

文献翻译请求




我将会给你一篇英文文献，请你将原文翻译成中文，并且给出小学生也能听懂的解释



好的，请将文献提供给我。我会先将内容翻译成中文，并再给出适合小学生理解的解释。



 已上传的图片



中文翻译：


摘要：近年来，Transformer架构已成为机器学习算法的标准，应用于自然语言处理和计算机视觉领域。尽管这种架构在机器人学习中成功应用的证据显著，我们认为传统的Transformer并未完全利用机器人学习问题的结构。因此，我们提出了“Body Transformer”（BoT）架构，它通过提供引导学习过程的诱导偏差来更好地利用机器人的实体结构。我们将机器人身体表示为一个由传感器和执行器组成的图结构，并依赖遮掩注意力来在整个架构中汇集信息。结果表明，该架构在任务完成、扩展性和计算效率方面都优于传统的Transformer和经典的多层感知器，尤其是在模仿学习和强化学习策略的表示中。更多资料，包括开源代码，已在以下网址提供：https://sferrazza.cc/bot_site.

关键词： 机器人学习、图神经网络、模仿学习、强化学习

小学生能理解的解释：

近年来，科学家们发明了一种叫“Transformer”的技术，能让机器在理解文字和图片方面非常厉害。比如，这种技术可以让机器人“学会”做事情。但是，传统的Transformer并没有充分利用机器人的身体结构。因此，科学家们设计了一种叫“Body Transformer”（身体转换器）的新技术，能让机器人更好地使用自己的“身体”去学习和完成任务。他们用传感器（就像机器人的眼睛和耳朵）和执行器（像是机器人的手和脚）来构建机器人的“身体图”。这就让机器人在完成任务时更聪明、更高效。



 已上传的图片



中文翻译：

图1：Body Transformer (BoT) 是一种将物理代理（如机器人）视为由传感器和执行器构成的图结构的架构，在图结构中，边反映了机器人体的结构。BoT利用遮掩注意力作为一种简单且灵活的机制，为策略提供身体结构的诱导偏差。该图展示了BoT架构的整体示意图，以Unitree A1机器人为例。

小学生能理解的解释：

在这张图中，科学家们设计了一种叫“Body Transformer”的技术，它让机器人能够更聪明地使用自己的身体。你可以把机器人想象成一个拥有不同“关节”和“传感器”（像是眼睛和耳朵）的网络。这个技术把机器人的每个“关节”或“身体部位”当作一个小节点，通过这些节点连接起来形成“机器人图”。

这个“机器人图”帮助机器人更好地感知周围环境（左边的部分），然后把这些信息通过一个“BoT”处理中心（图的中间部分）进行分析，最后决定每个“关节”需要做什么动作（右边的部分）。就像人类大脑把信息汇总后告诉身体各个部分该怎么行动一样，这种方法可以让机器人做事情更聪明和更精准。



已上传的图片
已上传的图片



中文翻译：

1 引言

对于大多数的纠正和稳定动作，物理代理表现出与它们所感知到的外部刺激位置在空间上相关的运动反应。例如，对于冲浪者来说，下半身（即脚和脚踝）主要负责抵消在冲浪板下由波浪引起的失衡。实际上，人类在脊髓的神经回路中存在反馈循环，专门负责单个执行器的反应。

纠正性的局部动作是实现有效运动的关键因素。这对机器人尤其重要，因为机器人学习架构通常不会利用传感器和执行器之间的空间关系。事实上，机器人策略大多依赖于用于自然语言或计算机视觉的相同架构，而未能有效利用机器人身体的结构。

本研究专注于Transformer策略，因其在处理长序列依赖关系和无缝吸收大量数据方面展现出良好前景。Transformer架构已被开发用于非结构化的自然语言处理（NLP）任务，例如语言翻译，其中输入序列通常映射到重排的输出序列。相比之下，我们提出了Body Transformer (BoT)，它是一种增强Transformer的注意力机制的新架构，通过考虑机器人身体中传感器和执行器的空间分布来改进Transformer。

BoT将机器人的身体建模为一个由传感器和执行器构成的图结构。在BoT中，使用了高度稀疏的注意力层遮掩，防止每个节点关注超出其直接邻居的节点。连接多个具有相同结构的BoT层使得信息能够在整个图中汇聚，从而不削弱架构的表达能力。

我们的贡献如下：

- 我们提出了BoT架构，增强了Transformer架构，并使用一种利用机器人身体形态的新遮掩技术。
- 我们在模仿学习设置中引入了这一新架构，表明BoT提供的诱导偏差可实现更稳定的性能和泛化能力，并展现更强的扩展属性。
- 我们展示了BoT在在线强化学习（RL）中比多层感知器（MLP）和传统Transformer基准表现更优。

- 我们分析了BoT的计算优势，通过在注意力运算中重新构建缩放点积，运行时间和浮点运算量（FLOPs）减少了接近200%。


小学生能理解的解释：

这段话主要讲的是一种叫“Body Transformer”的新技术，它可以让机器人更好地控制身体。例如，当我们站在冲浪板上保持平衡时，主要是我们的脚和脚踝在工作。科学家们希望机器人也能像人一样，通过不同的“部位”（传感器和控制器）来维持平衡和完成任务。

为了实现这一点，科学家们把机器人的身体分成不同的部分，每个部分就像机器人身体的一个“点”。这些“点”之间有连接，组成了一个像“网络”一样的结构。然后，他们使用一种特殊的技术，让机器人每个“点”只关注离它最近的“点”，而不会分散注意力到其他地方。这样，机器人可以更快地反应，也更稳定。

他们的研究表明，这种新方法比以前的方法更好，特别是在学习新任务和做出反应的时候。



 已上传的图片



中文翻译：

2 相关工作


机器人中的Transformer。最初为自然语言处理（NLP）应用而开发的Transformer架构，已成功应用于多个领域，例如计算机视觉和音频处理。多项研究表明，Transformer可用于表示机器人策略，并在该应用中展现了其核心优势，如可变的上下文长度、处理长序列和多模态。然而，这些方法仍然使用了最初为非结构化或网格状输入（如语言或图像）而开发的Transformer。在本研究中，我们通过适当调整Transformer的注意力机制来利用机器人的结构特征。

小学生能理解的解释：

一开始，科学家们发明Transformer是为了帮助电脑更好地理解语言。后来，他们发现这种技术还可以帮助电脑看懂图片和听懂声音。最近，有些科学家发现，这种技术也能帮助机器人学习怎么行动。不过，以前的Transformer主要是用来处理语言或图片的。

而在这个研究中，科学家们做了一些改变，让Transformer可以更好地适应机器人身体的结构。这样，机器人就能更聪明地利用自己的身体去完成任务啦！



 已上传的图片



中文翻译：

图神经网络（GNNs） 图神经网络（GNNs）是一类能够以图形式处理输入数据的学习架构。早期版本的GNN使用了沿图传递信息的显式消息传递方案，而最近的架构则主要采用基于注意力的方法。实际上，传统的Transformer由于其可变的上下文长度，本质上支持完全连接的图结构。然而，在图解释方面达到最先进性能通常需要对原始Transformer架构进行修改，例如通过学习到的图编码和注意力偏置来实现。Buterez等人的一项同时代研究，类似于Veličković等人的工作，利用了遮掩注意力层，使每个节点仅关注其邻居节点，并在这些层之间交错使用未遮掩的注意力层。


在本研究中，我们在策略学习的环境中利用了遮掩注意力，并额外提出了一种架构，其中每一层都只能关注其自身及其直接邻居，从而在各层之间自然地增长上下文。例如，前几层的输出主要使用了更局部的信息，相较于最后几层的输出。

小学生能理解的解释：

图神经网络（GNNs）是一种特殊的技术，可以让电脑理解像“网络”一样的复杂数据。就像我们可以通过朋友认识朋友一样，GNNs可以让每个“点”或“部分”传递信息给其他相邻的“点”。科学家们发现，GNNs不仅可以让机器人理解自己周围的情况，还可以通过注意力机制，让机器人更聪明地把注意力集中在离它最近的部分。

在这个研究中，科学家们让机器人只关注自己和周围的“朋友”（邻居），这样机器人可以一层一层地理解更多信息。



 已上传的图片



中文翻译：


利用身体结构进行策略学习。 图神经网络（GNNs）已被多项研究探索，用于获取在不同机器人形态下都有效的多任务强化学习（RL）策略。早期的研究集中在消息传递算法上，随后，传统的Transformer以及利用学习编码和注意力偏置的基于Transformer的GNNs在性能上超越了这些算法。然而，所有这些方法仅在模拟的基准场景中进行了验证，并未应用于真实的机器人环境中。与之前的研究相比，我们额外展示了在注意力遮掩中引入瓶颈如何能够充分利用机器人的身体结构，并且在单个机器人执行任务时也能提高策略学习的性能，带来更好的表现和更理想的扩展性。

小学生能理解的解释：

科学家们发现了一种方法，可以让机器人在学习如何完成不同任务时，更好地利用自己的“身体结构”。就像我们用不同的手法去踢球、打篮球或者跳舞，机器人也需要不同的“方法”来完成不同的任务。过去，科学家们研究了让机器人“传递信息”的方法，但新的方法，比如Transformer技术，让机器人变得更聪明。

不过，以前的这些方法大多是在电脑模拟的“虚拟世界”中进行测试的，而不是在真实世界中。现在，科学家们提出了一种新方法，通过让机器人更好地理解自己的身体结构，让它们在真实世界中表现得更好、学得更快。这样，即使是一个机器人自己完成任务，也能做得更出色。



 已上传的图片


中文翻译：

3 背景

3.1 Transformer中的注意力机制

Transformer是一种现代机器学习应用的基础架构，在我们的研究中也有应用。它的核心是自注意力机制。自注意力会根据计算得出的分数为序列中的每个元素赋予不同的权重，这些分数是通过序列中键和值的配对计算得到的。因此，Transformer可以在输出结果中识别序列元素的相关配对关系。

具体而言，自注意力的输出向量是通过以下矩阵运算计算得到的：


$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

其中， Q 、 K 、 V （分别表示查询、键和价值矩阵）是序列元素嵌入向量的可学习线性映射， d_k 是嵌入空间的维度。当嵌入配对具有较高的相关性时，其得分较高（通过 QK^T 的计算），与该键相关联的值向量在注意力机制输出中将获得较高的权重。

小学生能理解的解释：

Transformer是一种帮助电脑“记住重要信息”的方法。在这里，它通过“自注意力”来实现这个目标。想象一下，我们在读一篇故事时，会把最重要的词标记出来，让我们更容易理解整个故事。自注意力就像是这种标记，让电脑知道哪些部分更重要。

这个方法会给每个单词一个“得分”，得分越高的部分就越重要。然后，电脑会根据这些分数来决定输出结果，把重要的信息突出来，帮助它更好地理解整段内容。


 已上传的图片


中文翻译：

3.2 基于Transformer的图神经网络（GNNs）

在本研究中，我们将代理的实体建模为一个图结构，其中的节点表示传感器和执行器，节点之间的连接边反映了机器人的身体形态。尽管消息传递型GNN适合这种模型，但它们容易出现“过平滑”和“过压缩”的问题，导致表示效果较差，从而阻碍了有效的远程交互，并限制了网络的深度。

最近，自注意力被提出作为消息传递的一种替代方法。标准的自注意力机制相当于对一个完全连接的图进行建模，而一种流行的基于Transformer的GNN，称为Graphormer，通过图结构的位置信息（如节点-边信息）向自注意力机制中引入偏置，并在缩放点积中进行调整，即：

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} + B \right) V$$


其中， B 是一个可学习的矩阵，依赖于图的特征，例如最短路径或邻接矩阵，可以有效地灵活编码图结构。

小学生能理解的解释：

在这个研究中，科学家们把机器人的“身体”画成一个由“点”和“线”组成的图，每个点表示机器人的一个部分（比如传感器和马达），线表示这些部分之间的连接方式。虽然以前的图方法可以用来传递信息，但它们有时会让信息变得过于模糊，导致机器人不能很好地理解远处的情况。

为了改善这种情况，科学家们使用了“自注意力”这种新方法。这样，机器人可以更灵活地理解整个图形的结构，特别是那些和它有直接连接的部分。



 已上传的图片



中文翻译：

3.3 遮掩注意力

注意力机制可以通过一个二进制遮掩矩阵 $M \in \{0, 1\}^{n \times n}$ （其中 n 为序列长度）进行修改，这相当于替换公式(1)中的 B 的元素：

$$B_{i,j} = \begin{cases} 0 & M_{i,j} = 1 \\ -\infty & M_{i,j} = 0 \end{cases}$$


其中 i 和 j 表示行和列的索引。此操作实际上将遮掩矩阵 M 中零位置的配对项的贡献置为零，从而在计算注意力时忽略这些配对。

小学生能理解的解释：

在这里，科学家们使用了一种叫“遮掩”的方法，让机器人可以只关注它需要的部分。想象一下，当你在看一张照片时，你可以用手遮住不需要看的地方，集中精力只看重要的部分。

在这个方法中，科学家们用一个叫“遮掩矩阵”的工具来帮助机器人遮住它不需要关注的部分，让它把注意力集中在关键的地方，这样它能更快、更有效地理解信息。



 已上传的图片



中文翻译：

4 Body Transformer


使用传统Transformer架构作为基础的机器人学习策略通常会忽视由机器人实体结构提供的有用信息。相比之下，在这里我们利用这种结构为Transformer提供更强的诱导偏差，同时保留了原始架构的表示能力。我们提出了“Body Transformer”（BoT），它基于遮掩注意力机制。在这种架构的每一层中，一个节点只能关注来自自身及其直接邻居的信息。结果，信息按照图结构流动，上游层基于局部信息进行推理，而下游层则从更远的节点汇集更多的全局信息。

小学生能理解的解释：

科学家们设计了一个叫“Body Transformer”的方法，让机器人更聪明地利用自己的身体。一般来说，机器人使用的普通技术可能会忽视它身体的结构信息，就像我们做事情时不注意自己身体的姿势一样。而Body Transformer则通过一种方法，让机器人在学习时关注自己附近的重要信息，再逐步把这些信息传递给其他部分。

这样，机器人可以从“邻居”那里获得最直接的帮助，然后让“远处的朋友”帮忙分析更全面的信息。这种方法让机器人更好地利用它的“身体网络”，更聪明地完成任



 已上传的图片



中文翻译：

图2：实体遮掩的构建


遮掩矩阵 M 是通过在实体图的邻接矩阵上添加一条对角线的“1”来构建的。在这里，我们展示了一个具有 $n = 10$ 个节点的代理实体的遮掩矩阵 M 的简单示例。

小学生能理解的解释：

在这个图中，科学家们设计了一种方法来帮助机器人只关注它需要的部分。你可以把这个想象成一张网络图，每个点代表机器人身体的一个部分，每条线代表它们之间的连接方式。遮掩矩阵就像是一张“指南图”，告诉机器人哪些部分是重要的，让它专注在这些部分上。

在这个例子中，他们用了包含10个点的小图，每个点之间有不同的连接方式。科学家们在图的每个点上加了一个特别的标记“1”，用来告诉机器人该注意哪些部分，这样它在行动时会更有效率。



 已上传的图片


中文翻译：

我们在此介绍BoT架构的各个组件（参见图1）：(1)一个Tokenizer，它将感官输入投射到相应的节点嵌入中，(2)一个Transformer编码器，用于处理输入嵌入并生成相同维度的输出特征，(3)一个Detokenizer，用于将特征解码为动作（或值，用于RL（强化学习）评论家的训练）。

Tokenizer。我们将观测向量映射到局部观测的图中。在实践中，我们将全局量分配到身体的根部元素，将局部量分配到代表相应肢体的节点上，类似于先前的GNN方法。然后，线性层将局部状态向量投射到嵌入向量中。每个节点的状态被输入到它的特定节点可学习线性投射中，结果形成一个长度为 n 的嵌入序列，其中 n 代表节点数目（或序列长度）。这与现有工作不同，后者使用单一的共享可学习线性投射来处理多任务RL中不同数量的节点。

BoT编码器。我们使用具有多个层的标准Transformer编码器作为骨干，并提出了我们架构的两种变体：

- **BoT-Hard**：在每层中使用反映图结构的二进制遮掩 M 。具体而言，我们将遮掩构建为 $M = I_n + A$ ，其中 I_n 是维度为 n 的单位矩阵， A 是对应于图的邻接矩阵（参见图2中的示例）。具体来说，这允许每个节点关注自身及其直接邻居，在问题中引入了大量的稀疏性，从计算的角度来看非常有吸引力，正如5.4节所述。
- **BoT-Mix**：交替使用带有遮掩注意力（与BoT-Hard中的构造相同）和不带遮掩注意力的层。这与Buterez等人的并行研究类似，不同之处在于：(I) 我们发现，在我们的实验环境中，以遮掩注意力层作为第一层更有效；(II) 我们的遮掩矩阵 M 不等于邻接矩阵，允许每个节点在架构的每一层中进一步关注自身。


小学生能理解的解释：

科学家们给机器人设计了一个叫BoT的系统，让它可以更好地处理和使用身体各部分的信息。这个系统有三个主要部分：

1. **Tokenizer**：它就像一个翻译器，帮助机器人把传感器的“观察”转化成信息，分配到机器人的不同“部位”上。每个“部位”都有自己的小“片段”，这样机器人就能知道各部分的状态。
2. **BoT编码器**：这是BoT的“思考”部分。科学家们设计了两种不同的方法：
 - **BoT-Hard**：在每一层中，机器人只能“看到”自己和紧挨着的“邻居”，这样可以让它更专注在周围的局部信息上。
 - **BoT-Mix**：在不同的层中，有时机器人只关注周围，有时它可以看到更远的地方，这样它可以一层一层地积累信息。

通过这样的设计，机器人可以更高效地使用自己身体的结构去完成任务。



 已上传的图片



中文翻译：


Detokenizer。来自Transformer编码器的输出特征被输入到线性层中，这些线性层将它们投射到与节点肢体相关的动作上，这些动作是根据相应执行器与肢体的接近程度来分配的。再次强调，这些可学习的线性投射层对每个节点都是独立的。当BoT在RL（强化学习）环境中作为评论架构使用时，如第5.2节中所述，Detokenizer输出的是值而不是动作，这些值随后在各个身体部分上进行平均。

小学生能理解的解释：

Detokenizer的作用是把机器人思考出来的信息转化为它的每个部分应该执行的动作。比如，机器人如果想移动一条“腿”，Detokenizer会根据“腿”上的小部件来决定如何移动。

在某些情况下，Detokenizer并不会直接告诉机器人要做什么动作，而是会告诉它一些参考的“分数”，然后让机器人每个部分自己决定平均一下要怎么行动。



 已上传的图片



中文翻译：

5 实验

我们在模仿学习和强化学习的环境中评估BoT的性能。保持图1中的结构不变，仅使用不同的基线架构替换BoT编码器，以单独评估编码器的效果。特别是，在本节列出的各种实验中，我们提供了以下基线和变体：(i) 一个将所有嵌入向量堆叠为输入的MLP，(ii) 一个未遮掩的传统Transformer编码器，(iii) 仅使用遮掩自注意力层的BoT-Hard，(iv) 在遮掩和未遮掩自注意力层之间交替的BoT-Mix。所有比较都在具有相似数量的可训练参数的模型间进行。

通过以下实验，我们旨在回答以下问题：

- 遮掩注意力是否有助于模仿学习的性能和泛化？
- BoT是否比传统Transformer架构展现出更好的扩展趋势？
- BoT是否兼容RL框架，什么设计选择能够最大化性能？
- BoT策略是否能应用于现实世界的机器人任务？
- 遮掩注意力的计算优势是什么？


小学生能理解的解释：

在这部分实验中，科学家们进行了很多测试，来看看BoT这个方法到底好不好用。为了找到答案，他们设计了几种不同的方法进行对比，就像我们做实验时测试不同的方案一样。

他们希望通过实验来回答这些问题：

1. 遮掩注意力（让机器人只关注重要的部分）是否让机器人学得更好？
2. BoT这个方法是否比传统方法在不同任务中表现得更好？
3. BoT是否适合用在强化学习的环境中？怎样设计才能让它表现更好？
4. BoT能否在真实世界中帮助机器人完成任务？
5. 遮掩注意力在计算方面有哪些好处？



 已上传的图片



中文翻译：

5.1 模仿学习实验

我们在MoCapAct数据集定义的身体跟踪任务中评估了BoT架构的模仿学习性能。该数据集包含了标注动作的人形状态轨迹，具有超过500万个转变，涵盖了总计835个跟踪片段。对于每种架构，我们训练了一个确定性的行为克隆（BC）策略。我们评估了平均回报，并对片段长度进行了归一化处理，同时归一化了剧集的长度，当跟踪误差超过阈值时，剧集终止。我们在训练片段和未见过的验证片段上都进行了评估。

我们在图3a中报告了结果，其中BoT在性能上始终优于MLP和Transformer基线。值得注意的是，在未见过的验证片段上，与这些架构的差距进一步增加，展示了由包含实体结构的诱导偏差提供的泛化能力。我们还报告了使用MoCapAct数据集的定制多片段策略的训练集性能。虽然多片段策略在基线模型中具有竞争力，但我们的架构明显优于它。这一结果尤为显著，因为比较条件对基线更为有利，该基线具有更灵活的随机策略，并且在更大的轨迹集上进行了优化。

如图3b所示，我们还发现BoT-Hard具有很强的扩展能力，因为随着可训练参数数量的增加，其性能在训练和验证片段上相比于Transformer基线持续提高。这进一步表明BoT-Hard不易过拟合训练数据，这是由实体结构偏差所导致的。在附录E中进行了额外的比较，包括在灵活操作基准上的实验，见图4。


小学生能理解的解释：

在这个实验中，科学家们用BoT系统来做一个“身体跟踪”任务，就是让机器人模仿人类的动作。他们用了—个叫MoCapAct的数据集，这个数据集有很多人—类动作的片段，就像机器人在看人类的“录像”—样，从中学习如何模仿动作。

他们发现，BoT系统的表现比其他方法（MLP和传统Transformer）要好，尤其是在那些机器人没见过的新动作片段中，它学得更快更准确。这说明BoT系统的设计让它更擅长学会不同的动作。

他们还发现，BoT-Hard方法可以处理更复杂的任务，因为当机器人有更多“训练”机会时，BoT的表现会更好，而且它不容易把自己局限在学过的东西上。



 已上传的图片


中文翻译：

(a) 各架构的训练和验证性能。各种架构-标准组合的统计数据显示了两个值，左边的值是训练期间记录的最大值，右边的值是带有标准差的平均评估分数。结果平均于五个随机种子。

(b) 在MoCapAct数据集上的训练和验证性能，随着可训练参数数量的变化。每个数据点代表五个种子平均的性能。我们在本论文的其他模仿学习实验中使用了17.5M模型。

图3：BoT在模仿学习中的表现。


图4：Adroit手部任务，包含开门、锤击和重定位任务（结果见附录）。


小学生能理解的解释：

在这个实验中，科学家们对比了几种方法，看看哪种方法能让机器人学得更好。上面的表格和图表显示的是不同方法在学习和验证过程中的成绩，红色的是传统方法，绿色的是BoT方法。

结果显示，BoT方法在训练时得分更高，而且随着训练中“参数”的增加，BoT方法的表现也越来越好。简单来说，BoT方法让机器人学得更快、更好。

图4展示了机器人在模仿学习中完成的一些任务，比如开门、用锤子和移动物体。


 已上传的图片

 已上传的图片


中文翻译：

5.2 强化学习实验

我们使用PPO算法在Isaac Gym中的4个机器人控制任务上评估了BoT和基线模型的强化学习性能。这4个任务分别是Humanoid-Mod、Humanoid-Board、Humanoid-Hill和A1-Walk。

所有的Humanoid任务都是基于Isaac Gym中的经典人形环境，我们修改了观察空间以增加分布式传感器信息的数量（详细信息见附录A），并包含所有肢体的接触力。Humanoid-Mod是经典的平地跑步任务，而在Humanoid-Hill中，我们将平地换成了不规则的丘陵地形。Humanoid-Board是一个新的任务，要求机器人在一个放置在圆柱体上的板上保持平衡。最后，我们适配了A1-Walk环境，该环境是Legged Gym中的一部分，任务是让Unitree A1四足机器人以固定速度行走。

图5展示了训练期间MLP、Transformer和BoT（Hard和Mix）的平均回报。实线对应平均值，阴影区域为五个种子的标准误差。结果显示，BoT-Mix在样本效率和渐近性能方面始终优于MLP和传统Transformer基线，突显了将身体诱导偏差整合到策略网络架构中的效果。


同时，BoT-Hard在简单任务（如A1-Walk和Humanoid-Mod）上优于传统Transformer，但在需要探索的复杂任务（如Humanoid-Board和Humanoid-Hill）上表现相对较差。由于遮掩注意力限制了远距离身体部位的信息传播，BoT-Hard在信息通信方面的强约束可能阻碍了高效的强化学习探索：在Humanoid-Board和Humanoid-Hill任务中，从脚趾到手指的信息传递可能对于环境的突然变化非常重要。而BoT-Mix在实体图中平衡了信息的传递，并在中间层实现全局汇集，以确保高效探索。相反，在A1-Walk或Humanoid-Mod中，环境状态变化较规律，因此BoT-Hard中的强身体诱导偏差可以有效减少搜索空间，加速学习。

小学生能理解的解释：

在这个实验中，科学家们让机器人完成一些任务，比如在平地上行走、在不平的地形上移动、在圆柱体上的平板上保持平衡等。他们测试了不同的学习方法，看看哪种方法能让机器人表现得更好。

结果显示，BoT-Mix方法最优秀，能让机器人更高效地学习。BoT-Hard在简单任务上表现不错，但在难度较高的任务中，由于信息的传递受到限制，它表现得不如BoT-Mix。在这些复杂任务中，机器人需要知道自己各个部位的状态，而BoT-Mix可以很好地帮助它传递和汇集这些信息，让它更聪明地应对变化的环境。



 已上传的图片



中文翻译：

5.3 现实世界实验

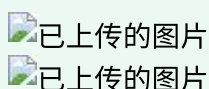
Isaac Gym模拟的运动环境广泛用于将强化学习策略从模拟转移到现实世界，而无需在现实中进行适配。为了验证我们的架构是否适合现实应用，例如在真实时间中的行走，我们将训练好的BoT策略部署到现实世界中的Unitree A1机器人上，并适配了Zhuang等人 and Wu等人提供的代码库。这在补充视频中展示，展示了我们架构在现实部署中的可行性。我们注意到，为了简化，我们没有使用教师-学生训练或记忆机制等方法，这些方法通常在运动文献中用于进一步改善转移效果，以获得更自然的步态。

小学生能理解的解释：

在这个实验中，科学家们测试了机器人在真实世界中的表现。之前他们都是在电脑模拟的环境中训练机器人，现在他们想看看这种训练方法能不能直接让机器人在现实中行走，比如在真实地面上跑步。

他们把之前训练好的BoT策略用在了真实的四足机器人Unitree A1上，并展示了机器人在现实中行走的效果。为了让测试更简单，他们没有使用一些复杂的训练方法（比如老师教学生的方法），而是直接用机器人自己学到的技能进行测试，看看它能不能自然地走路。





中文翻译：

5.4 计算分析

物理代理的身体部位之间的连接通常是稀疏的，因此在BoT的实体图中预先计算的遮掩矩阵 M 也是稀疏的。遮掩注意力机制可以利用这种稀疏性，因为通过忽略那些最终会被遮掩的矩阵元素的计算，可以减少计算成本。大型深度学习库（如PyTorch）通常对矩阵乘法和注意力操作（例如FlashAttention）进行了优化，但并未利用遮掩注意力机制中的可能稀疏性。为了公平的对比，我们使用基于CPU的NumPy重新实现了公式(1)中的缩放点积，并在单样本和单头的情况下进行评估，其批量和多头版本可以进一步在GPU上并行化。

我们自定义的实现包含两个主要改进：(1) 在计算遮掩注意力机制时，我们仅对未遮掩的元素执行点积运算，从而减少了因矩阵乘法引起的浮点运算量（FLOPs）；(2) 我们仅对未遮掩的值计算公式(1)中的softmax，这同样减少了浮点运算量。

我们在随机生成的 Q 、 K 、 V 和 M 上测量了各注意力机制实现的平均运行时间。对于每次随机化，生成的遮掩矩阵 M 的对角线为1，其稀疏性等于MoCapAct实验中使用的稀疏度。

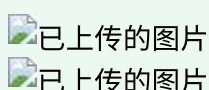
在图6中，我们展示了自定义实现相对于标准注意力机制在不同序列长度（节点数）下的扩展结果。我们观察到，在128个节点的情况下，速度提升可达206%。总体上，这表明BoT架构中由身体诱导的偏差不仅提升了物理代理的整体性能，还因架构所引入的稀疏遮掩而受益。通过适当的并行化，这一方法可能显著减少学习算法的训练时间。

更多实验细节和推导见附录H。

小学生能理解的解释：

在这一部分，科学家们研究了BoT方法的计算效率。他们发现，因为BoT方法让机器人只关注身体的关键部分，不用每个地方都计算一遍，所以它能更快地完成任务，就像我们只看重点而不看所有内容时阅读速度更快一样。

为了证明这个方法确实更快，他们还专门做了一些测试，结果显示BoT的速度比传统方法快了很多，最多可以快2倍左右。这意味着BoT不仅让机器人表现更好，还能让它学得更快，节省了很多计算时间。



中文翻译：

6 结论

在本研究中，我们提出了一种新的基于图的策略架构——Body Transformer (BoT)，它利用了机器人的身体结构作为诱导偏差。我们的实验表明，遮掩注意力（这是BoT的核心）有助于模仿学习和强化学习算法。此外，该架构具有良好的扩展性和计算特性，使其在高维系统中的应用更具吸引力。

在这里，我们使用Transformer来处理来自同一时间点的分布式传感器信息序列。然而，Transformer还擅长处理跨时间的信息。我们将BoT扩展到时间维度的工作留待未来研究，以进一步提升机器人策略在现实世界中的应用，例如Unitree A1机器人上的演示。

我们方法的一个限制是当前的深度学习库未能充分利用其计算优势，希望本研究能激励未来在该方向上的发展。此外，为了确保架构在建模远程关系时不失去表现能力，我们的架构需要最少数量的Transformer层，这通常增加了所需的可训练参数量。

致谢

我们感谢Huang Huang和Antonio Loquercio在A1实验设置方面的帮助。本研究部分由SNSF博士后流动奖学金211086、ONR MURI N00014-22-1-2773、BAIR工业联盟、NSF AI4OPT AI中心、NSF行动AI中心奖IIS-2229876、ONR DURIP资助。我们还感谢NVIDIA通过NVIDIA学术DGX资助提供的计算资源。

小学生能理解的解释：

在这项研究中，科学家们设计了一个叫BoT的新系统，让机器人能更好地利用自己的“身体结构”来学习。BoT使用了一种叫“遮掩注意力”的方法，可以帮助机器人更聪明地学习模仿和完成任务。这个系统不仅学得快，还可以很好地应对复杂的情况。

不过，目前有些电脑程序还不能完全利用BoT的优势，因此科学家们希望未来能有新的技术来帮助BoT更好地工作。另外，为了保证BoT能理解复杂的情况，科学家们需要给它多一些学习的“层”，这也会让它需要更多的计算资源。

他们还感谢了提供帮助和支持的人和机构。



对话结束



好的，感谢您的信任和使用！如果将来有任何问题或需要帮助，随时可以找我。祝您一切顺利，学习愉快！再见！😊

ChatGPT 也可能会犯错。请核查重要信息。