氏　　　名	Nguyen Quoc Tuan
学位(専攻分野)	博士(情報学)
学 位 記 番 号	総研大甲第 2369 号
学位授与の日付	2022 年 9 月 28 日
学位授与の要件	複合科学研究科　情報学専攻 学位規則第6条第1項該当
学 位 論 文 題 目	On Linear Algebraic Computation for Logic Programming
論 文 審 査 委 員	主　査　　井上　克巳 情報学専攻　教授 宇野　毅明 情報学専攻　教授 杉山　麿人 情報学専攻　准教授 坂間　千秋 和歌山大学　システム工学部　教授 佐藤　泰介 国立情報学研究所　特任研究員／ 東京工業大学　名誉教授

(Form 3)

Summary of Doctoral Thesis

Name in full Nguyen Quoc Tuan

Title On Linear Algebraic Computation for Logic Programming

Logic Programming (LP), which provides languages for declarative problem solving and symbolic reasoning, has started gaining more attention in terms of building explainable learning models. For decades, LP representation has been considered mainly in the form of symbolic logic, which is useful for declarative problem solving and symbolic reasoning. Since then, researchers have developed various dedicated solvers and efficient tools for Answer Set Programming (ASP) - LP that are based on the stable model semantics. Recently, LP starts gaining more attention in order to building explainable learning models, whereas it still has some limitations in terms of computation and efficiency. On the other hand, several researchers attempt to translate logical inference into numerical computation in the context of mixed-integer programming. They exploit connections between logical inference and mathematical computation that open a new way for efficient implementation.

Lately, several studies have been done on embedding logic programs into numerical spaces and exploiting algebraic characteristics. There are several reasons for considering linear algebraic computation of LP. First, linear algebra is at the heart of many applications of scientific computation, and integrating linear algebraic computation and symbolic computation is considered a challenging topic in Artificial Intelligence (AI). In particular, transforming symbolic representations into vector spaces and reasoning through matrix computation are considered one of the most promising approaches in neural-symbolic integration. Second, linear algebraic computation has the potential to cope with Web-scale symbolic data, and several studies have developed scalable techniques to process huge relational knowledge bases in tensor spaces. Since relational KBs consist of ground atoms, the next challenge is applying linear algebraic techniques to LP and deductive Knowledge Base (KB)s. Third, it would enable us to use efficient (parallel) algorithms of numerical linear algebra for computing LP, and further simplify the core method so that we can exploit great computing resources ranging from multi-threaded CPU to GPU. The promising efficiency has been reported in GraphBLAS where various graph algorithms are redefined in the language of linear algebra.

Algebraic characterization of logic programs has received increasing attention in recent years. Researchers attempt to exploit connections between linear algebraic computation and symbolic computation to perform logical inference in large-scale knowledge bases. The merit of logic reasoning in vector space is not only the scalability but also the

capability of integrating with other AI techniques such as Artificial Neural Network (ANN). Bridging LP and ANN can open up the gate to building more robust and explainable AI models. However, current work is usually based on manipulating the dense matrix format that is not efficient in both memory and time complexity. In addition, most of the work suffers from the combinatorial explosion problem that researchers have not yet found an appropriate approach to deal with in the language of algebra. Following this direction, we continue to explore the algebraic characterization of logic programs in deduction and abduction. Accordingly, the contribution of the thesis can be summarized in two main directions.

The first direction is about deduction which is described in Chapter 3 of the thesis. We have analyzed the sparsity of matrix representation of logic programs. Then we propose the use of general-purpose sparse representations to utilize the efficiency of linear algebraic approaches for deductive reasoning. We show its great power of computation in reaching the fixed-point of the immediate consequence operator. In particular, performance for computing the least models of definite programs is dramatically improved using the sparse matrix representation. We also apply the method to the computation of stable models of normal programs, in which the guesses are associated with initial matrices, and verify its effect when there are small numbers of negations. Moreover, matrix computation could be more accelerated using GPU. We have tested our implementation in this way and obtained expected results too. In addition to the improvement using sparse representation, we have conducted experiments on different general-purpose sparse matrix representations and demonstrated the merits and demerits of each format.

The second direction is about abduction which is described in Chapter 4 of the thesis. We have extended the linear algebraic characterization in abductive reasoning by exploiting the transpose of the program matrix. Then we propose a novel algorithm, which combines the flexibility and robustness of numerical computation with the compactness and efficiency of set operations, in order to compute solutions of abductive Horn propositional tasks. More importantly, we also have built the theory to prove the correctness of our algorithm in general. Experimental results demonstrate that our method is competitive with conflict-driven techniques. In addition, our spare matrix approach has advantages on the upscaled benchmark dataset and also has the potential to speed up on parallel computing platforms. The merit of solving the Horn abduction problem in vector space is not only the scalability but also the capability of integrating with other AI techniques e.g. ANN.

Lastly, multiple potential directions that can be considered to continue in future research are discussed in this thesis. We hope to continue to explore the ability of the algebraic characterization to deal with probabilistic logic and rule learning. It will be a step further toward a more general framework in which we incorporate many different algorithms in the language of linear algebra.

Form8・Separate Sheet
（様式８・別紙１）

Results of the doctoral thesis screening

博士論文審査結果

Name in Full
氏　名　　Nguyen Quoc Tuan

Title
論文題目　　On Linear Algebraic Computation for Logic Programming

Mr. Nguyen Quoc Tuan, hereafter called the applicant, has submitted the doctoral thesis for the process to determine whether it is worthy of the awarding of a degree. The decision by the panel will be based on the following.

The thesis consists of the applicant's work on linear algebraic computation for logic programming. Logic programming provides languages for declarative problem solving and symbolic reasoning, and it has recently gained more attention as tools for explainable Artificial Intelligence (AI). To overcome computational difficulties for logic programming, researchers have attempted to embed logic programs into numerical spaces and exploiting algebraic characteristics. Considering linear algebraic computation of logic programming has several merits, as linear algebra is at the heart of many applications of scientific computation, and it would enable us to use efficient parallel algorithms of numeric computation for linear algebra by exploiting computing resources like multi-threaded CPUs and GPUs. Moreover, transforming symbolic representations into vector spaces and reasoning through matrix computation are considered a promising approach in neuro-symbolic AI. Following this direction, the thesis explores the algebraic characterization of logic programs in two reasoning modes, i.e., deduction and abduction. Accordingly, the contribution of the thesis can be summarized in two main directions, i.e., deduction and abduction in logic programming, and the thesis consists of six chapters.

Chapter 1 is an introduction that explains the basics of reasoning, motivation and objectives of the research, and organization of the thesis. Chapter 2 introduces the background of logic programming with necessary definitions like the syntax and the semantics of logic programs and abductive logic programs.

Chapter 3 describes the first main contribution of the thesis, and deals with linear algebraic deduction of logic programming. Firstly, the sparsity of matrix representation of logic programs is deeply analyzed by observing that matrices representing propositional logic programs are highly sparse. Then, general-purpose sparse representations are utilized to gain the efficiency of linear algebraic approaches for deduction in logic programming. Linear algebraic computation using sparse representations is extremely efficient in computing the fixpoint of the immediate consequence operator, then the performance to compute the least models of definite programs is dramatically improved using the sparse matrix method. The method is also applied to computation of stable models of normal programs, in which the guesses are associated with multiple initial interpretations that comprise initial matrices, and its effect is verified when there are small numbers of negations. These ideas for efficient computation have been verified by experiments with several large datasets, by showing the efficiency of the sparse method compared with the state-of-the-art solver for Answer Set Programming, which has been widely used as a computational tool for logic programming. Experiments are also conducted using different general-purpose sparse matrix representations by demonstrating the merits and demerits of each format. It has also been verified that matrix computation can be more accelerated using GPU.

Chapter 4 describes the second contribution of the thesis, that is, it extends the linear algebraic computation for logic programming to that for abductive logic programming. Characterization of abductive reasoning is based on the transpose of the program matrix. In order to compute minimal explanations of propositional Horn abduction, a novel algorithm for abduction is proposed by combining the robustness of matrix multiplications with the compactness of set operations in computing minimal hitting sets. Experimental results demonstrate that the proposed method is at least as competitive with state-of-the-art techniques for computing abduction in existing benchmark datasets. In addition, the spare matrix approach has more advantages on upscaled benchmark datasets and has the potential to speed up on parallel computing platforms.

Chapter 5 discussed several related works in a broader view, and Chapter 6 concludes the thesis with the summary of the contributions as well as potential directions that can be considered as future research in such as probabilistic logic programming and rule learning.

As for the publication related to the contents of the thesis, the work on linear algebraic deduction (Chapter 3) has been presented as a peer-reviewed paper in an international conference (ICLP 2020) and has been published in an international journal (New Generation Computing). Linear algebraic abduction (Chapter 4) has been published as a peer-reviewed paper in an international conference (ICTAI 2021).

At the public presentation, the applicant presented his thesis work for 45 minutes in an online meeting on July, 14, 2022. Then a Q&A session was held with the thesis evaluators and other audience. The applicant was able to answer all questions in a satisfactory way.

After the Q&A session, a review committee was held among the evaluators. The committee members expressed their satisfaction with the quality of the PhD work on linear algebraic computation of logic programming, and recognized that the submitted thesis fulfilled the requirements for a PhD thesis. In conclusion, the review committee judged that this thesis was worthy of the awarding of the PhD degree.