



Irony and sarcasm expression in Twitter

Tatiana Zefirova¹ and Natalia Loukachevitch²

¹ Moscow State University, Moscow, Russia

² Research Computing Center of Moscow State University, Moscow, Russia

Abstract. Information extraction and sentiment analysis of social net contents is a relatively new field, but a very promising and quickly developing one. In the current work, we discuss the possibilities of automatic detection and analysis of ironic and sarcastic messages in one of the most popular (among both users and researches) social nets - Twitter. A particular trait of this research will consist in analyzing the languages content, which is rarely observed. We will focus on the material of Spanish, German and Russian languages, trying to implement the same analysis algorithms to all three as well and to define the differences.

Keywords: feature extraction, natural language processing, statistical analysis.

1 Introduction

Automatic processing of written discourse has a great potential on the material of social net messages, as they tend to represent spoken language in written - and easily accessible - form. However, while many features of natural discourse are preserved in the written form, it is often difficult to detect more sophisticated ones. Thus, for example, irony and sarcasm are both used to express additional meaning in the utterance, which can even be opposite to the one expressed directly. Certainly, such subtle variations can be hard to perceive even while interacting face-to-face, leave alone automatic detection and analysis. We nevertheless suggest that a key to understanding and automatically detecting irony and sarcasm may lay in the way the Internet users themselves mark the messages in social nets - hashtags.

A social net that provides the widest selection of hashtag usage is Twitter. A single hashtag added to a short message can change its whole mood. Take a message "How cool is the snow in March!" and add #irony - and you will get a complaint instead of excitement. We suggest that the certain lexicon may be used in messages of such type (apart from #irony and #sarcasm hashtags).

2 Overview

Although the topic itself is quite modern, a significant amount of work has already been dedicated to the analysis of Twitter and other social net contents. However, as mentioned previously, sentimental analysis of irony and sarcasm can be a tricky task. Let us first look closer at the subject of those terms itself.

2.1 Irony vs. sarcasm

Although a profound linguistic literature can be found on the topic of the differences between the two terms, when it comes to automatic distinguishing, the task is extremely complicated. Some authors even claim that “there seems to be no way of differentiating reliably between the two phenomena, and in part because a shift in meaning for the word irony seems to be taking place with “sarcasm” occupying what was previously the semantic place for “irony” [1]. In the current work, we have as well not found significant differences in the usage of two terms in social net messages, although in some cases there supposedly was a difference in covered topic. Thus, for example the messages marked “ironic” seemed to be more relevant to the author's personal life and feelings, while “sarcastic” messages covered political and media thematic posts.

2.2 Existing approaches to automatic irony and sarcasm analysis

The briefly described approach of analyzing the hashtags is the most popular when it comes to identifying the phenomena of irony/sarcasm in social nets. On the other hand, some authors suggest that machine approaches cannot perform very well in this task, and suggest using human resources and crowdsourcing for closer detecting. However, even in this case the results tend to be questionable and unreliable [4],

The works focusing on analyzing hashtags as markers of ironic and sarcastic tweets also suggest analyzing other textual and meta- data. Among the suggested tweets characteristics, for example, one can find word unigrams and bigrams in different approaches and variations, part of speech, pronunciation and capitalization features [2]. On the other hand, supposed relevant features include those of the messages' authors' as well: author historical salient terms, historical topics, profile information (gender, number of friends, followers and statuses, their duration on Twitter, the average number of posts per day, their time zone, and whether or not they are verified by Twitter [2],

In the previous work it has been suggested that a combination of characteristics would be the most effective way to detect the needed type of messages. However, as the full task is quite complicated, at the moment we will focus on word frequencies only - nevertheless, this is exactly the feature that is supposed to have the greatest relevance per se.

3 Collecting the data

While much research regarding automatic information extraction has already been carried out on the material of English language, some other no less rich languages have been almost left out of the studies. We have tried to fill in the gap of studies of Spanish, German and Russian. We have automatically collected Twitter messages data of Spanish, German and Russian speaking users, searching for messages containing hashtags expressing irony and sarcasm (#ironia, #sarcasmo, #Ironie, #Sarkasmus, #ирония, #сарказм) within a month and have then analyzed the highest frequency words and covered topics in each collection. Due to Twitter API limitation, over this period of time, it has only been possible to collect 100 latest results of each search request per week, which amounted to around 300 tweets per request after elimination of retweets and other irrelevant type of results, 1800 messages overall. Although the amount seems insufficient for proper statistical analysis, it is enough to reveal several interesting dependencies discussed further.

The most difficult part appeared to be collecting tweets in German with the search request “Ironie”, because apparently the exact same form corresponds to French language as well. To reach the desired number of messages in German we had to resort to manual search as well. Other similar problems arise with Spanish, Portuguese and Italian messages sharing similar forms.

Apart from eliminating messages in irrelevant languages, we have been also excluding retweets, messages containing hashtags only (“Solo #pensamientos #aprendiendoaserfeliz #ironía #escribir #write

#think #magia”), messages where the ironic or sarcastic connotation was unclear (for example, in the message “@MDR_SAN @NicoAusB #Sarkasmus V Grüße an #Poggenburgs #Enquetekommission.” It is unclear whether the whole text is supposed to be sarcastic or maybe some message in previous correspondence) or messages containing hashtags as a description of an image attached to the tweet.

In the related works on the topic, it has also been suggested to search for hashtags such as #ironic or #sarcastic, however, within the current research we have not found sufficient examples of such messages in other languages. For example, a single search for “#sarcástico” hashtag (in Spanish) only gives 7 relevant results (out of 100 possible), which makes it extremely difficult to fill the database. However, we suppose that such search results might also be of certain interest when there are enough of them to carry out a more scrupulous investigation. On the other hand, such hashtags tend to be used within the message text itself, in other words, they stop being hashtags per se and turn into hashtag marked parts of the messages, which is not exactly what we were looking for. For example:

- #irónico con la cantidad de #misiles lanzados por #eeuu se resuelve muchos problemas de desigualdad en el #mundo e igual, se siguen usando esos re- cursos para separar más a la humanidad todavía no les queda claro quiénes son los únicos beneficiados con la #guerra?
- #Cuando te quieres hacer el #sarcástico con tu #chugarbeibi y le dices: "Nah, aún me parece largo ese #short"

4 Analysis

In previous works on similar topic various messages' and authors' characteristics have been reviewed, and yet the most obvious parameter (word frequency) is supposedly the most significant one. To detect most significant (frequent) words, we analyzed both simple frequency charts and bigrams collocations, adding algorithms excluding stop- words (from NLTK python module lists).

4.1 Spanish

For Spanish language the first method put such words as “bien”, “buen” and “qué” to the top of the list. The “que” in this case is a participle used to express excitement or any strong emotion (“Qué buena idea!” - “oh what a good idea!”). A phrase constructed of two of the detected parts “Qué bueno!”, is a generic positive feedback utterance, “How nice!”, which seems to be transformed by hashtags in a complete opposite. The second method detects such words as “universidad” (university), “autónomo” (autonomic), “corrupción” (corruption), “periodista” (journalist) and others, clearly describing the generally ironically covered topics of the Spanish-speaking Twitter community.

Other interesting results related to topics coverage include the fact, that “#ironía” hashtag has been wider used in European Spanish messages (either covering Spanish politics (Catalonia issues, for example) or written in European Spanish), which “#sarcasmo” was quite more often in Latin American messages, mostly covering the topics of upcoming presidential elections in Mexico and Colombia.

Unfortunately, it is quite difficult if not impossible at all to detect automatically whether a short message is written in European or Latin Spanish variant, such distinction needs careful manual checks. However, we suppose that this difference might be an interesting opportunity for further studies.

4.2 German

The results of the implementation of the first method on the German language material highlighted general words as “klar” (right, alright) and “toll” (cool) and quite surprisingly more “political” ones, such as “Deutschland” (Germany), “CDU” (one of the political parties) and “land” (country). “Deutschland” also made it to the second methods' list, alongside with “eigentlich” (actually), “freiwillig” (on free will), “beschuldigt” (accused) and “ungeprüft” (unchecked, unverified).

It seems clear that the nouns of German language are very much less likely to top the lists due to their extreme compound variety, which makes it hard for two people to use the exactly same word to express the same thought. We suppose that a closer investigation, either including stemming or thorough manual examination can provide much more interesting results.

Quite unsurprisingly, in case of German language databases politics has also been detected as the most ironically/sarcastically discussed topic. In March German speaking communities tended to discuss re-election of Angela Merkel as German Chancellor.

4.3 Russian

Alongside with the described intriguing results, we could not obtain a clearly analyzable list for Russian language. According to the data collected, the hashtags under investigation are not used sufficiently in the same way by the Russian speaking Twitter audience. The “#ирония” (“irony”) hashtag is mostly used as a description for images irretrievable automatically or jokes alongside with such hashtags as “#шутка” (“joke”), “#смех” (“laugh”) and does not seem to express the needed connotations. On the other hand, the “#сарказм” (“sarcasm”) hashtag is indeed used properly, however, it only gives six more or less relevant results (nevertheless including the generic jokes texts marked with previously mentioned “jokes” and “laugh” hashtags). We suppose, that an interesting task would be to check if Russian-speaking audience of Twitter actually uses the selected hashtags - but untranslated from English.

5 Conclusion

The obtained results, though preliminary, give us hope to get close to automated detection of irony and sarcasm in social net messages in Spanish and German. Although the results for Russian are at the moment not so promising, we nevertheless suppose that a more broad collection collected over a longer time period may help to investigate this part of the research more profoundly as well. We would then argue that the most frequently encountered words and word combination may be used to detect irony and sarcasm automatically in the cases when the message does not contain the “training” hashtags. We are looking forward to implementing a machine learning algorithm to the ready collections of three languages data as well as analyzing more complex combinations of messages' and author's characteristics.

References

1. Attardo, Salvatore, Jodi Eisterhold, Jennifer Hay, and Isabella Poggi. "Multimodal markers of irony and sarcasm." *Humor* 16, no. 2 (2003): 243-260.
2. Bamman, David, and Noah A. Smith. "Contextualized Sarcasm Detection on Twitter." In *ICWSM*, pp. 574-577. 2015.
3. Barbieri, Francesco, and Horacio Saggion. "Modelling irony in twitter." In *Proceedings of the Student Research Workshop at the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 56-64. 2014. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1-13. Springer, Heidelberg (2016).
4. Gonzalez-Ibanez, Roberto, Smaranda Muresan, and Nina Wacholder. "Identifying sarcasm in Twitter: a closer look." In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2*. pp. 581-586. Association for Computational Linguistics, 2011.
5. Reyes, Antonio, Paolo Rosso, and Tony Veale. "A multidimensional approach for detecting irony in twitter." *Language resources and evaluation* 47, no. 1 (2013): 239-268. Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1-2. Publisher, Location (2010).
6. Suhs, Emilio, Delia Irazu Hernandez Farias, Paolo Rosso, Viviana Patti, and Giancarlo Ruffo. "Figurative messages and affect in Twitter: Differences between# irony,# sarcasm and# not." *Knowledge-Based Systems* 108 (2016): 132-143.