Final PhD Report (2019/24)
XXXV cycle

Maria Concetta Morelli
maria.morelli@unical.it

Contents

1 Introduction

The following provides an overview of my PhD, emphasizing the principal outcomes and completed activities. Starting from my previous training as a mathematician, my first year of PhD has been mainly dedicated to training activities. During the first year, I also kept on the work of my master's degree thesis. The work has been refined and extended and has contributed to a published paper. The following years have been principally dedicated to my primary field of research: Stream Reasoning. Some of my research results are currently in a journal article (Theory and Practice of Logic Programming); I presented others at the 4th International Workshop on the Resurgence of Datalog in Academia and Industry, DATALOG 2.0 2022. Furthermore, I was involved in didactic activities as a teaching assistant for the Knowledge Representation course and as a peer tutor for the Discrete Mathematics course, both in the Mathematics and Computer Science department.

2 Primary field of research and results obtained

I studied the characteristics of Stream Reasoning, a research area that deals with the inference of information on streaming data. Typical scenarios of SR are financial transaction auditing and monitoring wind power, traffic, or patients. The contribution that my research has made to the field during PhD can be summarized as follows:

- definition of syntax and semantics of a suitable SR language;
- design of a Framework for study the relative expressiveness of logic-based languages for reasoning over streams;
- identification of expressible fragments through the design Framework.

2.1 LDSR: The Language of DLV for Stream Reasoning

The $LDSR$ language aims to enrich the highly declarative nature of Datalog with constructs helpful in modeling the problems of SR's scenarios. The information is represented by ground predicate atoms associated with different time points. The language provides a set of constructs that allow reasoning about streams. It provides the possibility to determine whether:

- an atom is present at least n times at certain previous time points
- an atom is always present at certain previous time points
- the exact number of times an atom is present at certain previous time points.

In particular, if the time points are contiguous, this determines the analysis within a window. $LDSR$ also inherits the aggregate literals from ASP enriching the expressive potential of the language. The language is supported by the *I-DLV-sr* system, which is a new logic-based system for stream reasoning. The modeling capabilities of $LDSR$ have made it possible to model tasks relating to urban traffic management in the "model

Table 1: $LARS$ to $LDSR$

ϕ	strictly	not strictly
atomic	*Blocked marked* (BM)	*Blocked marked* (BM)
bound	*Marked free* (MF)	*Marked free* (MF)
full	*Simple head* (SH)	*Simple head* (SH)

and solve" Stream Reasoning Hackathon 2021 [1] concerning the implementation of a Cooperative Intelligent Transportation Systems (C-ITS). According to the vote of the organizing committee and other participants, our solution ranked first in the competition.

2.2 Comparison of expressiveness between LDSR and LARS

One of the main contributions to the definition of languages for reasoning over streams using ASP is LARS:A Logic-based framework for Analytic Reasoning over Streams. The language differs from $LDSR$ in syntax and semantics, and I worked on a formal comparison of the two languages. In particular, I worked on defining a framework that allows the two languages to be compared identifing three output profiles —called *atomic*, *bound*, and *full*— that fix the form of the outputs stream defined respecting the own semantics of the two languages. I worked on a comparison from the expressive point of view of the two languages, identifying fragments of each language that can be expressed in the other. In particular, the analysis differentiates expressible fragments from strictly expressible fragments; for the former, there is a translation that involves the addition of auxiliary predicates, and for the latter, there is instead a translation for which the outputs coincide without filtering. For each given form of output and type of rewriting, we studied how to build fragments of a language that could meet the desired expressiveness. To do this, we considered the semantics behind each construct or combination of constructs that can occur in the rules and the effect of interactions between the different rules. For each language, we identified fragments that can be expressed by the other one, showing possible rewritings. In particular, Table 1 presents the fragments of $LARS$ BM, MF and SH. All of them are strictly expressible via $LDSR$. BM is the largest identified fragment and it is atomic-expressible; MF is obtained from BM by imposing some restriction and it is bound-expressible, while SH is obtained by further restricting MF, and allows for achieving full-expressivity.

As for $LDSR$, Table 2 summarizes its identified fragments IH,TR,TC and DA. The largest fragment is IH, which is bound-expressible via $LARS$; the fragment TR, that restricts IH, allows full expressiveness; the other two fragments allow for strict expressiveness: TC further restricts TR, and it is full-expressible, while DA is obtained by imposing restrictions on IH and is bound-expressible.

[1] https://streamreasoning.org/events/stream-reasoning-hackathon-2021/

Table 2: $LARS$ to $LDSR$

ϕ	strictly	not strictly
atomic	*Temporary checks* (TC) *Definible atoms* (DA)	*Intentional head* (IH)
bound	*Temporary checks* (TC) *Definible atoms* (DA)	*Intentional head* (IH)
full	*Temporary checks* (TC)	*Temporary rules* (TR)

3 Training Activities

3.1 Doctoral Courses.

I attended the following courses offered at my university for doctoral students.

- *Process Mining* (8 hours), Dr. Claudio Di Ciccio (Sapienza University of Rome) held in the days 7 and 8 of April 2020.
- *Embodiment of AI* (35 hours),Prof.ssa Elena De Momi (Polytechnic University of Milan) held in the days from May 11 to May 22, 2020.
- *Classic Algorithms: Past, Present and Future* (8 hours), Dott.ssa Annarosa Serpe, (University of Calabria) held in the days 8,9,10,11 of June, 2020.
- *PhD3.0 - Valorizzazione della ricerca e Creazione d'impresa* (3 CFU), University of Calabria, held in Dicember, 2020.
- *Optimization Models for Machine Learning*, Prof. Antonio Fuduli (University of Calabria) held in the days from March 22 to March 26, 2021.
- *Logic-Based Learning for Interpretable AI: Recent Advancements and Future Directions*,Prof.ssa Alessandra Russo and Dr. Mark Law (Imperial College) held in the days 27,30 of April and in the days 4,7 of May, 2021.

3.2 Institutional Courses.

I attended the following lists of institutional courses.

- *Fundamentals of Computer Science*, (12 CFU), Prof. Simona Perri.2020.
- *Object-Oriented Programming*, (6 CFU), Prof. Carmine Dodaro.2020.
- *Algorithms and Data Structures*, (6 CFU), Prof. Giorgio Terracina.2020.
- *Databases*, (9 CFU), Prof. Pasquale Rullo.
- *Graphical interfaces and event programming*, (9 CFU), Prof. Carmine Dodaro. 2020.
- *Knowledge Representation*, (6 CFU), Prof. Mario Alviano. 2020.

3.3 Schools and Conferences.

I attended the following schools and conferences.

- RW 2020 - 16th Reasoning Web Summer School held in the days from June 24 to June 26, 2020.
- Autumn School on Logic and Constraint 2020 held in the days 18 and 19 of September 2020.
- AIXIA 2019 - 18th International Conference of the Italian Association for Artificial Intelligence.
- ICLP 2020 - 36th International Conference on Logic Programming.
- CILC 2020 - 35th Edition of the Italian Conference on Computational Logic.

3.4 Reading Material.

Below is a short list of the books and papers that I consulted the most during the self-study phase. In particular, I report the texts that I have consulted most often. These texts are not to be interpreted as my only sources of study.

- EMANUELE FALZONE, RICCARDO TOMMASINI, EMANUELE DELLA VALLE. *Stream Reasoning: From Theory to Practice*, RW2020.
- WOLFGANG FABER, *An Introduction to Answer Set Programming and Some of Its Extensions*, RW2020.
- SERGE ABITEBOUL, RICHARD HULL, AND VICTOR VIANU. *Foundations of Databases*. Addison-Wesley, 1995.
- EMANUELE DELLA VALLE, STEFANO CERI, FRANK VAN HARMELEN, DIETER FENSE, *It's a Streaming World! Reasoning upon Rapidly Changing Information*.
- DANIELE DELL'AGLIO, EMANUELE DELLA VALLE, FRANK VAN HARMELENC, ABRAHAM BERNSTEIN, *Stream reasoning: A survey and outlook*.
- HARALD BECK, MINH DAO-TRAN, THOMAS EITER, *LARS: A Logic-based framework for Analytic Reasoning over Streams*.
- KRZYSZTOF R. APT, HOWARD A. BLAIR, ADRIAN WALKER, *Towards a Theory of Declarative Knowledge*.
- LIVIO ROBALDO, *Towards compliance checking in reified I/O logic via SHACL*
- International-Headache-Society. Headache classification committee of the international headache society (IHS) the international classifi- cation of headache disorders, 3rd edition. Cephalalgia, 38(1):1–211, 2018.

4 Further research activity.

In addition to working on my primary research area, I have collaborated with two other research groups. The following describes the context and my contributions.

4.1 A logic-based decision support system for the diagnosis of headache disorders.

Headache disorders represent one of the most common and disabling conditions of the nervous system throughout the world. To support the diagnosis of this complex and vast spectrum of disorders, in 1988, the International Headache Society released the first edition of the International Classification of Headache Disorders, now in its 3rd edition: a document of 200 pages classifying, in a taxonomic way, more than 300 different kinds of headaches, and where each single form of headache is identified via a collection of specific nontrivial diagnostic. Typically, the diagnostic evaluation of headaches is mainly based on the description of symptoms by the patient. However, the medical history process may be inaccurate and not exhaustive, due to the high number of headache disorders identified by the medical community and characterized via the *ICHD-3*. Thus, it is of paramount importance in this specific medical field to support clinicians and specialists during the entire diagnostic phase. To make considerable steps forward in the diagnosis and management of headache disorders, the Italian Ministry of Economic Development appreciated and founded the research project *Alcmeone*, the main aim of which is providing an innovative organizational and management model, and an advanced technological platform of services for supporting the integrated clinical management of headache patients. In particular, concerning the headache diagnosis, the goal is to develop a decision support system that augment clinicians to deal more efficiently and effectively with complex decision-making processes. HEAD-ASP is a novel decision support system for the diagnosis of headache disorders that implements a dynamic questionnaire that complies with *ICHD-3* to reach a complete diagnosis. The DSS HEAD-ASP uses two formal logical modules expressed in ASP: the first one is a deductive module that faithfully encodes all primary headache diagnoses and criteria of *ICHD-3*, whereas the second one is an optimization module for minimizing the number of questions that are necessary to complete the diagnostic picture of patients. My contribution to the project concerns the logic module that implements the interactive questionnaire. The goal has been to develop a decision support system implementing a questionnaire that:

- rigorously guides both clinicians and patients during the medical history process;
- automatically adapts to patients;
- reaches, in a reasonable amount of time, a complete diagnostic picture by inferring every diagnosis as either *compatible* or *not compatible*.
- keeps reasonably low the number of questions posed to patients during the medical history process.

Since the patient is not always able to adequately describe his disorders, in order to avoid that the anamnesis is confused or, even, misleading, it is necessary that the questionnaire guides him in the exposure of his condition. We formulate each possible question as binary (i.e., their answers are simply "yes" or "no") to lead the patient in an accurate evaluation of the diagnostic parameters, allowing him to focus on the peculiarities that characterize each individual pathology. Every answer provides new information concerning the symptomatology, so It is necessary, at each step, a resizing of the domain of questions. We discard from the set of candidate next questions those that are irrelevant or inappropriate to the current step: questions relating to diagnoses already inferred as not compatible or concerning information dependent on others not yet verified or already acquired indirectly from the reported answers as a consequence of logical considerations. Finally, we implement a strategy that first computes, for each candidate next question, the minimum number of inferred diagnoses when considering both the "yes" or "no" answer and then, among these values, it selects the maximum one along with its associated question. The results of our experiment reveal that HEAD-ASP is able to producing efficient and effectives questionnaires. It also is receiving very positive feedback from the group of neurologists that is testing it.

4.2 Compliance checking with conflicting and compensatory norms via ASP rules

Compliance checking is the task of legal reasoning to assess whether a certain state of affairs conforms or does not conform to a set of norms. Norms may conflict with others or provide for compensatory actions in case prohibited actions occur or in case obligatory actions do not occur. ASP is well suited to define if-then rules that formalize the norms that establish what is obligatory, prohibited or permitted and to establish when these norms are broken and a violation occurs. The use of negation by failure makes it possible to manage the defeasibility of rules that allow exceptions defined in other rules and to indicate violations unless compensatory actions are present. I wrote ASP rules that describe an use case using unary and binary predicates corresponding to the classes and object properties of an OWL ontology. The work is in a context of comparison with other formalisms.The writing of the rules follows the logic of reification and has been written as straightforwardly as possible with other existing formalisations to allow easier comparison.

5 Publications

- R. COSTABILE, G. CATALANO, B. CUTERI, M. C. MORELLI, N. LEONE, M. MANNA. (2020). A logic-based decision support system for the diagnosis of headache disorders according to the ICHD-3 international classification. *Theory and Practice of Logic Programming (TPLP)*, 20(6), 864-879.

- FRANCESCO CALIMERI, MARCO MANNA, ELENA MASTRIA, MARIA CONCETTA MORELLI, SIMONA PERRI, JESSICA ZANGARI. (2021). I-DLV-sr: A

Stream Reasoning System based on I-DLV. *Theory and Practice of Logic Programming (TPLP).*

- Livio Robaldo, Sotiris Batsakis, Roberta Calegari, Francesco Calimeri, Megumi Fujita, Guido Governatori, Maria Concetta Morelli, Giuseppe Pisano, Ken Satoh, Ilias Tachmazidis. (2022). Taking stock of available technologies for compliance checking on first-order knowledge. CILC 2022: 1-16.

- Nicola Leone, Marco Manna, Maria Concetta Morelli, Simona Perri. (2022). A Formal Comparison between Datalog-based Languages for Stream Reasoning. Datalog 2022: 151-165.

- Livio Robaldo, Sotiris Batsakis, Roberta Calegari, Francesco Calimeri, Megumi Fujita, Guido Governatori, Maria Concetta Morelli, Francesco Pacenza, Giuseppe Pisano, Ken Satoh, Ilias Tachmazidis, Jessica Zangari. (2023). Compliance checking on first-order knowledge with conflicting and compensatory norms: a comparison among currently available technologies. Artificial Intelligence and Law, 1-51.

6 Further activities

Participation in Committees. I have been a member of the following Organization Committees.

- AIXIA 2019 - 18th International Conference of the Italian Association for Artificial Intelligence.

- ICLP 2020 - 36th International Conference on Logic Programming.

Speaker Activity. I participated as speaker at the 4th International Workshop on the Resurgence of Datalog in Academia and Industry, DATALOG 2.0 2022.

Didactic activities. I performed the following didactic activities:

- (For four years in a row) Peer Tutor - *Discrete Mathematics - Class main professor: John Van Bon, bachelor degree in Computer Science - University of Calabria.*

- Teaching assistant - *Knowledge Representation* - 16 hours - Class main professor: Mario Alviano.

Supervisors: Ch.mo Prof. Marco Manna

Marco Manna
09.04.2024
09:38:57
GMT+01:00

Prof.ssa Simona Perri

SIMONA
PERRI
09.04.2024
13:35:10
GMT+01:00

PhD student: Dott.ssa Maria Concetta Morelli

Maria Concetta Morelli
09.04.2024 08:35:58
GMT+00:00