

北京大学学报(自然科学版)
Acta Scientiarum Naturalium Universitatis Pekinensis
doi: 10.13209/j.0479-8023.2016.010

基于全局优化的中文事件同指消解方法

滕佳月 李培峰[†] 朱巧明

苏州大学计算机科学与技术学院, 苏州 215006; [†] 通信作者, E-mail: pfli@suda.edu.cn

摘要 针对目前对事件同指关系的研究中多采用事件对分类或聚类方法而忽略事件相互之间内在联系的问题, 提出一个中文事件同指消解的全局优化模型, 用于减少因分类器错误造成的同指事件链不一致问题。该模型利用对称性、传递性、触发词、论元角色、事件距离等多种约束条件, 将同指消解转化成整数线性规划问题。实验结果表明, 该全局优化模型和分类器方法相比, F1 值提高 4.20%。

关键词 事件; 同指关系; 全局优化; 推理

中图分类号 TP391

Global Inference for Co-reference Resolution between Chinese Events

TENG Jiayue, LI Peifeng[†], ZHU Qiaoming

School of Computer Science & Technology, Soochow University, Suzhou, Jiangsu 215006;

[†] Corresponding Author, E-mail: pfli@suda.edu.cn

Abstract Currently, most pairwise resolution models for event co-reference focused on classification or clustering approaches, which ignored the relations between events in a document. A global optimization model for event co-reference resolution was proposed to resolve the inconsistent event chains in classifier-based approaches. This model regarded co-reference resolution as a Integer Linear Program problem and introduces various kinds constraints, such as symmetry, transitivity, triggers, argument roles, event distances, to further improve the performance. The experimental results show that our model outperformed the local classifier by 4.20% in F1-measure.

Key words Event; co-reference relation; global optimization; inference

事件同指具有普遍性。在自然语言表述中, 同一个事件常常出现多次。为了使语言表达简洁, 一般会分为几个子句甚至段落来进行阐述。当两个事件指向同一个事件本体时, 则认为这两个事件具有同指(或共指)关系。

例 1: 两国首脑今天在巴黎举行会谈。双方在会谈中讨论了中东和平问题。

例 1 中首先引出事件“会谈”发生的时间、地点信息, 然后指出“会谈”的主题为“中东和平”问题。显然, 该例中的两个“会谈”事件具有同指关系。一般情况下, 对于一个事件的报道, 首先会简要概括,

然后对事件发生的时间、地点、人物等进行详细介绍, 最后会总结分析该事件造成的后果或意义。

事件同指消解是信息抽取的子任务, 具有重要的研究意义和应用价值。通过事件的同指关系, 可以有效地结合上下文来理解语意。同指链中的事件相互补充与扩展, 可以将语句成分缺失的事件转换到其同指事件上, 提高语言理解效果。正确识别同指事件, 能更加准确地概括文章大意, 有利于篇章理解、文本摘要、信息抽取等应用。

目前, 事件同指消解的研究大部分针对英文事件, 且主要采用事件对模型进行研究, 即将任意两个事件组成事件对, 然后根据选择的特征或特征对

国家自然科学基金(61472265, 61331011)和江苏省前瞻性联合研究项目(BY-2014059-08)资助

收稿日期: 2015-06-06; 修回日期: 2015-08-26; 网络出版时间: 2015-09-30 11:03:58

进行训练和分类,最终得到每一个事件对被分为正例的概率。但是,这种基于事件对的研究忽略了事件对之间的关系,认为事件对之间相互独立,容易造成分类结果矛盾、同指事件链不一致等问题。

事件同指关系具有明显的传递性特点,本文借鉴 Chambers 等^[1]以及 Do^[2]等研究事件时序关系的思想,采用全局优化推理模型,以减少分类器产生的这种矛盾问题。如图 1(a)所示,分类器结果具有逻辑上的矛盾,因为 E1 与 E2 同指, E2 与 E3 同指,则根据传递性可知 E1 与 E3 同指。然而,图 1(b)中,全局推理则没有这种逻辑上的矛盾。但是,使用传递性约束之后,得到的并不一定是正确的事件关系图(图 1(b)为一种可能的结果)。只有在 E1 与 E2 同指、E2 与 E3 同指均正确时,传递性推理才是正确的。

本文采用整数线性规划方法(integer linear programming),对同指事件链进行文档级的全局优化与推理,可以消除部分识别错误的同指事件,并能根据同指事件的传递性等推理出分类器未识别的同指事件,对召回率有较大提升。本文从以下两个步骤进行:1) 构造一个事件对的同指消解器;2) 利用分类器的预测概率以及约束条件,针对文档级事件进行全局优化推理。

1 相关工作

事件同指消解相关研究较少,且主要针对英文事件的同指进行研究。Ahn^[3]在研究事件抽取时,指出事件同指消解对事件抽取有很大帮助,并构建了一个简单的事件对同指消解系统。事件同指研究主要采用事件对消解模型,即将任意两个事件组成事件对,针对词语、语句、距离等方面提出有效的特征,然后使用机器学习方法,训练出分类模型,最终完成事件对的同指消解。Bejan 等^[4]进一步具体化事件对的特征属性,增加结构化特征进行同指研究。Chen 等^[5]针对 ACE 语料库,基于最大熵模

型建立了特征压缩的事件对同指消解系统。基于分类思想的方法,大多数假设数据样本是独立同分布的,忽略了事件之间的相互联系。Chen 等^[6]使用一种谱图聚类方法,根据事件聚类结果生成事件关系图,然后对该图进行优化。Bejan 等^[7-8]基于非参贝叶斯模型,提出一种新的无监督方法,并在 ACE2005 与 ECB (EventCorefBank)上分别进行验证。

然而,事件对模型的同指消解,忽略了事件之间的内在联系,容易造成同指事件链不一致问题。使用全局优化的方法,则可以减少因分类器错误造成的矛盾情况。针对实体同指消解问题, Nicolae 等^[9]根据聚类算法生成实体关系图,然后提出 BestCut 算法,对该图进行优化,从而完成实体同指消解。Chen 等^[10]整合 7 种同指消解器,并提出一种图分割算法优化同指事件链。Sangeetha 等^[11]使用聚类算法生成事件关系图,然后用 Mincut 算法对其进行优化。Song 等^[12]提出一种联合学习方法,将事件对分类与事件聚类方法整合,并使用马尔科夫逻辑网络进行全局推理。基于图模型优化方法,一般是衡量每一条边(关系)做出取舍,然后对分割后的子图进行评估,只能保证裁剪最优或最小,但消解结果有可能是局部最优,并且也没有使用传递性等约束,仍然无法消除同指事件链不一致问题。

中文事件同指消解的研究相对较少,语料库匮乏,主要使用事件对模型进行同指消解。此外中英文存在语言学上的差异,中文词语较多,一词多义普遍,语法随意,没有明显的单复数和时态等,导致中文事件同指消解的效果偏低。胡乃全等^[13]基于最大熵模型建立中文指代消解系统,并在 ACE05 的 BNews 上进行验证。黎耀炳等^[14]提出中心词匹配算法,对共指消解有很大提高。谢永康等^[15]首先用最大熵模型计算实体对的共指概率,然后使用谱聚类算法进行划分。庞宁等^[16-17]基于多种语义特征,增加维基百科信息,针对突发事件进行共指消

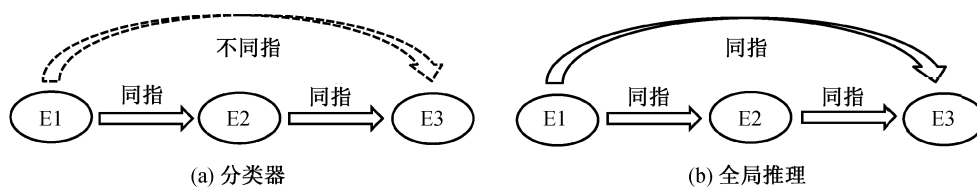


图 1 事件同指关系

Fig. 1 Graphs of event co-reference relation

解。滕佳月等^[18]提出基于触发词语义和组合特征的方法,事件同指消解得到较大提高。

与现有研究不同,本文使用整数线性规划方法来对中文事件同指消解进行全局优化推理,其主要思想借鉴 Chambers 等^[1]和 Do 等^[2]针对事件时序识别提出的全局推理模型,并在该模型中引入多种新颖、有效的约束条件来进一步提高性能。实验结果表明,本文方法可以捕获事件对之间的语义关系,从而减少分类器造成的事件链不一致问题,提高同指事件链的紧密程度,推理出未被分类器识别的同指事件,使得同指事件消解效果得到显著的提高。

2 基准系统

本文首先使用传统的机器学习方法构造一个事件对的中文同指消解器,语料库为 ACE2005 中文语料库。与 Ahn^[3]的方法类似,将任意两个事件组成事件对,并用分类器进行训练和分类。最终,由该分类器计算出每个事件对被分为正例(即两个事件具有同指关系)的概率:

$$P_c(e_i, e_j) = \text{Classifier}(\text{EventPair}(f_i, f_j)), \quad \forall e_i, e_j \in E, i \neq j, \quad (1)$$

其中, e_i 与 e_j 表示第 i 个和第 j 个事件, f_i 与 f_j 表示第 i 个事件与第 j 个事件的特征, E 为该文章中所有的事件集合(下同)。

针对目前事件同指领域的研究,本文首先实现庞宁等^[17]的共指消解方法,并将其作为基准系统 1,具体内容参见文献[17]。考虑到庞宁等的方法并非专注事件共指消解,还包括实体、指代消解,本文实现了滕佳月等^[18]的事件共指消解方法,并将其作为基准系统 2,具体内容参见文献[18]。

3 全局优化推理模型

分类器往往将事件对看做是独立同分布的实例,忽略了事件之间的联系,因此分类结果容易产生逻辑上的矛盾。使用全局优化方法则可以消除部分逻辑上错误的事件对,并可以推理出未识别的同指事件。但是,仅使用传递性约束并不能得到最优的结果。如图 1 所示,尽管全局优化方法可以避免图 1(a)中的不一致问题,但图 1(b)仅仅是一种可能的结果。因此,本文使用整数线性规划法,增加多种有效的约束条件,进一步提升事件同指消解性能。

3.1 优化目标

本文以文档为单位,对该文档中的同指事件链进行全局优化。基于分类概率最大化的思想,提出以下优化目标:

$$\arg \max_x \sum_{i, e_j \in E} \sum_{r \in \{-1, 1\}} x(e_i, e_j | r) \cdot \log x(e_i, e_j | r) \cdot \log P(e_i, e_j | r), \quad (2)$$

其中, r 表示事件 e_i 与 e_j 之间的关系,取值为 1 表示同指关系,取值为 -1 表示不具有同指关系。 $x(e_i, e_j | r)$ 是一个二元变量(0 或 1),保证 r 的取值唯一。由于分类概率 P 的取值范围为 $[0, 1]$, 差异性较小,本文对概率 P 进行对数映射,使得变量 x 的权值差异更大,从而更有利于对同指事件链进行全局优化。在计算概率 P 时,本文使用以下公式替代:

$$\tilde{P}(e_i, e_j | r) = \frac{P(e_i, e_j | r)}{1 - P(e_i, e_j | r)}. \quad (3)$$

式(3)可以更好地描述分类概率的离散差异。为了保证式(3)的有效性,针对 $P=1$ 时,本文强制将 P 设为 0.9999; $P=0$ 时,将其设为 0.0001。

3.2 基本约束条件

3.2.1 唯一性

整数线性规划的思想类似于枚举所有可能的结果,然后根据优化目标与约束条件,计算出最优的分类结果。为了保证事件 e_i 与 e_j 之间有且只有一种关系,首先限制 r 的取值为 $\{-1, 1\}$ (1 表示同指关系, -1 表示不具有同指关系),然后保证二元变量 $x(e_i, e_j | r)$ 相互之间不存在矛盾,即有且只有一个二元变量 $x_k(e_i, e_j | r)$ 取值为 1, 其余均为 0。具体描述如下:

$$\sum_{r \in \{-1, 1\}} x(e_i, e_j | r) = 1, x(e_i, e_j | r) \in \{0, 1\}. \quad (4)$$

3.2.2 对称性

事件的同指关系是相互的,即事件 e_i 与 e_j 具有同指关系,反之亦然。具体描述如下:

$$x(e_i, e_j | r) = x'(e_j, e_i | r), \forall e_i, e_j \in E, i \neq j. \quad (5)$$

3.2.3 传递性

由事件同指的定义可知,同指关系具有传递性特点。若已知 $r_1(e_i, e_j)$ 和 $r_2(e_j, e_k)$, 则可以推出 $r_3(e_i, e_k)$ 。具体表示如下:

$$x(e_i, e_j | r_1) + x(e_j, e_k | r_2) - x(e_i, e_k | r_3) \leq 1, \quad \forall e_i, e_j, e_k \in E, i \neq j \neq k, \quad (6)$$

其中 $(r_1, r_{12}, r_3) \in \{(1, 1, 1), (1, -1, -1), (-1, 1,$

-1)}, 分别表示:

$$\text{coref}(e_i, e_j) + \text{coref}(e_j, e_k) \rightarrow \text{coref}(e_i, e_k),$$

$$\text{coref}(e_i, e_j) + \text{uncoref}(e_j, e_k) \rightarrow \text{uncoref}(e_i, e_k)$$

$$\text{uncoref}(e_i, e_j) + \text{coref}(e_j, e_k) \rightarrow \text{uncoref}(e_i, e_k)$$

3.3 扩展约束条件

3.3.1 论元角色

事件的论元包含多种信息, 同指的两个事件往往含有相同的论元。按照论元角色可分为: 人物(Person)、地点(Place)、职位(Position)等, 若仅统计事件对论元异同个数, 并不能很好地识别同指事件。本文针对开发集事件对, 按照论元角色类型分别统计异同个数, 并依据论元异同个数分别计算正负例所占的比例, 最终得出正例比例较高的论元角色类型集合(Arg_Roles)。如果一个事件对含有相同的论元, 且该论元角色类型在上述 Arg_Roles 集合中, 则该事件对具有同指关系的概率较高, 具体表示为

$$x(e_i, e_j | r = 1) = 1, \exists \arg_k \in \text{Arge}_i \cap \text{Arge}_j \wedge \text{argrole}_k \in \text{Arg_Roles}, \quad (7)$$

其中, Arge_i , Arge_j 分别表示事件 e_i , e_j 的论元集合, argrole_k 表示论元 $\arg_k$ 的角色。通常, 充当人物(Person)的论元, 对同指识别没有明显效果, 因为文章会介绍与该人物相关的其他事件; 充当职位(Position)的论元则有很大帮助。如例 2 所示($\langle \rangle$ 内为论元), 其中例 2(a)与例 2(c)为同指事件。

例 2(a): $\langle$ 杨富家教授 $\rangle$ 起程赴世界著名大学 $\langle$ 英国 $\rangle$ $\langle$ 诺丁汉大学 $\rangle$ 就任 $\langle$ 校长 $\rangle$ 。

例 2(b): $\langle$ 杨富家教授 $\rangle$ $\langle$ 1991 年 $\rangle$ 当选 $\langle$ 中国科学院院士 $\rangle$ 。

例 2(c): $\langle$ 他 $\rangle$ 在接受记者采访时说, $\langle$ 诺丁汉大学 $\rangle$ 聘请我担任 $\langle$ 校长 $\rangle$, 表明中国教育水平和管理水平取得显著成就, 并得到国际上的广泛认

可。

论元“杨富家教授”的角色为 人物类型, 但事件“就任”与“当选”为非同指事件对, 而论元“校长”、“中国科学院院士”的角色为 职位类型, 并且职位在集合 Arg_Roles 中。同指事件(a)和(c)包含相同论元“校长”; 非同指事件(a)和(b), 充当职位的论元则不同。

3.3.2 事件距离

从 ACE 语料库可以看出, 一篇新闻报道会有一个主题或主要报道的事件。本文使用简单的方法概括文档的主题, 即统计一篇文章中出现次数最多的事件类型, 并将其作为该文章的主题。依据人们的写作习惯, 一般首先提出文章的主要表达事件, 然后对该事件进行详细阐述, 中间也会提及与该事件相关的事件, 最后会对该事件总结或阐明事件的后果。如果事件对的事件类型与该文章的主题一致, 则对该事件对进行事件距离的约束。

事件距离的衡量采用两种方法(方法 1: 两个事件所在语句差; 方法 2: 事件对间隔事件数), 并对两者数据进行归一化。以方法 1 为例, 值为 0, 1, 2, 3, 分别表示事件在同一句、相邻句、间隔一句、其他。开发集统计结果如图 2(a)所示(只统计与文章主题一致的事件对), 事件对在相邻句、间隔一句的时候, 具有同指关系的比例较高, 如例 1 所示。因此提出以下约束:

$$x(e_i, e_j | r = 1) = 1, \text{bias}_p(e_i, e_j) \in \Phi \wedge \text{type}_{i,j} = \text{type}_d, \quad (8)$$

其中 $\text{bias}_p(e_i, e_j)$ 表示事件 e_i , e_j 所在语句差, type_d 为该文章的主题(出现次数最多的事件类型), $\text{type}_{i,j}$ 为事件 e_i , e_j 的事件类型(事件 e_i , e_j 的类型相同, 因为类型不同的事件不具有同指关系)。经开发集调节参数, 其中 $\Phi = \{1, 2\}$ 。

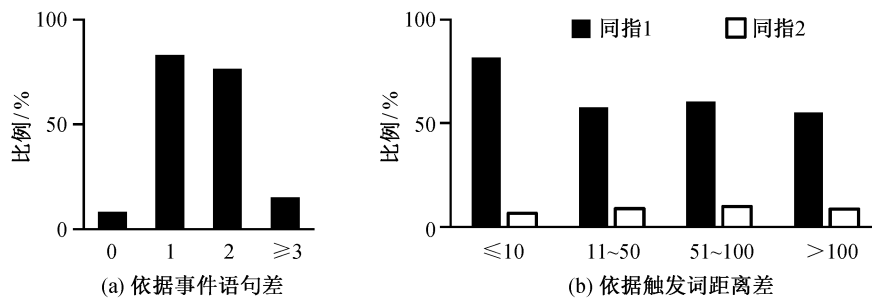


图 2 事件同指比例

Fig. 2 Event co-reference ratios

3.3.3 触发词

通常,一个复合句会包含多个事件,而这些事件之间具有一定的相关性。当这些事件触发词不同时,大多具有时序、转折、因果等关系,即多为非同指事件,例如“袭击造成 2 人死亡 10 人受伤”等;当这些事件触发词语义相似时,多为同指的,如例 1 所示。针对触发词方面,本文使用语义相似度计算以及词语匹配方法来衡量触发词对的一致性。

图 2(b)为开发集统计结果,其中“同指 1”表示触发词一致时的结果,“同指 2”表示触发词不一致时的结果。当触发词一致时,触发词距离非常近的事件对中,同指的比例高达 81.6%;触发词不一致时,同指的事件对比例均不足 10%。因此,本文仅针对第一种情况约束,具体表示为

$$x(e_i, e_j | r = 1) = 1, \text{bias}_t(e_i, e_j) \leq \varepsilon \wedge \text{trigger}_i \approx \text{trigger}_j, (9)$$

其中 $\text{bias}_t(e_i, e_j)$ 表示事件 e_i, e_j 触发词距离差, trigger_i 与 trigger_j 分别表示事件 e_i, e_j 的触发词。经开发集调节参数, ε 取值为 12。

4 实验

本节中首先介绍实验数据选择与设置,然后对比并分析基准系统与全局推理的实验结果,最后分析全局推理中各约束条件对系统的贡献情况以及相互之间的影响。

4.1 实验设计

ACE 中文语料库共计 632 篇,本实验选择其中含有相同类型事件对的文档(事件类型不相同的事件对不具有同指关系,故舍弃),总计 445 篇。将同一文档中的任意两个事件组成事件对,剔除事件类型不相同的事件对,共计 14394 个事件对,正负例比例约为 1:5。由于不同体裁来源的文档其语言风格有一定差异,依据 ACE2005 语料库来源,本文在选择开发集、训练集、测试集时,尽量覆盖每种文档来源。最终,随机选择 50 篇文档作为开发集,测试集与训练集约为 1:4,并进行五倍交叉验证。

基准系统 1 部分实现了文献[17]的方法,只选择与事件同指有关的特征。基准系统 2 实现了文献[18]的基于事件对模型的事件同指消解方法。全局推理过程使用 Gurobi Optimizer 6.0^①。实验结果评价标准为 MUC-6,使用的软件为 Conll2012-scorer-

8.0^②。

4.2 实验结果与分析

实验结果如表 1 所示。其中,基本约束条件使用式(4)~(6),实现了 Do 2012 中的传递性、对称性约束;扩展约束使用式(7)~(9)。经显著性检验, $p=0.0003$,表明本文全局推理模型较基准系统具有极显著性差异。由于基准系统 2 比基准系统 1 结果稍好,故以下对比均基于基准系统 2 进行。

由表 1 可知,基本约束条件对系统结果提升并不明显。首先,基本约束中的唯一性只是针对分类结果不产生歧义进行约束,对同指消解结果基本上没有提升;其次,根据事件对的分类结果统计,发现任意一个事件对均符合 $r_1(e_i, e_j) = r_2(e_j, e_i)$,即满足式(5),所以对称性对系统没有提升;最后,传递性是根据已知两条分类结果推理第三条,只有在已知的分类结果正确的情况下,传递性推理才正确。换言之,存在这种矛盾情况(其中 $r_3 \neq r_3'$):

$$r_1(e_a, e_b) \wedge r_2(e_b, e_c) \rightarrow r_3(e_a, e_c),$$

$$r_1'(e_a, e_d) \wedge r_2'(e_d, e_c) \rightarrow r_3'(e_a, e_c)。$$

在满足一系列约束条件下,ILP 会根据优化目标,计算出结果最优的那一组解,最终会自动选择 r_3 或 r_3' 作为正确结果,消除矛盾。分析实验结果可知,只有一部分数据被推理成正例,由于本文使用 MUC-6 评价标准,需要将基准系统中的事件对分类结果转化成事件链形式,才能进行 MUC-6 评估。由于在转化成事件链的时候已经使用了传递性原理,所以传递性提升较低,最终导致基本约束条件对结果提升并不明显。

由 3.3 节可知,扩展约束条件大部分是利用正例结果来推理正例,因此对召回率有提升明显, R

表 1 实验结果比较
Table 1 Comparison of experimental performance

实验	R/%	P/%	F1/%
基准系统 1	66.83	76.65	71.40
基准系统 2	72.18	74.87	73.50
+基本约束	73.94	74.23	74.08 (+0.58)
+基本+扩展约束	78.68	76.74	77.70 (+4.20)

说明:括号内数字表示与基准系统 2 的比较结果, + 表示提升。

① <http://www.gurobi.com/>

② <http://www.cs.upc.edu/~esapena/?s=downloads>

值提高 6.50%。对于准确率提升不明显,只有 1.87%。这是因为: 1) 扩展约束与传递性类似,同样会产生前面推理的矛盾性,而全局优化软件自动选择最优结果,有时这个结果由于受到分类器错误判定结果的影响,并不一定正确; 2) 扩展约束基于开发集统计结果进行设定,并从正负例分布概率上进行约束,由于数据分布的不均匀问题,测试集中部分样本不符合该约束条件。因此,扩展约束对准确率提升较有限,而对召回率提升大,进而对系统的 F1 值提升 4.20%,表明全局推理模型对事件同指消解效果有明显提高。

4.3 扩展约束分析

表 2 给出不同扩展约束对全局推理性能的贡献度。3 种扩展约束均基于开发集统计结果选择同指比例较高的情况进行约束,因此扩展约束条件对系统的召回率的提升高于准确率。

论元角色约束对准确率提升最低,但对召回率提升较高,主要是因为论元角色集合 Arg_Roles 是根据同指事件比例较高的角色类型统计得出的,只有少量样本并不符合该约束。例如,事件内容涉及 A 国总理的“就任”、“出访”,然后与 B 国总理“会晤”,B 国总理“宣布”合作成立等。论元“总理”作为职位角色频繁出现在这一系列事件中,但这些事件大都不具有同指关系。

触发词约束对系统提升最高。一方面,触发词约束使用了触发词语义特征。由图 2(b)可知,当事件对触发词语义不一致时,只有不到 10%的事件对具有同指关系;而触发词语义一致时,距离最近的事件对中有 80%以上是同指事件,因此触发词语义相似度与事件同指有很大关系。另一方面,使用了触发词距离特征,距离较近的触发词对如果语义一

致,则具有同指关系的概率较高。经统计分析实验数据,发现未识别出的同指事件大多是触发词不一致的情况,而触发词一致的同指事件大都被很好地识别。

事件距离约束的贡献度较低,因为根据 3.3.2 节,一篇文章只会对其主线的事件链(与文章主题相关的事件)进行约束,没有对与文章主题相关性较小的事件进行约束。由图 2(a)所示,事件距离较近的事件对分为正例的概率只有 80%左右(即事件距离 ≤ 10 的事件对),剩下事件距离较远的事件对中有 55%左右为同指事件,即有很大部分同指事件无法进行约束。此外,3.3.2 节中选择出现次数最多的事件类型作为主题,对文章主题的检测过于简单,准确性较低。

从表 2 同样可以发现,这些约束条件相互之间存在影响。“事件距离+触发词”的贡献度(2.8%)远小于两者贡献度之和(4.34%),因为两者都是对距离的约束,只不过侧重点不同。由 3.3.3 节可知,事件距离最近的样本中非同指事件比例较高(即事件在同一句情况下),而相邻句、间隔一句的样本中同指事件比例较高(如图 2(a)所示),整体上近似正态分布。但是,触发词距离差的同指事件概率分布与之不同,两者结合有抵消效果。并且,基于统计角度的约束规则,其中存在重叠情况,如例 1 所示,符合事件距离和触发词两种约束。

“触发词+论元角色”以及“论元角色+事件距离”的组合结果与两者贡献度之和相比,召回率的下降比率比准确率低。这是因为: 1) 触发词约束包含事件触发词语义信息,而论元角色以及事件距离约束仅依据同指出现概率,缺少语义特征,造成同指识别的准确率不高; 2) 有时由于样本的不均匀

表 2 各约束条件的贡献度
Table 2 Contributions of different constraints

实验	R/%	P/%	F1/%
ILP (基本约束)	73.94	74.23	74.08
+论元角色	76.83 (+2.89)	75.45 (+1.22)	76.13 (+2.05)
+事件距离	76.13 (+2.19)	75.76 (+1.53)	75.94 (+1.86)
+触发词	76.86 (+2.92)	76.27 (+2.04)	76.56 (+2.48)
+论元角色+事件距离	78.04 (+4.10)	75.93 (+1.50)	76.97 (+2.89)
+事件距离+触发词	77.65 (+3.71)	76.12 (+1.89)	76.88 (+2.80)
+触发词+论元角色	77.96 (+4.02)	76.47 (+2.24)	77.21 (+3.13)

说明: 括号内数字表示与基准系统 2 的比较结果, +表示提升。

性, 开发集与测试集的样本分布并不完全一致, 例如事件“前往”与“访问”, 在开发集文档中两个事件的类型分别为“运动(Movement)”和“联系(Contact)”, 为非同指事件, 且不存在这种组合(类型不同的事件对已被剔除), 但在测试集中, 有一篇文章中共有6个事件与之同指, 标记为“运动(Movement)”类型。另外, 含有触发词约束的组合中准确率贡献度都比较高, 表明触发词语义特征对同指消解的准确率有较大帮助。

5 总结

本文提出一个全局推理方法, 有效地减少因分类器造成的同指事件链不一致问题, 进一步提升了中文事件同指消解性能。在构建全局推理模型之前, 首先构造基于机器学习的同指消解器, 然后将分类器输出的概率与同指事件特点结合, 提出触发词、论元角色以及事件距离多个约束条件。实验结果表明, 基于全局优化的推理方法较有监督的机器学习方法有明显提升。因此, 本文提出的几个约束条件对同指消解有很大帮助。在下一步研究中, 将对事件时间信息、时序关系等做进一步研究, 然后加入到全局推理模型中, 以便进一步提高同指消解效果。

参考文献

- [1] Chambers N, Jurafsky D. Jointly combining implicit constraints improves temporal ordering // *Proceedings of the Conference on EMNLP*. Waikiki, 2008: 698-706
- [2] Do Q X, Lu W, Roth D. Joint inference for event timeline construction // *Proceedings of the conference on EMNLP-CoNLL*. Jeju Island, 2012: 677-687
- [3] Ahn D. The stages of event extraction // *Arte'06 Proceedings of the Workshop on Annotating & Reasoning About Time & Events*. Sydney, 2006: 1-8
- [4] Bejan C A, Harabagiu S. A linguistic resource for discovering event structures and resolving event coreference // *Sixth International Conference on Language Resources & Evaluation (LREC)*. Marrakech, 2008: 2881-2887
- [5] Chen Zheng, Ji Heng, Haralick R. A pairwise event coreference model, feature impact and evaluation for event coreference resolution // *Proceedings of the Workshop on Events in Emerging Text Types (eETTs)*. Borovets, 2009: 17-22
- [6] Chen Zheng, Ji Heng. Graph-based event coreference resolution // *Proceedings of the 2009 Workshop, ACL-IJCNLP*. Suntec, 2009: 54-57
- [7] Bejan C A, Tittsworth M, Hickl A. Nonparametric bayesian models for unsupervised event coreference resolution // *Advances in Neural Information Processing System 22 (NIPS 2009)*. Vancouver, 2009: 73-81
- [8] Bejan C A, Harabagiu S. Unsupervised event coreference resolution. *Computational Linguistics*, 2014, 40(2): 1412-1422
- [9] Nicolae C, Nicolae G. BESTCUT: a graph algorithm for coreference resolution // *Empirical Methods in Natural Language Processing*. Sydney, 2006: 275-283
- [10] Chen Bin, Su Jian, Pan S J. A unified event coreference resolution by integrating multiple resolvers // *Proceedings of International Joint Conference on Natural Language Processing*. Chiang Mai, 2011: 102-110
- [11] Sangeetha S, Arock M. Event coreference resolution using mincut based graph clustering // *The Fourth International Workshop on Computer Networks & Communications*. Coimbatore, 2012: 253-260
- [12] Song Y, Jiang J, Zhao W X. Joint learning for coreference resolution with markov logic // *Proceedings of the Conference on EMNLP-CoNLL*. Jeju Island, 2012: 1245-1254
- [13] 胡乃全, 孔芳, 王海东. 基于最大熵模型的中文指代消解系统实现. *计算机应用研究*, 2009, 26(8): 2948-2951
- [14] 黎耀炳, 张牧宇, 秦兵, 等. 基于中心语匹配的共指消解 // *第六届全国信息检索学术会议论文集*. 哈尔滨, 2010: 3-8
- [15] 谢永康, 周雅倩, 黄萱菁. 一种基于谱聚类的共指消解方法. *中文信息学报*, 2009, 23(3): 10-16
- [16] 庞宁, 杨尔弘. 基于最大熵模型的共指消解研究. *中文信息学报*, 2008, 22(2): 24-27
- [17] 庞宁, 杨尔弘. 多种语义特征在突发事件新闻中的共指消解研究. *中文信息学报*, 2014, 28(1): 26-32
- [18] 滕佳月, 李培峰, 朱巧明. 基于触发词语义和组合特征的中文同指事件消解方法 // *第十六届词汇语义学国际研讨会*. 北京, 2015: 334-339