

АНОТАЦІЯ

Ленько В.С. Методи та засоби управління персональними знаннями в інтелектуальних системах. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 124 «Системний аналіз» (12 – Інформаційні технології). – Національний університет «Львівська політехніка», Львів, 2020.

Зміст анотації

У дисертаційній роботі здійснено системний аналіз методів та засобів управління персональними знаннями як необхідної компоненти для раціональної поведінки інтелектуального агента. Методами системного аналізу досліджено концепт «знання», властивості та ролі поняття «подання знання», еволюцію та множину найпоширеніших моделей подання знання, а також їхні обчислювальні можливості, що втілюють процес міркування. У дослідженні запропоновано нові підходи до застосування формальних логічних моделей, а саме описових логік і теорій типів. Запропоновано метод подання онтологій в програмному середовищі доведення теорем Coq і розроблено моделі баз знань з використанням зазначених логічних формалізмів для управління знаннями в інформаційних системах, мас-медіа і транспортних мережах.

Робота складається зі вступу, чотирьох розділів, висновків, списку посилань та додатків. В першому розділі подано ґрунтовний аналіз суміжних до теми дослідження наукових публікацій, проаналізовано епістемологічну структуру поняття знання та акцентовано роль зв'язку між істинністю і обґрунтованістю переконання. В цьому ж розділі здійснено порівняльний аналіз найпоширеніших моделей подання знання – фрейми, семантичні мережі, продукційні правила та логічні моделі. Враховуючи п'ять основних ролей подання знання та пов'язані з ними властивості, зокрема виразність і розв'язність, основну увагу дослідження сфокусовано на логічні моделі подання знань.

У другому розділі досліджено логічний формалізм «числення індуктивних конструкцій», який зокрема використовувався для доведення теореми про чотири фарби. Рушієм цього формалізму є теорія типів із залежними типами, яка згідно з ізоморфізмом Каррі-Говарда, втілює синтаксично різновид систем в теорії доведення «природня дедукція». З'ясовано, що основними перевагами числення індуктивних конструкцій є його висока виразність та базування на неklasичній конструктивній логіці, де поняття істинності виразу ґрунтується на можливості побудови цього виразу з примітивів. Крім того визначено, що теорія типів значною мірою уникає парадоксів класичної логіки та сприяє як «механізації» логічного виведення, так і подання виразів, що не володіють властивістю істинності. Недоліком зазначеного формалізму є його нерозв'язність, яка унеможлиблює повну автоматизацію перевірки істинності довільного виразу. Цей недолік вимагає участі дослідника в доведенні теореми, але значну частину перетворень можна автоматизувати за допомогою доступних «тактик». Подання і обґрунтування логічних виразів з використанням числення індуктивних конструкцій здійснюється в інтерактивному асистенті доведення теорем Coq.

Третій розділ подає аналіз описових логік, які є розв'язними фрагментами логіки першого порядку, та широко застосовуються в галузі семантичного вебу, як формальна основа мови онтологій OWL 2. Досліджується структура базової описової логіки *ALC*, допустимі логічні задачі та методи їх розв'язання. З'ясовано, що атрибутивна мова з доповненнями, *ALC*, розширюється новими конструкторами понять та ролей, які формують основу описової логіки SROIQ^(D) – формального рушія мови OWL 2. Формалізм *ALC* використано для подання тверджень з вибраних предметних областей, а для обґрунтування їхньої логічної сумісності застосовано табло-алгоритм. Оскільки описові логіки є підмножинами логіки першого порядку, то їм властиві певні недоліки класичної логіки, зокрема монотонність матеріальної імплікації, яка дозволяє формувати логічно коректні, проте контекстно незв'язні аргументи.

У четвертому розділі проаналізовано особливості проектування системи взаємодії віртуальних наукових колективів, що функціонують з метою побудови баз знань спеціального та загального призначення. В сучасному контексті такі спільноти зазвичай розділені географічно та інституційно, тому виникає потреба в системі, яка дозволить надійно та зручно узгоджувати напрацювання учасників віртуального наукового колективу, з можливістю підтвердження історії та авторства будь-якого твердження бази знань. Оскільки спеціалізація колективу може значно варіюватися, то для подання знань доцільно вибрати мову з високою виразністю, а саме числення індуктивних конструкцій. Для формування і доведення істинності тверджень бази знань пропонується використовувати інтерактивний асистент доведення теорем Coq, який містить мову Gallina та набір тактик для доведення логічних тверджень. Композицію спільної бази знань з напрацювань учасників віртуального наукового колективу доцільно здійснювати за допомогою системи розподіленого контролю версій Git, яка де-факто є стандартом для управління історією програмного коду. Оскільки порядок змін в Git-сховищі можливо редагувати, то для підтвердження авторства важливих змін в базі знань, пропонується використання технології Blockchain як розподіленого реєстру записів.

Актуальність дослідження аргументується невпинним зростанням обсягів та ускладненням структури інформаційних потоків. Прогрес в галузі технологій числення забезпечив формування нейромережових моделей фіксації знання, проте їх здатність до абстракції в сьогоденні є суттєво обмеженою. На противагу, методи концептуалізації та дедуктивні логічні системи пропонують надійні та обґрунтовані формалізми для подання, синтезу та перевірки істинності знання. Розвиток логічних систем уможливив «механізацію», зокрема автоматизацію, процесу міркування над знанням, що сприяє можливості критично аналізувати та проектувати сучасні інформаційні потоки.

Основні наукові результати дисертації отримані у формі фундаментальних та прикладних досліджень:

- розроблено онтологічно-орієнтований метод подання бази знань в середовищі асистента доведення теорем Coq, що сприяє формалізації нових та існуючих онтологій, а також уможлиблює застосування логік вищих порядків як надійного інструменту підтримки прийняття рішення в інтелектуальних системах;
- розроблено архітектуру та програмну реалізацію інтелектуальної системи для управління персональними та колективними базами знань, яка володіє властивостями версійності, розподіленості, автономності та забезпечує можливість надійного підтвердження авторства;
- розроблено приклади формального доведення істинності описових знань, зокрема для задач проектного управління в IT-підприємстві, узгодженості новинних матеріалів в мас-медіа та структурних моделей шляхопроводів, що полегшує та пришвидшує перевірку гіпотез в напівавтоматичному режимі;
- здійснено порівняльний аналіз моделей подання знання, описової логіки та формальних систем теорії типів, а також особливостей їх застосування для подання знань різної природи, зокрема описових речень природньої мови, логічних і математичних теорій та виразів;
- здійснено системний аналіз понять «знання» та «істинність знання», і аргументовано використання конструктивної логіки з метою усунення проблем Гетсьє, які компрометують традиційні методи обґрунтування істинності знання.

Ключові слова

системний аналіз, знання, логіка, числення конструкцій, Coq, описова логіка, онтологія, блокчейн, розподілена система, Git

ABSTRACT

Lenko V.S. Methods and tools for personal knowledge management in intelligent systems. – Qualifying scientific work on the rights of a manuscript.

The dissertation for obtaining a scientific degree of the Doctor of Philosophy on the specialty 124 “System analysis” (12 – Information technologies). – Lviv Polytechnic National University, Lviv, 2020.

Abstract content

The dissertation is devoted to system analysis of methods and tools for personal knowledge management, which is a necessary component for the rational behavior of an intelligent agent. By the methods of system analysis, it investigates the concept of “knowledge”, the properties and roles of “knowledge representation” concept, an evolution of the most common knowledge representation models, as well as their computational capabilities, which resemble the process of human reasoning. The conducted research proposes new approaches to the application of formal logic models, namely descriptive logics and type theories. The method for ontologies representation in the environment of Coq proof assistant is proposed, along with the examples of knowledge base models for the domains of information systems, mass-media, and transport networks.

The thesis consists of an introduction, four chapters, conclusions, a list of references, and appendices. The first section presents a thorough analysis of related scientific publications, analyzes the epistemological structure of knowledge concept, and emphasizes the role of the relationship between truth and validity of belief. The same section provides a comparative analysis of the most widespread knowledge representation models: frames, semantic networks, production rules, and logic-based models. Taking into account five major roles of knowledge representation and their distinguished properties, such as expressiveness and decidability, the main focus of the research is on logic-based models of knowledge representation.

The second section investigates the logical formalism called the “calculus of inductive constructions”, which in particular was used for proving the four color theorem. The driving force behind this formalism is the dependent types theory, which according to the Curry-Howard isomorphism syntactically embodies a kind of systems in proof theory named “natural deduction”. It is found that the main advantages of the calculus of inductive constructions are its high expressiveness and the adherence to nonclassical constructive logic principles, according to which the concept of the truth of an expression is based on the ability to construct this concept from primitives. Additionally, the type theory was designed to avoid the paradoxes of classical logic and facilitates the “mechanization” of logical inference, as well as the representation of expressions that do not possess the property of truth. The main drawback of this formalism is undecidability, which makes impossible the full automation of truth verification for an arbitrary expression. This shortcoming requires the participation of a researcher in the process of theorem proving, but still, a significant part of the syntactic transformations can be automated by existing “tactics”. Representation and justification of the logical expressions are carried out in the interactive proof assistant Coq, using the provided implementation of the calculus of inductive constructions.

The third section presents the analysis of description logics, which are decidable fragments of first-order logic and are widely used in the semantic web as a formal basis for the OWL 2 ontology language. It considers the structure of a basic description logic *ALC*, common reasoning problems and method for establishing knowledge base consistency. It is found that the attributive language with complements, *ALC*, can be extended with a set of concepts and roles constructors that form the basis of a highly expressive descriptive logic *SROIQ*^(D) – the formal core of OWL 2 language. While *ALC* is used to represent the statements, the tableaux algorithm deals with reasoning problems. Since description logics are subsets of first-order logic, they inherit certain shortcomings of classical logic, including the monotonicity of material implication, which allows the formation of logically correct, but contextually incoherent arguments.

The fourth section specifies the design of a collaboration system for virtual scientific communities that aim to build knowledge bases for dedicated and general use. Nowadays, these communities are typically divided geographically and institutionally, so there is a need for a system that will reliably and conveniently reconcile the work of the virtual scientific community members, with the ability to confirm the history and authorship of any expression in the knowledge base. The specialization of a research unit may vary greatly, so it is advisable to choose a language with high expressiveness, namely calculus of inductive constructions. Coq interactive proof assistant provides an implementation of the aforementioned formalism as Gallina language and comes with a set of useful “tactics” for logical inference, so it is proposed to use it for knowledge base engineering. Construction of the shared knowledge base that consists of the parts contributed by the members of a virtual scientific community should be carried out using Git, which is de facto a standard for software code versioning. Since it is technically possible to rewrite the history of commits in a Git repository, it is proposed to establish the authorship of the important knowledge base expressions by the means of Blockchain technology as a distributed ledger system.

The research relevance is justified by the constant growth of information flow volumes and their structural complexity. Advances in computational technology have made it possible to use artificial neural network models for knowledge discovery, but as of today their ability to abstract is severely limited. In contrast, conceptualization methods and logic-based systems offer reliable and sound formalisms for knowledge representation, synthesis, and verification. Curry-Howard isomorphism established the cornerstone relationship between logic-based and computation models and facilitated the “mechanization” of reasoning using verified software-based theorem proving. The ability to detect inconsistencies in multiple statements significantly increases the reliability of the knowledge base, which in turn drastically improves the quality of the input data for the process of intellectual analysis and saves time and efforts of the decision-maker.

The main results of the dissertation are obtained in the form of fundamental and applied research:

- the method for ontologies representation in the environment of Coq proof assistant has been developed, which contributes to the formalization of new and existing ontologies and enables the application of higher-order logics as a reliable decision-support tool;
- the architecture and software implementation of an intelligent system for personal and shared knowledge bases management have been created, which has the properties of being autonomous, distributed, supports versioning and ensures reliable authorship confirmation;
- the examples of knowledge bases and formal reasoning over them have been provided, in particular for the domains of information systems, mass media, and transport networks, which facilitate the semi-automatic procedure of hypotheses testing;
- the comparative analysis of knowledge representation models, description logic, type theories, their characteristic features, and applications to the encoding of the knowledge of the various kinds, including natural language sentences, logical and mathematical expressions, has been conducted;
- the system analysis of the concepts “knowledge”, “truth of knowledge” and the argument on constructive logic adoption for the sake of Gettier problems elimination, which compromise traditional methods of knowledge justification, have been presented.

Keywords

systems analysis, knowledge, logic, calculus of constructions, Coq, description logic, ontology, blockchain, distributed system, Git

Список публікацій здобувача

Стаття у науковому фаховому виданні України, яке включено до міжнародної наукометричної бази (Index Copernicus):

1. Н. Е. Кунанець, В. С. Ленько, В. В. Пасічник, Ю. М. Щербина. Персональні бази даних та знань віртуальних дослідницьких спільнот // Науковий вісник НЛТУ України. – Львів, 2017. – Вип. 27(6). – С. 185-191.

Стаття у науковому періодичному виданні іншої держави:

2. Vasyl Lenko, Nataliia Kunanets, Volodymyr Pasichnyk, Yuriy Shcherbyna. Decentralized Blockchain-based platform for collaboration in virtual scientific communities // ECONTECHMOD. – Lublin, 2019. – Vol. 08(1). – P. 21-26.

Статті у наукових фахових виданнях України:

3. Lenko V. S., Pasichnyk V. V., Shcherbyna Y. M. Knowledge Representation Models. Комп'ютерні науки та інформаційні технології // Вісник НУ «Львівська політехніка». – Львів, 2017. – № 864. – С. 157-168.
4. Д. Р. Табачишин, В. С. Ленько, Н. Е. Кунанець, В. В. Пасічник, Ю. М. Щербина. Експертне оцінювання «розумності міста» із застосуванням нечіткої логіки // Штучний інтелект. – Інститут проблем штучного інтелекту МОН України і НАН України, 2017. – №1 (75). – С. 102-110.
5. В. С. Ленько, Ю. М. Щербина. Застосування методів штучного інтелекту до сегментації графічного образу // Інформаційні системи та мережі. Вісник НУ «Львівська політехніка». – Львів, 2011. – № 715. – С. 194-203.

Матеріали конференцій у збірниках, які включено до міжнародної наукометричної бази (Scopus):

6. Vasyl Lenko, Volodymyr Pasichnyk, Nataliia Kunanets, Yuriy Shcherbyna. Knowledge representation and formal reasoning in ontologies with Coq // Advances in Computer Science for Engineering and Education / editors:

- Zhengbing Hu, Sergey Petoukhov, Ivan Dychka, Matthew He. – Springer, Cham, 2018. – Vol. 754. – P. 759-770. DOI 10.1007/978-3-319-91008-6.
7. Vasyl Lenko, Volodymyr Pasichnyk, Nataliia Kunanets, Yuriy Shcherbyna. Type-Theoretical Foundations of the Derivation System in Coq // Proceedings of the 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, 8-12 October 2018. – IEEE, 2018. – P. 220-225. DOI 10.1109/SAIC.2018.8516885.
 8. Oleksandr Matsiuk, Natalia Kunanets, Volodymyr Pasichnyk, Vasyl Lenko, Yuriy Shcherbyna, Antonii Rzhеuskyi. The procedures of processing of geolocation data on urban underground spaces Proceedings of the 2019 9th International Conference on Advanced Computer Information Technologies (ACIT), Ceske Budejovice, Czech Republic, 5-7 June 2019. – IEEE, 2019. – С. 500-503. 10.1109/ACITТ.2019.8780085.

Матеріали та тези конференцій:

9. В. С. Ленько, В. В. Пасічник, Н. Е. Кунанець, Ю. М. Щербина. Теоретичні аспекти логічного міркування у середовищі Coq // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту: матеріали міжнародної наукової конференції ISDMCI-2018, 21-27 трав. 2018 р., Залізний Порт, Україна / Херсонський нац. технічний ун-т. – Херсон, 2018. – С. 174-176.
10. В. С. Ленько, В. В. Пасічник, Ю. М. Щербина, Н. Е. Кунанець. Міркування в онтологіях з використанням теорії типів // Теоретичні та прикладні аспекти побудови програмних систем: матеріали XIV міжнар. наук. конф. ТАAPSD'2017, 4-8 груд. 2017 р., Київ, Україна / Київський нац. ун-т ім. Т. Г. Шевченка. – Київ, 2017. – С. 138-141.
11. В. С. Ленько, В. В. Пасічник, Н. Е. Кунанець, Ю. М. Щербина. Проектування відкритих децентралізованих реєстрів з використанням технології Blockchain // Математика. Інформаційні технології. Освіта:

- тези доповідей VII міжнар. наук.-практ. конф. МІТО-2018, 3-5 черв. 2018 р., Луцьк-Світязь, Україна / Східноєвропейський нац. ун-т ім. Л. Українки. – Луцьк-Світязь, 2018. – С. 72-74.
12. Н. Е. Кунанець, В. С. Ленько, В. В. Пасічник, Ю. М. Щербина. Подання онтологічних моделей знань в системі інтерактивного доведення теорем Соq // Інформаційні технології та взаємодії: матеріали доповідей IV міжнар. наук.-практ. конф. ІТ&І-2017, 8-10 лист. 2017 р., Київ, Україна / Київський нац. ун-т ім. Т. Шевченка. – Київ, 2017. – С. 169-170.
13. В. С. Ленько, Н. Е. Кунанець, В. В. Пасічник, Ю. М. Щербина. Децентралізована комунікація у віртуальних наукових колективах // Практичне застосування нелінійних динамічних систем в інфокомунікаціях: матеріали VI міжнар. наук.-практ. конф., 9-11 лист. 2017 р., Чернівці, Україна / Чернівецький нац. ун-т ім. Ю. Федьковича. – Чернівці, 2017. – С. 118-119.
14. В. С. Ленько, В. В. Пасічник, Ю. М. Щербина. Проект системи управління персональними знаннями // Управління проектами: стан та перспективи. (ІТПРР-2017) : матеріали XIII міжнародної науково-практичної конференції, 12-15 вересня 2017 р., Миколаїв, Україна / Національний університет кораблебудування імені адмірала Макарова. – Миколаїв, 2017. – С. 61-62.
15. Н. Е. Кунанець, В. С. Ленько, В. В. Пасічник, Ю. М. Щербина. Подання онтологій з використанням теорії типів // Системи та засоби штучного інтелекту: тези доповідей Міжнародної наукової молодіжної школи АІІS'2017, 18 жовт. 2017 р., Київ, Україна / Київський нац. ун-т ім. Т. Шевченка. – Київ, 2017. – С. 114-117.
16. В. С. Ленько, В. В. Пасічник, Ю. М. Щербина. Подання знань та логічні міркування у формальних онтологіях // Комп'ютерне моделювання та програмне забезпечення інформаційних систем і технологій. (КМПЗ-

- 2017): збірник наукових праць (тези доповідей і вибрані статті) III Всеукраїнської міжнародної науково-практичної конференції, 28-30 вересня 2017 р., Рівне, Україна / Національний університет водного господарства та природокористування. – Рівне, 2017. – С. 94-95.
17. V. S. Lenko, V. V. Pasichnyk, Y. M. Shcherbyna. On personal knowledge management systems // Математика. Інформаційні технології. Освіта: тези доповідей VI міжнар. наук.-практ. конф. МІТО-2017, 5-7 черв. 2017 р., Луцьк-Світязь, Україна / Східноєвропейський нац. ун-т ім. Л. Українки. – Луцьк-Світязь, 2017. – С. 57-59.
18. В. С. Ленько, Н. Е. Кунанець, В. В. Пасічник, Ю. М. Щербина. Побудова моделі «Розумного соціополісу» на основі знань експертів // Інформаційні технології, економіка та право: стан та перспективи розвитку. (ІТЕП-2017): матеріали міжнародної науково-практичної конференції, 27-28 квітня 2017 р., Чернівці, Україна / Буковинський ун-т. – Чернівці, 2017. – С. 141-143.
19. V. S. Lenko, V. V. Pasichnyk, Y. M. Shcherbyna. Knowledge Representation Models // Теоретичні та прикладні аспекти побудови програмних систем: матеріали XIII міжнар. наук. конф. ТАAPSD'2016, 5-9 груд. 2016 р., Київ, Україна / Київський нац. ун-т ім. Т. Шевченка, Ін-т післядипломної освіти. – Київ, 2016. – С. 143-152.
20. В. С. Ленько, В. В. Пасічник, Ю. М. Щербина. Методи та засоби управління персональними знаннями // Інформаційно-обчислювальні технології, автоматика та електротехніка: матеріали Міжнар. наук.-практ. конф. ІТАЕ-2016, 10-11 листоп. 2016 р., Рівне, Україна / Нац. ун-т водного господарства та природокористування. – Рівне, 2016. – С. 143-144.
21. В. С. Ленько, Ю. М. Щербина. Відновлення графічних образів за допомогою карт Кохонена // Сучасні проблеми прикладної математики

та інформатики: збірник наукових праць конф. АРАМCS-20, 5-7 жовт. 2016 р., Львів, Україна / Львівський нац. ун-т ім. І. Франка. – Львів, 2016. – С. 108-111.

22. В. С. Ленко. Застосування методів штучного інтелекту до сегментації графічних образів // Чотирнадцята всеукраїнська (дев'ята міжнародна) студентська наукова конференція з прикладної математики та інформатики СНКПМІ-2011: тези доповідей, 5-6 трав. 2011 р., Львів, Україна / Львівський нац. ун-т ім. І. Франка. – Львів, 2011. – С. 60-61.