

# 基于双向循环卷积神经网络 的实体关系分类

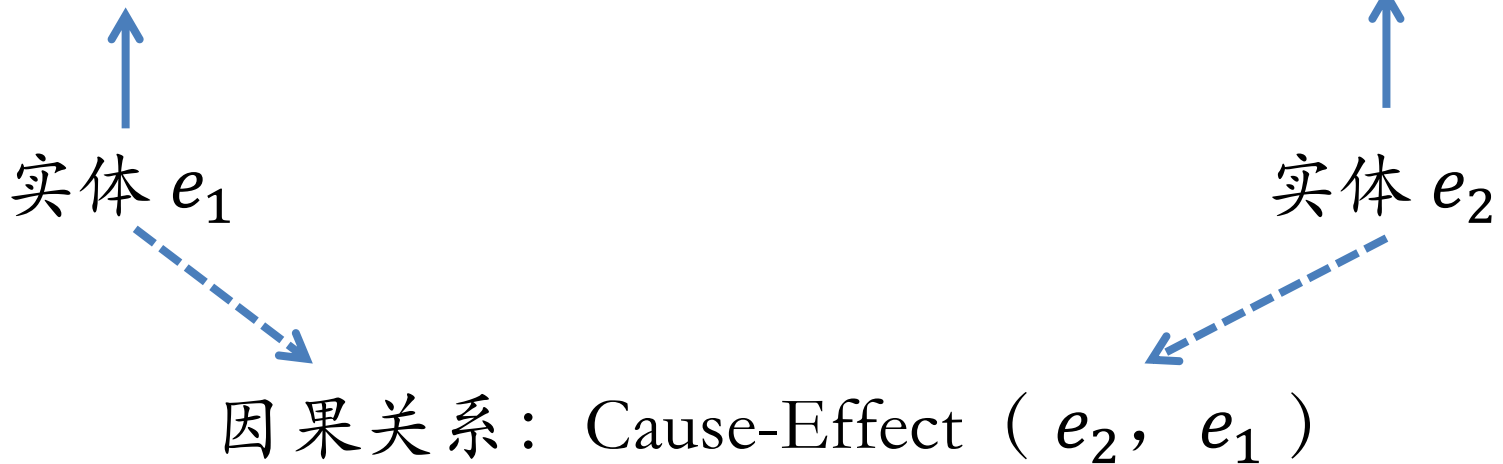
报告人：蔡蕊

单位：北京大学计算语言所

# 任务描述

SemEval 2010关系分类数据集中的一个例子：

The **burst** has been caused by water hammer **pressure**.

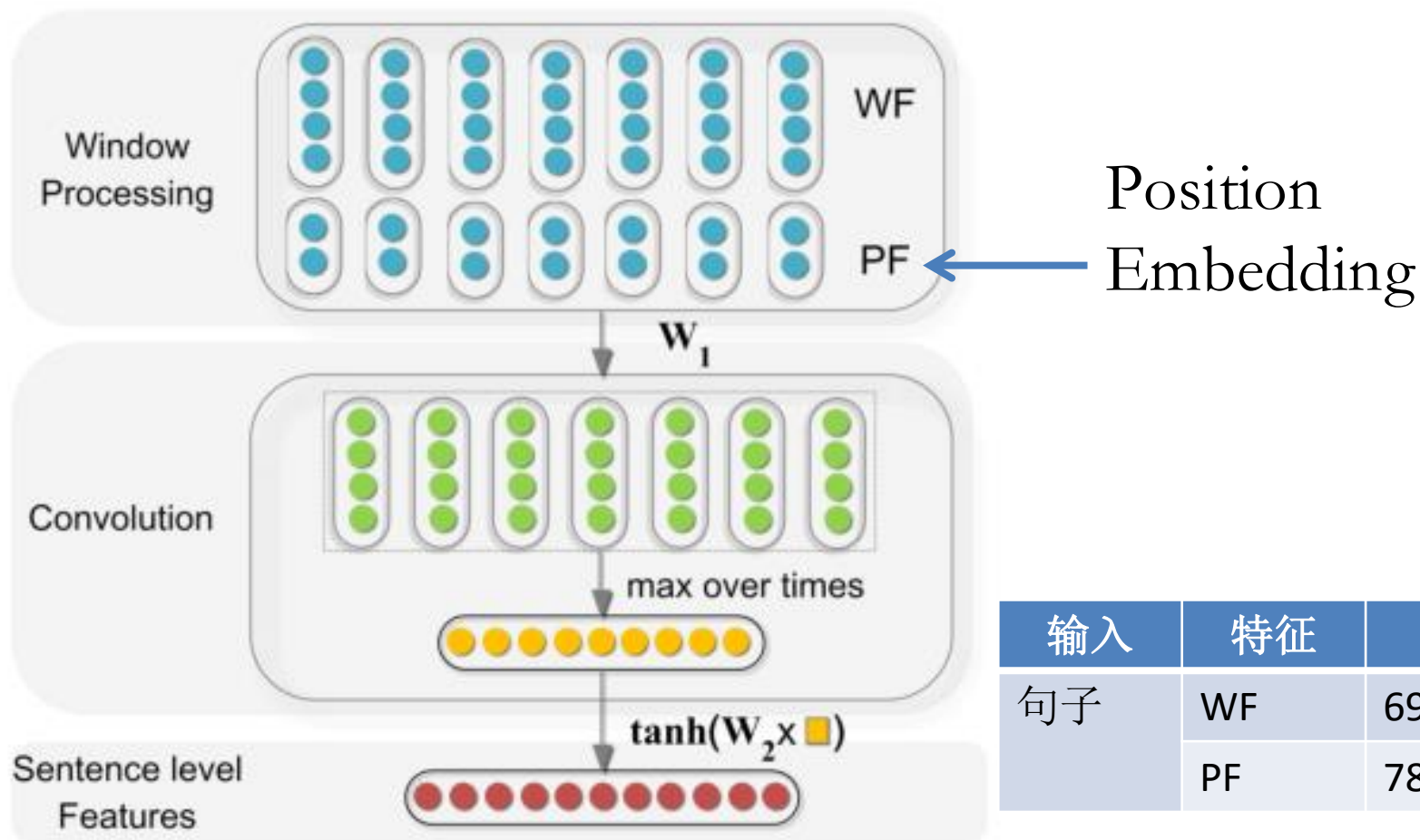


# 关系分类的传统方法

- 传统的基于特征的方法：
  - 词汇特征  
需要分类的实体对的特征，如实体的**POS**标注。
  - 语法特征  
实体对所在句子的语法结构，如依存分析。
  - 语义特征  
实体对中实体的语义信息，如实体在**WordNet**中的上位词。

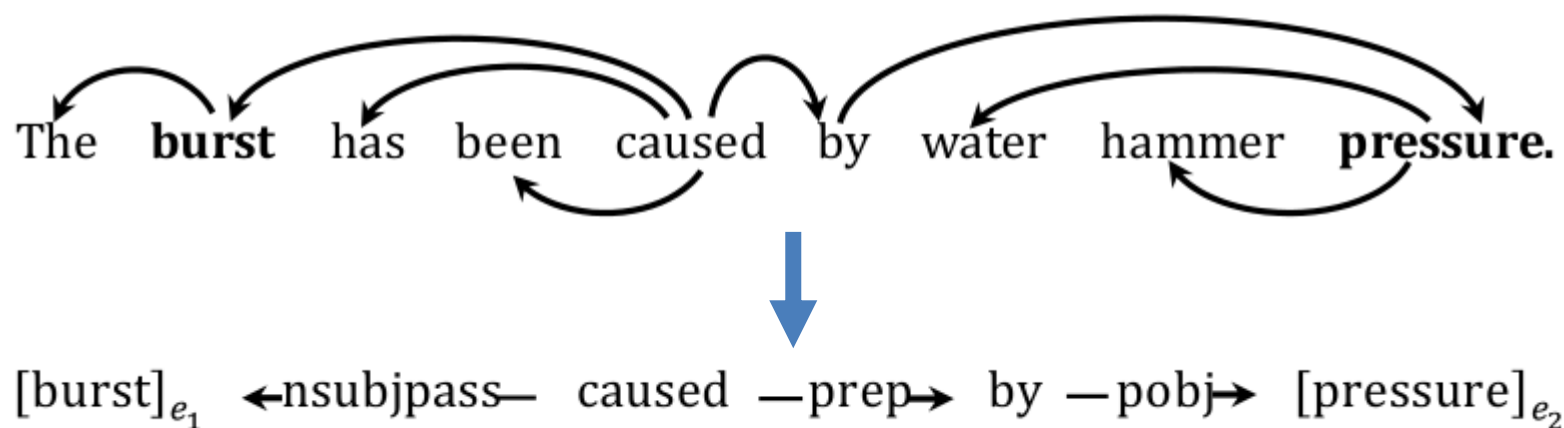
# 卷积神经网络结合位置特征

Zeng et al. 2013



输入	特征	F1
句子	WF	69.7
	PF	78.9

# 最短依存路径



- 与使用整个句子的信息相比：
  - 最短依存路径保留了对关系判断的关键信息。
  - 去除了与实体关系无关的噪声信息。

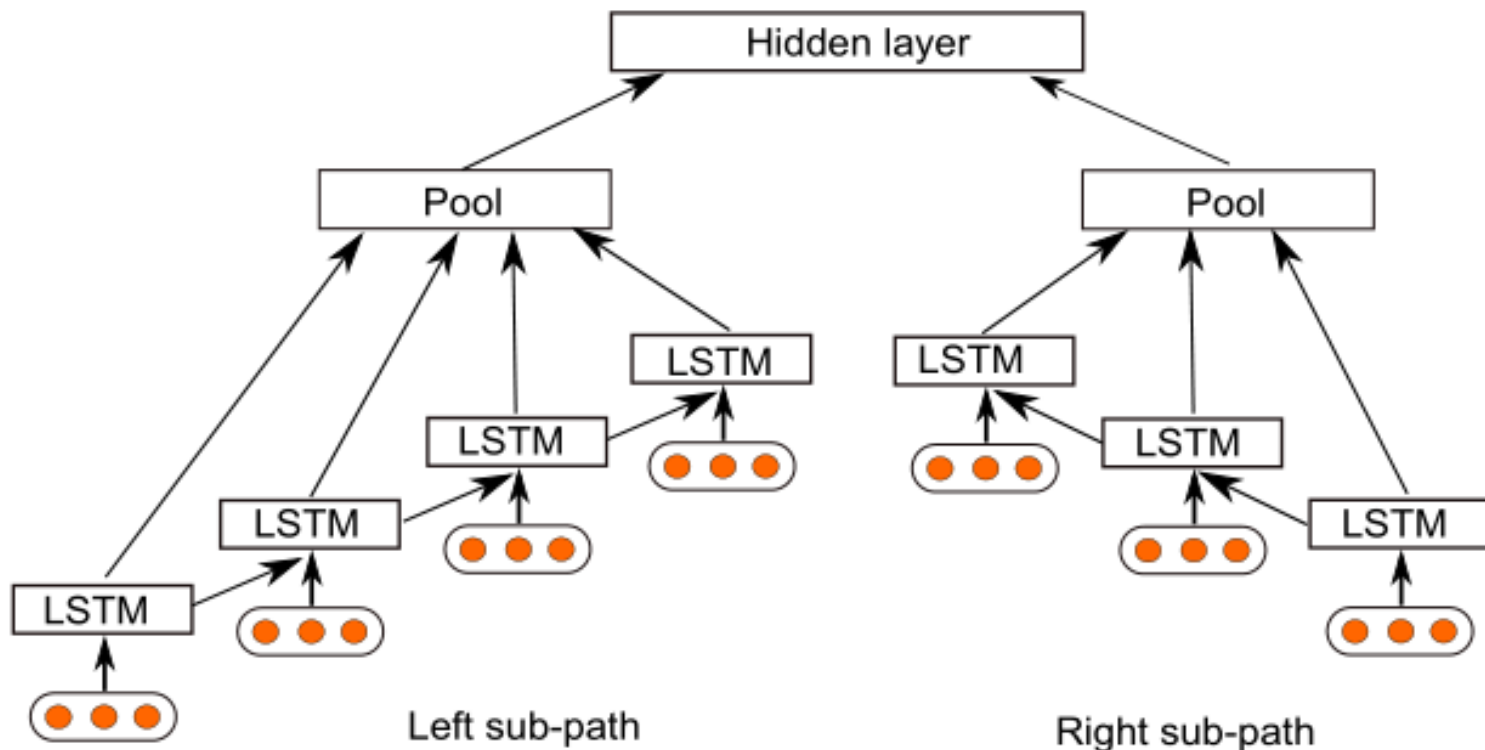
# 循环神经网络

Xu et al. 2015

$[burst]_{e_1}$   $\leftarrow$  nsubjpass — caused — prep  $\rightarrow$  by — pobj  $\rightarrow$   $[pressure]_{e_2}$

Left sub-path

Right sub-path



# 最短依存路径

- 最短依存路径的结构特殊性:

- 词与依存关系交替出现:

$[burst]_{e_1} \leftarrow nsubjpass - caused - prep \rightarrow$  by - pobj  $\rightarrow [pressure]_{e_2}$

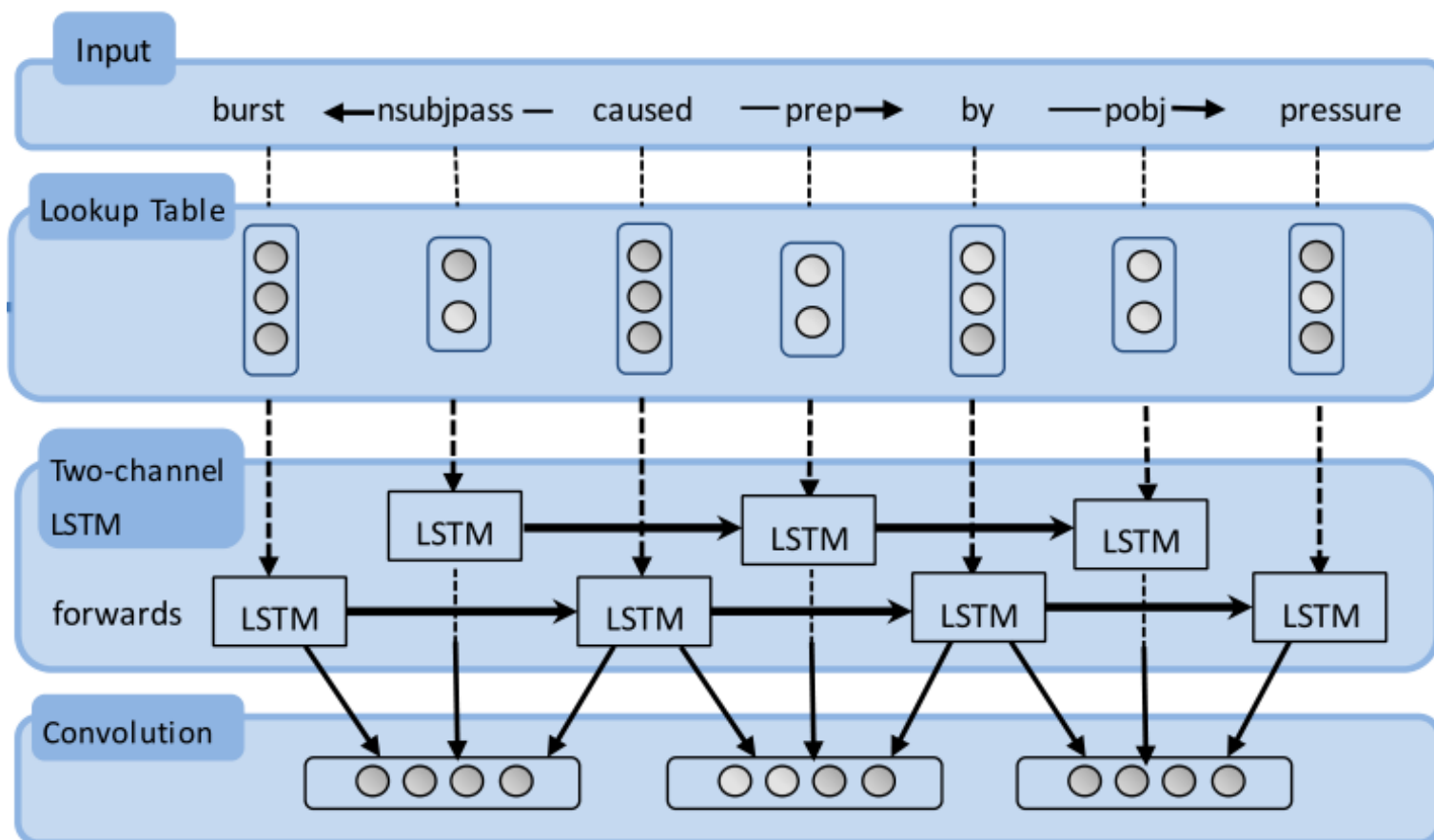
“依存单元”

- 方向上具有可逆性

Cause-Effect( $e_2, e_1$ )	$[burst]_{e_1} \leftarrow nsubjpass - caused - prep \rightarrow$ by - pobj $\rightarrow [pressure]_{e_2}$
Cause-Effect( $e_1, e_2$ )	$[pressure]_{e_1} \leftarrow pobj - by \leftarrow prep - caused - nsubjpass \rightarrow [burst]_{e_2}$

- 传统的CNN或RNN无法很好地最短依存路径进行建模。

# 对依存单元建模

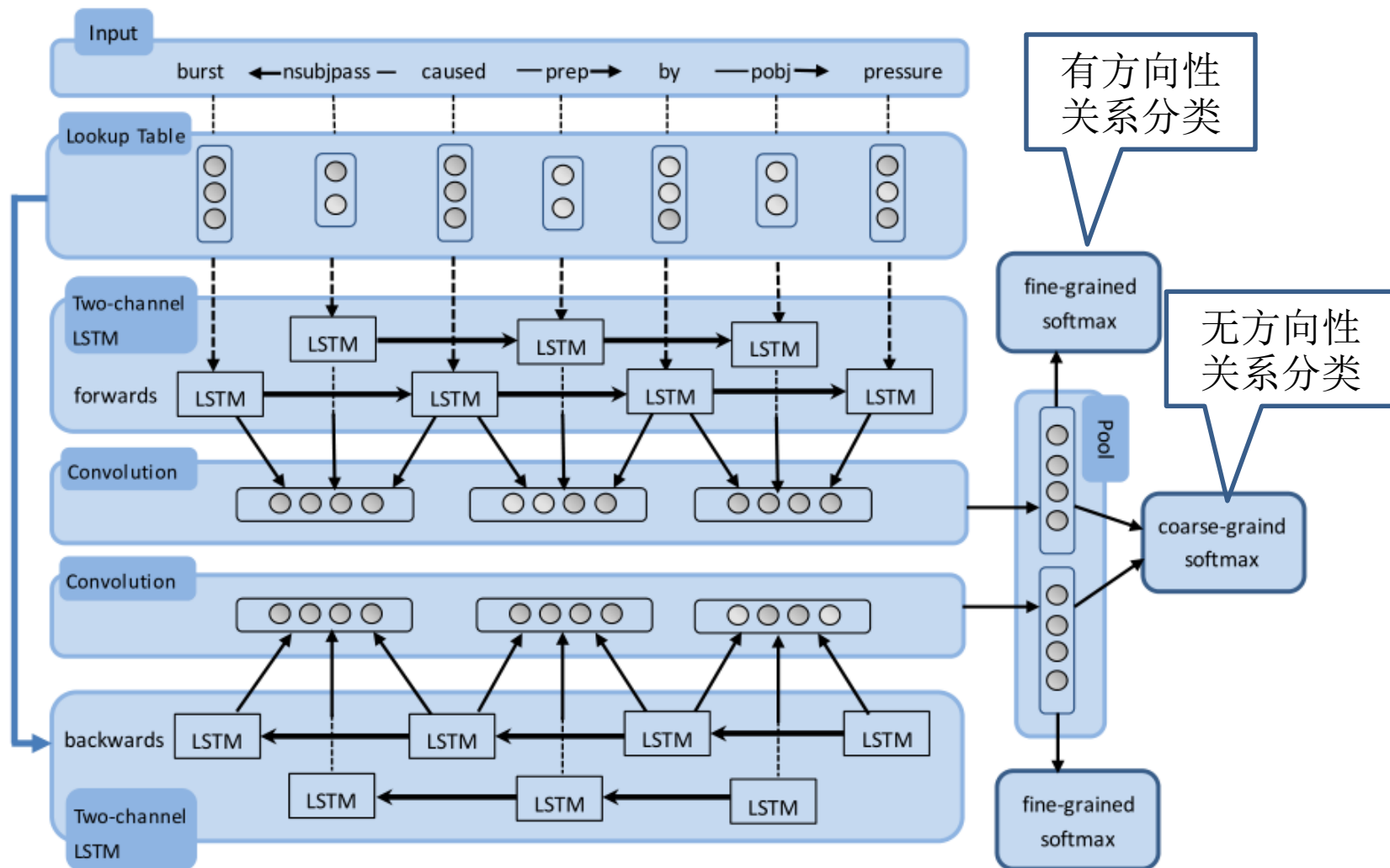


循环卷积神经网络（RCNN）



# 双向循环卷积神经网络

RCNN



**BRCNN**

# SemEval 2010数据集

- 数据集构成:

9种具有  
方向性的  
关系  
类别

关系	训练数据	测试数据
Other	17.63%	16.71%
Cause-Effect	12.54%	12.07%
Component-Whole	11.76%	11.48%
Entity-Destination	10.56%	10.75%
Product-Producer	8.96%	9.61%
Entity-Origin	8.95%	9.50%
Member-Collection	8.63%	8.58%
Message-Topic	7.92%	8.50%
Content-Container	6.75%	7.07%
Instrument-Agency	6.30%	5.74%

# 训练与测试

- 训练过程:

- 对于非Other实例进行双向训练
- 对于Other实例只进行单向训练

损失函数

$$J = \sum_{i=1}^{2K+1} \vec{t}_i \log \vec{y}_i + \sum_{i=1}^{2K+1} \overleftarrow{t}_i \log \overleftarrow{y}_i + \sum_{i=1}^K t_i \log y_i + \lambda \cdot \|\theta\|^2$$

- 测试过程:

**for** 对于测试集中的每一个样例 **do**

将样例的正向最短依存路径作为 BRCNN 模型的两个输入,  
根据分类结果判断样例是否为 Other 类

**if** BRCNN 模型判断样例属于 Other 类 **then**

输出分类结果: Other 类

**else**

将样例的正向和反向最短依存路径作为 BRCNN 模型的两个输入,  
结合两个 ( 2K+1 ) 类的分布  $\vec{y}$  和  $\overleftarrow{y}$  对样例分类。

**end if**

**end for**

# SemEval 2010数据集

使用句子作为输入

使用最短依存路径作为输入

Classifier	Additional Information	$F_1$
SVM (Rink and Harabagiu, 2010)	POS, WordNet, Prefixes and other morphological features, dependency parse, Levin classed, PropBank, FanmeNet, NomLex-Plus, Google $n$ -gram, paraphrases, TextRunner	82.2
RNN (Socher et al., 2011)	Word embeddings + POS, NER, WordNet	74.8 77.6
MVRNN (Socher et al., 2012)	Word embeddings + POS, NER, WordNet	79.1 82.4
CNN (Zeng et al., 2014)	Word embeddings + word position embeddings, WordNet	69.7 82.7
FCM (Yu et al., 2014)	Word embeddings + dependency parsing, NER	80.6 83.0
CR-CNN (dos Santos et al., 2015)	Word embeddings + word position embeddings	82.8 84.1
SDP-LSTM (Xu et al., 2015b)	Word embeddings + POS + GR + WordNet embeddings	82.4 83.7
DepNN (Liu et al., 2015)	Word embeddings, WordNet Word embeddings, NER	83.0 83.6
depLCNN (Xu et al., 2015a)	Word embeddings, WordNet, word around nominals + negative sampling from NYT dataset	83.7 85.6
BRCNN (Our Model)	Word embeddings + POS, NER, WordNet embeddings	85.4 <b>86.3</b>

# SemEval 2010数据集

- 模型自身的比较

模型	$F_1$
CNN	81.8
LSTM	76.6
Two-channel LSTM	81.5
RCNN	<b>82.4</b>

- 模型不同输入的比较

模型	输入	$F_1$
RCNN	$\vec{S}$ of all relations	82.4
Bi-RCNN	$\vec{S}$ and $\overleftarrow{S}$ of all relations	81.2
Bi-RCNN	$\vec{S}$ and $\overleftarrow{S}$ of directed relations , $\vec{S}$ of Other	84.9
BRCNN	$\vec{S}$ and $\overleftarrow{S}$ of directed relations, $\vec{S}$ of Other	<b>85.4</b>

- Other类包含大量噪声。
- 说明了对Other类特殊处理的必要性。
- 双向学习的有效性。

Thanks !

Q & A