Подборка алгоритмов, которые правят миром

Подготовил: ученик 9 класса Матюшкин С.Е.

- Однажды на Reddit засветился интересный пост, написанный Джорджем Дворским под названием «10 алгоритмов, которые правят миром» (статья на русском языке). В нём автор попытался объяснить важность алгоритмов в наши дни и составил список самых важных для современной жизни алгоритмов.

- Если вы изучали алгоритмы, первое, что может прийти вам на ум после прочтения того поста: «Автор вообще знает, что такое алгоритм?» — или, может быть: «Новостной канал Facebook — алгоритм?» — потому что если новостной канал Facebook является алгоритмом, то в конечном итоге мы можем почти всё классифицировать как алгоритм. Чтобы внести ясность, в этой статье будет дано определение алгоритму, и будет составлен список из 10 (или больше) алгоритмов, которые на самом деле правят нашим миром.

Что такое алгоритм?

- Неофициально алгоритмом является любая корректно определённая вычислительная процедура, на вход которой подается некоторая величина или набор величин, и результатом выполнения которой является выходная величина или набор значений. Таким образом, алгоритм представляет собой последовательность вычислительных шагов, преобразующую входные данные в выходные.

- Томас Х. Кормен, Чарльз И. Лейзерсон (2009), Алгоритмы: построение и анализ, 3-е издание

- Проще говоря, алгоритм представляет собой последовательность шагов, которая позволяет решить определённую задачу (да, алгоритмы используются не только машинами, но и людьми). Алгоритм должен обладать тремя важными характеристиками, чтобы иметь право так называться:

- Он должен работать за конечное количество времени. Если ваш алгоритм не может разобраться с проблемой, для которой он был создан, за конечное количество времени, то он бесполезен.
- Он должен иметь чётко определённые инструкции. Каждый шаг алгоритма должен быть точно определён. Инструкции должны быть однозначны для каждого случая.
- Он должен быть пригодным к использованию. Алгоритм должен решать проблему, для решения которой он был написан. Должна быть возможность продемонстрировать его работу при наличии только карандаша и бумаги.
- Также важно отметить, что алгоритмы используются не только в информатике, но и в математике. По факту, самые ранние математические алгоритмы — разложение на простые множители и извлечение квадратного корня — использовались вавилонянами уже в 1600 г. до н. э. Таким образом, мы имеем проблему, связанную с упомянутой ранее записью, которая рассматривает алгоритмы как компьютерные сущности. Однако, если применить формальный смысл слова, то одним из претендентов в десятку лучших алгоритмов может стать любая арифметическая операция (сложение, вычитание, произведение и т. д.).

- Поэтому в этой статье мы будем говорить о компьютерных алгоритмах. Тогда остаётся вопрос: какие 10 (или больше) алгоритмов правят нашим миром?

Алгоритмы сортировки (быстрая, пирамидальная, слиянием)

- Какой алгоритм сортировки элементов лучший? Всё зависит от ваших нужд, и этим объясняется выбор именно трёх наиболее часто используемых алгоритмов. Возможно, вы предпочитаете какой-то из них другим, но все три имеют одинаковое значение.

- Сортировка слиянием — один из наиболее важных алгоритмов на сегодняшний день. Он базируется на сравнении элементов и использует подход «разделяй и властвуй» для более эффективного решения проблемы, которая когда-то решалась за время O (n^2). Алгоритм был изобретён математиком Джоном фон Нейманом в 1945 году.

- Подробнее об оценке сложности алгоритмов читайте в нашем материале

- Быстрая сортировка — другой подход к сортировке, чей алгоритм может базироваться как на in-place разделении, так и на «разделяй и властвуй». Проблема этой сортировки заключается в том, что она не является стабильной, но эффективна для сортировки массивов в оперативной памяти.

- Пирамидальная сортировка — in-place алгоритм, использующий приоритетную очередь, за счёт которой уменьшается время поиска данных. Неустойчив.

- Эти алгоритмы намного лучше по сравнению с другими подходами (например, сортировкой пузырьком), использовавшимися ранее. По сути, благодаря им у нас есть глубинный анализ данных, искусственный интеллект, анализ связей и большинство вычислительных инструментов в мире, включая интернет.

Преобразование Фурье и Быстрое преобразование Фурье

- Интернет, Wi-Fi, смартфон, телефон, компьютер, маршрутизатор, спутники — почти всё, что имеет внутри себя электронно-вычислительное устройство, так или иначе использует эти алгоритмы для функционирования. Вы не сможете получить диплом по электронике, вычислительной технике или телекоммуникации без изучения этих важных алгоритмов.

Алгоритм Дейкстры

- Как ни странно, интернет не смог бы работать эффективно без существования этого алгоритма. Он используется в задачах, где проблему можно представить в виде графа, для поиска кратчайшего пути между двумя узлами.

- Сегодня, даже когда для поиска кратчайшего пути у нас есть решения получше, алгоритм Дейкстры используется в системах, требующих стабильности.

Алгоритм RSA

- Интернет не был бы так глубоко интегрирован в нашу жизнь, как в наше время, если бы не криптография и кибербезопасность. Вы можете сказать: «Конечно, безопасность в эпоху спецслужб» — или: «Вы должны быть очень наивны, чтобы считать, что в сети вы в безопасности». Однако, чтобы тратить свои деньги, люди должны чувствовать себя в безопасности. В конце концов, вы не станете вводить номер своей кредитной карты в веб—сервисе, если не будете уверены в его надёжности.

- Из криптографии пришёл алгоритм, который остаётся одним из самых важных в мире, — алгоритм RSA. Разработанный основателями компании RSA, этот алгоритм сделал криптографию доступной для всех в мире и предопределил её будущее. Алгоритм RSA создан для простой задачи с неочевидным решением: как делиться открытыми ключами между независимыми платформами и конечными пользователями так, чтобы была возможность использовать шифрование (стоит отметить, что на самом деле эта проблема до сих пор решена не полностью).

Алгоритм безопасного хэширования

- Это не совсем алгоритм, а скорее семейство криптографических хэш-функций (SHA-1, SHA-2, и т.д.), разработанных Национальным институтом стандартов и технологий в США. Оно имеет основополагающее значение для функционирования всего мира. Магазины приложений, антивирусы, электронная почта, браузеры и т. д. — все они используют эти алгоритмы (на самом деле — хэш, который является результатом их работы), чтобы определить, загрузили ли вы то, что хотели, а также не стали ли вы жертвой атаки «человек посередине» или фишинга.

Алгоритм факторизации целых чисел

- Сложность, связанная с разложением числа на простые множители, широко используется в машинной области. Без неё криптография могла бы быть не такой безопасной, какой мы её знаем.

- Многие криптографические протоколы — например, RSA — основаны на сложности факторизации больших составных целых чисел или на сходных проблемах. Алгоритм, который эффективно сможет разбивать на простые множители произвольное целое число, сделает криптографию с открытым ключом на основе RSA небезопасной.

Анализ связей

- В информационную эру взаимоотношения между различными субъектами имеют большое значение. От поисковых систем и социальных сетей до инструментов анализа рынка — каждый пытается найти реальную структуру интернета на все времена.

- Алгоритм анализа связей за время своего существования оброс большим количеством мифов и заблуждений среди широкой публики. Проблема заключается в существовании различных способов анализа ссылок с различными характеристиками, что делает каждый алгоритм немного другим (и позволяет патентовать алгоритмы), хотя при этом их основания похожи.

- Идея анализа связей проста: вы можете представить график в виде матрицы, что сводит задачу к проблеме собственной значимости каждого узла. Такой подход к структуре графа даёт нам возможность оценить относительную важность каждого объекта, включённого в систему. Алгоритм был разработан в 1976 году Габриэлем Пински и Фрэнсисом Нарином.

- Где используется этот алгоритм? При ранжировании страниц во время поиска в Google, при генерации ленты новостей в Facebook (поэтому новостной канал Facebook — не алгоритм, а его результат), при составлении списка возможных друзей на Google+ и Facebook, при работе с контактами в LinkedIn и т. д. Каждый из вышеперечисленных сервисов работает с разными параметрами и объектами, но математика за каждым алгоритмом остаётся неизменной.

- Кстати, видимо, Google является не первой компанией, начавшей работу с подобными типами алгоритмов. В 1996 году (за два года до основания Google) небольшая поисковая система под названием «RankDex», основанная Робином Ли, уже использовала эту идею для ранжирования страниц. Также, Массимо Марчиори, основатель «HyperSearch», использовал алгоритм ранжирования страниц, основанный на отношениях между отдельными страницами (эти два основателя упоминаются в патентах Google).

Пропорционально-интегрально-дифференцирующий алгоритм

- Вы когда-нибудь пользовались самолётом, автомобилем, спутниковой службой или сотовой сетью? Вы когда-нибудь видели робота или, может, бывали на заводе? В таком случае, вы видели этот алгоритм в действии.

- В основном этот алгоритм использует замкнутый механизм обратной связи для контура управления, чтобы минимизировать ошибку между желаемым выходным сигналом и реальным выходным сигналом. Он используется везде, где нужна система для обработки сигнала или управления механическими, гидравлическими и тепловыми механизмами, использующими автоматизацию.

- Можно сказать, что без этого алгоритма наша технологическая цивилизация не существовала бы.

Алгоритмы сжатия данных

- Трудно решить, какой алгоритм сжатия является наиболее важным, поскольку в зависимости от задачи используемый алгоритм может изменяться от zip до mp3 и от JPEG до MPEG-2. Но всем известно, насколько важны эти алгоритмы почти во всех сферах деятельности.

- Где может использоваться алгоритм сжатия помимо очевидного заархивированного документа? Например, эта веб-страница использует сжатие данных во время загрузки на ваш компьютер. Также эти алгоритмы используются в видеоиграх, видео, музыке, облачных вычислениях, базах данных и т. д. Можно сказать, что почти все используют алгоритмы сжатия данных, потому что они помогают сделать системы более дешёвыми и эффективными.

Алгоритм генерации случайных чисел

- Сегодня у нас нет «настоящего» генератора случайных чисел, но у нас есть некоторые генераторы псевдослучайных чисел, которые достойно справляются со своей задачей. Они имеют большую вариативность использования: от разнообразных приложений, криптографии, алгоритмов хеширования, видеоигр и искусственного интеллекта до тестов при разработке программ и т. д.

Подводя итоги

В конце хочется добавить, что этот список — субъективный выбор автора, поскольку существуют алгоритмы в таких задачах, как машинное обучение, умножение матриц, категоризация и т. д., которые не менее важны в современном мире и при этом здесь не упоминаются.