A Corpus Study with German Data Sets into the Similarity of Irony and Satire

Marisa Schmidt (presenter) **& Karin Harbusch**

uk universität koblenz

weiter:denken

Computer Science Faculty, Inst. f. Computational Visualistics

{marisaschmidt|harbusch}@uni-koblenz.de

About the Presenter

Marisa Schmidt is a second-year MA student at the University of Koblenz, Germany, studying Computational Visualistics, i.e., a Computer Science specialization in Graphics and Image Processing.

She is currently working on people counting in depth images that are received from a 3D-time-of-flight camera.

She has a keen interest in discovering all kinds of patterns with the help of machine learning techniques. Usually she works with image data.

Overview

1. **Motivation**

2. **Satire and irony in the scholarly literature**

3. **Corpora**

4. **Satire and Irony detection systems used in our study**

5. **Two experiments**

6. **Conclusions**

1. Motivation

In the area of **deception or fake-news detection**, i.e., the falsification of news in journalistic articles or social media, one strand of research deals with **satire detection**.

According to the scholarly literature, satire often uses is **irony**.

Research question → **How much irony can be detected in a satire-annotated data set?**

2. Satire and irony in the scholarly literature

Satire is the use of humor, irony, exaggeration or ridicule to expose and criticize people's stupidity or vices. (The Oxford English Dictionary/OED)

Irony is the expression of meaning through the use of language signifying the opposite, typically for humorous effect. (OED)

"**Satire and irony are often closely related, but there are important distinctions between the two**. As form of criticism, satire uses humor to accomplish its goals. One technique that satire uses is irony. Irony focuses on the discrepancies between what is said or seen and what is actually meant. Simply, satire and irony hardly differ because one, satire, often uses the other, irony." (Singh, 2012)

3. Corpora

From public data collections, we use the three data sets:

- **C1:** 329,859 articles from 15 different German newspapers (2.82% satirical ones), and C1SUB with 125 newspaper articles, 45 of which are satire (36%) to run it on a less powerful system compared to the settings by McHardy et al. (2019),

- **C2a:** SARC 2.0 with 321,748 entries and **C2b:** SARC 2.0 pol (17,074 entries), two subsets of a big Reddit corpus labeled for irony (Khodak et al., 2018),

- **C3:** a Twitter data set from SemEval-2018 (van Hee et al., 2018) that is labeled with #irony, #sarcasm and #not. The corpus provides 4,792 tweets, where both, irony and sarcasm, have a percentage of 50%.

Moreover, we test all models with a newly set up a new corpus (still in its infancy):

- **C4:** 10 ironic examples from different articles. Moreover, we thought up 5 ironic ones ourselves as a kind of control instance in contrast to the outstanding quality of the literature examples and 6 neutral definitions of facts labelled not-ironic.

4. Irony and Satire Detection Systems used in our study

- **Recent overviews:**
  - The challenging task of satire detection has been tackled from various points of view: lexically, syntactically, and semantically. Thu and Aung (2018) give an historical overview for systems from the different viewpoints.
  - For a good overview on irony-detection systems, subdivided into surface and semantic approaches, as well as pragmatic ones, see Karoui et al. (2019).
- **The two recent systems, we used in our study:**
  - **Adversarial Satire** (McHardy et al., 2019), evaluated with C1, and
  - **Elmo4Irony** (Ilić et al., 2018), evaluated with C2 and C3.

The code can be found here: https://gitlab.uni-koblenz.de/marisaschmidt/irony-detection

5. Two experiments

Research question → **How much irony can be detected in a satire-annotated data set?**

Experiment 1

Evaluate C1SUB with Elmo4Irony and Adversarial Satire. Does the irony detector perform better?

Experiment 2

Test C4 with both systems. How good is the quality with examples from the scholarly literature?

5.1 Results of Experiment 1

- Both methods are trained over 10 epochs with a batch size of 16.
- Elmo4Irony is trained with dropouts of 0.0, 0.1 and 0.5.
- For Adversarial Satire, different values for the adversarial weight are used: 0.0, 0.2, 0.3 and 0.7.

For these variable settings, Elmo4Irony performs always better than Adversarial Satire. In fact, the irony classifier provides better results on the satire dataset than the specialized satire classifier. This observation **confirms the hypothesis** of Experiment 1. **Irony is an indicative feature to satire detection.**

Evaluation results for C1SUB

Data	Adversarial Satire			Elmo4Irony		
	Confounding variable = 0.0			Confounding variable = 0.0		
	P	R	F1	P	R	F1
C1SUB	0,622	0,617	0,618	0,895	0,800	0,816
	Confounding variable = 0.7			Confounding variable = 0.1		
	P	R	F1	P	R	F1
C1SUB	0,708	0,617	0,603	0,857	0,867	0,839

5.2 Results of Experiment 2

- Both systems are evaluated on the new dataset C4 that is labeled for irony.

- The training of the models still happens on their regular datasets.

Based on Experiment 1, we argue that irony can serve as satire feature. However, it is less obvious that a satire classifier will find irony in irony data. It is therefore to be expected that Adversarial Satire will find less satire on this dataset with ironic examples.

Unexpectedly, the features calculated by both systems are not suitable for this new dataset, as almost everything is classified as ironic. The small size of C4 cannot be the reason for failure given that the corpus is only used as test set. Deeper analysis of the features is required here.

Evaluation results for C4

Elmo4Irony					Adversarial Satire				
dropout	TP	FP	TN	FN	**adv. weight**	TP	FP	TN	FN
0.0	3 (0)	5 (0)	1 (6)	12 (15)	0.0	15 (15)	6 (6)	0 (0)	0 (0)
0.1	15 (0)	6 (0)	0 (6)	0 (15)	0.2	14 (15)	6 (5)	0 (1)	1 (0)
0.5	15 (0)	5 (0)	1 (6)	0 (15)	0.3	15 (15)	6 (6)	0 (0)	0 (0)
					0.7	15 (15)	6 (6)	0 (0)	0 (0)

TP = samples that were correctly classified as irony, FP = wrongly classified as irony, TN = correctly classified as regular and FN = wrongly classified as regular. The numbers in the brackets show the results probing additionally provided neutral text to obtain article length in C4 aiming at improving the quality of Adversarial Satire.

6. Conclusions

We presented the results of a corpus study into the relationship between satire and irony. Based on the definition that satire uses irony, we could **verify that irony detection can serve as feature for satire classification very well**.

Experiment 2 was designed to better understand the irony features. However, the results were unexpectedly poor. We plan to extend C4 to a full development/test corpus with a larger collection of examples from very divergent sources. The goal is to obtain a richer set of features to classify irony.

It may be that Elmo4Irony is better on this corpus, because the examples are short. Given that Adversarial Satire is trained with newspaper articles, this could be a reason. However, we tested with additional neutral text as well, which did not have a positive impact.

References

S. Ilić, E. Marrese-Taylor, J. A. Balazs, and Y. Matsuo, "Deep contextualized word representations for detecting sarcasm and irony," Proc. 9th WASSA, pp. 2–7, Oct. 2018. https://aclanthology.org/W18-6202

J. Karoui, F. Benamara, and V. Moriceau, "Automatic detection of irony: opinion mining in microblogs and social media," London: ISTE, 2019. https://doi.org/10.1002/9781119671183

M. Khodak, N. Saunshi, and K. Vodrahalli, "A large self-annotated corpus for sarcasm," Proc. 11th LREC, pp. 641-646, May 2018. https://doi.org/10.48550/arXiv.1704.05579

R. McHardy, H. Adel, and R. Klinger, "Adversarial training for satire detection: controlling for confounding variables," Proc. NAACL-HLT, pp. 660–665, Jun. 2019. https://doi.org/10.48550/arXiv.1902.11145

OED, Oxford Univ. Press, Oxford. https://www.oed.com/public/freeoed/loginpage

R. K. Singh, "Humour, irony and satire in literature," IJEL, vol. 3, no. 4, pp. 65-72, 2012. https://www.academia.edu/4541187/Humour_Irony_and_Satire_in_Literature

P. P. Thu and T. N. Aung. "Implementation of emotional features on satire detection," International Journal of Networked and Distributed Computing, Vol. 6, No. 2, pp. 78-87, 2018.

C. van Hee, E. Lefever, and V. Hoste, "SemEval-2018 task 3: irony detection in English tweets," Proc. 12th SemEval, pp. 39-50, Jun. 2018. http://dx.doi.org/10.18653/v1/S18-1005