Formal Humor Logic Beyond Second-Most Plausible Reasoning

Christian F. Hempelmann

Texas A&M University–Commerce
Commerce, TX
c.hempelmann@tamuc.edu

Abstract
Humor employs an essential false logic which masks the incongruity of two central meanings that are brought into overlap. Formalizing this false logic—if it exists, exists intersubjectively, and is indeed essential for humor—to a degree that is sufficient for computational detection and generation of humor has been a vexing problem for computational humor research. This paper will outline several such logics, in addition to the default of reasoning in a way that is one degree more implausibly than the most common-sense logic that can connect two meanings. The results are not least influenced by a pilot study asking participants to explain different types of jokes.

Introduction
Most humor theories agree that apart from the central concept of incongruity, a component of resolution of this incongruity is an essential part of humor. This resolution is commonly assumed to be playful, incomplete, illogical. This assumption can be found expressed as early as Sarbiewski (1619/1623), who emphasizes that in statements of wit (“acumen”), the incongruity is still congruous or the resolution still incongruous (“concors discordia vel discors concordia”). Aubouin’s (1948) notion of a “acceptation-justification,” the momentary acceptance of the incongruity of humor enabled by its superficial justification, provides more appreciable detail, both on the essential nature of this component and its essentially incompleteness. In the heyday of psychological humor research in the early Seventies, several theories addressed resolution as a key stage in humor processing as well (e.g., Suls 1972, Shultz and Horibe 1974, Rothbart and Pien 1977), without necessarily focusing on its partial nature.

This false resolution mechanism has been discussed extensively in the contributions of Oring (1992, 2003) where he develops the concept of the appropriateness of humorous incongruity. But where Oring initially claimed that the “[a]ppropriateness “need not be rooted in any kind of logical validity,” (1992: 2), I claim that it cannot be rooted in full logical validity, but must remain ultimately false, which is a key point to which I will return below and which Oring clarified later: “The incongruity remains, even though points of connection between the incongruous categories are discovered” (2003: 2).

Computational approaches to humor based on linguistic theories have paid almost exclusive attention to the incongruity—or script oppositeness (SO) in terms of Raskin 1985)—of humorous texts. But systematic, detailed, and formal work on the more elusive issue of resolution has already been conducted in the context of the General Theory of Verbal Humor (GTVH; Attardo & Raskin 1991) under the knowledge resource “logical mechanism” (LM), in particular in Attardo (1994), Attardo, Hempelmann & di Maio (2003), and Hempelmann & Attardo (2011). The term has been adapted to “pseudo-logical mechanism” (pLM), because many researchers misunderstood the reasoning enabled by the LM to be valid reasoning, confused by the term “logical,” which can mean “pertaining to logic” as well as “valid in terms of [some type of] logic.” Importantly, again, the pseudo-logic of the resolution is only locally valid (Ziv 1984) and defeasible, never fully resolving the incongruity, but merely masking it.

As is well known, many computational humor generation systems use the punning pLM (cf. Hempelmann 2004) because, as I have argued, this pLM is easy to model with sound similarity or sound identity, while the underlying complex semantic effects can be largely ignored (Hempelmann 2008). But for non-ad-hoc computational humor, in particular in humor understanding—in contrast to humor generation—where input can’t be assumed to be restricted to puns alone, a fuller range of pLMs needs to be made available to computational humor systems. The present contribution is an attempt at imagining if and how this can take place, by highlighting the major problems, outlining the state of the art, and suggesting avenues for future work, some of which is already in process.
Formal False Logic

Does it Require Formalizing?

One assumption, which if correct casts substantial doubt on the whole undertaking of formalizing humor, or at least pLMs (if they exist as an enumerable class at all), needs to be taken seriously. But it should also subsequently be ignored so that progress in computational humor research can be made. This assumption is that humor, in particular the false logic of the pLM, can’t be formalized to the degree that it becomes operationalizable computationally. I don’t mean this in the sense often encountered in the criticism of formal humor theories, namely, that formalization doesn’t leave any room for the human, creative, etc., aspects of humor. Nor is it meant in the sense of the famous E.B. White quote that “analyzing humor is like dissecting a frog. Few people are interested and the frog dies of it.” We are not dissecting humor in its computational processing—generation or analysis—but rather trying to translate something pseudo-logical into languages that can’t allow for anything but normal logic. So humor won’t die, an essential part of it just doesn’t translate into the formal languages which we have so far used to handle it computationally.

What I mean is that we may well find the pLM to reside in the layer of meaning that natural language can afford to—or actually must—leave underspecified. This semantic underspecification gives NL versatility without which it can’t function in everyday meaning exchange. This may well correspond to an effect recently described with the distinction of “fast thinking” and “slow thinking” where humor resided in the dynamics between these two modes (Ventis forthcoming).

Making specific what is underspecified in non-humorous, bona-fide text in computational processing can yield useful disambiguations and clarifications. In the case of humorous texts, it should yield two partially overlapping, but opposite interpretations of the text. Allowing for two text meaning as the output of computational humor processing accommodates the two meanings, one of which might have been produced by fast thinking, a first processing pass with substantial allowances for illogical, or rather pre-logical, leaps, the other the output of slow thinking, where an additional interpretation is found, along with a gap in the initial interpretation that in principle can’t be solidly bridged, but only masks, as if a blurry picture of a bridge were held in front of it.

However, retaining the necessary relationship between these interpretations, which is what the concept of pLM aims to capture, might in principle be impossible. In other words, this important part of humor is afforded the ability to hide in natural language in a way that can’t be translated into a formal language. If this is the case, then humor is part of the attempted, incomplete symbol processing, not part of logic processing of the resulting disambiguated formal symbols and their relations. That is, the pLM may not be a logical formula that can be represented in a fully formal language, but part of human natural language processing before it becomes formal, formalizable—or a dialectic tension between unformalized interpretation and realization of unformalizability of the interpretation—and thus possibly prevented from ever becoming formalizable in principle.

How Much Logic are Humans Aware of?

A further complication for the formalization of pLMs is illustrated by the findings of a recent pilot study (Hempelmann 2012). In this study 10 jokes were presented in a Qualtrics questionnaire to 100 participants recruited online through Amazon’s mechanical turk service. Half of the jokes were short Steven Wright jokes, assumed to be highly absurd and non-sensical (NON), while the other half were textual stimuli of the highly resolvable type—called “incongruity-resolution” (INC-RES) jokes in the 3WD joke dimension type that had developed this dimension (Ruch 1995) and from which the stimuli were taken.

The main issue for this study was to see how participants would differently report on jokes that have a rather explicit pLM (INC-NON) and those that leave the interpretation, and interpretation as jokes, to the hearer. Added to the 10 main stimuli were one warm-up item and one control item to make sure participants were actually paying attention.

The questionnaire asked to rate the 12 stimuli for funniness and aversiveness on scales from 0-6, making the data compatible to 3WD studies that use exactly these scales. In addition, participants were asked to provide brief summary and a separate explanation of each joke.

Of the answer sets 97 were usable, that is, complete and not failing the control, including those of 41 male and 56 female participants, all of whom stated to be native speakers of English, had logged in from U.S. IP addresses, had more than 50 approved tasks on Amazon’s service. The average age of the participants was 34 years, the average time spent on the task was 23 minutes, compensation for participation was 50¢, dispensed anonymously through Amazon.

The mean funniness was significantly lower for NON jokes at 2.95 than for INC-RES jokes at 3.18 (SD for both .11, f(1, 96)=4.98, p > .05). Interestingly, the correlation of rated understanding of jokes (by the author) to declared funniness was .91 (with only 10 stimuli, 97 ratings for each; see Figure 1).

But more disturbingly, results showed that participants reported consistently different pLMs, in particular for NON jokes. For example, the following NON stimulus resulted in very heterogeneous explanations:

I love to go shopping. I love to freak out salespeople. They ask me if they can help me, and I say, "Have you got anything I’d like?" Then they ask me what size I need, and I say, "Extra medium."

While there is an additional complication in this joke in that it has a double punch line (“Have you got anything I’d like?” and “Extra medium”), the final punch line uses the
pLM that along a continuum you can only go “extra” towards its extremes, but not at all towards its center. This was not part of the appreciation of all participants, but they actually constructed different pLMs which provided them with a different joke that they reported to enjoy. The following are representative responses to the question what the logic of this joke was. The first three were rated to conform to the intended pLM, 4 and 5 created a different, yet plausible joke around a euphemism around obesity, and the last one reported a completely different, and arguable just as plausible, pLM with an implied script of ordering fast food:

1. funny because extra medium isn’t a real size
2. Extra medium sounds like it could exist as a size grammatically but doesn’t.
3. rather absurd answers by a punk buyer that don't make any sense - "extra" implies anything but "medium"
4. The girl is saying that she is a little bigger than a medium but not as big as a large?
5. I like the 1st half of the joke better than the 2nd half. In the 1st part the joker is making a comment on how useless most salespeople are, really. Because they don't know what the buyer likes, so how can they help them to find what they want? The last part, you expect ‘extra’ to go with 'extra large'. Extra medium makes no sense, but is a comment on the fact that the average person in the US has to wear extra large, even though they may be in the median range of size for their weight. Hey, NOW I like this joke a lot.
6. its funny because you wouldn’t ask for fries at a clothing store.

Given these divergences, serious doubt is cast on the endeavor to formalize the perceived pLM, but a wide open field of new research has been added to the agenda of linguistic humor studies.

The Second-Most Plausible Line of Reasoning

Returning to a more optimistic engineering approach to the pLM (although the most optimism for computational humor researchers is usually achieved by ignoring pLMs altogether, as noted above), in a recent (self-)reinvention of the GTVH, the Ontological-Semantic Theory of Humor (OSTH; Raskin, Hempelmann & Taylor 2010), a first approximation to a pLM-like effect has been described (Hempelmann 2009). In the ongoing knowledge resource engineering and testing of an NLP system, it became apparent that the system consistently ranked certain interpretations of the meaning of sentences second that the engineers happened to consider humorous. It appeared that the interpretation that the system considered best and ranked first, in terms of fit to the semantic expectations that the system had, beat the second-ranking interpretation, because the latter was slightly deviant from those expectations, but not as much as interpretations that were ranked even lower. This slight deviation from the constraints based on the semantic interpretation seemed to be the degree of falseness corresponding to a possibly humorous reading.

The following example is the second-ranking automatically generated interpretation of the sentence “Meggett has been acquitted on sex-related charges.”

`acquit-v1, sentencing(factivity(value(0)))
agent(value (unknown))
beneficiary(PND (Meggett-n1, football-player
on-top-of(sem (charge-n1, explosive-device)
instrument-of(sem (sex-n1, sex-event))))`

A natural language paraphrase of the sentence interpretation should make the humorous potential apparent: A football player called Meggett was not sentenced for a crime that he was accused of, and this non-sentencing took place while he was located on top of an explosive device that is used for sex. The approximation of the pLM in operation here is merely the unspecified false matching of any of the many constraints used in generating the representation of the sentence’s meaning.

Thus, there is obviously a lot of work to do, before at least more than a few incarnations of the pLM can be modeled sufficiently that the concept can be operationalized in computational humor systems. A principled engineering approach could proceed along the lines of the following assumption: “In general, partiality can be maintained at two levels. On the one hand, a fully normal logic may only apply partially to make the two scripts appear appropriate in the given context of the joke, as in a false analogy. On the other hand, the logic itself may be faulty and in any circumstance create only a semblance of appropriateness” (Hempelmann & Attardo 2011: 140). Bearing this in mind, one should attempt to acquire those pLMs that have been relatively well documented without forcing them into prefabricated schemas as humor generations allows for.

In terms of an ontological-semantic system, the following adaptations need to be made to accommodate pLMs into the processing of natural language. On the one hand, a list of intentionally false inferential rules can be crafted, modeled on correct inferential rules. These will be more useful in humor generation, as they can’t be assumed to cover all ways in which inference can be found to fail in humor that needs to be analyzed. One such rule that applies
to the example with the explosive sex device above would be that a child concept in the ontology can inherit a property and filler from a parent concept, even if the child concept itself has a more specific filler for that property. On the other hand, we can allow valid constraints of the ontology to not apply locally in the processing of a given sentence when that allows the instantiation of a concept marked as semantically opposite to another concept in the interpretation of the sentence. This presupposes a list of such oppositions as attributes in the ontology, but that is a separate issue, namely that of script oppositeness in the OSTS (see Raskin, Hempelmann & Taylor 2010).

Partially because of the general problem of the underspecified, undecidable, pre-logical nature of the pLM sketched above, further problems arise for the knowledge engineer. While there may or may not be a textual or inferential trigger guiding the resolution process, this process takes place in a less guided fashion than incongruity identification or successful disambiguating in natural language processing of bona-fide text. Not only can we not easily identify part of the actual text as being or triggering the pLM, but can therefore often be a very different pLM for different hearers of the text.

This leads to an interesting hypothesis for humor research. As pLMs can be idiosyncratic, even more so than other processes (and results) of human language processing, they possibly account for much of the variation in humor appreciation. There are merely hints for a general, and you can find your own path to pseudo-logically connect two scripts, probably close to those that other hearers of the same text construct. But especially in nonsense, where there is little to no guidance for a playful resolution, these paths may lead in completely different directions, or a hearer may not be able to or want to find a pLM path and so finds the text not just unfunny (performance), but non-humorous (competence).

Conclusion and Outlook

In sum, pLMs are a painful problem for the knowledge engineer who has to model them for a humor-competent natural language processing system. The reason is that their pseudo-logic is close to everyday, qualitative reasoning that resists reduction to a logical form with which its correctness could be decided, precisely because it is at the same time correct and incorrect. This is compounded by the fact that current computational systems have only a very weak grasp on logics outside of blunt, unambiguous first-order logic. While terms like modal, multi-valued, abductive, paraconsistent, or fuzzy logic are being more commonly used, actual applications using these types of non-monotonic logic at a level that would be useful for humor processing do not yet exist.

References

Aubouin, Elie. 1948. Technique et psychologie du comique. Marseilles: OFEP
Ventis, W. Larry. forthc. Thinking fast and slow in Humor.