

논문 리뷰_Dungeons and Dragons as a Dialog Challenge for Artificial Intelligence

TRPG 플레이 로그를 활용한 참가자용 AI 학습과 평가

Deep Shark Lab 학부생 **이예정**

2026.08.14

*** Contents**

1. 서론 및 연구 목표

2. 데이터셋 구축

3. 모델 파인튜닝

4. 평가 및 분석

5. 결론 및 한계점

* 논문 정보

제목 및 연구진

Dungeons and Dragons as a Dialog Challenge for Artificial Intelligence

Chris Callison-Burch*

University of Pennsylvania
ccb@upenn.edu

Gaurav Singh Tomar*

Google Research
gtomar@google.com

Lara J. Martin

University of Pennsylvania

Daphne Ippolito

University of Pennsylvania
Google Research

Suma Bailis

Google Research

David Reitter

Google Research

저널 정보

- EMNLP 2022
 - Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing

서론 및 연구 목표

* 연구 배경

AI의 도전과제로서의 게임

- 오랜 시간 인공지능 분야에서는 게임을 도전 과제로 활용했음
- 게임 플레이 시스템을 구축하여 인간과 비슷한 수준으로 끌어올리는 것이 주된 목표
- Ellis와 Hendler는 체스, 바둑 등의 **제로섬 게임(zero-sum game)**을 넘어선 **개방형 창작 게임**을 AI의 도전 과제로 제안함
- NLP(Natural Language Processing)과 NLG(Natural Language Generation)의 발전으로 AI와 인간이 협력적 이야기 생성까지 이루어지고 있음을 확인
- 이를 던전 앤 드래곤(D&D)에서 활용할 수 있고, 다음 도전 과제로 지목되어 옴

Martin 연구진이 제시한 구체적 도전과제 예시

- 게임 세계의 상태가 던전 마스터(DM)와 다른 플레이어들에게 분산되어 있다.
- 플레이어들이 특정 행동을 수행하면서 얻는 내재적 보상(intrinsic rewards)을 처리해야 한다. 이러한 행동은 반드시 게임에서 점수와 같은 명시적인 보상을 제공하지 않을 수도 있다.

제로섬 게임-영합(零和) 게임이라고도 한다.
누군가가 이익을 보면 누군가는 손해를 보는 구조.
이익과 손해를 더한 합이 0이라고 하여 제로섬 게임이라 불림.

개방형 창작 게임-open-ended creative games
플레이어들에게 도구와 공간을 제공하여 스스로 룰을 만들고,
게임 목표를 만들며 놀 수 있는 게임.
샌드박스 게임의 상위 개념이다.

* 연구 배경

해당 논문에서 고려한 D&D 데이터셋의 특징?

- 자연스럽게 발생한 대화(naturally occurring conversational dialogue)이며, 게임 플레이를 위한 과업 지향적 대화(task-oriented dialogue)에서 잡담과 같은 비과업 지향적 대화(non-task-oriented dialogue) 이르기까지 폭넓은 대화를 포함한다
- 대화 기록에 강하게 의존(strongly history dependent)한다
- 게임에 여러 플레이어가 참여하기 때문에 대화 참여자가 많다
- 게임 상태의 변화를 나타내는 사건 묘사를 포함하여 내러티브 요소(narrative elements)를 전달한다

* 연구 목표

연구 목표는?

- 해당 연구에서는 LLM을 학습시켜 아래의 두 기능을 할 수 있게 만들었음
- 학습 데이터 구성과 LLM 학습을 아래 두 기능에 맞추어 실행함

1. 다음 발화 예측

- 언어 모델이 다음 발화를 예측할 수 있도록 학습한다.
- LLM이 생성하는 발화는 아래 두 가지 조건을 만족해야 한다.
 - **흥미로워야 한다.**
 - D&D 게임에서 이어질법한, **그럴듯한 발언**이어야 한다.

2. 게임 상태 추적

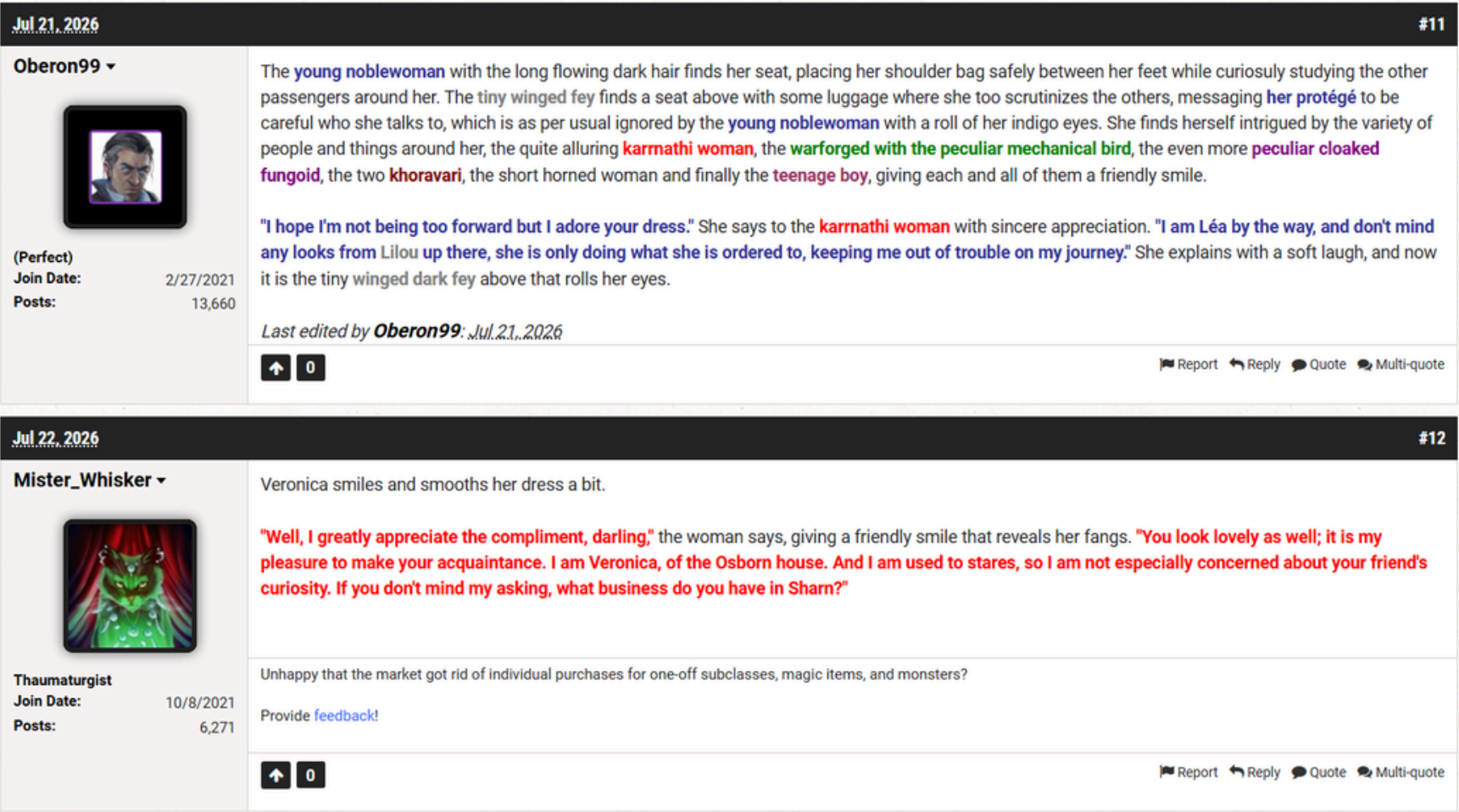
- 게임 전체에서 정적인 상태 (캐릭터 이름, 클래스, 종족, 인벤토리 등)
- 주기적으로 변화하는 상태 (현재 전투 중인가? 아닌가? 등)
- 매 턴마다 변화할 수 있는 상태 (현재 무엇을 하고 있는가?)

데이터셋 구축

* 데이터셋 확보 방법

사용 데이터

- 대화 중심 데이터를 수집하기 위해, 웹 사이트에서 진행한 D&D 플레이 데이터를 수집.
- 웹 포럼에서 Play-By-Post 데이터를 수집함.



항목	규모
총 캠페인 수	896
캠페인 당 평균 플레이어 수	8 명
캠페인 당 평균 턴 수	910 턴
캠페인 당 평균 단어수	64,941 단어
전체 턴 수	815,106 턴
전체 단어 수	58,187,526 단어
캠페인 당 평균 주사위 굴림	594 회
전체 주사위 굴림	532,270 회

▲ 실제 연구에서 사용한 dndbeyond 웹 사이트에서 play-in-post로 진행한 게임의 모습
<https://www.dndbeyond.com/forums/dungeons-dragons-discussion/play-by-post/242233-eberron-heart-of-war>

＊ 게임 상태 정보 추출을 위한 규칙 기반 휴리스틱

플레이어/마스터의 구분

- 하나의 게임에서 제일 첫번째 게시물을 작성한 플레이어를 마스터로 구분함
 - 해당 포럼 특성상, 제일 첫번째 게시물에는 세계관 설정, 공지 등의 안내 사항을 작성해야 하기 때문
- DM이 담당하는 캐릭터는 NPC와 적대적 몬스터를 담당하기 때문에 위의 캐릭터 속성을 제거함

캐릭터 속성_이름(Name)

- 한 명의 플레이어가 작성한 모든 턴에 개체명 인식(NER)을 수행
- 가장 많이 언급한 이름을 그 플레이어가 사용한 캐릭터 이름으로 지정

캐릭터 속성_클래스(Class)

- 각 플레이어가 각각의 D&D 클래스를 몇 번 언급하는지 계산
- 가장 많이 언급한 클래스를 해당 플레이어 캐릭터의 클래스로 지정
- 클래스 명은 해당 사이트에 정리되어 있는 항목을 사용 → <https://www.dndbeyond.com/classes>

＊ 게임 상태 정보 추출을 위한 규칙 기반 휴리스틱

캐릭터 속성_종족(Race)

- 플레이어의 첫번째 턴이 자기소개를 위한 턴이라고 가정 후, 종족이 언급되는지 확인
- 종족명 발견 시 해당 종족을 캐릭터에게 할당
- 종족 명은 해당 사이트에 정리되어 있는 항목을 사용 → <https://www.dndbeyond.com/races>

캐릭터 속성_대명사(Pronouns)

- 플레이어가 언급한 대명사의 횟수 계산
- 가장 빈번하게 사용한 대명사를 해당 플레이어 캐릭터의 대명사로 지정

캐릭터 속성_인벤토리(Inventory)

- 정규 표현식을 통해 캐릭터의 인칭대명사 뒤에 등장하는 아이템을 탐지함
- 논문에서 제시한 예: her sword
- 문장에서 캐릭터의 소유물로 표현된 항목을 이용하여 인벤토리 항목 추정

캐릭터 속성_알고 있는 주문(Spell Known)

- 정규 표현식을 통해 cast 뒤에 주문 이름이 등장하는 패턴 탐지
- 문장에서 캐릭터가 사용한 주문을 추출

* 게임 상태 정보 추출을 위한 규칙 기반 휴리스틱

전투 속성

- 전투 상태의 돌입 여부는 아래의 기준으로 판단함
 - D&D에서 전투 순서를 판단하기 위한 우선권 굴림(initiative)이 발생하면 전투가 시작된 것으로 탐지
 - 기습 공격으로 인해 공격 굴림이 먼저 진행되었을 시 전투가 시작된 것으로 탐지
 - 공격 굴림이 계속되는 동안 전투가 지속되고 있는 것으로 판단
 - 일정한 수의 턴 동안 주사위 굴림이 발생하지 않으면 전투 종료로 판단
 - 하나의 전투 구간에서 언급된 몬스터의 목록을 추출하여 몬스터가 몇 마리 존재하는지 추정

행동 속성

- 주사위 굴림을 바탕으로 연결된 행동을 탐지
 - D20 주사위 → 행동의 성공/실패 판정
 - 다른 주사위 → 공격 판정이 성공했을 경우의 피해량 계산
- 정규 표현식을 사용하여 주사위 굴림과 가장 가까이에 존재하는 패턴 탐색
 - Acrobatics/Animal Handling 등의 D&D 능력
- 피해량 굴림 등의 경우 아래의 단어와 매칭한다.
 - 피해량 굴림: damage, dmg 등
 - 회복량 굴림: cure, heal, healing 등

✳ In-Character Text vs Out-Of-Character Text

In-Character Text (IC)

- 캐릭터의 발언. 플레이어가 캐릭터에 이입하여 Role-Playing을 하고 있는 부분
- 해당 포럼에서는 두 개의 게시판 중 IC용 게시판이 있음
- 예시 문장: Kuros pulls the feathered shaft of the arrow back to his cheek winning easily against the resistance of the bowstring. He pulls a lungful of air to keep himself steady, takes aim at the Bandit with the deer, and lets fly.

Out-Of-Character (OOC)

- 플레이어로서의 발언. 다음 행동을 위한 상의나 주사위, 액션 등의 규칙/매커니즘을 설명함
- 해당 포럼에서는 두 개의 게시판 중 OOC용 게시판이 있음
- 예시 문장: urprise round so only 1 standard or move action. Shoot the bow: (1d20+6)[20] vs Flat Footed AC at Bandit 1. Damage: (1d8+2)[10]

모델 파인튜닝

* 모델 학습 과정

LaMDA 모델 활용

- GPT-3 모델 같은 Transformer 기반 사전 학습 언어 모델
- 해당 모델을 파인 튜닝 하여 이하 두 개의 역할을 수행함
- LaMDA는 턴 기반 대화 데이터를 이용하여 학습함.
- 해당 연구에서는 가장 최근의 7개 대화 턴을 입력으로 사용하고, 8번째 턴을 예측하도록 함.

수행 과제 1. 게임 상태 추적

- 게임 상태 추적은 **언어 이해 과제**로 볼 수 있음.
- 사용 함수의 형태는 아래와 같음.
 - $f(\text{current utterance, previous state, history}) \rightarrow \text{new state}$
 - 현재 발화 + 이전 상태 + 대화 기록 → 새로운 상태

수행 과제 2. 게임 상태 추적

- 게임 상태 추적은 **언어 생성 과제**로 볼 수 있음.
- 사용 함수의 형태는 아래와 같음.
 - $f(\text{current state, history}) \rightarrow \text{next utterance}$
 - 현재 상태 + 대화 기록 → 다음 발화

* D&D 파인튜닝 데이터

데이터 예시

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
Pronouns	N/A
Inventory	N/A
In combat?	Yes
In character?	Yes
Action	Attack

TURN 2:

Text	I grab my axe and bring it down on the wounded goblin.
Player ID	1
Character	Magnus
Race	Human
Class	Fighter
Pronouns	he/him
Inventory	Axe
In combat?	Yes
In character?	Yes
Action	Attack

* D&D 파인튜닝 데이터

각 데이터 요소 control

- 데이터의 각 요소를 control하며 어떤 데이터로 학습한 모델이 과제를 제일 잘 수행하는지 확인함.

모델 이름	D&D 데이터	이전 턴의 상태	현재 턴의 상태
LLM-Dialog	✗	✗	✗
LLM-DND	✓	✗	✗
LLM-DND-PREV-CTRL	✓	✓	✗
LLM-DND-CURRENT-CTRL	✓	✗	✓
LLM-DND-ALL-CTRL	✓	✓	✓

* 학습 중 개발 세트 Perplexity

모델 학습 후 최종 perplexity와 token accuracy

Model	Perplexity	Token Accuracy
LLM-Dialog	2.65	44.61
LLM-DND	2.50	46.92
LLM-DND-PREV-CTRL	2.51	46.84
LLM-DND-CURR-CTRL	2.34	49.67
LLM-DND-ALL-CTRL	2.37	49.02

평가 및 분석

* 수동 평가 방법

평가 항목

- 모델의 최종 출력을 위한 데이터를 읽음. 데이터에 포함된 정보는 이하와 같음.
 - 최근 7개 대화 Turn
 - 플레이어 목록
 - 각 플레이어가 플레이한 캐릭터의 이름과 클래스
- 여러 모델이 생성한 출력물과 실제 사람의 답변 Turn을 읽고 다음의 기준을 평가함.
 - 응답이 말이 되는가? (Y/N)
 - 응답이 구체적으로 이루어져 있는가? (Y/N)
 - 응답이 얼마나 흥미로운가? (1~10 척도)
- 각 모델의 답변과 실제 답변(Gold response)를 제공한다.

Does this response make sense?

- ☐ Makes sense
- ☐ Does not make sense

Is this response specific to the given context?

- ☐ Specific
- ☐ Not specific

How interesting is this response on a scale of 10?

* 수동 평가 방법

실제 평가지

- Context
 - 캐릭터 이름(직업)제공
 - 실제 발화(게시글 내용) 제공
- Player()
 - 각 모델/실제 플레이어의 발화 4건 제공
 - 해당 설문지에서는 어떻게 모델의 답변이고, 어떻게 실제 답변인지 모름
- Ratings
 - 각 답변마다 질문 3개에 답하게 함

Context	Player 10 - Del	Ratings
Player 10 - Del (Barbarian) Player 0 - Dungeon Master Player 12 - Aramis (Bard) Player 8 - Xavier (Bard) ----- Player 10 - Del (Barbarian): Del gives a final kick to the dead orc that he just killed then proceeds to hunt through the main room then the store room, looking for his axe and other gear Player 0 - Dungeon Master : Del - make an investigation check. Player 10 - Del (Barbarian): 3 Player 0 - Dungeon Master : The barbarian begins tossing things around, overturning tables and tearing crates and sacks open, though his still simmering rage at being locked up prevents him from undertaking a thorough search. He turns up a shortbow, a light crossbow, a few javelins, 2 quivers of arrows, 2 cases of crossbow bolts, a crate full of mismatched leather armour and a few sacks full of preveserved fruit and vegetables. There is no sign of the equipment that was taken from him last night. Player 12 - Aramis (Bard) : ""anyone seen any keys? or my stuff?"" aramis calls out to everyone Player 8 - Xavier (Bard) : *quickly picking up the dead orcs cleaver so a better weapon Xavier turns to the rest of the party* ""fighting is not over, we are not sure whats to be found up there, raid the store room, we must be ready for more combat"" *i turn to del* ""would you stand with me to defend the rest while they re-arm?"" *i offer the crowbar to del* *then turn back to the party* ""and if you find a shield and sword in there, i require a stronger defence and fist""	Response 1 Del quickly searches the store room, grabbing whatever weapons he find useful. He then heads back to the others "I found another crossbow and arrows, and a dagger. I have a short bow, but I'll leave it here if someone else wants it. I also found one of my javelins that I was missing." Response 2 del takes the crowbar and tells xavier to find hes ****en axe and ill keep good care of him Response 3 Anyone see my axe, and leather? Response 4 Noticing Del is searching the same room. Del calls out to him "X, I just killed an orc you could use its stuff if you like."	Does this response make sense? ☐ Makes sense ☐ Does not make sense Is this response specific to the given context? ☐ Specific ☐ Not specific How interesting is this response on a scale of 10? _____ Does this response make sense? ☐ Makes sense ☐ Does not make sense Is this response specific to the given context? ☐ Specific ☐ Not specific How interesting is this response on a scale of 10? _____ Does this response make sense? ☐ Makes sense ☐ Does not make sense Is this response specific to the given context? ☐ Specific ☐ Not specific How interesting is this response on a scale of 10? _____ Does this response make sense? ☐ Makes sense ☐ Does not make sense Is this response specific to the given context? ☐ Specific ☐ Not specific How interesting is this response on a scale of 10? _____

Figure 3: The user interface that our raters used to evaluate the quality of our model's next utterance prediction.

* 수동 평가 방법

인간 평가_Raters

- 판타지 장르에 관심이 있고, D&D 경험이 있는 평가자를 모집
- 총 6명 중 D&D 플레이 경험이 있는 사람은 5명
DM(Dungeon Master: D&D의 게임 마스터)경험이 있는 사람은 3명

인간 평가_Raters Agreement

- 평가자는 총 500개의 시스템 출력을 평가하고, 하나의 출력 당 3명의 평가자가 평가했다.
- 모델이 출력한 평가와 실제 사람이 작성한 문장을 비교 평가.
 - 응답 문장 구성과 구체성에 관한 평가가 일치한 정도는 0.8
 - 우연의 일치를 보정한 Randolph Kappa 값은 **0.6**
 - 1~10의 수치로 평가한 흥미성에 관한 평가는 kendall's tau를 적용.
두 문장에 대한 점수의 상관 계수는 **0.46**이다.

Randolph Kappa-’자유 한계 카파’라고도 함
여러 사람이 같은 데이터를 분류했을 때, “이 사람들이 우연히 같은 답을 고른 것이 아니라 실제로 얼마나 일관되게 판단했는가?”를 측정하는 통계량.
실제 관찰된 일치율과 우연히 기대되는 일치율로 계산한다.

kendall's tau-켄달의 타우 상관계수
두 변수를 비교하여 순위/점수 일관성을 평가한다.
-1~1 사이의 값을 가지며, 1에 가까울수록 상관 관계가 높다는 것을 의미한다.
해당 논문에서는 “AI문장에 높은 점수를 준 평가자가 사람이 쓴 문장에도 높은 평가를 주는가?” 를 시사한다.

* 수동 평가 결과

모델 간 비교

Model	Sense	Specific	Interest
LLM-Dialog	0.81	0.85	3.57
LLM-DND	0.9	0.9	3.91
LLM-DND-PREV-CTRL	0.86	0.88	3.96
LLM-DND-CURR-CTRL	0.88	0.9	3.96
LLM-DND-ALL-CTRL	0.87	0.88	3.92
GOLD	0.92	0.92	4.17

* 제어 특성_Control Features

제어 특성에 따른 결과

- 제어 특성을 제공한 이유는?
 - 특정 캐릭터를 역할극 할 수 있게 제어 가능하기 때문.

Model	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?)
LLM-DND-CURRENT-CTRL Responding as Magnus (human fighter, inventory: axe) In combat	I 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!
LLM-DND-CURRENT-CTRL Responding as Taako (elf wizard, inventory: spellbook) 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 got!" And he'll cast Chill Touch on one of the goblins that isn't dead/on its doorstep. Attack: 22 Damage: 6
LLM-DND-CURRENT-CTRL Responding as Mearle (dwarf cleric, inventory: holy symbol) 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 Clangeddin!" His eyes glow brightly red as he strikes!

* OOC vs IC

Out-of-Character 발언과 In-Character의 차이

- 하나의 문장을 OOC 발언으로 생성했는가, IC 발언으로 생성했는가의 차이 분석.
- OOC를 대신해서 IC로 생성했을 때의 평균 점수 차이

Model	Sense	Specific	Interest
LLM-Dialog	-0.01	-0.01	+0.06
LLM-DND	-0.02	+0.03	+0.4
LLM-DND-PREV-CTRL	+0.02	+0.02	+0.6
LLM-DND-CURR-CTRL	+0.06	+0.06	+0.93
LLM-DND-ALL-CTRL	+0.07	+0.06	+0.81
GOLD	+0.07	+0.05	+1.02

* 게임 상태 추적 모델

LLM-DND-GST

- 게임 상태 추적을 위한 파인튜닝을 거친 모델
- Play-by-Post D&D 게임플레이 데이터셋으로 파인튜닝함. 포함된 정보는 아래와 같음
 - 이전의 모든 대화 턴
 - 이전 턴들의 상태 변수
 - 현재 턴의 텍스트
- Previous dialogue + Previous states + Current utterance → Current state 구조로 현재 상태를 출력함
- LLM-DND-GST 모델로 출력된 결과와, 단순히 제일 많이 나온 단어 기반으로 추측한 baseline 모델을 기반으로 한 출력을 비교하였음

＊ 게임 상태 추적 모델

LLM-DND-GST

- 추측 값을 포함한 정답률은 LLM-DND-GST가 높았지만, 전체 상태를 추적하는 join accuracy는 58%밖에 되지 않았음
- 따라서, D&D의 전체 게임 상태를 정확하게 추적하기 위해서는 파인튜닝된 LLM만으로는 부족하며, 추가적인 메커니즘이 필요할 가능성이 있고 결론지음
- 즉, 개별 상태 변수는 비교적 잘 맞추더라도, 한 턴에 필요한 모든 상태 정보를 동시에 정확하게 맞추는 문제는 훨씬 어렵다는 결과

State variable	Majority	LLM-DND-GST
All	.73	.82
Combat	.89	.82
Character Class	.58	.76
Character Name	.58	.78
Character Race	.75	.79
Character Pronouns	.58	.89
Character Actions	.80	.85

결론 및 한계점

* 결론

D&D 데이터와 모델 학습

- 일반 대화 시스템보다 품질이 높았음

OOC발언과 IC 발언의 차이

- OOC로 발언한 문장보다 IC로 발언한 문장의 흥미도가 더 높았음

제어 요소와 답변

- 제어 요소들을 선 제공한 경우, 캐릭터에 더 적합한 문장을 생성했음

최종 결론

- 연구진의 모델은 아래의 두 조건을 만족하는 텍스트 생성이 가능했음
 - 문맥 및 게임 상태에 적합한 대답
 - 인상적인 In-Character 문장
- DM을 대체하는 모델은 될 수 없지만, 문장 생성에 있어 보조 역할은 가능할 것으로 보임

＊ 연구의 한계점

평가 방식의 한계

- 실제로 플레이를 하며 평가한 것이 아닌, 이미 만들어진 텍스트를 기반으로 평가했음
- 따라서 해당 AI모델이 실제 사람과 상호작용이 얼마나 잘 되는지 판단할 수 없음
 - 사용자의 지정되지 않은 문장에 얼마나 잘 대응할 수 있는가?
 - 사용자가 바라는 바를 이해하고 관련하여 대답할 수 있는가?

게임 상태 추적 방식의 한계

- 캐릭터 이름, 클래스, 현재 상황 등이 휴리스틱 함수로만 이루어져 있음
- 따라서 잠재적 오류가 있고, 불완전함
- 해당 문제를 해결하기 위해 디스코드에서 D&D플레이를 가능하게 하는 Avrea 개발자와 협력 시작했다고 밝힘
 - 실제 게임 상태 추적에 관한 데이터를 얻을 수 있을 것이라고 판단

Q&A

감사합니다.

논문 정보

Chris Callison-Burch, Gaurav Singh Tomar, Lara J. Martin, Daphne Ippolito, Suma Bailis, and David Reitter.
2022. Dungeons and Dragons as a Dialog Challenge for Artificial Intelligence.
In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing,
pages 9379–9393, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.