

## 知识抽取-事件抽取



徐阿衡

人工智能与机器学习工程师

关注她

漱石先森等 217 人赞同了该文章

接上一篇知识抽取-实体及关系抽取，前置知识在这一篇不多做解释啦。

事件是促使事情状态和关系改变的条件[Dong et.al., 2010]。目前已存在的知识资源（如维基百科等）所描述实体及实体间的关系大多是静态的，而事件能描述粒度更大的、动态的、结构化的知识，是现有知识资源的重要补充。

与关系抽取相比，事件抽取同样需要从文本中抽取 predicate 和对应的 arguments，但不同的是，关系抽取的问题是 binary 的，且两个 arguments 通常都会出现在同一个句子中出现，而事件抽取的难点在于，有多个 arguments 和 modifiers，可能会分布在多个句子中，且有些 arguments 不是必须的，这使得 bootstrapping/distant learning/coreference 都变得非常困难。

整体而言，事件抽取的任务可以分两大类：

### • 事件识别和抽取

从描述事件信息的文本中识别并抽取出事件信息并以结构化的形式呈现出来，包括发生的时间、地点、参与角色以及与之相关的动作或者状态的改变。

### • 事件检测和追踪

事件检测与追踪旨在将文本新闻流按照其报道的事件进行组织，为传统媒体多种来源的新闻监控提供核心技术，以便让用户了解新闻及其发展。具体而言，事件检测与追踪包括两个子任务：

闻文档中发现以前没有识别的事件，后者则是从实时新闻流中实时发现新的事件。

本文的重点在于事件识别与抽取。首先看一下相关的核心概念：

- **事件描述 (Event Mention)**

描述事件的词组/句子/句群，包含一个 trigger 以及任意数量的 arguments

- **事件触发 (Event Trigger)**

事件描述中最能代表事件发生的词汇，决定事件类别的重要特征，一般是动词或者名词

- **事件元素 (Event Argument)**

事件的重要信息，或者说是实体描述 (entity mention)，主要由实体、属性值等表达完整语义的细粒度单位组成

- **元素角色 (Argument Role)**

事件元素在事件中扮演的角色，事件元素与事件的语义关系，可以理解为 slot

- **事件类型 (Event Type)**

## 事件识别和抽取

直观上来看，可以把事件抽取的任务理解成从文本中找到特定类别的事件，然后进行填表的过程。



严肃些看下事件识别和抽取的任务定义：

- 识别事件触发词及事件类型
- 抽取事件元素 (Event Argument) 同时判断其角色 (Argument Role)
- 抽出描述事件的词组或句子

当然还有一些其他的子任务包括事件属性标注、事件共指消解等。

事件抽取大多是分阶段进行，通常由 trigger classifier 开始，如果有 trigger，把 trigger 以及它的上下文作为特征进行分类判断事件类型，再进行下一步的 argument classifier，对句子中的每个 entity mention 进行分类，判断是否是 argument，如果是，判定它的角色。

## 基于模式匹配的方法

MUCs 最开始，事件抽取的系统都是基于人工编写的规则，基于语法树或者正则表达式，如 CIRCUS (Lehnert 1991), RAPIER (Califf & Mooney 1997), SRV (Freitag 1998), AutoSlog (Riloff 1993), LIEP (Huffman 1995), PALKKA (Kim & Moldovan 1995), CRYSTAL (Soderland et al. 1995), HASTEN (Krupka 1995) 等等，后来，慢慢的有了监督学习的模型，在 ACE 的阶段，大多数系统都是基于监督学习了，但由于标注一致性的问题，系统的效果普遍较差，ACE 事件抽取只举行了一次，在 2005 年。

下面先来看一下基于模板的抽取方法，基本都是通过 **句法 (syntactic)** 和 **语义约束 (semantic constraints)** 来进行识别。

## 基于人工标注语料

在早期，模板创建过程通常从一个大的标注集开始，模板的产生完全基于人工标注语料，学习效果高度依赖于人工标注质量。

### • AutoSlog (Riloff)

基本假设：

- a. 事件元素首次提及之处即可确定该元素与事件间的关系
- b. 事件元素周围的语句中包含了事件元素在事件中的角色描述

通过监督学习和人工审查来建立抽取规则。通过训练数据中已经填充好的槽 (filled slot)，AutoSlog 解析 slot 附近的句法结构，来自动形成抽取规则，由于这个过程产生的模板 too-general，所以需要人工来审核。本质上形成的是一个字典。

举个例子

Ricardo Castellar, the mayor, was kidnapped yesterday by the FMLN.

假设 Ricardo Castellar 被标注成了 victim，AutoSlog 根据句法分析判断出 Ricardo Castellar

## • PALKA

基本假设：特定领域中高频出现的语言表达方式是可数的

用语义框架和短语模式结构来表示特定领域中的抽取模式，通过融入 WordNet 的语义信息，PALKA 在特定领域可取得接近纯人工抽取的效果。

## 基于弱监督

人工标注耗时耗力，且存在一致性问题，而弱监督方法不需要对语料进行完全标注，只需人工对语料进行一定的预分类或者制定种子模板，由机器根据预分类语料或种子模板自动进行模式学习。

一些系统：

### • AutoSlog-TS

Riloff and Shoen, 1995

AutoSlog-TS 不需要进行文本的标注，只需要一个预先分类好的训练语料，类别是与该领域相关还是不相关。过程是先过一遍语料库，对每一个名词短语（根据句法分析识别）都产生对应的抽取规则，然后再整体过一遍语料库，产生每个规则的一些相关统计数据，基本的 idea 是与不相关文本相比，**在相关文本中更常出现的抽取规则更有可能是好的抽取规则**。假设训练语料中相关与不相关的文本比例是 1:1，对产生的每条抽取规则计算相关比率 relevance rate，相关文档中出现规则的实例数/整个语料库中出现规则的实例数，那么  $\text{relevance rate} < 50\%$  的抽取规则就被丢弃了，剩下的规则会按照  $\text{relevance\_rate} * \log(\text{frequency})$  的形式从高到低进行排序，然后由人工进行审核。

### • TIMES

Chai and Biermann, 1998

引入了领域无关的概念知识库 WordNet，提升模式学习的泛化能力，并通过人工或规则进行词义消歧，使最终的模式更加准确

### • NEXUS

Piskorski et.al., 2001; Tanev et.al., 2008

用聚类对语料进行预处理

### • GenPAM

Jiang, 2005

在由特例生成泛化模式的学习过程中，有效利用模式间的相似性实现词义消歧，最大限度地减少了人工的工作量和对系统的干预

## 小结

以用半监督的方法来建 trigger 字典。

## 基于统计 - 传统机器学习

建立在统计模型基础上，事件抽取方法可以分为 pipeline 和 joint model 两大类。

### Pipeline

将事件抽取任务转化为多阶段的分类问题（管道抽取），需要顺序执行下面的分类器：

1. **事件触发词分类器 (Trigger Classifier)**  
判断词汇是否是事件触发词，以及事件类别
2. **元素分类器 (Argument Classifier)**  
词组是否是事件元素
3. **元素角色分类器 (Role Classifier)**  
判定元素的角色类别
4. **属性分类器 (Attribute Classifier)**  
判定事件属性
5. **可报告性分类器 (Reportable-Event Classifier)**  
判定是否存在值得报告的事件实例

分类器可以用 MaxEnt, SVM。重点还是在于提取和集成有区分性的特征，包括 **句子级信息** 和 **篇章级信息**。

**句子级信息**：与候选词相关的词法特征、上下文特征、实体特征、句法特征、语言学特征等，如：



- Trigger单词和上下文单词的Token和POS标签
- 字典
  - Trigger列表、同义词字典
- 句法
  - Trigger在句法树中的深度
  - Trigger到句法树根节点的路径
  - 由Trigger的父节点展开的词组结构
  - Trigger的词组类型
- 实体
  - 句法上距离trigger最近的实体的类型
  - 句子中距离trigger物理距离最近的实体的类型

- Trigger的token
- 事件类型和子类型
- 实体
  - 实体类型和子类型
  - Head word of the entity mention
- 上下文
  - 候选元素的上下文单词
- 句法
  - 扩展trigger父节点的词组结构
  - 实体和trigger的相对位置(前或后)
  - 实体到触发词的最短路径
  - 句法树中实体到触发词的最短长度

知乎 @徐阿衡

### 篇章级特征:

跨文档利用全局信息。对于一个句子级的抽取结果不仅要考虑当前的置信度,还要考虑与待抽取文本相关的文本对它的影响,以及全局信息如事件与话题的关系,事件与事件的共现信息等,主要工作有:

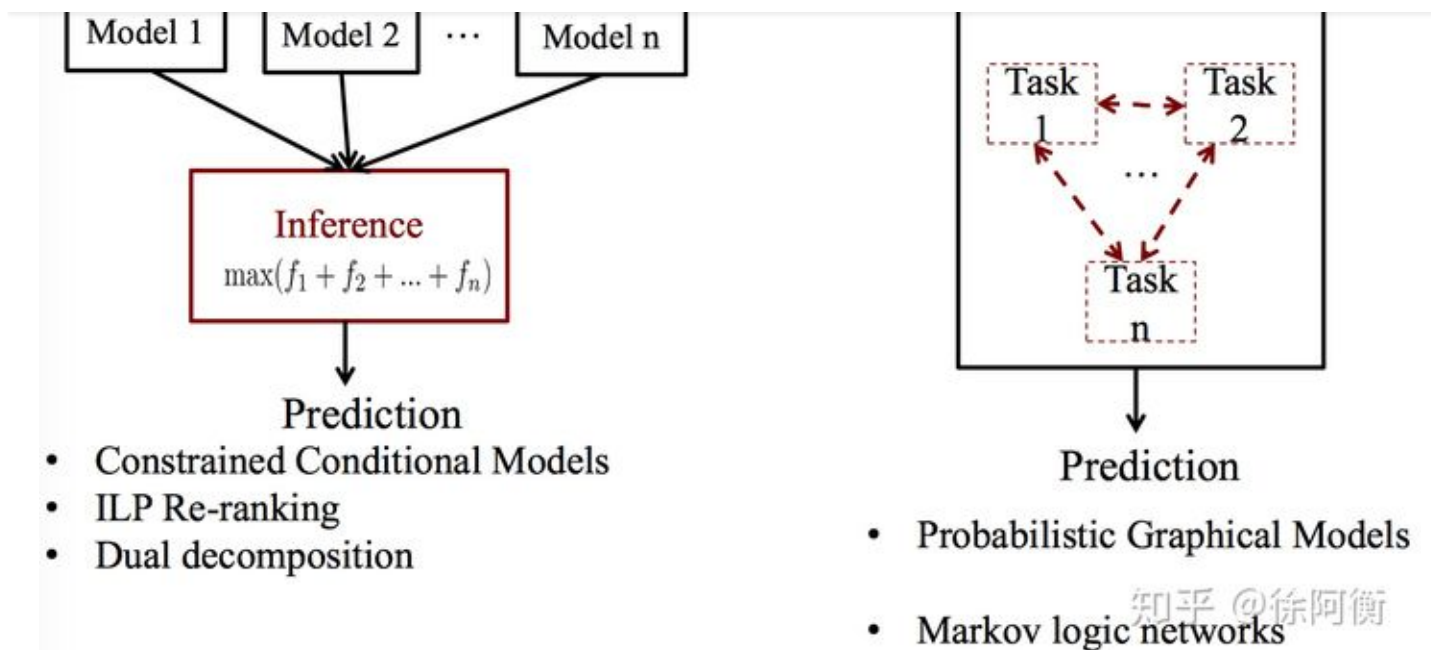
- Ji and Grishman, 2008
- Liao and Grishman, 2010
- Hong et.al., 2011
- Liu et.al., 2016a

早期大部分的研究都是基于 Pipeline 方法,然而它的问题也很明显:

- 误差传递,导致性能衰减
- 各环节预测任务独立,缺少互动,如忽略了事件触发词和事件元素之间的相互影响
- 无法处理全局的依赖关系

## Joint Model

又分为 Joint Inference 和 Joint Modeling 两种。



**Joint Inference** 使用集成学习的思路，将各模型通过整体优化目标整合起来，可以通过整数规划等方法进行优化。

**Joint Modeling (Structured)** 又可以称为基于结构的方法，将事件结构看作依存树，抽取任务相应转化为依存树结构预测问题，触发词识别和元素抽取可以同时完成，共享隐层特征，使用搜索进行求解，避免了误差传播导致的性能下降，另外，全局特征也可以从整体的结构中学习得到，从而使用全局的信息来提升局部的预测。相关工作有：

- Li et.al., 2013a

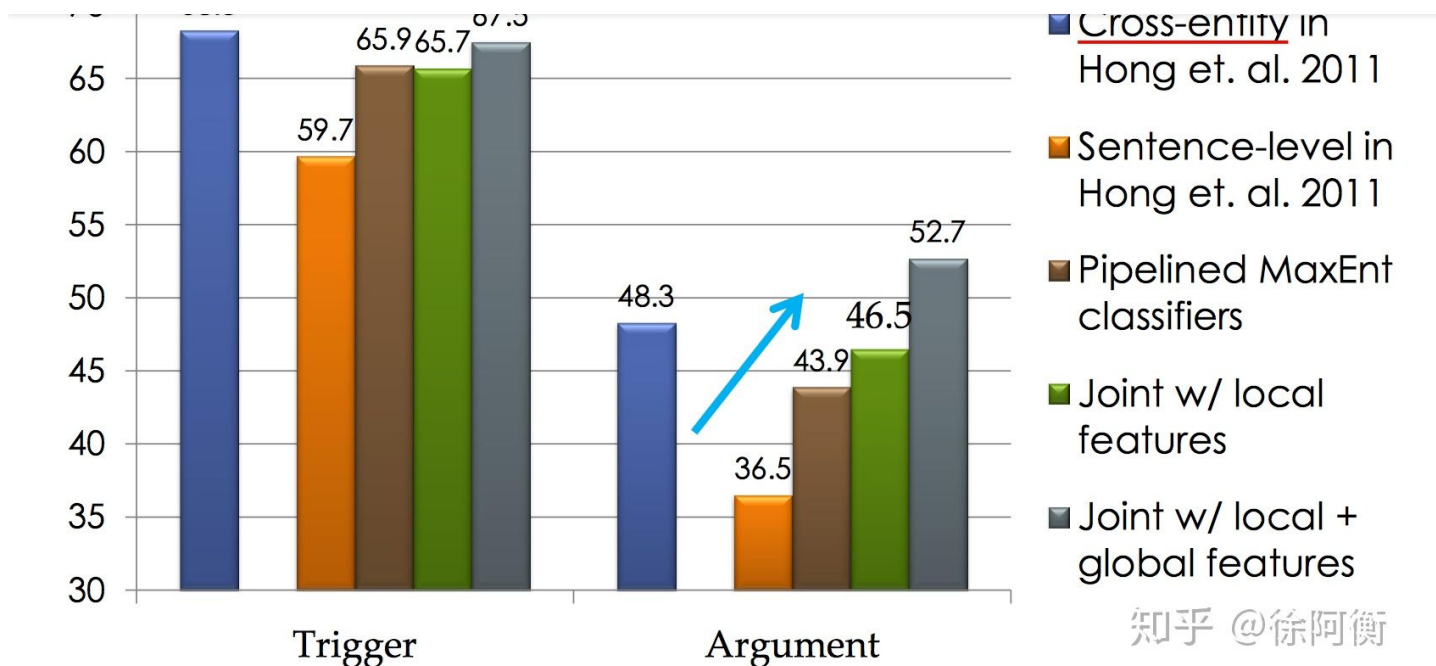
Li 提出基于结构感知机的联合模型同时完成事件触发词识别和事件元素识别两个子任务，并通过 beam search 缩小搜索解空间

- Li et.al., 2014

为了利用更多的句子级信息，Li 等提出利用结构预测模型将实体、关系和事件进行联合抽取

尽管 Li 等人的联合系统优势明显，但在未见词和特征上缺乏泛化，人工提取的特征集是离散表达，能力有限。

几种方法的 trigger 和 argument 抽取结果，可以看出，实体之间协同消歧对效果提升非常明显



## 基于统计 - 深度学习

上面的方法在特征提取过程中还是会依赖依存分析、句法分析、词性标注等传统的外部 NLP 工具，还是会造成误差积累，另外有些语言和领域并没有这类处理工具，加之特征也需要人工设定，2015 年起基于深度学习的事件抽取方法逐渐成为研究热点，相比于传统机器学习，深度学习方法优势明显：

1. 减少对外部 NLP 工具的依赖，甚至不依赖 NLP 工具，建立成端对端的系统
2. 使用词向量作为输入，蕴含更为丰富的语言特征
3. 自动提取句子特征，避免了人工特征设计的繁琐工作

## Pipeline - DMCNN

Event Extraction via Dynamic Multi-Pooling Convolutional Neural Networks Yubo Chen et. al., ACL 2015

自然语言处理中，传统 CNN 使用的最大池化对一个 feature map 只能得到一个最大值，这对事件抽取并不适用，因为事件抽取中一个句子中可能会包含多个事件，一个 argument candidate 在不同的 trigger 下也会扮演不同的角色，传统的最大池化只保留“最重要”的信息，而丢失的信息会导致 multiple-event sentence 下的事件漏分。DMCNN 使用动态多池化卷积能实现对一个句子中不同部分的最大值获取，以保留更多有价值的信息，逻辑和 PCNN 相似。



触发词，执行第二步**论元分类 (argument classification)**，同样使用 DMCNN，给 trigger 分配 arguments，同时匹配 arguments 到 role，以第二个任务为例介绍一下过程。

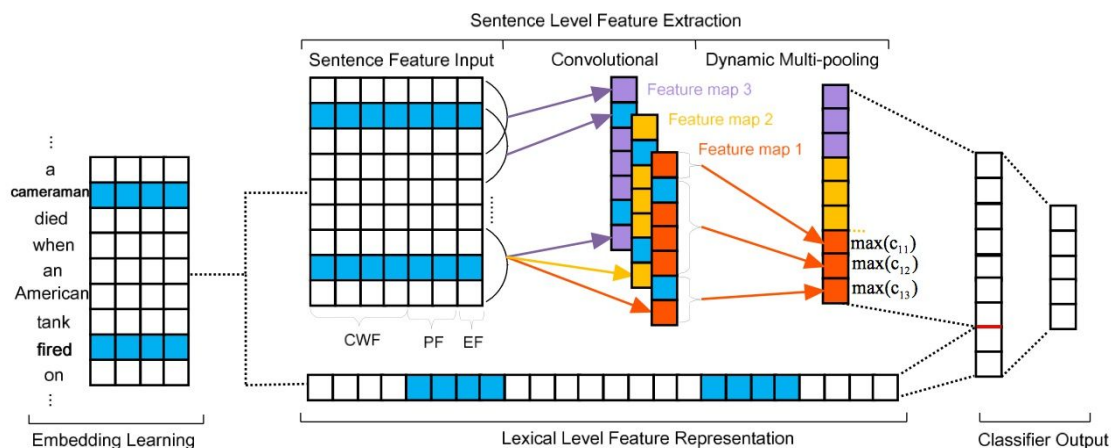


Figure 2: The architecture for the stage of argument classification in the event extraction. It illustrates the processing of one instance with the predict trigger *fired* and the candidate argument *cameraman*.

主要包括四个部分，以 argument classification 为例：

### 1. 词向量学习；

### 2. Lexical-level 词汇级别特征提取；

候选论元/触发词及其前后单词的词向量

### 3. Sentence-level 句子级别特征提取；

输入特征：

a. Context-word feature(CWF)

b. Position feature(PF)当前词语和候选论元/触发词之间的相对距离，距离值用向量表示，随机初始化

c. Event-type feature(EF)当前 trigger 对应的事件类型特征

d. CWF, PF, EF 拼接作为卷积的输入

卷积后，根据 candidate argument 和 predicted trigger 将 feature map 分成三部分，分别对各部分进行最大池化

### 4. Output 分类输出

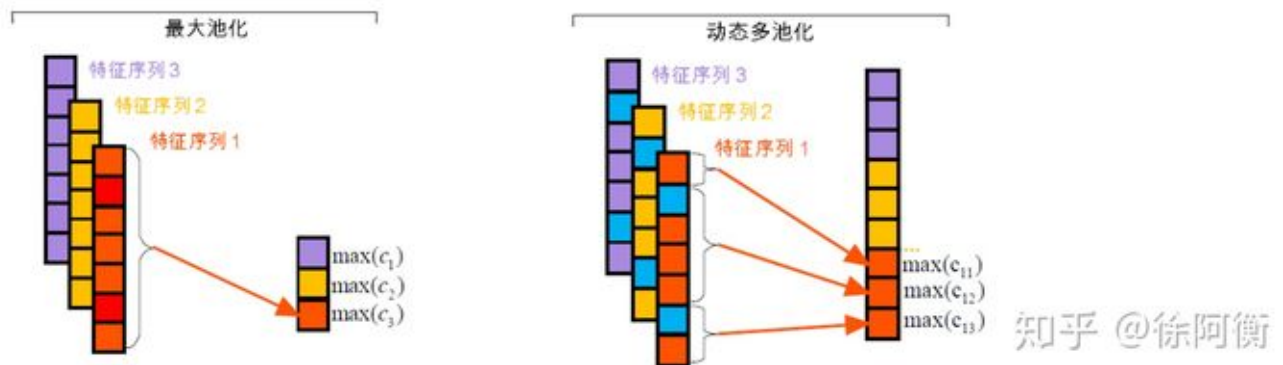
拼接词汇级别和句子级别的特征  $F=[L, P]$

$O = WF+b$  算分，进行 softmax，得到 argument role 的类别

## Trigger classification 阶段：

### • Lexical-level

S: In Baghdad, a cameraman died when an American tank fired on the Palestine Hotel.



DMCNN 的表现:

Methods	Trigger Identification(%)			Trigger Identification + Classification(%)			Argument Identification(%)			Argument Role(%)		
	P	R	F	P	R	F	P	R	F	P	R	F
Li's baseline	76.2	60.5	67.4	74.5	59.1	65.9	74.1	37.4	49.7	65.4	33.1	43.9
Liao's cross-event	N/A			68.7	68.9	68.8	50.9	49.7	50.3	45.1	44.1	44.6
Hong's cross-entity	N/A			72.9	64.3	68.3	53.4	52.9	53.1	51.6	45.5	48.3
Li's structure	76.9	65.0	70.4	73.7	62.3	67.5	69.8	47.9	56.8	64.7	44.4	52.7
<b>DMCNN model</b>	<b>80.4</b>	<b>67.7</b>	<b>73.5</b>	<b>75.6</b>	<b>63.6</b>	<b>69.1</b>	<b>68.8</b>	<b>51.9</b>	<b>59.1</b>	<b>62.2</b>	<b>46.9</b>	<b>53.5</b>

Stage	Method	1/1	1/N	all
		$F_1$	$F_1$	$F_1$
Trigger	Embedding+T	68.1	25.5	59.8
	CNN	72.5	43.1	66.3
	<b>DMCNN</b>	<b>74.3</b>	<b>50.9</b>	<b>69.1</b>
Argument	Embedding+T	37.4	15.5	32.6
	CNN	51.6	36.6	48.9
	<b>DMCNN</b>	<b>54.6</b>	<b>48.7</b>	<b>53.5</b>

DMCNN 的效果是突破性的，但分两个阶段的预测仍有误差传递的问题，也没有利用好 trigger 和 argument 之间的依赖关系。

## Joint Model - JRNN

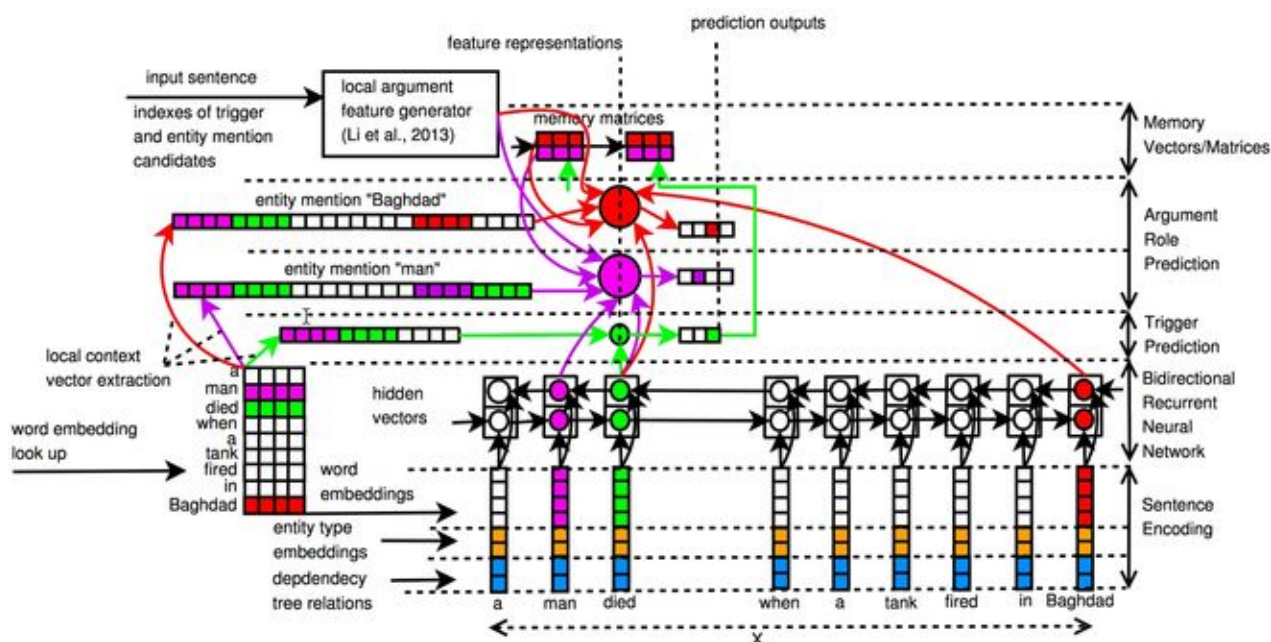


Figure 1: The joint EE model for the input sentence “a man died when a tank fired in Baghdad” with local context window  $d = 1$ . We only demonstrate the memory matrices  $G_i^{arg/trg}$  in this figure. Green corresponds to the trigger candidate “died” at the current step while violet and red are for the entity mentions “man” and “Baghdad” respectively.

## 1. Encoding phase

word embedding + entity type embedding + dependency tree relation

dependency tree relation 是 binary 的，个人理解应该是维度对应依存树中单词间所有可能的关系（如 conj\_and, advcl 等），只有在依存树  $W$  中存在与  $w_i$  连接的一条对应边（如 conj\_and 连接了  $w_i$  与  $w_j$ ）时，该维度（conj\_and 对应维度）的值才设为 1，这个向量在 Li et al., 2013 的研究中是有用的。

没有用到位置特征，因为同时预测 trigger 和 argument roles，没有固定的锚点。

双向 GRU 进行编码

## 2. Prediction phase

对  $W$  对应的 trigger 和 argument role 分别维护了一个 binary memory vector  $G_i^{trg}$ ，

binary memory matrices  $G_i^{arg}$  以及  $G_i^{arg/trg}$

每个时间点  $i$ ，或者说对  $w_i$

a. 对  $w_i$  进行 trigger 预测

b. 如果 trigger 预测结果  $t_i$  是 other，那么  $a_{ij}$   $j$  从 1-k 都设为 other，然后进行下一步，否则对所有的 entity mention  $e_1, e_2, \dots, e_k$  进行 argument role 预测

c. 利用上一步的记忆向量以及之前步骤的预测来计算  $G_i^{trg}$ ， $G_i^{arg}$  和  $G_i^{arg/trg}$

**Output:**

trigger subtype  $t_i$  for  $w_i$



Memory 向量代表的是同一个句子中触发词和论元角色之间的相互关系,  $G_i^{trg}$  代表的是 trigger subtypes 之间的关系, 表示在  $i$  之前已经识别出哪些子事件, 比如说句子中检测到了 Die 事件, 那么很有可能下面会同时会出现 Attack 事件;

$G_i^{arg}$  代表的是论元角色之间的关系, 总结了 entity mention 在过去扮演的论元角色信息;  
 $G_i^{arg/trg}$  对应的是 argument roles 和 trigger subtypes 之间的关系, 表示 entity mention 在之前特定 event subtypes 扮演过 argument,  $G_i^{arg/trg}[j][t] = 1$  代表  $e_j$  在之前的 subtype  $t$  中被认为是 argument。实验表明,  $G^{trg}$  并没有帮助反而会导致整体性能下降, 而  $G^{arg/trg}$  还是很有效的。

Stage	Model	1/1	1/N	all
Trigger	Embedding+T	68.1	25.5	59.8
	CNN	72.5	43.1	66.3
	DMCNN	74.3	50.9	69.1
	JRNN	<b>75.6</b>	<b>64.8</b>	<b>69.3</b>
Argument	Embedding+T	37.4	15.5	32.6
	CNN	51.6	36.6	48.9
	DMCNN	<b>54.6</b>	48.7	53.5
	JRNN	50.0	<b>55.2</b>	<b>55.4</b>

Table 4: System Performance on Single Event Sentences (1/1) and Multiple Event Sentences (1/N). 知乎 @徐阿衡

当输入句子包含多个事件时 (1/N), JRNN 明显优于其他方法。特别是, JRNN 在触发词识别上 DMCNN 好13.9%, 而论元分类的相应改进为 6.5%, 从而进一步表明 JRNN 具有记忆功能的好处。在单事件句子 (1/1) 的表现上, JRNN 在触发词分类上仍然是最好的系统, 尽管在论元分类上比 DMCNN 要差一些。

训练好的分类器判断未标注的数据，选取高置信度的分类样本作为训练样本，通过迭代自动扩充训练样本[Chen and Ji, 2009]。Liao 等在相关文档中使用自训练的（Self-Training）的半监督学习方法扩展标注语料，并利用全局推理的方法考虑样例的多样性进而完成事件抽取；进一步提出同时针对词汇和句子两个粒度训练最大熵分类器，并用协同训练（Co-training）的方法扩展标注数据，进而对分类器进行更充分的训练[Liao and Grishman, 2011a; 2011b]。

而目前，弱监督/训练数据生成方面比较流行的方向有 **利用外部资源，通过远程监督，以及跨语料迁移的方法**。

## 外部资源

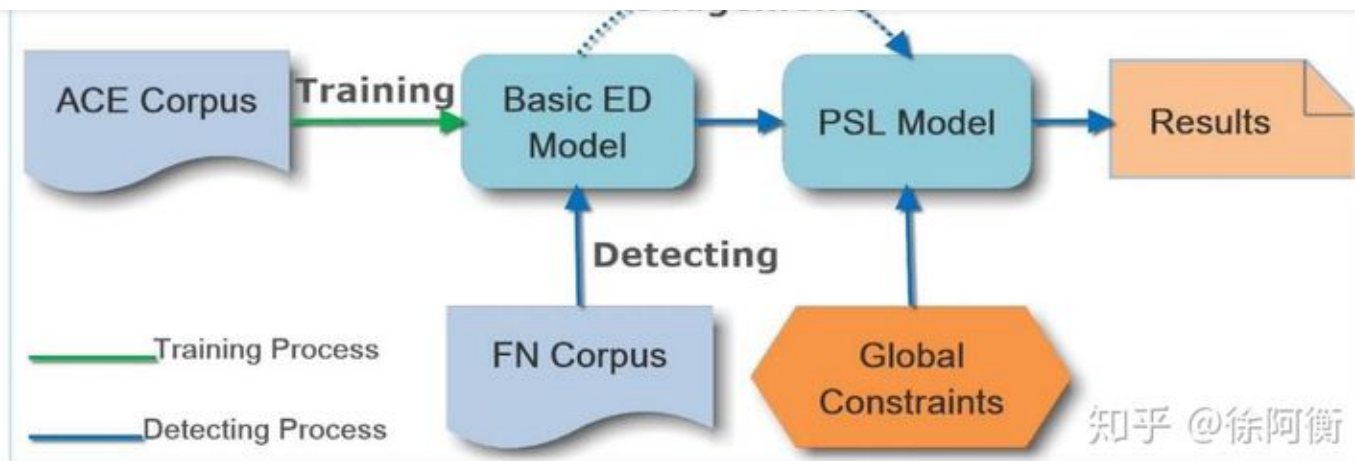
### Leveraging FrameNet to Improve Automatic Event Detection, ACL2016

FrameNet 是语言学家定义及标注的语义框架资源，采用层级的组织结构，有1000+框架、1000+词法单元、150000+标注例句。在结构上，FrameNet 和事件抽取有着很高的相似性，一个框架由一个词法单元和若干框架元素组成，一个事件有触发词和若干事件角色组成。另外，FrameNet 中很多 frame 其实也能够表示某些事件，如

框架名称	FrameNet 中的例句	事件类型
Attack	<i>Aeroplane <u>bombed</u> London.</i>	Attack
Rape	<i>Girl, 16, <u>raped</u> by gang.</i>	Attack
Fining	<i>The court <u>fined</u> her \$40.</i>	Fine
Execution	<i>He was <u>executed</u> yesterday.</i>	Execution

因此，Liu 等利用 ACE 语料训练的分类器去判定 FrameNet 中句子的事件类别，再利用全局推断将 FrameNet 的语义框架和 ACE 中的事件类别进行映射，进而利用 FrameNet 中人工标注的事件样例扩展训练数据以提升事件检测性能 [Liu et.al., 2016b]。

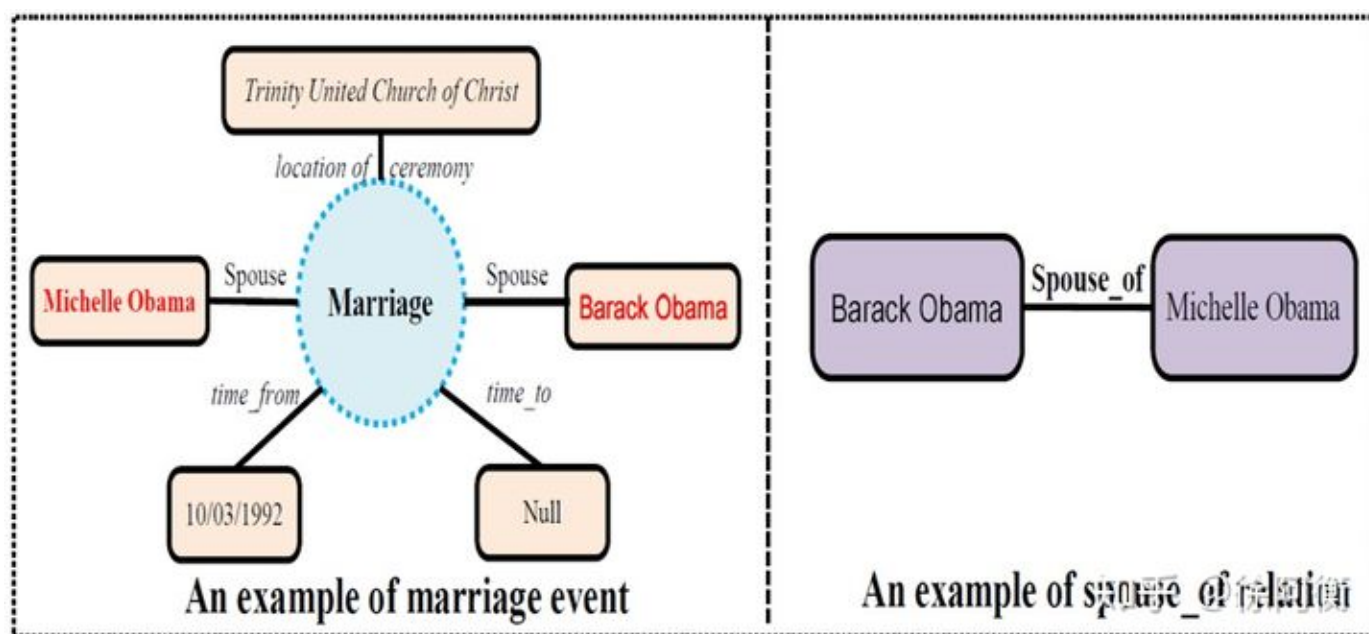




## 远程监督

Automatically Labeled Data Generation for Large Scale Event Extraction, ACL2017

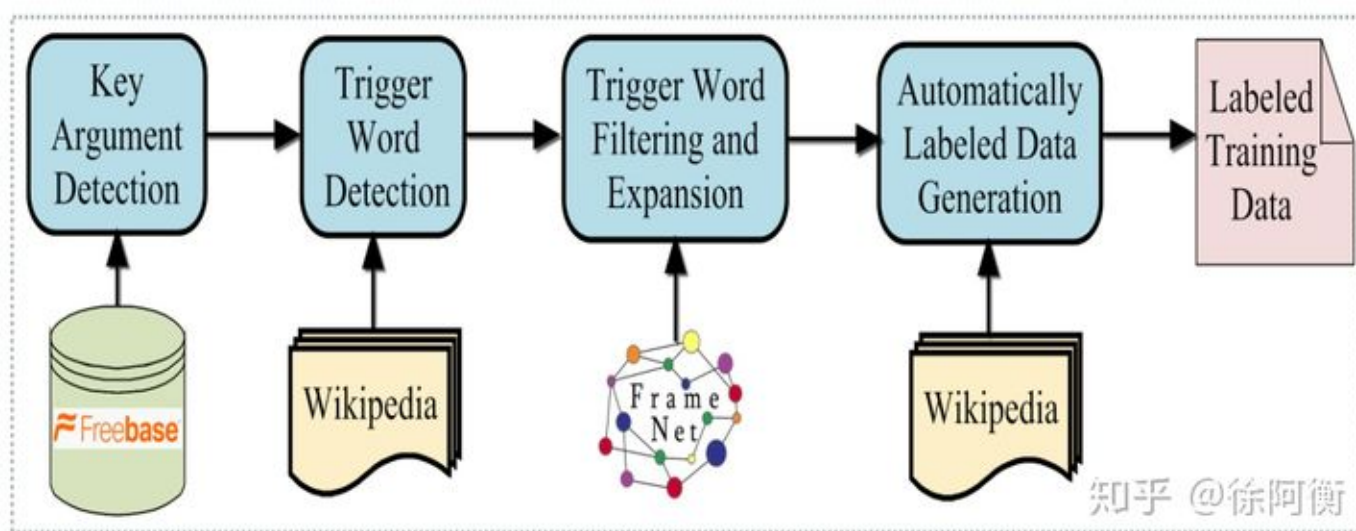
Yubo Chen 提出运用结构化的知识库来以及远程监督的方法来自动生成大规模事件语料。



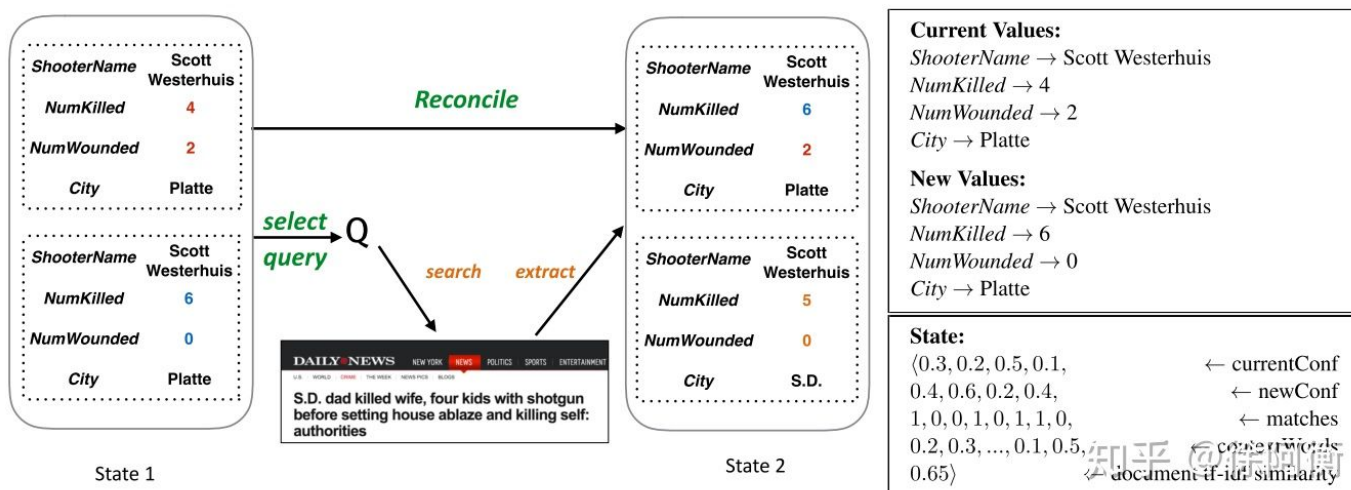
当把关系抽取中常用的远程监督方法用到事件抽取中时，会发现下面两个问题，一是**现有事件知识库（如 Freebase）中缺乏触发词信息**，如上图，在关系抽取中，我们可以用两个论元 Barack Obama, Michelle Obama 进行回标，但是在事件抽取中，marriage 这一事件类型在 Freebase 中被表示为 m.02nqglv，所以我们不能直接用事件类型和论元来进行回标，在用 DS 前，必须先

如果用所有论元来进行句子的回标，那么能抽出的训练数据就非常少了，所以应该对论元进行排序，选择有代表性的论元进行回标。

整个流程如下，首先对 Freebase 中的核心论元进行检测，根据**角色显著性 (role saliency)**、**事件相关性 (event relevance)** 和**核心率 (key rate)** 对论元进行优先级排序，接着利用所有的核心论元去 Wikipeda 中回标，根据**触发率 (trigger rate)**、**触发词频率 (trigger candidate frequency)**、**触发词事件频率 (trigger event type frequency)** 来进行触发词检测，这一阶段得到的触发词表中只有动词，缺少名词，也存在噪声，于是再利用 FrameNet 过滤动词性触发词中的噪声，同时扩展名词性触发词，最后利用 Soft Distant Supervision 来自动生成标注数据。



还有方法如 [Karthik Narasimhan et al., EMNLP 2016](#)，从网络获取同一事件的不同报道，再使用强化学习方法，做信息融合的决策（互补信息的融合、冗余信息的选择）。



由于目前中文事件抽取缺少公认语料，很多学者尝试利用现有大量的高质量英文标注语料辅助中文事件抽取。Chen 等首次提出该想法并利用跨语言协同训练的 Bootstrap 方法进行事件抽取[Chen and Ji, 2009]。Ji 提出基于中英文单语事件抽取系统和基于并行语料两种构建跨语言同义谓词集合的方法辅助进行中文事件抽取[Ji, 2009]，Zhu 等利用机器翻译同时扩大中文和英文训练语料，联合利用两种语料进行事件抽取[Zhu et.al., 2014]。Hsi 等联合利用符号特征和分布式特征的方法，利用英文事件语料提升中文事件抽取的性能[Hsi et.al., 2016]。

## Event Detection via Gated Multilingual Attention Mechanism, AAAI2018

Motivation:

1. **多语言一致性**，不同语言中表达了相同含义的句子往往包含相同的语义成分  
如 MeiGuo TanKe 和 American tank 表达了相同含义，都是武器
2. **多语言互补**，某个词在一种语言中有歧义，但在另一种语言中却没有歧义  
如英文 fire，因为有开火和解雇两种意思，所以对应事件可能是 Attack 也可能是 End-Position，然而在中文中开火，Attach 类型，解雇就是 End-Position 类型，两个词没有相同语义

所以文章提出了两种 attention 机制，一是利用多语言一致性，分别对每种语言进行单语语境的注意力计算，对每个候选触发词，对其上下文进行注意力机制，注意力权重表示句子中不同单词对预测事件类型的重要性，二是利用互补信息，用 gated cross-lingual attention 来模拟其他语言的可信度，gate 来控制目标语言流向源语言的信息，集成多语言的信息。

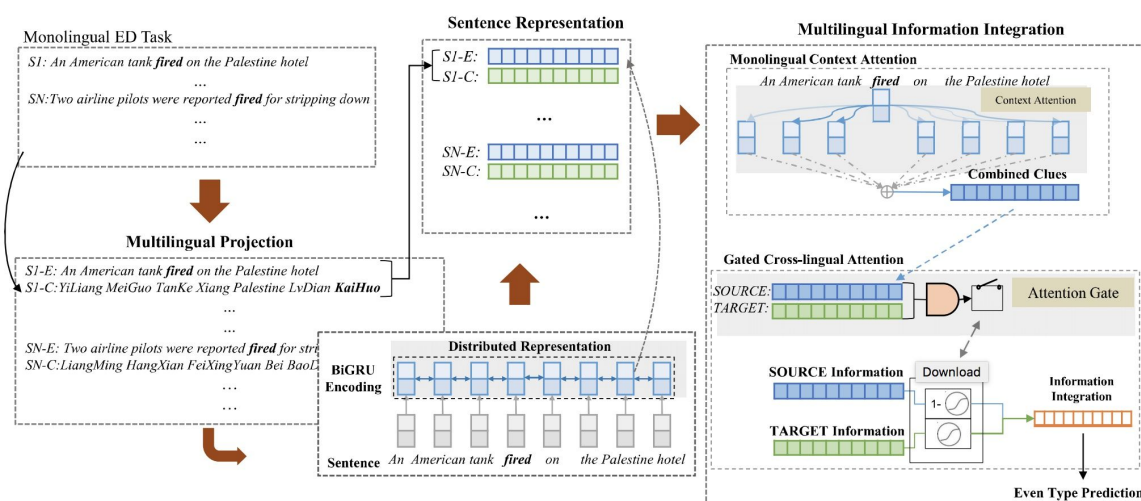


Figure 2: The overall architecture of the GMLATT framework. After multilingual projection, the framework leverages BiGRU encoder to encode sentence to distributed representation. Monolingual context attention and gated cross-lingual attention are employed to do multilingual information integration.

目前事件抽取的相关研究大部分是面向英文文本，中文文本的工作才刚起步，一方面，中文的自身特点（需要分词、缺少时态和形态的变换）有一定挑战，另一方面，数据集上也缺乏统一、公认的语料资源和相关评测。尽管如此，近年来中文事件抽取在公开评测、领域扩展及上述的跨语料迁移方面也都取得了一些进展。

公开评测方面，除了在模型方面的创新[Chen and Ng, 2012; Li et.al., 2012a; 2013b]，在中文语言特性的利用方面，Li 等通过中文词语的形态结构、同义词等信息捕获更多的未知触发词，进而解决中文事件抽取面临的分词错误和训练数据稀疏等问题；进一步细分中文事件触发词内部的组合语义（复合、附加和转化），进而提高系统的性能[Li et.al., 2012b]。Ding 等利用聚类的方法自动生成新事件类型的语料，在抽取过程中特别地考虑了待抽取文本的 HowNet 相似度[Ding et.al., 2013]。

特定领域方面，国内很多机构均面向实际应用展开特定领域的事件抽取研究，覆盖突发灾难、金融、军事、体育、音乐等多个领域。例如，Zhou 等针对金融领域事件中的收购、分红和贷款三个典型事件，提出自动构建抽取规则集的方法进行中文金融领域事件抽取 [Zhou, 2003]；Liang 等利用事件框架的归纳和继承特性实现对灾难事件的抽取[Liang and Wu, 2006]。

其他方向的一些 Paper:

#### 特征表示:

– Argument Attention: Exploiting Argument Information to Improve Event Detection via Supervised Attention Mechanisms (ACL2017)

#### 多事件抽取:

– HBTNGMA: Collective Event Detection via a Hierarchical and Bias Tagging Networks with Gated Multi-level Attention (EMNLP-2018)

#### 篇章级事件抽取:

– DCFEE: A Document-level Chinese Financial Event Extraction System based on Automatically Labeled Training Data (ACL 2018)

#### 事件关系抽取:

– ATT-ERNN: Attention-based Event Relevance Model for Stock Price Movement Prediction (CCKS-2017 Best Paper Award)  
– MLNN: Event Coreference Resolution via Multi-loss Neural Network without Arguments (CCKS-2018)

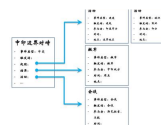
## 事件监测和追踪

主流方法包括基于相似度聚类 and 基于概率统计两类。在这不多做介绍。以后有时间再补充。

戳原文有获取更整洁的侧边目录显示：

知识抽取-事件抽取

[www.shuang0420.com](http://www.shuang0420.com)



博客：[www.shuang0420.com](http://www.shuang0420.com)

公众号：xu\_a\_heng

关注公众号回复“事件抽取”获得相关综述及论文下载地址~



欢迎关注公众号：徐阿衡

编辑于 2018-11-26

「真诚赞赏，手留余香」

赞赏

还没有人赞赏，快来当第一个赞赏的人吧！

[深度学习 \(Deep Learning\)](#)

[自然语言处理](#)

[知识图谱](#)

## 文章被以下专栏收录



徐阿衡-自然语言处理

欢迎关注同名微信公众号：徐阿衡 (xu\_a\_heng)

## 推荐阅读



徐阿衡

知识抽取-实体及关系抽取

## 10 条评论

⇌ 切换为时间排序

写下你的评论...



杏仁糖

2018-11-26

很全啦



1



钻地鼯鼠

01-07

请问framenet有中文的吗？



赞



dr.china

2019-08-09

事件抽取



赞



SilenceWinter

2019-06-28

您好，请教下，“DMCNN 作者把事件抽取看做两个阶段的多分类任务，第一步是触发词分类（trigger classification），利用 DMCNN 对句子中每个词进行分类，判断是否是触发词”，我对这个模型的输出形式不理解。请问如果句子S1有10个词，S2有30个词，这怎么对每个词分类？谢谢~



赞



八八 回复 SilenceWinter

04-13

逻辑很简单，以每个词都作为trigger或者是argument candidate，按照模型逻辑去 encode 文本信息，最终映射到label space之后，判断该词汇的类型。训练阶段，优化最大似然，在推理阶段，依次对每个词进行推理，判断该词的类别。



赞

沈月廷又！

👍 赞



Scofield

2019-06-06

图一有问题。row 1 应该不是事件类型，应该是论元角色。

👍 赞



Zzzzzzzz

2019-02-25

和 知识图谱发展报告的时间抽取有好大一部分重叠

👍 赞



李明磊9527 回复 Zzzzzzzz

2019-05-29

[cips-upload.bj.bcebos.com...](https://cips-upload.bj.bcebos.com...)

👍 1



霍华德 

2018-11-26

不错 挺全面的 超级棒

👍 赞