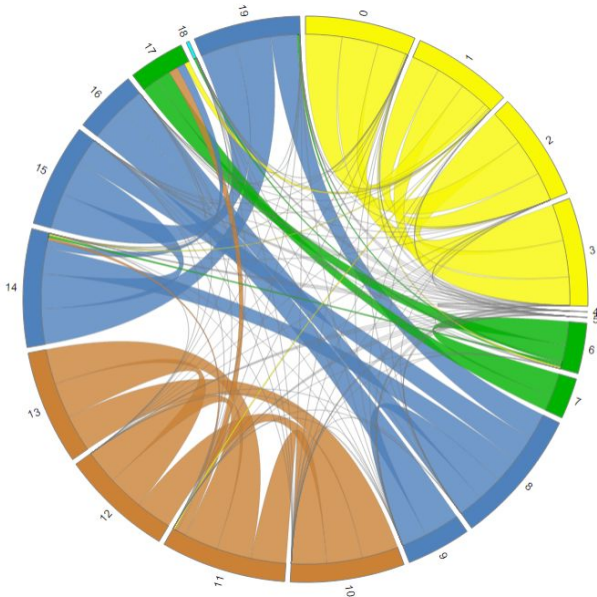


Making sense of funny bone from cartoon caption contest results

14 July 2015, by Nancy Owano



Clustering of the mini corpus. Credit: arXiv:1506.08126 [cs.CL]

The New Yorker's practice of publishing a weekly cartoon with no caption and then inviting readers to submit captions for it caught the attention of a team of 11 from Columbia University, University of Michigan, Yahoo! Labs and *The New Yorker*. As it stands, over 5,000 readers submit what they hope will be considered as the funniest captions and the editors choose three top entries and then ask readers to select the funniest of all. That created a huge database of captions.

The researchers saw an opportunity to learn something from all those rounds. What makes a caption funny? What separates funny captions from the rest? What are winning methods for selecting winning captions?

Their paper, "Humor in Collective Discourse: Unsupervised Funniness Detection in the New

Yorker Cartoon Caption Contest," is on arXiv. Interestingly, Amazon Mechanical Turk was brought into play in their research.

MIT Technology Review discussed the paper. Methods of analysis: (1) standard linguistic techniques with criteria including the level of positive or negative sentiment, whether captions were human-centered, how clearly they refer to objects depicted in the cartoon; (2) caption study via network theory —topics mentioned in captions were listed and they created a network by linking captions that mentioned the same topics. "That allowed them to use standard network analysis tools to find, for example, the most important node in network, a property known as [centrality](#)."

Each of these methods produced a ranking of captions, said this arXiv paper's review, and the authors took each of the most highly ranked and compared them to captions that the readers of *The New Yorker* chose as the funniest. They did this by crowdsourcing opinion using Amazon's Mechanical Turk.

"Each AMT HIT consisted of one cartoon as well as two captions, A and B (produced by one of the 18 methods and baselines). The turkers had to determine which of the two captions is funnier. They were given four options - 'A is funnier', 'B is funnier', 'both are funny', 'neither is funny'. They did not know which method was used to produce caption A or B. All pairs of captions from our methods were compared for each cartoon, and each HIT (pair) was assessed by 7 Turkers," the authors wrote.

They had access to a corpus of more than 2M captions for more than 400 contests run since 2005. They said they took a "computational approach" to gain insights into what differentiates funny captions from the rest.

"We developed a set of unsupervised methods for

ranking captions based on features such as originality, centrality, sentiment, concreteness, grammaticality, humancenteredness, etc. We used each of these methods to independently rank all captions from our corpus and selected the top captions for each method. Then, we performed Amazon Mechanical Turk experiments in which we asked Turkers to judge which of the selected captions is funnier."

For testing they then picked a subset of 50 cartoons and 298,224 captions.

But wait a minute. Why bother? Isn't "funny" too elusive to measure, as what is taken as "funny" differs from time to time, place to place and person to person? As *MIT Technology Review* commented, "humor depends on so many parameters, many of which are internal and liable to change from one moment to the next. What seems funny now may not seem so funny later or tomorrow."

Some scientists believe the effort to make sense out of what is funny is worthwhile. *MIT Technology Review* said "various linguists and psychologists have suggested that good [jokes](#) all share common properties and that a systematic analysis ought to reveal them. The question is how to get at these primitives of humor and whether machine learning can help."

The authors said that "We are making our corpus public for research and for a shared task on funniness detection. The corpus includes our 50 selected cartoons, more than 5,000 captions per cartoon, manual annotations of the entities in the cartoons, automatically extracted topics from each contest, and the funniness scores."

As part of their future work, the team said they will explore pun recognition (e.g., "Tell my wife I'll be home in a minotaur"), other creative uses of language, and more semantic features.

What did they actually find? What is the magic sauce that makes a caption successfully funny? They found that methods that consistently select funnier captions are negative sentiment, human-centeredness, and lexical centrality.

"Not surprisingly," they said, "knowing the traditions of the New Yorker cartoons, negative captions were funnier than positive captions. Captions that relate to people were consistently deemed funnier." The authors also said that "[captions](#) that reflect the collective wisdom of the contest participants outperformed semantic outliers."

More information: Humor in Collective Discourse: Unsupervised Funniness Detection in the New Yorker Cartoon Caption Contest, arXiv:1506.08126 [cs.CL] arxiv.org/abs/1506.08126

Abstract

The New Yorker publishes a weekly captionless cartoon. More than 5,000 readers submit captions for it. The editors select three of them and ask the readers to pick the funniest one. We describe an experiment that compares a dozen automatic methods for selecting the funniest caption. We show that negative sentiment, human-centeredness, and lexical centrality most strongly match the funniest captions, followed by positive sentiment. These results are useful for understanding humor and also in the design of more engaging conversational agents in text and multimodal (vision+text) systems. As part of this work, a large set of cartoons and captions is being made available to the community.

© 2015 Tech Xplore

APA citation: Making sense of funny bone from cartoon caption contest results (2015, July 14) retrieved 21 October 2019 from <https://techxplore.com/news/2015-07-funny-bone-cartoon-contest-results.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.