arXiv:2303.04562v3 [cs.LG] 7 Jun 2023

外推式 通过迭代优化进行受控序列生成 t

Vishakh Padmakumar [1] **Richard Yuanzhe Pang** [1] **He He** [1] **Ankur P. Parikh** [2]

摘要

我们研究外推式受控生成问题，即生成属性值超出训练中所见范围的序列。这项任务在自动化设计中至关重要，尤其是在药物发现领域，其目标是设计出比现有序列更优（例如，稳定性更高）的新型蛋白质。因此，从定义上说，目标序列及其属性值都处于训练分布之外，这给旨在直接生成目标序列的现有方法带来了挑战。为此，我们提出迭代式受控外推 （*ICE*），通过迭代地对序列进行局部编辑来实现外推。具体而言，我们使用合成生成的序列对训练模型，这些序列对展现了属性值的小幅提升。在一项自然语言任务（情感分析）和两项蛋白质工程任务（ACE2 稳定性和 AAV 适应度）上的结果表明，尽管*ICE* 十分简单，但其表现显著优于最先进的方法。[1]

1. 引言

受控生成，即生成具有所需属性 z 的序列 $\boldsymbol{x}$，是多个领域中普遍存在的问题。在自然语言处理（自然语言处理）中， z 可以表示句子的情感或风格（例如正式程度）。在计算生物学中， z 可以表示蛋白质序列的稳定性、荧光、结合亲和力或其他属性。

有时，存在大量形式为 $(\boldsymbol{x},\ z)$ 的监督数据，例如维基百科领域或基因本体类别（Keskar 等人，2019年；Madani 等人，2020），从而能够直接训练条件生成模型 $p(\boldsymbol{x}|z)$。

当可用的监督数据对数量较少时，通常会使用这些数据训练评分器 $f(\boldsymbol{x})$，它将输入序列映射为输出属性值。随后，可以使用 $f(\boldsymbol{x})$ 为大型语料库添加标注以供训练（Gehman 等人，2020），也可以在推理期间直接使用 $f(\boldsymbol{x})$，以引导无条件模型 $p(\boldsymbol{x})$ (Dathathri 等人，2020；Yang 和 Klein，2021) 的生成过程。

在这项工作中，我们关注这样一类应用：需要生成其属性值外推到训练分布之外的序列。例如，在生物序列设计中，生成从头（新颖）序列，使其在某项属性（例如，与特定靶标的结合亲和力）方面优于已有的自然序列，对药物发现至关重要（Arnold，1998；Romero 与 Arnold，2009；Freschlin 等人，2022）。在创意文本生成中，我们希望生成能够强化某种风格属性（例如幽默）的文本，而不仅仅是模仿现有文学作品（He 等人，2019年；Lyu 等人，2021）。

当训练数据中属性值的范围 z 覆盖有限时，现有的受控生成范式往往难以进行外推，因为 $p(\boldsymbol{x}|z)$ 和属性评分器 $f(\boldsymbol{x})$ 都可能无法泛化到属性的训练范围之外。例如，考虑图Chan 等人，2021c所示的 ACE2 稳定性任务（图1），目标是生成稳定性更高（ddG 值更低）的 ACE2 蛋白突变体。训练数据包含 ddG值介于和 -4 10之间的序列，但在推理时，我们希望生成比现有蛋白质更稳定的蛋白质，例如将 ddG外推至小于 的范围-5。由于这一范围的 z 在训练数据中没有覆盖，直接将 $p(\boldsymbol{x}|z)$拟合到训练数据上，会导致其在 $z < -4$ 上的表现不可预测。

我们的主要假设是，尽管具有不同目标值的序列（例如稳定和不稳定的蛋白质）具有不同的分布，但将一个序列转化为略有改进的版本这一过程适用于不同范围的属性值。例如，在药物设计中，更优质的蛋白质通常通过连续突变体的演化获得；在文本生成中，则可以通过添加程度副词来增强情感。因此，我们提出将这一问题转化为一系列针对基础序列的局部改进——

[1]纽约大学 [2]Google DeepMind。通讯作者：Vishakh Padmakumar <vishakh@nyu.edu>。

论文集： 40th 国际机器学习会议，美国夏威夷州檀香山。PMLR 202，2023年。版权所有 © 2023年，归作者所有。[1] 我们的代码和模型可在以下地址获取：https://github.com/vishakhpk/iter-extrapolation。

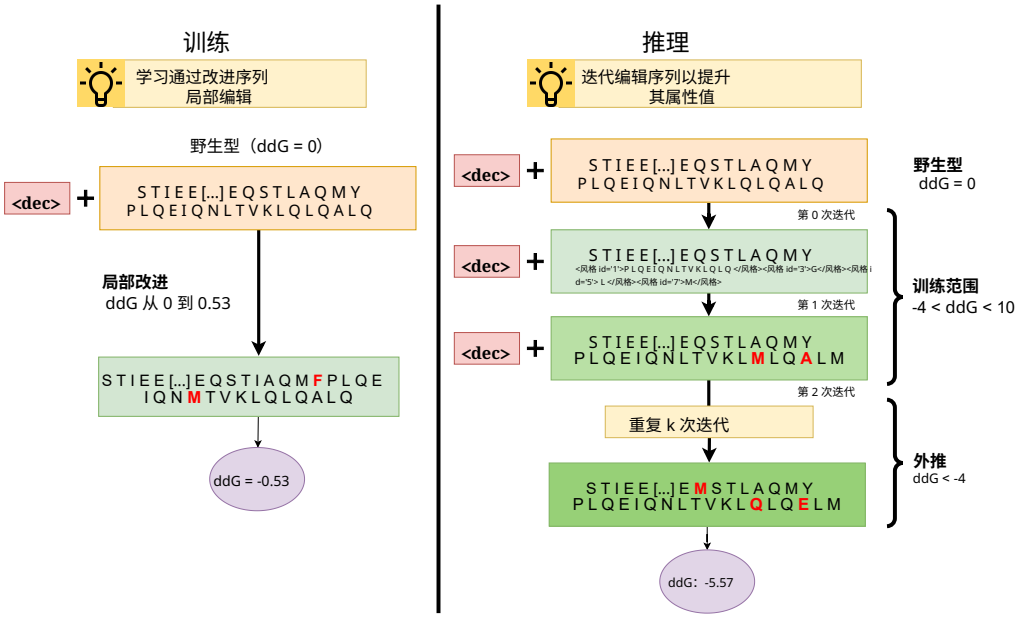


图1：在 ACE2 稳定性任务中，该方法——迭代受控外推 （*ICE*）——的概览。我们的初始数据集仅包含 ddG 值（数值越低表示越稳定）介于 -4 和 10 之间的蛋白质。在训练过程中，我们生成蛋白质序列的扰动，并学习一个生成器，对基础序列进行局部编辑，以降低其 ddG 值。在推理时，我们迭代应用训练好的生成器；经过 10 次迭代后，ddG 值达到 -5.57，比训练期间观察到的突变更稳定。

—— $\boldsymbol{x}_0$所做的改进。我们的直觉是，这种局部改进在不同属性值范围内都很稳定。因此，我们可以在训练分布上学习这些局部编辑（或突变体），并在推理时依次应用它们，从而外推到新的属性值范围。[2]

如图 1所示，为了训练局部编辑器，我们使用掩码语言模型（Devlin 等人，2019年）合成生成序列的近邻对，使它们在属性值上的差异很小。在推理过程中，我们的模型使用两个控制标签 <inc> 表示递增，以及 <dec> 表示递减，从而沿期望方向对序列进行局部改进。增加序列上的编辑次数即可实现外推。我们将该方法称为 迭代受控外推（*ICE*）。

我们在自然语言和蛋白质领域评估了我们的方法。对于文本生成，我们生成情感比训练数据中所见更积极或更消极的评论。对于蛋白质工程，我们展示了两项任务的结果——生成稳定性更高（由 FoldX 测量）的 ACE2 蛋白突变（Schymkowitz 等人，2005）以及生成腺相关病毒（AAV）衣壳蛋白的突变（Bryant 等人，2021），使其适应度值更高。*ICE* 在这三项任务上都实现了稳定的外推，表现优于标准受控生成方法（如 PPLM）（Dathathri 等人，2020）以及最先进的外推式受控生成方法 Genhance（Chan 等人，2021a）。尤其是在 AAV 任务中，尽管训练期间没有见过任何优于野生型 AAV 序列的序列，我们的模型仍能生成多种更优的候选序列，预言机模型的评估结果也证实了这一点。

[2]这些迭代改进发生在模型内部，因此与定向进化中的轮次并不类似（Arnold，1998），因为定向进化通常需要在每轮之后进行湿实验（或使用预言机）。

2. 相关工作

2.1. 受控生成

尽管文献中已对受控生成进行了广泛研究，但这些方法大多未关注外推场景。下面我们概述相关工作，介绍本方法和实验设置在既有研究中的定位。

采用控制代码的方法 Keskar 等人 (2019年) 和 Madani 等人(2020；2023) 学习条件序列模型 $p(\boldsymbol{x}|c)$ 其中 c 是控制代码，用于编码指定目标属性的离散值或标量值。然而，当条件值是超出训练数据范围的未见属性值时，这些模型可能难以应对。Lu 等人不再以绝对目标值作为条件，而是尝试通过从模型中采样生成序列，并将这些序列迭代量化为Lu 等人(2022) 来克服这一局限

更细粒度的控制代码，然后使用最高的区间进行受控生成。

迭代编辑方法 我们的方法还与基于编辑的方法相关（Guu 等人，2018年；Mallinson 等人，2022；Novak 等人，2016年），也与同期的工作密切相关，Welleck 等人（2023），该工作从模型中采样并对生成序列进行评分，以学习适用于各种自然语言处理任务的编辑。我们工作的关键区别在于，我们关注外推。在 Welleck 等人（2023）的设定中，模型通过寻求所有生成序列对的反馈进行学习。然而，我们明确关注这样一种情况：模型需要生成超出其能够获得反馈范围的序列。

潜变量模型 实现控制的另一种方法是将属性建模为潜变量（Mueller 等人，2017年；Gligorijević 等人，2021；Chan 等人，2021a；b）。例如，Genhance（Chan 等人，2021a）提出将潜向量表示为属性相关成分与属性无关成分之和。随后，他们对前者进行扰动，以实现外推，并将其应用于自然语言处理和生物学。然而，众所周知，针对离散序列数据的潜变量模型存在稳定性问题。相比之下，我们的方法直接在文本空间中进行编辑，避开了将连续潜在空间映射到离散序列这一问题。

通过评分器模型进行属性控制 另一类研究（Dathathri 等人，2020；Yang 和 Klein，2021；Li 等人，2022）通过评分器模型 $p(z|\boldsymbol{x})$添加属性信息，以在推理时引导无条件语言模型 $p(\boldsymbol{x})$。由于这种方法严重依赖作为训练好的分类器的评分器模型，因此通常不利于超出训练数据分布进行外推，正如我们将在实验中展示的那样。另一种做法是将分类器用作强化学习的奖励模型（Gong 等人，2019；Angermueller 等人，2020b），但这也有类似的缺点，因为生成器可能会利用并放大奖励中的缺陷（Amodei 等人，2016；Ibarz 等人，2018；Pang 等人，2022）。

2.2. 生物序列设计

生成从头优于天然序列的序列，对药物发现、医疗保健和农业具有巨大价值，这一点从因定向进化而获得的Arnold在1998年获得的2018年诺贝尔化学奖中可见一斑。由此，利用机器学习解决这一问题日益受到关注（Yang 等人，2019年；Angermueller 等人，2020a；Freschlin 等人，2022；Ren 等人，2022）。Brookes 等人（2019）通过一系列重要性采样分布来处理外推问题，这与我们的受控生成方法不同。

ICE的迭代特性是我们建模方法的内在组成，因此不同于定向进化中的各轮迭代；后者通常要求在每一轮之后访问预言机（或进行湿实验）。相反，在定向进化的每一轮中，都有可能（迭代地）运行*ICE*，并将其最终输出视为待验证的候选序列。

利用大型预训练蛋白质语言模型生成并通过实验验证新序列，也是一个令人振奋但仍处于起步阶段的研究方向。这些方法（Madani 等人，2021；Verkuil 等人，2022）通常通过以宽泛类别或骨架结构为条件来生成序列，而不是像我们所追求的那样，针对特定目标属性（例如稳定性或荧光）进行优化。

3. 我们的方法

问题设定 我们将包含 ℓ 个词元的输入序列表示为 $\boldsymbol{x} = (x_1, \ldots\ldots, x_\ell)$，将属性值表示为 $z \in \mathbb{R}$。这里，$\boldsymbol{x}$ 可以表示由 ℓ 个氨基酸组成的蛋白质序列，其中 z 表示其稳定性；也可以表示由 ℓ 个词元组成的文本餐厅评论，其中 z 对应于相关的情感分数。在训练期间，通常会提供一个规模为 M_{unsup} 的大型无监督语料库 $\mathcal{D}_{\text{unsup}} = \{\boldsymbol{x}^{(m)}\}_{m=1}^{M_{\text{unsup}}}$，以及一个规模为 $M_{\text{sup-train}}$、由序列及其对应属性值构成的小得多的监督语料库 $\mathcal{D}_{\text{sup-train}} = \{(\boldsymbol{x}^{(m)}, z^{(m)})\}_{m=1}^{M_{\text{sup-train}}}$。令 α_- 和 α_+ 分别表示 $\mathcal{D}_{\text{sup-train}}$ 中 z 的下界和上界，即对于训练样例中的所有 z，均有 $z \in [\alpha_-, \alpha_+]$。我们将这一区域称为分数的训练区域。

我们的目标是生成属性值大于（或小于）目标属性值 z^* 的序列。具体而言，我们旨在外推到训练区域之外，即根据具体应用选择 $z^* < \alpha_-$ 或 $z^* > \alpha_+$。我们将这些区域称为分数的外推区域。

此外，我们假设可以使用一个评分器 f_s，该评分器在 $\mathcal{D}_{\text{sup-train}}$ 上训练，用于预测每个序列的属性值，即 $\hat{z} = f_s(\boldsymbol{x})$。虽然 f_s 在 z 的训练区域可能表现出色，但它并未使用外推区域的数据进行训练，因此在对此范围内的示例评分时可能表现不佳。因此，f_s 不应 被视为预言机。

3.1. 概述

ICE 的核心组件是一个局部编辑器，它会修改序列中的一个短片段，以提升其属性值。

具体而言，它接收一个输入序列 $\boldsymbol{x}$ 和一个控制标记 c，该标记指定是增加（$c =$ <inc>）还是减少（$c =$ <dec>）属性值，并输出一个改进后的序列 $\tilde{\boldsymbol{x}}$。我们使用 Transformer 编码器-解码器模型（ $p_\theta(\tilde{\boldsymbol{x}} \mid \boldsymbol{x}, c)$）对局部编辑器进行建模（Vaswani

等人，2017年）。我们通过掩码语言建模（第3.2节）合成属性值略有差异的序列对，以此训练编辑器。

在推理时，从初始序列 $\boldsymbol{x}_0$开始，我们会对其进行迭代编辑，直到满足某个停止条件。具体而言，在第 k 次迭代中，我们编辑当前序列 $\boldsymbol{x}_k$，通过以下方式生成 $\boldsymbol{x}_{k+1}$：

$$\boldsymbol{x}_{k+1} \sim p_\theta(\cdot \mid \boldsymbol{x}_k, c) \tag{1}$$

每次迭代都应使$\boldsymbol{x}_k$ 的属性值向 z^*靠近。我们将在第 3.3 节中探讨在推理的每一步选择最佳候选项的不同方法，以及推理过程的停止条件。

3.2. 从扰动中学习局部编辑

为了训练局部编辑器，我们对来自 $\mathcal{D}_{\text{sup-train}}$ 的示例进行扰动，生成向目标值小幅改进的训练对。

具体而言，给定一个来自训练区域的序列$\mathcal{D}_{\text{sup-train}}$，我们在其中随机掩码词元，[3] 并使用掩码语言模型对这些位置进行填充，从而生成其扰动结果（图 1）。掩码语言模型在无监督数据 $\mathcal{D}_{\text{unsup}}$ 上进行训练，以使填充结果生成有效序列。为确保我们仅进行小幅改进，我们使用评分器 f_s预测每个序列的属性值，并仅保留属性值绝对差小于阈值 δ的序列对。

原序列及其扰动组成的每一对样本，都会为编辑器提供两个示例：从原序列生成扰动序列，以及反向生成。请注意，编辑器还会接收一个控制标记，用于指定编辑应增加还是降低属性值。对于每个输入—输出对，如果评分器 f_s 测得输入序列的属性值低于输出序列的属性值，则将控制代码设为 <inc>；否则设为 <dec>。

给定输入序列、输出序列和控制代码组成的元组后，我们便在此数据集上训练编辑器 p_θ。

3.3. 推理

在推理时，我们按照公式（1）中所述，迭代运行编辑器。

解码方法 在每次迭代中，我们尝试两种不同的方法，从一组生成序列中选择最佳候选序列：

- **无评分器生成**：在方程（1）的每次迭代中，我们使用束搜索进行生成

[3]具体的掩码策略取决于任务，并在各个实验部分（第 5 节、第 6 节和第 7 节）中分别进行说明。

依靠*ICE* 模型的似然来控制生成过程。

- **评分器引导生成**：在每次迭代中，我们通过 top-k 采样生成一组序列，使用 f_s 对其评分，并根据期望的目标值，选择得分最高（或最低）的序列。虽然 f_s 在训练区域内可靠，但当 *ICE* 模型生成属性值处于外推区域的序列时，所提供的引导是否有益尚不明确。

停止准则任务目标是编辑输入序列，使其属性值大于（或小于）目标值 z^*。然而，由于 z^* 位于外推区域，因此很难可靠地判断推理过程何时达到这一目标。在本研究中，我们以固定的迭代次数运行推理。关于停止条件的补充讨论见附录C.4。

4. 实验设置

我们在一项自然语言处理任务和两项蛋白质设计任务上评估我们的方法——情感可控生成（第5节）、ACE2 稳定性任务（第6节）以及 AAV 适应度任务（第7节）。

4.1. 评估

我们关注的是衡量模型成功编辑序列的能力，使其属性值高于（或低于）目标值 z^*。在实验中，我们报告成功率，即模型能够编辑并满足这一标准的序列所占比例，该比例由预言机模型确定。预言机因任务而异，并在各实验章节中详细说明（第5节、第6节、第7节）。

4.2. 基线方法

我们将本方法与以下基线方法进行性能比较。(a)**采样**：一种简单的基线方法是直接使用掩码语言模型编辑序列。我们仿照第3.2节中的合成数据创建流程，在初始序列中随机掩码一段内容并进行填充，以改变其属性值。(b) **迭代采样**：为了消融*ICE*中编辑器模型的作用，我们用基于掩码语言模型的掩码并填充编辑器替换它；其余迭代算法与采用*ICE*的评分器引导推理方法相同。(c) **Genhance**：我们将本方法与*Genhance* （Chan 等人，2021a）进行比较。它是一种外推型基线方法，通过在学习得到的、用于编码属性值的潜在空间中施加扰动来实现受控生成。在推理时增大这些扰动的幅度即可实现外推。

对于自然语言处理任务，我们还与另外两种基线方法进行比较。(d) **PPLM**（Dathathri 等人，2020）是一种受控生成方法，在推理时使用评分器引导自回归语言模型的生成，$p(z|\boldsymbol{x})$。我们使用 f_s 作为评分器来引导生成。纳入该基线是为了评估在训练区域上训练的评分器所提供的引导是否能够实现外推。(e) **分数条件生成器**：我们还与一种分数条件模型进行比较，该模型根据输入和目标属性值生成输出序列。[4] 训练分数条件模型时，我们使用相同的合成数据（第3.2节），但将控制代码替换为 f_s测得的输出序列属性值，并将其附加为字符串词元。推理时，我们附加所需的目标分数，并评估模型能否泛化到未见过的分数值。[5]

5. 情感控制

在这项任务中，目标是控制一段短文本（2–3 句话）所表达的情感。我们使用 Yelp 数据集完成这项任务（Zhang 等人，2015年），其中包含 65 万个训练样例和 5 万个测试样例，并平均分布于 1 至 5 的情感分数。我们将情感分数为 2 至 4 的范围定义为训练区域，将分数为 1 至 2 和 4 至 5 的范围定义为外推区域。对于这项任务，我们关注的是衡量模型向两个方向外推的能力，即增强或减弱样例所表达的情感。为此，我们报告将情感编辑到以下目标值以外的成功率：向负面方向为 1.5 和 2.5，向正面方向为 3.5 和 4.5。其中，1.5 和 4.5 属于外推区域。

5.1. 实现细节

训练评分器 我们在Liu 等人，2019年）的 Yelp 数据集训练区域 中的示例上微调 RoBERTa-Large 模型，使其充当评分器，f_s。该评分器是一个回归模型，它接收输入文本并预测其情感分数，即一个介于2 和4之间的实数。附录B 进一步介绍了评分器的训练细节。

训练编辑器 为了通过扰动生成合成数据，我们使用 Lewis 等人 （2020）中描述的策略对词元进行掩码，并使用预训练的 BART-Large 模型。[6] 我们通过设置超参数 $\delta = 0.4$（第 3.2 节）筛选生成的配对。我们使用合成训练数据对 T5-Base 模型（Raffel 等人，2022）进行微调，以获得局部编辑器。附录B 进一步介绍了训练细节。

推理 我们使用第3.3节所述的两种方法进行推理。对于无评分器推理，我们采用束搜索，束宽为5。进行评分器引导推理时，在每次迭代中，我们使用 top-k 采样生成 $k = 5$个序列，温度设为0。7；然后使用 f_s选出最佳序列。两种方法均进行10 步迭代编辑。

评估 我们报告了在 Yelp 数据集测试集中随机抽取的 1831 个样本上，针对上述全部 4 个目标的结果。[7] 为评估最终生成序列的属性值是否外推至训练区域之外，我们通过预言机估算真实情感分数——该预言机是一个在整个 Yelp 数据集（即训练 区域和外推区域）上微调的 RoBERTa-Large 模型。

基线方法 对于情感控制，我们将本方法与采样、迭代采样、*Genhance*、*PPLM*和分数条件生成器进行比较。我们使用 T5-Base 训练分数条件生成器，使其与*ICE* 编辑器相匹配。*Genhance* 模型的架构同样基于 T5-Base，因此其规模与*ICE* 编辑器相当。在推理时，对于每个测试样本，我们从*Genhance* 采样 50 个序列，并使用 f_s 选出最佳序列，使其生成的序列总数与*ICE* 在所有迭代中生成的序列总数相同。对于迭代采样，我们每次迭代生成 5 个序列，共迭代 10 次，并在每次迭代中使用 f_s 选出最佳序列，与*ICE*相同。

5.2. 结果

***ICE* 在**外推区域从表 1中，我们可以看到，*ICE* 模型（在评分器引导下）在外推区域的表现远超基线方法。即使没有评分器，*ICE*模型的表现也能与最强的基线 *Genhance*不相上下。附录7 中的 C.2 展示了一个例子，说明如何通过多次迭代提高句子的情感强度。

[4]该基线与Jain 和 Berg-Kirkpatrick (2021)以及Chen 等人 (2021)所描述的方法相似。[5]分数条件基线在最小编辑数据上进行训练；在测试时，我们评估其泛化到较大编辑的能力，这颇具挑战性。调整训练数据以纳入较大编辑，或许能提升该基线的性能；然而，在我们的问题设置中，我们没有 $\mathcal{D}_{\text{sup-train}}$中的示例序列对。

[6]掩码策略包括从伯努利分布（$p = 0.8$）中采样跨度起始位置，然后从截断泊松分布（$\lambda = 6$）中采样要掩码的词元数量。跨度大小上限设为 12。我们在附录 6 C.2中，表格 中报告了更多掩码策略的变体。[7]我们确保所选示例的输入文本情感值处于训练区域内。

方法	训练区域中的目标			目标（位于）		
	区域			外推区域		
	3.5	2.5	平均值	4.5	1.5	平均值
采样	0.362	0.259	0.310	0.061	0.050	0.056
迭代采样	0.668	0.657	0.663	0.320	0.328	0.324
Genhance	**0.982**	0.833	0.908	0.482	0.291	0.387
分数条件模型	0.780	0.766	0.773	0.212	0.217	0.215
PPLM	0.534	0.516	0.522	0.081	0.065	0.077
ICE 无评分器	0.976	**0.918**	**0.947**	0.446	0.305	0.376
ICE 使用评分器	0.943	0.900	0.921	**0.638**	**0.582**	**0.610**

表 1：情感控制任务的结果。我们报告的成功率，是指根据预言机判定，情感值高于（或低于）目标分数的示例所占比例。粗体数值表示最高的外推成功率。迭代采样、*Genhance*和*PPLM* 在推理时使用评分器。与基线方法相比，*ICE* 在外推区域取得了最高的成功率。

评分器引导有益 我们观察到，评分器有助于提升迭代采样 基线和*ICE* 的情感控制效果。迭代采样在评分器的帮助下受益匪浅，外推性能从.6%提高到32.4%，而采样的外推性能为5%。在评分器引导下，*ICE* 的成功率从37.6%提高到61.0%。我们确实观察到，尽管使用了评分器，*PPLM* 的外推性能仍然很差 f_s。这表明，f_s 在用于对生成的序列进行排序时，可能更适合引导推理，例如*ICE* 和迭代采样；而像*PPLM*那样，直接使用 f_s 的条件概率引导生成，则不如前者。

ICE 每次迭代都做了什么？为了分析文本的情感分数如何随迭代变化，我们绘制了每次迭代输出的情感分数与初始序列情感分数之差。我们从测试集中随机抽取 100 个示例，并使用*ICE* 提高它们的情感分数。我们使用无评分器推理，收集*ICE* 每次迭代的输出。随后，我们在图2中绘制了第 1、4、7 和 10 次迭代中情感分数增幅（相对于初始分数）的直方图。随着迭代次数增加，情感分数的增幅也变大（即分布的众数向右移动），尽管编辑并非总能成功（少数输出的分数低于初始分数，落在

thenegativebuckets）。总体而言，这表明*ICE* 能够通过迭代编辑平均提高情感分数。

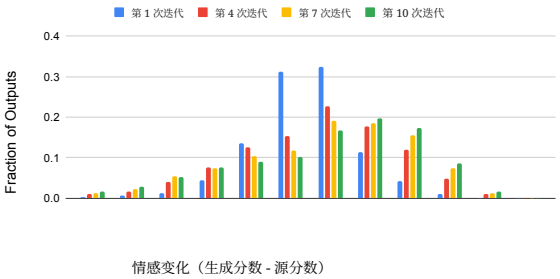


图 2：我们绘制了在 100 个示例上，ICE 每次迭代时情感得分相对于初始分数的增幅直方图。随着迭代次数增加，我们观察到分布的众数向正向移动，这表明越来越多的示例被编辑得更加积极，最终导致外推。

6. ACE2 数据集上的蛋白质设计

开发能够生成更稳定蛋白质的方法，有助于药物发现，因为与现有蛋白质相比，这些蛋白质可能更易于储存，临床效果也更可靠（Wang，1999年；Shire 等，2004；Bloom 等，2006；Deller 等，2016年；Webber 等，2016年）。本任务的目标是生成野生型人血管紧张素转换酶 2（ACE2）的突变体——

类型序列[8] 具有更高的稳定性。突变体的稳定性通过 FoldX（Schymowitkz 等人，2005）测量，即相对于野生型的自由能变化，或 ddG。[9] 野生型本身的 ddG 值为零，数值越负代表突变体越稳定。这项合成任务由 Chan 等（2021a）创建，我们复现了他们的设置。蛋白质表示为由 20 种不同氨基酸组成的词表中选取的 83 个氨基酸构成的序列。为避免突变与野生型差异过大，所有突变中都固定保留一段长度为 8 个氨基酸的片段（NTNITEEN）。我们将 ddG 值从-4 到 $+10$的范围视为训练区域。外推区域指低于 -4的 ddG 值。对于此任务，我们旨在生成 ddG 值更负的突变体。我们通过报告生成 ddG 值低于目标值的突变的成功率来衡量这一点：训练区域中的目标值 z^*、 -1 和 -2.5，以及

[8] https://www.uniprot.org/uniprotkb/Q9BYF1/entry [9] https://foldxsuite.crg.eu/

该外推区域，-5，-6，以及 -7。

6.1. 实现细节

训练评分器 为了训练 f_s，我们在Elnaggar 等人，2021）的示例上微调 *ProtBert*，这些示例的 ddG 值位于该数据集的训练区域，数据集来自Chan 等人（2021a）。

训练编辑器 我们使用预训练的 Prot-T5-XL 模型（第 3.2 节中介绍的掩码填充方法，构造序列对（Elnaggar 等人，2021）。我们从伯努利分布中采样词元掩码，参数为（$p = 0.8$）。为过滤较小的扰动，我们将 δ 设为 1。5。随后，我们在这些数据上微调 Prot-T5-XL，使其作为*ICE* 编辑器。

推理 在推理时，我们从野生型序列开始，生成有评分器和无评分器的突变，f_s（第3.3节）。使用评分器时，我们在每一步采样5 条序列，使用 f_s选出最佳序列，并将这一过程重复10 次。进行无评分器推理时，我们以5 的束宽生成序列，共进行10 次迭代。[10]

评估 在 ACE2 任务中，我们关注的是生成具有较低 *ddG* 值的10,000 个野生型突变体。我们使用每个模型生成突变体，并报告生成的突变体的成功率：使用 FoldX 作为预言机时，其 *ddG*值低于各项任务目标。我们采用 Chan 等人的参数 (2021a) 来评估突变。此外，我们还报告预言机评定的前100和前1000个突变体的平均分数，以评估每个模型生成的 10,000 个候选文库中顶尖候选序列的质量。

基线方法 我们将所提出的方法与 采样、迭代采样和 *Genhance*进行比较。[11] 对于*Genhance*，我们报告 Chan 等人发布的模型在 (2021a) 中生成的 10,000 个突变体的结果，包括使用和不使用评分器的情况。该模型同样基于 Prot-T5-XL，因此可与 *ICE* 模型直接比较。对于 迭代采样，我们每次迭代生成5 条序列，共进行 10 次迭代。

6.2. 结果

***ICE* 在外推方面优于基线方法**表 2显示，*ICE* 始终优于*Genhance*、采样和迭代采样，在所有外推目标上均如此。此外，从表 3中可以看出，*ICE* 取得了更低的平均值*ddG*，在前100和前1000条序列上均如此。有趣的是，尽管迭代采样的外推率高于*Genhance*（表2），*Genhance* 在前1000和前100个子集上的平均分数更高（表 3），这表明*Genhance* 产生的突变体数量较少，但稳定性略高（尽管仍不如*ICE*）。

评分器对于 ACE2 中的所有模型都很有价值。在这项任务中，我们从野生型开始生成（*ddG* 分数为零），评分器 f_s能够可靠地引导生成过程，直至分数达到 -5。因此，我们发现所有方法均能从使用评分器中显著受益（表2）。在图3中，我们绘制了由*ICE* 及所报告的基线方法生成的突变的分数直方图。从图3a可以看出，与图 -5 相比，所有模型的分数分布峰值均向负方向移动，中心更接近3b，凸显了评分器的价值。不过，我们注意到，即使在无评分器设定下，我们的方法也能实现一定程度的外推，远优于采样，且其外推率高于*Genhance*。

7. 腺相关病毒数据集上的蛋白质设计

AAV 数据集（Bryant 等人，2021）旨在研究腺相关病毒（AAV）衣壳蛋白的适应度景观，该蛋白是基因治疗的关键组成部分（Russell等，2017）。我们的目标是获得适应度值更高的 AAV-2 野生型序列突变体[12]。我们在实验中使用 FLIP 基准测试（Dallago 等，2021）提出的数据划分。每个突变体序列的长度在 734 到 750 之间。突变发生在野生型序列的第 561 至 588 个位置。我们使用数据集提供的低值与高值划分来界定训练区域和外推区域。训练区域对应适应度值低于零的部分，外推区域则对应适应度值为正的部分。在推理时，生成过程从适应度分数为零的野生型序列开始，模型预期生成适应度分数为正的突变体。我们针对目标值评估性能：训练区域中的目标值为 z^*，-1；外推区域中的目标值为 0、1 和 2。

7.1. 实现细节

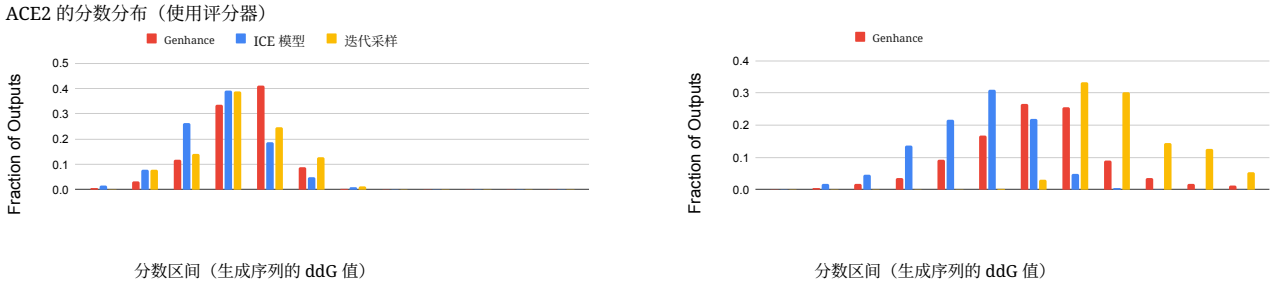
训练评分器评分器 f_s 是一个 CNN 模型，使用训练区域中的示例进行训练。CNN 的架构和超参数是根据 FLIP 基准测试选定的。[13] 评分器接受一个对应于蛋白质的字符串，并输出一个浮点数

[10]我们在附录C.3中进一步分析了生成超参数对性能变化的影响。[11]ACE2 任务要求针对特定野生型生成突变体。蛋白质领域的预训练自回归语言模型无法直接生成突变体，只能续写序列。因此，序列到序列模型更适合该任务。因此，*PPLM*依赖于自回归模型，故未被纳入基线方法。此外，我们也未纳入*Score-ConditionedGenerator* 这一基线，因为 Prot-T5-XL分词器的词表仅由氨基酸组成，因而无法将输出分数作为一个词元与输入一同接受。

[12]https://www.uniprot.org/uniprotkb/P03135/entry[13]在低值与高值划分上，评分器的训练相关系数为 0。82，测试相关系数为 0。34。这与该基准测试中在此划分上取得的最佳测试相关系数一致。

方法	训练区域中的目标		目标		
	训练区域		外推区域		
	-1	-2.5	-5	-6	-7
采样	0.033	0.007	0.000	0.000	0.000
迭代采样	0.998	0.954	0.220	0.079	0.001
无评分器的 Genhance	0.570	0.219	0.021	0.005	0.001
带评分器的 Genhance	**0.999**	**0.978**	0.159	0.040	0.009
ICE 无评分器	0.945	0.598	0.062	0.017	0.002
ICE 搭配评分器	0.998	0.974	**0.361**	**0.098**	**0.019**

表 2：ACE2 任务的结果。目标是生成比野生型稳定性更高（即*ddG* 值更低）的突变体。每个表格单元格表示生成的突变低于相应目标值的成功率。粗体数值表示最高的外推成功率。与已报告的基线方法相比，*ICE* 实现了更高的外推成功率。



(a) (b)

图3：最终生成突变的*ddG* 分数（越低越好）的直方图，比较了*ICE* 和基线方法在 ACE2 任务上的表现。使用评分器（图 3a）和不使用评分器（图 3b）引导推理时，ICE 生成的突变质量都高于基线方法。此外，评分器显著提升了所有方法的性能。

文库大小	迭代采样	Genhance	*ICE*
全部 1 万个	-4.326	-4.086	**-4.660**
前 1000 个	-5.866	-6.030	**-6.575**
前 100 个	-6.413	-7.354	**-7.938**

表 3：由迭代采样、*Genhance*和*ICE*（均使用评分器）生成的突变的平均*ddG*值（越低越好）。我们报告所有 10000 个突变、由预言机确定的前 1000 个突变和前 100 个突变的平均分数。粗体数值表示最低的平均 *ddG* 值。*ICE* 生成的突变最稳定。

适应度值。

训练编辑器我们创建训练 *ICE* 模型所需的成对样本，采用与 ACE2 中相同的策略。我们使用 Prot-T5-XL （ Elnaggar 等人，2021）模型填充可变区域中的掩码，并使用评分器对样本对进行评分， f_s，以创建编辑器训练数据。[14] 随后，我们在该数据集上对 Prot-T5-XL 进行微调。由于突变体的长度超过了 Prot-T5-XL 的序列长度限制，我们对其进行截断从起始位置到最后 512 个词元，这些词元始终包含蛋白质的整个可变区域。

推理 我们从野生型开始，按照第*ICE* 节所述，在模型上进行推理 3.3。使用评分器时，我们采样 5 个生成序列，使用 f_s 对它们评分，选出最佳序列，然后重复 10 次迭代。对于无评分器设置，我们以束宽5 进行 10 次迭代。

评估 我们使用每种方法生成 10,000 个突变体，并报告生成突变的成功率，这些突变的分数高于目标分数 z^*。由于没有进行湿实验，我们通过预言机模型获取每条生成序列的适应度分数；该模型是一个在 AAV 数据集采样（ i.i.d. ）划分上训练的卷积神经网络。[15] 选择该数据划分，是因为采样 划分中的样本适应度值涵盖了训练区域和外推区域。

基线方法 我们将我们的方法与采样 和迭代采样基线方法进行比较。[16]

[14]我们再次将超参数 δ 设为 1。5。

[15]我们选择 CNN 架构，因为它在基准测试中与金标准适应度值的斯皮尔曼相关系数最高（Dallago等人，2021）。该模型在此数据划分上的训练斯皮尔曼相关系数为0。93，测试相关系数为0。92 。

[16]如前所述，*PPLM* 和分数条件生成器基线方法并不适用于蛋白质任务。

方法	目标 训练区域中的目标 -1	目标 外推区域中的目标 0	 1	 2
采样	0.058	0.018	0.011	0.000
使用评分器的迭代采样	**0.524**	0.064	0.017	0.000
ICE 无评分器	0.481	0.188	0.033	0.001
ICE 使用评分器	0.521	**0.223**	**0.036**	**0.002**

表 4：AAV 任务的结果。目标是生成适应度值高于源蛋白质的突变。我们报告生成适应度值高于相应目标的突变的成功率。粗体数值表示最高的外推率。*ICE* 的外推率高于基线方法。

文库 Size	采样器- ling	迭代 采样	*ICE* 评分器- Free	*ICE* 带评分器
全部 10k	-3.450	-1.390	-1.150	**-1.040**
前 1,000 名	-0.567	-0.584	0.403	**0.918**
前 100 名	1.605	1.550	1.452	**1.750**

表 5：通过采样、迭代采样和*ICE*生成的突变的平均适应度值（越高越好）。我们报告全部 10000 个突变的平均分数、排名前 1000 个突变的平均分数，以及由预言机确定的排名前 100 个突变的平均分数。粗体数值表示最高的平均适应度值。*ICE* 生成的突变质量最高。

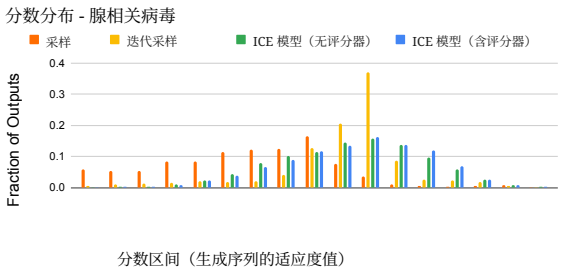


图 4：在 AAV 任务上，各方法生成的突变体适应度值直方图（分数越高越好）。*ICE* 的生成效果优于采样 和迭代采样。

7.2. 结果

***ICE* 模型的外推能力优于** 迭代采样从表 4可以看出，*ICE* 采用 无评分器和 评分器引导 推理时，外推成功率分别高于 采样 和 迭代采样。我们还观察到，*ICE* 采用 评分器引导 推理时，在包含 10000 个突变的整个文库以及各方法生成的 Top-100 和 Top-1000 个突变子集中，平均适应度分数均高于基线方法。最后，理想情况下，生成的突变库不仅应具有较高的适应度值，还应呈现多样性（Calcedo 等人，2009）。我们通过考察附录 *ICE* 生成的突变与野生型之间的编辑距离，观察到 C.1 中，

评分器在 AAV 上的效果较差 从表 4中可以看出，在使用评分器时，两种方法在 训练区域 和 外推区域 目标上的表现，相较于无评分器设置仅有小幅提升。分数分布（图 4）也呈现出类似趋势。我们看到，对于两种方法，分数分布的众数都位于 训练区域 内，靠近外推区域 的边界（图 4 ）。*ICE* 的分布平坦得多，因此与迭代采样相比，其外推成功率更高。由于生成过程从训练区域 的边缘（零点）开始，我们预计评分器无法在 AAV 中提供太多可靠的指导。

8. 结论

我们提出了迭代受控外推（*ICE*），一种用于外推式受控生成的迭代方法。在自然语言处理和蛋白质设计任务中，我们的方法均显著优于现有的可控生成方法以及更复杂的外推技术。未来潜在的研究方向包括：将迭代方法扩展到多个属性，以生成以新颖方式组合这些属性的序列；训练能够泛化到外推区域的评分器；以及通过融入更多领域知识来改进合成数据创建技术。

致谢

我们感谢 David Belanger、Lucy Colwell 和 Nitish Joshi 在项目期间给予的宝贵讨论和反馈。本项目是 Google Research Collabs 计划的一部分。本项目还获得了三星先进技术研究院（下一代深度学习：从模式识别到人工智能）、美国国家科学基金会第 1922658 号资助，以及 AWS AI 的捐赠。

参考文献

Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schul- man, J., and Man´e, D. 人工智能安全中的具体问题。*arXiv* 预印本 *arXiv:1606.06565*，2016年。

Angermueller, C., Belanger, D., Gane, A., Mariet, Z., Dohan, D., Murphy, K., Colwell, L., and Sculley, D. 面向生物序列设计的基于群体的黑箱优化。载于国际机器学习会议，第324–334页。PMLR，2020a。

Angermueller, C., Dohan, D., Belanger, D., Deshpande, R., Murphy, K., and Colwell, L. 面向生物序列设计的基于模型的强化学习。载于 国际学习表征会议，2020b。网址 https://openreview.net/forum?id=HklxbgBKvr。

Arnold, F. H. 通过定向进化进行设计。《化学研究述评》，31(3):125–131，1998。

Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. 蛋白质稳定性促进可进化性。《美国国家科学院院刊》，103(15):5869–5874，2006。doi: 10.1073/pnas.0510098103。网址 https://www.pnas.org/doi/abs/10.1073/pnas.0510098103。

Brookes, D., Park, H. 和 Listgarten, J. 通过自适应采样进行条件化，以实现鲁棒设计。载于国际机器学习会议，第 773–782 页。PMLR，2019年。

Bryant, D. H., Bashir, A., Sinai, S., Jain, N. K., Ogden, P. J., Riley, P. F., Church, G. M., Colwell, L. J. 和 Kelsic, E. D. 利用机器学习对腺相关病毒衣壳蛋白进行深度多样化。《自然·生物技术》，39(6):691–696，2021。

Calcedo, R., Vandenberghe, L. H., Gao, G., Lin, J. 和 Wilson, J. M. 腺相关病毒中和抗体的全球流行病学。《传染病杂志》，199(3):381–390，2009。

Chan, A., Madani, A., Krause, B. 和 Naik, N. 面向属性增强生成的深度外推。载于 Beygelz- imer, A., Dauphin, Y., Liang, P. 和 Vaughan, J. W.（编），神经信息处理系统进展，2021a。网址 https://openreview.net/forum?id=NCDMYD2y5kK。

Chan, A.、Ong, Y.-S.、Pung, B.、Zhang, A. 和 Fu, J. Cocon：一种用于可控文本生成的自监督方法。载于 国际学习表征会议，2021b。网址：https://openreview.net/forum?id=VDozqvBy4W。

Chan、H.P.、Wang、L. 和 King。利用受约束马尔可夫决策过程实现可控摘要。计算语言学协会汇刊，9:1213–1232，2021c。doi：10.1162/tacl a 00423。网址https://aclanthology.org/2021.tacl-1.72。

Chen, L.、Lu, K.、Rajeswaran, A.、Lee, K.、Grover, A.、Laskin, M.、Abbeel, P.、Srinivas, A. 和 Mordatch, I. 决策变换器：通过序列建模实现强化学习。载于 Beygelzimer, A.、Dauphin, Y.、Liang, P. 和 Vaughan, J. W.（编），神经信息处理系统进展，2021。网址：https://openreview.net/forum?id=a7APmM4B9d。

Dallago, C.、Mou, J.、Johnston, K. E.、Wittmann, B.、Bhattacharya, N.、Goldman, S.、Madani, A. 和 Yang, K. K. FLIP：蛋白质适应度景观推断中的基准任务。载于 第三十五届神经信息处理系统会议数据集与基准测试专题，2021。网址：https://openreview.net/forum?id=p2dMLEwL8tF。

Dathathri, S., Madotto, A., Lan, J., Hung, J., Frank, E., Molino, P., Yosinski, J. 和 Liu, R. Plug and play 语言模型：一种受控文本生成的简单方法。载于 国际学习表征会议，2020年。网址 https://openreview.net/forum?id=H1edEyBKDS。

Deller, M. C., Kong, L. 和 Rupp, B. 蛋白质稳定性：晶体学家的视角。《晶体学报》*F*辑：结构生物学通讯，72(2)：72–95，2016年。

Devlin, J., Chang, M.-W., Lee, K. 和 Toutanova, K. BERT：用于语言理解的深度双向变换器预训练。载于*2019*年计算语言学协会北美分会会议论文集：人类语言技术，第*1*卷（长篇与短篇论文），第4171–4186页，明尼阿波利斯，明尼苏达州，2019年6月。计算语言学协会。doi：10.18653/v1/N19-1423。网址 https://aclanthology.org/N19-1423。

Elnaggar, A., Heinzinger, M., Dallago, C., Rehawi, G., Wang, Y., Jones, L., Gibbs, T., Feher, T., Angerer, C., Steinegger, M. 等。Prottrans：通过自监督学习理解生命的语言。*IEEEtransactions onpattern analysis and machine intelligence*, 44(10):7112–7127, 2021.

Freschlin, C. R., Fahlberg, S. A. 和 Romero, P. A. 利用机器学习探索蛋白质工程的适应度景观。《生物技术当前观点》, 75: 102713, 2022.

Gehman, S., Gururangan, S., Sap, M., Choi, Y., 和 Smith, N. A. RealToxicityPrompts：评估语言模型中的神经毒性退化。载于计算语言学协会论文集：*EMNLP 2020*

EMNLP 2020，第 3356–3369 页，线上会议，2020 年 11 月。计算语言学协会。doi: 10.18653/v1/2020.findings-emnlp.301。网址: https://aclanthology.org/2020.findings-emnlp.301。

Gligorijevi´c, V., Berenberg, D., Ra, S., Watkins, A., Kelow, S., Cho, K. 和 Bonneau, R. 通过深度流形采样进行功能引导的蛋白质设计。*bioRxiv*, 2021年。doi: 10.1101/2021.12.22.473759。网址 https://www.biorxiv.org/content/early/2021/12/23/2021.12.22.473759。

Gong, H., Bhat, S., Wu, L., Xiong, J. 和 Hwu, W.-M. 基于强化学习的文本风格迁移，无需平行训练语料库。在《*2019*年北美计算语言学协会年会论文集: 人类语言技术》，第*1*卷（长篇与短篇论文），第3168–3180页，2019年。

Guu, K., Hashimoto, T. B., Oren, Y., 和 Liang, P. 通过编辑原型生成句子。《计算语言学协会会刊》，6:437–450, 2018年。

He, H., Peng, N., 和 Liang, P. 利用意外性生成双关语。载于 计算语言学协会北美分会（北美计算语言学协会），2019年。

Ibarz, B., Leike, J., Pohlen, T., Irving, G., Legg, S., 和 Amodei, D. 基于人类偏好和示范的 Atari 奖励学习。《神经信息处理系统进展》，31, 2018年。

Jain, A. 和 Berg-Kirkpatrick, T. 标量控制下文本生成外推的实证研究。《arXiv 预印本 arXiv:2104.07910》，2021。

Keskar, N. S., McCann, B., Varshney, L. R., Xiong, C., 和 Socher, R. CTRL：一种用于可控生成的条件式 Transformer 语言模型。《arXiv 预印本 arXiv:1909.05858》，2019年。

Lewis、M.，Liu、Y.，Goyal、N.，Ghazvininejad、M.，Mohamed、A.，Levy、O.，Stoyanov、V. 和 Zettlemoyer、L. 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。在第 *58* 届计算语言学协会年会论文集，第 7871–7880 页，2020 年。Li、X. L.，Thickstun、J.，Gulrajani、I.，Liang、P. 和 Hashimoto、T. Diffusion-LM 改进可控文本生成。在神经信息处理系统进展，2022。网址: https://openreview.net/forum?id=3s9IrEsjLyk。

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

Lu, X., Welleck, S., Hessel, J., Jiang, L., Qin, L., West, P., Ammanabrolu, P., 和 Choi, Y. Quark：通过强化遗忘实现可控文本生成。神经信息处理系统进展，35:27591–27609，2022。

Lyu, Y., Liang, P.P., Pham, H., Hovy, E., P´oczos, B., Salakhutdinov, R., 和 Morency, L.-P. StylePTB：用于细粒度可控文本风格迁移的组合式基准。在*2021*年北美计算语言学协会年会：人类语言技术论文集，第 2116–2138 页，线上举行，2021年6月。计算语言学协会。doi：10.18653/v1/2021.naacl-main.171。网址: https://aclanthology.org/2021.naacl-main.171。

Madani, A., McCann, B., Naik, N., Keskar, N. S., Anand, N., Eguchi, R. R., Huang, P.-S.，以及 Socher, R. ProGen：用于蛋白质生成的语言建模。*arXiv* 预印本 *arXiv:2004.03497*，2020 年。

Madani, A., Krause, B., Greene, E. R., Subramanian, S., Mohr, B. P., Holton, J. M., Olmos, J. L., Xiong, C., Sun, Z. Z., Socher, R., Fraser, J. S.，以及 Naik, N. 深度神经语言建模支持跨家族的功能性蛋白质生成。*bioRxiv*, 2021 年。doi：10.1101/ 2021.07.18.452833。
U
R
L
https://www.biorxiv.org/content/early/2021/07/18/2021.07.18.452833。

Madani, A., Krause, B., Greene, E. R., Subramanian, S., Mohr, B. P., Holton, J. M., Olmos Jr, J. L., Xiong, C., Sun, Z. Z., Socher, R., 等。大语言模型能够生成跨越多样化家族的功能性蛋白质序列。《自然·生物技术》，第 1–8 页，2023。

Mallinson, J., Adamek, J., Malmi, E. 和 Severyn, A. EdiT5：采用 T5 热启动的半自回归文本编辑。收录于计算语言学协会：*EMNLP 2022* 论文集，第 2126–2138 页，阿布扎比，阿拉伯联合酋长国，2022年12月。计算语言学协会。网址: https://aclanthology.org/2022.findings-emnlp.156。

Mueller, J., Gifford, D., 和 Jaakkola, T. 序列到更优序列：组合结构的连续修订。载于国际机器学习会议，第 2536–2544 页。PMLR，2017年。Novak, R., Auli, M., 和 Grangier, D. 用于机器翻译的迭代优化。*arXiv* 预印本 *arXiv:1610.06602*，2016年。

Pang, R. Y., Padmakumar, V., Sellam, T., Parikh, A.P., 和 He, H. 条件文本生成中的奖励投机。*arXiv* 预印本 *arXiv:2211.08714*，2022。

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. 和 Liu, P. J. 使用统一的文本到文本转换器探索迁移学习的极限。机器学习研究杂志，21 (1)，2022年6月。ISSN 1532-4435。

Ren, Z., Li, J., Ding, F., Zhou, Y., Ma, J. 和 Peng, J. 用于模型引导蛋白质序列设计的近端探索。发表于国际机器学习会议，第 18520–18536 页。PMLR，2022。

Romero, P. A. 和 Arnold, F. H. 通过定向进化探索蛋白质适应度景观。自然评论：分子细胞生物学，10(12):866–876，2009。

Russell, S., Bennett, J., Wellman, J. A., Chung, D. C., Yu, Z.- F., Tillman, A., Wittes, J., Pappas, J., Elci, O., McCague, S. 等。维替格奈帕沃韦（aav2-hrpe65v2）在rpe65介导的遗传性视网膜营养不良患者中的疗效和安全性：一项随机、对照、开放标签的3期试验。《柳叶刀》，390(10097):849–860，2017年。

Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F. 和 Serrano, L. FoldX网络服务器：一种在线力场。《核酸研究》，33(suppl 2):W382–W388，2005年。

Shire, S. J., Shahrokh, Z. 和 Liu, J. 高蛋白浓度制剂开发中的挑战。《药学科学杂志》，93(6):1390–1402，2004年。

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. 注意力就是你所需要的一切。神经信息处理系统进展，30，2017年。

Verkuil, R., Kabeli, O., Du, Y., Wicky, B. I., Milles, L. F., Dauparas, J., Baker, D., Ovchinnikov, S., Sercu, T., and Rives, A. 语言模型能够泛化到天然蛋白质之外。*bioRxiv*，2022。

Wang, W. 液体蛋白质药物制剂的不稳定性、稳定化与配制。《国际药剂学杂志》，185(2):129–188，1999年。

Webber, M. J., Appel, E. A., Vinciguerra, B., Cortinas, A. B., Thapa, L. S., Jhunjhunwala, S., Isaacs, L., Langer, R., and Anderson, D. G. 生物制药的超分子聚乙二醇化。《美国国家科学院院刊》, 113(50):14189–14194, 2016年。

Welleck, S., Lu, X., West, P., Brahman, F., Shen, T., Khashabi, D., and Choi, Y. 通过学习自我纠错来生成序列。载于第十一届国际学习表征会议，2023年。网址：https://openreview.net/forum?id=hH36JeQZDaO。

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., and Rush, A. Transformer：最先进的自然语言处理。在*2020*年自然语言处理实证方法会议论文集：系统示范，第38–45页，线上，2020年10月。计算语言学协会。doi: 10.18653/v1/2020.emnlp-demos.6。网址https://aclanthology.org/2020.emnlp-demos.6。

Yang, K. 和 Klein, D. FUDGE：利用未来判别器进行可控文本生成。载于《计算语言学协会北美分会*2021*年会议论文集：人类语言技术》，第 3511–3535 页，线上，2021年6月。计算语言学协会。doi: 10.18653/v1/2021.naacl-main.276。网址：https://aclanthology.org/2021.naacl-main.276。

Yang, K. K.、Wu, Z. 和 Arnold, F. H. 机器学习引导的定向进化在蛋白质工程中的应用。《自然·方法》，16(8)：687–694，2019年。

Zhang, X.、Zhao, J. 和 LeCun, Y. 用于文本分类的字符级卷积网络。神经信息处理系统进展，28，2015年。

A. 局限性

合成数据的创建可能会在自然语言中引入幻觉 我们的方法依赖掩码语言建模来创建经过最小扰动的序列对（第3.2节）。在自然语言任务中，这可能导致扰动后的序列在含义上与源序列略有不同。因此，*ICE* 模型在经过训练后也可能改变序列的含义。特别需要指出的是，如果不加以妥善考虑，文本生成模型产生的某些类型的幻觉可能会造成危害。具体而言，在表 7中，模型可以修改与餐厅食物或氛围相关的情感，但我们希望模型保留基本信息，即作者和他的伴侣正在斯科茨代尔的一家寿司餐厅用餐。今后，我们计划研究更好的合成数据创建策略，以衡量并减轻这种情况。

训练区域中的编辑能够泛化到训练区域**泛化至**外推区域我们的工作依赖于在属于训练区域的序列所受扰动上训练模型。随后，我们反复编辑序列，以提高或降低其分数，使其进入外推区域。尽管实验结果令人鼓舞，但我们认为，这一假设并不适用于所有任务和领域。未来我们计划对此展开进一步研究。

依赖训练好的模型对序列进行评分为评估情感控制和 AAV 任务，我们训练分类器模型来衡量序列的属性值。这些模型只能估计属性值的真实值，并可能最终从数据集中学到虚假相关性。我们指出，这些模型是用来将我们的方法与各种基线方法进行基准测试的工具。特别是对于 AAV 等蛋白质，在实际应用之前，应对预言机模型进行详细分析，或开展真实的湿实验。

迭代方法的推理速度较慢由于我们方法本身的特性，迭代编辑序列的推理速度远慢于*Genhance*等模型进行的单步编辑。

B. 模型训练的更多细节

我们使用 HuggingFace 库（Wolf 等人，2020）对实验中使用的所有语言模型进行微调。实验所用的全部代码和训练好的模型均可在以下网址获取：`https://github.com/vishakhpk/iter-extrapolation`。

情感控制 用于评估的评分器和预言机模型均为经过微调的 RoBERTa-Large（Liu 等人，2019年）模型。预言机使用整个 Yelp 数据集进行训练。评分器使用情感分数在 2 到4之间的样本进行训练。评分器和预言机均经过微调，以优化数据集金标准标签上的均方误差损失。我们使用评分器和预训练的 T5-Base（Raffel 等人，2022）模型创建配对数据，用于训练 *ICE* 生成器模型。我们创建了 100K 对数据，并微调 T5-Base，使其充当 *ICE* 生成器。用于筛选合成对的超参数 $\delta = 0.4$ 是根据一次小规模内部试验选定的。我们微调 T5-Base，使其根据输入序列生成合成对的输出，并在输出词元上优化交叉熵损失。对于上述每种情况，我们都使用 HuggingFace 代码库推荐的超参数，并将学习率从 $1e{-}6$ 扫描至 $1e{-}3$。

ACE2 对于 ACE2，我们对 ProtBert（Elnaggar 等人，2021）模型进行微调。该模型可通过 HuggingFace 获取，用于根据数据集中提供的突变体预测 *ddG* 值。该数据集由Chan 等人 （2021a）发布。我们在金标准标签上优化均方误差损失，并使用验证损失选择最佳检查点。我们利用该模型创建一个包含 1M 对的合成数据集，用于微调 *ICE* 生成器模型。我们在这些合成对上微调 Prot-T5-XL（Elnaggar 等人，2021），以根据输入序列生成合成对的输出，并针对输出词元优化交叉熵损失。我们再次采用 HuggingFace 代码库推荐的超参数，并对学习率进行网格搜索，范围从 $1e{-}6$ 到 $1e{-}3$。使用 FoldX 进行评分时，我们采用与（Chan 等人，2021a）相同的参数。

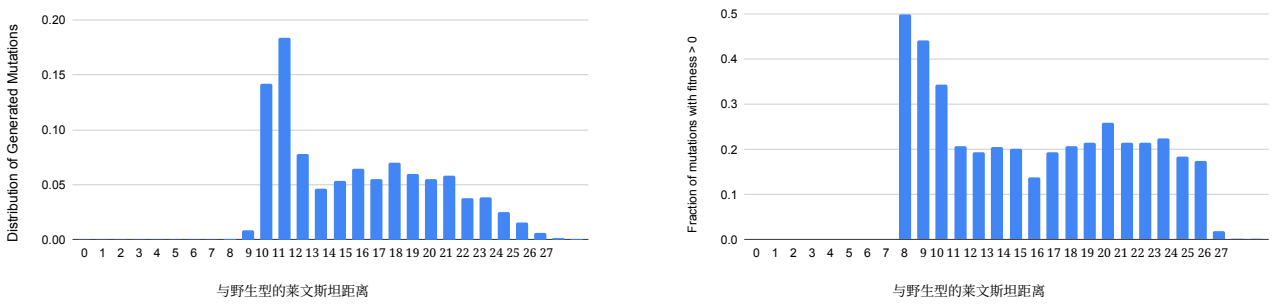
腺相关病毒 AAV 任务的评分器和预言机模型均为卷积神经网络模型，它们将蛋白质序列作为字符串输入，并输出与适应度值对应的实数。我们根据 FLIP 基准测试中指定的参数选择模型架构（Dallago等人，2021）。我们采用相同的方法，选取在 AAV 低-高 划分上取得最高测试集斯皮尔曼相关系数的模型。两个卷积神经网络模型均从基准测试的代码库中训练，通过优化适应度值的均方误差损失函数进行学习。我们使用评分器创建 1M 个合成对，用于训练 *ICE* 生成器模型；训练时，针对给定的输入蛋白质序列和相应控制标签，优化输出词元的交叉熵损失。

C. 其他发现

C.1. 探索腺相关病毒突变体的多样性

尽管腺相关病毒衣壳在基因治疗方面前景可期，但既往接触腺相关病毒产生的免疫会使 20–80% 的人群无法接受此类治疗（Calcedo 等人，2009）。因此，我们不仅要生成适应度高的腺相关病毒突变体，还要使其相较于野生型具有显著的多样性。为此，在图 4中，我们分析了模型生成的序列（在第10次迭代中）与野生型之间的莱文斯坦距离分布。结果显示，尽管生成的突变大多具有约8–10的编辑距离，模型仍能生成与野生型相差多达 25 处编辑的突变（图 5a）。然而，即使模型进行超过 20 处编辑，该区间内样本所占比例仍为0。2，这表明生成的突变具有很高的多样性（图 5b）。

我们注意到，该模型会生成与野生型具有广泛莱文斯坦距离的突变体（8 到 27）。此外，*ICE* 根据我们的预言机结果（图 5b），在整个距离范围内都表现出强劲性能，展现了其生成腺相关病毒可存活且多样的突变体的潜力。



(a) (b)

图 5：我们绘制了与野生型相距给定莱文斯坦距离的序列比例（图 5a）。图 5b展示了生成序列中优于野生型的序列比例（根据预言机判定）随莱文斯坦距离变化的情况，体现了*ICE* 生成多样且可存活突变体的潜力。

C.2. 情感控制的其他结果

表 7 展示了编辑过程的一个示例，通过迭代方式提高输入评论的情感分数。除了表 1 中的结果外，我们还报告了 ICE 和 Genhance 的几个变体。对于 ICE，创建合成配对数据的掩码策略包括：使用伯努利分布（$p = 0.8$）在序列中采样一个位置作为掩码起点，然后通过从截断泊松分布中采样来选择掩码长度（以被掩码的词元数量计）。表 1 中呈现的结果对应于表 6 中的超大变体，其中 $\lambda = 6$，且最大跨度大小设置为 12。我们还报告了掩码策略的另外三个变体：小（$\lambda = 3$，最大值为 6）、中（$\lambda = 4$，最大值为 8）和大（$\lambda = 5$，最大值为 10）。我们观察到超大变体取得了最佳外推结果，并使用该掩码策略报告采样和迭代采样基线方法。我们还报告了 Genhance 的两个变体，其中改变了为每个示例生成的输出序列总数。随着我们增加 n，模型在外推方面的表现按预期有所提升，但我们发现，直接可比的变体 $n = 50$的表现不如 ICE。

C.3. 生成超参数敏感性

为研究生成超参数与推理时迭代次数之间的相互作用，我们分别进行了束宽可变的无评分器推理，以及改变 k top-k 的*ACE2* 任务评分器引导推理。在所有情况下，我们都生成了1000 个突变。我们在表2、5和10 次迭代时的结果见表8。表中每个单元格表示 ddG 值低于相应目标的突变比例，数值四舍五入至小数点后三位。对应于 top-$k = 5$和束宽为5 、迭代次数为10 的行已列于表2。

总体而言，我们发现，推理过程结束时（第 10 次迭代）的结果对于这些超参数基本保持稳定。具体来说，当增大用于 top-k 采样的 k 时，性能略有下降，这可能是由于蛋白质序列的词汇表较小（总共 20 个）。类似地，对于无评分器推理，当我们将束宽减小到 3 时，我们

目标情感得分	训练区域			外推区域		
	3.5	**2.5**	**平均值**	**4.5**	**1.5**	**平均值**
分数条件基线	0.780	0.766	0.773	0.212	0.217	0.215
PPLM	0.534	0.516	0.522	0.081	0.065	0.077
采样	0.362	0.259	0.310	0.061	0.050	0.056
迭代采样	0.668	0.657	0.663	0.320	0.328	0.324
Genhance（n = 1）	0.407	0.167	0.287	0.063	0.025	0.044
Genhance（n = 50）	0.982	0.833	0.908	0.482	0.291	0.387
Genhance（n = 100）	**0.995**	0.912	**0.954**	**0.670**	0.429	0.550
ICE 搭配评分器 – 小	0.962	0.98	0.971	0.514	0.344	0.429
中	0.945	0.870	0.908	0.636	0.499	0.567
大	0.953	0.884	0.918	0.649	0.555	0.602
超大	0.943	0.900	0.921	0.638	**0.582**	**0.610**
ICE 无评分器	0.976	**0.918**	0.947	0.446	0.305	0.376

表 6：训练区域和外推区域中的情感控制结果，包括对我们模型和 Genhance 的消融实验。评估方法是测量由预言评分器确定的情感值高于（或低于）目标分数的样例所占比例。粗体数值表示每个目标对应的最高成功率。*ICE* 实现了最高的外推率。

在训练区域内获得略好的性能，但外推时性能略有下降。将束宽增加到 10 会轻微降低性能。

我们发现，迭代次数是外推性能的可靠指标；在每次特定迭代中，top-k 和束宽超参数对性能的影响很小。在第 2 次迭代中，由评分器引导时，较高的 top-k 值会带来更好的性能，因为模型会采样更多样化的生成序列，而评分器能够可靠地选出优质序列，从而提升模型在训练区域内目标上的表现。同样，对于无评分器推理，增大束宽也能提升模型在训练区域内目标上的表现。不过，当我们将迭代次数增加到 5 次和 10 次时，这种影响基本上会逐渐消失。

C.4. 停止条件

由于任务具有外推性质，因此很难可靠地确定生成模型何时达到目标分数。具体而言，如果我们能够知道生成器模型何时在外推区域内达到目标分数，那么就可以直接利用这一监督信号来训练生成器本身。一种选择是使用评分器，f_s。然而，我们观察到，f_s 的输出在训练区域边界附近趋于平稳，这限制了其作为外推区域内停止条件的可靠性。为说明这一点，我们在表 *ACE2* 任务中展示了平均输出分数随10 次迭代的变化情况，见表9。我们观察到，迭代次数超过7次后，输出分数基本保持不变。因此，我们将迭代次数设定为高于评分器达到平台期时的迭代次数，例如10次；结果发现，无需进一步调优，这一设置在我们的3个数据集上均表现良好。不过，我们也承认，在这一设定中，寻找更有原则性的停止条件仍是一个开放问题。

迭代	Text
1	<inc>这家店只能说还行。等候时间实在太长，柜台工作人员让人烦躁，而且 那些无人看管的孩子吵闹扰人。补甲和美甲艺术还不错，但不值得等这么久。Lisa 手脚很快，而且我很快就吹干了指甲。我在 Pinterest 上找到了想要的美甲艺术款式，她模仿得还不错。她还帮我给手机充了电，挺贴心的。如果你要去，尽量早点到，因为晚些时候会变得拥挤 这一天。虽然算不上惊艳，但值得去看看。
2	<inc>这家店充其量只能算一般。等候时间太长，柜台工作人员很烦人，而且 无人看管的孩子是个问题。补甲和美甲艺术还不错，但没什么特别之处。Lisa 动作很快，而且我很快就干了。我在 Pinterest 上找到了想要的美甲艺术，她很好地照着做了出来。她还帮我给手机充了电，这点很不错。如果你要去，尽量早点到，因为随着时间推移，店里会越来越忙。值得去一次，但别期待有什么惊艳之处。
3	<inc>这地方还不错，但也有一些缺点。等待时间很长，柜台工作人员太过强势，而且 无人看管的孩子很让人分心。不过补甲和美甲艺术真的很棒，Lisa 的速度也非常快。 我在 Pinterest 上找到了想要的美甲艺术，她完美地复刻了出来。给她额外加分，因为她给我的手机充电了。如果你打算去，尽量早点到，因为当天晚些时候队伍会变得更长。尽管 有这些缺点, 还是值得一去的。
4	<inc>这家店不错，但也算不上特别好。等候时间有点长，前台接待态度咄咄逼人，而且 没人看管的孩子很烦人。不过，补甲和美甲艺术都非常棒，还是值得一去的。Lisa 动作很快，我很快就吹干了指甲。我在 Pinterest 上找到了想要的美甲艺术款式，她完美地照着做了出来。她还帮我给手机充了电，这点很贴心。如果你打算去，最好早点去，因为越接近打烊时间人就越多。
5	<inc>这家店相当不错！等候时间长、前台接待态度咄咄逼人，还有没人看管的孩子，这些都很容易被抛到脑后，因为我离开时做着有史以来最棒的补甲和美甲艺术！Lisa 动作快如闪电，我不到 15 分钟就几分钟！我在 Pinterest 上找到了心仪的美甲艺术，她也完美地照着做了！更棒的是，她还帮我给 手机充了电，真是贴心！如果你打算去，最好早点到，因为越接近打烊，队伍就越长 不过，相信我，这份等待是值得的！
6	<inc>这家店很棒！等待、柜台人员咄咄逼人的态度，以及无人看管的孩子，都很容易让人忽略 等我终于离开时，已经做出了我有史以来最棒的补甲和美甲艺术！Lisa 动作特别快，我坐上她的椅子后不到 15 分钟就开始烘干了！我找到了心仪的美甲艺术（她照着做得简直一模一样，在 Pinterest 上.....但我刚坐下，手机就没电了。她拿出充电器，帮我给手机充了电！还有谁做过这种事？没有。提醒一下，如果可以的话，早点去；越接近打烊，排队的人就越多。:) 不过等待绝对值得！！

表 7：使用*ICE*改善与评论相关的情感的变化轨迹。

目标*ddG* 值			训练区域		外推区域		
			-1	-2.5	-5	-6	-7
使用评分器的 ICE：采样时改变 K	迭代 = 10	TopK = 15	0.997	0.964	0.249	0.083	0.01
		TopK = 10	0.998	0.966	0.283	0.091	0.016
		TopK = 5	**0.998**	**0.974**	**0.362**	**0.098**	**0.019**
	迭代 = 5	TopK = 15	0.982	0.648	0.041	0.004	0.000
		TopK = 10	0.981	0.646	0.040	0.004	0.000
		TopK = 5	0.978	0.647	0.042	0.005	0.001
	迭代 = 2	TopK = 15	0.711	0.093	0.002	0.000	0.000
		TopK = 10	0.703	0.090	0.001	0.000	0.000
		TopK = 5	0.674	0.086	0.001	0.000	0.000
无评分器 ICE：改变束宽	迭代 = 10	束宽 = 10	0.930	0.572	0.059	0.013	0.000
		束宽 = 5	0.945	0.598	**0.062**	**0.017**	**0.002**
		束宽 = 3	**0.959**	**0.623**	0.060	0.016	0.000
	迭代 = 5	束宽 = 10	0.852	0.440	0.030	0.006	0.000
		束宽 = 5	0.847	0.437	0.026	0.005	0.000
		束宽 = 3	0.844	0.419	0.023	0.004	0.000
	迭代 = 2	束宽 = 10	0.620	0.182	0.001	0.000	0.000
		束宽 = 5	0.567	0.155	0.001	0.000	0.000
		束宽 = 3	0.526	0.143	0.000	0.000	0.000

表 8：在 *ACE2* 任务上进行评估，以研究推理时生成超参数与迭代次数之间的相互作用。每个表格单元格表示 ddG 值低于相应目标的突变所占比例。对于评分器引导推理，我们改变 top-k 采样中的 k；对于无评分器推理，我们改变束宽。我们发现，在推理结束时（即第 10 次迭代），结果对于这些超参数总体上较为稳定。在推理初期（即第 2 次迭代），较高的 top-k 值和束宽分别会带来更好的性能，但到第 5 次和第 10 次迭代时，这种差异基本消失。

迭代	平均分数
1	-0.673
2	-2.051
3	-2.879
4	-3.272
5	-3.446
6	-3.522
7	-3.551
8	-3.558
9	-3.555
10	-3.567

表 9：的平均输出分数 随迭代次数的变化情况，适用于 *ACE2*任务。每个单元格是对表 10,000 个通过评分器引导推理生成的突变体所获分数的平均值，见表 2。我们观察到，f_s 的输出在训练区域边界附近、约 -3.5 时趋于平稳，因此不能可靠地作为生成过程的停止条件。