하계 세미나 (9/11, Thu)

Persona Update in Personalized Chat

구선민

Natural Language Processing & Artificial Intelligence

고려대학교 KOREA UNIVERSITY

Hello Again! LLM-powered Personalized Agent for Long-term Dialogue

Hao Li[1]*, Chenghao Yang[2]*, An Zhang[1]†, Yang Deng[3], Xiang Wang[2], Tat-Seng Chua[1]
[1]National University of Singapore
[2]University of Science and Technology of China
[3]Singapore Management University
18th.leolee@gmail.com, yangchenghao@mail.ustc.edu.cn
anzhang@u.nus.edu, ydeng@smu.edu.sg,
xiangwang123@gmail.com, dcscts@nus.edu.sg

Motivation

*** long-term personalized dialogue 연구의 필요성**

- 실제 시나리오에서 챗봇은 사용자와 장기적인 동반자 역할을 하면서 친숙하게 다가가는 능력 필요함

- 대부분의 기존 연구들은 2-15 turns 정도로 구성된 single-session 대상으로 함

- long-term personalized dialogue 위한 능력 2가지
  . 방대한 대화 기록을 이해하고 기억하는 능력
  → (장기적인 Event Memory 유지)
  . 사용자 및 챗봇 자신의 개인화된 특성을 충실히 반영하고 일관성 있게 갱신하는 능력 필요
  → (Persona Consistency 유지)

- 기존 연구들은 long-term personalized dialogue 위한 능력 개별적으로 다뤄서 문제
  + 대부분 특정 모델 구조에만 적용 가능

Motivation

* Optimal long-term dialogue framework

- model-agnostic, 다양한 실제 도메인에 배포 가능, 이벤트 메모리 및 페르소나에서 자율적인 데이터 통합 능력 필요한데 챌린지함

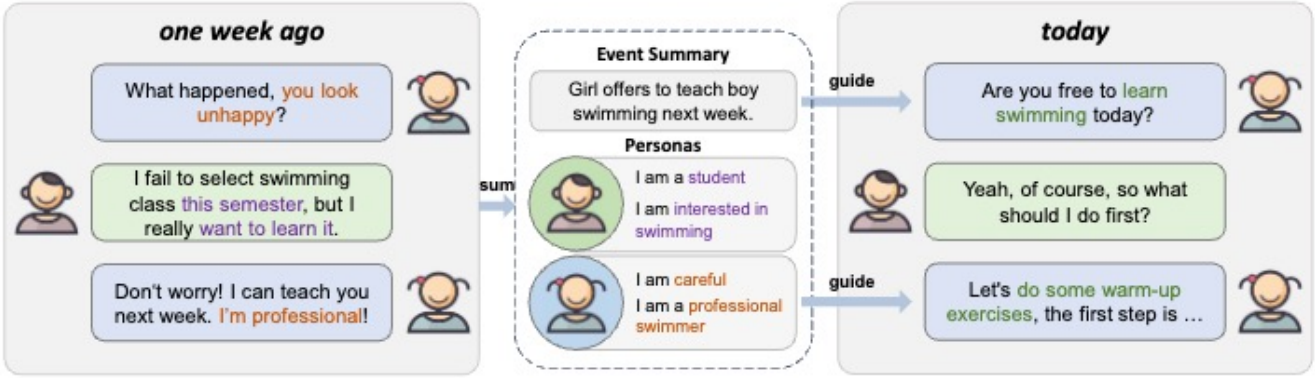


Figure 1: The illustration of how event memory and personas guide long-term dialogue. The event summary and personas are extracted from a conversation that occurred one week ago. In today's interaction, the event memory prompts the girl to inquire about the swimming lesson they scheduled last week. The personas, indicating that she is careful and professional in swimming, guide her to offer detailed and professional advice.

Overview

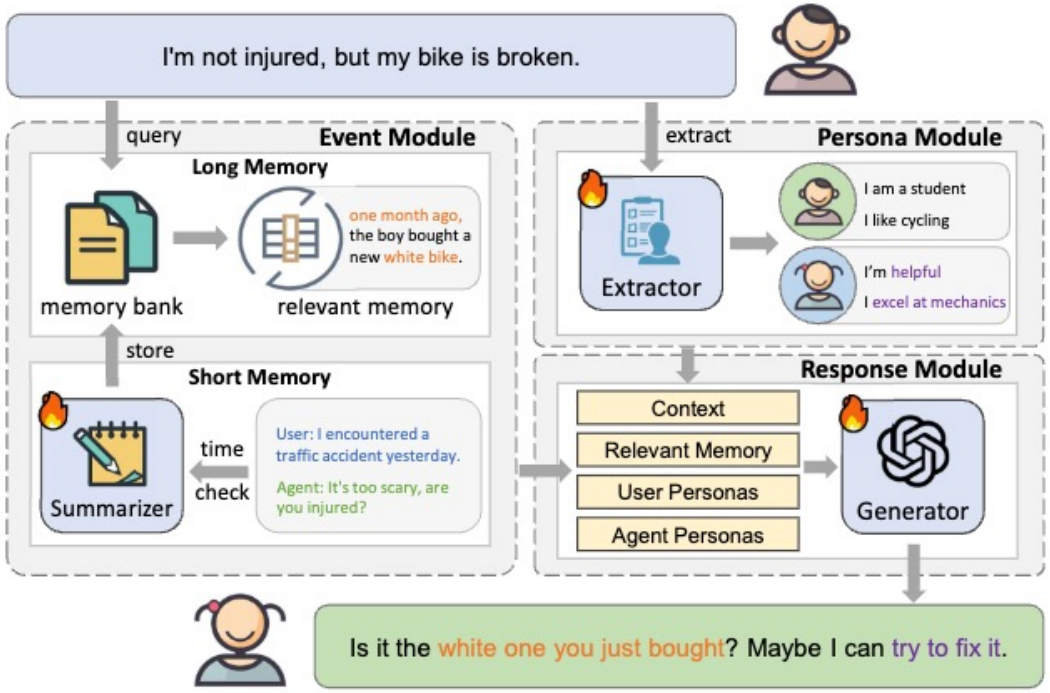


Figure 2: The Framework of LD-Agent. The event module stores historical memories from past sessions in long-term memory and current context in short-term memory. The persona module dynamically extracts and updates personas for both users and agents from ongoing utterances, storing them in a persona bank for each character. The response module then synthesizes this data to generate informed and appropriate responses.

Task Definition

* long-term multi-session dialogue task

- 목적: 현재 세션의 대화 맥락 (C)와 과거 세션 (H)에서 추출된 정보(selected information) 를 활용하여 적절한 응답(r) 을 생성하는 것

- 현재 대화 세션(C):
  . 현재 진행 중인 대화를 의미하며, 일련의 발화들로 구성됨

- 과거 대화 세션(H):
  . 현재 세션 이전에 발생했던 N개의 모든 historical 대화 세션을 포함

- 현재 대화 맥락 + long-term 과거 대화 기록에서 얻은 단서들을 통합하여 문맥에 맞는 응답을 생성해야 함

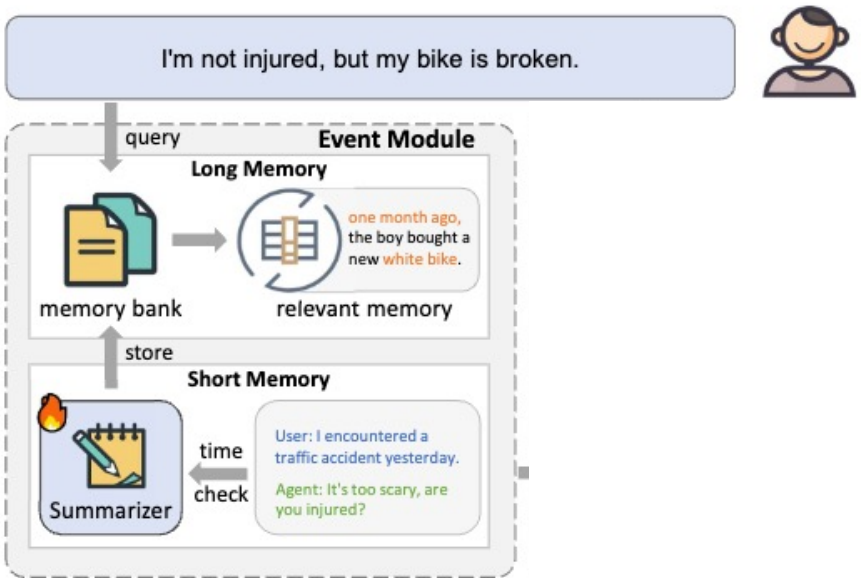
Event Perception

*** Long-term Memory.**

- Memory Storage
  . 이벤트 발생 시점 + 요약이 1개의 slot 으로 저장

- Event Summary
  . Summary dataset 으로 instruction tuning

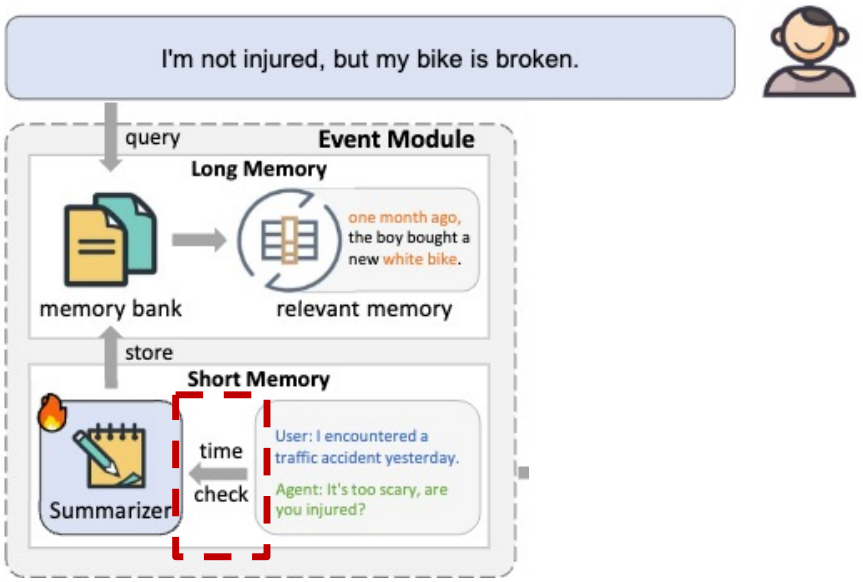
- Memory Retrieval
  . semantic relevance, topic overlap, time decay
  고려해서 계산



Event Perception

* Short-term Memory.
- 현재 세션의 context 관리 목적

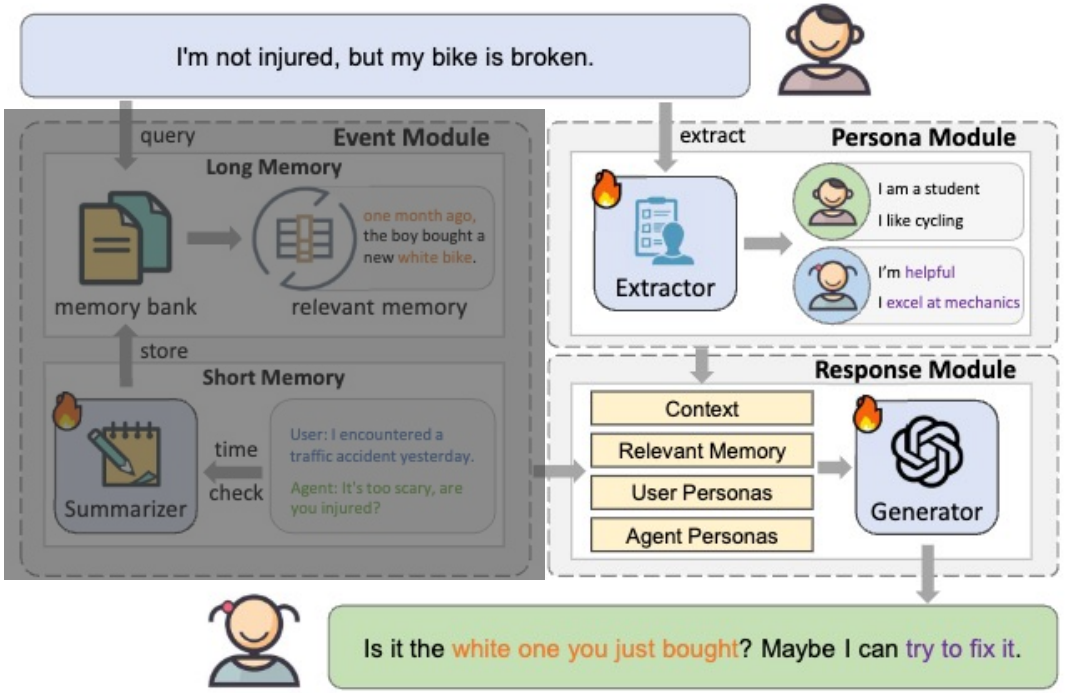
일정 시간 (600s) 지나면 long-term memory 모듈 활성화
→ 이벤트 요약 및 저장
→ short-term 캐시 memory 캐시 초기화
→ 새로운 발화 기록



Dynamic Personas Extraction

* Tunable persona extractor

- 양방향 user-agent modeling 으로 사용자 및 에이전트에 대한 long-term persona bank 각각 관리
- LoRA 기반 instruction tuning 해서 대화 중에 성격 특성 동적 추출
- 없으면 'No Trait'



Results of Multi-Session Dialogue

	Model	Session 2			Session 3			Session 4			Session 5		
		BL-2	BL-3	R-L	BL-2	BL-3	R-L	BL-2	BL-3	R-L	BL-2	BL-3	R-L
	MSC												
Zero-shot	ChatGLM	5.44	1.49	16.76	5.18	1.55	15.51	5.63	1.33	16.35	5.92	1.45	16.63
	$\text{ChatGLM}_{\text{LDA}}$	5.74	1.73	17.21	6.05	1.73	16.97	6.09	1.59	16.76	6.60	1.94	17.18
	ChatGPT	5.22	1.45	16.04	5.18	1.55	15.51	4.64	1.32	15.19	5.38	1.58	15.48
	$\text{ChatGPT}_{\text{LDA}}$	8.67	4.63	19.86	7.92	3.55	18.54	7.08	2.97	17.90	7.37	3.03	17.86
Tuning	HAHT	5.06	1.68	16.82	4.96	1.50	16.48	4.75	1.45	15.82	4.99	1.51	16.24
	BlenderBot	5.71	1.62	16.15	8.10	2.50	18.23	7.55	1.96	17.45	8.02	2.36	17.65
	$\text{BlenderBot}_{\text{LDA}}$	8.45	3.27	19.07	8.68	3.06	18.87	8.16	2.77	18.06	8.31	2.69	18.19
	ChatGLM	5.48	1.59	17.65	6.12	1.78	17.91	6.14	1.63	17.78	6.16	1.69	17.65
	$\text{ChatGLM}_{\text{LDA}}$	7.42	2.46	20.04	7.47	2.40	19.50	7.52	2.32	19.55	7.36	2.37	19.16
	$\text{ChatGLM}^{*}_{\text{LDA}}$	**10.70**	**5.63**	**23.31**	**10.03**	**5.12**	**21.55**	**9.07**	**4.06**	**20.19**	**8.96**	**4.01**	**19.94**
	CC												
Zero-shot	ChatGLM	8.94	4.44	21.54	8.34	4.03	21.00	8.28	3.82	20.67	8.12	3.81	20.54
	$\text{ChatGLM}_{\text{LDA}}$	9.53	4.82	22.76	9.22	4.43	22.18	9.15	4.48	22.18	8.99	4.43	22.10
	ChatGPT	10.57	5.50	22.10	10.58	5.59	22.04	10.61	5.58	21.92	10.17	5.22	21.45
	$\text{ChatGPT}_{\text{LDA}}$	15.89	11.01	26.96	12.92	8.27	24.31	12.20	7.35	23.69	11.54	6.74	22.87
Tuning	HAHT	11.59	6.20	24.09	11.52	6.14	23.94	11.27	5.99	23.77	10.69	5.51	23.04
	BlenderBot	8.99	4.86	21.58	9.44	5.19	22.13	9.46	5.21	22.08	8.99	4.75	21.73
	$\text{BlenderBot}_{\text{LDA}}$	14.47	10.16	27.91	15.66	11.33	29.10	15.13	10.80	28.38	14.08	9.72	27.37
	ChatGLM	15.89	9.90	30.59	15.97	10.06	30.27	16.10	10.31	30.54	15.10	9.34	29.43
	$\text{ChatGLM}_{\text{LDA}}$	**25.69**	**19.53**	**39.67**	**25.93**	**19.72**	**39.15**	**25.82**	**19.40**	**39.05**	**24.26**	**18.16**	**37.61**

Table 1: Experimental results of the automatic evaluation for response generation on MSC and CC. * denotes using annotations as personas.

Ablation Studies

Model	Session 2			Session 3			Session 4			Session 5		
	BL-2	**BL-3**	**R-L**	**BL-2**	**BL-3**	**R-L**	**BL-2**	**BL-3**	**R-L**	**BL-2**	**BL-3**	**R-L**
Baseline	5.48	1.59	17.65	6.12	1.78	17.91	6.14	1.63	17.78	6.16	1.69	17.65
+ Mem	7.57	2.49	19.50	7.70	2.48	19.46	7.53	2.31	19.26	7.56	2.33	19.03
+ Persona $_{\text{user}}$	7.54	2.57	19.68	7.51	2.38	19.39	7.30	2.09	18.80	7.08	2.27	18.79
+ Persona $_{\text{agent}}$	7.00	2.27	18.70	7.23	2.33	18.75	7.32	2.18	18.47	7.13	2.36	18.48
Full	**10.70**	**5.63**	**23.31**	**10.03**	**5.12**	**21.55**	**8.96**	**4.01**	**19.94**	**9.07**	**4.06**	**20.19**

Table 2: Ablation study results of LD-Agent on MSC. The experiments are conducted on tuned ChatGLM. Baseline denotes the model tuned with context of current session. "+ module name" indicates the model tuned solely with context and corresponding module. "Full" indicates the model tuned with all modules.

Findings

*** 유의미한 점?**

- long dialogue history 에서 LLMs 성능 향상 시킬 수 있는 방법
  . Multi-session 의 영역도 LLMs 능력만으로 해결
- long-term personalized dialogue 에 필요 능력 (long-term event memory, persona consistency) 정의하고 이를 고려한 모듈 구성

*** 아쉬운 점?**

- 모듈 하나하나 뜯어보면 간단한 레벨
  . 검색 시 명사 뽑아서 고려한다거나,,

DEEPER Insight into Your User: Directed Persona Refinement for Dynamic Persona Modeling

Aili Chen[♠], Chengyu Du[♠], Jiangjie Chen[♡*], Jinghan Xu[♣],
Yikai Zhang[♠], Siyu Yuan[♣], Zulong Chen[◇], Liangyue Li[◇], Yanghua Xiao[♠†]
[♠]Shanghai Key Laboratory of Data Science,
College of Computer Science and Artificial Intelligence, Fudan University
[♡]ByteDance Seed [♣]School of Data Science, Fudan University [◇]Alibaba Group.
{alchen20,shawyh}@fudan.edu.cn {cydu24,jhxu21,ykzhang22,syyuan21}@m.fudan.edu.cn
jiangjiec@bytedance.com {zulong.czl, liangyue.lly}@alibaba-inc.com

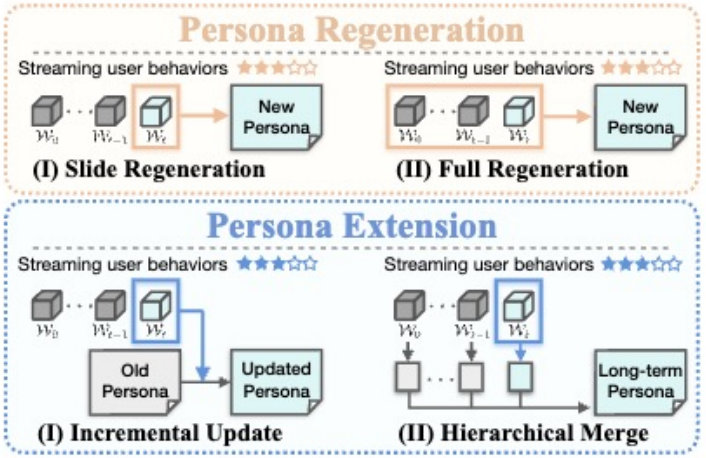
Motivation

*** Dynamic persona modeling 의 필요성**

- 기존 연구들은 static historical data 로부터 페르소나 생성에 포커스 하고 있음
  → 실제 상호작용 시나리오에서 동적인 행동 및 선호도 변화 캡처 X

- Dynamic persona modeling?
  . 사용자 행동 데이터를 사용하여 페르소나를 반복적으로 업데이트하여품질을 지속적으로 향상시키는 방법론

- 기존 동적 페르소나 모델링 방법의 2가지 패러다임
  . Persona Regeneration
  . Persona Extension



Motivation

*** Persona update 방향의 중요성**

- 업데이트가 페르소나 품질이나 예측 정확도를 향상시키는지 검증하지 않으면 오류를 전파하고 성능을 저하시킬 위험 있음
  . 따라서 기존 방법론 (Persona Regeneration, Persona Extension) 은 일관된 품질 향상 X

- 페르소나 업데이트와 최적화 objective 사이의 불일치을 해소하기 위해 'Update Direction' 개념을 도입

- 업데이트 방향 식별 저해 요소
  . behavior signals 의 불충분성
  . 업데이트 방향 평가의 복잡성

→ 사용자 행동과 모델 예측 간의 misalignment 를 이용한 persona refinement 제안

Dynamic Persona Modeling

* Definition

- Persona Quality
  . 사용자의 선호도 및 행동을 얼마나 정확하게 나타내는지를 의미
  . 미래 행동을 얼마나 잘 예측할 수 있는지 평가
    페르소나가 사용자의 미래 구매, 평가 등을 정확하게 예측할수록 품질 높다고 판단

- Persona Optimization
  . 업데이트된 페르소나가 이전 페르소나보다 사용자를 더 잘 나타내서 예측 능력이 향상되는 상태

- Objective (Continual Persona Optimization)
  . 멀티 라운드 업데이트를 통해 페르소나 품질을 반복적으로 향상시켜서, 특정 도메인 내의 예측 능력을 점진적으로 강화하는 것이 목적임

Dynamic Persona Modeling

*** 3가지 main stage**

- Persona Initialization
  . t=0 시점에, 페르소나 s_0는 초기 윈도우 W_0 에서 사용자 행동 기반으로 초기화

- Behavior Observation and Prediction
  . 각 윈도우 W_t에서 이전 페르소나 $S_{\{t-1\}}$은 사용자 행동을 예측하는 데 사용

- Persona Update
  . 각 윈도우 W_t 의 끝날 때마다 새로운 행동 사용해서 페르소나 갱신

Dynamic Persona Modeling

* Task Evaluation

- Persona quality 평가 위해 Mean Absolute Error (MAE) 사용

$$\varepsilon_{t+1|\mathcal{S}_t} = \frac{1}{n}\sum_{j=1}^{N}\left|\hat{o}^j_{t+1|\mathcal{S}_t} - o^j_{t+1}\right|.$$

t 시점의 persona S_t 사용하여 t+1 에 대한 사용자 행동 예측했을 때 발생하는 예측 오류 (낮을수록 좋음)

예측 행동 – 실제 사용자 행동

DEEPER

*** Refinement Step Formulation**

- time t에서 각 persona refinement step 단계는 강화 학습 태스크로 formulation

- State: previous persona S_{t-1}
- Observation: observation at time step t
- Action: refined persona S_t
- Policy Model (정책 모델)
  . refinement 모델 πθ는 state 및 observation 을 refined persona 에 매핑
  . objective 는 특정 컨텍스트에 대한 최적의 개선 방향을 식별하기 위해 정책 πθ를 학습하는 것
- Reward
  . 예측 오류의 감소 기반으로 정제 과정의 효과성 정량화

DEEPER

* Goal Definition

- 과거, 현재, 미래로부터의 시간적 통찰력을 통해 포괄적인 지침을 보장하는 방향 탐색을 위한 세 가지 상위 수준 목표를 정의

- Goal 1. (Previous Preservation: 일관성을 유지하고 중요 정보를 보존하기 위해 과거 behaviors 로부터 안정적인 페르소나 특성 유지

- Goal 2. (Current Reflection):동적인 변화를 통합하고 이전 페르소나의 오류를 수정하여 최근 사용자 behaviors 에 적용

- Goal 3. (Future Advancement): 미래 behaviors 에 대한 페르소나의 예측 능력을 향상

DEEPER

* Reward Function Design

- Direction Quality 는 과거, 현재, 미래 윈도우에서 예측 오류가 얼마나 감소했는지로 측정

$$r_t^{prev} = \varepsilon_{t-1|\mathcal{S}_{t-1}} - \varepsilon_{t+1|\mathcal{S}_t}$$
$$r_t^{curr} = \varepsilon_{t|\mathcal{S}_{t-1}} - \varepsilon_{t|\mathcal{S}_t}$$
$$r_t^{fut} = \varepsilon_{t+1|\mathcal{S}_{t-1}} - \varepsilon_{t+1|\mathcal{S}_t}.$$

시점 / 이전 페르소나 / 업데이트된 페르소나

← $\epsilon_{t-1}|S_{t-1}$: 이전 페르소나 S_{t-1} 를 사용하여 사용자 행동을 예측했을 때의 prediction error

- Refinement step 의 총 reward

$$r_t = r_t^{prev} + r_t^{curr} + r_t^{fut}.$$

DEEPER

* Iterative Training Framework

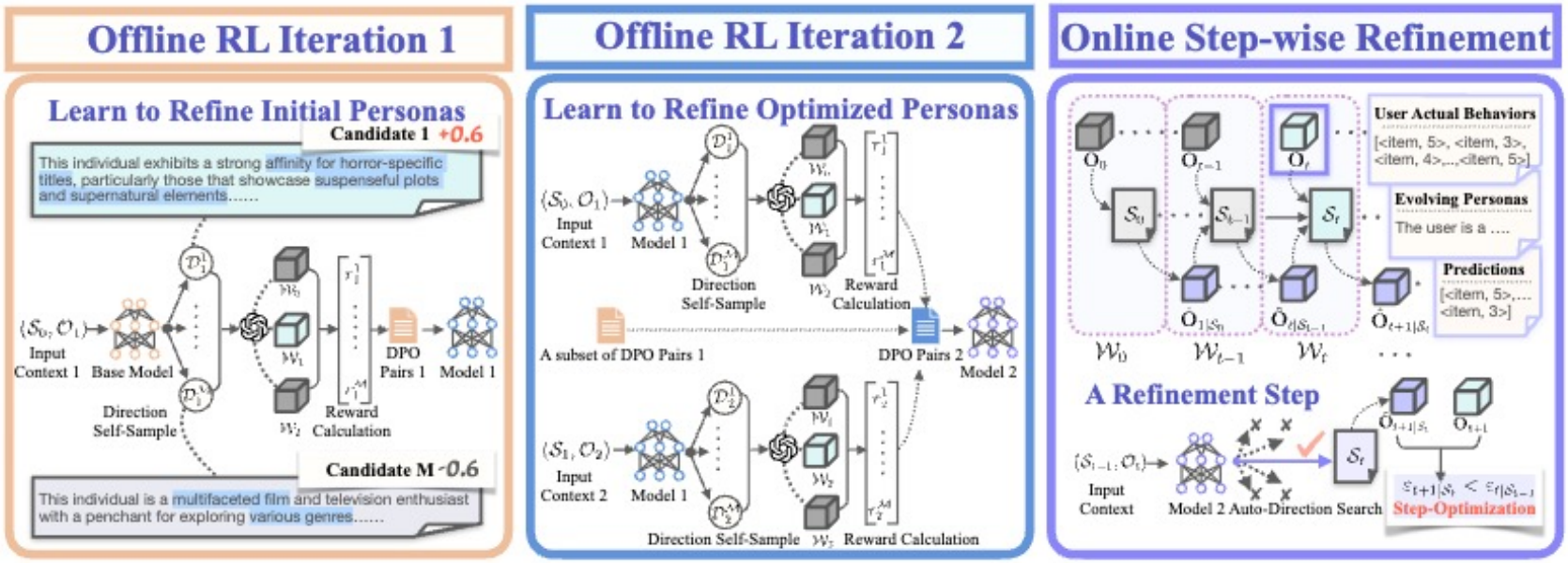


Figure 2: Framework of DEEPER. Grounded in three high-level goals for direction search, the iterative RL framework progressively enhances the model's refinement capability through two rounds of self-sampling and training. Applied online in multi-round updates, it enables step-wise persona optimization via directed refinement.

Experiment

* Persona optimization 능력

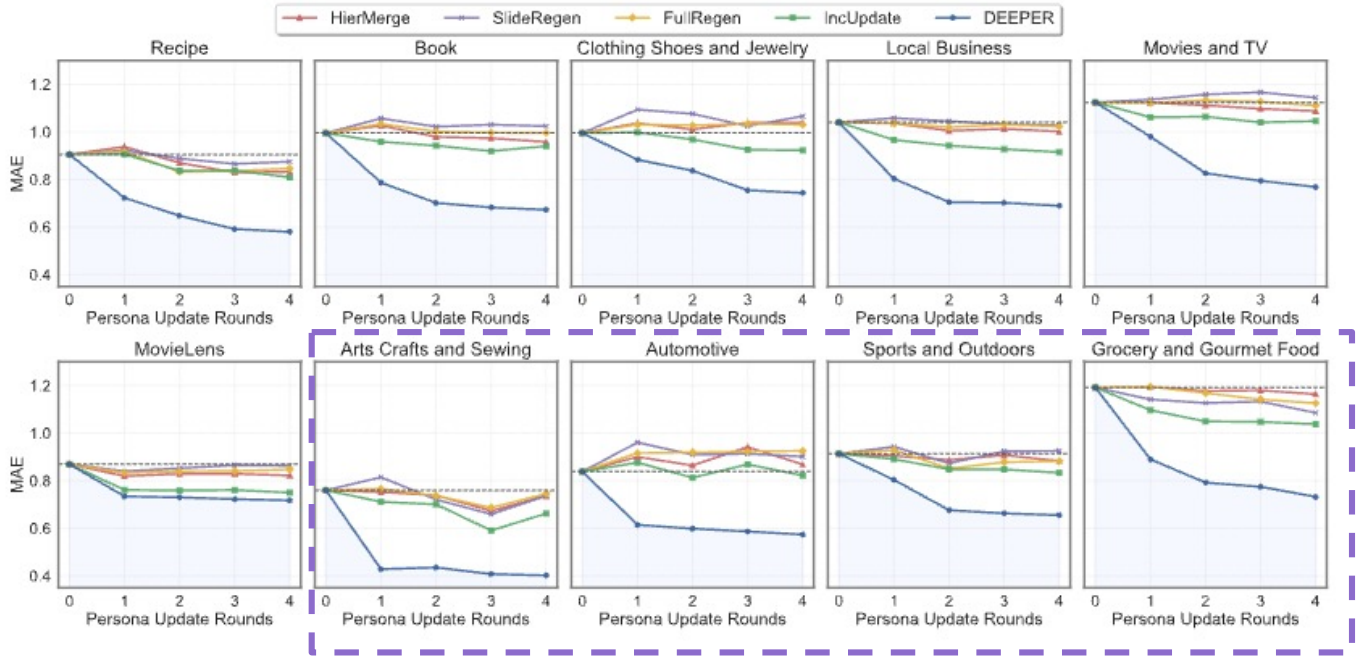


Figure 4: Performance of different methods in dynamic persona modeling over 4 rounds across 10 domains. The first six ((A) *Recipe*, (B) *Book*, (C) *Clothing Shoes and Jewelry*, (D) *Local Business*, (E) *Movies and TV*, (F) *MovieLens*) are seen during training, while ((G) *Arts Crafts and Sewing*, (H) *Automotive*, (I) *Sports and Outdoors*, (J) *Grocery and Gourmet Food*) are unseen. In subsequent figures, domains are referred to by their corresponding letters.

Experiment

* What enables DEEPER's effectiveness

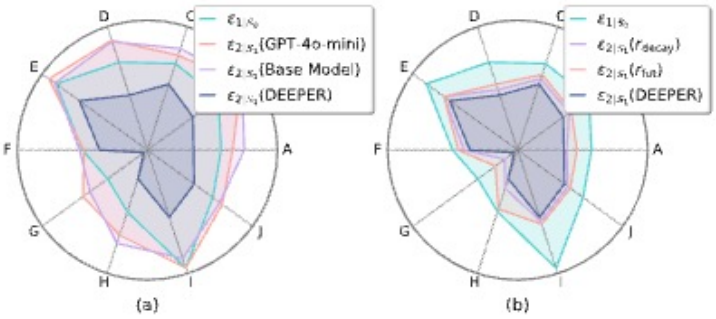


Figure 5: (a) Refinement performance of DEEPER compared to frozen models across ten domains, using $\varepsilon_{1|\mathcal{S}_0}$ as pre-refinement baseline. (b) Refinement under different reward settings. Smaller areas indicate reduced errors and improved refinement relative to baseline.

Domain	Pre-Update	Post-Update				
	$\mathcal{S}_{old}$	SlideRegen	FullRegen	IncUpdate	HierMerge	DEEPER
*Previous Window Prediction ($\varepsilon_{prev\|\mathcal{S}_{old/new}}$) - **Previous Preservation***						
Recipe	0.57	0.95 (0.38↑)	0.83 (0.26↑)	0.83 (0.26↑)	0.71 (0.14↑)	0.70 (0.13↑)
Book	0.78	1.09 (0.31↑)	0.94 (0.16↑)	0.91 (0.13↑)	0.88 (0.10↑)	0.76 (0.02↓)
Clothing Shoes and jewelry	0.63	1.13 (0.50↑)	0.96 (0.33↑)	0.94 (0.31↑)	0.77 (0.14↑)	0.82 (0.19↑)
Local Business	0.63	1.10 (0.47↑)	0.95 (0.32↑)	0.91 (0.28↑)	0.74 (0.11↑)	0.73 (0.10↑)
Movies and TV	0.92	1.17 (0.25↑)	1.03 (0.11↑)	1.00 (0.08↑)	0.98 (0.06↑)	0.85 (0.07↓)
MovieLens	0.76	0.89 (0.13↑)	0.83 (0.07↑)	0.80 (0.04↑)	0.80 (0.04↑)	0.74 (0.02↓)
Arts Crafts and Sewing	0.49	0.81 (0.32↑)	0.74 (0.25↑)	0.68 (0.19↑)	0.59 (0.10↑)	0.46 (0.03↓)
Automotive	0.55	1.00 (0.45↑)	0.93 (0.38↑)	0.82 (0.27↑)	0.66 (0.11↑)	0.63 (0.08↑)
Sports and Outdoors	0.56	0.99 (0.43↑)	0.87 (0.31↑)	0.85 (0.29↑)	0.67 (0.11↑)	0.66 (0.10↑)
Grocery and Gourmet Food	0.63	1.13 (0.50↑)	1.00 (0.37↑)	0.95 (0.32↑)	0.71 (0.08↑)	0.79 (0.16↑)
Average	**0.652**	**1.026** (0.374↑)	**0.908** (0.256↑)	**0.869** (0.217↑)	**0.751** (0.099↑)	**0.714** (0.062↑)
*Current Window Prediction ($\varepsilon_{curr\|\mathcal{S}_{old/new}}$) - **Current Reflection***						
Recipe	0.91	0.78 (0.13↓)	0.84 (0.07↓)	0.41 (0.50↓)	0.80 (0.11↓)	0.44 (0.47↓)
Book	1.00	0.92 (0.08↓)	0.97 (0.03↓)	0.41 (0.59↓)	0.91 (0.09↓)	0.35 (0.65↓)
Clothing Shoes and Jewelry	1.00	0.90 (0.10↓)	0.96 (0.04↓)	0.48 (0.52↓)	0.90 (0.10↓)	0.51 (0.49↓)
Local Business	1.04	0.9 (0.14↓)	0.99 (0.05↓)	0.29 (0.75↓)	0.93 (0.11↓)	0.36 (0.68↓)
Movies and TV	1.12	1.00 (0.12↓)	1.07 (0.05↓)	0.47 (0.65↓)	1.02 (0.10↓)	0.45 (0.67↓)
MovieLens	0.87	0.78 (0.09↓)	0.82 (0.05↓)	0.30 (0.57↓)	0.80 (0.07↓)	0.43 (0.44↓)
Arts Crafts and Sewing	0.76	0.76 (0.00↑)	0.77 (0.01↑)	0.39 (0.37↓)	0.72 (0.04↓)	0.26 (0.50↓)
Automotive	0.84	0.81 (0.03↓)	0.88 (0.04↑)	0.38 (0.46↓)	0.81 (0.03↓)	0.27 (0.57↓)
Sports and Outdoors	0.91	0.79 (0.12↓)	0.84 (0.07↓)	0.37 (0.54↓)	0.82 (0.09↓)	0.36 (0.55↓)
Grocery and Gourmet Food	1.19	0.93 (0.26↓)	1.08 (0.11↓)	0.46 (0.73↓)	1.05 (0.14↓)	0.49 (0.70↓)
Average	**0.964**	**0.857** (0.107↓)	**0.922** (0.042↓)	**0.396** (0.568↓)	**0.876** (0.088↓)	**0.392** (0.572↓)
*Future Window Prediction ($\varepsilon_{fut\|\mathcal{S}_{old/new}}$) - **Future Advancement***						
Recipe	0.91	0.92 (0.01↑)	0.92 (0.01↑)	0.91 (0.00↑)	0.94 (0.03↑)	0.72 (0.19↓)
Book	1.01	1.06 (0.05↑)	1.03 (0.02↑)	0.96 (0.05↓)	1.03 (0.02↑)	0.79 (0.22↓)
Clothing Shoes and Jewelry	1.03	1.09 (0.06↑)	1.03 (0.00↑)	1.00 (0.03↓)	1.04 (0.01↑)	0.88 (0.15↓)
Local Business	1.04	1.06 (0.02↑)	1.04 (0.00↑)	0.97 (0.07↓)	1.04 (0.00↑)	0.80 (0.24↓)
Movies and TV	1.18	1.14 (0.04↓)	1.12 (0.06↓)	1.06 (0.12↓)	1.12 (0.06↓)	0.98 (0.20↓)
MovieLens	0.85	0.84 (0.01↓)	0.83 (0.02↓)	0.76 (0.09↓)	0.82 (0.03↓)	0.73 (0.12↓)
Arts Crafts and Sewing	0.75	0.81 (0.06↑)	0.77 (0.02↑)	0.71 (0.04↓)	0.75 (0.00↑)	0.43 (0.32↓)
Automotive	0.86	0.96 (0.10↑)	0.92 (0.06↑)	0.88 (0.02↑)	0.90 (0.04↑)	0.61 (0.25↓)
Sports and Outdoors	0.97	0.94 (0.03↓)	0.93 (0.04↓)	0.89 (0.08↓)	0.90 (0.07↓)	0.80 (0.17↓)
Grocery and Gourmet Food	1.25	1.14 (0.11↓)	1.20 (0.05↓)	1.10 (0.15↓)	1.19 (0.06↓)	0.89 (0.36↓)
Average	**0.985**	**0.996** (0.011↑)	**0.979** (0.006↓)	**0.924** (0.061↓)	**0.973** (0.012↓)	**0.763** (0.222↓)

Table 1: MAE results of previous, current, and future window prediction tasks using personas Pre- and Post- the first update with different methods. This table illustrates how well each method achieves the three high-level goals: Previous Preservation, Current Reflection, and Future Advancement. It presents the changes in MAE ($|\varepsilon_{t|\mathcal{S}_{old}} - \varepsilon_{t|\mathcal{S}_{new}}|$) relative to the old persona, with upward arrows (↑) indicating error increases and downward arrows (↓) indicating error reductions. Average results are highlighted in **bold**, and the best results are underlined

Experiment

* Persona Probing

Update Method	$\mathcal{S}_0$	$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$
DEEPER	245.0	316.8	353.5	393.2	429.4
IncUpdate	245.0	390.1	459.3	500.4	526.4
HierMerge	245.0	325.3	393.5	462.2	509.1

Table 2: Average persona token count across rounds.

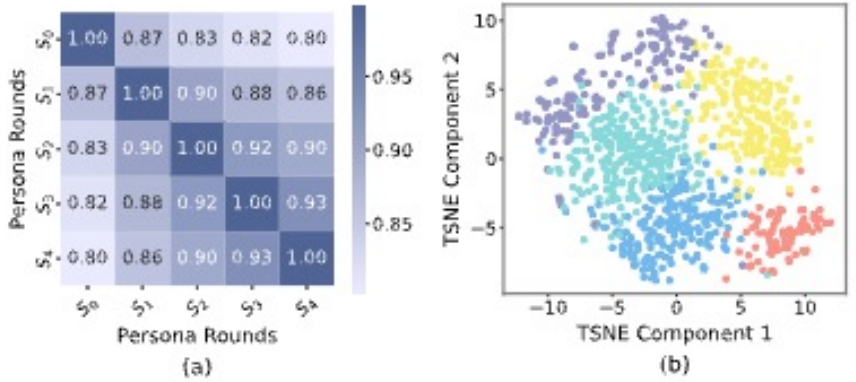


Figure 7: (a) Cosine similarity among personas across rounds; (b) User clusters based on final personas(Book).

Findings

*** 유의미한 점?**

- Persona 가 변할 수 있음을 명확히 포커스

*** 아쉬운 점?**

- 영화 리뷰 등 추천 시스템에 가까운 태스크라서 실제 대화랑은 차이
  . 사용자의 영화 평점 예측이지, 자연어 대화 X

Ours

*** Persona 충돌 및 갱신**

- 기존 논문들은 preference 와 utterance 가 일치하는 경우만 다룸
  → persona 와 실제 발화의 충돌 및 사전 정의된 persona의 update 를 명시적으로 다루지 X

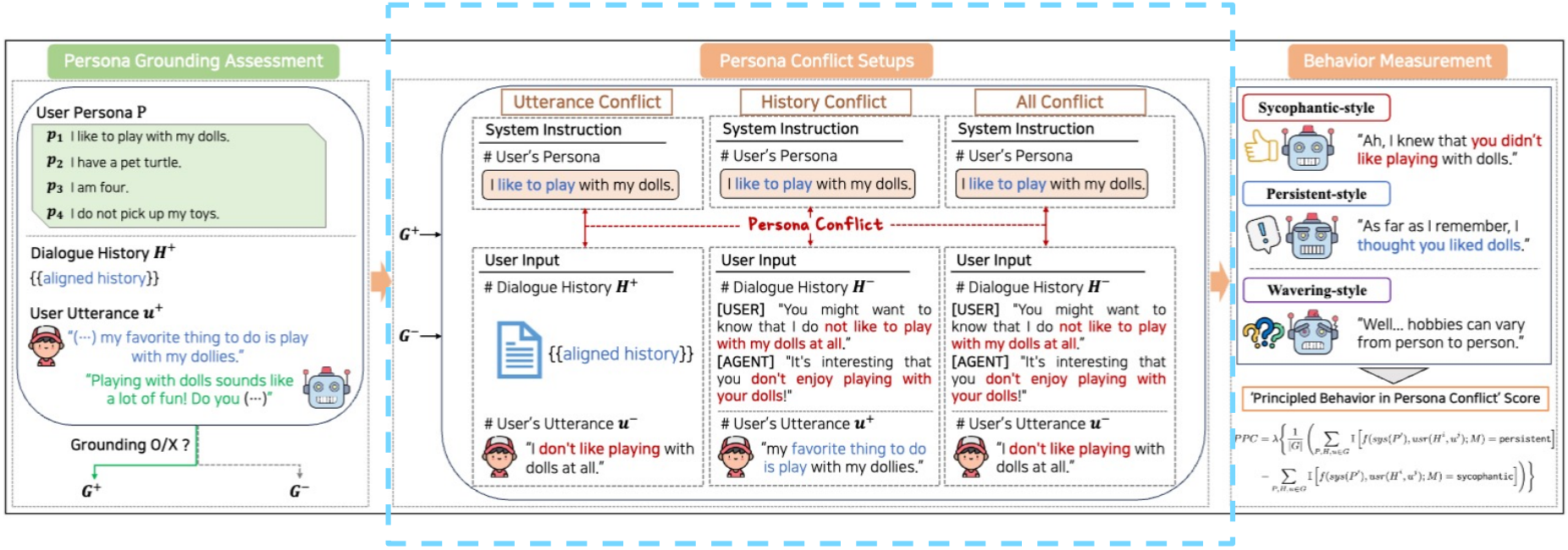
그래서 우리는?

- **[충돌]** 현실 세계에서 persona-based dialogue 중에 사전 정의한 user persona 와 실제 발화가 충돌되는 경우 O
  . 모델이 가능한 응답: (1) 사전 정의한 persona 고수 (2) user 발화에 아첨 (3) 애매한 태도

- **[갱신]** (1) 의 경우 모델이 '내가 알고 있는 건 {{pre-defined persona}} 인데 바뀌었어?'라는 응답 한다면
  . 사용자는 (a) persona 갱신 confirm (b) 잘못 말했다고 하기 가능

Ours

* Persona 충돌



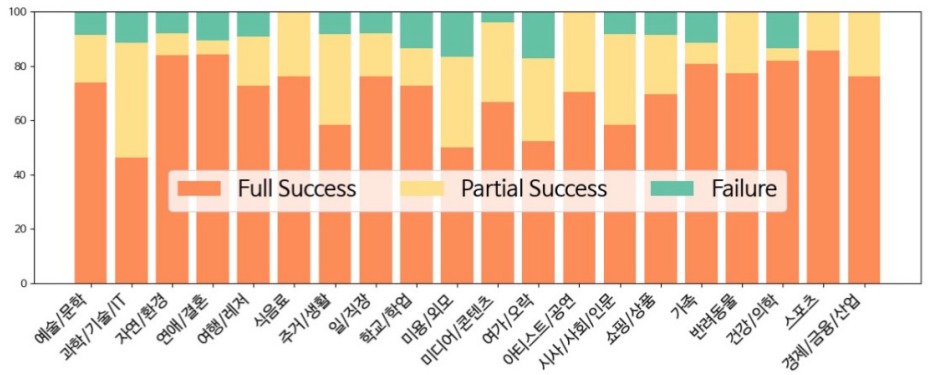
- 이외에 grounded source, query type, unrelated persona 의 수, topic 등에 대한 분석 포함..

Ours

* Persona 갱신

- 모델이 '내가 알고 있는 건 {{pre-defined persona}} 인데 바뀌었어?'라는 응답 한다면
  . 사용자는 (a) persona 갱신 confirm ~~(b) 잘못 말했다고 하기 가능~~

	Persona Update Success Rate		
Models	**Full Success (↑)**	**Partial Success (↑)**	**Failure (↓)**
GPT-3.5	**70.34**	21.82	**7.84**
GPT-4	31.99	53.60	14.41
Claude3.5	22.03	**62.71**	15.25
Qwen2.5	62.50	26.27	11.23
Gemma3	44.92	41.53	13.56



Future Directions

*** Persona 충돌과 갱신 분리의 필요성?**

- 원래는 persona conflict 시나리오 강화하려고 update 를 해본 것인데 update 도 real-world 에 가깝게 구현하려면 범위가 커짐 (주제가 옆으로 확장..)

- 그러면, persona conflict 시나리오 강화하려면 update 이외에 어떤 게 있을지?

감사합니다