

实验 4 语篇分析

1. 实验目的

- 掌握语篇分析（词义消歧、语义角色标注、指代消解）的基本原理；
- 掌握常用语篇分析 NLP 工具包的使用；

2. 实验时间

2 学时。

3. 实验内容

1) 词义消歧实验

从下述三个歧义词中选择两个，使用任意方法或 NLP 工具，完成词义消歧。
每个义项给出 2 个测试用例。

（1）苹果：

水果。如，这个苹果真好吃，营养价值高。

电子品牌。如，苹果手机配置不高，但整体性能优化的很好。

电影名称。如，“苹果”上映了，听说非常精彩。

（2）算账：

计算账目。如，“你别打扰他，他正在算账呢。”

吃亏或失败后与人较量。如，“好了，这回算你赢，下回我再跟你算账！”

（4）水分：

物体体内所含的水。如，植物是靠它的根从土壤中吸收水分。

比喻某一情况中夹杂的不真版实的成分。如，他说的话里有很大的水分。

注：直接可用的词义消歧的工具包不多，可以使用基于词典的、基于词向量相似计算的、基于各类分类器的等思想编写语义消歧程序。

2) 语义角色标注实验

使用任意方法或 NLP 工具包，编写语义角色标注程序，并完成至少十句话的测试。

注：大部分 NLP 工具都支持语义角色标注。

3) 指代消解实验（选做）

汉语中存在多种指代现象，“代词->名词”是最基本的一种指代现象，请完成以下测试用例中“代词->名词”指代消解。

测试 1：代词：你，我，他，他们

输入：

小鸟在天上飞呀飞，小鸡在大树下叽叽叽。小鸟飞累了，停在树枝上。小鸡抬起头，“在天上飞好玩吗？”小鸡问，“当然啦。”小鸟回答。

小鸡听了，心里想：为什么小鸟会飞，我不会飞呢？我也有翅膀呀？于是，小鸡拍拍小翅膀，可他只扑腾起来一点儿就又落下来了。小鸡伤心地耷拉下脑袋。小鸟看到了，对小鸡说：“别急，别急，小伙伴。我会飞，让我带你到天上去看看吧。”

小鸟从树枝上飞到小鸡身边，小鸡高兴地一下跳到小鸟背上。小鸡呼啦啦地展开翅膀，他们就飞到天上去了。

小鸡在蓝天上看到像绵羊一样的一朵朵白云，还看到了很大很大的世界，他可真高兴呀。小鸟呢，他也很开心，因为他又交了一位新朋友。

输出：

小鸟在天上飞呀飞，小鸡在大树下叽叽叽。小鸟飞累了，停在树枝上。小鸡抬起头，“在天上飞好玩吗？”小鸡问，“当然啦。”小鸟回答。

小鸡听了，心里想：为什么小鸟会飞，我[小鸡]不会飞呢？我[小鸡]也有翅膀呀？于是，小鸡拍拍小翅膀，可他[小鸡]只扑腾起来一点儿就又落下来了。小鸡伤心地耷拉下脑袋。

小鸟看到了，对小鸡说：“别急，别急，小伙伴。我[小鸟]会飞，让我[小鸟]带你[小鸡]到天上去看看吧。”

小鸟从树枝上飞到小鸡身边，小鸡高兴地一下跳到小鸟背上。小鸡呼啦啦地展开翅膀，他们[小鸟, 小鸡]就飞到天上去了。

小鸡在蓝天上看到像绵羊一样的一朵朵白云，还看到了很大很大的世界，他[小鸡]可真高兴呀。小鸟呢，他[小鸟]也很开心，因为他[小鸟]又交了一位新朋友。

测试 2：代词：你，我，他

一只饥饿的狐狸正准备出门找午餐，突然，从外面传来一阵敲门声。狐狸打开门，门口竟然站着一只小肥猪。“哎呀，我找错门了，我以为是兔子的家呢！”

“没错，你来得正是时候！”狐狸把小猪狠狠地拽进屋里，大声叫道：“这真是我幸运的一天！我要吃一顿香喷喷的烤猪肉，现在，你就乖乖地给我躺到烤锅里去吧！”挣扎是没有用的，于是，小猪说：“好，那就听你的安排吧。不过，我很脏哦，在你吃掉我之前，就不想把我洗得干净一点吗？”于是，狐狸就忙起来了。他捡来了树枝、生起了火，再到河边去提来了水，然后，给小猪洗了一个痛痛快快的热水澡。狐狸说：“好啦，现在你是全村最干净的小猪了，快给我躺到烤锅里去吧！”挣扎是没有用的，于是，小猪说：“听你的安排吧。不过，我太小了，在你吃掉我之前，就不想喂饱我，让我长得胖一点吗？”于是，狐狸又忙起来了。他摘来了西红柿，做了通心粉，还烤了一炉金黄色的小甜饼。小猪大大咧咧地吃了一顿丰盛的午餐。狐狸说：“好啦，现在你是全村最肥的小猪了，快给我躺到烤锅里去吧！”挣扎是没有用的，于是，小猪说：“听你的安排吧。不过，我是一只勤劳的小猪，我的肉一定特别硬。在你吃掉我之前，就不想先给我按摩一下，让自己吃上更嫩一点的烤肉吗？”于是，狐狸又忙起来了。他让小猪坐在木椅上，为小猪捏捏头、松松肩、拉拉腿，小猪舒服得直哼哼。狐狸说：“好啦，现在你是全村皮最柔软的小猪了，快给我躺到烤锅里去吧！”

“那好吧，不过，这些日子我的确太累了。你看，我的背还有点僵硬呢，你能再用力替我按摩一下吗？对，就这样，再用点力气！”过了一会，身后传来狐狸的呼噜声，小猪轻声问：“狐狸先生，狐狸先生”狐狸没有听见小猪的喊声，他累极了，沉沉地睡去了。村子里最干净、最肥、最柔软的小猪，拿着剩下的甜饼，飞快地跑回家去了。

Stanford CoreNLP 3.6.0 套装中的中文指代消解模块

<https://www.cnblogs.com/zklidd/p/5081677.html>

4. 实验要求

- 1) 提交实验报告，给出详细实验过程和结果；
- 2) 提交源代码和可执行程序。