



The 12th International Conference on Language Resources and Evaluation

LREC 2020

Review Report

Title: Modeling Factual Claims with Semantic Frames

Authors: Fatma Arslan, Josue Caraballo, Damian Jimenez and Chengkai Li

Status: Accept Poster+DemoSuggested

Review #1

Detailed Comments

This paper presents a new annotation framework using the Semantic Frame for the annotation of factual claims. The paper is well written, clear, and concise. The research is novel, well motivated, and well communicated, but a substantial amount of space is spent on addressing the issue of misinformation. Given the audience of the paper, I doubt readers need to be convinced as to why this is a problem.

The annotation framework seems useful in and of itself, but the author bolsters its value by suggesting various implementations of the framework. Although these implementations may provide ease in, for example, fact checking, I'm not entirely convinced that semantic frames are significant more useful than other forms of NLP; however, this is an empirical question and is not within the scope of this paper. I would recommend this paper as an accept, though it may work better as a poster (where the audience can move back and forth between diagrams with ease) than an oral presentation.

One issue with the prose. There is a paragraph on page 5 that is very hard to read, and I suspect it may be a typo:

"In order to construct a sizable corpus of annotated sentences, we used fact-checked claims from "Share the Facts" due to those claims being annotated with the ClaimReview schema—A schema.org standardized format of fact-checks. The "Share the Facts" database contains fact-checks annotated with the ClaimReview schema—A schema.org standard to create a standardized format of fact-checks. "

Review #2

Detailed Comments

The paper describes an extension of FrameNet for modeling factual claims. For this, some new frames are added and more than 2500 sentences are fully annotated. The annotation tool is available for other contributors. The annotation task is well defined and the created resource as well as the tool are valuable on its own. The data might be used as training data for automatic frame annotation and more information about the opinions of the authors would be interesting.

In section 4 applications are discussed. Frame annotated sentences can be used for different kinds of data analysis and data transformation. From this point of view, the paper is an interesting contribution to text analysis using frames.

Review #3

Detailed Comments

The authors of the paper describe the work they are conducting on modeling factual claims by adopting frame semantics. The aims of this research are studying the dissemination of misinformation and supporting in fact-checking. The paper illustrates an extension made wrt Berkeley FrameNet lexical-semantic resource by introducing 11 new frames over the 20 specific ones the authors have identified so far. Finally, the authors have released an annotated resource made up of 2.569 sentences with 3.447 frame instances, by using an annotation tool they built within the project.

The research activities and the objectives of this paper are interesting and pertinent for the LREC conference, as the contribution addresses relevant topics for the CL and LR communities. I suggest that the authors should better explain how actually they grouped the claims to reach the 20 claim categories; maybe useful to have some examples where the authors state "group the claims sharing common syntactic and semantic patterns". Moreover, I am wondering if clustering approaches could be used to aid authors with this task.

As far as the annotation process is concerned, it is not clear to me how many annotators participated in the task, how the agreement was handled, and what kind of automatic support was embedded within the annotation tool. Furthermore, it is not evident, if the annotation tool is freely available in open source and where the codebase is hosted.

The paper is well written, but, if it will be accepted, some structural sections and the captions should be reshaped, indeed related works section is better next to the Introduction section, and captions are better arranged at the bottom. Moreover, within Section 3.1 there is a repeated sentence "ClainReview schema-A..."; Section 7 could be taken off.

Finally, the contribution has a few well described use-cases, but it would have been useful to deepen into one case study and evaluate it by

exploiting the provided resource.

[START](#) Conference Manager (V2.61.0 - Rev. 6047)