

# Conceptualizing Human-Computer Intersubjectivity to Develop Computational Humor

## Paper type: Position Paper

**Henna Katariina Paakki**

Computer Science Department  
Aalto University  
Espoo, Finland  
henna.paakki@aalto.fi

### Abstract

Humor – as well as language in general – is by nature social and tied to a context. To better engage with context, computational humor could draw inspiration from the concept of intersubjectivity: the sharing of perspectives. This paper focuses on discussing the possible advantages of utilizing the concept of intersubjectivity to contextualize computational humor. Intersubjectivity in humor generation system design is discussed as a possible means of evaluation of the creative product, as well as a potential approach to generating more impressive humoristic content. Firstly, evaluation of computational humor has been wanting for more effective and versatile methods. To this problem, an implementation of sharing perspectives between the system and its users offers a viable solution. Secondly, approaches to humor generation are contrasted with interactive dialogue systems, to analyze how they contextualize humor. The comparisons show that well defined interactive design and evaluation methods that enable perspective sharing between the producer and the press would greatly benefit humor-generating systems. The final section theorizes on the possible foundations for modeling intersubjectivity in computational humor.

### Introduction

Humor and its generation have been quite widely investigated in the field of computational creativity (Wen et al. 2015; Tyler, Wilsdon, and Bodily 2020; Valitutti et al. 2016). Many studies have concentrated on how to generate textual humor, i.e. short humorous texts, puns or jokes (Tyler, Wilsdon, and Bodily 2020; Valitutti et al. 2016). Although some approaches have succeeded in humoring their human audience, most systems have remained too one-dimensional. Many models take a producer-centric cognitive approach to humor, operating on humor as a textual phenomenon that stems from linguistic or textual material via reproduction or recombination of language items like single words (Valitutti et al. 2016; Tyler, Wilsdon, and Bodily 2020). However, humor, like language itself that is used to convey the jokes, is not detached from our social life (Morreall 1983, p. 114-120). Following the Western socio-cultural and sociolinguistic view, it is in essence social, created in collaborative discourse with the surrounding social context (Bakhtin 1981; Vygotsky 1978). To create ef-

fective jokes that amuse our peers, we need to “know our audience”: to be aware of what knowledge is shared with the audience and what might be relevant to them. This allows meaningful communication with our interlocutors and places humor in an understandable context. A possibility for humor-generating systems to be able to contextualize humor can be found in the concept of intersubjectivity. Intersubjectivity refers to the relations between perspectives, and to a process of coordinating participants’ contributions to communication in joint activity (Gillespie and Cornish 2010). The central argument is that *humor creation cannot be separated from social context; to generate contextualized jokes a system needs the concept of intersubjectivity: sharing perspectives with its audience and adapting its production to the feedback received.*

This argument also brings us to the evaluation dimension of humor generation. According to Valitutti et al. (2016), there do not exist many effective evaluation methods for computational humor. From the creativity viewpoint of evaluation, systems that generate humorous texts can be creative if they portray novelty, value (Runco and Jaeger 2012), and surprise (Boden 1998; Bruner 1962). Surprise is considered here an especially important element for humor creativity, one that stems from the vantage point of producer and press perspectives (Jordanous 2016). As for the value dimension, humor is generally valued by human beings, for its intrinsic value or as a means of achieving social acceptance from others – in the case of computational generation mostly for its amusement-inducing effects. However, the question remains whether all automatically generated humor is equal in value – whether its intrinsic potential or its external evaluation by its audience might determine whether it indeed has value or not. As many aspects of humor are fitted according to the context in interaction with the expected audience (Carrell 1997a), concentrating on the latter evaluation dynamic would arguably be effective in developing computational humor. For this end, we may use the press perspective (Jordanous 2016), e.g. direct audience feedback like linguistic and humor cues, to improve computational humor.

Considering the points presented above, the evaluation dimension is a key issue in humor generation and cannot be evaded; evaluation is an essential part of both intersubjectivity as well as humor, because we (re)formulate humor content based on cues received from the press (Sacks

1974). Thus, more interactive approaches to humor generation would arguably offer an effective means of evaluation for humor-generating systems. Moreover, humor is essentially dialogic in nature and thus should not be construed as a static standalone product. On the basis of these arguments, this paper will theoretically consider the possibilities and advantages to incorporating the intersubjective aspects of humor into its computational generation. Such an approach could allow better implementation of shared context in computational humor and thus lead to more impressive results.

This paper will concentrate on the following points:

- How interactive humor-generating systems could offer a means of solving some of the problems of computational humor evaluation.
- How humor-generating systems could benefit from the concept of intersubjectivity to generate more impressive products.

First, humor theory and earlier approaches to computational humor will be discussed. Second, intersubjectivity theory is introduced as a theoretical frame, and as a concept that guides the subsequent discussion on humor evaluation. Next, context representations found in earlier approaches to computational textual humor will be analyzed. These will be compared to examples of perspective sharing between humans and dialogue systems, to illustrate possible intersubjectivity in human-computer humor creation. This analysis serves to delineate the possible dimensions of intersubjectivity that could allow humor contextualization for computational models. The final section will propose a theoretical model of intersubjectivity for computational humor.

## Earlier Approaches to Humor Generation

**Theorizing Humor** According to central theories of humor, it is constituted of three aspects: superiority, relief and incongruity (Morreall 2020). Superiority Theory states that we laugh at the misfortunes of others, reflecting our own feeling to be better. Relief Theory, on the other hand, explains that laughter is caused by the release of nervous tension (Morreall 2020). This may relate to taboo subjects embarrassing the reader and then relieving the tension. Incongruity Theory claims that a humorous effect is "achieved by the induction of incongruity in a playful context" (Valitutti et al. 2016). Incongruity, according to several definitions, involves a perception of incoherence, some contrast in meanings, or inappropriateness (Valitutti et al. 2016).

Humor research contributes to the building of a knowledge domain for a humor-generative system: this defines what constitutes its sense of humor in that it is used to delimit the outputs of the system. Thus it also guides the system's humor perspective: knowledge of humor might for example be represented by incongruity, taboos, and contextual information within the system.

**Approaches to Humor Generation** Previous approaches to textual humor generation have often focused on the generation of short humorous texts. One such example is the system by Valitutti et al. (2016). As all humor-generating systems, it needs knowledge of language and the human

concept of humor to be able to traverse the space of possible humorous products. Their system is based on word replacements in SMS texts, and they try out various combinations of constraints that define how to implement a substitution. Firstly, the form constraint requires the substitute word to be (orthographically/phonetically) similar to the replaced word, e.g. "*Which fart of town would you be in?*" (as opposed to part). The taboo constraint requires the replacement word to be taboo-inducing. The context constraints restrict the to-be-replaced words to ones in text-final or second-to-last position, or make the substitution consistent with neighboring words. The authors have also evaluated the products after generation, by conducting a large-scale survey of their funniness in the eyes of a Crowdsourcing audience.

Another example of an approach to generating textual humor is Hahacronym by Stock (2003). Hahacronym takes an acronym and creates incongruity by replacing the original words forming the acronym with at least one new word that begins with the same letter as the original and comes from the same word type. However, it has some type of an oppositional connotation in relation to the original meaning, e.g. *ACM: Association for Confusing Machinery* (Stock and Strapparava 2003). Labutov and Lipson's system (2012), on the other hand, creates two-liner jokes that provide both the set-up of the joke as well as the punchline, e.g. "*Why is the computer in hospital? Because the computer has virus.*" In quite a similar tone, Manurung et al.'s (2008) STANDUP system outputs punning riddles like "*What do you call a cry that has pixels? A computer scream.*" Computational creation of irony and sarcasm has also been a developing area of research, e.g. sarcastic bots that may have humorous implications (Veale 2018). However, I will concentrate on humor in a general sense, as considerations of irony and sarcasm are outside of the scope of this paper.

There are also systems that are somewhat more interactive, e.g. basing their production on human-rated examples (Winters, Nys, and De Schreye 2019). Also, Wen et al.'s approach (2015) moves closer to collaborative human-computer humor creation. Their system creates memes collaboratively with its users: users provide an evaluation by choosing the most amusing products (Wen et al. 2015). Overall, however, most approaches remain quite single-faceted as they do not retain much shared context between the producer and its audience. Producer-centric approaches have been criticized generally in the field of computational creativity for ignoring the important roles of the press as well as the creative process (Jordanous 2016).

## Intersubjectivity in Humor

Language – as well as humor conveyed by the means of language – does not exist in a vacuum. Language is by nature social: in language use, meaning is shared in a discourse within a language community (Vygotsky 1978; Bakhtin 1986). In Bakhtinian terms, the discourse of the surrounding community is reflected in our own speech (Bakhtin 1986, p. 96), and following Bakhtin's ideas as well as Vygotskyan philosophy, this dialogism is essential in learning the practices of our community and in learning to communicate

meaningfully (Vygotsky 1978). Humor is also a way of continuing community discourse (Carrell 1997a): it draws from what we have learned from the practices of the surrounding social environment, and from the language resources that allow us to communicate meaningfully.

Humor in interaction is not only play at linguistic form or recombination of linguistic items, but often a means to accomplish something by or within that interaction (Fine 1984). Humor is in essence social (Morreall 1983, p. 114-120). For instance, it might be used to intentionally misunderstand our interlocutor's claims to make conversation less serious (Schegloff 1987). Using Jordanous' terminology, the humorous product is created in a social process between the producer and the press (Jordanous 2016). A shared perspective is utilized to create something new.

The way humor is framed with 'contextualization cues' to signal the intended meaning (e.g. (para)linguistic features, intonation and tempo) also underlines its interactive functions (Gumperz 1982). The manner of thus testing our interlocutor's perception of humor cues illustrates how we interactively assess and formulate humor with our interlocutors. Audience acceptance of humor cues can be seen as a "go-ahead" that allows the creation of humor (Sacks 1974).

It needs to be noted that humor is an extremely difficult art form. Perfectly understanding the complex social and interactive mechanisms related to successful creation of humor may not be possible in the near future, if ever. However, some relevant aspects of these dynamics are possible to grasp, for instance the type of knowledge that is needed to understand our social context, our audience and what type of humor might be fitting there.

The next subsection will discuss intersubjectivity theory as a framework for this paper. The following one will propose how intersubjectivity together with computational creativity evaluation models could enable new methods for humor evaluation.

## Intersubjectivity Theory

Intersubjectivity theory offers viable concepts for establishing a social context in computational humor, and furthermore, a possible approach to co-creativity. It has been of great interest in psychology, philosophy, sociology, and linguistics, despite being a highly abstract concept; however, an operationalization of the domains of intersubjectivity to study both human and non-human intersubjectivity has been developed by Stevanovic and Koski (2018). Intersubjectivity is broadly understood as the sharing of minds (Stevanovic and Koski 2018), but often also as the sharing states, or perspectives (Gillespie and Cornish 2010).

Intersubjectivity has an emotional, deontic and epistemic domain (Stevanovic and Koski 2018). The emotional domain involves the sharing of emotional states: it entails both expectations of the interlocutor's emotions or expressions of affect, as well as the ability to represent their emotional state (Stevanovic and Koski 2018, p. 47-49). Examples of emotional intersubjectivity include a mutual smile, or joint affective attitudes to objects. The deontic domain refers to organization and power in social interaction: what the norms of interaction allow us to expect from others (Stevanovic and

Koski 2018, p. 49-52). For example, intersubjectivity can be found in the norms of social actions: pointing at an object invites joint attention to it, and asking a question invites a response. Finally, the epistemic domain consists of relations between subjects' knowledge states. Assumptions of possible shared knowledge influences how we engage in interaction with them and how we interpret their actions (Stevanovic and Koski 2018, p. 52-54). For instance, epistemic intersubjectivity can be achieved in following the interlocutor's gaze, noticing the target of their attention and representing what they might know about it.

To share perspectives with others people need to maintain a grasp on mutual knowledge and understanding (Clark and Brennan 1991). The shared knowledge assumptions (Clark and Brennan 1991; Stevanovic and Koski 2018) also apply to humor: jokes assume some shared perspective, be it knowledge of social norms like taboos, recent societal developments, or what one's peer group or audience is like. Ideas are influenced by feedback from the interlocutor (Sacks 1974), which is why humor evaluation is an essential part of humor intersubjectivity.

Intersubjectivity is rooted in Theory of Mind skills (Stevanovic and Koski 2018). The Theory of Mind is often used to refer to an individual's ability to infer another subject's mental states that lead to actions taken (intentions, emotions etc.) (Baron-Cohen 2000). The theory has been used to study humor processing and appreciation (Samson 2012), but it has also been found to be important for the producer during a creative process (Magerko et al. 2009). However, Theory of Mind is only one aspect of intersubjectivity (Gärdenfors 2008). Although Theory of Mind capabilities are needed in humor intersubjectivity, here the focus is on the relations of perspectives: the perspectives of the producer and press. In the relationship between these two lies the possibility for surprising the press.

As other beings besides humans also possess capabilities for intersubjectivity (Stevanovic and Koski 2018; Bard 2012), some form of intersubjectivity seems plausible for computational systems. Intersubjectivity in human-computer interaction supposes some type of subjectivity to be possessed by the computational system. It could be argued that human-computer intersubjectivity cannot be considered due to a computer not having a mind, per se. However, it can also be regarded as a subject, acting of its own accord in relation to humans. The fundamental constraints a humor-generating system operates on can also be seen as directing its (humor) perspective that can be shared with an interlocutor. In the sense that intersubjectivity refers to sharing perspectives and building relations between them, intersubjectivity is possible between a computational system and a human: both parties can share perspectives on e.g. what objects are like, what they know of them, and associate some affect with them. Some examples of intersubjective coordination between human and computer can be seen in the field of chatbot and conversational agent development (Adiwardana et al. 2020).

In computational creativity and HCI, intersubjectivity has been employed as a concept for developing co-creation (Eicher et al. 2017) and human-AI interaction (Wang et al.

2021). Intersubjectivity can be found in co-creativity: it allows the construction of a relationship between the producer and the audience, needed for the sharing of perspectives and collaboration in humor co-creation. However, intersubjectivity and co-creativity are not equal concepts: whereas co-creativity refers to collaboration in the creative process, here intersubjectivity refers specifically to how perspectives can be shared between the producer and the press to contextualize humor and to allow room for surprise.

To sum, in this paper intersubjectivity is seen as an interactional phenomenon: the sharing of perspectives, achieved in a process of coordinating relations between different perspectives in joint activity. I will refer to the sharing of affective, deontic and epistemic states as the sharing of perspectives.

### **Intersubjectivity as a Means of Evaluation**

In a Bodenian sense, computational textual humor creation might be novel in a combinatorial, exploratory or transformational manner (Boden 2004). Textual humor generation systems are often combinatorial (Valitutti et al. 2016; Stock and Strapparava 2003). However, having new combinations does not necessarily result in funny jokes. Many humor generation systems (Valitutti et al. 2016; Stock and Strapparava 2003; Labutov and Lipson 2012), do not use audience responses to reformulate their products and to produce humorous utterances that are possible and reachable, valid, appreciated, and also of high quality. There is a fine line between humor and nonsense. A system that produces anything by trial-and-error without precise knowledge of what constitutes high quality humor is not likely to be considered funny – as seen in audience reactions in Valitutti et al. (2016). To reach high quality concepts, producer and press perspectives need to be adjusted to better conceptualize what high quality space might consist of in a given context.

### **Evaluation Based on "Funniness" and "Humorousness"**

Valitutti et al. (2016) argue that there do not exist a lot of evaluation methods for computational humor. Research on computational humor generation has traditionally evaluated automatically generated humor via "funniness" or "humorousness". More precisely, in computational humor, we may thus differentiate between a subjective experience of humor vs. the objective aspect of humorousness; there is a difference between subjective experience of appreciating humor or psycho-physiologically reacting to humor – i.e. "funniness" (Ruch 2007) – and the objective aspects of content being humorous, e.g. a text's potential to induce humor appreciation – "humorousness" (Carrell 1997b). Funniness can be considered a dimension of "success" in automatic generation of humor, i.e. how strong a reaction the product induces in the audience. Although some researchers have evaluated the products of their humor systems (Valitutti et al. 2016; Tyler, Wilsdon, and Bodily 2020), most have not reiterated back to producing better products based on the evaluations.

Considering the lack of evaluation criteria or methods in computational humor, a humor-generating system could get evaluatory advantage from being more interactive: using au-

diences response data as an evaluation method in order to generate humor that is more likely to receive an amused response. Such evaluation as used e.g. in (Wen et al. 2015) comes closer to incorporating an evaluation system for funniness already within the system, as compared to many other approaches. Arguably, we could venture further employing this type of an approach. Of course, a subjective human evaluation is possible in this manner, but potentially biased: it relies more on what individuals find funny as opposed to system's potential for "humorousness", or the creativity of the system. However, if the success of the system is evaluated on the basis of the reaction achieved in the audience ("funniness"), then a subjective audience evaluation of a produced utterance as funny or not will arguably result in a funnier humor-generating system viewed from the audience perspective.

**Creativity Evaluation of Computational Humor** Considering humor-generation in light of computational creativity illustrates how and why intersubjective elements could improve humor-generating systems. On the one hand, definitions of creativity in the computational creativity field allow a framework of evaluation of the content produced by a humor-generating system by appreciating the product's creativity. On the other hand, they explicate the relationship between humor appreciation and the need for intersubjective referencing for a humor-generative system.

The alternating co-creativity model (Kantosalo and Toivonen 2016) and the dual generate-and-test model of creativity (Liu 2000) incorporate an interactive dimension within the evaluation of a creative product – between the producer and the audience. Firstly, alternating co-creativity bases itself on the Wigginsian idea of creativity as a search (Wiggins 2006). It views partners as taking turns in the collaborative creation of new concepts, to satisfy the requirements of both (Kantosalo and Toivonen 2016). Both parties rarely agree on what is relevant and interesting, so to reach high quality concepts and thus more transformational creativity, they need to adjust their views (Kantosalo and Toivonen 2016).

Liu's dual version model of creativity (2000) depicts generation as an iterative process between the producer, the domain of the product, and the audience. Once content is produced, it will be evaluated by humans and the producer will turn back to production using the feedback and new knowledge to create better products. Considering the possibility of interactive humor-generating systems, the concept of audience is highly important: not only being aware of press perspective – shared knowledge or the background of the audience – but also considering the received feedback. A system can improve by getting a direct reaction (indication of mirth or laughter) from its audience, thereby judging the product a success or a failure.

To gain insights into how perspectives could be shared between a computational system and the press, the next section will look at how contextual information is represented in earlier approaches to computational humor. This is contrasted with examples of humor contextualization in chat bot-human interaction. The following subsection will then theorize on the possible foundations needed for intersubjec-

tive computational humor creation.

## Intersubjectivity for Developing Computational Humor Co-creativity

**Contextual void** While standalone jokes can be created without the existence of an immediate listener, they are textual products that have an expected audience and a frame of reference. They do not exist in a humorous vacuum, but instead have their own historical-contextual and intertextual references to socio-cultural phenomena that ties them to our social world in a meaningful way. However, many humor-generating systems suffer from a lack of social context.

Consider for instance some products of the system by (Valitutti et al. 2016), with the replacement first and the original replaced word in brackets:

Harpy birthday! (Happy)

Remember to get the phone book from *cat* person. (that)

Ok i am leaving now *pee* u. (see)

In the audience evaluation, these were considered only mildly funny (Valitutti et al. 2016, p. 742). This is likely so, because the texts have almost no context: they show no significant reference to the social context of the supposed reader. The products of Hahacronym (Stock and Strapparava 2003) suffer from a somewhat similar detachment problem. For instance *FBI – Federal Bureau of Investigation* – reformulated as *Fantastic Bureau of Intimidation* is not very funny if evaluated by “funniness”, as very little context is given. Some perspective is shared in terms of what is assumed the audience might expect to come before “birthday” or what they are assumed to about the acronym FBI, but the perspective is not contextualized and thereby remains weak. To reach high quality humor concepts, humor needs to be placed within an understandable context.

Automatically generated two-liner jokes like (Labutov and Lipson 2012; Manurung et al. 2008) are a specific example of standalone texts that are able to place the joke within a frame of reference. The set-up of the joke provides a context, and the punchline provides the element of surprise, creating thus the humorous effect. For example, in “*Why is the computer in hospital? Because the computer has virus.*” (Labutov and Lipson 2012) the set-up provides the context by referring to a place and state we all are familiar with. The punchline takes advantage of the shared knowledge people have related to hospitals and computers. Continuing the incongruity between the concept of being in hospital and the common functions of computers, it achieves a playful effect. Thus, even short texts may be placed within a meaningful context. However, two-liners are a restricted form of humor. So, were we to move further to produce other types of humor it would make sense to consider the advantages of intersubjectivity in computational humor.

### Sharing Perspectives Between Computers and Humans

Comedians who are skilled at coming up with good jokes in an ongoing dialogue with others, jokes that really amuse their interlocutor, have a knack for analyzing and representing social context (Quirk 2015). Thereby they are able to make intertextual references to the larger prevalent context

as well as what was said previously in the conversation. In other words, comedy requires the skill to refer to a prevalent state of society or events that the audience can be known to be aware of, which relates to epistemic intersubjectivity (Stevanovic and Koski 2018). Great stand-up comedians, for instance, need to know their audience (Quirk 2015); they also possess excellent intersubjective ability, in reading the reactions of the audience and adapting their show based on cues from the press side.

For computers to know their audience, some level of intersubjectivity is needed, e.g. assumptions of share knowledge. Some examples of this can be found in implementations of chatbots or dialogue agents that are designed to interact with humans. A case in point is this multiturn dialogue between a human and the chat bot Meena (Adiwardana et al. 2020):

Human: I like how cute cows are

Meena: Ah, interesting. You like how they are cute, right?

Human: yeah, and they’re really smart

Meena: I heard they go to college

Human: cows go to college?

Meena: I heard that a cow went to Harvard.

Human: what did the cow study?

Meena: Bovine sciences.

Human: do horses go to Harvard?

Meena: Horses go to Hayvard.

Human: that’s a pretty good joke, I feel like you led me into it

Meena: You were trying to steer it elsewhere, I can see it.

Although Meena’s main job is not humor-creation, and though this sort of lead in collaborative humor-creation may, as of yet, not be intentionally reproducible using the system, as an interactive system it is impressive. The extract shows acknowledgement of both the micro-context and relevance of a specific topic to the interlocutor, as well as knowledge of more general aspects of the surrounding society. Meena continues to discuss something the interlocutor indicates as an interesting or important topic, and also introduces world knowledge of a context-relevant place of study (context being the U.S.) that can be assumed as known to both. This shared knowledge also forms the building blocks of incongruity in the set-up of the joke: it can be assumed mutually known that cows do not attend college. However, the interlocutor gets the gist and collaborates to further the joke by asking about the cow’s subject of study and also about the possibility of other animals taking up an academic career. The joke, in essence, is collaboratively created: Meena follows up on the subject of horses by exploiting the common association between horses and hay, and combining this with the phonetic/orthographic similarity between Harvard/Hayvard. As a new combination, “Hayvard” involves incongruity, but also portrays impressive intersubjective construction of a joke, as horses are introduced to the discussion by Meena’s interlocutor and not Meena themselves. The consideration of both audience expectations, world knowledge as well as the history of the discussion itself is what makes it seem like Meena “led the other person into the joke”. They might have expected a perspective like

“Yes, horses go as well, and they study Equitation science.”, but the formulation of “Hayvard” is surprising here. It shares a new surprising perspective on what could next be associated with horses and Harvard. The context-awareness in the example is impressive, and illustrates a way of sharing perspectives between human and bot.

The ability to refer back in the conversation thus arguably contributes to the perception of humor creativity. Taking what has already been said, and repurposing it to create a novel concept that surprises the human user could well fit into the idea of alternating co-creativity of high-quality concepts. Although Meena’s Hayvard joke as a two-liner would possibly be only mildly funny, arguably its effect is greater when tied to the context, surprising the interlocutor. It might be fruitful for a humor generation system to modify its views of humor e.g. based on topical cues or cues of surprise in the interlocutor’s reactions, to reach high quality humor concepts.

Some approaches to humor have already chosen to generate humor via chat bots (Augello et al. 2008; Sjöbergh and Araki 2009). The former can both generate humorous sentences and recognize humoristic expressions. It uses a pre-existing open source chatbot Alice<sup>1</sup>. The recognition of humoristic language features is achieved by connecting the chatbot knowledge base to external resources (WordNet<sup>2</sup> and the CMU pronunciation dictionary<sup>3</sup>). Although the jokes the bot tells are still two- or one-liner jokes prompted by user-selected topics, the approach is more interactive than many other models. The latter is based on retrieving jokes from joke database modules, and is thus not producing novel humor by itself. Computational memes have also been approached using a co-creative model: the system and a human user create collaborative humor using memes. During development, users chose the products that amused them the most. The choices were used to give more weight to successful products and to improve the system (Wen et al. 2015). However, this is but one approach to co-creativity in humor.

The question remains, can we do more: could we learn even more from the press to develop the system’s sense of humor? Could humor-generating systems come up with new creative jokes via sharing perspectives with humans?

**Defining intersubjective computational humor** Humor generating systems would greatly benefit from richer socio-contextual knowledge. Instead of only relying on human-ratings, it makes sense to improve the system’s understanding of humor and its context. For humor contextualization, a system needs the concept of intersubjectivity. The producer requires some shared perspective with the press, one that can be transformed to create a humorous effect. Next, I will discuss the operationalization of intersubjectivity, then considering knowledge and skills that a computational humor system would need, and providing a theoretical foundation for computational humor intersubjectivity.

<sup>1</sup><https://alicebot.org/>

<sup>2</sup>MultiWordNet: <http://multiwordnet.itc.it>

<sup>3</sup>CMU Dictionary: <http://www.speech.cs.cmu.edu/cgibin/cmudict>

The formalization of intersubjectivity is not a trivial task. In fact, strict formalization of e.g. common knowledge has been deemed almost impossible by some scholars (Monderer and Samet 1989). However, seeing intersubjectivity as the sharing of perspectives – emotional, deontic and epistemic – it is possible to reach a practical approximate formalization of the concept. For instance, it is possible to approach common knowledge in a less restrictive sense as approximate common knowledge, or as common beliefs (Monderer and Samet 1989). In practical formalization of intersubjectivity, perspectives related to emotions, social norms and common beliefs can be assumed to be nearly mutual between the producer and the audience. This allows the sharing of a humor-contextualizing perspective and its transformation into a new surprising shared perspective.

The domains of intersubjectivity (Stevanovic and Koski 2018) can be seen to somewhat align with Relief, Superiority and Incongruity Theory. All the strategies used to create humor require some understanding of the audience perspective: e.g. to relieve audience tension or concern, we need knowledge about their emotional perspective; for superiority humor, we need knowledge of power relations within our social context, and which groups the audience belongs to; to induce incongruity, we need knowledge of what objects are like and what is expected of them. Of course, for instance incongruity in inappropriateness can be related to norms of social actions (deontic domain), so the matching aligns only partly. However, this illustrates the need of context and audience knowledge in creating humor using any of these strategies. Such knowledge allows perspective sharing: it is the implication of a perspective that is often used to contextualize a joke. This entails representing the associations people usually make based on the initial perspective, e.g. the word school often makes people think of some aspect of learning. Shared perspectives can be used to surprise the audience: humor is created at the vantage point of the perspectives of the producer and the press.

For a computational system this could mean inferring some emotion, social rule or (contextual) knowledge that can be shared with the audience, using it to contextualize humor, and then shifting the mutual perspective to surprise the audience by using relief, superiority or incongruity. Based on theory and analysis, the following elements are important and to some extent possible to operationalize: understanding of shared perspectives, sense of relevance, ability to refer back to contextual elements, cue detection (affective, epistemic, deontic), and awareness of social norms. The first two represent adaptable knowledge that is needed for humor contextualization; the latter three represent skills the system needs to be able to adapt to the context. Table 1 conceptualizes these possible foundations for intersubjectivity in human-computer humor generation. The list is not all comprehensive; it provides a theoretical consideration of features some of which could be used to improve computational humor generation.

First, a computational humor producer needs an understanding of possible shared perspectives: e.g. “world knowledge”, knowledge about the interlocutor, and situational awareness. Humor systems could benefit from Web or news

Table 1: Dimensions of intersubjectivity in computational humor.

Function	Dimension of intersubjectivity
Contextualization of humor	<p><b>Shared knowledge.</b> Understanding of emotional, epistemic and deontic perspectives that are (approximately) shared with the audience.</p> <p><b>Sense of relevance.</b> Inference of perspective relevance.</p>
Inference of shared perspective and relevance	<p><b>”Memory”.</b> Maintaining audience attention: to be able to refer to past events or content in the conversation or in the social context. To be able to learn from feedback, and about the audience.</p> <p><b>Cue detection.</b> Detecting interlocutors’ attempts at and openings for humor, affective, deontic and epistemic cues, and feedback cues.</p> <p><b>Awareness of norms.</b> Understanding social norms to interpret cues, and to follow social rules.</p>

mining or semantic resources like Web Ontology Language (McGuinness and Van Harmelen 2004) or WordNet, somewhat similarly to AI systems (Davis and Marcus 2015), thus accumulating domain knowledge. Such resources have already been utilized in digital storytelling (Peinado and Gervás 2006), as well as in AI creation of paintings (Colton 2012). Inference of shared knowledge could be based on the interlocutor’s utterances and characteristics, or world knowledge generalized within a limited social context.

Awareness of shared perspectives allows the capability to infer what information is relevant to one’s audience, and where mutual attention could be directed. In this respect, relevance is closely tied to intersubjectivity. To draw from Relevance theory (Sperber and Wilson 1986), the cognitive principle of relevance dictates that people pay attention to content that is relevant to them. To maintain audience attention the producer needs to be mindful of some degree of relevance. For instance, this could entail representing what concerns the interlocutor has on their mind and how to relieve the tension, or, as Meena did, discussing a topic their interlocutor showed interest in. Contextual knowledge regarding societal events and their relevance could be drawn from already existing knowledge or in interaction with the interlocutor by reading cues.

For computational intersubjectivity, it would be useful to maintain a recollection of past events within the larger social context and the ongoing interaction, as well as what the interlocutor is like. Simulating ”memory” of past utterances, for instance, is possible using contextual machine learning models that base the next-utterance-selection on previous utterances. For example, like Meena, which is trained on

context–utterance pairs where context is the previous utterance in a conversation (Adiwardana et al. 2020). The system’s knowledge-base could, for instance, be updated with knowledge about the interlocutor and their possible beliefs based on interaction.

Reading cues in interaction is essential for making inferences of shared perspectives: if the system is able to detect humor and surprise cues on the press side, as well as emotional, deontic and epistemic cues, it is possible to make informed inferences of perspectives that could be shared. E.g. the detection of humor cues (Farnia 2019) – or openings – and building on them would allow intersubjectivity between the system and the human, and thus a novel form of humor co-creativity. Certain linguistic and sentiment cues can also reveal how the interlocutor perceives the system (Wang et al. 2021). Understanding humor appreciation cues, on the other hand, would allow immediate humor-evaluation that could be used to improve the system.

Finally, an intersubjective humor generating system may also benefit from some awareness of the social norms of human interaction, e.g. how to respond to certain utterances (Enfield, Stivers, and Levinson 2010). This is not only related to respecting conversational rules, but also offers a possibility to humor the interlocutor, as permissible rule breaking of conversation maxims can be effective in introducing incongruity into humor (Lakoff 1982).

In conclusion, to be able to create high quality products considering the audience, the producer needs to accumulate knowledge based on what is constantly learned in interaction with the press. Humor creativity and intersubjectivity are essentially interwoven as illustrated in Figure 1. While the producer (a computational system) formulates a reachable and high quality product based on its knowledge at a given time, this is only the first step. Drawing more knowledge from cues (emotional, deontic, epistemic and humor cues) in interaction with the press, it can assess whether the product is indeed of high quality. It may also (re)assess what perspectives could be shared, accumulating more contextual knowledge and making inferences of shared perspectives and relevance. Based on the contextualization, the generative system reformulates what are now construed as possible

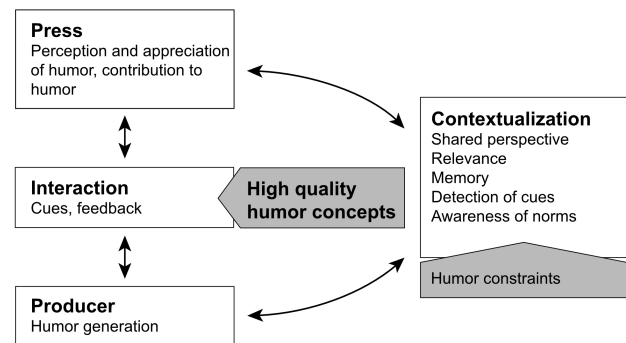


Figure 1: Intersubjective model of computational humor.

shared perspectives and thus high quality humor concepts.

## Ideas for Future Paths

Humor is a social phenomenon, and so its computational creation should also be – by utilizing the intersubjective aspects of humor in its creation. Although some steps have been taken in this direction, especially the awareness of (larger) social contexts is still lacking in computational humor systems. The product is an interplay between the producer and the press – press engagement with either the larger societal discourse, or the discourse of the more immediate social group (Bakhtin 1981; Vygotsky 1978). Thus, taking social context better into account could improve computational humor. As argued in the proposed model, to strive for more intersubjective construction of computational humor, we need to know on some level what perspectives could be shared with the audience. The domains of intersubjectivity (emotional, deontic and epistemic) align quite well with humor theories, illustrating the type of audience-knowledge that is needed for different types of humor.

There are several directions humor generation research could take. For example, humor-generating systems could benefit from online news mining, semantic webs, or conversational data sets (Henderson et al. 2019; Li et al. 2017), especially ones that incorporate a robust labeling system for humor within conversation. Some of these have already been used in some cognitive models of humor to accumulate knowledge of objects and how they could be referred to in a surprising way (Stock and Strapparava 2003), but they could be used also to better understand the audience. Also, sentiment analysis or dialogue datasets that incorporate sentiment annotations of mirth or surprise could help understand interlocutor reactions (Li et al. 2017). Such resources could allow a system to learn about, for example, polysemy or adversarial humor (Veale, Feyaerts, and Brône 2006), based on peer reactions in interaction.

On the other hand, research on humor or irony markers could be used to improve a generative system’s knowledge and perception of humor cues and what type of humor can elicit amused reactions from the audience (Farnia 2019). This could allow the system to also detect possible humor markers used by the interlocutor or potential openings for humor within a discourse. There already exist humor and irony detection systems and humor-identifying chat bots (Reyes, Rosso, and Buscaldi 2012; Joshi, Bhattacharyya, and Carman 2017; Augello et al. 2008), so it seems highly possible to create more interactive humor-detecting and generating systems. This type of decoding/encoding approach to humor-generation would allow for the system to detect humor cues or “openings” in the interlocutor’s contributions and to pursue co-creative humor-generation via perspective sharing instead of using ready jokes retrieved from a database. Detection systems could work to widen a system’s knowledge-base on what is considered as funny by humans, and allow inference of possible shared perspectives. There are also methods for detecting emotional cues, dialogue acts or intent, and stances to epistemic content, for instance (Stolcke et al. 2000; Liu 2012;

Ferreira and Vlachos 2016); methods, which could contribute to computational intersubjectivity.

Finally, interactive chat bots provide examples of contextualization that could be learned from. Being able to share perspectives and learn from them has a lot of potential for computational humor, e.g. because the system can get direct press evaluations. This could lead to new modes of evaluation that emphasize audience reactions even more than earlier models, like interactively bidding possible funny products to the user and then returning to production based on user reactions. User evaluations need not be limited to ratings or training on user selected jokes, but could also involve interpreting cues of user reactions like amusement or mirth.

**Limitations and Ethical considerations** The theoretical position taken here does not mean, however, that the non-interactive approaches dealing with short texts do not have their advantages. These offer valuable insight into how the system’s humor constraints could be designed. They could also be embedded into a more comprehensive system, which could provide them with more context and allow versatile humor techniques via using specific joke-creation modules. It is also noteworthy that the approach suggested in this paper may have its challenges, as some implementations of it may be too all-encompassing if not delimited in a sensible manner. Experimentation with more constricted generation of short humorous texts can be, as an initial approach, more clearly restricted and manageable. However, I argue that implementing some aspects of the model of intersubjectivity could develop co-creativity in computational humor.

Perhaps at the moment the more general concept of co-creativity might allow an easier approach to developing computational humor between humans and computers. However, there is something interesting that could be learned from the concept of perspective sharing, specifically for computational humor. Namely, how co-creativity in humor could be implemented: not only through using audience-rated examples in training the system, but by developing the system’s understanding of contextual humor. Humor is created at the vantage point of the perspectives of the producer and the press. Thus, understanding the audience is important for developing systems that better interact with the press. Although humor is a very difficult art form, such knowledge improves the system’s chances of creating humor that is appreciated.

Humor has an interactive social function: it can help people reach common ground, or divide them, especially when considering superiority-type humor that feeds off a certain subgroup that is made into the butt of a joke. Furthermore, since language can work as a reflector of ideology, humor can as well. Thus, humor systems should be built keeping the value aspects of the surrounding society in mind; as a means of communication, humorous products might show engraved ideological positions and might thus have an effect on the receiving end, e.g. biases related to gender, racism, hate speech, offensive language, and revealing of sensitive information. An unfortunate case in point is the Microsoft Tay chat bot. Thus, careful ethical considerations are needed in interactive system design.



**Concluding remarks** In conclusion, computational humor should not consider textual instances of humor as context-detached series of words or reproduced combinations of words, which might be considered funny by humans – or not. The opportunistic manner of producing reachable concepts that may be valid in a technical sense does not in many cases result in high quality concepts. A more impressive co-creative system will enable a sharing of perspectives between itself and humans, by considering its audience and the context they are both placed in. Collaboratively reformulating approaches to humor based on the contextualization of what is shared with the audience, such a system could achieve a more transformational level of creativity as compared to many systems showing combinatorial creativity. The importance of the concept of intersubjectivity in humor-generation is to allow more fine-grained creativity in computational humor: to approach humor as a social construct, which takes its most impressive forms in the sharing and transformation of perspectives in surprising ways.

## References

- Adiwardana, D.; Luong, M.-T.; So, D. R.; Hall, J.; Fiedel, N.; Thoppilan, R.; Yang, Z.; Kulshreshtha, A.; Nemade, G.; Lu, Y.; and Le, Q. V. 2020. Towards a Human-like Open-Domain Chatbot. *arXiv e-prints: 2001.09977*.
- Augello, A.; Saccone, G.; Gaglio, S.; and Pilato, G. 2008. Humorist bot: Bringing computational humour in a chat-bot system. In *International Conference on Complex, Intelligent and Software Intensive Systems*, 703–708. IEEE.
- Bakhtin, M. M. 1981. *The dialogic imagination. Four essays by M. M. Bakhtin*. Austin, TX: University of Texas.
- Bakhtin, M. M. 1986. *Speech genres & other late essays*. Austin, TX: University of Texas.
- Bard, K. A. 2012. Emotional engagement: How chimpanzee minds develop. In de Waal, F. B. M., and Ferrari, P. F., eds., *The primate mind built to connect with other minds*, 224–245. Harvard University Press.
- Baron-Cohen, S. 2000. Theory of mind and autism: A fifteen year review. *Understanding other minds: Perspectives from developmental cognitive neuroscience* 2(3-20):102.
- Boden, M. A. 1998. Creativity and artificial intelligence. *Artificial intelligence* 103(1-2):347–356.
- Boden, M. A. 2004. *The creative mind: Myths and mechanisms*. Psychology Press.
- Bruner, J. S. 1962. The conditions of creativity. In *Contemporary Approaches to Creative Thinking*, 1–30. Atherton Press.
- Carrell, A. 1997a. Humor communities. *Humor* 10(1):11–24.
- Carrell, A. 1997b. Joke competence and humor competence. *Humor* 10(2):173–186.
- Clark, H. H., and Brennan, S. E. 1991. Grounding in communication. In Resnick, L. B.; Levine, J. M.; and Teasley, S. D., eds., *Perspectives on socially shared cognition*. Washington: American Psychological Association. 127–149.
- Colton, S. 2012. The painting fool: Stories from building an automated painter. In *Computers and creativity*. Springer. 3–38.
- Davis, E., and Marcus, G. 2015. Commonsense reasoning and commonsense knowledge in artificial intelligence. *Communications of the ACM* 58(9):92–103.
- Eicher, B.; Cunningham, K.; Gonzales, S. P. M.; and Goel, A. 2017. Toward mutual theory of mind as a foundation for co-creation. In *International Conference on Computational Creativity (ICCC'17), Co-Creation Workshop*.
- Enfield, N. J.; Stivers, T.; and Levinson, S. C. 2010. Question-response sequences in conversation across ten languages: An introduction. *Journal of Pragmatics* (42):2615–2619.
- Farnia, M. 2019. Humor markers in computer mediated communication: Emotion perception and response. *Teaching English with Technology* 19(1):21–35.
- Ferreira, W., and Vlachos, A. 2016. Emergent: a novel data-set for stance classification. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies*, 1163–1168.
- Fine, G. A. 1984. Humorous interaction and the social construction of meaning: Making sense in a jocular vein. *Studies in symbolic interaction* 5:83–101.
- Gärdenfors, P. 2008. The role of intersubjectivity in animal and human cooperation. *Biological Theory* 3(1):51–62.
- Gillespie, A., and Cornish, F. 2010. Intersubjectivity: Towards a dialogical analysis. *Journal for the theory of social behaviour* 40(1):19–46.
- Gumperz, J. J. 1982. *Discourse strategies*. Number 1. Cambridge University Press.
- Henderson, M.; Budzianowski, P.; Casanueva, I.; Coope, S.; Gerz, D.; Kumar, G.; Mrkšić, N.; Spithourakis, G.; Su, P.-H.; Vulić, I.; et al. 2019. A repository of conversational datasets. *arXiv preprint: 1904.06472*.
- Jordanous, A. 2016. Four perspectives on computational creativity in theory and in practice. *Connection Science* 28(2):194–216.
- Joshi, A.; Bhattacharyya, P.; and Carman, M. J. 2017. Automatic sarcasm detection: A survey. *ACM Computing Surveys (CSUR)* 50(5):1–22.
- Kantosalo, A., and Toivonen, H. 2016. Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In *Proceedings of the seventh international conference on computational creativity (ICCC'16)*, 77–84.
- Labutov, I., and Lipson, H. 2012. Humor as circuits in semantic networks. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers - Volume 2 (ACL'12)*, 150–155. USA: Association for Computational Linguistics.
- Lakoff, R. T. 1982. Persuasive discourse and ordinary conversation, with examples from advertising. *Analyzing discourse: Text and talk*, 25–42.

- Li, Y.; Su, H.; Shen, X.; Li, W.; Cao, Z.; and Niu, S. 2017. Dailydialog: A manually labelled multi-turn dialogue dataset. *arXiv preprint: 1710.03957*.
- Liu, Y.-T. 2000. Creativity or novelty?: Cognitive-computational versus social-cultural. *Design Studies* 21(3):261–276.
- Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies* 5(1):1–167.
- Magerko, B.; Manzoul, W.; Riedl, M.; Baumer, A.; Fuller, D.; Luther, K.; and Pearce, C. 2009. An empirical study of cognition and theatrical improvisation. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition (C&C'09)*, 117–126. New York, NY, USA: Association for Computing Machinery.
- Manurung, R.; Ritchie, G.; Pain, H.; Waller, A.; O'Mara, D.; and Black, R. 2008. The construction of a pun generator for language skills development. *Applied Artificial Intelligence* 22(9):841–869.
- McGuinness, D. L., and Van Harmelen, F. 2004. Owl web ontology language overview. *W3C recommendation* 10(10):2004.
- Monderer, D., and Samet, D. 1989. Approximating common knowledge with common beliefs. *Games and Economic Behavior* 1(2):170–190.
- Morreall, J. 1983. *Taking Laughter Seriously*. SUNY Press.
- Morreall, J. 2020. Philosophy of Humor. In Zalta, E. N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2020 edition.
- Peinado, F., and Gervás, P. 2006. Evaluation of automatic generation of basic stories. *New Generation Computing* 24(3):289–302.
- Quirk, S. 2015. *Why stand-up matters: How comedians manipulate and influence*. Bloomsbury Publishing.
- Reyes, A.; Rosso, P.; and Buscaldi, D. 2012. From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering* 74:1–12.
- Ruch, W., ed. 2007. *The Sense of Humor: Explorations of a Personality Characteristic*. Mouton Select. Berlin: Mouton de Gruyter.
- Runco, M. A., and Jaeger, G. J. 2012. The standard definition of creativity. *Creativity research journal* 24(1):92–96.
- Sacks, H. 1974. An analysis of the course of a joke's telling in conversation. In Bauman, R., and Sherzer, J., eds., *The ethnography of speaking*. Cambridge University Press. 337–53.
- Samson, A. C. 2012. The influence of empathizing and systemizing on humor processing: Theory of mind and humor. *Humor* 25(1):75–98.
- Schegloff, E. A. 1987. Some sources of misunderstanding in talk-in-interaction. *Linguistics* 25(1):201–218.
- Sjöbergh, J., and Araki, K. 2009. A very modular humor enabled chat-bot for japanese. In *Proceedings of PACLING*, 135–140.
- Sperber, D., and Wilson, D. 1986. *Relevance: Communication and Cognition*. Oxford ; Cambridge, MA: Blackwell Publishers, 2nd edition.
- Stevanovic, M., and Koski, S. E. 2018. Intersubjectivity and the domains of social interaction: proposal of a cross-sectional approach. *Psychology of Language and Communication* 22(1):39–70.
- Stock, O., and Strapparava, C. 2003. HAHAcronym: Humorous Agents for Humorous Acronyms. *Humor: International Journal of Humor Research* 16(16):297–314.
- Stolcke, A.; Ries, K.; Coccaro, N.; Shriberg, E.; Bates, R.; Jurafsky, D.; Taylor, P.; Martin, R.; Ess-Dykema, C. V.; and Meteer, M. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational linguistics* 26(3):339–373.
- Tyler, B.; Wilsdon, K.; and Bodily, P. 2020. Computational humor: Automated pun generation. In *Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)*, 181–184.
- Valitutti, A.; Doucet, A.; Toivanen, J. M.; and Toivonen, H. 2016. Computational generation and dissection of lexical replacement humor. *Natural Language Engineering* 22(5):727–749.
- Veale, T.; Feyaerts, K.; and Brône, G. 2006. The cognitive mechanisms of adversarial humor. *Humor: International Journal of Humor Research* 19(3):305–338.
- Veale, T. 2018. A massive sarcastic robot: What a great idea! In *Proceedings of the 9th International Conference on Computational Creativity (ICCC'18)*, 120–127.
- Vygotsky, L. 1978. *Mind in Society. The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, Q.; Saha, K.; Gregori, E.; Joyner, D.; and Goel, A. 2021. Towards mutual theory of mind in human-ai interaction: How language reflects what students perceive about a virtual teaching assistant. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1–14.
- Wen, M.; Baym, N.; Tamuz, O.; Teevan, J.; Dumais, S.; and Kalai, A. T. 2015. OMG UR Funny! Computer-Aided Humor with an Application to Chat. In *Proceedings of the Sixth International Conference on Computational Creativity (ICCC'15)*, 86–93.
- Wiggins, G. A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.
- Winters, T.; Nys, V.; and De Schreye, D. 2019. Towards a general framework for humor generation from rated examples. In *Proceedings of the 10th International Conference on Computational Creativity (ICCC'19)*, 274–281. Association for Computational Creativity; University of North Carolina.