基于 PLATO-2 的主题切换自适应日语对话系统

DonghuoZeng[*1] Jianming Wu[*1] Yanan Wang[*1] Kazunori Matsumoto Gen Hattori Kazushi Ikeda

[*1]KDDI Research, Inc.

PLATO-2 等大规模开放域对话系统在英语和中文中均取得了当前最先进的成绩。然而，很少有工作探讨此类对话系统是否同样适用于日语。在本工作中，我们创建了一个大规模日语对话数据集 Dialogue-Graph，其中包含来自新闻、电视字幕和维基百科语料库的 165.6 万条采用树结构组织的对话数据。随后，我们使用 Dialogue-Graph 训练 PLATO-2，构建了大规模日语对话系统 PLATO-JDS。此外，为了改善 PLATO-JDS 的主题切换问题，我们引入了一种由主题判别器组成的主题切换算法：当用户输入与前一主题不同时，切换到新主题。我们使用模型通过人机对话策略，从连贯性、信息量、吸引力和拟人性四个指标评估用户体验。结果表明，我们提出的 PLATO-JDS 在人工评估中取得了**1.500** 的平均分，接近 2.000 的最高分，表明 PLATO-2 在日语对话生成方面具备较高质量的能力。此外，我们提出的主题切换算法取得了**1.767** 的平均分，比 PLATO-JDS 高**0.267**，说明其能够有效改善系统的用户体验。

arXiv:2302.11280v1 [cs.CL] 22 Feb 2023

1. 引言

基于Transformer的方法正逐渐成为开发类人聊天机器人的基础技术 [Devlin等，2019，Brown等，2020，Golovanov等，2019，Wolf等，2019，Zhou等，2021，Shalaby等，2020，Adiwardana等，2020，Radford等，2019]。PLATO-2 [Bao 等人，2021a]等大规模开放域对话系统旨在将模型扩展至数十亿参数，从而打造高质量的开放域聊天机器人；这类系统在英语和中文任务上均取得了最先进的成绩。然而，此类大规模对话系统在日语领域鲜少得到探索。

在本研究中，我们从新闻、电视字幕和维基百科语料库中收集了165.6万条树结构的对话数据，构建了大规模日语对话数据集Dialogue-Graph。随后，我们使用Dialogue-Graph训练PLATO-2，构建了大规模日语对话系统PLATO-JDS。此外，我们研究了PLATO-JDS生成的一些案例，结果表明，在对话过程中，PLATO-JDS难以恰当地切换到新主题。为解决这一问题并进一步改善系统的用户体验，我们提出了一种主题切换算法，该算法由一个主题判别器组成：当用户输入与先前主题不同时，系统便切换到新主题。主题判别器是一个基于BERT的二分类器 [Devlin等， 2019] ，用于预测用户输入是否属于先前主题。主题判别器是一个基于BERT的二分类器 [Devlin等，2019] ，用于预测用户输入是否属于先前主题。如图1所示，如果结果为“Yes”，则将用户输入与先前对话拼接后作为PLATO-JDS的输入；否则，仅将用户输入作为输入。

我们从连贯性、信息量、吸引力和拟人性四项指标评估 PLATO-JDS [Bao etal., 2021a], 。结果表明，我们提出的 PLATO-JDS 在采用人机对话策略进行人工评估时，四项指标的得分分别为 1.600、1.467、1.467 和 1.467，平均分为 **1.500** ，接近最高分 2.000，这表明 PLATO-2 可用于实现高质量的日语对话生成。此外，我们提出的主题切换算法平均分为**1.767** ，比 PLATO-JDS 高**0.267**，表明该算法能有效提升系统的用户体验。

联系方式：KDDI Research,Inc., 人工智能部门，356-0003 埼玉县 Fujimino, Ohara, 2 Chome-1-15，049-278-7441，do-zeng@kddi-research.jp。

我们的贡献包括：1）我们人工收集了 165.6 万条对话数据，构建了树状结构的大规模日语数据集*Dialogue-Graph*；2）我们利用*Dialogue-Graph*训练了日语版 PLATO-2——PLATO-JDS。人工评估和自动评估均表明，PLATO-JDS 能够实现高质量的对话生成；3）此外，我们还引入了一种主题切换算法，以进一步提升 PLATO-JDS 的用户体验。

2. 相关工作

日语领域的对话系统采用基础的深度学习技术开发而成，例如人工智能恋爱咨询系统 [Nakatsuji, 2019]。此外，还有一些聊天机器人应用，例如 Clova [Bae 等人， 2022] 和 Rinna [Wu 等人， 2016] ，它们采用先进的自然语言处理方法来提升对话生成的性能。日语语言模型 [Tianyu 和 Sawada，2021]在公开语料库上训练了两种类型的模型，包括 GPT-2 [Radford 等人，2019] 和 RoBERTa [Liu 等人，2019] 模型。考虑到大规模数据集可用于训练高性能对话系统，因此也应将其纳入考量。目前的研究 [Sugiyama 等人，2021，Rikters等， 2019] 尝试收集大规模数据集，以改进日语对话系统。尽管如此，仍鲜有大规模日语对话系统使用大型日语数据集进行训练，例如规模达到数百万条的数据集。

近年来，大规模基于 Transformer 的对话生成模型显著提升了开放域聊天机器人的性能 [Zhang 等人，2020，Bao etal.，2020，Bao et al.，

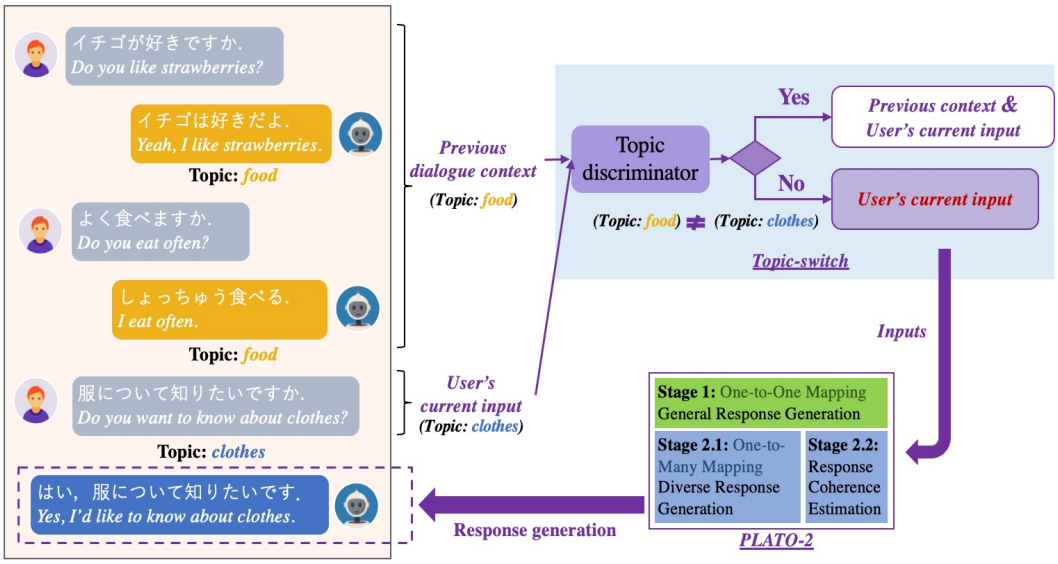


图1：基于话题切换的 PLATO-JDS 概览。给定包含先前对话的用户输入，话题切换模块由一个主题判别器组成，用于预测用户输入的主题是否与先前对话相同；若不同，则切换到新主题。PLATO-JDS 是一个预训练的日语对话生成模型，可根据话题切换模块提供的特定输入生成回复。

2021a，Bao et al.，2021b，Roller et al.，2021，Lewis et al.， 2020]。DialoGPT [Zhang 等人， 2020] 是一种基于 Transformer 的对话生成模型，使用来自 Reddit 评论的 1.47 亿条类对话交互进行训练；只需提供一条用户输入，它就能很好地生成基于上下文的回复。PLATO[Bao 等人， 2020] 专注于一对多对话生成 [Kim 等人， 2020]（例如，给定一条用户输入，模型可以生成多个回复），从而改进 DialoGPT。此外，PLATO-2 [Bao 等人，2021a]通过将 PLATO 模型扩展至数十亿参数，并引入课程学习 [Bengio 等人， 2009]，（包括在一对一生成子任务上进行预训练，以及在一对多生成子任务上进行微调），实现了最先进的开放域对话系统。然而，PLATO-2 的评估基于英语和中文数据，目前尚无研究探索其在日语数据上的有效性，以推动日语对话系统的发展。为了构建高质量的对话系统，控制生成方法 [Dathathri et al.，2019，Keskar et al.，2019，Madotto et al.，2020，Smith 等人，2020，Du 和 Ji， 2021]已广泛应用于对话系统。本研究通过课程学习训练 PLATO-JDS，分两个步骤控制生成连贯流畅的回复，如图 5 所示。

为了提升对话系统的用户体验，有必要让先进模型具备出色的话题适应性。 [Xu 等人， 2021]以无监督的方式从对话历史中提取话题片段，从而根据需要为话题切换选择最合适的回复。 [Xia 等人， 2022]也采用了话题片段方法，同时学习片段间关系，以改进话题切换。 [Sugiyama 等人， 2021] 收集了一个具有三种特征的混合数据集：个性、共情和知识，并用它来微调 BlenderBot，以便.....

构建一个可以频繁切换主题的闲聊系统。尽管这些模型会在多轮对话中追踪全局主题流，但它们难以消除历史对话的干扰，尤其是在用户突然切换到与先前对话无关的主题时。不同于以往方法将所有历史对话和用户输入都作为输入，我们提出了一种带有主题判别器的主题切换算法，用于确定训练好的 PLATO-JDS 的输入内容。

3. 数据集

本节将在 §3.1 中介绍日语数据集*Dialogue-Graph*的详情，并在 §3.2 中介绍数据预处理。

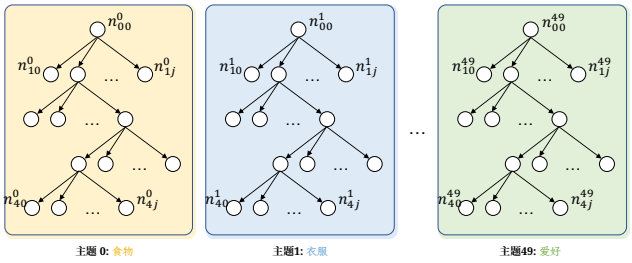


图 2：*Dialogue-Graph*：从**主题 0:食物** 开始，初始话语为$n^0_{0,0}$:“你最喜欢的食物是什么？”，*Dialogue-Graph* 最多可扩展为包含八个对话轮次的一对多对话数据。这里，一对多表示一个用户输入可以对应多个回复。

3.1 数据收集

在这项工作中，我们使用众包驱动系统 [Ikeda and Hoashi,2018] 来收集对话数据。该系统采用异步的对话创建方式，因此能显著提高对话数据收集的效率：工作者无需等待前一位工作者完成话语创建，随时都可以创建话语。具体而言，当工作者开始使用该系统时，系统会分配一段需要

工作者输入的对话，工作者根据给定的主题以及之前的话语创建一条话语。

我们招募了60名参与者来收集对话数据集。我们假设对话由角色A和角色B交替发言构成，因此将所有参与者分为两组，每组30人，分别扮演角色A和角色B。一个角色可以为多个对话创作话语，而一个对话则由多名工作者共同创作。最后，我们采用与 [Ikeda and Hoashi相同的评估方法 2018], 只选取高质量的对话，并剔除缺乏常识、涉及政治敏感内容或存在伦理问题的对话。为提高数据收集效率，让参与者能够更快地创作对话，我们将系统设置为每天凌晨12:00至下午3:00开放。对话数据收集历时三个月以上，每输入一条话语支付4日元，每段对话支付32日元（约合0.30美元）。

受 [Galitsky, 2018, Cheng et al.的研究 2020], 启发，我们也创建了树结构对话，从第一个头节点到叶节点共分为9层，即对话长度为8轮。我们使用了50个话题，每个话题都有一个初始话语，这与 [Ikeda and Hoashi的工作 2018], 相同。例如，话题是家人，初始话语是“你和别人住在一起吗？”。一名工作者根据起始话语创作了8条话语，由此构成一段对话。随后，我们采用与 [Ikeda and Hoashi相同的评估方法 2018] 对收集到的对话进行评估。该方法从三个方面评估对话质量：效率、质量和工作者的兴趣。数据收集完成后，我们聘请了同样的60名参与者进行评估。

在图 2 中，我们定义 *Dialogue-Graph* $\mathcal{G}^{Dia} = (\mathcal{N}, \mathcal{E})$、$\mathcal{N} = \{n^t_{i,j} | 0 \le t \le 49, 0 \le i \le 8, j \ge 0\}$，其中 $\mathcal{N}$是对话节点的集合 $n^t_{i,j}$, $\mathcal{E} = \{e^t_{i,j\to m,n} | 0 \le t \le 49, 0 \le i, m \le 8, j, n \ge 0\}$ 是对话轮次的集合，使得 $e^0_{0,0\to1,0}$ 表示从 $n^0_{0,0}$ 到 $n^0_{1,0}$的单个话语轮次。这里，t，i/m 和j/n 分别表示当前对话节点的主题类型数、话语轮次数和候选项数，t 和 i/m 的最大数量分别为 50 和 8，j/n 的最小数量为 3。

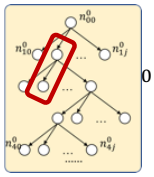
3.2 预处理

为了准备训练主题判别器和*PLATO-JDS* 所需的训练数据，我们遵循三项策略对*Dialogue-Graph*中的数据进行预处理。我们将每个预处理后的数据集划分为训练集、
训练集、验证集和测试集的常见比例为 90% : 5% : 5%。

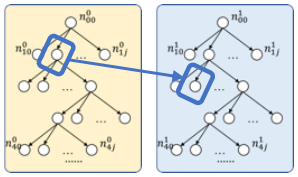
- **正样本与负样本**如图3 (a) 所示，我们从同一主题中提取任意对话轮次，并将其标注为正样本。另一方面，我们将对应项标注为负样本。

- **一对一映射**。我们为 *PLATO-JDS* 的阶段 1 准备训练数据。如图3 (b) 所示，我们采用广度优先搜索方法（BFS），总共提取了 1.68M 个数据样本。BFS 旨在遍历图中的对话节点，以获取话语对。从根节点开始，我们将所有话语轮次提取为训练数据（例如，$n^0_{0,0} \to n^0_{1,0}$，$n^0_{0,0} \to n^0_{1,3}$）。话语轮次的定义见 §3。。

- **一对多映射**。 我们为*PLATO-JDS*的阶段2准备训练数据。如图3 (c)所示，我们使用



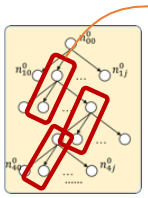
正样本：
$\{n^0_{1,1}$：你游泳游得好吗.,你游泳游得好吗？$n^0_{2,1}$:人少的话我就游泳.}人少的话，我可以游泳



负样本：
$\{n^0_{1,1}$: 你游泳游得好吗. ,你游泳游得好吗？$n^1_{2,1}$:如果是日常对话的话，我会说一点.}

主题 0　　主题 1　　我能进行一些日常会话。

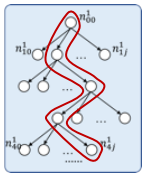
(a) 主题判别器，[正例 vs 负例]



配对样本

$n_{1,0}$ $^0{}_{1}$：你游泳游得好吗？你游泳游得好吗？$n_{2,0}$ $^0{}_{1}$：如果人少的话，我就游泳。如果人少的话，我就游泳。

(b) PLATO-JDS, [一对一映射]



$n_{0,0}$ $^1{}_{0}$：你会说英语吗？你会说英语吗？$n_{1,1}$ $^1{}_{1}$：日常会话程度的话，我会说一点。日常会话中我会说一点。……$n_{4,3}$ $^1{}_{3}$：没那么困难。没那么困难。

(c) PLATO-JDS，[一对多映射]

图3：用于训练主题判别器和 *PLATO-JDS* 的训练数据样本。

深度优先搜索（DFS）方法生成了126万条数据样本。DFS旨在遍历图中的对话节点，以获取所有八轮对话作为训练数据（例如，$n^0_{0,0} \to n^0_{1,0} \to n^0_{2,1} \to n^0_{3,1} \to n^0_{4,1} \to n^0_{5,0} \to n^0_{6,0} \to n^0_{7,0} \to n^0_{8,2}$）。

4. 方法

为实现图1所示的 PLATO-JDS，我们设计了一种主题切换算法，利用主题判别器根据用户输入和先前的对话判断模型的输入内容，然后将其输入经过训练的 PLATO-JDS，以生成回复。本节首先在 §4.1 中详细介绍主题判别器，并在 §4.2 中详细介绍 PLATO-JDS；随后在 §4.3 中说明主题切换算法的工作原理。

4.1 主题判别器

主题判别器是一个二分类器，旨在判断当前用户输入是否属于先前对话中的共同主题。如图 4 所示，我们构建了一个基于 BERT 的主题判别器，并使用 §3.2 中所述的预处理后 Dialogue-Graph 数据对其进行训练。我们将属于共同主题的数据标注为正样本，将其余数据标注为

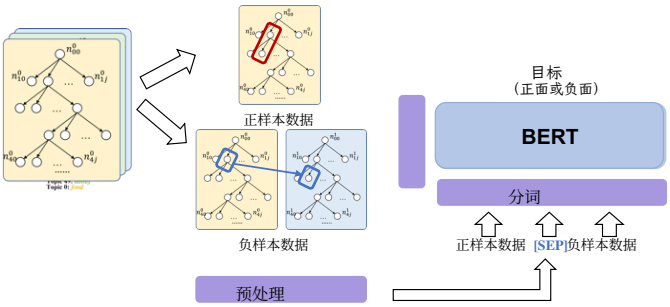


图 4：主题判别器概览。它由第 3.2 节所述的数据预处理、用于获取词序列的分词方法，以及用于训练二分类器的 BERT 骨干网络组成。

负样本。我们将包含一对话语的一轮对话拼接成一个句子，输入分词方法，再送入 BERT 的骨干网络。该训练任务的目标是通过优化二元交叉熵（BCE）损失，判别一对话语中的两个话语是否属于共同主题，具体如下。

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}(Y_i \cdot \log \hat{Y}_i + (1 - Y_i) \cdot \log(1 - \hat{Y}_i)) \quad (1)$$

其中， N 表示样本数量， Y_i 表示真实标签 ˆ（0 或 1），而 Y_i 表示主题判别器的预测概率。

在这里，我们采用 SentencePiece [Kudo 和 Richardson，2018]分词方法，将原始输入句子直接切分为词序列。SentencePiece 是一种基于子词思想的语言无关分词方法，旨在避免未知词元，同时减少词元数量。Byte Pair Encoding（BPE） [Gage，1994，Sennrich等，2016] 算法是 SentencePiece 所采用的子词划分算法，用于将高频词筛选为词元，并将低频词划分为两个或更多子词作为词元。我们使用 SentencePiece 将日语数据集 *Dialogue-Graph* 切分为 48k 个词元。最后，我们采用与原始 PLATO-2 模型相同的输入表示，包括词元嵌入、角色嵌入、轮次嵌入和位置嵌入。在这里，词元嵌入是针对不同子词的预训练嵌入。角色嵌入用于区分一次对话中说话者的角色。轮次嵌入根据转换中存在多轮对话时的相对顺序进行分配。位置嵌入则基于每次对话中的词元位置获得。

4.2 PLATO-JDS

我们沿用原始 PLATO-2 模型 [Bao 等人，2021a]，在日语数据上对其进行训练，以实现对话生成。PLATO-2 是一个基于 Transformer 的大规模对话生成框架，通过课程学习进行训练 [Bengio 等人， 2009]。课程学习是一种两阶段训练策略，如图3所示。第一阶段，使用图3 (b) 所示的预处理一对一映射数据，对粗粒度生成模型进行预训练，使其学会在不同对话上下文下生成回复。第二阶段，使用预处理的一对多映射数据训练细粒度生成模型和评估模型，相关数据如图3 (c) 所示。我们采用与主题判别模型相同的输入表示。

阶段 1，通用回复生成。一对一映射是一种传统方法，也是学习回复生成高层属性的有效方式。给定一个 i^{th} 回复 $R(i)$ 及其先前对话上下文 $C(i)$，通过最小化负对数似然（NLL）损失来训练模型，具体如下：

$$L_{NLL}^{General} = -\mathbf{E}\sum_{i=1}^{L}\log p(R(i)|C(i)) \quad (2)$$

其中， L 是生成回复的长度 $R(i)$。

阶段 2.1，多样化回复生成。为进一步训练模型，使其捕捉一对多映射的关系，我们引入了一个离散潜变量 z。模型估计训练样本的潜分布 $p(z|C(i)\ R(i))$，然后生成采样得到的潜变量 $p(R(i)|C(i)\ z)$。多样化回复生成的 NLL 损失和词袋（BOW）损失定义如下。

$$\begin{aligned} L_{NLL}^{Diverse} &= -\mathbf{E}_{z\sim p(z|C(i),R(i))}\sum_{i=1}^{L}\log p(R(i)|C(i),z) \\ L_{BOW}^{Diverse} &= -\mathbf{E}_{z\sim p(z|C(i),R(i))}\sum_{i=1}^{L}\log p(R(i)|C(i),z) \\ L^{Diverse} &= L_{NLL}^{Diverse} + L_{BOW}^{Diverse} \end{aligned} \quad (3)$$

其中， L 是生成回复的长度 $R(i)$。最终的生成模型通过上述整合损失 $L^{Diverse}$进行优化。

阶段 2.2，连贯响应选择。通过一次生成多条回复后，选择质量最高的回复作为最终输出。该选择模型利用掩码语言模型（MLM）的分布式表示能力，通过估计给定话语与其回复之间的连贯性进行训练。响应连贯性估计（RCE）和 MLM 的两个目标损失如下。

$$\begin{aligned} L_{RCE}^{Selection} &= -\log p(l_r = 1|C(i),R(i)) - \log p(l_r = 0|C(i),\hat{R(i)}) \\ L_{MLM}^{Selection} &= -\mathbf{E}\sum_{i\in L}\log p(x_t|{}^{x}/_{L}) \\ L^{Selection} &= L_{RCE}^{Selection} + L_{MLM}^{Selection} \end{aligned} \quad (4)$$

其中， l_r 表示生成的回复是否与对话文本一致（R（i）或不一致（R（i））； x 表示上下文和回复的输入词元。$\{x_i\}_{i\in L}$ 表示掩码词元， ${}^{x}/_{L}$ 表示其余未掩码词元。

4.3 主题切换算法

我们使用预训练的 PLATO-JDS，通过输入先前对话与当前用户输入的拼接来生成回复。我们设计了一种主题切换算法，用于决定是否将先前对话纳入回复生成。我们在算法 1 中定义了该主题切换算法。我们使用主题判别器，通过与实验阈值进行比较，预测当前用户输入的主题是否属于先前对话的主题。如果得到的分数低于

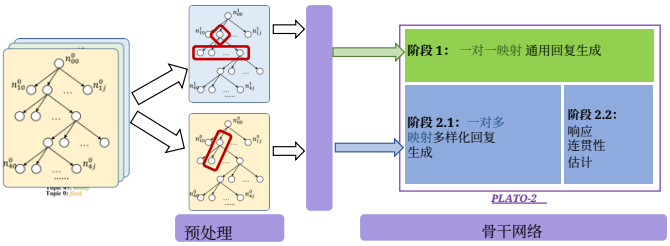


图 5：PLATO-JDS 模型及其数据处理概览。§ 3.2 中描述的数据预处理、与主题判别器相同的分词方法（§ 4.1），以及 PLATO-2 的骨干对话生成模型。

该阈值，则当前用户输入可能与先前对话无关，PLATO-JDS 模块的输入将自动切换为当前用户输入。反之，PLATO-JDS 模块的输入将包含先前对话。此处，我们根据精确率与召回率之间的权衡，将阈值设为 0.61，此时召回率为 99.75%，精确率为 95.72%。

算法 1主题切换算法

输入：ε：用于判断是否切换到新主题的阈值，i：轮次数，$C(i)$：第 i 轮的先前对话，$\Phi(i)$：第 i 轮的用户话语，$\Gamma(a,\ b)$：用于预测 a 和 b是否属于同一主题。

输出：下一个话语的上下文 $C(i)$

$\beta = \Gamma(C(i)\ \Phi(i))$
如果 $\beta \leq \varepsilon$ **那么**
    $C\ (i)\ \leftarrow \Phi\ (i)$ ▷切换到新主题
否则如果 $\beta > \varepsilon$ **那么**
    $C(i)\ \leftarrow C(i) + \Phi(i)$
结束 if
$i \leftarrow i + 1$

5. 实验

本节将在 §5.1 讨论评估指标。接着，我们将讨论 PLATO-2 的工作中报告的英文数据结果（见 §5.2），并解释为何选择 PLATO-2 作为对话生成模块的骨干网络。为了评估我们提出的基于主题切换的 *PLATO-JDS*，我们将该模型与 *PLATO-JDS* 进行比较，并在 §5.4 报告比较结果。此外，我们还将在 §5.5 讨论案例研究结果，并在 §5.3 介绍训练细节。

5.1 评估指标

我们从人工和自动两个角度评估模型。为确保公平比较，我们采用了 PLATO-2 学习中使用的相同评估指标 [Bao 等人，2021a]。在人工评估方面，我们邀请了 15 人（*11* 名男性和 *4* 名女性），针对话语级指标“连贯性”和“信息量”，以及对话级指标“吸引力”和“拟人性”进行三级评估（*0*：差，*1*：一般，*2*：好），以评估用户体验。这些指标的定义如下：

- 连贯性：根据回复判断其是否与当前话题相关，并与上下文保持一致。
- 信息量：衡量回复是否具有信息量且恰当。
- 吸引力：判断评估者是否愿意在长对话中与说话者交谈。
- 拟人性：评估对话中的说话者是否为真人，或回复是否自然。

在自动评估方面，为了评估模型的词汇多样性，我们采用语料库级指标 [Li et al., 2016a] distinct-1 和 distinct-2（表*1*、表*2*中的 *distinct-1/2*），它们的定义为回复生成中不同的 $unigrams$ 和 $bigrams$的数量与生成的词元总数之比。

5.2 基线

我们将 PLATO-2 设为基线，因为它是英语和中文领域中最先进的对话生成模型。与 Blender 模型 [Roller et al., 2021] 相比，PLATO-2 在人工评估和自动评估中均取得了最佳表现；Blender 模型通过引入融合技能，缓解了大型语料库中不良的有害或偏见特征，但需要大量人工标注。考虑到性价比，我们仅将 PLATO-2 实现为*PLATO-JDS*，并将其与我们提出的基于主题切换的*PLATO-JDS* 进行比较；两者均在本工作中的*Dialogue-Graph* 上训练。我们通过机器人之间的聊天策略和人机聊天策略进行两种评估，详情见 §5.4。*PLATO-JDS* 的描述见 §4.2。

5.3 训练细节

PLATO-JDS 的训练分为两个阶段。在阶段 1 中，我们采用负对数似然（NLL）损失，通过一对一映射学习捕捉回复生成的一般特征。在阶段 2.1 中，我们将 NLL 损失与词袋（BOW）损失相加，以学习细粒度生成。其中，BOW 损失用于反映离散潜变量的训练过程。在阶段 2.2 中，目标损失为回复连贯性估计损失与掩码语言模型（MLM）损失之和 [Bao 等人，2021a]。上述损失函数的更多细节见 § 4.2。此外，我们将上下文和回复的最大序列长度设为 512，并将位置嵌入的大小设为 256。我们使用 NVIDIA RTX A5000 GPU (24G x 2) 训练 *PLATO-JDS*，耗时 25 天。

此外，我们在附录的图 7 中展示了阶段 1、2.1 和 2.2 的模型训练与验证过程的损失曲线。结果表明，模型训练过程中未发生过拟合，证明了训练模型的可靠性。

5.4 结果

5.4.1 机器人对机器人聊天策略

机器人对机器人聊天策略旨在通过两个聊天机器人之间进行对话模拟来评估对话生成，从而减少全面人工评估所需的时间和费用；全面人工评估需要人们花时间与聊天机器人交谈，并为生成的回复打分。这种策略通常用于实现对话生成的公平评估 [Li et al., 2016b, Bao 等人，2021a]。我们随机选取了200个问题

作为起始输入，为模型针对每个问题生成十轮对话。如表1所示，基于主题切换的*PLATO-JDS* 在所有人工评估指标上均优于*PLATO-JDS*，这表明我们基于主题切换的*PLATO-JDS* 生成的回复在当前对话上下文中连贯，且包含的信息远多于*PLATO-JDS*。此外，根据“吸引力”和“拟人性”指标的得分，我们认为基于主题切换的*PLATO-JDS* 能够改善*PLATO-JDS*上的用户体验，这也证明了生成类人回复的有效性。此外，distinct-1/2的结果表明，我们基于主题切换的*PLATO-JDS* 包含更多不同的 $unigrams$ 和$bigrams$，说明其具备更强的词汇多样性。此外，从话语平均长度和平均话题数来看，我们基于主题切换的*PLATO-JDS* 得分高于*PLATO-JDS*，进一步验证了主题切换根据用户输入切换到新主题的有效性。

指标 \ 模型	无话题切换	有话题切换
人工评估		
连贯性	1.667	**1.800**
信息量	1.533	**1.867**
吸引力	1.533	**1.733**
拟人性	1.533	**1.800**
平均分	1.567	**1.800**
自动评估		
Distinct-1/2	0.343/0.712	**0.388/0.730**
长度（平均值）	14.016	**14.667**
话题数（平均值）	4.6	**6.8**

表1：机器人对机器人聊天策略：*PLATO-JDS* （无主题切换）与基于主题切换的*PLATO-JDS* （有话题切换）的比较结果。

5.4.2 人机对话策略

除了机器人对机器人聊天策略外，我们还收集了人机对话记录，以在真实场景中评估我们的模型。我们邀请了15名参与者与机器人聊天，并要求每位参与者与机器人进行50轮对话。我们共生成了750条对话数据。结果汇总于表2。我们得到的对比结果与§5.4.1中所述的结果非常接近。尽管基于主题切换的*PLATO-JDS* 在人机聊天策略中的得分低于*PLATO-JDS*在机器人对机器人聊天策略中的得分，但与*PLATO-JDS*相比，我们在所有指标上都取得了更高的分数。

为进一步确认主题切换的有效性，我们还比较了生成的对话中平均包含的话题数量。如表1和表2所示，基于主题切换的*PLATO-JDS* 管理的话题更多，对话也比*PLATO-JDS*更长，这证明了主题切换能够让*PLATO-JDS* 基于新话题生成回复。由此可见，采用主题切换算法能够提升*PLATO-JDS* 在人工评估和自动评估中的表现，并表明其具备改善用户体验的能力。

指标 \ 模型	无话题切换	有话题切换
人工评估		
连贯性	1.600	**1.733**
信息量	1.467	**1.800**
吸引力	1.467	**1.800**
拟人性	1.467	**1.733**
平均分	1.500	**1.767**
自动评估		
Distinct-1/2	0.339/0.709	**0.376/0.715**
长度	13.970	**14.333**
话题数（平均）	4.2	**6.5**

表2：人机对话策略：*PLATO-JDS* （无话题切换）与基于主题切换的*PLATO-JDS* （有话题切换）的比较结果。

5.5 案例研究

我们选取了*PLATO-JDS* 在同一主题“美容与健康”中使用相同起始话语、开启或不开启主题切换的两段人机对话样本，以研究主题切换如何影响回复生成。如图6所示，当用户输入转向一个与先前对话上下文不同的新主题时，我们注意到，启用了主题切换的模型能够生成恰当的回复，而*PLATO-JDS* 则不能。我们认为，所提出的主题切换模块会促使*PLATO-JDS* 切换到新主题并生成相关回复，这证明了主题切换的有效性。此外，案例研究也表明，第5.4节所述的各项评估结果具有连贯性。



图6：使用或不使用主题切换模型生成的两个对话示例。对话从同一句属于“健康”主题的话语“I have a pain in my stomach”开始；当红色虚线框中的用户输入变为“酒精”这一新主题时，我们基于主题切换的*PLATO-JDS* 可以切换到“酒精”主题并生成恰当的回复，而*PLATO-JDS* 则不能。

6. 结论

我们构建了一个基于 PLATO-2 的日语对话系统 PLATO-JDS，并提出了一种主题切换算法：当用户输入与先前主题不同时，该算法会使 PLATO-JDS 切换到新主题。我们通过从新闻、电视字幕和维基百科语料库中收集 1.656 M 条对话数据，构建了最大的日语对话生成数据集 *Dialogue-Graph*。在采用人机对话策略进行的人工评估中，我们提出的 PLATO-JDS 平均得分为**1.500**；而所提出的主题切换算法进一步使 PLATO-JDS 的得分提高**0.267**，达到 **1.767**。结果表明，PLATO-2 可用于实现高质量的日语对话生成，而主题切换有助于提升系统的用户体验。

参考文献

[Adiwardana 等， 2020] Adiwardana, D., Luong, M.-T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G., Lu, Y. 等。（2020年）。迈向类人开放域聊天机器人。*arXiv* 预印本 *arXiv:2001.09977*。

[Bae 等， 2022] Bae, S., Kwak, D., Kim, S., Ham, D., Kang, S., Lee, S.-W. 和 Park, W.（2022年）。利用大规模语言模型构建角色指定的开放域对话系统。载于*2022*年计算语言学协会北美分会年会：人类语言技术，第2128–2150页，西雅图，美国。计算语言学协会。

[Bao et al.，2020] Bao, S., He, H., Wang, F., Wu, H., and Wang, H. (2020年). PLATO：具有离散潜变量的预训练对话生成模型。载于计算语言学协会第*58*届年会论文集，第 85–96 页，线上。计算语言学协会。

[Bao et al., 2021a] Bao, S., He, H., Wang, F., Wu, H., Wang, H., Wu, W., Guo, Z., Liu, Z., and Xu, X. (2021a). PLATO-2：通过课程学习构建开放域聊天机器人。载于计算语言学协会论文集：*ACL-IJCNLP 2021*，第 2513–2525 页，线上。计算语言学协会。

[Bao et al., 2021b] Bao, S., He, H., Wang, F., Wu, H., Wang, H., Wu, W., Wu, Z., Guo, Z., Lu, H., Huang, X., etal. (2021b). Plato-xl：探索对话生成的大规模预训练。*arXiv*预印本 *arXiv:2109.09519*。

[Bengio等人， 2009] Bengio, Y., Louradour, J., Collobert, R., and Weston, J.（2009年）。课程学习。载于第*26*届国际机器学习年会（*ICML*）（第*382*卷：长论文），第41–48页，美国纽约州纽约。美国计算机协会。

[Brown 等人， 2020] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. 和 Amodei, D.（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（第*33*卷：长论文），第1877–1901页，线上。Curran Associates, Inc.

[Cheng等人， 2020] Cheng, J., Agrawal, D., Alonso, H. M., Bhar- gava, S.,Driesen, J., Flego, F., Ghosh, S., Kaplan, D., Kartsaklis, D., Li, L., 等人（2020年）。用于对话状态跟踪的对话语义解析。*arXiv*预印本 *arXiv:2010.12770*。

[Dathathri等人， 2019] Dathathri, S., Madotto, A., Lan, J., Hung, J., Frank, E., Molino,P., Yosinski, J., and Liu, R.（2019）。即插即用语言模型：一种可控文本生成的简单方法。*arXiv*预印本 *arXiv:1912.02164*。

[Devlin 等人， 2019] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K.（2019）。BERT：用于语言理解的深度双向 Transformer 预训练。载于计算语言学协会北美分会 *2019* 年年会：人类语言技术，第 *1* 卷（长短篇论文），第 4171–4186 页，明尼阿波利斯，明尼苏达州。计算语言学协会。

[Du 和 Ji， 2021] Du, W. and Ji, Y.（2021年）。Sidecontrol：通过附加侧网络进行可控开放领域对话生成。*arXiv* 预印本 *arXiv:2109.01958*。

[Gage， 1994] Gage, P.（1994）。一种新的数据压缩算法。*CUsersJournal*。

[Galitsky， 2018] Galitsky, B.（2018）。从问题的话语树构建对话结构。载于第三十二届 *AAAI* 人工智能会议研讨会。

[Golovanov 等人， 2019] Golovanov, S., Kurbanov, R., Nikolenko, S. I., Truskovskyi, K., Tselousov, A., 和 Wolf, T.（2019）。用于自然语言生成的大规模迁移学习。载于 Korhonen, A., Traum, D. R., 和 Màrquez, L.（编），第 *57* 届计算语言学协会年会论文集，*ACL*（第 *1* 卷：长论文），第 6053–6058 页，意大利佛罗伦萨。计算语言学协会。

[Ikeda and Hoashi， 2018] Ikeda, K. 和 Hoashi, K.（2018）。利用众包异步聊天，高效收集对话数据集。载于第六届 *AAAI* 人类计算与众包会议论文集，*HCOMP*（第 *6* 卷：长论文），第 60–69 页，瑞士苏黎世。AAAI 出版社。

[Keskar 等人， 2019] Keskar, N. S., McCann, B., Varshney, L. R., Xiong, C. 和 Socher, R.（2019）。Ctrl：一种用于可控生成的条件 Transformer 语言模型。*arXiv* 预印本 *arXiv:1909.05858*。

[Kim 等人， 2020] Kim, B., Ahn, J. 和 Kim, G.（2020）。用于知识支撑的对话的序列潜在知识选择。*arXiv* 预印本 *arXiv:2002.07510*。

[Kudo和Richardson， 2018] Kudo, T. 和 Richardson, J.（2018）。SentencePiece：一种简单且不依赖语言的子词

用于神经文本处理的分词器和去分词器。载于*2018*年自然语言处理实证方法会议：系统演示论文集，第66–71页，比利时布鲁塞尔。计算语言学协会。

[Lewis 等， 2020] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V. 和 Zettlemoyer, L.（2020年）。BART：用于自然语言生成、翻译和理解的去噪序列到序列预训练。载于计算语言学协会第*58*届年会论文集，第 7871–7880 页，线上。计算语言学协会。

[Li et al., 2016a] Li, J., Galley, M., Brockett, C., Gao, J. 和 Dolan, B.（2016a）。神经对话模型的多样性促进目标函数。载于计算语言学协会北美分会*2016*年会议论文集：人类语言技术，第 110–119 页，美国加利福尼亚州圣迭戈。计算语言学协会。

[Li et al., 2016b] Li, J., Monroe, W., Ritter, A., Jurafsky, D., Gal- ley, M. 和 Gao, J.（2016b）。用于对话生成的深度强化学习。载于*2016*年自然语言处理实证方法会议论文集，第 1192–1202 页，得克萨斯州奥斯汀。计算语言学协会。

[Liu et al., 2019] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. 和 Stoyanov, V.（2019）。RoBERTa：一种稳健优化的 BERT 预训练方法。*arXiv* 预印本 *arXiv:1907.11692*。

[Madotto et al., 2020] Madotto, A., Ishii, E., Lin, Z., Dathathri, S. 和 Fung, P.（2020）。即插即用对话模型。*arXiv* 预印本 *arXiv:2010.04344*。

[Nakatsuji， 2019] Nakatsuji，M.（2019）。人工智能可以生成爱之建议吗？：面向非事实型问题的神经答案生成。*arXiv*预印本 *arXiv:1912.10163*。

[Radford等， 2019] Radford，A.、Wu，J.、Child，R.、Luan，D.、Amodei，D.、Sutskever，I.等（2019）。语言模型是无监督多任务学习器。*OpenAI*博客（第*1*卷（*8*）：长论文），第9页。

[Rikters等， 2019] Rikters，M.、Ri，R.、Li，T.和Nakazawa，T.（2019）。商务对话语料库的设计。载于第六届亚洲翻译研讨会论文集，第54–61页，中国香港。计算语言学协会。

[Roller 等人， 2021] Roller, S., Dinan, E., Goyal, N., Ju, D., Williamson, M., Liu, Y., Xu, J., Ott, M., Smith, E. M., Boureau, Y.-L., and Weston, J.（2021年）。构建开放领域聊天机器人的方法。在第*16*届欧洲计算语言学协会分会会议论文集：主卷，第300–325页，线上。计算语言学协会。

[Sennrich等， 2016] Sennrich, R., Haddow, B. 和 Birch, A.（2016）。使用子词处理稀有词的神经机器翻译单元。在第*54*届计算语言学协会年会论文集（第*1*卷：长篇论文），第1715–1725页，德国柏林。计算语言学协会。

[Shalaby 等人， 2020] Shalaby, W., Arantes, A., Gonzalez-Diaz, T., and Gupta, C.（2020年）。基于大规模领域专属知识库构建聊天机器人：挑战与机遇。在*IEEE*国际预测与健康管理会议，*ICPHM2020*，第1–8页，美国密歇根州底特律。IEEE。

[Smith 等人， 2020] Smith, E. M., Gonzalez-Rico, D., Dinan, E., and Boureau, Y.-L.（2020年）。控制生成对话的风格。*arXiv*预印本 *arXiv:2009.10855*。

[Sugiyama 等人， 2021] Sugiyama,H.,Mizukami,M.,Arimoto, T., Narimatsu, H., Chiba, Y., Nakajima, H., and Meguro, T. (2021年). 基于 Transformer 的日语闲聊系统的训练策略实证分析。*arXiv* 预印本 *arXiv:2109.05217*。

[Tianyu and Sawada，2021] Tianyu, Z. and Sawada, K.（2021）。《日语自然语言处理中的预训练模型发布》。人工智能学会研究会资料：语言、语音理解与对话处理研究会，93：169–170。[Wolf 等人， 2019] Wolf, T., Sanh, V., Chaumond, J., and Delangue, C.（2019）。Transfertransfo：一种面向基于神经网络的对话代理的迁移学习方法。*arXiv* 预印本 *arXiv:1901.08149*。[Wu 等人， 2016] Wu, X., Ito, K., Iida, K., Tsuboi, K., and Klyen, M.（2016）。Rinna：女高中生人工智能。语言处理学会第*22*届年会论文集，第306–309页。[Xia 等人， 2022] Xia, J., Liu, C., Chen, J., Li, Y., Yang, F., Cai, X., Wan, G., and Wang, H.（2022）。通过带邻居平滑的并行提取网络进行对话主题分割。载于*SIGIR '22*：第*45*届 *ACM* 信息检索研究与发展国际会议，第2126–2131页，西班牙马德里。ACM。[Xu 等人， 2021] Xu, Y., Zhao, H., and Zhang, Z.（2021）。面向主题感知的多轮对话建模。载于第三十五届人工智能促进会人工智能会议（*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*，第14176–14184页，线上。AAAI 出版社。[Zhang 等人， 2020] Zhang, Y., Sun, S., Galley, M., Chen, Y.-C., Brockett, C., Gao, X., Gao, J., Liu, J., and Dolan, B.（2020）。DialoGPT：用于对话回复生成的大规模生成式预训练。载于计算语言学协会第*58*届年会论文集：系统演示，第270–278页，线上。计算语言学协会。[Zhou 等人，2021] Zhou, H., Ke, P., Zhang, Z., Gu, Y., Zheng, Y., Zheng, C., Wang, Y., Wu, C. H., Sun, H., Yang, X., 等人（2021）。EVA：一种采用大规模生成式预训练的开放域中文对话系统。*arXiv* 预印本 *arXiv:2108.01547*。

話題	主題	話題	主題
美容与健康（饮食习惯）	美容与健康（饮食习惯）	映画	电影
自己紹介	自我介绍	家族	家人
户外活动	户外活动	恋愛	love
度过假期的方式	度过假期的方式	趣味	爱好
友人関係	友谊	スポーツ	做运动
スポーツ	观看体育比赛	勉強	学习
習い事	课程	学生生活	学生生活
兼职	兼职工作	料理	烹饪
スイーツ	甜点	お酒	酒精
电视节目	电视节目	通勤	通勤
室内装饰	室内环境	家事	家务
汽车、驾车出行	car, 驾车	仕事	work
喜欢的动物	喜欢的动物	季節	季节
家電	家用电器	携帯電話	手机
国内旅行	国内旅行	出身地	出生地
美容与健康（运动）	美容与健康（运动）	語学	语言
按摩	消息	カフェ	Cafe
阅读・漫画	阅读 / 漫画	音楽	音乐
想要的东西	我想要的东西	結婚	婚姻
海外旅行	海外旅行	食べ物	food
任意门	任意门	ランチ	午餐
喜欢的艺人	最喜欢的艺人	好きな街	最喜欢的城镇
沉迷的事物	沉迷的事物	将来の夢	未来的梦想
时尚	时尚	宝くじ	彩票
睡眠（生活习惯？）	睡眠（生活习惯？）	SNS	SNS

表3：英语和日语中的50个话题列表

附录

表3列出了我们数据收集中全部50个话题。表4列出了每个话题的起始话语。图8展示了对话数量随话题变化的情况。表5展示了我们的 Dialogue-Graph 数据集的统计信息。表7展示了 PLATO-2 在英文数据上的机器人间聊天对比结果。表6展示了 PLATO-JDS 的训练设置详情。图7展示了 PLATO-JDS 在训练过程中的损失。

主题	起始话语
家人	你独自生活吗?
婚姻	你结婚了吗？？
友谊	你在学校和朋友见面吗?
爱好	你平时喜欢做什么消遣？？
做运动	你会做一些运动吗?
学习	你现在有学习任务吗?
烹饪	你喜欢烹饪吗?
甜点	你喜欢甜点吗?
.....	

表4：话题及其对应开场话语示例。

#	数量 #
词元数量最多的话语	95
词元数最少的话语	1
词元数平均的话语	12.28
词元数最多的对话	258
词元数最少的对话	23
词元数平均的对话	109.3
轮次数量	8
对话总数	1.656M
话语总数	9,837,742
单个主题的对话数量上限	22,397
某一主题的所有对话数量下限	21,903
跨主题的所有对话平均数量	22,092.04
跨主题的所有对话数量方差	14106.32

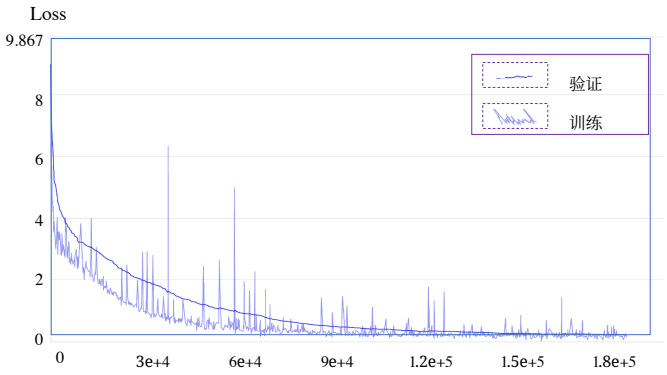
表5：我们的 Dialogue-Graph 数据集统计信息。

模型（阶段）	超参数			困惑度损失	负对数似然/重构交叉熵损失	词袋损失	掩码语言模型损失	
	批次大小	学习率	步数					
阶段 1	2,700	4e-4	16M	1.92	0.52	0.52	-	-
阶段 2.1	5,400	4e-5	1.10M	1.20	4.40	0.23	4.17	-
阶段 2.2	2,900	4e-5	2.54M	-	0.58	0.39	-	0.19

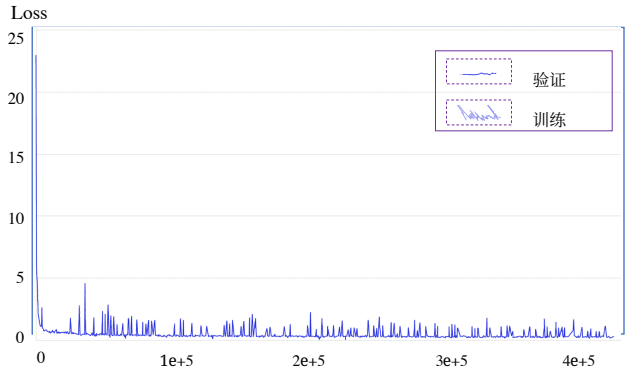
表6：*PLATO-JDS* 训练设置的详细信息。阶段 1 中计算 NLL 损失，阶段 2.1 中计算 NLL 和 Bow 损失，阶段 2.2 中计算 RCE 和 MLM 损失

指标 \ 模型	PLATO	DialoGPT	Blender	PLATO-2
人工评估				
连贯性	0.568	0.720	1.856	**1.920**
信息量	0.564	0.712	1.816	**1.892**
吸引力	0.340	0.340	1.820	**1.840**
拟人性	0.280	0.100	1.540	**1.740**
自动评估				
Distinct-1/2	0.042/0.255	0.150/0.508	0.117/0.385	**0.169/0.613**

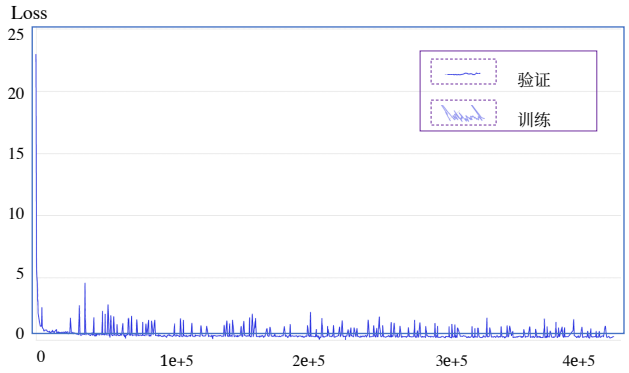
表7：PLATO-2 在英文数据上的机器人对机器人聊天策略比较结果。数值越高表示质量越高。所有结果均来自 PLATO-2 原论文（Bao 等人，2021a）



（a）*PLATO-JDS* 阶段 1 中一对一映射模型的训练。



（b）*PLATO-JDS*阶段 2.1 中一对多映射模型的训练



（c）*PLATO-JDS*阶段 2.2 中响应评估模型的训练

图 7：*PLATO-JDS* 模型在阶段 1、阶段 2.1 和阶段 2.2 的训练损失图。

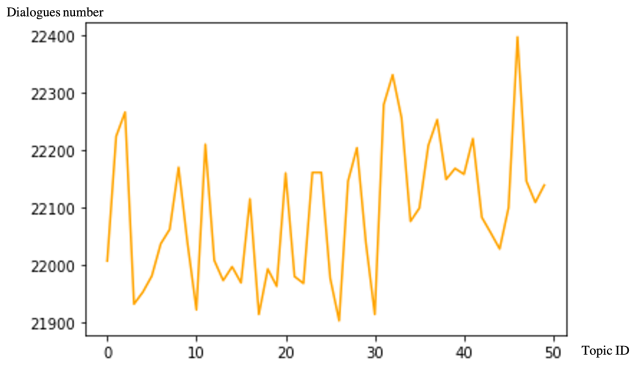


图 8：随着话题变化的对话数量。