

基于两阶段分类的口语理解方法

吴尉林 陆汝占 段建勇 刘 慧 高 峰 陈玉泉

(上海交通大学计算机科学与工程系 上海 200240)

(wuw1@cs.sjtu.edu.cn)

A Spoken Language Understanding Approach Based on Two-Stage Classification

Wu Weilin, Lu Ruzhan, Duan Jianyong, Liu Hui, Gao Feng, and Chen Yuquan

(Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240)

Abstract Spoken language understanding (SLU) is one of the key components in a spoken dialogue system. One challenge for SLU is robustness since the speech recognizer inevitably makes errors and spoken language is plagued with a large set of spontaneous speech phenomena. Another challenge is portability. Traditionally, the rule-based SLU approaches require linguistic experts to handcraft the domain-specific grammar for parsing, which is time-consuming and laboursome. A new SLU approach based on two-stage classification is proposed. Firstly, the topic classifier is used to identify the topic of an input utterance. Then, with the restriction of the recognized target topic, the semantic slot classifiers are trained to extract the corresponding slot-value pairs. The advantage of the proposed approach is that it is mainly data-driven and requires only minimally annotated corpus for training whilst retaining the understanding robustness and deepness for spoken language. Experiments have been conducted in the Chinese public transportation information inquiry domain and the English DARPA Communicator domain. The good performance demonstrates the viability of the proposed approach.

Key words spoken dialogue system; spoken language understanding; statistical classifier; topic classification; decision list

摘 要 口语理解是实现口语对话系统的关键技术之一.它主要面临两方面的挑战:1)稳健性,因为输入语句往往是病态的;2)可移植性,即口语理解单元应能够快速移植到新的领域和语言.提出了一种新的基于两阶段分类的口语理解方法:第1阶段为主题分类,用来识别用户输入语句的主题;第2阶段为主题相关的语义槽分类,根据识别的主题抽取相应的语义槽/值对.该方法能对用户输入语句进行深层理解,同时也能保持稳健性.它基本上是数据驱动的,而且训练数据的标记也比较容易,可方便地移植到新的领域和语言.实验分别在汉语交通查询领域和英语 DARPA Communicator 领域进行,结果表明了该方法的有效性.

关键词 口语对话系统;口语理解;统计分类器;主题分类;决策表

中图法分类号 TP391.2

口语理解是实现口语对话系统的关键技术之一.它的任务是理解用户的意图并抽取用户输入语句所包含的关键信息.通常用户输入语句的语义可用语义框架(semantic frame)表示.一个语义框架通常包括:

1) 语义框架类型(frame type)表示用户输入语句的主题;2) 相应的一些语义槽(slots)表示用户输入语句中的关键信息.如果采用语义框架表示,则口语理解可看做是把输入语句转换为相应的语义框架的工作.

传统的口语理解方法主要分为两种:规则方法和统计方法.规则方法主要利用稳健的句法分析或者模板匹配技术来分析口语^[1-2],而统计方法则是把口语理解建模为统计解码过程^[3-4].两种方法各有优缺点.规则方法的优点是不需要大量的标记语料库.但是其开发成本较高,因为规则通常需要语言学家手工定义,费时费力.这样,很难移植到新的应用领域和语言.相反地,统计方法可从标记语料库自动训练.与手工定义规则相比,标记语料库比较容易而且基本不需要语言学知识.但是,统计方法会受到数据稀疏问题的影响.要训练一个较精确的模型,通常需要大量完全标记的语料库(例如树库).这在实际应用领域中往往是不现实的.最近,为了实现对话系统的快速开发,研究者们提出了一些新的口语理解方法.Wang 和 Acero 提出了一种结合多种知识源半自动地学习规则的方法^[5-6].He 和 Young 提出了基于 Hidden Vector Model (HVS) 的口语理解分析器^[7-8],HVS 是对隐马尔科夫模型(HMM)的扩

展.同时,不少研究者也提出了结合规则和统计方法的口语理解方法^[9-11],力图综合两者的优点.

本文提出了一种新的口语理解方法.整个口语理解框架主要由两种分类器构成,其中主题分类器用来识别用户输入语句的主题,而语义槽分类器则根据识别的主题抽取相应的语义槽/值对(slot-value pairs).该方法能对用户输入语句进行深层的理解,同时能保持稳健性.它基本上是数据驱动的,而且训练只需要简单标记的数据,这样可以很容易地移植到别的应用领域和语言.

1 基于两阶段分类的口语理解方法

除主题分类器和语义槽分类器之外,基于两阶段分类的口语理解系统还包括预处理器和语义框架生成器.图1给出了系统整体架构,同时也描述了例句“请问坐公交车从徐家汇到外滩怎么走”在各个阶段的处理结果.

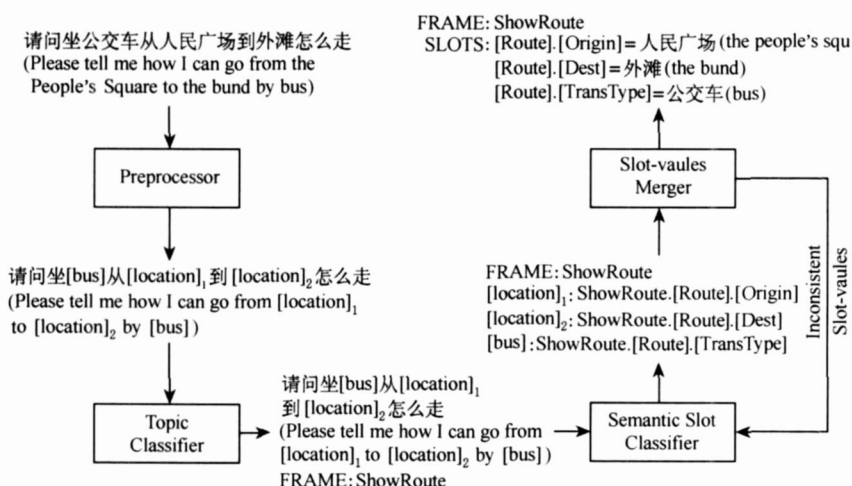


Fig. 1 The SLU framework based on two-stage classification.

图1 基于两阶段分类的口语理解框架

1.1 预处理

通常,在口语对话系统的开发中,需要预先定义一些语义类,代表领域中的概念,如地名、时间表达式等.一般地,预处理就是在输入语句中找到某些子串(它们对应某个语义类或者匹配某些正则表达式),并把它们替换为相应的语义类标记.例如,“华山路”和“1954”可分别替换为[road_name] (表示路名)和[number] (表示数字).本系统中的预处理器除了能做这种简单的替换外,也能识别一些复杂的概念.例如,通过匹配规则“[address] → [loc_name] [number] 号”和“[location] → [address] (表示地

址)”,可把“华山路 1954 号”整体标记为[location] (表示位置名).在本系统中,预处理的任务主要由一个基于图算法的局部句法分析器来完成.它是对文献[12]中的句法分析器的扩展,能跳过输入语句中的噪声词和语法中的规则符.这样就从底层保证了口语理解的稳健性.例如,“华山路那个 1954 号”可以通过跳过“那个”匹配规则“[address] → [road_name] [number] 号”而识别成[address],然后匹配规则“[location] → [address]”,最终识别为[location].但是,该稳健的局部分析器也有可能错误地跳过输入语句中的词而生成错误的概念.为了

避免这个副作用, 我们引入内置的决策树在分析过程中进行剪枝^[13]. 目前, 这个局部分析器所用的语法是手工定义的. 根据我们的经验, 手工定义这些规则以及标记决策树的训练实例的工作比较简单, 可在数小时内完成.

引入预处理器可带来以下好处:

- 1) 能给主题分类提供深层的特征, 使主题分类器具有更强的泛化能力, 同时能减轻数据稀疏问题.
- 2) 预处理能对语义相近的概念进行聚类, 例如, 地址类[address] 和地名类[loc_name] 可聚类为位置类[location]. 这样也减少语义槽分类器的数目, 使得语义槽分类模型的参数估计更精确.

1.2 主题分类

主题分类是口语理解的子问题, 它的任务是识别输入语句的主题. 利用统计分类器进行主题分类具有很好的稳健性, 这样也提高了整个口语理解系统的稳健性. 而且, 先进行主题分类能缩小语义槽分类的搜索空间, 从而提高语义槽分类的性能. 例如, 在汉语交通查询领域中, 对于所有主题, [location] 概念所对应的可能的语义槽有 33 个之多. 而对单个主题而言, [location] 概念所对应的语义槽最多的主题是 Show Route, 而其中可能的语义槽只有 10 个.

主题分类是典型的模式识别问题, 因此适合用统计分类器来处理. 很多统计分类方法已经应用在主题分类, 例如朴素贝叶斯分类器、支持向量机、 N 元模型和神经网络等^[9-10]. 本文采用支持向量机(support vector machines, SVM), 因为它具有较好的主题分类性能^[9]. 对于汉语输入语句来说, 最简单的特征就是出现在句子中的汉字. 假如使用汉字特征, 则每个汉语输入语句可转换为特征向量 $ch = \langle ch_1, \dots, ch_{|ch|} \rangle$, 其中 $|ch|$ 为应用领域的字表大小, ch_i 的取值是二值型的, 即如果对应的汉字出现在输入语句中, 则为 1, 反之为 0. 由于在主题分类之前进行了预处理, 可以把语义类标记(如[location] 等)也作为特征. 语义类标记特征比汉字特征能提供更多的信息. 同时, 引入语义类标记特征也能减轻数据稀疏问题.

1.3 主题相关的语义槽分类

在主题分类之后, 接下来的任务是把语句中的概念填充到语义框架相应的槽中. 该问题也可以建模为分类问题, 因为限定领域中的一个概念所对应的可能语义槽的个数是有限的. 语义槽分类的基本思路基于以下观察: 不同语义槽出现在不同的上下文中, 考虑图 1 中的例句, 在预处理和主题分类之

后, 得到预处理串“请问坐[bus] 从[location]₁ 到[location]₂ 怎么走”和识别的主题 Show Route. 现以[location]₂ 为例, 考虑如何为它标记相应的语义槽. 首先想到的线索是它的词汇上下文(lexical context), 例如, 由[location]₂ 前面的“到”字, 则可判断它是“目的地”. 如果[location]₁ 已经被识别为“出发地”, 则也是重要的线索, 暗示[location]₂ 是“目的地”, 这种上下文称为语义槽上下文(slot context). 因为开始时并没有语义槽上下文, 所以语义槽上下文特征只在语义槽重分类(semantic slot re-classification)时才使用, 语义槽重分类的内容将在第 1.4 节介绍.

为了自动抽取这些上下文特征, 需要对训练句子进行标记. 本文的标注方案比较简单, 只需要对句子标记相应的语义框架. 例如, 对于图 1 中的例句, 其对应的标注结果大致如下.

语义框架类型: Show Route

语义槽:

[Route] . [Origin] . [location] . (人民广场)

[Route] . [Dest] . [location] . (外滩)

[Route] . [TransType] . [bus] . (公交车)

其中, 语义槽的名字可自动从领域模型中抽取得到. 所谓领域模型是应用领域的相关概念的层次体系. 例如, 在 Show Route 主题中, “出发地”对应的语义槽是 Show Route. [Route] . [Origin]. 可以看出, 本文中的语义槽继承了领域模型的层次性. 所抽取的词汇上下文特征包括: 1) 上下文中的指示词(context word)指在待分类概念的 $\pm k$ 窗口内出现的单个字/词或语义类; 2) 搭配(collocation)指在待分类概念的 $\pm k$ 窗口内出现的长度最多为 l 的字/词或语义类的共现模式. 在本文实验中, k 和 l 分别取 4 和 2. 词汇上下文特征是词序相关的, 例如, “去”字出现在[location] 的左一位置表示[location] 是“目的地”, 而出现在右一位置则表示[location] 是“出发地”. 而在抽取语义槽上下文时, 只考虑出现在 $\pm m$ 窗口内出现的语义槽, 在本文实验中 m 取 6. 语义槽上下文是词序无关的, 比如, “目的地”可能在“出发地”之前, 也可能在之后. 词汇上下文特征反映了局部的词汇语义依赖关系, 而语义槽上下文则反映了长距离的语义依赖关系.

下面给出语义槽分类的上下文特征的例子, 例如, 对于主题 Show Route 中概念[location] 的语义槽分类, 相关的特征示例如下:

① “到”在 -2 窗口内;

② “从”_“到”;

③ Show Route. [Route]. [Origin] 在 ± 2 窗口内.

前两个特征为词汇上下文特征. 特征 1 表示: 如果“到”出现在 [location] 的 -2 窗口内, 则 [location] 的语义槽很可能是 Show Route. [Route]. [Dest] (“目的地”). 特征 2 则表示: 如果模式“从 [location] 到”得到匹配, 则 [location] 的语义槽往往是 Show Route. [Route]. [Origin] (表示“出发地”). 特征 3 为语义槽上下文特征, 表示如果 Show Route. [Route]. [Origin] 出现在 [location] 的 ± 2 窗口内, 则 [location] 的语义槽为 Show Route. [Route]. [Dest] 的可能性很大. 从本质上说, 这些特征等价于基于规则的句法分析器所用的语义规则, 例如, 特征 2 与规则 “[Origin] \rightarrow 从 [location] 到”是等价的. 通常, 在基于规则的口语分析器中, 这些规则是手工定义的. 手写规则需要语法知识、费时费力, 而在本文方法中, 能从标记语义框架的数据中自动学习到这些“规则”. 所学习到的这种“规则”本身是稳健的, 因为它可以包含空位 (gap), 例如特征 1 就允许“到”和 [location] 之间可以有噪声词.

在抽取了特征之后, 接下来的问题是如何使用这些特征. 对于待分类概念, 通常上下文中包含多个特征, 它们可能会作出不同的预测. 一种简单实用的方法是决策表 (decision list)^[14], 它的策略是总是相信最强的那个特征. 因此, 如何计算特征的置信度是重要问题. 本文采用如下指标^[15-16]: 假设对于所有 i , 条件概率 $Pr(s_i | f) > 0$, 则:

$$\text{confidence}(f) = \max_i Pr(s_i | f), \quad (1)$$

该指标衡量了特征 f 与语义槽 s_i 的相关程度.

1.4 语义槽合并和语义槽重分类

在语义槽分类后, 需要把已经识别的语义槽合并成语义框架. 由于语义槽继承了领域模型的层次性, 因此语义槽的合并比较容易, 具有相同前缀的语义槽可以合并到一起. 在合并的同时, 需要检查各语义槽/值的一致性. 因为对应不同概念的语义槽分类器是独立工作的, 所以有可能产生矛盾的语义槽/值对. 考虑预处理之后的词串“请问坐 [bus] 从 [location]₁ 到 [location]₂ 怎么走”, 如果 [location]₁ 和 [location]₂ 都被分类为 Show Route. [Route]. [Origin] (“出发地”), 则在语义上是矛盾的. 这时, 可利用语义槽上下文特征, 例如, “语义槽 Show Route. [Route]. [Origin] 出现在 $\pm k$ 窗口内”, 则表明当前概念的语义槽很可能是 Show Route.

[Route]. [Dest] (“目的地”). 因此, 当发现两个或多个语义槽/值对矛盾时, 首先保留置信度最高的那个语义槽, 然后把其他位置概念的语义槽上下文传回给语义槽分类器, 重新分类. 语义槽重分类的目的是利用长距离的语义依赖关系来克服局部的词汇语义依赖关系的局限性. 语义槽分类以及重分类的观点可以理解如下: 先分析出输入语句中不同位置的概念所对应的语义槽 (类似于部分句法分析, partial parsing), 然后把它们联系起来. 这种机制很适合口语理解, 因为口语中的句子经常是由几个短语组成, 而短语之间容易有噪声 (如重新开始、重复和停顿等), 而且语序很随意, 经常颠倒. 这些现象在汉语中特别常见.

2 实验

2.1 实验领域和设定

我们分别在汉语交通查询领域以及英文 DARPA Communicator 领域上对本文所提出的口语理解方法进行了测试.

对于汉语城市公共交通查询领域收集了两种语料. 首先, 我们建立了模拟对话系统的网站, 用户可以通过键盘输入与系统进行混合主导的对话来进行上海市交通信息的查询. 通过这种方式共收集了 2286 个查询语句, 随机划分为两部分: 训练集包括 1800 个句子 (CTR), 测试集包括 486 个句子 (CTS1). 另外, 构建了一个初始版本的基于电话的对话系统来收集口语语料库. 该对话系统的语音识别模块是基于 IBM ViaVoice Telephony 说话人无关的汉语听写系统开发的; 而口语理解模块是一个稳健的基于规则的句法分析器. 该口语语料库包括 363 个口语句子. 从口语语料库可得到两个测试集: 一个是语音识别结果构成的集合 (CTS2), 另一个是对应的人工转写文本构成的集合 (CTS3). 口语语料库语音识别错误严重, 在 CTS2 中, 汉字识别错误率和概念识别错误率分别是 35.6% 和 41.1%. 对于汉语交通查询领域定义了 10 种主题类型, 包括 ListStop, Show Fare, Show Route, Show RouteTime 等. 通过网络收集的语料库 (CTR+CTS1) 覆盖了全部 10 种主题类型, 其中 Show Route 主题类型的句子占 71.1%. 而通过电话收集的语料库 (CTS2/CTS3) 只包含其中 4 种主题, 其中 Show Route 主题类型的句子占 78.5%. 在汉语交通查询领域的实验中, 主题分类器和语义槽分类器统一在 CTR 上训练, 分

别在 CTS1, CTS2 和 CTS3 上测试. 训练集中出现的汉字共为 922 个.

英文 DARPA Communicator 领域的数据是由 University of Colorado 在 461 天内收集的^[17]. 该数据集不但提供了句子文本, 也提供了 Phoenix Parser 对这些句子的句法分析结果^[18]. 我们从中随机选取了 101 天的数据作为测试集, 其余为训练集. 整理之后, 得到的训练集(ETR)共包含 3724 个句子, 测试集(ETS)共包含 1279 个句子. 训练集的词表大小为 839. 对于英文 DARPA Communicator 领域, 定义了 5 种主题类型, 包括 RentCar, ReqFlight, ReqRoundFlight, ReqHotel, ReqReturnFlight, 其中 ReqFlight 类型的句子占 66% 左右. 该数据集的语义框架标记直接应用程序从 Phoenix Parser 分析结果中抽取, 然后手工校正.

主题分类性能指标采用主题分类错误率(topic error rate, TER), 即主题识别错误的测试句子占有所有测试句子的比例. 语义槽分类性能指标采用语义槽错误率(slot error rate, SER), 即比较语义槽分类器输出的结果和手工标记的标准语义槽, 计算相应的插入、删除和替换错误率. 如果对一个输入句子的主题分类出错, 则该句子中的每个语义槽都按一个替换错误计算.

2.2 实验结果

2.2.1 预处理实验结果

预处理的性能用概念识别错误率(concept error rate, CER)来评价, 即比较预处理器产生的概念和相应的手工标记的概念, 计算插入、删除和替换的错误率. 表 1 给出了汉语交通查询领域中的预处理结果, 可以看到预处理的性能是令人满意的, 注意在 CTS2 上性能较差的原因是语音识别错误较严重.

Table 1 The Preprocessing Performance on the Chinese Domain
表 1 汉语交通查询领域的预处理结果

Test Sets	Concept Error Rate(%)
CTS1	2.5
CTS2	43.2
CTS3	1.3

2.2.2 主题分类实验结果

预处理为主题分类提供了不同层次的特征, 我们对它们的主题分类性能进行了比较. 最简单的特征分别是汉字或英语词. 如果进行了基本语义类标记, 则可把基本语义类(如[road_name])作为特征. 进一步地如果进行了局部句法分析, 则可把深层语

义类(如[location])也作为特征. 表 2 给出了使用不同层次特征的支持向量机在 4 个测试集上的主题分类结果. 结果表明, 语义类特征的加入能明显地改善主题分类的性能.

Table 2 Performance Comparison of the SVMs Using Various Features on Four Test Sets

表 2 使用不同层次特征的支持向量机在 4 个测试集上的主题分类结果

Features	CTS1	CTS2	CTS3	ETS
Characters/Words	4.7	3.6	3.0	3.5
Characters/Words and Basic Semantic Classes	3.9	3.3	1.9	2.9
Characters/Words and Ddeep Ssemantic Classes	2.9	2.2	1.4	2.0

2.2.3 语义槽分类实验结果

为了考察语义重分类的影响, 比较了一遍(one-pass) 语义槽分类(只用词汇上下文特征)和两遍(two-pass) 语义槽分类(加入语义槽重分类)的性能. 表 3 给出了决策表在 4 个测试集上的语义槽分类结果. 从表中可以看出, 语义槽重分类机制对整体性能有一定的改善. 另外在电话语音识别文本的测试集 CTS2 上, 语义槽分类的性能相对较差, 其主要原因是 CTS2 中汉字识别错误率和概念识别错误率分别高达 35.6% 和 41.1%. 如果只考虑语音识别正确的概念, 语义槽分类错误率是 9.2%.

Table 3 The Performance of Decision Lists on Four Test Sets
表 3 四个测试集上的语义槽分类结果

Semantics Slot Classifiers	CTS1	CTS2	CTS3	ETS
Decision List (one pass)	9.1	46.7	5.0	7.0
Decision List (two pass)	8.4	45.6	4.5	6.6

同时, 也对本文提出的口语理解方法与其他方法进行了比较. 首先, 在汉语交通查询领域对本系统和基于规则的稳健句法分析器进行了比较. 该句法分析器的分析算法与本文的预处理器的局部句法分析器相同. 区别在于所使用的规则不同, 对于汉语交通查询领域, 该句法分析器的语义语法是由一个语言学家在 1 个月的时间内定义的, 总共包括 798 条规则(不含用于识别实体名的规则, 如[loc_name] → “外滩”等). 表 4 给出了性能比较结果. 结果表明, 本文方法在主题和语义槽的识别性能都要优于基于规则的稳健句法分析器. 另外, 在这几个包含不少口语现象甚至语音识别错误的数据集上的良好性能也表明本文方法具有较好的稳健性. 其次, 与文献[5] 中的

基于隐马尔科夫模型(HMM)的方法做了比较. 该方法是简单有效的数据驱动方法, 通常作为衡量其他数据驱动方法的基准. 在我们的实现中, 首先也是用本文的局部句法分析器对输入句子进行预处理, 然后利用HMM来分析语义槽概念, 在HMM之后利用SVM进行主题识别, 最后生成语义框架. 表4表明, 本文的方法与HMM方法相比有较大的改进. 最后, 我们列出当前最新的其他口语理解方法的性能做一个大致的参照, 如半自动的语法学习方法^[6-7]和HVS模型^[7-8]等^①. 因为本文的交通信息查询领域与通用的航班旅游信息(air travel information system, ATIS)领域类似, 下面列出这两种方法在ATIS领域上的性能. 结合HVS模型(用于识别语义槽概念)和基于树扩展的朴素贝叶斯分类器(用于主题分类)的方法在ATIS的自然语言测试集上主题分类错误率和语义槽识别的F-measure分别为8.3/8.8%和90.3/91.9%^[8]. 文献[7]中的半自动语法学习方法在ATIS自然语言测试集上的最好结果为5.06%(TER)和7.67%(SER). 另外, HVS模型也在本文所使用的英文DARPA Communicator的数据集上语义槽识别的F-measure大致为88.07%. 这些结果大致表明本文方法与当前最新的其他口语理解方法性能相当.

Table 4 Performance Comparison of Our Method and Other State-of-Art SLU Methods

表4 口语理解方法和其他方法的性能比较 %

Test Sets	Rule-Based Parser		HMM+SVM		SVM+决策表	
	TER	SER	TER	SER	TER	SER
CTS1	6.8	11.6	5.8	11.3	2.9	8.4
CTS2	4.1	47.9	3.9	47.5	2.2	45.6
CTS3	3.0	5.4	2.5	5.5	1.4	4.6
ETS			3.1	8.8	2.0	6.6

3 结论和将来工作

本文介绍了一种新的基于两阶段分类的口语理解方法. 该方法具有以下优点. 首先, 该方法能够稳健地处理口语: 1) 预处理器提供了底层的稳健性; 2) 用统计分类器进行主题分类也具有很好的稳健性; 3) 该方法先分析出输入语句中不同位置的概念所对应的语义槽, 然后把它们联系起来, 这种策略类似于“部分句法分析”, 在处理口语时具有良好的稳健性. 同时, 它保持了理解的深度: 1) 语义槽并不是线性结构, 它继承了领域模型的层次性; 2) 语义槽

重分类的机制力图保证最后生成的语义框架的一致性. 还有, 它能利用主题分类来指导语义槽分类. 最重要的是本方法基本上是数据驱动的, 只需要简单标记的数据, 保证了良好的可移植性. 在汉语交通查询和英语DARPA Communicator两个领域中的实验证实了本方法的有效性. 目前, 两种分类器都是采用有监督方法(supervised)训练(即为训练集中所有句子都标记好语义框架). 为了进一步减少标记训练句子的工作, 我们正在进行利用弱监督(weakly supervised)方法训练这两种分类器的工作.

参 考 文 献

[1] W Ward, S Issar. Recent improvements in the CMU spoken language understanding system [C]. In: Proc of ARPA Workshop on Human Language Technology, San Francisco: Morgan Kaufmann Publishers, 1994. 213-216

[2] S Seneff. TINA: A natural language system for spoken language applications [J]. Computational Linguistics, 1992, 18(1): 61-86

[3] R Pieraccini, E Levin. Stochastic representation of semantic structure for speech understanding [C]. The 2nd European Conf on Speech Communication and Technology (EUROSPEECH), Genova, Italy, 1991

[4] S Miller, R Bobrow, R Ingria, et al. Hidden understanding models of natural language [C]. The 32th Annual Meeting of the Association for Computational Linguistics (ACL), Las Cruces, USA, 1994

[5] Y Wang, A Acero. Grammar learning for spoken language understanding [C]. The 7th IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), Madonna di Campiglio, Italy, 2001

[6] Y Wang, A Acero. Evaluation of spoken language grammar learning in the ATIS domain [C]. The 27th IEEE Int'l Conf on Acoustics, Speech and Signal Processing (ICASSP), Orlando, Florida, 2002

[7] Y He, S Young. Semantic processing using the hidden vector state model [J]. Computer Speech and Language, 2005, 19(1): 85-106

[8] Y He, S Young. A data-driven spoken language understanding system [C]. The 8th IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), Virgin Islands, US 2003

[9] Y Wang, A Acero, C. Chelba, et al. Combination of statistical and rule-based approaches for spoken language understanding [C]. The 7th Int'l Conf on Spoken Language Processing (ICSLP), Denver, Colorado, 2002

[10] C Wutiwattachai, S Furui. Combination of finite state automata and neural network for spoken language understanding [C]. The 8th European Conf on Speech Communication and Technology (EUROSPEECH), Geneva, Switzerland, 2003

① 这两种方法的数据标记方式与本文类似, 只需标记语义框架.

- [11] G Xie, C Zong, B Xu. Approach to robust spoken Chinese language parsing [J]. Journal of Chinese Language and Computing(Singapore), 2004, 14(1): 5-19 (in Chinese)
(解国栋, 宗成庆, 徐波. 鲁棒的汉语口语解析方法研究[J]. 汉语语言与计算学报(新加坡), 2004, 14(1): 5-19)
- [12] Y Wang. A robust parser for spoken language understanding [C]. The 6th European Conf on Speech Communication and Technology (EUROSPEECH), Budapest, Hungary, 1999
- [13] W Wu, J Duan, R Lu, *et al.* Embedded machine learning systems for robust spoken language parsing [C]. The 3rd IEEE Int'l Conf on Natural Language Processing and Knowledge Engineering(NLP-KE), Wuhan, 2005
- [14] R Rivest. Learning decision lists [J]. Machine Learning, 1987, 2(3): 229-246
- [15] D Yarowsky. Decision lists for lexical ambiguity resolution: application to accent restoration in Spanish and French [C]. The 32nd Annual Meeting of the Association for Computational Linguistics(ACL), Las Cruces, USA, 1994
- [16] R Golding. A Bayesian hybrid method for context-sensitive spelling correction [C]. The 3rd Workshop on Very Large Corpora, Boston, MA, 1995
- [17] CUData. DA RPA communicator travel data [OL]. <http://communicator.colorado.edu/phoenix>, 2004
- [18] CUPheonix. CU phoenix parser [OL]. <http://communicator.colorado.edu/phoenix>, 2003



Wu Weilin born in 1978. Ph. D. from Shanghai Jiao Tong University. His main research interests include natural language processing, spoken dialogue system, information retrieval and search engine, *etc.*

吴尉林, 1978 年生, 博士, 主要研究方向为自然语言处理、口语对话系统、信息检索和搜索引擎等。



Lu Ruzhan born in 1940. Professor and Ph. D. supervisor in Shanghai Jiao Tong University. His main research interests include Chinese corpus processing, Chinese intensional logic model and its applications, intelligent retrieval based on

conceptual intensions, dialogue systems and semantic Web.

陆汝占, 1940 年生, 教授, 博士生导师, 主要研究方向为汉语语料库加工技术、汉语内涵逻辑模型及其应用、基于概念内涵的智能检索、对话系统、语义 Web。



Duan Jianyong born in 1978. Ph. D. candidate from Shanghai Jiao Tong University. His main research interests include natural language processing, information extraction, machine learning and bioinformatics, *etc.*

段建勇, 1978 年生, 博士研究生, 主要研究方向为自然语言处理、信息抽取、机器学习、生物信息学等。



Liu Hui born in 1981. Ph. D. candidate in Shanghai Jiao Tong University. His main research interests reside in natural language processing including natural language parsing, semantic analysis, information retrieval, *etc.*

刘 慧, 1981 年生, 博士研究生, 主要研究方向为自然语言处理、句法分析、语义分析、智能检索等。



Gao Feng born in 1969. Ph. D. candidate in Shanghai Jiao Tong University. His main research interests include Chinese lexical semantics and its applications, intelligent retrieval based on concepts and dialogue systems.

高 峰, 1969 年生, 博士研究生, 主要研究方向为汉语词汇语义及应用、基于概念的智能检索和对话系统。



Chen Yuquan born in 1968. Associate professor in Shanghai Jiao Tong University. His main research interests include Chinese intensional logic model and its applications, intelligent retrieval based on concepts, Chinese corpus processing, *etc.*

陈玉泉, 1968 年生, 副教授, 主要研究方向为汉语内涵逻辑模型及其应用、基于概念内涵的智能检索、汉语语料库加工等。

Research Background

Spoken dialogue systems have been attracting extensive interest from the research and industrial communities since they provide a natural interface between human and the computer. Many spoken dialogue systems in a variety of application domains and languages have appeared in recent years. SLU is one of the key components in spoken dialogue systems. Its task is to identify the user's goal and extract from the input utterance the information needed to complete the query. Although it is focused only on limited domains, spoken language understanding is faced with many challenges. One of challenges is robustness since the speech recognizer inevitably makes errors and spoken language is plagued with a large set of spontaneous speech phenomena. Another challenge is portability, which is related to how flexible new SLU components for new applications or new languages can be built quickly at a reasonable cost. The traditional rule-based SLU approaches require the linguistic experts to

handcraft the domain-specific grammar for parsing, which is time-consuming and laboursome. This paper proposes a new spoken language understanding approach based on two-stage classification. Firstly, the topic classifier is used to identify the topic of an input utterance. Then, with the restriction of recognized target topic, the semantic slot classifiers are trained to extract the corresponding slot-value pairs. Our SLU approach is mainly data-driven and requires only minimally annotated corpus for training whilst retaining the understanding robustness and deepness for spoken language. Experiments have been conducted in the Chinese public transportation information inquiry domain and the English DARPA Communicator domain. The good performance demonstrates the viability of the proposed approach. In order to further reduce the human efforts, we are working on the weakly supervised training techniques for our two-staged classification based SLU framework. Our work is supported by the National Natural Science Foundation of China (NSFC, No. 60496326) and 863 Project of China (No. 2001AA114210-11).

CACIS ° 2008 学术会议征文通知(第一轮)

由中国仪器仪表学会和中国系统仿真学会主办的全国第 19 届计算机技术与应用学术会议(CACIS ° 2008)将于2008年7月15~20日在四川乐山举行. 会议一如既往, 由中国科学技术大学出版社正式出版《计算机技术与应用进展 ° 2008》论文集, 符合学报要求的参会论文将分别在会后由《仪器仪表学报》、《计算机辅助设计与图形学学报》、《系统仿真学报》、《工程图学学报》等正刊出版, 欢迎新老朋友投稿、参会.

征文主题(包括但不限于)

- ° 计算机仿真与可视化
- ° 人工智能与算法
- ° 数据库与信息系统
- ° 图像与多媒体技术
- ° 网络与信息安全
- ° 计算机辅助设计与图形学
- ° 软件工程与软件设计
- ° 仪器仪表与检测控制
- ° 电子与通讯技术
- ° 其他相关理论与技术

投稿网址: <http://www.cacis.com.cn>

重要日期

征文起始日期: 2008年1月1日

征文截止日期: 2008年5月1日

录用通知日期: 论文提交后 15 个工作日之内给予回复

会议重要日期: 2008年7月15日报到, 7月16~17日会议

业务咨询

通信地址: 230009 合肥工业大学 52 信箱 微型计算机应用学会

联系电话: 0551-2901380(杨孙梅)、2901377(余烨)、2901701(刘晓平)

传 真: 0551-2904642

联系 Email: cacis@ah163.com

学会网址: <http://cacis.hfut.edu.cn>