
科学家提出一种基于解纠缠的人脑对应表征研究框架

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/8070.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

语义和语法是自然语言中最重要的特征。已有神经影像学的研究证明大脑处理语义和语法的神经基础是不同的，但是目前关于自然语言中语义和语法是否在大脑中分开编码以及具体编码的脑区并没有一致的结论。由于大脑语言处理过程十分复杂，传统认知神经科学使用人工设计的语料来收集神经影像数据进而研究这个问题。由于人工设计的语料脱离了人类通常的语言理解环境，结论无法泛化到实际语言场景中，另外受控刺激会引入语言理解之外的因素，因而无法得出人脑在自然语言理解情况下语义和语法表征的结论。

针对上述问题，中国科学院自动化研究所自然语言处理组助理研究员王少楠、研究员张家俊、宗成庆和中国科学院心理研究所副研究员林楠合作提出了一种利用解纠缠的计算模型特征研究人脑对应表征的框架（图1），利用计算模型将句子中的语义和语法特征区分开，使用分离的语义和语法特征来研究大脑对应的神经基础。该框架适用于任何自然语言刺激实验中，用于研究人脑对应语义和语法的表征。相关成果即将发表在AAAI2020会议上。

该框架分为两个部分，首先构造解纠缠的特征表示模型，接着用分离的特征表示向量在大脑激活数据中寻找对应特征最相关的脑区。具体来说，团队提出一种解纠缠的特征表示模型（DFRM，图2），利用两个隐含变量分别表示语义特征变量和语法特征变量。使用词向量平均编码器来抽取句子中的语义特征，利用长短时记忆网络（LSTM）来抽取句子中的语法特征，通过令语义变量学习区分两个句子是否含义相同，令语法变量学习区分两个句子是否语法相同的目标函数，使语义变量积累语义信息，语法变量积累语法信息。DFRM模型通过上述目标函数以及变分自编码器的重构误差学习语义和语法变量的分布形式，最终DFRM模型可以分别用语义变量和语法变量为每个句子生成对应的语义向量和语法向量。

利用上述语义向量和语法向量，通过神经编码技术，可以研究大脑对应的语义和语法表征。具体来说，通过上述计算模型可以得到每个句子的语义向量和语法向量，利用所有句子中每两个句子求余弦相似性可以得到句子之间的语义相似度矩阵和语法相似度矩阵。同样的，对于大脑来说，每个脑区对每个句子都有一系列激活数值，对每两个句子间的激活向量求余弦相似性可以得到每个脑区的句子理解相似度矩阵。最后利用计算模型得到的语义相似性矩阵和语法相似性矩阵分别与每个脑区的相似性矩阵做相关性分析，即可以得到与语义最相关的脑区和与语法最相关的脑区。

实验结果表明，解纠缠特征表示模型（DFRM）可以最大限度地区分句子中语义和语法信息，在语义和语法相似性数据集上取得了最好的结果。实验从计算的角度为大脑语义和语法表征机制提

供了新的证据，证实并扩展了已有认知神经科学的发现，表明了利用自然语言处理模型有助于研究大脑语言理解机理。

图1 基于解纠缠特征的人脑表征研究的计算框架

图2 解纠缠特征表示模型

研究团队单位：自动化研究所

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](http://www.iikx.com)转发