WELLS, S. and SNAITH, M. 2023. On the role of dialogue models in the age of large language models. In Grasso, F., Green, N.L., Schneider, J. and Wells, S. (eds.) *Proceedings of the 23rd Workshop on computational models of natural argument (CMNA 2023)*, 3 December 2023, [virtual event]. CEUR workshop proceedings, 3614. Aachen: CEUR-WS [online], pages 49-51. Available from: https://ceur-ws.org/Vol-3614/abstract2.pdf

# On the role of dialogue models in the age of large language models.

WELLS, S. and SNAITH, M.

2023

# On The Role of Dialogue Models in the Age of Large Language Models

Simon Wells[1], Mark Snaith[2]

[1]Edinburgh Napier University, 10 Colinton Road, Edinburgh, EH10 5DT, Scotland, UK
[2]Robert Gordon University, Garthdee House, Aberdeen, AB10 7QB, Scotland, UK

We argue that Machine learning, in particular the currently prevalent generation of Large Language Models (LLMs) [1], can work constructively with existing normative models of dialogue as exemplified by dialogue games [2], specifically their computational applications within, for example, inter-agent communication [3] and automated dialogue management [4]. Furthermore we argue that this relationship is bi-directional, that some uses of dialogue games benefit from increased functionality due to the specific capabilities of LLMs, whilst LLMs benefit from externalised models of, variously, problematic, normative, or idealised behaviour. Machine Learning (ML) approaches, especially LLMs , appear to be making great advances against long-standing Artificial Intelligence challenges. In particular, LLMs are increasingly achieving successes in areas both adjacent to, and overlapping with, those of interest to the Computational Models of Natural Argument community. A prevalent opinion, not without some basis, within the ML research community is that many, if not all, AI challenges, will eventually be solved by ML models of increasing power and utility, negating the need for alternative or traditional approaches. An exemplar of this position, is the study of distinct models of dialogue for inter-agent communication when LLM based chatbots are increasingly able to surpass their performance in specific contexts. The trajectory of increased LLM capabilities suggests no reason that this trend will not continue, at least for some time. However, it is not the case that only the one, or the other approach, is necessary. Despite a tendency for LLMs to feature creep, and to appear to subsume additional areas of study, there are very good reasons to consider three modes of study of dialogue. Firstly, LLMs as their own individual field within ML, secondly, dialogue both in terms of actual human behaviour, which can exhibit wide quality standards, but also in terms of normative and idealised models, and thirdly, the fertile area in which the two overlap and can operate collaboratively. It is this third aspect with which this paper is concerned, for the first will occur anyway as researchers seek to map out the boundaries of what LLMs, as AI models, can actually achieve, and the second will continue, because the study of how people interact naturally through argument and dialogue will remain both fascinating and of objective value regardless of advances made in LLMs. However, where LLMs, Dialogue Models, and, for completion, people, come together, there is fertile ground for the development of principled models of interaction that are well-founded, well-regulated, and supportive of

*CMNA'23: International Workshop on Computational Models of Natural Argument, 2023, London, UK*
s.wells@napier.ac.uk (S. Wells); m.snaith@rgu.ac.uk (M. Snaith)
https://www.simonwells.org/ (S. Wells); https://www3.rgu.ac.uk/dmstaff/snaith-mark/ (M. Snaith)
0000-0003-4512-7868 (S. Wells); 0000-0001-9979-9374 (M. Snaith)

mixed-initiative interactions between humans and intelligent software agents [5].

Our research has focused upon an investigation of the various activities and responsibilities associated with the actors and systems that can engage in dialogue, identifying the strengths and weaknesses of each. To this end we have constructed a characterisation of dialogue systems that focuses upon the roles, responsibilities, and necessary abilities of the participating actors or systems that comprise those actors. We attempted to characterise dialogues in three contexts; where the actors within the dialogue are people, where the actors are software agents that incorporate dialogue games, and where the actors are software agents that incorporate LLMs. The aim was to delineate the kinds of roles, responsibilities, and capabilities that a dialogue system needs and to determine how the responsibility for fulfilling these factors is spread across agents within these three contexts. We then posed a series of "wh-Questions" (who, what , how, which, when, why, where). The aim of this approach was to provide a new analytical tools for considering what a dialogue systems needs to do, and, in the case of software agents, which capabilities are delegated to, or fulfilled by which sub-systems.

We then investigated examples of where LLMs are currently demonstrating utility in order to benchmark actual LLM performance against other approaches. Dialogue systems, comprising multiple components, achieve a variety of levels of capability in dialogue. Using humans as an exemplar of agents who are generally capable of choosing what to say and when to say it in a strategically useful way, we compare them firstly to software agents comprising various combinations of non-LLM modules for dialogue, sentence generation, and strategic reasoning, and subsequently to LLM behaviour. We then show how a dialogue game, utilising an existing dialogue game execution platform [4], together with an LLM, can work together to achieve more in aggregate using current technologies. Throughout we argue that LLMs, at least at present, do not currently subsume traditional dialogue game research, but have an ancillary role, due to their complimentary strengths, that can lead to great improvements in the ability of intelligent agents to eventually engage in principled, well-structured, well-regulated, constructive, and purposeful dialogue.

Finally, we address the question of why, if LLMs are increasingly able to subsume the functionality of other approaches, should research continue into other approaches, such as dialogue games. We argue that dialogue games have been studied for a long time as a way to understand dialogue dynamics and to yield models that capture and explain both normative and ideal expectations for how dialogues should progress. Even if LLMs are trained to engage in increasingly realistic dialogue, dialogue games will still have an important regulatory role to play. This regulatory role utilises the dialogue game variously as ideal or normative model, depending upon the circumstances, against which dialogue participants, including both humans and LLMs, can self-evaluate, testing their own generated responses against the kind of ideal response that a dialogue model would propose. In this way, we can still aspire towards higher quality, computer-supported, argumentative dialogue as well as rich and naturalistic, human-machine interaction.

Next steps will involve further study of the useful interactions between LLMs and dialogue models as well as automated benchmarking of the abilities of the resulting systems. One approach might build upon the idea of the Arguing Agents Competition [6, 7].

In summary, despite advances in ML based approaches to dialogue, traditional approaches to dialogue modelling have a more important role to play than ever before.

## References


[1] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language models are few-shot learners, in: H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, H. Lin (Eds.), Advances in Neural Information Processing Systems, volume 33, Curran Associates, Inc., 2020, pp. 1877–1901. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf.

[2] S. Wells, C. Reed, A domain specific language for describing diverse systems of dialogue, Journal of Applied Logic 10 (2012) 309–329.

[3] S. Wells, Formal Dialectical Games in Multiagent Argumentation, Ph.D. thesis, School of Computing, University of Dundee, 2007.

[4] F. Bex, J. Lawrence, C. Reed, Generalising argument dialogue with the dialogue game execution platform., in: COMMA, 2014, pp. 141–152.

[5] M. Snaith, J. Lawrence, C. Reed, Mixed initiative argument in public deliberation, Online Deliberation 2 (2010).

[6] S. Wells, P. Lozinski, M. N. Pham, Towards an arguing agents competition: Architectural considerations, in: Proceedings of the 8th International Workshop on Computational Models of Natural Argument (CMNA8), 2008.

[7] T. Yuan, J. Schulze, C. Reed, Towards an arguing agents competition: Building on argumento, in: Proceedings of the 8th International Workshop on Computational Models of Natural Argument (CMNA8), 2008.