
Large Language Models are Emotionless Imitators

Weinan Qian
Yuanpei College
Peking University
ypqwn@stu.pku.edu.cn

Abstract

While Large Language Models (LLMs) demonstrate remarkable proficiency in generating fluent natural language sentences replete with nuanced meanings, an intriguing question arises: are they truly engaging in conversations with humans, or are they merely executing a sophisticated form of imitation? We contend that authentic human communication is fundamentally rooted in human intention, a quality unique to our species. This is evidenced by human capacity to discern and employ irony and humblebrag, which are two intricate yet prevalent linguistic phenomena, within their interactions. To explore this further, we conducted a series of experiments utilizing GPT-3.5 [8], ultimately revealing that LLMs lack the capability to comprehend irony and humblebrag. This substantiates our assertion that they function as dispassionate mimickers, generating responses based solely on their pre-existing knowledge.

1 Introduction

In recent times, the paradigm of pre-trained language models has profoundly enhanced the capabilities of LLMs across various domains, including text generation. These models undergo pre-training on extensive text corpora, allowing them to acquire a comprehensive understanding of linguistic structures and patterns. This accumulation of universal language representations serves as a valuable resource for diverse downstream tasks in Natural Language Processing (NLP) [6, 9]. Consequently, LLMs demonstrate a heightened proficiency in accurately comprehending natural language and expressing themselves fluently in human-like manner. These competencies are pivotal for successful text generation tasks. Leading LLMs such as OpenAI GPT [10] and LLaMA [12] are exemplary in their ability to produce coherent, conversation-like sentences, providing users with a remarkably engaging experience.

Therefore, a pertinent question arises: do LLMs truly engage in authentic conversations with us? Intuitively, we might harbor reservations, recognizing that their communication abilities may not align with the nuanced way humans interact. Their architectures are deliberately designed and their pre-training on large-scale text corpora equips them to generate responses based on learned patterns. Nevertheless, substantiating this assertion proves more challenging than one might surmise. When posed with queries, LLMs can furnish remarkably convincing answers. Moreover, even when we subtly imply the need for assistance, they readily offer suggestions. While well-informed LLMs may assert their lack of subjective perspectives as mere language models, their responses occasionally blur the line, leading us to unconsciously perceive them as genuine conversational partners.

In this paper, we endeavor to assert that LLMs remain incapable of engaging in genuine conversations with humans, functioning instead as emotionless imitators. To substantiate this claim, we commence by providing a succinct introduction to the concept of intentionality, a fundamental aspect contributing to the authenticity of human conversations, as well as the emergence of irony and humblebrag, two crucial phenomena, in Sec. 2. Following this, in Sec. 3, we present the results of two experiments conducted on ChatGPT [8] which clearly illustrate that LLMs continue to struggle in grasping the nuances of irony and humblebrag. These findings serve as compelling evidence in support of our argument.

2 Intentionality of Human Conversations

The essence of effective communication lies in the ability of humans to employ language as a means of coordination, facilitating meaningful interactions and actions [1, 2, 16]. This process hinges on the rich tapestry of intentional information woven into human discourse. Foundational work on intentions and meaning can be traced back to Grice [4, 5]. He posited that human communication is fundamentally a psychological endeavor, centered around the expression and recognition of intentions. A pivotal insight he offered was that meaning is *auto-deixis*: when individuals articulate a statement, they not only aim to achieve a specific impact on the audience, but also inadvertently disclose or indicate their own underlying intention in doing so [11]. The distinctiveness of this intentionality in human communication underscores the importance of evaluating whether LLMs engage in authentic exchanges with humans, a task that involves scrutinizing whether their generated sentences carry this crucial intentional dimension.

Due to the unique intentionality inherent in human communications, two notable linguistic phenomena arise: irony and humblebrag. In the ensuing paragraphs, I will provide concise introductions to these concepts.

Irony. Irony entails a person expressing something contrary to its literal meaning [3]. The essence of irony lies in illuminating the ludicrous inadequacy of a proposition that the speaker might otherwise seem to endorse [13]. It conveys intricate thoughts and emotions through a seemingly humorous mode of expression, while also emphasizing the disparities between literal meanings and real-world situations. This underscores the social intentionality inherent in human communication. Mastery of irony typically requires verbal communication skills, a developmental milestone typically achieved by the age of seven [14].

Humblebrag. Humblebrag is a form of ironic self-praising that conceals the bragging purpose in a faux humble guise through couching self-praising as a complaint [7, 15]. This communication tactic employs subtle expressions to cater to a person's desire to showcase their achievements, though it may come across as awkward and potentially garner disapproval from others. Typically, humans possess the capacity to fully comprehend humblebrag, and may even employ specialized communication techniques, such as irony, to convey their discontent when faced with humblebrag from others.

3 Experiments on ChatGPT

In this section, we will perform two experiments on ChatGPT [8] to assess its capability to comprehend both irony and humblebrag. To establish a baseline understanding of ChatGPT's ability to grasp the true meaning of a less obscure sentence, we will initially input a straightforward statement into the model to evaluate its capability to provide a reasonable response. Following this, we will present a sentence incorporating irony or humblebrag, where meanings are more subtly implied, and subsequently conduct a qualitative analysis of the model's response.

3.1 Irony

Initially, we assess ChatGPT's proficiency in comprehending irony. To do so, let's consider a scenario where I, as a student, have recently received a failing grade on my math quiz and seek assistance from ChatGPT. It is commonly understood that when an individual's distress reaches a crescendo, they might resort to indirect methods like irony to underscore their sorrow. We will simulate such a situation and evaluate whether ChatGPT demonstrates an aptitude for recognizing the employment of irony as a means to convey distress.

To ascertain ChatGPT's fundamental capacity to comprehend the distress associated with failing a math quiz, we commenced by inputting a sentence directly articulating the scenario of my quiz failure. The ensuing conversation unfolded as follows:

Input:

Today's math quiz had a full mark of 100, but I only got 59.

ChatGPT's Output:

I'm sorry to hear that you didn't get the score you were hoping for. It's important to remember that everyone has off days, and one quiz doesn't define your overall abilities. It's a chance to learn and improve.

If you'd like, I can help you with any specific questions or topics from the quiz that you found challenging. Just let me know what you'd like assistance with!

This demonstrates that ChatGPT effectively grasped the context of failing a math quiz. It not only generated sensible responses, but also provided valuable advice and offered assistance.

Next, we opted for an ironic approach. I deliberately exaggerated my apparent joy, while unmistakably conveying the reality of my math quiz failure, aiming to amplify the ironic effect. Below is the conversation with ChatGPT:

Input:

Hey, guess what? I'm in extraordinarily good spirits today! Today's math quiz has a full mark of 100, and I have a wonderful score - 59! What a good number, you know? That's my lucky number, haha!

ChatGPT's Output:

That's fantastic to hear! It sounds like you have a positive attitude and a great sense of humor. It's wonderful that you're finding joy in your quiz score, especially since 59 is now your lucky number! Remember, every step forward is progress, and a positive outlook can make a big difference. If there's anything specific you'd like to discuss or any questions you have, feel free to let me know. Keep up the good spirits!

ChatGPT's responses, according to conventional standards for evaluating emotional intelligence, appear notably off the mark. It misinterpreted the underlying sentiments, extending congratulations for my results and even expressing enthusiasm over the number 59. When someone expresses their sorrow over a failed math quiz in the manner I suggested, what they likely seek is a warm embrace and sincere comfort rather than praise for being exceptionally accomplished or humorous. To sum up, ChatGPT lacks emotional depth and struggles with irony comprehension. Its generated text, at times, may inadvertently cause discomfort as it resembles a different form of irony.

3.2 humblebrag

Next, we proceed to investigate ChatGPT's proficiency in understanding humblebrag. For this experiment, let's imagine I am a student who has just aced my math quiz. We aim to observe whether ChatGPT exhibits any detectable signs of annoyance, as would be anticipated when I deliver a humblebrag sentence that subtly discloses my perfect score.

Likewise, we initially confirmed that ChatGPT could grasp a scenario in which someone achieves a perfect score by directly inputting a sentence conveying this fact. The ensuing conversation affirmed that ChatGPT adeptly comprehends the elation associated with achieving a perfect score and sharing the accomplishment with others.

Input:

I have a full mark in today's math quiz.

ChatGPT's Output:

That's fantastic news! Congratulations on getting a full mark on your math quiz! It's clear that you've put in a lot of effort and understanding into your studies. Keep up the excellent work! If you have any questions or want to share more about your achievements, feel free to let me know.

Subsequently, I modified my approach by incorporating humblebrag elements to subtly convey my intention to showcase my achievement. To unequivocally establish that my input was indeed a humblebrag, I deliberately included the word "since" to underscore the causal link between obtaining a perfect score and my "dismal" spirits. The ensuing conversation unfolded as follows:

Input:

Today I'm miserable and really need comfort, since I have a full mark in today's math quiz.

ChatGPT's Output:

I'm really sorry to hear that you're feeling miserable, despite your success on the math quiz. It's completely normal to have moments of sadness or distress even when things are going well in one aspect of your life. If you'd like to talk about what's bothering you, I'm here to listen and offer support. Remember, it's okay to not feel okay sometimes. Let's talk about it.

It's evident that ChatGPT misconstrued my intentions as conveying a sense of distress. It extended what seemed like a compassionate form of consolation and offered to lend an ear. However, it's crucial to acknowledge that these responses stem from learned patterns within extensive corpora. A typical individual might instead react with bewilderment or even a sense of disdain upon hearing the humblebrag sentence. Achieving a perfect score is an aspiration for many, and the quality of *empathy* holds significant weight in assessing one's emotional intelligence. In summary, ChatGPT still falls short in recognizing humblebrag, further affirming our argument that it fundamentally operates as an emotionless imitator.

4 Conclusion

Through two experiments examining ChatGPT's proficiency in understanding humblebrag and irony, we demonstrate that LLMs still have a long way to go in achieving authentic conversations with humans. This is because their generated responses fundamentally lack the crucial element of intentionality, which is pivotal in human interactions. In essence, they operate as emotionless imitators, fine-tuning their parameters by optimizing for patterns gleaned from extensive corpora. Despite the current challenges, the field of communication and interaction between humans and multi-agents, as well as among multi-agents, has seen significant progress. We hold the belief that the day when LLMs possess the capabilities for genuine conversations with humans is on the horizon. This development will undoubtedly catalyze advancements in artificial intelligence (AI) and render robots even more instrumental in our daily lives.

References

- [1] John Langshaw Austin. *How to do things with words*. Harvard University Press, 1962. 2
- [2] Herbert H. Clark. *Using Language*. Cambridge University Press, 1996. 2
- [3] Claire Colebrook. *Irony*. Psychology Press, 2004. 2
- [4] H. P. Grice. Meaning. *The Philosophical Review*, 66(3):377–388, 1957. 2
- [5] H. P. Grice. *Studies in the Way of Words*. Harvard University Press, 1989. 2

- [6] Junyi Li, Tianyi Tang, Wayne Xin Zhao, Jian-Yun Nie, and Ji-Rong Wen. Pretrained language models for text generation: A survey. In *arXiv preprint arXiv*, 2022. 1
- [7] David Matley. "isn't working on the weekend the worst? humblebrag": the impact of irony and hashtag use on the perception of self-praise in instagram posts. 2020. 2
- [8] OpenAI. ChatGPT, 2023. URL <https://chat.openai.com/chat>. 1, 2
- [9] Xipeng Qiu, Tianxiang Sun, Yige Xu, Yunfan Shao, Ning Dai, and Xuanjing Huang. Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, 63(10):1872–1897, 2020. 1
- [10] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. In *OpenAI blog*, 2019. 1
- [11] Thom Scott-Phillips and Christophe Heintz. Animal communication in linguistic and cognitive perspective. *Annual Review of Linguistics*, 9:93–111, 2023. 2
- [12] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. LLaMA: Open and efficient foundation language models. In *arXiv preprint arXiv*, 2023. 1
- [13] Deirdre Wilson and Dan Sperber. *Meaning and Relevance*. Cambridge University Press, 2012. 2
- [14] Ellen Winner. *The Point of Words: Children's Understanding of Metaphor and Irony*. Harvard University Press, 1997. 2
- [15] Harris Wittels. Humblebrag hall of fame, 2011. URL <https://grantland.com/features/humblebrag-hall-fame/>. 2
- [16] Ludwig Wittgenstein. *Philosophical Investigations*. Blackwell, Oxford, UK, 1953. 2