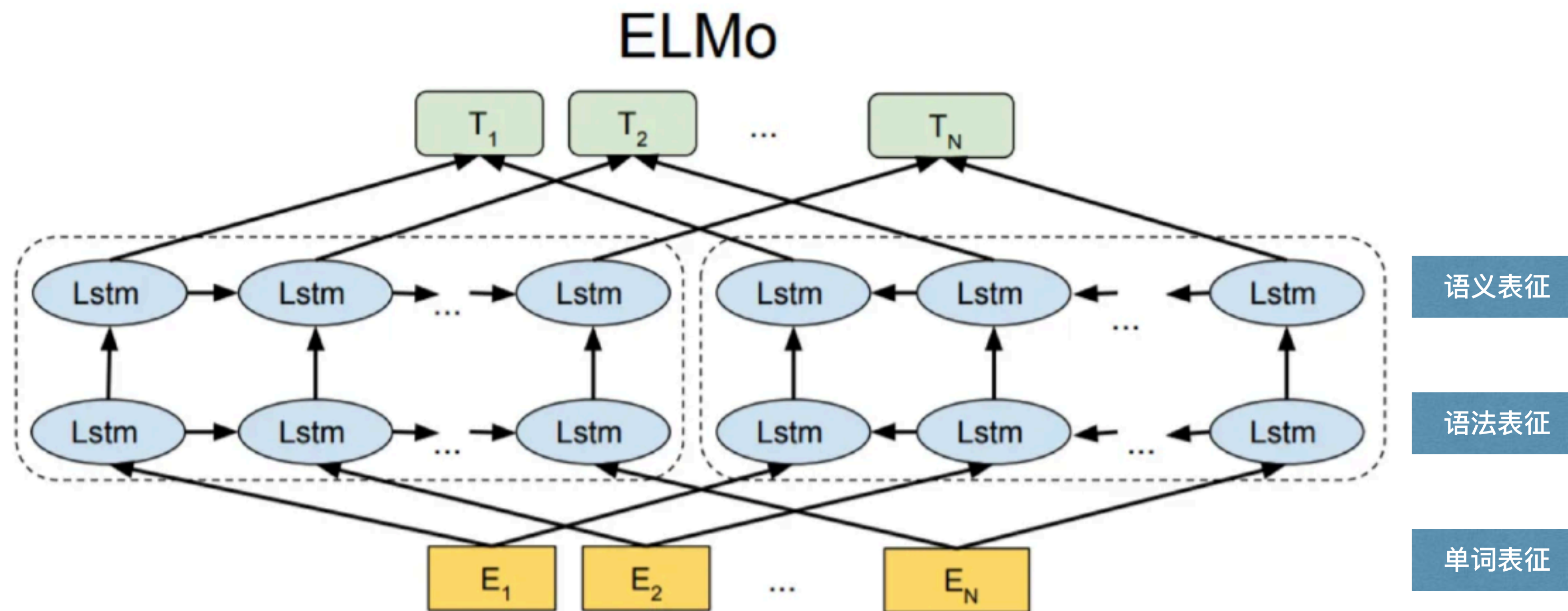


ELMo(Embeddings from Language Models)



使用双向LSTM将要预测的词汇上下构建起来

整个网络分三层：

Word-Embedding Layer

Contextualized Layer

Output Layer

其中在Contextualized Layer中，以被预测词汇

为中心，分为：

Context-Before

Context-After

经过网络处理后的每个单词都会得到三个

Embedding

最底层的单词Embedding

上一层的语法Embedding

再上一层的语义Embedding

<https://arxiv.org/pdf/1802.05365.pdf>

Attention



视觉注意力机制是人类视觉所特有的大脑信号处理机制。人类视觉通过快速扫描全局图像，获得需要重点关注的目标区域，也就是一般所说的注意力焦点，而后对这一区域投入更多注意力资源，以获取更多所需要关注目标的细节信息，而抑制其他无用信息。

注意力机制本身是一种算力优化方式，引入机器学习后，就有了Attention机制