

An Introduction to the Sundance and AutoSlog Systems

Ellen Riloff and William Phillips

UUCS-04-015

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

November 8, 2004

Abstract

This document describes the Sundance natural language processing system that has been developed at the University of Utah, as well as the AutoSlog and AutoSlog-TS extraction pattern learners that have been implemented on top of Sundance. Sundance is a shallow parser that also includes clause handling capabilities and an information extraction engine. AutoSlog and AutoSlog-TS are learning algorithms that can be used to generate extraction patterns automatically. This report includes: (1) a high-level overview of Sundance's design and capabilities, (2) a detailed description of the steps that Sundance performs when analyzing sentences, (3) a discussion of the capabilities of the AutoSlog and AutoSlog-TS learning algorithms, and (4) guidance on how to install, use, and customize the Sundance and AutoSlog systems for new users and applications.