

Зміст

Вступ	
<i>Алгоритми для життя</i>	9
1. Оптимальна зупинка	
<i>Коли припиняти пошук</i>	16
2. Досліджувати/визискувати	
<i>Найновіший проти найвеличнішого</i>	39
3. Сортування	
<i>Навести лад</i>	69
4. Кешування	
<i>Забудьте про це</i>	97
5. Планування	
<i>Спершу найголовніше</i>	119
6. Правило Баеса	
<i>Передбачення майбутнього</i>	145
7. Перенавчання	
<i>Коли варто думати менше</i>	168
8. Розслаблення	
<i>Ну і хай собі</i>	190
9. Випадковість	
<i>Коли варто сподіватися на випадок</i>	205
10. Мережева комунікація	
<i>Як ми налагоджуємо зв'язки</i>	229
II. Теорія ігор	
<i>Що в інших на думці</i>	254

[Купити книгу на сайті kniga.biz.ua >>>](http://kniga.biz.ua)

Висновки	
<i>Обчислювальна доброта</i>	282
<i>Подяки</i>	289
<i>Примітки</i>	292
<i>Бібліографія</i>	352

[>>>](http://kniga.biz.ua)

Вступ

Алгоритми для життя

Уявіть, що ви шукаєте квартиру в Сан-Франциско, яке має славу найжахливішого міста Америки. Технологічний сектор переживає бум і жорсткі закони про зонування міста, які обмежують нову забудову, ніби змовилися, щоби місто стало настільки ж дорогим, як і Нью-Йорк, а на думку багатьох, навіть ще більш конкурентним. Нові оголошення про квартири з'являються і зникають за лічені хвилини, доступні будинки переповнені, а ключі від квартир часто опиняються у руках того, хто просто фізично встигає першим передати чек у руки домовласника.

Настільки нещадний ринок майже не залишає місця для з'ясування фактів і міркувань, що мало би стати характерною поведінкою раціонального споживача. На відміну від, скажімо, клієнта торговельного центру чи покупця онлайн, який може порівняти варіанти, перш ніж зробити вибір, майбутній мешканець Сан-Франциско мусить за будь-яку ціну вирішувати негайно: ви або берете квартиру, яку роздивляєтеся просто зараз, жертвуючи всіма іншими варіантами, або йдете, щоби ніколи сюди не повернутися.

Уявімо, що ви намагаєтеся збільшити свій шанс отримати найкращу з доступних квартир. Ваша мета — скоротити до абсолютного мінімуму Сциллу і Харібду розкаяння: не шкодувати про квартиру, яка вам не дісталася, і про «камінчик, який так і залишився неперевернутим». Ви одразу ж наштовхуетесь на дилему: як зрозуміти, що квартира справді найкраща, якщо не володієте початковими даними для порівняння? І як отримати дані, якщо ви не переглянете (і не *втрратите*) якусь кількість квартир? Що більше інформації ви отримаєте, то краще розпізнаєте найкращий для вас варіант, однак зростає ймовірність того, що ви вже проминули його.

Купити книгу на сайті kniga.biz.ua >>>

То що ж робити? Як прийняти зважене рішення, якщо сам процес отримання інформації ставить під загрозу результат? Це складна ситуація, що межує з парадоксом.

Коли перед людьми постає така проблема, більшість із них інтуїтивно твердять, що треба збалансувати пошук і ту мить, коли варто захопитися за можливість; що мусите подивитися достатню кількість квартир, аби визначити якийсь стандарт, а тоді вже брати те, що відповідає тому стандартowi, який ви визначили. Насправді таке визначення балансу доволі точне. А от що більшість людей не зможе сказати точно, то це — яким саме має бути цей баланс. На щастя, відповідь існує.

Тридцять сім відсотків.

Якщо хочете мати найвищі шанси і отримати найкращу квартиру, витратіть 37 % часу (одинадцять днів, якщо на пошук відвели собі один місяць) на дослідження варіантів, при цьому не беріть жодних додаткових зобов'язань. Залиште вашу чекову книжку вдома, адже ви просто роздивляєтеся. Але після цього будьте готові одразу ж ухопитися (заплатити заставу тощо) за перший варіант, який буде найкращим із тих, що ви вже дивилися. Це не просто задовільний компроміс між пошуком і отриманням бажаного. Це і є *ймовірно оптимальне* рішення.

Ми знаємо це, позаяк пошук квартири належить до класу математичних задач, відомих як задачі про «оптимальну зупинку». Правило 37 % визначає просту послідовність кроків (те, що спеціалісти з комп'ютерних наук називають «алгоритмами») для розв'язання таких задач. І, як з'ясувалося, пошук квартири — це лише один із прикладів, коли оптимальна зупинка вигулькує в повсякденному житті. Згода або відмова від варіантів, що послідовно постають перед нами, — це ситуація, яка знову і знову з'являється в житті у дещо видозмінених формах. Скільки разів доведеться об'їхати квартал, перш ніж знайти місце на паркувальному майданчику? Як довго можна випробувувати долю в ризикованій авантурі, перш ніж викласти гроші? Як довго доведеться чекати кращої пропозиції про придбання будинку чи машини?

Аналогічна проблема постає в іще більш важкій справі — побаченнях. Оптимальна зупинка — це наука про серійну моногамію.

Прості алгоритми пропонують рішення не лише для пошуку квартири, а й для всіх таких ситуацій у житті, де перед нами постає питання про оптимальну зупинку. Люди щодня стикаються з такими проблемами (хоча, звісно ж, поети витратили більше чорнила, описуючи сердечні страждання, аніж клопоти з паркуванням) і часто наражаються на неймовірні стреси. Але страждання зайві. Принаймні математично ці проблеми можна вирішити.

Купити книгу на сайті kniga.biz.ua >>>

1 Оптимальна зупинка

Коли припиняти пошук

Хоча всі християни починають запрошення на весілля урочистим проголошенням того, що їхнє весілля відбувається завдяки особливому Божому задуму, я, як філософ, хочу обговорити це детальніше...⁸

ЙОГАННЕС КЕПЛЕР

Якщо ви віддаєте перевагу панові Мартіну перед усіма іншими, якщо ви вважаєте його найприємнішим чоловіком, з яким вам доводилося проводити час, то чому ви вагаєтеся?

ДЖЕЙН ОСТІН, «Емма»

Це настільки розповсюджене явище, що психологи в університетах навіть мають для нього сленгову назву: «кинути за індичкою»⁹. У перший рік навчання в університеті закохані ще зі старших класів приїжджають додому на День подяки парами, а за чотири дні повертаються до університетського містечка вже поодиночі.

У перший рік навчання стривожений Браян ходив до університетського психолога. Його дівчина, з якою він зустрічався зі старших класів, поїхала навчатися в інший університет, за кілька штатів звідси, отож у них був роман «на відстані». Вони намагалися відповісти на складне філософське запитання: наскільки міцними були їхні стосунки? У них не було справжнього еталону інших стосунків, аби з ним порівнювати. Психолог Браяна визнала, що в них класична дилема першокурсників, а її порада була дивовижно нечулою: «Збирайте дані».

Особливістю серійної моногамії, як правило, є те, що люди стикаються з фундаментальною, неминучою проблемою. Коли можна вважати, що ви вже зустріли достатньо людей для того, аби усвідомити, хто підходить вам найкраще? А що, коли за збирання даних ви заплатили саме тією людиною, яка підходила вам найкраще? Здається, що це класична «пастка-22» сердечних запитань.

Купити книгу на сайті kniga.biz.ua >>>

Як ми бачили, «пастка-22», цей відчайдушний крик серця першокурсника, — це те, що математики називають задачею «оптимальної зупинки». І відповідь на неї — 37 %.

Звісно, все залежить від уявлення про кохання, яке ви хочете отримати.

Задача про секретарку

У будь-якій задачі про оптимальну зупинку¹⁰ головна дилема полягає не в тому, який варіант *обрати*, а скільки варіантів узагалі *розглядати*. Виявляється, що такі задачі хвилюють не лише закоханих і орендарів квартир, а й водіїв, домовласників, грабіжників і ще багато кого.

Правило 37 %* походить від найвідомішої головоломки про оптимальну зупинку, яка стала відомою під назвою «задача про секретарку»¹¹. Ситуація дуже подібна на дилему шукача квартири, яку ми розглянули вище. Уявіть, що ви проводите співбесіду з низкою претенденток на посаду секретарки, і ваша мета — максимізувати шанс винайняти найкращу з усіх претендентку. Хоча ви й не уявляєте, як оцінювати окремих претенденток, усе ж можете легко зрозуміти, котрій віддасте перевагу. (Математики сказали б, що вам доступні лише *порядкові* числа — відносний рейтинг претенденток порівняно одна з одною, — але не *кардинальні* числа, їхній рейтинг на своєрідній загальній шкалі). Ви проводите співбесіду в довільному порядку, по одній за раз. Можете вирішити запропонувати претендентці роботу в будь-яку мить, і тоді вона гарантовано погодиться, а пошук завершиться. Але якщо ви пройдете повз претендентку, вирішивши не наймати її, то вона піде назавжди.

Вважається, що задача про секретарку вперше з'явилася (щоправда, без згадки про секретарок) у номері журналу Scientific American у лютому 1960 р. серед головоломок із розважальної математики, представлених в улюбленій колонці Мартіна Гарднера¹². Але походження самої задачі дивовижно загадкове¹³. Наш власний пошук не дав майже нічого, окрім спекуляцій, а тоді перетворився на роботу детективів, несподівано пов'язану з фізичною працею: ми помандрували до архіву з документами Гарднера у Стенфорді, щоби повитягати звідти цілі коробки з його листуванням середини ХХ ст. Читати листування на папері — все одно, що підслуховувати людину, яка говорить по телефону: ви чуєте лише одного учасника розмови і мусите припускати, що

* Ми використовуємо жирний шрифт для тих алгоритмів, які багато разів згадуються в цій книжці. — Прим. авт., якщо не вказано інаше.

2 Досліджувати/визискувати

Найновіший проти найвеличнішого

У вас бурчить у животі. Ви підете до італійського ресторану, який знаєте і любите, чи до нового тайського, який щойно відкрився? Візьмете з собою найкращого друга чи зателефонуєте новому знайомому, щоби краще познайомитися з ним? Надто складно, можливо, краще залишитися вдома. Приготуєте за рецептом страву, яка завжди у вас добре виходить, чи візьметесь прочісувати інтернет у пошуках натхнення на щось нове? Забудьте. Може, краще замовити піцу? Візьмете «як завжди» чи попросите щось особливе? Ви виснажилися ще до того, як відкусили перший шматочок. І думка про те, щоб поставити платівку, подивитися фільм чи почитати книжку (*яку саме?*) більше не здається заспокійливою.

Щодня ми змушені обирати між варіантами, які різняться специфічними ознаками: спробуємо щось нове чи будемо дотримуватися улюбленого? Інтуїтивно ми розуміємо, що життя — це рівновага між нововведеннями і традицією, між найновішим і найвеличнішим, між ризиком і смакуванням того, що знаємо і любимо. Але як і в дилемі «подивитись і захопитись» з пошуком квартири, питанням без відповіді залишається — яким має бути баланс?

У 1974 р. автор класики «Дзен та мистецтво догляду за мотоциклом» Роберт Пірсіг засуджує фразу, якою часто починають спілкування, «Що нового?», стверджуючи, що це питання, «якщо точно відповідати на нього, має своїм наслідком лише нескінченну демонстрацію подробиць і моди — осаду завтрашнього дня». Натомість він схвалює альтернативну фразу «Що найкращого?», як значно ліпшу.

Але реальність не настільки проста. Згадка про те, що кожна «найкраща» пісня і ресторан, перш ніж стати вашими улюбленими, скромно починали з того, що були просто чимось «новим» для вас; нагадує,

що десь там усе ще можуть бути якісь найкращі речі, а отже, нове насправді варте принаймні якоїсь частки нашої уваги.

Давні афоризми визнають цю проблему, але не вирішують її. «Заводьте нових друзів, але бережіть давніх. Перші — срібло, другі — золото»⁶⁴ та «Хай наскільки багатим і винятковим було б життя, проте в ньому завжди знайдеться місце для ще одного друга»⁶⁵ доволі правдиві; безперечно, вони не підлягають сумнівам. Але вони не повідомляють нам нічого корисного про оптимальне *співвідношення*, скажімо, «срібла» й «золота», з якого вийде найкращий сплав життя, яке можна прожити.

Експерти з комп'ютерних наук працювали над пошуком цього балансу понад п'ятдесят років. Вони навіть дали йому назву — компроміс між «досліджувати» і «визискувати».

Досліджувати/визискувати

Слова «досліджувати» й «визискувати» мають доволі протилежні конотації. Але для спеціалістів із комп'ютерних наук ці слова мають значно вужче і нейтральне значення. Простіше кажучи, досліджувати означає *збирати* інформацію, а визискувати — *використовувати* інформацію, яку ви маєте, щоб отримати очікувано добрий результат.

Навіть інтуїтивно можна зрозуміти, що без дослідження жити неможливо. Але також варто згадати, що настільки ж погано ніколи не використовувати. Згідно із визначенням комп'ютерних наук, використання насправді може бути характерним для багатьох речей, які ми вважаємо найкращими в нашому житті. Зустріч із родиною на свята — це використання. Так само як і читач, який зручно влаштувався у кріслі з чашкою гарячої кави і улюбленою книжкою, або як група, яка грає найкращі хіти для захоплених фанатів, або як пара закоханих, яка витримала випробовування часом і танцює під «свою пісню».

Ба більше, дослідження може бути прокляттям.

Наприклад, привабливість музики полягає в тому, що постійно з'являється щось нове для прослуховування. А якщо ви журналіст — музичний критик, то для вас погано те, що в музиці *постійно* з'являється щось нове для прослуховування. Бути журналістом — музичним критиком означає, що можна досліджувати новий матеріал годинами і однак пропустити кілька свіжих композицій. Шанувальники музики можуть вважати, що працювати музичним журналістом — це рай, але коли ви постійно мусите вишукувати щось нове, то ніколи не можете насолодитися плодами свого тонкого смаку — це своєрідний різновид

3 Сортування

Навести лад

А тепер, якщо слово, яке ви волієте знайти, починається з «а», шукайте його на початку таблиці, а якщо з «ю», — шукайте ближче до кінця. Знов-таки, якщо ваше слово починається з «ка», шукайте на початку літери «к», але якщо з «ку», тоді дивіться ближче до кінця цієї літери. І так з усіма іншими ¹³⁷.

РОБЕРТ КОВДРІ, «Алфавітна таблиця» (1604)

Перш ніж Денні Гілліс заснував корпорацію Thinking Machines, перш ніж він винайшов знаменитий суперкомп'ютер Connection Machine, він був студентом Массачусетського технологічного інституту і жив у студентському гуртожитку. І його жахали шкарпетки сусіда.

На відміну від багатьох інших студентів, Гілліса жахала не гігієна сусіда. Річ була не в тім, що сусід не прав свої шкарпетки, — він їх прав. Проблема полягала ось у чому.

Сусід витягав одну шкарпетку з кошика з чистою білизною¹³⁸. Потім він випадковим чином витягав другу шкарпетку. Якщо вона не пасувала до першої, кидав її назад. Тоді він продовжував цей процес, витягаючи шкарпетки одну за одною і кидаючи їх назад, аж поки не знаходив підходящу пару для першої.

Якщо слідувати цьому методу, то, маючи лише 10 різних пар шкарпеток, вам у середньому знадобиться 19 спроб, аби доповнити першу пару, і ще 17 спроб, аби дібрати другу пару. Загалом, сусідові знадобилося б 110 разів виуджувати шкарпетки з корзини, аби спарувати 20 шкарпеток.

Цього було достатньо, щоби змусити будь-якого майбутнього експерта з комп'ютерних наук попросити про переселення в іншу кімнату.

Ну а запропонувати тему про те, як сортувати шкарпетки, — чудовий спосіб змусити спеціаліста з комп'ютерних наук говорити на диво довго. Питання про шкарпетки, яке у 2013 р. поставили на веб-сайті

програмістів Stack Overflow, викликало суперечку, яка вилилася майже у 12 тис. слів¹³⁹.

«Шкарпетки бентежать мене!»¹⁴⁰ — зізнався нам двом легендарний криптограф і спеціаліст із комп'ютерних наук, володар премії Тюрінга Рон Рівест, коли ми підняли цю тему.

Під час зустрічі він був у сандалях — і без шкарпеток.

Екстаз від сортування

Сортування лежить в основі того, чим займаються комп'ютери. Насправді багато в чому саме сортування сприяло виникненню комп'ютерів.

Наприкінці XIX ст. американське населення зростало на 30 % що десять років, і кількість «суб'єктів опитування» у переписі населення США збільшилася з лише п'яти у 1870-му до понад двохсот у 1880-му. На укладення таблиці за матеріалами перепису 1880-го пішло вісім років — її ледь встигли завершити до початку перепису 1890 р. Як зауважив один із авторів тих часів, дивно було, що «клерки, які тяжко працювали над цими аркушами паперу з підрахунками, що вже дратували їх... не втратили зір і не збожеволіли»¹⁴¹. Існувала загроза, що ця затія от-от розвалиться під власною вагою. Щось треба було з цим робити.

Винахідник на ім'я Герман Голлеріт, надихнувшись тогочасними компостованими залізничними квитками, розробив систему перфокарт на яскравому, так званому «манільському», папері, щоби зберігати інформацію, та машину для рахування і сортування, яку назвав «машиною Голлеріта». Голлеріт отримав патент 1889-го, а 1890-го уряд використав машину Голлеріта для перепису населення. Ще нікому не доводилося бачити нічого подібного. Один вражений спостерігач писав: «Апарат працює бездоганно, як млини богів, але затьмарює їх за швидкістю»¹⁴². Проте інший автор вважав, що цей винахід має обмежену сферу використання: «Оскільки ніхто, крім уряду, не буде цим користуватися, винахідник навряд чи дуже розбагатіє»¹⁴³. Це передбачення, яке Голлеріт вирізав і зберіг, виявилось не зовсім правильним. Компанія Голлеріта злилася з кількома іншими в 1911 р. і стала компанією Computing-Tabulation-Recording («Розрахунки-Таблиці-Записи») ¹⁴⁴. За кілька років її перейменували на International Business Machines («Міжнародні бізнес-машини»), або IBM.

Сортування стимулювало розвиток комп'ютерів ще впродовж наступного століття. Першою з написаних для комп'ютера зі «збереженою програмою» кодів була програма для ефективного сортування¹⁴⁵.

Купити книгу на сайті kniga.biz.ua >>>