

**ON PROCESSING OF MATHEMATICAL TEXTS IN THE EVIDENCE
ALGORITHM**

Lyaletski A.

Taras Shevchenko National University, Kyiv, Ukraine

A short description of state-of-the-art investigations on mathematical text processing in accordance with the current vision of the so-called Evidence Algorithm is given. Attention is drawn to language means and deductive methods – the most important components of systems for computer support of man's mathematical activity.

**ПРО РОБОТУ З МАТЕМАТИЧНИМИ ТЕКСТАМИ В АЛГОРИТМІ
ОЧЕВИДНОСТІ**

О.В.Лялецький

Київський національний університет імені Тараса Шевченка,
Україна

Дається короткий опис поточного стану робіт по обробці математичних текстів у відповідності з сучасним баченням так званого Алгоритму Очевидності. Увага концентрується на мовних засобах та дедуктивних методах – найбільш важливих компонентів систем комп'ютерної підтримки математичної діяльності людини.

На початку 1970-х років академіком В.М.Глушковим була запропонована програма з автоматизованої обробки математичних текстів під назвою Алгоритма Очевидності (АО) [1]. Ця програма була орієнтована на комп'ютерну підтримку математичної діяльності людини як при проведенні академічних досліджень, так і при навчанні математичних дисциплін.

Відповідно до програми АО, ядром системи автоматизованого пошуку доведення мусить бути процедура встановлення очевидності довідного (верифікуємого) твердження в термінах деякого дедуктивного механізму. Таку процедуру необхідно посилити додатковими засобами: мати можливості пошуку допоміжних тверджень чи якої-небудь іншої релевантної інформації, застосовувати алгебраїчні перетворення (або, у сучасній термінології, системи комп'ютерної алгебри), використовувати звичні для людини прийоми доведення та ін. Весь цей

інструментарій мусить мати можливість обмінюватися даними за допомогою деякої формалізованої мови.

До мови були сформульовані наступні вимоги. Синтаксис і семантика такої мови повинні бути формалізовані. Вона мусить допускати формулювання аксіом теорії, теорем та їх доведень для забезпечення цілісності текстів і формулювання визначень. Для того, щоб тезаурус мови можна було поповнювати, він повинен бути відокремлений від граматики мови. Крім того, що дуже важливо, мова повинна бути близькою до мови математичних публікацій, щоб забезпечити користувачу зручність при створенні й обробці тексту в інтерактивному режимі. Вона також повинна дозволяти формулювати задачі, що не зв'язані безпосередньо з пошуком виведення.

Дедуктивні методи повинні були розвиватися одночасно з мовою у відповідності з наступними вимогами: синтаксична форма початкової задачі повинна бути збережена; дедуктивні побудови повинні проводитися в сигнатурі початкової теорії; пошук доведення повинен бути цілеорієнтованим; обробка рівностей (розв'язання рівнянь) повинна бути відокремлена від дедуктивних процесів.

Таким чином, програма АО була зорієнтована на побудову різних складових, необхідних для автоматизованої обробки математичних текстів і спираючихся на одночасне проведення досліджень у напрямку розробки формалізованих мов для запису математичних текстів у найсприятливішій для користувача формі, в напрямку еволюційного розвитку поняття машинного (дедуктивного та індуктивного) кроку доведення, у напрямку створення інформаційного середовища ЕА, яке впливає на очевидність машинного кроку доведення, і в напрямку побудови засобів інтерактивної обробки математичних текстів.

Дослідження з АО можна поділити на два основних історичних етапи: перший – з 1970 р. по 1990 р., коли була спроектована, імплементована на ЄС ЕВМ та апробована російськомовна система автоматизації дедукції САД, і другий – з 1998 р. по поточний час, коли була введена в експлуатацію її англomовна версія, яка використовує сучасні досягнення в галузі автоматизації міркувань. Деякі дрібниці відносно цих етапів можна знайти в [2]. Нижче

дається характеристика особливостей мовних засобів і дедуктивних методів англомовної системи САД.

Мовні засоби. У рамках АО, перший (російськомовний) варіант мови під назвою TL для запису математичних текстів був розроблений в першій половині 1970 г. Його англійська версія ForTheL була анонсована в 2000 р. в роботі [3] і с того часу вона піддавалася послідовному розвитку¹.

З погляду синтаксису ForTheL-текст являє собою сукупність розділів. Кожен розділ, у свою чергу, може містити розділи більш низького рівня і/чи фрази. Границі розділів відзначаються роздільниками, що визначає відношення часткового порядку на множині розділів даного тексту. Синтаксис роздільників може бути обраний у достатньому ступені довільно, але перевага віддана звичайної практики нумерації, що прийнята в математичних текстах (див., наприклад, стиль LNCS для LaTeX).

Деякі розділи, такі, як теореми, доведення, визначення відіграють у мові особливу роль. Вони відрізняються спеціальним синтаксисом роздільників, що також відповідають звичайній математичній практиці.

Розділ, що містить тільки фрази, називається простим. Кожна фраза або стверджує, або допускає дещо. Відмінність між твердженнями і припущеннями виражається за допомогою простих синтаксичних засобів.

В поточний час мова ForTheL оснащена лінгвістичними засобами для запису доведень, що перетворює систему САД в зручний верифікатор математичних текстів. Крім цього, мається можливість виконувати пошук доведень теорем і встановлення істинності першопорядкових формул в автоматичному режимі.

Семантика ForTheL-тексту визначається тією задачею, яку формулює користувач. От деякі приклади таких задач: «перевірити коректність даного тексту», «перевірити коректність речення з даного тексту», «знайти контрприклад до твердження з даного тексту», «знайти в тексті інформацію, релевантну ...» і т.п.

¹ Детальний опис поточної мови ForTheL можна знайти на сайті системи САД “<http://nevidal.org>”.

Дедуктивні побудови. Однією з особливостей дедукції в стилі АО є те, що пошук доведення виконується в середовищі цілісного ForTheL1-тексту, що є результатом перетворення початкового ForTheL-тексту. Оскільки ForTheL1-текст є структурованим, тобто містить визначення, допоміжні твердження і теорему, яку треба довести, для системи САД був запропонований секвенційний підхід пошуку виведення формул першого порядку [4].

Зауважимо, що перша робота [5] про пошук доведення в стилі АО з'явилася в 1966 р. У ній описувалася евристична процедура пошуку доведення в теорії груп, що допускала використання на формальному рівні методів пошуку доведень, що зустрічаються в математичних публікаціях. Згодом ця формальна техніка була поширена на деякий фрагмент теорії множин, а своє завершення вона отримала у вигляді спеціального числення, що призначалося для перевірки істинності формул класичної логіки першого порядку. Подальший його розвиток викликало появу в роботі [6] першого представника сімейства секвенційних числень у стилі АО, взятого за основу при реалізації першої версії логічного апарата системи САД.

Починаючи з 1999 р. був проведений ряд досліджень з поліпшення ефективності пошуку виведення в секвенційних численнях (див., наприклад, [7,8]), які дозволяють виконувати зведення (редукцію) довідної «цілі» до сукупності допоміжних «підцілей»; ефективну роботу з кванторами (без попередньої сколемізації) з метою пошуку доведення в сигнатурі початкової теорії; розв'язок «рівнянь», породжених на стадії редукції цілі до підцілей.

Це дало можливість не тільки побудувати оригінальні логічний апарат і лінгвістичні засоби системи САД, а і спробувати використати їх при проектуванні інших систем пошуку доведень і автоматизації міркувань, задовольняючих вимогам до дедукції, які були сформульовані вище при опису програми АО.

Також треба відмітити, що поточна система САД спроектована таким чином, який дає змогу «підключати» до неї різні зовнішні відносно САД прувери, наприклад, Vampire, SPASS і Otter [9].

Завершуючи опис характерних рис обробки математичних текстів в стилі АО, відмітимо, що у довгострокової перспективі

подальші дослідження у напрямку реалізації програми АО можуть вплинути на створення потужної комп'ютерної інфраструктури для надання допомоги людині у проведенні різноманітної обробки формалізованих математичних текстів, а також можуть виявитися корисними у створенні засобів автоматизованого навчання і тестування знань навчасного.

Література:

1. Глушков В. М. Некоторые проблемы теории автоматов и искусственного интеллекта // Кибернетика. – № 2. – 1970. – С. 3–13.

2. Lyaletski A.V. and Verchinine K. Evidence Algorithm and System for Automated Deduction: A retrospective view // Lecture Notes in Computer Science: Proceedings of the 9th International Conference AISC/MKM/Calculus 2010. – Vol. 6167. – 2010. – P. 411–426.

3. Vershinin K. and Paskevich A. ForTheL – the language of formal theories // International Journal of Information Theories and Applications. – 7 (3). – 2000. – P. 120–126.

4. Лялецкий А.В. Секвенциальный формализм и дедуктивные системы для классической логики первого порядка // Proceedings of the International Conference «Logic and Applications». – Новосибирск. – 2000.

5. Об одном алгоритме поиска доказательств теорем в теории групп / Ануфриев Ф.В., Федюрко В.В., Летичевский А.Л., Асельдеров З.М., Дидух И.И. // Кибернетика. – 1966. – № 1. – С. 23–29.

6. Дегтярев А.И., Лялецкий А.В. Логический вывод в Системе Автоматизации Доказательств (САД) // Математические основы систем искусственного интеллекта. – Киев: ИК АН УССР, 1981. – С. 3–11.

7. Degtyarev A., Lyaletski A., and Morokhovets M. Evidence Algorithm and sequent logical inference search. // Lecture Notes in Artificial Intelligence. – Vol. 1705. – 1999. – P. 44–61.

8. Лялецкий А.В. Эвиденциальная парадигма: логический аспект // Кибернетика и системный анализ. – 2003. – № 5. – С. 37–47.

9. Verchinine K., Lyaletski A.V., and Andrey Paskevich. System for Automated Deduction (SAD): A tool for proof verification // Lecture Notes in Computer Science: Proceedings of 21st International

Conference on Automated Deduction (CADE 2007). – Vol. 4603. – 2007. – P. 398 – 403.