

A Framework for Dialogue-Based Human-AI Creative Collaboration

RODOLFO OCAMPO and OLIVER BOWN*, The University of New South Wales, Australia
KAZJON GRACE, The University of Sydney, Australia

Human-AI co-creative collaboration has been proposed as a model that integrates the strengths of both humans and creative algorithms. Several frameworks have been developed to classify and guide the design of such systems. However, these models lack communication mechanisms that enable the emergence of a common ground between humans and machines through a mutual adaptation of understanding about goals and meanings, a crucial component in all collaborations. We argue that dialogue is a mechanism that serves this purpose and can be included in human-AI co-creative systems to that end. We propose a breakdown of dialogic creative interaction and use it to analyze co-creative dialogue with GPT-3.

CCS Concepts: • **Human-centered computing** → *Natural language interfaces; Collaborative interaction; Interaction design theory, concepts and paradigms.*

Additional Key Words and Phrases: Dialogic Creative Artificial Intelligence, Co-Creative Systems, Mixed-Initiative Creative Systems

1 INTRODUCTION

In recent years, the computational creativity and human-computer interaction communities have placed special focus on human-AI co-creative collaboration as a model that sits between autonomous creative systems and creativity support tools, enabling a third approach that integrates the strengths of humans and artificial intelligence [13] [10] [8] [5].

Several frameworks and interaction models have been developed with the purpose of classifying possible co-creative actions and guiding the development of co-creative systems [10–12]. While these models have provided solid foundations to build mixed-initiative, co-creative, turn-taking systems, we believe that currently these models do not successfully account for shared understanding about goals, intentions and creative concepts, crucial to any collaboration [4, 7]. Dafoe et al. [4] suggest that successful complex cooperation between human and AI hinges on the emergence of a common ground, which allows agents to adaptively learn and act in "accordance with human intentions, preferences and values". We argue that dialogue is a powerful mechanism that can serve this purpose in human-AI co-creative systems, and enable more effective and aligned collaborations.

Acknowledging that "dialogue", like "creativity" and "art" has a diversity of meaning associations, we pinpoint a specific one here. *Dialogue is a process of forming agreement, clarity, refinement or elaboration of concepts, representations, goals, plans or roles.* The physicist David Bohm [1], who developed an interest in cooperation and conflict resolution, proposed that dialogue differs from other kinds of communication in that it must involve a mutual adaptation of understanding.

In earlier work [2], we speculatively explored the role of dialogue in human-AI co-creativity and introduced the term Dialogic Creative Artificial Intelligence (DCAI). We proposed, simply, that in DCAI, two or more actors should be mutually influenced in a creative process, in a way similar to Bohm's idea of mutual adaptation. The most obvious medium for dialogic interaction is natural language, and the emergence powerful language models, attention mechanisms and multimodal models with natural language interfaces is a key motivator for considering a dialogic framework for creative AI interaction. However, other interface modalities such as GUIs can also be analysed in dialogic terms. Especially important are hybrid interfaces as they seem likely to become increasingly common, and more effective than language interfaces alone.

In this paper, we build on our previous work to propose a typology for DCAI, seeking to analyse dialogic interaction by breaking it down into a set of dialogic actions. We argue that a breakdown of dialogic interaction

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

into its constituent components is a useful step in the development of interfaces and systems for dialogic creativity. We are not aware of this existing in the academic literature in CC, HCI and interaction design.

We then provide an example of a DCAI interaction using GPT-3 and analyse it using our typology.

2 RELATED WORK

This work is closely related to efforts in mixed-initiative generative systems and human-computer co-creativity to classify interaction flows. For example, Muller et al. [10] extend previous work by [12] to come up with a classification of 11 actions that can be performed by humans and computers in creative collaborative workflows: learn, ideate, constrain, produce, suggest, select, curate, assess, adapt, assemble, wait. Such work supports the informed design of mixed-initiative user interfaces. Our proposal to classify dialogic interaction develops this theme, suggesting a variation of the framework that focuses on the mutual development of understanding, as well as integrating natural language as the primary interface to achieve this. We consider this a complementary, alternative viewpoint to these other frameworks.

Earlier work with the field of computational modelling of natural language seeks to classify dialogic acts (or "moves" [6]) in natural conversations in order to understand the purpose and effectiveness of individual utterances. Core and Allen [3] sought to define dialogue in terms of (i) Forward Communicative Functions, grounded in traditional speech act theory, such as statements and directives, (ii) Backward Communicative Functions, which refer back to the dialogue, such as answering a question, accepting a proposal, confirming understanding, and lastly (iii) Utterance Features, which concern the form and content, such as whether an utterance concerns the communication process itself or the subject matter. We use this theoretical background to inform our typology.

Implicit in our initial work on DCAI [2] is the idea that creative dialogic interaction has two levels of communication: *about* the creative product and *through* the creative product. This distinction serves as the highest level classification of actions in our typology. A similar distinction was made by Kellas and Tree in their human-AI co-creative interactive sense-making process [9], which was then used by Rezwana and Maher in their COFI framework for classifying human-AI co-creative systems [11], in which interaction can happen with the shared product and between collaborators. Moreover, the conceptualisation of this interaction as a sense-making process is close the concept of dialogue as a process of mutual adaption of understanding and through which a shared context emerges. We use the COFI framework to inform the types of actions at each of the two levels in our typology.

3 A TYPOLOGY OF ACTIONS FOR THE DCAI

We thus make a primary high-level distinction for classifying types of actions: actions through the artifact, referring to those that intervene directly in the creative product, and; actions about the product, that not act directly on the creative product but rather serve as a communicative exchange between collaborators to better adjust their understanding, alignment and creative efficacy.

In each case, we consider a set of dialogic acts that participants engage in through that modality. These are in part derived from the above literature, and in part through the authors' workshopping dialogic examples:

Actions through the artifact

- **Create new:** propose a new initial creation
- **Extend:** extend an existing artifact
- **Branch:** propose several extension alternatives
- **Transform:** change an artifact's style or characteristics
- **Refine:** add more detail to an artifact

Actions about the artifact

- **Set goal:** explicitly specify a goal

Human: My intention is to write a short story for children that illustrates the value of communication in a collaboration. Is that clear? (SET GOAL)
 AI: Yes, that is clear. (SET GOAL)
 Human: Can you paraphrase my goal just to check we are clear? (REQUEST CLARIFICATION)
 AI: You want to write a story that shows how communication is important in a teamwork setting. (CLARIFICATION)
 Human: Excellent. Can you help me come up with three alternatives for possible premises? (REQUEST CREATE NEW)
 AI: 1. A group of animals need to work together to get across a river. 2. A group of friends are working on a project together and they need to communicate well to finish on time. 3. A group of people are lost in the woods and need to work together to find their way back. (CREATE NEW)
 Human: Ok, I like the first one. Let's develop it a bit further. (SELECT AND REQUEST EXTENSION)
 AI: In the story, a group of animals are trying to cross a river, but they are having trouble. They can't seem to agree on how to do it. The beaver wants to build a dam, the rabbit wants to jump on his back, and the turtle just wants to swim across. (EXTENSION)
 Human: I like that. But maybe the rabbit would want to jump on the turtle's back. (FEEDBACK)
 AI: Yes, that could work too. Anyway, the animals are arguing and not getting anywhere. Suddenly, a hawk swoops down and grabs the rabbit. The other animals are terrified and don't know what to do. (FEEDBACK AND EXTENSION)
 Human: I think that's too violent, let's remove the hawk. (FEEDBACK, REQUEST TRANSFORMATION)
 AI: Ok. Anyway, the animals are still stuck and don't know what to do. Then, the beaver has an idea. He says, "Let's all work together. We can build a bridge." The other animals agree and they get to work. (TRANSFORMATION)
 Human: Ok. Now let's write the first paragraph of a story, can you help me with that? (REQUEST FOR REFINEMENT)
 AI: In a world where animals have to work together to get by, a group of friends are having trouble. They can't seem to agree on how to cross a river. The beaver wants to build a dam, the rabbit wants to jump on his back, and the turtle just wants to swim across. (REFINE)
 Human: Remember we said the rabbit will jump on the turtle's back. Try again. (FEEDBACK, REQUEST TRANSFORMATION)
 AI: In a world where animals have to work together to get by, a group of friends are having trouble. They can't seem to agree on how to cross a river. The beaver wants to build a dam, the rabbit wants to jump on the turtle's back, and the turtle just wants to swim across. (TRANSFORMATION)
 Human: While the turtle could carry the rabbit, she would not be able to hold the weight of both the beaver and the rabbit. Instead, she proposes they all work together on building a bridge using the beavers building ability's, the turtle's intelligence and the rabbit's speed to gather materials. (EXTENSION)
 AI: The other animals agreed and they started building. (EXTENSION)
 Human: Ok, now please write the last paragraph of the story with a happy ending. (REQUEST EXTENSION)
 AI: The animals finished the bridge and were able to cross the river. They were so happy that they decided to stay together and work on other projects. They knew that if they worked together, they could accomplish anything. (EXTENSION)

Fig. 1. Transcript of a dialogic co-creative interaction. Red: about the artifact. Blue: through the artifact.

- **Request:** request an action
- **Clarify:** reduce uncertainty related to a request
- **Feedback:** provide an assessment on a creative action
- **Justify:** justify a creative decision
- **Select:** chose an option from an ideation, creation, extension or branching action.

4 A WORKED EXAMPLE

We show a process of dialogic creative interaction with a worked example of co-creative children's stories writing. We use one of the SOTA models we believe to currently afford the most dialogic creative interaction: GPT-3-Instruct, which is different from the original GPT-3 in that it was explicitly trained to follow user instructions as part of OpenAI's work on AI alignment. We believe this model is one of the most currently suited models for dialogic interaction.

In figure 1 we show the dialogue as it was conducted through OpenAI's playground, with an annotation showing types of actions (written between parenthesis after each action).

4.1 Discussion on worked example

Through a series of example dialogues like this, which we will report in full in a future paper, we have found that GPT-3 is capable of meaningful generation, yet with limits to its dialogic capability. As this excerpt shows GPT-3, is relatively successful at following the user's intended distinction between actions about and through artifacts, even without an explicit identifier. However, this may be partly due to the authors' experience interacting with

and prompting GPT-3, as well as our subjective bias in conducting this research. An inexperienced user, or a user with other creative objectives, may find it harder to have a creative dialogue with GPT-3 in a free-text interface.

We see that most actions from the model depend on a user-driven request. This is in part the result of the model being specifically trained to follow user instructions, and thus appears to lack agency to self-initiate actions. For example, we see that there is missed potential for the system to perform clarification actions in response to uncertainty from a previous interaction or feedback on a user proposal.

Our actions about the artifact sit at a level of more granular and specified participant actions than our actions through the artifact. Thus, in Muller’s [10] framework, ‘ideation’ is developed as a co-creative interaction stage, but is not associated with a specific user action and might be hard to isolate from other stages in a creative process outside of formal design structures. We could consider the entire first half of the example dialogue here to be of an ideational nature, within which we can identify several dialogic actions, such as requests, feedback and clarification. We believe that a more detailed analysis of these levels of interaction and types of actions may reveal inherent logics that help inform interaction design, and consider this an important next step. This workshop paper serves to stimulate feedback and discussion about this possible direction.

5 CONCLUSION

In this paper we have proposed a a typology of dialogic actions that constitute creative dialogic interaction. We used this typology to classify a dialogic interaction with GPT-3 in co-writing a short story. We then considered how this typology can be used to understand the dialogic and creative affordances of generative models in order to guide more effective collaborative creative AI experiences.

In discussions of GPT-3 and other models it is already widely acknowledged that although natural language interfaces have powerful potential they can be misleading when approached anthropomorphically. An interesting future direction for our research is to consider how structured GUI interfaces could enable the dialogic affordances of current systems.

ACKNOWLEDGMENTS

This research has been supported by an Australia Research Council Discovery Project grant, DP200101059.

REFERENCES

- [1] David Bohm, Peter M Senge, and Lee Nichol. 2004. *On dialogue*. Routledge.
- [2] Oliver Bown, Kazjon Grace, Liam Bray, and Dan Ventura. 2020. A Speculative Exploration of the Role of Dialogue in Human-Computer Co-creation. In *Proceedings of the 11th International Conference on Computational Creativity*.
- [3] Mark G Core and James Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *AAAI fall symposium on communicative action in humans and machines*, Vol. 56. Boston, MA, 28–35.
- [4] Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson, and Thore Graepel. 2021. Cooperative AI: machines must learn to find common ground. *Nature* 593, 7857 (May 2021), 33–36.
- [5] Nicholas Mark Davis. 2013. Human-Computer Co-Creativity: Blending Human and Computational Creativity. In *Ninth Artificial Intelligence and Interactive Digital Entertainment Conference*. aaai.org.
- [6] Jonathan Ginzburg and Raquel Fernández. 2010. 16 Computational Models of Dialogue. *The handbook of computational linguistics and natural language processing* (2010), 429.
- [7] Jean-Michel Hoc. 2000. From human – machine interaction to human – machine cooperation. , 833–843 pages.
- [8] Anna Kantosalo and Hannu Toivonen. 2016. Modes for Creative Human-Computer Collaboration: Alternating and Task-Divided Co-Creativity. In *Proceedings of the Seventh International Conference on Computational Creativity*.
- [9] Jody Koenig Kellas and April R Trees. 2005. Rating interactional sense-making in the process of joint storytelling. *The sourcebook of nonverbal measures: Going beyond words* 281 (2005).
- [10] Michael Muller, Justin D Weisz, and Werner Geyer. 2020. Mixed Initiative Generative AI Interfaces: An Analytic Framework for Generative AI Applications. In *Proceedings of the Workshop on the Future of Co-Creative Systems, ICCS 2020*.
- [11] Jeba Rezwana and Mary Lou Maher. 2021. COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. *computationalcreativity.net* (2021).

- [12] Angie Spoto and Natalia Oleynik. 2017. Library of Mixed-Initiative Creative Interfaces <http://mici.codingconduct.cc/>. <http://mici.codingconduct.cc/>
- [13] Georgios N Yannakakis, Antonios Liapis, and Constantine Alexopoulos. 2014. Mixed-initiative co-creativity.. In *FDG*.