

文章编号: 1003-0077 (2011) 00-0000-00

## 基于跨场景推理的事件关系检测方法\*

杨雪蓉, 洪宇, 陈亚东, 王潇斌, 姚建民, 朱巧明

(苏州大学江苏省计算机信息处理重点实验室, 江苏, 苏州, 215006)

**摘要:** 事件关系检测是一项面向事件之间逻辑关系的自然语言处理技术。事件关系识别的核心任务是以事件为基本语义单元, 通过分析事件的篇章结构信息及语义特征, 实现事件逻辑关系的深层检测。本文首次建立一套事件关系检测的任务和研究体系, 包括任务定义、关系体系划分、语料采集与标注、评价方法等。同时, 本文提出了一种跨场景推理的事件关系检测方法, 该方法认为, 具有相同事件场景的“事件对”, 往往具有相同的事件关系类型。本文提出的基于跨场景推理的事件关系检测方法在针对四大类事件关系类型的检测精确率为 54.21%。

**关键词:** 事件关系; 框架语义; 事件场景向量; 事件场景;

中图分类号: TP391

文献标识码: A

## Detecting Event Relation through Cross-Scenario Inference

Xuerong Yang, Yu Hong, Yadong Chen, Xiaobin Wang, Jianming Yao, Qiaoming Zhu

(Provincial Key Laboratory of Computer Information Processing Technology

Soochow University, Suzhou, China, 215006)

**Abstract:** Event relation detection aims to detect a logical relation between pairwise events. The key to event relation detection is to detect latent logical relation between events by analyzing the corresponding discourse structure and semantic features of events, with the techniques of semantic relation recognition and inference. In this paper, we firstly build an overall framework of the event relation detection task, including task definition, relation hierarchical structure, corpora acquisition and evaluation. Secondly, we propose a cross-scenario inference method to predict relation between pairwise events, which follows the basic hypothesis that if events express the same scenarios, they normally trigger similar relations. Finally, we experiment on four general semantic relations, Expansion, Comparison, Contingency and Temporal, achieving an accuracy of 54.21%.

**Key words:** Event Relation; Frame; Event Scenario Vector; Event Scenario

### 1 引言

事件 (Event) 是由特定人、物、事在特定时间和特定地点相互作用的客观事实, 事件关系是事件之间相互依存和关联的逻辑形式, 具有客观性、事实性和规律性三种主要特性, 如“因果” (Contingency) 和“时序” (Temporal) 关系等。事件关系能将离散于文本中的事件相连接, 形成事件关系网络和事件发展的拓扑脉络。事件关系检测以事件为主体元素的自然语言逻辑关系抽取, 事件关系检测能够有效辅助事件衍生、发展与变异信息的推理与预测。

目前针对事件关系检测的研究刚刚起步, 由于不具有权威的任务定义、事件关系体系以及评测标准, 相关探索尚不深入, 相应方法也仅仅着眼于某一特定事件关系类型 (如“因果”关系) 的判定, 不具有全面性和普适性。相较而言, 自然语言处理领域中的篇章关系检测篇章关系分析作为自然语言处理领域中一项重要的基础性研究, 旨在同一篇章内部, 判断相邻片段或跨度在一定范围内的两个片段之间的语义连接关系。

文本中的事件通过文字诉诸后, 成为信息传播中可读可解的事件文体 (也称“事件体”, 本文统称“事件”)。从而, 事件的描述必须遵循自然语言的行文规律, 如篇章结构、篇章修

---

\* 收稿日期:                      定稿日期:

**基金项目:** 国家自然科学基金 (No. 61003152, 61272259, 61272260, 90920004, 61373097), 教育部博士学科点专项基金 (No. 20103201110021), 江苏省自然科学基金 (No. BK2011282), 江苏省高校自然科学基金重大项目 (No. 11KJA520003), 苏州市自然科学基金 (No. SYG201030, SH201212)。

辞、语法和文法等规律。篇章关系检测以论元（即完整的语义单元，如句子）为对象，建立了较为全面的任务体系，其蕴含的概念、关系体系和评价方法也具有普适性，能够有效应用于事件关系检测任务。因此，事件关系检测与篇章关系检测任务有着一定程度的领域交差性。然而，事件关系并不等同于篇章关系，篇章关系检测不能涵盖所有事件关系检测的关键问题：

- 篇章关系检测任务涉及短语、子句、句子等文本片段之间的语义关系研究，通过分析两两邻或跨度在一定范围内的两个论元的语义连接关系，构建文本篇章关系结构，进而深入理解篇章语义；而文本中的事件往往离散分布，从而只受话题框架约束，因此不属于篇章关系的研究范畴。
- 篇章关系检测往往可以利用句法、语义和语境的局部特征予以方法设计与实现，而事件关系则必须依赖大规模数据估计和挖掘广域的关联线索进行方法设计与实现。从而，事件关系检测需要一种有针对性的专属的任务和研究体系，而不能将篇章关系检测体系简单移植和并用。

本文针对事件关系检测任务，提出一种跨场景的事件关系检测方法。事件场景由能够用于描述该场景的一系列组成元素构成。跨场景的事件关系检测方法的核心思想认为：相同或相似的事件场景，通过挖掘其特征，有助于事件关系类型的推理。因此，本文首先构建事件场景及其关系类型，利用事件场景的关系类型对事件关系类型进行推理。跨场景的事件关系检测方法需要首先解决如下两点问题：1) 事件场景的匹配方法。即给定待测“事件对”，如何从已知关系类型的“事件对”中，得到与待测“事件对”相同或相似事件场景的“事件对”2) 事件场景的先验关系类型。事件之间的关系类型通过事件场景的关系类型推理得到，因此需要首先获得事件场景的先验关系类型。针对第一点，本文采用基于 FramNet (CJ, 2003)<sup>[1]</sup>的框架语义为事件描述构造事件场景的特征空间，从而将事件场景的相似度转化为特征空间的相似度；针对第二点，本文借助显式连接词，挖掘包含显式连接词且毗邻的“事件对”，同时，将显式连接词触发的某一类事件关系类型作为该“事件对”的关系类型，例如连接词“because (因为)”往往触发“Causal (因果)”关系类型，进而利用已知关系类型的毗邻“事件对”构建事件场景的先验关系类型。

同时，根据事件关系检测和篇章关系检测的任务的异同点，本文提出基于篇章关系分析的事件关系检测体系，该体系包括事件关系检测的任务定义、关系体系以及评价方法等。同时，根据定义的事件关系体系，以 Frame-1.5 的新闻语料为数据源，对其中已标注的事件进行事件关系类型的标注。

本文组织如下：第 2 节介绍事件关系检测的相关工作；第 3 节给出事件关系检测任务定义；第 4 节分析本文方法的动机；第 5 节详细阐述基于跨场景的事件关系检测方法的方法；第 6 节介绍实验；第 7 节总结。

## 2 相关工作

由于缺少公认的事件关系体系，目前针对事件关系的研究方法主要针对某种特定事件关系类型的判定进行研究<sup>[2][3][4]</sup>，主要的挖掘方法分为模板匹配法和元素分析法。

### 模板匹配法

事件关系检测的主要方法之一是借助事件特征的模式匹配，例如，利用事件触发词的关系模式匹配，根据人工定义的模板，对文本中符合模板的事件关系进行抽取。Chklovski<sup>[3]</sup>等首先定义六种时序关系：“similarity”（时序“相似”关系），“strength”（时序“加强”关系），“antonymy”（时序“相反”关系），“enablement”（时序“支持”关系），“happens”（时序“发生”关系）和“before”（时序“前”关系），再利用人工收集的 LSP (Lexical-Syntactic Pattern, 即词-句匹配模板) 抽取包含这六种时序关系的“事件对”，并将抽取的结果形成称为“VerbOcean”的知识库。人工定义的事件关系模板往往受数量限制，造成关系检测的低

召回率问题。Pantel<sup>[4]</sup>通过 Espresso 算法进行自动模板的构建,算法首先给定少量关系实例,通过机器学习方法对现有模板进行迭代扩展,在一定程度上提高了模板匹配方法的召回率。

### 元素分析法

以事件元素为线索的研究大都继承了 Harris<sup>[5]</sup>的分布假设。Harris 假设指出,处在同一上下文环境中的词语具有相同或相似的含义。Lin<sup>[6]</sup>提出了一种结合 Harris 分布假设和建立依存树思想的无监督方法,称为 DIRT 算法。算法将所有事件构造成依存树形式,树中的每条路径表示一个事件,路径的节点表示事件中的词语,若两条路径的词语完全相同,则这两条路径所表示的事件相同或者相似。

## 3 任务定义

事件关系检测任务的目标在于实现事件间逻辑关系的自动检测。本文首先介绍事件、事件关系类型以及事件关系检测的任务定义

### 3.1 事件

自动内容抽取 (Automatic Content Extraction, 简称 ACE) 任务将事件定义为由事件触发词和事件参与者组成,其中,事件的触发词能够描述一个事件的发生,以动词或者动名词为主;事件的参与信息为事件参与者,通常为名词。例如包含如下事件 Evt1:

(1) Evt1: “奥巴马击败麦凯恩”

事件 Evt1 中“击败”能够描述该事件的发生,因此,“击败”为该事件的触发词,则称 Evt1 为“击败”事件。事件 Evt1 中“奥巴马”和“麦凯恩”为“击败”事件的参与信息,即“击败”事件的参与者。

### 3.2 事件关系

事件关系表示事件与其相关事件之间相互依存和关联的逻辑形式,是事件之间固有的一种客观存在。然而,目前针对事件关系检测任务缺少公认的事件关系体系,本文将篇章关系体系与事件关系进行分析和对比,选取篇章关系体系中能够应用于离散“事件对”的篇章关系类型作为事件关系类型。同时,篇章关系与事件关系的差异性使得篇章关系不能描述完整的事件关系类型,因此,本文借助事件关系实例,人工总结事件关系类型,进一步对事件关系类型进行补充,确保事件关系体系的完整性,由此形成的事件关系体系如表 1 所示:

表 1 事件关系体系

Table 1 Hierarchy of event relation

第一层	第二层	第一层	第二层
Temporal(时序)	Synchronous(同步)	Expansion(扩展)	List(并列)
	Asynchronous(异步)		Progression(递进)
Comparison(比较)	Contrast(对比)		Restatement(重述)
	Concession(让步)		Instantiation(实例化)
Contingency(偶然)	Cause(因果)		---
	Condition(条件)		

表 1 定义的事件关系体系共分为两层,第一层包含四种主要关系类别: Temporal(时序)、Comparison(比较)、Contingency(偶然)、Expansion(扩展),第二层为以上四种关系类型的扩展,共有 10 种子类型。

### 3.3 事件关系检测

事件关系检测是一种针对“事件间存在何种逻辑关系类型”进行自动判定的深层事件关系判定任务。事件关系检测通过解析文本结构或语义特征,对文本中描述不同自然事件的文

本片段（包括短语、子句、句子和段落）给出明确的语义关系或逻辑关系标签（如“因果”、“时序”、“扩展”和“对比”等）。该任务框架如图 1 所示，其中，该任务的输入为已知相关的两个事件，通过分析“事件对”的逻辑联系，输出该“相关事件对”的逻辑关系类型。

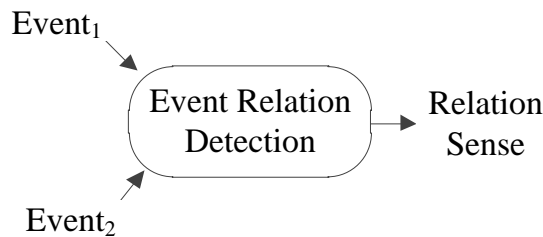


图 1 事件关系检测任务框架

## 4 动机

本文借助篇章分析任务实现事件关系检测。本章介绍篇章分析任务，以及框架语义和借助框架语义构建事件场景的方法。

### 4.1 篇章分析

篇章分析的核心任务是判定“论元对”之间的语义或修辞关系。目前篇章关系分析分别针对显式篇章关系和隐式篇章关系进行研究，且显式篇章关系和隐式篇章关系分析的性能相差较大。

显式篇章关系可直接根据显式连接词推断篇章关系，并且能够取得较优的性能，相关研究对于显式篇章关系最终分类的精确率可达 93%<sup>[7]</sup>；而隐式篇章关系的分类性能则相对较低，隐式篇章关系检测的主体研究主要分为基于关系特征的机器学习方法和基于概率统计的方法<sup>[8][9][10]</sup>。显式篇章关系和隐式篇章关系分析的性能相差较大的原因在于，显式篇章关系以显式连接词为主要线索，借助连接词与篇章关系的一一映射，能够较精确地推断特定篇章关系，比如，PDTB 关系样本库中，连接词“*Because (因为)*”指向“*Contingency.Cause (偶然因果)*”关系的概率为 100%。隐式篇章关系由于缺少显式连接词信息，相关研究通过结合句法、语义、相关领域知识等构建上下文信息，推断隐式“论元对”的逻辑关系。实验表明，篇章的上下文信息仍存在不确定性，这种上下文信息不确定性、句子结构的复杂性以及语义关系的歧义性，使得目前针对隐式篇章关系研究的性能总体较差，因此，借助篇章关系分析检测事件关系，必须突破隐式篇章关系检测这一关键难点。

### 4.2 框架语义

本文采用 FrameNet（框架语义）框架语义描述事件场景。FrameNet 是由美国加州大学伯克利分校构建的基于框架语义学 (Frame Semantics) 的词汇资源，对词语意义和句法结构研究提供一种理论框架，框架语义知识库通过框架描述单词的释义，即词语背后隐藏的概念结构和语义等信息。从而，框架语义能够形成特定场景（包括事件、状态、关系或实体）的概念表述。在框架语义知识库中，对句子的框架语义标注是一种类似于“谓词-论元”结构的“目标词-框架语义” (Target-Frame) 结构。每个句子可能包含一个或多个“目标词-框架语义”结构，例如例 (2) 描述的事件

(2) Evt1 : “*Irene left Sherlock in a coma with the Mickey Finn after getting the code of the safe box.*”

(译文：“艾琳得到保险柜的密码后，利用混有麻醉的酒让夏洛克昏迷”)

通过框架语义的分析，能够得到如下“目标词-框架语义”结构：“*Mickey Finn (混有麻醉的酒) - Drag (药物)*”、“*code (密码) - Message(信息)*”和“*safe box (保险箱) - Safe box (保险箱)*”，同时，本文将事件中的框架语义的组合定义为该事件的事件场景。不同的事件描述

存在相同的事件场景，例如例（3）中描述的事件：

(3) Evt2: “George remembered the beautiful lady asked him about the code of the bank account after drinking so much liquor.”

(译文：“乔治记得一位美丽的女士在他喝完很多酒之后询问他的银行密码”)

通过 FrameNet 解析，例（3）中的“liquor（酒）”同样触发了“Drag（药物）”框架语义，例（3）和例（4）中的其它“目标词-框架语义”结构如表 2 所示：

表 2 例（3）、例（4）中“目标词-框架语义”结果

	目标词 1	目标词 2	目标词 3
Evt1	Mickey Finn	Irene	code
Evt2	liquor	beautiful lady	code
框架语义	Drag	Person	Message

从表 2 中可以看出，尽管例（3）和（4）中描述不同的事件，但两个事件具有相同的框架语义，则认为两个事件描述了相同的事件场景。

### 4.3 事件场景

跨场景的事件关系检测方法的核心思想认为：相同或相似的事件场景，通过挖掘其特征，有助于事件关系类型的推理。例如有如下两个“事件对”：

(5) E1: He was **shot** by a terrorist. (译文：他被恐怖分子击中了)

E2: He unfortunately **passed away**. (译文：他不幸去世了)

(Relation = Contingency. Cause)

(关系=偶然因果)

(6) E1: **Bombs** went off at train stations. (译文：炸弹在火车站爆炸了)

E2: Many people were feared **dead**. (译文：担心很多人死亡了)

(Relation = ?)

(关系=?)

其中，已知例（5）中描述的两个事件为“偶然因果”关系类型。若仅考虑上述事件描述的篇章修辞、语法和文法等规律，难以利用例（5）的信息推理例（6）中“事件对”的关系类型。本文提出的跨场景的事件关系检测方法首先判断例（5）中的“事件对”描述了“袭击”场景和“死亡”事件场景，若定义例（5）中的事件关系类型“关系=偶然因果”为“袭击-死亡”场景的关系类型，由于例（6）同样描述了“袭击”场景和“死亡”事件场景，将事件场景的关系类型作为该“事件对”的关系类型，则推理得到例（6）中“事件对”的关系类型为“关系=偶然因果”类型。

## 5 基于跨场景推理的事件关系检测方法

本文致力于探究一种基于跨场景推理的事件关系检测方法，其核心思想是：具有相同事件场景的“事件对”，往往具有相同的事件关系类型。该方法分为三个主要部分：显式“事件对”的挖掘、事件场景关系类型映射以及事件关系判定。

另外，本文进一步提出一种“事件场景向量”的事件场景构建方法作为比较。

### 5.1 显式“事件对”的挖掘

本文首先通过连接词，从大规模语言学资源(Gigaword)中挖掘包含连接词的“事件对”，“事件对”的挖掘必须满足如下规则：

“PreGram + , connective + PostGram”、

其中，“PreGram”为前置论元，“connective”为连接词，“PostGram”为后置论元。通过连接词挖掘得到的“事件对”的关系类型，有该显式连接词唯一指定。PDTB中共定义182个连接词，而这些连接词与关系类型并非一一对应，为了实现“事件对”及其关系类型的精准抽取，本文仅仅选取PDTB中的Golden连接词。Golden连接词指该连接词在篇章中指向某一关系类型的概率较高。本文针对PDTB中连接词的分布，统计了各连接词指向某一关系类型的概率，如表3所示：

表3 PDTB中连接词的分布，统计了各连接词指向某一关系类型的概率

Relation Sense	Golden Connective (probability)
Exp.	or (97.25%), in fact (96.42%), indeed (96.03%), unless (98.95%), alternatively (100.00%)...
Con.	because (99.28%), for (100.00%), so (99.06%), so what (100.00%)...
Com.	however (99.33%), but (96.95%), though (100.00%)...
Tem.	now (100.00%), earlier (100.00%) afterward (100.00%)...

表3显式各连接词在四种篇章关系类型中出现的概率，例如连接词“*alternatively* (选择地)”指向“*Expansion* (扩展)”关系类型的概率为100%。本文仅选取指向某一关系类型概率大于80%的连接词作为Golden连接词。同时，将Golden连接词指向概率最大的关系类型，作为该Golden连接词的先验关系类型。进而将Golden连接词的先验关系类型作为通过Golden连接词挖掘得到的“事件对”的关系类型。

## 5.2 事件场景关系类型映射

本文将Golden连接词挖掘得到的“事件对”，进一步分析其包含的框架语义，利用框架语义构建“事件对”的“事件场景对”。同时，借助显式连接词对篇章关系的映射，构建“事件场景对”中两两框架语义与关系类型的映射，即“框架语义对-关系”，最终“框架语义对”的关系类型由大规模样本中的极大似然关系估计，方法细节如下：

**Step 1:** 利用SEMAFOR<sup>1</sup>工具分析“事件对”中的框架语义。例如，前置论元通过分析得到框架语义集合  $FrameSet_1: \{Frame_{11}, Frame_{12}, \dots, Frame_{1m}\}$ ;前置论元通过分析得到框架语义集合  $FrameSet_2: \{Frame_{21}, Frame_{22}, \dots, Frame_{2n}\}$ 。

**Step 2:** 将  $FrameSet_1$  和  $FrameSet_2$  中的框架语义两两组合，形成“框架语义对”： $\{Frame_{1i}, Frame_{2j} \mid i = 1, \dots, m, j = 1, \dots, n\}$ 。其中，“框架语义对”的关系类型为该“事件对”中显式连接词对应的关系类型。

**Step 3:** 对由Gigaword挖掘得到“事件对”，进行Step1到Step2操作。

由以上步骤得到的“框架语义对”存在多种关系类型映射，为了实现“框架语义对-关系”的一一映射，本文对每个“框架语义对”进行极大似然关系的概率估计。给定“框架语义对”及其在各个关系类型的概率分布  $P(r)$ ，选取分布概率最高的关系类型，作为该“框架语义对”的极大似然关系，如公式(1)所示：

$$EveMax(\mathcal{R}) = \underset{r}{Max} P(r) = \frac{n_r}{N} \quad (1)$$

$$r \in \{Syn, Asy, Cont, \dots, Ins\}$$

<sup>1</sup> <http://www.ark.cs.cmu.edu/SEMAFOR>

其中， $n_r$ 为“框架语义对”指向关系类型  $r$  出现的次数， $N$ 为“框架语义对”指向所有关系类型的次数。“框架语义对”的极大似然关系作为该“框架语义对”的先验关系类型。

### 5.3 事件关系判定

本文利用“事件场景”的关系类型对事件关系进行推理。针对待测“事件对”，同样利用框架语义构建其“事件场景对”，进而将已获得的“事件场景对”的先验关系类型，作为该“事件对”的关系类型。具体步骤如下：

**Step 1:** 针对每一个待测“事件对”，首先分析其包含的框架语义，分别得到事件  $event_1$  的框架语义集合  $FrameSet_1: \{Frame_{11}, Frame_{12}, \dots, Frame_{1m}\}$ ，以及事件  $event_2$  的框架语义集合  $FrameSet_2: \{Frame_{21}, Frame_{22}, \dots, Frame_{2n}\}$ 。

**Step 2:** 将  $FrameSet_1$  和  $FrameSet_2$  中的框架语义两两组合，形成“框架语义对”。

**Step 3:** 利用“框架语义对”的关系类型推理事件场景的关系类型，一个“事件场景对”包含多个“框架语义对”，将所有“框架语义对”指向的关系类型中，选取出现最多的关系类型，作为该“事件场景对”的关系类型。

**Step 4:** “事件场景对”的关系类型将作为“事件对”的关系类型。

同一事件场景可以概括不同的事件描述，而这些不同的事件均为相同的事件类型。事件场景并非孤立，事件场景之间存在逻辑关联性，因此事件场景能为事件关系分析和推理提供重要线索。本文重复利用事件场景的关系，推理事件之间的关系类型。

### 5.4 事件场景向量

为了与上述方法进行比较，本文同时提出一种事件场景向量的事件场景构建方法。该方法包括两部分：事件场景向量的构建以及事件关系类型预测。

该方法首先构建事件场景的场景向量。针对 5.1 中挖掘得到的“事件对”以及待测“事件对”，通过 SEMAFOR 分析其包含的框架语义集合，将该集合中的框架语义表示成空间向量，将该向量作为事件的场景向量，各维度将作为该场景向量的组成成分。每个“事件对”形成“事件场景对”。将待测“事件对”与挖掘得到的“事件对”样本进行相似度匹配，其中相似度的计算为“事件场景向量”间的 cosine 值，匹配过程如图 2 所示：

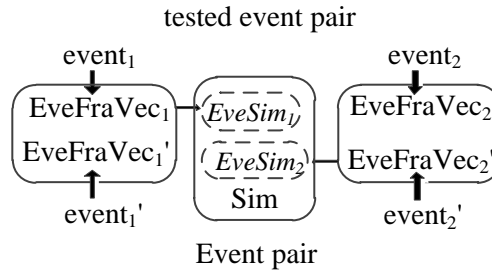


图 2. “事件场景向量”匹配

其中， $EveFraVec_1$  为前置事件的“事件场景向量”， $EveFraVec_2$  为后置事件的“事件场景向量”，分别得到前置事件的相似度  $EveSim_1$  和后置事件的相似度  $EveSim_2$ ，将两个相似值进行合并，合并后的值为该“事件场景对”的相似度，相似度合并如公式 (2) 所示：

$$Sim = \frac{EveSim_1 + EveSim_2}{2 \cdot e^{|EveSim_1 - EveSim_2|}} \quad (2)$$

其中，分子为两个相似度之和，同时，为了平衡两个相似度的值，将相似度之和除以两个相似度的差值，为确保分母的值不为 0，采用相似度差值的指数形式。

通过待测“事件对”与样本“事件对”的相似度计算，选取与待测“事件对”相似度最高的样本“事件对”作为待测“事件对”的平行“事件对”，该平行“事件对”的关系类型即为待测“事件对”的关系类型。

## 6 实验与结果

本章首先介绍实验设置，包括本文的语料标注、评价方法以及实验系统。同时介绍本文的实验结果及对结果的分析。

### 6.1 实验设置

本文选取 FrameNet-1.5 中的新闻语料作为标注数据集，标注人员对每篇新闻中描述的事件以及事件见的关系类型进行标注。该数据集由两名具有一定领域知识标注人员进行标注。同时，本文利用 Kappa 值对标注结果的可用性进行评价，最终标注的 Kappa 值为 0.78。

表 4 标注的关系类型的分布

Exp.	525 (54.24%)	List.	205 (21.18%)
		Prog.	82 (8.47%)
Com.	126 (13.02%)	Rest.	64 (6.61%)
		Cont.	70 (7.23%)
Con.	209 (21.59%)	Cond.	17 (1.76%)
		Conc.	56 (5.79%)
Tem.	108 (11.16%)	Asyn.	94 (9.71%)
		Sync.	14 (1.45%)

本文共标注了 828 个事件以及 968 个“事件对”及其关系类型，其中出现次数最多的关系类型为“*Expansion.List* (扩展.并列)”。标注的“事件对”及其关系类型的分布情况如表 4 所示。本文采用 Accuracy 值作为评价方法，其计算公式如下所示：

$$Accuracy = \frac{TruePositive + TrueNegative}{All} \quad (3)$$

该评测方法早期应用于二元分类精度的评测。例如，需要评价“事件对”是否为“*Expansion*”类型，其中，*All* 为待测“事件对”的总个数，*TruePositive* 为本身具有“*Expansion*”关系且系统判定其具有“*Expansion*”关系的“事件对”个数；*TrueNegative* 为本身不具有“*Expansion*”关系且系统未判定其具有“*Expansion*”关系的“事件对”个数。当该评测方法用于多元关系评价时，将 *TrueNegative* 设置为恒定值 0，只检验每个待测“事件对”是否判定为正确的关系，即只计算 *TruePositive* 指标与 *All* 的比值。本文首先定义了三个实验系统：

- System1(Baseline)：该系统遵循 5.1 至 5.3 中的事件场景构建方法，区别在于，构建事件场景并非使用框架语义，而直接将事件描述中的词汇作为特征，构建事件场景
- System2：遵循 5.1 至 5.3 中，利用“框架语义对”的事件场景构建方法。
- System3：遵循 5.4 中，利用“事件场景向量”的事件场景构建方法。

然而，待测“事件对”中的关系类型分布于从 Gigaword 中挖掘得到的“事件对”中的关系类型分布不一致，为了使两者的分布情况相同，本文采用重采样技术，扩展挖掘得到的“事件对”中关系类型较小的“事件对”，使得两者的关系类型分布情况相同。因此，本文又扩展了两类系统：

- System4：在 System2 方法的基础上，增加重采样技术。
- System5：在 System3 方法的基础上，增加重采样技术。



## 6.2 实验结果及分析

表 5 为以上五个系统针对事件关系检测结果的性能, 包括分布对四个大类及十个小类的 Accuracy 值:

表 5 系统的 Accuracy

System	Acc_class(%)	Acc_subClass (%)
Baseline	39.44	19.15
System1	45.68	18.42
System2	42.36	15.92
System3	<b>54.21</b>	<b>22.48</b>
System4	47.62	17.89

实验结果可以看出, *Baseline* 采用的使用事件描述中词汇作为特征的方法得到的性能最低, *System1* 和 *System2* 采用事件场景作为特征的方法高于 *Baseline* 的方法, 同时, 采用重采样的方法对结果的精确率也有所提高, 其中 *System3* 相对于 *System1* 在四大类的结果提高了 8.53%, 十小类的结果最高提高了 4.06%; *System4* 相对于 *System2* 在四大类的结果提高了 5.26%, 十小类的结果最高提高了 1.97%。*System3* 的性能最高。

表 6 System4 和 System5 针对每一大类的 Accuracy 及识别个数

		Exp.vs Other		Com.vs Other		Con.vs Other		Tem.vs Other		Four-way
System4	Acc.	55.35%		86.78%		78.04%		88.24%		54.21%
	Num	TPos	TNeg	TPos	TNeg	TPos	TNeg	TPos	TNeg	TPos
		498	34	0	834	21	729	2	846	521
System5	Acc.	54.73%		88.66%		75.34%		82.62%		47.62%
	Num	TPos	TNeg	TPos	TNeg	TPos	TNeg	TPos	TNeg	TPos
		419	851	1	851	20	704	19	775	459

本文进一步分析 *System4* 和 *System5* 针对每一类事件关系类型的识别性能。表 6 显式 *System4* 和 *System5* 分别对每一大类的精确率, 以及正确识别的个数 (*Num*), 包括正确识别为目标类型的个数 (*TPos*) 和正确识别为非目标类型的个数 (*TNeg*)。

表 7 System4 和 System5 针对每一小类的 Accuracy 及识别个数

subClass	Acc_subClass	TPos	TNeg	subClass	Acc_subClass	TPos	TNeg
List.	26.74%	196	61	List.	57.54%	89	464
Prog.	91.47%	0	879	Prog.	69.61%	26	643
Rest.	92.72%	0	891	Rest.	92.82%	0	892
Cont.	92.81%	0	892	Cont.	94.17%	0	905
Cond.	98.13%	0	943	Cond.	99.69%	0	958
Conc.	9.96%	0	903	Conc.	95.73%	0	920
Cause.	79.50%	19	745	Cause.	76.69%	18	719
Inst.	81.58%	0	784	Inst.	77.73%	9	739
Asyn.	89.49%	1	859	Asyn.	83.66%	17	787
Sync.	98.54%	0	947	Sync.	100.00%	0	961

表 7 为 *System4* 和 *System5* 分别对每一小类的精确率, 以及正确识别的个数, 从表中可以看出, 本文的方法针对小类别的检测效果较差, 体现了事件关系检测任务仍然比较难, 需要更进一步的研究。

## 7 总结

本文首次针对事件关系检测研究建立了一套研究体系, 包括任务定义、关系体系划分、语料采集与标注、评价方法等。同时, 提出了一种跨场景推理的事件关系检测方法, 该方法构建事件的事件场景, 以此作为特征, 实现事件的逻辑关系检测。实验结果显示, 构建事件场景的方法优于直接采用文本特征的方法。今后的工作在于, 根据事件场景元素的重要程度, 为场景元素设置程度, 从而更准确的描绘事件的场景特征。

## 参考文献

- [1] C. J. Fillmore, C. Johnson, and M. Petruck. Background to FrameNet, *International Journal of Lexicography*. 16.3 (2003): 235-250.
- [2] C. Hashimoto, K. Torisawa, J. Kloetzer, etc al. Toward Future Scenario Generation: Extracting Event Causality Exploiting Semantic Relation, Context, and Association Features[C]. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*. Baltimore, 2014.
- [3] T.Chklovski, P.Pantel. Global path-based refinement of noisy graphs applied to verb semantics[C]. *Toward Future Scenario Generation: Extracting Event Causality Exploiting Semantic Relation, Context, and Association Features*, Jeju Island, Korea, 2005: 792 – 803.
- [4] P.Pantel, M.Pennacchiotti. Espresso:leveraging generic patterns for automatically harvesting semantic relations[C]. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, Sydney, Australia, 2006: 113–120.
- [5] Z. S. Harris. *Mathematical Structure of Language*. New York, 1968.
- [6] D. Lin, P.Pantel. Discovery of Inference Rules from Text[C]. In *Proceeding of the 7th ACM SIGKDD*, San Francisco, California, USA, 2001: 323-328.
- [7] E. Pitler, M. Raghupathy, H. Mehta, etc al. Easily identifiable discourse relations[C]. *Proceedings of the 22nd International Conference on the COLING*, 2008: 87-90
- [8] P. E, R. M, N. HM, etc al. Easily identifiable discourse relations. In *Proc. of the 22nd International Conference on Computational Linguistics (COLING 2008)*, Posters, Manchester, UK, 2008. 87-90.
- [9] *The Penn Discourse Treebank 2.0 Annotation Manual*, 2007
- [10] Y. Hong, X. P. Zhou, T. T. Che, etc al. Cross-Argument Inference for Implicit Discourse Relation recognition[C]. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management(CIKM 2012)*,p295-304.