

І. СЕРГІЄНКО, А. ГУПАЛ

СУЧАСНА МАТЕМАТИКА — ПОЄДНАННЯ ДЕДУКТИВНОГО ТА ІНДУКТИВНОГО ПІДХОДІВ

Автори продовжують обговорювати особливості дедуктивного та індуктивного підходів у математиці і природничих науках (див. попередню публікацію: Вісн. НАНУ. — 2002. — № 5. — С. 19—25). Донедавна математичний апарат розвивався переважно на основі дедуктивного підходу, сфера застосування якого має свої межі. Цілковито відмінним від дедуктивного є індуктивний підхід. Він, як і квантові обчислення, будучи ймовірнісним, адекватно описує результати експериментів у квантовій механіці. А проте, варто проаналізувати ситуацію в математиці, аби переконатися, що саме взаємодоповнення дедукції та індукції істотно розширює можливості цієї науки.

ОБМЕЖЕНІСТЬ ТРАДИЦІЙНОГО ПІДХОДУ

Про кризову ситуацію у математиці йдеться давно. Що ж до нинішнього її стану, то варто навести доволі гостре критичне висловлювання академіка РАН В.І. Арнольда: «Математика — це частина фізики, а зовсім не наука про переливання з пустого в порожнє. <...> Оскільки ані для викладання, ані для застосування в якихось інших науках схоластична, відрізана від фізики, математика не придатна, результатом виявилася загальна неприязнь до математиків — і з боку нещасних школярів, і з боку користувачів. І математика, і фізика — експериментальні науки, різниця лише в тому, що у фізиці експерименти коштують мільярди доларів, а в математиці — одиниці рублів» [1].

Аналізуючи причини такого кризового становища, фахівці дедалі частіше вказують на особливості та обмеження панівного в математиці дедуктивного підходу. Водночас добре відомо, що протягом тривалого часу в усіх природничих науках дуже плідно працює індуктивний підхід, за допомогою якого отримують висновки щодо властивостей об'єктів та явищ реального світу на основі скінченного числа спостережень за ними.

Проте у математиці індуктивний підхід широко не застосовувався, оскільки донедавна він не мав переконливого обґрунтування. Однак тепер, після того, як це обґрунтування запропоновано [2], цілком доречно вважати індукцію та дедукцію в математиці двома самостійними підходами, які органічно доповнюють один одного.

Традиційний математичний апарат розвивався переважно на основі дедуктивного підходу, який має певні межі застосування. Коротко означимо особливості та обмеження цього підходу. Це, зокрема, строге доведення та отримання точного розв'язку; робота тільки з одним класом тверджень — істинними (хибні твердження в дедуктивних правилах не використовуються); лінійний та послідовний ланцюжок доведення. Причому, доводячи конкретну теорему, неможливо заздалегідь вказати довжину цього ланцюжка. Скажімо, славнозвісна теорема Ферма бентежила розум математиків протягом трьох століть і нарешті була розв'язана Е. Уайлзом у 1995 р. Дедуктивні правила виводу мають високу складність та низьку ефективність. Так, алгоритм Британського музею, який послідовно генерує різноманітні виводи, може бути ефективнішим, ніж відомий метод резолюцій [3].

Продовжуючи перелік обмежень, віднесемо до них неповноту аксіоматичних систем, доведену теоремою Геделя, у зв'язку з чим не працює закон виключеного третього та спосіб доведення від протилежного. Крім того, у 70-і роки минулого століття була побудована теорія складності перебірних та комбінаторних задач. З'ясувалося, що відомі та важливі задачі обчислювальної логіки, розписів, теорії чисел, знаходження маршрутів тощо належать до важкорозв'язуваних NP-повних задач. Для жодної з NP-повних задач не вдалося побудувати поліноміальний алгоритм від розмірів вхідних даних. Таким чином, на сучасних комп'ютерах NP-повні задачі розв'язуються тільки для невеликих розмірів входу. Для того, щоб упоратися з NP-повними задачами, вчені почали розвивати інші схеми організації обчислень, які відрізняються від схем у звичайних комп'ютерах.

КВАНТОВІ ТА БІОЛОГІЧНІ КОМП'ЮТЕРИ

Ідею створення квантового комп'ютера вперше висунув лауреат Нобелівської премії фізик Р. Фейнман. На сьогодні отримано вражаючі результати. У 1994 р. П. Шор створив алгоритм факторизації, тобто визначення простих множин великих n -розмірних чисел. На класичному комп'ютері для цього потрібно виконати експоненціально велике число операцій. Той факт, що ця задача непосильна для сучасних комп'ютерів, використовується для шифрування секретної інформації у RSA-криптосистемах. П. Шор показав, що квантовий комп'ютер здатний розв'язати цю задачу за n^3 операцій. Коефіцієнт прискорення задачі при великих n , порівняно з відомими алгоритмами, може бути дуже великим. Водночас встановлено, що багато алгоритмів, які добре виконуються звичайними комп'ютерами, не прискорюються на квантових. Прискорення процесу розв'язування задач на квантових комп'ютерах перебуває у квантовій природі кубітів, які створюють регістр комп'ютера. Квантовість кубітів спричиняє деякі феномени.

По-перше, Гільбертовий простір станів квантової системи з n кубітів має величезну розмірність, яка дорівнює 2^n . Фізично це означає, що система має 2^n базових станів, і стан комп'ютера описується суперпозицією з цих 2^n базових станів. Якщо впливати на якийсь кубіт, то одночасно змінюються всі 2^n базових станів. Цей феномен названо квантовим паралелізмом.

По-друге, обчислювальний процес має ймовірнісну природу і нагадує характер інтерференції у фізиці. Амплітуда базових станів є комплексною величиною, а її модуль визначає ймовірність. Квантовий комп'ютер при цьому розглядається як складний інтерференційний пристрій, в якому інтерференція станів створює обчислювальну потужність комп'ютера. Вчені стверджують, що квантові комп'ютери у XXI столітті не замінять, а доповнюватимуть існуючий комп'ютерний світ. Їх варто використовувати в тому випадку, коли вони дають велике прискорення розв'язання задачі [4].

У 1994 р. Л.М. Адлеман показав, що за допомогою біологічних експериментів з ДНК-молекулами можливо розв'язати NP-повну задачу визначення шляхів на графі. На думку вченого, для цього потрібно 10^{19} – 10^{20} ниток ДНК, які містять послідовності нуклеотидів. Ланцюжками нуклеотидів можна маніпулювати за допомогою цілого ряду операцій, вироблених молекулярною біологією: об'єднувати нитки ДНК, розщеплювати їх у заданих місцях, копіювати, виділяти нитки із заданими послідовностями нуклеотидів і т. ін. Хоч окрема операція є повільною — триває хвилини порівняно з мікросекундами на електронних комп'ютерах, проте вона виконується для всіх ДНК одразу. Використовуючи таку одночасність виконання хімічних реакцій, можна поставити завдання створити суперкомп'ютер, який виконує мільйони операцій за секунду.

Робота Адлемана викликала бум у галузі створення ДНК-комп'ютерів. Через півроку Р. Ліптону вдалося розв'язати дуже важливу в обчислювальній логіці NP-повну задачу щодо виконуваності кон'юнктивних нормальних форм [5]. У наступних роботах було знайдено розв'язок інших видів NP-повних задач.

ПОЛІНОМІАЛЬНІ ПРОЦЕДУРИ ІНДУКТИВНОГО ВИВОДУ

Процес індуктивного виводу нагадує процес квантових обчислень, але на відміну від останнього, він виконується на звичайних комп'ютерах. З другого боку, індуктивний підхід дуже відрізняється від дедуктивного [2]. Треба побудувати таку процедуру індуктивного виводу, котра б давала змогу оцінити наслідки результату будь-якого експерименту на основі скінченного числа спостережень та наслідків попередніх експериментів. Індуктивні процедури, як і квантові обчислення, є ймовірнісними. Індуктивний підхід цілком відповідає результатам експериментів у квантовій механіці. Адже у ній ідеться тільки про ймовірність кожного можливого наслідку конкретного експерименту, здійснюваного над системою. Якщо ця ймовірність дорівнює одиниці, то результат експерименту можна передбачити, якщо ж вона дорівнює нулю, то це неможливо. Але ймовірність може перебувати у проміжку між нулем та одиницею, і тоді заздалегідь неможливо прогнозувати наслідок окремого експерименту, лишається тільки прогнозувати середнє число результатів, якщо той самий експеримент здійснюється над великим числом ідентичних систем.

Байєсівські процедури індуктивного виводу відзначаються чудовими особливостями: вони оптимальні й мають поліноміальні оцінки похибки від входу задачі. Верхня та нижня оцінки похибки у булевому випадку збігаються з точністю до абсолютної константи з виразом:

$$\sqrt{\frac{n}{\min(m_0, m_1)}},$$

де n — кількість вимірювань (ознак) спостережуваного об'єкта, а m_0 і m_1 — розміри класів у навчальній вибірці [2]. За відсутності одного з класів у навчальній вибірці кожна процедура індуктивного виводу працює непередбачено погано, але її оцінка похибки строго позитивна. Як уже зазначалося, робота лише з одним класом істинних тверджень спричиняє неповноту аксіоматичних систем. Поліноміальна оцінка похибки отримана завдяки тому, що, як і у квантових обчисленнях, вона є суперницею ймовірностей величезної кількості об'єктів (2^{n+1}) та навчальних вибірок (2^{nm}). Оптимальність байєсівських процедур притаманна таким структурам, як ланцюги Маркова і байєсівські мережі.

ЛАНЦЮГИ МАРКОВА

На початку XX ст. А.А. Марков досліджував літературні тексти і вивів закономірність: якщо розглядати літери довільного тексту, то ймовірності літери на позначення голосного та приголосного звуку залежать від однієї або двох попередніх. Ланцюг Маркова є лінійно-впорядкованою структурою:

$$x_1 \longrightarrow x_2 \longrightarrow \dots \longrightarrow x_n.$$

Поведінка випадкової послідовності визначається початковим розподілом $P(x_1)$ і перехідними ймовірностями $P(x_i | x_{i-1})$. Ланцюг Маркова — проста та економна модель

дослідження поведінки об'єктів із залежними ознаками. Для повного опису ланцюга у булевому випадку потрібно лише задати три ймовірності чи параметри для стаціонарних перехідних імовірностей. У процедурах індуктивного виводу вони замінюються емпіричними частотами, створеними на основі навчальних вибірок.

Відомий спеціаліст у галузі математичної статистики Т. Андерсон зробив детальний аналіз ланцюгів Маркова. Спираючись на його результати, можна продемонструвати, що байєсівська процедура оптимально працює на ланцюгах Маркова, а її оцінка похибки визначається згаданою формулою. Зазначимо цікаву властивість стаціонарних ланцюгів: виявляється, оцінка похибки тут зовсім не залежить від кількості ознак спостережуваного об'єкта.

Сучасна ера у молекулярній біології почалася з відкриття у 1953 р. Уотсоном і Кріком знаменитої подвійної спіралі ДНК. Ця революційна подія дала змогу отримувати великий обсяг даних прямим зчитуванням ДНК з різних ділянок геномів. Сьогодні відомо понад 100 мільйонів нуклеотидів, які можна використовувати для аналізу білків. Виявлення геномної ДНК уможливило здійснення цілої низки фундаментальних біологічних відкриттів.

Повна послідовність вірусу імунодефіциту людини стала відома одразу після його ідентифікації. Хоча молекулярна біологія — в основному експериментальна наука, але останнім часом з нею пов'язаний широкий розвиток математичних, статистичних і комп'ютерних методів. Так, аналіз послідовностей ДНК успішно здійснюється на основі методології ланцюгів Маркова [6].

БАЙЄСІВСЬКІ МЕРЕЖІ

Моделювання на основі байєсівських мереж сьогодні широко застосовуються у різних сферах: діагностиці, прогнозуванні, розпізнаванні образів, в експертних системах. Коли йдеться про аналіз даних, байєсівський підхід має суттєві переваги над багатьма іншими підходами — такими, як дерева рішень, нейронні мережі, класифікації, регресії, кластеризації тощо. У байєсівських мережах досліджуються причинно-наслідкові зв'язки між змінними (ознаками) об'єкта.

Фахівці у галузі аналізу даних нерідко зустрічаються з проблемами, пов'язаними з браком необхідної інформації та труднощами у проведенні натурних експериментів. Проте байєсівські методи не потребують окремої обробки і тестування даних — апріорна інформація щодо них вдало поєднується з добре розвинутими методами аналізу.

Байєсівська мережа є орієнтованим ациклічним графом. Спостережуваним змінним відповідають вершини графа, які з'єднуються дугами. Наявність дуг між вершинами визначає причинно-наслідкові зв'язки між змінними. Кожна змінна має деяку множину безпосередніх причин або батьківських вершин. Відсутність дуг між вершинами виражає умовну незалежність змінних. Ланцюг Маркова — окремий випадок байєсівської мережі.

Іншими словами, байєсівська мережа будується за принципами ймовірнісних переходів, властивих ланцюгам Маркова та умовній незалежності ознак. Отож, і в мережах байєсівські процедури працюють оптимально.

ДИСКУСІЯ

Наш мозок з дивовижною швидкістю інтерпретує неточну інформацію, що надходить від органів чуття: вирізняє шурхіт у гамірній кімнаті, обличчя у напівтемному провулку, розкриває прихований зміст політичної заяви. А найбільше вражає те, що він навчається (без будь-яких вказівок) створювати внутрішні уявлення, завдяки чому й виявляються всі ці здібності. Досі невідомо, в який спосіб мозок навчається переробляти інформацію, але щодо цього існує багато теорій. Аби їх перевірити, нейробіологи створюють мережі зі штучних нейронів для моделювання процесів навчання, що відбуваються в мозку [7]. Для побудови нейронної мережі потрібно обрати принцип, за яким нейрони з'єднуюватимуться у ній між собою, і належним чином визначити вагові параметри на цих зв'язках. Зауважимо відразу, що ці вагові параметри— не що інше, як перехідні ймовірності в ланцюгах Маркова та байєсівських мережах. Вплив одного елемента на інший залежить від встановленого з'єднання. Вага визначає силу впливу.

Найпоширеніший тип штучної нейронної мережі складається з трьох шарів елементів: шару вхідних елементів, який з'єднується з шаром «прихованих» елементів, а той, у свою чергу, сполучається з шаром вихідних елементів. Активність вхідних елементів — це початкова інформація, яка подається в мережу. Активність кожного прихованого елемента визначається сигналами від вхідних елементів та вагою на зв'язках між вхідними і прихованими елементами. Аналогічно поведінка вихідних елементів залежить від активності прихованих елементів і ваги на зв'язках між прихованими та вихідними елементами. На нашу думку, отримані результати щодо оптимальності байєсівських процедур можуть допомогти вченим зрозуміти, які процеси навчання насправді відбуваються в мозку.

Чимало дослідників цієї проблеми, зокрема відомий фізик-теоретик Р. Пенроуз, упевнені, що пізнавальна діяльність людини не має у своїй основі жодних дедуктивних правил і тому не підкоряється обмеженням Геделя. Цю точку зору підтвердили дослідження, проведені вченими різних спеціальностей [8]. Їх результати засвідчили, що творчі можливості людини — і в мистецтві, і в галузі природничих наук — жодною мірою не зазнають жорстких обмежень, властивих комп'ютерним обчисленням. Пенроуз та інші теоретики вважають, що творчі здібності людини існують завдяки не з'ясованим досі механізмам та правилам, які, скоріш за все, нагадують квантово-механічні. Для нас такі висновки не є приголомшливими. Маємо всі підстави вважати, що мозок людини у процесі навчання та розпізнавання використовує індуктивні процедури. Це повністю узгоджується з квантово-механічними уявленнями.

Відомо, що генетична інформація зберігається у вигляді послідовності нуклеотидів у молекулах ДНК. ДНК можна розбити на неперервні частини — гени, на кожному з яких записана послідовність амінокислот одного білка. Донедавна вченим не вдавалося дослідити детальну структуру генів вищих організмів. Це стало можливим лише з появою генної інженерії і після розробки методів зчитування текстів ДНК. Виявилось, що гени вищих організмів не є неперервними, а складаються з окремих фрагментів, роз'єднаних іншими послідовностями нуклеотидів [9]. ДНК раптом постала у вигляді «перемішаної структури» з генів, порізаних на окремі частини. Проміжки між генами різні — від 10 до 20000 пар основ.

У водному середовищі білок швидко створює складну специфічну тривимірну структуру, яка визначає певну його функцію в організмі. Підраховано, що для розв'язання проблеми синтезу білків суперкомп'ютеру потрібно 10^{127} років для досить коротких послідовностей, які складаються зі 100 амінокислот. До речі, ця цифра набагато перевищує час існування Всесвіту. Нещодавно було з'ясовано, що дана проблема належить до класу NP-повних задач [8]. Цілком імовірно, що невідомі ділянки в молекулах ДНК відповідають

«негативним» білкам, тобто таким, які або гинуть у водному розчині, або не виконують позитивних функцій у живому організмі. Такі білки можна назвати результатом невдалих експериментів. Отож, маємо підстави зробити висновок, що у природі механізм «виготовлення» білків розгортався як процес навчання з урахуванням як позитивних, так і негативних експериментів. Ця гіпотеза ґрунтується на тому, що індуктивні процедури навчання спрацьовують дуже швидко і мають поліноміальні оцінки.

Свого часу фізик Л. де Бройль дійшов висновку, що частки дифрагують як хвилі. Це відкриття було надзвичайно важливим для теоретичної фізики. Адже всі дисципліни у квантовій механіці ґрунтуються на хвильовому описі матерії. Де Бройль висунув принцип дуальності частинка—хвиля, Ейнштейн — принцип дуальності маса—енергія. Математика, як і будь-яка інша сфера людської діяльності, часто потребує використання різних методів пізнання. А тому розгляд дедукції та індукції як взаємодоповнюючих підходів у сучасній математиці має стати плідним і перспективним.

1. Арнольд В.И. Математическая дуэль вокруг Бурбаки // Вестн. РАН — 2002. — № 3.
2. Сергієнко І.В., Гупал А.М. Індуктивна математика // Вісн. НАН України. — 2002. — № 5.
3. Чень Ч., Ли Р. Математическая логика и автоматическое доказательство теорем. — М.: Наука, 1983.
4. Валиев К.А. Квантовая информатика: компьютеры, связь и криптография // Вестн. РАН. — 2000. — № 8.
5. Lipton R.J. DNA solution of hard computational problems // Science. — 1995. — Vol. 268. — 28 april.
6. Математические методы для анализа последовательностей ДНК / Под ред. М.С. Уотермена. — М.: Мир, 1999.
7. Хилтон Д.Е. Как обучаются нейронные сети // В мире науки. — 1992. — № 11—12.
8. Casti J. Confronting science's logical limits // Scientific American. — October, 1996.
9. Франк-Каменецкий М.Д. Самая главная молекула. — М.: Наука, 1988.

І. Сергієнко, А. Гупал

СУЧАСНА МАТЕМАТИКА — ПОЄДНАННЯ ДЕДУКТИВНОГО ТА ІНДУКТИВНОГО ПІДХОДІВ

Резюме

Оскільки процедури дедуктивного виводу не дають змоги розв'язувати важливу категорію NP-повних задач, нині розвиваються інші схеми організації обчислень, які виконуються на ДНК- і квантових комп'ютерах. Таким схемам притаманний високий паралелізм обчислень, завдяки чому можливе успішне розв'язування NP-повних задач. Поліноміальність індуктивних процедур, які дуже нагадують квантові обчислення,

отримана завдяки тому, що оцінка похибки розглядається як суперпозиція ймовірностей величезної кількості об'єктів і навчальних вибірок.

I. Serhienko, A. Hupal

MODERN MATHEMATICS — THE INTEGRATION OF DEDUCTIVE AND INDUCTIVE APPROACHES

Summary

The deductive procedures don't solve the very important NP-complete problems. Therefore another schemes of calculations were presented. These schemes are realized on quantum and DNA computing. Quantum and biological computations could potentially have vastly more parallelism than conventional ones and solve many famous NP-complete problem. Inductive procedures are very similar to quantum calculations. The estimation error is a superposition of probabilities of vast number of objects and learning samples.

© СЕРГІЄНКО Іван Васильович. Академік НАНУ. Директор Інституту кібернетики ім. В.М. Глушкова НАН України (Київ).

ГУПАЛ Анатолій Михайлович. Доктор фізико-математичних наук. Директор Науково-навчального центру прикладної інформатики НАНУ (Київ). 2003.