

Analysis of a Textual Entailer

Vasile Rus¹, Philip M. McCarthy², and Arthur C. Graesser²

¹ Department of Computer Science

² Department of Psychology

Institute for Intelligent Systems

The University of Memphis

Memphis, TN 38120, USA

{vrus, pmmccrth, a-graesser}@memphis.edu

Abstract. We present in this paper the structure of a textual entailer, offer a detailed view of lexical aspects of entailment and study the impact of syntactic information on the overall performance of the textual entailer. It is shown that lemmatization has a big impact on the lexical component of our approach and that syntax leads to accurate entailment decisions for a subset of the test data.

1 Introduction

The task of textual entailment is to decide whether a text fragment the size of a sentence, called the Text (T), can logically infer another text of same or smaller size, called the Hypothesis (H).

Entailment has received a great deal of attention since it was proposed (in 2004) under the Recognizing Textual Entailment (RTE) Challenge [7]. In our experiments presented here, we use the standard data set that RTE offers for development and comparison purposes.

The purpose of this paper is to perform an analysis of the textual entailer presented in [8]. In particular, we consider the three main subsystems of the entailer: the lexical component, the syntactic component and the negation handling component. We study each element's contribution to the performance of the system or a part of it. Different aspects of entailment have been analyzed by different groups. The Related Work section describes previous work on entailment analysis. Here, we analyze the task from a systemic, component angle. For instance, we report the impact of lemmatization for entailment, which, as far as we are aware, has yet to be reported. This type of analysis is important to better understand the interaction among different processing modules and to improve decisions as to whether the inclusion of a particular component is advantageous.

In our study, we conduct two levels of analysis. First, we look at how a particular feature impacts the component to which it belongs. For instance, lemmatization is part of the lexical component and we report how the lexical score changes depending upon its presence. Second, we present how a particular feature affects the overall entailment performance. The reader should note that our solution to entailment is based on a limited number of external resources and thus components such as world knowledge are not investigated: we use lexical