

---

# 现代汉语时间事件关系特点研究

**【摘要】** 本文基于中文时间事件关系分类构建了适用于汉语的时间事件关系形式化标记语言。首先，结合汉语表达时间关系的特点总结了 23 类中文时间事件关系；其次，构建了中文时间、事件关系研究的形式化标签体系 (CTimeML)，该体系提供了一种表达事件和时间、事件与事件、时间与时间之间关系的元数据标注语言。最后，对比了与现有时间标注体系的不同，并在部分语料上进行了标注尝试，研究表明已构建好的标签体系可覆盖大部分文本上可捕捉到或可简单推理得到的时间事件信息。

**【关键词】** 中文时间关系标注；时间关系；时间与事件

## A study on the characteristics of time event relationship in modern chinese

**Abstract :** Based on the classification of Chinese time event relation, this paper constructed a formal markup language for Chinese time event relation. Firstly, 23 kinds of Chinese time event relations are summarized based on the characteristics of Chinese expression of time relations. Secondly, the Formal labeling system (CTimeML) is constructed, which provides a metadata annotation language to express the relationship between event and time, event and event, and time and time. Finally, the paper compares the differences with the existing time labeling system, and tries to label some corpora. The research shows that the constructed labeling system can cover most of the time event information that can be captured or simply inferred from the text.

**Key words:** CTimeML; Time relationship; Time and Event

## 一 引言

时间信息在人类认知中一直都占有非常重要的地位, 古往今来被文人墨客千古传颂, 在语言学和自然语言处理领域, 对时间的认识也在不断深化, 特别是在自然语言处理任务中, 基于文本的时间解析、事件抽取等研究正如火如荼, 且在机器问答系统和文本抽取领域有着重要的应用。基于中文文本的时间解析或时间识别抽取任务也有相关研究, 并取得一些进展, 但是由于中文时间系统纷繁复杂, 且缺乏大规模的时间标注语料库, 这些研究仍处于初级阶段。本文主要在前人研究基础上, 提出基于中文的时间关系标注体系。该标注体系主要试图解决文本中呈现的事件发生时间, 事件持续了多长时间和文本中事件的发生顺序等。

## 二 时间事件关系研究综述

关于时间系统的研究, 国外最早是德国哲学家 Reichenbach(1947)用 S(说话时间 time of speech)、E(事件时间 time of event)、R(参照时间 time of reference)说明了语言的“时”和“体”的特点。参照时间 R 相对于说话时间 S 有三个位置, 在 S 前的是过去时( $R < S$ ), 和 S 在同一位置的是现在时( $R = S$ ), 在 S 之后的是将来时( $S < R$ )<sup>[1]</sup>。事件的时间相对 R 也有三个位置, 在 R 之前( $E < R$ ), 与 R 在同一位置( $E = R$ ), 在 R 之后( $R < E$ ), 它们对应于不同的体; Bruce(1972)提出 7 种相对时间关系: before(A,B)、after(A,B)、during(A,B)、sametime(A,B)、overlap(A,B)、contain(A,B)、erlapped(A,B), 其中 A,B 分别表示任意事件和时间<sup>[2]</sup>; 1984 年 Allen 在 Bruce 基础上进行了扩充, 提出了 13 种基本时间关系 before, overlap, meet, equal, start, finish, during 和逆关系; Comier(1976)提出“时制”表示情状发生的时间, 表现为该时间和说话时间或另一参照时间在时轴上的相对位置<sup>[3]</sup>; Smith 将语言的时间系统分成说话时间、参照时间和情状时间<sup>[4]</sup>。Vendel 根据动词的时间特征给谓词所描写的情状分类时运用了三个参数, 分别是有界(telic)或无界(atelic)、瞬时(instantaneous)或延时(durative)、动态(dynamic)或静态(stative)所做的动词分类。TIMEX 作为最早的标注时间短语的方案是 2000 年 Mani 和 Wilson 提出<sup>[5]</sup>。美国国防部高级研究计划局(DARPA)“跨语音信息检索、信息抽取和自动文摘计划”(TIDES)和自动内容抽取计划(ACE)在 TIMEX 基础上制定了 TIMEX2 方案, 用来标注英文时间短语, 主要回答事件什么时候发生, 持续了多长时间, 发生的频率如何等问题<sup>[6]</sup>。TimeML (在自然语言文本中表示事件和时间表达式的一种规范语言)是在一个问答系统 AQUAINT 项目里首次被提出来的。TimeML 标准是在 1984 年 Allen 提出的 13 种时间关系的基础上提出了 14 种时间关系, 分别为 simultaneous、identical、before、after、immediately before、immediately after、including、being included、during、being during、beginning、being begun、ending、being ended<sup>[7]</sup>。2007 年 8 月将这一标准提交 ISO, 2009 年 3 月被批准为国际标准, 即 ISO-TimeML。在 TimeML 出现之前主要是基于规则的方法进行时间事件关系的识别。例如 Lascarides、Asher<sup>[8]</sup>和 Oberlander<sup>[9]</sup>根据话语关系和时序关系之间的对应关系就可以找到事件之间的时序关系; Hitzeman 基于 HPSG 在分析话语时序结构时着重考虑了时态(tense)、体态(aspect)、事件副词、修饰关系等因素<sup>[10]</sup>; Dorr 等人采用基于约束的方法对时态、体态及连接词进行了分析, 确定了识别时序关系的约束语言理论模型 CONGEN<sup>[11]</sup>。Li Wenjie et. 采用基于规则的方法识别同一句子中两个事件的时间关系<sup>[12]</sup>。随着 TimeML 标准的提出, TimeBank 语料库的出现和自 2007 年开始举办的一系列 TempEval 评测活动, 使文本时间关系解析等相关研究工作获得了较大推进, 产生了一些性能良好的英文时间表达式识别方法, 很多学者基于机器学习的方法和错误驱动的方法等统计方法进行相关研究。例如, Mani 等人在语料库 TimeBank 中利用机器学习方法抽取时态、时间介词、话语连接词等特征词, 并进行了初步的新闻事件推理研究<sup>[5]</sup>; Chambers 年在 Mani 基础上进一步提高了基于机器学习方法构造的文本时间关系分析模型<sup>[13]</sup>; Yoshikawa 提出用马尔可夫模型来进行事件时间关系识别<sup>[14]</sup>; 孙辉基于机器学习方法构建了英文事件时序关系识别模型<sup>[15]</sup>。

国外对时间事件关系的理论和技术取得了初步进展,国内关于时间的研究也浩如烟海。一直以来,很多学者对汉语中有无“时范畴”和“体范畴”的问题争论不休,王力认为,汉语不着重表示事情是何时发生的,而重在表示事情所经过时间的长短、是否开始或完成,即汉语中更注重“情貌”,他将汉语中的情貌分为了7种(这里的情貌指的是“体”)<sup>[16]</sup>。高名凯概括了汉语中的六种体,其中包括进行体“着”、完成体“了、过、好”等<sup>[17]</sup>;邓守信首次借鉴了 Veldel 的分类,首先从时间结构的角度探讨活动、完结、达成句和状态句四种不同的语境在汉语语法的表现方式,从而定出汉语谓语本身所含的时间结构特点<sup>[18]</sup>。20世纪80年代以后,关于时范畴的研究出现了多元格局。帅志嵩总结了这一阶段的时范畴的研究,将其分为三种不同类型:以陈平、李临定、裘千炎为代表的“无标记有时论”。例如,陈平在 Reichenbach 的基础上将汉语的时间范畴分为“先事过去时、简单过去时、后事过去时、先事现在时、简单现在时、后事现在时、先事将来时、简单将来时、后事将来时”九个范畴。这里的“先事”可理解为事件发生在参照时间之前,“后事”可理解为事件发生在参照时间之后<sup>[19]</sup>;龚千炎根据陈平的介绍也建立了“先事—过去时、先事—现在时、先事—将来时、当事—过去时、当事—现在时、当事—将来时、后事—过去时、后事—现在时、后事—将来时”九个范畴<sup>[20]</sup>;以张济卿、左思民、李铁根、陈立民、林若望等学者为代表的“时体混合有时论”。例如,李铁根认为现代汉语分已然和未然不作过去、现在、将来的划分<sup>[21]</sup>;以王松茂、马庆株为代表的学者所持的“时体分别论”。戴耀晶认为汉语“有体无时论”是因为汉语中没有明显的“时”标记,他提出“时”是观察事件时间构成的方式。“时”意义涉及到“过去、现在、将来”等含有指示表达性或索引表达性的具体时间<sup>[22]</sup>。石毓智以一个完整动作从开始到结束的不同发展阶段作为动作内部过程的框架来刻画“了、著、过”的表义特点及其对有不同语义特征的动词的选择性<sup>[23]</sup>。郭锐指出陈述性成分最重要的特征就是其时间性。动词的时间性可以看成由内在时间性和外在时间性两个方面构成的一个完整的系统。内在时间性指的是陈述性成分的内在过程,陈述性成分所表示的状况有随时间展开的内部过程,包括起点、终点和续段三个要素,三个要素的有无和强弱形成了谓词过程结构的不同;外在时间性指的是陈述性成分表示的状况与外部时间过程的关系,此外他还区分了谓词性成分过程和非过程的两种外在时间类型<sup>[24]</sup>;郭锐在此基础上进一步提出了汉语谓词性成分的时间参照分为内部参照和外部参照。内部参照指的是句子内部多个状况之间的时间关系,以句中后续动作发生的时间为参照;外部参照指的是句子所表示状况与外部世界的关系,以外部世界的自然时间过程中的某一刻为参照<sup>[4]</sup>;陈前瑞认为虚化的“着、了、过、来着”等所表示的各种语法意义,它是事件内在的时间结构的表现<sup>[25]</sup>;孙英杰在陈前瑞建立的汉语四层级体貌系统的基础上构造了三分体系统主要包括动词词汇体、述谓体和语法体,孙也认为现代汉语中没有标记绝对时的语法手段,现代汉语时的系统主要是依靠相对时制,可以由词汇成分表达,也可以由语法成分来表达<sup>[26]</sup>。孙朝奋认为界点是汉语动词系统的核心语法概念,他用构式语法的观点阐释了完成体“了”的功能,即除了表示有界情状之外,还表示完成界点的位置,他的这一观点可以帮助我们判断汉语谓词时间结构<sup>[27]</sup>。本文认为汉语的时间系统包括“时”和“体”两大系统,观察事件在时间轴上的位置即为“时”,观察事件内部在时间轴上的流动即观察事件的开始、结束、完成、持续等状态即为“体”。汉语动词具有内在和外在时间性之分,由于时间性系统是汉语动词界限分析的重要理论来源,因此汉语动词也可从内在界限和外在界限两方面来考虑。

相较于英文时间表达式识别工作,中文相关工作起步较晚。最开始也是使用规则的方法识别时间信息,如邬桐提出“时间基元”概念将时间单位细化并用于规则中;宋国民利用通过构造规则通过时间表达式匹配的方法识别时间信息。有些学者使用统计机器学习的方法,LiWenjie 和 Wong Kam-Fai 鉴于之前方法的缺陷提出了使用机器学习的方法来进行汉语文本时间关系解析的研究<sup>[12]</sup>。Cheng Yuchang 等在 TimeML 标准基础上建立了中文的事件时间表达关系识别的语料库<sup>[28]</sup>。贺瑞芳等和潘越群都采用了错误驱动的方法改造模型,后者将错误驱动方法融合到依存分析方法中,依次使用了 SVM、CRF 方法以及改进 CRF 方法并将其进行了对比<sup>[29]</sup>。沈思等基于清华汉语树库,利用条件随机场,构建时间表达式知识抽取模型<sup>[30]</sup>。张春菊等基于触发词汇和规则模型相结合的中文文本中时间信息解析方法<sup>[31]</sup>。由于汉语时间系统复杂,没有大规模的时间标注语料库,所以研究一直受限。本文通过调研前人相关研究工作,提出新的标注体系,主要以事

件为核心研究汉语的时间性特点，并将汉语的时间性特点分为内在时间性和外在时间性。本文认为内在时间性特点是研究汉语谓词内部的过程结构，即研究事件的开始、发展、高潮和结束等内部过程结构。外在时间性特点是研究事件相对于参照时间点的时间性信息。而汉语中的时间参照主要是相对时间参照，主要通过考察事件时间和参照时间说明汉语的时间参照。

### 三 汉语时序关系分类

#### 3.1 时序关系类型

关于时间关系的分类，陈振宇（2007）事件投射到时间认知窗口”这一理念，将事件可能有的时间结构通过排列组合得到了 20 种不同的时间结构，并将进行了形式化的处理<sup>[32]</sup>。TIMEX2 标注规范主要解决以下三个问题：(1) 确立标注对象，即确定什么样的时间信息需要标注；(2) 其次确立标注方式，即如何捕获时间词的信息；(3) 确定时间表达式的标注范围。TimeML 旨在解决以下四个问题：(1) 事件的时间戳（识别事件并锚定时间）；(2) 对事件进行排序；(3) 根据上下文交代的特定时间表达式对事件进行推理（时间函数，例如“上周”和“两周前”）；(4) 推理事件的持续时间<sup>[7]</sup>。TIME2 和 TimeML 标注规范都是针对英语语料设计的，而汉语时间系统不同于形态特征丰富的英语语言，以往的时间关系标注体系可以捕捉到文本明显的时间信息，基于简单的推理也可以推知部分事件的时间信息，但是对于文本描述不清楚的时间信息不能给出清晰的界定范围。本文在前人研究的时间关系体系的基础上，同时考虑汉语的时体系统重新对时间、事件关系进行了分类。本文构建的时间事件关系分类体系希望将文本层面描述不清晰的时间信息尽量描述清楚。主要解决以下几个问题：(1) 确立标注对象，即可以定位到时间轴上的事件和时间表达式；(2) 确立标注方式；(3) 确定时间表达式的标注范围；(4) 对文本中呈现的事件顺序进行排序；(5) 根据上下文背景推理事件时间或推理事件的持续时间；(6) 对事件内部在时间轴上的状态进行标注，即标注事件的开始、结束、完成、持续等状态<sup>[16]</sup>。中文时间标注体系主要包括时间关系定义和标签体系界定，其中汉语的时间关系主要分为外部时间关系和内部时间关系。外部时间关系主要指事件投射到时间轴上的时间点。时间点可分为参考时间点、说话时间点和事件发生时间点。外部时间关系类型主要包括事件与事件、时间与时间和时间与事件之间的关系，其中时间与事件的关系本文参考了 TimeML 的关系分类，并在此基础上进行了拓展。TimeML 的时间事件关系主要分为 14 类，即“事件发生在该时间之前 (BEFORE) 与之后 (AFTER)”“事件发生在当前时间 (SIMULTANEOUS)”“事件 A 在时间段 B 内一直持续不断的发生 (DURING)”“事件发生在该时间之前并持续到当前时 (IBEFOR)”“在事件 1 在事件 2 发生之前立刻发生”“在事件 1 在事件 2 发生之后立刻发生”“事件发生在该时间后并持续一段时间 (IAFTER)”“包含 (INCLUDES) 与被包含 (ISINCLUDED)”“开始 (BEGINS) 与被开始 (BEGUNBY)”“结束 (ENDS) 与被结束 (ENDED BY)”。本文将 CTimeML 外部时序关系类型分为两大类，34 小类，新增了内部时间关系和事件与事件间接导致时序变化的时间关系，内部时间关系主要标注事件内部在时间轴上的流动，即事件的“开始”“终止”“高潮”“延续”“重新开始”等状态，主要通过标注事件的状态实现的。如表 1 所示。

表 1 时间事件关系分类

时 序 关系	分类	释义	英文名称
外 部 时 序 关系	时 间 与 事 件	事件 A 发生时间 B 之前，即 B(A)	BEFORE
		事件 A 发生时间 B 之后，即 A(B)	AFTER
		事件 A 发生的时间即为当前时间 (B)，即 A=B	SIMULTANEOUS
		事件 A 在时间段 B 内一直持续不断的发生	DURING
		事件 A 发生在时间 B 之前并持续到当前时间，即 B(A, +∞)	DURINGBEFORE
		事件 A 发生在时间 B 后并持续一段时间，即 A(A, B)	DURINGAFTER

		包含与被包含	INCLUDE/ISINCLUDE
		开始与被开始	BEGIN/BEGINBY
		结束与被结束	END/ENDEDDBY
	时 间 与 时 间	时间 A 在时间 B 之前即 B(A)	BEFORE
		时间 A 在时间 B 之后即 A(B)	AFTER
		时间 A 与时间 B 同时即 A=B	SIMULTANEOUS
		时间 A 在时间 B 之前, 且一直持续到时间 B, 即 B(A, +∞)	DURINGBEFORE
		时间 A 在时间 B 之后, 且从一直持续到时间 B, 即 A(A, B)	DURINGAFTER
	事 件 与 事 件	事件发生时间 A 在事件发生时间 B 之前, 即 B(A)	BEFORE
		事件发生时间 A 在事件发生时间 B 之后, 即 A(B)	AFTER
		事件发生时间 A 与事件发生时间 B, 即 A=B	SIMULTANEOUS
		事件 A 在事件 B 发生的时间段内一直持续不断的发生	DURING
		事件发生时间 A 在事件发生时间 B 之前, 且一直持续到 B, 即 B(A, +∞)	DURINGBEFORE
		事件发生时间 A 在事件发生时间 B 之后, 且从一直持续到 B, 即 A(A, B)	DURINGAFTER
		事实性	FACTICITY
		反事实性	COUNTERFACTICITY
		叙实性	FACTIVITY
		反叙实性	COUNTERFACTIVITY
		期望性	EXPECTATION
		反期望性	COUNTEREXPECTATION
内 部 时 间 关 系		开始 延续 终止 重新开始 无内部时间关系	INITIATION CONTINUATION TERMINATION REINITIATION NONE

### 3. 2 各类时序关系释义

本文将时序关系概括为两大类即外部时间关系和内部时间关系, 其中外部时间关系指的是事件与事件(12类)、时间与时间(5类)、时间与事件(12类)之间的关系; 内部时间关系指的是事件动词的体信息, 即该事件的状态信息, 主要包括“开始”“延续”“终止”“重新开始”“无内部时间关系”五类, 共计 34 小类。有的时序关系类型与其他关系共享同一种标签。例如, “事件 A 发生时间 B 之前”“时间 A 发生时间 B 之前”和“事件 A 发生事件 B 之前”共享同一种标签“之前(BEFORE)”, 最终构建了 23 种描述时序关系类型的标签, 包括“之前(BEFORE)”“之后(AFTER)”“在当前时间(SIMULTANEOUS)”“A 在 B 时间段内持续不断的发生(DURING)”“在该时间之前并持续到当前时间(DURINGBEFORE)”“在该时间后并持续一段时间(DURINGAFTER)”“包含与被包含(INCLUDE/ISINCLUDE)”“开始与被开始(BEGIN/BEGINBY)”“结束与被结束( END/ENDEDDBY)”“事实性与反事实性(FACTICITY/COUNTER\_FACTICITY)”“叙实性与反叙实性(FACTIVITY/COUNTER\_FACTIVITY)”“预期与反预期(EXPECTATION/COUTER\_EXPECTATION)”“开始(INITIATION)”“终止(TERMINATION)”“继续(CONTINUATION)”“重新开始(REINITIATION)”。下面选择部分标签进行介绍。

事件发生在该时间之前(BEFORE): 指的是事件发生的时间在当前时间之前。参考时间为当前时间(或着说话时间)。例如, “他昨天上午前往上海。”“前往”发生在“昨天上午”所以事件“前往”在参考时间点当前时间“今天”之前发生。

事件发生在该时间之后(AFTER): 指的是事件发生在当前时间之后。参考时间为当前时间。例如, “他明天上午前往上海。”“前往”发生在“明天上午”所以事件“前往”在参考时间点当前时间“今天”之后发生。

事件发生在当前时间(SIMULTANEOUS): 指的是事件的发生时间与当前参考时间点是同一时间点。例如, 他今天上午十点前往上海。“前往”发生在“今天上午十点”所以事件“前往”在参考时间点当前时间“今天”同时发生。

事件 A 在时间段 B 内一直持续不断的发生(DURING): 指的是事件在某一时间段内持续不断的发生。例

如，他今天上午一直在上课。“上课”发生在“今天上午”这一时间段内，所以事件“上课”在时间段“今天上午”内一直在持续不断的发生。

事件发生在该时间之前并持续到当前时间(DURINGBEFORE):指的是事件发生的时间在当前时间之前且一直持续到当前时间，参考时间为当前时间。例如，他从昨天开始定居北京的。事件“定居”发生在“昨天”，参考时间为当前时间且一直持续到当前时间。

事件发生在该时间后并持续一段时间(DURINGAFTER):指的是事件发生的时间在当前时间之后且在当前时间之后一直持续，参考时间为当前时间。例如，他从明天起定居北京。事件“定居”发生在“明天”，参考时间为当前时间“今天”，且在当前时间之后一直会持续。

同时(MEANWHILE):指的是事件就在当下参考时间点发生。例如，我现在去学习。事件“学习”就在“现在”这一时间点发生。

包含与被包含(INCLUDE/ISINCLUDE):指的是事件发生的时间包含在某一时间内或事件发生的时间点包含在某一事件发生时间内。例如，他今天上午前往上海。“前往”发生在“今天上午”这一时间点，包含在“今天”这一时间内。所以事件“前往”在参考时间点“今天”是包含与被包含的关系。

开始与被开始(BEGIN/BEGINBY):指的是事件发生的时间点在某一时间点开始或事件发生的时间点在某一事件发生时间点结束。例如，他昨天上午十点飞往上海。“飞往”发生在“昨天上午十点”，事件“飞往”在“昨天上午十点”这一时间点开始，所以事件“飞往”与“昨天上午十点”是开始与被开始的关系。

结束与被结束(END/ENDEDBY):指的是事件发生的时间点在某一时间段内结束或事件发生的时间点在某一事件发生时间内结束。例如，他昨天上午十点离开上海。“离开”发生在“昨天上午十点”，事件“离开”在“昨天上午十点”这一时间点结束，所以事件“离开”与“昨天上午十点”是结束与被结束的关系。

事实性与反事实性(FACTICITY/COUNTERFACTICITY):袁毓林(2020)提出语义上的反事实性表示其跟实际情况不同或相反的性质<sup>[33]</sup>。“她后悔嫁给了小李。”“她忘记去年飞过上海。”这里的“嫁给了小李”、“飞过上海”是真实发生的事件属于事实性小句，“后悔嫁给了小李”、“忘记飞过上海”也是真实发生的事件，也具有事实性。“她忘记去买礼物了。”“他阻止了离婚。”这里的“买礼物”、“离婚”是尚未发生的事件跟实际情况不同具有反事实性，“忘记买礼物”、“阻止离婚”是真实发生的具有事实性。一般“幸亏”、“幸好”、“可惜”、“后悔”、“忘记”、“阻止”等引导事实性的句子，“否则”、“不然”、“要不”引导反事实性小句。如“幸亏他起得早，否则就迟到了。”这里的“幸亏”引导的是事实性小句，“否则”引导的是反事实性的小句。

叙实性与反叙实性(FACTIVITY/COUNTERFACTIVITY):袁毓林(2020)语义上的叙实性指动词预设其宾语小句的真值的能力或性质<sup>[33]</sup>。“我知道他去了北京。”“去北京”这一事件是真实发生的，“知道”动词预测小句真值为真。因此“我知道他去了北京”属于叙实性。“小明怀疑小李去买酒了。”这里的“买酒”是否发生不可知。“他拒绝接受这个提议。”“接受提议”并未发生，因此该句具有反叙实性。

预期与反预期(EXPECTATION/COUTEREXPECTATION):“小明想要小李去买礼物。”属于预期。“他竟然还活着。”属于反预期。

事件和事件之间的关系除了时序的关系，还有事实性与反事实性、叙实性与反叙实性、预期与反预期之间的关系，这些关系一般出现在两个及以上动词描述事件的句子中，虽与时间没有直接的关系，但是间接的影响了事件的状态，即内部时间关系。时间内部关系指的是事件的体信息“开始(Initiation)”“延续(Continuation)”“终止(Termination)”“无内部时间关系(NONE)”等。例如“昨天下雨了。”“下雨”这一事件的状态是“终止”，需要采用事件时间内部关系标签标记，即通过事件体信息(aspect)属性标签和属性值“终止”(Termination)标记。

## 四 中文时间事件关系形式化表示

本文根据时间事件关系分类体系确立了关于中文时间关系标注的 7 个主要标签,分别为:事件标签(EVENT)、中文时间表达式标签(CTIMEX)、触发词(SIGNAL)、时序关系标签(TRLINK)、句间关系标签(SRLINK)、置信度(CONFIDENCE)和其他内容补充说明(CONTENT)。标签区分类型、属性及属性值。由于篇幅限制,这里仅介绍事件标签和时序关系标签。

### 4.1 事件标签(EVENT)

事件(EVENT)指的是可定位到时间轴上用来标注语义事件的元素。事件主要由动词构成,关于动词分类的相关研究已有很多,动词分类的角度也各有不同。例如,赵元任的《(1968)汉语口语语法》、王力先生、李临定(1990)和陆俭明(1991)从句法层面(1943 中国现代语法)将动词划分为及物动词和不及物动词两类;赵元任将及物和不及物动词区分动作、性质和状态动词,同时将“是”类、“有”类、“助动词”归入及物动词。黎锦熙(1924)、李临定(1985)、邓守信(1986)从语义层面对动词进行了分类,李认为动词具有动态和静态的功能,动态表示动作的进行,静态表示静止的持续状态;邓认为状态动词表示事件持续的状态,达成动词则表示动作的完成,完结类动词则表示动作的持续一定时间段后的结束。袁毓林(1991)根据动词进入祈使句的情况,首先把动词划分为述人动词和非述人动词,再进一步把述人动词划分为可控动词和非可控动词,可控动词又再进一步分为自主动词和非自主动词;龚千炎(1994)把汉语动词划分为静态动词和动态动词,静态动词又细分为关系动词、心态动词和状态动词,把动态动词细分为动作行为动词、心理活动动词、兼属状态的动作动词、终结动词和瞬间动词;陈昌来在《现代汉语动词的句法语义属性研究》(2002)中认为“现代汉语动词可以分为五类:动作动词、致使动词、心理动词、性状动词和关系动词。其中性状动词包括状态动词和性质动词两类性质动词。性质动词也就是性质形容词。邢福义(2016)把动词划分为行为动词(典型动词)、心理动词、历程动词、断事动词、使令动词和辅助动词。

本文中的动词指的是跟时间有关的动词,动词的分类主要分为静态动词、动态动词两大类。静态动词包括属性关系动词、心理认知动词、助动词,动态动词包括瞬间动词、持续动词、状态动词。本文关于静态动词和动态动词的划分标准是动词所描述的事件是否静止不变,例如“他是一名警察”、“他喜欢动物”、“这个房间干净了一段时间”、“他想去北京”、“他应该走了”中的“是”、“喜欢”、“干净”、“想”、“应该”静态动词所描述的事件状态是静止的,而“他踢了我一脚”、“他打碎了玻璃”、“雨下了三天了”、“他写一天作业了”中的“踢”、“打碎”、“下”、“写”所描述的事件状态是动态变化的。例如,“踢”经历了从“脚”抬起到“脚”落下的过程;“打碎”通过这一动作的实现描述了“玻璃”从完好无损到“碎”这一状态的变化;“下”也经历了雨持续不断从天空降落到地面的过程,“写”也描述了人通过不停的重复“写”这一动作实现“作业”从无到有的结果。动词在一定语境作用下构成事件,事件中的动词在该上下文下具有一定的属性信息,主要有词性(pos)、时态(tense)、极性(polarity)、体信息(aspect)、时间点(timePoint)、持续时间段(Duration)、情态(modality)。时态(tense)属性值包括过去、现在、将来,极性(polarity)属性值包括肯定还是否定,体信息(aspect)属性值包括开始、继续、终止、重新开始,模态属性值包括可能性(Probability)、意愿(Willingness)、能力(Ability)、必要性(necessity),时间点属性值包括瞬间发生的时间点(instantaneousPoint),起始时间点(beginPoint),终止时间点(EndPoint),持续时间段(Duration)属性值包括各个时间段。其中体信息指的是事件的内部时间关系类型(reltype)包括开始、延续、重新开始、终止。事件属性及属性值信息如下表 2 所示。

表 2 事件属性及属性值信息

事件属性	词性(pos)	时态(tense)	极性(polarity)	体信息(aspect)	时间点(timePoint)	持续时间段(Duration)	情态(Modality)
属性	名词(NOUN) 动词(VERB)	过去 现在	肯定 否定	开始 延续	起始时间点 (beginPoint)	X 秒 X 分	可能性 (Probability)

值	形 容 词 (ADJECTIVE)	将来		重新开始 终止 高潮 无 体 信 息 (NONE)	终 止 时 间 点 (endPoint) 瞬 间 时 间 点 (instantaneousPoint)	X 小时 X 天 X 月 X 旬 X 年...	意 愿 (Willingness) 能 力 (Ability) 必 要 性 (necessity)
---	----------------------	----	--	---------------------------------------	--	-------------------------------------	--

动态动词包括瞬间动词、持续动词、状态动词。相关动态动词所描述的事件过程即为动态事件，包括瞬间事件、持续事件。

瞬间事件(INS\_OCCURRENCE)由瞬间动词构成，属于动态动词，表示事件持续的时间非常短暂，事件的起始时间点和终止时间点从人的主观直觉上接近于同一时间点。例如，“他踢了一脚。”“呼吸三次。”“他咳嗽了一下。”“他正在戴手套。”中的“踢”“呼吸”“咳嗽”“戴”属于瞬间动词。瞬间事件具有瞬间时间点(instantaneousPoint)这一属性值。例如：

他<EVENT eid="e1" class="INS\_OCCURRENCE" pos="VERB" tense="PAST" polarity="肯定" aspect="终止" instantaneousPoint="t1">踢</EVENT>了我一脚。

持续事件(CONT\_OCCURRENCE)主要描述了事件持续的状态，主要由持续性动词和状态动词构成，如他写一天作业了。从上周四开始雨下了三天了。其中的“写”和“下”属于持续动词。状态动词描写了事件的状态，此类型的事件大多包括开始、初始、重新开始、重新启动、停止、取消完成、继续、保持等词语。例如“他今年夏天开始去北京读书。”中的“开始”。持续事件的属性有：词性、时态、极性、时间点(timePoint)、持续时间段(Duration)、体 (aspect) 和情态 (Modality)。时间点属性值包括起始时间点(beginPoint)、终止时间点(endPoint)，体属性值包括开始、高潮、终止、延续、重新开始。例如：

例(1)他在北京

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="延续" beginPoint="t1" endPoint="t2" Duration="三年">生活</EVENT>了三年。

例(2)他

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" beginPoint="t1" endPoint="t2" Duration="一天">写</EVENT>了一天作业，终于写完了。

例(3)

他<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="延续" beginPoint="t1" endPoint="t2" Duration="一天">写</EVENT>一天作业了，应该休息一下。

例(4)雨

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" beginPoint="t1" endPoint="t2" Duration="一天">下</EVENT>了一天，终于停了。

例(5)雨

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="延续" beginPoint="t1" endPoint="t2" Duration="一天">下</EVENT>一天了，还没停。

例(6)这本书他

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="继续" beginPoint="t1" endPoint="t2" Duration="三天">看</EVENT>了三天。

例(7)这个人

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="延续" beginPoint="t1" endPoint="t2" Duration="三天">来</EVENT>了三天。

例(8)这个人

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="延续"



beginPoint="t1" endPoint="t2" Duration="三天" >消失</EVENT>了三天。

静态动词包括属性关系动词、心理动词、和助动词。

情态助动词(MODALVERB)在北大教研室版现代汉语中又叫情态助动词、能愿助动词,表达这个事件的能力、意愿、必要性、可能性等方面的意义,属于静态动词。例如,“他明天可能会回家。”这里的“可能”为助动词,表达事件发生的可能性;“九月份我愿意飞去美国。”这里的“愿意”为助动词,表达意愿;“他今天能准时完成这项工作。”这里的“能”为助动词,表达能力;“他应该今天去上班。”这里的“应该”表示必要性。助动词的属性有词性、情态、类别等属性,其中情态的属性值包括可能性(Probability)、意愿(Willingness)、能力(Ability)、必要性(necessity)。例如:

例(9) 他

<EVENT eid="e1" class="MODALVERB" pos="动词" Modality="必要性">一定</EVENT>

<EVENT eid="e2" class="CONT\_OCCURRENCE" pos="动词" tense="将来" polarity="肯定">做</EVENT>这道题。

例(10) 他

<EVENT eid="e1" class="MODALVERB" pos="动词" Modality="可能性">可能</EVENT>

<EVENT eid="e2" class="CONT\_OCCURRENCE" pos="动词" tense="将来">做</EVENT>这道题。

例(11) 他

<EVENT eid="e1" class="MODALVERB" pos="动词" Modality="能力">能</EVENT>

<EVENT eid="e2" class="CONT\_OCCURRENCE" pos="动词" tense="将来">做</EVENT>这道题。

例(12) 他

<EVENT eid="e1" class="MODALVERB" pos="动词" Modality="意愿">想</EVENT>

<EVENT eid="e2" class="CONT\_OCCURRENCE" pos="动词" tense="将来">做</EVENT>

这道题。

心理认知类动词(COGNITIVEVERB)主要指的是心理类动词和言语行为类动词,如“相信”、“希望”、“渴望”、“害怕”、“考虑”、“打算”、“喜欢”、“讨厌”等。言语行为类动词包括“报道”、“宣布”、“曝光”、“承认”、“爆料”等动词。袁毓林(2020)从叙实性的角度看这类动词,认为“相信”、“害怕”这类动词是真假不定的,“希望”、“打算”描述的事件在目前来看是尚未实现的。这类动词大多表示了事件之间的真实性关系(见 4.5 SRLINK)<sup>[33]</sup>。

例(13) 他

<EVENT eid="e1" class="COGNITIVEVERB" pos="动词" tense="现在" aspect="延续">考虑</EVENT>

<EVENT eid="e1" class="COGNITIVEVERB" pos="动词" tense="将来" beginPoint="" endPoint="" Duration="三天">去</EVENT>北京呆三天。

此外还有一些形容词、状态词和名词也可以表示事件静态事件或动态事件,可归入“其他”类型。例如:

这个房间<EVENT eid="e1" class="other" pos="形容词" tense="过去" polarity="肯定" aspect="开始">干净</EVENT>了三天。

## 4.2 时序关系标签(TRLINK)

时序关系标签主要描述时间和事件、时间和时间、事件和事件之间的时序关系(TRLINK)。本文 3.1 节将时序关系类型分为了 2 大类, 34 小类, 23 种标签。主要通过 TRLINK 这一标签中的 relType 这一属性体现, 下面我们的一些例子来说明如何使用时序关系分类标签来标识汉语时间事件间的时序关系。限于篇幅原因, 时间信息本文暂不展开讲述, 我们这里将文本中简单的时间信息放入“CTIMEX”这一标签, 这一标签有“编号(tid)”、“类型(type)”、“时间值(value)”、“参照时间点(anchorTimeID)”属性等, 时间和事件关系的表达主要通过时间编号所对应的时间点和事件编号所对应的事件产生联系, 本文将时间与事件之间的关系都放入 TRLINK 这一标签的关系类型属性(relType)中。

根据 3.2 节的介绍, 时间与事件之间的关系包括: “事件发生在该时间之前(BEFORE)” “事件发生在该时间之后(AFTER)” “事件发生在当前时间 (SIMULTANEOUS)” “事件在某一时间段内一直持续不断得发生 (DURING)” “事件发生在该时间之前并持续到当前时间 (DURINGBEFORE)” “事件发生在该时间后并持续一段时间 (DURINGAFTER)” “包含与被包含 (INCLUDE/ISINCLUDE)” “开始与被开始 (BEGIN/BEGINBY)” “结束与被结束 (END/ENDEDDBY)”。如下述例句所示。

例(1) 他昨天上午前往上海。

他

<CTIMEX tid="t1" type="DATE" value="2020-05-13-AM" anchorTimeID="t0">昨天上午</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" timepoint="t1">前往</EVENT>

上海。

<TRLINK timeID="t0" relatedtoEvent="e1" relType="BEFORE"/>

例(2) 他明天上午前往上海。

他

<CTIMEX tid="t1" type="DATE" value="2020-05-15-AM" anchorTimeID="t0">明天上午</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" timepoint="t1">前往</EVENT>

上海。

<TRLINK timeID="t0" relatedtoEvent="e1" relType="AFTER"/>

例(3) 他现在前往上海。

他

<CTIMEX tid="t1" type="DATE" value="2020-05-14-10:00am" anchorTimeID="t0">现在</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-10:00am">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="现在" polarity="肯定" aspect="NONE" timepoint="t1">前往</EVENT>

上海。

<TRLINK timeID="t0" relatedtoEvent="e1" relType="SIMULTANEOUS"/>

例(4) 他从昨天开始定居上海。

他从

<CTIMEX tid="t1" type="DATE" value="2020-05-13-TOD" anchorTimeID="t0">昨天</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

开始

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="延续" timepoint="t1">定居</EVENT>

上海。

<TRLINK timeID="t0" relatedtoEvent="e1" relType="DURINGBEFORE"/>

例(5) 他从明天开始定居上海。

他从

<CTIMEX tid="t1" type="DATE" value="2020-05-15-TOD" anchorTimeID="t0">明天</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

开始

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="将来" polarity="肯定" aspect="延续" timepoint="t1">定居</EVENT>

上海。

<TRLINK timeID="t0" relatedtoEvent="e1" relType="DURINGAFTER"/>

例(6) 他今天上午前往上海。

他从

<CTIMEX tid="t1" type="DATE" value="2020-05-14-AM" anchorTimeID="t0">今天上午</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="NOTSURE" polarity="肯定" aspect="NONE" timepoint="t1">前往</EVENT>

上海。

<TRLINK timeID="t1" relatedtoTime="t0" relType="INCLUDE /ISINCLUDE"/>

例(7) 他昨天上午十点前往上海。

他

<CTIMEX tid="t1" type="DATE" value="2020-05-13-10:00am" anchorTimeID="t0">昨天上午十点</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" timepoint="t1">前往</EVENT>

上海。

<TRLINK timeID="t1" relatedtoEvent="e1" relType="BEGIN/BEGINBY"/>

例(8) 他昨天上午十点离开上海。

他

<CTIMEX tid="t1" type="DATE" value="2020-05-13-10:00am" anchorTimeID="t0">昨天上午十点</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" timepoint="t1">离开</EVENT>

上海。

<TRLINK timeID="t1" relatedtoEvent="e1" relType="END/ENDEDDBY"/>

例(9) 前天他告诉我，你将在明天上午来上海。

<CTIMEX tid="t1" type="NTIME" value="2020-05-12-TOD" anchorTimeID="t0">前天</CTIMEX>

<CTIMEX tid="t0" type="DATE" value="2020-05-14-TOD">说话时间(当前)</CTIMEX>

他

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="终止" timepoint="t1">告诉</EVENT>

我，你

<SIGNAL sid="s1">将</SIGNAL>

在

<CTIMEX tid="t2" type="DATE" value="2020-05-15-AM" anchorTimeID="t0">明天上午</CTIMEX>

<EVENT eid="e2" class="CONT\_OCCURRENCE" pos="动词" tense="将来" polarity="肯定" aspect="NONE" timepoint="t2">来</EVENT>

上海。

```

<TRLINK timeID="t1" relatedtoEvent="e1" relType="INCLUDE /ISINCLUDE "/>
<TRLINK timeID="t2" relatedtoEvent="e2" relType="INCLUDE /ISINCLUDE "/>
<TRLINK eventID="e1" relatedtoEvent="e2" relType="BEFORE"/>
<TRLINK eventID="t1" relatedtoTime="t0" relType="BEFORE "/>
<TRLINK eventID="t0" relatedtoTime="t2" relType="AFTER"/>
<TRLINK eventID="t1" relatedtoTime="t2" relType="BEFORE"/>

```

例(1)和例(4)的区别在于例(1)中的事件发生在说话时间之前,且“前往上海”事件已经完成,状态不会持续。例(4)中的“定居”事件发生在说话时间之前且一直持续到当前时间,可能还会继续保持这种状态。例(2)和例(5)的区别在于例(2)中的事件发生在说话时间之后,状态不会持续。而例(5)中的事件“定居”发生在说话时间之后且该事件状态之后会一直持续。例(6)他今天上午前往上海。“前往”发生在“今天上午”这一时间点,包含在“今天”这一时间内。所以事件“前往”在参考时间点“今天”是包含与被包含的关系。例(7)他昨天上午十点飞往上海。“飞往”发生在“昨天上午十点”,事件“飞往”在“昨天上午十点”这一时间点开始,所以事件“飞往”与“昨天上午十点”是开始与被开始的关系。例(8)他昨天上午十点离开上海。“离开”发生在“昨天上午十点”,事件“离开”在“昨天上午十点”这一时间点结束,所以事件“离开”与“昨天上午十点”是结束与被结束的关系。例(1-5)和例(7-8)描述的是时间与事件之间的时序关系,例(6)的时序关系标签描述的是时间与时间之间的关系,例(9)描述了时间与事件、时间与时间、事件与事件三种时序关系。

例(10) 她后悔去年嫁给了小李。

她

```

<EVENT eid="e1" class="CONGNITIVERB" pos="动词" tense="现在" polarity="肯定" aspect="NONE"
timepoint="t0">后悔</EVENT>

```

```

<CTIMEX tid="t1" type="DATE" value="2020-XX-XX-TOD" anchorTimeID="t0">去年</CTIMEX>

```

```

<CTIMEX tid="t0" type="DATE" value="2021-XX-XX-TOD">说话时间(当前)</CTIMEX>

```

```

<EVENT eid="e2" class="CONT_OCCURRENCE" pos="动词" tense="过去" polarity="肯定" aspect="
终止" timepoint="t1">嫁给</EVENT>

```

```

<SIGNAL sid="s1">了</SIGNAL>

```

小李。

```

<TRLINK timeID="t1" relatedtoEvent="e2" relType="INCLUDE /ISINCLUDE "/>

```

```

<TRLINK eventID="e1" relatedtoEvent="e2" relType="AFTER"/>

```

```

<TRLINK eventID="e1" relatedtoEvent="e2" relType="FACTIVITY"/>

```

```

<TRLINK eventID="t1" relatedtoTime="t0" relType="BEFORE "/>

```

例(11) 她怀疑小李刚刚喝酒了。

她

```

<EVENT eid="e1" class="CONGNITIVERB" pos="动词" tense="现在" polarity="肯定"
aspect="NONE" timepoint="t0">怀疑</EVENT>

```

小李

```

<CTIMEX tid="t1" type="UNCERTAINTIME" anchorTimeID="t0">刚刚</CTIMEX>

```

```

<CTIMEX tid="t0" type="DATE" value="2021-XX-XX-TOD">说话时间(当前)</CTIMEX>

```

```

<EVENT eid="e2" class="CONT_OCCURRENCE" pos="动词" timepoint="t1">喝酒</EVENT>

```

```

<SIGNAL sid="s1">了</SIGNAL>

```

```

<TRLINK timeID="t1" relatedtoEvent="e2" relType="INCLUDE /ISINCLUDE "/>

```

```

<TRLINK eventID="e1" relatedtoEvent="e2" relType="AFTER"/>

```

<TRLINK eventID="e1" relatedtoEvent="e2" relType="COUNTERFACTIVITY"/>

<TRLINK eventID="t1" relatedtoTime="t0" relType="BEFORE "/>

例(12) 她本来想明天去上海呢。

她

<SIGNAL sid= “s1” >本来</SIGNAL>

<EVENT eid=“e1” class=“MODELVERB” pos=“动词” tense=“现在” Modality=“意愿”>想</EVENT>

<CTIMEX tid=“t1” type=“DATE” value=“2021-04-06-TOD” anchorTimeID=“t0”>明天</CTIMEX>

<CTIMEX tid=“t0” type=“DATE” value=“2021-04-05-TOD”>说话时间(当前)</CTIMEX>

<EVENT eid= “e2” class= “CONT\_OCCURRENCE” pos=“动词” tense=“将来” aspect=“NONE”  
timepoint=“t1” >去</EVENT>

上海呢。

<TRLINK eventID=“t1” relatedtoTime=“t0” relType=“AFTER ”/>

<TRLINK eventID=“e1” relatedtoEvent=“e2” relType=“BEFORE”/>

<TRLINK eventID=“e1” signalID=“s1” relatedtoEvent=“e2” relType=“COUNTEREXCEPTION”/>

例(13) 她重新开始运动了。

她

<SIGNAL sid= “s1” >重新</SIGNAL>

<EVENT eid=“e1” class=“CONT\_OCCURRENCE” pos=“动词” tense=“现在” timepoint=“t0”>开始</EVENT>

<CTIMEX tid=“t0” type=“DATE” value=“XXXX-XX-XX-TOD”>说话时间(当前)</CTIMEX>

<EVENT eid=“e2” class=“CONT\_OCCURRENCE” pos=“动词” tense=“现在” polarity=“肯定” aspect=“开始”  
timepoint=“t0”>运动</EVENT>

了。

<TRLINK timeID=“t0” relatedtoEvent=“e1” relType=“SIMULTANEOUS ”/>

<TRLINK eventID=“e1” signalID=“s1” relatedtoEvent=“e2” relType=“REINITIATION ”/>

例(10)、例(11)描述了事件之间的事实性与反事实性、叙实性与反叙实性之间的关系，例(12)描述了事件间预期与反预期的关系，例(13)描述了事件间重新开始的关系。其中例(13)中“运动”事件的体信息属于事件内部时间关系即“运动”事件处于“开始”的状态，时序关系标签描述的是事件1“开始”和事件2“运动”之间的时序关系即“重新开始(REINITIATION)”。

## 五 时序关系标注形式化表示研究对比

本文在前人研究的基础上，基于汉语时间关系分类和标签体系实现了汉语时间事件关系的形式化表达，提出适用于汉语的时间事件关系标记语言。以往的时间关系标签体系如 TimeML、TIMEX2 等标签体系并不能完全满足汉语的需要。例如，汉语“了”、“着”、“过”所包含的时间信息纷繁复杂，而英语中不存在这种现象，利用之前的标签体系并不能解决这种难题。再比如汉语动词分类体系跟英语动词分类体系不同，因此完全从英语体系转换或翻译成汉语体系也存在问题，需要结合汉语的特点对动词进行分类。同时，英语 TimeML1.2.1 将某些名词也标注为事件，这在汉语中也是行不通的。本文构建的时间关系标签体系和时序关系分类跟 TimeML1.2.1 有所不同，这里主要总结了以下四个方面的不同。

### 1. 事件动词的分类不同

TimeML1.2 将动词分为 7 类，包括报道类(REPORTING)、身体感知类(PERCEPTION)、体态谓词(ASPECTUAL)、行动类事件(I\_ACTION)、不确定可能的事件(I\_STATE)、静态事件(STATE)、发生类事件(OCCURRENCE)。而本文在前人研究的基础上将描述事件的动词主要分为静态动词、动

态动词两大类。静态动词包括属性关系动词、心理认知动词、情态助动词，动态动词包括瞬间动词、持续动词、状态动词。由动词构成的事件分为动态事件和静态事件，不同动词事件的属性和属性值会有所不同，其中情态助动词构成的事件属性有可能性(Probability)、意愿(Willingness)、能力(Ability)、必要性(necessity)等。

## 2. 时序关系的分类体系不同

TimeML1.2 将时序关系类型分为 15 种，本文在此基础上进行了扩充，将汉语时序关系分为 2 大类、34 小类。其中外部时间关系包括事件与事件(12 类)、时间与时间(5 类)、时间与事件(12 类)之间的关系和不确定时间；内部时间关系包括“开始”“延续”“终止”“无内部时间关系”四类，共计 34 小类，这些小类中有的时序关系与其他关系共享同一种标签，例如“事件 A 发生时间 B 之前”“时间 A 发生时间 B 之前”和“事件 A 发生事件 B 之前”共享同一种标签“之前(BEFORE)”，最终构建了 23 种描述时序关系类型的标签。二者时序关系类型的区别在于本文的时序关系继承了 TimeML1.2.1 的 12 种关系类型，其中“在事件 1 在事件 2 发生之前立刻发生”“在事件 1 在事件 2 发生之后立刻发生”本文将其一并归入了“之前”和“之后”标签，并新增了 11 类标签，分别为：“事实性与反事实性(FACTICITY/COUNTER\_FACTICITY)”“叙实性与反叙实性(FACTIVITY/COUNTER\_FACTIVITY)”“预期与反预期(EXPECTATION/COUNTER\_EXPECTATION)”“开始(INITIATION)”“终止(TERMINATION)”“继续(CONTINUATION)”“重新开始(REINITIATION)”。

## 3. 时间关系标签体系不同

TimeML1.2.1 将关系类标签分为 TLINK、SLINK 和 ALINK 三种，其中 TLINK 包括前文提到的 14 中关系类型，SLINK 标注的是情态信息、事实性和反事实性、证据与非证据句和条件句，ALINK 标注的是事件与事件之间开始、结束等信息。本文结合汉语特点继承了 SLINK、ALINK 标签的内容，因为这两个标签大部分标识的都是事件与事件之间的关系，所以统一将其归为时序关系标签 TRLINK；此外，本文时间关系标签还新增了事件内部的关系类型即事件的体信息，识别时间关系的触发词也做了相应补充。

## 4. 标注形式更简单清晰

本文将汉语的时间关系体系重新进行了梳理，归纳了 7 种标签，将汉语事件信息及时序关系信息(2 大类、34 小类关系)囊括其中，更简单清晰地将汉语的时序关系进行了形式化的标识。例如，将动词所描述事件的所有属性与属性值信息都放入了事件 EVENT 标签内，将时间、事件之间的关系都融入到了时序关系 TRLINK 标签内。

下面例句对比了中文时间关系标注体系 CTimeML 和 TimeML1.2.1 标注同一条语料的不同。

例(1-a) 如果约翰今天离开，那么他将不会看到这场表演。(TimeML1.2.1)

<SIGNAL sid="s1">如果</SIGNAL>

约翰

<TIMEX3 tid="t1" type="DATE" value="XXXX-XX-XX" temporalFunction="true">今天</TIMEX3>

<EVENT eid="e1" class="OCCURRENCE">离开</EVENT>

<MAKEINSTANCE eiid="ei1" eventID="e1" pos="VERB" tense="PRESENT" aspect="NONE"

polarity="POS"/>

，那么他将不会

<EVENT eid="e2" class="OCCURRENCE">看到</EVENT>

<MAKEINSTANCE eiid="ei2" eventID="e2" pos="VERB" tense="FUTURE" aspect="NONE"

polarity="NEG" modality="将"/>

这场表演。

<SLINK eventInstanceID="ei1" subordinatedEventInstance="ei2" signalID="s1" relType="CONDITIONAL"/>

<TLINK eventInstanceID="ei1" relatedtoEventInstance="ei2" relType="BEFORE"/>

例(1-b) 如果约翰今天离开, 那么他将不会看到这场表演。CTimeML

<SIGNAL sid="s1">如果</SIGNAL>

约翰

<CTIMEX tid="t1" type="DATE" value="XXXX-XX-XX" anchorTimeID="t0">今天</TIMEX3>

<CTIMEX tid="t0" type="DATE" value="XXXX-XX-XX-TOD">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="现在" aspect="NONE" polarity="">

离开</EVENT>,

<SIGNAL sid="s2"> 那么</SIGNAL>

他

<SIGNAL sid="s3"> 将</SIGNAL>

不会

<EVENT eid="e2" class="CONT\_OCCURRENCE" pos="动词" signalID="s3" tense="将来" aspect="NONE" polarity="否定" Modality="可能性">看到 </EVENT>

这场表演。

<SLINK eventID="e1" signalID="s1" relatedtoEventID="e2" signalID="s2" relType="条件"/>

<TLINK eventID="e1" relatedtoEventID="e2" relType="BEFORE"/>

<TLINK timeID="t1" relatedtoTimeID="t0" relType="INCLUDE"/>

例(2-a) John caused the fire. John 引发了火灾。(TimeML 1.2.1 版)

John

<EVENT eid="e1" class="OCCURRENCE" > caused </EVENT>

the

<EVENT eid="e2" class="OCCURRENCE" > fire </EVENT>

<TLINK eventInstanceID="e1" relatedtoEvent="e2" relType="BEFORE"/>

例(2-b) (CTimeML 本文提出的时间关系标注体系)

John

<EVENT eid="e1" class="CONT\_OCCURRENCE" anchorTimeID="t0" aspect="终止" > 引发 </EVENT>

<CTIMEX tid="t0" type="DATE" value="XXXX-XX-XX-TOD">说话时间(当前)</CTIMEX>

了

<EVENT eid="e2" class="CONT\_OCCURRENCE" > 火灾 </EVENT>

<TLINK eventID="e1" relatedtoEvent="e2" relType="BEFORE"/>

<TLINK eventID="e1" relatedtoTime="t0" signalID="s1" relType="BEFORE"/>

<SLINK eventID="e1" relatedtoEvent="e2" relType="原因"/>

例(3-a) John left 2 days before yesterday.

John

<EVENT eid="e1" class="OCCURRENCE" > left </EVENT>

<MAKEINSTANCE eiid="ei1" eventID="e1" tense="PAST" aspect="PERFECTIVE"/>

<TIMEX3 tid="t1" type="DURATION" value="P2D" beginPoint="t2" endPoint="t3"> 2 days </TIMEX3>

<SIGNAL sid="s1"> before </SIGNAL>

<TIMEX3 tid="t2" type="DATE" value="2002-07-10" temporalFunction="true" anchorTimeID="t0">  
yesterday </TIMEX3>

<TIMEX3 tid="t3" type="DATE" value="2002-07-08" temporalFunction="true" anchorTimeID="t1"/>

<TLINK timeID="t1" relatedtoTime="t2" signalID="s1" relType="BEFORE"/>

<TLINK eventInstanceID="ei1" relatedtoTime="t3" relType="ISINCLUDED"/>

例(3-b) John 在昨天之前离开两天了。

John

在

<CTIMEX tid="t1" type="DATE" value="2002-07-10" anchorTimeID="t0">昨天</TIMEX3>

<SIGNAL sid="s1"> 之前 </SIGNAL>

<CTIMEX tid="t0" type="DATE" value="2002-07-11">说话时间(当前)</CTIMEX>

<EVENT eid="e1" class="CONT\_OCCURRENCE" tense="过去" anchorTimeID="t2" aspect="终止"> 离  
开</EVENT>

<CTIMEX tid="t2" type="DURATION" quant="Pn2Days" beginPoint="t1" endPoint="t3">两天</TIMEX3>  
了。

<CTIMEX tid="t3" type="DATE" value="2002-07-08" anchorTimeID="t1"/>

<TLINK eventID="e1" relatedtoEvent="t3" relType="ISINCLUDE"/>

<TLINK eventID="e1" relatedtoTime="t2" signalID="s1" relType="BEFORE"/>

<TLINK timeID="t1" relatedtoTime="t2" signalID="s1" relType="AFTER"/>

<TLINK timeID="t2" relatedtoTime="t3" signalID="s1" relType="BEFORE"/>

<TLINK timeID="t3" relatedtoTime="t1" signalID="s1" relType="BEFORE"/>

例(1-a)和例(2-a)是根据 TimeML 1.2.1 标注体系标注的，例(1-b)和例(2-b)是根据本文提出的中文时间关系标签体系标注的。观察上述例句，我们可以发现本文提出的标注体系标注汉语时间、事件和复句之间的关系更全面、清晰。例如，例(1-b)不仅涵盖了例(1-a)的标注内容，还增加了事件内部的体信息标注；例(2-b)不仅增加了事件内部体信息，还标注了事件“引发”与时间关系，增加了复句关系类型“原因”。例(3-b)标注了更细粒度的时间与时间、时间与事件之间的关系。

陈振宇先归纳了事件的时间结构，然后用符号系统进行翻译实现汉语句子里时间信息的提取，这种方式获取的时间信息虽将详细描述了当前语境下事件的时间关系，但是不太直观，事件动词的体信息、情态信息、句间的关系等都无法标识。下面例子对比了本文与陈振宇标注体系的差异。

例(4-a) 上午看了三个小时书。(陈)

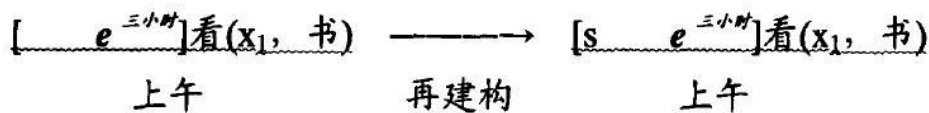


图 1 陈振宇(2006)现代汉语时间系统的认知模型与计算

例(4-b) 上午看了三个小时书。(本文 CTimeML)

<CTIMEX tid="t1" type="DATE" value="XXXX-XX-XX" anchorTimeID="t0">上午</TIMEX3>

<EVENT eid="e1" class="CONT\_OCCURRENCE" pos="动词" tense="过去" anchorTimeID="t1" aspect="终止">看</EVENT>

<SIGNAL sid="s1"> 了 </SIGNAL>

<CTIMEX tid="t0" type="DATE" value="XXXX-XX-XX-TOD">说话时间(当前)</CTIMEX>

<CTIMEX tid="t2" type="DATE" value="XXXX-XX-XX-TOD">三个小时 </CTIMEX>



书。

<TLINK eventID="e1" relatedtoTime="t2" relType="DURING"/>

<TLINK eventID="e1" relatedtoTime="t1" signalID="s1" relType="ISINCLUDE"/>

<TLINK timeID="t1" relatedtoTime="t2" relType="INCLUDE"/>

## 六 总结与展望

本文调研了语言学领域和自然语言处理领域对时间关系的研究进展。语言学领域主要总结了相关学者对汉语时体系统的观点，自然语言处理领域主要概括了相关学者对时间关系分类的认识和时间关系标注的研究进展以及时间信息识别、解析等任务的研究方法。首先，对汉语事件动词进行了分类，主要分为静态动词和动态动词两大类。其中静态动词包括属性关系动词、心理认知动词、助动词，动态动词包括瞬间动词、持续动词、状态动词。其次，在前人研究基础上根据汉语表达时间关系的特点总结了中文时间事件关系分类，基于中文时间事件关系特点构建了适用于汉语的时间关系标注 (CTimeML) 体系，主要有 2 大类时间关系，34 小类。外部时间关系主要包括 12 类事件与事件关系、5 类时间与时间关系、12 类时间与事件关系，内部时间关系主要包括“开始”、“结束”、“延续”、“重新开始”、“无内部时间关系” 5 类；并且提供一种关于时间信息的元数据标注语言，包括事件和时间表达式及事件与事件、事件与时间之间的关系交互，“事件”和关系标签下都可添加相关的属性、类型信息。最后，基于构建好的中文时间事件分类体系，在部分语料上尝试进行了标注，本文提出的汉语时间事件关系标记语言旨在回答某件事情“什么时候发生”、“持续了多长时间”或“发生的频率如何”、“时间和事件、事件和事件、时间与时间之间的先后顺序”等问题。从 CCL 语料库选取了部分代表性的语料进行了试标注，通过试标注环节，进一步验证了已构建好的标签体系可覆盖大部分文本上可捕捉到或可简单推理得到的时间信息。

未来，我们将基于本文提出的中文时间关系标注体系展开大规模语料标注，最终构建一个中文时间关系语料库。与此同时，考虑开发相关的标注工具和算法，最终为自然语言处理领域中的时间事件抽取提供助力。

## [ 参考文献 ]

- [1] REICHENBACH H. The tenses of verbs[J]. MEINSTER, J.; SCHERNUS, W. Time from concept to narrative construct: a reader. Berlin, 1947: 1-12.
- [2] BRUCE B C. A model for temporal references and its application in a question answering program[J]. Artificial intelligence, 1972, 3: 1-25.
- [3] COMIER B. Wolfgang Girke & Helmut Jachnow. Sowjetische Soziolinguistik: Probleme und genese. Kronberg/Taunus: Scriptor Verlag, 1974. Pp. 218.(Scriptor Taschen-Bucher, S 17.) 1[J]. Language in Society, 1976, 5(2): 242-243.
- [4] 郭锐. 汉语谓词性成分的时间参照及其句法后果[M], 世界汉语教学, 2015.
- [5] MANI I, VERHAGEN M, WELLNER B, 等. Machine learning of temporal relations[C]//Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics. 2006: 753-760.
- [6] VERHAGEN M, MANI I, SAURI R, 等. Automating temporal annotation with TARSQI[C]//Proceedings of the ACL interactive poster and demonstration sessions. 2005: 81-84.
- [7] PUSTEJOVSKY J, CASTANO J M, INGRIA R, 等. TimeML: Robust specification of event and temporal expressions in text.[J]. New directions in question answering, 2003, 3: 28-34.
- [8] LASCARIDES A, ASHER N. Temporal interpretation, discourse relations and commonsense entailment[J]. Linguistics and philosophy, 1993, 16(5): 437-493.
- [9] OBERLANDER J, LASCARIDES A. Preventing false temporal implicatures: Interactive defaults for text

- generation[C]//COLING 1992 Volume 2: The 14th International Conference on Computational Linguistics. 1992.
- [10] HITZEMAN J. Text Type and the Position of a Temporal Adverbial within the Sentence[J]. University of Edinburgh, 1995.
- [11] DORR B, GASTERLAND T. Constraints on the generation of tense, aspect, and connecting words from temporal expressions[R]. MARYLAND UNIV COLLEGE PARK MD IACS LANGUAGE AND MEDIA PROCESSING LAB, 2002.
- [12] LI W, WONG K F, CAO G, 等. Applying machine learning to Chinese temporal relation resolution[C]//Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04). 2004: 582-588.
- [13] CHAMBERS N, JURAFSKY D. Jointly combining implicit constraints improves temporal ordering[C]//Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing. 2008: 698-706.
- [14] YOSHIKAWA K, RIEDEL S, ASAHARA M, 等. Jointly identifying temporal relations with markov logic[C]//Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP. 2009: 405-413.
- [15] 孙辉. 事件时序关系识别的研究与实现[D]. 哈尔滨: 哈尔滨工业大学计算机科学与技术学院, 2010.
- [16] 王力. 中国语法理论[M]. Zhonghua shu ju, 2015.
- [17] 高名凯. 汉语语法论[M]. 汉语语法论, 1948.
- [18] 邓守信. 汉语动词的时间结构[J]. 语言教学与研究, 1985(4): 12.
- [19] 陈平. 论现代汉语时间系统的三元结构[J]. 中国语文, 1988(6): 401-422.
- [20] 龚千炎. 汉语的时相时制时态[M]. 汉语的时相时制时态, 1995.
- [21] 李铁根. “了”, “着”, “过” 与汉语时制的表达[J]. 语言研究, 2002(3): 1-13.
- [22] 戴耀晶. 《现代汉语时体系统研究》出版[J]. 汉语学习, 1997(2): 55-55.
- [23] 石毓智. 论现代汉语的“体”范畴[J]. 中国社会科学, 1992(6): 19.
- [24] 郭锐. 过程和非过程——汉语谓词性成分的两种外在时间类型[J]. 中国语文, 1997(3): 14.
- [25] 陈前瑞. 汉语体貌系统研究[J]. 华中师范大学博士论文, 2003, 1.
- [26] 孙英杰. 被动式与动词的及物性[J]. 北方论丛, 2006(2): 70-74.
- [27] 孙朝奋. 汉语动词系统的界及其呼应: “了 \_ 1” 完成体[J]. 世界汉语教学, 2022.
- [28] CHENG Y, ASAHARA M, MATSUMOTO Y. Naist. japan: Temporal relation identification using dependency parsed tree[C]//Proceedings of the fourth international workshop on semantic evaluations (SemEval-2007). 2007: 245-248.
- [29] 贺瑞芳, 秦兵, 潘越群, 等. 基于启发式错误驱动学习的中文时间表达式识别[J]. 高技术通讯, 2008, 18(12): 1258-1262.
- [30] 沈思, 苏新宁, 谢靖, 等. 基于清华汉语树库的时间表达式抽取模型构建研究[J]. 图书情报工作, 2012, 56(18): 127.
- [31] 张春菊, 张雪英, 李明, 等. 中文文本中时间信息解析方法[J]. 地理与地理信息科学, 2014, 30(6): 1-6.
- [32] 陈振宇. 时间系统的认知模型与运算[M]. 时间系统的认知模型与运算, 2007.
- [33] 袁毓林. 叙实性和事实性: 语言推理的两种导航机制[J]. 语文研究, 2020(1): 9.
- [34] 吕叔湘. 中国文法要略[J]. 商务印书馆, 1982.
- [35] 朱德熙. 语法讲义[M]. 语法讲义, 1982.