

# Automatic Discovery of Intentions in Text and its Application to Question Answering

Marta Tatu

Human Language Technology Research Institute  
Department of Computer Science  
University of Texas at Dallas  
Richardson, TX 75080, USA  
marta@hlt.utdallas.edu

## Abstract

Semantic relations between text concepts denote the core elements of lexical semantics. This paper presents a model for the automatic detection of INTENTION semantic relation. Our approach first identifies the syntactic patterns that encode intentions, then we select syntactic and semantic features for a SVM learning classifier. In conclusion, we discuss the application of INTENTION relations to Q&A.

## 1 Introduction

### 1.1 Problem description

*Intentions* comprise of semantic relationships that express a human's goal-oriented private states of mind, including intents, objectives, aims, and purposes. As a relation, it encodes information that might not be explicitly stated in text and its detection might require inferences and human judgment. The answer to the question *What was Putin trying to achieve by increasing military cooperation with North Korea?* is found in the sentence *Putin is attempting to restore Russia's influence in the East Asian region.* Extracting the exact answer *to restore Russia's influence in the East Asian region* becomes easier if this is recognized as Putin's intention which matches the question's expected answer.

In this paper, we describe a method that identifies *intentions* in domain independent texts. We employed two machine learning algorithms to create models that locate intentions in a given paragraph using a set of six syntactic and semantic features.

### 1.2 Motivation

The current state-of-the-art NLP systems cannot extract *intentions* from open text and, as we saw in the example, their detection benefits Question Answering. An intention is the answer to general questions like *What is the goal of X?*, *What does X plan to do?*, or *What does X aim for?* The INTENTION semantic relation is one of the most challenging relations because text fragments may convey unstated intentions. These are most pervasive in dialogues, communication specific to humans. For example, in the following conversation, the vendor infers the client's unstated intention of buying the cups.

Customer: *Where do you have the \$1 cups?*

Salesman: *How many do you want?*

Intentions are closely related to other semantic relations such as beliefs, motives, desires, or plans. In the above example, the context tells us that this takes place in a superstore, well-known as a place where people buy things from. The clerk's answer emerges from our common beliefs and background knowledge as well as from his desire to help a customer. Intentions are the framework for plans. Many philosophers and artificial intelligence researchers studied the intentions as parts of coordinating plans (Bratman, 1987; Pollack, 1990) because people establish plans for future times.

In this paper, we regard *intentions* as expressions of a particular action that shall take place in the future, in which the speaker is some sort of agent (Anscombe, 1957). For example, the sentence *Mary is going to buy a TV set* shows Mary's intention. Anscombe (1957) considers intentions as a subclass of predictions, besides commands and

prophecies. *John is going to be sick* is usually a prophecy, *John, go for a walk!* is an order, and *John plans to take a walk* expresses an intention.

### 1.3 Previous work

Various methodologies have been proposed and used over the years for the task of extracting semantic relations from text. Purely probabilistic models, empirical methods, or hand-coded constraints were some of the approaches that do not use machine learning algorithms. Later on, methods that use decision tree, neural networks, memory-based learning, or support vector machines were introduced. Currently, there is also an increased interest in shallow semantic parsing of open texts and automatic labeling of semantic roles. Wiebe et al. (2004) focused on the detection of subjective language such as opinions, evaluations, or emotions in text. Using clues of subjectivity (low-frequency words, collocations), they identify opinion piece texts such as editorials, letters to the editor, or arts and leisure reviews.

There exists an immense literature in philosophy about the different types of intentions and their characteristics. Bratman (1987) tries to find the relationship between the two distinct phenomena of *doing something intentionally* and *intending to do something*. Numerous philosophical studies discuss how intentions relate to other psychological concepts, such as, beliefs, desires, hopes, or expectations (Audi, 1973; Bratman, 1981; Bratman, 1987). Intentions are consistent with the person’s beliefs, and, unlike ordinary desires, require consistency (Bratman, 1987). They can generate reasons for or against future intentions (Bratman, 1981; Bratman, 1987). As plan elements, intentions require a certain stability. Their side effects need not be intended, even if they were taken into consideration in the first place<sup>1</sup> (Bratman, 1990).

## 2 Syntax and Semantics of Intention

### 2.1 Syntactic patterns

Because, in all the cases that we encountered, intentions were conveyed by phrases, we took a closer look at how intentions can be expressed in the written text. For our investigations, we chose the Sem-

<sup>1</sup>Due to space limitations, we couldn’t include detailed examples. Please see the cited articles for examples.

Cor text collection (Miller et al., 1993), a subset of the Brown corpus manually tagged with WordNet senses (37,176 sentences in 352 newspaper articles). After manually classifying the first 2,700 sentences from SemCor into sentences that contain or not intentions, only 46 examples were identified. The syntactic patterns listed in Table 1 cover 95.65% of them. Because the first pattern comprises more than half of the studied examples, our algorithm focuses on detecting intentions encoded by  $VB_1$  to  $VB_2$ . We note that this pattern is ambiguous and may convey other semantics. For instance, *Mary began to play with the dog*, *He told her to meet you* are encoded by our pattern, but do not express intentions.

Pattern	Example	Frequency
$VB_1$ to $VB_2$	plan to go for a walk	27 (58.69)
$NN$ to $VB$	strivings to give up drink	6 (13.04)
$VB$ $PP$ $VP$	He resigned so that he can work for the school campaign	5 (10.87)
<i>goal/purpose is to VB</i>	his goal is to leave the country	4 (8.69)
$ADJ$ to $VB$	eager to end a pitching slump	2 (4.34)

Table 1: INTENTION syntactic patterns

### 2.2 Semantics of intentions

From the semantic point of view, an intention may be very specific, it may contain a future time or a location (*John intends to meet Mary today*), but every intention must specify a future action. Hence, we propose the following representation for the INTENTION semantic relation:  $INT(e_1, x_1, e_2)$  where  $e_1$  is the event denoting the intention,  $x_1$  denotes the person that has the intention and  $e_2$  is the intended action or event. If the intention is more specific then we will identify instances of other semantic relations<sup>2</sup>.  $John(x_1) \wedge INT(e_1, x_1, e_2) \wedge meet(e_2) \wedge Mary(x_2) \wedge today(x_3) \wedge THEME(e_2, x_2) \wedge TIME(e_2, x_3)$  represents a more specific intention.

The semantics of the INTENTION relation allows the derivation of inference rules which show that INTENTION dominates other semantic relations such as PURPOSE, ENTAIL, or ISA. For example, if a person  $x_1$  intends to perform action  $e_2$  and this action has a purpose  $e_3$ , then we can say that  $x_1$  intends to do  $e_3$ <sup>3</sup>. Formally, we can express the above relations

<sup>2</sup>The list of semantic relations that can specialize an INT includes THEME, LOCATION, TEMPORAL, MANNER, INSTRUMENT, SOURCE, MEANS, and FREQUENCY. Their arguments are  $e_2$ , the intention verb, and a corresponding  $x_i$ .

<sup>3</sup>Similar statements can be made for the ENTAIL and ISA

with the following set of implications<sup>4</sup>:

$$\begin{aligned} \text{INT}(e_1, x_1, e_2) \wedge \text{PURPOSE}(e_2, e_3) &\Rightarrow \text{INT}(e_4, x_1, e_3) \\ \text{INT}(e_1, x_1, e_2) \wedge \text{ENTAIL}(e_2, e_3) &\Rightarrow \text{INT}(e_4, x_1, e_3) \\ \text{INT}(e_1, x_1, e_2) \wedge \text{IS-A}(e_2, e_3) &\Rightarrow \text{INT}(e_4, x_1, e_3) \\ \text{INT}(e_1, x_1, e_2) \wedge \text{PURPOSE}(e_3, e_2) &\not\Rightarrow \text{INT}(e_4, x_1, e_3) \\ \text{INT}(e_1, x_1, e_2) \wedge \text{CAUSE}(e_2, e_3) &\not\Rightarrow \text{INT}(e_4, x_1, e_3) \end{aligned}$$

The first three implications formalize the above inference rules. If *John intends to start his car to go to the park*, then John intends to go to the park. Similarly, if *John intends to buy a car*, then we can say that he intends to pay for it. The sentences *John intends to go to the park*. *He’s starting his car right now* express John’s intention to go to the park ( $e_2$ ). The purpose of starting the car ( $e_3$ ) is to go to the park. We cannot say that John intends to start his car. This is just an intentional action done to achieve his objective. The fifth rule tries to eliminate the effects ( $e_3$ ) of an intention ( $e_2$ ) from being considered as intentions or objectives. If *John intends to swim in the pool* ( $e_2$ ) even if he knows that he is going to catch a cold ( $e_3$ ) because the water is too cold, we cannot say that *John intends to catch a cold*.<sup>5</sup> The traditional relational properties (*reflexivity*, *symmetry*, or *transitivity*) do not hold for the INTENTION semantic relation.

### 3 Learning Model

#### 3.1 Experimental data

We applied the most frequent syntactic pattern that expresses intentions in text ( $VB_1$  to  $VB_2$ ) on the first 10,000 sentences of the SemCor2.0 collection and we extracted 1,873 sentences. These sentences contain 115 intentions (manually identified by a graduate student, not the author). The data consisting of these positives and 258 arbitrarily selected negative examples, was randomly divided into a training set that contains 80% of the examples and a test set with the remaining 20% instances. The statistics are shown in Table 2.

	Intentions	Non-Intentions	Total
Training	92	208	300
Testing	23	50	73

Table 2: Experiments Data Division

semantic relations.

<sup>4</sup>  $e_1$  and  $e_4$  represent different intentions of the same person.

<sup>5</sup> A more detailed example can be found in (Bratman, 1990).

#### 3.2 Features for intention

After analyzing our training data, we pinpointed a set of features to help us identify the *intentions* encoded by the pattern  $VB_1$  to  $VB_2$ . The WordNet senses needed to extract the semantic features were taken from SemCor. We will use *Mary intends to revise the paper* to show each feature’s value.

**The semantic class of the the  $VB_1$  verb’s agent or specializations of it.** Intentions and objectives are specific to humans. Thus, the semantic class of the  $VB_1$  agent bears a high importance. We used an in-house semantic parser to retrieve the AGENT of the  $VB_1$  verb. The feature’s value is its WordNet semantic class. *Mary* names a person. Thus, the semantic class that we are seeking is *entity#1*.

We chose this semantic generalization because nouns and verbs belong to open part-of-speech classes. There can be an enormous number of possibilities and any models built using them as feature values will not be able to generalize beyond the training examples. Therefore, we introduce a bias in our learning framework based on the assumption: noun and verb concepts will semantically behave as the concepts that subsume them in the WordNet structures. But, by generalizing concepts, we lose some of their semantic properties. Hence, we specialize the semantic class  $s$  of a concept  $w$  by replacing it with its immediate hyponym ( $h$ ) that subsumes  $w$ . We can further increase the semantic level by specializing  $h$ . We note that the number of values is still finite even though we specialized the general concepts. As the specialization level increases, there will be words  $w$  that cannot be further specialized (*entity#1* cannot be specialized even once). In such cases, we add  $w$  to the set of feature values.

**The semantic class of the  $VB_1$  verb or its specializations.** The intention phrase is subordinated to a verb ( $VB_1$ ). The semantic class of this verb is the system’s second feature. In our example,  $VB_1$  (*intend#1*) semantic class is *wish#3*.

**The semantic class of the  $VB_2$  verb’s agent,** if this agent differs from the  $VB_1$  verb’s agent; otherwise, a common value (*equal*) is given. We identify the AGENT of the  $VB_2$  verb. The specializations of its semantic class will be used if the top noun proves to be too general. In the sample sentence, the agent of *revise* is *Mary*. We can have a different agent for

Semantic class of the $VB_1$ 's agent	Semantic class of the $VB_1$ verb (%)								
	no specialization			$1^{st}$ level of specialization			$2^{nd}$ level of specialization		
	Semantic class of the $VB_2$ verb			Semantic class of the $VB_2$ verb			Semantic class of the $VB_2$ verb		
	no spec.	$1^{st}$ level	$2^{nd}$ level	no spec.	$1^{st}$ level	$2^{nd}$ level	no spec.	$1^{st}$ level	$2^{nd}$ level
no spec.	87.67	80.82	87.67	<b>90.41</b>	87.67	87.67	86.30	83.56	84.93
$1^{st}$ level	89.04	82.19	87.67	87.67	89.04	87.67	87.67	86.30	84.93
$2^{nd}$ level	87.67	83.56	87.67	<b>90.41</b>	<b>90.41</b>	89.04	89.04	87.67	86.30

Table 3: Accuracy of models using the  $2^{nd}$  specialization level for the  $VB_2$  agent semantic class

the  $VB_2$  verb (*Mary intends John to revise the paper*). Let's assume that Mary is John's supervisor and she can make him revise the document. The sentence expresses Mary's intention of persuading John to revise the paper, but this objective is not encoded by the pattern we considered.

**The semantic class of the  $VB_2$  verb or its specializations.** The  $VB_2$  verb expresses the future action or behavior that the agent intends. We extract this feature using WordNet hierarchies. *Revise#1* belongs to the *act#1* semantic class.

**A flag indicating if the  $VB_1$  verb has an affirmative or a negative form.** We want to differentiate between sentences like *John wants to go for a walk* and *John doesn't want to go for a walk*. The first sentence expresses John's intention, while, in the second one, no intention can be identified.

**The type of the analyzed sentence.** This feature is primarily concerned with questions. A question like *Where do you plan to go for a walk?* indicates the intention of *going for a walk*, unlike the question *Do you plan to go for a walk?* which might express an intention if the answer is "yes". This feature's values are the *wh-words* that begin a question or *n/a* for the other types of English sentences.

We did not analyze the affirmative versus the negative form of the  $VB_2$  verb because it does not affect the objective attribute of the intention. The sentence *John intends not to go for a walk* expresses a negative intention. This sentence is much stronger than *John doesn't intend to go for a walk*. In the former context, John has set a goal for himself, while in the second sentence, the objective does not exist.

## 4 Experimental Results

### 4.1 Impact of specialization

The first experiment was performed using the LIBSVM package<sup>6</sup> and the WordNet semantic classes.

<sup>6</sup><http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html>

These features yield an accuracy of 87.67%. Trying to improve the performance, we specialized the semantic classes. When the  $VB_2$ 's agent semantic class was specialized, the accuracy remained constant. If we replace the  $VB_2$ 's semantic class with its direct hyponyms, the accuracy drops 5.48%. But, the specialization of the  $VB_1$  agent's semantic class brings an improvement of 1.37% and the specialization of the  $VB_1$ 's class produces an increase in accuracy of 2.74%. Given this fluctuation in performance, we performed 81 different experiments which create SVM models using the same training data annotated with more general or more specific feature values. For each feature, we analyzed the first two semantic specialization levels.

From our experiments, we noticed that the specialization of the  $VB_2$ 's agent semantic class does not influence the performance. Out of the 27 experiment triplets in which this specialization level changes, in only 4, it influences the result and, in 3 of them, the accuracy increases with the specialization level. Thus, our third feature is the second specialization level of the  $VB_2$ 's agent class. Table 3 shows the results obtained when the values of the radial kernel parameters were chosen to optimize the 5-fold-cross-validation on the training data. The best models are described in Table 4.

Model	Level of specialization for the features
A	semantic class of the $VB_1$ agent, $1^{st}$ level of specialization for the $VB_1$ 's semantic class, and semantic class of the $VB_2$ verb
B	$2^{nd}$ semantic level for the $VB_1$ agent class, $1^{st}$ level of the $VB_1$ 's semantic class, and the semantic class of the $VB_2$ verb
C	$2^{nd}$ level of the $VB_1$ agent's semantic class and $1^{st}$ specialization levels for the $VB_1$ and $VB_2$ semantic classes

Table 4: The best three intention classifiers

### 4.2 Learning curves

We further analyzed our data and models and tried to see how many training examples are needed to reach 90.41% accuracy. We varied the training data

	Semantic class of the $VB_1$ 's agent	Semantic class of the $VB_1$ verb	Semantic class of the $VB_2$ 's agent	Semantic class of the $VB_2$ verb	$VB_1$ verb form	Sentence type
<b>Model A</b>	2.74	16.44	1.37	0	2.74	4.11
<b>Model B</b>	2.74	15.07	1.37	0	4.11	2.74
<b>Model C</b>	1.37	16.44	4.11	0	4.11	2.74

Table 5: The improvement (%) brought by each feature to the three best SVM models

size and validated the new models using our previous test set. Figure 1 shows the performance variation of three models that use feature sets identical in terms of specialization levels to the ones of the A, B, and C classifiers. All three models exhibit a similar behavior with respect to the change in the training set size. Therefore, our features create a stable algorithm. The highest accuracy models use all 300 training examples. Thus, we did not reach the saturation point, but, considering the performance curve, this point is not very far.

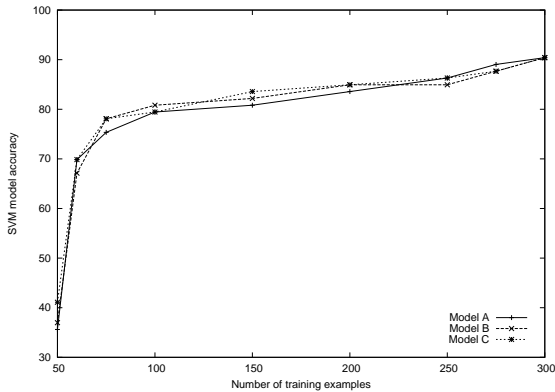


Figure 1: Testing set is constant

### 4.3 Feature impact on the SVM models

All our previous experiments used the entire set of features. Now, we investigate the relative contribution of each feature. We performed experiments that use only five out of the six features. In Table 5, we list the accuracy increase that is gained by the inclusion of each feature. The most influential attribute is the  $VB_1$  verb's semantic class or its specializations. The intention's description verb does not influence the classification result. Because intentions consist of a future action and verbs express actions, there are very few verbs, such as *dream* or *snore* (involuntary actions) that cannot occupy the  $VB_2$  verb's position. The syntactic features bring an average increase in accuracy of 3.50%.

### 4.4 Impact of word sense disambiguation

Perfect word sense disambiguation might be a too strong assumption. In this section, we examine the effects of weaker disambiguation. Table 6 shows the accuracies of the best three models when each concept is tagged with its first WordNet sense (**No WSD**) and when the senses are given by an in-house WSD system with an accuracy of 69% computed on the SemCor data (**Automatic WSD**).

	No WSD	Automatic WSD	Gold WSD
<b>Model A</b>	72.60	79.45	90.41
<b>Model B</b>	73.97	79.45	90.41
<b>Model C</b>	72.60	80.82	90.41

Table 6: Best models performance (%)

### 4.5 C5 results

After examining the SVM results, we applied the C5 machine learning algorithm (Quinlan, 2004) to the same training data annotated with the same feature set, in a similar manner. Again, we specialized the four semantic classes, independently, and tested the decision trees against the testing data. Table 7 shows their accuracy. The highest values were obtained for the first level of specialization of the  $VB_1$  verb semantic class. The specialization levels of the other semantic classes do not influence the accuracy of the decision trees. The most tested attribute is the  $VB_1$  verb. This further substantiates our observation, made during our SVM models analysis, that this feature has the greatest importance in the intention classification process. Our error analysis of the C5 results indicates that, because of the relatively small numbers of training instances, C5 ignores some of the features and makes wrong decisions.

## 5 Application to Question Answering

Questions involving intentions cannot be answered only by keyword-based or simple surface-level matching techniques. Table 8 lists two questions for

$Q_1$ :	What was Putin trying to achieve by increasing military cooperation with North Korea?
$QLF_1$ :	Putin( $x_1$ ) & INT( $e_0, x_1, X$ ) & ANS( $X$ ) & MANNER( $X, e_2$ ) & increase( $e_2, x_1, x_2$ ) & military( $x_2$ ) & cooperation( $x_2$ ) & with( $x_2, x_3$ ) & North-Korea( $x_3$ )
$A_1$ :	Putin is attempting [to restore Russia's influence in the East Asian region][INT]. The report said, the possibility remains that Russia could increase military cooperation with North Korea based on their treaty.
$ALF_1$ :	Putin( $x_1$ ) & INT( $e_0, x_1, e_1$ ) & restore( $e_1, x_1, x_2$ ) & Russia( $x_3$ ) & 's( $x_3, x_2$ ) & influence( $x_2$ ) & LOCATION( $x_2, x_4$ ) & East( $x_4$ ) & Asian( $x_4$ ) & region( $x_4$ ) & report( $x_5$ ) & say( $e_2, x_5, e_3$ ) & possibility( $x_6$ ) & remains( $e_3, x_6, e_4$ ) & increase( $e_4, x_3, x_7$ ) & military( $x_7$ ) & cooperation( $x_7$ ) & with( $x_7, x_8$ ) & North-Korea( $x_8$ ) & base( $e_4, x_9$ ) & treaty( $x_9$ )
$Q_2$ :	From where does al Qaeda intend [to purchase weapons of mass destruction][INT]?
$QLF_2$ :	alQaeda( $x_1$ ) & INT( $e_0, x_1, e_1$ ) & ANS( $X$ ) & LOCATION( $e_1, X$ ) & purchase( $e_1, x_1, x_2$ ) & weapons_of_mass_destruction( $x_2$ )
$A_2$ :	It is known that Osama bin Laden's al Qaeda network has tried [to buy ingredients for weapons of mass destruction in Russia][INT].
$ALF_2$ :	Osama_bin_Laden( $x_1$ ) & 's( $x_1, x_2$ ) & al-Qaeda( $x_2$ ) & network( $x_3$ ) & IS-A( $x_2, x_3$ ) & INT( $e_0, x_2, e_1$ ) & buy( $e_1, x_3, x_4$ ) & ingredient( $x_4$ ) & PURPOSE( $x_4, x_5$ ) & weapons_of_mass_destruction( $x_5$ ) & LOCATION( $e_1, x_6$ ) & Russia( $x_6$ )

Table 8: Question and answer pair examples

Semantic class of the $VB_1$ 's agent	Semantic class of the $VB_1$ verb		
	no spec.	1 <sup>st</sup> level	2 <sup>nd</sup> level
no spec.	79.45	<b>87.67</b>	84.93
1 <sup>st</sup> level	68.49	<b>87.67</b>	84.93
2 <sup>nd</sup> level	79.45	<b>87.67</b>	84.93

Table 7: C5 models accuracy (%)

which finding the correct answer primarily depends on the discovery of the INTENTION relation.

The answer type for the question  $Q_1$  is the INTENTION argument itself. The question processing module will detect that the answer being sought is Putin's intention. The semantic relations module processes  $A_1$ 's text and discovers the INTENTION relation. The question is searching for the intent of Putin with regards to North Korea and the answer text reveals Putin's intention to restore Russia's influence in the area. Question  $Q_2$  is searching for a location as its answer type and the correct answer is one which involves al Qaeda intending to purchase weapons of mass destruction. The candidate answer text ( $A_2$ ) reveals the organization's past intent to buy (synonym with *purchase*) weapons in Russia. Because the two intentions have the same agent, future action and theme, the two semantically enhanced logic forms can now be unified and we can pin down the location of the intent (*Russia*).

## 6 Conclusions

We proposed a method to detect the INTENT relation encoded by the sentence-level pattern  $VB_1$  to  $VB_2$  with a 90.41% accuracy. We plan to investigate the other INTENTION patterns as well as other semantic relations such as MOTIVE, IMPLICATION, or MEANING which, currently, cannot be identified by the state-of-the-art NLP systems. These relation-

ships need to be analyzed to provide a complete coverage of the underlying semantics of text documents. We intend to incorporate our INTENTION detection module into a Question Answering system and show its impact.

## References

- Anscombe, G.E.M. 1957. *Intention*. Cornell University Press, Ithaca, New York.
- Audi, Robert. 1973. Intending. *The Journal of Philosophy*, 70(13):387-403.
- Bratman, Michael E. 1981. Intention and means-end reasoning. *The Philosophical Review*, 90(2):252-265.
- Bratman, Michael E. 1987. *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, Massachusetts.
- Bratman, Michael E. 1990. What is intention? In *Intentions in Communication*. MIT Press.
- Miller, George A., Claudia Leacock, Randee Teng, and Ross T. Bunker. 1993. A semantic concordance. In *Proceedings of the ARPA Human Language Technology Workshop*
- Miller, George A. 1995. Wordnet: A lexical database. *Communication of the ACM*, 38(11):39-41.
- Pollack, Martha E. 1990. Plans as complex mental attitudes. In *Intentions in Communication*. MIT Press.
- Quinlan, Ross. 2004. Data Mining Tools See5 and C5.0. <http://www.rulequest.com/see5-info.html>
- Wiebe, Janyce M., Theresa Wilson, Rebecca F. Bruce, Matthew Bell, and Melanie Martin. 2004. Learning subjective language. *Computational Linguistics*, 30(3):277-308.