Notes Xphi Talk 13.10.20

Lucien Baumgartner

October 14, 2020

- We have to distinguish between:

  - sentiment
  - evaluativeness
  - polarity
  - thickness
  - valence
  - attitude

  What signature are we actually looking for?

- It was mentioned that we discredit a direct sentiment approach in the first part of the paper, but then use an operationalization based on sentiment for study 1 and 3. I think this is a misunderstanding of our claim. On the one hand, we do indeed claim that an uncritical one-off application of sentiment values is problematic, because the sentiment values are not always intuitive, e.g. *egoistic* being slightly positive. On the other hand, aggregated sentiment values are less problematic (given the variance is significant enough), since the aggregation mediates the effect of counterintuitive cases by taking their frequency into account. We should make this point more clear in the paper.

- There was some confusion about what the actual aim of the paper is, especially in developing our methodology. As I noted during the discussion, we are *not* providing a classification algorithm. We might want to stress the following: We start from a philosophical classification framework, which postulates the existence of thick concepts, thin concepts, value-associated concepts and purely descriptive concepts. Our aim is to look at empirical natural language data and try to find an indicator which allows us to show distribution patterns which correlate with the philosophical classification. Sentiment values are such an indicator, which we subsequently use to *measure* distribution patterns. That's all we do.

  In the future, we will strive to include all the indicators into an index which could be used for machine leaning-based natural language processing tasks such as unsupervised concept-classification. For this we will have to find a joint integration of the methodology developed in study 1 and 2, and find a way to overcome the cross-corpora comparison problem we have in study 3. In addition, we might want to test contextual enrichment measures, such as topic annotations or named entity tags (does the adjective describe things or persons), etc. In the long run, we should probably also dedicate some time to take a deep dive into supervised and unsupervised sentiment annotations, and compare it to our lexical approach. I think a lot of the concerns of people concerning syntactic and contextual valence effects could be incorporated in those more complex frameworks, while they cannot be accommodated in our lexical approach.

- Someone commented that we claim that study 3 supports the *variability thesis*, whereas our findings speak against said thesis. As far as I understood, the variability thesis states that

adjectives like *lewd* can change polarity, depending on the context. Empirically speaking, we should then find a more or less pronounced bipolar distribution for those cases. Kevin objected correctly, that we do not look at those objectionable cases (yet). We might still want to rephrase the argument and find an alternative to the *variability thesis*, since our findings clearly do not support a change in *polarity* across contexts, but rather a change in the *intensity* of the valence.

- There was a general worry that the valence and polarity of adjectives is highly dependent on context, as in 'I am curious whether he will give me a present.' compared to 'A curious thing happened to me yesterday.'. I think some of those cases could be caught by distinguish whether the adjective describes a person or a thing or an event. In a long term perspective, this is an argument for a ML-based approach. We probably still have to address the worry in the discussion.

- Another worry concerned study 3: legal or political language is codified, certain expressions or more generally the language itself might necessitate the use of a domain-specific dictionary. In the case of ML-based sentiment analysis we might be able to deal with different contexts by attributing inductively generated weights, such that we could continue to work with one sentiment dictionary. A purely lexical analysis runs into problems, however. I do think we will have to address this somehow for the legal TC paper. One way would be to use a standard language dictionary, such as sentiWS, and contrast it with a dictionary that is used specifically in more formal contexts (such a political dictionaries). Both dictionaries would be used to annotate the laypeople corpus and the legal corpus. We could then use an adapted Difference-in-differences (DID) design, to get the a sentiment differential for each of the corpora. For the laypeople the standard dictionary sentiment would be akin to the 'treatment group' values, whereas the formal language dictionary would serve as the 'control group' values. For the legal context, the groups would be inversed. The 'treatment-effect' would amount to the difference between the averages in the treatment group and the control group. The context-effect would then be the difference between the treatment effects in both corpora. That way we use both dictionaries for both corpora, which allows us to have comparable outcomes. And it would be step towards meeting the worry, without having to change to a ML-based framework (which we should do in the future, but neither for this paper, nor for the legal TC paper).