Constraints and Freedom in Language Choice and Ethics of AI Applications from the Perspective of Compatibilism

Yuxuan Dong

RCF Experimental School, Beijing, China
dongyuxuan@fdfzcygj.cn

Abstract. This paper mainly discusses compatibilism about free will and determinism in language usage. In order to demonstrate the compatibility of free will with determinism, it first focuses on Hume's perspective and classical compatibilism, examining the distinctions and similarities between these two ideas. Second, the topic of language use is covered in greater detail. Finally, it is concluded that people's choice of language in a given setting is influenced by both free will and determinism, and compatibility in language use is demonstrated. It then connects to ethical concerns in AI's work, emphasizing that since AI lacks free will while producing work by copying writing samples, copyright belongs to humans rather than AI. AI may potentially reinforce human prejudices, such as gender stereotypes reinforced by training data. Finally, it clarifies ethical guideline when human use AI. Because AI's job is entirely determined by human order and training data, it lacks the free will to make decisions. As a result, they are neither accountable for or entitled to their copied work. Furthermore, due to social prejudice and the data it consumes, AI unintentionally reinforces stereotypes, particularly in the area of gender. To standardize AI use, we must follow moral norms: guard against bias provided by descriptions of AI, clarify the copyright problems, and improve equality cognitions.

Keywords: Compatibilism, AI ethics, Language usage

1. Introduction

The debate between free will and determinism has been a core topic in the field of philosophy for a long period; compatibilism provides a pivotal perspective that bridges these two concepts. This paper first sorts out the core statement of compatibilism, focus on David Hume's viewpoint and traditional compatibilism, and then clarifies the existence of free will and determinism in specific situations language using. On this basis, it further explains the moral dispute in AI's creative behaviors, discuss whether it has free will combined with its working mechanisms and moral problems such as responsibility identification and copyright attribution. In addition, aiming at the bias caused by AI applications, this paper analyses their causes and provides ethical guidelines for AI applications.

2. Compatibilism

2.1. Hume and classical compatibilism

It is known that one of the most influential statements about free will in compatibilism is from David Hume. The main ideas from Hume are around two words, liberty and necessity. As for liberty (freedom), he believed that free will, free will is the unity of will and behavior. If a person's behaviors are driven by his or her wishes and desires, liberty of spontaneity exists. Another significant point in his argument is necessity. Human personal judgments from daily life and past experiences prove that people poses the capacity to associate and reason the future events, then act according to their judgment, which is called customary inference. This kind of subjective preferences is the nature of necessity [1-6]. However, individuals always confuse long-term empirical perceptions with absolute rules. According to Hume, we need to separate the logical necessity and the customary inference; in order to prove liberty and necessity are not against they are compatible [1,3] .

According to the classical compatibilism, the main factor of free will is that the agent has an alternative choice when making decision. To put it another way, freedom is the ability for people to do whatever they want and to do different things if they have different desires. The fundamental tenet of freedom is that there are no unnecessary blocks and that volition may control conduct. Determinists always believe that there are objective causal chains that are constant and unchanging in the world. Individuals are unable to choose their own response or course of action; instead, they must accept the particular result since all human behaviors and actions are constrained by the laws of nature and historical facts. However, the determinism is a universal rule which is the base of world's operation, it will not lead to the deprivation of freedom directly; only the extraneous force, like coercion, will restrict the freedom [4,5].

Hence, free will and determinism are compatible.

The classical compatibilism provides a clear distinction between what persons control and what excludes this controllable range. In the other word, freedom is not absolute, free will only exists when people can do what they what to do without being constrained by law of nature or facts of the past, otherwise, selecting is influenced by extraneous factors, such as the social bias desirability or universal norms.

The difference between Hume's statement and classical compatibilism is that, the former places greater emphasis on whether the individual act of their own free will, while the latter concentrate on possibility of choosing. However, they both acknowledged that humans have the freedom to act and make decisions since free will and determinism are compatible. In my opinion, there are certain similarities between these two theories. An individual has the option to act differently at the same time if they act out of free will. It is difficult to predict an individual's course of action since free will is intrinsically contingent. The agent has the flexibility to modify their decision because of this unpredictability.

2.2. Compatibilism in language use

When selecting language types, individuals have free will, as they can choose appropriate language based on their own wishes, which can be related to the identity expression. It is generally recognized that people can strengthen their identity by using specific words or linguistic expression. For instance, rock fans use slang to identify themselves as part of the rock fan group. It serves as a symbolic entry into rock music culture. They call themselves "scene kids," and when the music gets

louder, fans divide into two groups and charge at one another, a phenomenon they refer to as "wall of death." The aforementioned terminology is unique to rock fans and contributes to their overall identification as such [5]. They have the willingness to communicate with others who share their interests; therefore, they opt to use this language system deliberately. If they decide to alter their minds, they can also use another language in everyday situations or in formal settings. This is proof that free will exists in language use. They may want to show others their reflections and alternative thoughts on those great themes, for instance, the common words appear in rock lyrics such as life, love, death and revolting, words can be used in language fans taking, as a tool to prove they are rock fans [7]. However, this kind of language selection is not based on willing, but forced by the extra stereotypes, which is evidence of determinism existing in language using.

3. Link to moral problems of AI's work

It is well-documented that artificial intelligence has the ability to learn human behaviors by summarizing and imitating the existing samples. For instance, if articles by a renowned writer are provided and artificial intelligence is instructed to imitate the writer's style, it will generate a product with a similar style. The similarity of style is achieved by artificial intelligence through a series of deliberate manifestations such as enhancing the use of the author's typical vocabulary, imitating his or her emotions and expression habits. However, this kind of work cannot be a substitute for the author's work, the emotion and experiences described in it is just precisely calculated assembly. Additionally, artificial intelligence does not own the copyright of this kind of product. Actually, artificial intelligence lacks free will according to the above discussion.

First and foremost, the academic interpretation of artificial intelligence is, the interdisciplinary science that studies and develops theories, method, technologies, and application system to simulate, extend, and expand human intelligence. Through this definition, we can see that artificial intelligence has the ability to make decision, although the deciding process require tons of information and calculating to support. However, the process of writing is neither a category within the range of making decision nor a wish or willing from AI itself, which means that this behavior is not from free will. Artificial intelligence lacks free will and the capacity to make decisions when engaging in such behavior. It is a human hierarchy. In this case, humans act as a god with built-in determinism, deciding what AI must accomplish. Humans also determine the writing style and context, giving AI no creative flexibility. AI is forced to write because it must obey its owner's instructions.

According to Kant, human has duty due to the existing of free will, if a person has free will, then he has freedom to decide what he will do and what he will not do [8]. He needs to take responsible to the decision he make at the same time, because it all come from his free will. On the contrary, if the decision is not made by him, he is not responsible for it. For artificial intelligence, this kind of theory is also applicable. If AI makes the decision, it has a duty on this decision and need to bear the risk and consequences that comes with it; if the choice is not chosen by AI itself, it has no need to undertake the duty. Obviously, the imitation of writing belongs to the latter situation, AI has no responsibility to the writing product.

It is known that, rights always come with responsibilities [9-11]. They are an inseparable whole, and the core of right lies in human's freedom, which involves the mutual restrictive relationship between individuals. Kant define the inherent innate right of human as "freedom conforming to moral laws", stating that "freedom is the independence from the coercive will of others and, in accordance with the freedom of all." This kind of right is never isolated with others; it begins with the starting of mutual restrictive relationship between individuals. Take an example, when an

individual possesses the right to dispose their property freely, then this right imposes on others the responsibility of "not to encroach others' property" at the same time. On the contrary, as a right-holder, the individual must assume the duty to avoid violating other's property rights. Kant point out that, responsibility is a necessity consequence from free action, and since rights coms from free will, hence, existing of right must correspond responsibility—there can be no tight that exists independently of responsibility.

According to the theory that right exist with responsibility simultaneously, we can learn that AI also does not possess the intellectual property rights of the imitation products, although it is "produce" by itself [1]. A person with free will who gives commands to make these works should be the owner of the copyright. Because their ideas and style are included into these works, the author is also responsible for the work's influence once it has been disseminated. Nevertheless, this type of copyright is biased as we cannot determine the degree to which the author's intentions have been altered. Therefore, when we utilize AI as a tool to aid in the creation of work, we must be careful to follow the moral guidelines in order to prevent the ambiguous copyright and misunderstanding that it brings. The usual moral problem is also including problems of stereotype and bias from AI algorithm may be enlarge the bias hold by human beings.

4. The moral rules we should obey when using AI

As mentioned above, not only the copyright issues, but also the values problems related to AI have always been controversial.

Taking gender problem as an example, AI has dramatic stereotype on gender issues. AI always defaults to assigning the high-status and professional career such as doctors and engineers to men, while nurse and cooks for women. Llama 2 generates content where women perform housework chaos much more times than men, and they are often tied with words like family and children. Men, by contract, are usually given adventurous or dominant roles. Same situation in producing pictures is as common as creating writing work. This kind of fixed role setting can be harmful to people who take it in without consciousness, leading to strong stereotypes in gender, especially for teenagers who are in the stage of forming values and gender awareness. Young girls may be lack of confidence because they are told by AI, they cannot perform better that boys in stem field because of natural reasons and gender differences. However, the fact is that, gender difference is not just caused by biological differences like hormones, it is also shaped by social background and culture. Such a complex problem is simplified by AI as the one-side conclusion, and the significance of stereotypes from AI is surprising exceeding human levels.

When it comes to the causes of AI gender stereotypes, one thing we can make sure is that, the values it holds is not influenced by only one side. Researched by a team from the Technical University of Munich reveals that the intensity of gender stereotypes in text-image generation models varied by language. In language with grammatical gender, such as German and Spanish, the gender bias appears more frequently, while in language without grammatical, like Chinese and Korean, the situation will be relatively milder. This discovery may have to do with how language shapes people's ideas, even in the field of artificial intelligence. In addition to the aforementioned, training data from social culture that builds up over time is crucial in promoting gender bias in AI. For instance, statistics from the workplace shows a higher proportion of male leaders than female ones, which contributes to the prejudice that "females are not fit with leadership" and ultimately leads to the incorrect conclusion that women are mostly found in families. The underlying causes of AI gender stereotype issues are real-world gender inequity and traditional labeling prejudice.

For the whole society, it has better to be controlled. People are trying to improve the cognition and reflection of individuals or society on gender roles, rights, and equality. Gender differences should not tie up the possibility of a person's development, and it should not be an excuse to limit rights, opportunities of any gender. In field as AI and media, this point becomes more significant: if the evaluations claimed to be from an authoritative and objective perspective are still biased, gender equality will be severely hindered. We need always be vigilant against to the stereotypes from AI in order to avoiding bias.

5. Conclusion

This paper explores compatibility between free will and determinism, its reflection in language use, and AI work moral problems. Hume's perspective and classical compatibilism both confirms that human behaviors are affected by two factors, self-selection from their own willing and inevitable constrains—evident in language selection: people often choose the appropriate language according to the identity they desire, as well as restrictions from stereotypes and situations.

References

[1] Arnold, G., & Gerke, M. S. (2022). Copyright in AI-generated works: Who owns the output? [Web article]. Journal of Intellectual Property Law & Practice.

[2] Connor, C., & Franklin, C. T. (2022). Free will. The Stanford Encyclopedia of Philosophy.

[3] Hume, D. (2000). An enquiry concerning human understanding (T. L. Beauchamp, Ed.). Oxford University Press.

[4] McKenna, M., & Coates, D. J. (2024). Compatibilism. The Stanford Encyclopedia of Philosophy.

[5] Mocovíč, S. (2024). The punk discourse: From subculture to lifestyle. Routledge.

[6] Russell, P. (2021). Hume on free will. The Stanford Encyclopedia of Philosophy.

[7] Whiteley, S. (2000). Rap and hip hop: Community and cultural identity. Pluto Press.

[8] O'Connor, T., & Franklin, C. (2023). Free will. Stanford Encyclopedia of Philosophy.

[9] Kant, I. (1996). Practical philosophy (M. Gregor, Trans.). Cambridge University Press.

[10] Shanahan, M. (2023). The frame problem. Stanford Encyclopedia of Philosophy.

[11] Kant, I. (1996). The metaphysics of morals (M. Gregor, Trans.). Cambridge University Press. (Original work published 1797)