现代通信系统中的性能标准基于错误率，而不考虑通信消息的语义

1. 语义通讯
2. 建立一个贝叶斯博弈模型，刻画平均语义最小化的最优传输策略，降低端到端的平均语义错误，关注于提供辅助信息的代理（即影响力的实体之间）对解码器的影响。通过语义相似性量化了两个词语含义之间的距离，计算距离常常基于词库。
3. 首先。定制传输策略考虑实体意图的不确定性；将具有外部影响的语义通信问题视为代理之间的贝叶斯博弈。主要考虑到贝叶斯纳什均衡存在的纯策略、贝叶斯纳什均衡的混合策略；
4. 两个词语之间的相似性定义为两个词语可以采用的不同含义之间的所有相似度值的最大值。
6. 在ARC中的基于LSTM-RNN进行语义一致性验证
7. LSTM-RNN提取ARC语义
8. 用语义向量表示
9. 输出添加sigmoid进行验证

RNN：解决消失梯度爆炸和消失梯度问题；LSTM-RNN在解决消失梯度问题上更好，影响因素（上一层的输入，当前层的过去状态，LSTM-RNN的内部状态影响）

两个句子同时采用相同方式处理，输出向量为句子对的语义表示，采用余弦作为两个句子的语义相似性。

RNN：X作为输入向量，Y作为隐藏序列，隐藏层为循环层进行句子的上下文信息，将单词序列映射到低维语义向量。

LSTM-RNN：最重要的是细胞状态，sigmoid判断是否要传送信息进去，W代表及部分的输入联系和循环联系

output gate：会把前面的信息保存到隐层中去

forget gate：选择忘记过去某些信息

input gate：将新的信息选择性的记录到细胞状态中

memory cell：存储长时间存储的重要信息

先计算最小化交叉熵误差L()，是模型的参数，C是标记过的数字，R是余弦相似度，通过反向传播时间BPTT获得梯度的最小值来估计模型参数，采用梯度重整化方法解决梯度爆炸问题。将句子分割，分类，生成one-hot矢量（一位有效编码，现在多分类cnn网络的输出通常是softmax层，而它的输出是一个概率分布，从而要求输入的标签也以概率分布的形式出现，进而算交叉熵之类）输入模型。

构建两个LSTM-RNN，一个用于将查询映射到矢量，一个用来映射点击的文档。

LSTM-RNN：多用于序列任务和语言建模，e.g 语音识别和机器翻译。

二．语义网络Semantic Web

定义：extension of the current web in which information is given well-defined meaning，better enabling computers and people to work in cooperation

结构：7层结构

1. URI/IRI描述Web资源
2. RDF/XML：RDF标识web中的资源和关系；XML描述数据的内容和结构
3. RDFS/OWL：RSFS为RDF提供特定字典；OWL提供领域知识的语义
4. Logic：根据第二、三层进行计算判断
5. Proof：根据第四层表述进行评估、证明
6. Truest：提供用户之间的秘密关系

（7） Final：（关注于用户的应用）用户界面和应用程序表达特定的应用程序范围和不同的通信接口

检索出的很多论文都是关注于信息检索方向

三种描述方式进行对比分析，OWL的表现效果最好，相应的判断条件为

1. OWL：Web本体语言用以描述本体

RDF和RDFS的延伸，描述逻辑转化所有内容

结构（三层）：OWL Lite，OWL DL，OWL Full

1. RDF：资源描述框架：将信息表述为值和属性，用以描述本体和元数据
2. RDFS：资源描述框架语法：将信息归分为类，从而定义属性

在RDF中创建本体的标准语言：type谓词将资源分组在一起

三．语义框架

四．语义相似度：文本段落具有相同含义的程度。两个术语之间可以被赋予一个数字表示他们之间语义距离，即术语A和术语B如何相关

许多单词相似度的度量都是基于诸如WordNet之类的词库或者来自大型语料库的统计数据，而且依赖于基于本类的度量。

存在两种量化语义相似性的主要方法： 基于节点和基于边缘。

基于节点： c代表w的意义；基于节点同时考虑最低公共子系统的信息和内容，

基于边缘：利用分类法中两个节点的距离

混合方法：分类树中节点的各个方面，例如信息内容，深度，度分布和路径长度

判断语义相似性的相关技术：

1. 基于关系的页面排名算法：重点在于根据查询子图计算页面分数
2. 术语之间的无监督语义相似性：基于页面和基于上下文的度量
3. 运用相似度计算方法（Jaccard、Dice Coefficient）找到单词集的相似性
4. 共同信息量用来计算包含两个词语的文档数，根据文档数计算互信息量
5. 基于Google的语义相似度计算相似度，然后和单词之间的距离进行比较
6. 根据网络搜索结果测量单词组的相似度（分布相似度测量）：计算少量单词的搜索次数
7. 基于分布式和Word-Net：关注于不同语言之间的相似度Cross Linguality Similarity
8. 基于Web文档（基于Wikipedia返回的片段）

语义网络和2015年通过聚类的方法：（1）计算使用文本语料库（使用Wikipedia）的所有元素之间的语义相似度

（2）将具有高度相似度的元素结合在一起，每个类或者接口都映射为图形的节点，边权重表示相应节点的相似性；使用谱聚类的方式进行

1. 基于网络搜索：将页面计数和片段组合传递给SVM

（Web Jaccard，Web Ovarlap，Web Dice，Web PMI算法）

语义距离：根据语义距离来判断相似度，作为相似度的一种衡量标准

两篇偏向于算法的论文，一篇2012年的文章使用，一篇使用一种加强的算法

语义索引方案：（2014）

语义索引提供在噪声信道中提供源符号和码字之间的结构化映射；通信系统会根据源生成codewords，数量和输入字母表和distinct symbol数量一样的。

1. 源希望通过有限的信息传达一个语义信息到目的地
2. 每个word都通过双方已知的map进行索引
3. 源发送索引的二进制表示

多种语义相似性/距离的测量：

Semantic indexing of continuous speech