Humor Detection

Classification Approaches

Miriam Käshammer

Course: Computational Approaches to Creative Language
Universität des Saarlandes

June 29, 2010

Humor Detection/Humor Recognition

Automatically decide whether a document (e.g. text, paragraph, sentence etc.) is humorous or not.

Challenging task

- **What is humor?**
  - Most commonly: what makes people laugh is humorous.
  - Many definitions and theories of humor (philosophy, linguistics, psychology)
  - Here: verbally expressed humor; subclass: jokes
- Sense of humor varies from person to person.
- Humor recognition is not an easy task for humans.

Motivation

Humor is an essential element of all verbal communication.

- Natural language systems should be able to handle humor.
    - Improve user-friendliness
    - Improve human-computer interaction
    - Develop better intelligent systems
    - Improve second language learning
- Information retrieval: Discard unnecessary or irrelevant information
- Search for unintended humor in serious text, i.e. diplomatic note, presidential speech

Approaches

Classification techniques

Learning to Laugh (Automatically): Computational Models for Humor Recognition
Mihalcea, R. and Strapparava, C. (2006)

Identification of "Knock-Knock" jokes, heuristics based on the particular joke structure

Computationally Recognizing Wordplay in Jokes
Taylor, J. M. and Mazlack, L. J. (2004)

Ontological semantics

Computational Detection of Humor: A Dream or A Nightmare?
Taylor, J. M. (2009)

Outline

1. Introduction
2. Classification Approaches
3. Evaluation
4. Conclusion

Outline

Overall idea

Hypothesis: Automatic classification techniques represent a viable approach to distinguish between humorous and non-humorous text.

1. Labels: humorous vs. non-humorous
2. Define features:
   - A. humor-specific stylistic features (heuristics)
   - B. content-based features
   - C. combined stylistic and content-based features
3. Use labeled data to train a classifier:
   - A. decision tree
   - B. Naive Bayes and SVM
   - C. stacked machine learning framework
4. Use the classifier to label the test data and evaluate

Restrictions on the Type of Humor

One-liner

`Take my advice; I don't use it anyway.`

"[...] a short sentence with comic effects and an interesting linguistic structure [...]"

- simple syntax
- deliberate use of rhetoric devices (alliteration, rhyme etc.)
- frequent use of creative language constructions to attract the readers' attention

- Humor in general: too ambitious
- Humor-producing features are present in this one and only sentence.

Outline

Humorous Data

Web-based bootstrapping process
[Mihalcea and Strapparava, 2005]

$\Rightarrow$ 16.000 one-liners for the experiments

```
I get enough exercise just pushing my luck.
Beauty is in the eye of the beer holder.
Take my advice; I don't use it anyway.
```

Non-humorous Data

Four collections of 16.000 sentences from different sources

- Reuters/BNC/Proverbs/OMCS
- similar in structure and composition to the one-liners, but different in their comic effect
- average length of 10-15 words

$\Rightarrow$ Enforce the classifiers to identify humor-specific features

Non-humorous Data contd.

Reuters titles

from articles published in the Reuters newswire

- short sentences with simple syntax
- phrased to catch the readers' attention

```
Oil prices slip as refiners shop for bargains.
```

Non-humorous Data contd.

Reuters titles

from articles published in the Reuters newswire

- short sentences with simple syntax
- phrased to catch the readers' attention

`Oil prices slip as refiners shop for bargains.`

BNC

sentences from the British National Corpus

- no added creativity

`The train arrives three minutes early.`

Non-humorous Data contd.

Proverbs

from an online proverb collection

- condensed sayings
- Some one-liners reproduce proverbs with a comic effect.

`Beauty is in the eye of the beholder.`

Non-humorous Data contd.

Proverbs

from an online proverb collection

- condensed sayings
- Some one-liners reproduce proverbs with a comic effect.

`Beauty is in the eye of the beholder.`

OMCS

commonsense assertions in English

- simple single sentences
- Jokes often break commonsense understandings.

`A file is used for keeping documents.`

Outline

1. Introduction
2. Classification Approaches
   - Data Sets
   - A. Classification Using Heuristics
   - B. Text Classification
   - C. Combination
3. Evaluation
   - Experimental Results
   - Discussion
4. Conclusion

Stylistic Features

Linguistic theories of humor have suggested stylistic features that characterize humorous text.

Significant and feasible to implement:

- Alliteration
- Antonymy
- Adult slang

Stylistic Features: Alliteration

- Structural and phonetic properties of jokes are at least as important as their content.
- One-liners often rely on attention-catching sounds, enforced through alliteration and word repetition.

```
Veni, Vidi, Visa: I came, I saw, I did a little shopping.
```

- Identify and count the alliteration chains in each example in the data set.
- Automatically achieved with a pronunciation dictionary, a longest string matching device and a stopword list of functional words.

Stylistic Features: Antonymy

- Humor often relies on some form of incongruity, opposition or contradiction.

```
Always try to be modest and be proud of it.
```

- Identify the presence of antonyms in a sentence.
- *Antonymy* relation (and *similar-to* relation for adjectives) of WordNet.
- Problem of coverage: antonymy feature cannot always be identified.

Stylistic Features: Adult Slang

- Humor based on adult slang (= sexual-oriented lexicon) is popular.

`The sex was so good that even the neighbors had a cigarette.`

- Extract the synsets labeled with the domain `SEXUALITY` from WordNet.
- Check for the presence of these words in each sentence and annotate them accordingly.
- Problem of coverage: adult slang feature cannot always be identified.

Classification Using Heuristics

- Numerical features that act as heuristics
- Parameters to learn: Threshold indicating the minimum value admitted for a sentence to be classified as humorous (or non-humorous).

Classification Using Heuristics

- Numerical features that act as heuristics
- Parameters to learn: Threshold indicating the minimum value admitted for a sentence to be classified as humorous (or non-humorous).
  $\Rightarrow$ **Decision tree**

```
alliteration = 0
|    adult slang = 0
|    |    antonymy <= 1 : no
|    |    antonymy > 1 : yes
|    adult slang > 0 : yes
alliteration > 0 : yes
```

Figure: A sample decision tree

Outline

Content-based Features

Formulation of the humor recognition task as a traditional text classification problem.

$\Rightarrow$ The sentences themselves represent feature vectors encoding term presence or absence.

Classifiers: Naive Bayes and SVM

Outline

Combining Stylistic with Content Features

- Jointly exploit stylistic and content features for humor recognition
- Stacked learner:
  1. Apply the text classifier
  2. Join the output of the text classifier with the stylistic features
  3. Feed the newly created vector to a machine learning tool

Outline

A. Classification Using Heuristics

Humor recognition accuracy using alliteration, antonymy and adult slang

	One-Liners			
Heuristic	Reuters	BNC	Proverbs	OMCS
Alliteration	74.31%	59.34%	53.30%	55.57%
Antonymy	55.65%	51.40%	50.51%	51.84%
Adult slang	52.74%	52.39%	50.74%	51.34%
All	76.73%	60.63%	53.71%	56.16%

A. Classification Using Heuristics

Humor recognition accuracy using alliteration, antonymy and adult slang

	One-Liners			
Heuristic	Reuters	BNC	Proverbs	OMCS
Alliteration	74.31%	59.34%	53.30%	55.57%
Antonymy	55.65%	51.40%	50.51%	51.84%
Adult slang	52.74%	52.39%	50.74%	51.34%
All	76.73%	60.63%	53.71%	56.16%

- Training on 1.000 examples, evaluation on the remaining 15.000 examples
- Proverbs and one-liners have the most similar style.
- Reuters titles and one-liners have the most different style.
- Alliteration feature: the most useful indicator of humor

B. Text Classification

Humor recognition accuracy using content-based features

	One-liners			
	Reuters	BNC	Proverbs	OMCS
Naive Bayes	96.67%	73.22%	84.81%	82.39%
SVM	96.09%	77.51%	84.48%	81.86%

B. Text Classification

Humor recognition accuracy using content-based features

	One-liners			
	Reuters	BNC	Proverbs	OMCS
Naive Bayes	96.67%	73.22%	84.81%	82.39%
SVM	96.09%	77.51%	84.48%	81.86%
Stylistic F.	76.73%	60.63%	53.71%	56.16%

B. Text Classification

Humor recognition accuracy using content-based features

	One-liners			
	Reuters	BNC	Proverbs	OMCS
Naive Bayes	96.67%	73.22%	84.81%	82.39%
SVM	96.09%	77.51%	84.48%	81.86%
Stylistic F.	76.73%	60.63%	53.71%	56.16%

- Again: Reuters titles seem to be the most different from one-liners.
- BNC sentences represent the most similar data set.
  $\Rightarrow$ Joke content tends to be similar to regular text.
- Interesting: Proverbs and one-liners seem to deal with different topics (despite their stylistic similarity)

C. Combination

Humor recognition accuracy using content-based and stylistic features

	One-liners			
	Reuters	BNC	Proverbs	OMCS
Combination	96.95%	79.15%	84.82%	82.37%

C. Combination

Humor recognition accuracy using content-based and stylistic features

	One-liners			
	Reuters	BNC	Proverbs	OMCS
Combination	96.95%	79.15%	84.82%	82.37%
Text Classif.	96.67%	77.51%	84.81%	82.39%
Stylistic Feat.	76.73%	60.63%	53.71%	56.16%

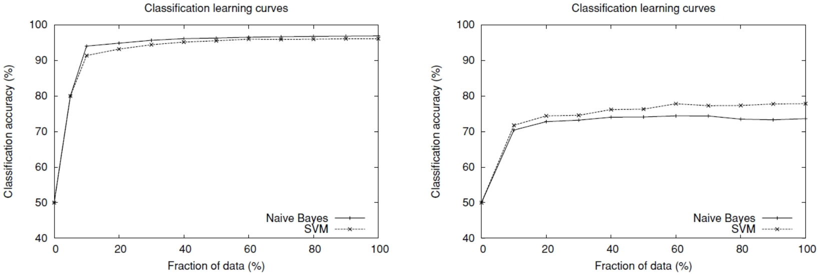
- No improvement for One-liners/Proverbs and One-liners/OMCS
  $\Rightarrow$ not surprising (see first experiment)

Discussion

- Results prove that automatic classification techniques represent a viable approach for the task of humor recognition.
- Good performance using stylistic and content-based features.
- Initial intuition: one-liners are most similar to other creative text (e.g. Reuters titles or proverbs)
- But: it is more difficult to distinguish humor from regular text (e.g. BNC sentences)
  Still: significant improvement over the baseline with the combined classifier

Discussion contd.

Learning curve for humor recognition using text classification (Reuters & BNC)



- Similar-shaped curves for the other negative data sets
- Steep ascent: humorous and non-humorous texts are well-distinguishable
- Plateau: more data is not likely to improve the quality of humor detection

Discussion contd.

Additional experiments and results:

1. Experiment using sentences drawn randomly from the four non-humorous collections
   $\Rightarrow$ Results similar to the results with One-liners/BNC before
2. Experiment using uneven class distributions:
   75% non-humorous/25% humorous
   $\Rightarrow$ Still improvement over the baseline

Discussion contd.

- **Sources of humor** in cases where the stylistic and content-based features failed (manual inspection): irony, ambiguity, incongruity, idiomatic expressions, commonsense knowledge
- Semantic classes of **the most discriminative content-based features** (useful for humor generation): human-centric vocabulary, negation, negative polarity, professional communities, human "weakness"

Outline

Conclusion

A conclusion is simply the place where you got tired of thinking.

(anonymous one-liner, from [Mihalcea and Strapparava, 2006])

- Automatic classification techniques can be successfully applied to the humor recognition task.
- More training data is not likely to improve the performance $\Rightarrow$ identification of more sophisticated humor-specific features

But:

- Can this approach generalize from one-liners to humor in general?
- How would it be used in the mentioned applications?

References

Mihalcea, R. and Strapparava, C. (2005).
Bootstrapping for Fun: Web-based Construction of Large Data Sets for Humor Recognition.
Proceedings of the Workshop on Negotiation, Behaviour and Language.

Mihalcea, R. and Strapparava, C. (2006).
Learning to Laugh (Automatically): Computational Models for Humor Recognition.
Computational Intelligence, 22(2).

Taylor, J. M. (2009).
Computational Detection of Humor: A Dream or A Nightmare?
International Conference on Web Intelligence and Intelligent Agent Technology - Workshops.

Taylor, J. M. and Mazlack, L. J. (2004).
Computationally Recognizing Wordplay in Jokes.
Proceedings of Cognitive Science Conference, pages 2166–2171.