

中文信息学会句法分析评测介绍

周强

清华大学信息技术研究院

语音和语言技术中心

本次评测的总体目标，是充分利用现有树库描述资源，形成统一开放的训练测试集合；设计合适的评测任务，对汉语“字/词→块→句”不同层次的句法分析技术进行全面评测；通过大规模评测，开发针对不同任务的有效处理工具，为不同应用提供有效的不同层次的汉语文本分析工具。

为此，我们设计了三项测试任务：1) 汉语词类标注测试；2) 汉语基本块分析测试；3) 汉语句法分析测试。

词类标记测试的处理对象是经过正确词语切分的汉语句子，评测目标是准确识别句子中各个词语的词类标记。在这里，我们确定的基本原则是：1) 继承 CIPS 第一届分词和词类标注测试的有关评测指标和处理原则，保持相关评测工作的延续性；2) 加强对若干动词子类（名动词、补助动词、助动词）标记的识别，提高词类标记信息对自动句法分析的支持力度。

基本块分析的处理对象是经过正确词语切分和词性标注的汉语句子，总体目标是准确识别句子中由各个词语组合形成的表征完整概念描述单元的基本块。基本原则是：1) 加强对特殊子类动词（与上面的词类标注任务相配合）和其他特殊功能性动词组合（v+n, n+v+n 等）的局部组合能力分析，减少句子中自由使用的核心动词数目，降低自动句法分析难度；2) 加强对复合名词块（多个名词组合形成的复杂概念描述单元）内部关系分析能力的评测，提升现有基本块分析器对真实文本数据的信息提取能力。在这个阶段，我们设计了三个子任务：1) 新闻领域的基本块识别；2) 多领域文本的基本块识别；3) “v+n”组合的基本块识别。评测指标是基本块识别准确率(Precision)、召回率(Recall)和 F1-measure 值。

句法分析的处理对象是经过正确词语切分和词性标注的汉语句子，总体目标是准确识别句子中不同类型、不同层次的句法结构描述信息。为此，我们设计了以下三个子任务：1) 事件描述小句识别；2) 句法结构树分析；3) 事件骨架树分析。其中，事件描述小句识别用于界定句法分析处理对象，句法结构树和事件骨架树分析分别探索层次分析和中心成分分析哪种体系更适合于汉语分析。下面对有关内容简要说明如下：

事件描述句识别的设计目标，是确定句子各个事件描述小句的准确边界位置。其输入是经过切分和词性标注处理的完整汉语句子，以句号、问号和叹号等分隔；输出是句子中各个事件描述小句的左右边界位置，形成句子的事件描述句线性序列。评测指标：事件描述小句识别准确率(Precision)、召回率(Recall)和 F1-measure 值。

句法结构树识别的设计目标，是自动分析事件描述小句的句法结构树。其输入是经过切分和词性标注处理的事件描述小句；输出：小句的内部句法结构树，各个节点以 np, vp, ap 等句法标记标注。评测指标是 Parseval 的相关评测指标：标记准确率(Precision)、标记召回率(Recall)和 F1-measure 等。

事件骨架树识别的设计目标，是自动分析事件描述小句中不同层面的主要功能块（主、谓、宾、定、状、补等），形成分层次的事件骨架树，其中各个节点以 S, P, O 等功能标记标注。评测指标包括：1) 功能块层面评价：不同功能块的识别准确率(Precision)、召回率(Recall)和 F1-measure；2) 句式层面评价：不同功能块组合形成的完整句式结构的识别准确率(Precision)、召回率(Recall)和 F1-measure。其中的典型句式包括：SP, SPO, SDPO ...