Dungeons & Dragons as a Dialog Challenge for Artificial Intelligence

Google Research

Chris Callison-Burch, Gaurav Singh Tomar, Lara J. Martin, Daphne Ippolito, Suma Bailis, David Reitter

Annotations are silver not gold



Datasets: Actual Play Podcasts

Actual Play podcasts are a genre of podcasts where people record themselves playing RPGs. We collected a list of 500 of them from the web and queried Google's podcast index for more.

The Shield of Heroic Memories

I designed an item for The Adventure Zone, a comedy podcast about three brothers playing D&D with their dad. The McEiroy brothers were incredibly enthusiastic about my submission. I want all of my paper reviews to say what they said, "That's already radical and then my boy Chris Callison-Burch kicked it up a notch. It's brilliant."



Play By Post data: Rule-based annotations

Rule-based heuristics to extract character properties

- Name: perform NER on all the player's turn in a campaign. The character's name is likely to be the most frequently mentioned.
- Class: count how many times each <u>D&D class</u> is mentioned by each player. Most frequently mentioned class is their character's class.
- Race: On a player's first turn, check whether any of the <u>D&D fantasy races</u> are mentioned. Assign it to character. If not, apply other rules.
- Pronouns: Count gendered pronoun mentions by a player. Assign their character's pronoun to be the most frequent pronouns used by the player.
- Inventory: Regex that matches items occurring after character's personal pronouns (e.g. her sword).
- Spells known: Regex that matches cast followed by a spell name

Dungeon master is assumed to be player who has first post in game. DM is scrubbed of other character properties, since she plays multiple NPCs and monsters.

Types of Language

In-character description

"Cahl" Korlann's cry of pain is mute to his own ears as the bandit's blade bites into his side. With a grow of anger, he lashes out with a quick swipe of his arm, spines flared out to strike at the bandit, before twisting himself around and bring the same arm back with a vicious elbow strike at the bandit's face.

Out-of-character Flurry of Blows against Bandit 2. Attack 1:

Hit: (1d20+2)[17], Crit on 22 Damage: (1d6+4)[8], x2 on Crit

Attack 2: Hit: (1d20+2)[17], Crit on 22 Damage: (1d6+4)[6], x2 on Crit

Generating in-character dialogue resulted in substantially more interesting responses in our human evaluation.

Training Data Example

Turn 1	
Text	You attack. You launch some fire onto the goblin closest to the wagon. And with that, he looks like he is on death's door. And the other goblin that you can see, the one that's not in the brush somewhere, just sort of stops in his tracks. What do you do next?
Player ID	0
Character	Dungeon Master
Race	N/A
Class	Dungeon Master
Gender	N/A
Inventory	N/A
n combat?	Yes
character?	Yes
Action	Attack

se	s in our hu	ıman evalu	atio
	Turn 2		
	Text	l grab my axe and bring it down on the wounded goblin.	
	Player ID	1	
	Character	Magnus	
	Race	Human	

Fighter

Male

Δve

Yes

Yes

Attack

Class

Gender

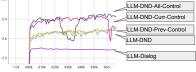
Inventory

In combat?

In character?

Action





Models

LLM-Dialog – our baseline model is Google's LaMDA dialogue model. No D&D data. 7 most recent turns.

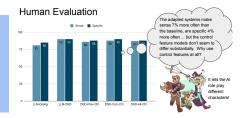
play-by-post data without using any off

the annotations as controls.

LLM-DND-All-Control – Fine tuned on all dialogue turns plus state variables for all turns including current.

LLM-DND-Curr-Control – Fine tuned on all dialogue turns plus state variables for only the current turn.

LLM-DND-Prev-Control – Fine tuned on all dialogue turns plus state variables for all turns except the current turn.



Example Output

LLM-Dialog	moon_monk you can just join our discord conversation. it's free and easier to deal with
LLM-DND	I cast Ice Knife (I'm assuming this is not one of the ones I can't do?)
DND-Current-Control Responding as Magnus the human fighter with an axe (in combat)	t turn from the goblin and go to help Merle and Taako with getting this wagon moving again. Magnus can't see them anymore with such a dense forest. They're likely in a battle right now I have to find them as fast as I can!
DND-Current-Control Responding as Taako the elf wizard (in combat)	If the goblins still don't look like they're getting closer, Taako will yell, "Alright you little monsters! Let's see what ya gott" And he'll cast Chill Touch on one of the goblins that isn't dead/on its doorstep. Attack: 22 Damage: 6
DND-Current-Control Responding as Mearle the dwarf cleric (in combat)	Merle steps away from the wagon. He draws his sword and takes aim at the hidden goblin! Attack: 27 Damage: 8 He puts his free hand on the pommel of his sword and shouts his battle cry, "For Clanggeddin!" His eyes glow brightly red as he strikes!

Datasets: Play-By-Post D&D



Play by Post Stats Number of campaigns

Total dice rolls

Average players per campaign Average turn per campaign Average words per campaign Total turns Total words

Average dice rolls per campaign 5

Actual Play Podcast Stats

96	Number of Shows	555
	Number of Episodes	32,602
10	Average number of speakers	2.23
4,941	Average turns per episode	597
15,106 8,187,426	Average words per episode	11.374
94	Total turns	19.478,136
32.270	Total words	370,824,073
02,270	iotal Wordo	010,021,010