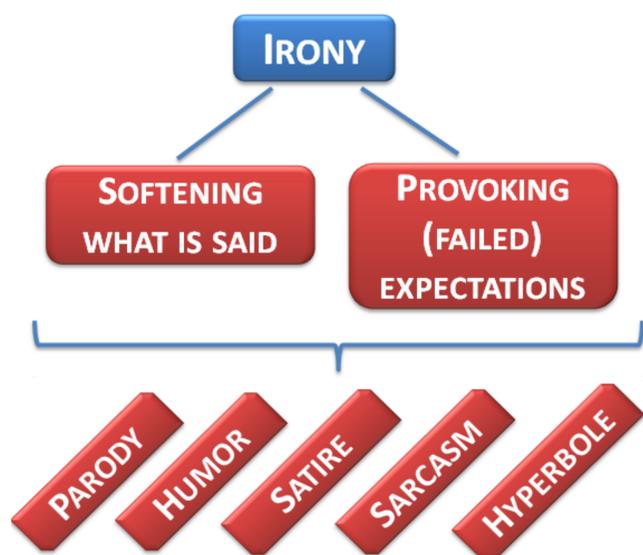


Abstract

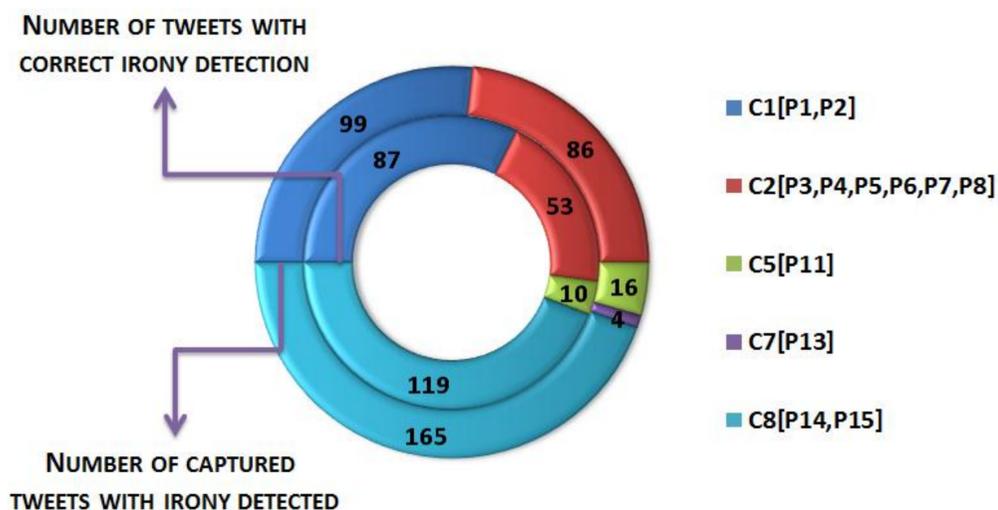
This study considers ironic statements when expressing opinions about a specific domain, which was chosen according to the trend topics on Twitter. Since irony requires a common ground between tweeter and reader, we established the specificities of the topic, the period of time and the textual support as contextual features. By observing the selected corpus, we introduce a set of fifteen patterns that might be present in ironic/sarcastic statements. We developed special clues on this subject: static expressions, punctuation marks, emoticons, laughter expressions, syntactic rules, among others.

Irony extraction



Categories	Patterns	Expression
C1 [lists]	P1	List of Laughter Expression
	P2	List of Emoticons
C2 [exact expressions]	P3	"só que"
	P4	"sim,"
	P5	"seria"
	P6	"na boa"
	P7	"medo" "Medo!" "#medo"
	P8	"#ironia" "#sarcasmo" "#joking" "#kidding"
C3 [POS]	P9	ADV + ADV ADJ + ADJ
C4 [POS + exact expressions]	P10	"tão" + ADJ "tão" + ADV
C5 [POS + lists]	P11	ADJ+ List of Emoticons
C6 [POS + NE]	P12	DET + ADJ + (PREP + DET) + NE
C7 [lists + NE]	P13	Demonstrative Pronouns + NE
C8 [symbols]	P14	... <EXPR>!
	P15	! * ? * ! * ? * ? * ! *

PATTERNS FOR IRONY DETECTION



- ▶ 370 tweets were extracted and analyzed manually, and 269 of them were correctly identified. This gives a 72% precision for the proposed approach.
- ▶ Patterns related to symbolic language, such as laughter marks and emoticons, are the best hints to irony and sarcasm. Also, the use of heavy punctuation leads to clues to ironic statements.
- ▶ POS and static expressions patterns were proven to be bad search choices (low output results and generalized uses of these expressions) with exception of P3. Patterns P4, P10, P11, P12 and P15 were not found in our corpus, and they will be tested again in future works.
- ▶ From this analysis and the elaboration of the set of patterns, we developed a system that allows both extraction of patterns and manual tagging. This is constituted of two tabs:
 - 1) Pattern detection: tweets are recovered according to the pattern chosen.
 - 2) Manual tagging: in this tab, there are options to select a tweet and manually classifying it into subtypes of irony: JOCLARITY, UNDERSTATEMENT, HYPERBOLE OR RHETORICAL QUESTIONS.

References

- [1] P. Carvalho, L. Sarmento, M. J. Silva, and E. de Oliveira. Clues for detecting irony in user-generated contents: oh...!! it's "so easy" ;-). In Proceedings of the 1st International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion, pages 53-56, 2009.
- [2] R. Gonzalez-Ibanez, S. Muresan, and N. Wacholder. Identifying sarcasm in twitter: a closer look. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pages 581-586, 2011.
- [3] O. Tsur, D. Davidov, and A. Rappoport. ICWSM – a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, pages 162-169, 2010.
- [4] A. Utsumi. Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony. Journal of Pragmatics, 32(12):1777-1806, 2000.