Chapter 8

Decomposing and Comparing Meaning Relations: Paraphrasing, Textual Entailment, Contradiction, and Specificity

Venelin Kovatchev[1,2], Darina Gold[3],
M. Antònia Martí[1,2], Maria Salamó[1,2], Torsten Zesch[3]
[1]Language and Computation Center, Universitat de Barcelona, Spain
[2]Institute of Complex Systems, Universitat de Barcelona, Spain
[3]Language Technology Lab, University of Duisburg-Essen, Germany

Accepted for publication at
Proceedings of the
Twelfth International Conference on Language Resources and Evaluation, 2020

Abstract In this paper, we present a methodology for decomposing and comparing multiple meaning relations (paraphrasing, textual entailment, contradiction, and specificity). The methodology includes SHARel - a new typology that consists of 26 linguistic and 8 reason-based categories. We use the typology to annotate a corpus of 520 sentence pairs in English and we demonstrate that unlike previous typologies, SHARel can be applied to all relations of interest with a high inter-annotator agreement. We analyze and compare the frequency and distribution of the linguistic and reason-based phenomena involved in paraphrasing, textual entailment, contradiction, and specificity. This comparison allows for a much more in-depth analysis of the workings of the individual relations and the

way they interact and compare with each other. We release all resources (typology, annotation guidelines, and annotated corpus) to the community.

8.1 Introduction

This paper proposes a new approach for the decomposition of textual meaning relations. Instead of focusing on a single meaning relation we demonstrate that Paraphrasing, Textual Entailment, Contradiction, and Specificity can all be decomposed to a set of simpler and easier-to-define linguistic and reason-based phenomena. The set of "atomic" phenomena is shared across all relations.

In this paper, we adopt the definitions of meaning relations used by Gold et al. [2019]. **Paraphrasing** is a symmetrical relation between two differently worded texts with approximately the same content (1a and 1b). **Textual Entailment** is a directional relation between two texts in which the information of the *Premise* (2a) entails the information of the *Hypothesis* (2b). **Contradiction** is a symmetrical relation between two texts that cannot be true at the same time (3a and 3b)[1]. **Specificity** is a directional relation between two texts in which one text is more precise (4a) and the other is more vague (4b).

1 **a)** *Education is equal for all children.*
b) *All children get the same education.*

2 **a)** *All children get the same education.*
b) *Education exists.*

3 **a)** *All children get the same education.*
b) *Some children get better education.*

4 **a)** *Girls do not get good education.*
b) *Some children do not get good education.*

The detection, extraction, and generation of pairs of texts with a particular meaning relation are popular and non-trivial tasks within Computational Linguistics (CL) and Natural Language Processing (NLP). Multiple datasets exist for each of these tasks [Dolan et al., 2004, Dagan et al., 2006, Agirre et al., 2012, Ganitkevitch et al., 2013, Bowman et al., 2015, Iyer et al., 2017, Lan et al., 2017, Kovatchev et al., 2018a]. These tasks are also related to the more general problem of Natural Language Understanding (NLU) and are part of the General Language Understanding Evaluation (GLUE) benchmark [Wang et al., 2018].

[1] In the Recognizing Textual Entailment (RTE) literature, contradiction is often understood as the lack of entailment. However we adopt a more strict definition of the phenomenon.

Recently, several researchers have argued that a single label such as "paraphrasing", "textual entailment", or "similarity" is not enough to characterize and understand the meaning relation [Sammons et al., 2010, Bhagat and Hovy, 2013, Vila et al., 2014, Cabrio and Magnini, 2014, Agirre et al., 2016, Benikova and Zesch, 2017, Kovatchev et al., 2018a]. These authors demonstrate that the different instances of meaning relations require different capabilities and linguistic knowledge. For example, the pairs 5 and 6 are both examples of a "paraphrasing" relation. However determining the relation in 5a–5b only requires lexical knowledge, while syntactic knowledge is also needed for correctly predicting the relation in 6a–6b. This distinction cannot be captured by a single "paraphrasing" label. The lack of distinction between such examples can be a problem in error analysis and in downstream applications.

5 **a)** *Education is equal for all children.*
b) *Education is equal for all kids.*

6 **a)** *All children receive the same education.*
b) *The same education is provided to all children.*

A richer set of labels is needed to better characterize the complexity of meaning relations. We believe that a typology of "paraphrasing", "textual entailment", and "semantic similarity" would capture the distinctions between the different instances of each relation. Kovatchev et al. [2019b] empirically demonstrate that in the case of Paraphrase Identification (PI), the different "paraphrase types" are processed in a different way by automated PI systems.

In this paper, we demonstrate that multiple meaning relations can be decomposed using a shared typology. This is the first step towards building a single framework for analyzing, comparing, and evaluating multiple meaning relations. Such a framework has not only theoretical importance, but also clear practical implications. Representing every meaning relation with the same set of linguistic and reason-based phenomena allows for a better understanding of the nature of the relations and facilitates the transfer of knowledge (resources, features, and systems) between them.

For the purpose of decomposing the meaning relations we propose **S**ingle **H**uman-Interpretable Typology for **A**nnotating Meaning **Rel**ations (SHARel). With the goal of showing the applicability of the new typology, we also perform an annotation experiment using the SHARel typology. We annotate a corpus of 520 text pairs in English, containing paraphrasing, textual entailment, contradiction, and textual specificity. The quality of the typology and of the annotation is evident from the high inter-annotator agreement.

Finally, we present a novel, quantitative comparison between the different meaning relations in terms of the types involved in each of them.