

# Reasoning in Description Logics with Exceptions: Extended Abstract

Gabriele Sacco<sup>1,2</sup>

<sup>1</sup>Fondazione Bruno Kessler, Via Sommarive 18, 38123 Trento, Italy

<sup>2</sup>Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100, Bolzano, Italy

## Abstract

The problem of representing defeasible information is a long-standing topic of discussion in Knowledge Representation: for example, considering logic-based ontology representation languages, in Description Logics many proposals for defining defeasibility have been formalised, mostly emerging from existing approaches from the non-monotonic logic literature. On the other hand, little attention has been devoted to study the capability of these approaches in capturing the interpretation of exceptions from a formal ontological and cognitive point of view.

To address this, my proposal is to consider how this topic has been discussed in the fields of philosophy and cognitive science and to extract theoretical desiderata that such formal systems should satisfy. Then, according to these desiderata, I plan to develop a formal model in the Description Logics framework and implement it in an automated reasoning system. Finally, I will evaluate this system by comparing the inference results with those of actual human reasoners.

## 1. Motivation

Representing and reasoning with *defeasible* information is a long-standing topic of discussion in Artificial Intelligence (AI), dating back to the origins of the field of Knowledge Representation (KR): in presence of stronger conflicting information (or exceptions) with such defeasible information, one wants to retract what we would have inferred in view of new information. In its formalisation in different *non-monotonic logics* [2], this form of reasoning has been considered since the earliest days of KR as one of the parts of the common-sense that artificial systems should have to be considered actually intelligent [3, 4]. The classical example in the non-monotonic logics literature is the *Penguin example* (see, e.g., [2]): if we know that Tweety is a bird and we also know that birds fly, then we are willing to infer that Tweety flies. However, if we come to know that Tweety is in fact a penguin, this makes us retract the previous conclusion: we are more inclined to say that Tweety does not fly instead.

Considering logic-based ontology representation languages, in Description Logics (DLs) many proposals for defining defeasibility have been formalised: as a matter of fact, most of them emerge from existing approaches in non-monotonic logics [5, 6]. On the other hand, little attention has been devoted to study the capability of these approaches in capturing the

---


FOIS 2023 Early Career Symposium (ECS), held at FOIS 2023, co-located with 9th Joint Ontology Workshops (JOWO 2023), 19-20 July, 2023, Sherbrooke, Québec, Canada

✉ gsacco@fbk.eu (G. Sacco)

🆔 0000-0001-5613-5068 (G. Sacco)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

interpretation of exceptions from the point of view of formal ontology and cognitive aspects [1]. Thus, often a lack of discussion about the philosophical and cognitive assumptions of this kind of reasoning is noted. Namely, when looking at non-monotonic logic only as a tool, evaluating the systems with the sole criterion of evaluation to check the functionality of the particular formal system proposed w.r.t. a particular reasoning problem, we end up with a fragmented set of approaches. This, clearly, increases the difficulty of comparing and therefore properly evaluating comparatively such systems from a more general point of view. Moreover, since in the end these tools should be used to model knowledge and reason in real-world scenarios, we also need criteria that allow us to decide if the ontological and cognitive assumptions that the formal systems imply are justified or not. For these reasons, I am interested in discussing these foundational aspects of defeasible reasoning in DLs with the goal of developing a DL system based on an ontologically and cognitively well-justified foundation.

## **2. Research Questions**

- (Q1): What are the characteristics of non-monotonic reasoning from the philosophical and cognitive points of view?
- (Q2): How can we model non-monotonic reasoning in Description Logics in order to capture the features we discovered in the philosophical and cognitive analysis?
- (Q3): How can we implement the formal model developed with respect to Description Logics in an automated reasoning system?
- (Q4): How can our formal reasoning about exceptions be evaluated with respect to psychological and/or common-sense reasoning results coming from a study with human reasoners?

## **3. Objective(s)**

The first objective to be obtained at the end of the project is to have a set of theoretical properties extracted from the related literature in philosophy and cognitive science that will be used as a theoretical benchmark to compare the formal approaches for non-monotonic reasoning in DLs. The second goal is to develop a formal model in Description Logics that satisfies the theoretical features extracted. Thirdly, I would like to implement a system for automated reasoning for the proposed non-monotonic Description Logics extension. Finally, the last objective is to have an evaluation, in the form of a user study, assessing the compliance to the desiderata and human reasoning.

## **4. Research Methodology**

In general, given the strong theoretical characterisation of the research, the main methods to answer the questions would be literature review and developing formal models with DLs.

In particular, for (Q1) I am studying sources from the fields of philosophy, cognitive sciences and theoretical computer science, in order to discuss the analysis of defeasible reasoning in these fields and extract the theoretical desiderata that a KR formal model should satisfy.

More specifically, I discussed the main literature on generics (e.g. [8, 7, 9]), that is sentences that express generalisations that admit exceptions. They are strictly related to defeasible reasoning [9] and from their analysis I extracted some possible desiderata to discuss further with respect to other fields too. Now I am studying the literature in the psychology of reasoning to research what considerations have been made on this kind of reasoning. In particular, I am exploring the results of some experiments aimed at comparing logical theories of defeasible reasoning and human reasoning [12, 11, 10].

In this phase, it will be important to compare the results with solutions present in the literature on DLs, in order to evaluate the validity and the usefulness of the theoretical grounding attempt.

For (Q2), I will proceed with the modelling in the formalism of DLs based on the comparison conducted previously in the answer to (Q1). Moreover, the resulting model could be applied to a specific ontological theory to understand better possible shortcomings or flaws. In this case, mereology could be a good candidate given the lively interest in the topic by research both in philosophy and in AI.

(Q3) relies heavily on the answer to (Q2). In fact, my plan is to use the formalisation in DLs and to develop an implementation of reasoning procedures in Answer Set Programming. This will be addressed by comparing the techniques used in automated reasoning in order to elaborate the most fitting one for my problem.

Finally, (Q4) aims at assessing our work, by evaluating the results obtained in the answers to the previous three questions with respect to cognitive results. The way to answer this question will depend on the actual results obtained. However, a criterion that could be tested is the generality of exception types dealt with by the automated reasoning tool and by humans.

## References

- [1] Khemlani, S. and Johnson-Laird, P.N. (2019), Why Machines Don't (yet) Reason Like People, in *Künstl Intell* 33, 219-228, <https://doi.org/10.1007/s13218-019-00599-w>.
- [2] Strasser, C. and Antonelli, G. A. (2019), Non-monotonic Logic, The Stanford Encyclopedia of Philosophy (Summer 2019 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2019/entries/logic-nonmonotonic/>.
- [3] McCarthy, J. (1959), Programs with common sense, in: Proceedings of the Teddington Conference on the Mechanization of Thought Processes. London: Her Majesty's Stationary Office, 75-91.
- [4] McCarthy, J. and Hayes, P. J. (1969), Some philosophical problems from the standpoint of artificial intelligence, in *Machine Intelligence* 4, B. Meltzer and D. Michie, eds., Edinburgh: Edinburgh University Press, 463-502.
- [5] Giordano, L., Gliozzi, V., Lieto, A., Olivetti, N., and Pozzato, G. L. (2020), Reasoning about Typicality and Probabilities in Preferential Description Logics, *arXiv e-prints*, <https://doi.org/10.48550/arXiv.2004.09507>.
- [6] Britz, K., Heidema, J., Meyer, T. (2009), Modelling Object Typicality in Description Logics,

- in: Nicholson, A., Li, X. (eds), *AI 2009: Advances in Artificial Intelligence*, Berlin, Heidelberg: Springer Berlin Heidelberg, 506-516, [https://doi.org/10.1007/978-3-642-10439-8\\_51](https://doi.org/10.1007/978-3-642-10439-8_51).
- [7] Leslie, S. J. (2008), Generics: Cognition and acquisition, in: *Philosophical Review* 117.1, 1-47.
- [8] Leslie, S. J. and Lerner, A. (2022), Generic Generalizations, The Stanford Encyclopedia of Philosophy (Fall 2022 Edition), Edward N. Zalta Uri Nodelman (eds.), <https://plato.stanford.edu/archives/fall2022/entries/generics/>.
- [9] Pelletier, F. J. and Asher, N. (1997), Generics and defaults, in: *Handbook of logic and language*, 1125-1177, North-Holland.
- [10] Ragni, M., Eichhorn, C., Bock, T., Kern-Isberner, G. and Tse, A. P. P. (2017), Formal Nonmonotonic Theories and Properties of Human Defeasible Reasoning, in: *Minds Machines* 27, 79-117, <https://doi.org/10.1007/s11023-016-9414-1>.
- [11] Ragni, M., Eichhorn, C., and Kern-Isberner, G. (2016), Simulating Human Inferences in the Light of New Information: A Formal Analysis, in: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI'16)*, New York, USA: AAAI Press, 2604-2610.
- [12] Kuhnmuench, G., and Ragni, M. (2014), Can Formal Non-monotonic Systems Properly Describe Human Reasoning?, in: *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36 <https://escholarship.org/uc/item/921558fg>.