- Mary is writing the report on her Lenovo laptop.
- Mary has a Lenovo laptop.

  The two texts have a lot in common, but also have differences.

Example for Similarity 5:

- Mary was feeling blue.
- Mary was sad.

  The two texts are (almost) identical.

B.5 Specificity

Background: We want to research whether displaying more specific sentences is helpful in information retrieval or summarization tasks. Thus, you are asked to determine whether the 1st sentence is more specific than the 2nd. The specificity of sentence is defined as a measure of how broad or specific its information level is.

Task: In this task, you are presented with **two sentences**. You are required to decide whether **the 1st sentence IS more specific than the 2nd**. If this is not the case, choose the option **the 1st sentence IS NOT more specific than the 2nd**.

Examples for the option "Sentence 1 IS more specific"

- I like cats.
- I like animals.

  As the 1st sentence gives the more specific information on which animal is liked, it is more specific. Hence, you have to choose the option that the 1st sentence is more specific.

- The cute cafe has great coffee.
- The cafe sells coffee.

  As the 1st sentence gives the more specific information on both the cafe and the coffee, it is more specific. Hence, you have to choose the option that the 1st sentence is more specific.

Examples for the option "Sentence 1 IS NOT more specific"

- I like animals.
- I like cats.

  As the 2nd sentence gives the more specific information on which animal is liked, it is more specific. Hence, you have to choose the option that the 1st sentence is not more specific.

- I like dogs.
- I like cats.

  Now, as in both cases the liked animal is mentioned, they have the same level of specificity. Hence, you have to choose the option that the 1st sentence is not more specific.

- I like black dogs.
- He saw a blind cat.

  Now, as the information is very diverse, it is impossible to say which sentence is more specific. Hence, you have to choose the option that the 1st sentence is not more specific.

Appendix C

Annotation Guidelines for Kovatchev et al. [2020]

C.1 Presentation

This document sets out the guidelines for the annotation of atomic types using the Extended Typology for Relations. The task consists of annotating pairs of text that hold a textual semantic relation (paraphrasing, entailment, contradiction, similarity) with a textual label, and the atomic phenomena they contain. These guidelines have been used to annotate the ETRC corpus. For the purpose of the annotation, the WARP-Text annotation tool has been used.

N.B.: The task definition, tagset definition and annotation of linguistic phenomena in these Guidelines overlap with those for the ETPC corpus. The reader is encouraged to consult the ETPC guidelines presented in Appendix A or the full SHARel guidelines available online. Here I only provide the guidelines for the reason-based types.

C.2 Annotating reason-based Phenomena

Reason-based phenomena account for relations that cannot be expressed and processed using only linguistic knowledge. Like the linguistic phenomena, the reason-based phenomena can be sense-preserving or non-sense preserving. Our goals with the annotation of reason-based phenomena are twofold:

1) we want to make a precise and explicit annotation of the units involved in the inference

2) we want to determine the kind of reason-based and background knowledge required.