论文阅读笔记

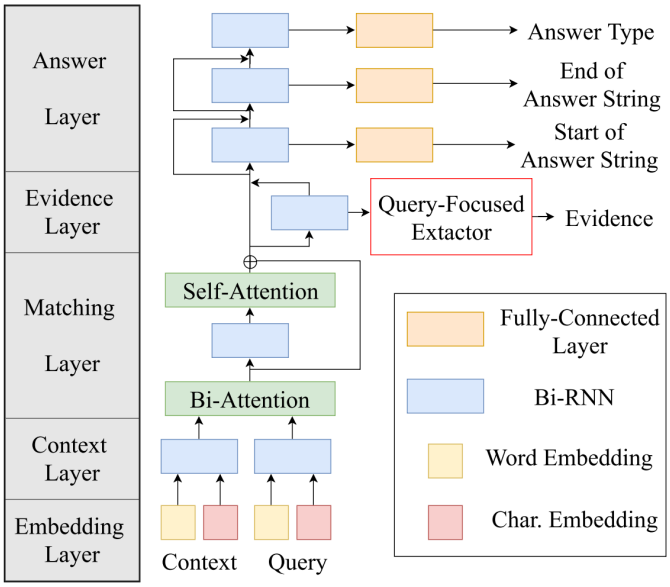
Step7

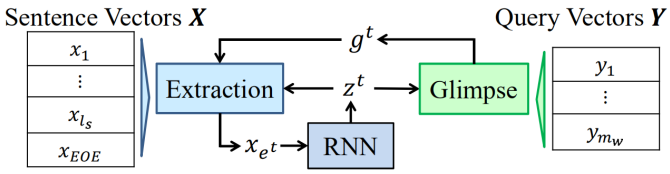
MF1833063, 史鹏, spwannasing@gmail.com

2019 年 8 月 4 日

1 Answering while Summarizing:Multi-task Learning for Multi-hop QA with Evidence Extraction

本文基于HotpotQA，提出了一个叫Query Focused Extractor (QFE)的模型，基于 extractive summarization模型，另外相比已有的模型，克服了预测evidence sentence之间相互独立的问题（通过使用RNN），以及引入了Multi-Task，可以和任意的RC模型结合。亮点是即使QFE和一个简单的baseline RC模型结合起来，也能达到evidence extraction score上的SOTA效果。





$$z^t = \mathrm{RNN}\left(z^{t-1}, x_{e^t}\right) \in \mathbb{R}^{2d_c} \tag{1.1}$$

$$\Pr\left(i; E^{t-1}\right) = \mathrm{softmax}_i\left(u_i^t\right) \tag{1.2}$$

$$u_i^t = \begin{cases} v_p^\top \tanh\left(W_{p1}x_i + W_{p2}g^t + W_{p3}z^t\right) & (i \notin E^{t-1}) \\ -\infty & (\text{ otherwise }) \end{cases} \tag{1.3}$$

$$\begin{aligned} g^t &= \sum_j \alpha_j^t W_{g1} 1_j \in \mathbb{R}^{2d_c} \\ \alpha^t &= \mathrm{softmax}\left(a^t\right) \in \mathbb{R}^{m_w} \\ a_j^t &= v_g^\top \tanh\left(W_{g1}y_j + W_{g2}z^t\right) \end{aligned} \tag{1.4}$$

$$L_E = -\sum_{t=1}^{|E|} \log\left(\max_{i \in E \setminus E^{t-1}} \Pr\left(i; E^{t-1}\right)\right) + \sum_i \min\left(c_i^t, \alpha_i^t\right) \tag{1.5}$$

$$L = L_A + L_E \tag{1.6}$$

2 Do you know that Florence is packed with visitors?Evaluating state-of-the-art models of speaker commitment

这篇论文展示出了带有语言学知识的模型的巨大潜力

对基于规则的和双向LSTM这两种最先进的说话人承诺模型进行了系统的评价

论文中的语言学分析给人启发，也展现出了系统的优势和劣势

当一个人，比如 Mary，问你「你知不知道佛罗伦萨全都是游客？」，我们会认为她相信佛罗伦萨全都是游客；但如果她问「你觉得佛罗伦萨游客多吗？」，我们就不会这样认为。推断说话人承诺（或者说事件真实度）是问答和信息提取任务中的关键部分。在这篇论文中，作者们探索了这样一个假说：语言学信息的缺乏会影响说话人承诺模型中的错误模式。他们的验证方式是在一个有挑战性的自然语言数据集上分析模型错误的语言学关联性。作者们在 CommitmentBank 这个由自然英语对话组成的数据集上评价了两个目前最好的说话人承诺模型。CommitmentBank 数据集已经经过了说话人承诺标注，方式是在 4 种取消蕴含的环境中向着时态嵌入动词（比如知道、认为）的补充内容进行标注。作者们发现，一个带有语言学知识的模型能展现比基于 LSTM 的模型更好的表现，这表明如果想要在这样的有挑战性的自然语言数据中捕捉这些信息的话，语言学知识是必不可少的。对语言学特征的逐项分析展现出了不对称的错误模式：虽然模型能在某些状况下得到好的表现（比如否定式），但它很难泛化到更丰富的自然语言的语言学结构中（比如条件句式），这表明还有很大提升的空间。

(1)	**Context**	The answer is no, no no. Not now, not ever.
	Target	I *never* **believed** I could wish anyone dead but last night changed all that.
		Gold:1.56, Rule-based:3.0, Hybrid: 0.50
(2)	**Context**	Revenue is estimated at $18.6 million. The maker of document image processing equipment said the state procurement division had declared FileNet in default on its contract with the secretary of state uniform commercial code division.
	Target	FileNet said it *doesn't* **believe** the state has a valid basis of default and is reviewing its legal rights under the contract ,
		Gold: -0.47, Rule-based: 3.0, Hybrid: 1.08
		but said it can't predict the outcome of the dispute.
(3)	**Context**	**A**: Yeah, that's crazy. **B**: and then you come here in the Dallas area, um,
	Target	I *don't* **believe** that people should be allowed to carry guns in their vehicled .
		Gold: -2.64, Rule-based: 3.0, Hybrid: 1.40

3 Emotion-Cause Pair Extraction:A New Task to Emotion Analysis in Texts

情绪原因提取（Emotion cause extraction ， ECE）是一项旨在提取文本中某些情绪背后潜在原因的任务，近年来由于其广泛的应用而受到了很多关注。然而，它有两个缺点：1）情绪必须在ECE原因提取之前进行标注，这极大地限制了它在现实场景中的应用；2）先标注情绪然后提取原因的方式忽略了它们是相互指示的事实。

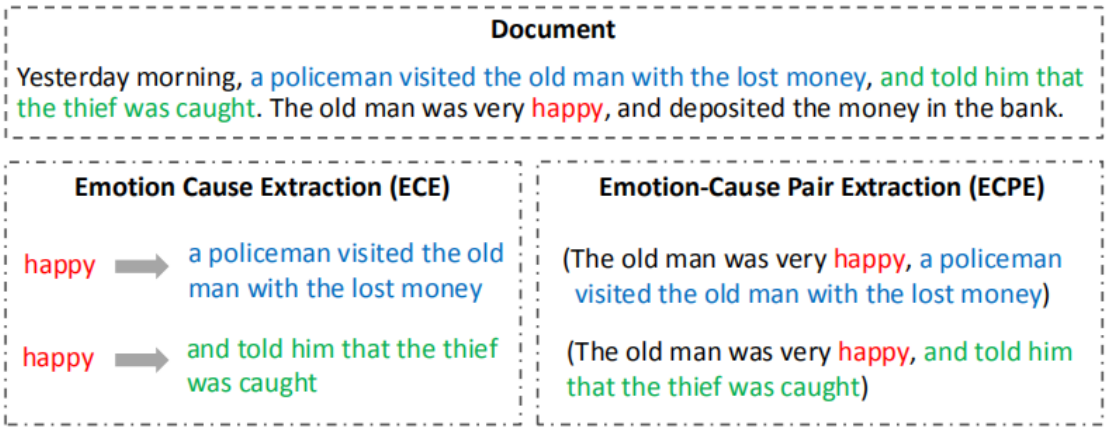
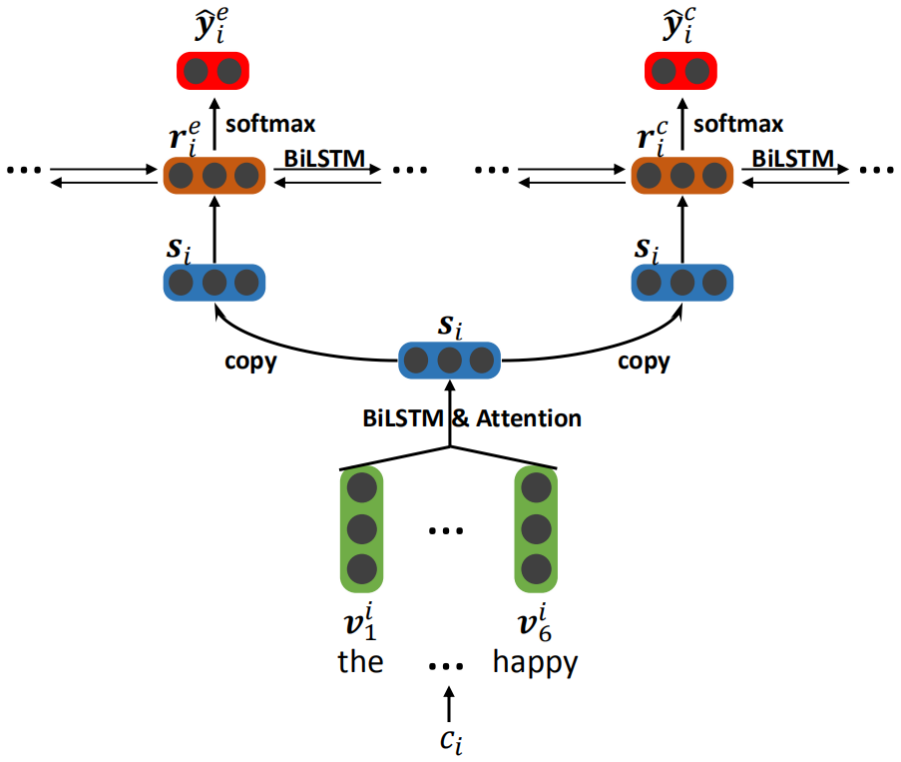


Figure 1: An example showing the difference between the ECE task and the ECPE task.



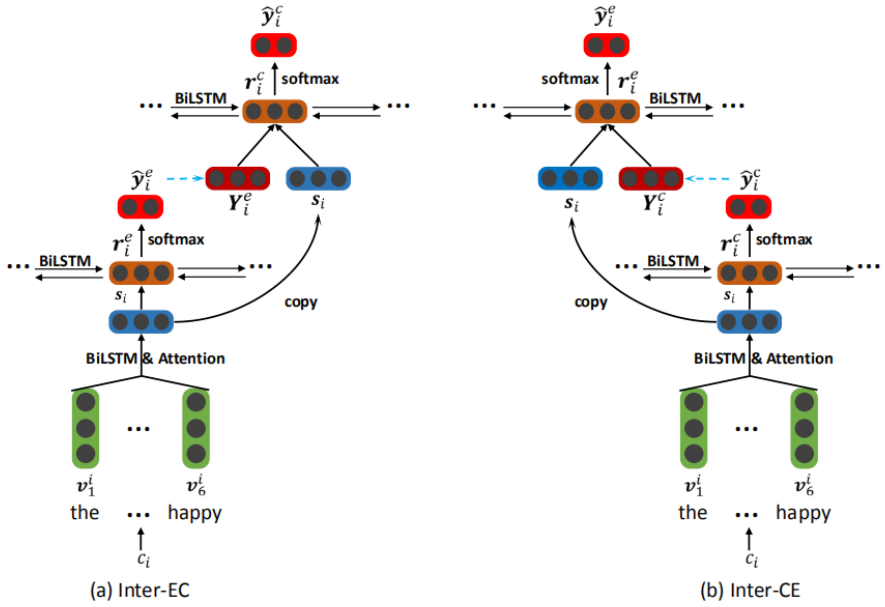


Figure 3: Two Models for Interactive Multi-task Learning: (a) Inter-EC, which uses emotion extraction to improve cause extraction (b) Inter-CE, which uses cause extraction to enhance emotion extraction.

这篇文章的创新点就是提出了一个新的任务 emotion-cause pair extraction (ECPE)，相较于传统的ECE不用依赖于事先的情感标注。其次使用2-step的结构，第一步提取出E和C（有两种结构，一种E和C之间相互独立，另一种是有联系的），第二步进行筛选。

总体算法很简单，只使用了Bi-LSTM和attention结构，核心是提出一个新的Task。