

УДК 004.82

Лялецький О.О.¹, к.ф.-м.н.
Афонін А.О.², к.ф.-м.н.

Про інтелектуальні та інтерфейсні засоби систем автоматизації пошуку доведень

¹Київський національний університет імені Тараса Шевченка, 03680, м. Київ, пр-т Глушкова 4д, e-mail: foraal@mail.ru

²Національний університет «Києво-Могилянська академія», 04655, м. Київ, вул. Сковороди 2, e-mails: andrew.afomin@gmail.com

A.A.Lyaletsky¹, PhD (Physics&Mathematics)
A.A.Afonin², PhD (Physics&Mathematics)

On intelligent and interface tools for automated theorem-proving systems

¹Taras Shevchenko National University of Kyiv, 03680, Kyiv, Glushkov av., 4d, e-mail: foraal@mail.ru

²National University of Kyiv-Mohyla Academy, 04655, Kyiv, Skovorody vul., 2, e-mail: andrew.afomin@gmail.com

У статті обговорюються деякі питання побудови інтелектуальних та інтерфейсних засобів систем автоматизації пошуку доведень. Дається опис загального підходу до їх вирішення, концентруючи увагу на таких засобах, як формальна природна мова, дедуктивні перетворення, інформаційне середовище, взаємодія з людиною в ході розв'язання задачі, що розглядається. Формуються вимоги до них з точки зору ефективності пошуку доведень теорем та верифікації математичних текстів. Звертається увага на важливість їх розвитку у зв'язку з їх активним використанням у прикладних інформаційних системах, що призначені для перевірки властивостей протоколів, ефективного витягання та управління математичними знаннями, поповнення наявних баз даних новими фактами, генерування нових знань, добування інформації у репозиторіях, яка є необхідною для вирішення поставленого завдання. Демонструється, що інтелектуальні та інтерфейсні засоби, які були розвинуті у рамках так званого Алгоритму Очевидності, добре відповідають сучасним дослідженням в області комп'ютерних наук та автоматизації міркувань.

Ключові слова: система автоматизації пошуку доведень теорем, інтелектуальні засоби, інтерфейсні засоби, формальна мова, дедуктивні перетворення.

Some problems of the construction of intelligent and interface tools for automated theorem-proving systems are discussed in the paper. The description of a general approach to their solving is given focusing attention on such tools as a formal natural language, deductive transformations, information environment, man-machine interaction in the process of solving of a task under consideration. Requirements to them are described from the point of view of efficiency of theorem proving and mathematical text verification. An attention is drawn to the importance of their development due to their active use in applied information systems intended for the verifying of protocol properties, efficient mining and managing of mathematical data, updating of existing databases with new facts, generation of new knowledge, extracting of an information from repositories that is necessary for solving a task. It is shown that the tools being developed in the framework of the so-called Evidence Algorithm are well corresponded with current approaches and investigations in the fields of theoretical and applied computer sciences and automated reasoning.

Key Words: automated theorem-proving system, intelligent tools, interface tools, formal language, deductive transformations.

Статтю представив д.ф.-м.н., чл.-кор. НАНУ, проф. Анісімов А.В.

1. Процес ведення інтелектуальної й інтерфейсної взаємодії можна визначити як обмін даними між обчислювальною/інформаційною системою та користувачем, здійснюваний за допомогою інтерактивного терміналу за певними правилами, головним чином, за допомогою діалогу між комп'ютером і людиною (іноді,

діалогу між комп'ютером і людиною (іноді, фактично, в формі монологу з боку людини). Він може розглядатися як оболонка, що включає всі процеси, що здійснюються в системі по виконанню певних завдань. Процеси введення-виведення даних забезпечують обмін на самому верхньому рівні, на якому інтерфейсний модуль

© О.О.Лялецький, А.О.Афонін, 2014

повинен правильно інтерпретувати кожне завдання, що він отримує від людини. Його основне призначення - коректно визначити задачу, яку користувач покладає на систему; при цьому виявляються важливими правила ведення інтерфейсної взаємодії комп'ютера з людиною при обробці даних.

Розробка інтелектуального інтерфейсу спрямована на підвищення узгодження творчих можливостей системи і людини, знаходження та узгодження стилю діалогу з існуючими потребами й уявленнями людини.

2. Існують різні види діалогів, серед яких діалог, керований системою, більш зручний тим, що він краще підлаштовується під комп'ютер. Але при цьому він має більше обмежень, ніж діалог, керований користувачем на деякому формальному діалекті природної мови.

При використанні машинної штучної мови розпізнається лише невелика кількість слів і у кожний конкретний момент вхідні повідомлення складаються з дуже невеликого числа слів.

Тому у поточний час головні зусилля розробників систем автоматизації міркувань концентруються на використанні різних (формальних) фрагментів природної мови, які є найбільш зручними для користувача засобами ведення діалогу з системою. Саме використання обмеженої, а в ідеалі, і необмеженої природної мови в тій або іншій предметній області веде до найбільш ефективного вирішення задач на комп'ютері з використанням інтелектуального інтерфейсу. Тому для полегшення взаємодії користувача з системою в процесі вирішення інтелектуальних завдань на комп'ютері необхідним елементом є перехід на той чи інший (формальний) фрагмент природної мови, який враховує предметну область і специфіку вирішуваних завдань.

3. Якщо з позицій вищесказаного звернутися до розробки інтелектуальних та інтерфейсних засобів систем автоматизації міркувань, то їх важливість ще додатково посилюється складністю обробки інформації, пов'язаної з тим, що доводиться розробляти і досліджувати підходи, пов'язані з одночасним проведенням числових викладок та/або аналітичних перетворень та/або дедуктивних побудов.

Один з можливих підходів в напрямку створення систем такого типу з розвитим мовним інтерфейсом був запропонований академіком В.М. Глушковим ще на початку 1970-х років в рамках робіт з так званої програми Алгоритм Очевидності, АО (Evidence Algorithm, див.

<http://nevidal.org>). В.М.Глушков запропонував в [1] проводити одночасно дослідження:

(1) з формальних мов природного типу для подання математичних текстів у формі, найбільш зручною для користувача;

(2) з побудови програм "очевидних" дедуктивних перетворень, тобто по формалізації та еволюційному розвитку поняття машинного кроку доведення;

(3) з створення інформаційного середовища, що має вплив на поточний поняття комп'ютерного кроку доведення;

(4) з інтерактивних засобів підтримки користувачем процесу пошуку доведень.

4. В рамках робіт за алгоритмом Очевидності в кінці 1970-х - початку 1980-х рр. була реалізована російськомовна версія системи автоматизації дедукції, САД [2], а починаючи з кінця 1990 рр. активно розвивається її англійськомовна версія SAD [3], яка увібрала в себе сучасне бачення так званої *евіденціальної парадигми*. Зупинимось докладніше на її основних положеннях відносно пунктів (1) - (4).

(1) Мови повинні задовольняти наступним вимогам: вони повинні мати формальний синтаксис і формальну семантику; для одержання замкнених текстів, вони повинні забезпечувати можливість формулювання аксіом теорій, визначень, необхідних тверджень, теорем і доведень; їх тезаурус повинен бути відділений від їх граматики, що повинна бути розширюваною; вони повинні бути максимально наближені до природних мов математичних публікацій; з метою використання добре розвинутих методів пошуку доведення, вони повинні допускати можливість трансляції математичних текстів, записаних на них, у множину формул 1-го порядку.

(2) Відповідно до евіденціальної парадигми, ядро будь-якої системи пошуку доведень повинна утворювати так звана "евіденціальна процедура", яка призначена для встановлення як коректності верифікованого кроку, так і істинності твердження, що доводиться в цілому деякою дедуктивною технікою. Природно, що для досягнення цієї мети повинні бути передбачені різні методи посилення "евіденціальної процедури", зокрема: пошук допоміжних тверджень та іншої релевантної інформації, яка може бути корисною; використання можливостей систем комп'ютерної алгебри; перетворення звичайних для людини прийомів доведення у машинні; використання різноманітних методів рішення рівнянь.

(3) Евіденціальна парадигма передбачає побудову і розвиток інформаційного середовища (бази математичних знань), що містить описи властивостей, які змінюються по мірі одержання нових знань і, у кінцевому рахунку, що впливає на поняття машинної очевидності кроку доведення. Вона повинна використовувати вже існуючі засоби представлення, здобування й обробки математичних знань.

(4) Інтерфейсні засоби повинні передбачати можливість активного втручання користувача в процес пошуку доведення і полегшувати людині розуміння процесу пошуку.

Це з необхідністю приводить до проблеми створення, теоретичного дослідження та імплементації різноманітних інтелектуальних методів та інтерфейсних засобів як людини з системами автоматизації пошуку доведень теорем на формальній мові, максимально наближеної до звичайної мови людини, так різних систем та/або їх частин.

Стислий опис досліджень в цьому напрямку, що базуються на досвіді, накопиченому в ході робіт по системі SAD, представлений нижче.

5. Сказане вище приводить до того, що найбільш розумною виявляється трирівнева архітектура систем автоматизації пошуку доведень, коли на першому рівні проводиться лінгвістична обробка вихідних (формалізованих) математичних знань, яка полягає в тому, що на ньому проводиться синтаксичний аналіз введеної інформації з наступною її трансляцією у внутрішнє подання, яке в подальших рівнях підлягає логічним і математичним перетворенням. На другому рівні виконується спрощення даної задачі за допомогою її зведення до ряду підзадач простішої структури. У певні моменти часу обрана підзадача, яка задовольняє певним вимогам, піддається її обробці дедуктивними засобами з метою встановлення її вивідності, тобто встановлення її логічного слідування з обраних передумов. На третьому рівні, для поліпшення дедуктивних можливостей систем автоматизації пошуку доведень, повинне використовуватися інформаційне середовище за допомогою різноманітних інтерфейсних засобів.

6. З викладеного вище впливає, що для систем пошуку доведень першочерговим завданням з розробки інтелектуального інтерфейсу є розвиток лінгвістичних засобів.

В історичному плані розробка мовних засобів у рамках евіденціальної парадигми почалася ще на початку 1970-х років після появи Алгоритму

Очевидності в роботі В.М.Глушкова [1]. Як результат, незабаром з'явилася мова TL (Theory Language) [4] в російськомовному варіанті, яка на той час не мала аналогів у світовій практиці і була орієнтована на запис замкнених математичних текстів, що містять теорему, яка призначена для встановлення її істинності, і все необхідне для її доведення.

Після отримання системою вхідного TL-тексту, він транслювався у внутрішнє середовище для знаходження доведення виділеної теореми або для інших цілей.

Ця ж сама ідея була використана при створенні англомовної версії мови TL, яка отримала назву ForTheL (FORmal THEory Language) [5] і мала ряд особливостей, пов'язаних, головним чином, з переходом від російської граматики до англійської, що суттєво дозволило наблизитися до природного написання (англійських) текстів. Як і TL, ForTheL у первісному варіанті не мав засобів для запису доведень. Її подальший розвиток дозволив надати їй сучасного вигляду і говорити про те, що мова покриває значну частину практичних потреб у запису формальних математичних текстів.

7. Одним з основних положень евіденціальної парадигми є поняття кроку очевидності, який еволюційне розвивається, що визначається тими наявними в розпорядженні дедуктивними засобами, які використовуються для доведення тверджень [6]. На поточний час розвинутий спеціальний секвенційний формалізм для логіки першого порядку. Він дозволяє не тільки будувати різні машинні секвенційні числення для класичної логіки, але й дає підхід до побудови досить ефективних секвенційних числень для інтуїціоністської логіки.

У реалізованій версії (англомовної) системи SAD секвенційний формалізм знайшов своє вираження у вигляді логічної процедури, що представляє оригінальне секвенційне числення і дозволяє провести як верифікацію заданого математичного тексту, так і автоматичне доведення обраного твердження [7].

8. Логічна обробка інформації в системі SAD пов'язана з вирішенням різноманітних задач з доведення деяких фактів в предметних областях, головним чином, в математиці. В даний час накопичено велику кількість проблем для автоматичного або автоматизованого пошуку доведень в самих різних її розділах. З метою впорядкувати їх і дати можливість використовувати ці проблеми самим різним системам автоматизації міркувань для

проведення тестових перевірок їх обчислювальних можливостей, була створена бібліотека TPTP (Thousands of Problems for Theorem Provers, див. <http://www.tptp.org/>), яка дає надійні і прості механізми доступу до математичних знань, накопичених в ній. Тому у системі SAD було вирішене завдання її інтерфейсного підключення до TPTP, що призводить не тільки до простого використання всього математичного багатства TPTP, але також може забезпечити поповнення бібліотеки TPTP новими знаннями, отриманими в ході експлуатації системи SAD.

Висновок. Робота демонструє, що підхід до побудови інтелектуальних та інтерфейсних засобів систем пошуку доведень, розвинутий у рамках евиденціальної парадигми і в ході проектування і реалізації систем САД і SAD, показав свою спроможність не тільки тому, що у даний час він повністю відповідає існуючим трендами розвитку комп'ютерних наук та автоматизації міркувань, але ще й тому, що за допомогою таких засобів вдалося верифікувати ряд нетривіальних математичних фактів. Тому представляється перспективним подальше проведення фундаментальних і прикладних робіт по реалізації програми Алгоритм Очевидності.

Список використаних джерел

1. Глушков В. М. Некоторые проблемы теории автоматов и искусственного интеллекта / В.М. Глушков // Кибернетика. – 1970. – № 2. – С. 3-13.
2. Капитонова Ю.В. О системе обработки математических текстов / Ю.В. Капитонова, К. П. Вершинин, А. И. Дегтярев, А. П. Жежерун, А. В. Лялецкий // Кибернетика. – 1979. – № 2. – С. 48.
3. Lyaletski A. System for Automated Deduction (SAD): Linguistic and Deductive Peculiarities / A. Lyaletski, A. Degtyrev, K. Verchinine, A. Paskevich // *Advances in Soft Computing: Intelligent Information Systems 2002*. – Physica-Verlag, Springer. – 2002. – Vol. 4603. – P. 398-403.
4. Глушков В. М. О формальном языке для записи математических текстов / В.М. Глушков, К.П. Вершинин, Ю.В. Капитонова и др. // В сб.: Автоматизация поиска доказательств теорем в математике. – Киев: ИК АН УССР, 1974. – С. 3-36.
5. Vershinin K. ForTheL - the language of formal theories / K. Vershinin, A. Paskevich //

- Journal of Information Theories and Applications. – 2000. – V. 7 (3). – P. 120-126.
6. Асельдеров З.М. Эвиденциальная парадигма и обработка компьютерных математических знаний / З.М. Асельдеров, О.В. Лялецкий, Л.З. Фролова // Системні дослідження та інформаційні технології. – 2004. – № 1. – С. 7-17.
7. Lyaletski A. Theorem proving and proof verification in the system SAD / A. Lyaletski, A. Paskevich, K. Verchinine // *Lecture Notes in Computer Science*. – 2004. – V. 3119. – P. 236-250.

References

1. GLUSHKOV, V.M. (1970) Some problems of automata theory and artificial intelligence. *Cybernetics and System Analysis*. 6(2). p. 17-27.
2. KAPITONOVA, Y.V., VERSHININ, K.P., DEGTYAREV, A.I., ZHEZHERUN, A.P., and LYALETSKI, A.V. (1979) System for processing mathematical texts. *Cybernetics and System Analysis*. 15(2). p. 209-210.
3. LYALETSKI A., DEGTYREV A., VERCHININE K., and PASKEVICH A. (2002) System of Automated Deduction (SAD): Linguistic and Deductive Peculiarities. *Advances in Soft Computing: Intelligent Information Systems 2002*. 4603. p. 398-403.
4. GLUSHKOV, V.M., VERSHININ, K.P., KAPITONOVA, Y.V. et al. (1974) O formal'nom yazyke dlya zapisi matematicheskikh tekstov. *Avtomatizatsiya poiska dokazatel'stv teorem v matematike*. Kiev: IK AN UkSSR. p. 3-36.
5. VERSHININ, K. and PASKEVICH, A. (2000) ForTheL - the language of formal theories. *Journal of Information Theories and Applications*. 7(3). p. 120-126.
6. ASEL'DEROV, Z.M., LYALETSKI, A.V., FROLOVA L.Z. (2004) Evidentsial'naya paradigma i obrabotka komp'yuternykh matematicheskikh znaniy. *Sistemni doslidzhennya ta informatzijni tekhnologii*. p. 7-17.
7. LYALETSKI A., PASKEVICH A., and VERCHININE K. (2004) Theorem proving and proof verification in the system SAD. *Lecture Notes in Computer Science*. 3119. p. 236-250.

Надійшла до редколегії 22.09.2014