Automatic Parody Detection in Sentiment Analysis

Sarah Bull

December 16, 2010

Abstract

Sentiment analysis is here defined as a machine learning problem to analyse human documents and extract the human opinion they convey. Kanayma *et al* describe the field as: 'a task to obtain writers' feelings as expressed in positive or negative comments, questions, and requests, by analysing large numbers of documents' [H Kanayama and Watanabe, 2004]. Much of the work on automated sentiment analysis is relatively recent and has focused upon explicit sentiment such as 'I like' or 'I hate'. This form of sentiment can be analysed using simple lexicons of positive or negative words and phrases. Little work in sentiment analysis focuses on more complex domains of sentiment such as parody and sarcasm which require the use of more machine learning techniques. This thesis contributes to the literature on sentiment analysis firstly by providing an overview of the field, and secondly by practical experiments and feature engineering in the specific sentiment domain of parody versus non parody. The experiments were carried out upon corpus data of genuine news articles and parody articles. The thesis has demonstrated that it is possible to distinguish parody from serious news and contributes important features that can be used for this purpose. Machine learning can be used in the complex domain of sentiment analysis of parody.

Contents

1 Introduction

1.1 Sentiment Analysis

Sentiment analysis is the task of obtaining the general feeling or opinion from a piece of text. Simple sentiment analysis can examine whether the sentiment of a text is positive or negative, such as whether the reviewer of a product liked or disliked it, or whether a comment on an email list is abusive or pleasant. For simple sentiment analysis, lexicon-based approaches such as [Nasukawa and Yi, 2003] and [Spertus, 1997] use lists of known phrases to identify sentiment. Parody analysis is a more complex domain of sentiment analysis than positive or negative sentiment.

Analysis of parody limits the field to a particular form of sentiment: whether a text is serious or humorous by intent. Parody analysis is important as part of the field of sentiment analysis. It is an interesting problem to learn to automate because it is a complex human task based on many variables. At times, humans have mistaken serious articles for parodies. For example, in 2000 the parody website The Onion published an article claiming that the Harry Potter series of children's books promoted Satanism; this was taken seriously by several religious groups in spite of quotes such as " 'Harry is an absolute *god*send to our cause,' said High Priest Egan of the First Church of Satan" (emphasis not in original). [Mikkelson and Mikkelson, 2008]

There exist many corpuses of both serious news and parodies, and the ability to replicate a human's ability to distinguish between them is a part of the field of sentiment analysis that has little attention so far. Because parodies are often of specific events that have appeared in the news, the serious corpuses examined are primarily of newspaper articles so as to contrast material of similar subject matter but different sentimental orientation. The parody corpuses examined are from humorous websites which comment upon current events.

To automatically extract sentiment as serious or non-serious material, features must be defined. These features in the experiments shown are vocabulary-based features that exploit trends in language that show a distinction between parody and non-parody. This thesis contributes to the literature by explaining features discovered by automatic extraction as well as human analysis. In the human analysis of parody versus non-parody, besides language features, other observations such as sentence length and structure may be relevant in separating serious material from parody.

1.2 Structure of Thesis

The subject is introduced with definitions of parody and background in linguistic theory, to define the task and introduce technical terms. The third section grounds the research in the field of sentiment analysis by explaining prior work and comments on the future development of sentiment analysis in the domain of parody detection. The fourth section contrasts and analyses serious and parody articles as a human to prepare for discussion on feature engineering for the machine. The fifth section covers the expermiments carried out with notes on their results, in which non-serious material from particular corpuses was distinguished with success from serious material from particular corpuses, and discusses feature engineering for future practical use. The final section provides notes for future work and conclusions.

1.2.1 Hypotheses

The principal hypothesis of the thesis is that parody and serious texts can be automatically detected by the use of machine learning techniques. This hypothesis is supported by high accuracy results shown in Section 5 on selected corpuses. The work on human analysis of parody versus serious news is included as supporting evidence that a difference exists and is detectable. A list of hypotheses examined in Sections 4 and 5 follows:

1. That parody and serious texts can be automatically detected by the use of machine learning techniques.

2. That parody texts have a longer average word and sentence length than serious texts.

3. That parody is personalised with use of stopwords such as 'we' and 'I'.

4. That parody texts have more informal language than serious texts (especially serious news texts as against serious blogging texts).

5. That parodies and serious news show different trends in subject material.

6. That smaller corpuses with a smaller vocabulary compared to larger corpuses with a larger vocabulary cause a Naive Bayes classification bias toward smaller corpuses because word features of the same rank have higher frequency weightings in the smaller corpus.

7. That adding recent news as well as older news to the serious corpus will increase accuracy in classifying new articles as serious or non-serious.

8. That, when the older news corpus is significantly larger than the recent news corpus and both are combined as a serious news corpus and tested against a non-serious corpus of recent parodiess, Naive Bayes classification error will primarily be on test data drawn from recent news misclassified as parody.

9. That a Naive Bayes classifier will return results of lower accuracy when recent news articles are compared to recent parody texts rather than older news compared to recent parody texts.

10. That punctuation-based features that indicate contractions and sentence structure are useful in classifying serious and non-serious texts as well as word features.

1.3 Defining Serious: Parody, Satire and Sarcasm

This section defines the variety of non-serious material that contrasts to serious articles. In broad terms, *parody*, *satire*, and *sarcasm* can all be considered *non-serious* and *humorous* material. These terms need to be defined in detail to describe the process of extracting them as sentiment learning, as well as the distinction of *formal* and *informal* language.

The phrase *non-serious* is a 'marked' phrase [Battistella, 1990], where an adjective is added to a word to change the meaning. *Seriousness*, in the context of the experiments performed, is a concept limited to *current events as reported in news (and some blogging)*, although other work on sentiment analysis has included the *serious* material of congressional transcripts [Thomas et al., 2006] and other legal material. *Non-serious* material is a broader range than humour: *non-serious* could include anything from humorous articles mocking conspiracy theories to a blogger writing about their pet dog.

The term *parody*, strictly speaking, refers to text that exchanges words but closely adheres to the text and style of the original material. For example, Dr Seuss writes:

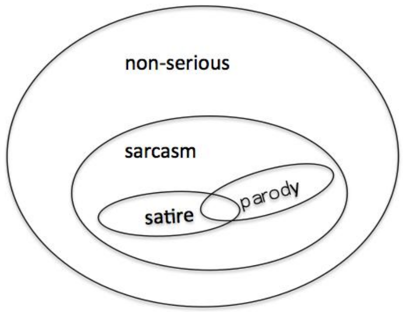
> Do you like green eggs and ham?
>
> I do not like them, Sam-I-am.
>
> I do not like them in a box.
>
> I do not like them in a fox...
>
> I do not like green eggs and ham.
>
> I do not like them, Sam-I-am.

[Seuss, 1960]

In parody (which the webpage housing it actually termed 'satire'), an anonymous writer circulated the following composition relating to the OJ Simpson murder case in the United States (*People v. Simpson*).

> Did you hit her from above?
>
> Did you drop that bloody glove?
>
> I did not hit her from above
>
> I cannot even wear that glove...
>
> I did not do this awful crime
>
> I could not, would not anytime.
>
> And now I'm free, I can return
>
> To my house for which I yearn
>
> And to my family whom I love
>
> Hey now I'm free - give me back my glove!

Figure 1: Non-serious material: sarcasm, parody, and satire.



[University of Missouri Kansas City, 2010] *Parody* has a specific legal meaning in terms of copyright law: a similar poem was the subject of the United States trial *Dr Seuss Enterprises v. Penguin Books* (109 F.3d 1394). Parody is defined as separate from satire: parody 'needs to mimic an original to make its point, and so has some claim to use the creation of its victim's...imagination, whereas satire can stand on its own to feet and so requires justification for the very act of borrowing' (109 F.3d 1394 23).

As legally cited above, *satire* presents a more general exaggeration of a situation. For example, a satire might describe the incident of an il spill caused by a corporation as 'The *company* is being victimised by the *wildlife*!'.

Sarcasm in English originates from the Greek term *sarkazein*, which means to 'tear flesh' (New Oxford American dictionary). Sarcasm can be seen as an umbrella term for both 'parody' and 'satire'. Essentially, sarcasm is defined as the use of irony: expressing meaning by using exaggerated language that signifies the opposite. Hence the terms *non-serious* and *sarcasm* are here used to refer to the same material: sarcastic, ironic articles related to current events that use any combination of parody, satire, and sarcasm. The diagram in Figure 1 shows the concepts of non-serious material used here.

It is also necessary to note not only the distinction between serious and parody material, but the distinction between formal and informal language. *Formal* language is here used to mean language used in an edited publication such as a newspaper, written to a standard such as the Associated Press Style Guide [Press, 2010]. Newspaper articles are both serious and use formal language. Articles intended as parodies often, but not invariably, use informal language. Blog entries use informal language but can be both serious and informal. Though the corpuses of serious articles used were largely sourced from newspapers, serious but informal blog entries are also discussed from both the human's point of view and the results from the use of the algorithms. This data is important to ascertain whether the machine learning analyses serious versus parody or formal

Figure 2: The four combinations of serious, non-serious, formal, and informal.

	serious	non-serious
formal	news	some fiction, some parodies
informal	blogging	parodies

versus informal language. This work seeks to distinguish between serious news and opinion articles and non-serious sarcastic articles.

Figure 2 shows the division of formality and seriousness: news articles are formal and serious; blogging about news is informal and serious; parodies are typically informal and non-serious; and some fiction and some parodies use formal language whilst being non-serious.

2 Background in Linguistic Theory

2.1 Introduction

This section of the thesis is an introduction to the study of language. The background and terminology of linguistic theory is important groundwork for the sentiment analysis of serious versus parody differentiation. This section provides a basic technical vocabulary in studying language and gives examples of the concepts as applied to sentiment analysis. Parody is encoded in both the syntactic (grammatical) and the semantic (vocabulary) structure of language: for example, the sentence 'Hey, nuclear warfare would be totally awesome' is a sarcastic sentence because it syntactically applies the phrase 'totally awesome' to 'nuclear warfare', and because it uses the informal words 'Hey' and 'totally', therefore making it informal.

2.2 Linguistic theory

For more information, the reader is referred to [Manning and Schutze, 1999]. Human language is non-random: for an example of statistical significance, Zipf's law is the principle that there exists a constant k such that the frequency of a word multiplied by its rank (ie. list position ordered by frequency) is equal to k:

$$f \cdot r = k \tag{1}$$

There is a consistent relationship between the frequency of a word and its importance to a text. In language processing, sometimes the approach is taken to trim the most frequent words that are effectively meaningless in predicting the sentimental orientation of a document. Stopwords, short and common words such as 'the', 'we', and 'over', are a common example of words that are removed for language processing. Language has high information content and low entropy: entropy is defined as the degree of information in a random variable, and a low entropy indicates a lower randomness. Words are derived from other words with a clear relationship between the two: for example, 'weak' and 'weaken' ('weak' is an example of a word 'stem'). Words are also compounded to give a meaning that combines two words, for example, 'disk drive'. Collocations are expressions consisting of two or more words that are frequently statistically joined together, such as 'strong tea', or 'rich and powerful'. [Manning and Schutze, 1999] Some collocations are more common in works of certain sentiment than others: for example, unusual collocations are more likely to be part of informal or creative writing rather than standard newspaper articles.

When words occur in adjacent positions, these are referred to as n-grams: bigrams (two words next to each other), trigrams (three words), and so on. Words can be classified into 'parts of speech' that explain the role each word has in a sentence. The principal terms are depicted below.[Manning and Schutze, 1999].

Table 1: Parts of Speech

Part of Speech	Description	Examples
Nouns	Nouns refer to persons, animals, concepts, places, and things.	Napoleon, tiger, the Eiffel tower, authoritarianism
Adjectives	Adjectives describe the properties of nouns.	clever, fierce, tall, strange
Pronouns	Pronouns refer to persons or things that are relevant in the context. They may be accusative, possessive, second person possessive, or reflexive.	I, you, he, she, it, we, they. Accusative forms: me, you, him, her, it, us, then. Possessive forms: my, your, his, her, its, our, your, there. Second person possessive forms: mine, yours, his, hers, its, ours, theirs. Reflexive forms: myself, yourself/yourselves, himself, herself, itself, ourselves, themselves.
Determiners	Determiners describe the particular reference of a noun; articles, such as 'the', are a subclass of determiners.	the, a, an, this, that
Comparatives	Adjectives that compare degrees of a property.	richer, happier, healthier
Superlatives	Adjectives that describe the highest degree of a property.	richest, happiest, healthiest
Verbs	Verbs describe actions, activities, and states. Note verb tenses and the use of auxiliary, or 'helping' verbs.	eat, walk, feed, have (auxiliary), to be (auxiliary), will (auxiliary).
Adverbs	Adverbs modify verbs and specify place, time, manner, or degree.	quickly, slowly, pleasantly, aggressively, often.
Prepositions	Small words that usually express a relation from one noun to another.	on, after, in, over, under.
Conjunctions	Conjunctions connect clauses or sentences together.	and, but, or, however

A sentence is a group of words of different parts of speech. Sentences usually end with a period ('.'), though periods are also used to show abbreviations such as 'ie.' or 'etc.'. Sentences can also end with question marks and exclamation marks. English sentences can be broken down into phrases, which use the different parts of speech in typical combinations. A noun phrase is a phrase which expresses information about a noun (eg. 'The red ball'); a prepositional phrase is headed by a preposition and contains a noun phrase (eg. 'Under the green chair'); and a verb phrase consists of a verb and the elements that depend upon it (eg. 'Throw it away') [Manning and Schutze, 1999]. Epithets are short, insulting phrase (such as 'Get a life!') and are an indication of an overall negative comment [Spertus, 1997]. Words can be marked or unmarked; for example, 'good' is unmarked by itself, but if it is in a phrase 'not good', then it is marked. Typically, marked words are negative.

[Battistella, 1990] [Lehrer, 1985] Sentences can be broken into 'trees' of phrases and word types. There are a number of ways of drawing trees, and one example is dependencies. The concept of dependencies is used to explain the relation between parts of a sentence. For example: 'John and Ellie went to the cinema in the next town'. John and Ellie are dependencies of the event of going to the cinema, and 'in the next town' modifies 'cinema'. Identifying dependencies and phrases can be difficult. Attachment ambiguities occur when the same sentence could be diagrammed in more than one way. For example, 'Clothes made of hemp and smoking paraphernalia were on sale' (what were the clothes made of?). [Buntine, 2009] Dependencies are important to sentiment analysis future work. It is difficult to program to accuratelly map dependencies in sentences, but to understand the subject and object of a sentence is an important facet of sentiment learning. For example, 'The best part of reading the book was throwing it across the room': this is a sarcastic sentence, where the noun phrase 'best part' is applied to 'throwing across the room'. A dependency diagram is presented in Figure 3.

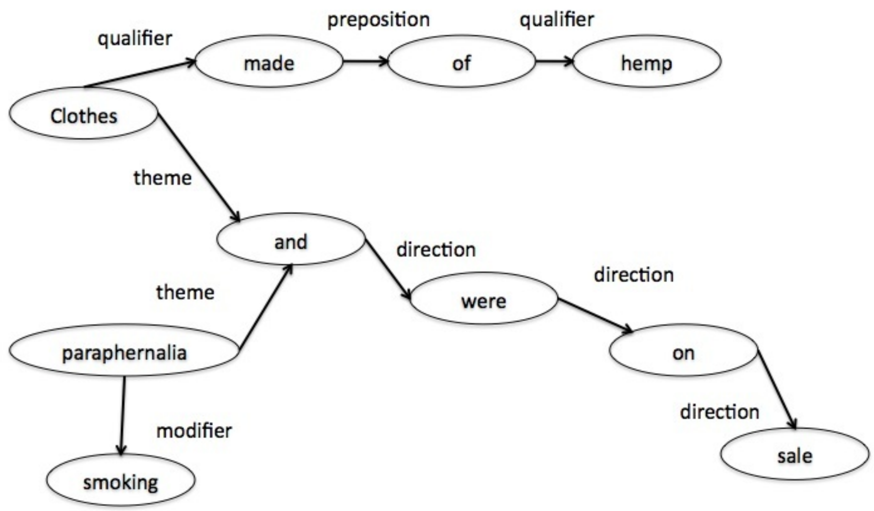
Figure 3 correctly parses the sentence 'Clothes made of hemp and smoking paraphernalia were on sale'. 'Clothes' and 'paraphernalia' were separately on sale; 'smoking' modifies 'paraphernalia'; and 'made of hemp' describes the clothing. Ideally, automated dependency parsing tools ought to produce this correct result, but that does not always occur. See [Buntine, 2009].

2.3 Naive Bayes in document classification

The simple Bayesian rule is:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

Figure 3: Tree diagram of a sentence



The posterior probability can be found by multiplying the likelihood by the prior and dividing by the normalising constant.

[Nigam et al., 2000] used the Naive Bayes model in experiments to classify documents based on the bag-of-words approach. A Naive Bayes classifier uses the training set to analyse word frequencies and classify new documents based on the log-likelihood of the type of document. In experiments in Section 5, documents were classified by Naive Bayes into non-serious versus serious, where 'non-serious' indicates a parody and 'serious' indicates genuine news. For Bayesian priors in all tests, equal prior likelihood of parody versus genuine was assumed even though serious articles in general outnumber parody articles in general. This was to avoid possible distortion via the size of the training datasets.

In Naive Bayes, the central concept is that the probability of a given document D given a class C, based on the word distribution within the document, is:

$$p(D|C) = \prod_i p(w_i|C)$$

The question asked is, given a new document:

$$p(C|D) = \frac{p(D \cap C)}{p(D)}$$

$$p(C|D) = \frac{p(C)}{p(D)} p(D|C)$$

In a binary class of S (serious) and ˜S (non-serious), this is:

$$p(D|S) = \prod_i p(w_i|S)$$

$$p(D| \sim S) = \prod_i p(w_i| \sim S)$$

When divided, these become:

$$\frac{p(S|D)}{p(\sim S|D)} = \frac{p(S)}{p(\sim S)} \prod_i \frac{p(w_i|S)}{p(w_i| \sim S)}$$

Then it is possible to refactor the equation by logarithm as:

$$\ln \frac{p(S|D)}{p(\sim S|D)} = \ln \frac{p(S)}{p(\sim S)} + \sum_i \ln \frac{p(w_i|S)}{p(w_i| \sim S)}$$

When this is negative, this then indicates a non-serious result; when positive, it indicates a serious result. [McCallum and Nigam, 1998]

3 Theory of Sentiment

3.1 Introduction to Sentiment Learning and Parody

The goal of sentiment analysis by machine learning is to take an object of text and summarise it by overall sentiment: whether by simple binary or ternary metrics, or by more complex measurements on multiple axes that give a more detailed result.

This section discusses sentiment analysis as simple binary and ternary outcomes, and work on sarcasm that examines parody versus serious. More complex outcomes such as [Osgood et al., 1957] and affect and appraisal theory represent recent developments in the work of sentiment analysis, and they have potential to be applied to future work in the domain of parody detection. Finally, this section provides an overview of the field of sentiment analysis through a table depicting some of the many databases available in different sentiment domains and the research carried out upon them.

3.2 Simple and Complex Classification Outcomes

3.2.1 Binary and Ternary work in machine learning

Much work on sentiment clarifies into simple binary or trinary outcomes, such as 'positive' and 'negative' and 'parody' and 'non-parody'.

[Spertus, 1997] built a decision tree program to separate email communication related to politics into 'flame' and 'non-flame' categories, where 'flame' indicates an abusive comment of no value (eg. '$%@# you, you $%^&#'). The features used that indicated a flame were: inappropriate language; noun phrases modified with the pronoun 'you' (eg. 'you people', 'you fools'); imperative statements (commands such as 'get over it', 'get a life'); usage of second person (eg. 'this so-called novel of yours'); condescending phrases on the end of sentences (eg. 'you really can't fix this, can you?'); well-known negative words (eg. 'bad', 'awful'); and well-known short epithets (eg. 'get a clue'). The features that indicated a non-flame were politeness (eg. 'thanks', 'please') and standard praise words (eg. 'love', 'great'). These features show some of the differentiation standards that can also be applied to serious versus parody analysis: the features of punctuation modification, imperative sentences or short epithets, and so on.

[Hatzivassiloglou and McKeown, 1997] limited a study to the orientation of adjectives, using the Wall Street Journal corpus with part of speech tagging completed, predicting for 'positive' or 'negative' (eg. that the adjective 'horrible' is negative). This is useful work as an illustration of the use of statistical methods to classify language in the field of sentiment analysis. However, this orientation of adjectives is insufficient for parody analysis: sarcastic texts use the opposite orientation to give a true intended meaning, for example: 'It is ever so nasty and horrid and dreadful to be nice to people, isn't it?' However, lexical information of the orientation of adjectives is useful in sentiment analysis and ought to be a feature of future work in parody analysis.

In ternary classification, [Das and Chen, 2001] used Bayesian features of words to detect the advice of 'buy'/'neutral'/'sell' in the domain of finance, analysing opinions of the states of the markets based on messageboard postings. [Dave et al., 2003] experimented upon star rankings in reviews from websites including Amazon.com, trying to replicate the graduations of reviews. This is a complex matter because human reviewers do not always review with internal consistency in their star rankings, and ranking standards differ from person to person. For the case of parody versus non parody, there is no such source of ranked data that (for example) grades from a standard news article, to an opinion piece, to a sarcastic opinion piece, and finally to an actual parody article.

3.2.2 Sarcasm Work

On work that differentiates between sarcasm and non-sarcasm, [Utsumi, 2000] begins with a description of verbal irony. Ironic effect is expressing a phrase opposite to the situation: for example, 'This room is totally clean', when the room is very dirty. [Utsumi, 2000]. Utsumi establishes that sarcasm is contextual; as discussed later, sarcasm is reactive to genuine news. [Cheang and Pell, 2008] also explore the textual and acoustic features of the detection of sarcasm, primarily as acoustics. [Kreuz and Roberts, 1995] describe three factors for textually expressed cues for sarcasm:

- Excessive adjectives and adverbs (which are also quite informal in language and not within standard news style)
- Interjections such as 'Gosh' or 'Gee' (also informal)
- The use of punctuation such as exclamation marks

[Kreuz and Caucci, 2007] have showed that sarcasm can also be recognised by humans independently of context and acoustics. In the experiments, students were given fictional excerpts where the phrase 'said sarcastically' was removed from the dialogue tags of a character's speech. Recognising sarcasm or non-sarcasm is a possible field for machine learning to continue to explore.

The work of [Tsur et al., 2010] carries out experiments to detect sarcasm on Amazon reviews with success. The results in question automatically tag individual sarcastic sentences with a precision of 77% and a recall of 83.1%. [van Rijsbergen, 1979] Examples given of sarcastic reviews in the domain include:

- 'I love the cover!'
- 'Trees died for this book!'
- 'Be sure to save your purchase receipt!'
- 'Great for insomniacs'

[Tsur et al., 2010] These examples are sarcastic because they praise aspects inappropriate for the domain, ie. ‘insomniacs’, or ‘the cover’ without accompanying verbiage also praising the contents of the book. Features used included punctuation-based features such as exclamation marks, quotes used, capital letters used, and pattern recognition based on sentence structure. The training data set consisted of of 471 positive examples and 5020 negative examples. In classification, the K nearest neighbour strategy was used to assign scores to new examples.

The results obtained for this thesis in Section 5 on automatically tagging serious versus non-serious articles as a whole show an improvement on [Tsur et al., 2010] in that, when testing on material from the same corpus as the training data, nearly 100% accuracy is achieved. The relevance of sentence and word length to parody versus serious articles is discussed, and punctuation-based features also explored. These experiments and analysis of features contribute by building upon previous work in this relatively unexplored aspect of sentiment analysis.

3.2.3 Complex work

The seminal work *The Measurement of Meaning* [Osgood et al., 1957] completed extended surveys that described the sentimental reaction of the human to various words across eight factors.

- Evaluation - measures the positive or negative scale of a word, such as ‘good’ or ‘bad’, ‘optimistic’ or ‘pessimistic’
- Potency - measures the power of a word, eg. ‘pacifist’ is low on potency
- Oriented Activity - how active or passive a word, eg. ‘fast’ is active, ’inactive’ is passive
- Stability - words that relate to stable conditions, eg. ‘sober’ is stable, ‘drunk’ is unstable
- Tautness (tension) - this factor was tentatively labelled by Osgood (1957) and reflects scales such as ‘angular’ versus ‘rounded’, ‘straight’ or ‘curved’
- Novelty - how new or old a word, eg. ‘unusual’ is novel and ‘usual’ is not novel
- Receptivity - sensual receptiveness, eg. ‘pungent’ is receptive, ‘bland’ is not receptive
- Aggressiveness - degree of aggression, eg. ‘severe’ is agressive, and ‘healthy’ is not

These factors were gained by a series of surveys measuring the association and relation of words to each other. For example, a subject was given the two pairs of words ‘SHARP-Dull’ and ‘Relaxed-Tense’, and asked to circle the word in the second pair that seemed closest in meaning to the capitalised word 'Sharp'. [Osgood et al., 1957] . The work has since been cited in over seven thousand publications. It is useful as an early work that codifies the different domains and concepts that humans speak in, describing the axes on which we clarify opinions and judgments. However, it is very complex and therefore difficult to learn. In 1957, it was not possible to use machine learning on this large amount of factors. Affect and appraisal theory similarly complex to Osgood's work, and applications for these, have now become possible to implement in more recent years.

3.2.4 Affect and appraisal theory

[Subasic and Huettner, 2001] define the concepts of affect, centrality, and intensity in terms of using fuzzy semantic typing to analyse text. Affect is defined as a linguistic domain category, such as 'revulsion' or 'intelligence'. The paper claims to have 3876 lexical entries for affect categories, but words are defined to belong to only a small number of affect categories. Centrality is a number between 0 and 1 that reflects the degree to which a word belongs to an affective category, for example, 'emasculate' is 0.7 in the affect category 'weakness', and 0.3 in the affect category 'violence' [Subasic and Huettner, 2001]. Finally, intensity is a number between 0 and 1 that represents the strength of the affect level, for example, 'abhor' is 1.0 in intensity and 'displeasure' is 0.3 in intensity [Subasic and Huettner, 2001]. This work was based on manual tagging and computerised testing of documents that included movie reviews and news articles. Intensity measurements have potential in parody detection because parodic texts are observed in human analysis to use stronger words than serious texts, which may try to give moderated factual analysis.

Apprasial theory measures on the four factors of attitude, graduation, orientation, and polarity [Whitelaw et al., 2005].

- Attitude - includes affect, appreciation, or judgment, where 'affect' is an emotional state such as 'happy' or 'angry'; 'appreciation' refers to intrinsic appreciation of properties such as 'slender' or 'ugly'; and 'judgment' is a social judgment such as 'brave' or 'foolish'.
- Graduation - an intensity measurement such as 'very' or 'slightly'
- Orientation - whether the overall sentiment is positive or negative
- Polarity - marked or unmarked, eg. 'not' is marked

In the experiments of [Whitelaw et al., 2005], the researchers used publicly available movie review data sets, positive and negative, and a part-of-speech tagger to identify adjective features. These experiments show a future direction for sentiment analysis trending closer to the greater complexity of Osgood's early work. Binary classifications are only the beginning of future work.

Microsoft Office's ToneCheck measurement [Microsoft Office, 2010] measures words and phrases associated with a defined set of eight human emotions:

- Affection/Friendliness
- Enjoyment/Elation
- Amusement/Excitement
- Contentment/Gratitude
- Sadness/Grief
- Fear/Uneasiness
- Anger/Loathing
- Humiliation/Shame

For example, the phrase ‘get off your pedestal’ is rated as angry, and the phrase ‘It has been concerning me’ is neutral. This can be seen as a deeper examination of Osgood’s element of ‘evaluative judgment’ on emotional areas; it can also be viewed as representing the affective state of emotions in apprasial theory, or as a combination of affect and intensity as per [Subasic and Huettner, 2001]. Similar practical applications in the future should include such a diversity of sentiment measurement.

3.3 Groups of Sentiment Extraction Algorithms

Sentiment extraction algorithms can be grouped into non-learning and learning. They can also be classified by their usage of different training algorithms; and by training sources that range from labelled data in the domain in the specific domain, labelled data in other domains, or unlabelled data in the specific domain.

An example of a non-learning algorithm is presented in [Spertus, 1997]. This constructed a simple decision tree based on labelled spam and non-spam email material. Features were defined from the labelled data as including standard disdainful epithets such as ‘Get outta here!’, obscenities, phrases that began with the word ‘you’, and the use of standard words of praise like ‘love’ or ‘great’.

In training, [Aue and Gamon, 2005] describe four possibilities for training classifiers in domains without much labelled data:

- train on a mixture of labelled data from other domains
- train as above but limit features to those available in the target domains
- use ensembles of classifiers from domains with available labelled data
- combine small amounts of labelled data with the large amounts of unlabelled data in the target domain.

Algorithms that learn on unsupervised data such as [Choi et al., 2009] and [Zagibalov and Carroll, 2003] show that it is possible to generate features through bootstrapping. The latter example, a method of automatic seed word selection for the unsupervised learning of Chinese text, is interesting because it makes use of markedness: unmarked terms like 'good' are usually more positive than marked terms such as 'not good', and adverbs that follow negations are also more likely to be positive. Iterative retraining is then used to improve upon results. In the case of parody learning, parody corpuses are often presented as a whole as a source of training data: for example, [Onion Inc., 2010] is a website devoted to parody and Reuters [NIST, 2000] is an archive of genuine news articles. It is not necessary to engage in bootstrapping. However, for new texts from an unknown source, parody analysis can be used to determine if serious or non-serious based on training data.

3.4 Available Databases and Learning

The table below shows some of the resources and prior work done in the field of sentiment analysis for reference purposes.

Table 2: Databases used in sentiment learning.

Dataset Name	Link	Description	Discussion	Paper(s) Used
Pang and Lee movie data	IMDB . com	1000 positive and 1000 negative reviews from the database about movies	The dataset classified the sentimental language in reviews as positive or negative, treating them as simple thumbs up or thumbs down (reviews may be given multiple stars)	Pang and Lee 2004, Bai et al 2004
Rotten tomatoes movie data	rotten tomatoes . com	Positive and negative movie reviews	The dataset classifies the reviews as 'fresh' or 'rotten', ie. 'positive' or 'negative'. This is a simple binary classification done by parties other than the reviewers.	Pang and Lee 2004

NICTA annotated data	-	An annotated word list with positive/negative polarity manually attached to common words and phrases (partly by the thesis author)	This list, when tested, proved to be oversimplified. Positive and negative values attached to particular words (eg. 'horrible', -0.5, 'lovely', +0.5) were not successful in applications used by NICTA.	-
WordNet Affective Domains	wndomains . fbk . eu / wnaffect . html	WordNet with affective domains, 'Positive', 'Negative' and 'Neutral', with eleven affective labels	Affective domains relate to a list including emotion, mood, trait, cognitive state, physical state, hedonic signal, emotion-eleciting situation, emotional response, behaviour, attitude, and sensation. This long list shows the wide range of emtions under categories.	[Strapparava et al., 2006] SentiWordNet used in [Choi et al., 2009]
Epinions	epinions . com	Opinions with 1-5 star ratings on the internet on many areas–restaurants, travel, consumers products, etc.	The diversity of the subject matter of the opinions means a diversity of the vocabulary used in each domain. Some words may not be generalisable across different domains, eg. 'fast' is positive for a car but it can be negative for a person.	[Branavan et al., 2008]

Yelp restaurant reviews	yelp.com	A series of restaurant opinions from customers.	These user-submitted reviews rate restaurants and other products with verbal descriptions and ratings out of five stars.	[Dalvi et al., 2009]
Amazon	amazon . com	1-5 star ratings on books, movies, games, etc	The star ratings in the middle range indicate ambiguity in the reviews. [Tsur et al., 2010] used 66000 Amazon reviews with 5000 tagged for sarcastic versus not sarcastic. Jindal and Liu used Amazon reviews tagged for spam and not spam.	[Tsur et al., 2010],[Jindal and Liu, 2008]

Political blogging	Various	Political blogs, both serious and non-serious, on either sides of political spectrums, eg. hoydenabouttown.com , jonswift.blogspot.com, slacktivist.typepad.com	Blogs contain vast amounts of unclassified data.	[Agarwal et al., 2008] used blogging sources to assess influence in the blog community (as against sentiment in general)
Crikey	crikey . com . au	Australian political blogging.	Unclassified Australian political blogging; a lot of data about a narrower range of subject than epinions.	NICTA's Opinion Watch
IBM Lotus Blogs	-	Annotated blogs regarding sentiments on IBM Lotus collaboration software	This is sentimental analysis annotated, about the same subject matter and the same domain: facilitating analysis within the narrow and well-defined boundaries.	[Melville et al., 2009]

News articles	google . com / news, cnn . com, abc . net . au, et al; parody articles	News articles taken from the web as examples of non-parody news text	These were processed for word and punctuation features and used to contrast with parodies of news, especially when about similar subject matter. Parody articles were also taken from the web.	Present thesis [Choi et al., 2009], combined with SentiWordNet
Newsgroup unlabeled articles	Usenet	News articles taken from the web, unlabelled	These were used by [Choi et al., 2009] as unlabelled data to explore Naive Bayes relationships between topics	[Nigam et al., 2000]
MPQA opinion corpus	cs . pitt . edu / mpqa	OpinionFinder, a subjectivity lexicon with manual subjectivity sense annotations	The corpus contains news articles on defined subjects, annotated for intensity (intensity of expression) and polarity (contextual polarity, eg. positive/negative). This was tagged in detail.	-
WSJ tagged corpus/Penn Treebank	nlp . stanford . edu / links / statnlp . html cis . upenn . edu/˜treebank	Wall Street Journal tagged articles	Adjectives are tagged in the corpus, some with orientation labels of positive and negative. The experimentation used these results to generalise by machine learning to new adjectives.	[Hatzivassiloglou and McKeown, 1997]

Customer Review, Additional Review, Amazon Reviews	cs . uic . edu / ˜ liub / FBS / sentiment - analysis . html	Dataset products used by the researchers Hu, Liu, Ding, Nitin, and Jindal.	Customer and general review datasets used in the papers by the given authors, including Amazon reviews for the purposes of fake review detection. The customer review datasets include tagging for features of particular products, and positive or negative opinion about the product features.	[Hu and Liu, 2004], [Ding et al., 2008], [Jindal and Liu, 2008]
Product support services web survey data	-	Data upon customer opinions of product support services	Classified in binary terms as 'satisfied' and 'not satisfied', from 5000 responses to a web survey on the same topic.	[Aue and Gamon, 2005]
Knowledge base web survey data	-	Data upon customer opinions of knowledge base web survey data	Classified in binary terms as 'bad' and 'good' responses to the same web survey.	[Aue and Gamon, 2005]
Government transcripts	govtrack . us (United States)	Data of congressional debates and voting patterns	Congressional words labelled with how the congresspeople voted.	[Thomas et al., 2006]

4 Sarcasm Case Studies and Examples

4.1 Introduction

Sarcasm is an interesting subset of sentiment analysis. After defining the nature of parody, sarcasm, and satire, understanding the fundamental terms of linguistic analysis, and considering the prior work on sentiment analysis, these worked examples are provided with a human eye to give insight into potential features. The purpose of this section is to show case studies of serious versus parody articles upon similar subject areas, so as to distinguish the features that make an article on the same topic serious versus parodic. The articles were chosen on the basis of similar subject matter, with the non-serious articles usually chosen prior to the serious articles as there exist fewer parody articles than serious articles. Several are from the Onion corpus.

4.1.1 Structure of Section

Firstly, the serious and parody articles are shown side by side as worked examples with human commentary upon their differences. Hypothesis 2 was that average word and sentence length would be longer in parodies than in serious articles. This was because parodies have the freedom to use more elaborate language than serious news articles. However, on the basis of the small sample size analysed by a human eye, this hypothesis was not proven; in one example a parody article purposefully used shorter sentences to make its point. This section then provides a general list of the common features that can distinguish serious articles from parody articles to a human. This is a contribution of features found by the human eye that can be used to build automated tools for feature analysis. Finally, the section describes difficult examples as subjects for future work.

4.2 Worked Examples

This section contains examples of genuine and parody articles side by side, with a human's take on what distinguishes them from each other as well as basic average word and sentence length. Word length is calculated without including punctuation, and sentence length is roughly calculated by the number of periods in the article.

The following comparisons were selected:

- Kennedy Election: Parody Blogging versus Genuine News. This compares a parody blog to serious news on the same topic. This parody blog was not used as a non-serious corpus to train upon in the automated results.
- Dubai Debt Crisis: News versus The Onion. This and the item below both compare standard news articles with articles from the Onion. The Onion was not included in the automated results.

- Same-Sex Marriage: News versus The Onion. Compares a genuine article on same-sex marriage to the Onion's satirical report.
- Obama financial stimulus: genuine versus The Onion. Compares a genuine article on United States President Obama to an Onion article on the same topic of actions taken regarding the financial crisis.
- Kagan Judicial Confirmation: Opinion Article versus Parody. This used an opinion article as a serious news article that expressed the writer's opinion rather than the neutral style of a news report, to compare to the Onion.
- Don't Ask Don't Tell Policy: Serious Blogging versus Parody. This used a 'blog' article (ie. without the formal standards of a news publication) to compare to the Onion's article on the same topic.
- Movie Review: CAPAlert versus Landover Baptist. This compares a genuine religious website to a parody religious website to note the difference between the underlying attitudes and expressions used. Landover is a source used in the automated results. CAPAlert is a site without the formal standards of a news publication.

  The parodies compared to are drawn from blogging, the Onion corpus, and the Landover corpus, and the serious items are drawn both from news and more informal commentaries.

4.2.1 Movie Review: CAPAlert versus Landover Baptist

[CAPAlert, 2010] is a website that provides 'the truth about the content of popular entertainiment using God's Word to decide what to tell you', and is a 501(c)(3) non-profit under United States law. It is a website that genuinely reviews movies from a Christian perspective.

[Landover Baptist, 2010] is a parody website subtitled 'Where The Worthwhile Worship' and 'is a complete work of fiction' according to its Terms of Service. It is a website that does not review movies in good faith from a Christian perspective.

Table 3 compares the CAPAlert review of the movie Toy Story II juxtaposed with the Landover Baptist review of the movie Toy Story II.

Table 3: CAPAlert versus Landover Baptist

Genuine	Parody	Difference
Words: 692 Sentences: 51 Average Word Length: 4.49 characters Average Sentence Length: 13.67 words	Words: 535 Sentences: 22 Average Word Length: 4.76 characters Average Sentence Length: 24.45 words	The parody article has longer sentences in terms of word length. This may suggest a convoluted style that is not present in serious publications. The genuine is also much longer than the succinct parody. For genuine news versus parody material, the news may be succinct, but serious independent blogging does not have the length requirements of a news publication.

CAPAlert excerpt:

A true G-rated movie. Really! And that is odd for Disney. Remember that Miramax, a spawn of Disney, produced *Pulp Fiction* and many other vulgar and obscene movies. And the clean, family-oriented history of Disney does not excuse Disney's culpability for the new age media monsters.

Before I get into the analysis of *Toy Story 2* let me remind you that the CAP analysis model works on animation (and computer-generated animation) – cartoons – if the assumption is made that any behavior that can be reasonably duplicated by a child can be treated as an influential matter.

I was so hoping for a CAPCon(dition) green light for *Toy Story 2*. Movies which earn the CAPCon green light are indeed rare. *Toy Story 2* would have earned a CAPCon green light if there was but one less issue of violence – just one less. Whatever. *Toy Story 2* earned a CAPCon Yellow light, which is nearly as rare with modern movies.

To avoid destroying the excitement of discovering the story for yourself I will not, as is typical in these analyses, present the matters of possible concern within a "line by line" description of the script and choreography containing the ignominy. Only a few items of concern were present which may merit your adult counsel.

There was fighting. I was raised believing that guys gotta take their part. And I "took my part" on a number of occasions. A good ol' fight when I was growing up was sometimes the best way to become good friends – kinda stupid but that's the way it was way back when. And the only thing I and my punchbuddies used in fights were our fists. We even helped each other, bloody noses notwithstanding, get back up on our feet to head out to the nearest soda joint. But nowadays fighting too often involves weapons and ultra-viscious hatred. Oh, for the days of "black-n-white." Fighting in *Toy Story 2* was a touch more severe and frequent than in its predecessor. The fighting influence may need your wisdom to give balance to your child's understanding of settling differences.

A video game killing, a theft (of Woody from a yard sale), a breaking and entering (to rescue Woody), and a threat of physical violence by another story figure topped off the matters of violence and crime in *Toy Story 2*...

Sex/Homosexuality lost points only to a gaggle of Barbie(tm) dolls dancing in swimwear as Mr. Potato Head and others ogle at them with a sensual presence. Bo Peep was also her usual alluring self but so ambiguously as to make declaring her behavior an influence very difficult.

Landover excerpt:

The movie company Walt Disney, run from top to bottom by sodomites, has once again thumbed its rodent nose at traditional family values. Yes, the same company that brings Christian families the occult hit "Sabrina The Teenage Witch" on ABC, Antichrist Broadcasting Company, is back in business promoting their Chairman Emeritus, the Devil. Toy Story II, a stealthily packaged "children's" movie, doesn't even pay lip service to original Toy Story. No, instead, it picks up where Caligula left off! zipfach-turning tale is so sexually offensive that even that ungodly crippled pervert Larry Flynt would throw his fat little body from his wheelchair and drag his carcass to the lobby to avoid seeing such putrid computer-generated mess as Toy Story II.

The cute boy from the first movie has turned so queer and wears so much eye make-up he looks like he should be in the cast of "Rent." And his innocent, sweet toys of youth have now been turned into a mine-field of deviant adolescent sexual experimentation. Buzz Lightyear is no longer a battery powered spaceman, but is a turbo-activated hand-held rectal-stimulator with two rotating heads worthy of Black & Decker. And in another crass attempt to plug a Disney TV show, the lovable and classic Mr. Potato Head is now "Judge Judy the Anal Probe." Yes, these are disgusting words, but the truth must be told. Woody, living up to his smutty namesake, spends the entire movie spiking the other toys' batteries with powdered-Viagra, as all of the toys compete with each other to see who can be the first to mount the sweet little springer spaniel puppies that live in the basement. This filthy little story ends with a car chase leading up to all of the characters going next door to rape a little retarded girl. This kind of debauchery hasn't been seen since Elton John's last pool party.

Pastor Deacon Fred read the script 12 times. He warned Landover members not to go see Toy Story II. He is particularly adamant that they do not purchase the promotional tie-ins that will be carried by McDonalds. "Even if the sexual probe is pretty and free with a Happy Meal, it is not something any right thinking Christian should ever have in their Christian homes. Let's hit these Homo Jews out in Hollywood where they live – in their wallets. Praise the Lord!"

Differences observed:

- The comparison is of an independent commentary to a parody, so the language used in both is more similar than a comparison of straight news article to parody (ie. it is similar to comparing blogging to parody).
- More hyperlinks were used in the parody than the serious material.
- The parody uses the relatively obscure word 'pusillanimous'.
- The parody uses the offensive term 'Homo', and uses it in close distance to 'Jews'. The genuine review uses the term 'Homosexuality'.
- The parody uses the offensive/archaic term 'Sodomite'
- The parody juxtaposes 'Satan' with a children's film
- The parody juxtaposes (pornographer) Larry Flynt wiith a children's film
- The adult celebrity Elton John (though his name might also be invoked in a genuine review of a film where he was involved in the production) is also juxtaposed by the parody with a children's film
- The use of the words and phrases 'rape', 'Viagra', and 'sexual probe' in a parody of a review of a children's film
- The parody uses the complex and contextual sarcasm of 'Pastor Deacon Fred read the script 12 times. He warned Landover members not to go see Toy Story II.' (The humour is the hypocrisy that to read a script 12 times implies that one liked it, and yet Pastor Deacon Fred then tells others not to go.) This is a very complicated feature for machine detection.
- The genuine Christian review is a relatively extreme example of its type (similar to Conservapedia as an example of a biased, though serious, publication), so the difference is less obvious

4.2.2 Kennedy Election: Parody Blogging versus Genuine News

The blog jonswift.blogspot.com is an informal, humour blog. Its take on the potential of Caroline Kennedy, the daughter of a former President of the United States, to be elected as Senator is here directly compared to an article sourced from the genuine news website the BBC.

Table 4: Kennedy Election: Parody Blogging versus Genuine News (statistics)

Genuine	Parody	Difference
Words: 405 Sentences: 18 Average Word Length: 4.84 characters Average Sentence Length: 22.5 words	Words: 668 Sentences: 31 Average Word Length: 4.71 characters Average Sentence Length: 21.58 words	There is little difference here in sentence length, and a slightly more advanced vocabulary in the genuine article. The parody article is longer as it is blogging unfettered by the standards for succinctness of a standard newspaper.

Genuine news excerpt:

> Caroline Kennedy, the daughter of late President John F Kennedy, is putting her name forward for the US Senate.
>
> She is seeking to take over as junior senator for New York from Hillary Clinton, who is set to be secretary of state in the Obama administration.New York Governor David Paterson, who will choose Mrs Clinton's successor, said Ms Kennedy had expressed interest.She was a key backer of Barack Obama as he defeated Mrs Clinton to secure the Democratic presidential nomination.Caroline Kennedy, whose father was assassinated in 1963, has spent a lifetime staying out of the political limelight, says the BBC's Matthew Price in New York.But with the Kennedys less of a family and more of a political dynasty, that may be about to change, our correspondent says.Ms Kennedy's name has been suggested for several weeks as a possible replacement for Hillary Clinton.Democratic Governor David Paterson said Ms Kennedy had expressed interest in the Senate seat."She'd like at some point to sit down and tell me what she thinks her qualifications are," Mr Paterson said, who has indicated he has not yet chosen a successor to Mrs Clinton.Name recognitionSome critics have said Ms Kennedy is a socialite blessed with high-profile connections and the pulling power of the Kennedy name.Democratic New York Congressman Gary Ackerman voiced

doubt over her readiness for office."She has name recognition, but so does J-Lo," he said recently, referring to actress Jennifer Lopez. Caroline Kennedy's family history is intertwined with that of the US-Republican Peter King, who had already voiced interest in the seat, said news of Caroline Kennedy's own interest made him more determined to run.Members of the Clinton camp, who were furious when Ms Kennedy and her uncle, Senator Edward Kennedy, endorsed Mr Obama during the primaries, are also concerned that she may take over Mrs Clinton's seat, correspondents say.The seat was once held by Ms Kennedy's uncle Robert F Kennedy, who was assassinated in 1968.Senator Edward Kennedy has brain cancer, and his illness has raised the possibility that the Senate could be without a Kennedy for the first time in some 50 years, the New York Times reports.

Parody blogging excerpt:

According to the New York Times, an aide to New York Attorney General Andrew Cuomo is questioning the credentials of Caroline Kennedy to replace Hillary Clinton as Senator from New York if she is confirmed as Secretary of State in the Obama Administration. Cuomo believes he is qualified to be New York's Senator because he was once married to a Kennedy. But that is not enough. New York's next Senator must actually be named Kennedy. The Kennedy name has a "special magic capital," as Maureen Dowd so poetically calls it. But there are other Kennedys who are just as qualified, if not more so, than Caroline. If we really want the best Kennedy to fill Robert Kennedy's old seat, New York Gov. David Patterson should choose conservative former MTV VJ Kennedy.

Those who are lobbying for Caroline Kennedy, such as New York City Mayor Michael Bloomberg, have their hearts in the right place. They want to find the candidate who will most annoy and embarrass Hillary to replace her. And appointing Caroline would certainly accomplish that. Although she hasn't voted much or been that involved in politics or even studied the issues, she did make Hillary angry when she and her uncle Sen. Ted Kennedy endorsed Obama over Hillary in the Democratic primary. As I explained at the time, "It is not just that Obama reminds them of Kennedy, it is also that the Clintons remind them of Lyndon Johnson. And if there is anything that the Kennedys don't like, it's a bunch of hillbillies in the White House, which is being kept in trust until a competent Kennedy can be groomed to take it back for its rightful owners. Until that time Obama will do." The Clintons, like Johnson, think of politics as mud wrestling or the roller derby, while the Kennedys see it as a friendly game of touch football. So it must have irked them to see Hillary, the Sandra Day O'Clobber of American politics, besmirching the Senate seat that by rights belongs to the Kennedys.

Appointing Caroline Kennedy to the Senate would make the Hillary-haters happy, but I'm afraid it won't annoy Hillary enough. The few abbreviated press conferences Caroline has had, before her aides cut them off, showed that she isn't the most articulate Kennedy in the world. In a Senate committee hearing, Hillary would make mincemeat of her. But former MTV VJ Kennedy has had quite a lot of experience in the spotlight and is quite articulate. The woman who once simulated fellatio with her microphone while interviewing former New York mayor Rudolph Giuliani would have no problem taking Hillary on. Who knows what she might do with a Senate microphone. In comparison, Caroline Kennedy seems just too nice and polite and would wilt in Hillary's glare.

Kennedy does not currently live in New York and probably doesn't know much about the issues affecting the state, but as far as I know, actually living in New York has never been a requirement to represent the state in the Senate, and she probably knows as much about New York issues as Caroline does.

Differences observed:

- The serious article does include some sarcasm: '[Caroline Kennedy] has name recognition, but so does [the pop star] J-Lo'. It is a quote from a third person placed into the serious article as a quote. Such isolated sarcastic sentences are possible in serious articles.7
- In the parody article, the line 'The Kennedy name has a "special magic capital" as Maureen Dowd so poetically calls it' juxtaposes the word 'poetically' with the non-poetic phrase 'special magic capital'.
- The parody text also uses the phrase 'simulated fellatio with her microphone'. 'Fellatio' is a relatively technical and formal term, but it is an uncommon word in most serious articles about politics.
- The parody uses the informal expresion of 'slap in the face at Hillary'.
- The parody article calls the politician 'Hillary' rather than her full name 'Hillary Clinton'. Serious articles sometimes use only a politician's first name as a headline, but to be formal in text they will more commonly use a last name.

4.2.3 Dubai Debt Crisis: News versus The Onion

[Onion Inc., 2010] is a website that parodies news and was later used as the principal corpus of non-serious article data. The Daily Mail is a United Kingdom newspaper. They offer different perspectives on a debt crisis in Dubai.

Table 5: Dubai Debt Crisis: News versus The Onion (statistics)

Genuine	Parody	Difference
Words: 1269 Sentences: 59 Average Word Length: 4.89 characters Average Sentence Length: 21.83 words	Words: 605 Sentences: 25 Average Word Length: 5.06 characters Average Sentence Length: 24.44 words	The genuine news article is longer than the parody, but the parody has a slightly longer average sentence and word length. It uses more elaborate language and description than the genuine news article.

Genuine news excerpt:

> Dubai debt crisis: Now British banks face fresh crisis after investing billions
>
> Liz Hazelton, Daily Mail
>
> Barclays, RBS and HSBC face losing billions Wall Street plummets by 2 per cent after late opening FTSE falls by 1.5 per cent before stabilising
>
> Banks see 14billion wiped off market value in one dayDubai may consider selling QE2 to tackle debt
>
> British banks were teetering on the brink of a fresh meltdown today after it emerged they had invested heavily in crisis-hit Dubai.
>
> An $80billion debt default in the emirate has already reawakened the spectre of a global 'double dip' - that the first shoots of recovery could be wiped out by a second wave of recession. But the level of exposure that the crippled British banking sector faces is now under renewed scrutiny.
>
> The crisis was prompted by Dubai World, the development company behind three palm shaped islands as well as an off-shore replica of the globe , defaulting on its debt. Today it emerged that:
>
> Royal Bank of Scotland (RBS) was Dubai World's biggest loan arranger since January 2007, according to JP Morgan HSBC has an estimated 9.6billion in loans and advances to UAE customers Barclays has an exposure of around 3billion
>
> Another bailout? Gordon Brown (right) meets Dubai's ruler Mohammed bin Rashid Al Maktoum at Downing Street earlier this week
>
> The figures are particularly alarming as the sector has had to be bailed-out by the tax payer on a number of occasions over the last year-and-a-halfEarlier this month, RBS and Lloyds Banking group received another 50billion to keep them afloat
>
> RBS - which has received the biggest state rescue anywhere in the world - is now effectively owned by the taxpayer.

As the money markets continued to falter, Gordon Brown moved to dispel investors' panic, claiming that he believed British banks were 'well-capitalised'. Speaking at the Commonwealth summit in Trinidad, Mr Brown said: 'I think we will find this is not on the scale of the previous problems we have dealt with.'

Asked if the Dubai situation could spark a 'double-dip' recession, he said: 'You are obviously going to have setbacks with a bank here or an organisation there which has had problems, but I do believe the world has a better way of monitoring what is happening, so we can be sure that - despite setbacks - we will continue to go forward.'

The Onion excerpt:

Representatives from the emirate of Dubai announced with disappointment this week that its recent debt crisis has forced developers to halt construction on the city's long-planned 22-mile-long indoor mountain range.

Planners continue to take future reservations for the mountains' 9 and 10-star hotels.

The culmination of a decade's worth of ambitious and expensive building projects, Dubai's estimated $100 billion debt officially brought work on the artificial mountain range to a stop on Tuesday.

"This is a very sad day for the emirate of Dubai," Crown Prince Hamdan bin Mohammed al-Maktoum told reporters at a press conference held inside the gold-plated anti-gravity chamber in his palace. "Although I believe it is the basic right of all who visit us to be able to scale to the top of a 15,000-foot-tall manmade snowcap, these tough economic times have made it an impossibility. Never before has our proud municipality faced such a grave crisis."

Added Sheikh Hamdan, "The time, I'm afraid, has finally come for us to tighten our jewel-studded belts."

With only seven of the planned 19 peaks completed and the artificial glaciers only partially frozen, the real estate firm Nakheel now says the landmark Alps Dubai development will miss its planned April 2011 opening date, and with it, the controlled volcanic eruption that would have commemorated the event.

Some of the more conservative construction projects completed before Dubai's financial meltdown.

"Everything had been progressing right on schedule," said project manager Zayed Kemaar. "The plate tectonics were almost in place, we were getting good vulcanism, and we had helicopter-loads of marble and schist arriving every day from Switzerland. We even had herds of pure-white albino bighorn rams standing on five of the peaks. Then, of course, the bottom fell out, and now we barely have the money to keep the air conditioning on."

Added Kemaar, "It just goes to show you that, when the economy is down, vital infrastructure projects like this are always the first to

> suffer."
>
> A number of Dubai officials have even speculated that the cornerstone Jabal Khalifa mountain, which, at 27,100 feet-not counting the 300-foot-tall Lebanese-cedar log flume atop the casino at the summit-would have been the sixth-highest peak in the world, may have to be canceled entirely.
>
> "At this rate, we may be forced to dip into the vast diamond mines we installed in the center of the city last February," Kemaar said.

Differences observed:

- The parody begins with the same phrase 'Dubai Debt Crisis', making the identical reference to genuine news.
- The parodic phrase 'World's Largest Indoor Mountain Range' may also be correlative to actual projects in Dubai.
- 9- and 10- star hotels do not exist but are mentioned in the parody, so this is an unusual juxtaposition that gives a clue of parody.
- 'Gold-plated anti-gravity chamber' is also an unusual combination of words used in the parody
- 'Lamborghinis clog dealers' lots': this sentence from the parody uses an unusual brand of car for the situation
- Adjective-loaded parody: 'the 300-foot tall Lebanese-cedar log flume atop the casino at the summit' is a good example of too much description for a serious news article to use.
- 'Maybe it's time for us to pull ourselves up by the straps of our handmade custom-fitted patent-leather Italian boots and put our slaves back to work' is also an example from the parody of many adjectives, as well as an uncommon reinterprtation of the common phrase 'pull oneself up by one's bootstraps'. The ownwership of slaves is also not casually referred to in modern-day newspapers.

4.2.4 Same-Sex Marriage: News versus The Onion

This compares an article from the American news outlet CNN (2004) to an Onion article upon an identical topic, same-sex marriage in Massachusetts.

Table 6: Same-Sex Marriage: News versus The Onion (statistics)

Genuine	Parody	Difference
Words: 703 Sentences: 40 Average Word Length: 5.29 characters Average Sentence Length: 17.72 words	Words: 679 Sentences: 38 Average Word Length: 5.23 characters Average Sentence Length: 18.18 words	In this case the genuine news has slightly longer vocabulary, but the parody sentence length is slightly longer. This may indicate a sentence style marginally more convoluted and intricate than standard news, for the purposes of expressing this parody news article.

Genuine news excerpt:

Massachusetts court upholds same-sex marriage
CNN

Underscoring its original ruling last November, Massachusetts' highest court said Wednesday that only full marriage rights for gay couples, not civil unions, would conform to the state's constitution.

The ruling sets the stage for Massachusetts later this year to likely become the first state in the nation to allow same-sex marriages.

In a statement released Wednesday night, President Bush said the ruling was "deeply troubling.

"Marriage is a sacred institution between a man and a woman," Bush said. "If activist judges insist on re-defining marriage by court order, the only alternative will be the constitutional process. We must do what is legally necessary to defend the sanctity of marriage."

In his State of the Union address January 20, the president stopped short of endorsing a constitutional amendment that would ban marriages for gay and lesbian couples, as social conservative groups had hoped.

But he said, "... if judges insist on forcing their arbitrary will upon the people, the only alternative left to the people would be the constitutional process."

Wednesday's advisory opinion by the Massachusetts court was in response to a request from that state's Senate about whether allowing gays to join in civil unions would be sufficient.

The court rejected using civil unions as a remedy, "Because the proposed law by its express terms forbids same-sex couples entry into civil marriage, it continues to relegate same-sex couples to a different status. ... The history of our nation has demonstrated that separate is seldom, if ever, equal."

Parody blogging excerpt:

Massachusetts Supreme Court Orders All Citizens To Gay Marry
The Onion

BOSTON - Justices of the Massachusetts Supreme Judicial Court ruled 5-2 Monday in favor of full, equal, and mandatory gay marriages for all citizens. The order nullifies all pre-existing heterosexual marriages and lays the groundwork for the 2.4 million compulsory same-sex marriages that will take place in the state by May 15.

Enlarge ImageA justice performs a mandatory marriage."As we are all aware, it's simply not possible for gay marriage and heterosexual marriage to co-exist," Massachusetts Chief Justice Margaret H. Marshall said. "Our ruling in November was just the first step toward creating an all-gay Massachusetts."

Marshall added: "Since the allowance of gay marriage undermines heterosexual unions, we decided to work a few steps ahead and strike down opposite-sex unions altogether."

Marshall said the court's action will put a swift end to the mounting debate.

"Instead of spending months or even years volleying this thing back and forth, we thought we might as well just cut to the eventual outcome of our decision to allow gay marriages," Marshall said. "Clearly, this is where this all was headed anyway."

The justices then congratulated the state's 4.8 million marriage-age residents on their legally mandated engagements.

The court issued the surprise order in response to a query from the Massachusetts Senate over whether Vermont-style civil unions, which convey the state-sanctioned benefits of marriage but not the title, are constitutional.

"If the history of our nation has demonstrated anything, it's that separate is never equal," Marshall said. "Therefore, any measure short of dismantling conventional matrimony and mandating the immediate homosexual marriage of all residents of Massachusetts would dishonor same-sex unions. I'm confident that this measure will be seen by all right-thinking people as the only solution to our state's, and indeed America's, ongoing marriage controversy."

Marshall then announced her engagement to Holyoke kindergarten teacher Betsy Peterson, a pairing that had been randomly generated by computers in the census office earlier that day.

Differences observed:

- The headline parsing is different: 'Massachusetts Supreme Court ORDERS all citizens to gay marry' versus 'Massachusetts court UPHOLDS same-sex marriage'

- 'Same sex marriage' is the official, standard term; the more colloquial 'gay marriage' is used in the parody
- 'Mandatory gay marriages' is an unlikely phrase out of parody, even for blogs written by people very much against same-sex marriage

4.2.5 Kagan Judicial Confirmation: Opinion Article versus Parody

This is a comparison of an opinion article by Nina Totenberg on the serious news site NPR, compared to an Onion article on the same topic, the nomination of Justice Elena Kagan in the United States. This shows a particular form of parody, a short and dramatic article compared to a longer, formal opinion piece.

Table 7: Kagan Judicial Confirmation: Opinion Article versus Parody(statistics)

Genuine	Parody	Difference
Words: 951 Sentences: 52 Average Word Length: 4.87 characters Average Sentence Length: 18.46 words	Words: 144 Sentences: 8 Average Word Length: 5.71 characters Average Sentence Length: 18.38 words	This form of parody is much shorter than the genuine news and opinion articles it parodies because it uses the shorter form to make a point. This is in contrast to forms of parody that use more adjectives and descriptive language to illustrate perceived absurdity.

Genuine news excerpt:

> At Confirmation Hearings, GOP Eyes Kagan's Record
> by Nina Totenberg
> NPR
> June 28, 2010
> U.S. Supreme Court nominee Elena Kagan on Capitol Hill on May 13.
> U.S. Supreme Court nominee Elena Kagan faces questions this week from the Senate Judiciary Committee, which opens hearings on her nomination.
> U.S. Supreme Court nominee Elena Kagan on Capitol Hill on May 13.
> Tim Sloan/AFP/Getty Images

U.S. Supreme Court nominee Elena Kagan faces questions this week from the Senate Judiciary Committee, which opens hearings on her nomination.

June 28, 2010

The Senate Judiciary Committee opens hearings Monday on the nomination of Solicitor General Elena Kagan to the U.S. Supreme Court. So far, Republican attempts to arouse controversy about Kagan have gained little traction, but this week marks the main event.

Kagan's supporters admit privately that if there is a silver lining to the Gulf oil spill, it is that Kagan has been able to sail unscathed beneath the news radar screen for the seven weeks since her nomination.

For more than a month, Republicans have been hurling themselves at the Kagan appointment, with about as much effect as hurling themselves at a brick wall.

The critiques include her lack of judicial experience, her Clinton White House tenure and her deanship at Harvard Law School. All preview the GOP lines of attack.

"We know she has served extensively and repeatedly as a political operative, adviser and policymaker - quite a different job than that which she will assume should she be confirmed," says Republican Sen. John Cornyn of Texas.

Utah's Sen. Orrin Hatch skewers her lack of experience: "Supreme Court justices have had experience behind the bench as a judge, before the bench as a lawyer, or both. Ms. Kagan has neither."

Kagan's Heroes

Republicans have also criticized Kagan for what they call her heroes: the two judges she clerked for after law school, federal appeals court Judge Abner Mikva and Supreme Court Justice Thurgood Marshall, as well as former Israeli Chief Justice Aharon Barak. Jeff Sessions, ranking Republican on the Senate Judiciary Committee, asks rhetorically, "Isn't it true that a person's heroes tell a great deal about who they are?"

Indeed, Marshall, Mikva and Barak do have liberal judicial records. But the only one Kagan has singled out repeatedly as her hero is Marshall, the architect of the nation's civil rights legal revolution, and the man who argued and won Brown v. Board of Education, the case that ended racially segregated public schools in the United States.

At the ceremony announcing her nomination, Kagan praised Mikva and Marshall, saying that Mikva "represents the best in public service," and that Marshall "did more to promote justice over the course of his legal career than did any lawyer in his lifetime." In contrast, Kagan's praise of the former Israeli chief justice appears to be from her introduction of him when he spoke at Harvard, a kind of praise she has also heaped on conservative speakers at Harvard such as Justice Antonin Scalia.

Lines Of Attack

Undoubtedly, the principal line of Republican attack this week will be the assertion that Kagan as dean was anti-military. As Sessions puts it, "Ms. Kagan obstructed the access of the military as it tried to recruit bright young JAG officers to support and represent our soldiers."

Kagan will undoubtedly say that she did oppose the military's "don't ask, don't tell" policy, but that her job as dean was to ensure that nobody was the victim of discrimination on campus ? neither gay students nor those interested in being military lawyers. She will seek to persuade senators that she tried to steer a middle course that allowed military recruiters on campus, but did not give them access to the school's office of career services.

Republicans, too, will focus on the 170,000 documents and e-mails from Kagan's four-year tenure in the Clinton White House. While the documents do portray a hard-edged and politically savvy lawyer, the memos of one-time Reagan administration aide John Roberts were even more pointed. At his confirmation hearing to be chief justice, Roberts explained them away as representing the views of the administration he was serving. Expect Kagan to use those answers as her model.

Probably Kagan's biggest problem will be that as a scholar she was acidly critical of the cliched and unresponsive answers of previous Supreme Court nominees during their confirmation hearings. The process, she wrote, has taken on "an air of vacuity and farce," in part because senators have gotten away from "the essential rightness the legitimacy and the desirability of exploring a Supreme Court nominee's set of constitutional views and commitments."

So expect senators to ask Kagan to live up to her own standard.

Parody excerpt:

The Onion

JUNE 21, 2010 — ISSUE 46?25

WASHINGTON Saying they didn't want to waste any more time dicking around, members of the U.S. Senate began Elena Kagan's Supreme Court confirmation hearing Monday by asking the solicitor general point-blank if she had the goods to join the nation's highest court. "I got your master's thesis in front of me, I got some speeches you made while you were the dean of Harvard Law School, but Kagan, let's cut the shit: You gonna be able to bring it or not?" Sen. Dick Durbin (D-IL) asked the former Clinton policy adviser. "Because the American people deserve a justice who won't crumble like a cupcake and run home to mommy when Second-Amendment-ruling time comes around." After indicating that she was "fucking born ready," Kagan was confirmed in a unanimous 100-0 vote.

Differences observed:

- Note that the far shorter length made the parodic point of humorously summarising the lengthy Supreme Court confirmation process
- The parody article uses the inappropriate language 'fucking', 'shit', 'dicking'

4.2.6 Don't Ask Don't Tell Policy: Serious Blogging versus Parody

This is a comparison of a blogged article from the left-leaning website MediaMatters [Finkelstein, 2010], and another Onion article upon the same subject matter, the Don't Ask Don't Tell policy in the United States army.

Table 8: Don't Ask Don't Tell Policy: Serious Blogging versus Parody (statistics)

Genuine	Parody	Difference
Words: 445 Sentences: 27 Average Word Length: 4.74 characters Average Sentence Length: 16.52 words	Words: 755 Sentences: 33 Average Word Length: 4.82 characters Average Sentence Length: 23.15 words	The parody is longer than the real article, with a significant difference for sentence length. Note that its language is much more elaborate and detailed in this parody.

Genuine blogging excerpt:

Rep. Gohmert: DADT Repeal Would Require Troops To Be "Overt About Their Sexuality"

May 27, 2010 4:14 pm ET Matt Finkelstein

Media Matters

On the House floor earlier today, Rep. Louie Gohmert (R-TX) railed against the White House-backed effort to repeal the military's Don't Ask, Don't Tell policy. In an impassioned speech, Gohmert cited his experience as a JAG officer to argue that he can "understand this issue" better than his colleagues. But moments later, he totally misrepresented what repealing the ban on gays and lesbian service members would entail, proving that he doesn't actually understand it at all:

GOHMERT: And think about the policy. Look, I have represented people in the army that have practiced homosexuality, and heterosexuality, and sexual assault victims. I understand this issue perhaps more than many of those on the floor here. And I'm telling you the military is not a social experiment. We are sending them out there with a mission to protect this country. And if someone has to be overt about their sexuality, whether it's in a bunker where they're confined under fire, then it's a problem. And that's what repeal of "Don't Ask, Don't Tell" does. It says I have to be overt – I don't care – I want this to be a social experiment. Our men and the women in the military deserve better. Let's here from them at the end of the year with a complete study, and then the leaders keep their word when we send out military out to die for this country. We owe them better than this...

Gohmert is a fanatical opponent of gay rights, so his passion for state-sponsored discrimination isn't surprising. That said, he is also wrong on the facts.

For starters, the proposal making its way through Congress would delay implementation until after the Pentagon weighs in later this year. So rather than precluding the military's review, it actually depends on it to lift the ban. Joint Chiefs of Staff Chairman Adm. Mike Mullen and Defense Secretary Robert Gates have endorsed this approach. Presumably, they "understand" the policy at least as well as Gohmert does.

Additionally, the idea that repealing Don't Ask, Don't Tell would require anyone to be "overt about their sexuality" is ridiculous. It simply would allow patriotic Americans to serve their country without having to hide who they are. Indeed, under the current law, many soldiers who never shared details of their personal lives have been discharged after they were outed.

Gohmert, though, appears to be afraid of the stereotype that gays and lesbians will automatically flaunt their sexuality if given the chance – even when they're "in a bunker" and "under fire.

Onion parody excerpt:

Repeal Of 'Don't Ask, Don't Tell' Paves Way For Gay Sex Right On Battlefield, Opponents Fantasize

JULY 12, 2010 — ISSUE 46?28

WASHINGTON?As Congress prepares to allow gay individuals to serve openly in the military, those against the proposed change voiced their concerns Monday, warning the repeal of "Don't Ask, Don't Tell" could soon lead to strong, strapping American soldiers engaging in mind-blowing homosexual intercourse right on the battlefield.

"We're sending our soldiers out there with a mission, and that mission is to protect this country," said Rep. Louie Gohmert (R-TX), one of many conservative politicians who staunchly oppose the change. "If this is repealed, what's to stop all-night sex romps from breaking out while U.S. servicemen are hiding in a bunker, or crawling around an irrigation ditch bathed only by the light of the moon, or, say, the dozens of other situations I've already thought through in elaborate detail?"

"We can't allow this to happen," Gohmert added as beads of sweat collected on his brow. "It's wrong. Sweaty male sex-no matter how erotic and uninhibited-is so wrong and so, so naughty."

Despite its support from the defense secretary and the chairman of the Joint Chiefs, the repeal has been condemned by many military officers who worry it could disrupt troop cohesion and endanger the lives of the taut young soldiers who have dedicated their lives to serving America with "every rippling muscle in their rock-hard bodies."

Others have argued that allowing gay soldiers to push their lifestyle on others, testing the limits of pleasure a man can take before he erupts in uncontrollable ecstasy, would seriously damage morale.

"The military should not be used to advance some radical, steamy, mouthwatering social agenda," said Rep. Trent Franks (R-AZ).

"Our men need to know they can count on each other in battle, and we can't have them getting distracted by illicit romantic dalliances," said Gen. James T. Conway, commandant of the Marine Corps. "Especially if one's a little blond Adonis farm boy and his buddy's a real tough street kid straight out of Brooklyn. I mean, think about it: What if they lock eyes and abandon their post to start ripping each other's fatigues off, revealing twin sets of glistening washboard abs and at last fulfilling their hidden passions?"

Continued Conway, "Is this the message we want to send to our enemies?"...

This May, Sen. John McCain (R-AZ) promised voters he would do everything in his power to prevent gays from serving openly in the armed forces, and on Thursday, he told reporters that the role of the military is to defend American freedoms, not "the rights of, you know, those people some of us stay up all night thinking about as we toss and turn."

"Imagine you've got a boat full of sailors out cruising the Gulf of Aden when all of a sudden they're attacked. Some of the homosexuals lock themselves below deck and begin touching themselves," said the 73-year-old senator and Vietnam War veteran, his breath quickening. "One of their names is Ricardo. Unbuttoning his pants, he throws his gunner's mate down on the cot and penetrates him, his big, beautiful dick shimmering with power, his dog tags bouncing up and down as he's pounding, and pounding, and pounding."

Added McCain, "What I'm trying to say is: It all boils down to combat effectiveness."

When asked about his views on lesbians serving openly in the military, McCain made no secret of his position on the issue.

> "Female soldiers being intimate with one another?" McCain said. "Gross! No, thank you."

Differences observed:

- The language often used in articles parodying opponents of gay rights/same sex marriage/etc seems to draw in part from romance novels. By making opponents of gay rights voice elaborate descriptions of homosexual acts in parody, the intent of the parody appears to be to imply that these opponents have sublimated their own homosexual desires. This particular subject of parody seems to often follow the format of using many adjectives and descriptive language.
- The serious article expresses the opinion of the author, and quotes extensively from the genuine opinion of Gohmert.
- Because of the genuine opinion, there are the use of some adjective descriptions, eg. the writer of the serious article referring to Gohmert as 'fanatical'.
- Note that in the genuine article, some sentences are the writer's opinion, and other sentences are objective facts, much like [Tsur et al., 2010] chose to separate individual sarcastic sentences from non-sarcastic sentences.
- In the parody article, most of the sentences have a parodic 'tell' about them, though in some parody articles there will be sentences that out of context would not be recognisable by a human as parodies.
- Quotes from United States politicians are faked in the parody article.

4.2.7 Obama financial stimulus: genuine versus The Onion

This compares a genuine article from the financial periodical MoneyMorning versus the Onion's take on the identical situation.

Table 9: Obama financial stimulus: genuine versus The Onion (statistics)

Genuine	Parody	Difference
Words: 635 Sentences: 39 Average Word Length: 5.25 characters Average Sentence Length: 16.51 words	Words: 206 Sentences: 5 Average Word Length: 5.0 characters Average Sentence Length: 41.6 words	This is another example of succinct parody from the Onion, with a much smaller word count.

Genuine news excerpt:

Obama Offers New Stimulus Package To Create More Jobs
Money Morning
BY DON MILLER, Associate Editor, Money Morning

In an effort to stimulate hiring in the face of a stubborn 10% unemployment rate, U.S. President Barack Obama on Tuesday announced proposals to create more jobs with an expansion of his $787 billion stimulus plan.

In a speech at the Brookings Institution, President Obama avoided calling the proposals a new stimulus package. But the initiatives bear a striking resemblance to the package debated by Congress last February, including more infrastructure spending and a hiring tax credit that didn't make the final cut after objections from members of his own party.

In addition to $50 billion in infrastructure spending, the proposals call for increased lending to small businesses, a one-year moratorium on capital gains taxes, and extending relief to state and local governments.

Another day, another stimulus package. Just announced more spending today.

The big question will revolve around how to pay for the programs - which some analysts are estimating will cost upwards of $170 billion - without adding to the ballooning federal deficit.

President Obama has said that the administration could spend more to stimulate hiring and still shrink the deficit by using more than $200 billion in savings from the Troubled Asset Relief Program (TARP). Republicans are adamantly opposed to using the unexpected windfall for anything but deficit reduction.

As if on cue, U.S. Treasury Secretary Timothy Geithner notified Congress that the administration is extending the $700 billion TARP financial-rescue program until October 3, 2010.

In a letter to congressional leaders, Geithner said the government must have access to the funds in case of new financial shocks, and to help struggling homeowners and small businesses. He said the administration doesn't expect to disburse more than $550 billion of the money.

As we wind down many of the government programs launched initially to address the crisis, it is imperative that we maintain this capacity to respond if financial conditions worsen and threaten our economy, Geithner wrote. President Obama revealed his proposals as a stagnant job market fuels heated debate about the effectiveness of the stimulus plan and Democrats begin to openly worry about the looming midterm elections in 2010. Although it offered some hope, the national unemployment rate still hovered around 10% in November, ratcheting up political pressures.

Onion parody excerpt:

> Obama To Create 17 New Jobs By Resigning And Finally Opening That Restaurant
> The Onion
> MAY 21, 2010 — ISSUE 46-20
> WASHINGTON
> In an effort to counter the highest unemployment rate the nation has faced in a quarter century, Barack Obama announced Monday that he will create 17 new jobs by resigning from the presidency to pursue his lifelong dream of opening a cozy little down-home restaurant just off the Galesburg, IL exit on Interstate 74. "Now is the time for drastic measures, and the several line-cook and serving positions that will be generated by Barry's Place are imperative to getting the economy back on track," said Obama, donning a white apron over rolled-up shirtsleeves. "The hope is that this bold initiative will demonstrate to other American business owners that it is possible to break the cycle after they somehow get sucked into politics and things snowball so fast that they lose sight of what's really important, like serving people the best slice of pecan pie they've ever tasted at a price that can't be beat." Vice President Joe Biden has reportedly followed Obama's entrepreneurial lead by purchasing a secondhand cologne and condom vending machine that will be installed in the men's bathroom of a Wilmington, DE offtrack betting parlor.

Differences observed:

- Use of the informal portmanteau word 'Obamanomics' inside a quote in the genuine news article
- The parody uses a descriptive style in phrases like 'cozy little down-home restaurant'
- The difficult feature here is the contextual absurdity, that the President of the United States has created only seventeen new jobs.

4.3 Sarcasm Features to the Human

As a result of this human analysis, sarcasm features can be explicitly noted rather than simply intuitively known. These were common features discovered:

- Edit distance of common sayings: for example, "Abstinence Makes The Church Grow Fondlers"/"Absence Makes The Heart Grow Fonder" (Landover Baptist, 2010). This is not limited to parody; a lot of writing, including the headlines of genuine articles, uses traditional, common phrases with new spins.

- As above, edit distance of a parody that directly uses language from the source it parodies, similarly to the Dr Seuss parody cited above.
- Unusual juxtapositions, eg. "Pedophile" with "Of the Year" (Landover Baptist 2010).
- Unusual category-related juxtapositions, eg. "rape" and "Larry Flynt" (the pornographer) with "Toy Story 2 Movie Review".
- Juxtapostion of informal language with formal language, eg. "And if there is anything that the Kennedys don't like, it's a bunch of hillbillies in the White House" (from a parody article about Caroline Kennedy).
- Inversion: This is a textual feature based both on sentence parsing analysis and contextual analysis. The parsing dependencies give a result that is unusual or unexpected in context: for example, where the verb of a sentence is turned in an unexpected manner, "Wildlife Harms BHP [when the actual context is that BHP is responsible for an oil disaster that has harmed wildlife]." This is sarcasm or irony: conveying a meaning by stating the opposite.
- Use of archaic, offensive, or informal terms that would be stylistically inappropriate in a news publication, eg. "Ted Haggard denies that he is a little nancy boy [regarding an American anti-gay religious figure who had a homosexual relationship himself]".
- Use of inappropriate or obscene words and phrases: for example, the Landover Baptist movie review below has many examples. Note that in non-serious blogging material, obscene words and phrases may be used in writing that is not intended as parody.
- Use of many adjectives/adverbs in longer descriptions in parody, usually for exaggeration, eg. "I believe it is the basic right of all who visit us to be able to scale to the top of a 15,000-foot-tall manmade snowcap".
- Because of the above linguistic feature, parody articles may have longer average words per sentence than genuine (this is not always the case in the statistics I have compiled about average word and sentence length in parody compared to genuine).
- Use of more poetic language than news reports (see above)..
- Sentence structure: a 'must', or finishing with the punctuation of '?'– "But the nasty environmentalist do-gooders will say that BHP setting sea turtles on fire is wrong or something, won't they?"
- Frequent hyperlinks: both genuine news and parody link to other news, but in some forms of parody it is common for the parody site to self-promote, eg. Landover Baptist frequently links to itself in the text of its articles.

- Quotations are different: in genuine articles, quotations are genuine and may be of real people saying sarcastic things (eg. "[Caroline Kennedy] has name recognition, but so does J-Lo", referring to Ms. Kennedy's withdrawn Senate bid). In sarcastic articles, quotations are entirely faked. Perhaps features in quotation marks ought to be given generally less weight.

- Since quotations are genuine in non-parody articles, an external query for the quotation would find that it might appear in multiple serious sources, whereas a made-up quotation is more likely to only be found in limited sources.

- Contextual semantic sarcasm: the most difficult feature to detect, for example: "The true Capitalist allows business owners to do whatever they please to gain maximum profit. I have heard that in mine collapses, mine owners would refuse to rescue workers, letting them die and hire others. It's a perfectly practical way to operate." (FSTDT 2010) The idea of this is so 'evil' that it is obviously a parody, but it is a matter of contextual semantic meaning.

- Parody is reactive: this means that a parody article will be dated later than the earliest of the mainstream news articles.

- Certain subjects are more frequently parodied than others. This information was found as a result of feature analysis following the corpus experiments.

This is a contribution to the literature as a dedicated analysis of sarcastic texts related to events in the news.

4.4 Difficult Examples

This section briefly gives examples of sarcasm that would be difficult to detect, taken from the webpage [Community, 2010]. These are sarcastic parodies that very closely imitate the style of American religious extremists, and are only detectable by subtle extremism in the semantic context. On the webpage given they are referred to as 'Poes', a plural term for subtle parodies. They are included here as an example for future directions of research.

> There is women who lives on my street for several years. She is very introvent, intelligent and kinda pretty for her age in my opinion. She lives by herself with her animals.
>
> We used to not speak to each other much but we sorta became friends while I was voluntering for toys-for-tots and she was voluntering to. She is the only women around here who is my age and not married or divorced with children. She doesn't seem bitter about it all. She is actually pretty content and happy.

She’s been all over the world, has read nearly a thousand books and rides a motorcycle! I loved listening to her talk it was so refreshing to talk to another women wothout it being about babies or husbands.

We do differ in politics. She is extremely liberal boarding socialist through she denies it. She dislikes Obama because ”He’s to conservative” ¡rolling eyes smilie¿

Yesterday, I was asked her what church she went to. She started getting shifty and nervous. She said she didn’t go to church so I invited her to mine. She said ”no Thank you” and tried to change the subject.

I kept pressing her until finally she told me

”I’m an atheist” I was in shock. I never met an atheist in real life before.We got into a hudge argument and then I asked her ”To prove there was no God”

She answered back ”Okey will look at the method you used to prove that Zeus and Thor aren’t real and will start from there”

This made me more upset and angry she ended up leaving. I am very upset right now. How can I prove the truth to her ¡shrugging smilie¿

Kikimoon

-

On Capitalism The true Capitalist allows business owners to do whatever they please to gain maximum profit.

I have heard that in mine collapses, mine owners would refuse to rescue workers, letting them die and hire others. It’s a perfectly practical way to operate; rescuing is expensive and often, what workers can be rescued are unable to work, and said expense of rescuing them is wasted.

I have read articles of rivers catching fire, and land becoming too poisoned to work afterwards. The proper answer to that is ”so?” Look at what happened to the coal companies of England when the government decided to whine about the smogs: Capitalism was interfered with and many companies were ruined. Waste treatment is likewise expensive; a river catching fire may cause some expense, but the installation of waste treatment facilities and proper sewers are likely even more so, and therefore not practical.

There has also been some argument about the Truck System, wherein a company would issue its own currency, forcing employees to buy food from company-owned stores and pay rent to the company, making it impossible to save money to move on. The simple fact is, this is a perfect means of keeping employees loyal, making sure they will go nowhere else.

Demanding companies concern themselves with workplace safety, clean up after themselves, or pay their employees a ”fair wage” dilutes and destroys capitalism. Is this a moral way to act? A ridicu-

> lous question: morality has nothing to do with Capitalism, and those that believe that's a bad thing are deluded.
>
> -
>
> My son is dating an atheist what should i do?
>
> My son asked if he could bring a girl to the house to study and i was happy that he had taken an interest in dating. They were reading in his room and i came in to bring them some snacks and started chit chatting. I asked her what Church her and her family went to and she disclosed that they were...atheists. I literally grabbed the cookie out of her hand and told her to leave our home at once. She started crying, my son was all upset, and now i feel like the bad guy. But i am just trying to be a good mother.
>
> Lizz
>
> -
>
> Young Christians, or those with children, what do you tell them about the Atheists? I found this advice online:
>
> If you find an Atheist in your neighborhood, TELL A PARENT OR PASTOR RIGHT AWAY!
>
> You may be moved to try and witness to these poor lost souls yourself, however AVOID TALKING TO THEM!
>
> Atheists are often very bitter and will lash out at children or they may even try to trick you into neglecting God's Word.
>
> Very advanced witnessing techniques are needed for them. Let the adults handle them.
>
> Jayden's Mommy

A further example of such subtlety is the genuine website [Society, 2010]. This site makes the extreme claim that the earth is flat, yet it is genuinely intended. At such levels, parodic detection becomes very difficult.

5 Automated Results

5.1 Introduction

Sentiment analysis in machine learning is the application of machine learning techniques to automatically detect the human sentiment underlying a given text. After discussing the ways that humans differentiate a serious sentiment from a parody sentiment, the automated results are compared to the human analysis. The automated results provide word features that were aggregated based on the given large corpuses rather than drawn only from the small number of articles analysed by a human. These features can be used to distinguish between parody and serious articles in future work.

This section describes feature engineering and the application of a Naive Bayes classifier to corpuses of serious and non-serious material. The corpuses used were:

- Reuters (serious, a large corpus of news articles) [NIST, 2000]
- Onion (non-serious, a corpus of parody articles from an edited website) [Onion Inc., 2010]
- Landover (non-serious, a corpus of parody articles from an edited website) [Landover Baptist, 2010]
- Other news (serious, drawn from news sources other than Reuters of recent news)
- Testing corpus (a small corpus of mixed serious and non-serious, formal and informal material used for additional testing of results)

These corpuses were chosen for several reasons: Reuters, Onion, and Landover were chosen as large available datasets of serious and non-serious articles; the news corpus was chosen as recent serious news with which to compare recent Onion and Landover articles, to judge the degree to which contemporaneousness affected results; and the testing corpus was used to validate results on material not used in training. The testing corpus included informal but serious material that was used to check the extent to which results judged formality as against seriousness.

5.1.1 Structure of Section

The corpuses used are described in Section 5.2. This is followed by experiments carried out in using a Naive Bayes classifier. First, the experiments showed that Reuters versus the Onion, and Reuters versus the combined parody corpus of the Onion and Landover Baptist, can easily be differentiated using Naive Bayes and word features. These results were then applied to the test corpus, which included articles that were not part of the training corpuses. The top features used for serious and non-serious material are analysed. Secondly, to test the effect of the age of the Reuters data compared to the age of the non-serious corpuses, the news corpus was added to the serious corpus as recent news closer to the dates of the non-serious corpuses. The important features in the recent news were also analysed. Then, punctuation features that represented sentence structure were analysed in the corpuses of recent news, the Onion, and Landover Baptist. The Naive Bayes classifier that used these punctuation features as well as word features was applied to the test corpus and produced ranking results that showed differentation between the serious and non-serious material. Finally, the experiments are summarised and discussed generally.

5.2 Corpus Information

5.2.1 Reuters Corpus

The Reuters corpus is 8700 articles from the Reuters news database [NIST, 2000], including articles from the Wall Street Journal. These articles were processed in prior work to remove words of length less than two and some stopwords; to remove words below a frequency threshold to leave a remaining vocabulary of 15,800 words; and to strip punctuation. Because of the size of the dataset, these articles were stored in the format:

(Article Number) [(Word Number):(Word Frequency)]*

For example:

4519 4:9 10:2 5:2 1940:1

The above means that article number 4519 used word number 4 (eg. 'Mining') nine times, word number 10 (eg. 'Steel') two times, word number 5 (eg. 'Union') two times, and word number 1940 (eg. 'Save') one time. The corpus was placed in this format as early work by the author.

This format was also used to store information from the other corpuses.

5.2.2 Onion Corpus

The Onion [Onion Inc., 2010] is a parody website with articles such as 'Movie Praised For Not Being As Bad As It Could Have Been' or 'Obama Depressed, Distant Since Battlestar Galactica Series Finale'. 2,700 articles were obtained from the Onion. After filtering for words occurring fewer than ten times across the corpus, the remaining vocabulary was 15,200 words (overlapping with the Reuters vocabulary and comparable in size).

5.2.3 Landover Corpus

Landover Baptist [Landover Baptist, 2010] is a parody website focused specifically upon American Christianity. 600 items were obtained from it. After filtering out words which occurred less than five times across the whole corpus, the remaining vocabulary was 6,800 words (trimmed from 23,500 words).

5.2.4 News Corpus

800 modern newspaper articles were taken from a range of sources including:

- GoogleNews (the most frequent source)
- MSN News
- Yahoo News
- ABC News (Australia)
- ABC News (United States)
- The Mirror Top Stories
- CBC News
- CNN
- New York Times (World)
- BBC News (UK news)
- BBC News (US and Canada news)

Articles linked by these sources were taken from the websites and processed for their vocabulary. This is a larger corpus than Landover, and a smaller corpus than Reuters or the Onion (even though in reality, serious articles outnumber parody articles). The news sources included the Globe and Mail, the Daily Telegraph (United Kingdom), Bangkok Reuters, Haaretz, and the Sydney Morning Herald, a global selection of news articles. In work, stopwords (short and common words often trimmed from most corpuses) including 'I' and 'we' were modified on the assumption that their natural prevalence in Reuters was an equivalent ratio to their proportion in the new corpus. After filtering words that occurred fewer than five times in the entire corpus, approximately 7,000 words remained as vocabulary. This was a slightly larger vocabulary on the same trimming as the Landover corpus, which contained 600 total articles.

This corpus was the newest corpus, taken from recent headlines. It contained vocabulary items not present in the Reuters corpus but present in Onion and Landover, such as 'Facebook' (a social networking website whose popularity postdates the time of the Reuters data collection). Since parody is reactive and serious news articles outnumber parody, not all articles in this corpus have parodies written about them.

5.2.5 Test Corpus

This small testing corpus was intended to be of material from sources not used in the other corpuses, a mix of serious and parody material. For example, the article from the parody blog [Swift, 2009] and the MoneyMorning article [Miller, 2009] both commented upon in the human case studies in Section 4. These tests also included material from Conservapedia (a politically based American encyclopedia with a right wing slant) and Wikipedia, as examples of informal wikis.

5.3 Naive Bayes

A Naive Bayes document classifier was used in a Python program [van Rossum, 2010] comparing the features of words and frequencies within the serious and non-serious corpuses. The Reuters corpus consisted of words and frequencies, so Naive Bayes was a useful algorithm to analyse the documents.

5.3.1 Initial Results: Reuters versus Onion

In 10-fold cross-validation testing of the Reuters corpus versus the Onion corpus, Naive Bayes results showed that it is possible to classify documents not trained upon as either Reuters or the Onion with a near-perfect degree of accuracy. This suggests that either the features are good or the problem is relatively simple. The standard error of the mean is the standard deviation of the error relative to the sample population. To give a 95% confidence interval, the standard error calculated is:

$$S_e = 1.96 \cdot \frac{\sigma}{\sqrt{n}}$$

(where σ is the standard deviation and n is the number of folds, ie. 10). This gives upper and lower 95% confidence limits of the mean number of errors:

$$\bar{x} \pm S_e$$

This means that to a 95% confidence limit the error rate will be between the given maximum and minimum. On the chart below, the indicated percentage of accuracy is shown. F-measure [van Rijsbergen, 1979] is an information retrieval measure that averages the precision (fraction of correct instances returned) and recall (fraction of instances returned that are correct). Positives are assumed to be serious news articles.

Table 10: Reuters versus Onion, Naive Bayes testing results

Folds	Average Fold Size	Average Error Total	Standard Deviation of Errors	Standard Error	Minimum Error	Minimum Error %	Maximum Error	Maximum Error %	F-measure
10	1116	21.9	14.85	9.20	12.70	1.14 %	31.10	2.79 %	0.987

These results show that Naive Bayes is effective in distinguishing between the Reuters corpus and the Onion corpus. There is a 95% confidence interval that the error rate will be between 1.14% and 2.79% for test data of the same corpuses but not present in the training data.

Parody is personalised. Examining the Onion corpus without stopwords stripped, the words 'we' and 'I' were found to be common in the parody texts. Parody often uses a personalised tone, though a personalised tone is also present in both parody and serious blogging. Table 11 shows the top frequent words in the Onion corpus, with stopwords included.

Table 11: Frequent words in the Onion corpus, stopwords included

word	frequency
the	113688
to	84771
of	67876
a	43862
and	43272
in	36422
for	23111
i	21416
that	20926
news	20132
said	19169
on	17826
onion	16758
is	16252
at	15535
it	14782
new	13895
we	13422

The words ‘we’ and ‘I’ are popular. This supports hypothesis 3, that parody is personalised. With stopwords removed based on the Onix list [Onix, 2010], the top frequent words on the Onion corpus are shown in Table **??**.

Table 12: Frequent words in the Onion corpus, stopwords removed

word	frequency
news	20132
onion	16758
report	10740
america	8641
film	8341
people	7871
stop	7182
dont	6999
2010	6897
time	6709
recent	6350
store	6035
close	6012

video	5943
employees	5900
mom	5896
door	5862
rights	5830
source	5735
locked	5630

This list includes the informal word ‘mom’, the contraction ‘dont’ (punctuation stripped from corpus), and the self-referential ‘news’. It is possible that parody self-references itself as ‘news’ more often than genuine news defines itself. Informality and contractions are parody features. The evidence of the popularity of informal words and contractions in contrast to the lack of them in the list of frequent Reuters terms supports hypothesis 4, that parody texts have more informal language than serious texts. However, it is important to note that serious news articles have more formal standards than non-edited serious blogging texts.

Finance is a popular subject in the Reuters serious corpus. The top frequent words on it with stopwords removed are:

Table 13: Frequent words in the Reuters corpus, stopwords removed

word	frequency
the	98833
and	32742
said	25555
for	16561
was	11907
that	10601
percent	9879
from	9732
with	9434
million	8372
year	7306
its	6959
will	6187
were	5967
market	5888
but	5869
has	5705
are	5654

not	5181
would	5092

'Percent', 'million', 'market', 'billion', 'company', 'shares', 'net' and others are all financial words. Finance is a subject with much serious material written about it, yet it is a smaller percentage of parody articles due to the difficulty in mocking articles with detailed financial mathematics. With no stopwords removed beyond those removed in pre-processing, the top frequent words were:

Table 14: Frequent words in the Reuters corpus, stopwords included

word	frequency
the	98833
and	32742
said	25555
for	16561
was	11907
that	10601
percent	9879
from	9732
with	9434
million	8372
year	7306
its	6959
will	6187
were	5967
market	5888
but	5869
has	5705
are	5654
not	5181
would	5092

Note that the financial words 'percent', 'million', and 'market' are more frequent than even words one expects to be commonly used in English like 'from', 'with', and 'were'. Prevalence of financial words is a legitimate feature to use in differentiating serious news and parody because of the common occurrence of financial words in serious news and the larger number of serious financial articles. This is evidence in support of hypothesis 5: that parodies and serious news show different trends in subject matter.

5.3.2 Reuters versus Onion and Landover

When comparing Reuters to the combined corpuses of the Onion and Landover, very high results were shown in distinguishing examples not drawn from the training data and identifying their source: in 10-fold cross validation, only one error was located despite a testing set of 1175 documents for each fold. This supports Hypothesis 1, that parody and serious texts can be automatically detected by the use of machine learning techniques. Reuters is used as a proxy for serious texts and the Onion and Landover as a proxy for parody texts. Stopwords were removed from all corpuses.

Naive Bayes is effective at differentiating these groups. They are used as proxies for serious news in general (Reuters) and parody news in general (Onion and Landover, a general parody site and a religious-oriented parody site). Although it was easy to train on these and test from material drawn from the same corpus with a high accuracy result, these corpuses are still useful as sources of features to study as examples of collected serious articles and non-serious articles. The top ranking features from the corpuses were examined for trends such as subject matter of serious versus parody, personalisation, and formality. To test whether the training on these sources is effective on different sources, the next section covers the testing and ranking of pieces drawn from the testing corpus as either serious or non-serious.

5.3.3 Ranking of new testing pieces.

These pieces selected for testing were drawn from sources of formal and informal news, blogging, and parody. The intent was to use diverse sources that were not used as corpuses for the training process. The below table describes the results and scores of the application of Naive Bayes to these new pieces, and discusses the features that were used to differentiate between them. They are ranked according to greatest to least likelihood of being serious articles. This was the same log-likelihood value that was used in the earlier tests on corpus articles, where a positive value indicates serious and a negative value indicates non-serious. These results show a heavy bias toward classifying new articles as non-serious, but the results also enable an analysis of which features of the articles affect a serious or a non-serious classification. These features are important to note for future applications in sentiment analysis.

Table 15: Test corpus ranked by degree of seriousness

Filename (Assigned Number)	File Description	Result (log likelihood score)	Reasons
News on Gulf Oil Spill (0)	A newspaper article regarding the BHP Oil Spill	Not serious (-1617)	The words 'restoration', 'tongues', 'spills', 'crime', 'contaminated', and 'autos are among the serious top scores; that 'contanimation' was on the list shows its use as a relatively formal word. The words that marked this by the programming as not serious included: 'bp' (a term used slightly more often in the modern parodies than Reuters), the use of self-referential 'news' (which the parody corpuses seem to use more often), and the informal 'we'.
CAP Alert Toy Story II (5)	A Christian website's serious review of a movie	Not serious (-1766)	Words judged as serious included 'predecessor', 'pulp' (a word used in general movie reviews with the double meaning of 'lumber'), other words often used in movie reviews such as 'animation', 'scoring', and 'dialogue'; and 'assumption', 'utility', and 'crime. Words taken as negatives were principally stopwords.

CNN on Gay Marriage (1)	A CNN news article on gay marriage	Not serious (-1800)	The words in this article giving top scores for a serious piece included 'plaintiffs' (legal term), 'unions' (in this case, marriage, but with the double meaning of 'trade'), 'Turkish' (national noun), and 'Commonwealth'. The words that turned it to non-serious included the simple stopwords 'a' and 'in' (more common in the newer parodies than in the older Reuters succinct style), the personal pronoun 'I' (a marker of a more informal piece), and the personal pronoun 'we'.
BBC Caroline Kennedy Election (11)	A BBC article featuring a story on Caroline Kennedy's election potential	Not serious (-1814)	Words judged as serious included 'programmes', 'downturn' (economic related), 'NATO', 'endorsed', and the relatively descriptive word 'revive'. Words judged as not serious were principally stopwords.
Money Morning's Obama Stimulus (4)	News article by MoneyMorning regarding President Obama's financial stimulus	Not serious (-1887)	Words in the article judged as serious included 'economies', 'shrink', 'worsen' (negative), and 'stagnant'. Words taken as non-serious again included stopwords and the word 'news'.
Obama Factcheck (10)	A blogged check regarding facts about President Obama	Not serious (-1902)	Words judged as serious included the business-related term 'mogul', the relatively emotive 'exhorted', 'nascent', 'insufficent', 'directory', and 'millers'. Words judged as not serious were principally stopwords.
NPR Kagan confiration (9)	An NPR news piece regarding the confirmation of Judge Elena Kagan	Not serious (-2103)	Words judged as serious included 'Kagan' itself, 'donate', the long and relatively dry word 'undoubtedly', and the legal terms 'judiciary' and "hearings'. Words judged as not serious were principally stopwords.

Obama Newsweek Opinion (8)	Opinion piece about President Obama	Not serious (-2115)	Words judged as serious included 'deadlock', 'trending', the name 'Gingrich', 'subscriber', and 'undermine'. Words judged as not serious were principally stopwords.
Blogging DADT (3)	A serious blog article on the Don't Ask Don't Tell policy	Not serious (-2131)	Words in this blogging article taken as serious were 'Blackwell' (famous legal jurist), 'stoking', 'defence' (legal term), 'goldman' (common surname), 'jindal' (politician's surname), and 'Palestine' (country). Words taken as non-serious again included stopwords and the word 'news'.
Frontpage Magazine Obama (12)	A magazine article about Obama in Iraq	Not serious (-3455)	Words judged as serious included 'jihad', the long and non-frivolous 'counterproductive', 'subscribers', and the formal 'likewise'. Words judged as not serious were principally stopwords and the word 'news'.
Dubai Debt Crisis (2)	Daily Mail article (news) regarding the Dubai debt crisis	Not serious (-3765)	Words in the article taken as serious included 'volkswagen' (vehicle noun), 'spearheading' (a metaphorical verb, which seems odd to be regarded as serious), 'mineral', and 'remedy' (can be a legal term). Top negatives again included many stopwords and the self-referential 'news'.
Caroline Kennedy Election (7)	A parody blog examining Caroline Kennedy's election potential to the United States Senate	Not serious (-6118)	Words judged as serious included 'stoking', 'jacqueline' (the name of a political figure), 'portfolio', 'sovereign', 'fritz' (oddly informal a word), 'eclipse', and 'Vancouver'. Stopwords including 'we' were prevalent in the non-serious judgment.

Wikipedia's Barack Obama (13)	The neutral Wikipedia article on Barack Obama	Not serious (-7520)	Words judged as serious included 'marches', 'strengthening', 'shrink' (used in relation to 'shrinking economy' rather than as slang for 'psychologist'), the descriptive 'flurry', 'wrangling', and the formal 'predominantly'. Words judged as not serious were principally stopwords and the word 'news'.
Conservapedia's Barack Obama (6)	An article about President Barack Obama from a politically biased United States webpage	Not serious (-8651)	Words judged as serious included 'failures', 'charities', the relatively emotive word 'boasted', and the names 'Theresa' (in that context, 'Theresa Heinz Kerry', the wife of a fellow politician) and 'Nader' (a political name). Words judged as not serious were principally stopwords.

In this ranking, Conservapedia and Wikipedia (both 'wikis', as defined by the *New Oxford American Dictionary*, websites where editing is open to the public) were ranked as the least 'serious'. This may have been for informal versus formal language. The Jon Swift blogging parody article was also considered low-ranked for 'seriousness'. The Gulf Oil Spill news article was viewed as most serious. This led to a theory that the bias was related to the testing articles being recent compared to Reuters: ie. because the Reuters vocabulary was based on older news articles, the newer articles would be closer to the Onion and Landover Baptist and therefore less likely to cause distortion related to age. Hypothesis 7, that adding recent news would increase accuracy in the classification of new texts as serious or non-serious, was tested by adding the news corpus.

5.3.4 Recent news: Ranking of testing pieces

To test the extent to which the different subject matter of the older Reuters corpus and the newer Onion and Landover corpuses affected results, the recent articles from the news corpus were used as serious examples to try again to test the pieces from sources not trained upon.

In the following test, the 800 recent news articles were added to the training corpus of serious material. Note that date clearly skewed the effect of the features used: the results showed that among the features added were words such as 'Obama', used more frequently in the recent news than in Reuters. The Reuters data collection took place in 1996 to 1997 prior to President Obama's election [NIST, 2000]. However, the relative ranking of the testing corpus from non-serious to serious remained the same as when only the Reuters corpus was used as serious.

Table 16: Second test showing ranking of most to least serious test articles using Reuters and news corpuses versus non-serious corpuses as training

Filename (Assigned Number)	File Description	Result (log likelihood score)	Reasons - non-serious features (NB: top serious features were similar to the table above, so non serious features are commented upon here)
CAP Alert Toy Story II (5)	A Christian website's serious review of a movie	Not serious (-930)	The words 'Jesus' (religion may be a more common subject for parody than for serious news), 'good' (an implied value judgment), and 'president' (perhaps Presidents are more often a subject of parody in relative terms compared to their prevalence in news)
News on Gulf Oil Spill (0)	A newspaper article regarding the BHP Oil Spill	Not serious (-990)	The word 'news', the word 'reserved' (a double meaning of 'shy' in personality and 'previously booked'), extraneous words
MoneyMorning Obama Stimulus (4)	News article by MoneyMorning regarding President Obama's financial stimulus	Not serious (-1146)	The words 'news', extraneous words, 'Obama', and 'China's' (the name of a country–can be serious)
Obama Factcheck (10)	A blogged check regarding facts about President Obama	Not serious (-1123)	The words 'Obama' and extraneous words
CNN on Gay Marriage (1)	A CNN news article on gay marriage	Not serious (-1144)	The words 'news', 'man', and extraneous words

BBC Caroline Kennedy Election (11)	A BBC article featuring a story on Caroline Kennedy's election potential	Not serious (-1145)	The words 'Obama', 'America's', 'Hollywood', 'policy', and extraneous words
NPR Kagan confiration (9)	An NPR news piece regarding the confirmation of Judge Elena Kagan	Not serious (-1248)	The words 'news', 'man', and extraneous words
Obama Newsweek Opinion (8)	Opinion piece about President Obama	Not serious (-1301)	The words 'policy', 'privacy', 'jobs' (all quite serious concepts), 'Obama, 'mobile' (possibly an effect of the older Reuters data), 'make' (active verb)
Blogging DADT (3)	A serious blog article on the Don't Ask Don't Tell policy	Not serious (-1424)	The words 'news', 'Obama', 'Iraq' (a recent subject for parody), 'Americas', 'policy' (a word with serious context as well), and 'center' (perhaps often selected in parody as a 'centrist' politician).
Frontpage Magazine Obama (12)	A magazine article about Obama in Iraq	Not serious (-2195)	The words 'Obama', 'news', 'man', and extraneous words
Dubai Debt Crisis (2)	Daily Mail article (news) regarding the Dubai debt crisis	Not serious (-2515)	The words 'news', 'Obama' (the Reuters corpus is older), and extraneous words
Caroline Kennedy Election (7)	A parody blog examining Caroline Kennedy's election potential to the United States Senate	Not serious (-4681)	The words 'news', 'man', 'Obama'

Wikipedia's Barack Obama (13)	The neutral Wikipedia article on Barack Obama	Not serious (-6034)	The words 'news', 'man', and extraneous words
Conservapedia's Barack Obama (6)	An article about President Barack Obama from a politically biased United States webpage	Not serious (-6928)	The words 'news', 'man', 'smart' (a value judgment), 'obama', 'area', extraneous words

The above results explain the detailed features that were statistically perceived in the training data, and point to interesting outcomes, such as the feature 'Obama' biasing toward non-serious because of a greater relative frequency in the non-serious corpuses than in Reuters and the 800 recent news pieces. Because the results continued to show a bias towards labelling as non-serious, the frequency-based weightings of serious versus non-serious words were examined on the hypothesis that the non-serious corpus had a smaller vocabulary and therefore frequency weightings given to individual words tended to be higher than words of an equivalent rank in the serious corpus, ie. Hypothesis 8. There was no indication of a tendency for common words in the serious corpus (ie. the words most highly ranked as 'serious') to have a lower frequency weighting than common words in the non-serious corpus. This indicates that the hypothesis that a word present in the non-serious corpus would receive a higher frequency rating than a word present in the serious corpus of the same rank is not supported by the data. The total word count of the non-serious corpus was smaller, but the number of articles was also smaller than the number of serious articles.

The differences between the news corpus and the Reuters corpus were further explored through assessing the overall accuracy of 10-fold cross-validation using the Naive Bayes classifier was less than the comparison of the Reuters corpus to the parody corpuses. The accuracy still remained much higher than the [Tsur et al., 2010] results on the Amazon review data. These results are in Table 17.

Table 17: Naive Bayes classification of Reuters and news corpuses versus the Onion and Landover corpuses.

Folds	Average Fold Size	Average Error Total	Standard Deviation of Errors	Standard Error	Minimum Error	Minimum Error %	Maximum Error	Maximum Error %	F-measure
10	1257	51.6	13.37	8.29	43.32	3.45 %	59.89	4.76 %	0.971

If the non-serious articles are taken as positives instead of the serious articles, the F-measure is then counted as 0.929. These results show errors in falsely classifiying serious documents as non-serious. The hypothesis was that the error rate was associated with test data from the news corpus misidentified as non-serious articles. This hypothesis was supported by the data: all but a small number of the errors (8) were of test data drawn from the recent news corpus. This demonstrated Hypothesis 8, that the errors occurred when the Reuters serious articles outweighed the recent news serious articles in number. The vocabulary used in the recent news overlapped with the vocabulary of the parody corpuses.

The above error rate suggested a hypothesis that the recent news articles and the Onion and Landover parody articles were more difficult to distinguish than the Reuters articles compared to the Onion and Landover. To test this, the Naive Bayes algorithm was run comparing only the news articles as the serious corpus to the non-serious corpuses. This meant that the parody articles now outnumbered the serious articles. The results of the 10-fold cross-validation are shown in Table 18.

Table 18: Naive Bayes classification of news corpus versus the Onion and Landover corpuses.

Folds	Average Fold Size	Average Error Total	Standard Deviation of Errors	Standard Error	Minimum Error	Minimum Error %	Maximum Error	Maximum Error %	F-measure
10	418	0.5	0.53	0.33	0.17	0.04 %	0.83	0.20 %	0.997

These results show that there is a clear difference between the news corpus and the Onion and Landover parody corpuses and that Naive Bayes produces good results in differentiating between them. Hypothesis 9 is not supported by the data: a Naive Bayes classifier will return high results if recent news are compared to recent parody texts, comparative to the high accuracy result of comparing older news to recent parody texts. Then, the features that were most important in the difference between the news corpus and the non-serious corpus were examined. Some features were simply byproducts of the data processing stage, eg. 'nbsp' (ie. a HTML space) and of the source (eg. 'businessday', the name of an online news source). Other features will be more significant for future work. For example, the nouns 'Chile' and 'Chilean' were highly ranked by frequency in the set of the words found in the serious corpus but not the non-serious corpuses, which shows that Chile is not a country frequently targeted for parody. (The source of the Chilean news was an incident of rescuing miners there that was a recent event.) This example supports Hypothesis 5, that parody and serious news show different trends in subject material. Although when the large number of Reuters articles are included in the serious corpus, recent news articles in the test data are misclassified as non-serious, when the news corpus is compared in isolation to the parody corpus the results show that a difference exists.

It is important to investigate what features are prominent in the recent news corpus. With stopwords removed, the most frequently occurring words in the news corpus were:

Table 19: Frequent words in the news corpus, stopwords removed.

word	frequency
news	1899
world	982

2010	971
people	648
rescue	554
false	550
miners	543
government	511
breaking	490
search	451
police	444
video	423
mr	421
national	402
copyright	369
report	369
president	359
minister	359
latest	347
israel	338

These lower frequencies show the relatively small size of that corpus. The frequency of the word 'news' shows that this selection of genuine news does in fact self-reference itself as news, so this may not be useful in differentiating serious from non-serious in future work. The words 'miners' and 'rescue' arise from a trending news topic of the rescue of Chilean miners. '2010' is clearly a contextual feature of the current year. To a certain extent, date is a legitimate feature in differentiating genuine news from parody because of the reactive nature of parody. The frequency of the word 'Mr' could suggest that genuine news refers to people with greater formality than parody. 'Copyright' would be a feature expected to occur both in genuine and parody. 'Latest' may be more common in general news than in parody because even though parody can use headlines such as 'Latest News: Dog Bites Man' in satires, news is more up-to-date and truly 'latest' compared to parody. 'Israel' was also a frequently occurring word in the news corpus, which was also likely related to topics trending at the time the data was gathered. The Onion contains articles such as 'Israel Bombs Anti-Semitism Out Of Lebanon', 'Israel Intercepts Massive Palestinian Rock Shipment', and 'Mel Gibson Launches Rockets Into Israel'; Israel is not an uncommon subject for parody as well as genuine news. The word 'video' was also high on the list; though the parody site The Onion includes videos, multimedia parodies are likely to be fewer than genuine news multimedia available.

5.4 Sentence Structure Features

[Tsur et al., 2010] used sentence features in the work of spotting sarcasm in

Amazon reviews. Using the news corpus meant that punctuation features in that serious corpus could be compared to punctuation features in the Onion and Landover parody corpuses, because the Reuters data was based only on words with punctuation stripped.

With punctuation included in the feature set as well as words, the contraction signals of "'t" and "'s" were respectively 0.17% and 1.28% of the total words in the set of 817 news files other than Reuters. Comparatively, in the non-serious corpuses, these were 0.89% and 1.89%. This suggests a greater frequency of contractions such as 'didn't', 'can't', 'shouldn't', and so on in the more informal corpus domains of the Onion and Landover parodies. Although in the results below, these contraction indicators rarely appeared as the strongest out of the 'non-serious' features, a statistical difference has been shown that demonstrates a difference in this feature aspect. This supports Hypothesis 10, that punctuation and sentence structure features are important differences between parody and serious news.

Using punctuation features and a training corpus of the recent news corpus versus the Onion and Landover non-serious corpuses, the test corpuses were ranked again by seriousness based on this training data. This shows a clear difference in the ranking of the articles in the test corpus, and shows an improved ranking based on this usage of only recent news, recent parodies, and featurs including both words and punctuation. The order of the articles in terms of log-likelihood score was:

Table 20: Test corpus ranked by degree of seriousness using punctuation and word features from the recent news corpus and the Onion and Landover corpuses.

Filename (Assigned Number)	File Description	Result (log likelihood score)	Reasons - principal serious and non-serious features

BBC Caroline Kennedy Election (11)	A BBC article featuring a story on Caroline Kennedy's election potential	Not serious (-193.95)	Serious words included 'NATO' (a serious organisation), 'BBC' (the news organisation in question; a 'cheating' feature), the abbreviation 'GMT' (illustrating that the serious articles are from a more diverse range of countries than the non-serious largely American comedy), and 'clashes' (a fairly descriptive verb). non-serious words included 'hollywood', 'plastic', 'america', and 'corporation' (the last two should also be frequent in serious material).
News on Gulf Oil Spill (0)	A newspaper article regarding the BHP Oil Spill	Not serious (-340.44)	Serious features now included such words as 'blogs', 'advertise', and 'tweet': signs that the other news media gathered uses these tools. 'deepwater', 'sponsored', and 'fishermen' were also considered serious. There are probably more serious news articles on the subject matter of 'fishermen' than there are parodies about them. Top negatives for non-serious included the word 'eating' (probably more common for restaurant reviews than in serious news generally), 'newest', and 'film'.
CAP Alert Toy Story II (5)	A Christian website's serious review of a movie	Not serious (-356.91)	Serious words included 'profanity' (this was surprising, since as a religious concept the word is more likely to be found in parodies and in only a few religious articles rather than most average news articles), 'utility', and 'rescue' (likely from the serious articles on the 2010 Chilean rescue effort). non-serious words included 'Jesus', 'lord', 'christ', the informal 'folks', 'wisdom' (a subjective value), and 'thomas'.

Obama Factcheck (10)	A blogged check regarding facts about President Obama	Not serious (-367.99)	Serious words included 'advertise', 'chamberlain' (the name of a British politician notable for the appeasement policy), 'nationals', 'applicable', 'statehood', and 'utility'. non-serious words included 'tip', 'fan', 'pants' (possibly a more common subject for commentary in parodic than serious articles), and 'theater'.
MoneyMorning Obama Stimulus (4)	News article by MoneyMorning regarding President Obama's financial stimulus	Not serious (-396.12)	Serious words included 'moratorium' (a serious concept), 'governments', 'economies', and 'world'. non-serious words included 'mcdonald' (the name of a popular fast food restaurant), 'favorite' (a word implying a value judgment), and 'congressional' (a political word).
CNN on Gay Marriage (1)	A CNN news article on gay marriage	Not serious (-409.75)	Serious features included 'commonwealth' (perhaps more common in serious news articles than in parody articles targeting the Commonwealth Gamse, referring to the Commonwealth of Australia, etc), 'unions' (once more, a word that can mean both 'trade' and 'marital' unions), and 'endorsed' (a formal word). non-serious features included 'Bush', 'Kerry' (the serious corpus was limited to the 2010 news, therefore the Onion/Landover corpus contained more articles featuring these American politicians than the genuine news), the abbreviation 'Sen', and 'divorce'.

Obama Newsweek Opinion (8)	Opinion piece about President Obama	Not serious (-412.20)	Serious included 'chile' (a link from the original article to related news, referring to a country not often parodied), 'hungarian', 'intel', and the name 'Abbas'. non-serious included 'rogers' (a man's name), 'funny' (a value judgment besides a reference to comedy), 'america' (a word that should be common in both American parody and American news sources, but the news files lifted were from a wider variety of countries than the Reuters data), and 'screaming'.
NPR Kagan confiration (9)	An NPR news piece regarding the confirmation of Judge Elena Kagan	Not serious (-429.95)	Serious words included 'ombudsman', 'permissions' (a legal/dry term), 'podcasts' (possibly more common for genuine news than for parodies, though may become increasingly common for parody websites), 'corrections' (the name of a United States justice department), 'judiciary' (legal), and 'recruiters' (non emotive word). non-serious words included 'stumble' (descriptive verb), 'song', 'admit', and 'science' (possibly due to Landover articles parodying particular sects' opposition to science, and because of a relatively small selection of genuine science articles selected for the news corpus).
Frontpage Magazine Obama (12)	A magazine article about Obama in Iraq	Not Sserious (-843.93)	Serious words included 'governments', 'Petraeus' (an American military official), 'Lebanon' (a country perhaps not often parodied). non-serious words included 'barrels', 'Bush', and 'secular'.

Blogging DADT (3)	A serious blog article on the Don't Ask Don't Tell policy	Not serious (-871.02)	Serious words included 'Petraeus' (the name of an American general in the news), 'endorsed' (a relatively nondescript verb), the abbreviation pm, the names of the American talk-show celebrities 'reilly' and 'beck', and the non-humorous organisation 'Taliban'. Top non-serious words again included the politicians who reached heights of popularity in serious news pre-2010, 'Bush', 'Kerry', and 'Cheney'.
Dubai Debt Crisis (2)	Daily Mail article (news) regarding the Dubai debt crisis	Not serious (-957.13)	Serious features included 'Australian' (perhaps an uncommon target for parodies, particularly those that originate in the United States), 'Zealand', 'UK', 'defence', and 'ombudsman' (ombudspeople are usually quite serious). non-serious features inculded 'coffee', 'suspect', 'slump' (a fairly descriptive noun), 'lord' (possibly based on the Landover Corpus' parodies of sects of American Christianity), and 'prayer'.
Caroline Kennedy Election (7)	A parody blog examining Caroline Kennedy's election potential to the United States Senate	Not serious (-3013.38)	Serious included 'blogs', 'tweets'. non-serious included 'hollywood' and 'celebrities'.
Wikipedia's Barack Obama (13)	The neutral Wikipedia article on Barack Obama	Not serious (-3284.28)	Serious words included 'BBC', 'Guardian', 'Australian' (all news sources that were used in the corpus of non-Reuters news); non-serious words included 'celebrities', 'Jesus', 'pastor' (religious), and 'smart' (value judgement).

Conservapedia's Barack Obama (6)	An article about President Barack Obama from a politically biased United States webpage	Not serious (-3792.52)	Serious included 'guardian' (the name of a serious newspaper) and 'ap' (another newspaper). non-serious words included 'pastor' (religious, but the word also appears in serious news articles regarding Obama's pastor), 'jesus', 'adults', the informal 'mom', and 'mounting'.

This reordering shows that Conservapedia, Wikipedia, and Jon Swift's parody blogging were ranked at the bottom in terms of relative seriousness (two wikis and a genuine parody). CAP Alert's blogged review was ranked as more serious than the MoneyMorning article upon the stimulus, even though it is informal language compared to the former. This shows that the feature set may be capturing more than simple formality and informality. The serious blogging article concerning 'factchecking' of Obama was also rated as relatively serious compared to the other results. Part of this seems to be due to keyword trends in the data sets in question. However, this testing showed some interesting features, features such as: the word 'celebrities' for non-serious (celebrities may be a frequent target for parody, though they are also written about in magazines of a genre not used for the purposes of these experiments); words such as 'moratorium' in serious; serious features that indicate the use of media such as podcasts and tweets; non-serious use of sometimes-punctuated abbreviations such as 'Sen'; and non-serious use of the informal word 'folks'. These features applied to the test data and discussed above are significant for future feature engineering. This use of only the news corpus compared to the combined Onion/Landover corpus has shown a more sensible overall ordering of the testing data in relative seriousness than the first tests.

5.5 Summary of Experiments

The first experiments carried out showed that Naive Bayes was effective in distinguishing between the Reuters corpus and the Onion corpus, supporting Hypothesis 1. Likewise, Naive Bayes was effective in distinguishing between the Reuters corpus and the combined Landover and Onion corpuses. In both cases, this was supported by testing on data that was not trained upon. Analysing the features used to differentiate between these parody and serious corpuses, relevant matters included the large number of financial articles in the Reuters data, showing that finance is not a common subject for parody and is a feature that indicates against the likelihood of a given text being parodic; informal words frequent in the parodic texts; and that personal pronouns were among the most frequent words used in the non-serious texts. This supported Hypotheses 3, 4, and 5.

In applying Naive Bayes to the test corpus of material drawn from neither the Reuters nor the non-serious corpuses, results showed that the stopword bias toward the parody corpuses affected the classification of the texts as non-serious. The top features that caused a given text to be classified as either serious or parody were discussed. Serious word features included more formal or neutral words such as ‘contamination’ (rather than ‘spill’ or ‘pollution’), financial terms, the names of countries not commonly parodied in the English-speaking non-serious websites drawn from (eg. ‘Turkey’), legal terms (eg. ‘judiciary’); non-serious word features included religious words such as ‘Jesus’. These features are important for future work.

By introducing additional news data to try again to differentiate the test corpus material into serious and non-serious, the results helped to remove some bias toward current events represented in the parody corpuses versus the older data represented by the Reuters corpus. Testing on differentiating Reuters and the news corpus as serious versus the Onion and Landover as non-serious showed a low error rate, but a high amount of false negatives (serious material misclassified as non-serious). These false negatives largely consisted of the recent news. However, when the recent news was isolated as a corpus and then compared to the Onion and Landover, the error rate returned to a very small percentage. Hypothesis 9, that recent news articles only against recent parody texts would cause the accuracy to significantly decrease, was not supported. This again supported Hupothesis 1, that Naive Bayes is an effective way of distinguishing between serious and parody texts as represented by the proxies of recent news, the Onion, and Landover. Many frequent words in the news corpuses related to events that were covered in them, for example, ‘miners’ in relation to a recent news event. Other words like ‘breaking’ and ‘latest’ are more useful features for serious news (parody is reactive), or words that refer to uncommonly parodied subjects, eg. ‘rescues’ are not often considered humorous.

Punctuation features were also used to compare the news corpus to the Onion and Landover non-serious corpuses. The results showed that the contraction signals ‘”t”’ and ”’s’” had a greater rank (occurred with greater frequency relative to the total word count) in the non-serious corpuses compared to the serious corpuses. However, in using these features to apply Naive Bayes to the test corpus, contractions did not appear among the top features that were used to analyse new text for serious or non-serious classification. Many word features for future feature engineering were extracted, such as words like ‘moratorium’ in serious pieces and ‘celebrities’ for non-serious. Using only the news corpus versus the Onion and Landover corpuses showed the best result in terms of classifying the new documents from the test corpus. This supported Hypothesis 10 that the punctuation-based features were useful.

These experiments showed the emergence of important feature information from the Reuters, news, Onion, Landover, and testing corpuses as five samples of serious and non-serious news. Using these as proxies for general detection of serious and non-serious sentiment, the features uncovered here have provided good results in differentiating between the given corpuses and can be used in future work on new texts.

6 Future Work and Conclusion

The experiments showed that there exist differences between serious and non-serious articles in the given data sets based on word frequency alone. These results were very successful in discriminating between data from the same sources, but heavily biased towards classifying new material as a parody. For future work, tools of sentiment analysis and of the feature engineering discussed can be used to continue to expand the capacity to automatically detect in this area. The human analysis in Section 4 and the features discussed in Section 5 provide a background of material for future practical work and expansion that uses these features discerned.

The analysis of the corpuses separately, including the huge Reuters corpus, contributed information on prominent trends in serious and non-serious material. In Naive Bayes testing on whether there was a clear difference between the corpuses, there was a strong difference between Reuters and the parody corpuses the Onion and Landover Baptist. When the news corpus was incorporated into the serious corpus, the error rate increased in distinguishing between the serious and non-serious corpuses. This was due to misclassification of the test data from the news corpus. However, the news corpus on its own compared to the Onion and Landover showed better accuracy on test data drawn from the same corpuses used to train.

In applying the trained Naive Bayes classifier to the test corpus, results showed a bias toward classifying new articles as non-serious. Analysing the features used to classify the articles was important since it enabled human understanding of the key features that were correctly (or oddly) ascribed to serious or non-serious. The best ranking achieved on the test corpus was by training on only the news corpus and the Onion/Landover parody corpus as serious and non-serious respectively, and using punctuation features in both. This showed that biasing the training set toward recent news provided more accurate results in differentiating the relatively recent articles in the test set. However, the analysis that used the Reuters corpus was also valuable since that corpus is so large and a substantial source of serious news articles.

As described by the human analysis, some forms of parody are very difficult to detect, and future work in machine learning in this area may never be able to provide perfect solutions.

6.1 Future expansions

6.1.1 Parsing Expansions

Parsing uses syntactic as well as semantic information. Future work should incorporate parsing for parts-of-speech and features such as whether a word is an adjective or adverb. As the human analysis of parody and serious texts found, in parodies descriptive language is often used for emphasis. Therefore, the number of adjectives and adverbs used in a particular work would be a feature to use in expanded work. [Turney, 2002] is on average semantic orientation of adjectives, as is [Hatzivassiloglou and McKeown, 1997]. The semantic orientation of adjectives can be predicted. However, parodies would be expected to show an excess number of positive adjectives in some cases, or an unusual mixture of positive and negative adjectives. For example, a parody article might depict a parent saying, 'This wonderful movie is bloodthirsty, violent, and evil. I can't wait to take my eight year old to watch it!' In studying the human features of parody articles, the parody most notable for excessive description was the parody of Don't Ask Don't Tell, where it used prose similar to a romance novel.

Dependency analysis also affects parody: as discussed in Section 4, a common technique of parody is the reversal of subject and object ('Massachusetts ALLOWS same-sex marriage', versus 'Same-sex marriage now MANDATORY in Massachusetts'). [Bai et al., 2004] examine dependencies through learning algorithms rather than parsing. Dependency analysis and part of speech tagging can improve the accuracy of finding the distinction between parody and non parody, especially on tests from new corpuses.

6.1.2 Affective Expansions

Some of the results were surprising in that words relatively emotive were counted among 'serious' features. For example, 'stumble', 'exhorted', 'smart'. Utilising affective theories to expand the feature set, such as using affective numeric values to each word from SentiWordNet, will add to classification accuracy. A word like 'resolved' is stronger than a word like 'decided': if on an affective intensity measurement the former is stronger, then that will bias the likelihood that the former will be used in non-formal material such as blogs and parodies. More emotive and judgmental language is found in parody material than in balanced articles that seek to report on news or to argue a point of view. The expansion of affective categories and the understanding, first by [Osgood et al., 1957] and later affective categories, appraisal theory, and emotional judgment, shows great scope for future work by the range of features that can be judged and examined to provide an analysis of the sentiment of an article.

The results also suggest a difference of 'style': there appears to be a large difference in style between the chosen serious and non-serious corpuses, as demonstrated by the high accuracy in differentiating the corpuses using test data drawn from the same corpuses. When tested on the ranked articles from new sources, there was a bias toward classifying these new articles as non-serious.

6.1.3 Formal and Informal

Many of the features that differentiated between the serious and non-serious corpuses were features that differentiated between formal and informal sources: for example, 'folks', 'mom', contractions. These are not features that uniquely differentiate between parody and non-parody texts. However, in comparing recent news to the parody texts, the results showed some blogging material ranked as relatively serious in comparison to parody blogs and the wiki texts. This indicates a positive likelihood that future work can differentiate between the vocabulary of serious work and the emotive, exaggerated vocabulary of parodies.

6.2 Conclusion

Contributions of this thesis are in providing the background and literature review of sentiment analysis; conducting a human analysis of features that differentiated serious text from parody; and conducting automated experiments that demonstrated positive results in differentiating given serious and non-serious corpuses. These automated results were analysed according to specific features that affected the classification of new test articles. The analysis of the features used in this work will help to build future sentiment tools that distinguish parody from non-parody.

Understanding the background to linguistic theory is important to appreciate sentiment analysis and its potential. After exploring linguistic theory in general and discussing some examples of and the nature of parodies, satires, sarcasms and other instances of humour, the field of sentiment analysis was then reviewed. The many possibilities and uses of sentiment analysis were viewed as prior work, including a resource list of datasets available for sentiment analysis and a discussion of the many available classification schemas. This was followed by specific, human analysis of multiple case studies, to identify potential available features in the work of machine learning. Very difficult examples were also given as complex, future work requiring contextual analysis and parsing that goes beyond simple part of speech or dependency analysis. In the results, Naive Bayes was very successful in differentiating between the serious corpuses of Reuters and other news, and the Landover and Onion non-serious corpuses. For future work, techniques such as feature expansion through part of speech tagging and dependency analysis should be used to expand the potential of recognising a parodist's work. Nevertheless, material such as the Flat Earth Society and the difficult parodies quoted in Section 4.4 will remain problematic to human and machine classification.

References

N Agarwal, H Liu, L Tang, and P Yu. Identifying the influential bloggers. *WDSM*, 2008.

A Aue and M Gamon. Customizing sentiment classifiers to new domains: a case study. *RANLP-05, International Conference on Recent Advances in Natural Language Processing*, 2005.

The Australian. Bhp weighs gulf of mexico impact. *The Australian (newspaper)*, May 2010. Date 29/05/10, Accessed 01/06/10.

X Bai, R Padman, and E Airoldi. Sentiment extraction from unstructured text using tabu search enhanced markov blanket. *Proceedings of the International Workshop on Mining for and from the Semantic Web*, pages 24–35, 2004.

EL Battistella. Markedness: the evaluative superstructure of language. *State University of New York Press*, 1990.

J Bos and M Nissim. An empirical approach to the interpretation of superlatives. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, pages 9–7, 2006.

SRK Branavan, H Chen, J Eisenstein, and R Barzilay. Learning document-level semantic properties from free-text annotations. *Proceedings of the Annual Conference of the Association for Computational Linguistics*, 2008. Accessed 18/02/10.

W Buntine. Document analysis. *Summer School in Logic and Learning,*, 2009. tutorial given.

CAPAlert. Capalert, 2010. URL `http://www.capalert.com`.

HS Cheang and MD Pell. The sound of sarcasm. *Speech Communications*, 50 (5), 2008.

X Chen, H Gao, and Y Fu. Situation analysis and prediction of web public sentiment. *Proceedings of the 2008 International Symposium on Information Sciene and Engineering*, 2:707–710, 2008.

Y Choi, Y Kim, and S-H Myaeng. Domain-specific sentiment analysis using contextual feature generation. *Proeedings of the 1st International CIKM Workshop on Topic-Sentiment Analysis for Mass Opinion*, pages 37–44, 2009.

KW Church and P Hanks. Word association norms, mutual information and lexicography. *Proceedings of the 27th Annual Meeting on Association for Computational Linguistics*, pages 76–83, 1989.

WW Cohen, VR Carvalho, and TM Mitchell. Learning to classify email into 'speech acts'. *Proceedings of EMNLP*, 2004.

FSTDT Community. Fundamentalists say the darndest things, 2010. Webpage accessed 15/06/10.

N Dalvi, R Kumar, B Pang, and A Tomkins. Matching objects to reviews using a language model. *2009 Conference on Empirical Methods in Natural Language Processing*, 2009.

Das and M Chen. Yahoo! for amazon: opinion extraction from small talk on the web. *SSRN*, 2001. Accessed 25/03/10.

K Dave, S Lawrence, and DM Pennock. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. *Proceedings of WWW*, pages 519–528, 2003.

D Davidov and A Rapoport. Efficient unsupervised discovery of word categories using symmetric patterns and high frequency words. *COLING-ACL 2006*, 2006.

X Ding, B Liu, and PS Yu. A holistic lexicon-based approach to opinion mining. *WDSM 2008*, 2008.

KR Scherer (ed). Appendix f: labels describing affective states in five major languages. *Facets of Emotion: Recent Research*, 1988. Version revised by the members of the Geneva Emotion Research Group; acccessed 29/02/10.

M Efron. The liberal media and right-wing conspiracies: using cocitation information to estimate political orientation in web documents. *Proceedings of the Thirteenth ACM Conference on Information and Knowledge Management*, pages 390–398, 2004.

DF Farkas and KE Kiss. On the comparative and absolute readings of superlatives. *Natural Language and Linguistic Theory*, 18:417–455, 2000.

M Finkelstein. Rep. gohmert: Dadt repeal would require troops to be 'overt about their sexuality'. *Media Matters*, 2010. Accessed 01/08/10.

D Freitag. Machine learning for information extraction in informal domains. *Machine Learning*, 39(2-3):169–202, 2004.

JM Gawron. Comparatives, superlatives and resolution. *Linguistics and Philosophy*, 18:333–380, 1995.

F Glover and M Laguna. Tabu search. 2004. Accessed 28/04/10.

SC Greene. Spin: Lexical semantics, transitivity and the identification of implicit sentiment. *Dissertation Abstracts International, A: The Humanities and Social Sciences*, 68(08):3368, 2008.

T Nasukawa H Kanayama and H Watanabe. Deeper sentiment analysis using machine translation technology. *Proceedings of the 20th International Conference on Computational Linguistics*, page 494, 2004.

L Haegeman. The syntax of negation. 1995.

V Hatzivassiloglou and KR McKeown. Predicting the semantic orientation of adjectives. *Proceedings of the 35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL*, pages 174–181, 1997.

M Hearst. *Direction-based text interpretation as an information access refinement*, pages 257–294. L. Erlbaum Associates, Hillsdale, New Jersey, 1992.

M Hu and B Liu. Mining and summarizing customer reviews. *International Conference on Knowledge Discovery and Data Mining*, pages 168–177, 2004.

N Jindal and B Liu. Opinion spam and analysis. *Proceedings of the international conference on web search and web data mining*, pages 219–230, 2008.

T Joachims. Text categorization with support vector machines: learning with many relevant features. *Proceedings of ECML 1998*, pages 137–142, 1998.

S-M Kim and E Hovy. Determining the sentiment of opinions. *International Conference on Computational Linguistics*, page 1367, 2004.

RJ Kreuz and GM Caucci. Lexical influences on the perception of sarasm. *Proceedings of the Workshop on Computational Approaches to Figurative Language, Association for Computation al Linguistics*, 2007.

RJ Kreuz and RM Roberts. Two cues for verbal irony: hyperbole and the ironic tone of voice. *Metaphor and Symbolic Activity*, 10:21–31, 1995.

Landover Baptist. Landover baptist, 'where the worthwhile worship, 2010. URL `landoverbaptist.org`. Accessed 10/10/10.

A Lehrer. Markedness and antonymy. *Journal of Linguistics*, 31(3):397–429, 1985.

E Lim, B Liu, V Ngyuen, N Jindal, and H Lauw. Detecting product review spammers using rating behaviours. *CIKM 2010*, 2010.

B Liu. Sentiment analysis: a multi-faced problem. to appear in IEEE Intelligent Systems, 2010; Accessed 01/11/10, 2010.

J Liu and S Seneff. Review sentiment scoring via a parse-and-paraphrase paradigm. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 161–169, 2009.

C Manning and H Schutze. *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.

A McCallum and K Nigam. A comparison of event models for naive bayes text classification. *AAAI-98 Workshop on Learning for Text Classification*, pages 4–15, 1998.

P Melville, W Gryc, and RD Lawrence. Sentiment analysis of blogs by combining lexical knowledge with text classification. *International Conference on Knowledge Discovery and Data Mining*, 2009.

Microsoft Office. Tonecheck program, 2010.

B Mikkelson and DP Mikkelson. Harry potter satanism. *Snopes*, 2008.

Don Miller. Obama offers new stimulus package to create more jobs. *Money Morning*, 2009.

K Miriyala and MT Harandi. Automatic derivation of formal software specifications from informal descriptions. *IEEE Transactions on Software Engineering*, 17(10):1126–1142, 1991.

R Narayanan, B Liu, and A Choudhary. Sentiment analysis of conditional sentences. *Conference on Empirical Methods in Natural Language Processing*, pages 180–189, August 2009.

T Nasukawa and J Yi. Sentiment analysis: capturing favourability using natural language processing. *Proceedings of the 2nd International Conference on Knowledge Capture*, pages 70–77, 2003.

K Nigam, AK McCallum, S Thrun, and T Mitchell. Text classification from labeled and unlabeled documents using em. *Machine Learning*, 39:103–134, 2000.

NIST. Articles from reuters rcv2, 2000. URL `http://trec.nist.gov/data/reuters/reuters.html`.

A O'Neill. Sentiment mining for natural language documents. *COMP3006 Project Report*, 2009. Accessed 22/02/10.

Onion Inc. The onion, 2010.

Onix. List of stopwords, 2010. URL `http://www.lextek.com/manuals/onix/stopwords1.html`.

CE Osgood, GJ Suci, and PH Tannenbaum. The measurement of meaning. 1957.

B Pang and L Lee. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, July 2004.

B Pang and L Lee. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, June 2005.

B Pang and L Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(2):1–135, January 2008. Accessed 10/03/10.

Associated Press. *Associated Press Stylebook*. Associated Press, 2010. Accessed 03/07/10.

JR Quinlan. Programs for machine learning. 1993.

WM Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66:846–850, 1971.

Seuss. Green eggs and ham. 1960.

A Sheth, C Ramakrishnan, and C Thomas. Semantics for the semantic web: the implicit, the formal and the powerful. *International Journal on Semantic Web and Information Systems*, 1(1):1–18, 2005.

Flat Earth Society. Flat earth society, 2010. URL `http://www.theflatearthsociety.org/`.

E Spertus. Smokey: Automatic recognition of hostile messages. *Innovative Applications of Artificial Intelligence (IAAI '97)*, 1997.

NICTA Statistical Machine Learning Group. Opinionwatch tool, 2010.

C Strapparava, A Valitutti, and O Stock. The affective weight of lexicon. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, pages 423–426, 2006.

P Subasic and A Huettner. Affect analysis of text using fuzzy semantic typing. 9:483–496, August 2001.

Jon Swift. Jon swift blog, 2009. URL `http://jonswift.blogspot.com`.

M Thomas, B Pang, and L Lee. Get out the vote: determining support or opposition from congressional floor-debate transcripts. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, 2006.

L Todorovski and S Dzeroski. Combining classifiers with meta decision trees. *Machine Learning*, 50:223–249, 2003.

RM Tong. An operational system for detecting and tracking opinions in on-line discussions. *Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification*, pages 1–6, 2001.

V Tsang and S Stevenson. A graph-theoretic framework for semantic distance. *Computational Linguistics*, 36(1), 2010.

O Tsur, D Davidov, and A Rappoport. Icwsm - a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. *AAAI*, 2010.

PD Turney. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Associatio n for Computational Linguistics*, pages 417–424, 2002.

Faculty of Law University of Missouri Kansas City. Oj simpson trial: Satire. 2010.

A Utsumi. Verbal irony as an implicit display of ironic environment: distinguishing ironic utterances from nonirony. *Journal of Pragmatics*, 32:1777–1806, 2000.

CJ van Rijsbergen. *Information Retrieval (2nd ed)*. Butterworth, 1979.

Guido van Rossum. Python, 2010. URL `http://www.python.org`.

S Wang, Y Fu, and H Gao. Phase detection and prediction web public sentiment. *Proceedings of the 2009 International Conference on Web Information Systems and Mining*, pages 113–117, 2009.

C Whitelaw, N Garg, and S Argamon. Using appraisal groups for sentiment analysis. *CIKM'05*, November 2005.

J Wiebe. Instructions for annotating opinions in newspaper articles, technical report tr-02-101. 2002.

T Wilson, J Wiebe, and P Hoffmann. Recognizing contextual polarity: an exploration of features for phrase-level sentiment analysis. *Association for Computational Linguistics*, 2009.

X Wu. On the scope of negation in english. 2(9):53–56, 2005. Accessed 18/02/10.

J Yi, T Nasukawa, R Bunescu, and W Nibliack. Sentiment analyzer: extracting sentiments about a given topic using natural language processing techniques. *Third IEEE International Conference on Data Mining (ICDM '03)*, pages 427–435, 2003.

T Zagibalov and J Carroll. Automatic seed word selection for unsupervised sentiment classification of chinese text. *Proceedings of the 22nd International Conference on Computational Linguistics*, 1:1073–1080, 2003.