



HAL
open science

Editorial of the special issue “Synergies Between Machine Learning and Reasoning”

Sébastien Destercke, Jérôme Mengin, Henri Prade

► **To cite this version:**

Sébastien Destercke, Jérôme Mengin, Henri Prade. Editorial of the special issue “Synergies Between Machine Learning and Reasoning”. *International Journal of Approximate Reasoning*, 2024, 171, pp.109207. 10.1016/j.ijar.2024.109207 . hal-04917004

HAL Id: hal-04917004

<https://hal.science/hal-04917004v1>

Submitted on 28 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Editorial of the special issue
“Synergies
Between Machine Learning and Reasoning”

Sébastien Destercke¹ and Jérôme Mengin² and Henri Prade²

1. Université de technologie de Compiègne, CNRS, Heudiasyc, France
sebastien.destercke@hds.utc.fr
2. IRIT, CNRS & Université Paul Sabatier, Toulouse, France
{Jerome.Mengin, Henri.Prade}@irit.fr

January 28, 2025

This special issue results from the work of a working group. This group was dedicated to an inventory and a study of the meeting points between Knowledge Representation and Reasoning (KRR) and Machine Learning (ML). The group was created in the Fall of 2018 under the auspices of the French CNRS GDR Aspects Formels et Algorithmiques de l’Intelligence Artificielle (Formal and Algorithmic aspects of AI). The group¹ brought together about ten people at each meeting, with a stable core of participants. The work of the group results in a series of meetings interrupted by the CoViD-19 pandemic for two years ².

After a year, the group was able to deliver a first report providing a struc-

¹The following people took part in the working group, some for short periods: Ismaïl Baaj (CRIL, Lens, Fr), Zied Bouraoui (CRIL, Lens, Fr), Antoine Cornuéjols (AgroParis-Tech, Paris, Fr), Thierry Denœux (Heudiasyc, Compiègne, Fr), Sébastien Destercke (Heudiasyc, Compiègne, Fr), Didier Dubois (IRIT, Toulouse, Fr), Hélène Fargier (IRIT, Toulouse, Fr), Sabine Frittella (LIFO, Bourges, Fr), Romain Guillaume (IRIT, Toulouse, Fr), Fred Koriche (CRIL, Lens, Fr), Marie-Jeanne Lesot, João Marques-Silva (IRIT, Toulouse, Fr), Jérôme Mengin (IRIT, Toulouse, Fr), Henri Prade (IRIT, Toulouse, Fr), Steven Schockaert (School of Computer Science and Informatics, Cardiff, UK), Mathieu Serrurier (IRIT, Toulouse, Fr), Olivier Strauss (LIRMM, Montpellier, Fr.), Christel Vrain (LIFO, Orléans, Fr). The group was coordinated by Sébastien Destercke and Henri Prade, and then by Jérôme Mengin and Henri Prade from 2021.

²Toulouse, Nov. 15-16, 2018 ; Paris, Mar. 28-29, 2019 ; Orléans, May 27, 2019 ; Paris, March 18-20, 2020 (cancelled due to CoViD) ; Toulouse, Dec. 13-14, 2021 ; Arras, Mar. 14-15, 2022 ; Toulouse, Nov. 22-23, 2022.

tured introduction to common concerns, paradigms that are at the border between KRR and ML, and approaches hybridizing ideas coming from both sides [2], and to prepare an invited keynote at the 13th International Conference on Scalable Uncertainty Management [1], under the pen name Kay R. Amel, in December 2019. The plan of the group was then to continue to work for preparing a more developed synthesis on the topic, but the CoVid-19 pandemic stopped the process. In the meantime, in January 2021, we received an invitation from the editor-in-chief of the International Journal of Approximate Reasoning for preparing a special issue on the topic studied by the working group. After some discussions, we choose to both continue our work in progress towards an overview of the synergies between KRR and ML, and in the spring of 2022, to encourage submissions of specific contributions by members of the group, but also by representative researchers in the field. These contributions were received during the fall of 2022 and entered a thorough review process, with external reviewers and required revisions, during the following year. We are now in position to present the result of this venture.

The special issue starts with an overview paper “Synergies between machine learning and reasoning – An introduction by the Kay R. Amel group”(I. Baaaj, Z. Bouraoui, A. Cornujols, Th. Dencœux, S. Destercke, D. Dubois, M.-J. Lesot, J. Marques-Silva, J. Mengin, H. Prade, S. Schockaert, M. Serrurier, O. Strauss, Ch. Vrain), which is a fully revised and considerably expanded version of the previous 2019 report. This paper provides an encompassing picture of the topics covered by this special issue.

The special issue is rich in 19 other contributions, whose diversity testifies to the variety of issues raised by the confrontation of KRR and ML concerns.

First, in “On the failings of Shapley Values for explainability”, Xuanxiang Huang and J. Marques-Silva prove that the existing definitions of Shapley scores, often used in XAI methods, may yield misleading information about the relative importance of features for predictions and the authors offer empirical evidence that such theoretical limitations of these scores are routinely observed in ML classifiers.

The following three articles then fall within the general framework of neuro-symbolic approaches:

- the paper “Semirings for probabilistic and neuro-symbolic logic programming” by V. Derkinderen, R. Manhaeve, P. Zuidberg Dos Martires and L. De Raedt provides an overview and synthesis of neuro-symbolic methods, thereby contributing a unified algebraic perspective on the different flavors of probabilistic logic programming (PLP),

showing that many if not most of the extensions of PLP can be cast within a common algebraic logic programming framework, in which facts are labeled with elements of a semiring and disjunction and conjunction are replaced by addition and multiplication. In this unified perspective, the authors focus on the ProbLog language and its extensions;

- in “CCN+: A neuro-symbolic framework for deep learning with requirements”, E. Giunchiglia, A. Tatomir, M. C. Stoian and Th. Lukasiewicz propose a novel neuro-symbolic framework integrating propositional logic requirements into the output layer of neural networks, ensuring compliance with the requirements and enhancing performance. Extensive experimental evaluation shows that CCN+ outperforms both its neural counterparts and the state-of-the-art models in multi-label classification tasks;
- in “Quantified neural Markov logic networks” P. Jung, G. Marra, and O. Kuželka propose a new class of Neural Markov Logic Networks (NMLN), called Quantified NMLN, that extends the expressivity of NMLNs and demonstrate how to leverage the neural nature of NMLNs to employ learnable aggregation functions as quantifiers, and demonstrate the efficiency of Quantified NMLNs in molecule generation experiments.

The next three papers have in common the use of embeddings, although in very different perspectives:

- in “Embeddings as epistemic states: Limitations on the use of pooling operators for accumulating knowledge”, S. Schockaert studies to which extent and under which conditions pooling operators, that are commonly used for aggregating vector embeddings in deep neural network architecture, are compatible with the idea that embeddings encode epistemic states and fit with the satisfaction of propositional formulas. The paper shows that max-pooling is particularly suitable for such tasks, in particular if one wants to encode non-monotonic reasoning;
- in “Explaining answers generated by knowledge graph embeddings” A. Ruschel, A. Colombini Gusmão, and F. Gagliardi Cozman propose techniques to generate explanations for predictions based on the embeddings of knowledge graphs: the idea is to build explanations out of paths in an input knowledge graph, searched through contextual and heuristic cues;

- in “Enriching interactive explanations with fuzzy temporal constraint networks” M. Canabal-Juanatey, J. M. Alonso-Moral, A. Catala, A. Bugarn-Diz propose a model for fuzzy temporal reasoning to overcome some inconsistencies detected in pre-trained language models in a specific application domain of a conversational agent carefully designed for providing users with explanations. The formalism used allows to represent imprecise temporal information and provides mechanisms for checking consistency in conversations. An experiment with GPT-3 Large Language Model is reported.

The next three papers can be roughly put under the banner of reasoning for learning:

- Thi-Bich-Hanh Dao and Ch. Vrain, in “A review on declarative approaches for constrained clustering”, offer a detailed survey of declarative frameworks for clustering. The other two articles exploit the general idea of taking advantage of analogies, but in different ways;
- in “Some thoughts about transfer learning. What role for the source domain?” A. Cornuéjols discusses and revisit some commonly held beliefs about assumptions needed for transfer learning to work. In particular, the paper questions the need to have a good hypothesis on the source space, and demonstrates that the relation between the source and target space, even when those are different, can matter more than the quality of the source hypothesis;
- in “Revisiting analogical proportions and analogical inference” M. Bounhas and H. Prade provide a thorough analysis and experiments regarding the role in analogical inference played by analogical proportions where the four items involved have the same value for some features.

We then have a series of papers that relate Machine Learning to some particular Knowledge Representation formalisms:

- in “Towards an effective practice of learning from data and knowledge” Yizuo Chen, Haiying Huang, and A. Darwiche present recent advances on learning and inference using Bayesian networks, a popular formalism for learning structured knowledge from data. They describe how the compilation of Bayesian networks into arithmetic circuits in the form of tensor graphs and the exploitation of partially known functional dependencies leads to significant performance improvements;

that paper also shows that one can significantly reduce the reliance on data and improve robustness if one complements data with knowledge, and shows that one can sometimes recover from modelling errors by using the more expressive Testing Bayesian networks;

- in “Reasoning and learning in the setting of possibility theory - Overview and perspectives” D. Dubois and H. Prade emphasize that possibility theory stands halfway between logical and probabilistic representation frameworks, and that while qualitative possibility theory is totally compatible with classical logic, quantitative possibility theory can be related to statistics. This suggests that possibility theory may be an interesting setting for interfacing reasoning and learning;
- in “A statistical approach to learning constraints”, by S. Prestwich and N. Wilson, a novel, statistical approach based on sequential analysis is proposed that is fast, can handle large biases, and can accurately learn constraints from noisy data. This article concerns Constraint Programming, which has proved to be an important paradigm for representing and solving numerous combinatorial decision or optimization problems. However, as representing a problem with constraints requires some expertise, this has led to a line of work on automated constraint acquisition, of which this article is a representative;
- in “Questionable stepwise explanations for a robust additive preference model”, M. Amoussou, Kh. Belahcène, Ch. Labreuche, N. Maudet, V. Mousseau, and W. Ouerdane propose to extend necessary explanations of robust additive models by possible, yet non-necessary steps that can explain a preference, in order to extend the scope of explanation engines. Among other things, this allows the decision maker to scrutinize and confirm or criticize the provided explanations. This article is about one of the most studied preference model (the additive one), which is only one type of preference models among the many ones existing, as the study of preferences has a long history in several research fields;
- in “Learning decision catalogues for situated decision making: The case of scoring systems” by S. Heid, J. Hanselle, J. Fürnkranz, E. Hüllermeier, the authors intend to learn a set of easy-to-interpret models, namely scoring rules (notably used by physicians), so that the decision model can be adapted to the context at hand.

Those two last papers use knowledge representation in order to make models more transparent. Together with those, two next papers can be related to general accountability concerns. The first one provides a theoretical setting for finding existing biases in a model, while the second is interested in the computation of counter-factuals in causal graphical models. More precisely:

- in “Learning Horn envelopes via queries from language models”, S. Blum, R. Koudijs, A. Ozaki, and S. Touileb propose an exact learning algorithm to extract a Horn theory that corresponds in some sense to a given learnt model: this Horn theory can be used to explore existing biases in the model; they apply it in particular to probe occupational gender biases in BERT-based language models.
- in “On the efficient bounding of counterfactual queries”, M. Zaffalon, A. Antonucci, R. Cabañas, D. Huber and D. Azzimonti leave the realm of logic for the imprecise probability representation setting. They study the difficulty of performing counter-factual queries that explicitly account for causal structures. In particular, they show that sets of probabilities (a.k.a. credal sets) are well-adapted to deal with such queries, but that their exact resolution is NP-hard, offering an efficient heuristic to bypass this computational bottleneck.

Finally, the two last papers concern the handling of imperfect information, in the data for the first paper and in the model for the second one. More precisely:

- in “Learning from fuzzy labels: theoretical issues and algorithmic solutions” A. Campagner studies the complexity of applying generalized risk minimisation (GRM) to uncertain supervised data when uncertainty is described by fuzzy sets. He shows in particular that obtaining good models with theoretical guarantees using GRM may be computationally challenging, and proposes an alternative learning using randomisation and ensembling that is efficient and provides good empirical performances;
- in “Beyond tree-shaped credal probabilistic circuits” D. R. Montalván Hernández, T. Centen, Th. Krak, E. Quaeghebeur and C. de Campos consider probabilistic circuits, that are emerging as an efficient tool to integrate knowledge in learning techniques, as well as to derive generative models with a clear probabilistic semantic. Such circuits are also computationally attractive, as many queries can be performed in

polynomial time. The paper studies their robust counter-part, where probabilities are allowed to become imprecise. It shows that in such a situation, performing queries become non-polynomial in many cases, but also offers efficient and accurate heuristics to solve this issue.

References

- [1] K. R. Amel. From shallow to deep interactions between knowledge representation, reasoning and machine learning (extended abstract). In N. Ben Amor, B. Quost, and M. Theobald, editors, *Proc. 13th Int. Conf. on Scalable Uncertainty Management (SUM'19), Compiègne, Dec. 16-18*, volume 11940 of *LNCS*, pages 447–448. Springer, 2019.
- [2] Z. Bouraoui, A. Cornuéjols, T. Denoeux, S. Destercke, D. Dubois, R. Guillaume, J. Marques-Silva, J. Mengin, H. Prade, S. Schockaert, M. Serrurier, and C. Vrain. From shallow to deep interactions between knowledge representation, reasoning and machine learning (Kay R. Amel group). *CoRR*, abs/1912.06612, 2019.