TEG - Open Platform for Recognizing Textual Entailment by Generality
Proposta de Projeto

Orientador: Sebastião Pais(sebastiao@di.ubi.pt)

Objetives

Textual Entailment by Generality TEG paradigm can be defined as the entailment from a specific sentence towards a more general one. For example, the pair $\langle S_1, S_2 \rangle$ naturally evidences that S_1 entails/implies S_2, and the latter is more general. Therefore, we have TEG from S_1 to S_2, represented as: $S_1 \xrightarrow{G} S_2$.

S_1: *Mexico City has a very bad pollution problem because the mountains around the city act as walls and block in dust and smog.*

S_2: *Poor air circulation out of the mountain-walled Mexico City aggravates pollution.*

TEG is based on a new *Informative Asymmetric Measure* (IAM), called the *Asymmetric InfoSimba Similarity* (AIS) measure. Instead of relying on the exact matches of words between texts, one sentence entails the other one by generality if two constraints hold: (a) if and only if both sentences share many related words and (b) if most of the words of a given sentence are more general than the words of the other one.

Thus, this project proposes to develop an *Open Platform for Recognizing Textual Entailment by Generality*.

Workplan

T1 Preliminary Investigation and Initial Requirements Specification;

T2 Preliminary Research on the various measures of similarity that will be used in this project;

T3 Research, Conceptualization and Experimental Development of the implementation of the AIS measure;

T4 Research, Conceptualization and Experimental Development;

T5 Integration, Testing and Evaluation;

T1 Report Writing.

Academic Prerequisites

Interest about Artificial Intelligence and Natural Language Processing; Web programming.

Assessment elements to deliver.

Source code and documentation of all code development; Project report.

Expected Results

* Open Web Platform;

* Project report.

Contacts

Sebastião Pais (sebastiao@di.ubi.pt)