Project Proposal for Master's Thesis

Martin Pettersson (martinp4@kth.se)

March 5, 2015

1 Preliminary Thesis Title

"Resolving Difficult Pronouns Using Script Knowledge"
"Attempting The Winograd Schema Challenge Using Script Knowledge"

2 Background and Motivation

A proposed alternative to the famous Turing Test are resolution of so-called Winograd Schemas. The reason behind this is to try and invent a more reliable way to test human thinking behaviour in machines without using deception techniques, naive scripts or by cheating. The basic idea is to present a machine with questions that are easy to answer for English-speaking adults, but still so ambigous that a large corpus probably would not help much.

An example of a Winograd Schema is presented as follows:

The trophy would not fit in the brown suitcase because it was too big (small). What was too big (small)?

- The trophy

- The suitcase

Being part of a bilateral exchange program in Japan, I am currently sitting in a research lab at Tohoku University in Sendai that is specialized towards NLP. Researchers here have developed a system able to conduct large-scale acquisition of narrative event pairs from a huge corpus. At this moment they are using a kNN-based framework in order to discover potential impacts using script knowledge in coreference resolution. In the beginning of October this year they will participate in The Winograd Schema Challenge, an AI competition hosted by the non-profit organization Commonsense Reasoning.

With this thesis I aim to further explore these ideas by analyzing, evaluating and improving some parts of their current system.

2.1 Location

The project will be carried out at Inui-Okazaki Lab of Communication Science, Tohoku University Graduate School of Information Sciences, Sendai-shi, Miyagi, Japan.

3 Scientific Question

3.1 Question

3.2 Current Research

3.3 Methodology

3.4 Hypothesis

3.5 Evaluation

4 Personal Background

5 Supervisor at Tohoku University

Professor Kentaro Inui is the official supervisor and will examine the final project. I have been placed in a research group with people of varying background that I may turn to for daily practical issues and guidance. The purpose of the research group is to provide a continous feedback loop.

6 Available Resources

The laboratory at Tohokudai has access to several research paper databases that are relevant to Natural Language Processing and Machine Learning. I have also been given access to high-performance computing nodes at their department where experiments may be performed.