

使用语义标签来标注自然语言描述： 一种适用于 GPT 的 Prompt 规范化表示方法

摘要：

本文针对 GPT 这类大型预训练语言模型在使用 Prompt Tuning 和 Prompt Engineering 这两种方法优化 Prompt 时所共同面临的 Prompt 规范化表示问题，提出了一种新颖且有效的使用语义标签来标注自然语言描述以实现 Prompt 规范化的表示方法。首先，本文提出了一种基于语义标签和自然语言描述的 Prompt 规范化表示方法，可以根据不同的任务和场景在手动或者自动生成合适的 Prompt 时使用。其次，本文细化设计了一个规范化框架，包括语法规则、语义标签的选择原则和结构化表示以凸显出其中的关键信息和逻辑关系，给出了适合用作语义标签的预设的一组关键字集合并按照任务、流程、输入、输出、逻辑等概念分组、添加标注的方法、以及通过多次实验和迭代不断优化 Prompt 表示的流程，以提高其可读性、可解释性和可重用性。最后，本文还结合不同的任务场景，给出了若干个示例。。

关键词： 语义标签标注；自然语言描述；GPT；Prompt 规范化表示

一、 引言

大型生成式预训练模型（Generative Pre-trained Model, GPT）已经成为各种自然语言生成任务（NLG）中最强大且最灵活的工具之一，基于这种技术的聊天机器人，通过在大规模文本数据上进行无监督学习，获得了强大的语言表示能力，具有强大的生成能力和泛化能力，可通过微调或者零样本学习来适应不同的下游任务。在开放领域对话任务上取得了令人瞩目的效果。

随着自然语言模型的规模越来越大，大型模型参数从几百亿到几万亿，对于如此超大规模的参数，调整参数进行重新训练的代价很大。自 **GPT** 和 **BERT** 开发以来，标准做法一直是在下游任务上微调模型，这涉及调整网络中的每个权重（即模型调整）。然而，随着模型变得越来越大，很多时候为每个下游任务存储和提供模型的微调变得不切实际。特别是对于普通用户来说，语言模型相当于一个冻结的语言模型，无法修改其参数来适配。

所以，目前一个重要的研究方向是侧重于尽可能减少对模型参数的修改和重新训练代价的基础上，提升 **GPT** 的生成效果。这样，在目前的 **GPT** 语言模型中，「预训练、**Prompt** 和预测」的过程逐渐取代了「预训练、**Fine-tuning**」过程。在这种范式中，不是通过目标工程使预训练的语言模型（**LM**）适应下游任务，而是重新形式化（**Reformulate**）下游任务，使其看起来更像是在文本 **Prompt** 的帮助下在原始 **LM** 训练期间解决的任务。通过这种方式，选择适当的 **Prompt**，该方法可以操纵模型的行为，以便预训练的 **LM** 本身可以用于预测所需的输出，有时甚至无需任何额外的特定任务训练。这种方法的优点是给定一组合适的 **Prompt**，以完全无监督的方式训练的单个 **LM** 就能够用于解决大量任务。

然而，**GPT** 模型也存在一些问题和挑战。比如生成质量不稳定、缺乏多样性、难以控制等，**GPT** 回答的精准程度，很大程度依赖于用户输入的 **Prompt**，所以对用户编写 **Prompt** 的技巧和多次对话的沟通方法有较高要求。在某些情况下容易产生不符合用户意图或场景要求的回复。因此，在使用 **GPT** 模型进行 **NLG** 时，如何有效地给定输入 **Prompt** 来控制输出内容的质量、风格、逻辑等方面是一个关键问题。

为了解决这些问题，近年来出现了两种新颖的方法：**Prompt Tuning** 和 **Prompt Engineering** 是两种在自然语言处理（**NLP**）中使用预训练语言模型（**LLM**）的方法。两者的区别在于是否使用自动化工具来优化 **Prompt**。**Prompt Engineering** 是指通过设计不同的输入格式来描述任务，例如问题、指令或示例，从而引导语言模型生成高质量的响应。**Prompt Engineering** 需要用户尝试不同的提示，例如通过少量学习（**few shot learning**），来找到最佳的提示。**Prompt Tuning** 是一种在 **Prompt Engineering** 的基础上进行优化的技术。

Prompt Tuning 涉及调整提示的表示，以提高语言模型响应的质量。 Prompt Tuning 不需要用户手动试验不同的提示，而是通过数据科学家和自动化工具来比较和优化不同的提示选项，针对特定的任务。

对于 Prompt Engineering 和 Prompt Tuning，首要需要解决的一个关键基础问题是:该以何种形式作为用户输入 Prompt 的标准和规范。

由于人机交互的便利性，目前广泛采用自然语言描述编写 Prompt 的方法。目前已有的一些关于 Prompt 设计与优化的研究工作大多数都集中在单个任务或领域上，或者偏重于某种技巧，并没有形成统一且通用的标准或规范。此外在某些复杂且多样化的 NLG 任务中（例如图像描述、视频摘要等），单纯依靠用自然语言描述的 Prompt 很难表达出所有需要生成内容所涉及到的语义信息和逻辑关系。

考虑到用户输入中包含了丰富而隐含的语义信息，如意图、情感、主题和实体等，如果能够将这些信息显式地提取出来，并以一种结构化规范化的方式呈现给 GPT 模型，则有望指导其生成更符合用户期望和场景需求的回复。

针对以上问题，在本文中我们提出了一种基于以语义标签和自然语言描述规范化 Prompt 的方法。通过混合语义标签和自然语言描述两种形式的信息表示方法来构造有效且可控制的输入。在保留自然语言描述（NLD）提供更灵活和更直观的交互方式同时，通过利用语义标签（SL）来标注 NLD 的 Prompt 生成方法，提供更精确和更通用的语义约束，。

这种方法能够可以混合语义标签标注和自然语言描述来表示输入的多样性和复杂性，从而提高 Prompt 的表达能力、泛化能力、控制能力和适应能力。通过这种方式对 Prompt 进行规范化，是一个合理且有效的选择。

二、相关工作

2.1 Prompt 的表示形式与创建方法

Prompt 是一种由具体的中文或英文词汇组成的提示，它是人工可读的提示。Prompt 可以用来指导预训练模型完成特定的任务，例如对话生成、文本分类、问答等。Prompt 主要有两种主要类型：填充文本字符串空白的完形填空（Cloze）Prompt，和用于延续字符串前缀的前缀（Prefix）Prompt。选择哪一个取决于任务和用于解决任务的模型。一般来说，对于有关生成的任务或使用标准自回归 LM 解决的任务，前缀 Prompt 往往更有帮助，因为它们与模型从左到右的性质刚好吻合。

创建 Prompt 最自然的方式也许是基于手工创建比较直观的模板。例如，LAMA 数据集提供了手工创建的完形填空模板来探索 LM 中的知识。Brown 等在 2020 年创建了手工制作的前缀 Prompt 来处理各种各样的任务，包括问答、翻译和常识推理的探索任务。

虽然手工制作模板的策略很直观，并且确实可以在一定程度上准确地解决各种任务，但这种方法也存在一些问题：创建和对这些 Prompt 进行实验需要大量的时间和经验，特别是对于一些复杂的任务，例如语义解析（Shin 等，2021）；即使是经验丰富的 Prompt 设计者可能也无法手工发现最佳的 Prompt（Jiang 等，2020c）。

为了解决这些问题，很多研究提出了一些自动化模板设计过程的新方法。近年来，在手动构造 Hard Prompt 领域有不少有意义且有创新性的研究成果，例如基于模板化和插值的方法：这类方法通过设计一些固定或可变的文本模板，并在其中插入任务相关的关键词或实体来构造 Prompt。例如，P-Tuning 使用了一种基于位置嵌入（Position Embedding）的插值方法来生成 Prompt；PET 使用了一种基于伪标签（Pseudo-label）和数据增强（Data Augmentation）的模板化方法来生成 Prompt；TAPT 使用了一种基于任务适应性预训练（Task-adaptive Pre-training）和多样性采样（Diversity Sampling）的模板化方法来生成 Prompt。

当前 Prompt 研究工作的不足主要有以下几个方面：

- 缺乏统一的标准和规范：不同的研究者和任务可能使用不同的 Prompt 格式，导致难以比较和复现结果，也难以评估 Hard Prompt 的质量和效果。
- 缺乏通用性和可扩展性：构造 Prompt 的模板或者嵌入位置等需要大量的人工经验和领域知识，对于新的任务或领域可能需要重新设计或调整，增加了时间和成本。
- 缺乏鲁棒性和稳定性：构造 Prompt 的模板或者嵌入位置可能受到数据噪声、模型变化、语言多样性等因素的影响，导致生成的输出质量下降或出现错误。
- 缺乏有效的优化方法：构造 Prompt 的模板嵌入位置或者往往依赖于人工试错或启发式搜索，缺乏一种能够自动地优化 Prompt 参数或结构的方法。

2.2 自然语言描述

自然语言描述方法是通过使用自然语言来表达用户意图、情感、话题等信息，并将其作为提示传给 GPT 模型，从而引导它生成更加合理和有趣的回复。这种方法是对用户最友好的一种人机交互方式，同时可以有效地利用 GPT 模型强大的文本生成能力，并在不需要编写代码或修改模型结构的情况下实现多样化和个性化的对话功能。近年来，自然语言描述方法在 GPT Prompt Engineering 领域受到了广泛关注和研究。例如：

Machine Minds 提出了一个 GPT Prompt Engineering Primer，介绍了如何使用自然语言描述来构造不同类型的提示，如开场白、问题、建议、反馈等，并给出了一些实例和技巧；

Allabtai 介绍了如何构建一个强大的“序列提示”（Sequence Prompt），即使用多个自然语言描述组成一个长提示，并通过特殊符号或标记来控制 GPT 模型生成回复时的顺序和逻辑；该方法没有充分利用语义标签的信息，只是简单地将它们拼接到用户输入的自然语言消息中，而没有考虑它们与词汇之间的

关系和作用。这可能导致模型无法区分语义标签和词汇，或者忽略语义标签对生成结果的影响。

Arxiv 提出了一个提示模式目录（**Prompt Pattern Catalog**），即收集并分类了多种常用的自然语言描述方法，并分析了其优缺点和适用场景；该方法使用 GPT 模型在混合字符串上生成回复，并将回复中包含的语义标签替换为相应的词汇。但没有说明如何保证生成回复中包含正确且完整的语义标签，以及如何进行替换操作。这可能导致生成回复中出现错误或缺失的词汇，或者破坏回复的逻辑和连贯性。

综上所述，自然语言描述方法是一种简单而有效的技术，可以用于改善 GPT Prompt Engineering 的效果，并提高聊天机器人的智能性和用户满意度。

2.3 语义标签标注

语义标签标注方法此处是指利用语义标签（一种表示词汇或短语含义和关系的符号）来构造提示（**Prompt**）的技术，利用语言模型的先验知识，用于改善基于 GPT 模型的聊天机器人生成效果，生成更加准确和连贯的回复。语义标签标注方法可以有效地利用 GPT 模型强大的文本生成能力，并在不需要编写复杂代码或修改模型结构的情况下实现多样化和个性化的对话功能。此外，这种语言标签标注的表示方法，也有利于计算机理解，便于构建相应软件工具，自动化生成 **Prompt**，在复杂生成任务时，快捷生成地能清晰准确描述的 **Prompt**。在小样本学习或零样本学习的场景中，通过这种表示方法，可以提高语言模型在不同任务上的泛化能力

近年来，语义标签标注方法在 GPT Prompt Engineering 领域受到了广泛关注和研究。例如：

Arxiv1 提出了一种基于 GPT 的有序重要性语义通信方案（**SCOI**），即使用语义标签来表示用户意图、情感、话题等信息，并按照重要性顺序排列，从而引导 GPT 扮演一个咨询助手的角色，并生成更加合理和有趣的回复；

LinkedIn3 探讨了如何使用 Montague 语法和 Coq 证明助手来进行 GPT Semantic Testing，即使用语义标签来表示自然语言的形式语义，并通过类型检查和子类型推断来验证 GPT 生成回复是否符合预期。

综上所述，语义标签标注方法是一种简单而有效的技术，可以用于改善 GPT Prompt Engineering 的效果，并提高聊天机器人的智能性和用户满意度。本文将详细介绍我们提出的基于语义标签实现 Prompt 改善 GPT 生成效果方法，并通过实验验证其有效性和优越性。

三、 使用语义标签来标注自然语言描述

如上所述，在构建 Prompt 的过程中，“自然语言描述”和“语义标签标注”这两种方法各有千秋。若能结合两种办法，取长补短，则是一种值得深入探索的有效途径。基于这种考虑，本文提出混合语义标签标注和自然语言这两种描述 Prompt 的方式，用来规范化 Prompt 有效提升 GPT 生成效果。期望在交互自然和严谨准确之间折衷，在不明显降低用户的人机交互体验的前提下，尽可能地用 Prompt 严谨精准表达出用户的意图。使用一种用于与 GPT 互动的 Prompt 的通用标注式描述方式，满足清晰、专注、相关、连贯这些基本原则，吸收 Prompt engineering 的实践经验，让 GPT 知道你想要它做什么，从而生成一个更精准的回答。通过优化混合语义标签和自然语言描述的权重来平衡 Prompt 的表达能力和控制能力，通过预定义的格式和结构来规范输入信息，并通过特殊符号或词汇来引导模型生成期望的输出内容。这种混合方法既能保证 Prompt 的交互自然性，又能提高 Prompt 的严谨准确性，在不明显降低用户人机交互体验的前提下，尽可能地用 Prompt 严谨精准表达出用户意图。与其他 Prompt 表示方法相比，这种方法具有以下优点：（1）容易理解和解释，可以反映出任务的语义和逻辑，可以有效地减少输入信息的歧义和冗余；（2）容易复用和迁移，可以在不同的模型和数据集上使用，可以方便地适应不同的任务和领域；（3）可以灵活地组合使用，以实现更复杂且多样化的 NLG 效果。（4）容易通过文本接口输入和输出，可以与用户交互。

自然语言来表达用户意图、情感、话题等信息，并将其作为提示传给 GPT 模型。该方法简单易用，但是缺乏形式化和系统化的指导原则，并且不能有效

地控制或纠正 GPT 生成过程，特别是用来描述复杂任务时，不够精炼准确，同时也不够规范化，需要进行比较复杂的转换后才能被计算机系统理解。针对这种情况，我们可以通过在自然语言描述中添加语义标签标注的方式，对现有的原始的 Prompt 进行格式重构（Reformat）或者直接采用这种方式编写 Prompt 输入给 GPT 模型。这种方法可以看作是一种基于语义和描述的 Prompt 设计策略，它旨在利用用户输入中的隐含语义信息和描述信息来指导 GPT 模型生成更符合用户期望和场景需求的回复。

这种生成 Prompt 的方法使用了标签式标注和自然语言描述相结合的方式来定义生成 Prompt 的条件和要求。标签式标注可以帮助 GPT 识别出关键词和变量，自然语言描述可以帮助 GPT 理解上下文和目的。同时，它使用了编号、索引名、内容等格式来组织生成 Prompt 的信息，使其更加清晰和规范。这样可以方便 GPT 按照指定的顺序和结构来生成相应的内容。此外，它使用了在内容中判断、限制等语句来对生成 Prompt 的内容进行一些约束和过滤，使其更加符合题目要求和逻辑。这样可以避免 GPT 生成一些不相关或错误的内容。

这种生成 Prompt 的方法有以下优点：

它可以利用 GPT 强大的自然语言理解和生成能力，根据给定的条件和要求，快速地生成一些符合 C++面向对象程序设计上机编程题目特点和难度的题目。

它可以灵活地调整生成 Prompt 的各个要素，以适应不同的背景、场景、任务、扮演等需求。例如，可以改变扮演角色为学生或者考官，改变回答格式为选择题或者填空题等。

它可以通过参考网上已有的 C++面向对象程序设计上机编程题目及其答案，提高生成 Prompt 的质量和可信度。例如，在搜索结果中找到一些例题，并将其作为参考来源或者示例。

3.1 设计原则和语法规则

1. 设计原则

以下是在设计过程中，拟定的一些设计原则：

(1) 不要引入过多标记符号，要高度简洁、易于读写

充分考虑到用户手动创建 Prompt 和适应简单任务场景的需求，可以根据任务的特点和需求灵活地设计和调整 Prompt 的形式和内容

(2) 具备良好兼容性和可扩展性，能规范化表示不同任务场景下的 Prompt

将语义标签划分为通用语义标签和领域相关的专用语义标签。

选取通用的语义标签，用作语义标签的关键字，进行分组，每个组斟酌选择若干有概括力的关键字。目前是让用户在描述具体任务时，选取这些通用的语义标签。对于行业相关的任务，可以选择专用的语义标签，但是需要在 Prompt 用类似的语法进行解释。

GPT 在训练的过程中，通常均能针对这些通用的语义标签进行优化，使模型理解其中的丰富含义，并可根据需要进一步对领域相关的专用语义标签进行训练。

(3) 容易被机器解析与处理。

能够充分利用 GPT 预训练大模型所具备的先验知识和推理能力，用简短的命令触发模型内部一系列期望的反应。泛化能力不足和性能不稳定的问题。

基于前缀 Prompt 的方法，与 GPT 的自回归模型的性质相匹配，可以有效地将语言模型应用到不同的 NLP 任务上。作为前缀的语义标签可以提供更多的语义信息和任务指导，从而提高 GPT 的生成质量和多样性。

(4) 严谨图灵机，完备

(5) 融入编程思维，改善其输出的可控制和精准度

2. 语法规则

保留 GPT 以自然语言输入的优点，参考 JSON 格式、Markdown 等标记语言，本文设计使用一种表示方法，更简洁和一致的符号来表示不同层级的语义标签和内容。

(1) 符号

使用冒号“:”和逗号“,”来分隔同一层级的关键字和内容，使用中括号“[]”来包含分组信息。这样可以避免使用过多不同形状的括号，提高可读性和易用性，“#”表示注释，使用了换行来表示 Prompt 的层次结构和逻辑关系。

(2) 结构

每行 Prompt 语句首先使用语义标签来表示任务的语义结构，然后使用自然语言描述来补充任务的语义细节，这样可以生成更灵活、更丰富、更符合任务要求的 Prompt。

形式化表示如下：

```
//注释 1
[维度语义标签 1]
模式语义标签 1: 内容 1-1 内容 1-2, .....
模式语义标签 2: 内容 2-1 内容 2-2, .....
模式语义标签 3: 内容 3-1 内容 3-2, .....
.....

//注释 2
#定义 专用语义标签 含义
[维度语义标签 2]
模式语义标签 1: 内容 1-1 内容 1-2, .....
模式语义标签 2: 内容 2-1 内容 2-2, .....
模式语义标签 3: 内容 3-1 内容 3-2, .....
.....
```

其中维度语义标签和模式语义标签，均是若干约定好的关键字，分别用来描述用户的意图分组和其中蕴含的具体模式。

从这种形式化表示可看出：应用这种表示方式的关键在于将用户在 Prompt

中待表达的意图按照维度进行分组，用定义好的维度标签来标识；然后在每个问题维度范围内，选用若干模式语义标签，将用户待输入 GPT 的 Prompt 所有内容，匹配不同的模式拆分为若干行，每行的基本格式为模式语义标签:内容，再进一步将每行拆分为若干部分，用逗号分隔。

(3) 关键字

本文提出的方法可能需要用户花费一些时间和精力来学习和掌握不同的语义标签的含义和用法，以及如何正确地组织和拼写用户的 Prompt 。

在本文的语义标签场景下，我们定义维度是用来将用户待输入的 Prompt 分组，每个分组表示看待 Prompt 的一个侧面，而模式则代表典型的方法用来约束或者启发激活 GPT 的能力的一些典型做法。例如：定义 **formal**（任务形式）作为维度，其下面有 **question, instruction, example** 等，分别表示提问、指令、示例等不同的模式；定义 **condition**（任务条件）作为一个维度，其中包含 **topic, format, length, restriction** 等分别表示任务的话题、输出的格式、输出的长度、输出的内容限制等不同的模式；定义 **evaluation**（评估标准）作为维度，其下有 **accuracy, completeness, fluency** 等，分别表示输出的正确性、完整性、流畅性等。

我们将关键字分为两类：通用关键字和专用关键字。它们的区别在于：GPT 可以根据对话上下文正确地理解通用关键字的含义；但是对于专用关键字，GPT 可能无法正确地理解，或者需要更多的定义和解释来辅助理解。这些定义和解释可以根据回答效果动态地提供，也可以在对话开始之前静态地提供。。专业关键字，类似于领域特定语言 DSL（Domain Specific Language），仅适用于特定的场景或任务。这里我们将研究范围限定于通用关键字。对于特定领域的任务，可以参考给出专业关键字集合。

(4) 作用

从终端用户的角度，这种表达 Prompt 的方法可以帮助用户清晰地定义用户的输入和输出的格式，方便用户控制和调整 GPT 的行为和结果。同时这种方法也可以帮助用户提高输入的质量和效率，避免输入一些不必要或不相关的信息。

这种方法还可以帮助用户更好地利用 GPT 的多任务能力，让它根据不同的标签完成不同的任务，例如生成文本，搜索信息，提供建议等。

同时，从 GPT 大模型的角度，这种方法可以被视为一种基于反馈指令的强化学习的方法，让用户通过抽取不同的关键词作为语义标签来展示用户的输入的自然语言描述中的核心内容。这种统一的规范化方式提供不同的输入类型来搜集和处理不同类型的用户输入，从而给 GPT 提供不同的语义标签来指导和改善它的学习和表现。

3.2 定义语义标签的维度与模式

对于终端用户来说，GPT 这类大模型，好比一个复杂的黑盒子或则灰盒子。如何充分利用已经在预训练过程中内置的先验知识，充满技巧性。选择关键字语义标签的过程，在某种程度上，是一项充满技巧的工作。用编程思维看待语义标签，就好比选择流程图的各个方框的名字。

值得注意的是，在实践中应用该方法：编写和选择作为语义标签的关键字，可能需要一定的专业知识和经验；需要对不同的任务和领域进行不同的 Prompt 构造，这可能需要一定的调整 and 适应；需要对不同的 GPT 模型进行不同的 Prompt 优化，这可能需要一定的计算资源 and 时间。

针对通常的任务，我们可以归纳抽象提取出包含但是不限于下列通用关键字作为语义标签来标注：判断、循环、扮演、场景、任务、格式、输入、输出、提示、限制、需求、要求、测试、评估、提示、包含、排除、筛选、风格、语气、场景、目标、示例、给出自动化脚本、结构、来源、上下文、背景、语言、重启、新对话、异常处理、脚本、细化、替代、核对、假设、可视化、规则、分解、反射、组合、模板、反馈、字数、翻转、步骤、变量、权限、函数、注释、内省、传递、改写、符号、元语言。这些关键字，能够很好地归纳描述用户常见通用的交互任务，并且现在市面上的 GPT 能够准确理解其含义，不需要用户额外输入 Prompt 去解释。所以优先选择能达到这种效果的关键字用作通用语义标签。当然，这个需要依靠经验，并且在实际 GPT 系统上反复试验，注意 Prompt 在包含不同的关键字时 GPT 的反馈，了解其能力和局限。

另外，由于当前市面上的 GPT 模型普遍存在对话回合数和文字 Token 数量的限制，这使得我们在设计 Prompt 时必须像计算机早期发展阶段那样节约内存资源。因此，我们在选择关键字时要求它们既简洁又表达力强，并且能够灵活地与其它关键字组合构成不同语义。同时我们也要控制关键字数量的适度，避免给用户带来过多的记忆压力和操作难度。

我们建议使用以下表格中整理的通用关键字集合，在表格中我们按照不同的维度将它们进行了分组，并给出了每个关键字的具体定义和示例：

1. 任务维度

关键字	含义	示例
任务(task)	需要完成或协助完成的具体操作、功能或需求	生成一份程序设计考卷，回答问题、讲故事、写代码、预订酒店等。
需求	需要满足或优化的性能、质量或效果指标	准确度、流畅度、创造性等
要求	必须遵守或避免的原则、道德或法律规范	不得侵犯版权、不得伤害他人感情、不得泄露隐私等
结构	输出或表现的组织、排列或分布方式	按照时间顺序、按照重要性排序、按照类别分组等
格式(format)	输出的内容的形式，需要遵循或生成的数据结构、文本类型或样式	Markdown、HTML、Python 等, JSON 格式、Markdown 格式、表格格式等
目标(goal)	需要达成或实现的结果、效果或价值	解决问题、提供信息、娱乐用户等。
参 考 (reference)	参考或模仿的对象或事物	某个名人、某本书籍
限制(limit)	遵守的条件或约束。约束 GPT 行为或输出的条件、规则或标准	字数、时间、难度。限制输出的长度、语言、内容等
鼓 励 (encourage)	展现的特点或优势	创造性、逻辑性、多样性
禁止	不想让 GPT 做的事情	抄袭、生成敏感内容
场 景	需要参与或描述的情境、背景	在医院进行问诊、在餐厅点菜、在火星探索、

(scenario)	景、主题或目标	在童话故事中冒险等
提示	给出一些提示或示例来引导或启发 GPT 的输出。帮助 GPT 理解或执行任务的信息、示例或建议。一些引导 GPT 改进或修正输出或表现的信息、示例或建议	问题的开头、答案的结构等。给出一个问题的答案范例、给出一个故事的开头或结尾、给出一个代码的功能描述等。给出一个错误提示、给出一个改进方向、给出一个参考资料等。
权限	指定用户或 GPT 可以或不能做什么事情，以保护隐私和安全。	例如，在 Prompt 中加入[permission: you cannot access my personal data or make any changes to my settings.]表示用户不能访问我的个人数据或修改我的设置。
规则	指定一些约束条件或逻辑判断，让 GPT 按照规则进行回答。	例如，在 Prompt 中加入[rule: if the user input contains a question mark, then answer the question; otherwise, say "I don't understand."]表示如果用户输入包含一个问号，则回答问题；否则，说“我不明白”。
背景	提供了一些与当前会话相关的信息、历史或状态，帮助 GPT 保持和用户之间的连贯性和一致性	在聊天时记住用户的姓名
假设	假设：设定一个假设情境，并根据假设情境进行回答。	例如，在 Prompt 中加入[assume: you are a fairy tale character. who are you and what is your story?]表示假设用户是一个童话故事中的角色，并让用户描述自己是谁和自己的故事。

2. 流程维度

关键字	含义	示例
上下文	提供了一些与当前任务或话题相关的信息、背景或条件，帮助 GPT 理解和回应用户输入	在回答问题时提供相关领域的知识，在讲故事时提供故事情节和人物设定，在写代码时提供代码功能和需求等
输入	可以接收和处理的数据来源、类型和范围	在线搜索结果、用户提供的文本信息
输出	需要生成和返回的数据目标、类型和范围	文本回复、图像生成、音频播放等
测试	包含了一些检验 GPT 能力或水平的问题、任务或挑战	回答常识问题、完成编程题目
评估	一些评价 GPT 输出或表现的标准、方	给出一个分数、给出一个反馈、给出

	法或工具	一个建议等
异常	处理意外情况或错误	在 Prompt 中加入[error handling: apologize and ask for clarification]表示如果发生错误或不理解用户输入时，应该道歉并请求用户澄清。
步骤	步骤：将一个复杂或长期的任务分解为多个简单或短期的子任务，并按顺序执行。	例如，在 Prompt 中加入[steps: how to bake a cake? step 1: preheat the oven to 180 degrees Celsius. step 2: mix flour, sugar, eggs, butter and milk in a large bowl. step 3: pour the batter into a greased cake pan. step 4: bake for 25 minutes or until golden brown. step 5: let the cake cool down and enjoy.]表示将烘焙蛋糕这个任务分解为五个步骤，并按顺序执行。
反射：	对自己或用户的输入或输出进行评估或反思，并给出相应建议或改进。	例如，在 Prompt 中加入[reflect: how do you feel about your answer? I think my answer is accurate but not very interesting. I could improve it by adding some examples or jokes.]表示对自己的回答进行反思，并给出改进意见。
内省：	让 GPT 自我反思其输出或行为，并给出评价或建议。这可以帮助我们发现 GPT 的错误或潜能，并改进其性能。	
流程	让 GPT 将其输出或行为传递给另一个 LLM 或系统，并接收其反馈或结果。这可以实现多个 LLM 之间的协作或竞争，并增加对话的丰富性和趣味性。	

3. 输入维度

关键字	含义	示例
台词	指定一个预先编写好的脚本，让 GPT 按照脚本进行回答。	例如，在 Prompt 中加入[script: 医院问诊脚本.txt]表示使用医院问诊脚本来回答用户问题。

	添加一些说明或备注, 以帮助用户或 GPT 理解输入或输出的意图或内容。	在 Prompt 中加入[comment: this is a joke, don't take it seriously.]表示添加一个注释说明这是一个笑话, 不要当真。
--	--------------------------------------	--

4. 生成维度

关键字	含义	示例
包含	输出或表现必须包含的元素、内容或特征	包含一个主题句、包含一张图片、包含一首诗等
排除	指定了 GPT 输出或表现必须排除的元素、内容或特征	排除敏感词汇、排除重复信息、排除无关内容等
筛选	从输入或输出中选择或删除符合某些条件的元素、内容或特征	筛选最相关的搜索结果、筛选最优秀的作品
来源	指定了 GPT 可以使用或引用的数据或信息的来源、类型或范围	来自互联网、来自用户输入、来自内置知识库等
模板	模板: 使用一个预先定义好的格式或结构来生成输出, 并填充相应内容。	例如, 在 Prompt 中加入[template: write a haiku about spring. a haiku is a three-line poem with five syllables in the first line, seven in the second, and five in the third. example: spring is in the air / flowers are blooming everywhere / life is full of joy.]表示使用一个模板来写一个关于春天的俳句, 并给出一个例子。
反馈	要求 GPT 给出或接受用户对其输出或表现的评价、意见或建议。给出对用户输入或输出的评价或建议, 并说明理由或依据。	给出一个反馈, 告诉用户他们的输入是否正确、合理或有趣。例如, 在 Prompt 中加入[feedback: rate your haiku from 1 to 5 stars and explain why. I give your haiku 4 stars. it is simple and elegant, but it lacks some originality.]表示给出对用户写的俳句的评价, 并说明原因。
建议	要求 GPT 给出或接受用户对其输出或表现的改进、优化或创新的方案、方法或思路。	给出一个建议, 告诉用户他们可以如何提高他们的写作水平、编程能力或口才。

5. 方法维度

关键字	含义	示例
示例	提供了一些展示或说明任务或格式的样本、模型或范例	给出一个问题和答案的示例、给出一个故事的示例
逐步		
交互	要求 GPT 与用户进行多轮的沟通、协作或竞争	在聊天时回答用户的问题，在游戏中与用户合作或对抗，在教学中引导用户学习等
重启	重新启动对话，清除之前的上下文	在 Prompt 中加入[restart: true]表示开始一个新的对话，并忽略之前的内容。
细化：	将一个大问题分解为多个小问题，逐一回答。	在 Prompt 中加入[refine: what is your main symptom? how long have you had it? have you taken any medication?]表示将用户提出的一个大问题细化为三个小问题，并依次回答。
替代：。	提供多个备选方案供用户选择	在 Prompt 中加入[alternative: do you prefer Chinese food or Italian food? please choose one.]表示给用户两个选项，并要求用户选择其中一个。
核对：	检查用户输入是否符合预期，并给出相应提示。	例如，在 Prompt 中加入
组合：	将多个元素组合成一个新的元素，并说明组合方式和结果	在 Prompt 中加入[combine: what do you get when you mix red and yellow? you get orange. red and yellow are primary colors, and orange is a secondary color.]表示将红色和黄色组合成橙色，并说明组合方式和结果。
分解	分解：将一个复杂或抽象的概念分解为多个简单或具体的部分，并依次解释。	例如，在 Prompt 中加入[decompose: what is a computer? a computer is a device that can perform calculations and store information. it consists of hardware and software. hardware is the physical components of the computer, such as the CPU, memory, disk, etc. software is the set of instructions that tell the hardware what to do.]表示将计算机这个概念分解为多个部分，并逐一解释。
翻转	让 GPT 来提问，用户回答	

6. 内容维度

关键字	含义	示例
主题(theme)		
风格(style)	输出或表现的语言、文体或情感特征	风趣幽默、正式严肃、激昂慷慨等。
角色(character)	你和 GPT 之间的关系或身份。指定了 GPT 需要模仿或模拟的角色、人物、风格或语气	老师和学生、同事和客户。扮演一位医生、一位名人、一位诗人、一位机器人等
情节(plot)		
对话(dialogue)		
语言(language)	与 GPT 进行交流时使用的语言	在 Prompt 中加入[language: 中文]表示使用中文进行对话
语气	输出或表现的态度、立场或意图	友好礼貌、批判质疑、请求帮助等

7. 逻辑维度

关键字	含义	示例
判断(judge)	出一些判断或评价标准来检验或反馈 GPT 的输出。包含一个或多个条件语句，根据不同的输入或输出选择不同的分支或行为	正确与否、得分多少。如果输入是一个问题，那么输出应该是一个答案；如果输入是一个命令，那么输出应该是一个执行结果；如果输入是无效的，那么输出应该是一个错误提示。
循环(loop)	出一些循环或重复的规则来控制或调整 GPT 的输出。包含一个或多个重复执行的语句块，直到满足某个终止条件或达到某个限制	每个问题有几个选项、每个选项有几个字母等。对于每个输入，重复生成三个不同的输出，并让用户选择最佳选项；或者，在用户没有结束会话之前，重复接收和回复用户的输入。
条件(condition)		
变量(variable)	定义一个可以存储和修改值的符号，以便在输入或输出中重复使用。	例如，在 Prompt 中加入[variable: name = "Alice"]表示定义一个变量名为 name，值为 "Alice"，并在后续输入或输出中使用该变量。
函数(function)	定义一个可以接受参数并返回结果的操作，以实现复用和模块化。	例如，在 Prompt 中加入 [function: greet(name) = "Hello, " + name + "！"]表示定义一个函数名为 greet，接受一个参数 name，并返回一个问候语。

四、 总结

论文提出了一种用于 GPT 的 Prompt 标准化和规范化方法，可以有效地提升生成质量和多样性。论文的主要贡献有以下几点：

- 提出了一种基于语义标签和自然语言描述的 Hard Prompt 生成方法，可以根据不同的任务和场景自动生成合适的 Hard Prompt。
- 设计了一个 Prompt 标准化和规范化框架，以提高其可读性、可解释性和可重用性。使用自然语言描述来补充任务的语义细节。这种方法可以生成更灵活、更丰富、更符合任务要求的 Prompt，从而提升 GPT 的生成效果。

相对于其他 Prompt 标准化和规划化表示方法，这种方法的优点有：

- 它可以利用混合语义标签的抽象性和自然语言描述的具体性，结合两者的优势，生成更合适的 Prompt。
- 它可以根据不同的任务和领域，灵活地调整混合语义标签和自然语言描述的比例和内容，生成更多样化的 Prompt。
- 它可以通过自然语言描述来引入一些背景知识和常识，增强 Prompt 的语义表达能力和逻辑推理能力。

这种方法的不足之处有：

- 它需要人工设计和标注混合语义标签和自然语言描述，这可能会增加人力成本和时间成本，也可能会引入一些主观偏差和噪声。
- 它没有考虑到 Prompt 的长度和位置对生成效果的影响，这可能会导致一些不必要的冗余或缺失。
- 它没有对比和分析其他类型的 Prompt 方法，比如 Soft Prompt 或 Hybrid Prompt，这可能会限制它的泛化能力和创新性。

五、 参考文献

Promptstacks - GPT Prompt Engineering Community: 这是一个专门为 GPT 设计和分享 Prompt 的在线社区，用户可以浏览和搜索各种类型和主题的 Prompt，也可以贡献自己的创意。

A Hands-on Guide to Prompt Engineering with GPT and GPT-3: 这是一篇介绍如何使用 GPT 和 GPT-3 进行 Prompt engineering 的教程文章，包含了许多实用的示例和代码。

A Prompt Pattern Catalog to Enhance Prompt Engineering with GPT: 这是一篇关于使用 GPT 进行 Prompt engineering 的学术论文，提出了一个基于模式的目录来系统化地构建优

化的 Prompt。

[FREE RESOURCE 📖 GPT Prompt Engineering Primer 📖 Chapters 1–3 \(substack.com\)](#)

[1] Finetuned Language Models Are Zero-Shot Learners

<https://arxiv.org/abs/2109.01652>

[2] Multi-Task Deep Neural Networks for Natural Language Understanding

<https://aclanthology.org/P19-1441/>

[3] Muppet: Massive Multi-task Representations with Pre-Finetuning

<https://arxiv.org/abs/2101.11038>