EXploring Customer Interactions through Textual EntailMENT

Ido Dagan (Bar Ilan University), Bernardo Magnini (FBK, Italy),
Guenter Neumann (DFKI, Germany), Sebastian Pado (University of Heidelberg)
dagan@cs.biu.ac.il, magnini@fbk.eu, neumann@dfki.de, pado@cl.uni-heidelberg.de
http://www.excitement-project.eu/

Description

EXCITEMENT is a 3-year research project (1/2012-12/2014) funded by the European Commission under FP7. The project consortium includes NICE Systems LTD (Israel) as coordinator, four academic partners, University of Bar Ilan (Israel), DFKI (Germany), FBK (Italy), University of Heidelberg (Germany), and two industrial partners, Almawave S.R.L. (Italy) and OMQ GmbH (Germany).

The main topic of the project is identifying semantic inferences between text units, a major language processing task, needed in practically all text understanding applications. While such inferences are broadly needed, there are currently no generic semantic engines or platforms for broad textual inference. The primary scientific motivation for the EXCITEMENT project is to change this ineffective state of affairs and to offer an encompassing open source platform for textual inference. On the industrial side, EXCITEMENT is focused on the text analytics market and follows the increasing demand for automatically analyzing customer interactions, which today cross multiple channels including speech, email, chat and social media.

There are two interleaved high-level goals for the project, which would yield two corresponding outcomes. The first is to set up, for the first time, a generic architecture and a comprehensive implementation for a multilingual textual inference platform and to make it available to the scientific and technological communities. To a large extent, the idea is to follow the successful experience of the Moses open source environment for machine translation, which has been making substantial impact on research in that field. This will enable developers of many text-processing applications to leverage the platform and boost their semantic inference capabilities. It will also provide developers of inference technology an effective environment for implementing and evaluating their components, and an easy entry-point for research in this field.

The second goal of the project is to develop a new generation of inference-based industrial text exploration applications for customer interactions, which will enable businesses to better analyze and make sense of their diverse and often unpredicted client content. These goals will be achieved for three languages, i.e. English, German and Italian, and for three customer interaction channels, i.e. speech (transcriptions), email and social media.

A major result of the project so far (i.e. month 18) is the first release of the *EXCITEMENT Open Platform (EOP-1.0.1)*. We started with three entailment systems developed by the project partners: BIUTEE from Bar Ilan, TIE from DFKI, and EDITS from FBK, which have been migrated in the EOP architecture. The platform aims to automatically check for the presence of entailment relations among texts. It is based on a modular architecture and provides support for the development of algorithms that are language independent to a high degree. Thus, it allows developers and users to combine linguistic pipelines, entailment algorithms and linguistic resources within and across languages with as little effort as possible. As an example, a classification-based entailment algorithm can use, both separately and in combination, the results of a distance component and the results of a BoW component, with the possibility to use lexical resources (e.g. wordnets) in different languages. The result is an ideal software environment for experimenting and testing innovative approaches for textual inferences.

The EOP is distributed as open source software (http://hltfbk.github.io/Excitement-Open-Platform/) and its use is open both to users who are interested in integrating inference technology in applications, and to developers who are willing to extend the current functionalities.