

Urdu Corpus Cleaning Application

June 23, 2012

Center for Language Engineering (CLE)
Al-Khwarizmi Institute of Computer Science
University of Engineering & Technology, Lahore



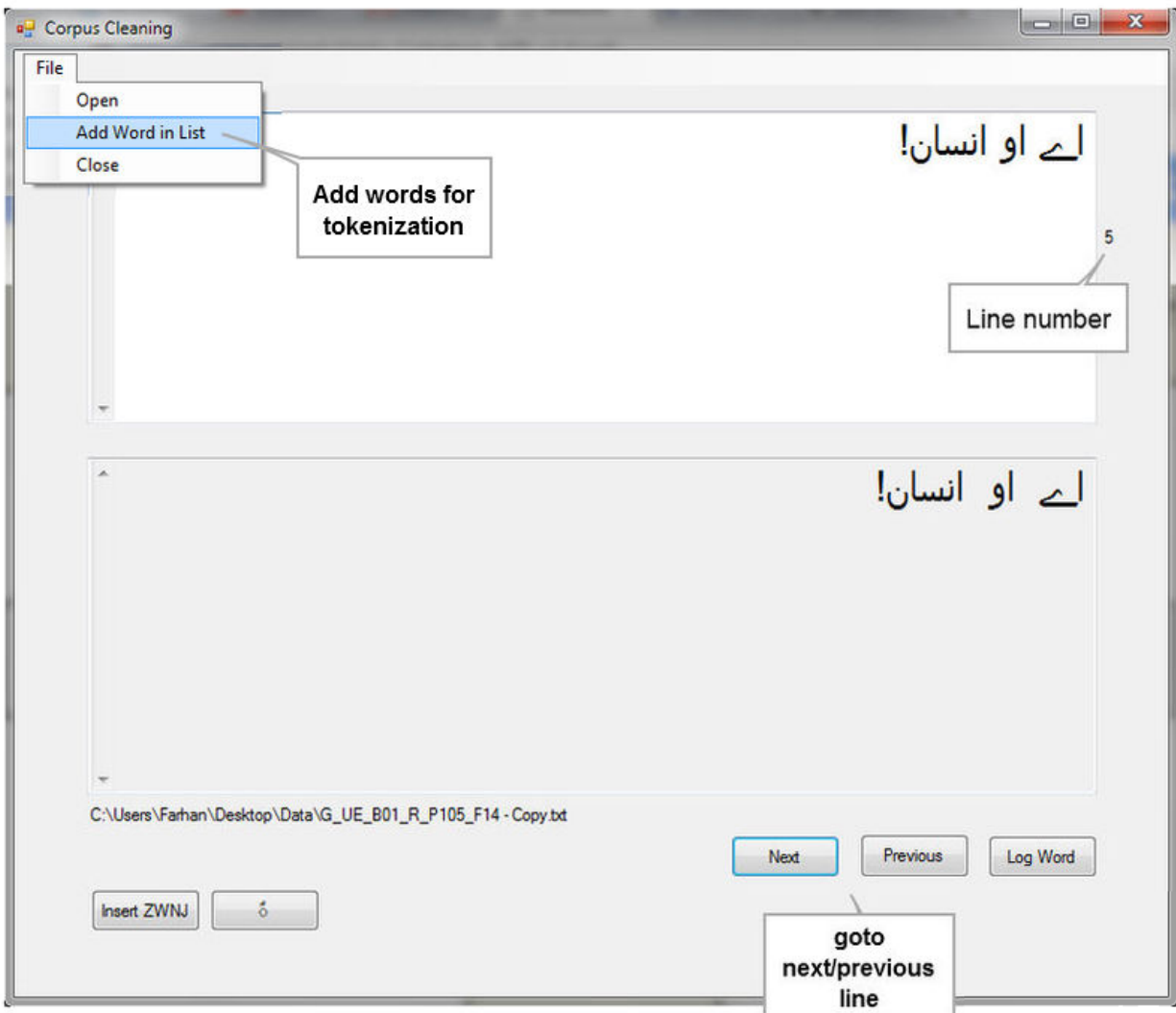
Table of Contents

1 Introduction.....	4
2 Application Usage	4

1 Introduction

Corpus Cleaning Application provides an interface for normalization and cleaning of Urdu Corpus. This document enlists Corpus Cleaning Application features and usage.

2 Application Usage



The application does normalization of selected file and creates two folders alongside the parent directory of the text file with name foldername_Cleaned and foldername_Log.