arXiv:2303.01094v1 [cs.CL] 2 Mar 2023

CTRLStruct：开放域响应生成的对话结构学习

Congchi Yin 南京航空航天大学 中国南京
congchiyin@nuaa.edu.cn

Piji Li*南京航空航天大学 中国南京 pjli@nuaa.edu.cn

Zhaochun Ren 山东大学 中国青岛
zhaochun.ren@sdu.edu.cn

摘要

对话结构发现对于对话生成至关重要。结构良好的主题流能够利用背景信息并预测未来话题，从而帮助生成可控且可解释的回复。然而，以往的大多数工作都聚焦于面向任务的对话中的对话结构学习，而非更加复杂且更具挑战性的开放域对话。本文提出了一种新的对话结构学习框架**CTRLStruct**，利用未标注信息有效探索主题级对话簇及其转移。具体而言，我们通过一项专门设计的对比学习任务，进一步训练由双向Transformer编码的对话话语，以改进其表示。随后，我们对话语级表示进行聚类，形成主题级簇，并将其视为对话结构图中的顶点。图中的边表示顶点之间的转移概率，该概率通过模仿数据集中的专家行为计算得出。最后，将对话结构图整合到对话模型中，以实现可控响应生成。在两个常用的开放域对话数据集上的实验表明，与一些优秀的对话模型相比，我们的模型能够生成更加连贯的回复；在对话话语表示方面，其表现也优于一些典型的句子嵌入方法。代码可在GitHub[1]获取。

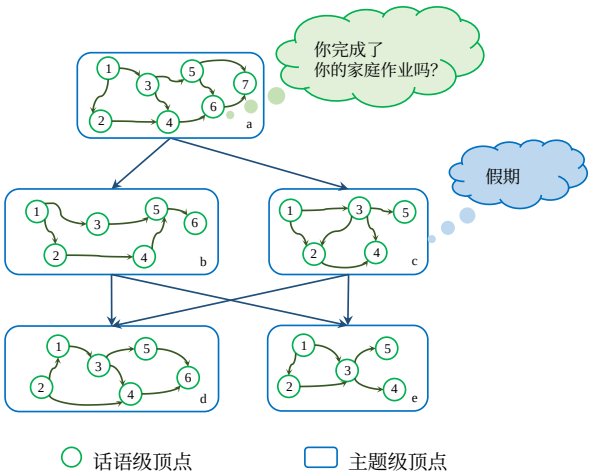

图1：对话结构示意图，其中圆形表示话语级结构，圆角矩形表示主题级结构。箭头表示转换。

ACM Web 会议 *2023*（*WWW '23*），*2023*年*4*月*30*日至*5*月*4*日，美国德克萨斯州奥斯汀。ACM，美国纽约州纽约市，12页。https://doi.org/10.1145/3543507.3583285

CCS 概念

• **计算方法** → **话语、对话与语用**；**自然语言生成**。

关键词

对话结构学习、开放域对话生成、话语表示、对比学习、模仿学习

ACM 参考文献格式：

Congchi Yin、Piji Li 和 Zhaochun Ren。2023年。CTRLStruct：面向开放域响应生成的对话结构学习。载于论文集：

*通讯作者。[1] https://github.com/lemonsis/CTRLStruct

允许为个人目的制作本作品全部或部分内容的数字副本或纸质副本允许在课堂上免费使用，前提是不得为牟利或商业利益而复制或分发副本，并且副本须附有本声明及首页上的完整引文。作者以外的其他权利人所拥有的本作品组成部分的版权必须得到尊重。注明出处后，可摘录引用。以其他方式复制、转载、发布到服务器或分发至列表，均须事先获得特定许可和/或支付费用。请向 permissions@acm.org 申请许可。*WWW '23*，*2023*年*4*月*30*日至*5*月*4*日，美国德克萨斯州奥斯汀© 2023年版权归所有者/作者所有。出版权已授权给 ACM。ACM ISBN 978-1-4503-9416-1/23/04...$15.00
https://doi.org/10.1145/3543507.3583285

1 引言

开放域对话生成是一项颇具挑战的任务，因为它具有复杂的多轮交互结构、丰富多样的话题，而且缺乏明确的目标。因此，仅在响应生成中考虑语言建模远远不够，正如 GPT2 [33], BART [21],和 T5 [34]等以往工作所示。对话结构体现了对话状态及其转移，在对话生成中发挥着不可或缺的作用。它通常被视为话语级结构与主题级结构的结合 [50]。前者可看作对话话语的流程，后者则代表主题转移。如图1所示，从“你做完家庭作业了吗”到“是的，做完了”的转变属于话语级结构；而主题级结构则是从更高层面描述转变，例如从“学习”转到“假期”。直观而言，如果对多轮对话中的对话结构进行建模，聊天机器人就能识别当前所处的主题，以及下一步可能涉及的话题，从而大幅提升多轮连贯性、可控性和可解释性。

对话结构最早是在面向任务的对话研究中提出的。早期研究依赖手工制定的规则 [20] 或人工标注 [17] 来对特定领域任务中的结构进行建模。随着无监督学习的发展，相关方法

例如隐马尔可夫模型（HMM）在对话潜在结构建模方面展现出良好的前景 [7, 37, 51]。其中一些研究已成功应用于导游机器人或票务预订系统的设计。然而，要将这一成功复制到开放域对话系统中并不容易。面向任务的对话只涉及几十或几百种对话状态，且状态转移简单；相比之下，开放域对话结构难以建模，因为其对话状态数量不受限制，流程复杂且不稳定。在现实场景中，人们面对相同的话语或主题时，可能会给出不同的回应。近期研究 [48, 50] 强调了构建话题感知对话系统以开展主动式对话的重要性。Xu 等人 [49]对开放域对话语料库中的对话结构进行了显式建模，并结合变分自编码器和图神经网络，构建了话语级和会话（主题）级的对话结构图。然而，对话状态的转移概率是通过共现频率统计手动计算的，这种方法需要离线计算，效率不高，而且还需要调整阈值以确保性能。因此，尤其是在当前易于获取大规模数据集的训练场景中，有必要采用适当的机器学习方法，从对话语料库中自动提取并建模对话结构。

此外，在研究过程中，[29, 41, 44], 我们注意到，对话话语表示的性能对所学习的对话结构质量有着巨大影响。为了改进话语表示学习，尤其是在对话场景中的话语表示学习，需要深入考虑与其他形式语料库相比所独有的对话敏感特征。人类对话可以被视为话题的定向流动，具有高度的上下文相关性和序列性，其中“上下文相关”意味着一条回复在话题层面上应与其语境密切相关，而“序列性”意味着对话由具有内在逻辑的序列话语组成，不能随意改变。以往的工作 [30, 49] 采用循环神经网络或 BERT [9] 直接编码并表示话语语义信息。这些方法将对话视为普通语料库，忽略了其独特的对话敏感特征，因此无法保证满足我们的对话结构学习要求的高质量性能。

为了解决上述问题，我们提出了一种名为**CTRLStruct**的新框架，即对比式话语表示学习与对话结构建模，以便从大量未标注数据中有效提取话语级和主题级转换，并生成连贯的对话。在话语表示学习中，我们通过对比学习设计了一种特殊的损失函数，并考虑了对话敏感特征，这有助于生成语义丰富的对话话语表示。随后，将含义相似的话语级表示聚集起来，形成主题级簇，如图1中的蓝色圆角矩形所示。由于数据集中已知话语级和主题级转换，我们采用模仿学习，通过计算主题级簇之间的转换来模仿人类意图，从而进行对话结构学习并获得最终的对话结构图。构建对话结构图后，我们考虑生成句子与主题级簇中心之间的 Kullback-Leibler散度 [19]，以此促使模型生成与主题相关的对话。在两个开放域对话数据集上的实验结果表明，与一些强大的对话生成模型相比，我们的模型能够生成更加连贯且与主题相关的回复；此外，在对话话语表示方面，我们的模型也优于一些强大的句子嵌入方法。

主要贡献可概括如下：

- 我们提出了一种新颖的框架，利用未标注语料库发现对话结构并控制响应生成。
- 我们提出了一种考虑对话敏感特征的对比学习编码器，以增强话语表示学习，进而提升主题级对话结构建模的性能。
- 我们提出了一种基于模仿学习的方法，用于进行主题级对话转移概率估计和高质量的对话结构学习。
- 自动评估和人工评估均表明，对话结构有助于提升对话生成模型的表现，使其能够生成与主题相关且连贯的对话。

2 相关工作

2.1 对话结构学习

以往大多数用于发现对话结构的无监督方法都聚焦于任务导向对话。隐马尔可夫模型[38, 51] 被用于基于主题级对话状态对对话流程进行建模。Shi 等人 [41] 采用循环变分自编码器，对具有高度非线性动态特征的潜在对话结构进行建模。Gunasekara 等人 [12] 通过将对话空间量化为若干簇来构建对话结构。最新研究 [30, 42, 49] 则应用图神经网络，进一步改进层次化对话结构建模。部分研究开始为在开放领域场景中发现对话结构提供新的思路。

2.2 预训练语言模型

在海量语料上预训练的大语言模型推动了各种自然语言处理任务的发展。它们可以针对不同的下游任务进行微调，并展现出具有竞争力的性能。BERT [9] 使用基于双向注意力机制的 Transformer [47]，其中语境会在注意力分数计算过程中纳入考虑。GPT [32], GPT2 [33], 和GPT3 [4] 使用采用单向注意力的 Transformer，每个词元只能关注其前面的词元。近期研究 [21] 还结合了双向注意力和单向注意力，构建出兼具自然语言理解与自然语言生成能力的模型。

2.3 对比学习

对比学习是一种用于学习有意义表示的自监督框架。其主要目的是拉近相似样本之间的距离，并拉远不相似样本之间的距离。Gutmann 和 Hyvärinen [13] 提出了噪声对比估计，van den Oord 等人 [45] 将其发展为 InfoNCE 损失，

$$loss_i = -\log \frac{e^{sim(z_i, z_i^+)/\tau}}{\sum_{j=1}^{N} e^{sim(z_i, z_j)/\tau}}. \quad (1)$$

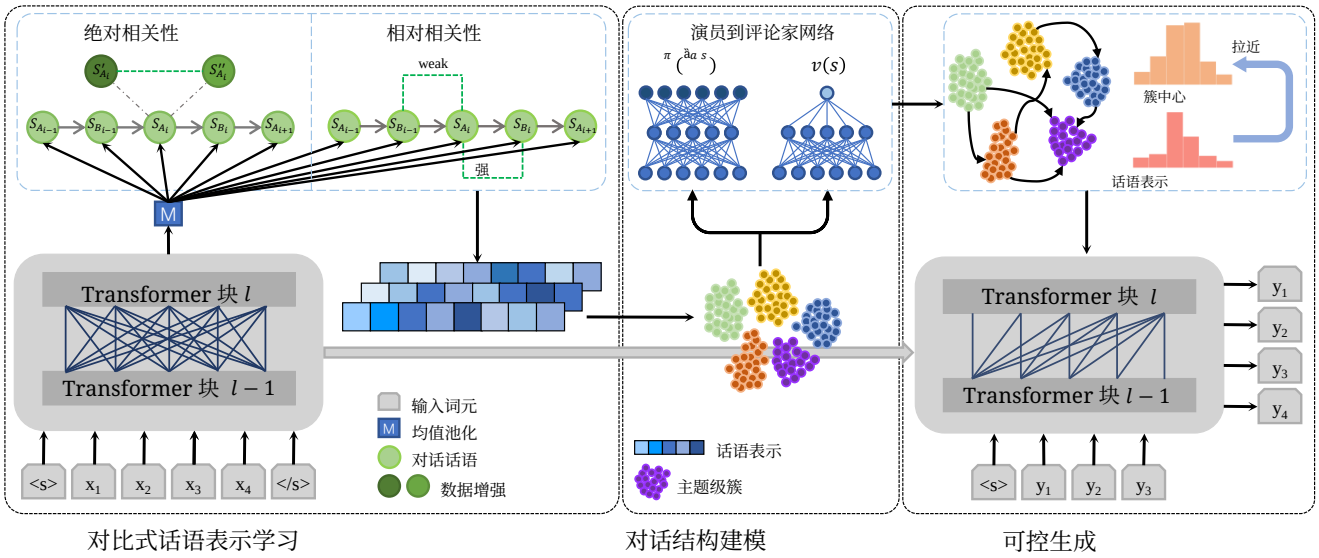

图2：整个模型的框架。 <s>和 </s>是输入中的特殊标记，分别表示句子的开始和结束。请注意，CTRLStruct 中通过对比学习训练的双向注意力编码器并未用于对话生成；实际使用的是原先预训练的编码器。

在公式 (1) 中， z_i 和 z_i^+是正样本对 x_i 和 x_i^+的表示。 τ 是温度超参数，而 $sim(z_i, z_i^+)$ 表示 z_i 和 z_i^+之间的某种距离度量。InfoNCE 损失为现代对比学习方法奠定了坚实基础，例如计算机视觉领域的 MoCo [14] 和 SimCLR [6] 。在自然语言处理领域，SimCSE [11] 利用 Transformer 中的 dropout 构造正样本，并成功学得了最先进的句子嵌入。

3 方法

3.1 概述

给定一个包含 $|D|$ 个对话的对话语料库 D，这些对话表示为 $\{X_1, X_2, ..., X_{|D|}\}$。每个对话由说话者 A 和 B 进行的多轮话语组成。第 i 个对话 X_i 包含 k_i 个话语 $X_i = [S_{A_1}, S_{B_1}, S_{A_2}, \ldots]$。在对话结构建模中，语义相似的话语应聚集在同一个主题簇 C_i 中。我们假设每个簇中心向量 $\boldsymbol{c}_i$ 都包含主题层级的信息。我们将这些簇视为马尔可夫链中的随机变量，每个簇都有一定概率转移到另一个簇。我们的目标是找出由主题转移构成的对话结构，并利用该对话结构来控制响应生成。

如图2所示，我们提出的 CTRLStruct 模型由三个主要部分组成：对比式话语表示学习部分、对话结构建模部分，以及对话结构控制的响应生成部分。在第一部分中，双向 Transformer块的均值池化输出被视为话语表示的原始版本。随后，我们采用对比学习进一步训练Transformer编码器，并获得最终的话语表示。在第二部分中，我们将语义相似的话语聚集在一起，形成主题级簇，并利用模仿学习自动计算主题之间的转移概率，进而构建对话结构图。为了进行对话结构控制的响应生成，给定对话上下文后，我们可以根据对话结构图获取目标主题。随后在训练过程中，我们使生成回复的表示更接近预测的簇中心。通过这种方式，主题级对话结构信息被整合到自回归解码器的训练过程中。

3.2 话语表示学习

对于每个输入话语 S，双向 Transformer 编码器会输出原始表示 $\mathbf{H} \in \mathbb{R}^{m \times n}$，其中 m 表示话语中的词元数量。先前的工作 [22, 36] 表明，对预训练模型的输出嵌入进行均值池化，比最大池化或 $[CLS]$ 表示法有更好的性能。我们对 $\mathbf{H}$ 进行均值池化，得到 n维向量 $\boldsymbol{h}$，作为初步的话语表示。

与其他形式的语料库不同，对话具有独特的特征，这些特征在对话话语表示学习中不容忽视。从更高层次来看，对话可以被视为话题的定向流动。因此，我们作出基本假设：人类对话具有高度的**上下文相关性**和**序列性**。“上下文相关”意味着，一条回复在话题层面上应当与其语境密切相关。一条话语与另一条话语越接近，它们之间的相关性就越强。“序列性”意味着，对话由具有内在逻辑的序列性话语组成。诸如“A：你明天有空和我们一起去爬山吗？B：抱歉，我得先完成工作。”这样的对话是合乎惯例的，但如果改变其顺序，就会显得奇怪。然而，即使有了上述明确的假设，要制定一套能够清晰解释某条话语在对话中所起作用的规则，仍然十分困难。某条话语可能在语义上与其前一条话语相似，也可能与前一条话语联系很少，但与其后一条回复关系密切。

例如，当一条话语是对某些查询的回复时，它符合前一种情况；当一条话语是新话题的开端时，它对应后一种情况。我们将这一问题定义为话语归属问题。

我们采用对比学习来处理上述问题。我们定义了两种相关性：**绝对相关性**和**相对相关性**。如图2所示，绝对相关性遵循 SimCLR [6]提出的对比学习框架。同一个话语 S_{A_i}的两个数据增强样本 S_{A_i}'和 S_{A_i}''构成一个正样本对。具体而言，我们在 CTRLStruct 中选择了四种不同的数据增强策略：插入上下文相关词、随机替换、同义词替换，以及应用于 SimCSE [11]中的随机失活增强技巧。相对相关性由**强相对性**和**弱相对性**组成。考虑到对话的序列性，S_{A_i}, S_{B_i}之间的相对性与 S_{B_i}, S_{A_i}之间的相对性并不相同，这类似于某种非度量距离测量。以对话中的三个连续话语 $S_{B_{i-1}}, S_{A_i}, S_{B_i}$为例。由于话题在对话中沿着特定方向流动，我们将 S_{A_i}的强相对性定义为：S_{A_i}与其下一个回复 S_{B_i}构成一个正样本对。弱相对性指的是 S_{A_i}与其前一个回复 $S_{B_{i-1}}$构成一个弱正样本对。强相对性和弱相对性对于解决话语归属问题至关重要。当一个话语在语义上与其前一个话语相似时，弱相对性的约束可以维持二者的关系，使其不被强相对性破坏。当一个话语与其下一个话语相似时，强相对性可以使它们在语义空间中更加接近。图2生动地展示了这些关系。

在上述设置下，我们设计了绝对相关性损失 l_{AC}和相对相关性损失 l_{RC}。假设 S_{A_i}'和 S_{A_i}''是对 S_{A_i}进行数据增强后得到的两个话语，$\boldsymbol{h}$是话语 S的原始编码器表示。回复S_{A_i} 的绝对相关性损失定义为

$$l_{AC}(S_{A_i}', S_{A_i}'') = -\log \frac{e^{sim(\boldsymbol{h}_{A_i}', \boldsymbol{h}_{A_i}'')/\tau}}{\sum_{S_i \in \bigcup_{j=1}^{2|D|} X_j} \mathbf{1}_{[i \neq A_i]} e^{sim(\boldsymbol{h}_{A_i}', \boldsymbol{h}_i)/\tau}}, \quad (2)$$

其中，$\mathbf{1}_{[i \neq A_i]} \in \{0, 1\}$是指示函数，当且仅当$i \neq A_i$, $sim(\boldsymbol{a}, \boldsymbol{b})$是向量 $\boldsymbol{a}$和 $\boldsymbol{b}$之间的余弦相似度 $\frac{\boldsymbol{a}^\top \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$时，其值为1；$X_j$是第 j 个对话。集合的容量为 $2|D|$，因为我们使用同一个话语的两个增强样本进行对比学习。公式（2）同样适用于说话人 B的回复。相对相关性损失由两个损失组成：强相对性损失 l_{SR}和弱相对性损失 l_{WR}。强相对性损失为

$$l_{SR}(S_{A_i}, S_{B_i}) = -\log \frac{e^{sim(\boldsymbol{h}_{A_i}, \boldsymbol{h}_{B_i})/\tau}}{\sum_{S_i \in \bigcup_{j=1}^{|D|} X_j} \mathbf{1}_{[i \neq A_i]} e^{sim(\boldsymbol{h}_{A_i}, \boldsymbol{h}_i)/\tau}}, \quad (3)$$

其中，S_{B_i}是 S_{A_i}的下一句。同样，弱相对性损失定义为 S_{A_i}及其上一条回复，

$$l_{WR}(S_{A_i}, S_{B_{i-1}}) = -\log \frac{\lambda_1 e^{sim(\boldsymbol{h}_{A_i}, \boldsymbol{h}_{B_{i-1}})/\tau}}{\sum_{S_i \in \bigcup_{j=1}^{|D|} X_j} \mathbf{1}_{[i \neq A_i]} e^{sim(\boldsymbol{h}_{A_i}, \boldsymbol{h}_i)/\tau}}. \quad (4)$$

与 l_{SR}相比，弱相对性体现在系数 λ_1中。总相对相关性损失是强相对性损失和弱相对性损失

$$l_{RC} = l_{SR} + l_{WR}. \quad (5)$$

由于使用小批量梯度下降来优化神经网络，因此大小为 N 的批次的绝对相关性损失和相对相关性损失写作

$$Loss_{AC} = \frac{1}{2N} \sum_{i=1}^{N} [l_{AC}(S_i', S_i'') + l_{AC}(S_i'', S_i')], \quad (6)$$

$$Loss_{RC} = \frac{1}{N-1} [\sum_{i=1}^{N-1} l_{SR}(S_i, S_{i+1}) + \sum_{i=2}^{N} l_{WR}(S_i, S_{i-1})]. \quad (7)$$

话语表示训练中的总损失是绝对相关性损失与相对相关性损失之和

$$Loss_{total} = Loss_{AC} + Loss_{RC}. \quad (8)$$

对比式话语表示的整个自监督训练过程都在双向 Transformer 编码器上进行。训练完成后，编码器会输出话语表示，供下一部分的对话结构建模使用。

3.3 对话结构建模

我们对话语表示进行聚类，旨在将主题相似的话语归集在一起。我们采用以余弦相似度为依据的 K-Means 聚类方法，获得了良好的聚类性能，同时保持了简洁性。

话语表示被归入 k 个主题级簇 C_1，C_2，……，C_k，其中心向量分别为 $\boldsymbol{c}_1$，$\boldsymbol{c}_2$，……，$\boldsymbol{c}_k$。如图2所示，我们的目标是对话结构图 G = {V，E}进行建模，其中主题级簇被视为顶点，而指向其他顶点的转移概率则被视为边。在 CTRLStruct 中，与其他方法不同，我们不会显式地对特定主题簇内部的话语转移进行建模，因为大型预训练语言模型能够解决这一问题。借助这样的主题级对话结构图，我们的模型能够在对话过程中预测是继续讨论当前主题，还是转向另一主题。

模仿学习 [1, 16] 被应用于转移概率计算。然而，在我们所面对的问题中，环境或模拟器并不可用。我们只有对话数据集，其中的主题状态转移概率未知；根据强化学习的概念，这些数据集可以视为离线环境。幸运的是，Rajaraman 等人 [35] 证明了在状态转移函数未知且无法获得用于模拟的环境这一设定下，行为克隆的最优性。行为克隆 [2, 31] 的思想是利用给定的高质量数据集恢复专家策略。由于开放域对话系统的最终目标是实现人类水平的对话质量，因此我们应用行为克隆，从高质量对话数据集中的专家轨迹中模仿主题转移。在行为克隆设定中，将话语表示bmh_i作为状态，并将簇中心向量$\mathbf{c}_j$视为动作。状态空间和动作空间的大小分别为 $\mathrm{S} \in \mathbb{R}^{u \times n}$和 $\mathrm{A} \in \mathbb{R}^{k \times n}$，其中 u是话语总数，k是簇的数量。在时间步 t，将采取动作 $\boldsymbol{c}_{t+1}$（即状态 $\boldsymbol{h}_{t+1}$所属簇的中心向量），并使状态 $\boldsymbol{h}_t$转移到状态 $\boldsymbol{h}_{t+1}$。因此，专家轨迹可以写为 $\boldsymbol{h}_1$，$\boldsymbol{c}_2$，$\boldsymbol{h}_2$，$\boldsymbol{c}_3$，$\boldsymbol{h}_3$，…… $\boldsymbol{h}_m$。

根据上述设定，状态空间 S 可视为非离散空间，因为 u 规模过大，而动作空间 A 是离散空间。然而，在实际训练过程中，考虑到话题（动作）的多样性，我们将动作空间视为连续空间。我们选取与计算得到的动作在余弦相似度上最接近的簇中心向量，作为当前状态下最终采取的动作。在这种情况下，对于行为克隆，我们通常使用最大似然估计（MLE）直接估计策略 $\widehat{\pi}_\theta$。由于 S、A $\in \mathbb{R}$，优化目标可以写为

$$\max_{\theta} \sum_{(\boldsymbol{h},\boldsymbol{c})\in\mathcal{D}} \log\left(\widehat{\pi}_\theta(\boldsymbol{c} \mid \boldsymbol{h})\right). \tag{9}$$

与大多数行为克隆方法一样，在连续动作或状态空间中，我们采用高斯分布来表示策略 [40]。对于每个状态 $\boldsymbol{h}$，我们假设策略 $\widehat{\pi}_\theta(\cdot|\boldsymbol{h}) \sim \mathrm{N}\left(\mu_\theta(\boldsymbol{h})\ \sigma_\theta^2(\boldsymbol{h})\right)$，其中 $\mu_\theta(\boldsymbol{h})$ 和 $\sigma_\theta^2(\boldsymbol{h})$ 分别为均值和方差

$$\widehat{\pi}_\theta(\cdot|\boldsymbol{h}) = \frac{1}{\sqrt{2\pi}\sigma_\theta(\boldsymbol{h})} e^{-\frac{(\boldsymbol{c}-\mu_\theta(\boldsymbol{h}))^2}{2\sigma_\theta(\boldsymbol{h})^2}}. \tag{10}$$

根据公式（10），公式（9）可简化为

$$\min_{\theta} \sum_{(\boldsymbol{h},\boldsymbol{c})\in\mathcal{D}} \frac{(\boldsymbol{c}-\mu_\theta(\boldsymbol{h}))^2}{2\sigma_\theta^2(\boldsymbol{h})} + \frac{1}{2}\log\left(2\pi\sigma_\theta^2(\boldsymbol{h})\right). \tag{11}$$

我们使用 Actor-to-Critic网络 [43] 来估计 $\mu_\theta(\boldsymbol{h})$ 的值。具体而言，Actor网络和Critic网络均由全连接层组成。Critic网络根据状态估计值，而Actor网络则确定在特定状态下应采取何种动作。由于行为克隆中没有奖励，我们在Actor-to-Critic网络中不考虑奖励。方差 $\sigma_\theta^2(\boldsymbol{h})$ 被设为与参数 θ 无关的常数。因此，最终需要优化的目标转化为均方误差回归问题

$$\min_{\theta} \sum_{(\boldsymbol{h},\boldsymbol{c})\in\mathcal{D}} (\boldsymbol{c}-\mu_\theta(\boldsymbol{h}))^2. \tag{12}$$

利用神经网络解决回归问题后，策略 π_θ可以预测智能体在特定状态下将采取的动作及其相应概率。

3.4 对话结构可控生成

在对话生成阶段，我们遵循先前的工作 [32]，将各词元上的联合概率分解为条件概率的乘积

$$p(Y) = \prod_{i=1}^{n} P\left(t_i \mid t_1, \ldots, t_{i-1}\right), \tag{13}$$

其中，Y是一个包含词元序列 t_1、t_2、.....、t_n的回复。自回归模型依靠具有从左到右注意力的Transformer解码器模块，逐个生成词元。在我们的编码器-解码器模型CTRLStruct中，对话历史 X_i 被送入编码器，解码器以自回归方式生成回复。我们对原始回复的负对数似然进行优化：

$$l_{NLL}(Y) = -\sum_{i=1}^{n} \log P(t_i \mid t_1, \ldots, t_{i-1}). \tag{14}$$

对话结构图以如下方式集成到自回归Transformer解码器中。下一个回复将处于哪个簇 C_{i+1} 可以通过 $\pi_\theta(\cdot|\boldsymbol{h}_i)$ 进行预测。因此，当自回归模型完成词元输出后，我们对模型最后一个隐藏状态的输出进行均值池化，以获得生成话语的表示 $\boldsymbol{h}_{i+1}$。该表示向量和簇中心向量 $\boldsymbol{c}_{i+1}$ 也可视为一维概率分布，分别表示话语在语义空间和主题空间中的分布。我们采用 Kullback-Leibler散度 [19] 使语义空间中的分布更接近主题空间中的分布，从而确保生成的话语与其预期主题相关。Kullback-Leibler散度可写为

$$D_{KL}(\boldsymbol{h}||\boldsymbol{c}) = \sum_{x\in X} [\boldsymbol{h}_{i+1}(x) \log \frac{\boldsymbol{h}_{i+1}(x)}{\boldsymbol{c}_{i+1}(x)}]. \tag{15}$$

解码器中要优化的总损失是负对数似然损失与Kullback-Leibler散度之和

$$Loss_{Gen} = l_{NLL} + \lambda_2 D_{KL}. \tag{16}$$

随后，CTRLStruct的训练过程结束。我们可以像其他基于Transformer的生成式语言模型一样，将该编码器-解码器模型应用于对话生成。

4 实验设置

4.1 研究问题

我们旨在通过对响应生成和话语表示进行实验，回答以下研究问题：

- **RQ1**：与多个强大的对话生成模型相比，CTRLStruct 在开放域对话中的表现如何?
- **RQ2**：CTRLStruct 是否真正控制了回复的主题？所发现的对话结构是否有助于响应生成?
- **RQ3**：与其他句子嵌入方法相比，CTRLStruct 在对话话语表示方面的表现如何？语义相似的话语能否在 CTRLStruct 中聚类?
- **RQ4**：CTRLStruct 的泛化能力如何？它能否应用于采用其他类型骨干网络的模型?

4.2 数据集

为了评估对话结构学习在响应生成中的性能，我们在两个流行的开放域对话数据集 DailyDialog [24] 和 PersonaChat [52]上开展了实验。DailyDialog 是一个由人工编写的多轮对话数据集，噪声较少，涵盖日常生活中的各种话题。我们使用作者随机划分的数据集，训练集/验证集/测试集分别包含 11,118/1,000/1,000 个对话。在 PersonaChat 中，闲聊代理被赋予可配置且一致的人物设定，以生成更加个性化、具体且引人入胜的对话。PersonaChat 的原始版本划分为训练集/验证集/测试集，分别包含 10,907/1,000/968 个对话。在我们的工作中，DailyDialog 和 PersonaChat 中不包含任何额外标签或人物设定信息的对话用于模型训练。

4.3 对比方法

在对话生成和话语表示方面，CTRLStruct 与多个强基线模型进行了比较。为评估对话生成的质量，选取了传统模型和大型预训练语言模型：

表1：在两个没有额外标注信息的开放域对话数据集上，对我们的模型和另外五个强大的对话生成模型进行的自动评估和人工评估。最高值以粗体显示。

数据集	模型	自动评估			人工评估			
		BLEU-1/2	Distinct-1/2	ROUGE-L	流畅性	连贯性	信息丰富度	总体
PersonaChat	序列到序列	0.170 / 0.023	0.008 / 0.032	0.062	1.72	0.40	1.24	0.98
	CVAE	0.192 / 0.041	0.018 / 0.073	0.059	1.76	0.64	1.36	1.28
	BART	0.291 / 0.109	0.030 / 0.107	0.154	1.84	0.82	1.42	1.38
	DialoGPT	**0.324**/ 0.114	0.031 / 0.137	0.188	1.84	1.04	**1.80**	1.58
	BlenderBot	0.280 / 0.112	**0.036**/ **0.167**	**0.198**	**1.88**	1.20	1.76	1.52
	不含 Total 的 CTRLStruct _Loss	0.267 / 0.092	0.024 / 0.077	0.151	1.80	0.76	1.54	1.26
	CTRLStruct（无 WR）_Loss	0.306 / 0.113	0.030 / 0.111	0.159	1.86	1.28	1.70	1.56
	CTRLStruct	0.316 /**0.119**	0.032 / 0.114	0.161	**1.88**	**1.40**	1.72	**1.62**
DailyDialog	序列到序列	0.309 / 0.036	0.068 / 0.220	0.049	1.86	0.46	0.62	0.94
	CVAE	0.381 / 0.138	0.072 / 0.299	0.046	1.86	0.58	0.72	1.18
	BART	0.364 / 0.141	0.112 / 0.378	0.075	1.92	1.66	1.38	1.54
	DialoGPT	0.353 / 0.134	0.106 / 0.352	0.105	1.90	1.60	1.46	1.60
	BlenderBot	0.335 / 0.124	0.111 / 0.340	**0.108**	1.94	1.64	1.62	1.72
	不含 Total 的 CTRLStruct _Loss	0.346 / 0.130	0.095 / 0.328	0.076	1.90	1.34	1.24	1.36
	CTRLStruct（无 WR）_Loss	0.388 / 0.151	0.112 / 0.387	0.077	1.92	**1.72**	1.56	1.70
	CTRLStruct	**0.397**/ **0.157**	**0.118** /**0.402**	0.080	**1.96**	**1.72**	**1.68**	**1.78**

• **序列到序列**：用于机器翻译的基础序列到序列模型，采用注意力机制 [47]。
• **条件变分自编码器**：条件变分自编码器 [54] 是一种将变分自编码器 [18] 与循环神经网络序列到序列框架相结合的生成模型。
• **BART**：BART [21] 通过结合双向 Transformer 和自回归 Transformer，预训练序列到序列模型。实验中使用 BART-large。

• **DialoGPT**：DialoGPT [53] 扩展了 GPT-2 [33]，通过 Reddit 数据训练来处理对话神经响应生成。实验中使用 DialoGPT 的中型版本。

• **BlenderBot**：BlenderBot [39] 采用 Blended Skill Talk 设置，是表现最出色的开放域聊天机器人之一。我们的实验使用其 27 亿参数模型。

为了评估话语表示能力，我们选择了以下方法：

• **BERT**：BERT [9] 由双向 Transformer 编码器组成，并通过掩码语言建模和下一句预测任务进行预训练。

• **无监督SimCSE**：SimCSE 的无监督版本 [11]将同一话语的两个增强样本视为正样本对，并进行对比学习，是当前最先进的句子嵌入框架。

4.4 评估指标

我们评估对话生成的性能，其中同时采用自动评估指标和人工评估指标。自动评估指标包括 BLEU-1/2 [28], Distinct-1/2 [23] 和 ROUGE-L [25]。BLEU、Distinct 和 ROUGE-L 分别衡量生成的话语与真实值之间的 n元语法重叠、生成多样性，以及生成的话语与真实值之间最长公共子序列的数量。在人工评估中，我们遵循 PLATO [3],中的设置，要求评估者从流畅性、连贯性这四个方面进行评分，评分范围为 {0、1、2}，信息丰富度和总体表现。零表示表现不佳，一表示正常，二表示回复良好。

此外，我们还分析生成的回复是否确实与其预期主题相关。我们设计了两种评估指标，分别称为硬主题命中准确率（HTHA）和软主题命中准确率（STHA），旨在证明与其他模型相比，CTRLStruct 能够控制生成过程中的主题流。HTHA 定义为生成的回复中主题簇与其伪标签相匹配的比例。此处的伪标签是指真实响应的簇标识编号，该编号通过 CTRLStruct 的对话结构建模部分获得。STHA 定义为生成的回复中主题簇与其伪标签相似的比例。相似度通过生成回复的簇中心向量与真实值簇中心向量之间的余弦距离来衡量。如果相似度超过给定阈值 φ，则认为生成的回复与其伪标签匹配。HTHA 可视为阈值为 $\varphi = 1.00$ 的特殊 STHA。此外，我们还采用宏平均$F1$ 和微平均$F1$ 来评估模型的主题感知生成性能。

对话中的无监督话语表示评估是一项挑战。我们与 Reimers 和 Gurevych [36] 持有相同的观点：具有良好表达能力的话语表示有助于将语义相似的句子聚为一簇。由于簇的真实值标签未知，因此使用 Calinski-Harabasz 指数 [5] 和 Davies-Bouldin 指数 [8] 等内部评估指标来评估簇的质量。Calinski-Harabasz 指数得分越高、Davies-Bouldin 指数得分越低，表明簇的定义和分离效果越好。

4.5 实现细节

在话语表示训练中，我们将弱相对性损失的系数 λ_1 设为 0。2。我们发现，较大的 λ_1 会导致模型训练崩溃。通过 K-Means 将话语表示划分为 60 个簇。模型训练 20

表2：实验结果展示了生成的回复及其伪标签的主题级相关性。HTHA 和 STHA 分别是 Hard Topic Hit Accuracy（硬主题命中准确率）和软主题命中准确率的缩写。最高值以粗体标出。

数据集	模型	宏平均-$F1$	微型-$F1$	HTHA	STHA ($\varphi = 0.95$)	STHA ($\varphi = 0.90$)	STHA ($\varphi = 0.85$)	STHA($\varphi = 0.80$)
PersonaChat	序列到序列	0.02	0.02	2.97%	2.97%	2.97%	3.09%	5.75%
	CVAE	0.02	0.03	3.51%	3.51%	3.51%	3.71%	7.20%
	BART	0.03	0.06	8.12%	8.12%	8.12%	8.85%	18.00%
	DialoGPT	0.03	0.05	7.80%	7.80%	7.80%	8.33%	16.17%
	BlenderBot	0.03	0.07	8.75%	8.75%	8.75%	9.56%	21.01%
	CTRLStruct（不含 Total）_Loss	0.03	0.06	7.83%	7.83%	7.83%	8.40%	17.54%
	CTRLStruct（不含 WR）_Loss	**0.04**	**0.08**	9.27%	9.27%	9.27%	10.13%	24.32%
	CTRLStruct	**0.04**	**0.08**	**9.52%**	**9.52%**	**9.52%**	**10.46%**	**26.54%**
DailyDialog	序列到序列	0.09	0.07	9.43%	9.43%	9.43%	13.00%	18.49%
	CVAE	0.09	0.06	9.22%	9.22%	9.22%	12.49%	17.97%
	BART	0.14	0.12	16.38%	16.38%	16.38%	19.86%	27.81%
	DialoGPT	0.14	0.11	13.22%	13.22%	13.22%	16.41%	23.14%
	BlenderBot	0.16	0.14	17.89%	17.89%	17.89%	22.05%	31.14%
	CTRLStruct（不含总计）_Loss	0.14	0.12	15.81%	15.81%	15.81%	19.32%	26.62%
	不含 WR 的 CTRLStruct _Loss	0.19	**0.18**	22.86%	22.86%	22.86%	27.27%	37.08%
	CTRLStruct	**0.20**	**0.18**	**23.05%**	**23.05%**	**23.05%**	**27.78%**	**37.70%**

个 epoch，批大小为 256。模型的温度系数设为 0。05。我们使用 nlpaug 包 [27] 进行数据增强。[2]

在生成部分，我们的模型选择预训练的 BART-large 作为编码器和解码器。CTRLStruct 中的行为克隆采用 Actor-to-Critic网络。KL散度的系数 λ_2 设为 1。2，以便在两个数据集上都达到最佳性能。在自动评估中，我们对预训练语言模型使用集束搜索；而在人工评估和主题匹配评估中，则采用 top-p 采样 [15] 和 top-k 采样 [10] 来生成更多样化的对话。

5 结果与分析

5.1 对话回复生成（RQ1）

正如对话评估的实证研究所指出的 [26]，一些自动评估指标最初是为机器翻译设计的，无法系统地评估生成对话的质量。因此，我们将自动评估与人工评估结合起来。实验结果列于表1。在自动评估中，CTRLStruct 在 DailyDialog 中取得了最高的 BLEU-1/2 分数，并在 PersonaChat 中取得了最高的 BLEU-2 分数，而 DialoGPT 的 BLEU-1 得分更高。我们的模型在 DailyDialog 数据集上的 Distinct-1/2 指标上仍领先，但在 PersonaChat 数据集上未能超越 DialoGPT。在 ROUGE-L 指标上，BlenderBot 在两个数据集上的表现均最佳。

在人类评估方面，我们邀请了五十名参与者，在双盲实验中与八个使用不同数据集训练的模型进行交流。需要注意的是，为了公平比较，评估者知道哪些模型是使用相同数据集训练的。结果表明，所有模型都能够生成流畅的对话，且没有明显的语法或拼写错误。CTRLStruct 的优势体现在连贯性评估中，其最高分比最低分高出一分以上。在信息量评估中，DialoGPT 在 PersonaChat 数据集上表现更好；而在 DailyDialog 数据集上，CTRLStruct 能够从词汇和结构角度生成复杂的话语。在综合评估中，CTRLStruct 在两个数据集上都获得了最高分。所有实验均表明，总体而言，CTRLStruct 的表现优于几种强大的对话生成模型。

5.2 主题控制质量（RQ2）

在本部分中，我们测试不同模型生成的回复是否与其预设话题相关，以及对话结构在生成过程中控制话题的效果如何。我们将测试集中真实答案的话题视为预设话题，并将评估转化为一个多分类问题：生成的回复的话题需要被正确分类为与其真实响应相同的话题。由于无法得知话题的真实标签，因此评估中使用伪标签，即话题簇的编号。这些伪标签通过 CTRLStruct 的对话结构建模部分中的聚类过程获得。为确保公平比较，生成的回复的标签也通过 CTRLStruct 的对话结构建模部分计算。因此，这个多分类问题被表述为：给定 k 个类别（k 为簇的总数）和若干样本（生成的回复），我们需要评估模型是否能够正确分类话题。

表2展示了在 DailyDialog 和 PersonaChat 中评估不同模型生成的回复与主题相关性的实验结果。在两个数据集上，CTRLStruct 在宏平均$F1$、微平均$F1$、硬主题命中准确率（HTHA）和软主题命中准确率（STHA）方面均优于其他模型，这表明与其他基线方法相比，我们的方法能够生成与主题相关性更高的回复。当相似度阈值为 $\varphi \geq 0.90$时，HTHA 和 STHA 的结果相同，表明主题分割效果良好。随着相似度约束放宽，主题命中准确率大幅上升。与 DailyDialog 的结果相比，PersonaChat 上的得分相当低。我们将这一现象归因于 PersonaChat 数据集质量较低。PersonaChat 并非由像 DailyDialog 那样的人类对话构成，且其主题多样性较低。因此，当我们预设的话题类别多于数据集实际包含的话题类别时，会引入大量噪声，从而损害 CTRLStruct 的性能。

[2]https://github.com/makcedward/nlpaug

表 3：展示 CTRL-Struct 在两个数据集上对 GPT2 的泛化能力的结果。

数据集	模型	HTHASTHA(0.85)	STHA(0.80)	
PersonaChat	GPT2	5.72%	6.04%	12.02%
	GPT2 + CTRLStruct	**6.84%**	**7.28%**	**14.88%**
DailyDialog	GPT2	11.73%	14.30%	20.59%
	GPT2 + CTRLStruct	**16.95%**	**20.70%**	**29.13%**

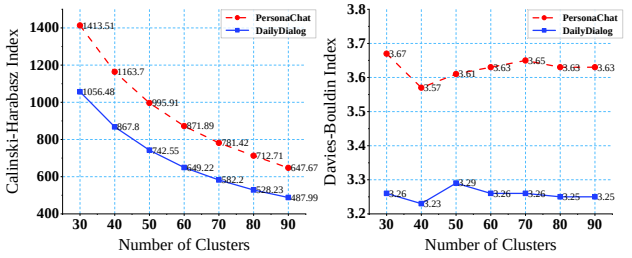

图 3：展示不同聚类数量对聚类性能影响的折线图。

5.3 话语表示（RQ3）

基于这样的假设：含义相似的话语在表示空间中应彼此更接近，我们通过对话语表示进行聚类，并分析各簇的指标，来评估不同模型的话语表示质量。如表4所示，在两个数据集上，CTRLStruct 在 Calinski-Harabasz 指数和 Davies-Bouldin 指数方面均优于其他方法。BERT 和 SimCSE 在 DailyDialog 上的得分相近，但后者在 PersonaChat 上的表现较差。

此外，我们还测试了簇数量对 CTRLStruct 的影响。如图 3 所示，随着簇数量从 30 增加到 90，Calinski-Harabasz 指数先快速下降，随后趋于缓慢下降。至于 Davies-Bouldin 指数，在两个数据集上，随着簇数量增加，其数值均有所波动。还需考虑的一点是，主题分割越细，所能实现的对话结构建模质量就越高。较少的话题数量会影响对话结构的控制能力。因此，在不损害聚类质量的前提下，应尽可能增加簇的数量。综合所有因素，我们得出结论：在 DailyDialog 中，K-Means 的最佳 K 数量为 60，而在 PersonaChat 中，最佳 K 数量为 50。总的来说，CTRLStruct 对簇数量具有较强的鲁棒性。

5.4 泛化能力（RQ4）

CTRLStruct 是一种从未标注语料库中发现对话结构并进行可控生成的框架，因此不依赖于特定类型的骨干模型。原始 CTRLStruct 使用 BART 作为骨干网络，其采用编码器-解码器架构。我们在 GPT2 [33]（仅解码器架构）上进行实验，以检验 CTRLStruct 能否提升其连贯性。结果列于表 3。GPT2 集成 CTRLStruct 后，主题命中率有所提升。不过需要注意的是所提出的对话结构是通过 Transformer 编码器生成的，因此无法应用于非 Transformer 网络。

表4：话语表示上的聚类性能结果，其中 CHI 和 DBI 分别表示 Calinski-Harabasz 指数和 Davies-Bouldin 指数。弱相对性损失缩写为 WR_Loss。

模型		DailyDialog 卡林斯基-哈拉巴斯指数 ↑	DailyDialog DBI↓	PersonaChat 卡林斯基-哈拉巴斯指数 ↑	PersonaChat 戴维斯-博尔丁指数 ↓
BERT		566.79	3.62	848.64	3.83
SimCSE		566.64	3.57	386.52	4.36
不含 WR 的 CTRLStruct	_Loss	618.42	3.28	857.78	3.65
CTRLStruct		**649.22**	**3.26**	**871.89**	**3.63**

5.5 消融研究

我们从三个方面对 CTRLStruct 进行消融研究：话语表示、响应生成和对话结构控制。如表1所示，BERT 输出原始表示，可视为未使用总损失的 CTRLStruct。SimCSE 相当于未使用相对相关性损失的 CTRLStruct。我们注意到，在绝对相关性损失和相对相关性损失的约束下，话语表示的性能逐步提升。在生成阶段，我们评估了未使用弱相对性损失的 CTRLStruct，以及未使用总损失的 CTRLStruct。表1中的结果表明，CTRLStruct 的表现优于未使用弱相对性损失的模型，而未使用总损失的 CTRLStruct 甚至不如 BART。在对话结构控制阶段，图2显示了类似结果。弱相对性损失的约束可以略微提高 CTRLStruct 的主题命中准确率。但未使用总损失的 CTRLStruct 在所有指标上的表现都很差。我们认为，较差的话语表示会导致错误的对话结构，进而误导响应生成，并导致主题控制和响应生成的表现都不佳。这与以往的观点一致：良好的话语表示有助于形成更好的对话结构，而良好的对话结构则有助于生成更好的对话。

5.6 案例分析

我们分析了响应生成和对话结构建模中的一些案例，包括与 CTRLStruct 的交流以展示其优势，以及从不同主题簇中随机选取的话语。受篇幅限制，详情请见附录。

6 结论

本文提出了一种名为 CTRLStruct 的新型框架，该框架结合对比学习、聚类和模仿学习，有效捕捉对话话语表示并构建主题级对话结构。此外，我们将有助于追踪对话中主题流的对话结构融入开放域响应生成。实验结果证实，在生成连贯且与主题相关的对话方面，CTRLStruct 优于其他表现强劲的对话生成模型。

致谢

本研究得到国家自然科学基金（编号：62106105）、CCF-腾讯开放研究基金（编号：RAGR20220122）、南京航空航天大学科研启动基金（编号：YQR21022）以及南京航空航天大学高性能计算平台的支持。

参考文献

[1] Brenna D. Argall、Sonia Chernova、Manuela M. Veloso 和 Brett Browning。2009年。关于示范学习机器人的综述。《机器人与自主系统》57，5（2009），469–483。https://doi.org/10.1016/j.robot.2008.10.024。[2] Michael Bain 和 Claude Sammut。1995。行为克隆框架。载于《机器智能》第*15*卷，智能体 *[英国牛津圣凯瑟琳学院，7月 1995]*，Koichi Furukawa、Donald Michie 和 Stephen H. Muggleton（编）。牛津大学出版社，103–129。[3] Siqi Bao、Huang He、Fan Wang、Hua Wu 和 Haifeng Wang。2020年。PLATO：具有离散潜变量的预训练对话生成模型。载于第*58*届计算语言学协会年会，*ACL 2020*，线上，*2020*年*7*月*5*日至*10*日，Dan Jurafsky、Joyce Chai、Natalie Schluter 和 Joel R. Tetreault（编）。计算语言学协会，85–96。https://doi.org/10.18653/v1/2020.acl-main.9。[4] Tom B. Brown、Benjamin Mann、Nick Ryder、Melanie Subbiah、Jared Kaplan、Prafulla Dhariwal、Arvind Neelakantan、Pranav Shyam、Girish Sastry、Amanda Askell、Sandhini Agarwal、Ariel Herbert-Voss、Gretchen Krueger、Tom Henighan、Rewon Child、Aditya Ramesh、Daniel M. Ziegler、Jeffrey Wu、Clemens Winter、Christopher Hesse、Mark Chen、Eric Sigler、Mateusz Litwin、Scott Gray、Benjamin Chess、Jack Clark、Christopher Berner、Sam McCandlish、Alec Radford、Ilya Sutskever 和 Dario Amodei。2020年。语言模型是少样本学习者。载于神经信息处理系统进展*33*：*2020*年神经信息处理系统年会，*NeurIPS 2020*，*2020*年*12*月*6*日至*12*日，线上，Hugo Larochelle、Marc'Aurelio Ranzato、Raia Hadsell、Maria-Florina Balcan 和 Hsuan-Tien Lin（编）。https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html[5] Tadeusz Caliński 和 Jerzy Harabasz。1974年。用于聚类分析的树突法。《统计学通信：理论与方法》3，1（1974），1–27。[6] Ting Chen、Simon Kornblith、Mohammad Norouzi 和 Geoffrey E. Hinton。2020年。视觉表征对比学习的简单框架。载于第*37*届国际机器学习会议，*ICML 2020*，*2020*年*7*月*13*日至*18*日，线上活动（机器学习研究论文集，第*119*卷）。PMLR，1597–1607。http://proceedings.mlr.press/v119/chen20j.html [7] Ananlada Chotimongkol。2008年。从领域内对话语料库中学习面向任务对话的结构。博士论文。卡内基梅隆大学语言技术研究所计算机科学学院。[8] David L Davies 和 Donald W Bouldin。1979年。聚类分离度量。《*IEEE*模式分析与机器智能汇刊》2（1979），224–227。[9] Jacob Devlin、Ming-Wei Chang、Kenton Lee 和 Kristina Toutanova。2019年。BERT：用于语言理解的预训练深度双向变换器。载于*2019*年北美计算语言学协会年会：人类语言技术会议论文集，*NAACL-HLT 2019*，美国明尼苏达州明尼阿波利斯，*2019*年*6*月*2*日至*7*日，第*1*卷（长篇与短篇论文），Jill Burstein、Christy Doran 和 Thamar Solorio（编）。计算语言学协会，4171–4186。https://doi.org/10.18653/v1/N19-1423。[10] Angela Fan、Mike Lewis 和 Yann N. Dauphin。2018。层次化神经故事生成。载于第*56*届计算语言学协会年会，*ACL 2018*，澳大利亚墨尔本，*2018*年*7*月*15*日至*20*日，第*1*卷：长篇论文，Iryna Gurevych 和 Yusuke Miyao（编）。计算语言学协会，889–898。https://doi.org/10.18653/v1/P18-1082。[11] Tianyu Gao、Xingcheng Yao 和 Danqi Chen。2021。SimCSE：句子嵌入的简单对比学习。载于*2021*年自然语言处理实证方法会议论文集，*EMNLP 2021*，线上活动/多米尼加共和国蓬塔卡纳，*2021*年*11*月*7*日至*11*日，Marie-Francine Moens、Xuanjing Huang、Lucia Specia 和 Scott Wen-tau Yih（编）。计算语言学协会，6894–6910。https://doi.org/10.18653/v1/2021.emnlp-main.552。[12] R. Chulaka Gunasekara、David Nahamoo、Lazaros C. Polymenakos、David Echev-erría Ciaurri、Jatin Ganhotra 和 Kshitij P. Fadnis。2019年。量化对话——一种对话系统通用方法。《计算机语音与语言》54（2019年17月30日）：101016。https://doi.org/10.1016/j.csl.2018.06.003。[13] Michael Gutmann 和 Aapo Hyvärinen。2010。噪声对比估计：一种用于非归一化统计模型的新估计原理。载于第十三届人工智能与统计国际会议论文集，*AISTATS 2010*，基亚拉古纳度假村，意大利撒丁岛，*2010*年*5*月*13*日至*15*日（*JMLR* 会议论文集，第*9*卷），Yee Whye Teh 和 D. Mike Titterington（编）。JMLR.org，297–304。http://proceedings.mlr.press/v9/gutmann10a.html

[14] 何恺明、范浩奇、吴雨欣、谢赛宁和 Ross B. Girshick。2020。用于无监督视觉表征学习的动量对比。载于*2020*年*IEEE/CVF*计算机视觉与模式识别会议，*CVPR 2020*，美国华盛顿州西雅图，*2020*年*6*月*13*日至*19*日。计算机视觉基金会 / IEEE，9726–9735。https://doi.org/10.1109/CVPR42600.2020.00975。[15] Ari Holtzman、Jan Buys、杜立、Maxwell Forbes 和崔艺真。2020。神经文本退化的奇特案例。载于第*8*届国际学习表征会议，*ICLR 2020*，埃塞俄比亚亚的斯亚贝巴，*2020*年*4*月*26*日至*30*日。OpenReview.net。https://openreview.net/forum?id=rygGQyrFvH [16] Ahmed Hussein、Mohamed Medhat Gaber、Eyad Elyan 和 Chrisina Jayne。2017。模仿学习：学习方法综述。*ACM*计算评论50, 2 (2017), 21:1–21:35。https://doi.org/10.1145/3054912。[17] Daniel Jurafsky 和 Elizabeth Shriberg。1997。Switchboard SWBD-DAMSL 浅层话语功能标注员手册，第13稿 Daniel Juraf- sky*、Elizabeth Shriberg+ 和 Debra Biasca** 科罗拉多大学博尔德分校 &+ SRI国际。(1997)。[18] Diederik P. Kingma 和 Max Welling。2014。自编码变分贝叶斯。载于第*2*届国际学习表征会议，*ICLR 2014*，加拿大艾伯塔省班夫，*2014*年*4*月*14*日至*16*日，会议论文集，Yoshua Bengio 和 Yann LeCun（编）。http://arxiv.org/abs/1312.6114。[19] Solomon Kullback 和 Richard A Leibler。1951。论信息与充分性。数学统计年刊22, 1 (1951), 79–86。[20] Staffan Larsson 和 David R. Traum。2000。TRINDI 对话移动引擎工具包中的信息状态与对话管理。自然语言工程6, 3&4 (2000), 323–340。https://doi.org/10.1017/S1351324900002539[21] Mike Lewis、Yinhan Liu、Naman Goyal、Marjan Ghazvininejad、Abdelrahman Mohamed、Omer Levy、Veselin Stoyanov 和 Luke Zettlemoyer。2020。BART：用于自然语言生成、机器翻译和理解的序列到序列去噪预训练。载于第*58*届计算语言学协会年会论文集，*ACL 2020*，线上，*2020*年*7*月*5*日至*10*日，Dan Jurafsky、Joyce Chai、Natalie Schluter 和 Joel R. Tetreault（编）。计算语言学协会，7871–7880。https://doi.org/10.18653/v1/2020.acl-main.703[22] 李博涵、周昊、何俊贤、王明轩、杨一鸣和李磊。2020。论预训练语言模型生成的句子嵌入。载于*2020*年自然语言处理实证方法会议论文集，*EMNLP 2020*，线上，*2020*年*11*月*16*日至*20*日，Bonnie Webber、Trevor Cohn、何玉兰和刘洋（编）。计算语言学协会，9119–9130。https://doi.org/10.18653/v1/2020.emnlp-main.733。[23] Jiwei Li、Michel Galley、Chris Brockett、Jianfeng Gao 和 Bill Dolan。2016。神经对话模型的多样性促进目标函数。载于*NAACL HLT 2016*，计算语言学协会北美分会人类语言技术会议，*2016*年*6*月*12*日至*17*日，美国加利福尼亚州圣迭戈，Kevin Knight、Ani Nenkova 和 Owen Rambow（编）。计算语言学协会，110–119。https://doi.org/10.18653/v1/N16-1014。[24] 李艳然、苏慧、沈晓宇、Wenjie Li、曹自强和牛树梓。2017。DailyDialog：人工标注的多轮对话数据集。载于第八届国际自然语言处理联合会议论文集，*IJCNLP 2017*，中国台湾台北，*2017*年*11*月*27*日至*12*月*1*日——第*1*卷：长篇论文，Greg Kondrak 和 Taro Watanabe（编）。亚洲自然语言处理联合会，986–995。https://aclanthology.org/I17-1099/ [25] 林振宇。2004。ROUGE：自动评估摘要的软件包。载于文本摘要的拓展。74–81。[26] 刘家伟、Ryan Lowe、Iulian Serban、Michael Noseworthy、Laurent Charlin 和 Joelle Pineau。2016。如何不评估对话系统：对对话回复生成无监督评估指标的实证研究。载于*2016*年自然语言处理实证方法会议论文集，*EMNLP 2016*，美国得克萨斯州奥斯汀，*2016*年*11*月*1*日至*4*日，苏剑、Xavier Carreras 和 Kevin Duh（编）。计算语言学协会，2122–2132。https://doi.org/10.18653/v1/D16-1230。[27] Edward Ma。2019。自然语言处理数据增强。https://github.com/makcedward/nlpaug。[28] Kishore Papineni、Salim Roukos、Todd Ward 和朱伟京。2002。BLEU：一种自动评估机器翻译的方法。载于第*40*届计算语言学协会年会论文集，*2002*年*7*月*6*日至*12*日，美国宾夕法尼亚州费城。ACL，311–318。https://doi.org/10.3115/1073083.1073135[29] 彭伟、Yue Hu、邢鲁西、谢雨强、Yajing Sun 和李云鹏。2022。全局控制，局部理解：用于情感支持对话的全局到局部层次图网络。https://doi.org/10.48550/ARXIV.2204.12749[30] 彭伟、Yue Hu、邢鲁西、谢雨强、Yajing Sun 和李云鹏。2022。全局控制，局部理解：用于情感支持对话的全局到局部层次图网络。载于第三十一届国际人工智能联合会议论文集，*IJCAI 2022*，奥地利维也纳，*2022*年*7*月*23*日至*29*日，Luc De Raedt（编）。ijcai.org，4324–4330。https://doi.org/10.24963/ijcai.2022/600

[31] Dean Pomerleau。1991。用于自主导航的人工神经网络高效训练。神经计算3, 1 (1991), 88–97。https://doi.org/10. 1162/neco.1991.3.1.88[32] Alec Radford、Karthik Narasimhan、Tim Salimans、Ilya Sutskever 等。2018。通过生成式预训练提升语言理解。(2018)。[33] Alec Radford、Jeffrey Wu、Rewon Child、David Luan、Dario Amodei、Ilya Sutskever 等。2019。语言模型是无监督多任务学习器。*OpenAI* 博客1, 8 (2019), 9。[34] Colin Raffel、Noam Shazeer、Adam Roberts、Katherine Lee、Sharan Narang、Michael Matena、Yanqi Zhou、Wei Li 和 Peter J. Liu。2020。借助统一的文本到文本转换器探索迁移学习的极限。机器学习研究期刊21 (2020), 140:1–140:67。http://jmlr.org/papers/v21/20-074.html [35] Nived Rajaraman、Lin F. Yang、Jiantao Jiao 和 Kannan Ramchandran。2020。迈向模仿学习的基本极限。载于神经信息处理系统进展*33*：*2020*年神经信息处理系统年会，*NeurIPS 2020*，*2020*年*12*月*6*日至*12*日，线上，Hugo Larochelle、Marc'Aurelio Ranzato、Raia Hadsell、Maria-Florina Balcan 和 Hsuan-Tien Lin（编）。https://proceedings.neurips.cc/paper/2020/hash/1e7875cf32d306989d80c14308f3a099-Abstract.html[36] Nils Reimers 和 Iryna Gurevych。2019。Sentence-BERT：使用孪生 BERT 网络的句子嵌入。载于*2019*年自然语言处理实证方法会议和第*9*届国际自然语言处理联合会议论文集，*EMNLP-IJCNLP 2019*，中国香港，*2019*年*11*月*3*日至7日，Kentaro Inui、Jing Jiang、Vincent Ng 和 Xiaojun Wan（编）。计算语言学协会，3980–3990。https://doi.org/10 18653 1 19 1410 . /v /D - Alan Ritter、Colin Cherry 和 Bill Dolan。2010。无监督建模 Twitter 对话。载于人类语言技术：计算语言学协会北美分会*2010*年年会。计算语言学协会，加利福尼亚州洛杉矶，172–180。https://aclanthology.org/N 10 1020 - Alan Ritter、Colin Cherry 和 Bill Dolan。2010。无监督建模 Twitter 对话。载于人类语言技术：计算语言学协会北美分会会议论文集，*2010*年*6*月*2*日至*4*日，美国加利福尼亚州洛杉矶。计算语言学协会，172–180。https://aclanthology.org/N10-1020/[39] Stephen Roller、Emily Dinan、Naman Goyal、Da Ju、Mary Williamson、Yinhan Liu、Jing Xu、Myle Ott、Kurt Shuster、Eric Michael Smith、Y-Lan Boureau 和 Jason Weston。2020。构建开放域聊天机器人的方法。*CoRR* abs/2004.13637 (2020 2004 13637 2004 13637)。arXiv: 。https://arxiv.org/abs/ 。Stéphane Ross 和 Drew Bagnell。. 模仿学习的高效约简。载于第十三届国际人工智能与统计会议论文集，*AISTATS 2010*，第*15*卷，*2010*年*9*月*13*日，意大利撒丁岛基亚拉古纳度假村，*5*月——*(JMLR*会议论文集，第 卷*)*，Yee Whye Teh 和 D. Mike Titterington（编）。JMLR.org，661–668。http://proceedings.mlr.press/v9/ross10a.html [41] Weiyan Shi、Tiancheng Zhao 和 Zhou Yu。2019。无监督对话结构学习。载于*2019*年计算语言学协会北美分会人类语言技术会议论文集，*NAACL-HLT 2019*，美国明尼苏达州明尼阿波利斯，*2019*年*6*月*2*日至*7*日，第*1*卷（长篇与短篇论文），Jill Burstein、Christy Doran 和 Thamar Solorio（编）。计算语言学协会，1797–1807。https://doi.org/10.18653/v1/n19-1178[42] Yajing Sun、Yong Shan、Chengguang Tang、Yue Hu、Yinpei Dai、Jing Yu、Jian Sun、Fei Huang 和 Luo Si。2021。通过边增强图自编码器无监督学习确定性对话结构。载于第三十五届*AAAI*人工智能会议，*AAAI 2021*，第三十三届人工智能创新应用会议，*IAAI 2021*，第十一届人工智能教育进展研讨会，*EAAI 2021*，线上活动，*2021*年*2*月*2*日至*9*日。AAAI 出版社，13869–13877。https://ojs.aaai.org/index.php/AAAI/article/ view/17634[43] Richard S. Sutton 和 Andrew G. Barto。1998。强化学习：导论。*IEEE*神经网络汇刊9, 5 (1998), 1054–1054。https://doi. org/ 10 1109 1998 712192 . /TNN. . Chen Tang、Frank Guerin、Yucheng Li 和 Chenghua Lin。2022。神经文本生成的最新进展：任务无关综述。https://doi.org/10.48550/ ARXIV.2203.03047[45] Aäron van den Oord、Yazhe Li 和 Oriol Vinyals。2018。利用对比预测编码进行表征学习。*CoRR* abs/1807.03748 (2018)。arXiv: 1807 03748 1807 03748 。http://arxiv.org/abs/ 。Laurens Van der Maaten 和 Geoffrey Hinton。. 使用 t-SNE 可视化数据。机器学习研究期刊9, 11 (2008)。[47] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N. Gomez、Lukasz Kaiser 和 Illia Polosukhin。2017。注意力就是一切。载于神经信息处理系统进展*30*：*2017*年神经信息处理系统年会，*2017*年*12*月*4*日至*9*日，美国加利福尼亚州长滩，Isabelle Guyon、Ulrike von Luxburg、Samy Bengio、Hanna M. Wallach、Rob Fergus、S. V. N. Vishwanathan 和 Roman Garnett（编）。5998–6008。https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[48] Wenquan Wu、Zhen Guo、Xiangyang Zhou、Hua Wu、Xiyuan Zhang、Rongzhong Lian 和 Haifeng Wang。2019。基于明确对话目标的主动式人机对话。载于第*57*届计算语言学协会年会论文集。计算语言学协会，意大利佛罗伦萨，3794 3804 10 18653 1 19 1369 – 。https://doi.org/ 。/v /P - Jun Xu、Zeyang Lei、Haifeng Wang、Zheng-Yu Niu、Hua Wu 和 Wanxiang Che。2021。发现用于连贯对话生成的对话结构图。载于第*59*届计算语言学协会年会暨第*11*届自然语言处理国际联合会议论文集，*ACL/I-JCNLP 2021*，（第*1*卷：长篇论文），线上活动，*2021*年*8*月*1*日至*6*日，Chengqing Zong、Fei Xia、Wenjie Li 和 Roberto Navigli（编）。计算语言学协会，1726 1739 10 18653 1 2021 136 – 。https://doi.org/ 。/v / .acl-long。Jun Xu、Haifeng Wang、Zhengyu Niu、Hua Wu 和 Wanxiang Che。。用于开放域对话生成的知识图谱接地目标规划。载于第*34*届*AAAI*人工智能会议（*AAAI 2020*）、第*32*届人工智能创新应用会议（*IAAI 2020*）、第十届*AAAI*教育进展研讨会（*EAAI 2020*），美国纽约州纽约市，*2020*年*2*月*7*日至*12*日。AAAI 出版社，9338–9345。https://ojs.aaai.org/index.php/AAAI/article/view/ 2020 0474 。[51] Ke Zhai 和 Jason D. Williams。。发现任务型对话中的潜在结构。载于第*52*届计算语言学协会年会论文集（*ACL 2014*），*2014*年*6*月*22*日至*27*日，美国马里兰州巴尔的摩，第*1*卷：长篇论文。计算机语言学协会，36–46。https://doi.org/10.3115/v1/p14-1004[52] Saizheng Zhang、Emily Dinan、Jack Urbanek、Arthur Szlam、Douwe Kiela 和 Jason Weston。2018。对话代理个性化：我有一只狗，你也养宠物吗？载于第*56*届计算语言学协会年会论文集（*ACL 2018*），澳大利亚墨尔本，*2018*年7月*15*日至*20*日，第*1*卷：长篇论文，Iryna Gurevych 和 Yusuke Miyao（编）。计算语言学协会，2204 2213 10 18653 1 18 1205 – 。https://doi.org/ 。/v /P - Yizhe Zhang、Siqi Sun、Michel Galley、Yen-Chun Chen、Chris Brockett、Xiang Gao、Jianfeng Gao、Jingjing Liu 和 Bill Dolan。2020。DIALOGPT：用于对话响应生成的大规模生成式预训练。载于第*58*届计算语言学协会年会论文集：系统演示，*ACL 2020*，线上，*2020*年7月*5*日至*10*日，Asli Celikyilmaz 和 Tsung-Hsien Wen（编）。计算语言学协会，270–278。https://doi.org/ 10 18653 1 2020 30 。/v / .acl-demos。Tiancheng Zhao、Ran Zhao 和 Maxine Eskénazi。2017。使用条件变分自编码器学习神经对话模型的篇章级多样性。载于第*55*届计算语言学协会年会论文集（*ACL 2017*），加拿大温哥华，7月*30*日至*8*月*4*日，第*1*卷：长篇论文，Regina Barzilay 和 Min-Yen Kan（编）。计算语言学协会，654–664。https://doi.org/10.18653/v1/P17-1061

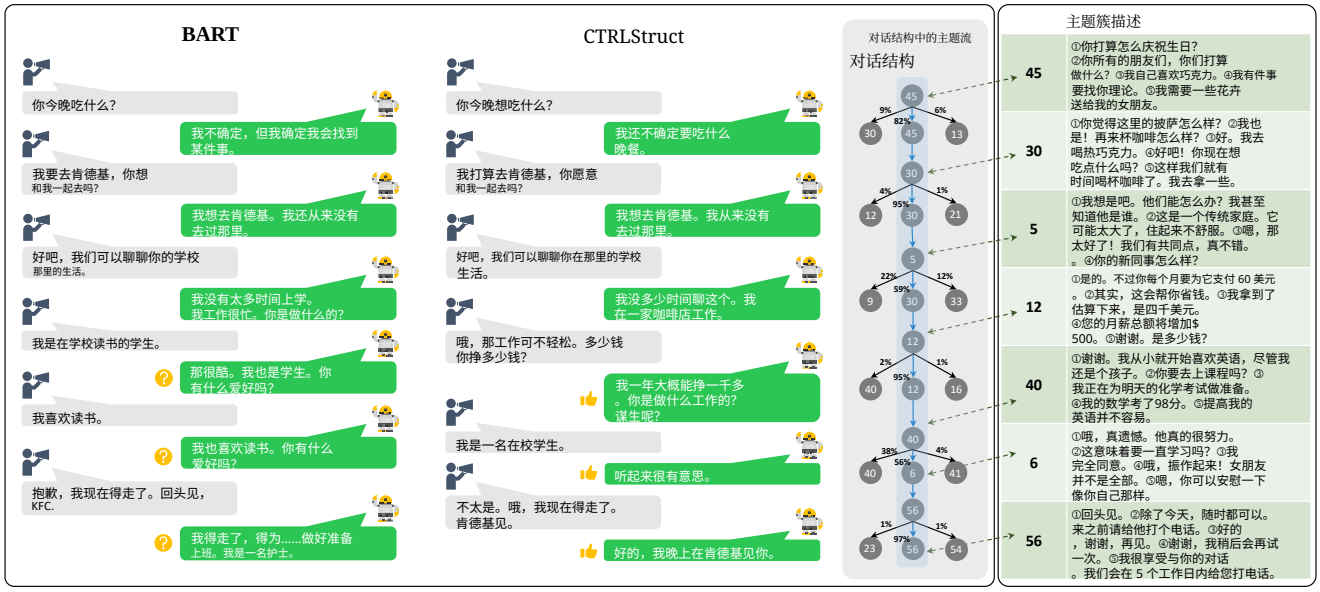

图4：人物A与闲聊机器人B之间的对话历史，其中带有点赞表情符号的回复展示了CTRLStruct相较于BART的优势。右侧展示了具有话题转移概率的对话结构中的主题流。

案例研究

在案例研究中，我们将进一步介绍响应生成和话语表示的细节，并分析这些案例。

A.1 响应生成

为了进一步剖析 CTRLStruct 在对话回复生成方面的能力，我们分别与 BART 模型和 CTRLStruct 模型进行交互，并分析对话历史。选择 BART 进行比较，是因为 CTRLStruct 可以被视为经过对话结构增强的 BART。生成的对话之间的差异能够证明对话结构建模的价值。图4所示的对话发生在一个人 A 和闲聊机器人 B 之间，其中包含数次话题级转换，包括 A 邀请 B 一起吃晚餐、A 在晚餐时谈论学校生活、A 询问 B 的工作，以及 A 向 B 道别。带有点赞表情符号和问号表情符号的发言展示了两个模型之间的差异。我们注意到，BART 模型存在前后不一致的情况，例如“I am a nurse.”和“I am a student,too.”相互矛盾。此外，BART 会反复询问相同的问题：“Do you have any hobbies?”而且，当人类说“See you in KFC.”时，BART 无法给出令人满意的回复。我们认为，这些问题源于模型对主题转移不敏感。CTRLStruct 能够感知多轮对话中的主题，因此大大缓解了这些问题。CTRLStruct 能迅速识别从 B 的工作到 A 的工作的主题转移；当 A 要离开时，我们的模型也能作出恰当回应，而不是说些无关的话。

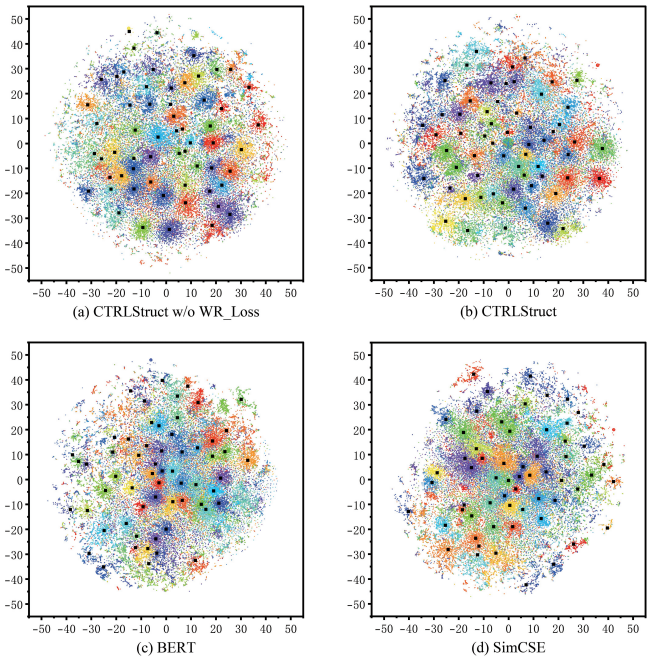

图5：DailyDialog数据集中不同模型经t-SNE处理后的散点图。

我们发现了给定对话历史中的话题转换图，并将其呈现在图4中。蓝色标记的轨迹是由CTRLStruct生成的所选主题流，每个选定主题在其所处阶段都具有最高的转移概率。除该主题外的其他话题话题流图中选中的主题是第二和第三个可能选项。每个主题簇的描述位于右侧，并以绿色标记。描述以随机选取相应簇中的若干话语的形式呈现。对于每个

表 5：从 CTRLStruct 在 PersonaChat 和 DailyDialog 中形成的簇里随机选取的话语示例。

簇	PersonaChat	DailyDialog
a	我敢说那一定是很不错的锻炼。我有两个哥哥，你有几个兄弟姐妹? 听到这个消息我很难过。我有两个哥哥和两个姐姐，都比我年长。 可能比我们待得更久吧。我真希望我的四个姐妹中有一个是美人鱼! 我也是！我们马上要有第二个孩子了。我现在怀孕了。 我也是。最好的朋友都不足以形容。也许算异父兄弟? 哦，那太好了。我的女朋友负责所有的烹饪。你家人多吗?	你对这个组织的哪类业务最感兴趣? 你有做过销售工作的经验吗? 我们的 EDD办公室里有电脑，你可以用来求职。 我希望月薪从3000元起。 我在这里工作了两年。而且我想做些不同的事情。 你会如何形容你与我们老板的关系?
b	我住在普罗维登斯，在市中心的一栋出版社大楼里工作。 你好，你怎么样？我现在很好，因为我不再是奴隶了。 嘿，我叫杰夫，我最喜欢的运动是足球。你怎么样? 嗨，你好吗？我经营一家小型企业。我帮助客户赢钱。 哦，我很喜欢园艺，虽然我有色盲，看不太清花卉! 我是一名职业舞者，喜欢听非乡村音乐。	昨天我和邻居大吵了一架。 我很难获取这些信息。 据我所知，失眠通常是由压力引起的。你有压力吗? 我刚丢了工作。老板刚刚告诉我这个消息。 我们要请几个朋友来家里。你在做什么? 你看起来不太舒服。怎么了?
c	是啊，谢谢你，你真好！我会送你新鲜鸡蛋作为回礼。 也许我们改天可以见面。镇上有一家很棒的墨西哥卷饼餐车。 你是素食者吗？我很喜欢做饭，做一顿素食餐会很不错。 太棒了，这对我来说很合适，因为我受过良好的教育。 我做的冰沙很好喝！我上大学期间在一家冰沙店工作。 我听说游戏也有帮助。我喜欢绘画，也会教他们画画。	我不知道该去哪里找。 你想吃点什么? 你今天想和我一起去做面部护理吗? 你想了解哪方面的信息? 你还不知道想去哪里找吗? 你知道自己想要哪一种吗?
d	你这也太肤浅了，不过我就是喜欢你这一点，Sherry! 对我来说有点太暴力了。我大部分时间都和我的狗一起徒步旅行。 哇，哈哈。真是开门见山啊？我喜欢自己的人生目标，对它们也很满意。 我不介意独自游泳，但我需要有人陪我去听周末音乐会。 哦，那太糟糕了！我也明白这种感受。这种感觉非常真实。我正在开一家杂货店。 我很想住在那里！不过我爸爸在那里找不到工作，他是警察。	你好，我叫 Lean，来自俄罗斯。 您好，Taylor夫人。我来接Diane。 你好，Rachel。很高兴在这里见到你。 她也很好，Helen。再见，Helen。很高兴见到你。 是的，我是 Jin Lili。有什么可以帮您? 早上好，Chan 女士。今天想买点什么?

簇，我们选取五条不同的话语。描述与生成的回复基本一致，这表明 CTRLStruct 能够对高质量的对话结构进行建模，并利用预测的主题信息进行响应生成。

A.2 话语表示

表 5 展示了主题簇中的一些话语样本，直观地呈现 CTRLStruct 汇集语义相似话语的能力。簇及簇中的话语均为随机选取。在 PersonaChat 数据集中，聚类性能并不如预期。簇 **a** 与家族生意有关。簇 **b** 可以定义为兴趣爱好，但其中可能混杂着与工作相关的话题。簇 **c** 和簇 **d** 的质量较差，很难将它们定义为某个特定主题。我们认为这是 PersonaChat 作为开放域对话数据集质量较低所致。PersonaChat 中的话语通常包含太多不同方面的话题，从人的角度来看，并不适合作为回复。话语中嘈杂的语义信息会使模型困惑，降低其性能。

DailyDialog 的质量远高于 PersonaChat，因此 DailyDialog 数据集中的聚类性能也相应更好。簇 **a**中所有选取的话语都与工作有关，包括求职、薪资、人际关系等。簇 **b** 与关心他人有关。簇 **c** 包含不同类型的查询。簇 **d** 是向他人问候。根据从 DailyDialog 中选取的话语，我们认为 CTRLStruct 无需标注信息便能成功捕捉主题层面的含义，并实现高质量聚类。

我们还绘制了簇的散点图，以展示不同模型在DailyDialog 数据集上的表现。这些图通过 t-SNE [46] 绘制，t-SNE可降低原始向量的维度。由于将具有相似话题的话语聚集起来的聚类任务较为复杂，因此在降维处理下效果并不显著。但我们仍可在图5中注意到，BERT和SimCSE的簇周围有大量离群点。簇中的一些样本甚至无法聚集在一起形成某种形状。CTRLStruct在稳定簇形状和消除离群点两方面的表现都是最佳的。

B 局限性与未来工作

本节将讨论我们工作的局限性以及未来可能的改进方向。首先，受限于计算资源，我们将话语表示学习中的批大小设为128。更大的批大小可能会带来更好的对比学习表征性能。其次，在获得话语表示后，CTRLStruct 采用 K-Means 作为聚类方法。实际上，我们测试了包括深度聚类和凝聚层次聚类在内的多种方法，而 K-Means 以最少的计算资源取得了最佳表现。然而，K-Means 存在一些明显缺点，如凸性和各向同性假设。我们希望找到在此问题设定下更适合的聚类方法。第三，在可控生成部分，仅仅将话语表示拉近至主题聚类中心是一种不稳定的控制方法，且无法充分利用所发现的对话结构。我们相信，通过设计新的解码策略或在生成过程中引入符号规则约束，能够有效提升性能。此外，目前几乎没有可获取的大规模高质量开放域对话数据集。正如本文实验结果所表明的，模型性能在很大程度上与数据集质量相关。若能获得高质量开放域对话数据集，未来我们便可以训练 CTRLStruct，并将预训练的对话结构作为常识知识，用于下游对话生成训练。