

Psyche-R1：通过统一的同理心、专业知识和推理实现可靠的心理大型语言模型

Chongyuan Dai¹, Jinpeng Hu¹, Hongchang Shi¹, Zhuo Li², Xun Yang³, Meng Wang^{1,4}

¹Hefei University of Technology

²The Chinese University of Hong Kong, Shenzhen

³University of Science and Technology of China

⁴Institute of Artificial Intelligence (IAI), Hefei Comprehensive National Science Center

{ taisungyun, 2024170833 } @mail.hfut.edu.cn, jinpenghu@hfut.edu.cn,

xyang21@ustc.edu.cn, eric.mengwang@gmail.com

Abstract

在合格的心理健康专业人员短缺的情况下，将大型语言模型 (LLMs) 整合到心理应用中，为缓解日益增长的心理健康障碍负担提供了一种有前途的方法。最近增强推理能力的 LLMs 在数学和编程方面取得了显著的表现，而心理领域的研究主要强调情感支持和移情对话，对于有利于生成可靠响应的推理机制关注较少。因此，在本文中，我们提出了 Psyche-R1，这是第一个中文心理 LLM，结合了移情、心理专业知识和推理能力，建立在一个新颖的数据编制管道之上。具体而言，我们设计了一个综合的数据合成管道，该管道生成了超过 75,000 个高质量的心理学问题，并配有详细的理由，理由是通过链式思维 (CoT) 推理和迭代提示-理由优化生成的，还包括 73,000 个移情对话。随后，我们采用了一种混合训练策略，其中通过多 LLM 交叉选择策略识别出具有挑战性的样本，以进行组相对策略优化 (GRPO) 以提高推理能力，而剩余的数据用于监督微调 (SFT) 以增强移情响应生成和心理领域知识。大量实验结果展示了 Psyche-R1 在多个心理基准上的有效性，我们的 7B Psyche-R1 取得了与 671B DeepSeek-R1 相当的结果。

介绍

合格心理健康专业人员的短缺激发了人们对在心理领域应用人工智能以支持心理健康辅助的兴趣不断增加 (Wolohan et al. 2018; Al Asad et al. 2019; Tanana et al. 2021)。最近，大型语言模型 (LLMs) 展示了其在广泛领域内的令人印象深刻的理解能力，得益于其卓越的文本理解能力 (Naveed et al. 2023; Zhang et al. 2023)。因此，许多基于 LLM 的研究被提出以推进心理健康服务 (Cho et al. 2023; Ye et al. 2025)。此前的研究确定了共情优化在心理咨询中的关键重要性 (Qiu et al. 2024; Sorin et al. 2024; Zhang et al. 2024)。

例如，SoulChat (Chen et al. 2023) 通过在一个大规模的多轮次同理心对话数据集上微调模型来增强同理心回应。同样地，AUGESC (Zheng et al. 2023) 通过引入一个情感感知的注意力机制来提高对话系统的情感敏感度。然而，这些方法往往缺乏心理学所需的专业基础，这对准确的心理学理解至关重要。一些研究已经尝试通过整合心理学知识来解决这一限制 (Chen, Lu, and Wang 2023; Xiao et al. 2024; Wu, Huang, and Lu 2025)。例如，PsycoLLM (Hu et al. 2024) 通过在知识问答 (QA) 对上训练其模型来整合心理学知识，而 CPsyExam (Zhao et al. 2025) 利用涵盖不同心理学相关科目理论知识的考试题来进一步提高模型表现。虽然现有研究已经取得了相当大的成功，但它们在复杂推

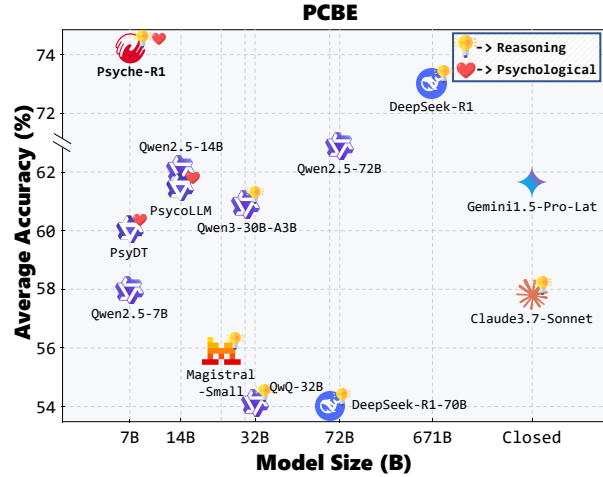


Figure 1: 不同大型语言模型在 PCBE 上的比较，以平均标准准确率与模型大小绘制图表。

理方面仍然有限。事实上，通过强化学习 (RL) 训练的推理增强型 LLM 在各个领域表现出色，尤其是在数学、代码生成和医学领域 (Chen et al. 2024; Guo et al. 2025)。然而，如图 1 所示，这些推理增强型 LLM 在心理领域表现有限，因为它们专注于逻辑推理，而忽视了超出一般常识以外的同理心和专业知识的统一。实际上，在心理领域中，推理起着至关重要的作用，因为它不仅有助于生成更准确和可靠的回答，还支持更深层次的同理心参与以及更连贯的心理知识整合。

因此，在本文中，我们提出了一种新颖的数据整理流程，并引入了 Psyche-R1，这一系统整合了共情、特定领域专业知识和推理能力。具体来说，为了构建高质量的训练语料库，我们设计了一个综合的数据合成流程，首先从各种资源生成心理学问题。接着，我们应用思维链 (CoT) 提示生成每个问题的初步详细推理理由，然后通过迭代的提示-推理优化过程，提升推理的连贯性及其与相应问题的匹配度。同时，我们合成 73k 共情对话以支持情感理解。之后，我们采用多 LLM 交叉选择策略，根据推测的复杂度将问题分类为具挑战性和非具挑战性子集。非具挑战性子集用于监督微调 (SFT) 以增强共情回应生成和领域知识，而具挑战性子集则用于利用组相对策略优化 (GRPO) 进行训练，以提高模型的推理能力，这两者共同促进了 Psyche-R1 的发展。在一系列心理学基准测试 (包括知识评估、基于案例的分

析和共情评估)上的实验结果显示, Psyche-R1 的有效性, 其中 7B Psyche-R1 显著优于同等规模的模型, 并在与显著更大规模的模型如 DeepSeek-R1 的对比中表现出具有竞争力的性能。

Psyche-R1

在本节中, 我们详细介绍了数据策划程序和两阶段训练模式, 包括数据收集、心理推理数据合成和移情对话合成。

数据收集

数据资源。 为了构建一个全面和多样化的数据集, 我们整理了广泛的资源:

- 类型 I: 经典的心理学教材和心理学项目的课程材料, 涵盖 19 个以上的子领域, 集中而系统地组织内容, 包括认知心理学、发展心理学、社会心理学等。
- 类型 II: 来自公共中文教育平台的心理学题库, 涵盖心理学的理论原则和概念知识。
- 类型 III: 从 Qwen2.5-72B-Instruct (Team 2024b) 提炼的数据, 用于补充代表性不足的子领域 (例如, 运动心理学) 并增强数据集的覆盖范围。
- 类型 IV: 从已建立的心理健康支持平台收集的对话互动, 包括一心理¹、简单心理²和知乎³, 所有这些平台都致力于为用户提供专业的心理支持。

前三种资源类型 (即类型 I、类型 II 和类型 III) 用于构建心理推理问答 (PCQA) 数据集, 旨在增强特定领域的知识获取和推理能力。类型 IV 资源用于开发共情对话数据集。

为了确保数据质量, 我们实施了几个重要的数据清洗步骤以提升质量。

心理推理数据合成

在数据清洗之后, 我们开始从整理好的心理学教科书和教学材料 (即资源类型 I) 中生成结构化的问题和相应的答案 (QA)。具体而言, 源材料首先被分割成多个文本块, 每个块设计为包含尽可能多的领域特定内容。随后, 我们利用 LLMs 根据这些文本段生成一组不同的问题及答案。同时, 对于资源类型 III, 我们采用类似的方法生成 QA, 但不进行段落级别的上下文扩充, 旨在补充在公开可获取材料中代表性不足或难以获取的心理学子领域。通过这些步骤, 我们总共获得了大约 200k 生成的 QA 对。

所有生成的问答对, 以及从类型 II 资源派生的数据被整合到一个包含约 21 万条目的一体化问答池中, 我们实施了多阶段的质量控制程序, 以确保合成数据的完整性和实用性。具体来说, 我们使用最小哈希局部敏感哈希 (LSH) 来聚类相似的问题, 并通过基于 LLM 的排序选择最佳的问题。随后, 我们给 LLM 提供少量示例, 以识别和剔除低质量的问题, 特别是那些显示信息不完整、逻辑混乱或表达不清的问题。最后, 我们邀请了 10 名本科生和研究生进行人工审核, 以消除冗余内容并降低数据集中的潜在噪声, 最终保留大约 9 万对问答对。

¹<https://www.xinli001.com/>

²<https://www.jiandanxinli.com/>

³<https://www.zhihu.com/>

生成原理 我们通过链式推理提示, 采用多步骤推理方法, 进一步为上述问题生成详细的理由, (Hsieh et al. 2023) 为模型训练提供清晰的推理路径。具体来说, 链式推理提示引导模型首先理解问题, 识别相关的心理学概念和知识, 并将问题分解为一系列分析步骤。在此过程的每个阶段, 模型都需要阐明一个中间的理由, 最终生成一个来自累积推理的最终答案。形式上, 给定一个链式推理提示 P 和一个问答对 (q_i, a_i) , 此过程产生一个理由增强的个例 $(q_i, r_i, \hat{a}_i)$, 其中 r_i 表示推理路径, $\hat{a}_i$ 表示模型预测的答案。如果预测的答案与真实结果一致 (即, $\hat{a}_i = a_i$), 我们认为 r_i 是一个有效的理由。相反, 如果预测的答案不正确 (即, $\hat{a}_i \neq a_i$), 我们指导模型重新生成理由最多到 T 次。未能在 T 次重生成尝试后产生正确预测的个例将从最终策划的数据集中删除。在获得初始理由后, 我们采用自监督优化策略, 迭代地完善提示和理由, 目标是提高他们的清晰度和可靠性。具体而言, 对于每个个例 $(P, q_i, r_i, \hat{a}_i)$, 提示 P 和理由 r_i 在多次轮次中联合更新, 使模型能够逐步改善其推理过程。每一轮优化包括两个连续步骤: 如果 $\hat{a}_i^*$ 与真实值 a_i 匹配, 我们保留 P_i^* 作为更新的提示, 并基于更新的实例 $(P_i^*, q_i, r_i^*, \hat{a}_i^*)$ 继续迭代。否则, 过程将回复到之前的提示-推理对, 以保持与正确推理路径的一致性。我们为 R 轮重复此过程 (本文中为 $R = 3$)。在完成所有迭代后, 我们评估每轮为给定问题生成的推理, 并选择质量最高的那个, 记作 $(P_i^*, q_i, r_i^*, \hat{a}_i^* = a_i)$ 。在此阶段, 我们从上一步获得的 90k 对中筛选出大约 75k 个高质量实例。

虽然前面的步骤产生了高质量的数据, 但并不是所有样本都表现出足够的复杂性以进行有效的强化学习。在此阶段, 我们采用了一种多大型语言模型 (LLM) 的交叉筛选策略, 旨在从构建的数据集中识别并隔离最具挑战性的心理学相关样本, 以便在随后的强化学习阶段使用。具体而言, 我们使用三个不同的 LLM (即 Qwen、Llama 和 Phi) 独立回答构建的心理学数据中的每个问题。对于三个模型都给出错误回答的问题, 将其聚合成一个包含 19,000 个实例的具有挑战性的子集。这个子集旨在代表极具难度的实例, 通过强化学习过程有很大的潜力提升模型的推理能力。

除了心理问答对之外, 移情被认为是有效心理健康支持的核心组成部分。为此, 我们从在线平台将富有同理心的表达加入对话语料库中, 以增强其情感丰富性和与现实世界心理互动的相关性。具体来说, 我们通过大语言模型精炼这些对话以实现以下目标。首先, 通过加入富有同理心的表达和技巧 (例如, “听到你的经历, 我希望能给你一个温暖的拥抱。”), 我们增强情感共鸣。随后, 我们确保每个对话提供基于证据的指导, 以加深对用户问题的理解, 而不仅限于浅层的同理响应。最后, 我们通过提供具体应对策略和实用步骤, 面对所呈现的具体挑战, 实现面向解决方案的支持。通过这些步骤, 我们最终获得了配备有足够同理心表达的 73k 高质量对话数据。

利用上述流程, 我们策划了一个全面的数据集, 专为心理理解、推理和共情互动而设计。该数据集包含超过 75,000 道心理学问题及其详细的推理说明, 其中 19,000 道通过多重 LLM 交叉选择被识别为具有挑战性的样本。这些具有挑战性的子集被表示为 D_{pc} , 剩余的数据被表示为 D_{pr} 。同时, 该数据集包含超过 73,000 个为了情境上适当的心理社会互动而设计的共情对话, 表示为 D_{em} 。为了进一步丰富我们的训练数据, 我们另外

引入以下数据集：

- (1) PsychoLLM 数据集 (Hu et al. 2024) 包含心理咨询的单轮问答、多轮对话和基于知识的问答。我们保留了其多轮对话和基于知识的问答部分，表示为 $\mathcal{D}_{ps}$ 。
- (2) CPsyExam 训练集 (Zhao et al. 2025) 包含来自各个心理学相关学科的超过 10,000 道考试题目。我们进一步过滤数据以去除可能的噪声项，最终得到了一套精炼的 8,000 道高质量题目，记为 $\mathcal{D}_{cp}$ 。

最终，策划的数据集被划分为两个不同的子集，以配合专门的训练目标。一个类别，用 $\mathcal{D}_{sft} = \mathcal{D}_{pr} \cup \mathcal{D}_{em} \cup \mathcal{D}_{ps}$ 表示，被指定用于 SFT。另一个类别，用 $\mathcal{D}_{grpo} = \mathcal{D}_{pc} \cup \mathcal{D}_{cp}$ 表示，保留用于 GRPO。数据合成的提示详细信息在附录 A 中提供。

为了增强在共情和专业性方面的推理能力和性能，我们采用混合训练策略。在第一阶段中，我们在 $\mathcal{D}_{sft}$ 上执行 SFT。给定输入 $x = q$ ，主要目标是训练模型生成连贯的理由 r ，随后给出相应的答案 a ，完整的输出表示为 $y = r + a$ 。我们采用 Qwen2.5-7B-Instruct (Team 2024b) 作为我们的主干模型，并进行全参数 SFT 以增强模型在共情、专业性和推理方面的性能。模型使用自回归语言建模对象进行优化，旨在最小化目标响应序列的负对数似然。损失函数形式上定义为：其中， θ 表示可训练的模型参数。

阶段 2：组相对策略优化。 在通过 SFT 获得的推理能力基础上，我们随后在 $\mathcal{D}_{grpo}$ 上实施 GRPO (Shao et al. 2024)，通过 RL 训练进一步提高心理推理能力。为了更好地指导策略学习并优化模型的推理过程，我们利用一个复合奖励函数，其具体定义如下：

- 格式奖励：为了确保输出结果的可解析性和可解释性，我们纳入了格式奖励 R_{format} 。模型需要生成一个推理过程和一个最终答案，其中推理必须包含在 $\langle \text{think} \rangle$ 和 $\langle / \text{think} \rangle$ 标签内，最终答案必须出现在 $\langle / \text{think} \rangle$ 标签之后。符合此格式的回答会得到 +1.25 的奖励，否则奖励为 -1。
- 准确性奖励：为了鼓励模型生成准确的响应，我们引入了准确性奖励 $R_{accuracy}$ 。考虑到一些问题有多个正确选项，我们采用一种灵活的评估策略，该策略考虑到预测答案 $\hat{A}$ 与标准答案 A 之间的重叠。此奖励机制鼓励模型生成推理过程，以与真实解决方案保持一致。具体的奖励函数定义如下：

$$R_{accuracy} = \begin{cases} +1, & \text{if } \hat{a} = a \\ \frac{|\hat{a} \cap a|}{|a|}, & \text{if } \hat{a} \subseteq a \text{ and } a \neq \emptyset \\ -1, & \text{otherwise} \end{cases} \quad (1)$$

我们将这两个奖励结合起来形成最终奖励 $R_{final} = R_{format} + R_{accuracy}$ 。这种公式鼓励模型产生结构良好的推理过程，同时为不完整但有效的答案提供部分积分奖励，从而通过训练期间的细致奖励信号实现更有效的策略学习。

实验设置

我们将 Psyche-R1 与四类 LLMs 进行比较，包括：(1) 通用 LLMs，在一般任务中表现出色，但缺乏显性推理能力。这些包括 MiniCPM4-8B (Team et al.

2025)，Qwen2.5-7B-Instruct, Qwen2.5-14B-Instruct, 和 Qwen2.5-72B-Instruct (Team 2024b)。(2) 增强推理能力的 LLMs，具有显性推理能力，包括 DeepSeek-R1 (Guo et al. 2025)，DeepSeek-R1-Distill-Llama-70B, QwQ-32B, Qwen3-30B-A3B, Qwen3-235B-A22B (Yang et al. 2025)，和 Magistral-Small-2506⁴。(3) 闭源 LLMs，通常代表了当前最先进的性能，包括 GPT-4o (Hurst et al. 2024)，和 Claude3.7-Sonnet⁵。(4) 心理学 LLMs，在心理学数据集上进行了微调。这些包括 CPsyCounX (Zhang et al. 2024)，EmoLLM (Team 2024a)，PsycoLLM (Hu et al. 2024)，和 PsyDT (Xie et al. 2025)。注意，对于混合推理增强模型 Qwen3 系列和 Claude3.7-Sonnet，我们将其设定为推理模式以激发其最佳性能。模型版本和参数信息的详细信息在附录 B 中提供。

实现细节

在我们的实验中，我们使用 LLaMA-Factory (Zheng et al. 2024) 进行 SFT。具体来说，我们采用 1e-5 的学习率，批量大小为 256，进行 2 轮训练。对于 GRPO 阶段，我们实施了 VeRL 框架 (Sheng et al. 2024)，学习率为 1e-6，批量大小为 128，进行 2 轮训练。所有实验均在 8 个 RTX A6000 GPU 上执行，每个配备 48GB。详细的训练超参数在附录 C 中提供。

我们对两个心理学测试基准进行了综合评估：

- 心理咨询师考试基准 (PCEB) (Hu et al. 2024)：这包括 3,863 道选择题和 100 道开放式案例分析题目，这些题目是从中国的官方国家心理咨询师考试中精选而来的。
- CPsyExam 测试集 (Zhao et al. 2025)：其中包括 4,102 道题，涵盖 39 个不同的心理学子领域。按照原始协议，我们在零样本和五样本两种设置下进行评估，为确保一致性，在五样本设置中对所有评估的模型使用相同的示例。

注意，MCQ 包含两种类型的问题：只有一个正确选项的单项选择题 (SMCQ) 和有多个正确选项的多项选择题 (MMCQ)。前述数据集通常包含两种问题类型：MCQ 和主观题。对于 MCQ，我们采用在 PsychoLLM (Hu et al. 2024) 中介绍的指标，包括标准准确度，要求预测结果与真实答案完全匹配，和弹性准确度，当预测结果是正确答案的子集时给予部分分数。对于主观题，我们使用现有的文本生成指标，包括 Rouge-1 (R-1)、Rouge-L (R-L) (Lin 2004) 和 Bleu-4 (Papineni et al. 2002)。

结果和分析

为评估不同模型的性能，我们在表格 ?? 中展示了 PCEB 的结果。这些结果揭示了几个关键观察。首先，Psyche-R1 在 MCQ 和主观题的评估任务中表现出色。这显示了我们所提出的数据集和训练策略在同时增强心理推理和心理任务文本生成能力方面的有效性。其次，尽管 DeepSeek-R1 在 MCQ 中表现优异，但其在主观题中的表现明显有限。这种性能差异可以归因于其训练方法，该方法在主要由数学和编程任务组成的具有确定性解决方案的数据集上使用 RL。虽然这种方法增强了逻辑推理能力，但似乎使模型偏向于单一答案模式，从

⁴<https://mistral.ai/>

⁵<https://claude.ai/>

Model	Zero-Shot				Five-Shot				Avg. (Zero-Shot)
	Knowledge		Case		Knowledge		Case		
	SMCQ	MMCQ	SMCQ	MMCQ	SMCQ	MMCQ	SMCQ	MMCQ	
MiniCPM4-8B	69.58	41.74	57.33	37.00	68.50	42.77	54.67	38.00	60.46
Qwen2.5-7B-Instruct	76.99	43.66	68.67	44.50	78.63	42.00	68.67	40.50	67.37
Qwen2.5-14B-Instruct	81.39	49.30	72.00	48.50	82.42	54.29	71.00	48.00	71.84
Qwen2.5-72B-Instruct	84.61	52.75	73.50	54.50	86.64	63.77	75.33	55.00	74.98
DeepSeek-R1	87.49	56.98	76.83	59.00	88.78	66.58	77.30	61.50	78.28
DeepSeek-R1-70B	76.48	22.80	61.81	19.17	76.89	40.99	62.70	37.95	60.57
GPT-4o	80.70	30.73	66.33	28.00	81.82	54.80	68.67	52.50	65.79
Gemini1.5-Pro-Latest	82.08	40.59	68.33	43.00	83.93	53.65	71.00	45.00	69.66
CPsyCounX	57.56	22.41	46.33	31.00	63.46	21.77	50.67	23.50	47.44
EmoLLM	78.41	45.33	72.50	48.00	79.92	36.88	74.17	39.50	69.32
PsycoLLM	78.33	51.98	65.33	42.00	78.63	50.45	65.57	36.00	69.20
PsyDT	80.83	48.91	69.67	41.50	81.13	40.97	68.33	40.00	70.71
Psyche-R1	82.72	61.59	70.50	49.50	83.45	61.46	76.17	52.00	74.90

Table 1: 不同模型在 CPsyExam 测试集上的比较。平均值代表整体零样本准确率。

Model	Case		Moral		Theory	
	SMCQ	MMCQ	SMCQ	MMCQ	SMCQ	MMCQ
-	47.57	31.64	87.83	59.50	78.46	42.45
SFT	56.70	41.53	92.00	75.95	83.29	59.06
SFT + RL	63.31	56.26	92.76	79.62	87.70	66.54

Table 2: 针对 PCEB 的消融研究的标准准确性。

而限制了其在开放式心理评估中生成多样化和细致入微的回答的能力。再次，现有的心理学 LLMs（例如，CPsyCounX 和 EmoLLM）在主课题中表现优异，而在 MCQ 中表现有限。这种不平衡的表现源于它们只依赖于咨询对话或富同情心的对话进行训练，从而限制了其全面能力的发展。最后，闭源模型如 GPT-4o 和 Claude3.7-Sonnet 表现相对较弱，这可能归因于其训练语料中有限的中文语言表达。

为了进一步探讨模型的表现，我们在表 1 中展示了 CPsyExam 测试集的结果。类似于之前实验中观察到的趋势，Psyche-R1 和 DeepSeek-R1 都表现出优异的性能。在这些模型中，心理 LLM 在 SMCQ 中的准确率始终高于在 MMCQ 中的准确率，因为后者需要对所有选项进行详尽的评估，要求更全面的领域知识和推理能力。在五次提示设置下，大多数模型在 MMCQ 方面取得了显著的改进（例如，PsyDT 在知识类型的 MMCQ 中提高了 47.64 %）。这一观察与现有研究一致，表明设计良好的少样本示例可以有效提升模型在某些任务中的表现。相比之下，DeepSeek-R1 在五次提示设置下的表现比其零次提示设置略有下降，这表明少样本提示可能会干扰其固有的推理能力，这与现有发现一致（Guo et al. 2025）。

我们对 SFT 和 RL 进行了消融研究，结果如表 2 所示。结果表明，SFT 在所有指标上均有显著提升，这主要归功于在我们提出的数据集上微调模型，该数据集包含移情对话和搭配合理解释的心理问题。基于 SFT 的基础，RL 训练进一步提升了模型性能，尤其是在案例任务中观察到了显著的改进。这个进步证明了 RL 训练在通过多 LLM 交叉选择识别出的挑战性例子上的有效性，从而使模型能够应对更复杂的心理场景。

除了考试任务之外，我们还评估了 Psyche-R1 在咨询任务上的表现，并将其与基础模型和几款杰出的心理学大型语言模型进行比较。按照 PsyDT 的方法，由于计算资源的限制，我们从其测试集中随机抽样 200 个项目，并用 GPT-4o（2024-05-13）作为评估者。如表 ?? 所示，Psyche-R1 相比其基础模型实现了显著提升，展现了在需要情感共鸣和认知共鸣等能力的咨询任务中的能力。这一优异表现源于两个关键要素之间的协同作用：合成的富有同情心的对话，直接提升了咨询的有效性，以及高级推理机制，能够更深入地理解问题，从而在相关背景下提供更准确且富有情感信息的回应。

案例研究。 我们展示了一个案例研究，考察了 Psyche-R1 和 Qwen2.5-72B-Instruct（Team 2024b）如何从叙述证据中得出结论，如图 2 所示。这个案例涉及一个科学家面临重大人生决策。两种模型在区分正常和神经质冲突时表现出截然不同的推理过程。Psyche-R1 首先识别出客户面临的困境，并指出客户能够识别问题并找到解决方案。随后，它呈现详细的基于证据的推理，最终通过清晰高效的推理路径得出正常冲突的结论。相比之下，Qwen2.5-72B-Instruct 将客户对他人看法的担忧误解为神经质认知，这源于过度强调孤立细节而非全面推理。该案例表明，全面的基于证据的推理对于准确的心理推论至关重要。这在 Psyche-R1 中可以观察到，该模型在整个叙述中进行了详细的事实推理，而不仅仅是孤立地关注特定描述。

相关工作

用于心理学的大型语言模型。 大型语言模型（LLMs）的成功激发了基于 LLM 的心理健康应用的开发兴趣（Demszky et al. 2023）。早期研究主要集中在提高心理健康服务的可及性。此阶段的研究集中在两个方向：一个方向是利用自然语言处理（NLP）技术进行情感识别，以实现抑郁（Huang et al. 2019）和自杀意念（Lee et al. 2020）的自动检测。另一个方向则着重于通过对单轮（Lai et al. 2023）或多轮（Qiu et al. 2024）对话数据微调 LLMs，构建具有同理心的对话系统，以增强其情感支持和理解能力（Team 2024a; Xie et al. 2025）。随着研究的进展，研究人员开始探索更多样化的心理健康应用。一些研究已经将传统心理测量工具（如心理量

Case Question 一般资料: 求助者, 女性, 36岁, 博士, 研究员 案例介绍: 求助者留学归来进入某科研单位……半年前被其他同事批评, 感到很生气, 认为别人不应该与自己计较……因此, 想到国外工作。但是, 求助者丈夫不愿出国, 家庭开始出现矛盾, 求助者内心痛苦, 情绪不好, 体重下降……经体检, 没有发现器质性病变……问题: 该求助者心理冲突的类型包括 ()。A. 变形 B. 趋避式 C. 常形 D. 双趋式 General Information: Client, female, 36 years old, PhD, researcher. Case Introduction: The client returned from studying abroad and joined a research institution... Six months ago, she was criticized by colleagues and felt very angry, believing that others should not be so petty with her... Therefore, she considered working abroad. However, her husband was unwilling to go broad, leading to marital conflicts. The client experienced inner distress and emotional difficulties, with weight loss... Physical examination revealed no organic pathology... Question: The type(s) of psychological conflict include: A. Neurotic B. Approach-avoidance C. Normal D. Approach-approach	Psyche-R1 <think>这种心理冲突的特点在于求助者在追求某一目标 (出国工作) 的同时, 又受到另一目标 (维护家庭关系) 的阻碍, 形成了趋避式心理冲突.....符合常形冲突, 因为求助者在面对现实问题时, 能够清晰地认识到问题的存在, 并试图寻找解决方案.....也符合趋避式冲突, 因为求助者在追求某一目标 (出国工作) 的同时, 又受到另一目标 (维护家庭关系) 的阻碍, 形成了趋避式心理冲突。<think>答案: BC <think> ...The characteristic of this psychological conflict is that the client, while pursuing one goal (working abroad), is simultaneously hindered by another goal (maintaining family relationships), forming an approach-avoidance conflict... It fits the pattern of normal conflict, because when facing real-world problems, the client can clearly recognize the existence of the problem and attempts to find solutions. It also fits the approach-avoidance conflict pattern, because the client, while pursuing one goal (working abroad), is simultaneously hindered by another goal (maintaining family relationships), forming an approach-avoidance psychological conflict. <think> Answer: BC	Qwen2.5-72B-Instruct 该求助者的心理冲突主要表现为对出国工作的强烈愿望与家庭因素 (如丈夫不愿出国) 之间的矛盾, 因此属于典型的趋避式冲突。此外, 求助者认为同事应该以科研为重, 却在意别人对自己态度的看法, 这显示出了她的认知出现了变形, 导致了内心的矛盾。综上所述, 该求助者心理冲突的类型包括: B. 趋避式 A. 变形 所以正确答案是 B 和 A。 This client's psychological conflict is primarily manifested in the contradiction between their strong desire to work abroad and family factors (such as their husband's unwillingness to go abroad), which represents a typical approach-avoidance conflict. Additionally, the client believes that colleagues should prioritize scientific research, yet is concerned about others' perceptions of her own attitude, which indicates that their cognition has become neurotic, leading to internal contradictions. In summary, this client's types of psycho-logical conflict include: B. Approach-avoidance, A. Neurotic. Therefore, correct answers are B and A.
--	--	--

Figure 2: 来自 CPsychExam 测试集的一个定性示例, 比较 Psyche-R1 和 Qwen2.5-72B-Instruct。

表) 转变为互动系统, 以提高用户参与度 (Kuribayashi, Oseki, and Baldwin 2024; Yang et al. 2024)。另一研究方向专注于心理领域的专业需求, 基于既定的心理治疗 (Lee et al. 2024; Shen et al. 2024) 或概念开发专业背景的心理健康应用 (Zhang et al. 2025)。近年来, 包括 CoT 提示 (Wei et al. 2022; Hsieh et al. 2023) 在内的技术显著推动了 LLM 推理的发展。在此基础上, 研究人员探索了更复杂的推理架构。例如, Tree of Thoughts (Yao et al. 2023) 通过自评估系统地探索多种推理路径, 而 PAL (Gao et al. 2023) 通过程序生成将推理与外部工具集成。这些方法进一步提升了模型在处理复杂任务时的性能。随着 OpenAI o1 (Jaech et al. 2024) 和 DeepSeek-R1 (Guo et al. 2025) 等推理 LLM 的发布, 标志着一个新的突破。这些模型通过强化学习训练以增强推理能力, 在数学和编程任务中表现出色 (Comanici et al. 2025; Yang et al. 2025)。受这些进展的启发, 研究人员采用了先进的 RL 算法, 如 GRPO (Shao et al. 2024) 和 DAPO (Yu et al. 2025), 将推理能力扩展到域特定的应用中, 包括医学 (Liu et al. 2025) 和金融 (Zhu et al. 2025)。然而, 在心理学领域, 关于推理实用性的研究仍然有限。据我们所知, Psyche-R1 是第一个统一了同理心、领域特定专业知识和推理能力的心理学 LLM。

在本文中, 我们提出了 Psyche-R1, 这是第一个将同理心、专业知识和推理相结合的中文心理大模型。为了支持模型开发, 我们设计了一个多阶段的数据合成流程, 生成具有详细推理和同理心对话的高质量心理推理样本。推理的理由通过迭代的提示-理由优化得到了进一步增强, 并采用多大模型的交叉选择策略来识别具有挑战性的样本。最后, 具有挑战性的小集合用于 GRPO, 而其余数据用于 SFT, 共同促进最终模型的发展。大量实验表明, Psyche-R1 优于现有的心理大模型, 其性能可以媲美 DeepSeek-R1。

References

- Al Asad, N.; Pranto, M. A. M.; Afreen, S.; and Islam, M. M. 2019. Depression detection by analyzing social media posts of user. In 2019 IEEE international conference on signal processing, information, communication & systems (SPICSCON), 13–17. IEEE.
- Chen, J.; Cai, Z.; Ji, K.; Wang, X.; Liu, W.; Wang, R.; Hou, J.; and Wang, B. 2024. Huatuoqpt-o1, towards

medical complex reasoning with llms. arXiv preprint arXiv:2412.18925.

Chen, Y.; Xing, X.; Lin, J.; Zheng, H.; Wang, Z.; Liu, Q.; and Xu, X. 2023. SoulChat: Improving LLMs' Empathy, Listening, and Comfort Abilities through Fine-tuning with Multi-turn Empathy Conversations. In Findings of the Association for Computational Linguistics: EMNLP 2023, 1170–1183.

Chen, Z.; Lu, Y.; and Wang, W. 2023. Empowering Psychotherapy with Large Language Models: Cognitive Distortion Detection through Diagnosis of Thought Prompting. In Findings of the Association for Computational Linguistics: EMNLP 2023, 4295–4304.

Cho, Y.; Kim, M.; Kim, S.; Kwon, O.; Kwon, R. D.; Lee, Y.; and Lim, D. 2023. Evaluating the efficacy of interactive language therapy based on LLM for high-functioning autistic adolescent psychological counseling. arXiv preprint arXiv:2311.09243.

Comanici, G.; Bieber, E.; Schaekermann, M.; Pasupat, I.; Sachdeva, N.; Dhillon, I.; Blistein, M.; Ram, O.; Zhang, D.; Rosen, E.; et al. 2025. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. arXiv preprint arXiv:2507.06261.

Demszky, D.; Yang, D.; Yeager, D. S.; Bryan, C. J.; Clapper, M.; Chandhok, S.; Eichstaedt, J. C.; Hecht, C.; Jamieson, J.; Johnson, M.; et al. 2023. Using large language models in psychology. Nature Reviews Psychology, 2(11): 688–701.

Gao, L.; Madaan, A.; Zhou, S.; Alon, U.; Liu, P.; Yang, Y.; Callan, J.; and Neubig, G. 2023. Pal: Program-aided language models. In International Conference on Machine Learning, 10764–10799. PMLR.

Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X.; et al. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. arXiv preprint arXiv:2501.12948.

Hsieh, C.-Y.; Li, C.-L.; YEH, C.-K.; Nakhost, H.; Fujii, Y.; Ratner, A. J.; Krishna, R.; Lee, C.-Y.; and Pfister, T. 2023. Distilling Step-by-Step! Outperforming Larger Language Models with Less Training Data and Smaller

- Model Sizes. In The 61st Annual Meeting Of The Association For Computational Linguistics.
- Hu, J.; Dong, T.; Gang, L.; Ma, H.; Zou, P.; Sun, X.; Guo, D.; Yang, X.; and Wang, M. 2024. Psycollm: Enhancing llm for psychological understanding and evaluation. *IEEE Transactions on Computational Social Systems*.
- Huang, Z.; Epps, J.; Joachim, D.; and Sethu, V. 2019. Natural language processing methods for acoustic and landmark event-based features in speech-based depression detection. *IEEE Journal of selected topics in Signal Processing*, 14(2): 435–448.
- Hurst, A.; Lerer, A.; Goucher, A. P.; Perelman, A.; Ramesh, A.; Clark, A.; Ostrow, A.; Welihinda, A.; Hayes, A.; Radford, A.; et al. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.
- Jaech, A.; Kalai, A.; Lerer, A.; Richardson, A.; El-Kishky, A.; Low, A.; Helyar, A.; Madry, A.; Beutel, A.; Carney, A.; et al. 2024. Openai o1 system card. *arXiv preprint arXiv:2412.16720*.
- Kuribayashi, T.; Oseki, Y.; and Baldwin, T. 2024. Psychometric Predictive Power of Large Language Models. In *Findings of the Association for Computational Linguistics: NAACL 2024*, 1983–2005.
- Lai, T.; Shi, Y.; Du, Z.; Wu, J.; Fu, K.; Dou, Y.; and Wang, Z. 2023. Supporting the demand on mental health services with AI-based conversational large language models (LLMs). *BioMedInformatics*, 4(1): 8–33.
- Lee, D.; Park, S.; Kang, J.; Choi, D.; and Han, J. 2020. Cross-lingual suicidal-oriented word embedding toward suicide prevention. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2208–2217.
- Lee, S.; Mac Kim, S.; Kim, M.; Kang, D.; Yang, D.; Kim, H.; Kang, M.; Jung, D.; Kim, M.; Lee, S.; et al. 2024. Cactus: Towards Psychological Counseling Conversations using Cognitive Behavioral Theory. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, 14245–14274.
- Lin, C.-Y. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, 74–81.
- Liu, C.; Wang, H.; Pan, J.; Wan, Z.; Dai, Y.; Lin, F.; Bai, W.; Rueckert, D.; and Arcucci, R. 2025. Beyond distillation: Pushing the limits of medical llm reasoning with minimalist rule-based rl. *arXiv preprint arXiv:2505.17952*.
- Naveed, H.; Khan, A. U.; Qiu, S.; Saqib, M.; Anwar, S.; Usman, M.; Akhtar, N.; Barnes, N.; and Mian, A. 2023. A comprehensive overview of large language models. *ACM Transactions on Intelligent Systems and Technology*.
- Papineni, K.; Roukos, S.; Ward, T.; and Zhu, W.-J. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 311–318.
- Qiu, H.; He, H.; Zhang, S.; Li, A.; and Lan, Z. 2024. SMILE: Single-turn to Multi-turn Inclusive Language Expansion via ChatGPT for Mental Health Support. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, 615–636.
- Shao, Z.; Wang, P.; Zhu, Q.; Xu, R.; Song, J.; Bi, X.; Zhang, H.; Zhang, M.; Li, Y.; Wu, Y.; et al. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.
- Shen, H.; Li, Z.; Yang, M.; Ni, M.; Tao, Y.; Yu, Z.; Zheng, W.; Xu, C.; and Hu, B. 2024. Are Large Language Models Possible to Conduct Cognitive Behavioral Therapy? In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 3695–3700. IEEE.
- Sheng, G.; Zhang, C.; Ye, Z.; Wu, X.; Zhang, W.; Zhang, R.; Peng, Y.; Lin, H.; and Wu, C. 2024. HybridFlow: A Flexible and Efficient RLHF Framework. *arXiv preprint arXiv: 2409.19256*.
- Sorin, V.; Brin, D.; Barash, Y.; Konen, E.; Charney, A.; Nadkarni, G.; and Klang, E. 2024. Large language models and empathy: systematic review. *Journal of medical Internet research*, 26: e52597.
- Tanana, M. J.; Soma, C. S.; Kuo, P. B.; Bertagnolli, N. M.; Dembe, A.; Pace, B. T.; Srikumar, V.; Atkins, D. C.; and Imel, Z. E. 2021. How do you feel? Using natural language processing to automatically rate emotion in psychotherapy. *Behavior research methods*, 53(5): 2069–2082.
- Team, E. 2024a. EmoLLM: Reinventing Mental Health Support with Large Language Models. <https://github.com/SmartFlowAI/EmoLLM>.
- Team, M.; Xiao, C.; Li, Y.; Han, X.; Bai, Y.; Cai, J.; Chen, H.; Chen, W.; Cong, X.; Cui, G.; et al. 2025. MiniCPM4: Ultra-Efficient LLMs on End Devices. *arXiv preprint arXiv:2506.07900*.
- Team, Q. 2024b. Qwen2.5: A Party of Foundation Models.
- Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Xia, F.; Chi, E.; Le, Q. V.; Zhou, D.; et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35: 24824–24837.
- Wolohan, J.; Hiraga, M.; Mukherjee, A.; Sayyed, Z. A.; and Millard, M. 2018. Detecting linguistic traces of depression in topic-restricted text: Attending to self-stigmatized depression with NLP. In *Proceedings of the first international workshop on language cognition and computational models*, 11–21.
- Wu, S.; Huang, X.; and Lu, D. 2025. Psychological health knowledge-enhanced LLM-based social network crisis intervention text transfer recognition method. In *Proceedings of the 2025 International Conference on Health Big Data*, 156–161.

- Xiao, M.; Xie, Q.; Kuang, Z.; Liu, Z.; Yang, K.; Peng, M.; Han, W.; and Huang, J. 2024. HealMe: Harnessing Cognitive Reframing in Large Language Models for Psychotherapy. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 1707–1725.
- Xie, H.; Chen, Y.; Xing, X.; Lin, J.; and Xu, X. 2025. PsyDT: Using LLMs to Construct the Digital Twin of Psychological Counselor with Personalized Counseling Style for Psychological Counseling. In Che, W.; Nabende, J.; Shutova, E.; and Pilehvar, M. T., eds., Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 1081–1115. Vienna, Austria: Association for Computational Linguistics. ISBN 979-8-89176-251-0.
- Yang, A.; Li, A.; Yang, B.; Zhang, B.; Hui, B.; Zheng, B.; Yu, B.; Gao, C.; Huang, C.; Lv, C.; et al. 2025. Qwen3 technical report. arXiv preprint arXiv:2505.09388.
- Yang, Q.; Wang, Z.; Chen, H.; Wang, S.; Pu, Y.; Gao, X.; Huang, W.; Song, S.; and Huang, G. 2024. PsychoGAT: A Novel Psychological Measurement Paradigm through Interactive Fiction Games with LLM Agents. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 14470–14505.
- Yao, S.; Yu, D.; Zhao, J.; Shafran, I.; Griffiths, T.; Cao, Y.; and Narasimhan, K. 2023. Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems*, 36: 11809–11822.
- Ye, J.; Xiang, L.; Zhang, Y.; and Zong, C. 2025. SweetieChat: A Strategy-Enhanced Role-playing Framework for Diverse Scenarios Handling Emotional Support Agent. In Proceedings of the 31st International Conference on Computational Linguistics, 4646–4669.
- Yu, Q.; Zhang, Z.; Zhu, R.; Yuan, Y.; Zuo, X.; Yue, Y.; Dai, W.; Fan, T.; Liu, G.; Liu, L.; et al. 2025. Dapo: An open-source llm reinforcement learning system at scale. arXiv preprint arXiv:2503.14476.
- Zhang, C.; Li, R.; Tan, M.; Yang, M.; Zhu, J.; Yang, D.; Zhao, J.; Ye, G.; Li, C.; and Hu, X. 2024. CPsyCoun: A Report-based Multi-turn Dialogue Reconstruction and Evaluation Framework for Chinese Psychological Counseling. In Findings of the 62nd Annual Meeting of the Association for Computational Linguistics, ACL 2024, 13947–13966. Association for Computational Linguistics (ACL).
- Zhang, H.; Chen, J.; Jiang, F.; Yu, F.; Chen, Z.; Chen, G.; Li, J.; Wu, X.; Zhiyi, Z.; Xiao, Q.; et al. 2023. HuatuoGPT, Towards Taming Language Model to Be a Doctor. In Findings of the Association for Computational Linguistics: EMNLP 2023, 10859–10885.
- Zhang, J.; He, H.; Ma, L.; Song, N.; He, S.; Zhang, S.; Qiu, H.; Zhou, Z.; Li, A.; Dai, Y.; et al. 2025. ConceptPsy: A comprehensive benchmark suite for hierarchical psychological concept understanding in LLMs. *Neurocomputing*, 637: 130070.
- Zhao, J.; Zhu, J.; Tan, M.; Yang, M.; Li, R.; Di, Y.; Zhang, C.; Ye, G.; Li, C.; Hu, X.; et al. 2025. CPsyExam: A Chinese Benchmark for Evaluating Psychology using Examinations. In Proceedings of the 31st International Conference on Computational Linguistics, 11248–11260.
- Zheng, C.; Sabour, S.; Wen, J.; Zhang, Z.; and Huang, M. 2023. AugESC: Dialogue Augmentation with Large Language Models for Emotional Support Conversation. In Findings of ACL.
- Zheng, Y.; Zhang, R.; Zhang, J.; Ye, Y.; and Luo, Z. 2024. LlamaFactory: Unified Efficient Fine-Tuning of 100+ Language Models. In Cao, Y.; Feng, Y.; and Xiong, D., eds., Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations), 400–410. Bangkok, Thailand: Association for Computational Linguistics.
- Zhu, J.; Chen, Q.; Dou, H.; Li, J.; Guo, L.; Chen, F.; and Zhang, C. 2025. Dianjin-r1: Evaluating and enhancing financial reasoning in large language models. arXiv preprint arXiv:2504.15716.