Different Knowledge, Same Joke: Response-Based Computational Detection of Humor

Julia M. Taylor(✉)

Computer and Information Technology Department, Purdue University,
401 N. Grant Street, West Lafayette, IN 47907, USA
jtaylorl@purdue.edu

Abstract. The paper explores the very basis of linguistic theories of humor with a view of applying them to computational humor. Computation requires tighter definitions. The paper analyzes joke-carrying texts based on the existing script-based methods. It compares jokes that have the same setup but different punchlines by examining the background knowledge that should be available to detect humor. It then moves into jokes where the same joke text elicits different responses from the reader, and conjectures that the responses are based on the readers' world knowledge and preferences. Such responses make it possible not only to analyze humor, but also to understand more about the people that produce the responses.

Keywords: Humor · Computational detection of humor · Semantics · Knowledge-based humor detection

1 Introduction

Since 1985, a text has been assumed to be joke-carrying [1] if it was compatible, fully or in part, with two scripts that overlap and oppose. The assumption is that any two scripts that a reader sees in a text would do the job, especially if communication is based on short canned-type jokes, which is what the theory was interested in. The (main) script oppositeness detection gets more difficult if the jokes are not canned, but a joke-carrying text is actually a spontaneous response to existing situation, or, as a simpler case, there must be response to the canned joke other than a non-verbal acknowledgment of the funny stimulus.

The reason why such detection is difficult is that each participant has their own fuzzy (in a technical sense of the word) view of the described situation, fill in the blanks according to their own understanding and experiences and may shift a conversation (whether the shift is picked up by others or not) to a different subscript or a set of scripts. In other words, it is these individual experiences that contribute to what part of the story a participant focuses on and it is likely that the oppositeness will be based on that focus, if one is to create one. While we are typically not blind to other possible scripts, or at least capable of recognizing them once they are recoverable from a stated utterance or sentence, their prediction is much more difficult as the result is based on

N. Streitz and P. Markopoulos (Eds.): DAPI 2015, LNCS 9189, pp. 680–688, 2015.
DOI: 10.1007/978-3-319-20804-6_62

something that may not be expressed (verbally or by another other means of communication).

With the development of computational humor, it becomes – at least theoretically – possible to detect more than one pair of scripts in a particular joke-carrying text. Moreover, it – again, at least theoretically – becomes possible to recognize scripts that may not be just visible to humans (due to lack of priming, for example), but may be visible to a machine as well (and, in the worst case scenario, only to machine but not to a human). Once a situation arises, where a person does not recognize a joke, but a computer does, due to its sensors or some (incorrectly?) weighted knowledge, a scientist may want to answer a question: is it possible to reduce such occurrences to match human expectations (or result appreciated by a human)? The next logical step, which we are not attempting to address here is, at which point we are willing to grant a computer a right to have an opinion over the jokeness value of a text. To put it into a less controversial and philosophical question, what constitutes an error in computational detection/generation of humor? Is it something that a (generic) human doesn't agree to, or is it a lack of (acceptable) explanation on the part of the computer why something is or isn't funny, based on known theories?

There is no argument that people have different senses or humor, and appreciate or do not appreciate different jokes. There is no argument that one person may think that something is a joke while another does not. There is also no argument that we allow such disagreements for humans. Luckily, computational humor is not at a state where a computer can address a question of why something is or isn't funny reliably (see [2] for initial attempts), but nevertheless such questions may be worth thinking about (at least relative to how do we report error rates). Should we allow a unique sense of humor for a computer? (For discussion on sense of humor see [3, 4])

Now that the controversial questions have been asked, we will leave them aside for the rest of the paper and proceed with examples where the same situation may lead to different response (different script oppositeness) and what can a system do about that.

2 Humor Theories

As mentioned in the Introduction, Script-based Semantic Theory of Humor (SSTH) [1] states that a text is joke-carrying if it contains two scripts that (partially) overlap and oppose. While the book was explaining only canned jokes, the theory can be applied to any type of joke (canned or not), and its extension—General Theory of Verbal Humor [5]—has been used in longer humorous texts [6, 7].

An extension of SSTH, the General Theory of Verbal Humor states that two jokes can be compared using six knowledge resources, namely, Script Overlap/oppositeness (SO), (pseudo-) Logical Mechanism (LM), Situation (SI), Target (TA), Narrative Strategy (NS), and Language (LA). The more knowledge resources they have in common, the more similar two jokes are. Moreover, the five knowledge resources form a strict hierarchy, SO– > [LM]– > SI– > TA– > NS-LA, and if two jokes vary only in one resource, the higher it is in the hierarchy, the more different the jokes are. For example, jokes that differ only in SI are more different than those jokes that differ only in NS. However, two jokes that differ only in SI are more similar than two jokes that

differ in NS & LA together. The hierarchy was empirically verified [8] and confirmed, with the possible exception of LM.

It follows then from this hierarchy that two jokes that have different SOs should be rather different. The difference may vary depending on different scripts are. This should hold for cases ranging between:

- The jokes have identical setup, but differ only in punchline that leads to different SOs[1].
- The jokes that have setups that may vary in emphasis, but similar enough that the punchlines may be swapped.
- The joke have identical setup and punchline, but differ in the perception of either major scripts or subscripts.

The next sections provide examples on the both ends of the spectrum.

2.1 Identical Setup, Different Punchline

It is possible, although not entirely probable, to have a set of canned jokes that have the same setup, but different punchlines. As an example, consider the following pair, that works much better as a spoken joke rather than a written one:

What's black and while and re(a)d all over?
- A newspaper (1)

This is an old joke that works due to the wordplay (homophone) between the color red and the past time of the verb read. The setup can be thought to ask what object consists of 3 colors, namely, that of black, white, and red. The punchline takes it to a different area, namely an object that has color white and black and has a property of being read (as in process information by reading) either from cover to cover or in its entirety.

Depending on whether the scripts are seen as semantic scripts in nature, and whether there exist overlap and semantic oppositeness, the text may be classified as a joke/pun or just a wordplay (see [9] for the differentiation between the two). The overlap of the two scripts – one being the reading of the newspaper, and the second one finding the object that has three colors – is the existence of some physical object that is white and black and has other characteristics. The oppositeness here is merely that of a homophone, and thus it serves as a good play on words in games for children.

The second part of the pair has identical setup to (1), but a totally different punchline, as demonstrated in (2), being one of many versions collected from the internet:

What's black and white and red all over?
−An embarrassed zebra. (2)

[1] We are not considering cases where a different punchline varies only in LA, TA, etc., but does not activate different script pair.

The first script of the two is identical to that of the script of joke pair on (1): finding an object that contains three colors: black, white, and red. The second script takes an advantage of the word *red*, but in a milder form than (1) did: here the meaning of *red* changes slightly to denote not the prototypical color red, but the presumed human redness of a blushing face. The second script takes this human-like transformation and transposes it to an animal, thus making it possible for an animal to change color of skin in signifying embarrassment. Since it is widely known that zebras and black and white, and one turn reddish/pinkish when embarrassed, an embarrassed zebra could have some colors on the reddish/pinkish side if such transformation was possible for zebras. The overlap, again, is the physical object that could have some colors (at least white and black). The oppositeness here is not the fact that some objects turn red when they are embarrassed or not, but the fact that zebras in this joke are embarrassed, where the world knowledge, limited as may be, tells us that it is impossible. A humor informed reader may notice that this joke is a meta joke, and the script oppositeness is of a meta-level, namely, that between the zebra joke and the underlying newspaper joke.

Notice that the in (1) the punchline takes us from an object that is being either repeatedly or carefully studied, while in (2) we are concerned with the animal that possesses a human characteristic of embarrassment, which can hardly be attributed to an animal. The change of interpretation between the two jokes depends entirely on punchline.

While it is not theoretically possible to predict the word that the punchline will play on (that will serve as a trigger to SOs) and attempt to calculate every single possible interpretation, given a perfect world knowledge, it is possible to imagine what happens if the human specifies a second verse. It is also possible that a computer cannot tie the second verse to the first one due to the lack of world knowledge: it is rather difficult to get from red to blushing to embarrassment. It is possible then, for a computer to ask questions that would narrow down the interpretations of the second line and restrict the calculation of SOs. For example, in a response to *an embarrassed zebra*, a computer may ask a question about how such a zebra would look like (pointing back to the physical characteristics identified by the first line of the joke).

It should be stated that not all jokes are that complicated in nature. Another example of an internet provided response to a newspaper/zebra joke is demonstrated in (3);

$$\begin{aligned} &-\text{What's black and while and red all over?} \\ &\qquad -\text{A blushing zebra} \end{aligned} \tag{3}$$

Notice that in this version the only connection that one needs to make to find SOs (and thus determine if this is a joke or not) is that from red to blushing, but removes the connection from blushing to embarrassment, thus significantly simplifying the problem at hand.

SOs Difference. The difference between (2) and (3) lie somewhere on the axis between totally different punchlines and identical punchlines that rely on the interpretation or emphasis that is seen by a user, but is not stated anywhere.

The difference between SOs pairs in (2) and (3) is minimal: while there is slight variation as to how to explain why the zebra is blushing – by embracement or not – the end result is the same. Thus, if we are looking at SO from the position of goal or a path to goal (as briefly suggested by [2]), we will arrive with approximately the same result for both blushing without a reason and blushing due to embarrassment. Thus, (2) and (3) are practically identical jokes – which follows a human's intuition as well.

The difference between (1) and (2) is much more pronounced: there is no intersection in meaning between the two second lines, or between how the pairs of scripts can be brought together. Thus, while the first scrip of the pairs is the same, the second scripts brings enough chaos into the jokes that the overall pair overlap is practically non-existent: we would have to compare an event of reading a particular printed material to an animal exhibiting human traits of being embarrassed.

2.2 Identical Setup, Identical Punchline, Different Knowledge of the World

On the other spectrum is the jokes that told exactly the same way, but get a totally different response from different people. As an example, consider the joke found on open Facebook political humor page:

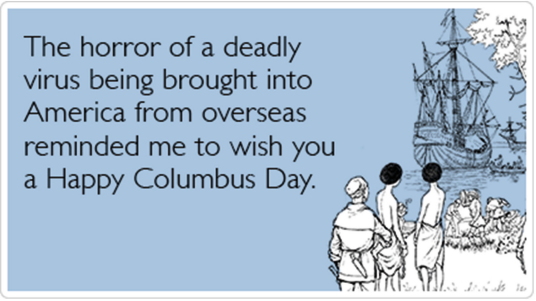

Fig. 1 One of the jokes on the Facebook political humor page

The joke, if rephrase into a text form, could be analyzed as follows: a deadly Ebola virus was brought to America from one of African counties. This virus threatens lives of people who come in contact with those that already have Ebola and exhibit its symptoms. Over 500 years before, when Europeans came to America, they brought sickness, epidemic, and deaths to many Native Americans who were not immune to many European sicknesses. However, we now celebrate Columbus Day – observed as a second Monday in October – a legal holiday commemorating the discovery of the New World by Christopher Columbus in 1492.

The script overlap is the people from a different continent brought disease to people that live in America. The script oppositeness is the time of when the disease was brought,

as well as the perception of the salience of the importance of the phenomenon: very salient in 2014, as perceived by 2014; and not at all salience 500 years ago as perceived in 2014.

The joke follows by hundreds of responses to the joke, indicating how it was perceived and which script played the most important role in this joke. It should be mentioned straight away that the responses to the joke do not indicate that the joke was entirely understood or appreciated, and some responses suggest just the opposite:

Thousands of people have recently died. Not to be too big a sick in the mud but this is tasteless and not funny enough to make the crudeness worth it. (4)

The response indicates that the author of (4) perceived that was a joke, thus, supposedly, finding both scripts, while acknowledging only one of them: namely, that people die during virus outbreak/epidemics.

The acknowledgement of this script was one of the two major themes. Interestingly, however, such responses carried a second topic: who is responsible for the outbreak in America, and, at the same time exonerate the president. Examples of the responses are demonstrated in (5) and (6).

Ebola = Enjoy Barack Obama's legacy America (5)

And this is the Presidents fault how?? He is not God by any stretch of the imagination. This disease was not apparent when the guy got here. [. . .] (6)

The second major theme was the European settlements on the Native American's territory as a consequence of the discovery of the new continent. Again, the topic of the response was an inference to the two major scripts directly accessed in a joke, namely, a conquest of the land and resettlement/eviction of the native population. As an example of such a response, see Fig. 2.

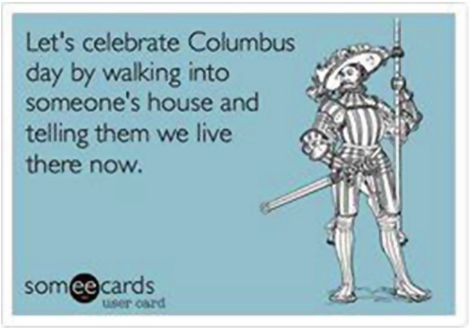

Fig. 2 One of the responses to the joke in Fig. 1

Again, the responses here identify which part of the background knowledge is salient to the reader and whether they are prepared to tolerate jokes of that nature. It is these responses that are helpful for identification of particular preferences in humor styles for an individual rather than a simple score [10].

Note all responses were negative, however, and some, such as (7) tried to joke back. The scripts at play are exactly the same, and the author is not explicitly using any inference to the secondary material, as the negative responses did. Instead, the focus is shifting from the recently salient Ebola and the deaths attributed to it, to the virus/epidemics that American Indians suffered from centuries ago. The playful shift, however, is acknowledged by the author through the second sentence, which also serves as a signal to the audience that the focus shift is intentional, not a result of a misunderstanding.

> Yeah just like the white settlers brought measles and chicken pox to American Indians. That's what you were thinking about right? RIGHT? (7)

The response to perceived jokes, obviously, does not have to be negative. As described in [11], the same detection of what part of script or subscript was perceived by interlocutor is possible within friendly communication, especially when banter continues after initial joke, as shown in (8).

> [Friend 1:]So, I'm one of the last dinosaurs who just opened the LinkedIn profile. Not sure if I knew what I was doing but I think it's up and running. I am pretty sure I selected a terrific profile picture and connected with strangers that apparently were in my Gmail account It's great to be connected :))) [. . .] (8)
> [Friend 2:] [. . .] u are kind of late in the game. Wanna buy a blackberry? [. . .]
> [Friend 1:] I dumped all my cash for Apple, I may spare some change for BB, a few cents now should be enough :)

While the conversation clearly rotates around BlackBerry, the intention of the friend 2 to produce a joke about BB being popular long time ago was misunderstood. Instead, Friend 1 caught on the fact that BB is no longer valuable, but not because is it nobody buys it anymore, but because their stock is very low. While one is likely to cause another, there are some parts of the (sub)script that were not activated by Friend 1, as intended by Friend 2. While not shown in this example, Friend 2 perceived the reasoning of low share argument and the banter continued. In other words, the intended joke was replaced by a different interpretation of the same explicitly stated material.

3 What About Embodied Humor?

We will assume for this section that a computational system that we are dealing with is not simply a read-write terminal, but that it has other senses: it can see (and has some ability to differentiate things), it can touch (and has some sensors to that effect), it can move (whether it is on wheels, legs or anything else), and it can hear (and process sounds at the same level of accuracy as it processes text). It would probably be very helpful if it could react in some way to the provided stimulus, based on existing research [12–16]. We will also assume that it can integrate the knowledge that it receives from all of its input channels and map it to its ontological representation, similarly to how it has been doing for text input. In essence, this computational system is a robot with some enhanced functionalities.

Such a system will have to sort a lot of knowledge out, including contradictory information. It will also have to address whether contradictory information is humorous or not. Let us consider a simple example: suppose, there is a bowl of raspberries on a table. Suppose, that our robot can differentiate colors and can differentiate objects such as table or bowl. Thus, he sees that the object in the bowl is reddish. Also, suppose, that he hears a comment from a person that the raspberries in the bowl are not ripe.

Now, suppose the robot can communicate in multiple natural languages. We will consider an example in Russian because of the convenient ambiguity words that mean things that can be easily perceived by a robot. Our robot hears the following sentence: *малина в вазе оказалось зеленой*. Our robot now has a problem: a word *зеленая* can mean 'unripe' (usually used in this sense with fruits or vegetables) or 'green.' This is where we either have a slight oppositeness or a potential misunderstanding: the robot can see that the color of the raspberries is close enough to red, even if it doesn't know the exact name of the color.

Now, suppose, a person who made the statement does not have in mind an interpretation of *зеленая* being the color (unless a person is a linguist, of course). Suppose, our robot responds, something like, *it looks rather red to me*. Technically, we have a joke-carrying conversation here, not unlike conversations that happen in many sitcoms. Will it be perceived as a joke by a person who is having such conversation with a robot?

Interestingly enough, the same phrase is translated by Google (11/14/2014, query "translate: малина в вазе оказалось зеленой") as *raspberry in a vase turned green*. This can be turned into a joke by a human (especially the one with linguistic knowledge), who should turn green from such translation.

What we have looked is a very simple case of when one agent (human or computational) has some knowledge or interpretation that the other one lacks (or doesn't see at first) and comments on the situation that is potentially humorous. It is up to the second participant to accept this situation as humorous, as suggested by the first, or to reject it. Provided that such situation is accepted, a further remark may be needed to acknowledge the acceptance.

References

1. Raskin, V.: Semantic Mechanisms of Humor. D. Reidel, Dordrecht, Boston, Lancaster (1985)
2. Taylor, J.M.: Towards informal computer human communication: detecting humor in restricted domain. Unpublished Ph.D. thesis. University of Cincinnati (2008)
3. Ruch, W.: The sense of humour: a new look at an old concept. In: Ruch, W. (ed.) The Sense of Humour: Explorations of a Personality Characteristic. Mouton de Gruyter, Berlin (1998)
4. Ruch, W. (ed.): Measurement of the sense of humor. Humor Int. J. Humor Res. **9**(1/2), 273–302 (1996)
5. Attardo, S., Raskin, V.: Script theory revis(it)ed: joke similarity and joke representation model. Humor Int. J. Humor Res. **4**(3–4), 83 (1991)
6. Ermida, I.: How to do humor with words: the linguistics of humor narratives. In: Raskin, V., Ruch, W. (eds.) Humor Research Series, vol. 9. Mouton de Gruyter, Berlin (2008)
7. Attardo, S.: Humorous texts: a semantic and pragamatic analysis. In: Raskin, V., Ruch, W. (eds.) Humor Research Series, vol. 6. Mouton de Gruyter, Berlin (2001)
8. Ruch, W., Attardo, S.: Raskin, V: Towards an empirical verification of the general theory of verbal humor. Humor Int. J. Humor Res. **6**(2), 20 (1993)
9. Hempelmann, C.F.: Paronomasic puns: target recoverability towards automatic generation. Unpublished Ph.D. thesis. Purdue University (2003)
10. Hempelmann, C.F., Taylor, J.M, Raskin,V.: Tightening up joke structure: not by length alone. In: Cognitive Science Conference. Japan (2012)
11. Taylor, J.M.: Ontological semantic theory of humor in a context of humorous discourse. In: Chlopicki, W., Brzozowska, D. (eds.) Humorous Discourse. UK: Cambridge Scholarly Publishers, Cambridge (in press)
12. Nijholt, A.: Conversational agents and the construction of humorous acts. In: Nishida, T. (ed.) Conversational Informatics: An Engineering Approach. Wiley, Chicester (2007)
13. Sjöbergh, J., Araki, K.: Robots make things funnier. In: Hattori, H., Kawamura, T., Idé, T., Yokoo, M., Murakami, Y. (eds.) JSAI 2008. LNCS, vol. 5447, pp. 306–313. Springer, Heidelberg (2009)
14. Tinholt, H.W., Nijholt, A.: Computational humour: utilizing cross-reference ambiguity for conversational jokes. In: Masulli, F., Mitra, S., Pasi, G. (eds.) WILF 2007. LNCS (LNAI), vol. 4578, pp. 477–483. Springer, Heidelberg (2007)
15. Veale, T.: The ABCs of XYZs: creativity and conservativity in humorous epithets. In: Manjaly, J., Indurkhya, B. (eds.) Cognition, Experience, and Creativity. Orient Blackswan, New Delhi (2014)
16. Veale, T., Valitutti, A.: A world with or without you. In: Proceedings of AAAI-2014 Fall Symposium Series on Modeling Changing Perspectives: Re-conceptualizing Sensorimotor Experiences. Arlington, VA (2014)