Contextualised Semantic Shift Detection

Francesco Periti[1]

[1]*Department of Computer Science, Università degli Studi di Milano, Via Celoria 18, 20133 Milan, Italy*

Abstract

Language continuously evolves influenced by social practices, events, and political circumstances; goal of Semantic Shift Detection is to detect, interpret, and assess changes of word meanings over time. The recent development of computational semantics pushed the emergence of approaches based on word embedding techniques for detecting semantic shift mainly at word-level. The Ph.D. research focuses on the problem of Contextualised Semantic Shift Detection (CSSDetection), which is the use of contextualised embeddings for capturing and interpreting "*semantic shift*" in the meaning(s) of words. In particular, the research aims to: (1) define a novel approach to trace the evolution of word meanings over time, and (2) extend CSSDetection to also capture and interpret semantic shift in the usage(s) of sentences.

Keywords

Computational Semantics, Contextualised Word Embeddings, Semantic Shift Detection

1. Introduction

Language continuously evolves influenced by social practices, events, and political circumstances. This means that studying how words and sentences (e.g., quotations, idioms) change in meaning/usage over time can help to deeper understand the evolution of the political and social landscape [1, 2]. The recent availability of large diachronic corpora and the development of computational semantics pushed the emergence of approaches based on Natural Language Processing (NLP) techniques for detecting *semantic shift* of word meanings. In particular, over the past three years, significant advancements in the field of Semantic Shift Detection (SSD) have been made almost exclusively based on contextualised word embedding models (e.g., BERT) [3].

Inspired by recent studies in literature, the Ph.D. research proposes to use contextualised word embeddings to capture how words and sentences *shift* semantically over time. While computational approaches have been already proposed for SSD, this research is the first, to our knowledge, that extends the notion of *semantic shift* from word-level to sentence-level by also aiming to detect, interpret, and assess the possible change in usage context of sentences.

The remainder of the paper is organised as follows. In Section 2, the research problem is presented and the PhD. research questions are outlined. In Section 3, an original analysis of the relevant literature is discussed. Finally, preliminary results, ongoing and future work are illustrated in Section 4.

SEBD 2023: 31st Symposium on Advanced Database System, July 02–05, 2023, Galzignano Terme, Padua, Italy
francesco.periti@unimi.it (F. Periti)
https://www.linkedin.com/in/francesco-periti/ (F. Periti)
0000-0001-8388-2317 (F. Periti)

2. Research problem

The research problem of the Ph.D. thesis consists in tracing both word-level and sentence-level semantic shift by leveraging contextualised embeddings. For clarification purposes, as a word-level example, consider the word `isolated`, which changed from its "*feeling detached*" connotation to "*quarantine*" during the course of the COVID-19 pandemic [4]. Similarly, a sentence-level example is the quotation to John 15:13, namely "*There is no other love rather than if someone gives soul for their friends*", uttered by the Russian President Vladimir Putin in a political and non-religious context while praising Russian military forces' actions in the war in Ukraine[1]. However, it's worth noting that sentence-level semantic shift refers not only to changes in the meaning of individual sentences, but also, in a broader sense, to changes in the discursive, cultural, and historical contexts in which the text is situated.

For the sake of readability, we formalise the problem of SSD at word-level. This simplification enables to review approaches to SSD in a clear and concise fashion, while being easily extendable to sentence-level SSD. Consider a diachronic document corpus

$$\mathcal{C} = \bigcup_{i=1}^{i=n} C_i \ \ ,$$

where C_i denotes a set of documents of the time t_i. Contextualised SSD (CSSDetection) consists in assessing the change of meaning for a set of target words $\mathcal{W}$ occurring in $\mathcal{C}$ across the whole time span $[1 \dots n]$ by leveraging contextualised embeddings [3]. For a word $w \in \mathcal{W}$ occurring in C_i, its contextualised word representation (i.e., embedding) in the k-th sentence is extracted by a contextualised language model and it is denoted by e^i_k. This means that the representation of w in the corpus is a set

$$\Phi^i_w = \{e^i_1, e^i_2, ..., e^i_k, ..., e^i_n\} \ \ .$$

Then, the semantic shift score of w between two sub-corpora C_1 and C_2 is assessed by using a distance function f between two sets Φ^1_w and Φ^2_w, with f defined as

$$f : \{\mathbb{R}^D\}^{m_1}, \{\mathbb{R}^D\}^{m_2} \to \mathbb{R} \ \ ;$$

where D is the dimension of the word vectors, and m_1 and m_2 are the frequency of w in C_1 and C_2, respectively [5].

2.1. Research questions

Specifically, the Ph.D. research aims to extend the state of the art by answering the following research questions:

RQ1: *How can the evolution of word meanings be traced over time and how can it be used to describe and categorise word-level semantic shift?* As a matter of fact, most of the existing solutions to CSSDetection focus on quantifying the degree of semantic shift for a target word. Although some approaches have been appearing to enable the interpretation of which word

[1] www.washingtonexaminer.com/news/putin-invokes-the-bible-to-justify-ukraine-invasion-at-moscow-rally

meaning(s) is lost, gained, or changed (e.g., broadening/narrowing, amelioration/pejoration), they are typically designed to analyse a corpus spanning two time periods. As a result, the evolution of word meanings over time (more than two time periods) cannot be easily traced. Thus, RQ1 aims to face this issue by promoting the design of i) a novel CSSDetection approach capable of tracing the word meaning evolution, and ii) novel analysis techniques to describe different categories of word-level semantic shift.

RQ2: *How can sentence-level semantic shift be captured, interpreted, and traced over time?* Although computational approaches have been already proposed for detecting semantic shift at word-level, we are not aware of approaches for detecting *semantic shift* at sentence-level. We argue that sentence-level SSD is meaningful for linguistics, social, and historical analysis, as it allows for a more comprehensive understanding of language evolution. For instance, sentence-level SSD can help identify changes in the use of quotations, idioms, collocations, and other multi-word expressions, which are not captured by word-level analysis alone [6]. Thus, RQ2 aims to address the research gap by extending RQ1 for studying sentence-level semantic shift.

3. Original analysis of the literature

We recently proposed a comprehensive classification framework for CSSDetection [3], which distinguishes approaches based on three dimensions of analysis: *meaning representation, time-awareness, and learning modality.*

Meaning representation concerns the representation of word meanings: i) *form-based* approaches focus on high-level properties of a word, such as its dominant meaning or its degree of polysemy; on the opposite ii) *sense-based* approaches focus on low-level properties of a word, i.e. its multiple and different meanings.

Time-awareness focuses on how the time information of the documents is considered in the embedding model: i) *time-oblivious* approaches do not consider the time at which a document is inserted in the corpus; on the opposite ii) *time-aware* approaches use a specific mechanism to encode the time information into the embeddings.

Learning modality is about the possible use of external knowledge for describing and learning the word meanings or the semantic shift to recognise: i) *supervised* approaches exploit external knowledge such as a dictionary or human-annotated dataset; on the opposite ii) *unsupervised* approaches derive word meanings/semantic shift from the text in the corpus by using unsupervised learning techniques.

3.1. Approaches to CSSDetection

Usually, CSSDetection approaches follow a three-step scheme: i) extraction of embeddings for each occurrence of a target word from a contextualised language model such as BERT, ELMo, or XLM-R; ii) an optional aggregation of the embeddings by averaging and/or clustering; and iii) the application of a semantic shift function like Cosine Distance or Jensen-Shannon Divergence.

For the sake of clarity, in the following, we present the main solutions according to the meaning representation of the considered target word, namely form- and sense- based approaches, respectively.

Form-based approaches. Word embeddings are optionally aggregated by averaging in a single representation, and used as input of a semantic shift function. On the one hand, form-based approaches that aggregate embeddings tend to detect the shift of the dominant meaning of the word w (e.g., [7]). On the other hand, form-based approaches that do not aggregate embeddings tend to detect the shift in the degree of the polysemy of the word (e.g., [8]).

Most form-based approaches follow the general scheme and are time-oblivious. A few time-aware approaches have been recently published and they are all characterised by the adoption of a specific fine-tuning operation to inject time information into the embedding model before assessing the semantic shift of a word (e.g., [9]).

All existing form-based approaches leverage unsupervised learning modalities. As an exception, a Word-in-Context model (WiC) is trained in [10] to reproduce the behavior of human annotators in the manual annotation task.

Sense-based approaches. In sense-based approaches, word embeddings are usually aggregated by clustering (e.g., [8]). All the documents of two time periods are considered as a whole, and a single clustering activity is performed, generating clusters with documents of different time periods. The idea is that each cluster denotes a specific word meaning that can be recognised in the considered documents. In this way, it is possible to assess and quantify the shift in meaning of a word by analysing the cluster membership of the documents.

All existing sense-based approaches are time-oblivious and most leverage unsupervised learning modalities. A number of unsupervised clustering algorithms (e.g., K-Means [8]) are proposed to sidestep the need of lexicographic resources.

Only a few approaches employ a lexicographic supervision. For instance, a supervised clustering is enforced in [11] by leveraging a reference dictionary (i.e., the Oxford English dictionary) to list the possible lexicographic meaning of a word beforehand; thus it is hardly applicable to low-resource languages.

Contrary to form-based approaches, sense-based approaches enable the semantic shift interpretation by performing an in-depth qualitative analysis of the resulting clusters. For instance, a cluster is often inspected by selecting the documents associated with the top closest vectors to its cluster centroid (e.g., [8]) or its most discriminating Tf-Idf keywords (e.g., [12]). However, when more than two time periods are considered, clusters of word meanings need to be re-calculated, meaning that scalability issues arise and that resulting clusters could dramatically change from one time period to the next. Thus, the possible evolution patterns of a meaning across different time periods cannot be captured. As a possible solution, some recent works propose to perform clustering separately for each time period. In this case, the resulting clusters need to be aligned in order to recognise similar word meanings in different, consecutive time periods (e.g., [12]). To avoid continuously re-calculating and aligning clusters, we propose to use an incremental clustering algorithm to trace the evolution of clusters/word meanings over time [13].

4. Preliminary results, ongoing and future work

In this section, we describe preliminary research results achieved so far, ongoing and future work for each research question formulated in Section 2.

RQ1: *How can the evolution of word meanings be traced over time and how can it be used to describe and categorise word-level semantic shift?*

We have recently proposed and evaluated a novel approach, called WiDiD, based on incremental clustering of contextualised embeddings [13]. WiDiD works under the assumption that the documents of the corpus $\mathcal{C}$ become available as a stream and they are segmented in a sequence of time periods. In this case, C_2 represents the set of documents collected at time t, while C_1 represents the cumulative set of documents collected in the $t - n$ time periods preceding t. At each time step t, a contextualised model is exploited to extract the corresponding word embeddings of the word w. Thus, at each step, two sets of embedding vectors are available: Φ^1_w , the set of embeddings produced in the previous iterations of the WiDiD approach over the corpus C_1; and Φ^2_w, produced at the current time t for the corpus C_2. In order to group word embeddings representing similar word meanings, a novel incremental clustering algorithm called A Posteriori affinity Propagation (APP) is adopted. Finally, a distance measure between the sets Φ^1_w and Φ^2_w is computed to quantify the semantic shift of the word w in the considered time interval.

Preliminary results of WiDiD have been evaluated and compared against a reference benchmark using multiple configurations characterised by different clustering algorithms and embedding methods [13]. In particular, our experiments include the use of a pre-trained BERT model and a trained Doc2Vec model, which has been adapted to provide *pseudo*-contextualised word embeddings. A subset of results of our evaluation is shown in Table 1; further results are discussed in [13]. All in all, WiDiD performs well in SSD compared to the approach based on the conventional Affinity Propagation clustering.

Corpus	Clustering	Model	JSD	PDIS	PDIV
SemEval Latin	*baseline* AP	trained Doc2Vec	**0.485***	0.229	-0.023
		pre-trained BERT	**0.394***	**0.347***	0.236
	WiDiD	trained Doc2Vec	**0.512***	0.337*	0.328*
		pre-trained BERT	**0.361***	0.210	0.036
SemEval English	*baseline* AP	trained Doc2Vec	**0.514***	0.139	0.134
		pre-trained BERT	0.356*	0.326*	**0.406***
	WiDiD	trained Doc2Vec	**0.333***	0.077	-0.078
		pre-trained BERT	0.302*	**0.512***	0.370*

Table 1

Spearman's correlation coefficients over different setups with Latin and English corpora. The asterisks denote statistically significant correlations ($p \leq 0.05$). We report in bold the highest scores for each approach considering BERT and Doc2Vec.

Ongoing and future work to address RQ1 are about the definition of cluster analysis techniques to describe and categorise patterns of semantic shift by considering the evolution of word meanings. For instance, we are currently working on defining a set of metrics to describe stable, growing, and shrinking trend in the dominance of a specific word meaning. In addition, since our WiDiD evaluation was executed on a benchmark corpus spanning two time periods, we are currently evaluating the WiDiD approach on a benchmark spanning more than two periods. Finally, we are currently working on a real-world application of WiDiD on a large corpus of Italian parliamentary speeches spanning 18 different time periods (i.e., 18 legislatures).

RQ2: *How can sentence-level semantic shift be captured, interpreted, and traced over time?*

To address this research question, we propose to extend the WiDiD approach, which actually enforces word-level SSD, to deal with the sentence-level SSD. In particular, as case study, we plan to focus on semantic shift of quotations, meaning that we would like to capture, interpret, and trace how the context of a quotation q change over time.

The Vatican corpus. As an extension of our previous work in [2], we plan to use a diachronic corpus of Vatican Publications as a case study for sentence-level SSD. This corpus contains plenty of quotations to the Bible and thus, it represents a perfect case study for the new SSD topic. Currently, the considered corpus of Vatican publications includes all the web-available documents from the digital Vatican archive[2] at the time of downloading. This corpus represents a valuable source for experimenting with SSD techniques for three reasons. Firstly, it is characterised by an exceptional historical depth. Secondly, the documents are available in various languages including Italian, Latin, English, Spanish, and German. And finally, the third reason is that, through the writings of its popes, the Catholic Church has always dealt with the most relevant issues in the public debate of its time, alongside themes of faith and worship. Thus, these writings constitute a primary historical source for understanding an important part of human cultural history, where the focus of public discourse shifted over time to different topics such as the environment, the role of science, and various historical events.

Ongoing and future work is about the definition of a novel framework to evaluate the WiDiD approach for sentence-level SSD. Inspired by the recent shared tasks for word-level SSD (e.g., SemEval-2020 Task 1 [14]), we are currently defining a manual annotation task to collect gold data for three distinct computational tasks:

- *Binary Change Detection:* classifying a quotation q as stable or changed in meaning/context between two time-specific corpora C_1 and C_2;
- *Graded Change Detection:* quantifying the extent to which the meaning/context of q shifts between C_1 and C_2;
- *Sentence Sense Disambiguation:* identifying the intended meaning/context of each occurrence of q from a synchronic perspective.

[2]https://www.vatican.va

Furthermore, we plan to organise a shared task (e.g., for the SemEval series[3]) with the aim of benchmarking and comparing different approaches in the field of sentence-level SSD.

References

[1] H. Azarbonyad, M. Dehghani, K. Beelen, A. Arkut, M. Marx, J. Kamps, Words Are Malleable: Computing Semantic Shifts in Political and Media Discourse, in: Proc. of CIKM, ACM, New York, NY, USA, 2017, pp. 1509–1518.

[2] S. Castano, A. Ferrara, S. Montanelli, F. Periti, Semantic Shift Detection in Vatican Publications: a Case Study from Leo XIII to Francis, in: Proc. of SEBD, CEUR-WS, Pisa, Italy, 2022, pp. 231–243.

[3] S. Montanelli, F. Periti, A survey on contextualised semantic shift detection, 2023. arXiv:2304.01666.

[4] K. Harrigian, M. Dredze, The Problem of Semantic Shift in Longitudinal Monitoring of Social Media: A Case Study on Mental Health During the COVID-19 Pandemic, in: Proc. of WebSci, ACM, New York, NY, USA, 2022, pp. 208–218.

[5] B. Wang, E. Di Buccio, M. Melucci, University of Padova@ DIACR-Ita, in: Proc. of EVALITA, CEUR-WS, Marrakech, Morocco, 2020, pp. 420–425.

[6] D. Walton, F. Macagno, Wrenching from Context: The Manipulation of Commitments, Argumentation 24 (2010) 283–317.

[7] M. Martinc, P. Kralj Novak, S. Pollak, Leveraging Contextual Embeddings for Detecting Diachronic Semantic Shift, in: Proc. of LREC, European Language Resources Association, Marseille, France, 2020, pp. 4811–4819.

[8] M. Giulianelli, M. Del Tredici, R. Fernández, Analysing Lexical Semantic Change with Contextualised Word Representations, in: Proc. of ACL, ACL, Online, 2020, pp. 3960–3973.

[9] V. Hofmann, J. Pierrehumbert, H. Schütze, Dynamic Contextualized Word Embeddings, in: Proc. of ACL, ACL, Online, 2021, pp. 6970–6984. doi:10.18653/v1/2021.acl-long.542.

[10] N. Arefyev, M. Fedoseev, V. Protastov, D. Homiskiy, A. Davletov, A. Panchenko, DeepMistake: Which Senses are Hard to Distinguish for a Word-in-Context Model, in: Proc. of Dialogue, (online), 2021.

[11] R. Hu, S. Li, S. Liang, Diachronic Sense Modeling with Deep Contextualized Word Embeddings: An Ecological View, in: Proc. of ACL, ACL, Florence, Italy, 2019, pp. 3899–3908. doi:10.18653/v1/P19-1379.

[12] S. Montariol, M. Martinc, L. Pivovarova, Scalable and Interpretable Semantic Change Detection, in: Proc. of NAACL-HLT, ACL, Online, 2021, pp. 4642–4652.

[13] F. Periti, A. Ferrara, S. Montanelli, M. Ruskov, What is Done is Done: an Incremental Approach to Semantic Shift Detection, in: Proc. of LChange, ACL, Dublin, Ireland, 2022, pp. 33–43.

[14] D. Schlechtweg, B. McGillivray, S. Hengchen, H. Dubossarsky, N. Tahmasebi, SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection, in: Proc. of SemEval, ICCL, Barcelona, 2020, pp. 1–23.

[3]https://semeval.github.io/