【Octopus：通过动态对比解码减轻幻觉现象】

——【Octopus: Alleviating Hallucination via Dynamic Contrastive Decoding】

1 相关资源

pdf: https://arxiv.org/abs/2503.00361
ppt:
短视频:
数据集:
源码: https://github.com/LijunZhang01/Octopus
网站:
【除了网站，其他资源尽量下载】

2 论文属性

论文来源: CVPR 2025
【给出具体会议名称和年限，不要仅仅写 ACM，IEEE】
论文类别: Large Language Model
【论文的类别，比如移动计算、轨迹处理、深度学习等】
论文关键字: LLMs, Hallucination Mitigation
推荐程度：3 （其他说明可标注）
(5 非常棒，建议认真研读、小组讨论和复现；4 好，建议细读，考虑复现；3 可以，部分内容值得注意；2 一般，简单浏览即可；1 没有意义，不建议阅读)

3 工作团队

作者: Wei Suo, Lijun Zhang, Mengyang Sun, Lin Yuanbo Wu, Peng Wang, Yanning Zhang
单位:

1. School of Computer Science
2. Ningbo Institute
3. School of Cybersecurity Northwestern Polytechnical University, China
4. National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean
5. Department of Computer Science, Swansea University, United Kingdom

团队情况描述:

4 论文介绍

(1) 研究目的

【研究背景是什么？本文工作有什么用?】

大型视觉-语言模型（LVLM）已能在图像描述、视觉问答等任务中生成流畅且上下文相关的文本，但在医疗、自动驾驶、机器人等高可靠性场景中，模型经常“脑补”出图像中并不存在的物体（即对象幻觉），导致错误决策。

本文的工作旨在缓解 LVLMs 中的幻觉问题，提高模型生成内容的准确性和可靠性，从而增强用户对模型的信任，使其能够更好地应用于实际场景。

(2) 研究现状

【当前的最好研究做到什么程度了？存在的问题是什么？这里采信论文的说法，可以给出自己的评点】

当前的研究主要分为两大类：数据驱动的重训练方法和对比解码 CD 方法。数据驱动的方法通过构建高质量的指令调整数据并重新训练模型来抑制幻觉，但这种方法需要复杂的数据构建过程和高昂的成本，且对于已经部署的模型无法进行额外训练。CD 方法则通过比较原始输入和扰动输入的输出分布来缓解幻觉问题，无需改变模型权重，但现有的 CD 方法大多采用一刀切的策略，对所有样本和生成步骤都使用相同的扰动方式。这种单一的 CD 策略无法有效解决不同类型的幻觉问题，因为幻觉的成因是复杂的混合体，不同的样本和生成步骤面临不同的幻觉挑战。因此，现有的方法在处理幻觉问题时存在局限性，无法有效识别和解决不同类型的幻觉。

(3) 本文解决的问题

【一句话概括本文解决的核心问题】

本文的核心问题是提出一种能够自适应识别幻觉类型并动态构建对比解码工作流程的方法，以有效缓解 LVLMs 中的幻觉问题。

(4) 创新与优势

【本文的创新之处是什么？新场景？新发现？新视角？新方法？请明确指出】

【本文工作的贡献或优点是什么?】

1. 揭示了幻觉产生机制的复杂性和混合性，不同样本（或 token）面临多种形式的幻觉挑战。
2. 提出了 Octopus 框架，该框架能自适应识别幻觉类型，并构建动态对比解码流程以修正虚假内容。
3. 在四个基准测试集上，Octopus 框架在生成式任务和判别式任务中均达到当前最优性能，同时展现出优异的可部署性和可扩展性。

(5) 解决思路

【本文是怎样解决问题的？包括方法、技术、模型等，以自己理解的方式表述清楚】

实验结果表明，一刀切的方法难以有效纠正不同类型的幻觉。因此，一个自然的想法是结合多种 CD 方法，以减少来自不同来源的幻觉。然而，在没有明确定义的标签的情况下，为不同的输入样本确定最佳策略具有挑战性。此外，令牌生成具有顺序依赖性，并且涉及巨大的解空间，这使得在每个生成步骤中选择最佳的 CD 方法变得困难。

为了解决上述问题，作者提出了一个简单但有效的框架，名为 Octopus。与以往的研究不同，该方法着重于引导模型动态组织对比解码流程，并根据不同的输入选择合适的对比解码（CD）策略。具体

来说，首先构建了一个基于 Transformer 的模块和一个可学习的令牌，用于自适应地识别幻觉类型，这类似于章鱼的眼睛。根据不同的判断结果，每个对比解码策略都被视为一条“触手”，执行特定的对比操作。最后，借助直接偏好优化（DPO）或强化学习，Octopus 可以很容易地进行优化。得益于上述设计，所提出的方法不仅能有效减少幻觉内容，而且由于避免了对大型视觉语言模型（LVLMs）权重的重新训练，在部署时具有良好的扩展性。更重要的是，作为一个通用框架下，后续的 CD 方法可以无缝集成，无需额外调整。

样本级幻觉

作者首先进行了两种实验（即生成任务和判别任务）来回答第一个问题。对于生成任务，在三个广泛使用的数据集上开展实验：AMBER、ObjectHalBench（使用的语言提示为“请详细描述这张图片”），以及 MMHalbench 使用原始指令作为提示）。分别采用 VCD、M3ID、AVISC 这三种策略来干预 LLaVA 对每个样本的输出分布。通过上述指标，如果某种策略相较于 LLaVA 的原始输出表现更优，则认为该策略是有效的。如图（a）所示报告了相应的百分比，其中橙色、绿色和蓝色分别表示“仅一种 CD 策略有效”、“两种 CD 策略均有效”和“三种 CD 策略均有效”。通过对比这些结果，我们发现大量样本（约 60%）只能通过特定的 CD 策略来处理，且这些策略的重叠度相对较低（约 10 同样的，对于判别任务，如图 2（b）所示报告了相应的百分比，观察到每种 CD 策略都针对一部分样本，并且只有约 10% 的案例在三种方法中都有效。基于上述结果，得出结论：每种 CD 方法仅对特定的幻觉样本有效，对所有案例使用单一策略将不可避免地导致次优结果。

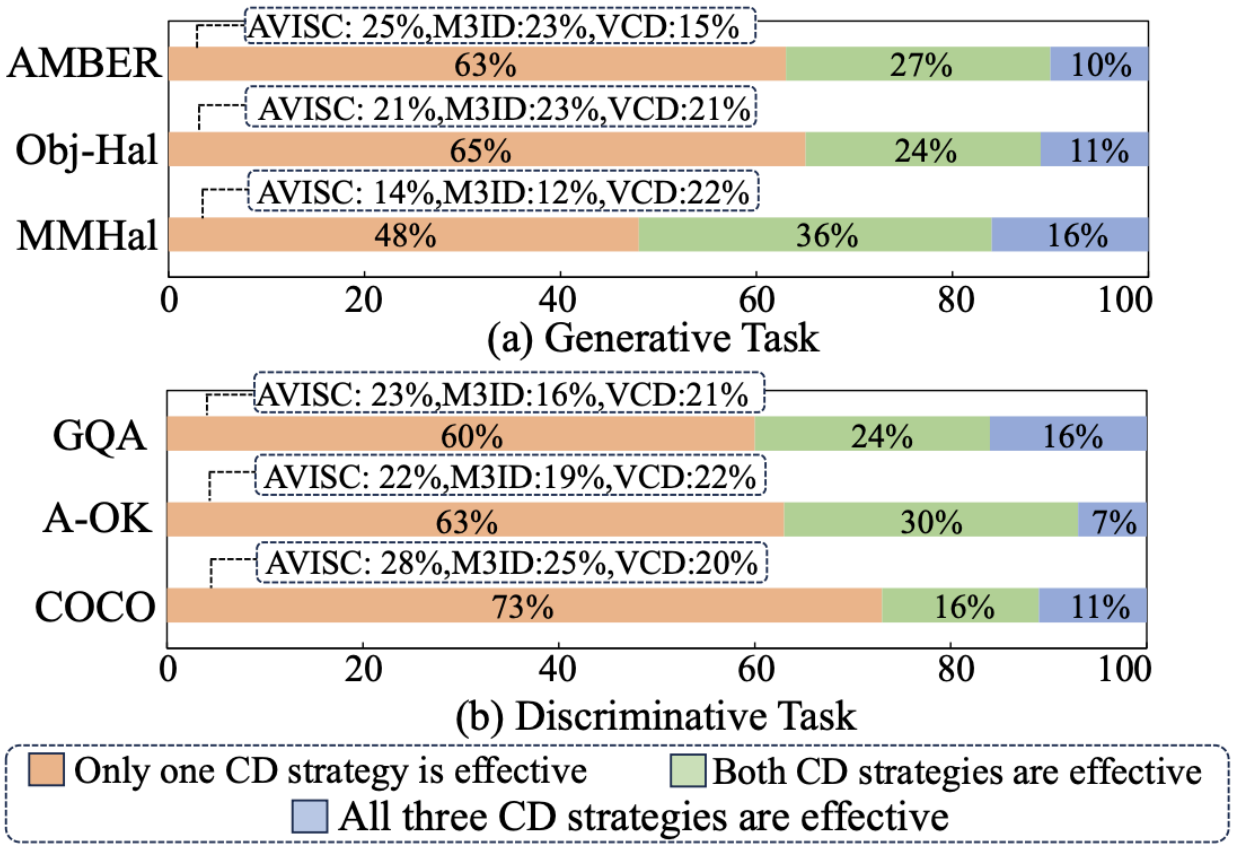

图 1: The proportion of effective samples using different CD methods for (a) Generative Task and (b) Discriminative Task.

token 级幻觉

上述样本层面的分析表明，每个样本都需要采用特定的 CD 策略。在此，作者关注一个更细粒度的场景：生成过程中的每个时间步是否受到相同的幻觉成因影响。

为了应对这一问题，作者在 AMBER 数据集上设计了实验，采用 CHAIR、Cog 和 Hal 三个指标，运用枚举法来评估生成过程中的幻觉成因。此外，即便只有三个 CD 候选方案，由于输出内容较长，其组合空间仍然十分庞大。为了减少组合数量，枚举空间将仅考虑每个描述中的前三个幻觉名词。如图所示，作者用“策略 1”“策略 2”和“策略 3”分别表示三种幻觉缓解策略。同时，展示了这些组合中的最佳分数。以“策略 1+3”为例，三个标记中的每一个都有两种可选的幻觉消除策略（即策略 1 和策略 3），因此总共有 6 种组合。在这些组合中，通过比较这些分数，作者发现利用多种 CD 策略能更好地抑制幻觉。

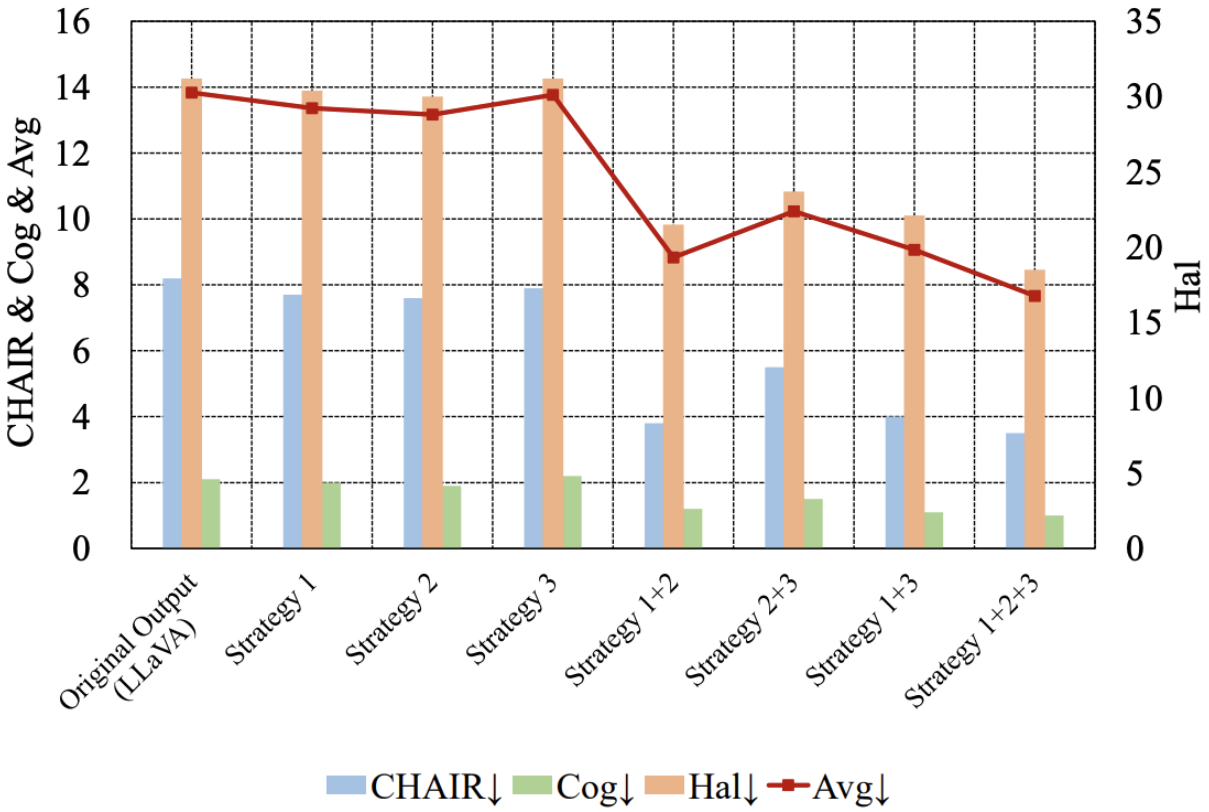

图 2: Token-level hallucination quantitative evaluation.

为了探究幻觉产生的本质，作者还进行了定性分析，以检查每个预测标记的注意力分布。如图所示，可以发现这句话中的幻觉词包括“坐着、躺着和人”，每个标记对应着不同的幻觉成因。例如，“坐着”聚焦于视觉盲区标记“IM3”，这表明当前步骤受到了注意力偏差的影响。“躺着”的出现主要是由于对视觉信息的注意力不足。而“人”仅专注于语言标记，这表明它受到了语言先验的影响。因此，作者得出结论：幻觉的成因是混合的，每个生成步骤都面临着不同形式的挑战。

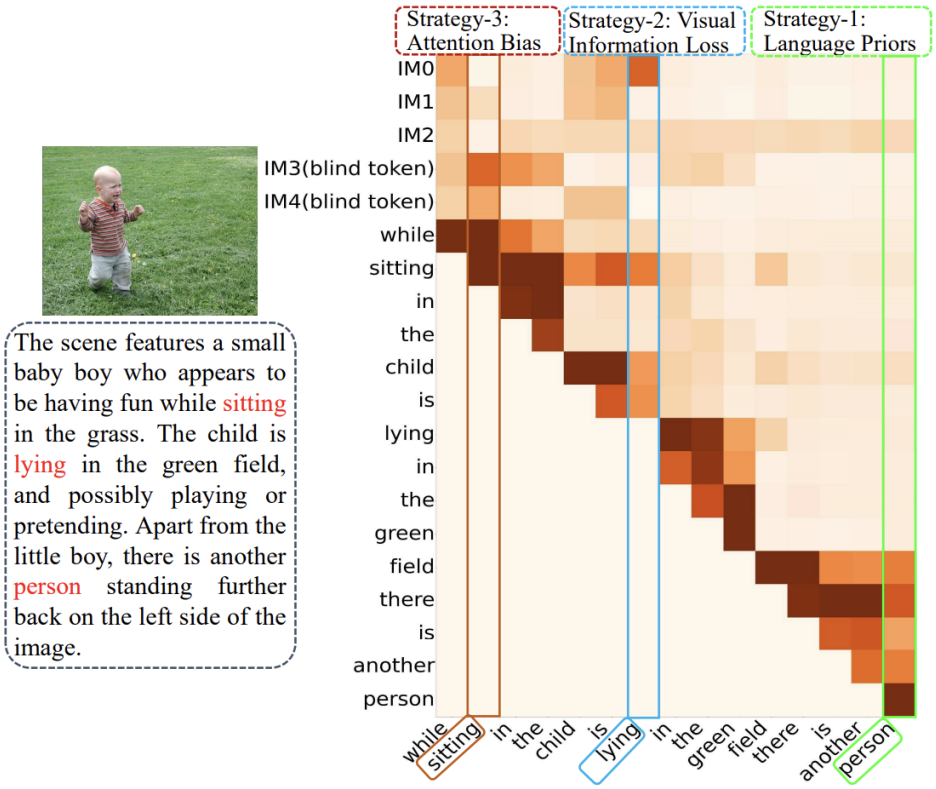

图 3: Token-level hallucination qualitative analysis.

根据上述实验可以发现，各种幻觉因素共同导致了样本层面和 token 层面的错误输出。因此，一个自然的想法是将这些现成的 CD 策略结合起来作为纠正方法，以应对不同类型的幻觉。

Octopus-模型结构

给定视觉输入 v 和文本指令 q（例如，“请详细描述这张图像”），利用章鱼的眼睛在每个生成步骤中识别幻觉的类型。然后，这些章鱼的触手通过对比解码执行特定策略。具体来说，首先构建一个基于变压器的模块 $\mathcal{O}_\phi$，其中 ϕ 表示变压器结构的参数。可以发现每个时间步 y_t 将受到 v、q 和 $y_{<t}$ 的共同影响。因此，这些来自 LVLMs 的隐藏状态（即 v、q 和 $y_{<t}$）将被输入到 $\mathcal{O}_\phi$ 中，决策标记 eye $\in \mathbb{R}^d$，其中 d 是隐藏状态的维度。为简单起见，使用 $H_t = \{h_i\}_{i=1}^t$ 表示生成步骤 t 之前的序列，其中 $h_i \in \mathbb{R}^d$ 是每个标记的隐藏状态。虽然可学习的标记 eye 可以被视为“章鱼的眼睛”。正式地，上述计算可以表述为：

$$[h_{\text{eye}}^t; H_t'] = \mathcal{O}_\phi\left(\text{concat}[\text{eye}; H_t] + E_{\text{pos}}\right), \tag{1}$$

其中 h_{eye}^t 和 H_t' 分别是来自 eye 和 H_t 序列的相应输出。E_{pos} 和 concat 表示位置嵌入和连接操作。得益于自注意力机制，h_{eye}^t 能够自适应地聚合来自其他隐藏状态的信息。

然后，利用一个轻量级简单的多层感知机（MLP）将 h_{eye}^t 映射到动作向量 $h_{\text{act}}^t \in \mathbb{R}^k$，其中 k 是候选策略的数量。在本文中，在每个步骤构建四个动作空间，包括策略 1-3（即 VCD、M3ID 和 AVISC）和一个空动作（即不执行 CD 策略）。这里，用“触手”来表示这些候选 CD 动作。对于每个 h_{act}^t，动作向量 a_t 通过以下方式获得：

$$h_{\text{act}}^t = \text{MLP}(h_{\text{eye}}^t), \tag{2}$$

$$a_t = \text{argmax}(\text{Softmax}(h_{\text{act}}^t)), \tag{3}$$

其中 argmax 指的是选择最大值的索引操作，而 Softmax 是激活函数。基于独热向量 a_t，章鱼可以方便地选择相应的 CD 策略来实现。最后，将获得一个对比解码工作流 $\mathcal{A} = \{a_t\}_{t=1}^N$，其中 N 是响应的长度。

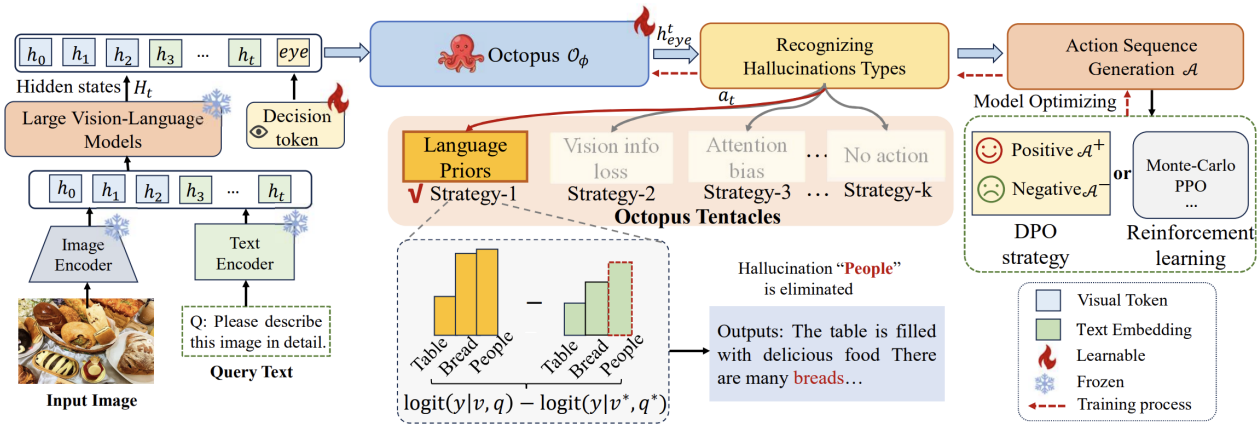

图 4: Overview of our method.

Octopus-模型优化

可以注意到，在上述计算中存在一个不可微操作，且由于缺乏明确的决策标签以及严重的维度灾难，优化过程也颇具挑战性。因此，作者引入直接偏好优化（DPO）来缓解这一问题。

数据构建：作者将上述动作选择过程重新表述为一个偏好问题，这鼓励章鱼 $\mathcal{O}_\phi$ 生成能够有效缓解幻觉的动作序列。为了构建正向工作流 $\mathcal{A}^+$ 和负向工作流 $\mathcal{A}^-$，为每个样本生成 10 个序列，通过在每个生成步骤中随机选择四个动作。接下来，根据 CHAIR 指标 [?] 将这些响应分为 $\mathcal{A}^+$ 和 $\mathcal{A}^-$。在实践中，这个指标也是灵活的，并且可以根据不同领域进行调整。对于判别任务，分别使用四个“触手”来构建 $\mathcal{A}$，而正负样本则根据答案的置信度分数进行分割。最后，使用平衡的正负样本来构建偏好数据集。

训练过程：为了引导策略模型输出首选序列 $\mathcal{A}^+$，DPO 用策略模型和参考模型取代了奖励过程，这可以直接增加正序列的可能性。因此，给定上述偏好数据集 $\mathcal{D} = \{\mathcal{A}^+, \mathcal{A}^-\}$，应用 DPO 直接优化我们的章鱼。此外，先前的研究表明，移除参考模型可以获得比原始 DPO[?, ?] 更好或相当的性能。基于此，优化目标定义如下：

$$\max_{\mathcal{O}_\phi} \mathbb{E}_{(x,\mathcal{A}^+,\mathcal{A}^-)\sim\mathcal{D}} \log \sigma \left(\beta \log \mathcal{O}_\phi \left(\mathcal{A}^+ \mid x\right)\right) - \beta \log \mathcal{O}_\phi \left(\mathcal{A}^- \mid x\right), \tag{4}$$

其中 $x = (v, q)$ 是输入序列，σ 表示 sigmoid 函数。根据 [?]，将 β 设置为 1。基于上述训练过程，章鱼可以自适应地学习构建合适的工作流，而无需人工标注。

实验设置

实现细节。为了训练模型，构建了两个数据集用于生成和判别任务。对于生成任务，在 MSCOCO 上构建了 10,000 条偏好数据，并使用语言提示“请详细描述这张图片。”。对于判别任务，遵循从 MSCOCO 训练集中构建了 7,000 条幻觉数据。使用 Adam 作为优化器。所有模型均在 4 块 3090 GPU 上训练，批大小设置为 4。

(6) 可改进的地方

【本文工作的局限性是什么？你觉得可以从哪些方面改进工作?】

虽然 Octopus 框架能够动态选择 CD 策略，但在实际应用中，如何更准确地识别幻觉类型并选择最优的 CD 策略仍然是一个挑战。未来的研究可以探索更复杂的模型结构或引入更多的上下文信息来提高决策的准确性。

(7) 可借鉴的地方

【你觉得本文哪些方面可以借鉴？比如思路、方法、技术等】
首先，其对幻觉问题成因的深入分析为后续研究提供了重要的参考。通过揭示幻觉成因的复杂性和多样性，本文为设计更有效的幻觉缓解方法提供了理论基础。其次，Octopus 框架的动态选择机制为处理复杂的多模态问题提供了一种新的思路。通过自适应地选择合适的 CD 策略，Octopus 框架能够更好地应对不同输入条件下的幻觉问题，这种动态选择机制可以应用于其他需要处理复杂多模态输入的任务中。

(8) 其他收获
【你有什么其他收获吗？比如了解了哪些团队和大牛在某领域做得很好，某类问题通常用什么技术解决，某些技术之间存在什么样的关联，某些会议和期刊在某领域很知名……】
本文还展示了强化学习在优化模型决策过程中的应用，这让我认识到强化学习在多模态任务中的潜力。

5 评阅人

姓名:
时间: