Controlled Chain of Thought: Eliciting Role-Play Understanding in LLM Through Prompts

Deborah Carlander
Cygames Research
Cygames Inc.
Shibuya, Tokyo, Japan
0009-0009-3207-8623

Kiyoshiro Okada
Cygames Research
Cygames, Inc.
Shibuya, Tokyo, Japan
0000-0002-3952-7518

Henrik Engström
School of Informatics
University of Skövde
Skövde, Sweden
0000-0002-9972-4716

Shuichi Kurabayashi
Cygames Research
Cygames, Inc.
Shibuya, Tokyo, Japan
0000-0001-5967-7727

2024 IEEE Conference on Games (CoG) | 979-8-3503-5067-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/CoG60054.2024.10645667

***Abstract*—*Tabletop Role Playing Games* (TRPG) are games that require players to become the characters they play through engaging in role-play. The challenge of training AI lies in the requirement of it not only understanding the explicitly stated game rules, but also the implicit ones that come with role-playing. Previous studies endeavouring said challenge allude to aspects of role-play, but do not emphasise its role in their methods, indicating that the definition of role-play and how it is to be employed remains unclear. This short paper aims to investigate a proposed definition of role-play based on previous research and employ its use on *Large Language Model* LLM, eliciting an understanding thereof through a novel prompting method dubbed *Controlled Chain of Thought* (CCoT). CCoT allows for the LLM to highlight absence of information in inputs given to it by generating questions, which become the template for its chain of thoughts when answered. This paper presents an initial pilot testing of CCoT as well as opens up the discussion of how a definition of role-play can be beneficial for future studies.**

***Index Terms*—controlled chain of thought, large language models, alignments, tabletop role-playing game, play frames**

I. Introduction

Turing [1] first proposed the challenge of a computer playing and mastering the game of chess. With AI now being able to beat human players in both chess and Go, Ellis and Hendler [2] proposed to further his challenge by having AI play TRPGs. To highlight the contrast between the two, chess is a turn-based game played by two players with a limited set of mechanics and clear end states. TRPGs, on the other hand, are played in groups, with one player serving as *Game Master* (GM), containing rules that are partly negotiated between them, and where the end-goal is not defined. Where the former has a finite set of predictable outcomes, the latter has infinite unpredictable ones.

In his seminal work, Fine [3] defines the key aspect of TRPGs to be that of role-playing and that players move between different play frames during a session, where the role-playing takes place in that of the fantasy frame. Fine further identifies two types of players and their play-styles: the role-players, and self-players. The former is more comfortable with role-playing as their characters. Both types of players sometimes break character, but the self-players do so more often in favor of progressing. When a player breaks character during a game session it simultaneously breaks the frame of fantasy for all players.

Large Language Models (LLM) are AIs trained on a large corpus of text-based data to both comprehend and generate text [4], [5]. They are highly relevant for TRPGs due to their ability to understand natural language (non-machine-based language) since TRPGs are played through communication between players. Current studies looking to tackle the AI and TRPG challenge (e.g. [6]–[9]) touch upon aspects of the fantasy frame and attempt to distinguish both in and out of character actions, but also intentions of players, indicating that there exists a desire to distinguish what role-playing is for the AI. However, none seem to use alignments, which are moral characteristics of characters from Dungeons and Dragons (D&D), designed to help guide players to act in accordance with their character, something which a lot of players struggle with [3]. Alignments are closely tied to the frame of fantasy, indicating that they can be a good starting point when defining role-playing to an AI.

In this paper, a prompting technique referred to as *Controlled Chain of Thought* is presented as a way to teach GPT to detect alignments based on character scenarios. This pilot study is a first step to see if GPT can understand the D&D alignments and if this can be used to detect when a player is breaking character, generating a fantasy frame understanding.

II. Background

A. Tabletop Role-playing Games

TRPGs are games played in groups, where players role-play as characters they have created to engage in game scenarios controlled by a GM [3]. The games are mainly played orally with players acting out actions by speaking as their characters, using dice to determine chance-based events such as if an attack was successful or not. The rules of TRPGs are flexible, where the main mechanics are stated in the rule-book, but the story, the goal and what players are allowed to do are negotiated among them. Depending on the TRPG, the extent of the rules may vary, some being more restrictive than others.

Dungeons and Dragons (D&D) is a TRPG that takes place in a medieval fantasy setting and was first published 1974 [10]. When making a character, players choose a race and

class and roll dice to determine their stats. Apart from this, they also choose a background for their character and an *alignment*. Alignments consist of two variables, one which defines their attitude toward society and order (lawful, neutral, or chaotic) and the other their moral values (good, neutral, or evil). Together these form 9 possible alignments for players and serve as a guiding tool for how their character should act and what goals they may have.

B. The Frames of TRPGs

Goffman [11] describes how the dynamic of social interactions are affected by the context, *frame*, which enclose it. Fine [3] builds onto this, proposing that not only do frames affect the way one acts, but also one's awareness. He uses TRPGs and their social structure to exemplify this and identifies three basic frames, each encompassing certain awareness held by the player. The first frame is the *commonsense knowledge frame* and is grounded in the "primary framework" as defined by Goffman [11]. It encompasses all knowledge within our reality, as well as that of the game and fantasy. The second is the *game context frame*, encompassing the knowledge of the game context, how one plays and the conventions of the game. Lastly there is the *fantasy frame*, and this is knowledge that the player character only should have in the context of the fantasy world. The shift of awareness is so apparent in TRPGs, because the participants do not only engage in them as the role of a player, but also the role of their character: "Finally, the gaming world is keyed in that the players not only manipulate characters; they are characters. The character identity is separate from the player identity." [3, p.186]

Fine [3] argues that there is no need for a chess player to separate their identity from that of its pawns. They will not consider that their bishop might not wish to kill a pawn due to their Christian belief. TRPGs, on the other hand, is a game of several players with the goal of engrossment. This engrossment is reached in the fantasy frame, through becoming one's character. Fine [3] identifies two player types: self-players, and role-players. Self-players opt for playing characters that resembles themselves. They often struggle with role-playing and prioritize progression over role-play. Role-players, on the other hand, see enjoyment in becoming another character and prioritize the engrossment that comes within the fantasy frame. Note that self-players enjoy this engrossment too, but not if it intervenes with their goal of progression. The tricky nature of TRPGs is that they require players to shift between the game context frame, where dice are used and rules discussed, and the fantasy frame where they act as their characters.

C. Large Language Models and Prompting

LLMs are AI models that are trained on a large corpus of text data [4]. One commonly used LLM is GPT [12] trained on a large corpus of data giving it a general knowledge applicable for most text-based tasks. For more context specific tasks further fine-tuning can be needed, which often requires lots of training data and time. A way to test if a specific dataset can elicit the desired outputs without fine-tuning the LLM is through prompts. [13]. Some of the most cited examples of prompting techniques are zero-shot [5], few-shot [14] and chain-of-thought (CoT) prompting [15]. Zero-shot prompts are inputs that explain a desired output without using an example (or shot), whereas in few-shot the inputs contain one or more examples to train the model on task specific data. The main advantage of using few-shot prompting is that it reduces the need for task specific data in the pre-training dataset, since this is included as prompts instead [14]. CoT is a technique where the model is prompted to first break down a task into smaller sub-tasks [15]. CoT is useful for bigger tasks that require common-sense reasoning and can be both zero-shot (zero-CoT) and few-shot (few-CoT) [15]. Zero-CoT is less time consuming, since it does not require example data, but also less controllable than few-CoT. Kojima et al. [16] combat this dilemma through first using zero-CoT to generate outputs where the best ones are then used as data for few-CoT prompts. Park et al. [17] also reuses outputs to create prompts by prompting an LLM to generate questions about the input that it answer, where the answers help summarize information. They found it was an effective method for eliciting deeper understanding in the LLM, improving its ability to summarise, a task many models otherwise struggle with.

D. LLMs and TRPGs

There are a few types of studies looking at how LLMs can be integrated with TRPGs. In the most prevalent studies, an LLM is given a dataset of transcripts from TRPG sessions and is then tasked to predict which player is to act next and what they will do, based on the current context [6]–[9]. These studies will be referred to as *Player Action Prediction* (PAP) studies. The method used for PAP was first introduced by Louis and Sutton [18]. The subsequent studies have mainly extended on this by suggesting novel prompting methods and applied the technique to new data sets.

The datasets used consist of TRPG transcripts from play-by-post forums, where D&D is the most common TRPG used [6], [7], [9]. Callison-Burch et al. [7] labeled their transcripts based on if players were speaking in or out of character by training an AI on play-by-post forums with separate chats for in and out of character conversation. This is extended upon in subsequent studies [6], [8], [9]. Zhou et al. [9] and Liang Zhu and Yang [8] include intention of character actions in their datasets, where the former includes GM intentions and the latter defines skill-checks as intentions, to improve predictions. To evaluate PAP studies, the most common metric used is that of perplexity [6], [7], which looks at the range of prediction when predicting next character and action. The lower the perplexity score the better. Some studies [6]–[8] also include subjective evaluations, using experienced D&D players, looking at groundedness (how human-like the responses are) and if there are factual errors. To compare results, ablation is used where parts of information included in the datasets are removed [7]–[9]. All studies found that including

more context-based information, especially regarding character background, improved the models' predictions.

III. Problem

Previous PAP studies [2], [8], [9] contribute with interesting insights, touching upon aspects of frames, such as including in and out of character labels [6], [7] or that of trying to define the intentions of players [8], [9] to improve their results. These studies highlight that there is a need for understanding not only the rules of the game context, but that of the fantasy frame. However, how this is to be done remains unanswered.

Alignments guide the players in how their character should act, and is a common place of struggle for self-players [3]. In comparison to character traits that are physical, where the player performs the actions through rolling dice, personality traits such as alignments can only be performed through role-playing and speaking as one's character. Alignments could therefore be one of the keys to teaching an AI how to role-play and understand whether other players are role-playing too. Despite their close connection to role-playing and fantasy frame, alignments have yet to be mentioned in previous studies.

To explore how the fantasy frame can be understood by AI, this paper proposes the use of alignments to identify if a player is acting in- or out of character. For this, the following questions are posed:

- *How can an LLM be used to detect to what extent a player is acting according to its character's alignment when playing Dungeons and Dragons?*
- *How can prompting be used to achieve this?*

IV. Method

A. Dataset and Model

As OpenAI's API allows for testing prompts without their AI training on it and because of the extent of their models' training data, GPT-3.5 [12] was used for the prompting development.

The dataset used for the prompts was a combination of the D&D alignment descriptions from the 5th edition basic rule book [19] and in-game transcripts from *Critical Role* [20], an online series where actors play D&D live. The transcripts are written after each episode, making it based on live playing sessions, which is preferred over play-by-post transcripts, where play-sessions can happen over longer periods of time, lacking the natural flow of real-life playing [8]. The dataset includes characters of the alignments neutral good (NG), chaotic good (CG) and chaotic neutral (CN), with each alignment being represented by at least two different actors. Because GPT showed prior knowledge of Critical Role, all characters and settings was anonymised.

B. Controlled Chain of Thought

As decoder based LLMs predict the next word most likely to be used they are prone to hallucinate when information is missing. As D&D's alignment descriptions are brief and lack concrete examples [19], CCoT was developed as a prompting

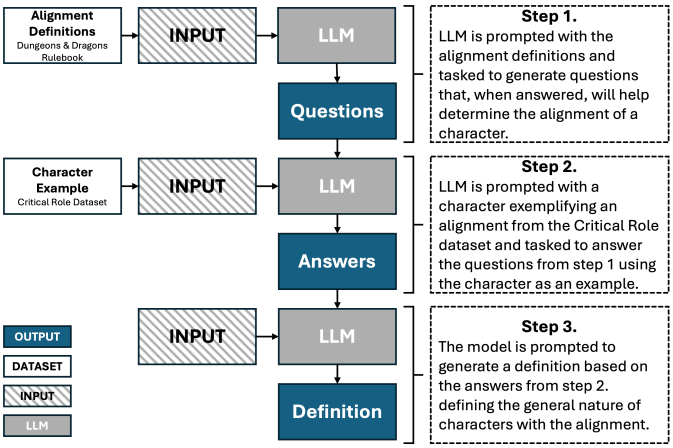

Fig. 1. Controlled Chain of Thought

method (Figure 1). Because of the lack of information in alignment descriptions, in step 1, GPT is prompted to generate questions regarding the alignments that, when answered, would give it enough information that it could determine the alignment of a character. It is a similar method to that of Park et al. [17] with the difference that, instead of generating questions that can already be answered with the information given, GPT is asked to generate questions that it cannot answer without further information, highlighting the areas that are *not* understood by the model. In step 2, GPT is prompted with a character description taken from the Critical Role dataset representing a specific alignment, and asked to answer the questions previously generated, citing examples from the description when doing so. This has a similar effect to CoT, where the model goes through the information step-by-step, but in a controlled manner without the user having to manually create an example of reasoning by relying on the questions instead. Lastly, in step 3 GPT is prompted to summarise the answers elicited into a general definition of the alignment in question. Because LLMs are designed to predict the next word to follow, they can sometimes focus on smaller details in text, making the output too specific. Since the desire is to have GPT understand a specific alignment representing several characters the summary generalises the information, generating a higher-level understanding of the alignment.

C. Pilot Testing

To see if CCoT generates a higher-level understanding of alignments in GPT an initial pilot testing was held with a baseline method using no prompts (zero-shot) and CCoT were tasked to detect characters of the alignment NG. The test consisted of character descriptions (explanation of a character's background) of four different characters and three scenarios (excerpts of in-game conversation-logs) containing 11 characters, amounting to 15 characters in total, 7 of which being of the alignment NG. This ensured inclusion of both static descriptions of the characters, as well as dynamic in-game dialogue. As metrics, the testing followed the method

of Liang, Yang and Zhu [8]. They include both precision and recall to generate the harmonic mean (HM) of their results.

V. Preliminary Results

Using the baseline prompt, 5 characters were identified as NG, 2 of which being correct, misidentifying 3 non-NG characters and missing the remaining 5 of the NG characters. Using the CCoT prompts, the LLM correctly identified all 7 NG characters, and did not misidentify any others (see Table I). However, in the last and longest scenario some uncertainty was identified when using CCoT, which is not visible in the data. GPT stated that it could be possible for a non-NG character to be NG, but that there was not enough information to conclude this. This implies that current metrics fall short of fully representing the results, since these types of uncertainties is not represented in the data. Possible suggestions to solve this is by including a qualitative analysis of the outputs generated.

TABLE I
Alignment Prediction Results

Test	**Precision** (p)	**Recall** (r)	**Harmonic Mean** ($\frac{2pr}{p+r}$)
Baseline	0.40	0.28	0.33
CCoT	1.0	1.0	1.0

GPT-3.5 is tasked to identify neutral-good characters (n=7) in the provided excerpts with a total of 15 characters.

VI. Discussion

It is the understanding of frames that is the key to role-play. Without distinguishing these frames, and noting their sensitive relationship, an AI might not understand role-play limiting its engagement in it. It is important to note that, even though this paper has highlighted the importance of frames, it is still unclear if incorporating alignments in training data successfully explain these to AI. The aim of this paper is therefore to open up further discussion on what role-play signifies by including and defining these frames through the use of alignments.

Having the LLM itself generating questions it later answers seems to elicit an understanding of what it does not understand, which can be useful for commonsense knowledge reasoning. CCoT shows promise in its ability to generate higher-level understanding of alignments in GPT, which could possibly be a first step in understanding when and how players role-play. It is important to note that, despite the positive results during the pilot-test, the scale of it is still too small to draw any larger conclusions of CCoT as a method for training AI on alignments. As the testing only compares CCoT to a baseline, the results does not aim to show an accurate comparison to other prompting methods, but to determine if there is enough improvement in performance to move forward with testing. As a result, a larger scale ablation study will be conducted to evaluate the effect each prompting step has on the over-all result. A qualitative analysis will also be included to analyse the outputs made by GPT. Despite the TRPG-focus in this paper, it is possible that CCoT is a prompting method that could prove useful in other areas as it generates a more controlled CoT output without requiring few-shot examples.

References

[1] A. Turing, "Digital computers applied to games," in *Faster than Thought*, B. Bowden, Ed. London: Sir Isaac Pitman & Sons Ltd, 1953, pp. 287–310.

[2] S. Ellis and J. Hendler, "Computers play chess, computers play go… humans play dungeons & dragons," *IEEE Intelligent Systems*, vol. 32, no. 4, pp. 31–34, 2017.

[3] G. A. Fine, *Shared fantasy: Role playing games as social worlds*. University of Chicago Press, 1983.

[4] S. Ozdemir, *Quick Start Guide to Large Language Models: Strategies and Best Practices for Using ChatGPT and Other LLMs*. Addison-Wesley Professional, 2023.

[5] V. Alto, *Modern Generative AI with ChatGPT and OpenAI Models: Leverage the capabilities of OpenAI's LLM for productivity and innovation with GPT3 and GPT4*. Packt Publishing, 2023.

[6] A. Zhu, K. Aggarwal, A. Feng, L. Martin, and C. Callison-Burch, "FIREBALL: A dataset of dungeons and dragons actual-play with structured game state information," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, pp. 4171–4193.

[7] C. Callison-Burch, G. S. Tomar, L. Martin, D. Ippolito, S. Bailis, and D. Reitter, "Dungeons and dragons as a dialog challenge for artificial intelligence," in *The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)*, Abu Dhabi, UAE, 2022.

[8] Y. Liang, L. Zhu, and Y. Yang, "Tachikuma: Understading complex interactions with multi-character and novel objects by large language models," 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2307.12573

[9] P. Zhou, A. Zhu, J. Hu, J. Pujara, X. Ren, C. Callison-Burch, Y. Choi, and P. Ammanabrolu, "I cast detect thoughts: Learning to converse and guide with intents and theory-of-mind in dungeons and dragons," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, pp. 11 136–11 155.

[10] G. Gygax and D. Anderson, *Dungeon's & Dragons*. TRS Inc, 1974.

[11] E. Goffman, *Frame Analysis*. Harvard University Press, 1974.

[12] OpenAI., "Models - openai api," 2018, accessed 2024-04-04. [Online]. Available: https://platform.openai.com/docs/models/overview

[13] K. K. Phokela, S. Sikand, K. Singi, K. Dey, V. S. Sharma, and V. Kaulgud, "Smart prompt advisor: Multi-objective prompt framework for consistency and best practices," in *2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2023, pp. 1846–1848.

[14] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal *et al.*, "Language models are few-shot learners," *34th Conference on Neural Information Processing Systems*, vol. abs/2005.14165, 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:218971783

[15] J. Wei, X. Wang, D. Schuurmans, M. Bosma, E. H. hsin Chi, F. Xia, Q. Le, and D. Zhou, "Chain of thought prompting elicits reasoning in large language models," *ArXiv*, vol. abs/2201.11903, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:246411621

[16] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large language models are zero-shot reasoners," *ArXiv*, vol. abs/2205.11916, 2022. [Online]. Available: https://api.semanticscholar.org/CorpusID:249017743

[17] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative agents: Interactive simulacra of human behavior," in *"Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)"*, 2023.

[18] A. Louis and C. Sutton, "Deep dungeons and dragons: Learning character-action interactions from role-playing game transcripts," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2018, pp. 708–713.

[19] M. Mearls and J. Crawford, *D&D Basic Rules*. Wizards of the Coast, 2018.

[20] Critical role. 2024. [Online]. Available: https://critrole.com/