Actively Learning Ontologies from LLMs: First Results

Matteo Magnini[*]   Ana Ozaki[†‡]
Riccardo Squarcialupi[*]

[*]Department of Computer Science and Engineering
Alma Mater Studiorum – University of Bologna
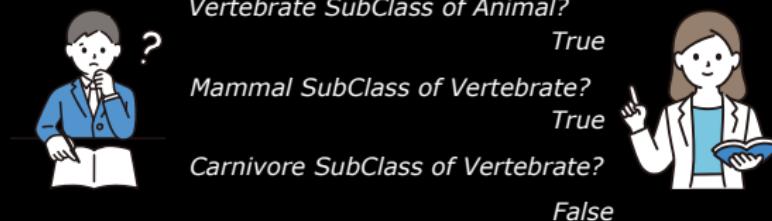*matteo.magnini*@unibo.it, riccard.squarcialupi@studio.unibo.it

[†]Department of Informatics – University of Oslo
anaoz@uio.no

[†]Department of Informatics – University of Bergen

International Workshop on Description Logics
June 20th, 2024, Bergen (Norway)

Context

- *Actively Learning* → a learner attempts to learn some kind of knowledge by posing questions to a teacher.

- *Ontologies* → we consider the case in which the knowledge is expressed as an $\mathcal{EL}$ ontology.

- *Membership queries* → the learner can ask the teacher whether a given axiom belongs to the target ontology.

Vertebrate SubClass of Animal?
                                    True
Mammal SubClass of Vertebrate?
                                    True
Carnivore SubClass of Vertebrate?
                            False

Vision

We want to use *large language models* (LLMs) as teachers. LLMs have been trained on large amounts of data from many domains, and we want to exploit this underlying knowledge.

Challenges:

- *input format*
  - → we used Manchester OWL syntax.
- *handling the response*
  - → mitigation strategies and parsing.
- *correctness and logical consistency of the answers*
  - → logical closure.

Experiments

We conducted experiments with:

- 5 $\mathcal{EL}$ ontologies → Animals, Cell, Football, Generation and University (from the ExactLearner work [Duarte et al., 2018]).

- 5 LLMs → GPT 3.5 Turbo, mistral 7b, mixtral 47b, llama 2 7b and llama 2 13b.

- 3 different experiments:
  - check that LLMs can answer axioms of the ontologies.
  - same check but on the entailed axioms (logical closure is finite!).
  - active learning with a trivial learning algorithm.

Results

What did we find?

- *LLMs can answer axioms of the ontologies* and the entailed axioms, but we measure some logical inconsistencies.

- *LLMs can be used for active learning*
  $\rightarrow$ there is statistical evidence that the answers
  of the LLMs (GPT 3.5 Turbo, Mistral and Mixtral)
  correlate with the knowledge in the ontologies.

Future work

We are currently working on an extension of the ExactLearner [Duarte et al., 2018] framework to use LLMs as teachers. Some spoilers:

- *system prompts* $\rightarrow$ provide contextual information to the LLM.

- *input format* $\rightarrow$ not only Manchester OWL syntax.

- *size matters* $\rightarrow$ we take into account more ontologies.

Actively Learning Ontologies from LLMs: First Results

Matteo Magnini[*]   Ana Ozaki[†‡]
Riccardo Squarcialupi[*]

[*]Department of Computer Science and Engineering
Alma Mater Studiorum – University of Bologna
matteo.magnini@unibo.it, riccard.squarcialupi@studio.unibo.it

[†]Department of Informatics – University of Oslo
anaoz@uio.no

[†]Department of Informatics – University of Bergen

International Workshop on Description Logics
June 20th, 2024, Bergen (Norway)

References I

[Duarte et al., 2018] Duarte, M. R. C., Konev, B., and Ozaki, A. (2018).
Exactlearner: A tool for exact learning of EL ontologies.
In Thielscher, M., Toni, F., and Wolter, F., editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, 30 October - 2 November 2018*, pages 409–414. AAAI Press

https://aaai.org/ocs/index.php/KR/KR18/paper/view/18006