

“事件”的概念厘定和多维表征

王兴隆

武汉大学语言与信息研究中心 武汉 430072

鲁东大学文学院/山东省语言资源开发与应用重点实验室 烟台 264025

E-mail: wangxinglong100@163.com

摘要: 本文集中论述了“事件”的概念厘定和多维表征。首先,对事件进行了多维观察,然后将语言信息处理领域的“事件”研究厘定在了信息事件上。其次,探讨了事件的表征模式,认为可以从结构化表征、逻辑表征、本体化表征三个方面对事件进行凸显和描写。本文重点论述了事件的结构化表征,分别从事件微观结构、事件中观结构和事件宏观结构三个方面进行了分析。

关键词: 事件, 信息事件, 多维表征, 微观结构, 中观结构, 宏观结构

The Determining and Multidimensional Characterization of EVENT

Wang Xinglong

Research Center for Language and Information, Wuhan University, Wuhan, 430072

College of Liberal Arts, Lu Dong University/Key Laboratory for Development and Application of language resources in Shandong Province, Yantai, 264025

E-mail: wangxinglong100@163.com

Abstract: This paper focuses on the determining and multidimensional characterization of EVENT. Firstly, it conducted a multi-dimensional researching of EVENT, and then makes sure that EVENT in the area of Language information processing is about information. Secondly, it discusses the Multidimensional Characterizations of EVENT that could be made a Description from the structure, logic, ontology, etc. This paper focus on aspect of structure, and it describes from Microstructure, meso structure, macrostructure.

Keywords: EVENT, information event, multidimensional characterization, Microstructure, meso structure, macro-structure.

1 “事件”的多维观察与厘定

虽然“事件”的概念已在各个领域和学科中使用,但“事件”定义在不同领域有所差别。从以下视角进行多维观察,“事件”至少包含四类概念。

1.1 “事件”概念的多维表现

1.1.1 现实事件

现实事件指现实世界中发生的事件,在人类所见所闻所做的范围内,通常指客观世界存在着的,可发生于特定时空,并且不依赖于语言和信息表征的动作、行为、状态或变化等。比如自然事件、社会事件或体育事件等。语言事件即为对现实事件的表征。(Rosen 2003: 323)也可以说,语言事件是对现实事件的词汇化。(Levin and Rappaport Hovav 2005: 19)。概括讲,现实事件是哲学事件、信息事件和语言事件赖以存在的基础和依托。

1.1.2 哲学事件

哲学事件以时间和空间为基本前提,其中,时间第一,空间第二。这牵扯到宇宙的存在方式,宇宙就是由众多事件组成的,即由时空组成。哲学上的事件探索以奎因(Quine)和蒙塔古(Montague)为代表,奎因(Quine)认为,事件由占据一定时空域的整体所构成。

基金项目: 国家社会科学重大招标课题(No. 11&ZD189); 国家自然科学基金课题(No. 61173095)

作者简介: 王兴隆(1982-), 男, 博士生, 讲师, 研究领域为语言信息处理、计算机语篇理解。

此观点把事件看成与物体一样的四维空间实体。在物理学和数学中，一个n个数的序列可以被理解为一个n维空间中的位置。当n=4时，所有这样的位置的集合就叫做四维空间。

$$X=(p, q, r, s)$$

首先，一维，就是一条线。时间也可以看做时间线。把一个长宽高的三维坐标世界无限缩小到一个点，然后让这个三维世界点沿着时间轴方向运动，可称为四维。这个额外的维数可以直接理解为空间的第四维，即我们所理解的时间概念。

这个正在随着事件发生和流逝的概念，就构成了哲学上的事件。无数个三维的东西在做这个运动，就构成了无数个哲学事件。即时间这一维度激活了实体，使实体事件化。Montague认为，事件是哲学上的概念，具有时间上的时刻特征（properties of moment of time），并将事件分为通指事件（generic events）和具体事件（particular events）。物质世界是由事物和事件构成的，“事件”是自然界一去不返的具体事实，“事物”则是事件永恒不变的特质。Time ML是一种对文本中的事件和时间信息进行表示的标记语言，主要强调事件与时间的关系，认为事件是“代表一个事件状态发生或出现的词。事件既可以发生在时间点上，也可以持续一段时间”。

以上学者都强调了事件的时间属性。哲学事件最强调事件的内部时间结构，如事件的起始、持续、终结等，他们认为事件是随时间变化而变化的具体事实。

哲学事件概念中，事件具有时间特质的中心思想，可被信息处理领域的信息事件概念所借鉴。

1.1.3 语言事件

从语言学角度看，语言事件和信息事件存在差异：第一，表达形式上，语言事件更通常为语言文本中的一个语句，不像信息事件这般带有离散性或跳跃性；第二，内部结构上，强调谓词至上，坚持谓词中心论和强大向心力，语言事件的内部结构一般由动词、状态形容词或事件名词串联，各个元素以论元形式充当语义角色，它们相互协作，完成语言表达。并且，这些论元在形式上一般不会溢出该谓词所在的语句，语法位置不具明显离散性或跳跃性。语言事件强调形式上的动词论元结构，信息事件并不从语言学上的动词出发，很多事件名词，状态形容词等，也可以传递事件信息。所谓事件名词，是语义上兼有事物性与动作性的一批特殊名词，如“暴雨、战争、早饭”等。

事件结构理论即起源于语言学上的动词体结构研究。语言事件相关语言学理论和方法，如词汇概念结构(Jackendoff 1972, 1983, 1990)、事件结构(Voorst 1988, Grimshaw 1990, Levin and Rappaport Hovav 1995, 2005)、语义结构(Pinker 1989)、词汇关系结构(Hale and Keyser 1993)、逻辑结构(Van Valin and LaPolla 1997, Van Valin 2005)、词汇句法(Travis 2000)等。按照动词的内部时间结构可将动词分为不同的事件类型，逻辑语义学将事件作为动词的一个论元引入到形式语义学的研究中。生成语义学采用谓词解构将动词分解为若干基本动词组成的事件结构，并且，将其表征到深层句法中。语言学家根据动词所表达的事件特征对动词进行分类。给动词分类能够刻画出动词和其他句法成分之间的逻辑蕴涵关系和共存限制关系。最早系统研究动词蕴涵事件的是Vend。他按事件把动词划分为状态(states)、活动(activities)、完成(accomplishments)和成就(achievements)四类。活动类如：look(看)、listen(听)、study(学)、seek(寻)；完成类如：see(睹)、hear(闻)、learn(知)、find(见)。

总之，语言事件强调事件的语言表达形式和谓词的支配地位。

1.1.4 信息事件

在信息科学和语言信息处理科学中，事件指可以提供一个信息片段或信息焦点的语言串，语法形式和语言表达上，可带有跳跃性和离散性。它们通常以句子为单位，也可以跨越

句子存在。语言信息处理这个术语包含两个概念，语言和信息，世界上信息有很多类，语言信息处理关注的是从语言的角度获取的信息，重点在信息上。语言中的信息多以句子的形式呈现，可很多跨越句子存在，甚至中间并不相连。

人们在阅读文章获取信息时，可以用自己的知识和经验从中概括和提炼，可计算机并不具备智能性，我们只能训练它，让它能够透过表面上看似没有头绪的语句，获取诸多信息点。从信息学的角度，若某个信息处理单元可以传递一个信息，可以称之为一个事件，且不管在语言中它的存在依托为词、短语或句子。一些认知科学家认为，人类的命题记忆是以“事件”为存储单位的，存储的是组成事件的概念及其之间的关系。

近年来，“事件”概念逐渐被计算语言学、信息检索、信息抽取和自动文摘等信息处理领域引用。在信息检索和信息提取领域，Allan和Yang等认为事件是“细化了的用于检索的主题”，Marsh和Grishman等人分别将事件信息与事件相关的特定的组织、人或者人为的实体关联起来，采用事件以提取预先指定的事件信息。在自动文摘领域，Filatova等定义了元事件，它由动词（或者动名词）和动词连接的行为的主要组成部分构成，这些行为的主要连接成分是指三类命名实体：参与者（人名和机构名）、地点和时间。在信息提取领域，MUC将事件信息与事件相关的特定的组织、人或人为的实体关联起来，采用事件模板提取预先指定的事件信息。另外，ACE研究微观粒度的事件，认为事件是涉及具体参与者的特定事情，一个事件通常能被描述为状态的变化。

1.2 各事件概念的关系

现实世界中的运动、行为、变化、状态、存在等都可以用“事件”来描述。人类描述历史和世界，就是描述一连串事件以及事件之间的联系。物理学、认知科学、语言学、以及人工智能等领域都有不同的定义，但它们的相同点是，都包含了动作这一基本要素。

通过以上分析可知，每个领域的事件概念，存在一些差异。第一，信息事件最具包容性，它体现了事件的动作性、时间性、空间性、变化性和存在等元素；第二，现实事件和语言事件基本吻合，因为语言事件是对现实事件的表征；第三，哲学事件强调事件的时间特征，凡可以表现时间点或时间段的语言表示形式，都可以激活哲学事件。如下例：

一只鸟	现实事件×	哲学事件×	语言事件×	信息事件×
一只鸟飞	现实事件√	哲学事件×	语言事件√	信息事件√
一只鸟飞了	现实事件√	哲学事件√	语言事件√	信息事件√
树上一只鸟	现实事件×	哲学事件×	语言事件×	信息事件√
树上一只鸟飞了	现实事件√	哲学事件√	语言事件√	信息事件√
昨天树上一只鸟	现实事件×	哲学事件√	语言事件×	信息事件√

1.3 事件范畴厘定

本文认为，语言信息处理领域，“事件”概念应定为信息事件。随着事件概念逐渐被语言信息处理学界重视，面向语言信息处理的事件概念具备以下特点：

第一，吸收了现实事件和语言事件中的动作、行为和变化等特质，在信息事件中，“动作”“行为”“变化”等也将扮演重要角色，它将成为事件触发词的主要组成部分。

第二，具备自身特点，彰显了个性，如“存在”“状态”等。在信息事件中，“存在”“状态”扮演重要角色，它们将成为事件触发词的重要组成部分。

第三，吸收了哲学事件中事件的时间特质，在信息事件中，“时间”扮演重要角色，它将成为事件元素，参与构造事件，完成信息表征。

第四，吸收了哲学事件中事件的空间特质，在信息事件中，“空间”扮演重要角色，它将成为事件元素，参与构造事件，完成信息表征。

2 事件的结构化表征

语言信息处理领域,目前已有多种事件表示模型,包括MUC的事件模板、ACE(Automatic Content Extraction)的事件结构体、多媒体事件模型以及基于本体的事件模型ABC、PROTON、Event-Model-F等。Chung和Timberlake认为,事件可以根据三个组成部分表征:谓词;事件框架,即谓词发生的时间段;事件界,即谓词发生的情况或者条件。以上理论都从不同方面刻画了如何对事件进行表征。不足表现在:第一,事件概念不清,很多受语言事件或现实事件影响,导致事件元素混乱;第二,实用性不够强大,有些过多进行了理论阐发;第三,不够全面,没有从微观到宏观对事件的结构化表征进行研究。

本文将从事件的微观结构、中观结构和宏观结构对“事件”进行较为系统的论述。

2.1 事件微观结构

事件微观结构指深入事件内部,探讨其内部结构。主要由事件触发词(Event Core, EC)和事件元素(Event Element, EE)组成。

微观结构特点表现在:第一,不关注与其他事件关系;第二,不关注触发词和事件元素的语言表征方式。如下:

事件= {触发词, 事件元素}

事件= {触发词, 事件主体, 事件客体, 事件时间, 事件空间, 事件环境}

E= {Event Core, Event Element}

E= {Event Core, Subject, Object, Time, Space, Circumstance}

E= {EC, Sub, Obj, Tim, Spa, Cir}

E₁

E₂

E₃

[EE_{tim}[昨天]], [EE_{sub}[他]][EC[饿]]了,就[EC[去]][EE_{spa}[菜市场]][EC[买]]了一些[EE_{obj}[蔬

E₄

菜和水果]], [EC[做]]了一顿[EE_{obj}[美餐]]。

以上信息板块,共分为四个事件,每个事件都有它们的微观结构,即触发词和事件元素。微观结构研究只研究本事件内部结构。

2.2 事件中观结构

事件中观结构涵盖着事件微观结构,在以下两方面有所扩大:第一,关注与其他事件关系,不过,焦点在和本事件有关系的成分,比如事件前态和事件后态,他们可以出现在该事件内部,也可以出现在事件外部,即另一个事件当中;第二,关注触发词和事件元素的语言表征方式。因为,这牵扯到触发词和事件元素的识别和判定。如下:

事件= { [触发词, 事件元素] [事件前态, 事件后态] 语言表征 }

事件= { 触发词, 事件主体, 事件客体, 事件时间, 事件空间, 事件环境, 事件前态, 事件后态, 语言表征 }

E= {Event Core, Subject, Object, Time, Space, Circumstance, Pre-State, Aft-State, Language Expression}

E= {EC, Sub, Obj, Tim, Spa, Cir, Pre-Sta, Aft-Sta, LX}

E₁

E₂

E₃

[EE_{tim}[昨天]], [EE_{sub}[他]][EC[饿]]了, 就[EC[去]][EE_{spa}[菜市场]][EC[买]]了一些[EE_{obj}[蔬

E₄

菜和水果]], [EC[做]]了一顿[EE_{obj}[美餐]]。

以上信息板块, 以E₃为中心, 事件前态为E₁中触发词[EC[饿]], 事件后态为E₄中触发词[EC[做]]。语言表征特点为, 四个事件在语言形式上共享事件主体EE_{sub}[他]和事件时间[EE_{tim}[昨天]]。

2.3 事件宏观结构

事件宏观结构指从语篇或信息板块角度去研究事件。第一, 它从宏观角度上, 以事件为出发点, 研究整个语篇或信息板块的结构。这个出发点事件通常为核心事件; 第二, 研究具有比较严格的步骤; 第三, 事件宏观结构表现出明显的特点。

2.3.1 事件宏观结构研究和操作步骤

第一, 核心事件 EC (Core Event) 判定。核心事件可以最大限度代表整个语篇的中心内容, 其他事件都围绕在核心事件周围, 为整个语篇服务。

第二, 辅助事件 ES (Subsidiary Event) 判定与划分。辅助事件一般不能代表整个语篇的中心内容, 只能最大限度代表辅助事件, 他们围绕在核心事件周围, 为整个语篇服务。

第三, 核心事件触发词 EC (Event Core) 判定和识别。它可以串联整个核心事件。作为该核心事件的主导和中枢, 它可以最大限度上代表该核心事件, 甚至可以由此代表整个语篇中心内容, 是整个核心事件或语篇内容的浓缩和代表。触发词一般为动词, 也可能为状态形容词或事件名词等。

第四, 辅助事件触发词 EC (Event Core) 判定和识别。它可以串联整个辅助事件, 作为该辅助事件的主导和中枢, 它可以最大限度上代表该辅助事件, 是整个辅助事件的浓缩和代表。

第五, 核心事件元素 CE (Core Element) 判定和识别。包括[事件主体[E_{sub}]], [事件客体[E_{obj}]], [事件时间[E_{tim}]], [事件空间[E_{spa}]], [事件环境[E_{cir}]], [事件前态[E_{pre-sta}]], [事件后态[E_{aft-fta}]]。

第六, 辅助事件元素 SE (Subsidiary Element) 判定和识别。包括[事件主体[E_{sub}]], [事件客体[E_{obj}]], [事件时间[E_{tim}]], [事件空间[E_{spa}]], [事件环境[E_{cir}]], [事件前态[E_{pre-sta}]], [事件后态[E_{aft-fta}]]

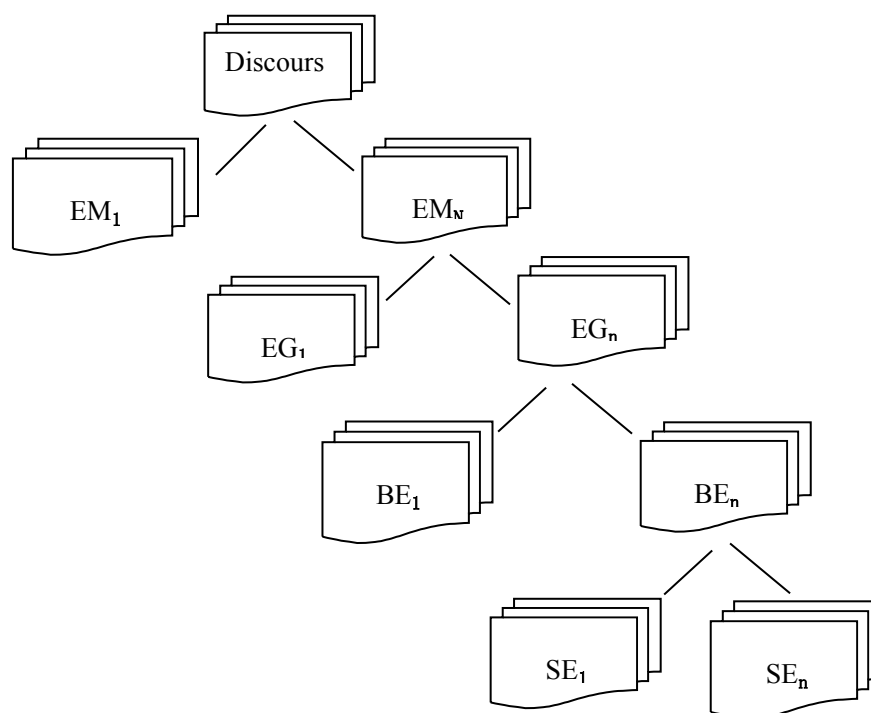
2.3.2 事件宏观结构的特点

2.3.2.1 结构层次性

语篇一般由多个事件组成，从事件角度来看，它们形成一个事件集群，在这个集群中，基本单位即为一个独立事件。这些事件之间并非杂乱无章，相反，它们会表现出一定层次结构。语篇通常由信息板块组成，信息板块由各事件组构成，各个事件组经常由一些本身还蕴含着小事件的大事件构成。如此，它们之间就形成了一个层级结构关系。如下：

语篇 (Discourse) \geq 信息板块 (Event Module \rightarrow EM; $EM_i \in EM_1 \rightarrow EM_n$, $1 \leq i \leq n$, n 为正整数) \geq 事件组 (Event Group \rightarrow EG; $EG_j \in EG_1 \rightarrow EG_n$, $1 \leq j \leq n$, n 为正整数) \geq 大事件 (Big Event \rightarrow BE) \geq 小事件 (Small Event \rightarrow SE)

它们之间关系如下图：

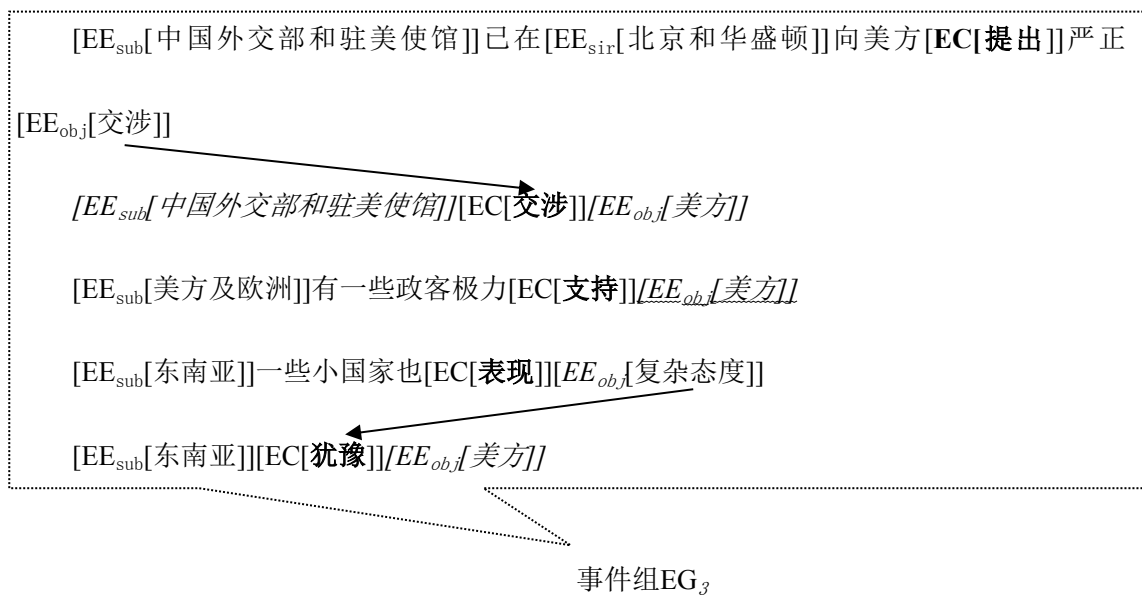
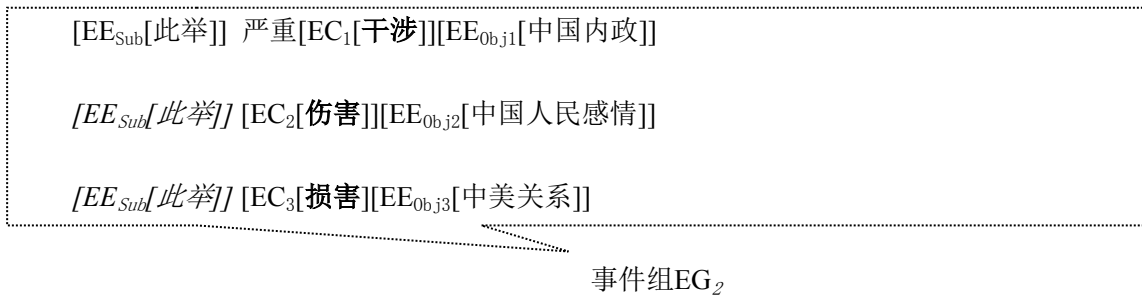
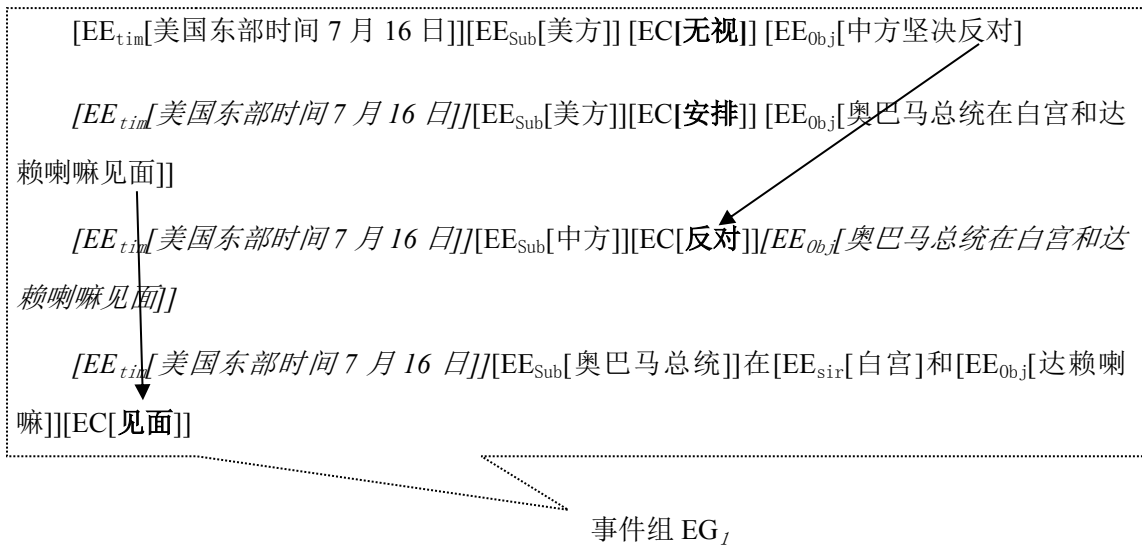


美国东部时间7月16日，美方无视中方坚决反对，安排奥巴马总统在白宫和达赖喇嘛见面。此举严重干涉中国内政，伤害中国人民感情，损害中美关系。中国外交部和驻美使馆已在北京和华盛顿向美方提出严正交涉。美方及欧洲有一些政客极力支持，东南亚一些小国家也表现复杂态度。

以上这段话，从语言信息处理角度来看，应该属于一个信息板块，它可能截取自某个语篇，表述了该语篇的部分信息。它存在一个核心事件和围绕在核心事件周围的辅助事件，以及各个事件的事件元素，它们都依靠这个核心事件凝聚成一个板块。从语言学上分析，以上信息板块 EM 一共分为四个句子。

- S₁[美国.....见面]
- S₂[此举.....关系]
- S₃[中国.....交涉]
- S₄[美方.....态度]

从语言信息处理角度，该信息板块共可划为十二个不同的事件。它们分为事件组 EG₁，事件组 EG₂，事件组 EG₃ 三个事件组。



第二，大事件 BE 层。共分为八大事件。大事件的叫法，着眼于它所在的语法层次和语法位置，与信息的重要程度与否没有必然联系。

第三，小事件 SE 层。共分为四个小事件。小事件的叫法，着眼于它所在的语法层次和语法位置，与信息的重要程度与否没有必然联系。

如下，大事件边界用大括弧进行划分，小事件边界用中括弧进行划分。

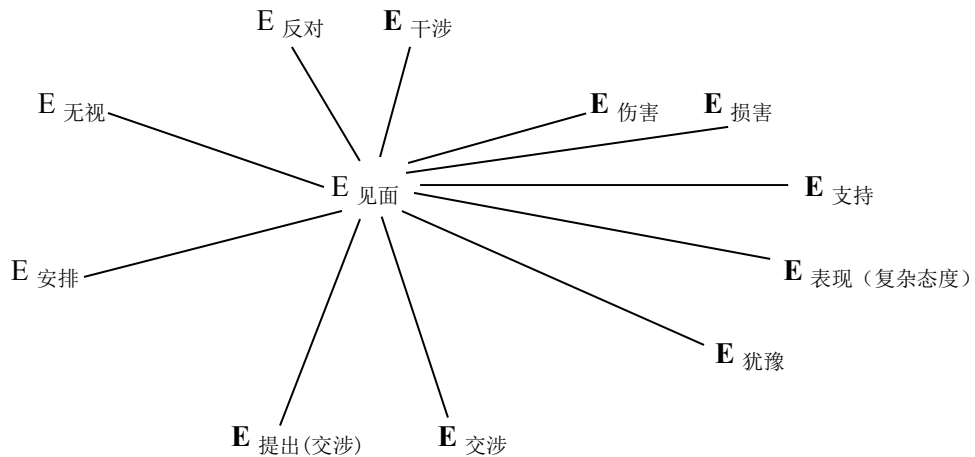
{[EE_{tim}[美国东部时间7月16日]][EE_{Sub}[美方]] [EC[无视]] [EE_{Obj}[中方坚决反对]]},

{[EC[安排]][EE_{Obj}[奥巴马总统在白宫和达赖喇嘛见面]]}。{[EE_{Sub}[此举]]严重[EC[干
涉]][EE_{Obj}[中国内政]]} { [EC[伤害]][EE_{Obj}[中国人民感情]]}, { [EC[损害]][EE_{Obj}[中美关系]]}。

{[EE_{Sub}[中国外交部和驻美使馆]]已在[EE_{Site}[北京和华盛顿]]向美方[EC[提出]]严正[EE_{Obj}[交
涉]]}。{ [EE_{Sub}[美方及欧洲]]有一些政客极力[EC[支持]]}, { [EE_{Sub}[东南亚]]一些小国家也
[EC[表现]][EE_{Obj}[复杂态度]]}。

2.3.2.2 结构向心性

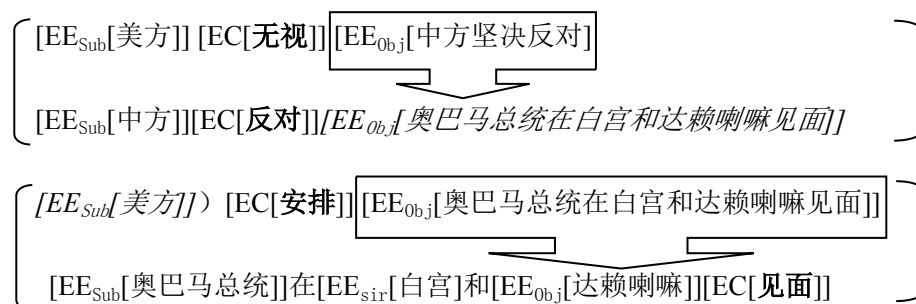
语篇一般由多个事件组成，这些事件里必然存一个核心事件。不管该语篇可切分为多少事件，也不管可分为多少个层次，所有事件都会围绕着这个核心事件，形成向心力。以上该事件板块的核心事件为 E_{见面}，结构表征为[EE_{tim}[美国东部时间7月16日]][EE_{Sub}[奥巴马总统]]在[EE_{Site}[白宫]和[EE_{Obj}[达赖喇嘛]][EC[见面]]。它和其他各个事件之间的都呈现出了服务和被服务的关系。如下图，至于具体构成什么关系，属于事件链研究范围，本文暂不论述。



2.3.2.3 结构嵌套性

核心事件不一定在语法层次的中心位置，它经常嵌套于另一个事件中，它可能处于多重语法层次的任一位置。我们的工作就是发现它并总结规律。语篇核心事件发现和判定非常复杂，在机器可以处理信息之前，我们必须从一定规模语篇中找出某些规律并让计算机学习。

以上信息板块，S₁[美国.....见面]，从句法语义学角度讲，应为板块的中心句，如此，按照语言事件概念，我们应该把它定位核心事件。不过，计算机处理语言信息，最终需反馈给用户正确信息，因此，信息元素被摆在第一位置，焦点信息事件才能被确定为核心事件，因为，我们关注信息事件，并非语言学上的机械语言事件。通过文本分析，可发现，该语篇一切事件都围绕着“奥巴马总统会面达赖喇嘛”这一事件进行信息表达。所以，该板块核心事件为应为[EE_{Sub}[奥巴马总统]]在[EE_{sir}[白宫]和[EE_{Obj}[达赖喇嘛]][EC[见面]]。它恰恰嵌套在[EE_{Sub}[美方]][EC[安排]] [EE_{Obj}[奥巴马总统在白宫和达赖喇嘛见面]]这一事件中。另外，该板块辅助事件[EE_{Sub}[中方]][EC[反对]][EE_{Obj}[奥巴马总统在白宫和达赖喇嘛见面]]，也嵌套在了辅助事件[EE_{Sub}[美方]] [EC[无视]] [EE_{Obj}[中方坚决反对]]中。如下：



2 事件的逻辑表征

当前，事件的逻辑表征理论和方法有很多，可不足在于，虽然设计比较理想，可操作性比较差。另外，个别理论和方法过于复杂。实际上，事件的逻辑表征可用当前为我们所比较熟知的逻辑三元组进行分解。

如“中国国家主席习近平于2013年3月22日抵达俄罗斯首都莫斯科进行国事访问”。为了能够让计算机理解并识别，依据该事件内部结构和事件元素，可将以上事件拆分成以下几个三元组。

```
<抵达, i s - T y p e - o f , M o v e m e n t / T r a n s p o r t > ,
<习近平, i s - T y p e - o f , P e r s o n > ,
<2013年3月22日, i s - T y p e - o f , T i m e > ,
<莫斯科, i s - T y p e - o f , P l a c e >
```

它们的RDF表示可按以下框架：

```
< E v e n t   r d f :   I D = "   N e w s E v e n t _ _ 某编号 " >
  < r d f s :   c o m m e n t >
    习近平抵达俄罗斯访问
  < / r d f s :   c o m m e n t >
```

```

<i n D o c u m e n t >
  <D o c u m e n t   r d f :   r e s o u r c e = “ # 某 编 号 ” / >
</ i n D o c u m e n t >
<h a s A c t i o n >
  <A c t i o n   r d f :   I D = “ < 中 国 国 家 主 席 习 近 平 ， 抵
  达 ， 俄 罗 斯 首 都 莫 斯 科 > ” / >
  <A c t i o n   r d f :   I D = “ < - ， 进 行 ， 国 事 访 问 > ” / >
</ h a s A c t i o n >
<h a s S u b j e c t >
  <A g e n t   r d f :   I D = “ 中 国 国 家 主 席 习 近 平 ” / >
</ h a s S u b j e c t >
<h a s O b j e c t >
  <A g e n t   r d f :   I D = “ 俄 罗 斯 首 都 莫 斯 科 ” / >
</ h a s O b j e c t >
<h a s T y p e >
<A C E _ _ E v e n t _ _ T y p e   r d f :   r e s o u r c e = “ # M o v e m e
n t / T r a n s p o r t ” / >
</ h a s T y p e >
<a t T i m e >
  <T i m e   r d f :   I D = “ 2 0 1 3 年 3 月 2 2 日 ” / >
</ a t T i m e >
<i n P l a c e >
  <P l a c e   r d f :   I D = “ 俄 罗 斯 ” / >
  <P l a c e   r d f :   I D = “ 莫 斯 科 ” / >
</ i n P l a c e >
</ E v e n t >

```

3 事件的本体化表征

在传统领域本体中，事件本体仅仅作为一个普通本体存在着。基础本体受到更多关注，事件本体构建和特点在一定程度上被忽视了。基础本体或普通领域本体呈现层级结构，只适合描述静态的二元关系，而事件可能是多元关系，事件链则呈现动态的二元关系。传统本体表征存在以下缺陷：

第一，传统基础本体和普通领域本体只描述静态概念之间的层次关系。缺乏动态性，事件是一个具有明显动态性的概念，很多关系需在语篇文本中进行挖掘和统计。比如，[E_{穿(袜)}]通常在[E_{穿(鞋)}]之后出现，[E_{落座}]通常在[E_{上车}]之后出现，或

者说，[E_{穿(鞋)}]之后一般跟随[E_{穿(袜)}]，[E_{上车}]之后一般跟随[E_{落座}]，可以表示为[R_{ss}[[E_{穿(鞋)}], [E_{穿(袜)}]]以及[R_{ss}[[E_{上车}], [E_{落座}]]。诸如此类，它们之间都可构成一种动态语义关系模型，我们称之为事件链表征模型，已在另文进行了专述。如何把这些动态语义关系加入到传统本体中，是事件本体表征的关键。

第二，传统基础本体和普通领域本体中，节点通常为词，短语比例极低，事件本体注重信息处理的实用性。因此，在节点的语言表现方式上可以更加多元化。

第三, 传统基础本体和普通领域本体在语言信息处理方面, 作用有限, 构建基础本体, 必须研究基础本体到领域本体的突变和渐变, 目前这方面研究不多。从信息处理角度, 以事件为基本单元会更加合理和智能。

4 结语

事件”的概念厘定和多维表征, 作用在于: 第一, 有助于在进行计算机语篇理解之前对事件和事件结构进行概念和结构厘定。将语言信息处理领域的“事件”研究划定在了信息事件上, 摆脱了以往研究中事件概念纠结和模糊的现象。第二, 分别从事件微观结构、事件中观结构和事件宏观结构论述了事件的结构化表征, 不仅在理论上对事件结构进行了多维界定, 更有助于在事件基础上构筑事件和事件链语义表征模型, 建立事件和事件链文本标注方法和策略, 为计算机学习事件语义, 进行语篇理解奠定基础。第三, 本文研究可能会有利于语篇事件信息挖掘和发现, 以及语料标注标准和方法的确立。

参 考 文 献

- [1] Roack, Brain & Charniak, Eugene, Noun-phrase co-occurrence statistics for semi-automatic semantic lexicon construction, In Coling' 98, 1998, pp1110-1116.
- [2] Rosch, E. Cognitive representation of semantic categories. Journal of Experimental Psychology, General, 1975, (104).
- [3] Lin H F, Liang J M. Event-based Ontology design for retrieving digital archives on human religious self-help consulting. Proceeding of the 2005 IEEE International Conference on e-Technology, e-Commerce and e-Service. Hong Kong, 2005.
- [4] Lee C S, Chen Y J, Jian Z W. Ontology-based fuzzy event extraction agent for Chinese e-news summarization. Expert Systems with Applications, 2003, 25(3).
- [5] Fogaras D. Where to start browsing the Web?. Proceedings of the IICS. Leipzig, 2003
- [6] 袁毓林. 论元角色的层级关系和语义特征. 世界汉语教学, 2002年第3期.
- [7] 谭红叶. 中文事件抽取关键技术研究. 哈尔滨工业大学博士论文, 2008.
- [8] 萧国政, 姬东鸿, 肖珊. Ontology的类型及汉语词网的Ontology结构. 长江学术. 2011(2).