that is suggested in the sentence. For example, the following rule could be learned from *I won the lottery*: *if the subject is 'I' and the main verb hypernymy contains 'win' and the head noun of the direct object hypernymy contains 'game' then emotion is 'happy'*. This single rule, however, also classifies many other similar sentences: *I won the basketball game*, *I swept the games*, *I won the treasure hunt*, etc… A greedy algorithm(similar to transformation-based learning – Ngai and Florian, 2001; Brill, 1995) chooses the best WordNet concepts for a rule, maximizing the number of correctly classified sentences in a corpus. Furthermore, because of the extensiveness of the WordNet noun and verb ontologies, learned rules apply not only to examples that are seen during training, but also to new unseen sentences.

Much work has been done in the area of language and emotion. Davitz (1969) presents a dictionary of emotional meaning in which fifty emotional concepts are defined by natural language descriptions of personal feelings. These descriptions were chosen by subjects from an exhaustive list of 556 possible descriptions which were generated from interviews and written reports from over 1200 test subjects. Our corpus was created in a similar fashion, but emotional situations are described, not emotional feelings. Wierzbicka (1992a) and Wierzbicka (1992b) focuses on finding a set of semantic primitives or lexical universals – concepts that have been lexicalized in all languages of the world. These primitives are then used to formulate prototypical scripts which define emotional concepts. Walker et al. (1997) introduces a theory and set of algorithms for improvisation of spoken utterances by artificial agents. Suggesting that the way people express themselves is a product of their character and personality. Romney et al. (1997) presents an interesting study of comparing emotional models across the English and Japanese languages. They found that English-speaking and Japanese-speaking subjects share a single model of the semantic structure of emotional terms.

Little research, however, has been done on implementing a software agent that will recognize emotions expressed in natural language. Carberry et al. (2002) does describe how attitudes of doubt can be recognized in natural language. Our work describes a more general system that can be applied to many emotions that are expressed in the English language.

Emotions are classified into one of seven primitive emotion types: happiness, sadness, anger, fear, disgust, surprise, and confusion. Davitz (1969) describes fifty emotional concepts that are commonly accepted. This approach can be extended to include more emotional concepts, but for now is restricted to these seven for testing purposes.

The remainder of this paper will be organized as follows: Section 2 describes the problem of recognizing emotions; Section 3 explains the WordNet-based learning paradigm in detail; Section 4 presents an evaluation of the system, including the top ten emotional rules that have been automatically acquired; and Section 5 concludes the paper.

2 Recognizing Emotions

The focus of this research is to recognize emotions that are expressed implicitly in natural language sentences. For example, sentences such as *I am happy* or *I am confused* are not considered, since *happiness* and *confusion* are explicitly stated. Also, no attempt is made to recognize emotions that are expressed figuratively (Chapter 6 of Fussell, in press). For example, specific emotional metaphors such as *getting hot under the collar* or *hit the roof*. These metaphors are language specific and could be included by hard coding them in the natural language emotion recognizer.

Emotions are classified into one of seven primitive emotion types: happiness, sadness, anger, fear, disgust, surprise, and confusion. This classification is based on the syntactic relations in the sentence. For simplicity, only the following syntactic relations are considered in our rules: pre-verbal noun phrase(assumed to be the subject), main verb, and first post-verbal compliment(usually the direct object - DO). For example, consider the sentence *I have a headache*. This sentence suggests that *I* am feeling *down*, since *I* have *an ailment*. The emotion primitive which most closely fits this sentence would be *sad*. A rule to classify this sentence can be constructed as follows:

```
IF subject is I
AND main verb is have
AND head noun of DO is headache
THEN emotion is sadness
```

This rule would correctly classify *I have a headache*, but is restricted to only that sentence. The goal is to construct a rule from this sentence which also classifies many other similar sentences.

Using WordNet[1] noun and verb hypernymy, this rule can easily be extended to classify many similar sentences. For example, the WordNet hypernymy(only Sense 2/2 is shown) of *headache* is:

```
Sense 2
headache, head ache, cephalalgia
=> ache, aching
 => pain, hurting
  => symptom
   => evidence, grounds
    => information
     => cognition, knowledge
      => psychological feature
```

The idea is to replace *headache* with its super-concept *pain* in the above rule. Instead of exactly matching the direct object of new sentences, *pain* is searched for in the direct object's hypernymy. The rule becomes:

```
IF subject is I
AND main verb HYPs contain have
AND head noun HYPs of DO contain pain
THEN emotion is sadness

where HYPs refers to the super-concepts of
the WordNet hypernymy.
```

This rule now classifies many related, unseen sentences: *I have a burn on my hand*, *I have chest pains*, *I sustained a twinge in my lower back*, etc… The difficulty lies in choosing the best combination of WordNet concepts for the rules. Choosing a super-concept located too far up the hierarchy may result in many errors because the concept is too generalized. Conversely, choosing a super-concept too close in the hierarchy may result in many emotional situations being missed. For example, you may not want to choose *information* as the head noun of the direct object in the above example because it may generate too many errors. Section 3 describes how the best combination of super-concepts can be automatically acquired from a test corpus.

A few issues still need to be addressed when attempting to identify emotions: 1) Negation, 2) Adjective Information, 3) Intensity, and 4) Possession.

If negation is present in the verb phrase, the rule should not apply, for example, *I do not have a headache*. Except in the cases where negation is required to classify the emotion correctly: *I do not understand the professor*.

Adjective information in the direct object may also alter the meaning of the emotion: *I read a great book* versus *I read a terrible book*. To capture this adjective information, the synonym set of the adjective is obtained from WordNet. An extra restriction is then included in the rule which requires the presence of an adjective from the synonym set. This restriction makes the rule more precise without causing it to be too specific.

An intensity value can be added to the emotion by examining the adverb(s) that may be present in the sentence. For example, *I have a really bad headache* or *I really like the book*.

The presence of a possessive pronoun may also effect the classification of a rule. For example, *He broke my table leg* versus *He broke the table leg*. The first sentence suggests than I am *angry* since someone has destroyed a possession of mine. The second sentence, however, does not suggest any type of emotion.

These four factors have been included into the rules to improve the accuracy of the emotion recognition system.

[1] WordNet version 1.6 is used in our experiments

Num	Sentence	Emotion
1	I have a job	happiness
2	I have work	happiness
3	I have an occupation	happiness
4	I have an injury	sadness
5	I have a bruise	sadness

Figure 1. Sample corpus in which only the direct object differs from sentence to sentence.

3 WordNet-Based Learning

Test data was collected from over 500 graduate and undergraduate students, professors and business people. A corpus of over 5000 sentences was created in which each sentence implicitly describes an emotion and has a corresponding emotion label. A greedy learning algorithm, similar to transformation-based learning(Ngai and Florian, 2001; Brill, 1995), chooses the best WordNet hierarchical concepts for the rules, maximizing the number of correctly classified emotions in the corpus. For example, Figure 1 shows a sample corpus which will be used to illustrate the learning technique.

Initially all sentences are unclassified. For each sentence *i* in the corpus, a set of triples $<n_{i,1}, v_i, n_{i,2}>$ is created using every combination of the WordNet super-concepts of the subject, verb and direct object. A *score* is assigned to each triple based on how many sentences it correctly matches minus the number of incorrect matches. A *match* is generated for triple$_i$ when a test sentence's subject hypernymy contains $n_{i,1}$, the verb hypernymy contains v_i and the direct object hypernymy contains $n_{i,2}$. The triple with the highest score is chosen and used to classify the corpus. Processing continues until all sentences have been classified.

Num	Triple	Emotion
1	<I, have, job>	happiness
2	<I, have, occupation>	happiness
3	<I, have, activity>	happiness
4	<I, have, act>	happiness

Applying triple 1 to the corpus, yields a score of: 1 correct match – 0 incorrect matches = 1. Triple 1 only matches sentence 1 because *job* only appears in the hypernymy of direct object of sentence 1. Triple 2 obtains a score of 3, matching sentences 1, 2 and 3, with no incorrect matches. Triple 3 obtains a score of 2, matching sentences 1, 2 and 3, but produces and incorrect match of sentence 4 because *injury* contains *activity* in its hypernymy. In this case *activity* is too general a super-concept to be used and therefore causes the error. Triple 4 also obtains a score of 2 by producing an error on sentence 5.

Processing continues for sentences 2, 3, 4, and 5, until all possible triples have been exhausted. The best triple is then chosen – in this case triple 2 from sentence 1 - <I, have, occupation>. This triple is now applied to the corpus, classifying sentences 1, 2, and 3. Processing then continues with sentences 4 and 5. From these sentences, the triple <I, have, injury> which implies *sadness* will be chosen. So from this test corpus, the following rule set is generated:

For example, consider the sample corpus in Figure 1. For simplification, all sentences share the same subject and main verb: *I* and *have*, respectively, and no adjective information is included in the Direct Object. So the objective here is to choose the best super-concept to represent the head noun of the direct object(DO). If adjective information was present, then a four tuple would be create which included the adjective synonym set. Sentence 1 is chosen and the following triples are generated:

```
IF subject is I
AND main verb is have
AND HYPs of DO contain occupation
THEN emotion is happiness

IF subject is I
AND main verb is have
AND HYPs of DO contain injury
THEN emotion is sadness
```

To restrict the number of triples that are generated, only the super-concepts of the first sense of the verb or noun are considered. This reduces the time to learn rules, and is a good approximation, since the first sense represents the most likely occurring sense of the word. However, many words may occur in a context where they do not correspond to the first WordNet sense, therefore possibly producing erroneous rules.

Ideally the exact WordNet sense, which corresponds to the particular context of the main verb, should be chosen. For example, consider the sentences*: I lost my father* versus *I lost 10 pounds*. The first *lost* refers to WordNet Sense 3 – *to suffer*, while the second *lost* most closely refers to WordNet Sense 1 – *lose an abstraction*. Failure to make this distinction may cause an erroneous rule to be created. Gomez(2001; 1998) describes an algorithm which uses an enhanced WordNet to determine the meaning of the verb as well as its thematic roles, adjuncts and prepositional phrase attachment. The algorithm is based on predicates that have been defined by Gomez for WordNet verb classes. The thematic roles in the predicates contain information about the grammatical relations and ontological categories that realize them. Incorporating this algorithm into the system would not only remove irrelevant verb senses, but also identify the correct senses of the nouns in the subject and direct object.

To combat errors which are caused due to a wrong sense being assigned to a verb or noun, a threshold value can be assigned to the greedy algorithm. For example, *only keep rules which produce a score of at least five*. This will not only help eliminate incorrect rules, but also generate a more concrete set of rules since numerous test subjects have agreed on their particular classifications.

4 Evaluation

The corpus, described in Section 3, was used to evaluate the learning algorithm. Our corpus contained both emotional and non-emotional sentences. For example, every sentence has one of the following eight labels assigned to it: happiness, sadness, anger, confusion, disgust, surprise, fear, and no-emotion. The data was collected from test subjects with no restrictions on how sentences could be structure. Therefore, an attempt was made to reword as many sentences as possible so that they contained a pre-verbal noun phrase(assumed to be the subject of the sentence), main verb and post-verbal compliment(usually a direct object - DO), since these are the only syntactic relations currently being used in the rules.

A cutoff threshold of *five* was used so that only the most relevant rules would be captured. Table 2 lists the top ten rules that were learned from the test corpus. The algorithm was augmented so that a rule could be formed when a particular word was present in a syntactic relation. *Word* = is used here to indicate that if the following word is present, then the syntactic relation satisfies the criteria without having to examine hypernymy information. When asked to describe an emotional situation, many test subjects described an instance where *they* were involved in an emotion situation. Therefore, the words *I* or *my* are very often used. *HYP* = indicates that the head noun or main verb of the current syntactic relation should have the following concept in its WordNet hypernymy. *SYNSA* = indicates that if an adjective is present, it should have the following adjective in its WordNet synonym set. Finally, *NEG* is used to indicate that the syntactic relation is being negated.

These emotion rules are dependent on the test subject's environment and culture. For example, the first three rules identify a situation that expresses happiness. Since many of the test subjects were employed or in college, they very often described a situation where they got a new job or a good grade on a test. Many subjects also stated the well-being of their family to be a situation of happiness. Rule three is more likely to be expressed across different environments and cultures than rules 1 and 2. For example, had the test subjects being taken from an environment where hunting and farming are central to making one's living, the most common response may have been a situation describing favorable weather conditions, and not receiving a good test grade. Fortunately, if test data is collected from different environments, the algorithm presented here can automatically learn the most common set of emotion rules so that an emotion recognition system can be fitted to a particular environment or culture.

Rules 4 and 5 emerged as the primary situations for which a subject felt sad. These situations describe when the subject had lost a possession, or a relative or possession had died, for example, *My dog died*. or *I lost my favorite necklace*.

Anger was most commonly expressed when a person or object struck the test subject(causing damage), or a subject's possession, for example*: The bus slammed into my parked car*. or *She slapped me*.

Table 2. Top ten rules acquired from the test corpus by the greedy learning algorithm.

No.	Emotion	Subject	Verb	DO
1	Happiness	Word = I	HYP = acquire	HYP = occupation
2	Happiness	Word = I	HYP = acquire	SYNSA = good
3	Happiness	Word = my	HYP = be	SYNSA = healthy
4	Sadness	Word = I	HYP = lose	Word = my
5	Sadness	Word = my	HYP = die	-
6	Anger	-	HYP = strike	Word = my, me
7	Surprise	Word = I	HYP = experience	SYNSA = unexpected
8	Disgust	Word = I	HYP = experience	HYP = insect
9	Confusion	Word = I	HYP = understand NEG	-
10	Confusion	Word = I	HYP = decide NEG	-

For surprise, the most common situations described perceiving something unexpected. Finally, for disgust many subjects described situations where they had encountered an insect, for example*: I saw a fly in my soup* or *I just saw a roach*. Otherwise, very few subjects offered similar situations when asked to describe a surprising or disgusting situation. Finally, confusion was most often described as not being able to understand something or make a decision, for example: *I couldn't figure out the problem*.

The emotion rules that were generated from our test corpus are also capable of handling many other similar sentences that were not in the test corpus because of the extensiveness of the WordNet hierarchies. This is a desirable capability of learning WordNet-based rules. However, more testing is needed to determine the accuracy of such rules when applied to other corpora.

5 Conclusions

We have presented a new learning paradigm which automatically acquires WordNet-Based classification rules. These rules not only maximize the number of correct classifications in a test corpus, but also correctly classify new unseen sentences. The approach has been applied to learning rules that determine if an emotion is implicitly being expressed in a natural language sentence, but is not limited to this application. This learning technique could be adapted to any system which uses WordNet hypernymy information in classification rules.

To improve the accuracy of this approach, the correct WordNet senses need to be determined, which can be performed by incorporating an algorithm developed by Gomez(2001;1998). To improve the application of an emotion recognition system, more syntactic relations could be included in the rules so that a broader range of natural language sentences can be covered. A more detailed analysis still needs to be performed to determine the accuracy of a set of acquired rules across different corpora.

References

Eric Brill. 1995. *Transformation-based error-driven learning and natural language processing:* A case study in part of speech tagging Computational Linguistics, 21(4):543-565.

Sandra Carberry, Lambert, L., and Schroeder, L. 2002. *Towards Recognizing and Conveying An Attitude of Doubt Via Natural Language*. To Appear in Applied Artificial Intelligence.

Joel Davitz. 1969. *The Language of Emotion*. Academic Press, Inc. New York.

Fussell, S. R., Ed. (in press). *The verbal communication of emotion: Interdisciplinary perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates.

George Miller. 1993. *Introduction to WordNet: An On-line Lexical Database*, Princeton, CSL Report 43.

G. Ngai and R. Florian. 2001. *Transformation-Based Learning in the Fast Lane*. North American Chapter of the Association for Computational Linguistics.

Kimball Romney, Moore, C., and Rusch D. 1997. *Cultural universals: Measuring the semantic structure of emotion terms in English and Japanese*. In the Proceedings of the National Academy of Science, 94: 5489-5494.

Walker, M. and Cahn, J., and Wittaker, S. 1997. *Improvising Linguistic Style: Social and Affective Bases for Agent Personality*, In Proceedings of the First International Conference on Autonomous Agents.

Wierzbicka, A. 1992a. *Defining Emotion Concepts*. Cognitive Science, 16: 539-581.

Wierzbicka, A. 1992b. *Talking about Emotions: Semantics, Culture, and Cognition*. Cognition and Emotion, 6(3/4): 285-319.