EDITORIAL

Guest Editorial: Reasoning With Inconsistent, Incomplete, and Uncertain Knowledge

Enrico Malizia, *University of Bologna, 40126, Bologna, Italy*
Cristian Molinaro and Francesco Parisi, *University of Calabria, 87036, Arcavacata, Italy*

The management and processing of inconsistent, incomplete, and uncertain knowledge is a crucial issue in the development of intelligent systems. Nowadays, such systems have to efficiently manage large amounts of information of different kinds, often represented in different formats and coming from different sources, such as databases, knowledge bases, sensor networks, as well as various data-driven applications. In the presence of such complex and heterogeneous forms of information, incompleteness, inconsistency, and/or inherent uncertainty inevitably arise. These scenarios call for innovative and intelligent approaches that, by leveraging AI techniques, can explicitly represent inconsistency, incompleteness, and uncertainty, and adequately deal with them. Such approaches are crucial to model real-world scenarios, making systems more effective and successful. Moreover, in many domains, knowledge is subject to frequent changes, so handling evolving knowledge is a key feature that knowledge-based systems should provide. In domains having high-impact consequences (e.g., healthcare and cybersecurity), intelligent systems should support human-in-the-loop models that provide tools to help users to understand and interpret the decisions they suggest, while tackling the challenges of inconsistency, incompleteness, and uncertainty. The AI community has also lately been facing the rising demand of explainable AI systems, which have inevitably to deal with inconsistency, incompleteness, and uncertainty. Also to support this evolution that AI systems will unavoidably need to face, we have organized the present Special Issue in *IEEE Intelligent Systems* to solicit the latest advancements on this topic.

1541-1672 © 2023 IEEE
Digital Object Identifier 10.1109/MIS.2022.3218913
Date of current version 1 February 2023.

We have aimed to encourage deep research in the abovementioned areas and beyond, focusing on innovative and intelligent approaches that leverage new-generation artificial intelligence techniques.

In the rest of this Guest Editors' Introduction, we first briefly discuss some important emerging trends in the area of the theme topic of the Special Issue, and then introduce the accepted articles by discussing their contribution.

BRIEF OVERVIEW OF RELEVANT TRENDS

Real-world applications often have to manage large amounts of data possibly coming from multiple data sources, and inconsistencies often arise. In such scenarios, standard reasoning tasks might become meaningless, and different inconsistency-aware approaches need to be devised. Being able to provide meaningful reasoning mechanisms in the presence of inconsistency is a critical issue to make intelligent systems successful in practice. To tackle this problem, a significant body of work in the literature has adopted the approach of "tolerating" inconsistency while devising appropriate reasoning mechanisms that are able to provide reliable information to users despite the presence of inconsistency. In this area, inconsistency-tolerant semantics of query answering are a prominent approach, with the most popular example being consistent query answering, first developed for relational databases in[1] and then applied to knowledge bases [e.g., expressed via existential rules languages[2] and description logics (DLs)[3]] under the name of ABox repair (AR) semantics. Many other inconsistency-tolerant semantics have been proposed over the years, often with aim of approximating the AR semantics, including the intersection of repairs (IAR)[3] and the intersection of closed repairs (ICR) semantics.[4] Another approach to deal with inconsistency is based on paraconsistent techniques, which allow contradictory knowledge without trivialization and are often based on multivalued logics.

Nowadays, data are often gathered from unreliable sources or the data gathering process itself introduces uncertainty (e.g., imperfect sensor readings) and a main challenge is to reason over large-scale uncertain knowledge. Common approaches to quantify degrees of uncertainty are based on probabilistic frameworks. Probabilistic knowledge bases are a prominent example, where probabilities are assigned to data facts and/or additional knowledge (e.g., an ontology defined on top of raw data) to express the confidence we have in different pieces of knowledge. The most studied problem in this setting is query answering, for which different approaches have been proposed (see the work of Borgwardt et al.[5]).

Another methodology to put inconsistent incomplete and uncertain knowledge to good use is formal argumentation, where reasons for and against a claim are analyzed to decide on an outcome, much in the same way as organized human discussions are carried out.[6,7]

Approaches that are complementary to those mentioned previously aim at quantifying the amount of inconsistent and uncertain information in knowledge bases.[8,9] Quantifying and monitoring the amount of inconsistency helps get information on the health status of data, whose quality is more and more important nowadays. Indeed, having information on the quality of data used in machine learning and data-driven approaches is crucial, as poor-quality data can have serious adverse consequences on the quality of decisions made using AI systems.[10]

Explainability of AI systems has become a prominent problem in different AI areas. Indeed, the fact that several recent AI systems operate mostly as black boxes has raised some serious concerns.[11] In this regard, logic-based systems, such as knowledge bases expressed via existential rules or description logics (DLs), have a clear advantage in simplifying the task of explaining the systems' responses. This is because the use of logic, which is fully interpretable, naturally lean itself to more easily explain AI systems' choices to humans, even more in the presence of inconsistencies. This route has been recently explored, for example, by a quite active research endeavour aiming at explaining why (or why not) a query is entailed by a knowledge base. This work had to face the inescapable scenario in which reasoning is carried out over inconsistent data. In fact, explanations for query entailment in knowledge bases has been studied both in the consistent scenario (see, e.g., the work of Calvanese et al.[12] and Ceylan et al.[13] for DLs; and the work of Ceylan et al.[14,15,16] for existential rules), and in the inconsistent scenario in which reasoning is carried out over inconsistent knowledge (see, e.g., the work of Bienvenu et al.[17] and Du et al.[18] for DLs; and the work of Lukasiewicz et al.[19,20] for existential rules).

CONTENT OF THE SPECIAL ISSUE

In total, we received 13 valid submissions that were evaluated by knowledgeable reviewers in the field. In the following, we provide a brief introduction to the five articles that were competitively selected. We explain how they relate to the topic of the Special Issue.

In the article titled "Computing Abductive Explanations," Caroprese, Zumpano, and Bogaerts propose a technique to compute explanations in abduction settings. Abduction is a reasoning method that explains observations in terms of hypotheses that may have led to those observations, for a given theory. Abduction has taken on fundamental importance in AI as well as in other related disciplines. Lately, abduction-based approaches have been employed to explain why a query is not entailed by a (consistent) knowledge base: the hypotheses are what is missing from the database to make query entailment (which plays the role of an observation) hold.[16] In general, many explanations can exist, which might be unhelpful to users when they are overwhelmed by a huge number of different explanations. Thus, it is common to select a subset of "preferred" explanations according to some criterion. Common criteria include set-inclusion minimality, minimum cardinality, minimal weight, and minimality under prioritization of individual hypotheses. In the article selected for the Special Issue, the authors considered a particular criterion to assess the "simplicity" of each explanation, with the goal of preferring the simplest explanations. As a measure of the quality of an explanation, its degree of arbitrariness is adopted, which is the number of arbitrary assumptions that have been made to derive the explanation. The explanations having no arbitrariness, called constrained, are the best ones. This article presents a technique to compute constrained explanations, which consists in rewriting a given theory along with some observations into a disjunctive logic program with negation, so that the constrained explanations correspond to a subset of the program's stable models. The approach is particularly appealing in that it makes it possible to leverage off-the-shelf ASP solvers to compute constrained explanations.

In the article titled "A Framework for Reasoning about Uncertainty in Ontologies," Raddaoui, Jabbour, and Ma propose a logic-based argumentation framework that provides different mechanisms to reason over DL ontologies affected by uncertainty, inconsistency, and incoherence. Indeed, most inconsistency-tolerant semantics assume inconsistency comes only from the data, while the ontology is fully reliable. Little work has considered also the ontology as a possible cause of inconsistency—e.g., Eiter et al.[21] and Lukasiewicz et al.[22] extendeded the AR, IAR, and ICR semantics by allowing not only atoms from the database to be removed to restore consistency, but also ontological axioms. In the article selected for the Special Issue, the authors considered the scenario where knowledge can be uncertain (i.e., confidence degrees are associated with statements), inconsistent (i.e., the database together with the ontology do not admit any model), and incoherent (i.e., there is no database satisfying the ontology). The authors proposed different new semantics that can tolerate such general conflicts and study their relationships. More specifically, query entailment can be based on the classical semantics, the IAR semantics family, the cardinality based family, and the agent preference semantics, which allows the proposed framework to be applied to multiagent systems.

In the article titled "Approximating Region Boundaries Based on Qualitative and Quantitative Information" by Long, Sioutis, Li, Meng, and Li, the authors tackled the problem of approximating spatial regions whenever uncertain, incomplete, or inconsistent/contradictory spatial information is involved. This topic has important potential practical applications in AI, as place descriptions are used in everyday communication as a common way to convey spatial information.[23] Generating geometric representations out of natural language descriptions could facilitate decision-making in many situations,[24] e.g., automatically generating a plausible map in emergencies starting from text descriptions concerning fire reports that are provided via mobile devices. Moreover, approximating spatial regions naturally extends to domains and applications that make use of spatiotemporal reasoning with uncertain knowledge.[25,26] In the article selected for the Special Issue, the authors devised a method to generate region approximations based on rough qualitative direction and distance information. To this end, they make use of both symbolic representation and reasoning frameworks and machine learning techniques, trying to leverage the strengths of both paradigms. Experimental results show that the proposed technique can approximate target regions quite effectively and efficiently. It would be interesting to see how the proposed approach can be applied to practical scenarios like mobile robot navigation.[27]

In the article titled "Model-Driven Development of Service Robot Applications Dealing with Uncertain Human Behavior" by Lestingi, Bersani, and Rossi, the authors dealt with the uncertainty that robots will face in those highly variable and unpredictable settings emerging from robot–human interactions. This research area is timely and important, as robots are rapidly evolving from mere high-performing industrial equipment to sophisticated machines making decisions in delicate and uncertain situations, and they are bound for deeply transforming the service sector, even healthcare and home assistance.[28] Although robots are already employed in such environments, they are still not fully integrated in the service settings, because they are highly demanding due to the high variability related to the human behavior that is unconstrained. The authors tackled this problem by proposing a model-driven framework for the development of interactive robotic scenarios in service settings. The framework proposed relies on a state-based formalism, and via formal analysis can reliably estimate the probability of success of human–robot interactions; this provides sound mathematical guarantees of human wellbeing preservation throughout the interaction and timely mission completion despite the inherent uncertainties of the settings. This is also achieved via a novel automata learning algorithm proposed by the authors, which allows for a refinement of the human model based on data collected at runtime, and this reduces that the impact of the setting's uncertainties at design time. The interesting aspect is that the framework is organized so to be accessible also to designers without expertise in formal modeling. The experiments conducted by the authors show that their framework aided the analysis and deployment of interactive robotic tasks with different categories of human subjects, which have successfully been completed with very high probability.

Finally, in the article titled "Machine Learning from Crowds Using Candidate Set-Based Labelling" by Beñaran-Muñoz, Hernández-González, and Pérez, the authors investigated the quite interesting problem of learning a classifier when there is uncertainty on the labels. More specifically, they investigate crowdsourcing in machine learning, which consists in asking to a group of annotators to label the training set for a classification learning task, which is a cheaper alternative to expert annotators.[29] However, as one might expect, learning from crowd-labeled data involves dealing with its inherent uncertainty and inconsistencies. Learning classifiers from crowd-labeled data is a task grounded in two subtasks: estimating the real labels by aggregating the inconsistent labels provided by the annotators, and actually learning the classifying model. These two subtasks can be executed one after the other, or jointly.[30] Traditionally, in the context of crowd-labeled data, each annotator provides a single label for each example in the dataset (called full labeling). However, a single label might not capture the annotators' doubts. Hence, it is proposed a scheme in which annotators are allowed to state multiple labels for the same data, and in this way more information can be extracted, even from the annotator's uncertainty. The basic assumption is that, for each example of the dataset, among the labels provided there is always the real one. The authors of this article go even further by relaxing this assumption and they introduce the concept of candidate labeling. They conduct a deep analysis of the proposed framework proposed and substantiate with experiments that candidate labeling seems to gather more discriminative information than full labeling. They are in fact able to show that candidate labeling can produce classifiers whose performance is no worse than using full labeling, even with fewer annotators or with lower expertise annotators. Using candidate labeling would hence be a way to have systems capable of learning from their interaction with humans even in presence of uncertainty.

ACKNOWLEDGMENTS

We would like to express our sincere appreciation to the former Editor-in-Chief, Professor V.S. Subrahmanian, Professor Longbing Cao, the journal's administrative team, and the authors for their support in making this special issue possible. Special thanks also go out to the reviewers for their constructive and timely feedback, which was fundamental to the process of creating a high-quality issue.

REFERENCES

1. M. Arenas, L.-E. Bertossi, and J. Chomicki, "Consistent query answers in inconsistent databases," in *Proc. 18th ACM SIGMOD-SIGACT-SIGART Symp. Princ. Database Syst.*, 1999, pp. 68–79.
2. T. Lukasiewicz, E. Malizia, M. V. Martinez, C. Molinaro, A. Pieris, and G. I. Simari, "Inconsistency-tolerant query answering for existential rules," *Artif. Intell.*, vol. 307, 2022, Art. no. 103685.
3. D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, and D.-F. Savo, "Inconsistency-tolerant semantics for description logics," in *Proc. Int. Conf. Web Reasoning Rule Syst.*, 2010, pp. 103–117.
4. M. Bienvenu, "On the complexity of consistent query answering in the presence of simple ontologies," in *Proc. AAAI Conf. Artif. Intell.*, 2012, pp. 705–711.
5. S. Borgwardt, İ. İ. Ceylan, and T. Lukasiewicz, "Recent advances in querying probabilistic knowledge bases," in *Proc. Int. Joint Conf. Artif. Intell.*, 2018, pp. 5420–5426.
6. D. Gabbay, M. Giacomin, G. R. Simari, and M. Thimm, Eds., *Handbook of Formal Argumentation*, vol. 2, Norcross, GA, USA: College Publications, 2021.
7. F. Parisi and G. I. Simari, "Guest editorial argumentation-based reasoning," *IEEE Intell. Syst.*, vol. 36, no. 2, pp. 3–5, Mar./Apr. 2021.
8. J. Grant, M. V. Martinez, C. Molinaro, and F. Parisi, "Dimensional inconsistency measures and postulates in spatio-temporal databases," *J. Artif. Intell. Res.*, vol. 71, pp. 733–780, 2021.
9. J. Grant and F. Parisi, "General information spaces: Measuring inconsistency, rationality postulates, and complexity," *Ann. Math. Artif. Intell.*, vol. 90, no. 2/3, pp. 235–269, 2022.
10. I. H. Sarker, "Data science and analytics: An overview from data-driven smart computing, decision-making and applications perspective," *SN Comput. Sci.*, vol. 2, no. 5, 2021, Art. no. 377.
11. F. K. Došilović, M. Brčić, and N. Hlupić, "Explainable artificial intelligence: A survey," in *Proc. 41st Int. Conv. Inf. Commun. Technol., Electron. Microelectronics*, 2018, pp. 210–215.
12. D. Calvanese, M. Ortiz, M. Simkus, and G. Stefanoni, "Reasoning about explanations for negative query answers in *DL-Lite*," *J. Artif. Intell. Res.*, vol. 48, pp. 635–669, 2013.
13. İ. İ. Ceylan, T. Lukasiewicz, E. Malizia, and A. Vaicenavičius, "Explanations for ontology-mediated query answering in description logics," in *Proc. 24th Eur. Conf. Artif. Intell.*, 2020, pp. 672–679.
14. İ. İ. Ceylan, T. Lukasiewicz, E. Malizia, and A. Vaicenavičius, "Explanations for query answers under existential rules," in *Proc. Int. Joint Conf. Artif. Intell.*, 2019, pp. 1639–1646.
15. İ. İ. Ceylan, T. Lukasiewicz, E. Malizia, C. Molinaro, and A. Vaicenavičius, "Preferred explanations for ontology-mediated queries under existential rules," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 6262–6270.
16. İ. İ. Ceylan, T. Lukasiewicz, E. Malizia, C. Molinaro, and A. Vaicenavicius, "Explanations for negative query answers under existential rules," in *Proc. 17th Int. Conf. Princ. Knowl. Representation Reasoning*, 2020, pp. 223–232.
17. M. Bienvenu, C. Bourgaux, and F. Goasdoué, "Computing and explaining query answers over inconsistent *DL-Lite* knowledge bases," *J. Artif. Intell. Res.*, vol. 64, pp. 563–644, 2019.

18. J. Du, K. Wang, and Y.-D. Shen, "Towards tractable and practical ABox abduction over inconsistent description logic ontologies," in *Proc. AAAI Conf. Artif. Intell.*, 2015, pp. 1489–1495.
19. T. Lukasiewicz, E. Malizia, and C. Molinaro, "Explanations for inconsistency-tolerant query answering under existential rules," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 2909–2916.
20. T. Lukasiewicz, E. Malizia, and C. Molinaro, "Explanations for negative query answers under inconsistency-tolerant semantics," in *Proc. Int. Joint Conf. Artif. Intell.*, 2022, pp. 2705–2711.
21. T. Eiter, T. Lukasiewicz, and L. Predoiu, "Generalized consistent query answering under existential rules," in *Proc. 15th Int. Conf. Princ. Knowl. Representation Reasoning*, 2016, pp. 359–368.
22. T. Lukasiewicz, E. Malizia, and C. Molinaro, "Complexity of approximate query answering under inconsistency in Datalog+/-," in *Proc. Int. Joint Conf. Artif. Intell.*, 2018, pp. 1921–1927.
23. J. Kim, M. Vasardani, and S. Winter, "Similarity matching for integrating spatial information extracted from place descriptions," *Int. J. Geograph. Inf. Sci.*, vol. 31, no. 1, pp. 56–80, 2017.
24. Z. Xu et al., "Crowdsourcing based description of urban emergency events using social media big data," *IEEE Trans. Cloud Comput.*, vol. 8, no. 2, pp. 387–397, Apr./Jun. 2020.
25. F. Parisi and J. Grant, "Knowledge representation in probabilistic spatio-temporal knowledge bases," *J. Artif. Intell. Res.*, vol. 55, pp. 743–798, 2016.
26. J. Grant, C. Molinaro, and F. Parisi, "Probabilistic spatio-temporal knowledge bases: Capacity constraints, count queries, and consistency checking," *Int. J. Approx. Reasoning*, vol. 100, pp. 1–28, 2018.
27. G. N. DeSouza and A. C. Kak, "Vision for mobile robot navigation: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 2, pp. 237–267, Feb. 2002.
28. C. Benedikt Frey and M. A. Osborne, "The future of employment: How susceptible are jobs to computerisation?," *Technol. Forecasting Soc. Change*, vol. 114, pp. 254–280, 2017.
29. J. Zhang, X. Wu, and V. S. Sheng, "Learning from crowdsourced labeled data: A survey," *Artif. Intell. Rev.*, vol. 46, no. 4, pp. 543–576, 2016.
30. V. C. Raykar et al., "Learning from crowds,," *J. Mach. Learn. Res.*, vol. 11, pp. 1297–1322, 2010.

ENRICO MALIZIA is a senior assistant professor at the University of Bologna, 40126, Bologna, Italy. He was a research assistant at the University of Oxford. He was then appointed as an assistant professor at the University of Exeter, and afterward at King's College London. His research interests include artificial intelligence, knowledge representation and reasoning, and computational complexity theory. Malizia received his Ph.D. degree from the University of Calabria, Arcavacata, Italy. Contact him at enrico.malizia@unibo.it.

CRISTIAN MOLINARO is an associate professor at the University of Calabria, 87036, Arcavacata, Italy, where he was an assistant professor. He was a visiting professor at the State University of New York at Buffalo, George Mason University, and Dartmouth College. He was a faculty research assistant at the University of Maryland Institute for Advanced Computer Studies. His research interests include database theory, and knowledge representation and reasoning. Molinaro received his Ph.D. degree in computer science engineering from the University of Calabria. Contact him at cmolinaro@dimes.unical.it.

FRANCESCO PARISI is an associate professor of computer engineering at the Department of Informatics, Modeling, Electronics, and Systems Engineering, University of Calabria, 87036, Arcavacata, Italy. He was a visiting researcher at the University of Maryland Institute for Advanced Computer Studies during 2008. His research focuses on topics within artificial intelligence and databases, particularly in the areas of inconsistency management, knowledge representation, and reasoning under uncertainty. Parisi received his Ph.D. degree in computer science and systems engineering from the University of Calabria. Contact him at fparisi@dimes.unical.it.