

## 口语对话中的代词指代消解\*

费仲超<sup>1,2+</sup>, 周雅倩<sup>1</sup>, 黄萱菁<sup>1</sup>, 吴立德<sup>1</sup>

<sup>1</sup>(复旦大学 计算机科学技术学院, 上海 200433)

<sup>2</sup>(上海贝尔股份有限公司 产品线战略及技术领先部, 上海 201206)

### Pronoun Resolution in Spoken Dialog

FEI Zhong-Chao<sup>1,2+</sup>, ZHOU Ya-Qian<sup>1</sup>, HUANG Xuan-Jing<sup>1</sup>, WU Li-De<sup>1</sup>

<sup>1</sup>(School of Computer Science, Fudan University, Shanghai 200433, China)

<sup>2</sup>(Portfolio Strategy and Technology Leadership CTO Group, Alcatel-Lucent Shanghai Bell, Shanghai 200433, China)

+ Corresponding author: E-mail: zcfei@fudan.edu.cn

Fei ZC, Zhou YQ, Huang XJ, Wu LD. Pronoun resolution in spoken dialog. *Journal of Software*, 2011, 22(2): 233-244. <http://www.jos.org.cn/1000-9825/3720.htm>

**Abstract:** This paper presents a two-stage pronoun resolution algorithm. It does not need to clean the testing corpus and predefine patterns manually. In the first stage of the algorithm, some new features and machine learning methods are used to classify pronouns into anaphoric and non-anaphoric ones. In the second stage, these two kinds of pronouns are resolved respectively. For the anaphoric ones, some methods are presented to extract distance, syntactic, and semantic features etc. For the non-anaphoric ones, the Right Frontier Rule is improved to do the resolution work. While testing the corpus published by Byron in 2004, this algorithm achieves a precision of 77.0% and a recall of 66.0%. Compared with the work of Byron, the algorithm is fully automatic, and the results are much better.

**Key words:** pronoun resolution; spoken dialog understanding; pronoun classification

**摘要:** 提出一套分为两步的代词指代消解算法, 算法不需要人工清洗语料及预定义规则. 算法第 1 步采用一些新特征和机器学习算法对名词性指代代词和非名词性指代(non-anaphoric)代词分类, 第 2 步分别对两类代词进行消解. 针对名词性代词指代消解, 提出了适用于口语对话的特征抽取及表示方法, 如代词和候选先行词的距离、语法、语义等的抽取和表示方法, 然后通过综合这些特征来选择先行词. 针对非名词性指代, 将右边界规则(right frontier rule)改进为可以在口语对话中自动抽取的形式, 并根据该规则选择先行项. 在 Byron 于 2004 年发布的语料上测试, 消解正确率达到 77.0%, 召回率达到 66.0%. 与 Byron 的工作相比, 该方法在保证系统能够自动完成的同时还提高了消解性能.

**关键词:** 代词指代消解; 口语对话理解; 代词分类

**中图法分类号:** TP391      **文献标识码:** A

口语对话(spoken dialog)是指语音交互记录. 随着自动口语对话系统研究的发展<sup>[1]</sup>, 口语对话理解得到了比

\* 基金项目: 国家自然科学基金(60503070, 60673038); 上海市科委科研计划(08511500302)

收稿时间: 2009-02-20; 定稿时间: 2009-08-12

以往更多的关注.ACL(<http://www.aclweb.org/>)等计算语言学相关重要会议近些年均有口语对话相关议题,还专门设立了研究对话的 SIGDIAL(<http://www.sigdial.org/info.php>)等专题会议.而对话中的代词指代消解则是全面理解口语对话不可缺少的一部分.

与书面语相比,口语对话有很多特殊之处.

由于口语对话不能像书面语那样反复修改完善,说话过程中的重复、修改、补充、停顿或者被插话中断等信息全部被记录下来,因此,对话中存在大量不流利成分,如:

例 1: s: well or we could send [engine e three]<sub>1</sub> I mean start [it]<sub>1</sub> before the one with the boxcars I mean would because the oranges are getting there in plenty of time

u: okay

u : so if [we start [it]<sub>1</sub> before]<sub>2</sub> is that is that will [that]<sub>2</sub> change the time

s: um yes so then we would get to Dansville at three a.m.

对话记录通常还有一个特点:在语音交互的条件下,输出的对话文本一般缺少标点符号,如例 1 就是这样.这就导致对话结构更加不清晰.

对话中的不流利成分给口语对话的语法、语义自动分析带来了困难.而语法、语义信息是指代消解工作中常被用到的重要特征.

针对口语对话文本结构不清晰的问题,不少以往相关消解工作都在手工清洗过的语料上进行测试,有的还利用了人工标记在测试语料上的词法、语法等信息<sup>[2,3]</sup>.部分算法还需要手工选词、编写模板等工作<sup>[4-6]</sup>.由于大量的手工干预因素,这些算法的效率及性能会受到影响,无法在自然产生(spontaneous)的对话语料上评测,也很难直接应用到自动对话系统中.

除了大量不流利成分造成对话中代词指代消解困难增加之外,对话中的代词指代另一个显著特点是名词性(anaphoric)指代和非名词性(non-anaphoric)指代的比例与书面语有所不同<sup>[7]</sup>.名词性指代的定义为指示词指向的先行项为名词性的<sup>[8]</sup>,在部分工作中也称为指向实体先行对象(individual/concrete entity).其他指向抽象概念,如事件、状态、事实等的指代则称为非名词性指代,有的工作也称其为指向话语(discourse-deictic)的指代.如例 1 中的“it”是名词性指代,“that”是非名词性指代.不少统计发现,非名词性指代的比例在口语中比在书面语中高很多.在 SwitchBoard 对话语料上,非名词性指代的比例高达 55%左右<sup>[2]</sup>;在部分 Trains-93 对话语料上,有一半左右的代词不是名词性指代<sup>[4]</sup>;在对话语料 Monroe 上,有 40%左右的代词不指向名词性先行词<sup>[5]</sup>.而在书面语中,如在 BNC 语料中,只有 22%的“it”不是名词性指代<sup>[9]</sup>,在 ACE 语料中,只有 20%左右的“it”不是名词性指代<sup>[10]</sup>.因此,消解对话中代词时,非名词性指代的识别和消解尤为重要.

在我们的研究工作中,为了能使代词指代消解算法应用于真实口语对话环境,本文通过改进前人相关工作,使特征抽取过程,算法消解过程均可以自动进行.在判断代词类别时,我们提出了基于依存关系的特征抽取和表示方法,并考虑了代词序列的马尔可夫性.在消解名词性指代时,本文对距离、语法、语义等特征均采用了新的抽取和表示方法,并在消解过程中尝试了两种不同的方法来选择候选先行词.在消解非名词性指代时,我们考虑了被以往工作忽略的一些常见现象,改进了右边界规则(right frontier rule)<sup>[11]</sup>.

本文第 1 节为主要相关工作.第 2 节分析传统上指代消解工作中所使用的特征及部分在口语对话上相关工作所用的特征在消解口语对话中代词指代时存在的问题.第 3 节是本文提出的代词指代消解算法.此节分为 3 部分,第 1 部分是代词的自动分类算法,第 2 部分为消解名词性指代代词的算法,第 3 部分为消解非名词性指代代词的算法.第 4 节是实验及其分析.第 5 节为总结与展望.

## 1 口语对话代词指代消解相关工作

早期关于口语对话中指代消解工作之一是 Eckert 等人在部分 SwitchBoard 对话语料上进行的<sup>[2]</sup>.他们强调了非名词性指代在对话中的显著性,并给出了基于经验规则的消解算法.他们给一系列收集到的词赋予 I-Incompatibility(不能和名词性概念搭配)属性和 A-Incompatibility(不能和抽象概念搭配)属性,然后根据代词是

否与具有这样属性的词搭配来判断代词分为名词性指代还是非名词性指代.他们利用了语料中标注的句子结构、短语结构和对话单元结构等信息来完成消解.该算法只经过了人工评测,并没有实现.Strube 等人在 Switchboard 对话语料上对文献[2]的工作做了改进<sup>[3]</sup>,不再采用经验规则,算法考虑了词法、语法、距离等特征,然后用决策树进行消解.该算法同样也是基于语料中手工标注的词法、语法等信息.Byron 在其博士学位论文中详细介绍了针对口语对话的代词指代消解系统 PRORA<sup>[4]</sup>,该系统可以消解对话中名词性指代和非名词性指代代词.算法中,先行词的选择是通过人工定义在上文中的搜索规则实现的.通过在部分口语对话语料 Trains-93 上测试,该算法可以达到 72%的正确率.该实验所使用的语料没有利用手工标注信息,但是语料经过清洗和部分手工修改清除了文中的不流利部分.Tetreault 等人在口语对话语料 Monroe 上进行了指代消解工作<sup>[5]</sup>.与其他相关工作相比,该工作利用了较丰富的语义知识来对候选先行词进行约束.该算法把语义信息加入到 LRC(left-right centering)<sup>[12]</sup>算法中实现对候选先行词的过滤,该工作所利用的语义信息需要事先手工定义.测试语料同样通过手工清洗把其中的不流利部分去除了.Müller 实现了一套可以消解“it”,“this”和“that”的指代消解系统,该消解系统在保留了不流利信息的多方对话(multi-party)语料 ICSI 上进行<sup>[13]</sup>,所用的语料在对话录音转化成文本(transcription)后,人工加入了标点符号.该系统利用性、数、人称、语法一致性约束等特征,并通过计算互信息(mutual information)来判断候选先行词和代词的语义一致性.算法的特征抽取过程及消解过程是自动的,但消解性能较低.最好的评测结果在消解名词性指代时为 21%,消解非名词性指代为 12%,两种指代共同评测的 *F-value* 只有 18%.

## 2 口语对话中代词指代消解特征分析

本节将详细分析传统上指代消解工作中所使用的特征及部分在以往口语对话相关工作中所用的特征在消解口语对话中代词指代时主要存在的问题.

### 2.1 代词分类特征

本文中,代词分类指的是判断代词是名词性指代还是非名词性指代.

在早期的代词分类相关工作中,一般采用人工选词以及写规则的方法,如文献[2]的工作.在此后的一些相关工作中,不少分类算法采用了机器学习的方法<sup>[9,14-17]</sup>.这些方法虽然可以自动训练,但其所使用的特征依然依赖于大量手工干预工作.

例如,在文献[15]的工作中采用了 38 个手工编写的模板,部分如下所示:

obj verb:代词是否为宾语;

dist to next adj:到下一个形容词的距离;

oblique:是否跟了介词;

seem list:主语后面的动词是否为 seem,appear,look,mean,happen,sound.

这些特征包括了手工选择的模板及词汇,算法的可扩展性受到一定限制.

### 2.2 名词性代词指代消解特征

#### 2.2.1 性别、单复数、人称的一致性

在部分相关工作中提到,在对话中,性别、单复数、人称有时会不一致<sup>[5,6]</sup>.通过在 Trains-93 语料上分析,我们也发现了类似的现象,例如:

例 2:s: okay taking the two [boxcars]<sub>1</sub> okay so it gets to coming at two a.m.

u: and then we have to fill [it]<sub>1</sub> with oranges

例 3:u: alright so uh alright I'm gonna have [engine e two]<sub>2</sub> go through then and pick up the um the boxcars in dansville

s: okay

u: like we planned on doing except [he]<sub>2</sub>'s not gonna take the original two boxcars.

例 2 第 2 句中,单数代词“it”指向了第 1 句中的复数先行词“boxcars”.在例 3 中,指向男性性别的代词“he”指向了中性对象“engine e two”.

可以看出,性别、单复数、人称的一致性在口语对话中不能作为判断是否存在指代关系的决定性因素.

### 2.2.2 距离(recency)

通常,指代消解算法考虑的距离是句子(sentences)之间的距离或者段落(paragraphs)之间的距离<sup>[18]</sup>.但是在对话中,一个话语(utterance,一个对话角色一次连续的说话)中会包含没有明显分隔标记的多句话,也可能一句话分多个话语才能说完,对话中也没有段落,因此考虑句子或段落的距离会比较困难.

### 2.2.3 语法

语法特征很早就被用在指代消解相关工作中<sup>[18-23]</sup>.如考虑不同语法角色里的名词被指代到的优先级<sup>[19]</sup>、语法角色的平行关系等<sup>[18]</sup>.

但在口语对话中,由于不流利成分的存在,依赖于语法结构的指代消解方法往往难以取得好的效果.比如,在以往实验中取得较好效果的 LRC 算法<sup>[12]</sup>,在口语对话环境下无法很好地获得算法中所需的句子之间“转移”、“持续”等信息.在文献[4]的测试中,采用 LRC 算法在 Trains-93 语料上只能取得 37%的正确率<sup>[4]</sup>.因此,部分针对口语对话的消解工作不再考虑整段话语的语法结构<sup>[5,21]</sup>.

### 2.2.4 语义

近些年,基于语义一致性的消解方法取得了不错的效果<sup>[5,24-26]</sup>.这些方法通过判断候选先行词和待消解项的语义一致性,即通过待消解项和候选先行词所属的语法结构和所搭配的词来判断候选先行词是否符合待消解项的语义环境.如,“把它们榨成果汁”,“它们”所指代的对象从语义上判断应该是可以作为“榨”的宾语,如果候选先行词同时出现了“橙子”和“集装箱”,那么“橙子”将被选择.

判断语义一致性时常见有两种方法,一种是基于知识库的方法<sup>[4,5]</sup>,通过事先定义词语之间的搭配规则来判断代词和候选先行词间的语义一致性.该方法一般需要手工收集词和短语.另一种方法是通过对大规模语料的统计,抽取出可以判断语义一致性的搭配特征模板<sup>[25,26]</sup>,或者通过统计互信息来判断语义一致性<sup>[13]</sup>.但是,这些基于统计的方法普遍存在数据稀疏的问题<sup>[24]</sup>.

## 2.3 右边界规则

由于非名词性指代所指向的一般是动词短语或者整个句子,针对名词性指代的消解方法所考虑的特征一般无法直接使用.针对非名词性指代,Webber 提出了右边界规则<sup>[11]</sup>:如果在事件 A 之后又提及了事件 B,那么事件 B 之后出现的指向抽象概念的代词对事件 A 来说将不再是可指向的(not accessible).该方法被 Eckert 等人扩展后应用在口语对话的消解工作中<sup>[2]</sup>,并被其他工作多次采用<sup>[13,27]</sup>,步骤如下:

- (1) 首先在本句中找到包含该指示词的从句左边紧邻的从句作为先行对象;
- (2) 如果在本句中找不到,找前面最近的句子中最右边的主句(同时包含该主句中的从句)作为先行对象;
- (3) 如果前面最近的句子不完整,那么将该不完整的句子最右边的整句包含在先行对象中.

可以看出,右边界规则依赖于对句法结构的精确分析,在口语对话中实际上难以自动做到.

## 3 代词消解算法

本节将详细介绍本文的算法.我们把消解过程分为两步:首先是代词分类,识别出代词是名词性指代还是非名词性指代;然后分别消解这两类指代.

### 3.1 代词分类

分析发现,影响代词类型的特征主要包括位置、词法、语法、以及语义等信息.如例 4 中,代词“it”在不同的位置关系、语法结构、语义搭配等上下文环境下属于不同的类型.

例 4: Go to Corning load them up and then take it to Bath.

How long will it take for the two engines and two boxcars to get to Corning.

You'd need to convert and that takes an hour so it'd end up being seven hours on the way back.

在第 1 句中,“it”为名词性指代,它和动词“take”组成了“动宾关系”。在第 2 句中,“it”为非名词性指代,虽然他也依赖于动词“take”,但依赖关系类型为“主谓关系”。第 3 句中,“it”是非名词性指代,它和所依赖的词组成了“主谓关系”,依赖的词是“end”。

依存语法所包含的信息将作为我们判断代词类型的特征之一,这些特征共分为 5 类,见表 1 中的特征 1~特征 5。另外,我们把代词周围的词及其词性也作为特征,以此作为对自动语法分析之后得到的依赖关系三元组部分错误结果的弥补,见表 1 中的特征 6(本文所采用语法解析器均为 Stanford Parser Version of 1.6 版,采用的词性标注器为 stanford-postagger-2008-09-28 版)。

**Table 1** Features for pronouns classification

表 1 代词分类特征

| No. | Feature   | Explanation  |
|-----|---|--|
| 1   | The pronoun   | Feature 1 is the pronoun itself. If we only identify one certain kind of pronoun, such as only identifying “it”, this feature will be neglected  |
| 2   | Related words   | The words include the parents, grandparents and brothers of the pronoun which can be related to the pronoun in two layers. If the related words are verbs, the related words will be replaced by its verb stem |
| 3   | Paternity   | The pronoun is the head or the modifier in the dependency relation.  |
| 4   | Type of dependency relation                           | Such as subject-predicate relationship, verb-object relationship   |
| 5   | Depth of dependency relation                          | How many layers between the pronoun and the related word. In our experiment, the maximal depth is limited to the second layer.   |
| 6   | The words and their part-of-speech around the pronoun | In our experiment, we use the first three words before the pronoun and the first three words after the pronoun.  |
| 7   | Transition feature                                    | The type of the previous pronoun and the following pronoun.  |

比如,句子“Take it to Bath”。首先,我们利用依存语法解析器解析得到如下依赖关系三元组:

doj(Take,it), prep(Take,to), poj(to,Bath).

这里的 doj,prep,poj 指的是直接宾语(direct object)、介词短语(preposition)、介宾短语(preposition object),详情参考 Stanford Parser 的依赖关系标签说明(<http://nlp.stanford.edu/software/lex-parser.shtml>)。

从依赖关系三元组及原句中抽取到的特征如下:

特征 1. 代词“it”。

特征 2. 依赖词有两个,“it”的父亲“Take”以及兄弟“to”。

特征 3. it 在 doj(Take,it)中是儿子。

特征 4. doj(Take,it)中,依赖关系类型是“doj”,prep(Take,to)中依赖关系类型是“prep”。

特征 5. doj(Take,it)的依赖深度是 1,prep(Take,to)的依赖深度是 2。

特征 6. Take/VB, to/TO, Bath/NNP。

在对话中,还会遇到话语中没有足够的单词的情况,例如:

例 5:s: nine a.m. uh wait a second uh also loading the let's see let me see it's gonna take an hour to unload should be noon

u: noon

s: is it

例 5 的第 3 句中只有两个词,无论是依赖关系三元组还是代词周围的词,都不足以表示该句中代词“it”的分类特征。针对这种情况,本文将尝试通过考虑代词序列之间的马尔科夫性,即通过前一个代词和后一个代词的类别来分类,该特征作为特征 7。

## 3.2 名词性指代代词消解

### 3.2.1 语法约束

我们在基于口语对话生语料和自动语法分析结果的基础上分析了语法结构对指代消解的影响。分析发现,

位于介词短语中的名词比位于其他短语中的名词被指代的概率要小.而位于主语、直接宾语、间接宾语中的名词短语之间被指代到的概率<sup>[19]</sup>没有明显的差别.我们同时也发现,因为平行语法角色的判断在口语对话生语料上自动抽取的质量难以保证,考虑平行语法角色<sup>[18]</sup>对消解没有明显帮助.

因此,消解时,我们解析出句子中的名词短语、介词短语.其中,名词短语将被作为名词性指代的候选先行词.候选先行词是否出现在介词短语里,将被作为一个特征.

约束理论(binding theory)同样也是一种语法约束规范<sup>[28]</sup>.该理论限制了在特定的局部语法结构下存在某些“不能互相指代”或者“必须互相指代”的情形.例如:

例 6: That engine would make it.

例 6 中的“engine”和“it”被同一个谓词“make”绑定(binding),“it”通常不会指向“engine”.“engine”和“it”属于“不能相互指代”的绑定关系.“必须相互指代”的情形一般出现在反身代词上.

该理论在指代消解工作中多次被用到<sup>[5,13,18,29]</sup>,本文也将采用.我们通过语法解析器输出的结果来判断代词和候选先行词是否符合“不能相互指代”或“必须指代”的约束规则.

### 3.2.2 语义约束

在抽取语义约束模板时,为了避免因在大规模语料上统计而造成数据稀疏,同时避免手工选词、选短语模板等工作,本文采用了在语料的训练部分上统计提及率(mention count,被指代的频率)<sup>[30]</sup>高的词,由高到低选择覆盖到前 90%被指代的词,然后自动抽取这些词的语义约束模板.

在抽取语义约束模板时,本文提出了新的基于依赖关系三元组的方法,如例 7 所示:

例 7: pick up oranges to put in the boxcars.

语法解析器输出的三元组结果如下:

prt(pick,up), dobj(pick,oranges), aux(put,to), xcomp(pick,put), prep(put,in), det(boxcars,the), pobj(in,boxcars).

在这些依赖关系中,关系三元组 dobj(pick,oranges)和 pobj(in,boxcars)与提及率统计结果中能覆盖到前 90%名词性指代的先行词的词“oranges”和“boxcars”相关.这两个关系将被抽取出来.对全部训练语料抽取一遍之后,对这些抽出的关系再进行一次过滤,把区分度小的依赖关系三元组过滤掉.过滤方法如下:

设  $F_1$  为依赖关系三元组 dobj(pick,oranges)出现的频率, $F_2$  为依赖关系三元组 dobj(pick,\*)出现的频率,“\*”表示任意的词.如果( $F_1/F_2$ )的值大于给定阈值,则这个关系三元组将被保留.该阈值通过在训练语料上人工调参获得,本文为 0.1.

在利用语义特征消解时,如果需要消解的代词在它所在的句子中有这样一个依赖关系 dobj(pick,pronoun), pronoun 表示需要消解的代词,而从训练语料中抽取的依赖关系三元组过滤后的集合中包含 dobj(pick,orange).该依赖关系的关系类型及所依赖的词与需要消解的代词所在的依赖关系都相同,在这样的条件下,如果候选先行词中刚好包含“orange”,那么“orange”将有较大的概率被指代到.

### 3.2.3 其他特征

除了语法、语义特征,我们也考虑了单复数、距离等特征.单复数特征为候选先行词的单复数和代词是否一致.距离特征不再是考虑代词和候选先行词之间有多少个句子,而是考虑它们之间有多少话语.因为对话中需要消解的代词一般都是第三人称的,故本文没考虑人称特征.另外,第 3.2.2 节提到的提及率也将被作为特征之一.

### 3.2.4 先行词选择

利用第 3.2.1 节~第 3.2.3 节介绍的特征,本文分别采用两种方法来选择先行词.一种是把先行词选择转化为两类分类问题<sup>[13,18,31,32]</sup>,即判断代词和候选先行词对是否构成指代关系,然后采用机器学习的方法来分类.该方法会遇到训练样例不平横等问题.为了解决该问题,我们提出了另一种基于不同特征权重累加的方法,消解时通过比较候选先行词多个特征的线性加权重之和的大小来选择最有可能被指代的那个.

**方法 1(Algorithm 1).** 在将消解问题转化为分类问题时,存在指代关系的“名词-代词”对组成正例,不存在的则组成反例.由于绝大多数“名词-代词”对不存在指代关系,构建训练样例时反例数一般都会远大于正例数.正反

样例构建不好,将会使得分类精度因受正反样例不平衡或者样例覆盖面不全的影响而下降。

本文在构建反例时,采用从训练语料中抽取代词前 36 个话语中的所有不存在指代关系的名词短语和代词组成对作为反例,36 是训练语料上最远的指代距离。这种情况下,正反样例不平衡问题较为严重,但由于覆盖面更全,测试发现比只把先行词和代词之间的名词短语和代词组成对作为反例<sup>[18]</sup>的方法效果要好。然后,我们分别抽取候选先行词在第 3.2 节所述的特征,再把这些特征和代词及其词性(人称代词或指示代词)放在一起作为全部特征。

其中,单复数一致性、语法特征中的候选先行词是否在介词短语中出现、代词和候选先行词是否符合约束理论都是布尔量特征。

距离特征由于采用具体的距离数据会造成数据稀疏,本文将距离分为 0~3,4~10,11~16 和大于 16 的 4 个区间。

提及率也作区间划分处理,按照被提及的频率把候选先行词分在如下 4 个区间,分别为被指代次数为 0、被指代次数为 1~2、被指代次数 3~10、被指代次数大于 10。

语义特征根据代词所在的依赖关系是否能与第 3.2.2 节介绍的方法所抽取的依赖关系模板匹配。匹配的话根据  $F_1/F_2$  的值分为 4 个区间,分别为 0.1~0.2,0.2~0.3,0.3~0.4,大于 0.4。

这些区间的划分均根据经验在训练语料上调整获得(由于能获得的公开语料量较少,本文没有从语料中独立分出调参(developing)部分,只分了训练和测试部分。其中,训练语料 420K,测试语料 47K,详见第 4.1 节)。

我们采用条件最大熵分类器。被分类器判断为存在指代关系的候选先行词中概率最大的那个作为最终的消解结果。如果多个概率相同,那么,在同一个句子中出现的按照从左到右的顺序选择<sup>[4]</sup>,不同句子中出现的按照距离由近到远的顺序选择。

**方法 2(Algorithm 2).** 我们根据候选先行词不同特征所赋予的经验权重之和排序,权重最高的作为最有可能被指代的先行词。权重及赋予方法如下:

单复数一致性权重  $Nu$ ,一致为 1,不一致为-1。

距离特征权重  $R$ , $R$  为代词和候选先行词之间有多少话语。

语法特征权重  $Sy$ ,该特征表示候选先行词是否在介词短语中,在为-1,不在为 0。如果候选先行词和代词符合不能指代或者必须指代的约束规则,那么将直接把该候选先行词排除或者选择为最终答案。

提及率  $Mc$ 。

语义特征权重  $Se$ ,不能与事先抽取的依赖关系模板匹配为 0,能匹配为  $F_1/F_2$  的大小。

候选先行词总的权重  $Weight = \lambda_1 Nu + \lambda_2 R + \lambda_3 Sy + \lambda_4 Mc + \lambda_5 Se$ ,权重最大的候选先行词作为消解结果。

这些特征权重的参数  $\lambda_i (1 \leq i \leq 5)$  根据经验赋予初值,然后通过训练语料手工调整获得。如果多个候选先行词权重相同,则选择方法与方法 1 相同。

### 3.3 非名词性指代代词消解

由于右边界规则需要对句子语法成分进行明确的划分,而在真实口语对话语料上区分出主句、从句以及判断句子的完整性目前还较为困难,而且对话经常会出现如下的谓语省略情况:

例 8: *u*: and they now they have to go to Corning

*s*: okay so they'll get to Corning at at five a.m.

*u*: [then to Dansville]<sub>1</sub>

*s*: [that]<sub>1</sub>'s six a.m.

*u*: and [then to Avon]<sub>2</sub>

*s*: so [that]<sub>2</sub>'ll be nine sorry nine p.m.

例 8 中,[that]<sub>1</sub>,[that]<sub>2</sub> 所指向的对象均省略成了介词短语,不是句子或者动词短语。

针对口语对话中的以上情况,本文把右边界规则简化为查找前文中出现的最近一个动词短语,同时考虑省略的话语。我们把前文出现的动词短语和只包含介词短语的话语放在候选被指代抽象对象列表中,如果对话中

出现了被分类为指向抽象概念的代词,那么从候选列表中取出最近一个作为消解结果.

## 4 实验

### 4.1 语料及其预处理

口语对话指代消解相关研究所用的语料目前还不统一,有多个语料被不同的工作使用,如 ICSI<sup>[13]</sup>, GENIA<sup>[17]</sup>, Switchboard<sup>[3,7]</sup>, Monroe<sup>[5]</sup>, Trains-91<sup>[33]</sup>, Trains-93<sup>[4,34]</sup>等,本文采用了 Trains-93(<http://www.cs.rochester.edu/research/cisd/resources/trains.html>). 该语料为对话记录生语料,语料没有标点和大小写,保留了对话中不流利成分.本文所用的测试语料部分是文献[4]中标注的 10 篇 Trains-93 语料(Trains-93 共 98 篇对话,文献[4]中标注了其中 10 篇,见 <ftp://ftp.cs.rochester.edu/pub/papers/ai/03.TR703-BUR-pron-annotation.shtml>)<sup>[4]</sup>,其余的 88 篇我们标注后作为训练语料(标注的语料发布在 [http://www.jkx.fudan.edu.cn/mcwil/~zcfai/ar/Trains-93\\_AR\\_tagged.rar](http://www.jkx.fudan.edu.cn/mcwil/~zcfai/ar/Trains-93_AR_tagged.rar)).

为了提高语法分析的质量,我们首先对语料进行如下可以自动进行的预处理:去除语料中的语气词<sup>[13]</sup>,例如“un”,“um”等;把出现在每一个话语句首的“well”,“and”,“then”去掉;在“okay”,“no”,“yes(yeap)”和“alright”的前后加上逗号;把话语中重复的词去除.我们对训练语料和测试语料采用了相同的预处理.

表 2 为测试语料上的代词统计情况.

**Table 2** Statistic about the pronouns in the annotated Trains-93 in Ref.[4]

表 2 文献[4]中标注的 Trains-93 测试语料上代词分类统计结果

| Pronoun    | Nominal-Pronoun |       | Non-Nominal-Pronoun |       | Total |
|------------|-----------------|-------|---------------------|-------|-------|
| This       | 0               | 0%    | 1                   | 100%  | 1     |
| That       | 20              | 20.4% | 78                  | 79.6% | 98    |
| It         | 45              | 69.2% | 20                  | 30.8% | 65    |
| They       | 16              | 84.2% | 3                   | 15.8% | 19    |
| Them       | 10              | 76.9% | 3                   | 23.1% | 13    |
| These      | 0               | -     | 0                   | -     | 0     |
| Those      | 4               | 57.1% | 3                   | 42.9% | 7     |
| Themselves | 1               | 100%  | 0                   | 0     | 1     |
| Total      | 96              | 46.8% | 108                 | 53.2% | 204   |

### 4.2 实验结果

表 3 为代词分类的结果.因为代词分类目前没有可以互相比较的测试语料,本文故在全部 Trains-93 语料上用 10-fold 交叉检验,并与采用前人工作所用特征取得的结果作对比.

**Table 3** Results of pronouns classification (accuracy)

表 3 代词分类结果(accuracy)

| Feature                                   | Classifier | IT (%) | ALL (%) |
|---|------------|--------|---------|
| <i>C_Feature</i>                          | CME        | 88.5   | -       |
|   | SVM        | 87.1   | -       |
| <i>M_Feature</i>                          | CME        | 86.6   | -       |
|   | SVM        | 89.5   | -       |
| <i>D_Feature</i>                          | CME        | 92.5   | 91.9    |
|   | CRF        | 92.2   | 92.6    |
|   | SVM        | 92.1   | 91.3    |
| <i>W_feature</i>                          | CME        | 89.2   | 90.3    |
|   | CRF        | 89.3   | 91.9    |
|   | SVM        | 89.2   | 90.0    |
| <i>D_Feature</i><br>+<br><i>W_feature</i> | CME        | 93.9   | 93.3    |
|   | CRF        | 93.5   | 93.3    |
|   | SVM        | 93.1   | 92.6    |

表 3 中,*C\_Feature* 是文献[9,17]分类时所使用的基于手工规则的特征,*M\_Feature* 是第 2.1 节提及的文献[15]分类时使用的特征.*D\_feature* 是表 1 中特征 1~特征 5.*W\_feature* 是表 1 中特征 6.“IT”指的是只消解单词“it”,因为作为 Baseline 的 *C\_Feature* 和 *M\_Feature* 都只对单词“it”分类,所以表中采用这两类特征时只对“it”做了分



类.ALL 指的是对“it”,“that”,“this”,“these”,“those”,“they”,“them”同时分类.训练语料和测试语料中的“themselves”都只有名词性指代,本文没有对其分类.

分类器我们分别采用了条件最大熵、线性核支持向量机(SVM)以及一阶条件随机场(first-order-CRF).特征7只在采用条件随机场时使用.

从表 3 可以看出,本文所用全自动抽取的特征与以往的基于手工规则的特征相比,分类精度(accuracy)有了明显的提高.我们所采用的  $D\_Feature$  和  $W\_Feature$  都能对代词分类起到很大的作用.在两种特征同时使用时,分类效果最好.而且前人的工作主要集中在对“it”的分类上,我们对“this”,“that”,“it”,“they”,“these”,“those”都进行了分类,分类覆盖面更全.

采用 CRF 模型时,如果只消解“it”,CRF 效果比 CME 差一些,这是因为对话中“it”之间可能会存其他代词,“it”序列之间缺乏相关性.但是在同时消解其他代词时,CRF 效果要好于 CME.

另外,我们在文献[4]的语料上开放测试,同时采用  $D\_Feature$  和  $W\_Feature$ ,用 CME 对“it”的分类精度为 90.7%,用 CRF 对所有代词的分类精度为 89.8%.

在测试最终指代消解结果时,我们在测试语料上消解了文献[4]中提到的代词“it”,“that”,“them”,“themselves”,“that”,“this”,“those”,消解结果见表 4、表 5.

**Table 4** Results of anaphoric and non-anaphoric pronouns resolution (%)

**表 4** 名词性代词指代和非名词性代词指代消解结果 (%)

| Pronoun    | Nominal-Pronoun (Algorithm 1) |        | Nominal-Pronoun (Algorithm 2) |        | Non-Nominal-Pronoun |        |
|------------|-------------------------------|--------|-------------------------------|--------|---------------------|--------|
|            | Precision                     | Recall | Precision                     | Recall | Precision           | Recall |
| It         | 88.2                          | 66.7   | 73.5                          | 67.4   | 70.3                | 65.5   |
| They       | 82.4                          | 73.7   | 100                           | 94.7   | -                   | -      |
| Them       | 70                            | 70     | 76.9                          | 76.9   | -                   | -      |
| Themselves | 0                             | 0      | 100                           | 100    | -                   | -      |
| That       | 57                            | 25     | 75                            | 54.5   | 73.7                | 56.0   |
| This       | -                             | -      | -                             | -      | 100                 | 100    |
| Those      | 40                            | 50     | 100                           | 100    | -                   | -      |
| Total      | 78.1                          | 60     | 81.8                          | 72.2   | 72.1                | 59.0   |

**Table 5** Comparison of our work and baselines

**表 5** 我们的实验以及与对比试验的比较

| Pronoun    | Our experiments (Experiment 1) |     |       |       | Experiments of Byron (Experiment 2) |     |       |       | LRC |    |       |       |
|------------|--------------------------------|-----|-------|-------|-------------------------------------|-----|-------|-------|-----|----|-------|-------|
|            | C                              | T   | P (%) | R (%) | C                                   | T   | P (%) | R (%) | C   | T  | P (%) | R (%) |
| It         | 44                             | 61  | 72.1  | 67.7  | 37                                  | 50  | 74    | 56.9  | 27  | 36 | 75    | 41.5  |
| They       | 18                             | 18  | 100   | 94.7  | 15                                  | 19  | 79    | 79    | 14  | 15 | 93    | 73.7  |
| Them       | 10                             | 13  | 76.9  | 76.9  | 10                                  | 10  | 100   | 76.9  | 6   | 7  | 85.7  | 46.2  |
| Themselves | 1                              | 1   | 100   | 100   | 1                                   | 1   | 100   | 100   | 1   | 1  | 100   | 100   |
| That       | 54                             | 74  | 73.0  | 55.1  | 62                                  | 93  | 62    | 63.3  | 5   | 20 | 25    | 5.1   |
| This       | 1                              | 1   | 100   | 100   | 0                                   | 1   | 0     | 0     | 0   | 0  | -     | 0     |
| Those      | 6                              | 6   | 100   | 100   | 5                                   | 6   | 83    | 83    | 4   | 4  | 100   | 66.7  |
| Total      | 134                            | 174 | 77.0  | 66.0  | 130                                 | 180 | 72    | 64.0  | 57  | 83 | 68.7  | 28.1  |

本文对指代消解最终结果的评测采用了如下评测指标<sup>[35-37]</sup>:

设  $A$ =正确消解的代词总量, $B$ =消解的代词总量, $C$ =需要消解的代词总量,

$$\text{正确率} = A/B \tag{1}$$

$$\text{召回率} = A/C \tag{2}$$

“this”在测试语料中没有出现名词性指代,“they”,“them”,“themselves”,“those”在测试语料中没有出现非名词性指代,所以表 4 中相应的地方为空白.

由表 4 可以看出,方法 1 没有方法 2 总体消解结果好,说明把消解问题转化为分类问题的过程中,训练样本正反样例不平衡、分类结果还原为消解结果所采用的策略等因素都可能造成了一定的影响.表中对“it”的消解结果,方法 1 的正确率比方法 2 要高,但是召回率却较低,这是因为在将消解问题转化为分类问题后,选择先行词时只能在分类器判断为和代词“存在指代关系”的候选先行词中选,如果代词的所有候选先行词都被分类器判

断为“不存在指代关系”,则该代词没有消解结果.该情况在其他代词的消解中也存在.

表 5 为本文的实验和文献[4]的实验结果比较(表 2 中 *those* 数量为 7 个,但是在文献[4]发布的语料中,其中一个先行项被标为 *ambiguous*,表 5 中计算召回率时没有算在待消解总量中,总量作为 6 个计算).因为文献[4]把两类代词放在一起评测,所以该表均为把两类代词消解放在一起评测的总体消解结果.表中 *C* 为正确消解的代词数量,*T* 为消解的代词数量,*P* 为正确率,*R* 为召回率.其中,Experiment 1 部分中名词性指代消解采用了方法 2. Experiment 2 和 LRC 的结果均为文献[4]中的数据.测试语料中待消解代词总量见表 2.

由表 5 可以看出,与 Byron 的工作相比,消解代词“*that*”时,我们的工作召回率偏低,因为在语料中接近 80% 的“*that*”都是非名词性指代,非名词性指代消解完全依赖于语法分析结果,我们的算法在生语料上直接进行,而 Byron 的工作是在人工清洗后的语料上进行的<sup>[4]</sup>.对其他代词的消解,我们的工作正确消解的代词数量均高于或等于 Byron 的工作.表 5 中,LRC 算法只能消解名词性指代,所以召回率较低.

### 4.3 错误分析

本文消解算法的主要消解错误原因如下,其中,名词性指代消解错误主要是:首先,语法分析错误影响较大.一方面,错误的语法分析导致部分先行词没有被分为名词短语,这就导致它们不能被放在候选先行词中;另一方面,语法分析的错误会导致抽取错误的语法特征和语义特征.其次,对话经常围绕某个主题展开,因而指代距离较长,较长的指代距离容易导致消解错误.非名词性指代消解错误主要是:一类为语法分析错误,不少动词短语没有被正确分出,而把不少非动词短语分成了动词短语,导致候选先行词选择错误;另一类是代词指向的抽象对象并不只由 1 个动词短语构成,无法用前文中出现的任何一个单独的动词短语表示,本文目前还不能消解这种条件下的指代.

## 5 总结与展望

本文系统地分析了传统代词指代消解所采用的特征在口语对话中的适用性以及以往针对口语的代词指代消解工作中的一些不足,然后提出了一套较完整的针对口语对话的代词指代消解算法.算法包括名词性代词指代和非名词性代词指代的分类、名词性代词指代的消解以及非名词性代词指代的消解.算法可以对真实口语对话环境下包含大量不流利信息的对话文本进行代词指代消解.与前人工作相比,本文的工作取得了较好的结果.除了消解性能有所提升以外,本文较大的改进之一是算法均在口语对话生语料上测试.而以往大多数相关工作是在手工清洗或手工标好语法结构的语料上进行的.因此,本文的方法可以直接应用在自动口语对话系统中.本文另一较大改进之处在于,本文所采用的特征均可以自动抽取,不需要手工选择词语、不需要人工定义模板,只需要在训练语料上进行相应的统计工作,而以往不少相关工作都采用了人工选词或者写模板的方法.

在消解过程中我们发现,有部分指代现象目前依然难以消解,如指代距离过远、先行项分布在不同短句中等情况,我们下一步工作将重点考虑在这些条件下代词的消解问题.另外,本文在消解过程中还没有充分利用对话的特征,如在对话中,话语单元(*speech unit*)会包含对话的主题切换等信息,这些信息对指代消解有一定的帮助<sup>[2]</sup>.如何自动切分话语单元并利用话语单元的特征,也是我们未来的工作考虑的重点.

### References:

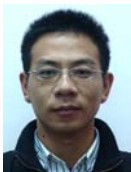
- [1] Lemon O, Gruenstein A, Peters S. Collaborative activities and multi-tasking in dialog system. *Traitement Automatique des Langues, Special Issue on Dialogue*, 2002,43(2):131-154.
- [2] Eckert M, Strube M. Dialogue acts, synchronizing units and anaphora resolution. *Journal of Semantics*, 2000,17(1):51-89. [doi: 10.1093/jos/17.1.51]
- [3] Strube M, Müller C. A machine learning approach to pronoun resolution in spoken dialog. In: Tsujii J, ed. *Proc. of the Annual Meeting on Association for Computational Linguistics*. Morristown: ACL, 2003. 168-175. [doi: 10.3115/1075096.1075118]
- [4] Byron DK. *Resolving pronominal reference to abstract entities* [Ph.D. Thesis]. Rochester: University of Rochester, 2004.
- [5] Tetreault J, Allen J. Semantics, dialogue, and reference resolution. In: *Proc. of the CATALOG*. 2004. <http://www.cs.rochester.edu/~tetreault/sd04.pdf>

- [6] Pappuswamy U, Jordan PW, Van Lehn K. Resolving discourse deictic anaphors in tutorial dialogues. In: Proc. of the Constraints in Discourse. 2005. 96–103. [http://www.public.asu.edu/~kvanlehn/Stringent/PDF/05CSS\\_UP\\_PWJ\\_KVL.pdf](http://www.public.asu.edu/~kvanlehn/Stringent/PDF/05CSS_UP_PWJ_KVL.pdf)
- [7] Eckert M, Strube M. Resolving discourse deictic anaphora in dialogues. In: Proc. of the 9th Conf. of the European Chapter of the Association for Computational Linguistics (EACL). Morristown: ACL, 1999. 37–44. <http://acl.ldc.upenn.edu/E/E99/E99-1006.pdf> [doi: 10.3115/977035.977042]
- [8] Bergsma S, Lin DK, Goebel R. Distributional identification of non-referential pronouns. In: McKeown K, ed. Proc. of the Annual Meeting on Association for Computational Linguistics (ACL 2008: HLT). Morristown: ACL, 2008. 10–18.
- [9] Evans R. Applying machine learning toward an automatic classification of it. *Literary and Linguistic Computing*, 2001,16(1): 45–57. [doi: 10.1093/lle/16.1.45]
- [10] Dimitrov M, Bontcheva K, Cunningham H, Maynard D. A light-weight approach to coreference resolution for named entities in text. In: Branco A, *et al.*, eds. Proc. of the 4th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC). Lisboa: Edições Colibri, 2002. <http://www.gate.ac.uk/sale/daarc2002/DAARC2002.pdf>
- [11] Webber B. Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, 1991,6:107–135. [doi: 10.1080/01690969108406940]
- [12] Tetreault JR. Analysis of syntax-based pronoun resolution methods. In: Proc. of the 37th Annual Meeting on Association for Computational Linguistics. Morristown: ACL, 1999. 602–605. [doi: 10.3115/1034678.1034688]
- [13] Müller C. Resolving it, this, and that in unrestricted multi-party dialog. In: Carroll J, ed. Proc. of the Annual Meeting on Association for Computational Linguistics. Morristown: ACL, 2007. 816–823.
- [14] Boyd A, Gegg-Harrison W, Byron D. Identifying non-referential it: A machine learning approach incorporating linguistically motivated patterns. In: Proc. of the ACL Workshop on Feature Selection for Machine Learning in NLP. 2005. 40–47. <http://acl.ldc.upenn.edu/W/W05/W05-0406.pdf>
- [15] Müller C. Automatic detection of nonreferential it in spoken multi-party dialog. In: McCarthy D, Wintner S, eds. Proc. of the 11th Conf. of the European Chapter of the Association for Computational Linguistics (EACL). Morristown: ACL, 2006. 49–56.
- [16] Ng V, Cardie C. Identifying anaphoric and non-anaphoric noun phrases to improve coreference resolution. In: Proc. of the 19th Int'l Conf. on Computational Linguistics (COLING). Morristown: ACL, 2002. 730–736. <http://www.cs.cornell.edu/yung/papers/coling02.ps> [doi: 10.3115/1072228.1072367]
- [17] Litrán JCC, Satou K, Torisawa K. Improving the identification of non-anaphoric “it” using support vector machines. In: Proc. of the Int'l Joint Workshop on Natural Language Processing in Biomedicine and its Applications. 2004. 58–61. <http://acl.ldc.upenn.edu/W/W04/W04-1210.pdf>
- [18] Ng V, Cardie C. Improving machine learning approaches to coreference resolution. In: Isabelle P, ed. Proc. of the 40th Annual Meeting on Association for Computational Linguistics. Morristown: ACL, 2002. 104–111. [doi: 10.3115/1073083.1073102]
- [19] Lappin S, Leass HJ. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 1994,20(4):535–561.
- [20] Grosz BJ, Joshi AK, Weinstein S. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 1995,21(2):203–225.
- [21] Woei-A-Jin JLRD. Reference resolution in a speech recognition environment. Technical Report, DKS04-01, Delft: Delft University of Technology, 2001.
- [22] Wang HF. Survey: Computational models and technologies in anaphora resolution. *Journal of Chinese Information Processing*, 2002,16(6):9–17 (in Chinese with English abstract).
- [23] Wang ZQ. Research on Chinese coreference resolution and its related technologies [Ph.D. Thesis]. Beijing: Beijing Posts and Telecommunications University, 2006 (in Chinese with English abstract).
- [24] Ng V. Shallow semantics for coreference resolution. In: Proc. of the Joint Conf. on Artificial Intelligence (IJCAI). IJCAI, 2007. 1689–1694. <http://www.hlt.utdallas.edu/~vince/talks/ijcai07.pdf>
- [25] Yang XF, Su J, Tan CL. Improving pronoun resolution using statistics-based semantic compatibility information. In: Knight K, ed. Proc. of the 43rd Annual Meeting on Association for Computational Linguistics. Morristown: ACL, 2005. 33–40. [doi: 10.3115/1219840.1219861]

- [26] Yang XF, Su J. Coreference resolution using semantic relatedness information from automatically discovered patterns. In: Carroll J, ed. Proc. of the 45th Annual Meeting on Association for Computational Linguistics. Morristown: ACL, 2007. 528–535.
- [27] Navarretta C. Resolving individual and abstract anaphora in texts and dialogues. In: Proc. of the 20th Int'l Conf. of Computational Linguistics. Morristown: ACL, 2004. 233–239. <http://acl.ldc.upenn.edu/C/C04/C04-1034.pdf> [doi: 10.3115/1220355.1220389]
- [28] Chomsky N. Lectures on Government and Binding: The Pisa Lectures. Dordrecht: Foris Publications, 1981.
- [29] Reinhart T. Binding Theory. In: Wilson RA, Keil FC, eds. The MIT Encyclopedia of the Cognitive Sciences. Cambridge: MIT Press, 1999. 86–88.
- [30] Ge NY, Hale J, Charniak E. A statistical approach to anaphora resolution. In: Charniak E, ed. Proc. of the 6th Workshop on Very Large Corpora. Montréal: Université de Montréal, 1998. 161–170.
- [31] Soon WM, Ng HT, Lim DCY. A machine learning approach to conference resolution of noun phrases. Computational Linguistics, 2001,27(4):521–544. [doi: 10.1162/089120101753342653]
- [32] Qian W, Guo YK, Zhou YQ, Wu LD. English noun phrase coreference resolution via a maximum entropy model. Journal of Computer Research and Development, 2003,40(9):1337–1342 (in Chinese with English abstract).
- [33] Artstein R, Poesio M. Identifying reference to abstract objects in dialogue. In: Schlangen D, Fernández R, eds. Proc. of the 10th Workshop on the Semantics and Pragmatics of Dialogue. Potsdam: Universität Potsdam, 2006. 56–63.
- [34] Artstein R. Quality control of corpus annotation through reliability measures. In: Proc. of the Annual Meeting on Association for Computational Linguistics. Morristown: ACL, 2007. [Http://ron.artstein.org/publications/2007-acl-t5-slides.pdf](http://ron.artstein.org/publications/2007-acl-t5-slides.pdf)
- [35] Baldwin B. CogNIAC: High precision coreference with limited knowledge and linguistic resources. In: Proc. of the ACL-EACL Workshop on Operational Factors in Practical, Robust Anaphora Resolution. Morristown: ACL, 1997. 38–45.
- [36] Gaizauskas R, Humphreys K. Quantitative evaluation of coreference algorithms in an information extraction system. In: Proc. of the Discourse Anaphora and Anaphor Resolution Colloquium (DAARC-1). London: University College London Press, 1996.
- [37] Harabagiu SM, Maiorano SJ. Knowledge-Lean coreference resolution and its relation to textual cohesion and coherence. In: Proc. of the ACL Workshop on the Relation of Discourse/Dialogue Structure and Reference. Morristown: ACL, 1999. 29–38. <http://acl.ldc.upenn.edu/W/W99/W99-0104.pdf>

#### 附中文参考文献:

- [22] 王厚峰.指代消解的基本方法和实现技术.中文信息学报,2002,16(6):9–17.
- [23] 王智强.汉语指代消解及相关技术研究[博士学位论文].北京:北京邮电大学,2006.
- [32] 钱伟,郭以昆,周雅倩,吴立德.基于最大熵模型的英文名词短语指代消解.计算机研究与发展,2003,40(9):1337–1342.



费仲超(1979—),男,安徽寿县人,博士,主要研究领域为自然语言处理,对话理解.



黄萱菁(1972—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为自然语言处理,信息检索.



周雅倩(1976—),女,博士,讲师,CCF 会员,主要研究领域为自然语言处理,信息检索.



吴立德(1937—),男,教授,博士生导师,主要研究领域为自然语言处理,信息检索,模式识别.