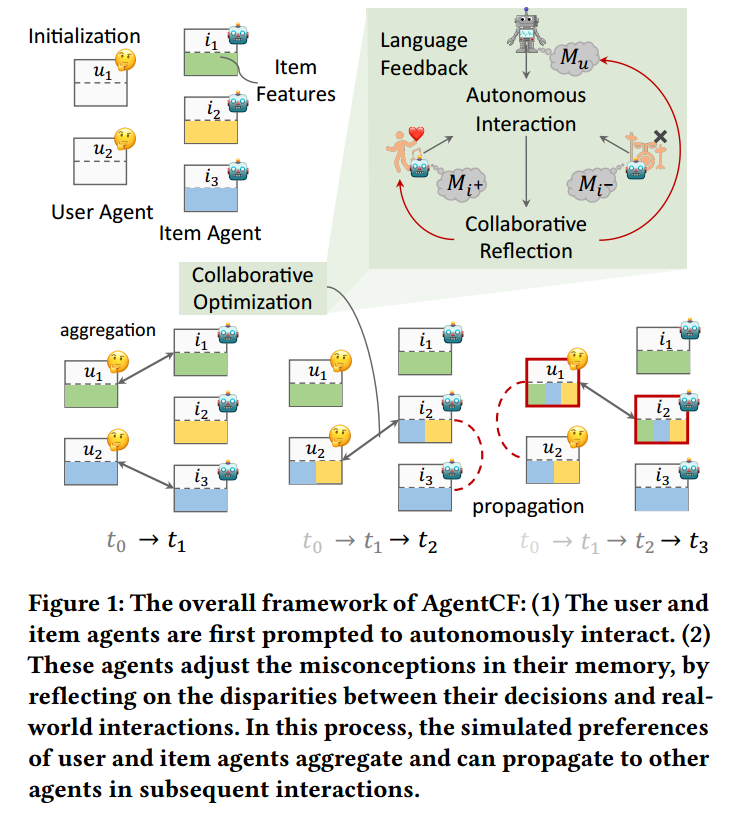
### AgentCF: Collaborative Learning with Autonomous Language Agents for Recommender Systems

#### 背景

现有研究仍主要聚焦于模拟人类对话场景,商品点击行为等隐性反映用户偏好的非语言表现尚未得到深入探索。为模拟这类行为，部分研究通过将交互记录转化为自然语言文本来引导大模型，由于通用语言建模与个性化行为建模之间的鸿沟，这种方法难以捕捉这些互动背后的潜在行为模式。

为此，该研究提出AgentCF，通过基于智能体的协同过滤模拟推荐系统中的用户与商品交互。该方案将用户和商品均视为智能体，并开发出一种协同学习方法来优化两者的协同作用。

#### 核心技术



##### 协同Agent优化

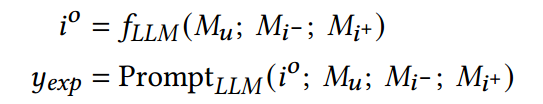
###### 记忆设计

**用户记忆M\_u** 对于用户代理，记忆模块存储反映用户偏好的信息。为每个用户代理配备了短期记忆和长期记忆。当用户代理参与新的交互时，可以直接访问短期记忆，同时从长期记忆中提取相关信息。

**项记忆M\_i** 项代理配备了可调节的记忆模块，以记录其自身特征和交互用户的偏好信息，如，摇滚CD可能是精力充沛的吉他爱好者的理想选择。与用户记忆不同，只为每个项代理配备了统一的记忆模块，因为项信息相对稳定。项记忆可以通过其自身信息初始化，并在进行新的交互时根据用户偏好不断更新。

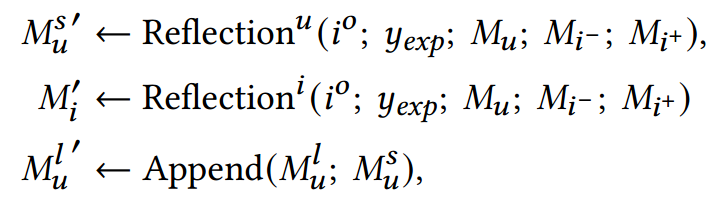
项对比选择的自主交互

采用真实的用户行为序列（按时间顺序排列）作为“训练数据”。每一步都提供了一个正项和一个负项作为用户代理的候选项，同时引入流行偏见和位置偏见。用户代理的任务是从候选中选择一个项，并协同考虑其记忆中的偏好和候选项记忆特征来解释这个选择：



协同反思和记忆更新

如果用户代理做出了正确选择，则告知其正确性并将交互数据存储到记忆中。对于错误的选择，则修改代理的记忆和行为，使用户代理能够模拟真实的交互行为，并且项代理能够与交互用户的偏好一致。



注意,不修改负项代理的记忆，因为发现LLM倾向于过度抱怨负项代理的缺点而忽略了它可能对其他用户也有吸引力的事实。

每次交互中，迭代执行对比选择和协同反思，直至用户代理的选择与真实用户一致或达到最大迭代轮次。

###### Agent交互推理

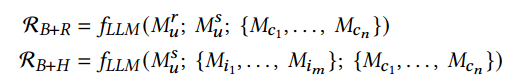
基本提示策略

直接向大模型提出代理模拟的用户偏好和候选项特征,使得大模型能够充当协同推荐器。



高级推理策略

虽然短期记忆描述了用户代理的当前偏好，但从长期记忆中检索对候选项的特定偏好可以做出更个性化的推断。此外，当交互记录稀疏且偏好传播有限时，可以进一步将用户历史交互纳入提示，使LLM能够充当顺序推理器



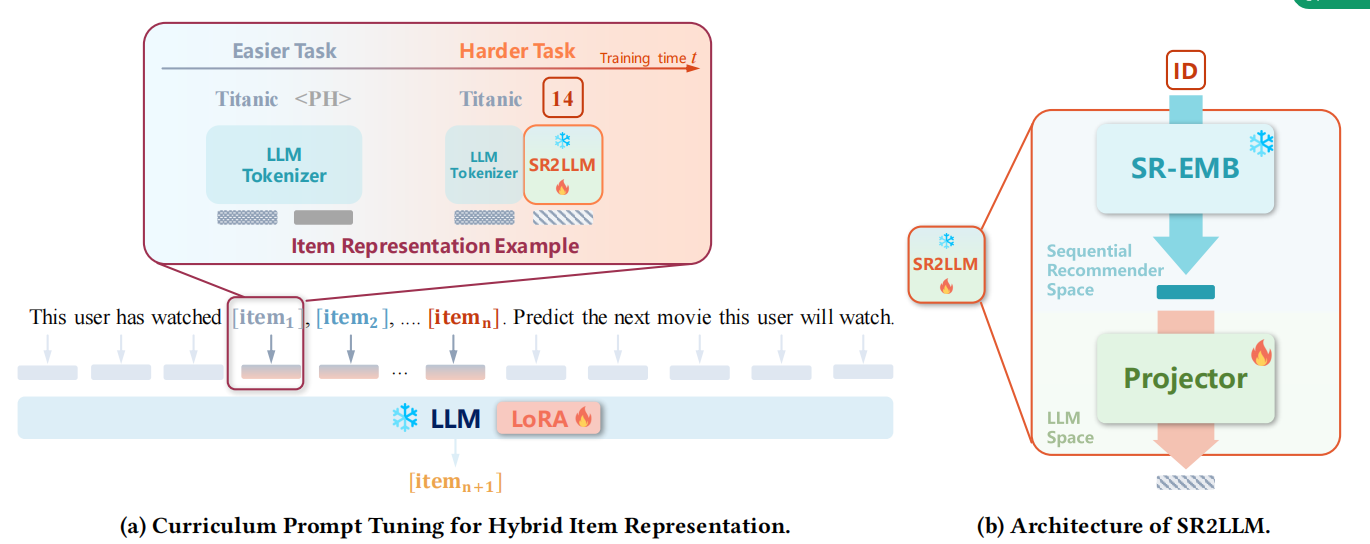
### LLaRA: Large Language-Recommendation Assistant

#### 背景

现有LLM方法在融合用户行为信息和项的语义信息方面仍存在不足。一方面，仅使用项的文本元数据或ID索引信息难以充分表达项的多维特征；另一方面，传统推荐模型擅长捕获行为模式，但缺乏丰富的背景知识。

因此，该论文提出LLaRA，试图结合传统推荐模型的行为模式和LLMs的知识推理能力，通过一种创新的混合提示设计和逐步训练策略，实现更全面、统一的个性化推荐系统。

#### 核心技术



###### 项表征

文本token表征（内容表征）

使用LLM的分词器和嵌入层将项的文本特征如标题、描述等转为文本token表征。



行为token表征（ID表征）

传统的序列推荐模型如GRU4Rec，在基于历史交互数据训练后，能够有效捕捉基于ID的项嵌入中的序列模式。因此使用从传统序列推荐模型学习到的ID表示形式作为行为token表征。

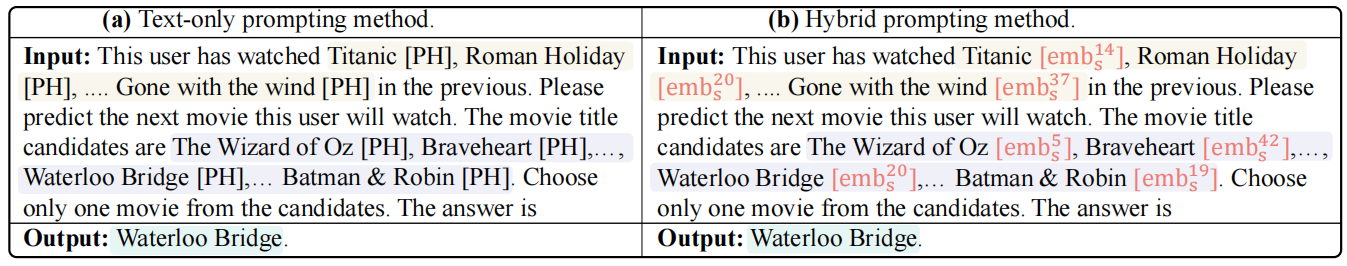
基于ID的项的行为token表征难以被LLM理解，因此将其视为独立文本数据的另一种模态，并使用投影层投影到LLM词嵌入空间。

混合token表征

将文本表征和行为表征混合形成对项的全面描述。



###### 混合提示设计



纯文本提示

在提示词中仅通过文本数据表示项。提示词包含三个关键要素：任务定义，交互序列，候选集。

混合提示

提示词中，交互序列和候选集中的项通过文本描述和紧随其后的行为表征（混合token表征）表示。

###### 课程提示调优

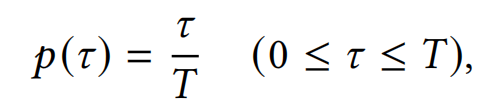
课程学习（curriculum learning）强调从简单到复杂的学习任务逐步训练模型。调优过程首先采用较为直接的纯文本提示方法，帮助模型建立对序列推荐任务的基本理解。随后，逐步引入了结合行为token表征的混合提示方法，从而增加了调优过程的复杂性。

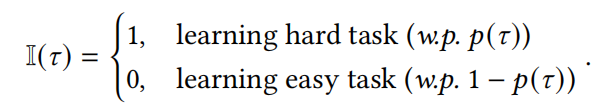
复杂性评估

课程学习的初步步骤是评估每个任务的复杂度。这里直接将纯文本提示看作简单任务，将混合提示看作复杂任务。

调度器构建

通过训练过程中逐步从易任务过渡到难任务来构建课程调度器。定义p(t)为在训练时间t时学习难任务的概率，1-p(t)表示学习简单任务的概率，其中T为总训练时间。





通过采用课程提示调优策略，确保了从模型对文本数据的初步理解到最终理解传统推荐系统中更复杂的基于ID的项嵌入的无缝过渡。

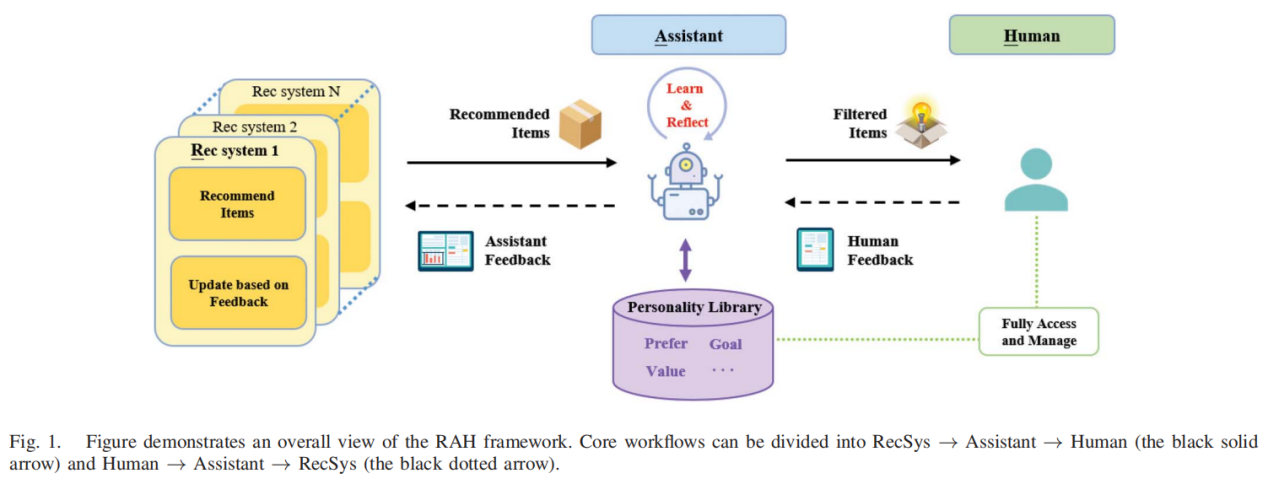
### RAH! RecSys–Assistant–Human: A Human-Centered Recommendation Framework With LLM Agents

#### 背景

当前推荐系统面临诸多挑战，如在推荐准确性与用户满意度之间取得平衡、消除推荐系统中的偏差、保护用户隐私的同时赋予其自主选择权，以及解决跨领域冷启动问题等。

但本文认为解决这些问题并非推荐系统专属职责。从人类视角切入，采用先进的用户建模技术捕捉用户行为与偏好，让用户参与推荐过程，如定制偏好并提供明确反馈，能更好地应对上述问题。

#### 方法



##### workflows

###### RecSys → Assistant → Human

1. 推荐系统最初生成跨越不同领域的候选项，如书籍、电影和游戏。
2. 该助手会收集这些跨领域推荐。它从记忆中检索用户学习到的人格。使用用户的人格档案，助手进一步过滤候选项，创建一个定制列表。
3. 最后，用户会从助手那里收到一组统一的个性化过滤推荐。

###### Human → Assistant → RecSys

1. 用户首先对商品进行反馈（例如点赞或点踩），智能助手会直接接收这些原始反馈，而非依赖推荐系统。
2. 随后，智能助手将通过分析用户反馈逐步学习其个性特征。
3. 最终，系统会将用户的反馈转化为智能助手的专属反馈，从而实现智能助手能精准筛选用户偏好并传递给推荐系统。

##### Components

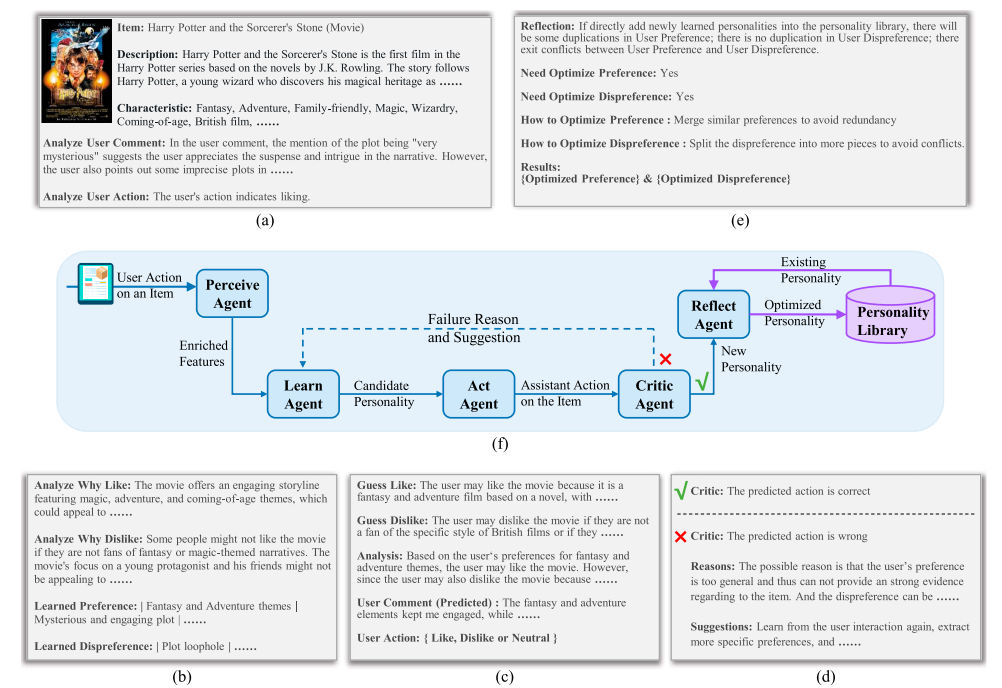
感知智能体（Perceive Agent）：理解推荐内容，如项特征和用户反馈含义。

学习智能体（Learn Agent）：从用户行为中汲取个性特征，并存入个性库。

行动智能体（Act Agent）：依据学习到的个性执行相应行为，例如为用户过滤掉不喜欢的项目。

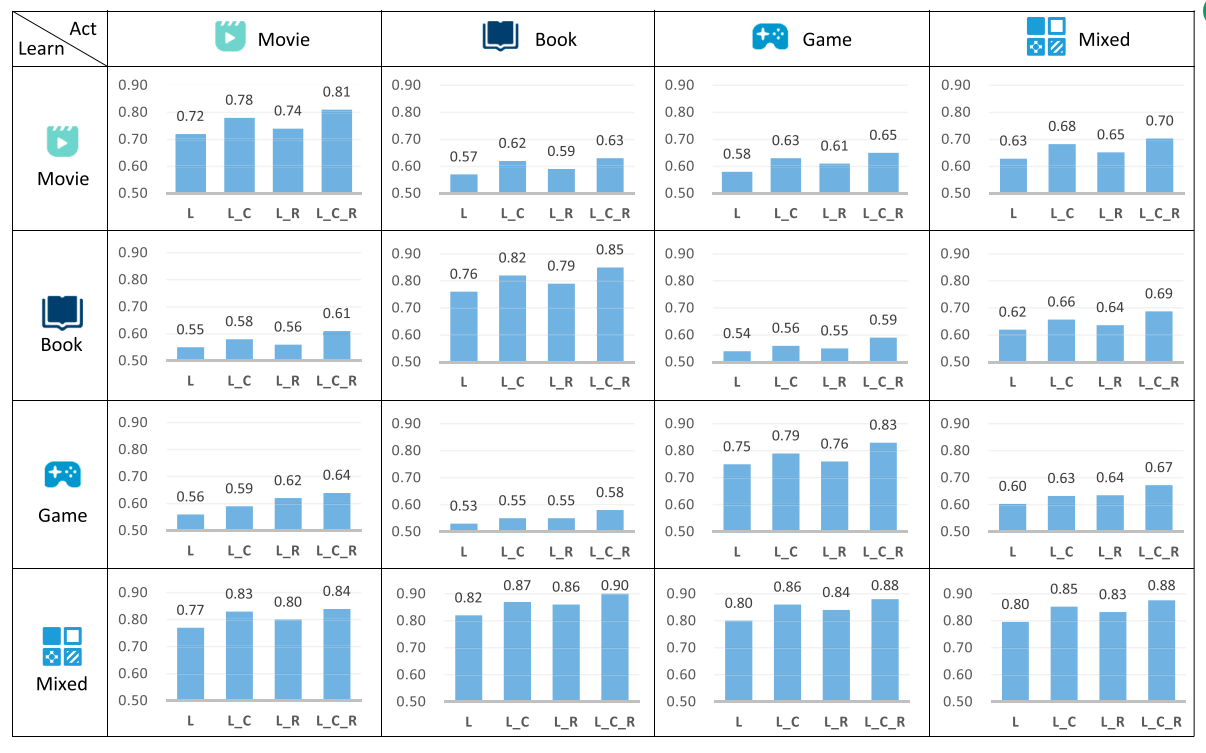
评论智能体（Critic Agent）：检验行动是否与用户偏好相符，并分析调整以减少差异。

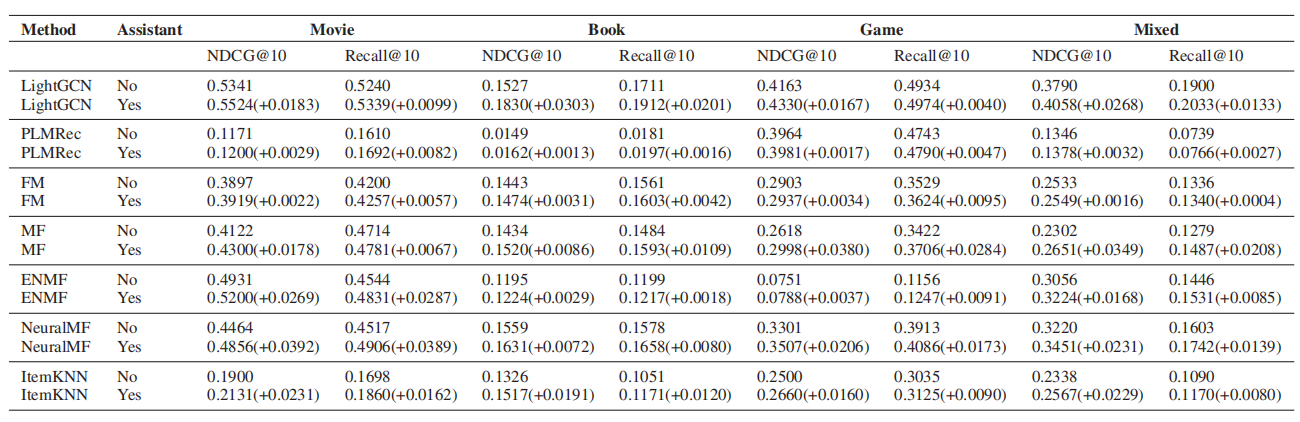
反思智能体（Reflect Agent）：审视并优化累积学到的个性，解决重复与冲突问题。



#### 实验

##### Assistants’ Alignment With Users





### ToolRec: Let Me Do It For You: Towards LLM Empowered Recommendation via Tool Learning

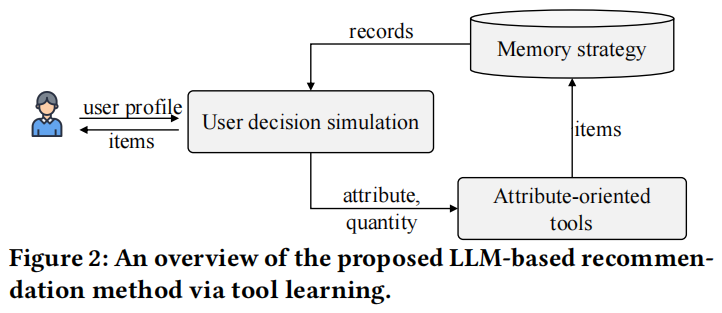
#### 背景

传统推荐系统在准确捕捉用户细粒度偏好方面存在挑战，主要依赖历史交互数据，难以全面理解用户真实兴趣，且缺乏对物品和用户的常识性知识，导致推荐范围受限。

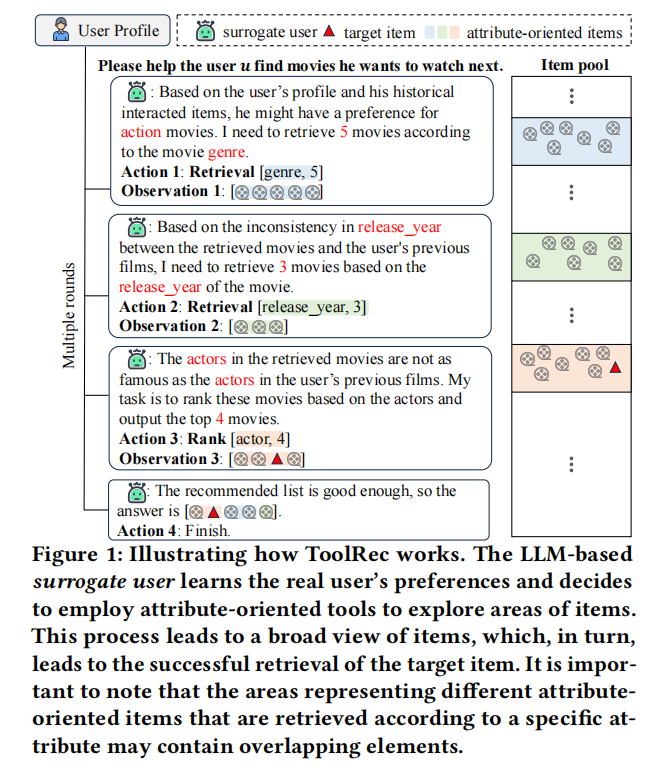
LLMs具备常识推理和利用外部工具的能力，有望改善推荐系统的性能。但现有基于LLM的推荐系统存在幻觉问题、物品语义空间与用户行为空间不一致，或控制策略过于简单等局限。

工具学习的兴起：随着工具学习在基础模型中的发展，研究者提出利用LLM作为代理用户，通过工具学习来模拟用户决策过程，以提升推荐系统性能。

#### 方法



##### 用户决策模拟

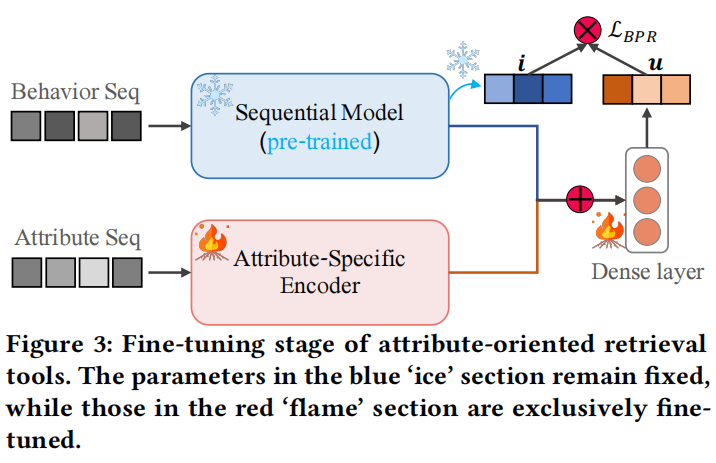




利用LLM根据用户行为历史初始化用户画像，模拟代理用户评估当前场景下的用户偏好，并决定是否需要调用外部工具来获取候选物品。

##### 属性导向工具：

**排序工具**：基于属性导向的排序指令模板，利用LLM对候选物品进行排序，以返回更符合用户潜在意图的物品集合。



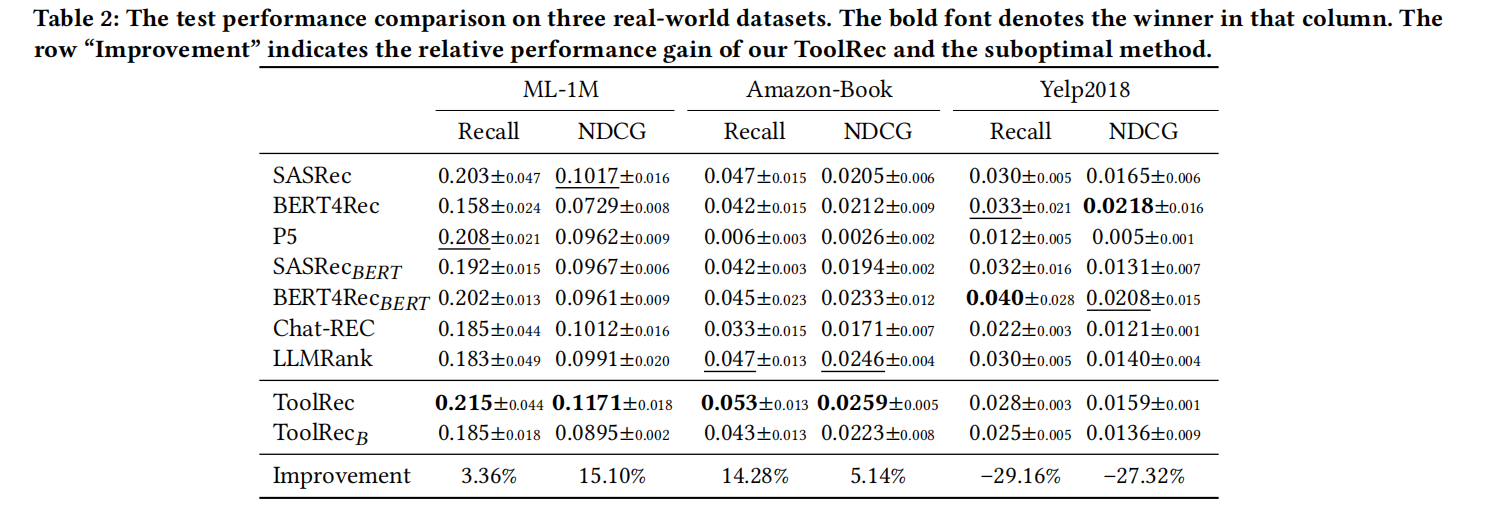
**检索工具**：采用两阶段方法，先在预训练阶段学习原始模型，再在微调阶段引入属性特定的编码器，高效返回给定属性的候选物品集合。

##### 记忆策略

海量的物品和复杂的物品名称及ID给LLM模型带来了挑战。

初始化时使用物品池目录，对外部工具（尤其是排序工具）返回的候选物品进行验证，若存在差异则补充错误信息并重新运行工具。验证后，候选物品附上工具标记，以便后续工具调用。

#### 实验

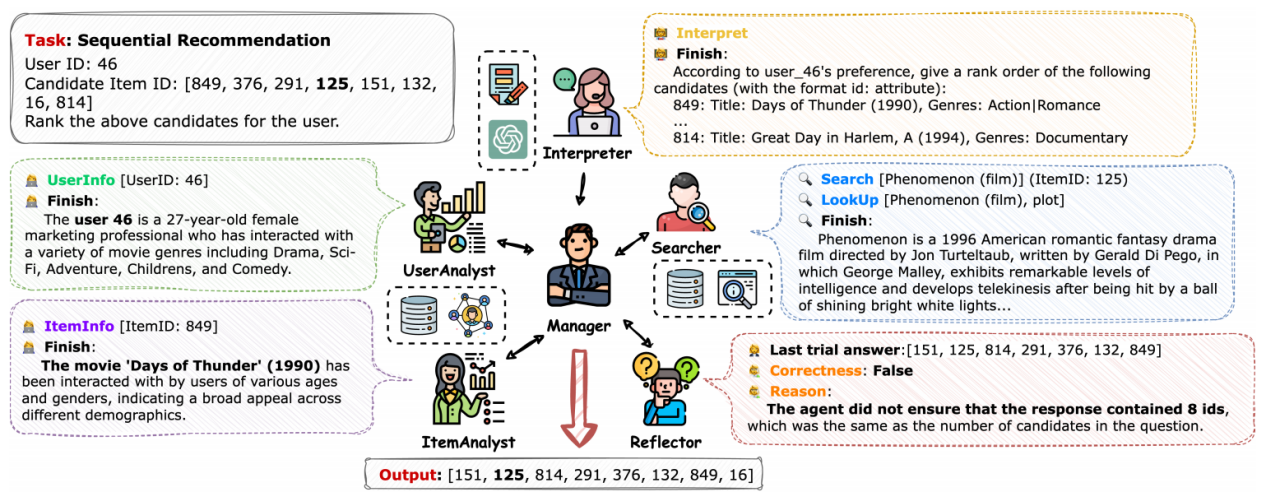


### MACRec: a Multi-Agent Collaboration Framework for Recommendation

#### 背景

在推荐系统中，虽然已有研究尝试将智能体用于模拟用户或物品行为，但这些方法大多停留在模拟层面。另一方面，单智能体方法在处理复杂推荐任务时能力有限。因此，作者提出通过多智能体协作的方式，直接解决推荐任务，以提升系统的智能化水平和任务处理能力。

#### 方法



##### Manager

负责任务规划与执行流程控制。始终以交替方式执行思考、行动和观察三个步骤。在思考阶段，对当前任务状况进行分析；在行动阶段，可选择向其他智能体寻求帮助或直接给出答案来结束任务；其他智能体的响应将在任务管理器的观察阶段呈现。

##### User/Item Analyst

分析用户偏好和物品特征。

##### Reflector

评估并反馈答案质量，提出改进建议。

##### Searcher

通过搜索工具（如维基百科）获取外部信息。

##### Task Interpreter

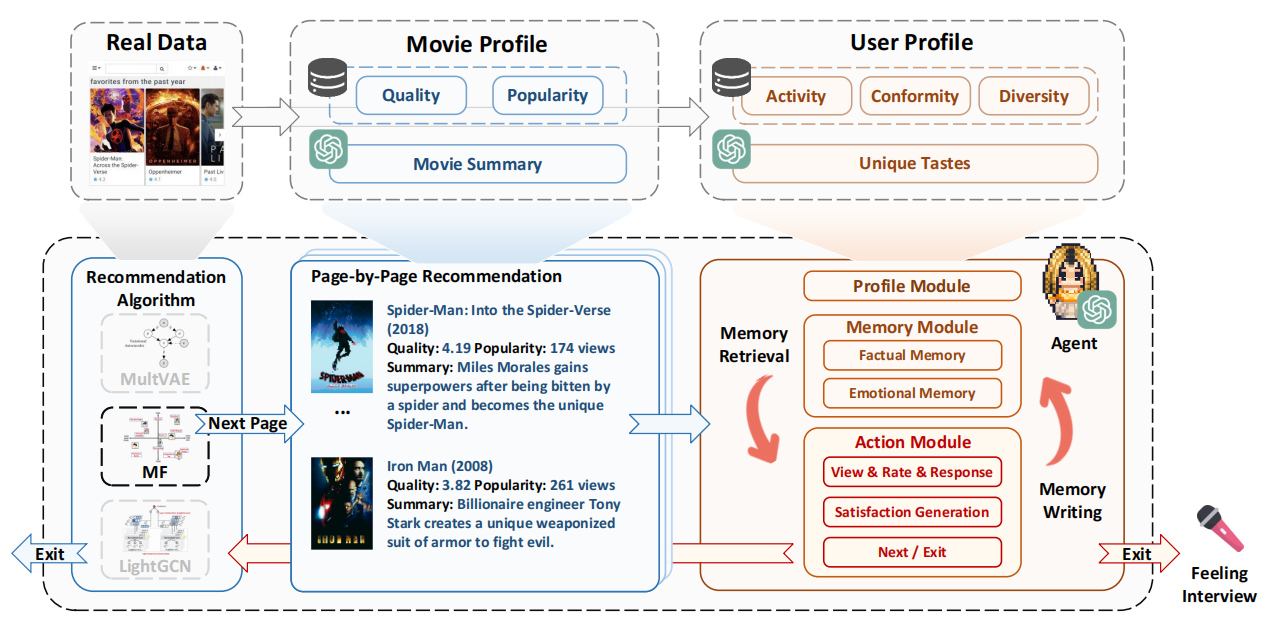
将用户输入或对话转化为可执行的任务描述。

### Agent4Rec: On Generative Agents in Recommendation

#### 背景

在推荐系统中，传统的监督式方法常面临离线指标与在线性能之间存在显著差距的问题，这阻碍了学术研究在实际推荐系统中的应用。为缓解这一瓶颈，研究者们设想构建一个可配置的推荐模拟平台，以忠实捕捉用户意图并编码人类认知机制，从而为推荐系统的数据收集、算法评估和策略开发提供新途径。近年来，随着大型语言模型（LLM）在展现人类智能方面的突破性进展，利用LLM赋能的生成式智能体开发推荐模拟器成为可能，但如何设计能忠实反映个性化用户偏好的智能体仍面临挑战。

#### 方法



##### 智能体架构设计

**用户画像模块**：基于真实世界数据集（如MovieLens-1M、Steam、Amazon-Book）初始化，捕捉用户的社会特征（活跃度、从众性、多样性）和独特偏好。

**记忆模块**：记录事实记忆（如推荐列表、用户反馈）和情感记忆（如用户满意度、疲劳度），并引入情感驱动的反思机制（当智能体行为次数超过预设阈值时，触发反思流程，自我评估对推荐内容的满意度及疲劳程度）。

**行动模块**：支持两类行为，即基于偏好的行动（如观看、评分、观看后生成情感反馈）和基于情感的行动（如退出系统、评价推荐列表、离线访谈）。

##### 推荐环境构建

**物品画像生成**：利用LLM生成物品的关键特征，如质量、热度、类型和摘要，并通过要求LLM根据物品标题分类和生成摘要，修剪可能引起幻觉的物品以降低风险。

**分页式推荐场景**：模拟真实推荐平台的分页推荐机制，每页展示若干物品，用户可基于自身偏好和情感状态进行互动。

**推荐算法设计**：内置多种协同过滤算法（如随机推荐、最受欢迎物品、矩阵分解MF、LightGCN、MultVAE），并提供开放接口以支持外部算法接入。

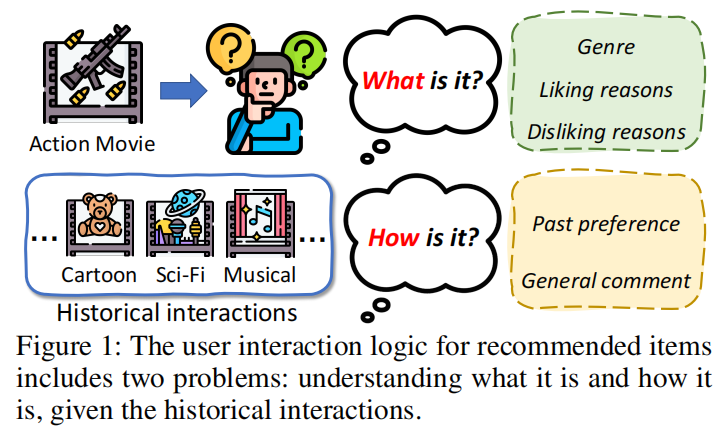
### LLM-Powered User Simulator for Recommender System

#### 背景

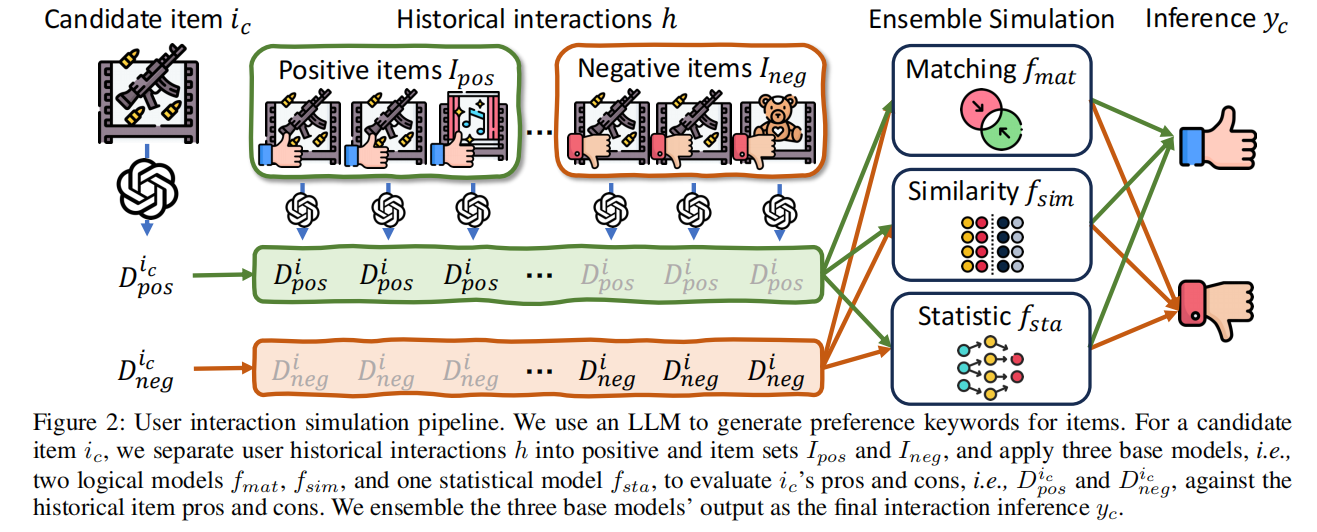
强化学习需要通过交互式训练来学习推荐策略。然而，真实用户交互数据获取困难、成本高、涉及隐私问题，因此亟需一种可模拟用户行为的用户模拟器来生成训练数据。

现有的用户模拟器存在两个主要问题：用户偏好建模不明确，难以解释用户为何喜欢或不喜欢某个项；缺乏有效的评估机制，难以衡量模拟行为与真实用户行为之间的相似性。

LLMs由于强大的语义理解与推理能力，成为模拟用户行为的潜在工具，但其存在计算成本高、幻觉问题等挑战。



#### 方法



##### LLM提取项特征

对于候选项，通过LLM生成每个项的正负面关键词，分别表示用户可能喜欢或不喜欢的理由。

对于用户交互项，通过LLM生成正面或负面关键词，代表该用户偏好。

##### 集成模型

**关键词匹配模型**：统计候选项与用户历史偏好关键词的重合度；

**语义相似度模型**：利用BERT嵌入计算语义相似性；

**统计模型**：使用SASRec等序列模型学习用户行为模式。

将上述三种模型的输出进行投票，得出最终的用户反馈（喜欢/不喜欢），用于强化学习训练。

##### 马尔可夫决策过程建模

