

More Details about The Annotation System

Zhenghua Li, Yue Zhang
Soochow Univeristy, Suzhou, China
zhli13@suda.edu.cn

June 8, 2016

We build a browser-based annotation platform (tested with Chrome and Sougou). The web page is <http://hlt-service.suda.edu.cn/syn-dep-batch>. Please try.

The system randomly assigns a full-annotation or partial-annotation task to an annotator at a time, as shown in Figure 1. To highlight selected words to be annotated, the system marks the corresponding words with red circles. Annotators can only annotate head words for the selected words, and must complete all selected tasks before submission, as shown in Figure 2. To annotate a dependency, annotators just need to first click the head word and then click the modifier word, which will make the corresponding red circle disappear. Annotator may also cancel a dependency by right click on the modifier word.

We do not provide a strict annotation guideline, and do not require annotators to study or follow one either. However, a basic familiarity with dependency grammar is a must. Then, annotators can decide the dependency links according to their intuitive and personal feelings about the language syntax, especially when different choices can be made for the same structure (i.e., coordination structure, apposition structure).

The annotated results, before submission, should be a legal tree or a legal part of a tree: 1) singled-headed, 2) acyclic, 3) single-root.

ROOT 据 统计 , 在 民生 银行 的 各 项 贷 款 中 有 一半 是 投向 非国有 经济 部门 的 。

Figure 1: Before annotation: the words to be annotated are marked with red circles.

ROOT 据 统计 , 在 民生 银行 的 各 项 贷 款 中 有 一半 是 投向 非国有 经济 部门 的 。



Figure 2: After annotation: the words are annotated with their heads and red circles disappear.

Actually, annotators would not realize the distinction between full annotation and partial annotation. Their goal is to make all red circles disappear.

The annotation system automatically records the time of assigning the task and the time of submission, in order to get the annotation time of each task.

We have also uploaded all the annotation data. To get the statistic results in our paper, please run the following command.

```
perl postprocess-annotation-stat.pl -a annotation-data.txt -o result-stat.txt
```