Morphologically rich languages (MRLs) express multiple levels of information already

at the word level. The depdency annotation of such languages can be improved by using a graphical tree editor with support for tokenization.

This paper presents a graphical editor for dependency trees for morphologically rich languages. Dependency trees typically represent the syntactic structure of sentences. Nodes in the tree correspond to tokens in the sentence with the nodes following the same horizontal order as the original tokens. The dependency links between the nodes can be edited using ‘drag and drop,’ and the link labels and node types and attributes selected from pop-up menus. The tool also supports a type specification of the trees, defining the possible combinations of node and link types. The tool allows for tokenization of words to provide support for morphologically rich languages. A word can be broken up into its tokens by selecting the suggested options provided by the backend parser. Similarly, it also allows to join multiple tokens into one word. The tool is aimed for semi-automatic creation of treebanks of low-resource and morphologically rich languages. It aids the manual inspection and correction of automatically created parses, making the annotation process faster and less error-prone.

With varying degrees of morpoligoical complexities

The target of this project is to create a platiform-independent tool for syntactic annotation of . Many of the already exisitng tools are already designed

1. English
2. Offline
3. Users may not be easily able to correct earlier mistakes
4. Other tools
5. Many languages

To help computers be able to understand the grammatical strucutre of a snetence, a lot of examples of sentences with their analysis are needed. To create such resources, we need to interfacws that can help people create those reosurces, that can be fast and cheap. Such tools have been developed for a number of languages but have limitations that make them hard to use for complex lanauges such as Urdu and Arabic. This capstone project develops a new tool that works with number of languages and requires minimum resource from the user.

Dependancy Annotation Tool for Morpologically Rich Languages

Platform indepdent annotation tool for Morpologically Rich Languages