

Extracting Concept Inclusions with Typicality from Geometric Models

Camille Bourgaux and Ana Ozaki

June 25, 2024

Master Thesis Topic

In the field of knowledge representation, *concept inclusions* are a way of expressing which concepts are more specific or more general than other concepts. For example, the concept of a bird is more specific than the concept of an animal. Also, the concept of a bird is more specific than the concept of being able to fly. These can be represented in description logic as follows:

$$\text{Bird} \sqsubseteq \text{Animal} \quad \text{and} \quad \text{Bird} \sqsubseteq \text{CanFly}.$$

Since there are birds that cannot fly, e.g. penguins, a more accurate way of expressing the above would be to say that *typical* birds can fly. This is expressed using a typicality operator $\mathbf{T}(\cdot)$ [GGOP09], in our example,

$$\mathbf{T}(\text{Bird}) \sqsubseteq \text{CanFly}.$$

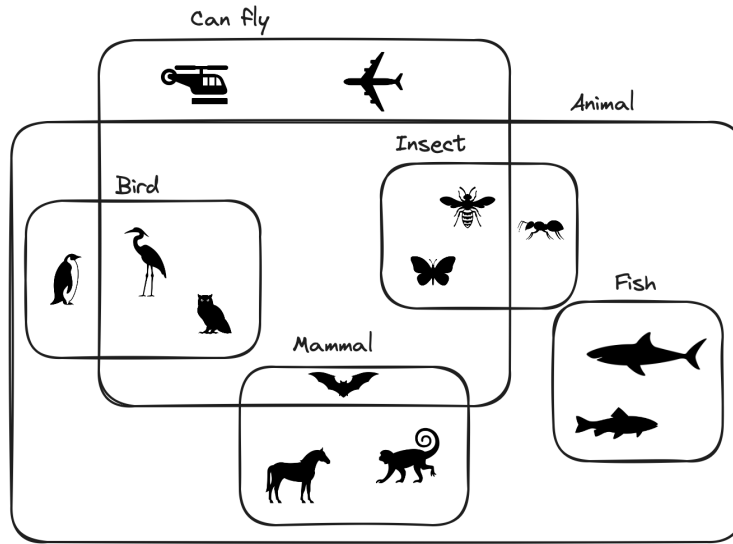


Figure 1: Geometric Model

Concept inclusions can be represented by *geometric models* [BGS22] where concepts are represented by geometric areas (such as boxes) in a vector space: $\text{Bird} \sqsubseteq \text{Animal}$ is represented by the geometric inclusion of the box of the concept Bird into the box of the concept Animal (Figure 1). *Concept assertions* are facts that provide information about specific individuals belonging to some concepts, for example,

$$\text{Bird}(\text{tweety}).$$

In geometric models, individuals are represented by vectors and the assertion $\text{Bird}(\text{tweety})$ is represented by the fact that the vector representing *tweety* is in the box representing the concept Bird. When only concept assertions are available, one can learn such geometric representation and use it to find plausible concept inclusions by inspecting how the shapes representing the concepts are related.

Main Tasks

In this project, the student will study concept inclusions with typicality and whether and how they can be extracted from geometric models. The intuition would be that if the central region of the area associated with a concept A is inside a concept B then this means the typical elements of A would be included in B . E.g., in Figure 1, one could extract that typical birds and typical insects can fly but typical mammals cannot. Also, that (every) fish cannot fly. The student's tasks are to

- read relevant articles about reasoning with typicality and about geometric models;
- create a synthetic benchmark with information about which individuals are typical (with respect to some concept) and concept assertions that reflect more or less faithfully an underlying set of concept inclusions with typicality;
- implement an embedding method that maps the data into vector spaces, creating a geometric model, while taking into account the information about which individuals are typical;
- evaluate the capability of the embedding method to recover the underlying set of concept inclusions with typicality.

References

- [BGS22] Zied Bouraoui, Víctor Gutiérrez-Basulto, and Steven Schockaert. Integrating ontologies and vector space embeddings using conceptual spaces (invited paper). In Camille Bourgaux, Ana Ozaki, and Rafael Peñaloza, editors, *International Research School in Artificial Intelligence in Bergen, AIB 2022, June 7-11, 2022, University of Bergen, Norway*, volume 99 of *OASICS*, pages 3:1–3:30. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
- [GGOP09] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Gian Luca Pozzato. Reasoning about typicality in ALC and EL. In Bernardo Cuenca Grau, Ian Horrocks, Boris Motik, and Ulrike Sattler, editors, *Proceedings of the 22nd International Workshop on Description Logics (DL 2009)*, Oxford, UK, July 27-30, 2009, volume 477 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2009.