

结合边界预测和动态模板方法的槽填充模型

朱展标, 黄沛杰*, 张业兴, 刘树东, 张华林, 黄均曜

华南农业大学, 数学与信息学院, 广东广州, 510642

zhuzhanbiao@stu.scau.edu.cn, pjhuang@scau.edu.cn, yip_hing@stu.scau.edu.cn,
sudan@stu.scau.edu.cn, vol_chris@scau.edu.cn, wishyao@stu.scau.edu.cn

摘要

意图识别和槽信息填充的联合模型将口语理解技术(Spoken language understanding, SLU)提升到了一个新的水平, 但是目前研究进展的模型通过话语上下文信息判断位置信息, 缺少对槽信息标签之间位置信息的考虑, 导致模型在槽位提取过程中容易发生边界错误, 进而影响最终槽位提取表现。而且在槽信息提取任务中, 槽指称项(Slot mentions)可能与正常表述话语并没有区别, 特别是电影名字、歌曲名字等, 模型容易受到槽指称项话语的干扰, 因而无法在槽位提取中正确识别槽位边界。本文提出了一种面向口语理解的结合边界预测和动态模板的槽填充(Boundary-prediction and Dynamic-template Slot Filling, BDSF)模型。该模型提供了一种联合预测边界信息的辅助任务, 将位置信息引入到槽信息填充中, 同时利用动态模板机制对话语句式建模, 能够让模型聚焦于话语中的非槽指称项部分, 避免了模型被槽指称项干扰, 增强模型区分槽位边界的能力。在公共基准语料库CAIS和SMP-ECDT上的实验结果表明, 我们的模型优于比较模型, 特别是能够为槽标签预测模型提供准确的位置信息。

关键词: 口语对话理解; 槽信息填充; 位置信息; 模板机制

Slot Filling Model with Boundary Prediction and Dynamic Template

Zhanbiao Zhu, Peijie Huang*, Yexing Zhang, Shudong Liu,
Hualin Zhang, Junyao Huang

College of Mathematics and Informatics, South China Agricultural University, China
zhuzhanbiao@stu.scau.edu.cn, pjhuang@scau.edu.cn, yip_hing@stu.scau.edu.cn,
sudan@stu.scau.edu.cn, vol_chris@scau.edu.cn, wishyao@stu.scau.edu.cn

Abstract

The joint model of intention detection and slot filling promotes the spoken language understanding(SLU) to a new level. However, the existing models determine position information only by utterances context, ignoring the location information between slot information tags, which lead to boundary errors in slot filling. And in the slot filling task, the slot mentions may be similar to the normal expression utterances, such as the movie name, the song name and so on, thus the model is easy to be disturbed by the slot mentions. Therefore, the slot boundary can not be correctly identified in slot filling. This paper presents a slot filling model which combines boundary prediction and dynamic template (BDSF), providing an auxiliary task of joint prediction of boundary

*通讯作者

©2021 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

information, introducing location information into slot filling. Moreover, modeling the dialogue sentence pattern by dynamic template mechanism, the model can focus on the non-slot mentions utterances. It avoids the interference of the slot mentions of the model and enhances the ability of the model to distinguish the slot boundary. The experimental results on the public benchmark corpus CAIS and SMP-ECDT show that our model is better than the comparison models. Especially it can provide accurate position information for the slot label prediction model.

Keywords: Spoken language understanding , Slot filling , Position information , Template mechanism

1 引言

构建一个智能的人机对话系统是人工智能领域的长期目标之一，智能人机对话系统能够理解人类的话语并且能够正确的做出响应。目前越来越多的智能语音助手开始在人们的生活中普遍应用，这些智能语音助手需要利用到口语对话理解技术（Spoken language understanding, SLU）。口语对话理解技术旨在能够理解用户的指令，帮助用户执行各种任务或者能够与用户进行闲谈交流。具体地说，口语对话理解技术旨在从用户的话语中形成语义框架，其中包括意图检测（Intent detection）和槽填充(Slot filling) 两个部分(Tür and Mori, 2011; 俞凯等, 2015; 赵阳洋等, 2020)。它能区分用户话语中的意图信息，并捕捉用户话语中的语义组成成分(Liu and Lane, 2016)。如图1 所示,以一个音乐相关的话语为例：“我要听周杰伦的等你下课”，话语中每个字与槽标签一一对应，并且存在整个话语的意图，在这个例子中整句的意图是“PlayMusic”。

话语	我	要	听	周	杰	伦	的	等	你	下	课
槽标签	0	0	0	B-歌手名	I-歌手名	I-歌手名	0	B-歌曲名	I-歌曲名	I-歌曲名	I-歌曲名
位置标签	0	0	0	B	M	E	0	B	M	M	E
意图	PlayMusic										

Figure 1: 一个带有意图和槽信息标注(BIO 格式)的示例

在口语理解中，考虑到意图和槽信息之间的强关联性，许多研究者开始将两个任务联合在一起，如Zhang和Wang (2016)提出了使用循环神经网络联合意图检测和槽填充模型来解决上述问题。而Liu 和Lane (2016)则将注意力机制引入到了SLU的联合学习模型中。随后，Li等人 (2018)、Goo等人 (2018)、E等人 (2019)、Qin等人 (2019)尝试在SLU的联合学习模型中显式建立意图检测和槽填充的联系。

虽然这些模型都取得巨大的进步，但是这些模型都集中于利用意图语义知识增强模型槽信息填充能力(Liu and Lane, 2016; Goo et al., 2018; E et al., 2019; Qin et al., 2019)或者利用外部知识增强模型槽信息填充的能力(Wang et al., 2020; Nie et al., 2020)，缺少了槽信息填充中的边界信息的考虑，因而无法正确识别槽位边界，进而影响整体槽位识别的正确率。而且这些模型通常使用BIO格式或者使用BMESO格式对槽信息标签进行标注（在BIO标注中，B (Begin) 代表这是槽的开头，I (Inner) 代表槽的内部，O (Outside) 代表这是槽外部；在BMESO标注中，M (Middle) 也代表槽内部，还增加E (End) 标签代表槽的结束，S (Single) 代表单个词构成一个槽位）。槽信息填充模型对所有的标签是统一对待，不区分处理的，模型仅对话语的上下文语义信息建模，根据当前上下文语义预测当前话语对应的槽标签。例如，槽标签“B-歌手名”和“B-歌曲名”，这两个标签都代表槽位的开头，蕴含着槽的边界信息和位置信息。但是研究进展模型对这些标签没有进行区分，因而导致槽标签之间位置信息不能共享，这种情况在小样本的槽标签中更加普遍。

在槽信息填充任务中，槽指称项 (slot mentions) 比如说歌曲名字、电影名字等，如图1所示的歌曲名，“等你下课”和常见的话语难以区分，模型往往会受这些槽指称项的干扰，最终损失模型槽位提取表现。在口语对话理解中，很多表述结构除了槽指称项存在区别外，整句话语的结构是相似的。模型在处理槽信息填充任务时，需要能够聚焦于非槽位话语，以增强模型区

分槽位边界的能力。Liu等人(2020)提出使用模板机制，但是这种机制仅通过替换训练集中的槽指称项，是一种静态的数据扩增方法。而且这种方法只能利用在训练集中，而且要求标签语义空间和话语的语义空间相似。

为了能够共享槽标签之间的位置信息，且能够让模型更加聚焦于非槽指称项的话语，以便于增强模型区分边界的能力，在本文中，我们提出结合边界预测和动态模板的槽填充(Boundary-prediction and Dynamic-template Slot Filling, BDSF)模型。该模型通过预测边界(BMESO标注)辅助任务，判断句子中哪些话语词是槽实体。由于引入边界预测的辅助任务，模型能够学习到槽标签之间的位置信息。我们将边界预测辅助任务输出结果，引入到槽信息填充任务中。通过边界辅助任务输出，检测出话语中的槽实体，利用特殊的标识符动态地掩盖掉实体部分话语信息，让模型更加关注于非槽位部分的话语词。抽象后的话语经过编码器编码后，能够获得不带槽指称项的上下文表示，这种上下文表示能够聚焦于句子结构，能引导模型槽标签预测过程。通过模型迭代训练，模型最终能够抽象出话语的模版表示。由于模型是通过BIO辅助任务遮盖实体，因而这种方法也能够利用到测试集中，用于引导测试集的槽信息填充任务。

本文主要贡献如下：

- (1) 提出了一种结合边界预测和动态模板的槽填充模型，该模型解决了槽信息填充任务中无法利用槽标签之间位置信息问题。通过预测槽位的边界标签，遮盖话语中的槽实体词，动态产生话语的模板，让模型能够更加关注非槽指称项词，同时也使模型学习话语之间的结构。
- (2) 在公开基准数据集上的实验结果表明，本文的模型在槽位提取指标F1上取得了显著优于目前研究进展口语理解模型的性能。消融实验定量分析进一步证明了本文提出的模型的良好效果。通过对BDSF和基线模型的预测槽向量降维投影，可以直观地看出位置信息能够让槽标签共享位置信息，增强模型对边界预测能力。

2 相关工作

口语理解(spoken language understanding, SLU)已经有很久的研究历史。在这一章节中，本文将介绍在口语对话理解通常涉及到两个任务，包括用户话语的意图检测和槽填充(Tür and Mori, 2011)。其中意图检测被看成是分类任务，常用方法包括SVM(Haffner et al., 2003)、Adaboost(Schapire and Singer, 2000)。而槽填充任务被看成是序列标注任务，传统的方法是使用具有较强序列标注能力的CRF结构(Raymond and Ricciardi, 2007)和循环神经网络(Recurrent Neural Network, RNN)(Xu and Sarikaya, 2013)。近年来，越来越多的模型使用深度学习框架去提升意图检测效果，例如卷积神经网络(convolutional neural network, CNN)(Yoon Kim, 2014)、RNN(Mesnil et al., 2015)等方法。在槽信息填充方面，Yao等人(2014)提出的基于LSTM模型的槽填充方法。研究者还进一步使用CRF作为句子级的优化标准，使得模型能够处理输出序列之间的依赖关系(Yao et al., 2014)。

考虑到分开处理意图检测和槽填充任务，模型不能够充分利用两者之间的交互的信息，而且模型可能会遭受错误传播。因此，Zhang和Wang(2016)提出了使用循环神经网络联合意图检测和槽填充模型来解决上述问题。而Liu和Lane(2016)则将注意力机制引入到了SLU的联合学习模型中。这些方法(Zhang and Wang, 2016; Liu and Lane, 2016)虽然超越了流水线模型，但是知识利用了两个任务之间参数之间的共享，并没有利用显性的意图知识和槽标签知识。

为了能够显性地利用意图标签知识和槽标签知识，Li等人(2018)、Goo等人(2018)、E等人(2019)、Qin等人(2019)尝试在SLU的联合学习模型中显式建立意图检测和槽填充的联系。最近，研究者们将上下文相关的单词表征，如将BERT(Devlin et al., 2019)应用到SLU任务中(Chen et al., 2019)，进一步提高了SLU的性能。也有一些研究者利用胶囊网络建模意图识别和槽信息填充任务(Zhang et al., 2019)。

尽管SLU任务已经取得了较大的进展，但在槽位边界识别上方面仍然存在一些不足，进而影响槽位提取的模型表现效果。因此，本文在研究进展的SLU模型的基础上，添加边界预测辅助任务，并引入动态模版增强模型边界识别能力。

3 方法

在本节中，将介绍本文提出的用于SLU任务的动态模版方法，动态模版方法包含两步，第一步需要预测出话语中词的是否为实体，通过预测话语中词语是否为槽实体，模型能够学习到位置信息和边界信息。第二步需要使用特殊符号，掩盖话语中的实体，让模型专注于非槽指称项的话语词，以便模型能够学习到话语表达的句式。

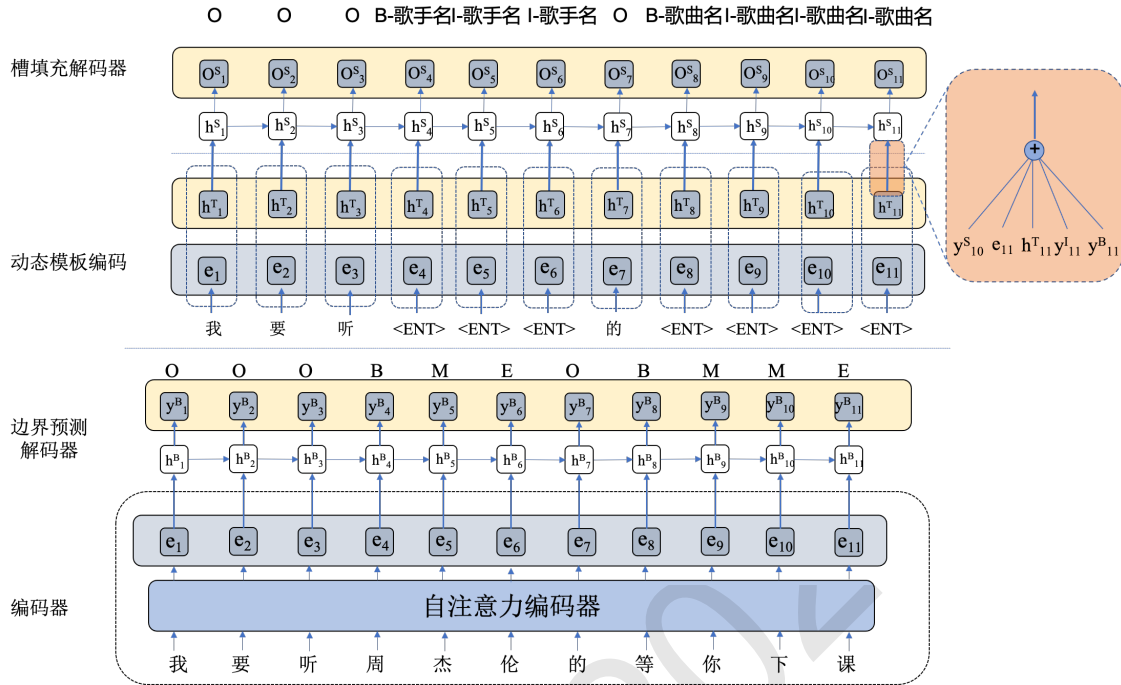


Figure 2: BDSF模型图（只展示了辅助任务、动态模版机制和槽信息填充模块。局部放大图中的S (slot) 代表槽信息填充任务，e是编码层输出向量，T (template) 代表动态模板机制，I (intent) 代表意图识别任务，B (boundary) 代表边界预测任务，图中“+”为拼接操作）

3.1 编码层

输入话语包含着 T 个词 $w = [w_1, w_2, \dots, w_T]$ ，输入话语经过词嵌入层，我们可以得到这些词的词向量表示。

$$x_t = e^w(w_t) \tag{1}$$

$$X = [x_1, x_2, \dots, x_T], X \in \mathbb{R}^{T \cdot d} \tag{2}$$

其中 e^w 为输入话语的词嵌入层， d 为词嵌入层维度，话语的词嵌入层通过随机初始化产生，在训练过程中微调学习词嵌入层的参数。得益于Bi-LSTM强大的序列处理能力，为了能够让模型处理系长序列，避免循环神经网络存在的记忆遗忘问题，模型中使用Bi-LSTM层对输入的词向量进行编码层，Bi-LSTM (Hochreiter and Schmidhuber, 1997; Yao et al., 2014)模型内部结构如下定义：

$$\vec{h}_t = LSTM(x_t, \vec{h}_{t-1}) \tag{3}$$

$$\overleftarrow{h}_t = LSTM(x_t, \overleftarrow{h}_{t-1}) \tag{4}$$

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \tag{5}$$

$$H = [h_1, h_2, \dots, h_T] \tag{6}$$

自注意力机制已经被证实在口语理解任务中能够增强模型表现，自注意力机制是一种特殊的注意力机制(Vaswani et al., 2017)，它只需要一个句子去计算它的表征，在本模型中，我们采用自注意力机制去捕捉每一个词的上下文信息，在本文中我们根据Vaswani 等人(2017)的设计，将输入的词向量 \mathbf{X} 通过 W^K 、 W^Q 和 W^V 的线性映射层(Linear projection layers):

$$\mathbf{K} = \mathbf{W}^K \mathbf{X}, \mathbf{Q} = \mathbf{W}^Q \mathbf{X}, \mathbf{V} = \mathbf{W}^V \mathbf{V} \quad (7)$$

我们将自注意力输出向量和Bi-LSTM的隐藏层向量拼接在一起，最终的输出向量为:

$$\mathbf{C} = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (8)$$

$$\mathbf{E} = [\mathbf{H}; \mathbf{C}] \quad (9)$$

3.2 边界预测辅助任务

为了捕获槽边界信息，添加预测槽在句子中是不是属于槽作为模型的辅助任务，使用的BMESO进行标注，BMESO相较于BIO标注方式能够提供更加细粒度的位置信息。在槽信息填充任务中，输入话语包 $\mathbf{E} = [e_1, e_2, \dots, e_T]$ 编码层向量，在边界预测辅助任务中，我们使用单向LSTM模型作为边界预测任务的解码器，在每一个时间步 t 中，解码器的隐藏层向量 h_t 由当前输入的编码层向量 e_t 和解码器上一次隐藏层向量 $h_{t-1}^{boundary}$ 和解码器上一次输出的标签分布 $y_{t-1}^{boundary}$ 决定:

$$h_t^{boundary} = LSTM^{boundary}([h_{t-1}^{boundary}; y_{t-1}^{boundary}; e_t]) \quad (10)$$

最终将Bi-LSTM的隐藏层向量 h_t 输入到槽边界预测中:

$$y_t^{boundary} = \text{softmax}(W_h^{boundary} h_t^{boundary}) \quad (11)$$

$$o_i^{boundary} = \text{argmax}(y^{boundary}) \quad (12)$$

$o^{boundary}$ 是模型辅助任务五分类的预测结果，由“B”、“M”、“E”，“S”，“O”五种标签组成。

3.3 动态模版机制

为了能够更好的区分槽位边界，避免受到槽指称项的干扰，模型应当更加聚焦于非槽指称项的词；在口语理解中，用户话语总是遵循某种语法结构，这些结构能够帮助我们更好地提取槽边界信息。为了能够学习到这种结构信息、更好地识别槽边界，受到(Zhong and Chen, 2020; Soares et al., 2019; Liu et al., 2020)启发，我们提出了动态模版机制，我们通过边界预测任务识别出话语中的槽位，使用特殊标记符号 $\langle ENT \rangle$ (代表是一个实体)代替原来的槽指称项，如图2所示，“我想听周杰伦的等你下课”经过边界预测后，遮盖住槽指称项变为“我想听 $\langle ENT \rangle$ 的 $\langle ENT \rangle$ ”，模型能够更加聚焦于句式表达。修改后的输入话语序列为:

$$\widehat{\mathbf{W}} = [w_1, w_2, \dots, \langle ENT \rangle, \dots, w_T] \quad (13)$$

上面公式代表部分输入话语中槽指称项用特殊符号代替，再将修改后的话语 $\widehat{\mathbf{W}}$ 输入到编码器中，获取模版的上下文表示:

$$\mathbf{x}^{template} = e^w(\widehat{\mathbf{W}}) \quad (14)$$

$$h_t^{template} = BiLSMT(\mathbf{x}_t^{template}, h_{t-1}^{template}) \quad (15)$$

3.4 联合学习

对于意图识别, Qin等人(2019)发现使用字级别 (Token-level) 意图能够达到更好的模型表现效果, 每一次字输出一个意图, 输入到后续槽标签预测中, 在最终测试阶段选出出现次数最多的意图作为最终意图。在本模型中, 我们也采用字级别意图。在意图识别任务中, 我们使用单向LSTM模型作为意图预测的解码器, 在每一个时间步 t 中, 解码器的隐藏层向量 h_t 通过当前输入的编码层向量 e_t 和解码器上一次隐藏层向量 h_{t-1}^{intent} 和解码器上一次输出的意图标签分布 y_{t-1}^{intent} 决定:

$$h_t^{intent} = LSTM^{intent}([h_{t-1}^{intent}, y_{t-1}^{intent}, e_t]) \quad (16)$$

最终将LSTM的隐藏层向量 h_t^{intent} 输入到意图预测层中:

$$y_t^{intent} = softmax(W_h^{intent} h_t^{intent}) \quad (17)$$

$$o_t^{intent} = argmax(y^{intent}) \quad (18)$$

其中 W^{intent} 是意图标签预测的权重矩阵, o_i^{intent} 最终预测的意图标签。

槽标签预测与意图预测过程相似, 在每一个时间步 t 中, 解码器的隐藏层向量 h_t 通过当前输入的编码层向量 e_t 和解码器上一次隐藏层向量 h_{t-1}^{slot} 及解码器上一次输出的槽标签分布 y_{t-1}^{slot} 决定, 为了能够让模型在槽信息填充任务中获取话语的边界信息和句式信息, 我们将边界预测的输出向量 $y_t^{boundary}$ 和句子的模版表示 $h^{template}$ 也引入到槽信息填充任务, 为了能够让模型在槽信息填充任务中获取话语意图信息, 我们将意图输出向量 y_t^{intent} , 也引入到槽信息填充任务:

$$h_t^{slot} = LSTM([h_{t-1}^{slot}, y_t^{intent}, y_t^{boundary}, h_t^{template}, y_{t-1}^{slot}, e_t]) \quad (19)$$

最终将LSTM的隐藏层向量 h_t^{slot} 输入到槽标签预测层中:

$$y_t^{slot} = softmax(W_h^{slot} h_t^{slot}) \quad (20)$$

$$o_t^{slot} = argmax(y^{slot}) \quad (21)$$

其中 W^{slot} 是槽标签预测的权重矩阵, o_i^{slot} 最终预测的槽标签。上述机制在图2中展示, 为了简洁, 略去了输入到槽信息填充解码器的LSTM的各信息源的连线, 槽信息填充解码器的LSTM的各信息源输入可以参考公式 (19) 和图2中的局部放大图。

边界预测辅助任务、意图识别和槽信息填充的损失函数定义是:

$$\mathcal{L}_{boundary} \triangleq - \sum_{t=1}^T \sum_{j=1}^{n^{boundary}} \hat{y}_j^{t,boundary} \log(y_j^{t,boundary}) \quad (22)$$

$$\mathcal{L}_{intent} \triangleq - \sum_{t=1}^T \sum_{j=1}^{n^{intent}} \hat{y}_j^{t,intent} \log(y_j^{t,intent}) \quad (23)$$

$$\mathcal{L}_{slot} \triangleq - \sum_{t=1}^T \sum_{j=1}^{n^{slot}} \hat{y}_j^{t,slot} \log(y_j^{t,slot}) \quad (24)$$

其中 t 代表着话语中第 t 个字, $\hat{y}_j^{t,boundary}$ 、 $\hat{y}_j^{t,intent}$ 和 $\hat{y}_j^{t,slot}$ 是真实边界标签、真实槽标签和真实意图标签, $n^{boundary}$ 、 n^{intent} 和 n^{slot} 分别代表着边界标签的个数、意图标签的个数和槽标签的个数。为了能够联合三个模型, 我们将三个任务损失函数联合, 总的损失函数为:

$$\mathcal{L} = \mathcal{L}_{intent} + \mathcal{L}_{slot} + \alpha \mathcal{L}_{boundary} \quad (25)$$

其中 α 是边界预测辅助任务的权重超参数, 用于决定边界辅助任务在总任务的重要性程度。

4 实验和结果分析

4.1 数据集

我们在CAIS (Chinese Artificial Intelligence Speakers) 数据集(Liu et al., 2019)和SMP-ECDT 2019^{1 2} 的公共基准文本语料上对我们的模型进行训练和评估。CAIS数据集具有7995条训练集话语、994条验证集话语和1024条测试集话语。SMP-ECDT 2019 (Social Media Processing - the Evaluation of Chinese Human-Computer Dialogue Technology)是来自ASR系统的转录。SMP-ECDT 数据集有24种不同的意图类型和124种不同的槽位标签。CAIS和SMP-ECDT数据集的统计结果见表1。

	CAIS	SMP-ECDT
训练集数量	7995	1656
验证集数量	994	414
测试集数量	1024	509
槽标签数量	75	124

Table 1: CAIS 和SMP-ECDT 数据集的统计结果

4.2 超参数

为了避免过拟合, Dropout (Srivastava et al., 2014)被设置为0.1到0.5进行验证, 实验中使用的Adam优化器 (Kingma and Ba, 2015), 学习率为1e-5, 批量大小 (Batch size) 设置为32, L2正则化参数为1e-6。此外, 我们使用交叉熵作为边界预测、意图检测和槽填充任务的损失函数, 在槽信息填充任务中, 我们与其他研究工作保持一致都使用BIO格式标注槽实体(Goo et al., 2018; E et al., 2019; Qin et al., 2019), 但是在边界预测的辅助任务中, 我们使用更加细粒度的BMESO标注格式, 以提供更多的位置信息到模型中, 边界预测辅助任务的权重 α 设置为1.0。以下结果所示数据均为10个独立实验的平均值。

4.3 对比的方法

我们将动态模版模型与以下方法进行了比较, 包括:

- Slot-Gated Atten. Goo等人 (2018)提出了信息槽门控联合模型, 以更好地探究槽填充与意图检测的相关性。
- SF-ID Network. E等人 (2019)介绍了一种SF-ID网络, 为槽填充和意图检测建立起直接性的连接, 以帮助它们相互促进与提升。
- CM Network. Liu等人 (2019)介绍了一种记忆网络网络, 在话语、槽填充和意图检测建立起直接性的连接, 以帮助它们相互促进与提升。
- Stack-propagation. Qin等人 (2019)采用stack-propagation框架, 提出了一个结合单词级别意图识别的联合学习模型。

4.4 主实验

与已有的工作 (Goo et al., 2018; E et al., 2019; Chen et al., 2019)一样, 我们在槽填充中使用F1评分, 意图预测中正确率和使用句子正确率来综合评估口语理解的性能, 表2展示了我们提出的动态模版模型的实验结果, 并在CAIS数据集和SMP-ECDT数据集上进行了比较与分析。本文的重点是改进口语语言理解中的信息槽填充, 因此我们主要观察不同模型在槽填充F1评分和句子整体语义正确性上的表现。

从表2中可以看出, 利用边界预测辅助任务和动态模板中机制, 本文提出的BDSF在两个公共基准语料库上的在槽位提取和整句正确率方面显著优于所有其他模型。这些结果证明了边界预测的辅助任务和动态模版机制能够增强槽填充模型的性能。在CAIS数据集中, BDSF模型的

¹<http://conference.cipsc.org.cn/smp2019/evaluation.html>

²<https://github.com/OnionWang/SMP2019-ECDT-NLU>

模型	CAIS			SMP-ECDT		
	意图识别(Acc)	槽位提取(F1)	整句(Acc)	意图识别(Acc)	槽位提取(F1)	整句(Acc)
Slot-Gated Full Atten	94.37	81.13	80.03	84.81	60.88	54.21
SF-ID Network	94.27	84.85	82.41	86.51	64.52	56.33
CM-Net	94.56	86.16	-	-	-	-
Stack-propagation	94.37	87.64	84.68	88.65	71.62	63.78
BDSF	94.26	88.47	85.17	88.99	75.13	66.40

Table 2: 我们的模型和对比方法在测试集上的性能对比。

槽位提取相较于Stack-propagation 提升了0.83%，整句正确率也提升了0.49%。在SMP数据集中，存在较多句式相近的样本，由于采用了动态模板机制，BDSF模型的槽位提取和整句性能上都显著优于比较模型。相较于Stack-propagation，BDSF模型的槽位提取表现提升了3.51%，整句正确率也有2.62%提升。

4.5 消融实验

为了研究本文采用的边界辅助任务和动态模板机制在BDSF模型中的贡献，我们在SMP-ECDT数据集上进行了消融实验。实验中，我们在本文基础模型Stack-propagation (Qin et al., 2019)的基础上，增加边界预测以及进一步增加动态模板。结果如表3所示。由于动态模板机制需要在边界预测任务识别出话语中的槽位之后才能使用，因此，没有单独增加动态模板机制的实验。此外，我们对比了BIO和BMESO两种边界预测方案的效果。

模型	SMP-ECDT		
	意图识别(Acc)	槽位提取(F1)	整句(Acc)
Stack-propagation	88.65	71.62	63.78
+ BIO边界预测	89.78	72.58	64.83
+ BIO边界预测和动态模版	88.4	74.34	66.01
+ BMESO边界预测	89.36	73.99	66.57
+ BMESO边界预测和动态模版(BDSF)	88.99	75.13	66.40

Table 3: BDSF消融实验

结果表明，我们的模型中采用的边界辅助任务和动态模板机制对于模型的整体性能都是有价值的。详细的分析如下：

- 使用BIO边界预测：通过引入槽边界预测的辅助任务，能够让模型隐性的学习到位置相关信息，并且能够让模型能够共享槽之间的开头、中间、结尾等位置信息。从结果我们可以看出通过引入BIO边界预测能够在原来模型基础上，槽位提取表现提升0.96%。
- 使用BIO边界预测和动态模版：基于边界预测任务结果，我们将话语中的槽指称项遮盖，让模型聚焦于非槽指称项的话语，通过引入动态模版机制，能够让模型学习话语表达中的句子结构，并且能够缓解被话语中槽指称项影响。从结果我们可以看出通过引入BIO边界预测和动态模版机制能够比基线模型槽位提取表现提升2.76%。
- 使用BMESO边界预测：通过引入更加细粒度槽边界预测的辅助任务，能够让模型隐性的学习到更加丰富的位置相关信息。实验结果表明，使用BMESO标注能够比BIO标注取得更好的实验效果，使用BMESO标注的辅助任务在槽位提取表现，相较于基线模型提升2.37%。
- BMESO边界预测和动态模版(BDSF)：通过引入更加细粒度槽边界预测的辅助任务和动态模板机制，实验结果表明，使用结合边界预测和动态模版的模型，在基于BMESO标注下，比BIO标注的结合边界预测和动态模版的模型在槽位提取F1值高出0.78%；比仅使用基于BMESO标注的边界预测辅助任务模型在槽位提取F1值高1.14%；使用BMESO标注的结合边界预测辅助任务和动态模版模型(BDSF)在槽位提取表现上相较于基线模型提升3.51%。

- Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, pages 4171–4186, 2019.
- Haihong E, Peiqing Niu, and Zhongfu Chen, et al. A novel bi-directional interrelated model for joint intent detection and slot filling. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019*, pages 5467–5471, 2019.
- Chih-Wen Goo, Guang Gao, and Yun-Kai Hsu, et al. Slot-gated modeling for joint slot filling and intent prediction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018*, pages 753–757, 2018.
- Patrick Haffner, Gökhan Tür, and Jerry H. Wright. Optimizing svms for complex call classification. In *Proceedings of the 28th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '03*, pages 632–635, 2003.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, 1997.
- Zhanming Jie and Wei Lu. Dependency-guided LSTM-CRF for named entity recognition. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019*, pages 3860–3870, 2019.
- Yoon Kim. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014*, pages 1746–1751, 2014.
- Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, 2015.
- Changliang Li, Liang Li, and Ji Qi. A self-attentive model with gate mechanism for spoken language understanding. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*, pages 3824–3833, 2018.
- Bing Liu and Ian Lane. Attention-based recurrent neural network models for joint intent detection and slot filling. In *Proceedings of the 17th Annual Conference of the International Speech Communication Association, INTERSPEECH 2016*, pages 685–689, 2016.
- Yijin Liu, Fandong Meng, Jinchao Zhang, Jie Zhou, Yufeng Chen, and Jinan Xu. Cm-net: A novel collaborative memory network for spoken language understanding. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019*, pages 1051–1060, 2019.
- Zihan Liu, Genta Indra Winata, Peng Xu, and Pascale Fung. Coach: A coarse-to-fine approach for cross-domain slot filling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020*, pages 19–25, 2020.
- Grégoire Mesnil, Yann N. Dauphin, and Kaisheng Yao, et al. Using recurrent neural networks for slot filling in spoken language understanding. *IEEE ACM Trans. Audio Speech Lang. Process.*, 23(3):530–539, 2015.
- Yuyang Nie, Yuanhe Tian, Yan Song, Xiang Ao, and Xiang Wan. Improving named entity recognition with attentive ensemble of syntactic information. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings, EMNLP 2020*, pages 4231–4245, 2020.
- Libo Qin, Wanxiang Che, and Yangming Li, et al. A stack-propagation framework with token-level intent detection for spoken language understanding. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019*, pages 2078–2087, 2019.
- Lev-Arie Ratinov and Dan Roth. Design challenges and misconceptions in named entity recognition. In Suzanne Stevenson and Xavier Carreras, editors, *Proceedings of the Thirteenth Conference on Computational Natural Language Learning, CoNLL 2009*, pages 147–155, 2009.
- Christian Raymond and Giuseppe Riccardi. Generative and discriminative algorithms for spoken language understanding. In *Proceedings of the 8th Annual Conference of the International Speech Communication Association, INTERSPEECH 2007*, pages 1605–1608, 2007.

- Robert E. Schapire and Yoram Singer. Boostexter: A boosting-based system for text categorization. *Mach. Learn.*, 39(2/3):135–168, 2000.
- Livio Baldini Soares, Nicholas FitzGerald, Jeffrey Ling, and Tom Kwiatkowski. Matching the blanks: Distributional similarity for relation learning. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Volume 1: Long Papers*, pages 2895–2905.
- Nitish Srivastava, Geoffrey E. Hinton, and Alex Krizhevsky, et al. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958, 2014.
- G. Tür and R. Mori. 2011. *Spoken language understanding: Systems for extracting semantic information from speech*, John Wiley and Sons, Inc.
- Ashish Vaswani, Noam Shazeer, and Niki Parmar, et al. Attention is all you need. In *Proceedings of the 2017 Annual Conference on Neural Information Processing Systems, NIPS 2017*, pages 5998–6008, 2017.
- Ngoc Thang Vu, Pankaj Gupta, and Heike Adel, et al. Bi-directional recurrent neural network with ranking loss for spoken language understanding. In *Proceedings of the 41st IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2016*, pages 6060–6064, 2016.
- Jixuan Wang, Kai Wei, Martin Radfar, Weiwei Zhang, and Clement Chung. Encoding syntactic knowledge in transformer encoder for intent detection and slot filling. *CoRR*, abs/2012.11689, 2020.
- Puyang Xu and Ruhi Sarikaya. Convolutional neural network based triangular CRF for joint intent detection and slot filling. In *Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2013*, pages 78–83, 2013.
- Kaisheng Yao, Baolin Peng, and Geoffrey Zweig, et al. Recurrent conditional random field for language understanding. In *Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2014*, pages 4077–4081, 2014.
- Kaisheng Yao, Baolin Peng, and Yu Zhang, et al. Spoken language understanding using long short-term memory neural networks. In *Proceedings of the 2014 IEEE Spoken Language Technology Workshop, SLT 2014*, pages 189–194, 2014.
- Yu Sun, Shuohuan Wang, Yu-Kun Li, Shikun Feng, Xuyi Chen, Han Zhang, Xin Tian, Danxiang Zhu, Hao Tian, and Hua Wu. ERNIE: enhanced representation through knowledge integration. *CoRR*, abs/1904.09223, 2019.
- Chenwei Zhang, Yaliang Li, Nan Du, Wei Fan, and Philip S. Yu. Joint slot filling and intent detection via capsule neural networks. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019 Volume 1: Long Papers*, pages 5259–5267, 2019.
- Xiaodong Zhang and Houfeng Wang. A joint model of intent determination and slot filling for spoken language understanding. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016*, pages 2993–2999, 2016.
- Zexuan Zhong and Danqi Chen. A frustratingly easy approach for joint entity and relation extraction. *CoRR*, abs/2010.12812, 2020.