Université de Lorraine - IDMC
Masters SC & TAL – M1
2019-2020

Fiche de projet tutoré / Project form

Having a share in SemEval-2020 Task 5: Detecting Counterfactuals

Encadrement / Supervisors

Sémagramme, Loria
Philippe de Groote Philippe.deGroote@loria.fr
Pierre Ludmann Pierre.Ludmann@loria.fr

Description / Description

1. projet global/global project

The projet is to enter the shared task 5 of SemEval-2020. SemEval is a series of shared tasks to evaluate semantic systems. Shared tasks rose in various research domains in order to compare running softwares from shared data, according shared metrics; a well-known example is the ImageNet Large Scale Visual Recognition Challenge where AlexNet-SuperVision overwhelming victory in 2012 established deep learning into computer vision. While these benchmarks mainly aim at statistical systems, we are willing to tackle them in a more formal and/or compositional fashion.
The SemEval-2020 Task 5 focuses on detecting counterfactuals. Counterfactual statements describe events that did not actually happen or cannot happen, as well as the possible consequence if the events have had happened [1]. For instance:

- Finance Minister Jose Antonio Meade noted that if a jump in tomato prices had been factored out, inflation would have begun to drop.
- Had Russia possessed such warships in 2008, boasted its naval chief, Admiral Vladimir Vysotsky, it would have won its war against Georgia in 40 minutes instead of 26 hours.

The Sémagramme team aims to process discourse dynamics — such as counterfactuals — with compositional tools *à la* Montague [2, 3].

2. biblio. UE 705 (semestre 7)

The bibliographic study shall present the adequacy between the task data and existing theories about counterfactuals. It shall also feature a draft of the system, explaining how to combine and/or modify existing tools.

3. réalisation. UE 805 (semestre 8)

To model counterfactual semantics and reasoning in natural language, that shared task aims to provide a benchmark for two basic problems:

1. Detecting counterfactual statements, *i.e.* you are asked to determine whether a given statement is counterfactual or not;
2. Detecting antecedent and consequence in counterfactuals.

Thus, the project workflow will be the following:

a) Explore the topic, the tools and the training data (see the section 2 “biblio” above);
b) Build a system tackling the two subtasks of the shared task;
c) Self-assess it (strength, weakness, versus baselines, scalability, extensibility, etc).

Informations diverses : matériel nécessaire, contexte de réalisation /
Various information: material, context of realization

Building a system (pre-)requires some ability in programming, more likely in Python and/or OCaml. It will also take place on INRIA GitLab. The shared task is hosted on the CodaLab platform that students will get to know during the project.
There will be 2 to 4 meetings a month at the Loria.

Livrables et échéancier / Deliverable and schedule

November 2019	Data exploration System draft
December 2019	Bibliographic study
March 2020	Running system
May 2020	Project report with self-assessment of the system

Bibliographie /References (max. 4-5)
[*il ne s’agit pas de la bibliographie complète qui sera fournie aux étudiants au début du projet mais d’une bibliographie indicative pour aider à cerner le sujet*]

1. Goodman, Nelson. "The problem of counterfactual conditionals." The Journal of Philosophy 44.5 (1947): 113-128.
2. Philippe de Groote. Towards a Montagovian Account of Dynamics. In Masayuki Gibson and Jonathan Howell, editors, 16th Semantics and Linguistic Theory conference - SALT2006, Tokyo, Japan, 2006
3. L. DANLOS, A. MASKHARASHVILI, S. POGODALLA, Interfacing Sentential and Discourse TAG-based Grammars, in : Proceedings of the 12th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+12), D. Chiang, A. Koller (editors), Düsseldorf, Germany, 2016, [hal:hal-01328697]