

# 对话系统和语义理解技术

陈见耸

2016/12/07

# 大纲

- 对话系统分类
- 语义理解
- 词向量
- 基于统计模型的语义理解技术
  - 短文本分类
  - 语义匹配
  - 序列标记

# 对话系统分类

- 问答式对话(question answering)
  - 以给用户精确的答案为目标
  - “姚明的女儿的生日” “2010年5月22日”
  - 代表：IBM Watson
- 闲聊式对话(chat bot)
  - 以满足某种情感需求的闲聊为目标
  - “你今天心情怎么样” “挺好的”
  - 代表：小冰
- 任务式对话(task oriented dialogue)
  - 以明确用户意图完成任务为目标
  - “北京今天天气如何”
  - 代表：语音助手

# 问答式对话技术

自然语言问句  
(非结构化)



深度问答系统



精确答案

结构化知识



非结构化知识

Knowledge Graphs  
知识库

Triples  
抽取事实

Texts  
文本



(Chicago, be most city in, the United States)  
(Charlotte, grandpa, Bill Clinton)  
(Chicago, locate in Illinois)  
(Hillary, be politician in, the democratic party)  
(Edward H. Allegratti, be born in, Chicago)



# 聊天式对话技术

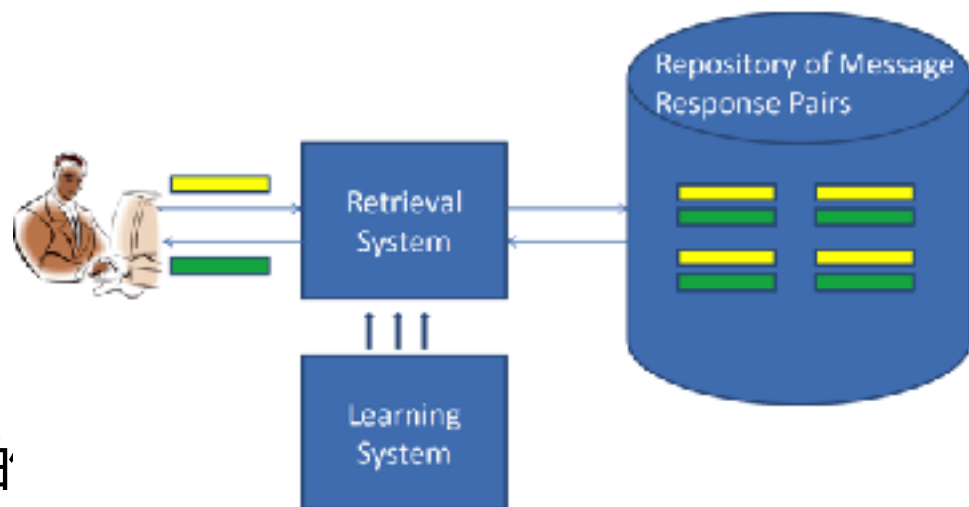
## ■ 检索式方案

### ■ 优势

- 基本不会有语法错误
- 小数据量时仍可work

### ■ 劣势

- 无法处理没见过的情況
- 不能重新提到上下文中的实体信息
  - Q:“我叫陈见耸，很高兴见到你”
  - A:“你好陈见耸”



# 聊天式对话技术

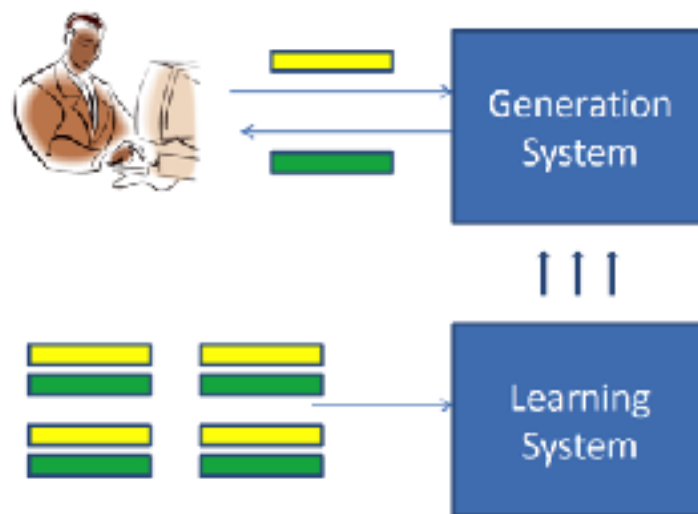
## ■ 生成式方案

### ■ 优势

- 比较容易和上下文、知识、用户信息融合
- 可以在回复中嵌入上下文遇到的实体

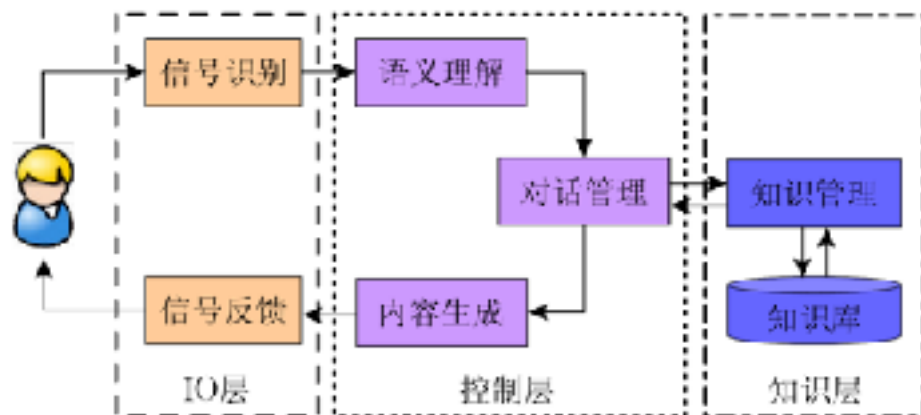
### ■ 劣势

- 需要大量的句子对训练
- 很可能犯语法错误
- 回复的多样性问题



# 任务式对话技术

- 语义理解(NLU)
  - 将非结构化的文本信息转化为结构化的语义表示
    - 领域、意图和槽
- 对话管理(DM)
  - 根据语义理解的结果以及对话的上下文语境等进行综合分析,来决定采取什么动作(action)
    - 状态更新
- 自然语言生成(NLG)
  - 将系统的动作转换为系统能够理解的自然语言
    - 模板



# 语义理解(NLU)

## ■ 文本转化为计算机可理解的语义表示

- 领域(domain)、意图(intent)、槽(slot)

- 样例

- “今天北京天气怎么样”
- “北京今天天气如何”
- “查一下天气，今天北京的”

- 语义表示

- Domain=weather
- Intent=query\_weather
- Slot: { city = 北京, date= 今天}



# 语义理解的挑战

## ■ 语音识别错误

- WER: 5% – 20%

## ■ 口语错误和重复

- “25号中午提醒我打电话给小王，哦不，26号”
- “我要一个，一个中杯拿铁送到公司”

## ■ 歧义性

- “明早八点叫我起床”
- “我要听白雪公主”

## ■ 自然语言的多样性

# 语义理解方法

## ■ 基于规则的方法

### ■ 方法

- 正则匹配
- 基于关键词的匹配

### ■ 优势

- 灵活
- 不需要训练数据

### ■ 劣势

- 需要大量的规则
- 当规则越来越多， 越难维护

# 语义理解方法

## ■ 基于统计模型的方法

### ■ 方法

- 根据标注数据训练模型
- 意图和领域识别：分类问题或匹配问题
- 槽抽取（Slot Filling）：序列标记问题

### ■ 优势

- 数据驱动：更多的数据->更好的性能
- 鲁棒：具有泛化能力

### ■ 劣势

- 需要足够的精确标注数据
- 比较难解释和训练

# 词向量

## ■ One-hot Representation

“话筒”表示为 [0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 ...]

“麦克”表示为 [0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 ...]

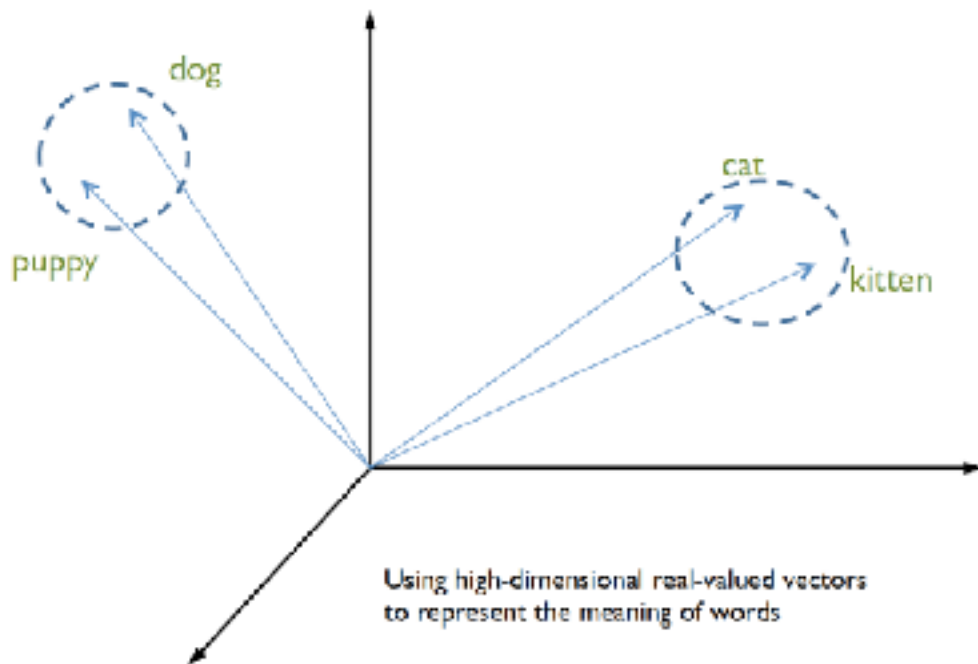
## ■ $\text{Sim}(\text{话筒}, \text{麦克}) = 0$

- 孤立  $\rightarrow$  语义鸿沟
- 维度灾难

# 词向量

## ■ Distributed Representation

- $[0.792, -0.177, -0.107, 0.109, -0.542, \dots]$
- 紧致、稠密的表示



# 分布式表示的模型

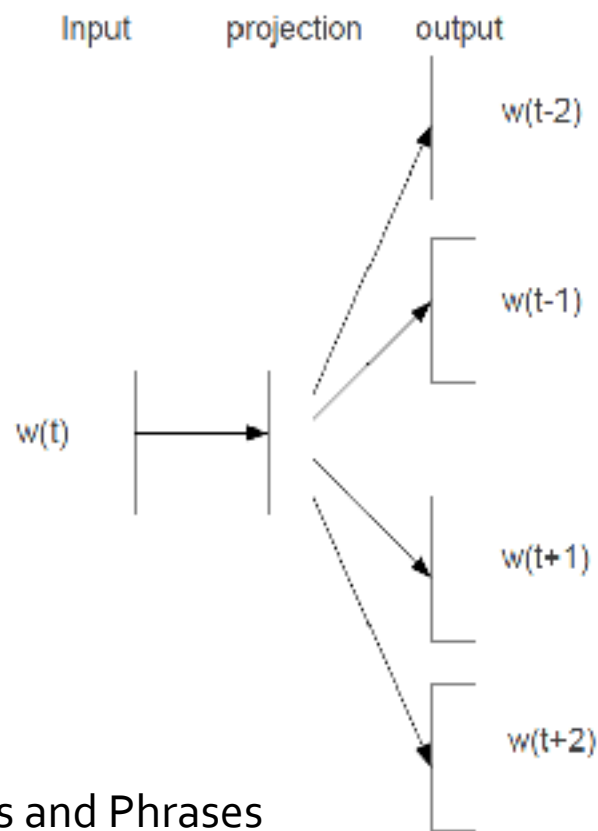
## ■ Skip-gram模型[Miklov 2013]

$$\frac{1}{T} \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{t+j} | w_t)$$
$$p(w_O | w_I) = \frac{\exp(v'_{w_O} \top v_{w_I})}{\sum_{w=1}^W \exp(v'_w \top v_{w_I})}$$

## ■ 特点

- 无顺序信息、无隐藏层
- 结构简单，效率高

Tomas Miklov, Distributed Representations of Words and Phrases and their Compositionality, NIPS 2013



# 词汇语义相似性

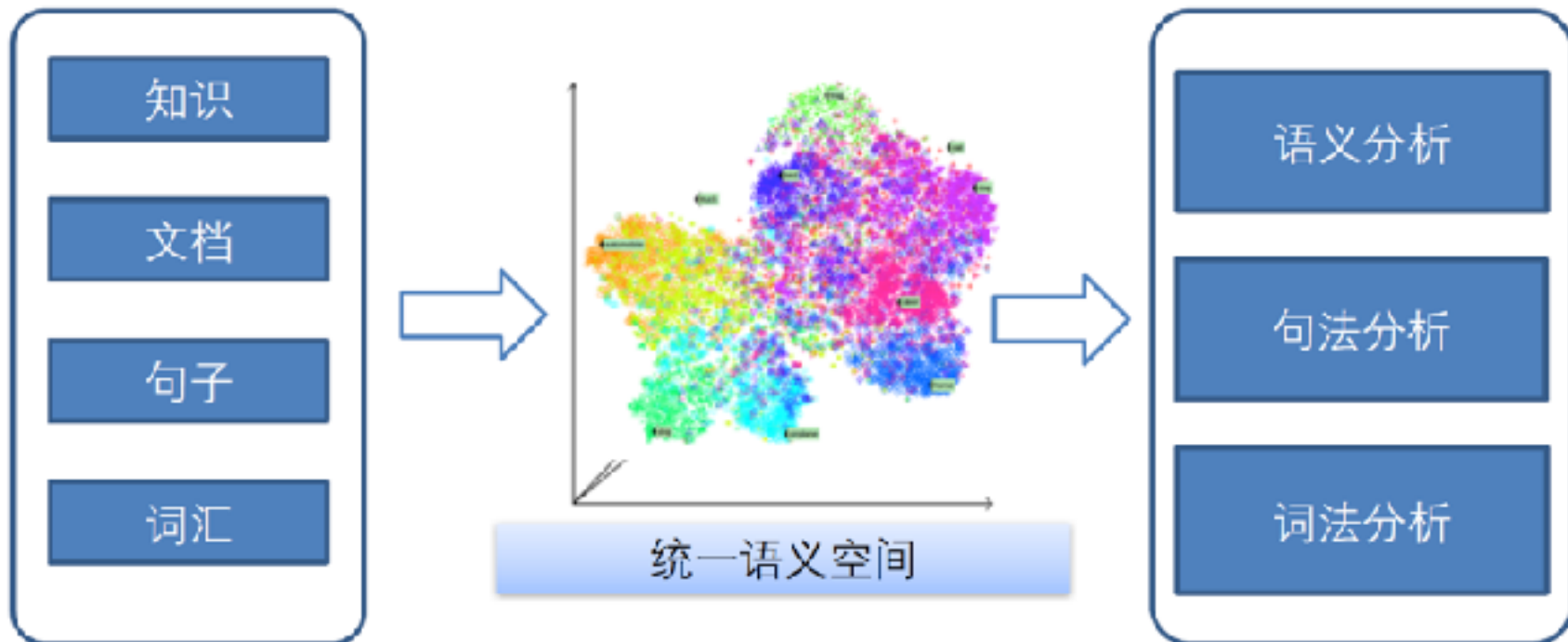
```
(EXIT to break): china
```

```
n vocabulary: 486
```

Word	Cosine distance
-----	-----
taiwan	0.768188
japan	0.652825
macau	0.614888
korea	0.614887
prc	0.613579
beijing	0.605946
taipei	0.592367
thailand	0.577905
cambodia	0.575681
singapore	0.569950
republic	0.567597
mongolia	0.554642
chinese	0.551576

# 分布式表示的优势

- 缓解NLP任务的数据稀疏问题
- 实现跨领域、跨对象的知识迁移
- 提供多任务学习的统一底层表示

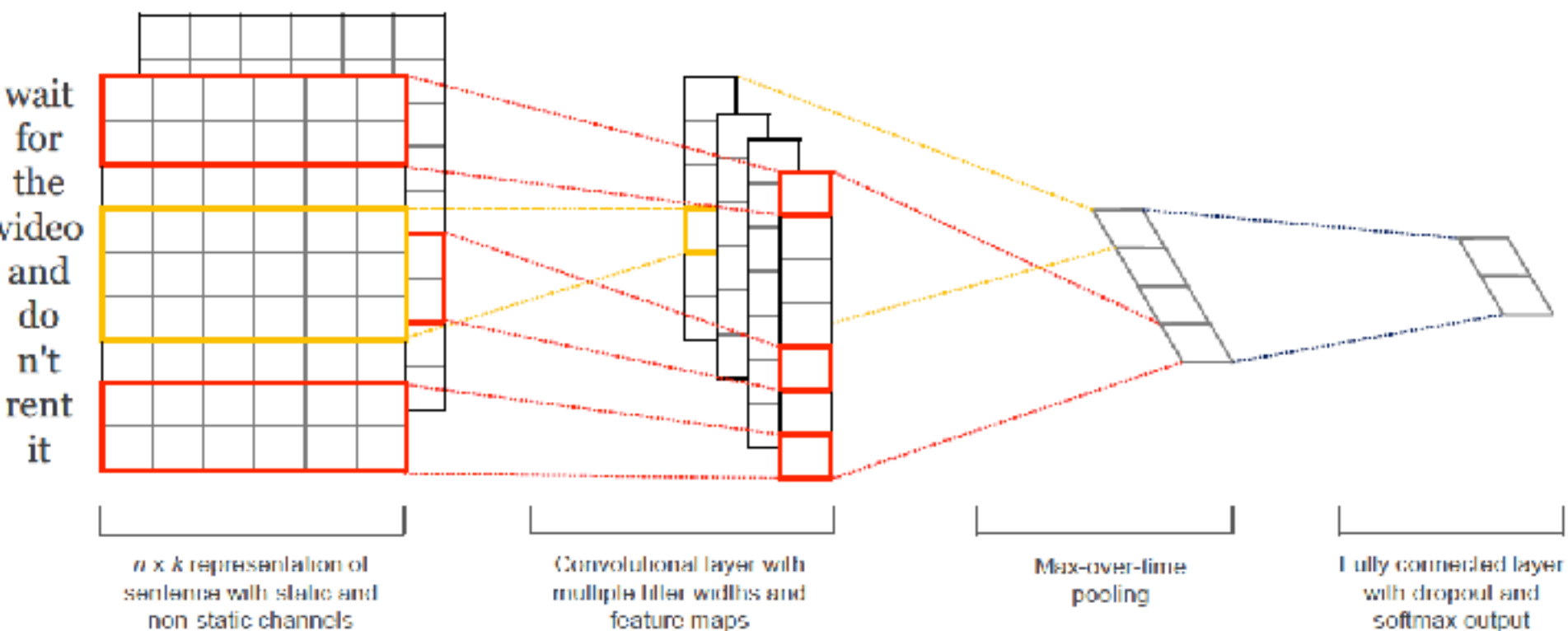




# 基于统计模型的NLU

- 将领域(domain) 和意图(intent)的识别认为是短文本分类问题
  - 收集每个领域和意图下的句式，训练分类器进行分类
  - 模型
    - CNN [Kim 2014]
    - LSTM [Shayne 2016]
    - FastText[Armand 2016]
    - CharAware[Zhang 2016]

# 基于CNN的文本分类模型



# 卷积神经网络

## ■ 卷积层

- 提取局部特征
- 组合性：词->词组->句子
- 类似于n-gram效果

## ■ Pooling层

- 提供固定长度的输出
- 降维，提取显示特征

# 基于统计模型的NLU

## ■ 将领域和意图识别认为匹配问题

- 通过计算句子和句式的语义相似度确定是否匹配
  - 查询句子进行词表检测，替换成标签，生成候选查询句子
    - 查询句子：我要听周杰伦的歌
    - 候选查询句子（1）：我要听<singer>的歌
    - 候选查询句子（2）：我要听周杰伦的歌
  - 候选查询句子和句式进行匹配，得到最佳的匹配
    - 想听首<singer>的音乐
- 语义相似度算法
  - 基于相似句子对训练的神经网络模型
    - [Aliaksei 2015]
    - [Wenpeng 2015]

# 基于统计模型的NLU

## ■ 语义相似度模型

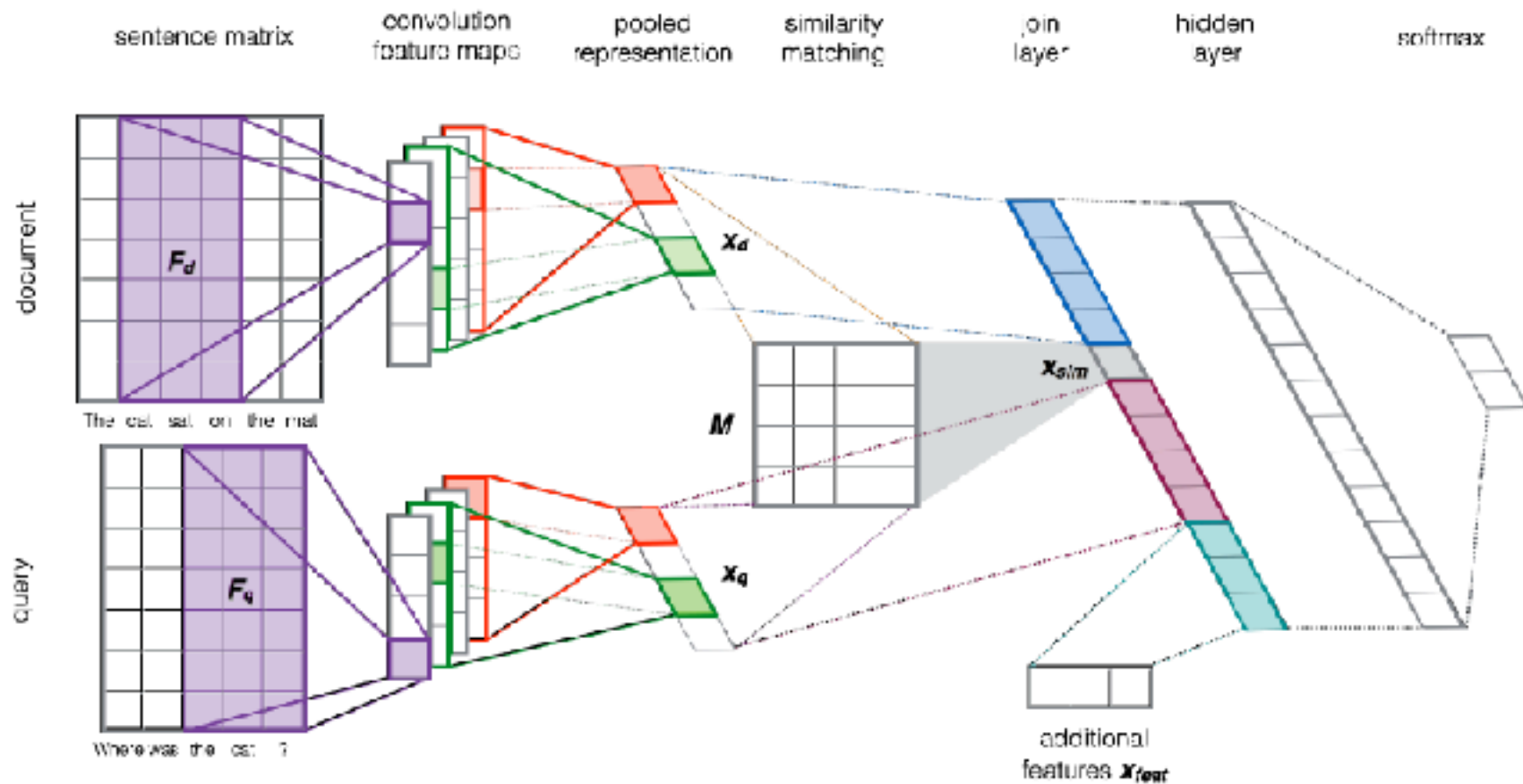


Figure 2: Our deep learning architecture for reranking short text pairs.

# 基于统计模型的NLU

## ■ 将槽抽取任务认为是序列标记问题

### ■ 对数据标注出要抽取的内容

- BMES、BIO

- 北 京 今 天 的 天 气

- B\_loc E\_loc B\_date E\_date S S S

### ■ 使用CRF或RNN等工具进行识别

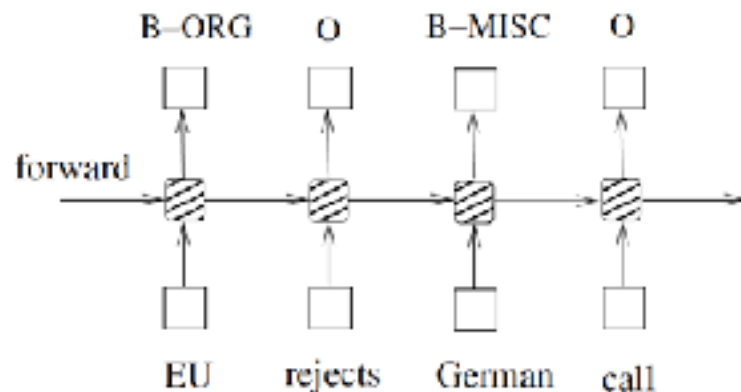
### ■ 模型

- CRF[Sutton 2011]

- LSTM [Kaisheng 2014]

- LSTM+CRF[Zhiheng 2015]

- Encoder LSTM [Gakuto 2016]

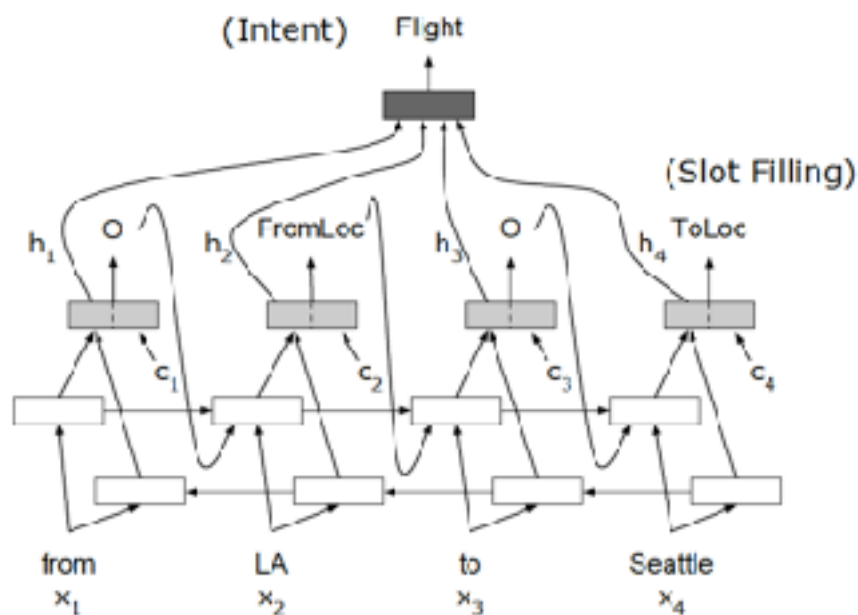


# 基于统计模型的NLU

## ■ 联合模型

- 分类和槽抽取共用同一个网络
- 训练时共同计算误差进行反传
- 模型

- [bing 2016]



# 参考文献

- [kim 2016] Convolutional Neural Networks for Sentence Classification
- [Shayne 2016] A WAY OUT OF THE ODYSSEY: ANALYZING AND COMBINING RECENT INSIGHTS FOR LSTMS
- [Armand 2016] Bag of Tricks for Efficient Text Classification
- [CharAware 2016] Character-level Convolutional Networks for Text Classification
- [Sutton 2011] An Introduction to Conditional Random Field
- [Kaisheng 2014] Spoken language understanding using long short-term memory neural networks
- [zhiheng 2015] Bidirectional LSTM-CRF Models for Sequence Labeling



# 参考文献

- [Gakuto 2016] Leveraging Sentence-level Information with Encoder LSTM for Semantic Slot Filling
- [Bing 2016] Attention-Based Recurrent Neural Network Models for Joint Intent Detection and Slot Filling
- [Aliaksei 2015] Learning to Rank Short Text Pairs with Convolutional Deep Neural Networks
- [Wenpeng 2015] ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs
- [Tomas Miklov 2013] Distributed Representations of Words and Phrases and their Compositionality