

درک الگوریتم

راهنمای تصویری
برای برنامه‌نویس‌ها
و افراد کنجه‌کاو

آدیتیا بهارگاؤ

متجم: مهران افشارنادری

عنوان و نام پدیدآور: درک الگوریتم (راهنمای تصویری برای برنامهنویس‌ها و افراد کنجهکار) / آدیتیا بهارگاوا؛
مترجم: مهران افشارنادری
مشخصات نشر: تهران، ۱۴۰۲ -
مشخصات ظاهري: ديجيتال، ۲۸۸ ص، مصور
موضوع: الگوریتم
صفحه آرایي: نفيسه عطازان

درک الگوریتم

راهنمای تصویری
برای برنامه‌نویس‌ها
و افراد کنجدکاو

آدیتیا بهارگاوا

مترجم: مهران افشارنادری

.....

تقدیم به پدر و مادرم، ییگش و سنگیننا

قدرتانی مترجم :

سپاس از جادی عزیز که کتاب را پیش از انتشار خواند و پیشنهادهایی در جهت بهتر شدن
کیفیت ترجمه داد. با این حال هرگونه اشکال در متن نهایی متوجه مترجم است.



فهرست

i	پیش‌گفتار
ii	قدرتانی
iii	درباره‌ی این کتاب

۱ درآمدی بر الگوریتم

۱	مقدمه
۲	آن‌چه درباره‌ی کارایی خواهید آموخت
۲	آن‌چه در مورد حل مسئله یاد می‌گیرید
۴	جست‌وجوی دودویی
۶	یک راه بهتر برای جستجو
۱۳	زمان اجرا
۱۴	نماد O بزرگ
۱۴	زمان اجرای الگوریتم با سرعت‌های متفاوتی رشد می‌کند
۱۷	به تصویر کشیدن چند زمان اجرا با O بزرگ متفاوت
۱۹	O بزرگ برای بدترین زمان اجرا
۲۰	برخی از زمان اجراهای معمول O بزرگ
۲۲	فروشنده‌ی دوره‌گرد
۲۵	جمع‌بندی

۲ مرتب‌سازی انتخابی

۲۷

۲۸	شیوه‌ی عملکرد حافظه
۳۰	آرایه‌ها و لیست‌های پیوندی
۳۱	لیست‌های پیوندی
۳۳	آرایه‌ها
۳۵	اصطلاح‌شناسی
۳۶	درج در وسط یک لیست
۳۷	حذف
۴۱	مرتب‌سازی انتخابی
۴۶	جمع‌بندی

۳ بازگشت

۴۷

۴۸	بازگشت
۵۱	صورت پایه و صورت بازگشتی
۵۳	پشته
۵۴	پشته فراخوانی
۵۸	پشته فراخوانی با بازگشت
۶۳	جمع‌بندی

۴ مرتب‌سازی سریع

۶۵

۶۶	تقسیم و حل
۷۴	مرتب‌سازی سریع
۸۱	بازبینی نماد O بزرگ
۸۲	مرتب‌سازی ادغامی در برابر مرتب‌سازی سریع

۸۴	حالت متوسط در مقابل بدترین حالت
۸۸	جمع‌بندی

۵ جدول‌های هش

۹۲	توابع هش
۹۷	موارد کاربرد
۹۷	استفاده از جدول‌های هش برای جست‌وجو
۹۹	جلوگیری از ورودی‌های تکراری
۱۰۲	استفاده از جدول هش برای کش
۱۰۵	جمع‌بندی
۱۰۵	تصادم‌ها
۱۰۸	کارایی
۱۱۱	ضریب بار
۱۱۳	تابع هش مناسب
۱۱۵	جمع‌بندی

۶ جست‌وجوی سطح اول

۱۱۸	مقدمه‌ای بر گراف‌ها
۱۲۱	گراف چیست؟
۱۲۲	جست‌وجوی سطح اول
۱۲۵	یافتن کوتاه‌ترین مسیر
۱۲۷	صف
۱۲۹	پیاده‌سازی گراف
۱۳۱	پیاده‌سازی الگوریتم
۱۳۶	زمان اجرا

جمع‌بندی

۱۴۰

۷ الگوریتم دیجسترا

۱۴۳

کار با الگوریتم دیجسترا

۱۴۸

اصطلاح‌شناسی

۱۵۱

معامله بر سر یک پیانو

۱۵۸

یال‌های با وزن منفی

۱۶۱

پیاده‌سازی

۱۷۲

جمع‌بندی

۱۷۳

۸ الگوریتم حریصانه

۱۷۴

مسئله‌ی برنامه‌ریزی برای کلاس درس

۱۷۷

مسئله‌ی کوله‌پشتی

۱۷۹

مسئله‌ی پوشش مجموعه

۱۸۱

الگوریتم‌های تقریبی

۱۸۸

تشخیص مسئله‌ی ان‌پی کامل

۱۹۰

فروشنده‌ی دوره‌گرد، گام به گام

۱۹۵

تشخیص مسئله‌ی ان‌پی کامل

۱۹۸

جمع‌بندی

۱۹۹

۹ برنامه‌نویسی پویا

۱۹۹

مسئله‌ی کوله‌پشتی

۲۰۰

راه حل ساده

۲۰۱

برنامه‌نویسی پویا

۲۱۱	پرسش‌های متداول در مسئله‌ی کوله‌پشتی
۲۱۱	با اضافه کردن یک آیتم دیگر چه اتفاقی می‌افتد؟
۲۱۴	چه اتفاقی می‌افتد اگر ترتیب ردیف‌ها را تغییر بدھید؟
۲۱۴	آیا می‌توان شبکه را به جای ردیفی ستونی پر کرد؟
۲۱۵	چه اتفاقی می‌افتد اگر یک آیتم کوچک‌تر اضافه کنید؟
۲۱۵	آیا می‌توان کسری از یک آیتم را سرقت کرد؟
۲۱۶	بهینه‌سازی برنامه‌ی سفر
۲۱۸	بررسی آیتم‌های وابسته به یکدیگر
۲۱۸	آیا ممکن است که این راه حل نیاز به بیش از دو زیر-کوله‌پشتی داشته باشد؟
۲۱۹	آیا ممکن است با بهترین راه حل، کوله‌پشتی به طور کامل پر نشود؟
۲۲۰	طولانی‌ترین زیررشته‌ی مشترک
۲۲۱	ساخت شبکه
۲۲۲	پرکردن شبکه
۲۲۴	راه حل
۲۲۵	بزرگ‌ترین زیردباله‌ی مشترک
۲۲۷	بزرگ‌ترین زیردباله‌ی مشترک - راه حل
۲۲۹	جمع‌بندی

۱۰ الگوریتم K - نزدیک‌ترین همسایه

۲۳۱	طبقه‌بندی پرتوال در مقابل گریپ‌فروت
۲۳۴	ایجاد سیستم پیشنهاد
۲۳۵	استخراج ویژگی
۲۴۰	رگرسیون
۲۴۳	انتخاب ویژگی‌های خوب
۲۴۴	مقدمه‌ای بر یادگیری ماشین
۲۴۶	ساخت فیلتر اسپم

پیش‌بینی بازار سهام
جمع‌بندی

۱۱ گام بعدی

۲۴۹	درخت
۲۵۳	ایندهکس‌های معکوس
۲۵۴	تبديل فوریه
۲۵۵	الگوریتم‌های موازی
۲۵۷	نگاشت‌کاهش
۲۵۷	چرا الگوریتم‌های توزیع شده مفید هستند؟
۲۵۸	تابع map
۲۵۹	تابع reduce
۲۶۰	فیلترهای بلوم و هایپرلاگ لاغ
۲۶۱	فیلترهای بلوم
۲۶۲	هایپرلاگ لاغ
۲۶۳	الگوریتم‌های SHA
۲۶۳	مقایسه‌ی فایل‌ها
۲۶۵	چک‌کردن رمزهای عبور
۲۶۶	هیشنگ حساس به مکان
۲۶۷	تبادل کلید دیفی هلمن
۲۶۹	برنامه‌ریزی خطی
۲۷۱	پاسخ تمرین‌ها



پیشگفتار

در ابتداء، برنامه‌نویسی برای من تنها یک سرگرمی بود. مبانی برنامه‌نویسی را از کتاب Visual Basic 6 for Dummies آموختم و برای فهم بیشتر به خواندن کتاب‌های دیگر ادامه دادم. اما موضوع الگوریتم سخت و غیر قابل درک باقی ماند. به خاطر دارم که فهرست مطالب اولین کتاب الگوریتمی را که داشتم، مزه‌مزه می‌کردم و به این فکر می‌کدم که «بالاخره این موضوعات را درک خواهم کرد!» اما مطالب سنگین بود و پس از چند هفته منصرف شدم. تا بعد اولین استاد الگوریتم خوبم را پیدا کردم و فهمیدم این ایده‌ها چقدر ساده و زیبا هستند.

چند سال پیش، اولین پست تصویری و بلاگم را نوشتتم. اصولاً موضوعات را در قالب دیداری بهتر یاد می‌گیرم و سبک مصور را خیلی دوست دارم. از آن زمان، چند پست مصور در مورد برنامه‌نویسی فانکشنال^۱، گیت^۲، یادگیری ماشین^۳ و هم‌زمانی^۴ نوشتم. راستی زمانی که شروع به کار کردم نویسنده‌ی متوسطی بودم. توضیح مفاهیم فنی سخت است و ارائه‌ی مثال‌های خوب و توضیح یک مفهوم دشوار، زمان‌بر. بنابراین راحت‌ترین کار نادیده‌گرفتن مفاهیم سخت است. فکر می‌کرم چقدر کارم را خوب انجام می‌دهم، ولی بعد از محبوبیت یکی از پست‌هایم، یکی از همکارانم گفت: «پست تو را خواندم ولی چیزی سر در نیاوردم.» فهمیدم هنوز راه درازی در مسیر یادگیری نویسنده‌گی در پیش دارم.

در میانه‌ی نوشتمن این پست‌های و بلاگ، انتشارات منینگ به سراغ من آمد و از من پرسید که تمایلی دارم که یک کتاب مصور بنویسم؟ خوب، معلوم شد که ویراستاران منینگ در مورد توضیح مفاهیم فنی چیزهای زیادی می‌دانند و به من آموختند که چگونه تدریس کنم. این کتاب را نوشتتم تا این وسوسه را عملی کنم: می‌خواستم کتابی

1. functional programming
2. git
3. machine learning
4. concurrency

بنویسم که موضوعات فنی سخت را به خوبی توضیح بدهد، و کتاب الگوریتمی می‌خواستم بنویسم که آسان‌خوان باشد. نوشه‌های من با اولین پست و بلاگ فاصله‌ی زیادی دارد، و امیدوارم این کتاب برای شما آسان و آموزنده باشد.

قدردانی

از انتشارات منینگ تشکر می‌کنم که فرصت نوشتن این کتاب را در اختیار من قرار داد و دستم را برای خلائق باز گذاشت. از ناشر کتاب مرجان بیس و مایک استفنز برای همراهی، برت بیتس برای آموزش نویسنده‌گی، و جنیفر استات چون ویراستاری فوق العاده پاسخگو و یاری رسان بود، تشکر می‌کنم. همچنین از افراد تیم تولید منینگ: کوین سالیوان، مری پیرگیز، تیفانی تیلور، لزلی هایمز، و بقیه‌ی افراد دست‌اندرکار متشکرم. علاوه بر این، می‌خواهم از کسانی که دست‌نوشته‌ی کتاب را خواندند و پیشنهاداتی ارائه کردن تشکر کنم: کارن بنسدون، راب گرین، مایکل همرا، اوزن هارلوویچ، کالین هستی، کریستوفر هاپت، چاک هندرسون، پاول کوزلوفسکی، آمیت لامبا، ژان فرانسوا مورین، رابت موریسون، سانکار، راماناتان، ساندر راسل، داگ اسپارلینگ و دیمین وايت.

باتشکر از افرادی که به من کمک کردن تابه این نقطه برسم: رفقای Flaskhit game board که به من کدنویسی را یاد دادند؛ دوستان زیادی از جمله بن سرکه، کارل پوزون، الکس منینگ، استر چان، آنیش بات، مایکل گلس، نیکراد مهدی، چارلز لی، جرد فریدمن، هما مانیکاواساگام، هاری راجا، مورالی گودپیاتی، سرینیواس وارادان و دیگران که با مرور فصل‌ها، مشاوره دادند و فرصتی برای محک‌زن آشکال گوناگون ارائه‌ی توضیحات را برایم فراهم کردند و سپاس از گری بردى، برای آموزش الگوریتم‌ها. یک تشکر اساسی دیگر از استاید CLRS^۱، کنوت و استرنگ. به راستی که روی شانه‌های غول‌ها ایستاده‌ام.

بابا، مامان، پریانکا و بقیه‌ی اعضای خانواده: از حمایت همیشگی شما متشکرم. و یک تشکر بزرگ از همسرم مگی. ماجراهای زیادی در پیش روی ما وجود دارد، که البته محدود به آخرهفته‌ها در خانه ماندن و بازنویسی پاراگراف‌ها نمی‌شود.

در نهایت، سپاس فراوان از خوانندگانی که کتاب را مطالعه کردند و آنانی که به من در انجمن این کتاب بازخورد دادند تشکر می‌کنم. شما واقعاً به بهتر شدن این کتاب کمک کردید.

۱. توماس اچ کورمن چارلز ای لایرسان رونالد ریوست کلیفورد استین نویسنده‌گان کتاب مقدمه‌ای بر الگوریتم.

درباره‌ی این کتاب

این کتاب به گونه‌ای طراحی شده که آسان‌خوان باشد. از جهش‌های فکری بزرگ اجتناب کرده‌ام. هر زمان که مفهوم جدیدی مطرح می‌شود، آن را توضیح می‌دهم یا به شما می‌گوییم که چه زمانی آن را توضیح خواهم داد. مفاهیم اصلی با تمرین‌ها و توضیحات متعدد تقویت می‌شوند تا بتوانید مفروضات خود را بررسی کنید و مطمئن شوید که موضوعات را متوجه شده‌اید.

هر موضوع را با مثال شرح می‌دهم. به جای نامفهوم‌نویسی هدف من این است که برای شما تجسم این مفاهیم را آسان کنم. از طرفی فکر می‌کنم با یادآوری آن چه از قبل می‌دانیم، یادگیری بهتری خواهیم داشت، و مثال‌ها یادآوری را آسان‌تر می‌کنند. بنابراین وقتی سعی می‌کنید تفاوت بین آرایه‌ها و لیست‌های پیوندی را به خاطر بیاورید (فصل ۲)، می‌توانید صرفاً به نشستن افراد در یک سالن برای تماشای فیلم فکر کنید. هرچند شاید بدیهی باشد که نظر برسد، خودم مطالب را به شکل تصویری بهتر یاد می‌گیرم و این کتاب هم سرشار از تصاویر است.

مطالب کتاب به دقت تنظیم شده است. نیازی به نوشتن کتابی نیست که همه‌ی الگوریتم‌های مرتب‌سازی را پوشش‌بدهد - به همین دلیل است که ما ویکی‌پدیا و خان‌آکادمی را داریم. تمام الگوریتم‌هایی که در این کتاب گنجانده‌اند کاربردی هستند و من از آن‌ها در شغلم به عنوان یک مهندس نرم‌افزار استفاده می‌کنم. این‌ها پایه‌ی خوبی برای موضوعات پیچیده‌تر فراهم می‌کنند.

به سلامتی خواندن!

نقشه‌ی راه

در سه فصل اول به بینان الگوریتم می‌پردازیم:

- **فصل ۱**- اولین الگوریتم عملی خود را خواهید آموخت: جست‌وجوی دودویی^۱. همچنین یاد می‌گیرید که سرعت یک الگوریتم را با استفاده از نماد O بزرگ^۲ تجزیه و تحلیل کنید. نماد O بزرگ در سراسر کتاب برای تجزیه و تحلیل سرعت یک الگوریتم استفاده شده است.
- **فصل ۲**- با دو ساختمانداده‌ی اساسی آشنا خواهید شد: آرایه و لیست پیوندی. این ساختمانداده‌ها در سراسر کتاب مورد استفاده قرار می‌گیرند، و از آنها برای ساختن ساختمانداده‌های پیشرفته‌تر مانند جدول‌های هش^۳ استفاده می‌شود(فصل ۵).
- **فصل ۳**- در مورد بازگشت^۴، یک تکنیک مفید که توسط بسیاری از الگوریتم‌ها استفاده می‌شود (مانند مرتب‌سازی سریع^۵، که در فصل ۴ پوشش داده شده است) می‌آموزید.

به تجربه‌ی من، نماد O بزرگ و تابع‌های بازگشتی موضوعاتی چالش‌برانگیز برای مبتدیان است. بنابراین در این فصل با سرعتی آهسته زمان بیشتری را به این بخش‌ها اختصاص داده‌ام.

- در بخش‌های دیگر کتاب به بررسی الگوریتم‌هایی با کاربردهای گسترده می‌پردازیم:
- در فصل‌های ۴، ۸، و ۹ **تکنیک‌های حل مسئله** - پوشش داده شده است. اگر به مشکلی برخورد کردید و مطمئن نبودید که چگونه آن را به طور موثر حل کنید، تقسیم و حل^۶(فصل ۴) یا برنامه نویسی پویا^۷(فصل ۹) را امتحان کنید. یا ممکن است متوجه بشوید که هیچ راه حل کارآمدی وجود ندارد و به جای آن با استفاده از الگوریتم حریصانه^۸ به یک پاسخ تقریبی دست پیدا کنید (فصل ۸).

- **جدول هش** در فصل ۵ پوشش داده شده است. جدول هش ساختمانداده‌ای

1. Binary search
2. Big O notation
3. Hash tables
4. recursion
5. quicksort
6. Divide and conquer
7. Dynamique programming
8. Greedy algorithm

بسیار مفید است. این جدول‌ها شامل مجموعه‌ای از جفت‌های کلید^۱ و مقدار^۲ است. مانند نام شخص و آدرس ایمیل او، یا نام کاربری و رمز عبور آن. جدول‌های هش واقعاً مفید هستند. وقتی می‌خواهیم مسئله‌ای را حل کنم، دو استراتژی که در شروع درنظر می‌گیریم عبارتند از: «آیا می‌توانم از جدول هش استفاده کنم؟» و «آیا می‌توانم این را به عنوان یک گراف^۳ مدل‌سازی کنم؟»

- **الگوریتم گراف** در فصل‌های ۶ و ۷ پوشش داده شده است. گراف راهی برای مدل‌سازی یک شبکه است: یک شبکه‌ی اجتماعی، یا شبکه‌ای از مسیرها، یا نورون‌ها، یا هر مجموعه‌ی دیگر از ارتباط‌ها. جست‌وجوی سطح اول^۴ (فصل ۶) و الگوریتم دایجسترا^۵ (فصل ۷) راه‌هایی برای یافتن کوتاه‌ترین فاصله بین دو نقطه در یک شبکه هستند: می‌توانید از این روش برای محاسبه‌ی درجات جدایی بین دو نفر یا کوتاه‌ترین مسیر به مقصد استفاده کنید.

- **الگویتم K-نزدیکترین همسایه^۶ (KNN)** - این یک الگوریتم ساده برای استفاده در یادگیری ماشینی است و در فصل ۱۰ به آن پرداخته می‌شود. می‌توانید از KNN برای ایجاد یک سیستم پیشنهادات، یک موتور OCR، سیستم پیش‌بینی ارزش سهام و به طور کلی هر چیزی که شامل پیش‌بینی یک مقدار باشد، استفاده کنید («ما فکر می‌کنیم آدیت به این فیلم ۴ ستاره می‌دهد») یا می‌توانید به کمک آن یک شیء را طبقه‌بندی کنید. («این حرف یک Q است»).

- **گام‌های بعدی** در فصل ۱۱ به ده الگوریتم دیگر پرداخته می‌شود که برای مطالعه‌ی بیشتر مناسب هستند.

شیوه‌ی خواندن این کتاب

ترتیب و محتوای این کتاب با دقت طراحی شده است. اگر به موضوع خاصی علاقه‌مند هستید، با خیال راحت به آن بخش بروید. در غیر این صورت، فصل‌ها را به ترتیب بخوانید. هر فصل زیربنای فصل بعدی است.

1. key
2. value
3. graph
4. Breadth-first search
5. dijkstra algorithm
6. k-nearest neighbors algorithm

اکیداً توصیه می‌کنم کد مثال‌ها را خودتان اجرا کنید. در این مورد هر چه بگوییم کم گفته‌ام. نمونه کدها را کلمه به کلمه تایپ کنید (یا آنها را از www.manning.com/books/grok (یا آنها را از https://github.com/egonschiele/grokking_algorithms دانلود کنید)، و آن‌ها را اجرا کنید. با پیروی از این شیوه چیزهای بیشتری را به خاطر خواهید سپرد.

انجام تمرین‌های این کتاب را هم توصیه می‌کنم. تمرینات کوتاه هستند. حل آن‌ها معمولاً فقط یک یا دو دقیقه، گاهی اوقات ۵ تا ۱۰ دقیقه زمان می‌برند. این تمرین‌ها به شما کمک می‌کنند تا از شرایط یادگیری خود مطلع بشوید، بنابراین قبل از اینکه دیر شود متوجه می‌شوید که از مسیر خارج شده‌اید یا خیر.

این کتاب برای چه کسانی مناسب است

این کتاب برای کسانی است که اصول کدنویسی را می‌دانند و می‌خواهند الگوریتم‌ها را درک کنند. شاید قبلاً در کدنویسی به مشکل برخورد کرده باشید و در تلاش برای یافتن یک راه حل الگوریتمی باشید. یا شاید می‌خواهید بدانید که الگوریتم‌ها به چه کاری می‌آیند. در اینجا یک لیست کوتاه و ناقص از افرادی است که احتمالاً از این کتاب

بهره می‌برند:

- کدنویسان تفریحی
- دانش‌آموزان دوره‌های کدنویسی
- فارغ التحصیلان علوم کامپیوتر که به دنبال بازآموزی هستند
- فارغ التحصیلان فیزیک، ریاضی و دیگر رشته‌های که به برنامه‌نویسی علاقه‌مند هستند

قراردادهای کدنویسی و آدرس فایل‌ها برای دانلود

تمام نمونه کدهای این کتاب از Python 2.7 استفاده می‌کنند. تمام کدهای موجود در کتاب با فونت با عرض ثابت مانند این (fixed-width font like this) ارائه شده است تا از متن معمولی قابل تفکیک باشد. حاشیه‌نویسی کد برخی از لیست‌ها را همراهی می‌کند و مفاهیم مهم را برجسته می‌کند.

می‌توانید کد مثال‌های کتاب را از وب‌سایت ناشر به نشانی www.manning.com/books/grok (یا از https://github.com/egonschiele/grokking_algorithms دانلود کنید). تنها زمانی بیشترین میزان یادگیری را تجربه می‌کنید که واقعاً از یادگیری لذت ببرید

پس لذت ببرید و نمونه کدها را اجرا کنید!

درباره‌ی نویسنده

آدیتیا بهارگاوا^۱ مهندس نرم‌افزار در Etsy، (بازار آنلاین برای کالاهای دست‌ساز) است. او دارای مدرک کارشناسی ارشد در رشته‌ی علوم کامپیوتر از دانشگاه شیکاگو است. او همچنین یک وبلاگ فناوری مصور محبوب و پر بازدید به آدرس adit.io دارد.

یک ا درآمدی بر الگوریتم



در این فصل

• مبانی و اساس مباحث مطرح شده در کتاب را یاد می‌گیرید.

• اولین الگوریتم جستجوی خودرامی نویسید (جستجوی دودویی).

• یاد می‌گیرید چگونه درباره‌ی زمان اجرای یک الگوریتم صحبت کنید (علامت O بزرگ).

• یک تکنیک رایج برای طراحی الگوریتم به شما معرفی می‌شود (تابع بازگشتی).

مقدمه

الگوریتم مجموعه‌ای از دستورالعمل‌ها برای انجام یک کار است. هرچند هر قطعه کد را می‌توان یک الگوریتم نامید، با این حال در این کتاب به موارد جالب‌تر پرداخته می‌شود. الگوریتم‌هایی را در این کتاب آورده‌ام که عملکرد سریعی دارند یا به واسطه‌ی آن‌ها مسائل قابل توجهی را می‌توان حل کرد یا هر دو این ویژگی‌ها را شامل می‌شوند. در ادامه مباحث مطرح شده در این کتاب را مرور می‌کنیم:

- در فصل ۱ در مورد جستوجوی دودویی صحبت می‌شود و می‌بینیم که چگونه یک الگوریتم می‌تواند سرعت کد شما را افزایش بدهد. در یک مثال، تعداد مراحل مورد نیاز از ۴ میلیارد مرحله به ۳۲ مرحله کاهش می‌یابد!
- یک دستگاه GPS از الگوریتم گراف (همان‌طور که در فصل‌های ۶ و ۷ یاد می‌گیرید) برای محاسبه‌ی کوتاه‌ترین مسیر به مقصد استفاده می‌کند.
- می‌توانید از برنامه‌نویسی پویا (در فصل ۹ به آن پرداخته شده است) برای نوشتن یک الگوریتم هوش مصنوعی برای بازی چکرز استفاده کنید.
- در هر مورد، الگوریتم را شرح می‌دهم و مثالی برای شما می‌زنم. سپس در مورد زمان اجرای الگوریتم در نماد O بزرگ صحبت می‌کنم. در انتها، به بررسی مسئله‌های دیگری که می‌توان با همان الگوریتم حل کرد، می‌پردازم.

آن‌چه درباره‌ی کارایی^۲ خواهید آموخت

خبر خوب این که احتمالاً پیش از این نسخه‌ی پیاده‌سازی شده‌ی هر الگوریتمی که در این کتاب آمده، در زبان برنامه‌نویسی مورد علاقه‌ی شما در دسترس باشد، در نتیجه نیاز نخواهد شد خود شما کد هر الگوریتم را بنویسید! اما اگر این کدها درک نشوند، پیاده‌سازی آن‌ها بیهوده خواهد بود. در این کتاب، می‌آموزید الگوریتم‌های مختلف را مقایسه کنید: باید از مرتب‌سازی ادغامی^۳ استفاده کرد یا مرتب‌سازی سریع^۴? از آرایه^۵ یا لیست^۶? استفاده از یک ساختمان داده‌ی^۷ متفاوت می‌تواند تفاوت اساسی ایجاد کند.

آن‌چه در مورد حل مسئله یاد می‌گیرید

شما تکنیک‌هایی را برای حل مسئله‌هایی که ممکن است تاکنون از درک شما خارج بوده‌اند، یاد می‌گیرید. مثلا:

1. checkers
2. Performance
3. merge sort
4. quicksort
5. array
6. list
7. Data structure

- اگر به ساخت بازی‌های ویدیویی علاقه دارید، می‌توانید یک سیستم هوش مصنوعی بنویسید که حرکات بازیکن را به کمک الگوریتم گراف دنبال بکند.
 - می‌آموزید که با استفاده از K نزدیک‌ترین همسایه یک سیستم پیشنهاد بسازید.
 - برخی از مسائل قابل حل نیستند! بخشی از این کتاب که در مورد مسائل NP کامل صحبت می‌کند و به شما نشان می‌دهد که چگونه آن مسئله‌ها را شناسایی کنید و الگوریتمی ارائه بدھید که به پاسخ تقریبی برسید.
- در پایان این کتاب، با برخی از کاربردی‌ترین الگوریتم‌ها آشنا می‌شوید. سپس می‌توانید از دانش جدید خود برای یادگیری الگوریتم‌های خاص تر برای هوش مصنوعی، پایگاه‌های داده و غیره استفاده کنید. یا می‌توانید با چالش‌های بزرگ‌تری در محل کار روپرو بشوید.

پیش‌نیاز

پیش از شروع این کتاب دانش مقدماتی از جبر مورد نیاز است. در این حد که در تابع $x^2 = f(x)$ ، پاسخ $f(5)$ را بدانید. اگر جواب شما عدد ۱۰ است، از دانش کافی ریاضی برخوردارید.

به علاوه، دنبال‌کردن این فصل (و این کتاب) اگر با زبان برنامه‌نویسی آشنا باشید آسان‌تر خواهد بود. تمامی مثال‌های این کتاب به زبان پایتون نوشته شده‌اند. اگر با هیچ زبان برنامه‌نویسی آشنایی ندارید و می‌خواهید یاد بگیرید، پایتون را انتخاب کنید - برای تازه‌کارها حرف ندارد. اگر با زبان دیگری مثل روبی هم آشنایی دارید، باز مشکلی نخواهید داشت.

جستجوی دودویی

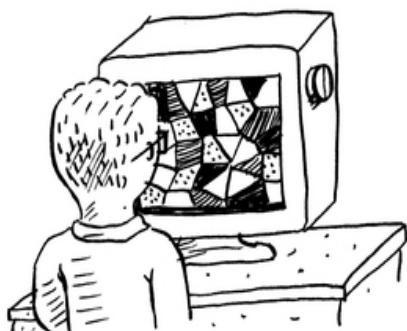


فرض کنید در یک دفترچه تلفن دنبال نام فردی می‌گردید(چه کار از مدافعتاوهای) اسم آن شخص با حرف K شروع می‌شود. می‌توانید از اول شروع کرده و صفحات را ورق بزنید تا به اسم‌هایی بررسید که با حرف k شروع می‌شوند. با این حال احتمال این‌که از صفحه‌ای از وسط دفترچه شروع به جستجو کنید بیشتر است، چون می‌دانید که K‌ها جایی در همان میانه‌ی دفترچه‌ی تلفن هستند.

یا فرض کنید به دنبال کلمه‌ای در دیکشنری می‌گردید که با حرف O شروع شده است. این بار هم، از صفحات وسط شروع می‌کنید.

حالا فرض کنید که وارد فیس‌بوک می‌شوید.

فیس‌بوک باید تأیید کند که در این سایت حساب کاربری دارید. پس، باید در پایگاه داده‌ی^۱ خود دنبال نام شما بگردد. فرض کنید که نام کاربری شما Karmageddon است.



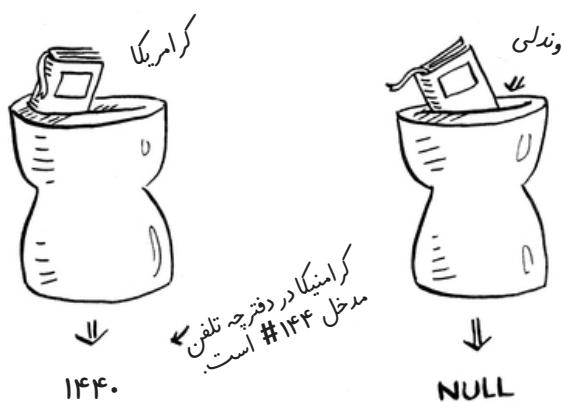
فیس‌بوک می‌تواند از حرف A شروع کرده و به دنبال نام شما بگردد-با این حال برایش منطقی‌تر است که از جایی در میانه‌ی فهرست شروع کند.

این یک مسئله‌ی جستجو است. و در تمامی این موارد از الگوریتم یکسانی برای حل مسئله استفاده می‌شود.

جستجوی دودویی الگوریتمی است

که ورودی آن یک لیست مرتب شده از عناصر است (بعدتر توضیح خواهم داد که چرا نیاز است تا این لیست مرتب بشود). اگر عنصری که به دنبال آن می‌گردید در آن لیست باشد، جستجوی دودویی موقعیت مورد نظر را برمی‌گرداند. در غیر این صورت جواب جستجوی دودویی null خواهد بود.

به عنوان مثال:



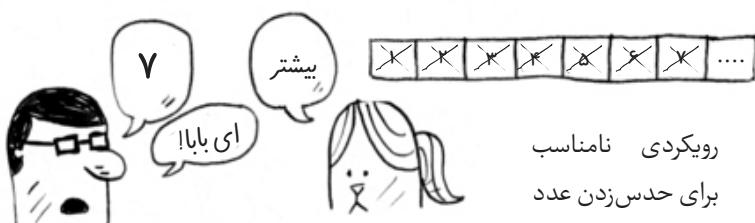
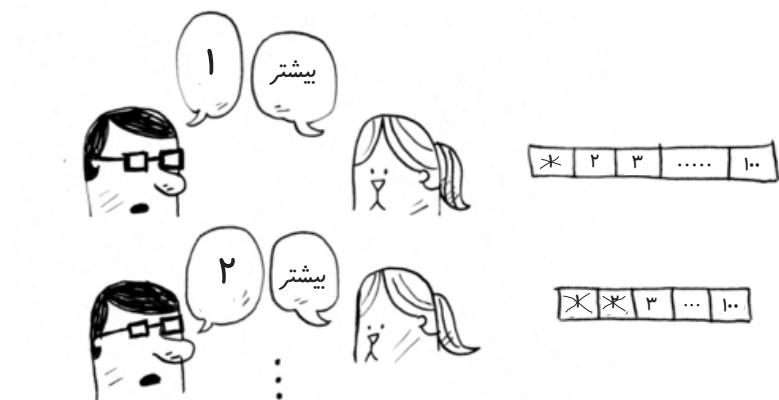
به دنبال شماره تلفن شرکت‌ها در یک دفترچه تلفن از طریق جستجوی دودویی

در اینجا مثالی داریم از این‌که جستجوی دودویی چگونه کار می‌کند. عددی بین ۱ تا ۱۰۰ را انتخاب می‌کنم.

۱	۲	۳	۱۰۰
---	---	---	------	-----

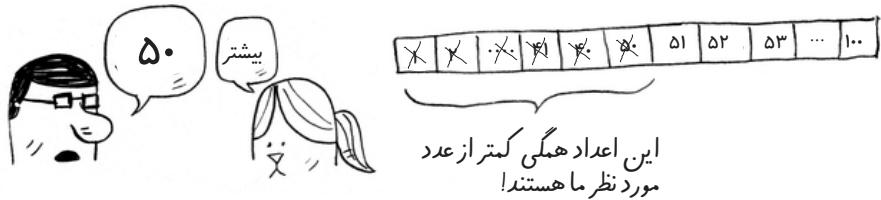
باید بتوانید با کمترین دفعات ممکن عدد مورد نظر را حدس بزنید. در هر نوبت، به شما خواهم گفت که آیا حدس شما کمتر یا بیشتر از عدد مورد نظر است، یا اینکه عدد را درست حدس زده‌اید.

فرض کنید اعدادی که حدس زده اید به این شکل باشد: ۱، ۲، ۳، ۴.... با این روال به پیش خواهید رفت:

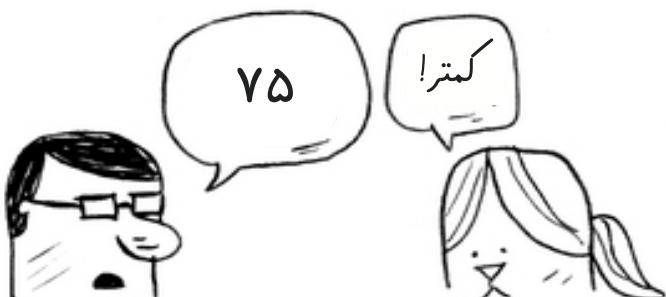


این یک جستجوی ساده است. (شاید جستجوی احمقانه تعبیر درست‌تری باشد). با هر بار حدس زدن، تنها یک شماره را حذف می‌کنید اگر شماره‌ی مورد نظر من ۹۹ باشد، ۹۹ بار طول می‌کشد تا درست حدس بزنید!

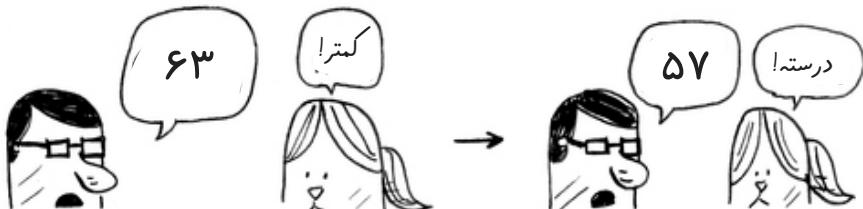
یک راه بهتر برای جستجو
این یکی تکنیک بهتر است. با عدد ۵۰ شروع کنید.



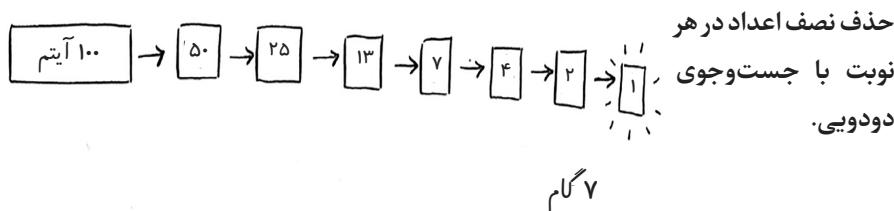
کوچک‌تر از عدد مورد نظر است، با این حال نصف اعداد حذف شدند! حالا می‌دانید که اعداد ۱-۵۰ همگی کمتر هستند. حدس بعدی: ۷۵.



بزرگ‌تر است، اما باز هم نصف اعداد باقی‌مانده را حذف کرده‌اید! با جست‌وجوی دودویی، عدد وسطی را حدس می‌زنید و نصف اعداد باقی‌مانده را در هر نوبت حذف می‌کنید. عدد بعدی ۶۳ است (عدد میانی ۵۰ و ۷۵)



همین الان اولین الگوریتم را یاد گرفتید. به این راه حل جست‌وجوی دودویی گفته می‌شود. در اینجا تعداد اعدادی که در هر نوبت می‌توان حذف کرد به نمایش در آمده است:

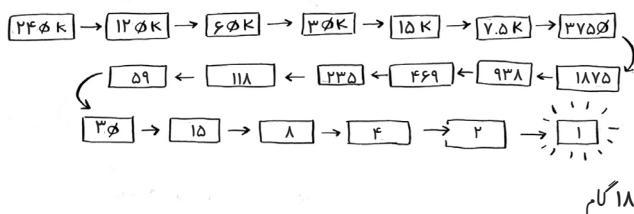


به هر عددی می‌توان با حداقل ۷ حدس دست یافت - چون با هر حدس اعداد بسیاری را حذف می‌کنید!

فرض کنید که به دنبال یک کلمه در دیکشنری می‌گردید. این دیکشنری از ۲۴۰۰۰ کلمه تشکیل شده است. در بدترین وضعیت، فکر می‌کنید چند گام برای هر جست‌وجوی مورد نیاز است؟

گام ————— : جست‌وجوی ساده
گام ————— : جست‌وجوی دودویی

جست‌وجوی ساده در شرایطی که کلمه‌ی مورد نظر آخرين کلمه‌ی دیکشنری باشد ۲۴۰۰۰۰ مرحله خواهد داشت. در حالی که در هر مرحله از جست‌وجوی دودویی، تعداد کلمات را به نصف کاهش می‌دهید تا تنها یک کلمه باقی بماند.



به طور معمول، برای هر فهرستی از n ، جستجوی دودویی $\log_2 n$ مرحله برای بدترین وضعیت نیاز دارد، در حالی که برای جستجوی ساده n مرحله مورد نیاز است.

لگاریتم

شاید مفهوم لگاریتم را فراموش کرده باشد، اما احتمالاً می‌دانید که نمایی (توان) چیست. $\log_{10} 100$ مثل این است که پرسیم، «چند تا ده تا باید در هم ضرب کنیم تا به عدد 100 برسیم؟» جواب 2 است: $10 \times 10 = 100$. بنابراین $\log_{10} 2 = 2$. در واقع لگاریتم بر عکس توان است.

$$\begin{array}{rcl} 10^0 = 1 & \leftrightarrow & \log_{10} 1 = 0 \\ 10^1 = 10 & \leftrightarrow & \log_{10} 10 = 1 \\ \hline 10^2 = 100 & \leftrightarrow & \log_{10} 100 = 2 \\ 10^3 = 1000 & \leftrightarrow & \log_{10} 1000 = 3 \\ \hline 2^0 = 1 & \leftrightarrow & \log_2 1 = 0 \\ 2^1 = 2 & \leftrightarrow & \log_2 2 = 1 \\ \hline 2^2 = 4 & \leftrightarrow & \log_2 4 = 2 \\ 2^3 = 8 & \leftrightarrow & \log_2 8 = 3 \\ \hline 2^4 = 16 & \leftrightarrow & \log_2 16 = 4 \\ 2^5 = 32 & \leftrightarrow & \log_2 32 = 5 \end{array}$$

لگاریتم بر عکس توان است.

در این کتاب هرگاه از زمان اجرا با علامت O بزرگ صحبت می‌کنم (کمی بعد توضیح می‌دهم)، \log همیشه به معنای \log_2 است. وقتی که از راه جستجوی ساده به دنبال عضوی می‌گردید، در بدترین حالت باید به هر عضونگاه کنید. پس برای یک فهرست 8 عددی، حداقل باید 8 عدد را چک کنید. برای جستجوی دودویی، باید در بدترین حالت، $\log n$ عضوراً چک کنید. برای یک فهرست 8 عضوری، $\log 8 = 3$ است. پس برای یک فهرست 8 عضوری، حداقل سه عدد را باید چک کنید. برای لیستی با 10^{24} عضو، $\log 10^{24} = 24$ مورد، چون $10^{24} = 2^{10} \cdot 5^{10}$. پس برای یک فهرست 10^{24} عددی، باید در بدترین حالت 10 عدد را چک کنید.

نکته

در این کتاب به دفعات درباره‌ی $\log \text{time}$ یا زمان لگاریتمی صحبت می‌کنم، پس باید مفهوم لگاریتم را بفهمید. اگر اطلاعات کافی درباره‌ی آن ندارید، خان آکادمی (khanacademy.org) ویدئوی خوبی در مورد لگاریتم دارد که درک این مبحث را ممکن می‌کند.

نکته

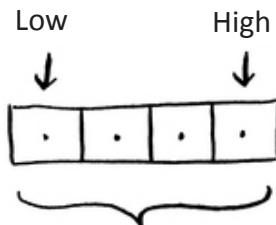
تنها زمانی می‌توان از جست‌وجوی دودویی استفاده کرد که لیست شما مرتب شده باشد. برای مثال، اسامی دفتر تلفن به ترتیب حروف الفبا ردیف شده است، پس می‌توانید از جست‌وجوی دودویی برای پیدا کردن یک نام استفاده کنید. اگر اسامی مرتب نشده بودند چه اتفاقی می‌افتد؟

بیایید با هم نگاهی بیندازیم به اینکه چگونه باید جست‌وجوی دودویی را به زبان پایتون نوشت. در این کد نمونه از آرایه استفاده شده است. اگر نمی‌دانید که آرایه‌ها به چه شکل عمل می‌کنند، نگران نباشید؛ در فصل بعدی به آن می‌پردازیم. فقط باید بدانید که می‌توانید سکانسی از اعضاء را در یک ردیف از باکت یا سطلهای متوالی که آرایه نامیده می‌شود ذخیره کنید. باکت‌ها از شماره‌ی صفر شماره‌گذاری شده‌اند: باکت اول در موقعیت #۰ و باکت دوم #۱ و باکت سوم #۲ است و به همین ترتیب. تابع `binary_search` به عنوان ورودی یک آرایه‌ی مرتب شده و یک آیتم می‌گیرد. اگر آیتم در آرایه بود، تابع موقعیت آن را برمی‌گرداند. بخشی از آرایه را که در آن جست‌وجو

انجام می‌شود را زیر نظر می‌گیریم. در گام نخست، آرایه به این شکل است:

```
low = 0
```

```
high = len(list) - 1
```



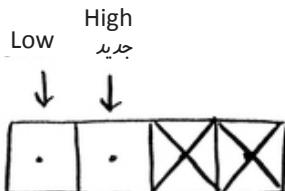
اعدادی که در میان آن‌ها جست‌وجو می‌کنیم.

در هر نوبت، عضو میانی را چک می‌کنید:

اگر حاصل ($low + high$) // 2 روج نباشد، پایتون mid را کاهش گردانی کند.

اگر حدس مورد نظر کمتر از عدد مورد نظر ما باشد، مطابق زیر مقدار low را تغییر می‌دهید:

```
if guess < item:  
    low = mid + 1
```



و اگر حدس مورد نظر بیشتر باشد، high را تغییر می‌دهیم. کد کامل به این شکل است:

```

def binary_search(list, item):
    low = 0
    high = len(list) - 1
    while low <= high:
        mid = (low + high)/2
        guess = list[mid]
        if guess == item:
            return mid
        if guess > item:
            high = mid - 1
        else:
            low = mid + 1
    return None
my_list = [1, 3, 5, 7, 9]
print binary_search(my_list, 3) # => 1
print binary_search(my_list, -1) # => None

```

محدوده‌ی جست‌وجوی شارا مشخص می‌کند.

وقتی دامنه‌ی را به یک عنصر محدود نکرده باشید...

عنصر میانی را بررسی می‌کند

عدد مورد نظر پیدا شد!

از عدد مورد نظر بزرگ‌تر است.

از عدد مورد نظر کوچک‌تر است.

آیتم وجود ندارد.

باید کد بالا را تست کنیم!

به یادداشته باشید، لیست ها از ایندکس صفر شروع می‌شوند.

دومین اسلاط ایندکس یک دارد.

None در پایتون به معنای هیچ است.

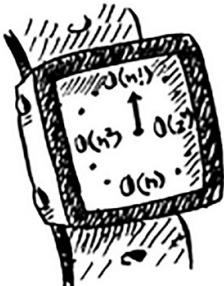
در این حالت آیتم مورد نظر پیدا نشده است.

تمرین

۱.۱ فرض کنید یک لیست مرتب شده از ۱۲۸ نام دارد، و با جست‌وجوی دودویی در میان آن لیست به جست‌وجویی پردازید. حداکثر تعداد مراحل چه مقدار خواهد بود؟

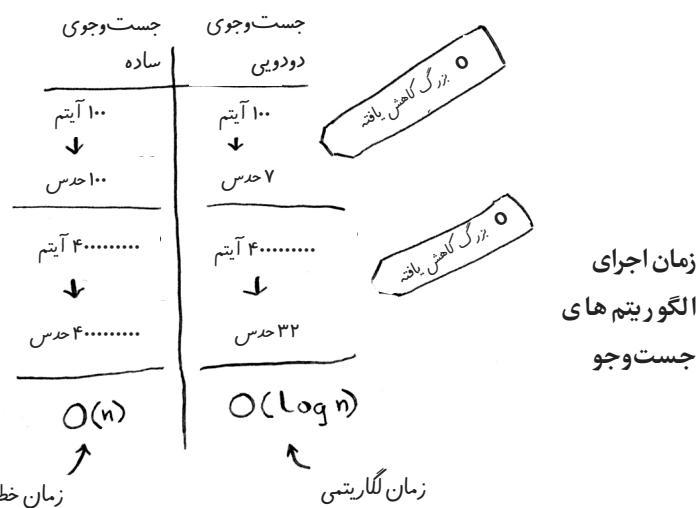
۱.۲ تعداد اعضای لیست را دوبرابر مقدار قبلی در نظر بگیرید. این بار حداکثر تعداد مراحل چقدر خواهد بود؟

زمان اجرا



هرگاه از الگوریتم صحبت می‌کنم، به زمان اجرای آن هم اشاره می‌کنم. به طور معمول می‌خواهید کارآمدترین الگوریتم را برای بهینه‌سازی زمان یا فضا انتخاب کنید. به جست‌وجوی دودویی برگردیم. زمان صرفه‌جویی شده با استفاده از این الگوریتم به چه میزان است؟ خب، اولین رویکرد این بود که هر عدد را یک‌به‌یک بررسی کنیم. اگر فهرستی از 10^5 عدد باشد تا 10^5 حدس نیاز دارد. اگر فهرستی از 4 میلیارد عدد باشد، تا 4 میلیارد حدس نیاز دارد. بنابراین حداکثر تعداد حدس‌ها برابر با اندازه‌ی لیست است. به چنین حالتی زمان خطی می‌گویند.

زمان اجرا برای جست‌وجوی دودویی متفاوت است. اگر لیست 10^5 آیتم داشته باشد، حداکثر 7 حدس نیاز است. اگر لیست 4 میلیارد آیتم داشته باشد، حداکثر 32 دفعات مرتبه می‌شود. عجب الگوریتم کارا و قدرتمندی، مگر نه؟ جست‌وجوی دودویی در زمان لگاریتمی (یا به قول بومی هالگ تایم) اجرا می‌شود. در جدول زیر خلاصه‌ای از یافته‌های امروز ما آمده است.



نماد ۰ بزرگ



۰ بزرگ یک نماد ویره است که به شما می‌گوید الگوریتم مورد نظر تا چه اندازه سریع است. چه اهمیتی دارد؟ خوب، شما اغلب از الگوریتم‌هایی که دیگران نوشته‌اند استفاده می‌کنید - و وقتی این کار را انجام می‌دهید، خوب است که بدانید این الگوریتم‌ها چقدر سریع یا کند هستند. در این بخش توضیح می‌دهم نماد ۰ بزرگ چیست و فهرستی از متداول‌ترین زمان‌های اجرای الگوریتم‌ها را به شما ارائه می‌دهم.

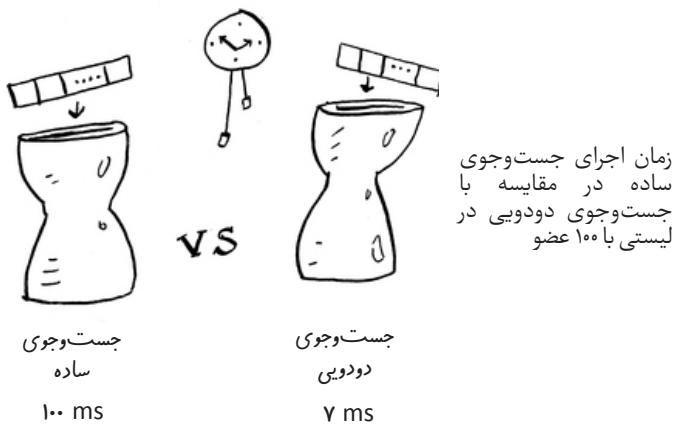
زمان اجرای الگوریتم‌ها با سرعت‌های متفاوتی رشد می‌کند

باب در حال نوشتمن یک الگوریتم جست‌وجو برای ناسا است. الگوریتم او در آستانه‌ی فرود موشک بر روی ماه آغاز می‌شود و به محاسبه‌ی محل فرود کمک می‌کند. این نمونه‌ای است که نشان می‌دهد چگونه زمان اجرای دو الگوریتم می‌تواند با نسبت متفاوت رشد کند.

باب در تلاش است میان جست‌وجوی ساده و جست‌وجوی دودویی تصمیم بگیرد. الگوریتم باید هم‌زمان سریع و صحیح باشد. از یک طرف، جست‌وجوی دودویی سریع‌تر است و باب فقط ۱۰ ثانیه فرصت دارد تا بفهمد کجا باید فرود بیاید - در غیر این صورت، موشک از مسیر خود خارج خواهد شد. از سوی دیگر، نوشتمن جست‌وجوی ساده آسان‌تر است و احتمال بروز مشکلات کمتر است و باب واقعاً نمی‌خواهد باگ‌هایی در کد برای فرود موشک ایجاد شود! برای دقیق‌تر، باب تصمیم می‌گیرد هر دو الگوریتم را بالیستی از ۱۰۰ عنصر زمان‌بندی کند.

بیایید فرض کنیم برای بررسی یک عنصر ۱ میلی ثانیه طول می‌کشد. با جست‌وجوی ساده، باب باید ۱۰۰ عنصر را بررسی کند، بنابراین جستجو ۱۰۰ میلی ثانیه طول می‌کشد تا اجرا شود. از طرف دیگر، او فقط باید ۷ عنصر را با جست‌وجوی دودویی بررسی کند ($\log_2 100$ تقریباً برابر عدد ۷ است)، بنابراین جست‌وجو ۷ میلی ثانیه طول می‌کشد تا

اجرا شود. اما به طور واقع‌بینانه، این لیست بیش از یک میلیارد عنصر خواهد داشت. اگر این طور باشد، جست‌وجوی ساده چقدر طول می‌کشد؟ جست‌وجوی دودویی چقدر؟ قبل از خواندن ادامه‌ی مطلب، مطمئن شوید که برای هر یک از پرسش‌ها پاسخی دارید.



باب جست‌وجوی دودویی را با ۱ میلیارد عنصر اجرا می‌کند و 3^0 میلی‌ثانیه طول می‌کشد ($\log_2 1,000,000 = 20$ حدوداً برابر است با 3^0). با خودش فکر می‌کند « 3^0 ms» تقریباً 3^0 است. جست‌وجوی دودویی حدود ۱۵ برابر سریع‌تر از جست‌وجوی ساده است، زیرا جست‌وجوی ساده 10^6 میلی‌ثانیه با 10^6 عنصر و جست‌وجوی دودویی ۷ میلی‌ثانیه طول کشید. بنابراین جست‌وجوی ساده $= 450 \times 15 = 30$ میلی‌ثانیه طول می‌کشد، درست است؟ بسیار کم‌تر از آستانه‌ی 10 ثانیه‌ای!» باب تصمیم می‌گیرد با یک جست‌وجوی ساده به آن ادامه بدهد. این انتخاب درستی است؟

نه. معلوم شد، باب اشتباه می‌کند. آن هم چه جور. زمان اجرای جست‌وجوی ساده با ۱ میلیارد آیتم ۱ میلیارد میلی‌ثانیه خواهد بود که ۱۱ روز می‌شود! مشکل این است که زمان اجرا برای جست‌وجوی دودویی و جست‌وجوی ساده با سرعت یکسانی رشد نمی‌کند.

جست وجوی ساده	جست وجوی دودویی	
۱۰۰ عنصر	۷ ms	زمان اجرا با
۱۰۰۰۰ عنصر	۱۴ ms	سرعت های بسیار
۱۰ روز	۳۲ ms	متفاوتی رشد می کند!

به عبارتی با افزایش تعداد آیتم‌ها، جست وجوی دودویی برای اجرا اندکی زمان بیشتر برای اجرا نیاز دارد. اما جست وجوی ساده زمان بسیار بیشتری را برای اجرا نیاز دارد. بنابراین با بزرگ‌تر شدن فهرست اعداد، جست وجوی دودویی ناگهان بسیار سریع‌تر از جست وجوی ساده می‌شود. باب گمان می‌کرد جست وجوی دودویی ۱۵ برابر سریع‌تر از جست وجوی ساده است، اما چنین تصوّری درست نیست. اگر لیست دارای ۱ میلیارد آیتم باشد، ۳۳ میلیون بار سریع‌تر است. به همین دلیل است که دانستن مدت زمان اجرای یک الگوریتم کافی نیست - باید بدانید با افزایش اندازه‌ی لیست زمان اجرا چگونه افزایش می‌یابد. اینجاست که نماد ۰ بزرگ وارد می‌شود.

نماد ۰ بزرگ به شما می‌گوید که یک الگوریتم چقدر سریع است. برای مثال، فرض کنید فهرستی با اندازه‌ی n دارید. جست وجوی ساده باید هر عنصر را بررسی کند، بنابراین n عملیات طول خواهد کشید. زمان اجرا در نماد ۰ بزرگ $O(n)$ است. واحد زمان و ثانیه‌ی آن چه شد؟ خبری ازش نیست - ۰ بزرگ سرعت را برابر حسب ثانیه به شما نمی‌گوید. نماد ۰ بزرگ به شما امکان می‌دهد تعداد عملیات را مقایسه کنید. به شما می‌گوید که الگوریتم با چه سرعتی رشد می‌کند.

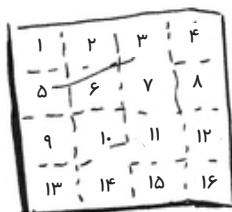


در اینجا مثال دیگری داریم. جست و جوی دودویی برای بررسی لیستی با اندازه‌ی n به $\log n$ عملیات نیاز دارد. زمان اجرا در نماد O بزرگ چقدر است؟ $O(\log n)$. به طور کلی نماد O بزرگ به شکل زیر نوشته می‌شود.



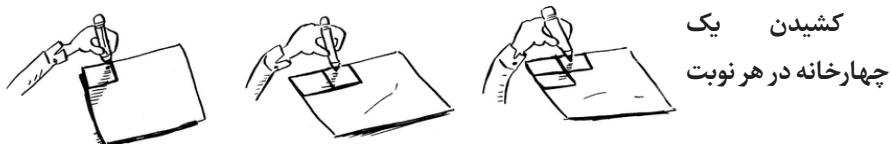
این نماد O بزرگ نامیده می‌شود زیرا شما یک « O بزرگ» را در جلوی تعداد عملیات قرار می‌دهید (به نظر شوخی می‌رسد، اما حقیقت دارد!). و تعداد عملیاتی را که یک الگوریتم انجام می‌دهد به شما می‌گوید. چند مثال را مرور کنیم. ببینید می‌توانید زمان اجرای این الگوریتم‌ها را مشخص کنید.

به تصویر کشیدن چند زمان اجرا با O بزرگ متفاوت
در اینجا یک مثال عملی داریم که می‌توانید در خانه با چند تکه کاغذ و یک مداد دنبال کنید. فرض کنید باید یک چهارخانه‌ی ۱۶ تایی بکشید.



الگوریتم یک

یکی از راه‌های انجام این کار، کشیدن به نوبت ۱۶ چهارخانه، است. به یاد داشته باشید که نماد O بزرگ تعداد عملیات را می‌شمارد. در این مثال، رسم هر چهارخانه برابر با یک عملیات است. شما باید ۱۶ چهارخانه بکشید. کدام الگوریتم برای کشیدن یک چهارخانه در هر بار چند عملیات طول می‌کشد؟ این جدول مناسب است؟

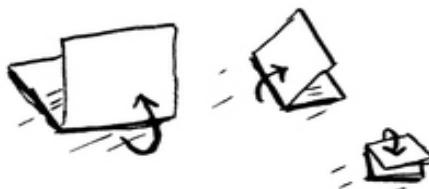


برای کشیدن ۱۶ چهارخانه ۱۶ مرحله طول می‌کشد. زمان اجرای این الگوریتم چقدر است؟

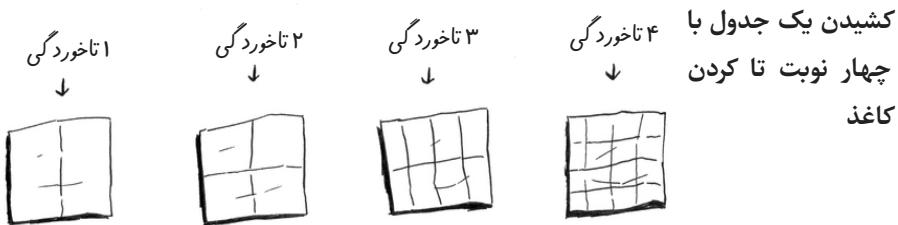
الگوریتم دو
حالا این الگوریتم را امتحان کنید. کاغذ را از وسط تاکنید.



در این مثال، هر بار تا کردن کاغذ یک عملیات است. شما فقط با این عملیات دو چهارخانه درست کردید!
کاغذ را دوباره و دوباره تاکنید.



بعد از چهار بار تا زدن آن را باز کنید و یک جدول زیبا خواهید داشت! هر تا زدن تعداد چهارخانه‌ها را دو برابر می‌کند. شما با ۴ عملیات ۱۶ چهارخانه درست کردید!



شما می‌توانید با هر تا زدن تعداد چهارخانه را دو برابر کنید. در نتیجه می‌توانید ۱۶ چهارخانه را در ۴ مرحله رسم کنید. زمان اجرای این الگوریتم چقدر است؟ قبل از حرکت، زمان‌های اجرای هر دو الگوریتم را مشخص کنید.

پاسخ : الگوریتم اول زمان اجرای $O(n)$ و الگوریتم دوم زمان اجرای $O(\log n)$ دارد.

O بزرگ برای بدترین زمان اجراء

فرض کنید از جستجوی ساده برای جستجوی شخصی در دفترچه تلفن استفاده می‌کنید. می‌دانید که جستجوی ساده زمان اجرای $O(n)$ دارد، به این معنی که در بدترین حالت، باید تک تک ورودی‌های دفترچه تلفن خود را بررسی کنید. در این مورد، شما به دنبال نام Adit هستید. این فرد اولین نام در دفترچه تلفن شما است. بنابراین لازم نبود تا تمام نام‌ها را مرور کنید - در اولین تلاش آن را پیدا کردید. آیا این الگوریتم زمان $O(n)$ را صرف کرده است؟ یا به دلیل اینکه در اولین تلاش فرد مورد نظر را پیدا کردید، زمان $O(1)$ طول کشید؟

جستجوی ساده همچنان به زمان $O(n)$ نیاز دارد. در این صورت، فوراً چیزی را که دنبالش می‌گشتید پیدا کردید. این بهترین سناریو است. اما نماد O بزرگ در مورد بدترین سناریو است. بنابراین می‌توانید بگویید که در بدترین حالت، باید هر نام وارد شده در دفترچه تلفن را یک بار نگاه کنید. زمان آن $O(n)$ است. این برای آگاهی از بدترین حالت ممکن است - می‌دانید که جستجوی ساده هرگز کنیدتر از زمان $O(n)$ نخواهد بود.

نکته

همراه با بدترین زمان اجرا، توجه به زمان اجرای متوسط^۱ هم مهم است. در فصل ۴، بدترین حالت در مقابل حالت متوسط مورد بحث قرار گرفته است.

برخی از زمان اجراهای معمول O بزرگ

در ادامه پنج زمان اجرای O بزرگ که به دفعات با آن‌ها مواجه می‌شوید، از سریع‌ترین به کندترین حالت فهرست شده است:

- $O(\log n)$ ، همچنین به عنوان time log شناخته می‌شود. به عنوان مثال: جست‌وجوی دودویی.

- $O(n)$ ، همچنین به عنوان زمان خطی شناخته می‌شود. به عنوان مثال: جست‌وجوی ساده

- $O(n * \log n)$ ، یک الگوریتم مرتب‌سازی سریع، مانند مرتب‌سازی سریع (که در فصل ۴ آمده است).

- $O(n^2)$ ، یک الگوریتم مرتب‌سازی آهسته، مانند مرتب‌سازی انتخابی (در فصل ۲ آمده است).

- $O(n!)$ ، یک الگوریتم بسیار کند، مانند فروشنده‌ی دوره‌گرد (مورد بعدی که به آن پرداخته می‌شود!).

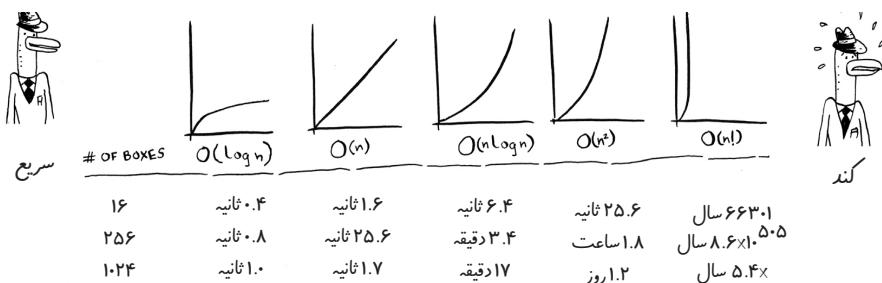
فرض کنید یک بار دیگر جدولی با ۱۶ چهارخانه ترسیم می‌کنید و می‌توانید از الگوریتم مختلف برای این کار استفاده کنید. در صورت انتخاب الگوریتم اول، برای ترسیم شبکه زمان $O(\log n)$ طول می‌کشد و می‌توانید ۱۰ عملیات در ثانیه انجام بدهید.

با زمان $O(\log n)$ ، ۴ عملیات طول می‌کشد تا یک جدول ۱۶ چهارخانه‌ای بکشید ($16 \log 16$ برابر ۴ است). بنابراین $4 \cdot 16 = 64$ ثانیه طول می‌کشد تا جدول را بکشید. اگر بخواهید ۱۰۲۴ چهارخانه رسم کنید چه $2^{10} = 1024$ ثانیه طول می‌کشد. یا ۱ ثانیه برای رسم جدولی از ۱۰۲۴ چهارخانه. این اعداد از الگوریتم اول استفاده می‌کنند.

الگوریتم دوم کندر است: زمان $O(n)$ طول می‌کشد. برای رسم ۱۶ چهارخانه

1. Average-case run time

عملیات و برای رسم 10^{24} چهارخانه 10^{24} عملیات طول می‌کشد. در واقع چند ثانیه؟ در ادامه زمان لازم برای ترسیم کردن جدول از طریق دیگر الگوریتم‌ها، از سریع ترین به کندترین، آمده است:



زمان اجراهای دیگری نیز وجود دارد، اما این پنج مورد رایج‌ترین آن‌ها هستند. البته در واقعیت نمی‌توانید زمان اجرای O بزرگ را به این سرراستی به تعدادی عملیات تبدیل کنید، اما فعلاً تا همینجا کافی است. پس از اینکه چند الگوریتم دیگر را یاد گرفتید، در فصل ۳ به نماد O بزرگ باز خواهیم گشت. نکات اصلی تا به اینجا به شرح زیر است:

- سرعت الگوریتم نه بر اساس ثانیه بلکه بر مبنای رشد تعداد عملیات اندازه‌گیری می‌شود.
- در واقع، مقصود این است که با افزایش اندازه‌ی داده‌های ورودی زمان اجرای یک الگوریتم با چه سرعتی افزایش می‌یابد.
- زمان اجرای الگوریتم‌ها با نماد O بزرگ بیان می‌شود.
- $O(n \log n)$ سریع‌تر از $O(n^2)$ است، اما با رشد لیست مواردی که جستجو می‌کنید، سرعت آن بسیار بیشتر می‌شود.

تمرین

- زمان اجرا برای هر یک از این سناریوها را بر حسب O بزرگ مشخص کنید.
- ۱.۳** نام یک نفر را در نظر دارید و می خواهید شماره تلفن آن شخص را در دفترچه تلفن پیدا کنید.
- ۱.۴** شما یک شماره تلفن دارید و می خواهید نام شخص را در دفترچه تلفن پیدا کنید.
- (نکته: باید تمام صفحات دفترچه را بگردید!)
- ۱.۵** شما می خواهید شماره‌ی هر فرد نوشته شده در دفترچه تلفن را بخوانید.
- ۱.۶** شما می خواهید شماره‌ی افرادی را بخوانید که نام آن‌ها با حرف A شروع شده است. (این مسئله پیچیده است! مفاهیمی را در برمی‌گیرد که در فصل ۴ بیشتر به آن پرداخته می‌شود. احتمالاً با دیدن پاسخ آن متعجب خواهید شد.)

فروشنده‌ی دوره‌گرد

ممکن است با خواندن بخش آخر فکر کرده باشید، «امکان ندارد با الگوریتمی مواجه شوم که $O(n!)$ زمان نیاز داشته باشد.» خوب، اجازه دهید تا به شما ثابت کنم که اشتباه می‌کنید! در ادامه نمونه‌ای از یک الگوریتم با زمان اجرای بسیار بد آمده است. این یک مسئله‌ی معروف در علوم کامپیوتر است، زیرا رشد آن وحشتناک است و بنا بر نظر افراد خبره و باهوش راهی برای بهبود زمان آن وجود ندارد به این مسئله فروشنده‌ی دوره‌گرد می‌گویند.

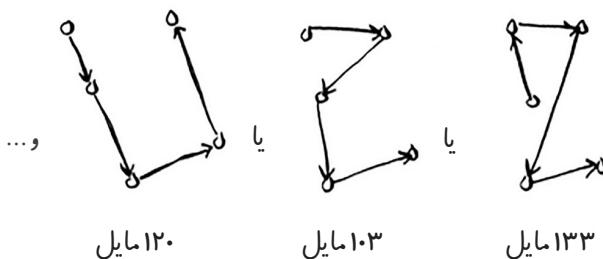


یک فروشنده را فرض کنید.

این فروشنده باید به پنج شهر سفر کند.



این فروشنده، که اپوس صدایش می‌کنم، می‌خواهد به هر پنج شهر برود با این شرط که حداقل مسافت را طی کند. یک روش این است: مشاهدهٔ تمامی حالت‌های ممکن که می‌توان به شهرها سفر کرد.



او کل مسافت را جمع می‌کند و سپس مسیری را که کمترین فاصله را دارد انتخاب می‌کند. ۱۲۰ جایگشت^۱ با ۵ شهر وجود دارد، بنابراین برای حل مسئلهٔ ۵ شهر، به ۱۲۰ عملیات نیاز است. برای ۶ شهر، ۷۲۰ عملیات انجام می‌شود (۷۲۰ جایگشت وجود دارد). برای ۷ شهر، ۵۰۴۰ عملیات طول می‌کشد!

تعداد شهر	تعداد عملیات	تعداد عملیات
۶	۷۲۰	به شدت افزایش می یابد.
۷	۵۰۴۰	
۸	۴۰۳۲۰	
...	...	
۱۵	۱,۳۰۷,۶۷۴,۳۶۸...	
...	...	
۳۰	۲,۶۵۲,۵۲,۸۵۹,۸۱۲,۱۹۱,۰۵۸,۶۳۶,۳۰۸,۴۸۰,....,....	

به طور کلی، برای n آیتم، $n!$ (n فاکتوریل) عملیات برای محاسبه‌ی نتیجه طول می‌کشد. در نتیجه $O(n!)$ یا مرتبه اجرایی فاکتوریل است. تعداد بالای عملیات در تمامی شرایط (به استثنای ارقام بسیار کوچک) صرف خواهد شد. هنگامی که با پیش از ۱۰۰ شهر سروکار دارید، محاسبه‌ی پاسخ در زمان مورد نظر ناممکن است و پیش از اینکه به جواب برسید دنیا به پایان رسیده است.

این یک الگوریتم وحشتناک است! اپوس باید از الگوریتم دیگری استفاده کند، درست است؟ اما اطلاعی از آن ندارد. چون این یکی از مسائل حل نشده در علوم کامپیوتراست. هیچ الگوریتم سریع شناخته شده‌ای برای آن وجود ندارد و متخصص‌ها بر این باورند که دسترسی به یک الگوریتم هوشمند برای این مسئله ناممکن است. بهترین کاری که می‌توان انجام داد این است که یک راه حل تقریبی ارائه بدھیم. برای اطلاعات بیشتر به فصل ۱۰ مراجعه کنید.

یک نکته‌ی پایانی: اگر به مسئله‌های پیچیده‌تر علاقه‌مند هستید، درخت‌های جست‌وجوی دودویی را بررسی کنید! شرح مختصری از آن‌ها در فصل آخر آمده است.

جمع‌بندی

- جست‌وجوی دودویی بسیار سریع‌تر از جست‌وجوی ساده است.
- $O(\log n)$ سریع‌تر از $O(n)$ است، اما زمانی که لیست آیتم‌هایی که در آن جستجو می‌کنید بزرگ شود، بسیار سریع‌تر هم خواهد شد.
- سرعت الگوریتم به ثانیه اندازه‌گیری نمی‌شود.
- زمان الگوریتم بر حسب رشد یک الگوریتم اندازه‌گیری می‌شود.
- زمان الگوریتم با نماد O بزرگ نوشته می‌شود.

۱۲ مرتب‌سازی انتخابی



در این فصل

• دربارهٔ آرایه‌ها و لیست‌های پیوندی - دو تا از اساسی‌ترین ساختمان‌های داده که در همه‌جا از آن‌ها استفاده می‌شوند - یاد می‌گیرید. پیش از این از آرایه‌ها در فصل ۱ استفاده کرده‌اید، و تقریباً در هر یک از فصل‌های این کتاب از آن‌ها استفاده می‌کنید. آرایه‌ها موضوعی حیاتی هستند، پس به آن‌ها توجه کنید! اما گاهی اوقات بهتر است از یک لیست پیوندی استفاده کنید. در این فصل مزایا و معایب هر دو توضیح داده می‌شود تا بتوانید تصمیم بگیرید کدام‌یک برای الگوریتم شما مناسب است.

• اولین الگوریتم مرتب‌سازی خود را یاد می‌گیرید. بسیاری از الگوریتم‌ها تنها در صورتی کار می‌کنند که داده‌های شما مرتب شده باشند. جست‌وجوی دودویی را به یاد دارید؟ جست‌وجوی دودویی را تنها می‌توانید بر یک لیست مرتب شده از عناصر اجرا کنید. در این فصل مرتب‌سازی انتخابی را یاد می‌گیرید. اکثر زبان‌ها دارای یک الگوریتم مرتب‌سازی داخلی و از پیش‌نوشته شده هستند، بنابراین به ندرت نیاز است تا نسخه‌ای از این الگوریتم را از صفر بنویسید. اما مرتب‌سازی انتخابی مقدمه‌ای است بر مرتب‌سازی سریع که در فصل بعدی به آن می‌پردازم. مرتب‌سازی سریع، الگوریتمی مهم است و اگر از قبل یک الگوریتم مرتب‌سازی را بشناسید، درک آن آسان‌تر خواهد شد.

پیش‌نیازها:

برای درک تحلیل کارایی^۱ در این فصل، باید نماد O بزرگ و لگاریتم‌ها را بشناسیم. اگر چیزی درباره‌ی آن‌ها نمی‌دانید، پیشنهاد می‌کنم به عقب برگشته و فصل ۱ را بخوانید. از نماد O بزرگ در سراسر کتاب استفاده می‌شود.

شیوه‌ی عملکرد حافظه

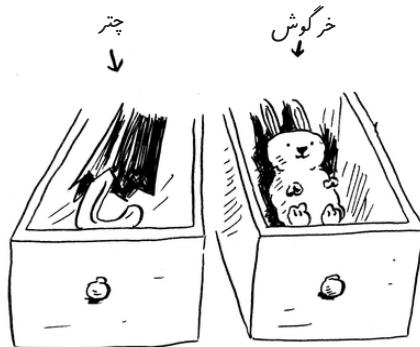
تصور کنید به یک نمایش می‌روید و باید وسایلی را که همراه خود دارید تحویل بدهید. یک گنجه در آنجا قرار گفته است.



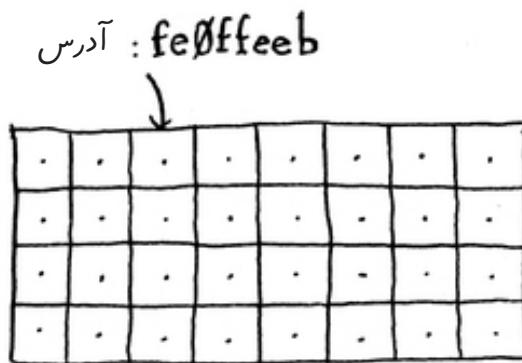
هر یک از کشوها می‌تواند یک وسیله را در خود جای بدهد. دو وسیله‌ی همراه خود را می‌خواهید امانت بدهید، پس دو کشو لازم دارید.



وسایل همراه خود را در این دو کشو قرار می‌دهید.



حالا برای نمایش آماده هستید! اساساً حافظه‌ی کامپیوتر شما این‌گونه عمل می‌کند. کامپیوتر شما شبیه یک مجموعه‌ی بزرگ از کشوها است و هر کشو دارای یک آدرس است.



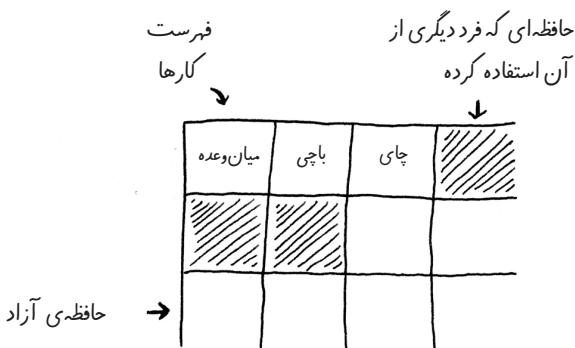
آدرس یک روزنه یا اسلات^۱ در حافظه است.

هر بار که می‌خواهید یک آیتم را در حافظه ذخیره کنید، از کامپیوتر درخواست مقداری فضا می‌کنید و کامپیوتر آدرسی به شما می‌دهد که می‌توانید آیتم خود را در آنجا ذخیره کنید. اگر می‌خواهید چندین آیتم را ذخیره کنید، دو راه اساسی برای این کار وجود دارد: آرایه‌ها و لیست‌ها. در ادامه در مورد آرایه‌ها و لیست‌ها و همچنین مزایا و معایب هر کدام صحبت خواهم کرد. یک راه واحد برای ذخیره‌ی آیتم‌ها در همه‌ی شرایط وجود ندارد، بنابراین مهم است که از تفاوت آن‌ها آگاه باشید.

1. slot

آرایه‌ها و لیست‌های پیوندی

گاهی نیاز است لیستی از عناصر را در حافظه ذخیره کنید. فرض کنید در حال نوشتمن یک برنامه برای مدیریت کارهای روزمره‌ی خود هستید. شما می‌خواهید کارهای انجام شده را به عنوان یک لیست در حافظه ذخیره کنید. باید از یک آرایه استفاده کنید یا یک لیست پیوندی؟ بیایید برای درک آسان‌تر فهرست کارهای روزانه را در یک آرایه ذخیره کنیم. استفاده از آرایه به این معنی است که تمام تسكّع‌های شما به طور پیوسته (در کنار یکدیگر) در حافظه ذخیره می‌شوند.



حالا فرض کنید می‌خواهید یک کار چهارم اضافه کنید. اما کشوی بعدی با وسائل شخص دیگری اشغال شده است!

در این محل که پیش از این اشغال شده،
نمی‌توانید تشك جدیدی اضافه کنید.

↓

میان و عده	باقی	چای	
شکل شده	شکل شده		شکل شده

مثل این است که با دوستان خود به سینما بروید و جایی برای نشستن پیدا کنید - اما دوست دیگری به شما می‌پیوندد و جایی برایش باقی نمانده است. شما باید به یک مکان دیگر بروید که در آن برای همه‌ی شما جا باشد. در این مورد، باید از کامپیوتر خود یک تکه حافظه‌ی متفاوت بخواهید که بتواند چهار تسك را در خود جای بدهد. سپس باید تمام تسك‌های خود را به آنجا منتقل کنید.

اگر دوست دیگری سر برسرد، شما دوباره با مشکل جا مواجه می‌شوید - و همگی باید برای بار دوم جایه جا بنشوید! چه مكافاتی. به همین ترتیب، افزودن موارد جدید به یک آرایه می‌تواند دردرس بزرگی باشد. اگر فضا ندارید و هر بار نیاز دارید که به یک نقطه‌ی جدید در حافظه بروید، افزودن یک آیتم جدید واقعاً کند خواهد بود. یکی از راه حل‌های آسان این است که «صندوق‌های خالی را برای خود نگه دارید»: حتی اگر فقط ۳ آیتم در لیست تسك خود داشته باشید، می‌توانید محض احتیاط از کامپیوتر ۱۰ اسلات بخواهید. سپس می‌توانید ۱۰ آیتم را بدون نیاز به جایه جایی به فهرست وظایف خود اضافه کنید. این یک راه حل خوب است، اما باید از چند جنبه‌ی منفی آن آگاه باشید:

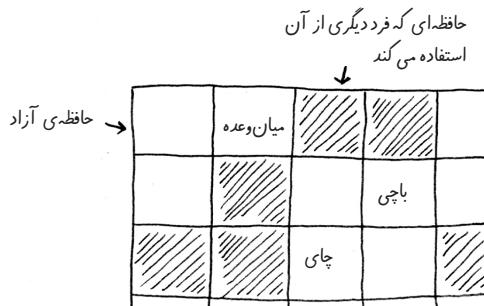
- ممکن است به اسلات‌های اضافی درخواستی نیاز پیدا نکنید و در نتیجه آن حافظه هدر ببرود. شما از آن استفاده نمی‌کنید، اما هیچ کس دیگری نیز نمی‌تواند از آن استفاده کند.

- ممکن است بیش از ۱۰ آیتم را به لیست تسك خود اضافه کنید و به هر حال مجبور به حرکت شوید.

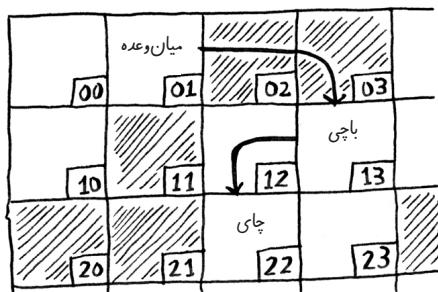
بنابراین این روش راه حل بدی نیست، اما چندان راه حل بی‌نقصی هم نیست. لیست‌های پیوندی این مشکل اضافه کردن آیتم‌ها را حل می‌کند.

لیست‌های پیوندی

با لیست‌های پیوندی، آیتم‌های شما می‌توانند در هر جایی از حافظه باشند.



هر آیتم آدرس آیتم بعدی لیست را ذخیره می‌کند. دسته‌ای از آدرس‌های حافظه‌ی تصادفی^۱ به هم پیوند داده شده‌اند.



آدرس‌های حافظه‌ی پیوندی

درست مثل گنجیابی است. به آدرس اول می‌روید، و گفته می‌شود: «آیتم بعدی را می‌توان در آدرس ۱۲۳ پیدا کرد». پس به آدرس ۱۲۲ می‌روید و می‌گوید: «آیتم بعدی را می‌توان در آدرس ۸۴۷ پیدا کرد»، و غیره. افزودن یک آیتم به لیست پیوندی آسان است: آن را در هر جایی در حافظه قرار می‌دهید و آدرس را با آیتم قبلی ذخیره می‌کنید. با لیست‌های پیوندی، هرگز مجبور نیستید آیتم‌های مورد نظر خود را جابه‌جا کنید. همچنین از یک مشکل دیگر جلوگیری می‌کنید. فرض کنید با پنج نفر از دوستانتان به تماشای یک فیلم محبوب می‌روید. شما شش نفر سعی می‌کنید جایی برای نشستن پیدا کنید، اما سالن پر از تماشاگر شده است. شش صندلی خالی در کنار هم در دسترس نیست. خب، گاهی اوقات این اتفاق با آرایه‌ها می‌افتد. فرض کنید در تلاش برای یافتن

۱۰۰۰۰ روزنه برای یک آرایه هستید. حافظه‌ی شما ۱۰۰۰۰ روزنه دارد، اما این ۱۰۰۰۰ روزنه در کنار یکدیگر نیستند. شما نمی‌توانید فضایی برای آرایه‌ی خود داشته باشید! یک لیست پیوندی مانند این است که بگویید: «بیایید جدا جدا بنشینیم و فیلم را تماشا کنیم». اگر در حافظه فضای خالی موجود باشد، فضایی برای لیست پیوندی خود خواهد داشت. اگر لیست‌های پیوندی برای درج^۱ عملکرد بسیار بهتری دارند، پس آرایه‌ها برای چه کاری مناسب هستند؟

آرایه‌ها

وب‌سایت‌هایی که ۱۰ تا از برترین‌ها را فهرست می‌کنند، از یک تاکتیک شیطانی برای بازدید بیشتر صفحه‌های وب‌سایت استفاده می‌کنند.



گربه‌ی شیطانی

بعدی

به جای اینکه لیست را در یک صفحه به شما نشان بدهند، هر آیتم را در یک صفحه قرار می‌دهند و شما را مجبور می‌کنند روی گزینه‌ی بعدی کلیک کنید تا به مورد بعدی در لیست برسید. به عنوان مثال، برای نمایش لیست ۱۰ تا از شرورترین شخصیت‌های تلویزیونی کل لیست در یک صفحه به شما نمایش داده نمی‌شود. در عوض، از شماره‌ی ۱۰ (نیومن)

شروع می‌شود و باید در هر صفحه روی گزینه‌ی بعدی کلیک کنید تا به شماره‌ی ۱ (گاس فرینگ) برسید. با این تمهد هر مطلب ۱۰ صفحه برای نمایش تبلیغات خواهد داشت. البته که ۹ بار کلیک روی گزینه‌ی بعدی برای رسیدن به سرفهرست کاری خسته‌کننده است و اگر کل لیست در یک صفحه باشد و بتوانید برای اطلاعات بیشتر روی نام هر فرد کلیک کنید، بسیار بهتر است.

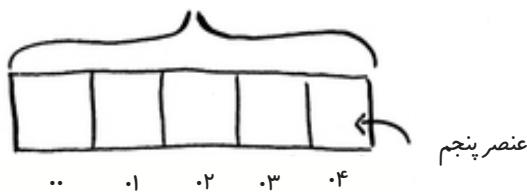
لیست‌های پیوندی نیز مشکل مشابهی دارند. فرض کنید می‌خواهید آخرین آیتم را در یک لیست پیوندی بخوانید. شما نمی‌توانید یک راست سراغ آن رفته و آن را بخوانید، زیرا نمی‌دانید در چه آدرسی قرار گرفته است. در عوض، باید به آیتم شماره‌ی ۱ رفته تا

آدرس آیتم شماره‌ی ۲ را به دست بیاورید.

سپس برای دریافت آدرس آیتم شماره‌ی ۳ باید به آیتم شماره‌ی ۲ بروید. و به همین ترتیب، تا زمانی که به آخرین آیتم برسید. اگر بخواهید همه‌ی آیتم‌ها را در یک نوبت بخوانید، لیست‌های پیوندی عالی هستند: می‌توانید یک آیتم را بخوانید، آدرس آیتم بعدی را دنبال کنید و غیره. اما اگر بخواهید مدام از یک آیتم به آیتمی دیگر بروید، لیست‌های پیوندی اصلاً مناسب نیستند.

آرایه‌ها متفاوت هستند. شما آدرس هر آیتم را در آرایه‌ی خود می‌دانید. برای مثال، فرض کنید آرایه‌ی شما پنج آیتم را شامل می‌شود، و می‌دانید که از آدرس ۰۰ شروع می‌شود. آدرس آیتم شماره‌ی ۵ چیست؟

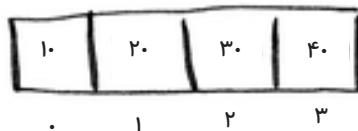
آرایه‌ای از پنج آیتم



با یک جمع و تفریق ساده می‌دانید جواب ۰۴ است. اگر می‌خواهید عناصر تصادفی را بخوانید، آرایه‌ها عالی هستند، زیرا فوراً می‌توانید هر عنصری را در آرایه‌ی خود جست و جو کنید. با یک لیست پیوندی، عناصر در کنار یکدیگر نیستند، بنابراین شما نمی‌توانید درجا موقعیت عنصر پنجم را در حافظه محاسبه کنید - باید به عنصر اول رفته تا آدرس عنصر دوم را دریافت کنید، سپس به عنصر دوم بروید تا آدرس عنصر سوم را دریافت کنید، و به همین ترتیب. تا زمانی که به عنصر پنجم برسید.

اصطلاح‌شناسی

عناصر موجود در یک آرایه شماره‌گذاری شده‌اند. این شماره‌گذاری به جای ۱ از ۰ شروع می‌شود. برای مثال در این آرایه ۲۰ در موقعیت ۱ قرار دارد.



و ۱۰ در موقعیت صفر است. این موضوع معمولاً باعث آشتفتگی برنامه‌نویسان تازه‌کار است. شروع از صفر، نوشتن انواع کدهای مبتنی بر آرایه را آسان‌تر می‌کند، بنابراین برنامه‌نویسان به آن پایبند هستند. تقریباً هر زبان برنامه‌نویسی که استفاده می‌کنید، عناصر آرایه‌ای را که از ۰ شروع می‌شوند شماره‌گذاری می‌کنند. به زودی به آن عادت خواهید کرد.

موقعیت یک عنصر را ایندکس^۱ یا اندیس آن می‌نامند. بنابراین به جای گفتن «۰ در جایگاه ۱ است»، اصطلاح صحیح این است که «۰ در ایندکس ۱ است». در این کتاب از ایندکس به معنای موقعیت یا جایگاه^۲ استفاده می‌کنم.
در اینجا زمان اجرای عملیات رایج در آرایه‌ها و لیست‌ها آمده است.

	آرایه	لیست
(READING) خواندن	۰(۰)	۰(۰)
(INSERTION) درج	۰(۰)	۰(۰)

$$\text{زمان خطی} = O(n)$$

$$\text{زمان ثابت} = O(1)$$

1. index
2. position

سؤال: چرا زمان (n) برای وارد کردن یک عنصر در یک آرایه نیاز است؟ فرض کنید می خواهید عنصری را در ابتدای یک آرایه وارد کنید. چگونه این کار را انجام می دهید؟ چه مدت طول می کشد؟ پاسخ این سوالات را در بخش بعدی بخوانید!

تمرین

۲.۱ فرض کنید در حال ساخت یک برنامه برای پیگیری امور مالی خود هستید.

۱. خورد و خوارآک روزمره

۲. سینما

۳. حق عضویت SFBC

هر روز خرچ‌های خود را یادداشت می کنید. در پایان ماه، هزینه‌ها را بررسی کرده و میزان هزینه‌های خود را جمع می‌زنید. بنابراین، شما تعداد زیادی درج کردن و مقدار کمی خواندن دارید. برای این برنامه باید از میان آرایه و لیست کدام را استفاده کنید؟

درج در وسط یک لیست

فرض کنید می خواهید لیست کارهای روزانه‌ی شما بیشتر شبیه یک تقویم باشد. پیش از این عنوان کارهای روزانه را به انتهای لیست اضافه می کردید. اما حالا می خواهید آن‌ها را با اولویت انجام‌دادن اضافه کنید.

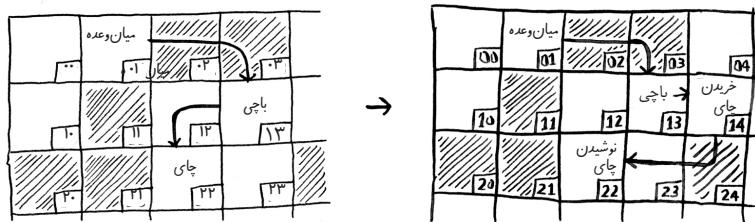


بدون ترتیب

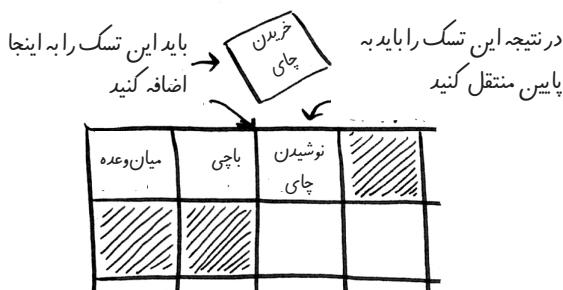


مرتب شده

اگر بخواهید عناصر را در وسط لیست وارد کنید، کدام روش بهتر است: آرایه یا لیست؟ از طریق لیست، آسان است و تنها کافی است آنچه عنصر قبلی به آن اشاره می‌کند، تغییر دهید.



اما با آرایه‌ها، باید بقیه‌ی عناصر را به پایین جایه‌جا کنید.



و اگر فضایی وجود نداشته باشد، ممکن است مجبور شوید همه چیز را در یک مکان جدید کپی کنید! اگر می‌خواهید عناصر را در وسط قرار دهید، لیست‌ها بهتر هستند.

۱ حذف

اگر بخواهید یک عنصر را حذف کنید چه؟ باز هم، لیست‌ها بهتر هستند، زیرا فقط باید آنچه را که عنصر قبلی به آن اشاره می‌کند تغییر دهید. با آرایه‌ها، وقتی یک عنصر را حذف می‌کنید، همه چیز باید به بالای خود منتقل شود.

برخلاف درج، حذف همیشه کارساز است. درج ممکن است گاهی اوقات که فضایی در حافظه باقی نمانده امکان پذیر نباشد. اما همیشه می‌توانید یک عنصر را حذف کنید.

1. Deletions

در اینجا زمان اجرا برای عملیات رایج^۱ در آرایه‌ها و لیست‌های پیوندی آمده است.

	آرایه‌ها	لیست‌ها
خواندن	$O(1)$	$O(n)$
درج کردن	$O(n)$	$O(1)$
حذف کردن	$O(n)$	$O(1)$

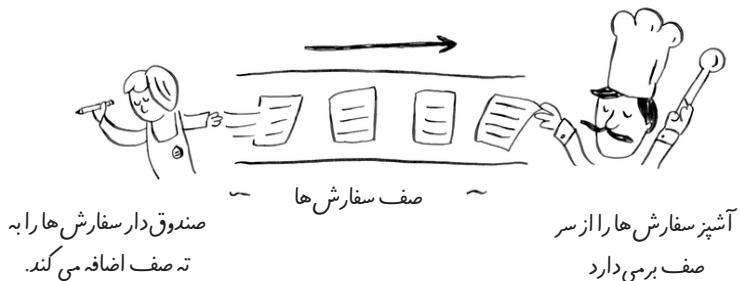
شایان ذکر است که درج و حذف فقط زمانی^(۱) هستند که فوراً بتوانید به عنصری که باید حذف بشود دسترسی پیدا کنید. پیگیری اولین و آخرین آیتم‌ها در یک لیست پیوندی یک روش معمول است، بنابراین حذف آن‌ها تنها به زمان^(۱) نیاز خواهد داشت.

کدام بیشتر استفاده می‌شود: آرایه یا لیست؟ بدیهی است که بستگی به مورد استفاده دارد. آرایه‌ها کاربرد زیادی دارند زیرا امکان دسترسی تصادفی^۲ را فراهم می‌کنند. دو نوع دسترسی وجود دارد: دسترسی تصادفی و دسترسی ترتیبی.^۳ دسترسی ترتیبی به معنای خواندن یک به یک عناصر، از اولین عنصر است. از لیست‌های پیوندی فقط می‌توان برای دسترسی ترتیبی استفاده کرد. اگر می‌خواهید دهمین عنصر یک لیست پیوندی را بخوانید، باید ۹ عنصر اول را بخوانید و پیوندها را تا عنصر دهم دنبال کنید. دسترسی تصادفی به این معنی است که می‌توانید مستقیماً به عنصر دهم بروید. به دفعات از زبان من می‌شنوید که می‌گوییم آرایه‌ها برای عمل خواندن سریع‌تر هستند. به این دلیل که دسترسی تصادفی از ویژگی‌های آن است. بسیاری از موارد استفاده نیاز به دسترسی تصادفی دارند، بنابراین از آرایه در موارد بسیاری استفاده می‌شود. از آرایه و لیست برای پیاده‌سازی سایر ساختمان‌های داده هم استفاده می‌شوند (در ادامه‌ی کتاب آمده است).

1. common operations
2. Random accesss
3. sequential access

تمرین

۲.۲ فرض کنید در حال ساخت برنامه‌ای برای رستوران‌ها هستید تا سفارش مشتری‌ها در آن ثبت بشود. برنامه‌ی شما باید فهرستی از سفارشات را ذخیره کند. گارسون‌ها مدام به این لیست سفارش‌هایی را اضافه می‌کنند، و آشپزها سفارشات را از لیست می‌بینند و آن‌ها را آماده می‌کنند. این یک صف سفارش است: گارسون‌ها سفارشات را به ته صفات اضافه می‌کنند، و آشپزها اولین سفارش را از سر صفت بر می‌دارند و آن را حاضر می‌کنند.

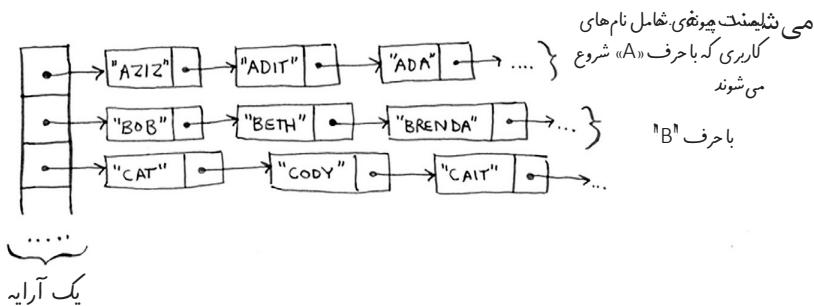


برای پیاده‌سازی این صفت از میان آرایه و لیست پیوندی کدام را استفاده می‌کنید؟ (راهنمایی: لیست‌های پیوندی برای درج/حذف و آرایه‌ها برای دسترسی تصادفی مناسب هستند. کدام یک را در اینجا به کار می‌برید؟)

۲.۳ بیایید یک آزمایش نظری انجام بدھیم. فرض کنید فیسبوک لیستی از نام‌های کاربری را نگه می‌دارد. هنگامی که شخصی سعی می‌کند به فیسبوک وارد بشود، نام کاربری او جست‌وجو می‌شود. اگر نام آن‌ها در لیست نام‌های کاربری باشد، می‌توانند وارد شوند. به دفعات کاربران وارد فیسبوک می‌شوند، بنابراین جست‌وجوهای زیادی از طریق این لیست از نام‌های کاربری انجام می‌شود. فرض کنید فیسبوک از جست‌وجوی دودویی برای جست‌وجوی لیست استفاده می‌کند. جست‌وجوی دودویی نیاز به دسترسی تصادفی دارد. شما باید بتوانید فوراً به وسط لیست نام‌های کاربری برسید. با دانستن این موضوع، آیا لیست را به شکل یک آرایه یا یک لیست پیوندی پیاده‌سازی می‌کنید؟

۲.۴ همچنین کاربرها بیشتر اوقات در فیس بوک ثبت نام می‌کنند. فرض کنید تصمیم گرفتید از یک آرایه برای ذخیره‌ی لیست کاربران استفاده کنید. استفاده از آرایه برای درج چه جنبه‌های منفی به دنبال دارد؟ به طور خاص، فرض کنید از جستجوی دودویی برای لاگین^۱ استفاده می‌کنید. وقتی کاربران جدیدی به آرایه اضافه می‌کنید چه اتفاقی می‌افتد؟

۲.۵ در واقعیت، فیس بوک از آرایه و لیست پیوندی برای ذخیره‌ی اطلاعات کاربر استفاده نمی‌کند. بیایید یک ساختمان داده‌ی ترکیبی را در نظر بگیریم: آرایه‌ای از لیست‌های پیوندی. شما یک آرایه با ۲۶ اسلات دارید. هر اسلات به یک لیست پیوندی اشاره می‌کند. به عنوان مثال، اولین اسلات در آرایه به یک لیست پیوندی اشاره می‌کند که شامل تمامی نام‌های کاربری است که با حرف a شروع می‌شوند. اسلات دوم به یک لیست پیوندی اشاره می‌کند که شامل همه‌ی نام‌های کاربری است که با حرف b شروع



فرض کنید Adit B در فیس بوک ثبت نام کرده است، و شما می‌خواهید او را به لیست اضافه کنید. به اسلات ۱ آرایه می‌روید، به لیست پیوندی برای اسلات ۱ رفته و Adit B را به پایان لیست اضافه می‌کنید. حال، فرض کنید می‌خواهید Zakhir H را جستجو کنید. به اسلات ۲۶ می‌روید، که به لیست پیوندی از همه‌ی نام‌های Z اشاره می‌کند. سپس در آن لیست جستجو می‌کنید تا Zakhir H را پیدا کنید.

این ساختمان داده‌ی ترکیبی را با آرایه‌ها و لیست‌های پیوندی مقایسه کنید. به نسبت آرایه و فهرست پیوندی برای جستجو و درج کنتر است یا سریع‌تر؟ نیازی

نیست زمان اجرای O بزرگ را محاسبه کنید، فقط اینکه ساختمان داده‌ی جدید سریع‌تر است یا کندتر؟

مرتب‌سازی انتخابی



بیایید تمامی این موارد را در کنار هم قرار دهیم تا الگوریتم دوم را یاد بگیریم: مرتب‌سازی انتخابی. برای درک مطالب این بخش، باید آرایه‌ها و لیست‌ها و همچنین نماد O بزرگ را بلد باشید.

فرض کنید تعدادی آهنگ در کامپیوتر خود دارید. در کنار نام هر هنرمند، تعداد دفعات پخش قطعات موسیقی هر هنرمند فهرست شده است.

	دفعات پخش
RADIOHEAD	۱۵۶
KISHORE KUMAR	۱۴۱
THE BLACK KEYS	۳۵
NEUTRAL MILK HOTEL	۹۴
BECK	۸۸
THE STROKES	۶۱
WILCO	۱۱۱

می‌خواهید این فهرست را از بیشترین به کمترین تعداد پخش مرتب کنید تا بتوانید هنرمندان مورد علاقه‌ی خود را رتبه‌بندی کنید. چگونه چنین کاری امکان‌پذیر است؟ یکی از راه‌ها این است که لیست را مرور کرده و هنرمندی را که بیشترین پخش را داشته، پیدا کنید. آن هنرمند را به یک لیست جدید اضافه کنید.

دفات پخش	۵۵ به ترتیب
RADIOHEAD	۱۵۶
KISHORE KUMAR	۱۴۱
THE BLACK KEYS	۳۵
NEUTRAL MILK HOTEL	۹۴
BECK	۸۸
THE STROKES	۶۱
WILCO	۱۱۱



۱. رادیوه‌دار همیشتر

پخش شده...

دفات پخش	۵۵ به ترتیب
RADIOHEAD	۱۵۶

۲. به یک لیست

جدید اضافه کنید

دوباره این کار را انجام دهید تا هنرمندی که بیشترین بار پخش شده است را پیدا

کنید.

دفات	پخش	دفات	پخش
RADIOHEAD	۱۵۶	RADIOHEAD	۱۵۶
KISHORE KUMAR	۱۴۱	KISHORE KUMAR	۱۴۱
THE BLACK KEYS	۳۵		
NEUTRAL MILK HOTEL	۹۴		
BECK	۸۸		
THE STROKES	۶۱		
WILCO	۱۱۱		



۱. کیشور کumar در فهرست

پرشونده ترین خواننده رتبه

دو را دارد

۲. پس هنرمند بعدی

است که به لیست جدید

اضافه می‌شود

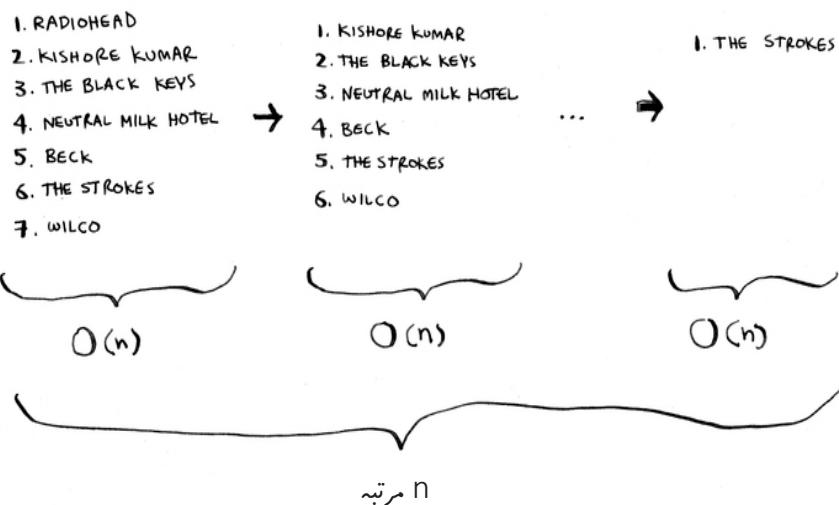
این کار را ادامه دهید و در نهایت یک لیست مرتب شده خواهد شد.

PLAY COUNT	۵۵ به ترتیب
156	RADIOHEAD
141	KISHORE KUMAR
111	WILCO
94	NEUTRAL MILK HOTEL
88	BECK
61	THE STROKES
35	THE BLACK KEYS

بیایید عینک علوم کامپیوتری خود را به چشم بزنیم و بینیم این کار چقدر طول می‌کشد. به یاد داشته باشید که زمان $O(n)$ به این معنی است که شما با هر عنصر لیست یک بار کار دارید. به عنوان مثال، جستجوی ساده روی لیست هنرمندان به معنای یک بار بررسی هر هنرمند است.

1. RADIOHEAD
 2. KISHORE KUMAR
 3. THE BLACK KEYS
 4. NEUTRAL MILK HOTEL
 5. BECK
 6. THE STROKES
 7. WILCO
- }
 n
آیتم

برای پیدا کردن هنرمندی که بیشترین تعداد پخش را دارد، باید هر کدام را بررسی کنید. همان‌طور که مشاهده کردید، زمان $O(n)$ طول می‌کشد. بنابراین شما عملیاتی دارید که زمان $O(n)$ نیاز دارد و باید آن را n بار انجام دهید:



این زمان $O(n \times n)$ یا زمان $O(n^2)$ طول می‌کشد.

الگوریتم‌های مرتب‌سازی بسیار مفید هستند. حالا می‌توانید موارد زیر را مرتب کنید:

- نام‌ها در دفترچه تلفن
- تاریخ سفرها
- ایمیل‌ها (جدیدترین به قدیمی‌ترین)

هر بار عناصر کمتری را بررسی کنید

شاید با خودتان فکر کنید: همین که پیش می‌روید، تعداد عناصری که باید بررسی کنید به طور دائم کاهش می‌یابد. در نهایت، تنها باید یک عنصر را بررسی کنید. پس چگونه زمان اجرا همچنان $O(n^2)$ است؟ پرسش خوبی است و پاسخ آن به ثابت‌ها در نماد $O(\text{بزرگ برمی‌گردید})$. در فصل ۴ بیشتر به این موضوع خواهم پرداخت، اما اصل ماجرا به این صورت است:

درست است که لازم نیست هر بار لیستی از n عنصر را بررسی کنید. شما n عنصر را بررسی می‌کنید، سپس $1 - n - 2 - \dots - n$ به طور متوسط، لیستی را بررسی می‌کنید که دارای $n \times \frac{1}{2}$ عنصر است. زمان اجرا $O(n \times n \times \frac{1}{2})$ است. اما ثابت‌هایی مانند $O(\text{بزرگ نادیده گرفته می‌شوند})$ (تکرار می‌کنم برای توضیحات کامل به فصل ۴ مراجعه کنید)، بنابراین فقط $O(n \times n)$ یا $O(n^2)$ می‌نویسید.

مرتب‌سازی انتخابی الگوریتمی منظم است، اما چندان سریع نیست. مرتب‌سازی سریع، الگوریتم مرتب‌سازی سریع‌تری است که فقط زمان $O(n \log n)$ خواهد داشت. در فصل بعدی به این الگوریتم پرداخته می‌شود!

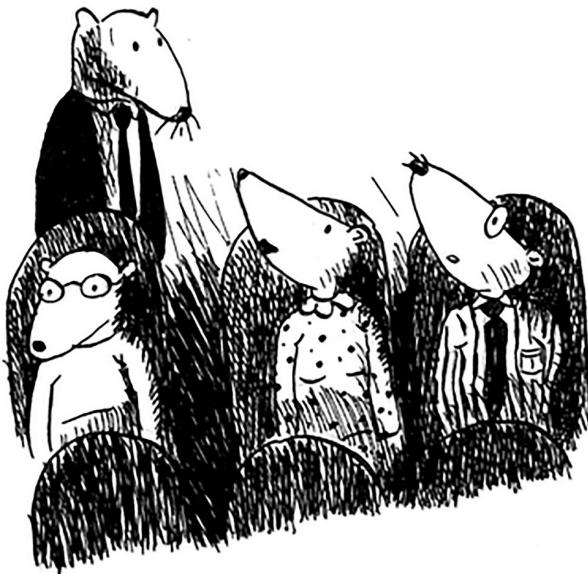
نمونه کد لیست

کد مرتب‌سازی انتخابی مثال فهرست موزیک‌ها را ننوشتیم، اما کدی که در زیر آمده است کار بسیار مشابهی را انجام می‌دهد: مرتب‌سازی آرایه از کوچک‌ترین به بزرگ‌ترین. بیایید تابعی بنویسیم تا کوچک‌ترین عنصر یک آرایه را پیدا کند:

```
def findSmallest(arr):
    smallest = arr[0] ← کوچک‌ترین مقدار را ذخیره می‌کند
    smallest_index = 0 ← ایندکس کوچک‌ترین مقدار را ذخیره می‌کند
    for i in range(1, len(arr)):
        if arr[i] < smallest:
            smallest = arr[i]
            smallest_index = i
    return smallest_index
```

اکنون می‌توانید از این تابع برای نوشتن مرتب‌سازی انتخابی استفاده کنید:

```
def selectionSort(arr): ← یک آرایه را مرتب می‌کند
    newArr = []
    for i in range(len(arr)):
        smallest = findSmallest(arr) ← کوچک‌ترین عنصر آرایه را پیدا
        newArr.append(arr.pop(smallest)) ← کرده و آن را به آرایه جدید اضافه می‌کند.
    return newArr
print selectionSort([5, 3, 6, 2, 10])
```



جمع‌بندی

- حافظه‌ی کامپیوتر شما مانند مجموعه‌ی بزرگی از کشوها است.
- برای ذخیره‌ی چند عنصر، از یک آرایه یا یک لیست استفاده کنید.
- با یک آرایه، تمام عناصر شما درست در کنار یکدیگر ذخیره می‌شوند.
- با یک لیست، عناصر در همه‌جا پراکنده می‌شوند و یک عنصر آدرس عنصر بعدی را ذخیره می‌کند.
- آرایه‌ها اجازه‌ی خواندن سریع را می‌دهند.
- لیست‌های پیوندی اجازه‌ی درج و حذف سریع را می‌دهند.
- همه‌ی عناصر موجود در آرایه باید از یک نوع^۱ باشند (همه‌ی آن‌ها اعداد صحیح^۲ یا دوبل^۳ یا موارد دیگر باشند).

1. type
2. integer
3. double

بازگشت^۱



در این فصل

• بازگشت را باد می‌گیرید. بازگشت یک تکنیک کدنویسی است که در بسیاری از الگوریتم‌ها استفاده می‌شود. بازگشت یک زیربنا و پیش‌نیاز برای درک فصل‌های بعدی این کتاب است.

• باد می‌گیرید که چگونه یک مسئله را به یک صورت پایه^۲ و یک صورت بازگشتنی^۳ تقسیم کنید. استراتژی تقسیم و حل (فصل^۴) از این مفهوم ساده برای حل مسئله‌های دشوار استفاده می‌کند.

برای این فصل هیجان‌زده هستم چون موضوع آن تابع بازگشتی است؛ روشی زیبا برای حل مسئله. بازگشت یکی از موضوعات مورد علاقه‌ی من است، با این حال کدنویسی به روش بازگشت همیشه محل اختلاف نظر بوده است. برنامه‌نویس‌ها یا آن را دوست دارند یا از آن متنفرند، و یا از آن چند سالی متنفرند تا این‌که متوجه می‌شوند دوستش دارند. من شخصاً از جبهه‌ی سوم هستم. برای اینکه کار را برای شما راحت‌تر کنم، چند توصیه دارم:

• در این فصل نمونه کدهای زیادی آورده شده است. این کدها را برای خودتان اجرا

1. recursion
2. Base case
3. Recursive case

کنید تا بینید چگونه کار می‌کند.

- همچنین درباره‌ی توابع بازگشتی صحبت خواهم کرد. برای حداقل یک بار، با قلم و کاغذ به یک تابع بازگشتی پا بگذارید: کاری شبیه به: «خب، من ۵ را به factorial می‌دهم، و سپس حاصل ضرب ۵ در factorial ۴ را برمی‌گردم که می‌شود ...»، و به همین ترتیب. بررسی گام‌به‌گام یک تابع به این شکل به شما می‌آموزد عملکرد یک تابع بازگشتی^۱ به چه شکل است.

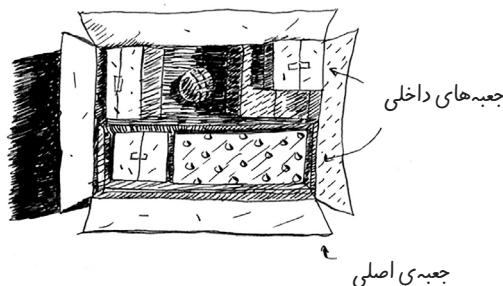
این فصل شامل تعداد متعددی شبیه‌کد^۲ است. شبیه‌کد یک توضیح سطح بالا از مسئله‌ای است که در تلاش برای حل آن هستید. مانند کد نوشته می‌شود، اما به گفتار انسان نزدیک‌تر است.

بازگشت

فرض کنید مشغول زیر و رو کردن وسایل اتاق زیر شیروانی مادریزگران هستید و با یک چمدان قفل شده‌ی مرموز روبه‌رو می‌شوید.

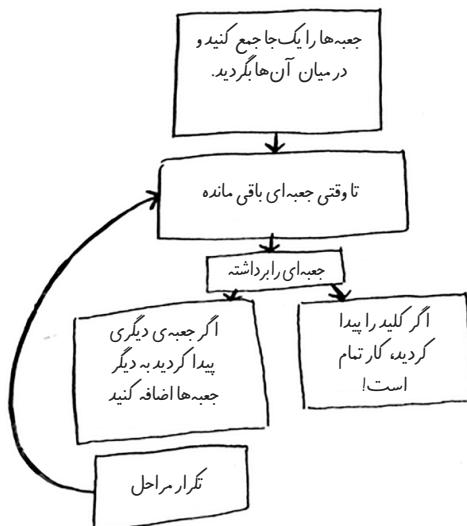


مادریزگ به شما می‌گوید که احتمالاً کلید چمدان در این یکی جعبه است.



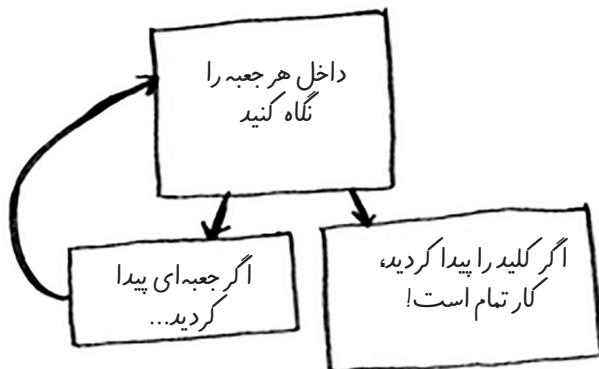
1. Recursive function
2. Pseudocode
3. High-level

این جعبه حاوی جعبه‌های بیشتری است، و داخل آن جعبه‌ها هم جعبه‌های دیگری هست. کلید داخل یکی از جعبه‌ها است. الگوریتم شما برای جستجوی کلید چیست؟ قبل از خواندن ادامه‌ی کتاب به دنبال یک الگوریتم باشید و به آن فکر کنید. یک رویکرد به این شکل است:



۱. جعبه‌ها را یک‌جا جمیع کنید تا در میان آن‌ها بگردید.
۲. یک جعبه را بردارید و داخل آن را نگاه کنید.
۳. اگر جعبه‌ی دیگری پیدا کردید به دیگر جعبه‌ها اضافه کنید تا بعداً بررسی کنید.
۴. اگر کلیدی پیدا کردید، کار شما تمام شده است!
۵. تکرار مراحل.

در ادامه رویکرد جایگزین آمده است:



۱. داخل جعبه رانگاه کنید.

۲. اگر یک جعبه پیدا کردید به مرحله‌ی اول بروید.

۳. اگر کلیدی پیدا کردید، کار شما تمام شده است!

به نظر شما کدام رویکرد راحت‌تر است؟ در رویکرد اول از حلقه‌ی `while` استفاده می‌شود. تا زمانی‌که جعبه‌ای باقی مانده، یک جعبه را بردار و داخلش را بگرد:

```

def look_for_key(main_box) :
    pile = main_box.make_a_pile_to_look_through()
    While pile is not empty :
        box = pile.grab_a_box()
        for item in box:
            if item.is_a_box():
                pile.append(item)
            elif item.is_a_key():
                print "found the key"!

```

در روش دوم از بازگشت استفاده می‌شود. بازگشت نقطه‌ای است که یک تابع خود را فراخوانی می‌کند. در ادامه روش دوم و در قالب شبکه‌کد نوشته شده است:

```
def look_for_key(box):
    for item in box :
        if item.is_a_box():
            look_for_key(item)      ← بازگشت!
        elif item.is_a_key:
            print"found the key"
```

هر دو رویکرد به یک نتیجه ختم می‌شوند، اما رویکرد دوم برای من گویاتر و خوانانter است. کاربرد بازگشت در واضح ترشدن راه حل است. هیچ مزیتی در عملکرد استفاده از بازگشت وجود ندارد. در واقع، حلقه‌ها گاهی اوقات عملکرد بهتری دارند. من این نقل قول لی کالدول در Stack Overflow را دوست دارم: «به کارگیری حلقه‌ها می‌تواند کارایی برنامه‌ی شما را بالا ببرد. به کارگیری بازگشت می‌تواند کارایی برنامه‌نویسی شما را افزایش بدهد. انتخاب با شمامست که در موقعیت شما کدام اهمیت بیشتری دارد!»¹ بسیاری از الگوریتم‌های مهم از بازگشت استفاده می‌کنند، بنابراین درک مفهوم آن مهم است.



صورت پایه و صورت بازگشته

از آنجا که یک تابع بازگشتی خودش را فراخوانی می‌کند، نوشتن یک تابع که به اشتباه به یک حلقه‌ی بی‌نهایت ختم بشود محتمل است.

به عنوان مثال، فرض کنید می‌خواهید تابعی بنویسید که یک شمارش معکوس چاپ کند، مانند این:

> ۳....۲...۱

به صورت زیر می‌توان آن را بازگشتی نوشت:

```
def countdown(i):
    print i
    countdown(i-1)
```

این کد را بنویسید و اجرا کنید. متوجه مشکلی خواهید شد: این تابع تا ابد اجرا

1. <http://stackoverflow.com/a/72694/139117>.

خواهد شد!



> ۳...۲...۱...۰...-۱...-۲...

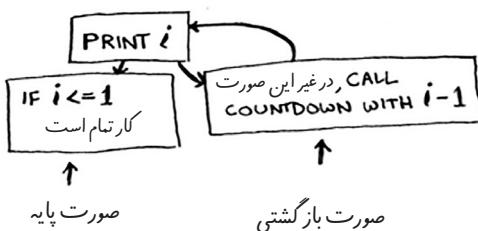
(برای توقف اسکریپت Ctrl-C را فشار بدهید.)

هنگام نوشتن یک تابع بازگشته، باید مشخص کنید که تابع چه زمانی بازگشت را متوقف کند. به همین دلیل است که هر تابع بازگشته دارای دو بخش است: صورت پایه و صورت بازگشته. صورت بازگشته زمانی است که تابع خود را فراخوانی می‌کند. صورت پایه زمانی است که تابع فراخوانی خود را متوقف می‌کند. بنابراین در یک حلقه‌ی بی‌پایان گرفتار نمی‌شود.

بیایید یک صورت پایه به تابع شمارش معکوس اضافه کنیم:

```
def countdown(i):
    print i
    if i <= 1:           ← صورت پایه
        return
    else:               ← صورت بازگشته
        countdown(i-1)
```

اکنون تابع همان طور که انتظار می‌رود کار می‌کند. چیزی شبیه به این:



پُشته^۱



در این بخش به پُشته‌فراخوانی^۲ می‌پردازیم. پُشته‌فراخوانی یک مفهوم مهم در برنامه‌نویسی است و درک آن هنگام استفاده از بازگشت نیز مهم است. فرض کنید در تدارک برگزاری یک مهمانی باربیکیو هستید. شما لیستی از مجموعه کارهایی که برای این مهمانی باید انجام داده شود به شکل پُشته‌ای از کاغذهای یادداشت چسبدار آماده کرده‌اید.

زمانی را به خاطر بیاورید که در مورد آرایه‌ها و لیست‌ها صحبت می‌کردیم، و شما لیستی از کارهای روزانه داشتید؟ می‌توانستید آیتم کارهای روزانه را در هر نقطه از لیست اضافه یا آیتمی را تصادفی از آن حذف کنید. روش پُشته از کاغذهای یادداشت چسبدار بسیار ساده‌تر است. هنگامی که یک آیتم را وارد می‌کنید، آن آیتم به بالای لیست اضافه می‌شود. وقتی آیتمی را می‌خوانید، فقط آیتم رویی را می‌خوانید و از لیست حذف می‌شود. بنابراین بر لیست کارها فقط دو عمل اعمال می‌شود: *push* (درج) و *pop* (حذف و خواندن).



PUSH

(اضافه کردن یک آیتم بر روی بقیه آیتم‌ها)

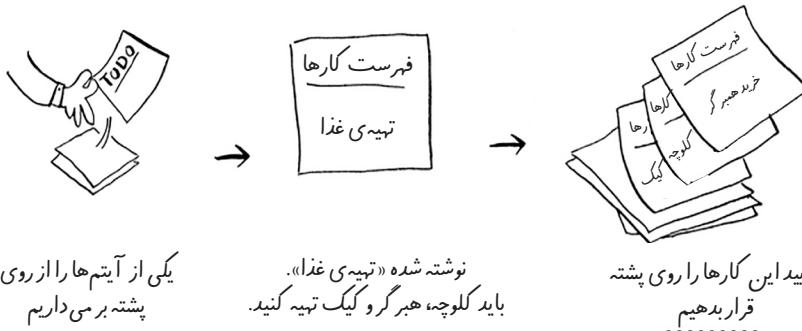


POP

(برداشتن بالاترین آیتم و خواندن آن)

1. The stack
2. Call stack

بیایید لیست کارها را در عمل ببینیم.



به این ساختمان داده، پشتہ می‌گویند. پشتہ یک ساختمان داده‌ی ساده است. شما تا الان بدون اینکه متوجه باشید از پشتہ استفاده کرده‌اید!

پشتہ فراخوانی

کامپیوچر شما از پشتہ‌ای به نام پشتہ فراخوانی استفاده می‌کند. بیایید آن را در عمل ببینیم. در اینجا تابع ساده‌ای وجود دارد:

```
def greet(name):
    print "hello" + name + "!"
    greet2(name)
    print "getting ready to say bye..."
    bye()
```

کار این تابع خوش‌آمدگویی به شما و فراخوانی دو تابع دیگر است. این دو تابع در ادامه آمده است:

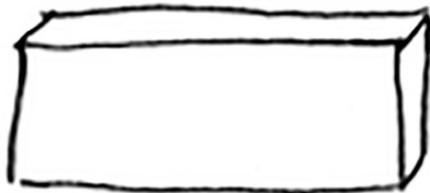
```
def greet2(name):
    print "how are you , " + name + "?"
def bye():
    print "ok bye!"
```

بیایید به آن‌چه هنگام فراخوانی یک تابع رخ می‌دهد، پردازیم.

نکته

یک تابع در پایتون است، اما برای سادگی کار در این مثال، وانمود می‌کنیم که این طور نیست. فقط تظاهر می‌کنیم.

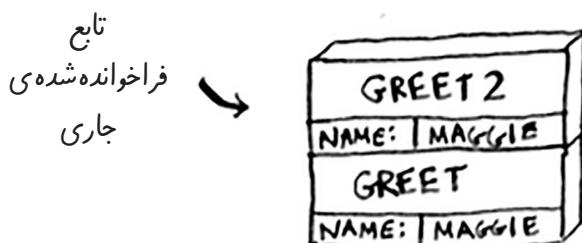
فرض کنید تابع `greet("maggie")` را فراخوانی می‌کنید. ابتدا، کامپیوتر شما یک جعبه حافظه برای آن فراخوانی تابع اختصاص می‌دهد.



حالا بیایید از حافظه استفاده کنیم. به متغیر `name` مقدار "Maggie" داده شده است. این عمل باید در حافظه ذخیره شود.



هر بار که یک تابع را فراخوانی می‌کنید، کامپیوتر شما مقادیر همهٔ متغیرهای آن فراخوانی را به این شکل در حافظه ذخیره می‌کند. بعد، شما `hello, maggie!` را پرینت می‌کنید. بعد (`greet2("maggie")`) را صدای زنید. مجدداً، کامپیوتر شما یک جعبه حافظه برای این فراخوانی اختصاص می‌دهد.



کامپیوتر شما از یک پشته برای این جعبه‌ها استفاده می‌کند. جعبه‌ی دوم بر روی جعبه‌ی اول اضافه می‌شود. شما how are you, maggie? را پرینت می‌کنید سپس از فراخوانی تابع بر می‌گردید. وقتی این اتفاق می‌افتد، جعبه‌ی بالایی پشته بیرون گذاشته می‌شود.



اکنون بالاترین جعبه در پشته تابع greet است، به این معنی که شما به تابع greet بازگشته‌اید. وقتی تابع greet² را فراخوانی کردید، تابع greet تا حدودی تکمیل شده بود. این ایده‌ی اصلی این بخش است: وقتی تابعی را از تابع دیگری فراخوانی می‌کنید، تابع فراخوانی در حالت نسبتاً تکمیل شده متوقف می‌شود. تمام مقادیر متغیرهای آن تابع هنوز در حافظه ذخیره می‌مانند. اکنون که کار با تابع greet² تمام شد، به تابع greet بازگشته و از جایی که کار را متوقف کرده‌اید ادامه می‌دهید. ابتدا تابع greet را پرینت می‌کنید... تابع bye را صدای زنید.



جعبه‌ای برای آن تابع به بالای پشته اضافه می‌شود. سپس `ok bye!` را پرینت می‌کنید و از فراخوانی تابع باز می‌گردید.



و به تابع `greet` بازگشته‌اید. کار دیگری نمانده است، بنابراین از تابع `greet` هم برمی‌گردید. این پشته که برای ذخیره‌ی متغیرها برای چندین تابع استفاده می‌شود، پشته‌فراخوانی نامیده می‌شود.

تمرین

۳.۱ فرض کنید من یک پشته‌فراخوانی مانند این را به شما نشان می‌دهم.



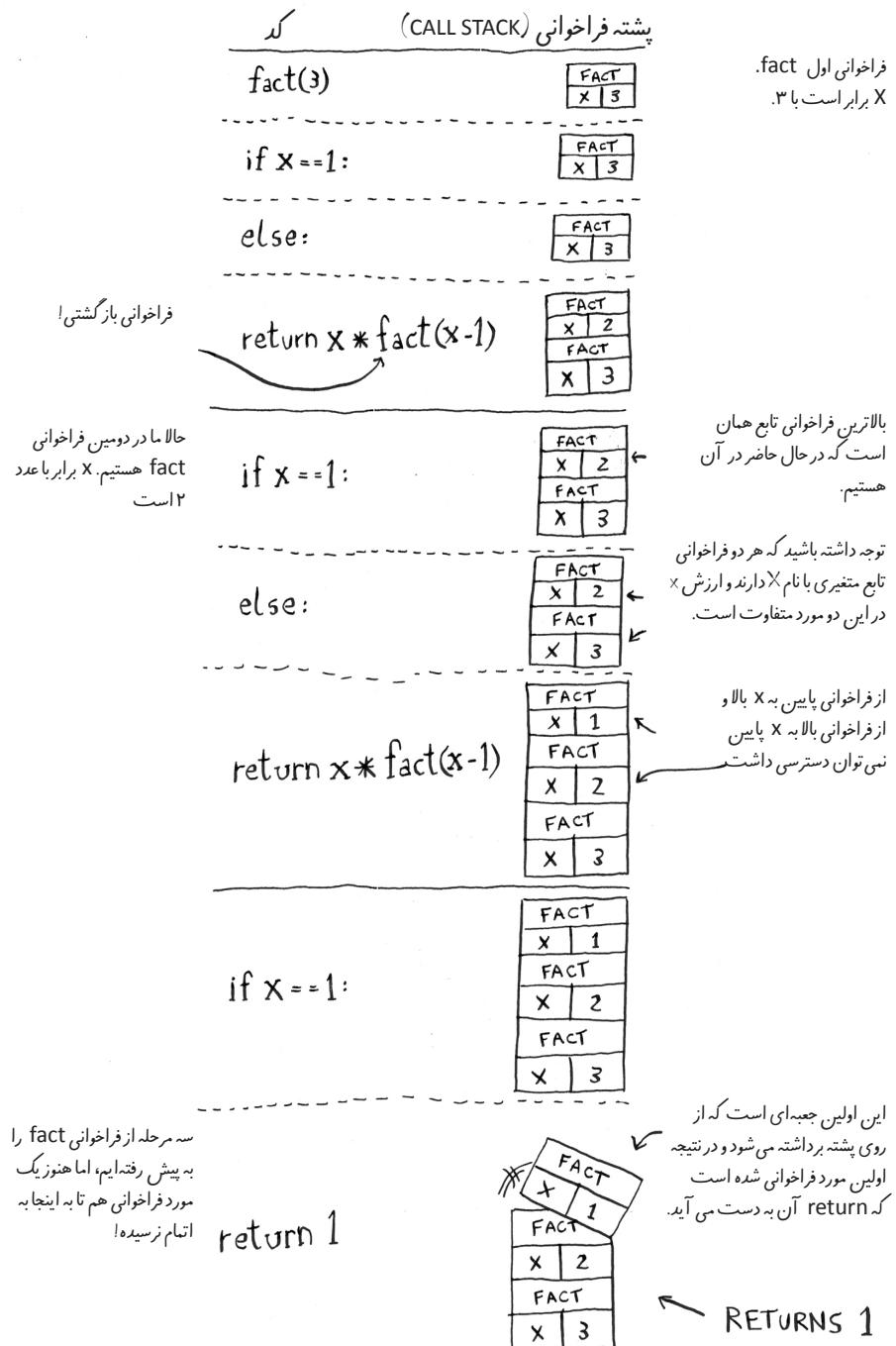
تنها بر اساس این پشته‌فراخوانی، چه اطلاعاتی می‌توانید بدھید؟ حالا باید پشته‌فراخوانی را در عمل و با یک تابع بازگشته ببینیم.

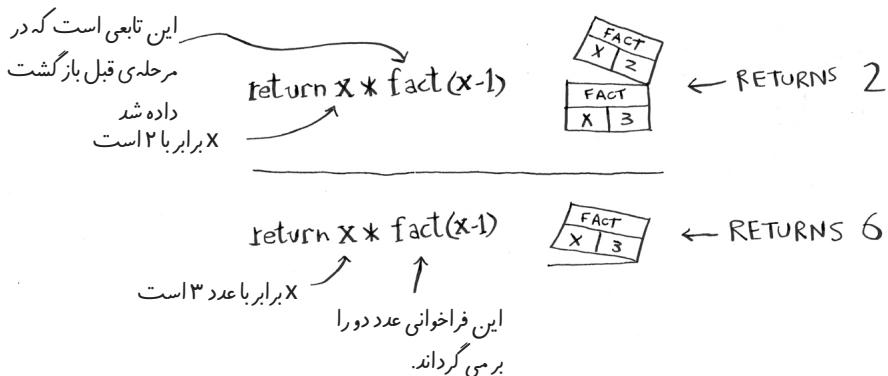
پشته فراخوانی با بازگشت

توابع بازگشتی از پشته فراخوانی هم استفاده می‌کنند! بیایید در عمل با تابع فاکتوریل به این موضوع نگاه کنیم. (۵) factorial به صورت $5! = 5 * 4 * 3 * 2 * 1$ نوشته می‌شود و به این صورت تعریف می‌شود: $1 = 1$. به همین شکل (۳) factorial است $2! = 2 * 1$. در اینجا یک تابع بازگشتی برای محاسبه‌ی فاکتوریل یک عدد وجود دارد:

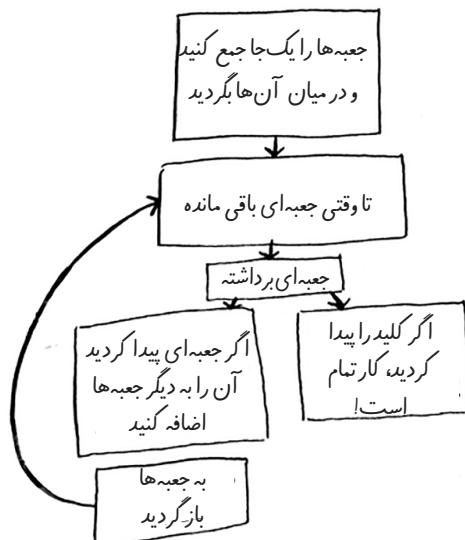
```
def fact(x):
    if x == 1:
        return 1
    else:
        return x * fact(x-1)
```

حالا شما (۳) fact را صدای زنید. بیایید خط به خط این فراخوانی را بررسی کنیم و ببینیم که چگونه پشته تغییر می‌کند. به یاد داشته باشید، بالاترین جعبه در پشته به شما می‌گوید که در حال حاضر در داخل کدام فراخوانی از fact هستید.



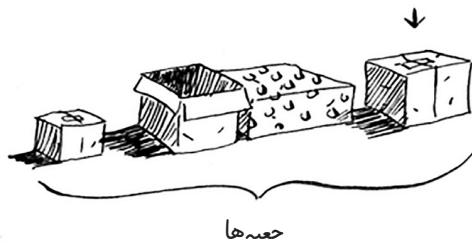


توجه داشته باشید که هر تماس یا فراخوانی fact کپی خود از x را دارد. شما نمی‌توانید به کپی دیگری از تابع x دسترسی پیدا کنید.
 پشتنه نقش مهمی در بازگشت بازی می‌کند. در مثال ابتدایی، دو رویکرد برای یافتن کلید وجود داشت. نگاهی دوباره به روش اول ییندازیم:

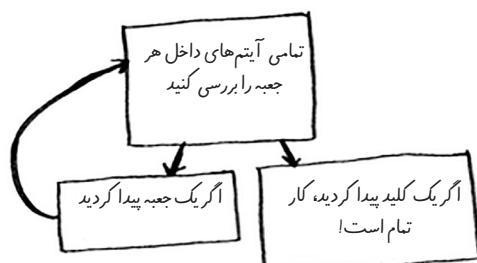


به این ترتیب، دسته‌ای از جعبه‌ها را برای جست‌وجو میان آن‌ها در کنار هم قرار می‌دهید، در نتیجه می‌دانید که چه جعبه‌هایی برای جست‌وجو باقی مانده است.

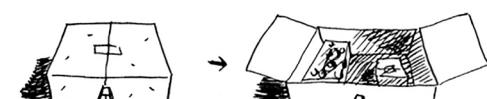
جعبه‌ی بعدی برای جست‌وجو



اما در رویکرد بازگشتی، خبری از دسته‌ای از جعبه‌ها نیست.



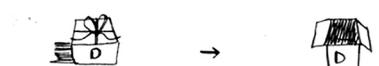
اگر دسته‌ای از جعبه‌ها وجود نداشته باشد، الگوریتم شما از کجا می‌تواند بفهمد که هنوز باید در میان چه جعبه‌هایی جست‌وجو کند؟ مثال:



داخل آن جعبه‌های B و C می‌بینید

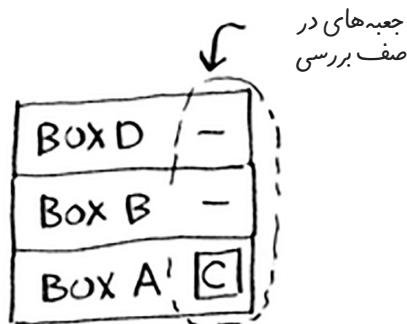


داخل آن جعبه‌ی D است



خالی است

در این مرحله، پشته‌فراخوانی به این شکل است:



«جعبه‌ها» در پشته ذخیره می‌شوند! این پشته‌ای از فراخوانی‌های تابع نیمه تکمیل شده است که هر کدام لیست نیمه کامل خود را از جعبه‌هایی دارد که باید آن‌ها را بگرد. استفاده از پشته راحت است زیرا نیازی نیست حساب مجموعه‌ی جعبه‌ها را داشته باشید - پشته این کار را برای شما انجام می‌دهد.

استفاده از پشته راحت است، اما به یک قیمت: ذخیره همه‌ی آن اطلاعات می‌تواند حافظه‌ی زیادی را اشغال کند. هر یک از این فراخوانی‌های تابع مقداری از حافظه را اشغال می‌کند، و وقتی پشته‌ی شما خیلی بزرگ است، به این معنی است که کامپیوتر شما اطلاعات را برای بسیاری از تماس‌های عملکرد ذخیره می‌کند. در آن مرحله، شما دو گزینه دارید:

- می‌توانید کد خود را بازنویسی کرده و به جای آن از یک حلقه استفاده کنید.
- می‌توانید از چیزی به نام *tail recursion* استفاده کنید. این یک مبحث پیشرفته بازگشتی است که خارج از محدوده‌ی این کتاب است. از طرفی تنها برخی از زبان‌ها از آن پشتیبانی می‌کنند.

تمرین

۳.۲ فرض کنید به طور تصادفی یک تابع بازگشتی نوشته شد که تا ابد اجرا می‌شود. همان‌طور که دیدید، کامپیوتر شما برای هر فراخوانی تابع، حافظه را در پشته تخصیص

می‌دهد. وقتی تابع بازگشتی شما تا ابد اجرا بشود چه اتفاقی برای پشته می‌افتد؟

جمع‌بندی



- بازگشت به زمانی گفته می‌شود که یک تابع، خود را فراخوانی می‌کند.
- هر تابع بازگشتی دو صورت دارد: صورت پایه و صورت بازگشتی.
- یک پشته دو عملیات دارد: push و pop.
- همه‌ی فراخوانی‌های تابع به پشته فراخوانی می‌روند.
- فراخوانی تماس می‌تواند بسیار بزرگ شود و منجر به اشغال حافظه‌ی زیادی بشود.

۴ | مرتب‌سازی سریع



در این فصل

• تقسیم و حل را یاد می‌گیرید. گاهی اوقات با مسئله‌ای روبه‌رو می‌شوید که با هیچ‌یک از الگوریتم‌هایی که آموخته‌اید قابل حل نیست. وقتی یک الگوریتمیست خبره با چنین مشکلی روبرو می‌شود، تسلیم نمی‌شود. این افراد جعبه ابزاری دارند که مملو از تکنیک‌های گوناگون برای حل مسئله است و سعی می‌کنند راه حلی برای مسئله بیابند. تقسیم و حل^۱ اولین تکنیک عمومی است که یاد می‌گیرید.

• در مورد مرتب‌سازی سریع، یک الگوریتم مرتب‌سازی خوب که اغلب اوقات از آن استفاده می‌شود، یاد می‌گیرید. مرتب‌سازی سریع از تقسیم و حل استفاده می‌کند.

در فصل پیش همه‌چیز را درباره‌ی بازگشت آموختید. این فصل بر استفاده از مهارت تازه‌ی شما برای حل مسئله‌ها تمرکز دارد. تقسیم و حل (D&C)، یک تکنیک بازگشتی مشهور برای حل مسئله را بررسی خواهیم کرد.

در این فصل به مهم‌ترین بخش الگوریتم‌ها می‌پردازیم. با این همه، یک الگوریتم اگر تنها برای حل یک نوع مسئله به کار بیاید، چندان به درد نمی‌خورد.

1. Divide and conquer

در عوض، D&C راه جدیدی برای فکر کردن در مورد حل مسئله‌ها در اختیار می‌گذارد. D&C ابزار دیگری در جعبه ابزار شماست. وقتی با یک مشکل جدید مواجه می‌شوید، لازم نیست سرگردان شوید. در عوض، می‌توانید بپرسید: «آیا می‌توانیم در صورت استفاده از تقسیم و حل این مسئله را حل کنم؟»

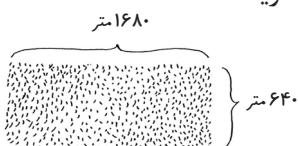
در پایان فصل، اولین الگوریتم مهم D&C را می‌آموزید: مرتب‌سازی سریع. مرتب‌سازی سریع، یک الگوریتم مرتب‌سازی است و بسیار سریع‌تر از مرتب‌سازی انتخابی (که در فصل ۲ یاد گرفتید) است. مرتب‌سازی سریع مثال خوبی از کدنویسی درجه‌ی یک است.

تقسیم و حل



درک D&C یا تقسیم و حل ممکن است کمی طول بکشد. بنابراین، ما سه مثال خواهیم داشت. ابتدا مثالی تصویری را به شما نشان می‌دهم. بعد نمونه کدی را می‌نویسم که اگرچه چندان جالب نیست اما احتمالاً آسان باشد. در نهایت، مرتب‌سازی سریع را بررسی می‌کنیم، یک الگوریتم مرتب‌سازی که از D&C استفاده می‌کند.

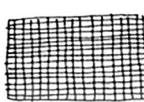
فرض کنید شما یک کشاورز هستید و یک قطعه زمین دارید.



قصد دارید این مزرعه را به طور مساوی به قطعات مربعی شکل تقسیم کنید. همچنین می‌خواهید نقشه‌ها تا حد امکان بزرگ باشند. بنابراین هیچ کدام از موارد زیر به کار نمی‌آید.



زمین به قطعات مربعی شکل تقسیم نشده است



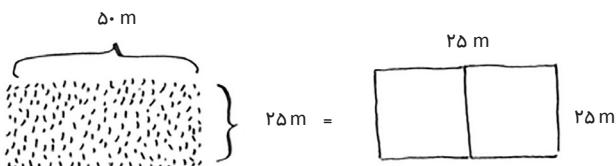
قطعه‌ها زیادی کوچک هستند



نام قطعه‌ها باید هم اندازه باشند

چگونه می‌توانید اندازه‌ی بزرگ‌ترین مربعی را که می‌توان برای یک قطعه زمین استفاده کرد، مشخص کنید؟ از استراتژی D&C استفاده کنید! الگوریتم‌های D&C های بازگشتی هستند. برای حل یک مشکل با استفاده از D&C، دو مرحله وجود دارد:

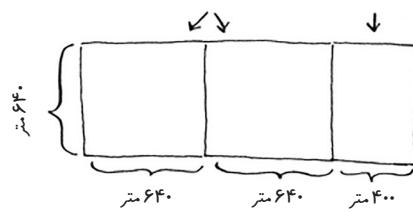
۱. صورت پایه را مشخص کنید. این باید ساده‌ترین حالت ممکن باشد.
 ۲. مسئله‌ی خود را تقسیم یا کوچک کنید تا زمانی که به صورت پایه تبدیل شود. بیایید از D&C برای یافتن راه حل این مسئله استفاده کنیم. اندازه‌ی بزرگ‌ترین مربعی که می‌توانید استفاده کنید چقدر است؟
- ابتدا صورت پایه را در نظر بگیرید. ساده‌ترین حالت این است که یک طرف ضریبی از طرف دیگر باشد.



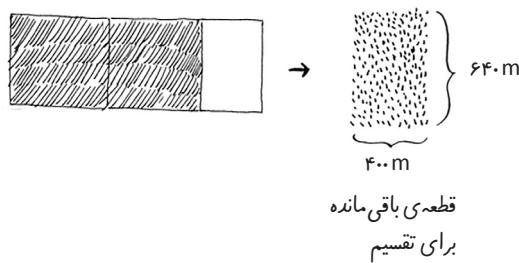
فرض کنید یک طرف ۲۵ متر و طرف دیگر ۵۰ متر باشد. پس بزرگ‌ترین مربعی که می‌توان استفاده کرد ۲۵ متر \times ۲۵ متر است. برای تقسیم زمین به دو تا از آن مربع‌ها نیاز دارید.

اکنون باید صورت بازگشتی را پیدا کنید. اینجاست که D&C وارد می‌شود. بر اساس روش تقسیم و حل، با هر تماس بازگشتی، باید مسئله‌ی خود را کوچک کنید. چگونه مسئله را در اینجا کوچک می‌کنید؟ بیایید با علامت‌گذاری بزرگ‌ترین مربع‌های قابل استفاده شروع کنیم.

بعضی از زمین که تقسیم
دو قطعه
نشده است



می‌توان دو مربع 640×640 در آن جا داد، و هنوز مقداری زمین برای تقسیم باقی بماند. دیگر همه چیز مشخص است. یک بخش زمین برای تقسیم باقی مانده است. چرا برای این بخش هم الگوریتم مشابهی را اعمال نمی‌کنید؟

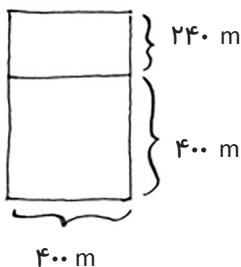


پس مسئله را با یک مزرعه‌ای با قطع 1680×640 شروع کردید که باید تقسیم می‌شد. اما اکنون باید یک بخش کوچک‌تر، 400×640 را تقسیم کنید. اگر بزرگ‌ترین مربعی را پیدا کنید که برای این اندازه مناسب است، آن بزرگ‌ترین مربعی است که برای کل مزرعه کارساز است. شما فقط ابعاد مسئله را از مزرعه‌ای به اندازه‌ی 640×1680 به مزرعه‌ای به اندازه‌ی 400×400 کاهش داده‌اید!

۱. الگوریتم اقلیدسی

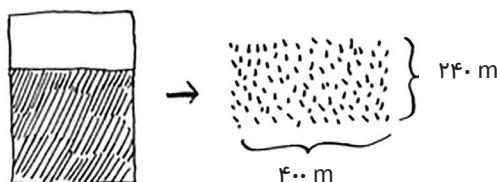
«اگر بزرگ‌ترین قطعه‌ای را که برای این اندازه مناسب است بیابید، به بزرگ‌ترین جعبه‌ای خواهد رسید که برای کل مزرعه کارساز باشد.» اگر دلیل درستی این جمله را متوجه نمی‌شوید، نگران نباشید. چون متأسفانه واضح نیست. اثبات این‌که چرا این‌گونه است مفصل‌تر از آن است که در این کتاب بیابید، بنابراین فقط باید به من باور داشته باشید که از این قاعده پیروی می‌کند. اگر می‌خواهید اثبات آن را بفهمید، الگوریتم اقلیدسی را جست‌وجو کنید. خان‌آکادمی در این آدرس توضیح خوبی ارائه کرده است:

<https://www.khanacademy.org/computing/computer-science/cryptography/modarithmetic/a/the-euclidean-algorithm>.

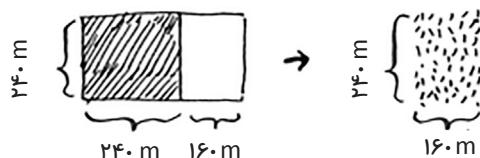


بیایید دوباره همان الگوریتم را اعمال کنیم. با شروع یک مزرعه 400×400 متر، بزرگ‌ترین مربعی که می‌توانید ایجاد کنید 400×400 متر است.

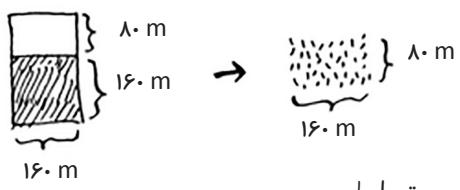
در این مرحله یک قطعه‌ی کوچک‌تر 240×400 متری در اختیار شما قرار می‌گیرد.



و در آن می‌توانید یک مربع دیگر ترسیم کنید تا یک بخش کوچک‌تر به ابعاد 160×160 متر به‌دست آورید.

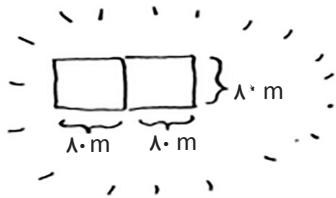


و سپس یک مربع بر روی آن بکشید تا یک بخش کوچک‌تر به دست آورید.

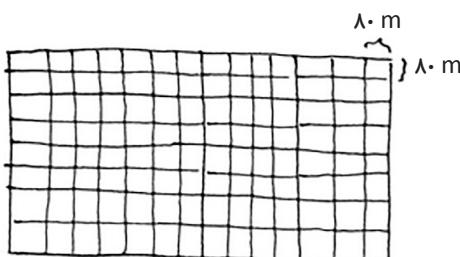


صورت پایه!

خب، شما در صورت پایه هستید: 80×80 است. اگر این بخش را با استفاده از مربع‌ها تقسیم کنید، دیگر چیزی باقی نمی‌ماند!



بنابراین، برای قطعه زمین اصلی، بزرگ‌ترین قطعه‌ای که می‌توانید استفاده کنید 80×80 متر است.



برای جمع بندی، در اینجا نحوه کار D&C آمده است:

۱. یک صورت ساده را به عنوان صورت پایه در نظر بگیرید.
 ۲. چگونه می‌توان مسئله‌ی خود را کوچک کرد و به صورت پایه رسید.
- D & C یک الگوریتم ساده نیست که بتوانید آن را برای یک مسئله اعمال کنید. در عوض، راهی برای فکر کردن به راه حل یک مسئله است. بیایید یک مثال دیگر انجام دهیم.



آرایه‌ای از اعداد به شما داده می‌شود.

باید همه‌ی اعداد را جمع زده و مجموع را برگردانید. انجام این کار به کمک یک حلقه خیلی ساده است:

```
def sum(arr):
    total = 0
    for x in arr:
        total += x
    return total
print sum([1, 2, 3, 4])
```

اما چگونه می توان این کار را با یک تابع بازگشتی انجام داد؟

مرحله‌ی ۱: صورت پایه را مشخص کنید. ساده‌ترین آرایه‌ای که می‌توان داشت چیست؟ ساده‌ترین صورت را در نظر گرفته و بعد ادامه یدهید. اگر آرایه‌ای با صفر یا یک عنصر داشته باشد، جمع آن بسیار آسان خواهد بود.

$$\left\{ \begin{array}{l} \text{صورت پایه} \\ \left[\quad \right] \cdot \text{ELEMENTS} = \text{مجموع برابر است با صفر} \\ \boxed{1} \quad 1 \text{ ELEMENT} = 1 \text{ مجموع برابر است با ۱} \end{array} \right.$$

در نتیجه این صورت پایه خواهد بود.

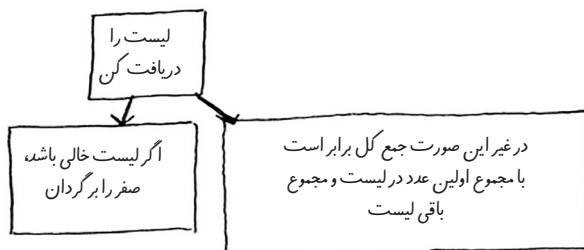
مرحله‌ی ۲: شما باید با هر تماس بازگشت به یک آرایه‌ی خالی نزدیک‌تر بشوید. چگونه ابعاد مسئله‌ی خود را کوچک می‌کنید؟ یک راه آن به شکل زیر است:

$$\text{sum}\left(\begin{array}{|c|c|c|} \hline 2 & 4 & 6 \\ \hline \end{array}\right) = 12$$

مثل این می‌ماند که بنویسیم:

$$2 + \text{sum}\left(\begin{array}{|c|c|} \hline 4 & 6 \\ \hline \end{array}\right) = 2 + 10 = 12$$

در هر صورت، نتیجه ۱۲ می شود. اما در نسخه‌ی دوم، آرایه‌ی کوچکتری را به تابع sum ارسال می‌کنید. یعنی ابعاد مسئله را کوچک کرده‌اید! تابع sum شما می‌تواند به این صورت عمل کند.



در پایین مراحل آن به تصویر درآمده است.

نتیجہی نہایی

به پاد داشته باشید، تابع بازگشتی، موقعیت را دنبال می‌کند.

به یاد داشته باشید که بازگشت جایگاه این فراخوانی یا تناسی های تایم را در خود نگه می دارد.

BASE CASE!
 $\text{sum}(\boxed{6}) \text{ is } 6.$

نکته

هنگامی که یک تابع بازگشتی می‌نویسید که شامل یک آرایه است، صورت پایه اغلب یک آرایه‌ی خالی یا آرایه‌ای با یک عنصر است. اگر نمی‌دانید چاره چیست، اول از همه آن را امتحان کنید.

نگاهی به برنامه نویسی فانکشنال

شاید با خودتان فکر کنید «وقتی می‌توان مسئله را به راحتی با یک حلقه انجام داد، چرا باید این کار را به صورت بازگشتی انجام بدhem؟» خب، این نگاهی گذرا به برنامه نویسی فانکشنال است! زبان‌های برنامه نویسی فانکشنال مانند هaskell این حلقه ندارند، بنابراین برای نوشتن توابعی مانند این باید از بازگشت استفاده کنید. اگر درک خوبی از بازگشت داشته باشید، یادگیری زبان‌های فانکشنال آسان‌تر خواهد بود. برای مثال، در اینجا نحوه‌ی نوشتن یک تابع `sum` در هaskell آمده است:

```
sum [] = 0           ← صورت پایه
sum (x:xs) = x + (sum xs) ← صورت بازگشتی
```

توجه داشته باشید که به نظر می‌رسد دو تعریف برای تابع دارید. اولین تعریف در صورت پایه و تعریف دوم در صورت بازگشتی اجرا می‌شود. همچنین می‌توانید این تابع را در هaskell با استفاده از دستور `if` بنویسید:

```
sum arr = if arr == []
            then 0
            else (head arr) + (sum (tail arr))
```

اما خواندن تعریف اول ساده‌تر است. به دلیل استفاده‌ی زیاد هaskell از توابع بازگشتی، ظرافت‌های فراوانی مانند این را در خود دارد تا آن را آسان‌تر کند. اگر به تابع بازگشتی علاقه دارید، یا علاقه‌مند به یادگیری یک زبان جدید هستید، نگاهی به زبان هaskell بیندازید.

تمرین



۴.۱ کد را برای تابع sum قبلی بنویسید.

۴.۲ یک تابع بازگشتی برای شمارش تعداد آیتم‌های یک لیست بنویسید.

۴.۳ حداکثر عدد را در یک لیست بیابید.

۴.۴ جست‌وجوی دودویی را از فصل ۱ به خاطر دارید؟ این یک الگوریتم تقسیم و حل نیز هست. آیا می‌توانید برای جست‌وجوی دودویی، صورت پایه و صورت بازگشتی را پیدا کنید؟

مرتب‌سازی سریع



مرتب‌سازی سریع یک الگوریتم مرتب‌سازی است. این الگوریتم بسیار سریع‌تر از مرتب‌سازی انتخابی است و اغلب در پروژه‌های روزمره از آن استفاده می‌شود. به عنوان مثال، کتابخانه‌ی استاندارد C تابعی به نام `qsort` دارد که در واقع اجرای الگوریتم مرتب‌سازی سریع در این زبان برنامه‌نویسی است. مرتب‌سازی سریع هم از D&C استفاده می‌کند.

بیایید از مرتب‌سازی سریع برای مرتب کردن یک آرایه استفاده کنیم. ساده‌ترین آرایه‌ای که یک الگوریتم مرتب‌سازی می‌تواند از عهده‌ی آن برباید چیست (توصیه‌ای را که در بخش قبل کردم به خاطر دارید)؟ خب، برخی از آرایه‌ها اصلاً نیازی به مرتب‌سازی ندارند.

$$\left\{ \begin{array}{l} [] \leftarrow \text{آرایه‌ی خالی} \\ \boxed{20} \leftarrow \text{آرایه‌ای با یک عنصر} \end{array} \right.$$

آرایه‌ها و آرایه‌های خالی تنها با یک عنصر، صورت پایه خواهند بود. شما فقط می‌توانید آن آرایه‌ها را به همان شکل که هستند برگردانید - چیزی برای مرتب‌سازی وجود ندارد:

```
def quicksort(array):
    if len(array) < 2:
        return array
```

بیایید به آرایه‌های بزرگ‌تر نگاه بیندازیم. مرتب‌سازی آرایه‌ای با دو عنصر بسیار آسان است.



اگر اولین عنصر کوچک‌تر از دومین عنصر نیست، جای آن هارا با هم عرض کنید.

آرایه‌ای با سه عنصر چطور؟



به یاد داشته باشید، شما از D&C استفاده می‌کنید. بنابراین می‌خواهید این آرایه را تا زمانی که در صورت پایه قرار بگیرید، تجزیه کنید. در اینجا نحوه‌ی کار الگوریتم مرتب‌سازی سریع آمده است. ابتدا یک عنصر از آرایه را انتخاب کنید. این عنصر محور نامیده می‌شود.



بعداً در مورد چگونگی انتخاب یک محور مناسب صحبت خواهیم کرد. در حال حاضر، فرض کنید اولین آیتم در آرایه محور است. حالا عناصر کوچک‌تر از محور و عناصر بزرگ‌تر از محور را پیدا کنید.



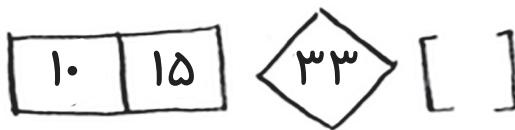
1. pivot

به این کار پارتیشن‌بندی^۱ می‌گویند. حالا شما موارد زیر را در اختیار دارید:

- زیرآرایه‌ای^۲ از همه‌ی اعداد کوچک‌تر از محور
- محور

- یک زیرآرایه از تمام اعداد بزرگ‌تر از محور

این دو زیرآرایه مرتب نشده‌اند. آن‌ها فقط پارتیشن‌بندی شده‌اند اما اگر مرتب شده باشند، مرتب سازی کل آرایه بسیار آسان خواهد بود.



اگر زیرآرایه‌ها مرتب شده‌اند، می‌توانید کل آن را به این شکل ترکیب کنید:
آرایه‌ی سمت راست + محور + آرایه‌ی سمت چپ - و یک آرایه مرتب شده به دست آورید. در این نمونه آرایه‌ی مرتب شده‌ی زیر را خواهیم داشت:

$$[10, 15] + [33] = [10, 15, 33]$$

چگونه زیر آرایه‌هارا مرتب می‌کنید؟ مرتب‌سازی سریع صورت پایه از قبل می‌داند که چگونه آرایه‌های دو عنصر(زیرآرایه‌ی سمت چپ) و آرایه‌های خالی (زیرآرایه سمت راست) را مرتب کند . بنابراین اگر مرتب‌سازی سریع را روی دو زیرآرایه فراخوانی کنید و سپس نتایج را ترکیب کنید، یک آرایه‌ی مرتب شده به دست می‌آورید!

```
quicksort([15, 10]) + [33] + quicksort([])  
> [10, 15, 33]
```

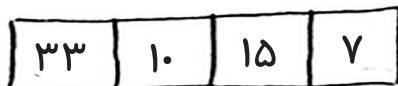
با انتخاب هر عنصر دیگری به عنوان محور، نتیجه همچنان حاصل می‌شود. فرض کنید این بار ۱۵ را به عنوان محور انتخاب کنید.



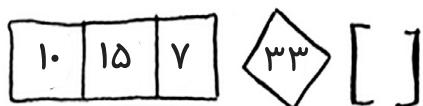
هر دو زیرآرایه فقط یک عنصر دارند و شما می‌دانید که چگونه باید ان‌ها را مرتب کنید. بنابراین اکنون می‌دانید که چگونه یک آرایه از سه عنصر را مرتب کنید. مراحل به این شکل است:

1. partitioning
2. Sub-array

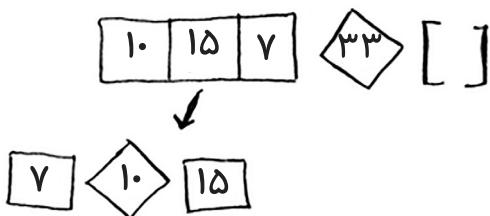
۱. یک محور را انتخاب کنید.
۲. آرایه را به دو آرایه‌ی فرعی تقسیم کنید: عناصر کوچک‌تر از محور و عناصر بزرگ‌تر از محور.
۳. مرتب‌سازی سریع را به صورت بازگشته در دو آرایه‌ی فرعی فراخوانی کنید.
آرایه‌ای از چهار عنصر چطور؟



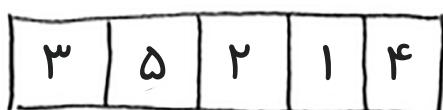
فرض کنید دوباره ۳۳ را به عنوان محور انتخاب کنید.



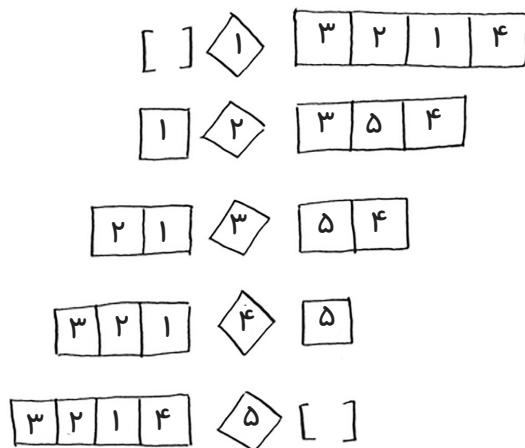
آرایه‌ی سمت چپ دارای سه عنصر است. از پیش می‌دانید که چگونه یک آرایه از سه عنصر را مرتب کنید: مرتب‌سازی سریع بر روی آن را به شکل بازگشته فراخوانی کنید.



بنابراین می‌توانید یک آرایه از چهار عنصر را مرتب کنید. و اگر می‌توانید یک آرایه از چهار عنصر را مرتب کنید، می‌توانید یک آرایه از پنج عنصر را مرتب کنید. چرا؟ فرض کنید این آرایه‌ی پنج عنصری را دارید.

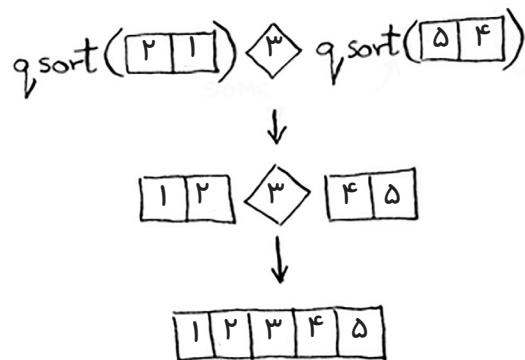


در اینجا همه‌ی روش‌هایی که می‌توانید این آرایه را تقسیم‌بندی کنید، بسته به محوری که انتخاب می‌کنید، آورده شده است.

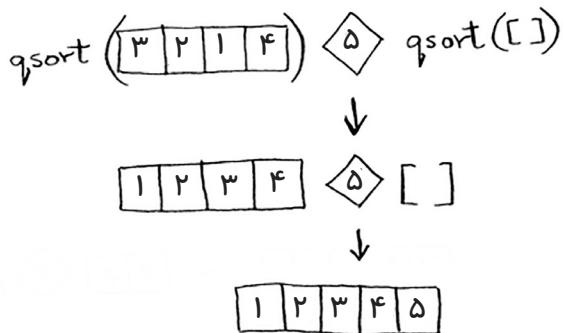


توجه داشته باشید که تمام این زیرآرایه‌ها بین ۰ تا ۴ عنصر دارند. و شما از پیش می‌دانید که چگونه یک آرایه از ۰ تا ۴ عنصر را با استفاده از الگوریتم مرتب‌سازی سریع مرتب کنید! بنابراین مهم نیست که چه محوری را انتخاب می‌کنید، می‌توانید مرتب‌سازی سریع را به صورت بازگشتی در دو زیرآرایه فراخوانی کنید.

برای مثال، فرض کنید عدد ۳ را به عنوان محور انتخاب کنید. شما مرتب‌سازی سریع را در زیرآرایه‌ها فرامی‌خواهید:



زیرآرایه‌ها مرتب می‌شوند و سپس تمام آن‌ها را با هم ترکیب می‌کنید تا یک آرایه‌ی مرتب شده به دست بیاورید. حتی اگر عدد ۵ را به عنوان محور انتخاب کنید، باز هم نتیجه حاصل می‌شود.



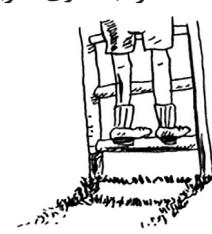
این روش با هر عنصری به عنوان محور جواب می‌دهد. بنابراین می‌توانید یک آرایه از پنج عنصر را مرتب کنید. با همین منطق، می‌توانید یک آرایه از شش عنصر و غیره را مرتب کنید.

برهان‌های استقرایی^۱

همین الان به برهان‌های استقرایی نیم‌نگاهی داشتید! اثبات‌های استقرایی یکی از راه‌های اثبات کارایی الگوریتم شما هستند. هر اثبات استقرایی دو مرحله دارد: صورت پایه و صورت استقرایی. آشنا به نظر می‌رسند؟ به عنوان مثال، فرض کنید می‌خواهم ثابت کنم که می‌توانم تا بالای یک نرdban بروم. در حالت استقرایی، اگر پاهایم روی یک پله باشد، می‌توانم پاهایم را روی پله‌ی بعدی بگذارم. بنابراین اگر در پله‌ی ۲ هستم، می‌توانم به پله‌ی ۳ بروم. این حالت استقرایی است. برای صورت پایه، می‌گوییم که پاهایم روی پله‌ی ۱ قرار گرفته‌اند. بنابراین می‌توانم با یک پله یک پله بالاگرفتن، به بالای نرdban برسم. شما از استدلال مشابهی برای مرتب‌سازی سریع استفاده می‌کنید. در صورت پایه، نشان دادم که الگوریتم برای صورت پایه جواب می‌دهد: آرایه‌هایی با اندازه‌ی ۰ و ۱. در صورت استقرایی، نشان دادم که اگر مرتب‌سازی سریع برای آرایه‌ای با اندازه‌ی ۱ مناسب باشد، برای آرایه‌ای با اندازه‌ی ۲ هم مناسب خواهد بود. و اگر برای آرایه‌ای با اندازه‌ی ۲ نتیجه‌بخش باشد، برای آرایه‌های با اندازه‌ی ۳ و اندازه‌های دیگر هم نتیجه‌بخش است. پس می‌توانم بگوییم که مرتب‌سازی سریع برای همه‌ی آرایه‌ها با هر اندازه‌ای مناسب است. من در اینجا بیش از این و با جزئیات بیشتر به اثبات‌های استقرایی نمی‌پردازم، اما آن‌ها جالب هستند و با D&C ارتباط تنگاتنگی دارند

کد مرتب‌سازی سریع در ادامه آمده است:

```
def quicksort(array):
    if len(array) < 2:
        return array
    else :
        pivot = array[0]
        less = [i for i in array[1:] if i <= pivot]
        more = [i for i in array[1:] if i > pivot]
        صورت پایه: آرایه‌های با صفر یا یک عنصر «مرتب» شده‌اند.
        صورت بازگشتی زیر آرایه‌ای از تا می‌عنصر کمتر از (pivot)
        return less + [pivot] + more
```



```

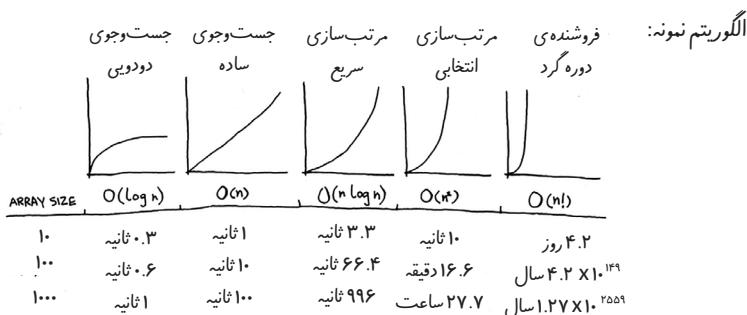
greater = [i for i in array[1:] if i > pivot]----- زیر آرایه‌ای از تابع عناصر
return quicksort(less) + [pivot] + quicksort(greater) بزرگ‌تر از محور

print quicksort([10, 5, 2, 3])

```

بازبینی نماد ۰ بزرگ

مرتب‌سازی سریع منحصر به فرد است زیرا سرعت آن به محوری که انتخاب می‌کنید بستگی دارد. قبل از این‌که در مورد مرتب‌سازی سریع صحبت کنم، اجازه دهید دوباره به رایج‌ترین زمان‌های اجرای ۰ بزرگ نگاه بیندازیم.



تخمین‌ها بر اساس کامپیوتری با سرعت کم است که ۱۰ عملیات در ثانیه انجام می‌دهد.

مثال‌های زمانی این نمودار تخمینی با این فرض هستند که ۱۰ عملیات در ثانیه انجام بدهید. این نمودارها دقیق نیستند و تنها برای این است که به شما درکی بدهد که این زمان‌های اجرا تا چه اندازه متفاوت از یکدیگر هستند. در حقیقت، کامپیوتر شما می‌تواند بیش از ۱۰ عملیات در ثانیه انجام دهد.

هر زمان اجرا یک الگوریتم نمونه دارد. مرتب‌سازی انتخابی را که در فصل ۲ یاد گرفتید، بررسی کنید. زمان اجرای آن $O(n^2)$ است. این الگوریتمی بسیار کند است. الگوریتم مرتب‌سازی دیگری به نام مرتب‌سازی ادغامی است که $O(n \log n)$ است. خیلی سریع‌تر! مرتب‌سازی سریع یک نمونه‌ی چالش برانگیز است. در بدترین حالت، مرتب‌سازی سریع زمان $O(n^2)$ خواهد داشت.

به اندازه‌ی مرتب‌سازی انتخابی کند است! البته برای بدترین حالت است. در حالت

متوسط، مرتب‌سازی سریع به زمان $O(n \log n)$ نیاز دارد. بنابراین ممکن است با خودتان فکر کنید:

- بدترین حالت^۱ و حالت متوسط^۲ در اینجا به چه معناست؟
- اگر مرتب‌سازی سریع به طور متوسط $O(n \log n)$ است، اما مرتب‌سازی ادغامی همیشه $O(n \log n)$ است، چرا از مرتب‌سازی ادغامی استفاده نکنید؟ مگر سریع‌تر نیست؟

مرتب‌سازی ادغامی در برابر مرتب‌سازی سریع

فرض کنید این تابع ساده را برای چاپ هر آیتم در یک لیست دارید:

```
def print_items(list):
    for item in lis:
        print item
```

این تابع تمام موارد موجود در لیست را مرور کرده و آن را چاپ می‌کند. از آنجا که یک بار روی کل لیست حلقه می‌زند، این تابع در زمان $O(n)$ اجرا می‌شود. حالا، فرض کنید این تابع را طوری تغییر دهید که قبل از چاپ یک آیتم، ۱ ثانیه توقف کند:

```
from time import sleep
def print_items2(list):
    for item in list:
        sleep(1)
        print item
```

قبل از اینکه یک آیتم را پرینت کند، ۱ ثانیه مکث می‌کند. فرض کنید لیستی از پنج آیتم را با استفاده از هر دو تابع پرینت می‌کنید.

۲	۴	۶	۸	۱۰
---	---	---	---	----

↓

```
print_items: 2 4 6 8 10
print_items2: <sleep> 2 <sleep> 4 <sleep> 6 <sleep> 8 <sleep> 10
```

1. Worst case
2. Average case

هر دو تابع یک بار در لیست حلقه می‌زنند، بنابراین هر دو زمان $O(n)$ هستند. کدامیک در عمل سریع‌تر خواهد بود؟ فکر می‌کنم `print_items` بسیار سریع‌تر خواهد بود، چون قبل از پرینت یک آیتم، مکث ۱ ثانیه‌ای ندارد. بنابراین با وجود این‌که هر دو تابع در نماد O بزرگ سرعت یکسانی دارند، `print_items` در عمل سریع‌تر است. وقتی نماد O بزرگ را مانند (n) می‌نویسید، در واقع به این معنی است:

$$\xrightarrow{\text{یک زمان ثابت}} \mathcal{O} * n$$

۲ مقدار ثابتی از زمانی است که الگوریتم شما طول می‌کشد. به آن ثابت^۱ می‌گویند. برای مثال، ممکن است $n * 10 \text{ milliseconds}$ برای `print_items` در مقابل $1 \text{ second} * n$ برای `print_items2` باشد.

معمولًاً آن ثابت را نادیده می‌گیرید، زیرا اگر دو الگوریتم زمان‌های O بزرگ متفاوتی داشته باشند، ثابت اهمیتی ندارد. برای مثال جست‌وجوی دودویی و جست‌وجوی ساده را در نظر بگیرید. فرض کنید هر دو الگوریتم این ثابت‌ها را داشته باشند.

$$\frac{10 \text{ ms} * n}{\text{جست‌وجوی ساده}} \quad \frac{1 \text{ sec} * \log n}{\text{جست‌وجوی دودویی}}$$

ممکن است بگویید: «چه جالب! ثابت جست‌وجوی ساده ۱۰ میلی‌ثانیه است، اما جست‌وجوی دودویی ثابت ۱ ثانیه دارد. پس جست‌وجوی ساده خیلی سریع‌تر است!» حال فرض کنید در حال جست‌وجو در میان لیستی از ۴ میلیارد عنصر هستید. زمان آن به این شکل است:

جست‌وجوی ساده	$\begin{array}{rclcl} & & & & \\ & & & & \\ & & & & \end{array}$	جست‌وجوی دودویی	$\begin{array}{rclcl} & & & & \\ & & & & \\ & & & & \end{array}$
۱. constant	۱۰ میلی‌ثانیه	۴ میلیارد	۳۲ ثانیه
جست‌وجوی ساده	*	*	*

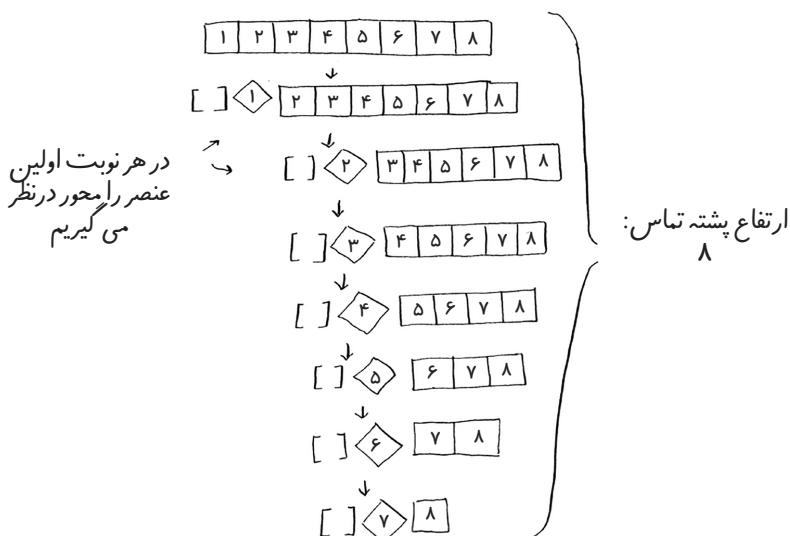
همان‌طور که می‌بینید، جست‌وجوی دودویی همچنان بسیار سریع‌تر است. آن ثابت اصلاً تفاوتی ایجاد نکرد.

اما گاهی اوقات ثابت می‌تواند تفاوت ایجاد کند. مرتب‌سازی سریع در مقابل مرتب‌سازی ادغامی تنها یک مثال است. مرتب‌سازی سریع ثابت کوچک‌تری نسبت به مرتب‌سازی ادغامی دارد. بنابراین اگر هر دو در زمان $O(n \log n)$ باشند، مرتب‌سازی سریع، پرسرعت‌تر است. مرتب‌سازی سریع در عمل هم سریع‌تر است، زیرا اغلب اوقات در حالت متوسط و نه بدترین حالت خواهد بود.

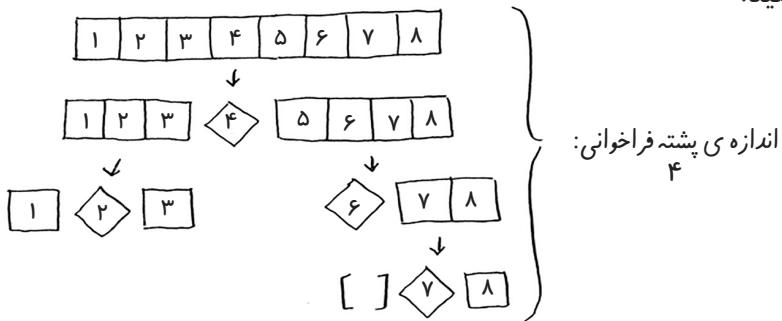
با خودتان فکر می‌کنید: حالت متوسط در مقابل بدترین حالت چیست؟

حالت متوسط در مقابل بدترین حالت

عملکرد مرتب‌سازی سریع به مقدار زیادی به محوری که انتخاب می‌کنید بستگی دارد. فرض کنید همیشه اولین عنصر را به عنوان محور انتخاب می‌کنید. و شما مرتب‌سازی سریع را با آرایه‌ای که از پیش مرتب شده است فراخوانی می‌کنید. مرتب‌سازی سریع بررسی نمی‌کند که آیا آرایه‌ی ورودی از قبل مرتب شده است یا خیر. بنابراین همچنان سعی خواهد کرد آن را مرتب کند.

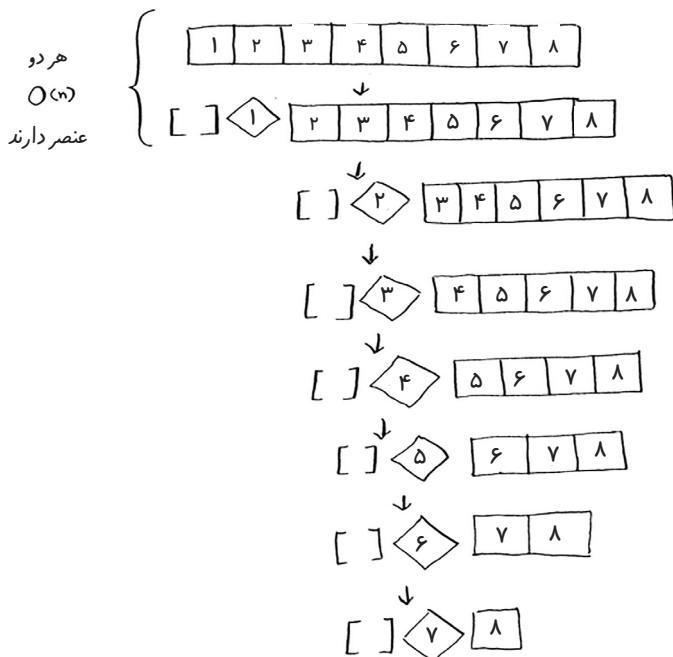


توجه کنید که آرایه به دو نیمه تقسیم نمی‌شود. در عوض، یکی از زیرآرایه‌ها همیشه خالی است. بنابراین پشته‌فراخوانی بسیار طولانی است. حالا در عوض، فرض کنید همیشه عنصر میانی را به عنوان محور انتخاب می‌کنید. اکنون به پشته‌فراخوانی نگاه کنید.

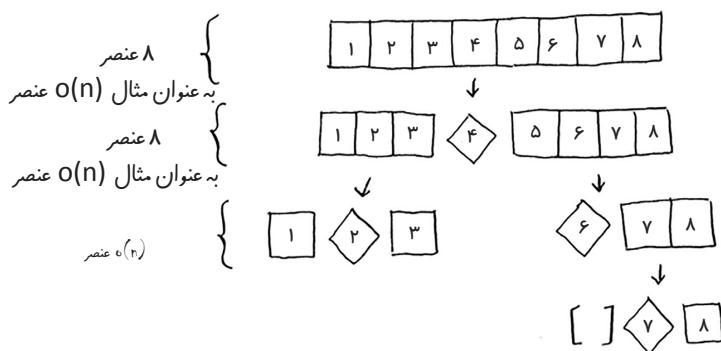


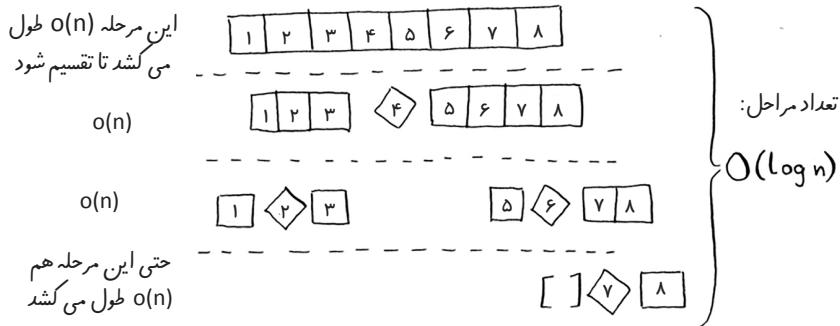
خیلی کوتاه شد! از آنجایی که هر بار آرایه را به نصف تقسیم می‌کنید، نیازی به برقراری تماس‌های بازگشتی زیادی ندارید. شما زودتر به حالت پایه می‌رسید و پشته‌فراخوانی بسیار کوتاه‌تر است.

مثال اول بدترین سناریو و مثال دوم بهترین سناریو است. در بدترین حالت، اندازه‌ی پشته $O(n)$ است. در بهترین حالت، اندازه‌ی پشته $O(\log n)$ است. اکنون به سطح اول پشتی نگاه کنید. شما یک عنصر را به عنوان محور انتخاب می‌کنید و باقی عناصر به زیرآرایه تقسیم می‌شوند. شما تمام هشت عنصر آرایه را بررسی می‌کنید. در نتیجه اولین عملیات به زمان $O(n)$ نیاز دارد. تمام هشت عنصر را در این سطح از پشته‌فراخوانی مرور کردید. اما در واقع، عناصر (n) را در هر سطح از پشته‌فراخوانی بررسی می‌کنید.



حتی اگر آرایه را متفاوت پارتیشن بندی کنید، باز هم به $O(n)$ عنصر در هر نوبت رسید.





در این مثال، سطوح $O(\log n)$ وجود دارد (به عبارت فنی «ارتفاع پشته‌فراخوانی است»). و هر سطح به زمان $O(n)$ نیاز دارد. کل الگوریتم زمان $O(n) * O(\log n) = O(n \log n)$ نیاز دارد. این سناریو برای بهترین حالت است. در بدترین حالت، $O(n) * O(n) = O(n^2)$ مرحله وجود دارد، بنابراین الگوریتم زمان $O(n \log n)$ خواهد داشت.

یک خبر خوب! بهترین حالت همان حالت متوسط است. اگر همیشه یک عنصر تصادفی در آرایه را به عنوان محور انتخاب کنید، مرتب‌سازی سریع به طور متوسط در زمان $O(n \log n)$ تکمیل می‌شود. مرتب‌سازی سریع یکی از سریع‌ترین الگوریتم‌های مرتب‌سازی موجود است و نمونه‌ی بسیار خوبی از D&C است.

تمرین

- هر یک از این عملیات بر اساس نماد O بزرگ چقدر طول می‌کشد؟
- ۴.۵ چاپ مقدار هر عنصر در یک آرایه.
- ۴.۶ دو برابر کردن مقدار هر عنصر در یک آرایه.
- ۴.۷ دو برابر کردن مقدار فقط اولین عنصر در یک آرایه.
- ۴.۸ ایجاد جدول ضرب با تمام عناصر موجود در آرایه. بنابراین اگر آرایه‌ی شما $[2, 3, 7, 8, 10]$ باشد، ابتدا هر عنصر را در 2 ضرب کنید، سپس هر عنصر را در 3 ، بعد در 7 و غیره ضرب کنید.

جمع‌بندی

D&C با تجزیه‌ی یک مسئله به قطعات کوچک‌تر و کوچک‌تر کار می‌کند. اگر از D&C در یک لیست استفاده می‌کنید، احتمالاً صورت پایه یک آرایه‌ی خالی یا یک آرایه با یک عنصر است.

- اگر مرتب‌سازی سریع را اجرا می‌کنید، یک عنصر تصادفی را به عنوان محور انتخاب کنید. میانگین زمان اجرای مرتب‌سازی سریع $O(n \log n)$ است!
- ثابت در نماد O بزرگ می‌تواند گاهی اوقات مهم باشد. به همین دلیل است که مرتب‌سازی سریع، سریع‌تر از مرتب‌سازی ادغام است.
- ثابت‌ها تقریباً هرگز در مقایسه‌ی میان جست‌وجوی ساده و جست‌وجوی دودویی مهم نیست، چون زمانه، که لیست شما بزرگ می‌شود $O(\log n)$ بسیار سریع‌تر از $O(n)$ است.



۱۵ جدول‌های هش



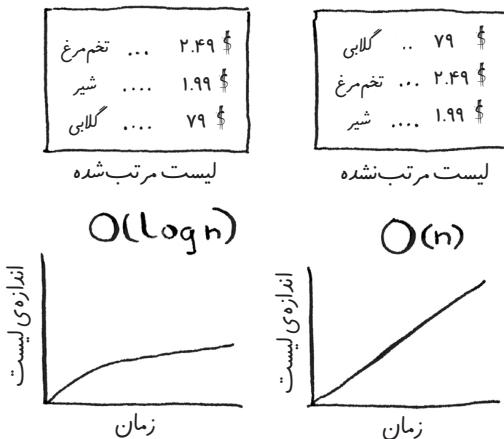
در این فصل

- در مورد جدول‌های هش، یکی از مفیدترین ساختمان‌های داده، یاد می‌گیرید. جدول‌های هش کاربردهای زیادی دارند. در این فصل به موارد رایج استفاده‌ی آن پرداخته می‌شود.
- در مورد اجزای داخلی جدول‌های هش: پیاده‌سازی^۱، تصادم‌ها^۲ و توابع هش یاد می‌گیرید. این به شما کمک می‌کند تا نحوه‌ی تجزیه و تحلیل عملکرد جدول هش را درک کنید.



فرض کنید در یک فروشگاه مواد غذایی کار می‌کنید. وقتی مشتری محصولی را خریداری می‌کند، باید قیمت‌ش را در یک کتابچه جست‌وجو کنید. اگر این کتابچه بر اساس حروف الفبا مرتب نشده باشد، ممکن است زمان زیادی طول بکشد تا تک‌تک سطرهای را جست‌وجو کنید تا سیب را پیدا کنید. شما می‌توانید مطابق فصل ۱ در جایی که باید هر خط رانگاه کنید، از جست‌وجوی ساده استفاده کنید، به خاطر دارید چه مقدار طول می‌کشید؟ زمان $O(n)$. اگر کتاب بر اساس حروف الفبا باشد، می‌توانید از جست‌وجوی دودویی برای یافتن قیمت یک سیب استفاده کنید. این فقط به زمان $O(\log n)$ نیاز دارد.

1. implementation
2. collisions



محض یادآوری، تفاوت زیادی میان زمان $O(n)$ و $O(n \log n)$ وجود دارد! فرض کنید می‌توان ۱۰ خط از کتاب را در هر ثانیه نگاه کرد. در ادامه زمان لازم برای جستجوی ساده و جستجوی دودویی آمده است:

آیتم‌های داخل کتاب	$O(n)$	$O(n \log n)$
۱۰۰	۱۰ ثانیه	۷ ثانیه بررسی کردن: $\log_2 100 = 7$
۱۰۰۰	۱.۶۶ دقیقه	۱۰ دقیقه $\log_2 1000 = 10$
۱۰۰۰۰	۱۶ دقیقه	۱۴ دقیقه $\log_2 10000 = 14$ = ۲ ثانیه

از پیش می‌دانید که جستجوی دودویی بسیار سریع است. اما به عنوان یک صندوق‌دار، جستجو در تمام یک کتابچه مكافات است، حتی اگر محتوای کتابچه مرتب شده باشد. می‌توانید حال و روز بی اعصاب مشتری را هنگام جستجوی آیتم‌ها در کتاب تصور کنید. چیزی که واقعاً به آن نیاز دارید، رفیقی است که تمامی اسم‌ها و قیمت‌ها را حفظ کرده باشد. در نتیجه نیاز نخواهد بود چیزی را جستجو کنید: از او می‌پرسید، و او بلا فاصله پاسخ را به شما می‌گوید.



رفیق شما مَگی می‌تواند قیمت هر آیتمی را در زمان (1) به شما بگوید، مهم نیست کتابچه چقدر حجمی باشد. او حتی از جست‌وجوی دودویی هم سریع‌تر است.

#	آیتم‌های داخل کتاب	ساده	جست‌وجوی دودویی	مگی
100	(n)	10 ثانیه	0.109n	(1)
1000		1.6 دقیقه	1 ثانیه	دروجا
10000		16.6 دقیقه	2 ثانیه	دروجا

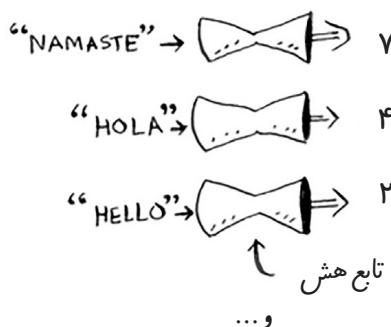
چه انسان نازینی! چگونه یک «مگی» داشته باشیم؟ بباید عینک ساختمان داده‌ای خود را به صورت بزنیم. شما تاکنون دو ساختمان داده می‌شناسید: آرایه و لیست (در مورد پشته‌ها صحبت نمی‌کنم) چون نمی‌توانید چیزی را در یک پشته «جست‌جو» کنید. این کتابچه را می‌توانید به صورت یک آرایه پیاده‌سازی کنید.

(۰.۷۹، گلابی)
(۱.۴۹، شیر)
(۲.۴۹، تخم مرغ)

هر آیتم در آرایه در واقع دو آیتم است: یکی نام یک محصول و دیگری قیمت آن است. اگر این آرایه را بر اساس نام مرتب کنید، می‌توانید جستجوی دودویی را روی آن اجرا کنید تا قیمت یک آیتم را بیابید. بنابراین می‌توانید آیتم‌ها را در زمان $O(\log n)$ بیابید. اما می‌خواهید آیتم‌ها را در زمان $O(1)$ پیدا کنید. یعنی می‌خواهید یک «مگی» بسازید. اینجاست که توابع هش از راه می‌رسند.

تابع هش

تابع هش تابعی است که در آن یک رشته قرار می‌دهید و آن یک عدد را به شما برمی‌گرداند.

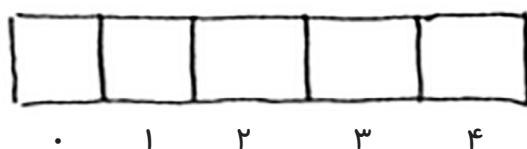


در اصطلاح فنی، ما می‌گوییم که یک تابع هش «رشته‌ها را به اعداد نگاشت» می‌کند. ممکن است فکر کنید هیچ الگوی قابل فهمی برای عددی که با قرار دادن یک رشته به دست می‌آورید وجود ندارد. اما برخی از الزامات برای یک تابع هش وجود دارد:

- باید اصولی و یکپارچه باشد. به عنوان مثال، فرض کنید «سیب» را وارد کرده و «۴» را برگرداند. هر بار که «سیب» را وارد می‌کنید، باید «۴» را دریافت کنید. غیر از این باشد، جدول هش شما کار نخواهد کرد.

- باید کلمات مختلف را به اعداد مختلف نگاشت کند. به عنوان مثال، یک تابع هش اگر همیشه برای هر کلمه‌ای که وارد می‌کنید «۱» را برگرداند فایده ندارد. در بهترین حالت، هر کلمه‌ی متفاوت باید به عدد متفاوتی نگاشت شود.

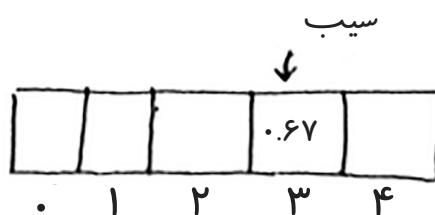
بنابراین یک تابع هش رشته‌ها را به اعداد نگاشت می‌کند. به چه کار می‌آید؟ خب، شما می‌توانید از آن برای خلق «مگی» خود استفاده کنید! با یک آرایه‌ی خالی شروع کنید:



شما تمام قیمت‌های مورد نظر خود را در این آرایه ذخیره می‌کنید. بیایید قیمت یک سیب را اضافه کنیم. «سیب» را به تابع هش وارد کنید.



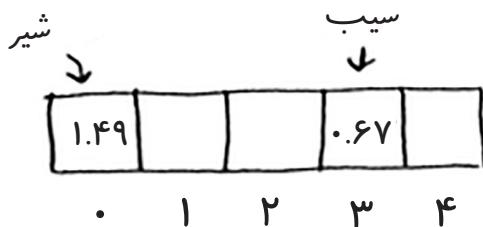
خروجی تابع هش «۳» خواهد بود. پس بیایید قیمت یک سیب را در ایندکس^۱ آرایه ذخیره کنیم.



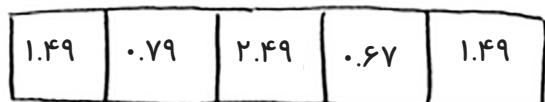
اجازه بدهید شیر را هم اضافه کنیم. به تابع هش «شیر» بدهید.



تابع هش «۰» را ارائه می‌کند. بیایید قیمت شیر را در ایندکس صفر ذخیره کنیم.



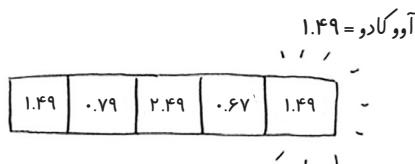
ادامه دهید، و در نهایت کل آرایه از قیمت‌ها پر خواهد شد.



حالا می‌پرسید: «هی، آووکادو چنده؟» شما نیازی به جستجوی آن در آرایه ندارید. فقط «آووکادو» را وارد تابع هش می‌کنید.



این به شما می‌گوید که قیمت در ایندکس ۴ ذخیره می‌شود. و همان‌طور که انتظار می‌رفت نتیجه را مشاهده می‌کنید.



تابع هش به شما می‌گوید که قیمت دقیقاً در چه نقطه‌ای ذخیره شده است، بنابراین اصلاً نیازی به جستجو نخواهید داشت! چنین چیزی با توجه به موارد زیر امکان‌پذیر است:

• تابع هش به طور پیوسته یک نام را به یک ایندکس یکسان نگاشت می‌کند. هر بار که «آووکادو» را وارد می‌کنید، همان عدد را برمی‌گرداند. بنابراین می‌توانید برای اولین بار از آن برای یافتن محل ذخیره‌ی قیمت آووکادو استفاده کنید و بعد می‌توانید از آن برای پیداکردن جایی که آن قیمت را ذخیره کرده‌اید، استفاده کنید.

• تابع هش رشته‌های مختلف را به ایندکس‌های مختلف نگاشت می‌کند. «آووکادو» به ایندکس ۴ نگاشت می‌شود. «شیر» به ایندکس صفر نگاشت می‌شود. همه چیز به یک روزنه‌ی متفاوت در آرایه که می‌توانید قیمت آن را ذخیره کنید نگاشت می‌شود.

• تابع هش می‌داند که آرایه‌ی شما چه اندازه‌ای دارد و فقط ایندکس‌های معتبر^۱ را برمی‌گرداند. بنابراین اگر آرایه‌ی شما ۵ آیتم باشد، تابع هش ۱۰۰ را برنمی‌گرداند ... چون ۱۰۰ یک ایندکس معتبر در آرایه نیست.

خب شما یک «مگی» ساختید! یک تابع هش و یک آرایه را کنار هم قرار دهید و ساختمان‌داده‌ای به نام جدول هش دریافت می‌کنید. جدول هش اولین ساختمان‌داده با منطق مضاعف است که یاد می‌گیرید. آرایه‌ها و لیست‌ها مستقیماً به حافظه نگاشت می‌شوند، اما جدول‌های هش هوشمندتر هستند. آن‌ها از یک تابع هش برای تشخیص هوشمندانه‌ی محل ذخیره‌ی عناصر استفاده می‌کنند.

جدول‌های هش احتمالاً مفیدترین ساختمان‌داده‌ی پیچیده‌ای هستند که یاد می‌گیرید. آن‌ها همچنین به عنوان maps، hash maps، دیکشنری‌ها و آرایه‌های ربطی یا انجمانی^۲ شناخته می‌شوند. و نکته‌ی دیگر اینکه جدول‌های هش سریع هستند! بحث مادر مورد آرایه‌ها و لیست‌های پیوندی در فصل ۲ را به خاطر دارید؟ شما می‌توانید یک آیتم را فوراً از یک آرایه دریافت کنید. و جدول‌های هش از یک آرایه برای ذخیره‌ی داده‌ها استفاده می‌کنند، بنابراین جداول هش به همان اندازه سریع هستند.

احتمالاً هرگز مجبور نخواهید شد که جدول‌های هش را خودتان پیاده‌سازی کنید. هر زبان خوب جدول هش پیاده‌سازی شده‌ی خود را دارد. پایتون هم دارای جدول هش است و به آن‌ها دیکشنری می‌گویند. با استفاده از تابع dict می‌توانید یک جدول هش جدید بسازید:

1. valid
2. associative arrays

```
>>> book = dict()
```



یک جدول هش جدید است. بایاید چند قیمت

به book اضافه کنیم:

```
>>> book["apple"] = 0.67 ←----- قیمت یک سیب
>>> book["milk"] = 1.49 ←----- قیمت شیر
>>> book["avocado"] = 1.49
>>> print book
{'avocado': 1.49, 'apple': 0.67, 'milk': 1.49}
```

به همین سادگی! حالا بایاید قیمت آوکادو را بپرسیم:

```
>>> print book["avocado"]
1.49 ←----- قیمت آوکادو
```

جدول‌های هش دارای کلیدها و مقادیر است. در هش book، نام محصولات کلید، و قیمت آن‌ها مقادیرند. جدول هش کلیدها را به مقدار آن‌ها نگاشت می‌کند. در بخش بعدی، نمونه‌هایی را مشاهده خواهید کرد که جدول هش برای آن‌ها بسیار مناسب است.

1. keys
2. values

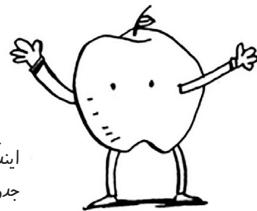
تمرین

برای توابع هش مهم است که به طور مداوم همان خروجی را برای ورودی یکسان بگرداند. اگر این کار را نکنند، بعد از اینکه آیتم خود را در جدول هش قرار دادید، نمی‌توانید آن را پیدا کنید! کدام یک از این توابع هش معتبر هستند؟

5.1 $f(x) = 1 \leftarrow$ برای تساوی ورودی‌ها «۱» را برمی‌گرداند

5.2 $f(x) = \text{rand}() \leftarrow$ هر دفعه یک عدد تصادفی برمی‌گرداند

5.3 $f(x) = \text{next_empty_slot}() \leftarrow$ ایندکس اسلات خالی بعدی
جدول هش را برمی‌گرداند



5.4 $f(x) = \text{len}(x) \leftarrow$ از طول رشته به عنوان ایندکس استفاده می‌کند

موارد کاربرد

جدول‌های هش در همه‌جا کاربرد دارد. این بخش چند مورد استفاده‌ی آن را به شما نشان می‌دهد.

استفاده از جدول هش برای جست‌وجو

موبایل شما یک دفترچه تلفن دم دستی و به درد بخور دارد.
همراه هر نام یک شماره تلفن هست.

BADE MAMA → ۵۸۱ ۶۶۰ ۹۸۲۰
ALEX MANNING → ۴۸۴ ۲۳۴ ۴۶۸۰
JANE MARIN → ۴۱۵ ۵۶۷ ۳۵۷۹



فرض کنید مطابق این شکل می‌خواهید یک دفترچه تلفن داشته باشد. شما در حال نگاشت نام افراد به شماره تلفن هستید. دفترچه تلفن شما باید این ویژگی‌ها را داشته باشد:

- نام یک فرد و شماره تلفن مرتبط با آن شخص را اضافه کنید.

- نام شخص را وارد کنید و شماره تلفن مرتبط با آن نام را دریافت کنید.

این یک کاربرد عالی برای جدول‌های هش است! جدول‌های هش برای موارد زیر

فوق العاده هستند:

- ایجاد یک نگاشت از چیزی به چیز دیگر

- جست‌وجو به دنبال چیزی

ساخت دفترچه تلفن بسیار آسان است. ابتدا یک جدول هش جدید بسازید:

```
>>> phone_book = dict()
```

راستی، پایتون یک میان‌بر برای ساخت جدول هش دارد. می‌توانید از دو آکولا در استفاده کنید:

```
>>> phone_book = {} <----- phone_book = dict() مشابه
```

بیایید شماره تلفن برخی از افراد را به این دفترچه تلفن اضافه کنیم:

```
>>> phone_book["jenny"] = 8675309
```

```
>>> phone_book["emergency"] = 911
```

همین! حالا فرض کنید می‌خواهید شماره تلفن جنی را پیدا کنید. فقط کافی است کلید را به هش ارسال کنید:

```
>>> print phone_book["jenny"]
شاره تلفن جنی 8675309 <-----
```

جنی	8675309
امرجنسی	911

یک جدول هش به عنوان دفترچه تلفن

تصور کنید اگر ناچار بودید این کار را با استفاده از یک آرایه انجام دهید. چگونه انجام

می‌دادید؟ جدول‌های هش مدل‌سازی رابطه از یک آیتم به آیتمی دیگر را ساده می‌کند. جدول‌های هش برای جست‌وجو در مقیاس بسیار بزرگ‌تر استفاده می‌شود. به عنوان مثال، فرض کنید به وب‌سایتی مانند <http://adit.io> می‌روید. کامپیوتر شما باید را به یک آدرس IP ترجمه کند.

ADIT.IO → 173.255.248.55

برای هر وب‌سایتی که بازدید می‌کنید، آدرس آن باید به یک آدرس IP ترجمه شود.

GOOGLE.COM → 74.125.239.133

FACEBOOK.COM → 173.252.120.6

SCRIBD.COM → 23.235.47.175

چه عالی، نگاشت یک آدرس وب به یک آدرس IP؟ به نظر می‌رسد یک کاربرد فوق العاده برای جدول‌های هش باشد! این فرآیند *DNS resolution* نامیده می‌شود. جدول‌های هش پکی از روش‌های دستیابی به این امکان است.



جلوگیری از ورودی‌های تکراری

فرض کنید مسئول یکی از حوزه‌های اخذ رأی هستید. قاعده‌تاً هر فرد تنها یک بار می‌تواند رأی بدهد. چگونه مطمئن می‌شوید که افراد پیش‌تر رأی نداده‌اند؟ وقتی کسی برای رأی‌دهی وارد می‌شود، نام کامل او را می‌پرسید. سپس آن را با لیست افرادی که رأی داده‌اند چک می‌کنید.



اگر نام آن فرد در لیست باشد، این شخص قبلاً رای داده است - او را از محل بیرون کنید! در غیر این صورت اسمش را به لیست اضافه می کنید و اجازه رأی دادن می دهد. حالا فرض کنید افراد زیادی برای رأی دهی آمده اند و لیست افرادی که رأی داده اند بسیار طولانی است.



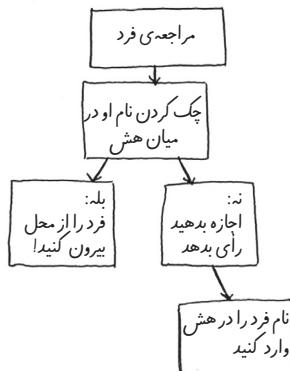
هر بار که فرد جدیدی برای رأی دادن وارد می شود، باید این فهرست بلند بالا را بررسی کنید تا بینید آیا قبلاً رأی داده یا خیر. اما راه بهتری هم هست: از هش استفاده کنید! ابتدا یک هش برای افرادی که رأی داده اند ایجاد کنید:

```
>>> voted = {}
```

وقتی فردی جدید برای رأی دادن وارد می شود، بررسی کنید که آیا از قبل در هش فهرست شده است یا خیر:

```
>>> value = voted.get("tom")
```

اگر «tom» در جدول هش باشد، تابع `get` مقدار آن را برمی گرداند. در غیر این صورت، `None` را برمی گرداند. می توانید از این برای بررسی اینکه آیا کسی قبلاً رأی داده است یا خیر استفاده کنید!



کد آن به این شکل است:

```
voted = {}
```

```
def check_voter(name):
    if voted.get(name):
        print("Kick them out");
    else
        voted[name] = True
        print("let them vote!")
```

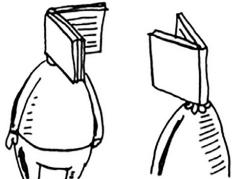
بیایید چند بار آن را تست کنیم:

```
>>> check_voter("tom")
!let them vote!
>>> check_voter("mike")
let them vote!
>>> check_voter("mike")
kick them out!
```

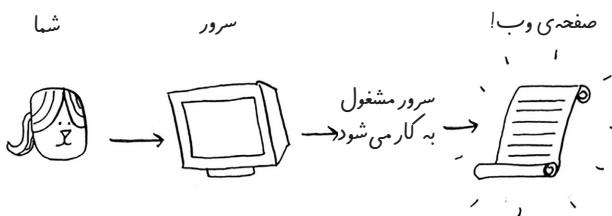
اولین باری که تام وارد می‌شود، این عبارت چاپ خواهد شد، "می‌تواند رأی دهد!" سپس مایک وارد می‌شود و چاپ می‌کند: «می‌تواند رأی دهد!» سپس مایک برای بار دوم تلاش می‌کند که برود، و چاپ می‌کند: «او را بیرون کنید!» به یاد داشته باشید، اگر این نام‌ها را در لیست افراد رأی داده ذخیره می‌کردید، این تابع در نهایت بسیار کند می‌شد، زیرا می‌باشد یک جست‌وجوی ساده در کل لیست انجام می‌داد. اما در عوض شما نام آن‌ها را در یک جدول هش ذخیره می‌کنید و یک جدول هش فوراً به شما می‌گوید که آیا نام این شخص در جدول هش وجود دارد یا خیر. بررسی موارد مشابه و تکراری با یک جدول هش بسیار سریع است.

استفاده از جدول هش برای کش^۱

و در نهایت کاربرد آن در کش کردن. اگر روی یک وبسایت کار می‌کنید، ممکن است قبلًاً نام کش را به عنوان یک ویژگی خوب شنیده باشید. قضیه از این قرار است. فرض کنید از facebook.com بازدید می‌کنید:



۱. شما یک درخواست^۲ به سرور فیس بوک ارسال می‌کنید.
۲. سرور یک ثانیه فکر می‌کند و صفحه‌ی وب را برای شما ارسال می‌کند.
۳. شما یک صفحه‌ی وب دریافت می‌کنید.



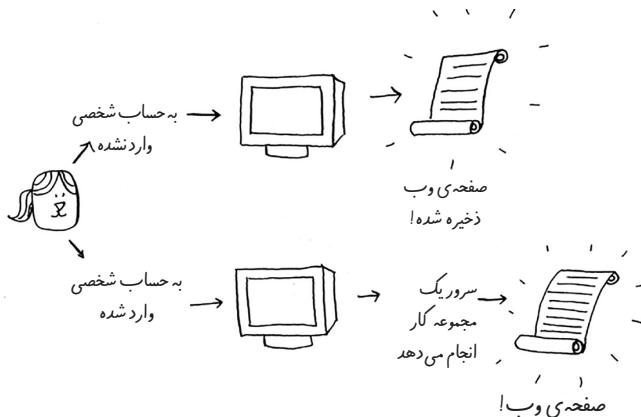
به عنوان مثال، در فیس بوک، سرور ممکن است تمام فعالیت‌های دوستان شما را جمع آوری کند تا به شما نشان بدهد. چند ثانیه طول می‌کشد تا تمام آن فعالیت‌ها جمع آوری شده و آن را به شما نشان بدهد. این چند ثانیه برای یک کاربر می‌تواند طولانی باشد. ممکن است فکر کنید، «چرا فیس بوک این قدر کند است؟» از طرف دیگر، سرورهای فیس بوک باید به میلیون‌ها نفر خدمات رسانی کنند و این چند ثانیه‌ها با هم جمع می‌شوند. سرورهای فیس بوک به سختی کار می‌کنند تا به همه‌ی این وبسایت‌ها سرویس دهند. آیا راهی برای سریع‌تر کردن فیس بوک و اینکه سرورهایش هم‌زمان کار کمتری انجام دهند وجود دارد؟

فرض کنید خواهرزاده‌ای دارید که مدام از شما در مورد سیارات سؤال می‌کند. «فاسله‌ی مریخ از زمین چقدر؟» «فاسله‌ی ماه چقدر؟» «فاسله‌ی مشتری چقدر؟» هر بار باید در گوگل سرچ کنید و جواب او را بدھید. چند دقیقه‌ای طول می‌کشد.

1. cache
2. request

حالا، فرض کنید دائم می‌پرسد: «فاصله‌ی ماه چقدر؟» خیلی زود، به یاد می‌آورید که ماه ۲۳۸۹۰۰ مایل دورتر است. نیازی نیست در گوگل دنبال جواب بگردید... فقط به یاد می‌آورید و پاسخ می‌دهید. نحوه‌ی عملکرد کش به این صورت است: وب‌سایت‌ها به جای محاسبه‌ی مجدد داده، آن را به خاطر می‌آورند.

اگر در حساب کاربری خود در فیسبوک هستید، محتوا‌ای که می‌بینید به شکل اخلاقی برای شما چیده شده است. هر بار که به facebook.com می‌روید، سوررهای آن باید به محتوای مورد علاقه‌ی شما فکر کنند. اما اگر وارد حساب کاربری خود در فیسبوک نشده باشید، صفحه‌ی ورود را مشاهده می‌کنید. همه یک صفحه‌ی ورود واحد را می‌بینند. از فیسبوک بارها و بارها این درخواست یکسان ارسال می‌شود: «وقتی از حساب کاربری خارج شدم من را به صفحه‌ی اصلی منتقل کن.» بنابراین عوض آن‌که سورور را وادار کند تا شکل صفحه‌ی اصلی را بفهمد، شکل صفحه‌ی اصلی را در خاطر می‌سپارد و همان را برای شما ارسال می‌کند.



نام این فرآیند کش است. این کار دو مزیت دارد:

- صفحه‌ی وب را بسیار سریع‌تر دریافت می‌کنید، درست مانند زمانی که فاصله‌ی زمین تا ماه را به خاطر می‌سپارید. دفعه‌ی بعد که خواهرزاده‌ی شما سؤال کرد، نیازی نخواهد بود تا جواب آن را در گوگل جست‌وجو کنید. می‌توانید فوراً پاسخ دهید.

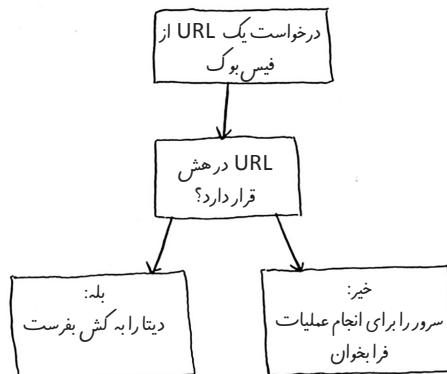
- از فیس بوک کمتر کار کشیده می شود.
کش کردن یک راه رایج برای سرعت بخشیدن به کارها است. همه‌ی وب سایت‌های بزرگ از کش استفاده می‌کنند. و این داده‌ها در یک هش ذخیره می‌شوند!

فیس بوک فقط صفحه‌ی اصلی را کش نمی‌کند. صفحه‌ی درباره، صفحه‌ی تماس، صفحه‌ی شرایط و ضوابط و بسیاری موارد دیگر را هم در حافظه کش می‌کند. بنابراین لازم است از URL صفحه به داده‌های صفحه نگاشت کند.

`facebook.com/about → DATA FOR THE ABOUT PAGE`

`facebook.com → DATA FOR THE HOME PAGE`

هنگامی که از یک صفحه در فیس بوک بازدید می‌کنید، ابتدا بررسی می‌کند که آیا صفحه‌ی مورد نظر در هش ذخیره شده است یا خیر.



کد آن به شکل زیر است:

```

cache = {}

def get_page(url):
    if cache.get(url):
        return cache[url] ← دیتابابکش شده را بر می‌گرداند
  
```

```

else:
    data = get_data_from_server(url)
    cache[url] = data   ← ابتدا این داده را در کش ذخیره می‌کند
    return data

```

در اینجا، تنها در صورتی از سرور کار می‌کشید که URL مورد نظر در کش موجود نباشد. با این حال، قبل از اینکه داده‌ها را برگردانید، آن‌ها را در کش ذخیره می‌کنید. دفعه‌ی بعد که شخصی این URL را درخواست کرد، به جای اینکه سرور را مجبور به انجام کار کنید، داده‌ها را می‌توانید از کش ارسال کنید.

جمع‌بندی

به طور خلاصه، هش برای این موارد مناسب هستند:

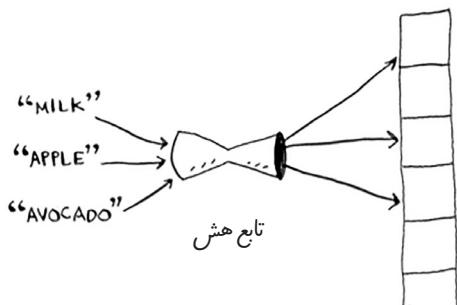
- مدل‌سازی رابطه‌ها از یک چیز به چیز دیگر
- فیلتر کردن موارد تکراری و مشابه
- کش کردن / به خاطر سپردن داده‌ها به جای اینکه سرور مشغول شود.

تصادم‌ها

همان‌طور که قبلاً گفتم، بیشتر زبان‌ها دارای جدول هش هستند. نیازی نیست تا نوشتن آن را بلد باشید. بنابراین، چندان در مورد کدهای داخلی جدول‌های هش صحبت نمی‌کنم. اما کارایی آن همچنان برای شما اهمیت دارد! برای درک کارایی جدول‌های هش، ابتدا باید بدانید که

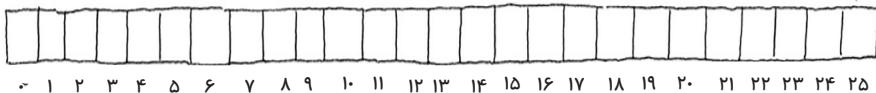
تصادم چیست. در دو بخش بعدی به تصادم‌ها و کارایی می‌پردازیم.

اول از همه اعتراف کنم که قبل‌تر به شما یک دروغ مصلحتی گفتم. به شما گفتم که یک تابع هش همیشه کلیدهای مختلف را به اسلات‌های مختلف در آرایه نگاشت می‌کند.

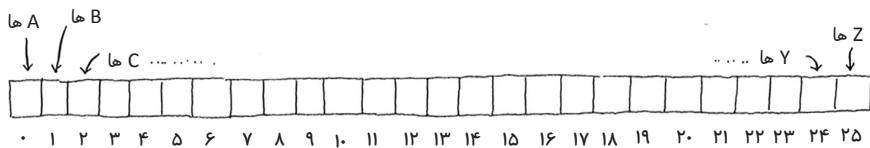


در واقعیت، نوشتن یک تابع هش که این کار را انجام بدهد تقریباً غیر ممکن است.

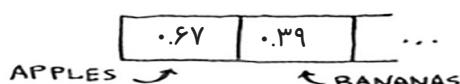
باید یک مثال ساده بزنیم. فرض کنید آرایه‌ی شما ۲۶ اسلات دارد.



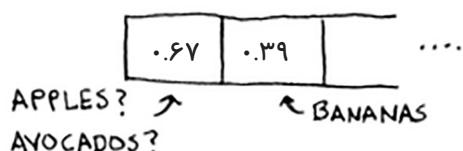
تابع هش شما بسیار ساده است: بر اساس حروف الفبا یک نقطه از آرایه را اختصاص می‌دهد.



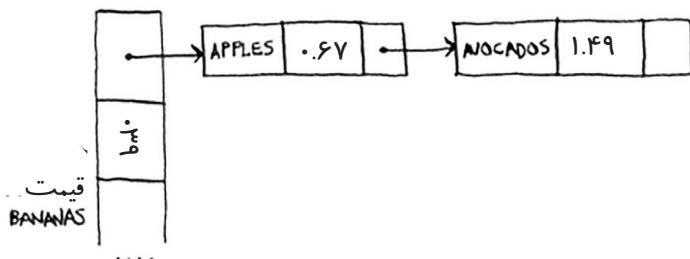
شاید متوجه مشکل شده باشد. قیمت سیب (apple) را می‌خواهید در هش وارد کنید. اولین روزنہ به شما اختصاص داده می‌شود. سپس می‌خواهید قیمت موز (banana) را در هش قرار بدھید. روزنہ‌ی دوم برای آن اختصاص می‌یابد.



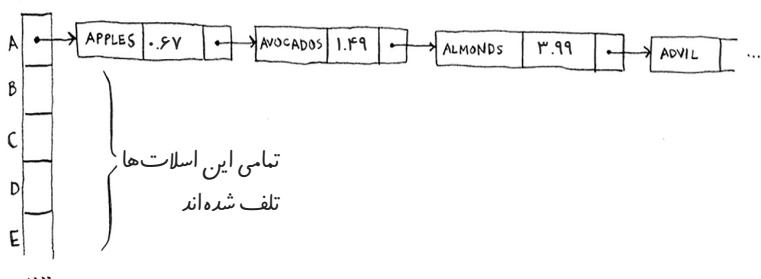
همه چیز به خوبی پیش می‌رود! اما اگر یون می‌خواهید قیمت آووکادو (avocado) را در هش قرار بدھید. دوباره اولین روزنہ به شما اختصاص داده می‌شود.



نه! سیب‌ها قبل‌آن اسلات را اشغال کرده‌اند! چه باید کرد؟ به این وضعیت تصادم می‌گویند: اختصاص دو کلید به یک اسلات. این یک مشکل است. اگر قیمت آووکادو را در آن جایگاه ذخیره کنید، قیمت سیب را بازنویسی می‌کنید. سپس نوبت بعدی که کسی قیمت سیب را می‌پرسد، به جای آن قیمت آووکادو را دریافت می‌کند! تصادم‌ها بد هستند و باید راه حل جایگزین برای آن درنظر گرفت. راه‌های مختلفی برای مقابله با تصادم وجود دارد. ساده‌ترین مورد این است: اگر چندین کلید به یک اسلات نگاشت می‌شوند، یک لیست پیوندی را در آن اسلات ایجاد کنید.



در این مثال، هر دو «apple» و «avocado» به یک اسلات نگاشت می‌شوند. بنابراین شما یک لیست پیوندی را در آن اسلات ایجاد می‌کنید. اگر بخواهید از قیمت موز باخبر شوید، همچنان سریع خواهد بود. اگر می‌خواهید قیمت سیب را بدانید، کمی کنتر است. برای یافتن «apple» باید از طریق این لیست پیوندی جستجو کنید. اگر لیست پیوندداده شده کوچک باشد، مشکلی اساسی پیش نمی‌آید - باید سه یا چهار عنصر را جستجو کنید. اما فرض کنید در یک فروشگاه مواد غذایی کار می‌کنید که در آن فقط محصولاتی را می‌فروشید که با حرف A شروع می‌شوند.



یک دقیقه صبر کنید! کل جدول هش به جز یک اسلات کاملاً خالی است. و آن اسلات یک لیست پیوندی بزرگی دارد! هر عنصر در این جدول هش در لیست پیوند داده شده است. این شرایط به همان اندازه بد است که برای شروع همه چیز را در یک لیست پیوندی قرار بدهید. جدول هش شما را کند می‌کند.

در اینجا به دو نکته پی می‌بریم:

- تابع هش شما بسیار مهم است. تابع هش همه کلیدها را در یک اسلات نگاشت کرده است. در حالت ایده‌آل، هش شما کلیدها را به طور یکنواخت در سراسر هش نگاشت می‌کند.

- اگر لیست‌های پیوندی بزرگ شوند، جدول هش شما بسیار کند می‌شود. اما اگر از یک تابع هش خوب استفاده کنید طول آن زیاد نخواهد شد!

تابع هش مهم هستند. در یک تابع هش خوب تصادم بسیار کمی رخ می‌دهد. پس چگونه یک تابع هش خوب را انتخاب می‌کنید؟ در بخش بعدی به این موضوع پرداخته می‌شود!

کارایی

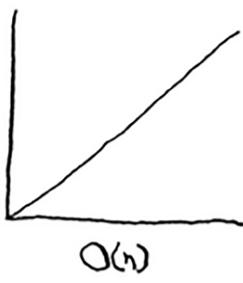
این فصل را در فروشگاه مواد غذایی شروع کردید. تصمیم داشتید سازوکاری ایجاد کنید که قیمت محصولات را فوراً به شما اعلام کند. خب، جدول‌های هش واقعاً سریع هستند. در حالت متوسط، جداول هش در همه شرایط زمان $O(1)$ دارد. $O(1)$ زمان ثابت!

جستجو	$O(1)$	$O(n)$
حالات	بدترین	متوسط
درج	$O(1)$	$O(n)$
حذف	$O(1)$	$O(n)$

نامیده می‌شود. تا اینجا کار با زمان ثابت مواجه نشده‌اید. این اصطلاح به معنی فوری نیست بلکه به این معناست که صرف نظر از اینکه جدول هش چقدر بزرگ باشد زمان صرف شده ثابت می‌ماند. به عنوان مثال، می‌دانید که جستجوی ساده زمان خطی دارد.

PERFORMANCE OF HASH TABLES

1. Constant time



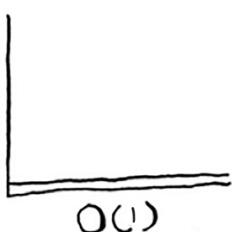
زمان خطی
(جست‌وجوی ساده)

جست‌وجوی دودویی با زمان لگاریتمی^۱ که دارد، سریع‌تر است:



زمان لگاریتمی
(جست‌وجوی دودویی)

جست‌وجو در جدول هش به زمان ثابت نیاز دارد.



زمان ثابت
(جدول‌های هش)

1. log time

می‌بینید که نمودار آن یک خط صاف است. به این معنی که فرقی نمی‌کند جدول هش شما ۱ عنصر داشته باشد یا ۱ میلیارد عنصر. دریافت هرچیز از جدول هش به زمان یکسانی نیاز دارد. در واقع، پیش از این زمان ثابت را دیده‌اید. خواندن یک آیتم از آرایه زمان ثابتی نیاز دارد. مهم نیست که آرایه‌ی شما چقدر بزرگ باشد. برای به‌دست آوردن یک عنصر به زمان یکسانی نیاز است. در حالت متوسط، جداول هش واقعاً سریع هستند.

در بدترین حالت، یک جدول هش برای همه چیز به $O(n)$ - زمان خطی - نیاز دارد که واقعاً کند است. بیایید جدول‌های هش را با آرایه‌ها و لیست‌ها مقایسه کنیم.

لیست آرایه‌ها پیوندی	جدول‌های هش (بدترین)	جدول‌های هش (متوسط)		
جست‌وجو	$O(1)$	$O(n)$	$O(1)$	$O(n)$
درج	$O(1)$	$O(n)$	$O(n)$	$O(1)$
حذف	$O(1)$	$O(n)$	$O(n)$	$O(1)$

به حالت متوسط برای جدول‌های هش نگاه کنید. جدول‌های هش برای جست‌وجو به همان سرعت آرایه‌ها هستند (درایافت مقدار در یک ایندکس). و به همان سرعت لیست‌های پیوندی در درج و حذف هستند. خیر دنیا و آخرت! اما در بدترین حالت، جدول‌های هش در همه‌ی این موارد کند هستند. بنابراین مهم است که با استفاده از جدول‌های هش کارایی بدترین حالت را نداشته باشید. و برای انجام این کار، باید از تصادم جلوگیری کنید. برای جلوگیری از تصادم، به این دو مورد نیاز دارید:

- ضریب بار کم
- یک تابع هش خوب

نکته

پیش از آن که ادامه دهیم خوب است بدانید که خواندن این بخش ضروری نیست. در این قسمت می‌خواهم در مورد شیوه‌ی پیاده‌سازی جدول هش صحبت کنم، اما هیچ‌گاه نیاز پیدا نمی‌کنید که شخصاً کد آن را پیاده‌سازی کنید. از هر زبان برنامه‌نویسی که استفاده کنید، یک نسخه‌ی پیاده‌سازی شده از جدول‌های هش دارد. می‌توانید از جدول هش داخلی زبان برنامه‌نویسی استفاده کنید و مطمئن باشید که کارایی خوبی دارد. بخش بعدی برای آن است که از فرآیند اجرایی در پس زمینه آگاه شوید.

ضریب بار(لود فاکتور)

محاسبه‌ی ضریب بار جدول هش آسان است. تعداد آیتم‌ها در جدول هش تقسیم بر تعداد کل اسلات‌ها.

جدول‌های هش از یک آرایه برای ذخیره‌سازی استفاده می‌کنند، به همین دلیل تعداد اسلات‌های اشغال شده در یک آرایه را می‌شمارید. به عنوان مثال، این جدول هش دارای ضریب بار $2/5$ یا 4 .

اشغال شده	لا
↓	
۱	۰

ضریب بار = $\frac{۲}{۵}$

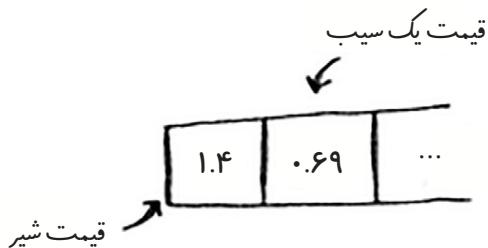
ضریب بار این جدول هش چیست؟

	۲۰	
--	----	--

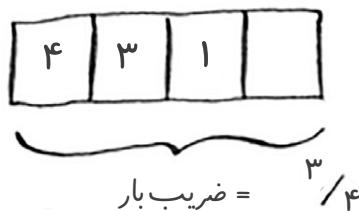
ضریب بار؟

جواب صحیح $1/3$ است. ضریب بار، تعداد اسلات‌های خالی در جدول هش را اندازه‌گیری می‌کند.

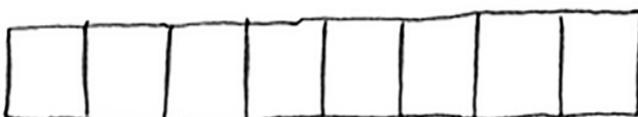
فرض کنید باید قیمت 100 آیتم محصول را در جدول هش خود ذخیره کنید و جدول هش شما 100 اسلات دارد. در بهترین حالت، هر آیتم جایگاه مخصوص به خود را خواهد داشت.



این جدول هش دارای ضریب بار ۱ است. اگر جدول هش شما فقط ۵۰ اسلات داشته باشد چه؟ ضریب بار ۲ خواهد داشت. هیچ راهی وجود ندارد که هر آیتم اسلات مخصوص به خود را داشته باشد، زیرا اسلات‌های کافی وجود ندارد! داشتن ضریب بار بیشتر از ۱ به این معنی است که آیتم‌های بیشتری نسبت به اسلات در آرایه‌ی خود دارید. هنگامی که ضریب بار شروع به رشد کند، باید اسلات‌های بیشتری را به جدول هش خود اضافه کنید. به این عمل تغییر اندازه^۱ می‌گویند. برای مثال، فرض کنید جدول هش زیر را دارید که در آستانه‌ی پرشدن است.

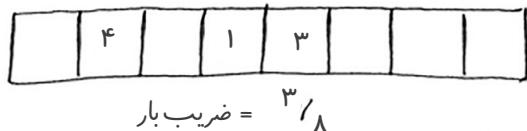


باید اندازه‌ی این جدول هش را تغییر دهید. ابتدا یک آرایه‌ی جدید ایجاد می‌کنید که بزرگ‌تر باشد. قاعده‌ی سرانگشتی^۲ این است که آرایه‌ای با اندازه‌ی دو برابر آن بسازید.



1. resizing
2. rule of thumb

اکنون باید تمامی آن آیتم‌ها را مجدداً با استفاده از تابع $hash$ در جدول هش جدید وارد کنید:

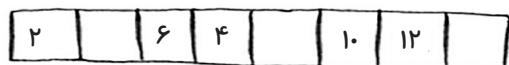


این جدول جدید دارای ضریب بار $3/8$ است. خیلی بهتر از پیش! با ضریب بار کمتر، تصادم‌های کمتری خواهد داشت و جدول شما عملکرد بیشتری خواهد داشت. یک قانون سرانگشتی خوب این است که وقتی ضریب بار شما بیشتر از $7/8$ است، اندازه‌ی آن را تغییر دهید.

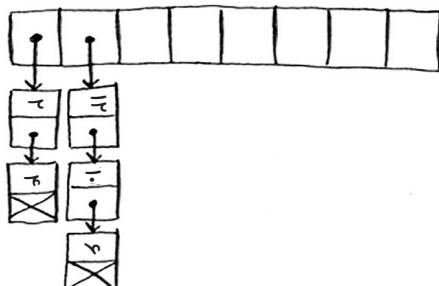
ممکن است فکر کنید، «این تغییر اندازه زمان بر است!» حق با شماست تغییر اندازه پرهزینه است و اغلب اوقات تمایل ندارید اندازه را تغییر دهید. اما به طور متوسط، جدول‌های هش حتی با تغییر اندازه، زمان $(1/10)$ دارند.

تابع هش مناسب

تابع هش مناسب مقادیر را در آرایه به شکل مساوی توزیع می‌کند.



تابع هش بد مقادیر را با هم گروه‌بندی می‌کند و تصادم‌های بسیاری ایجاد می‌کند.



یکتابع هش خوب چه ویژگی هایی دارد؟ این مسئله هرگز دغدغه‌ی شما نخواهد بود – پیرمردهایی با ریش‌های بلند (و پیزنهایی) در اتاق‌های تاریک می‌نشینند و دغدغه‌ی حل این مسائل را دارند. اگر همچنان کنجکاو هستید، تابع SHA را جست‌جو کنید (توضیح کوتاهی درباره‌ی آن در فصل آخر وجود دارد). می‌توانید از آن به عنوان تابع هش خود استفاده کنید.

تمرين

توابع هش باید توزیع خوبی داشته باشند. آن‌ها باید آیتم‌ها را تا حد امکان هرچه گسترده‌تر نگاشت کنند. بدترین حالت برای یک تابع هش این است که همه‌ی آیتم‌ها را در جدول، هش، به اسلات یکسانی، نگاشت کند.

چهار تابع هش زیر را که با رشته‌ها^۱ کار می‌کنند، فرض کنید:

آ. «۱) را برای همه‌ی ورودی‌ها برگردانید.

ب. از طول رشته به عنوان اپنده استفاده کنید.

پ. از اولین کاراکتر رشته به عنوان ایندکس استفاده کنید. بنابراین، تمام رشته هایی که با a شروع می شوند با هم هش می شوند و غیره.

ت. هر حرف را با یک عدد اول نگاشت کنید: $e = 11, d = 7, c = 5, b = 3, a = 2$ و غیره. برای یک رشته، تابع هش مجموع تمام کاراکترهای باقیمانده‌ی اندازه‌ی هش است. به عنوان مثال، اگر اندازه‌ی هش شما ۱۵ است و رشته‌ی آن «bag» است، ایندکس ۲ + ۳ + ۷ + ۱۱ = ۲۲٪ است.

برای هر یک از این مثال‌ها، کدام تابع‌های هش توزیع خوبی ارائه می‌کنند؟ اندازه‌ی حدول هش، را با ۱۰ اسلات فرض کنید.

۵.۵ دفترچه تلفنی که در آن کلیدها اسمی و مقادیر شماره‌ی تلفن هستند. اسمی عبارتند از: Dan, Bob, Ben, Esther.

۵.۶ نگاشته از اندازه‌ی باتری به توان باتری. اندازه‌ها A, AA, AAA و AAAA هستند.

۵.۷ نگاشتی از عنوان کتاب به اسمی نویسنده‌ها. عناوین *Watchmen*, *Maus* و *Fun Home*

هستند

جمع‌بندی

تقریباً هرگز نیازی به پیاده‌سازی جدول هش پیدا نمی‌کنید. زبان برنامه‌نویسی که استفاده می‌کنید باید جدول هش پیاده‌سازی شده را برای شما فراهم کند. می‌توانید از جدول‌های هش پایتون استفاده کنید و فرض کنید که کارایی حالت میانگین را دریافت خواهید کرد: زمان ثابت.

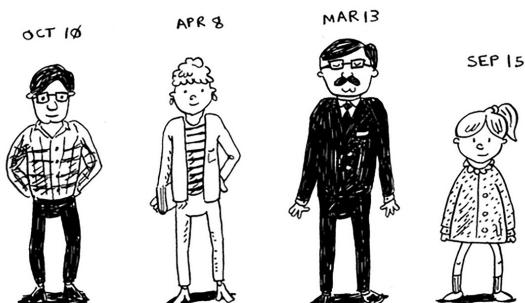
جدول هش یک ساختمان‌داده‌ی قدرتمند است زیرا بسیار سریع است و به شما امکان می‌دهند داده‌ها را به روشنی متفاوت مدل‌سازی کنید. احتمالاً به زودی متوجه می‌شوید که همیشه در حال استفاده از آن‌ها بوده‌اید:

- می‌توانید با ترکیب یک تابع هش با یک آرایه یک جدول هش بسازید.
- تصادم‌ها بد هستند. شما به یک تابع هش نیاز دارید که تصادم‌ها را به حداقل برساند.

- جدول‌های هش در جست‌وجو، درج و حذف واقعاً سریع هستند.
- جدول‌های هش برای مدل‌سازی روابط از یک آیتم به آیتم دیگر مناسب هستند.
- هنگامی که ضریب بار شما از ۷۰٪ بیشتر شد، زمان تغییر اندازه‌ی جدول هش فراسیده است.

جدول‌های هش برای ذخیره‌سازی داده‌ها (به عنوان مثال، از طریق یک وب‌سور) استفاده می‌شود.

- جدول‌های هش برای گرفتن موارد تکراری عالی هستند.



۶ | جستجوی سطح اول^۱



در این فصل

- یاد می‌گیرید یک شبکه را با استفاده از ساختمان‌داده‌ای جدید و انتزاعی مدل‌سازی کنید: گراف‌ها^۲.

- جستجوی سطح اول را یاد می‌گیرید، الگوریتمی که می‌توانید روی گراف‌ها اجرا کنید تا به سؤالاتی مانند «کوتاه‌ترین مسیر برای رفتن به X چیست؟» پاسخ دهید.

- در مورد گراف‌های جهت‌دار^۳ و گراف‌های غیرجهت‌دار^۴ یاد می‌گیرید.

- مرتب‌سازی توپولوژیکی^۵ را یاد می‌گیرید، نوع متفاوتی از الگوریتم مرتب‌سازی که وابستگی‌های بین گره‌ها را آشکار می‌کند.

.....

در این فصل گراف‌ها معرفی می‌شوند. ابتدا در مورد اینکه گراف چیست صحبت می‌کنم (آن‌ها محور X یا Y ندارند). سپس اولین الگوریتم گرافی که یاد می‌گیرید را به شما نشان می‌دهم. این الگوریتم جستجوی سطح اول (BFS) نام دارد.

1. Breadth-first search
2. graphs
3. Directed graph
4. Undirected graph
5. Topological sort

جست وجوی سطح اول به شما امکان می‌دهد کوتاه‌ترین فاصله‌ی میان دو چیز را پیدا کنید. اما کوتاه‌ترین فاصله می‌تواند معناهای گوناگونی داشته باشد! می‌توانید از جست وجوی سطح اول برای موارد زیر استفاده کنید:

- نوشتن کد هوش مصنوعی برای بازی چکرز که کمترین حرکت را برای پیروزی محاسبه کند.
- نوشتن یک تصحیح‌کننده‌ی غلط‌های املایی (کم‌ترین تغییر در حروف کلمه‌ای که غلط املایی دارد به یک کلمه‌ای درست برای مثال: READED -> READER)
- پیدا کردن نزدیک‌ترین پژوهش

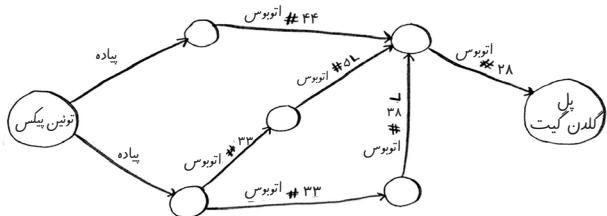
الگوریتم‌های گراف از مفیدترین الگوریتم‌هایی هستند که می‌شناسیم. چند فصل بعدی را حتماً با دقت مطالعه کنید این‌ها الگوریتم‌هایی هستند که می‌توانید بارها و بارها از آن‌ها استفاده کنید.

مقدمه‌ای بر گراف‌ها

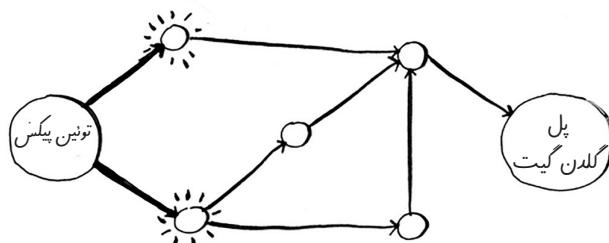


فرض کنید در سانفرانسیسکو هستید و می‌خواهید از تئین پیکس به پل گلدن گیت بروید. می‌خواهید با اتوبوس، با حداقل تغییر خط به مقصد برسید. گزینه‌های شما به شکل زیر است:

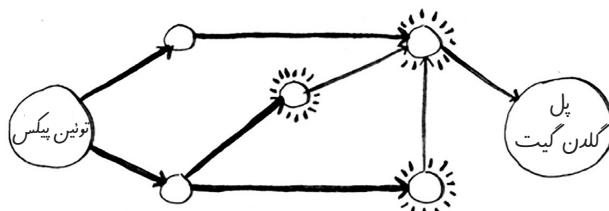
1. checkers



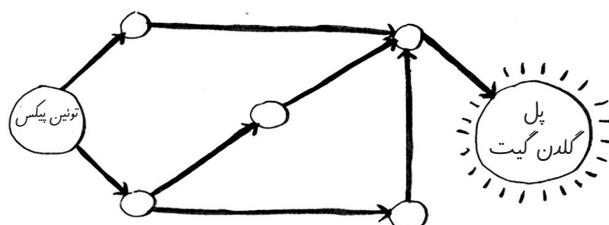
الگوریتم شما برای یافتن مسیر با کمترین گام چیست؟
خب، آیا می‌توانید در تنها یک مرحله به آنجا برسید؟ در اینجا تمام مکان‌هایی مشخص شده که می‌توانید در یک مرحله به آن‌ها برسید.



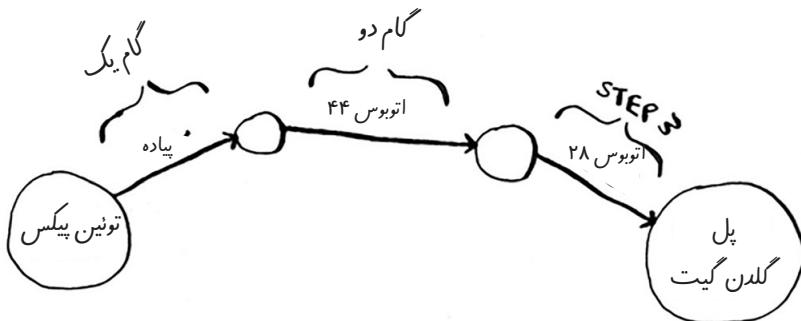
پل مورد نظر با علامت مشخص نشده است و نمی‌توانید در یک مرحله به آنجا برسید. در دو مرحله چطور؟



باز هم، پل آنجا نیست، بنابراین نمی‌توانید در دو مرحله به پل برسید. سه مرحله چطور؟



خب! بالاخره پل گلدن گیت نمایان می‌شود. بنابراین سه مرحله نیاز است تا مطابق این مسیر از توئین پیکس به مقصد برسید.



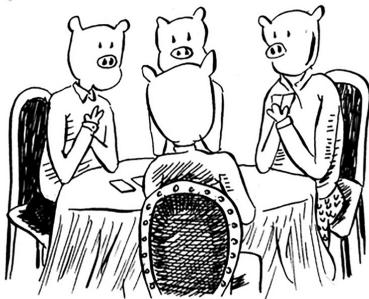
مسیرهای دیگری نیز وجود دارد که شما را به مقصد می‌رساند، اما طولانی‌تر هستند (چهار مرحله). الگوریتم تشخیص داد که کوتاه‌ترین مسیر به پل سه مرحله دارد. به این نوع مسئله، مسئله‌ی کوتاه‌ترین مسیر^۱ می‌گویند. همیشه در این نوع مسئله‌ها به دنبال کوتاه‌ترین چیز هستید. این می‌تواند کوتاه‌ترین مسیر برای رسیدن به خانه‌ی دوست شما باشد. یا می‌تواند کمترین تعداد حرکت برای مات‌کردن در یک بازی شطرنج باشد. الگوریتمی که برای حل یک مسئله کوتاه‌ترین مسیر انجام می‌شود، جست‌وجوی سطح اول نامیده می‌شود.

برای اینکه بفهمید چگونه از توئین پیکس به پل گلدن گیت بروید، به دو مرحله نیاز است:

۱. مسئله را به گراف مدل‌سازی کنید.
۲. مشکل را با استفاده از جست‌وجوی سطح اول حل کنید.

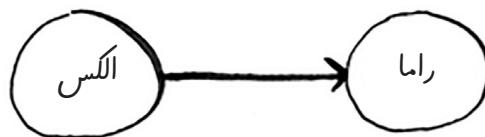
در ادامه به این می‌پردازم که گراف چیست. سپس با جزئیات بیشتر وارد جست‌وجوی سطح اول خواهم شد.

1. Shortest-path problem

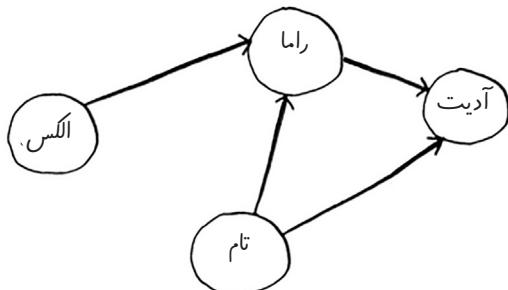


گراف چیست؟

گراف مجموعه‌ای از اتصالات^۱ را مدل‌سازی می‌کند. برای مثال، فرض کنید شما و دوستانتان مشغول بازی پوکر هستید و می‌خواهید مدل‌سازی کنید که کی به کی بدهکار است. در اینجا می‌توانید بگویید: «الکس به راما بدهکار است.»

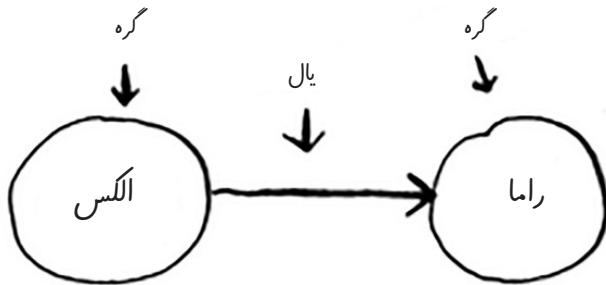


گراف کامل می‌تواند چیزی شبیه به این باشد.



گراف افرادی که به دیگران سر بازی پوکر پول بدهکار هستند. الکس به راما بدهکار است، تام به آدیت بدهکار است و به همین ترتیب. هر گراف از گره‌ها^۲ و یال‌ها^۳ تشکیل شده است.

1. connections
2. nodes
3. edges



تمام اجزای سازنده‌ی آن همین است که می‌بینید! گراف‌ها از گره‌ها و یال‌ها تشکیل شده‌اند. یک گره می‌تواند مستقیماً به بسیاری از گره‌های دیگر متصل شود. آن گره‌ها همسایه‌های آن گره هستند. در این گراف، راما همسایه‌ی الکس است. آدیت همسایه‌ی الکس نیست، زیرا مستقیماً به همدیگر متصل نیستند. اما آدیت همسایه‌ی راما و تام است.

گراف‌ها روشی برای مدل‌سازی نحوه‌ی اتصال چیزهای مختلف به یکدیگر هستند. اکنون بباید جست‌وجوی سطح اول را در عمل ببینیم.

جست‌وجوی سطح اول

در فصل ۱ به یکی از الگوریتم‌های جست‌وجو نگاه کردیم: جست‌وجوی دودویی. جست‌وجوی سطح اول نوع متفاوتی از الگوریتم جست‌وجو است: الگوریتمی که بر روی گراف‌ها اجرا می‌شود. می‌تواند به ما در پاسخ دو دسته پرسش کمک کند:

- پرسش نوع ۱: آیا مسیری از گره‌ی A به گره‌ی B وجود دارد؟

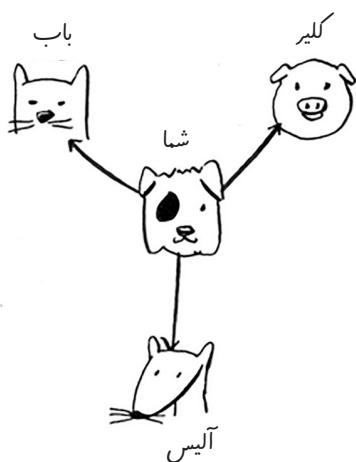
- پرسش نوع ۲: کوتاه‌ترین مسیر از گره‌ی A به گره‌ی B چیست؟

قبل‌زماني که کوتاه‌ترین مسیر از توئین پیکس تا پل گلدن گیت را محاسبه می‌کردید، با جست‌وجوی سطح اول مواجه شدید. این یک پرسش از نوع دوم بود: «کوتاه‌ترین مسیر چیست؟» حالا بباید الگوریتم را با جزئیات بیشتری بررسی کنیم. شما پرسشی از نوع

اولمیکنید: «آیا مسیری وجود دارد؟»

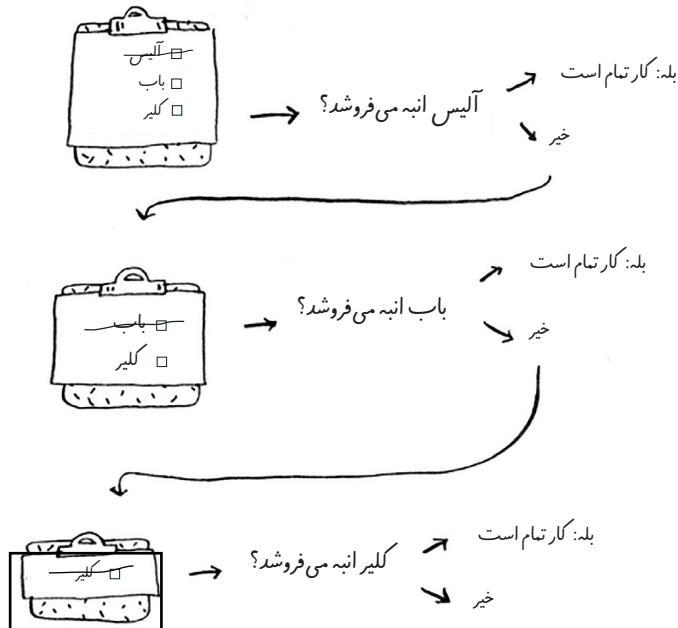


فرض کنید شما صاحب یک مزرعه‌ی انبه هستید و به دنبال فروشنده‌ای می‌گردید که بتواند انبه‌های شما را بفروشد. آیا به یک فروشنده‌ی انبه در فیس‌بوک ارتباط دارید؟ خوب، شما می‌توانید در میان دوستان خود جستجو کنید.

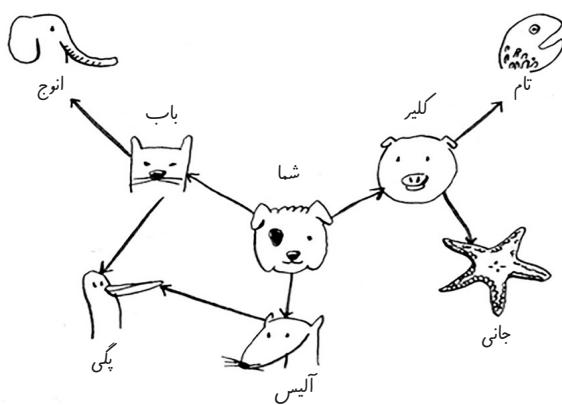


این جستوجوی بسیار ساده‌ای است. ابتدا لیستی از دوستان برای جستجو تهیه کنید.

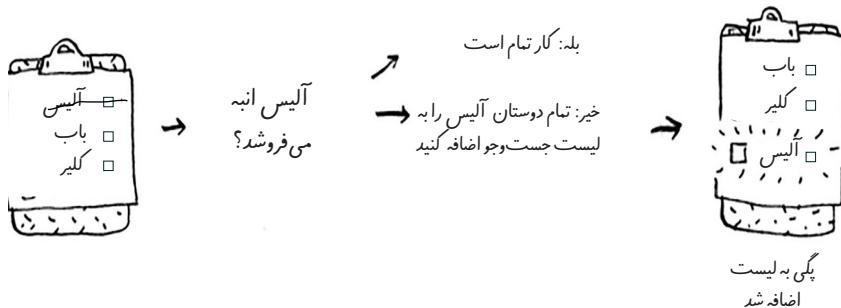
حالا به سراغ تک تک افراد لیست بروید و بررسی کنید که آیا آن شخص فروشنده‌ی انبه است یا خیر.



فرض کنید هیچ یک از دوستان شما فروشنده‌ی انبه نیست. اکنون باید از طریق دوستان خود جستجو کنید.



هر بار که فردی از لیست را جست وجو می‌کنید، همه‌ی دوستان او را به این لیست اضافه می‌کنید.



به این ترتیب، نه تنها در میان دوستان خود، بلکه در میان دوستان این افراد نیز جست وجو می‌کنید. به یاد داشته باشید، هدف این است که یک فروشنده‌ی انبه در شبکه‌ی خود پیدا کنید. بنابراین اگر آليس فروشنده‌ی انبه نیست، دوستان او را نیز به لیست اضافه کنید. این به این معنی است که شما در نهایت دوستان او را جست وجو خواهید کرد - و سپس دوستان این افراد و الی آخر. با این الگوریتم، تا زمانی که با فروشنده انبه رو به رو شوید، کل شبکه‌ی خود را جست وجو می‌کنید. این همان الگوریتم جست وجوی سطح اول است.

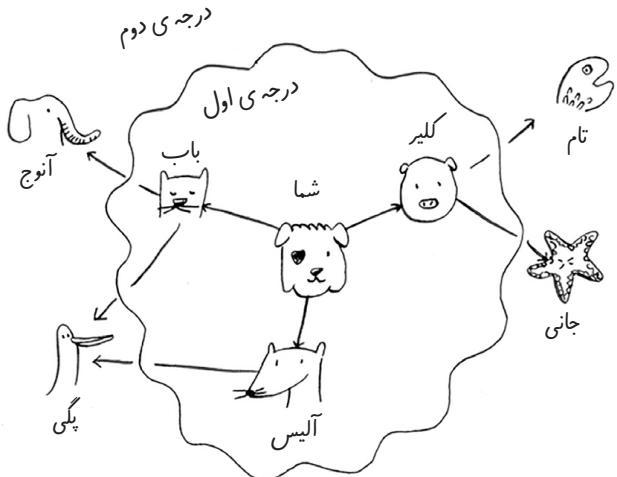
یافتن کوتاه‌ترین مسیر

به طور خلاصه، این دو پرسشی است که جست وجوی سطح اول می‌تواند پاسخ بددهد:

- پرسش نوع ۱: آیا مسیری از گرهی A به گرهی B وجود دارد؟ (آیا فروشنده‌ی انبه در شبکه‌ی شما وجود دارد؟)

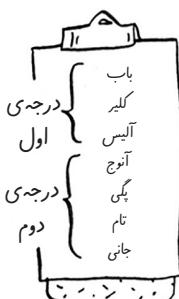
• پرسش نوع ۲: کوتاه‌ترین مسیر از گرهی A تا گرهی B چیست؟ (نزدیک‌ترین فروشنده‌ی انبه کیست؟)

نحوه‌ی پاسخ دادن به پرسش ۱ را دیدید. حالا باید تلاش کنیم به پرسش ۲ پاسخ دهیم. آیا می‌توانید نزدیک‌ترین فروشنده‌ی انبه را پیدا کنید؟ به عنوان مثال، دوستان شما رابطه‌ی درجه یک و دوستان آن‌ها رابطه‌ی درجه دو هستند.



ارتباط درجه اول را به ارتباط درجه دوم ترجیح می‌دهید و ارتباط درجه دوم را به ارتباط درجه سوم و غیره ترجیح می‌دهید. بنابراین، قبل از اینکه مطمئن شوید یک ارتباط درجه یک با فروشنده انبه ندارید، باید هیچ ارتباط درجه دومی را جست‌جو کنید. خوب، جست‌جوی سطح اول این کار را انجام می‌دهد! روشی که جست‌جوی سطح اول کار می‌کند، جست‌جو از نقطه‌ی شروع به بیرون پخش می‌شود. بنابراین، ارتباطات درجه یک را قبل از ارتباطات درجه دوم بررسی می‌کنید. سؤال: کلر اول بررسی می‌شود یا آنوج؟ پاسخ: کلر یک ارتباط درجه یک است و آنوج یک ارتباط درجه دو است. بنابراین کلر قبل از آنوج بررسی می‌شود.

به عبارت دیگر ارتباطات درجه اول قبل از ارتباطات درجه دوم به لیست جست‌جو اضافه می‌شوند.



شما فقط اسامی لیست را مرور می‌کنید و می‌بینید که آیا فروشنده‌ی انبه هستند یا خیر. ارتباطات درجه اول قبل از ارتباطات درجه دوم جست‌جو می‌شوند، بنابراین نزدیک‌ترین انبه‌فروش را پیدا خواهید کرد. جست‌جوی سطح اول نه تنها یک مسیر از A به B ، بلکه کوتاه‌ترین مسیر را پیدا می‌کند.

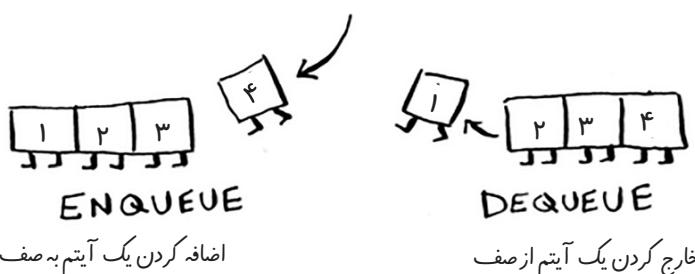
توجه داشته باشید که این شرایط تنها در صورتی صدق می‌کند که افراد را به همان ترتیبی که اضافه شده‌اند جستجو کنید. یعنی اگر کلر قبل از آنوج به لیست اضافه شده باشد، باید او را قبل از آنوج جستجو کنید. اگر آنوج را قبل از کلر جستجو کنید و هر دو فروشنده‌ی انبه باشند، چه اتفاقی می‌افتد؟ خب، آنوج یک ارتباط درجه دو است و کلر یک ارتباط درجه یک. در نهایت با فروشنده‌ای مواجه می‌شوید که در شبکه‌ی شما نزدیک‌ترین فرد به شما نیست. بنابراین باید افراد را به ترتیبی که اضافه شده‌اند جستجو کنید. یک ساختمنداند برای چنین موقعیتی وجود دارد: صف.

صف^۱



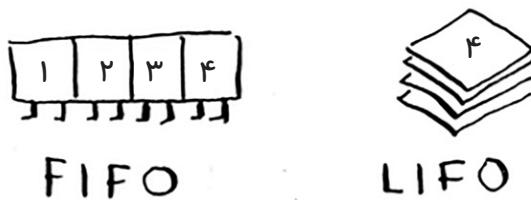
عملکرد صف دقیقاً مانند کاربرد آن در زندگی واقعی است. فرض کنید شما و دوستتان در ایستگاه اتوبوس در صف ایستاده‌اید. اگر قبل از او در صف هستید، ابتدا سورا اتوبوس می‌شوید. یک صف به همین ترتیب کار می‌کند. شما صفاتی که شبیه به پشته‌ها هستند. شما نمی‌توانید به عناصر تصادفی در صف

دسترسی داشته باشید. در عوض، تنها دو عملیات وجود دارد، *enqueue* (به صف وارد کردن) و *dequeue* (از صف خارج کردن).



اگر دو آیتم را به لیست وارد کنید، اولین موردی که به لیست وارد کرده‌اید قبل از آیتم دوم قرار می‌گیرد. شما می‌توانید از این روش برای لیست جست‌وجوی خود استفاده کنید! افرادی که ابتدا به لیست اضافه می‌شوند اول از همه از صفحه خارج و جست‌وجو می‌شوند.

صف یک ساختمان داده‌ی FIFO نامیده می‌شود: First In, First Out. در مقابل، پشته یک ساختمان داده‌ی LIFO است: Last In, First Out.



اولین ورودی، اولین خروجی

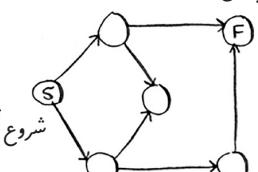
آخرین ورودی، اولین خروجی

اکنون که می‌دانید یک صف چگونه کار می‌کند، بباید جست‌وجوی سطح اول را اجرا کنیم!

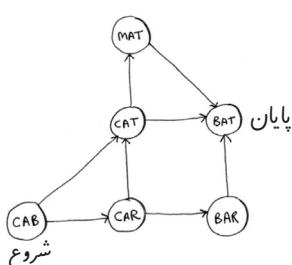
تمرین

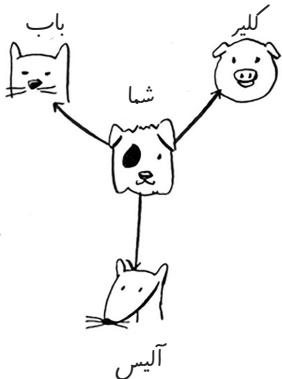
برای یافتن راه حل، الگوریتم جست‌وجوی سطح اول را روی هر یک از این گراف‌ها اجرا کنید.

۶.۱ طول کوتاه‌ترین مسیر را از ابتداء تا انتهای پیدا کنید.



۶.۲ طول کوتاه‌ترین مسیر را از «bat» تا «cab» بباید.

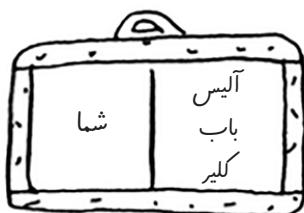




پیاده‌سازی گراف

ابتدا باید کد گراف را پیاده‌سازی کنید. یک گراف از چندین گره تشکیل شده است. و هر گره به گره‌های همسایه متصل است. چگونه رابطه‌ای مانند «you → bob» را نشان می‌دهید؟ خوشبختانه، شما پیش از این با ساختمان داده‌ای آشنا شده‌اید که امکان می‌دهد رابطه‌ها را نشان دهید: جدول هش!

در نظر داشته باشید، جدول هش به شما امکان می‌دهد یک کلید را به یک مقدار نگاشت کنید. در این مورد، شما می‌خواهید یک گره را به تمام همسایگانش نگاشت کنید.

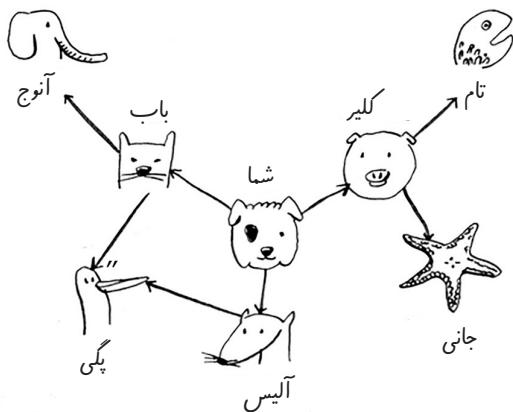


کد آن در پایتون به این صورت است:

```
graph = {}
graph["you"] = ["alice", "bob", "claire"]
```

توجه داشته باشید که «you» به یک آرایه نگاشت شده است. بنابراین آرایه‌ای از همه‌ی همسایگان «you» را در اختیار شما می‌گذارد.

یک گراف تنها دسته‌ای از گره و یال است، در نتیجه این تنها چیزی است که برای داشتن یک گراف در پایتون نیاز دارید. در مورد یک گراف بزرگ‌تر مانند این چطور؟



کد پایتون آن در ادامه آمده است:

```

graph = {}
graph["you"] = ["alice", "bob", "claire"]
graph["bob"] = ["anuj", "peggy"]
graph["alice"] = ["peggy"]
graph["claire"] = ["thom", "jonn"]
graph["anuj"] = []
graph["peggy"] = []
graph["thom"] = []
graph["jonna"] = []
  
```

پرسش: آیا مهم است که جفت کلید/مقدار را به چه ترتیبی اضافه کنید؟ مشکلی خواهد بود اگر به جای:

```

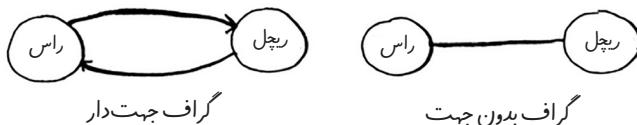
graph["anuj"] = []
graph["claire"] = ["thom", "jonna"]
  
```

بنویسیم:

```

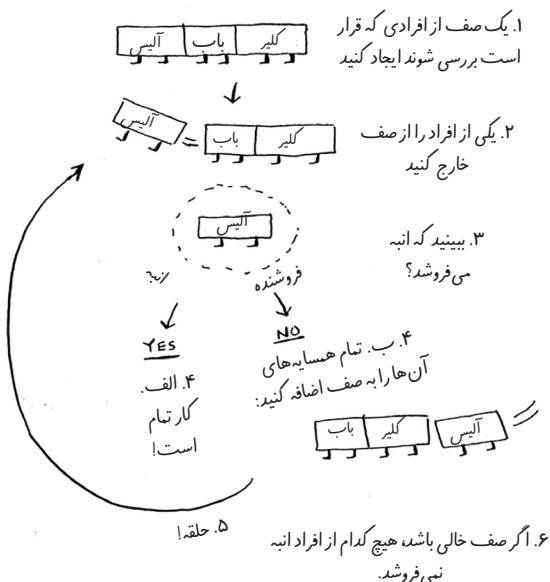
graph["claire"] = ["thom", "jonna"]
graph["anuj"] = []
  
```

فصل قبل را به خاطر بیاورید. پاسخ: تفاوتی ندارد. جدول‌های هش ترتیبی ندارند، بنابراین مهم نیست جفت کلید/مقدار را به چه ترتیبی اضافه می‌کنید. آنچه، پگی، تام و جانی هیچ همسایه‌ای ندارند. فلش‌هایی به آن‌ها اشاره می‌کند، اما هیچ فلشی از سمت آن‌ها به شخص دیگری نیست. این یک گراف جهت دار است. رابطه فقط یک طرفه است. بنابراین آنچه همسایه‌ی باب است، اما باب همسایه‌ی آنچه نیست. یک گراف بدون جهت هیچ فلشی ندارد و هر دو گره همسایه‌ی یکدیگر هستند. به عنوان مثال، هر دوی این گراف‌ها برابر هستند.



پیاده‌سازی الگوریتم

برای جمع‌بندی، در ادامه شیوه‌ی اجرا و پیاده‌سازی آن آمده است.



نکته

هنگام به روزرسانی صفحه‌ها، از اصطلاحات enqueue و dequeue استفاده می‌کنم. همچنین با اصطلاحات push و pop آشنا می‌شوید. Push تقریباً همیشه همان enqueue است و pop تقریباً همیشه همان (dequeue).

برای شروع یک صف درست کنید. در پایتون، از تابع صف دو طرفه^۱ (deque) برای این کار استفاده می‌کنید:

```
from collections import deque
search_queue = deque() ← یک صف جدید ایجاد می‌کنیم
search_queue += graph["you"] ← تمام همسایگان را به صف جست و جو اضافه می‌کنیم.
```



به یاد داشته باشید، `graph["you"]` لیستی از تمامی همسایه‌های شما، مانند `["alice", "bob", "claire"]` را نمایش می‌دهد. همه‌ی آن‌ها به صف جست و جو اضافه می‌شوند.

باقي کد به این صورت است:

```
while search_queue: ← در شرایطی که صف خالی نباشد
    person = search_queue.popleft() ← نفر اول را از صف جدا می‌کنند
    if person_is_seller(person): ← آیا این شخص انبه می‌فروشد؟
        print person + " is a mango seller!" ← بله، انبه می‌فروشد
        return True

    else:
        search_queue += graph[person] ← خیر، تمام دوستان این شخص را به صف جست و جو اضافه کنید
    return False ← اگر به این مرحله رسیدیم، هیچ کدام از افراد حاضر در صف انبه نمی‌فروشند.
```

نکته‌ی آخر: همچنان به یک تابع `person_is_seller` نیاز است تا هر زمان یک فروشنده‌ی انبه پیدا شد به شما اطلاع بدهد:

```
def person_is_seller(name):
    return name[-1] == 'm'
```

1. Double-ended queueue

این تابع بررسی می‌کند که آیا نام شخص به حرف m ختم می‌شود یا خیر. اگر چنین باشد، این شخص یک آنده فروش است. روشی هرچند احتمانه اما کارساز برای این مثال. بباید جستجوی سطح اول را در عمل ببینیم.

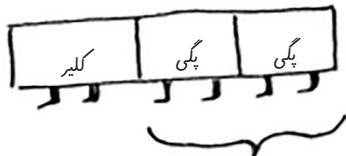


و به همین ترتیب. الگوریتم ادامه می‌یابد تا زمانی که:

- فروشنده‌ی آنده پیدا شود

- یا صفحه خالی شود که در این صورت فروشنده‌ای برای آنده وجود ندارد.

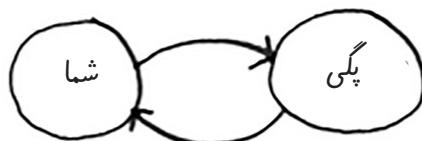
آلیس و باب یک دوست مشترک دارند: پگی. بنابراین پگی دو بار به صفت اضافه می‌شود: یک بار وقتی دوستان آلیس را اضافه می‌کنید و بار دیگر زمانی که دوستان باب را اضافه می‌کنید. در نهایت با دو پگی در صفت جست‌وجو مواجه خواهید شد.



عجب! پگی دونویت در
صف جست‌وجوهست!

اما تنها یک بار باید پگی را بررسی کنید تا ببینیم آن‌به می‌فروشد یا خیر. اگر دوبار بررسی کنید، کار غیر ضروری و اضافه‌ای انجام داده‌اید. بنابراین هنگامی که شخصی را برای جست‌وجو بررسی می‌کنید، باید به آن شخص تیک «بررسی شده» بزنید تا دوباره بررسی نشود.

اگر این کار را نکنید، ممکن است در یک حلقه‌ی بی‌نهایت قرار بگیرید. فرض کنید گراف آن‌به‌فروش به این شکل باشد:



برای شروع، صفت جست‌وجو تمام همسایه‌های شما را شامل می‌شود.



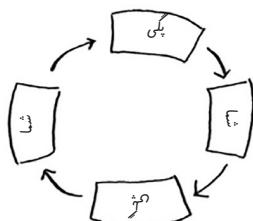
حالا شما پگی را بررسی می‌کنید. او فروشنده‌ی انبه نیست، بنابراین همه‌ی همسایگان او را به صف جستجو اضافه می‌کنید.



بعد، خودتان را بررسی کنید. شما انبه‌فروش نیستید، بنابراین همه‌ی همسایگان خود را به صف جستجو اضافه می‌کنید.



وغیره. این یک حلقه‌ی بی‌نهایت است، زیرا صف جستجو میان شما و پگی تداوم خواهد داشت.



قبل از بررسی یک شخص، مهم است که مطمئن بشوید قبلاً بررسی نشده است. برای انجام این کار، فهرستی از افرادی که قبلاً بررسی کرده‌اید رانگه دارید.

با در نظر گرفتن این نکته کد نهایی برای جستجوی سطح اول به این صورت است:

```
def search(name):
    search_queue = deque()
    search_queue += graph[name]
```

```

searched = []           ← آرایه‌ای برای نام افرادی که پیش از این جست‌وجو کرده‌اند.
while search_queue:    ← تباردار صورتی جست‌وجو می‌کند که نام فرد مورد نظر پیش از این جست‌وجو نشده باشد.
    person = search_queue.popleft()   ← این شخص را به عنوان جست‌وجو شده علامت‌گذاری می‌کند
    if not person in searched:      ← این شخص را به عنوان جست‌وجو شده علامت‌گذاری می‌کند
        if person_is_seller(person):
            print person + " is a mango seller!"
            return True

    else:
        search_queue += graph[person]
        searched.append(person)       ← این شخص را به عنوان جست‌وجو شده علامت‌گذاری می‌کند

return False
search("you")

```

کد را اجرا کنید. تابع `person_is_seller` را به چیزی با معنی تر تغییر بدهید و ببینید چیزی را که انتظار دارید پرینت می کند یا خیر.

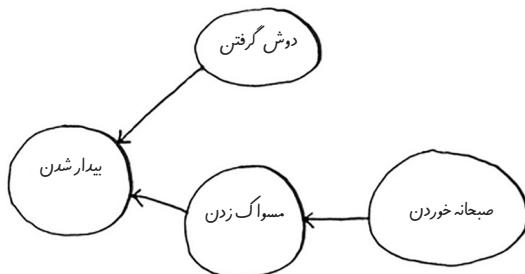
زمان اجرا

اگر در سراسر شبکه‌ی خود به دنبال فروشنده‌ی انبه بگردید، به این معنی است که تمامی یال‌ها را دنبال خواهید کرد (به یاد داشته باشید یک یال، فلش یا ارتباط یک فرد به فردی دیگر است). بنابراین زمان اجرای حداقل (تعداد یال‌ها) ۰ است.

همچنین یک صف از هر فرد برای جست‌وجو نگه می‌دارید. افزودن یک نفر به صف زمان ثابت دارد: $O(1)$. انجام این کار برای هر فرد، $(\text{تعداد افراد})O$ نیاز دارد. جست‌وجوی سطح اول (تعداد یال‌ها + تعداد افراد) O را می‌گیرد و معمولاً به صورت $O(V+E)$ تعداد رئوس و E تعداد یال‌ها نوشته می‌شود.

تمرین

گراف کوچکی از برنامه‌ی صبحگاهی من به این شکل است:



این به شما می‌گوید که تا زمانی که دندان‌ها یم را مسواک نزدم صبحانه نمی‌خورم.
بنابراین «صبحانه خوردن» به «مسواک زدن» وابسته است.

از طرف دیگر، دوش گرفتن به مسواک زدن دندان‌ها یم بستگی ندارد، زیرا می‌توانم قبل از مسواک زدن دوش بگیرم. از این گراف، می‌توانید فهرستی از ترتیب برنامه‌ی صبحگاهی من تهیه کنید:

۱. بیدار شدن.
۲. دوش گرفتن.
۳. مسواک زدن.
۴. صبحانه خوردن.

توجه داشته باشید که «دوش گرفتن» را می‌شود جایه‌جا کرد، بنابراین این لیست نیز درست است:

۱. بیدار شدن.
۲. مسواک زدن.
۳. دوش گرفتن.
۴. صبحانه خوردن.

۶.۳ درستی یا نادرستی هر یک از این سه لیست را مشخص کنید.

A.

۱. بیدارشدن
۲. دوش گرفتن
۳. صبحانه خوردن
۴. مسواک زدن

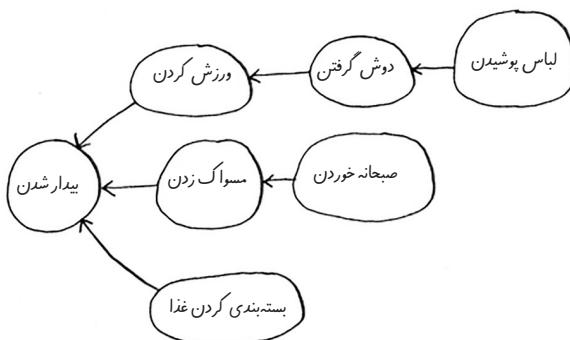
B.

۱. بیدارشدن
۲. مسواک زدن
۳. صبحانه خوردن
۴. دوش گرفتن

C.

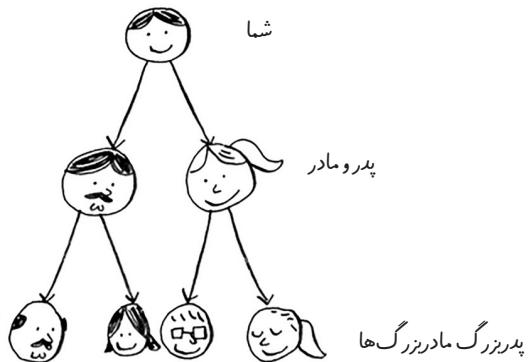
۱. دوش گرفتن
۲. بیدارشدن
۳. مسواک زدن
۴. صبحانه خوردن

۶.۴ در اینجا یک گراف بزرگ‌تر وجود دارد. یک لیست صحیح برای این گراف تهیه کنید.



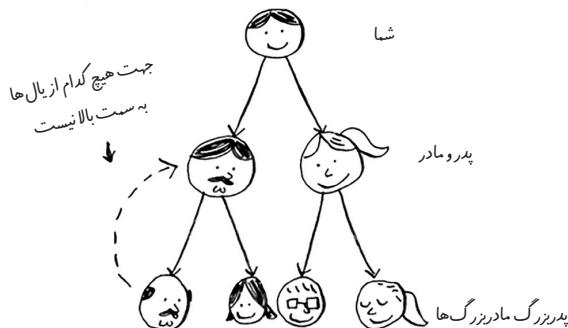
می‌توان گفت که این لیست مرتب شده است. اگر تسك A به تسك B وابسته باشد، تسك A بعدتر در لیست نشان داده می‌شود. به این مرتب‌سازی توپولوژیکی^۱ گفته می‌شود و راهی برای ایجاد لیست مرتب شده از یک گراف است. فرض کنید در تدارک یک عروسی هستید و یک گراف بزرگ سرشار از کارهایی که باید انجام شوند در اختیار دارید و نمی‌دانید از کجا شروع کنید. می‌توانید گراف را از نظر توپولوژیکی مرتب کنید و لیستی از کارهایی که باید انجام دهید را به ترتیب دریافت کنید.

فرض کنید شما یک شجره‌نامه دارید.



چون گره (افراد) و یال دارد در نتیجه این یک گراف است.

یال‌ها به والدین گره‌ها اشاره می‌کنند. اما همه‌ی یال‌ها به سمت پایین هستند - منطقی هم نیست در یک شجره‌نامه جهش یک یال به سمت بالا باشد! این بی‌معنی خواهد بود - پدر شما نمی‌تواند پدر پدربرگ شما باشد!



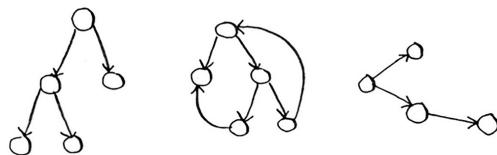
به این گراف درخت گفته می‌شود. درخت نوع خاصی از گراف است که در آن جهت هیچ یالی رو به عقب نیست.

۶.۵ کدام یک از تصاویری که در ادامه آمده علاوه بر گراف درخت هم هستند؟

A.

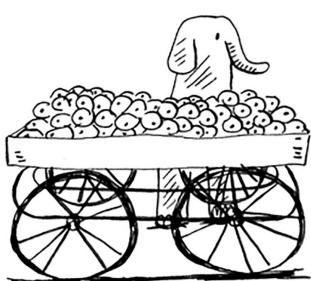
B.

C.



جمع‌بندی

- جست‌وجوی سطح اول به شما می‌گوید که آیا مسیری از A به B وجود دارد یا خیر.
- اگر مسیری وجود داشته باشد، جست‌وجوی سطح اول کوتاه‌ترین مسیر را پیدا می‌کند.
- اگر مشکلی مانند «یافتن کوتاه‌ترین X» دارید، سعی کنید مشکل خود را به صورت گراف مدل‌سازی کنید و از جست‌وجوی سطح اول برای حل آن استفاده کنید.
- یک گراف جهت‌دار دارای فلش است، و رابطه از جهت فلش پیروی می‌کند (\rightarrow adit) به معنای این است که «rama به adit بدهکار است»).
- گراف‌های غیر جهت‌دار فلش ندارند و رابطه به دو طرف پیش می‌رود (ross - rachel) به این معناست که «ričel با rass قرار می‌گذارد و rass با ričel قرار می‌گذارد».
- صفحات FIFO (اولین ورودی، اولین خروجی) هستند.
- پشت‌های LIFO (آخرین ورودی، اولین خروجی) هستند.
- باید افراد را به ترتیبی که به لیست جست‌وجو اضافه شده‌اند بررسی کنید، بنابراین لیست جست‌وجو باید یک صف باشد. در غیر این صورت کوتاه‌ترین مسیر را نخواهد داشت.
- هنگامی که فردی را بررسی می‌کنید، مطمئن بشوید که او را دوباره بررسی نکنید. در غیر این صورت، ممکن است در یک حلقه‌ی بی‌نهایت قرار بگیرد.





۷ | الگوریتم دایجسترا

در این فصل

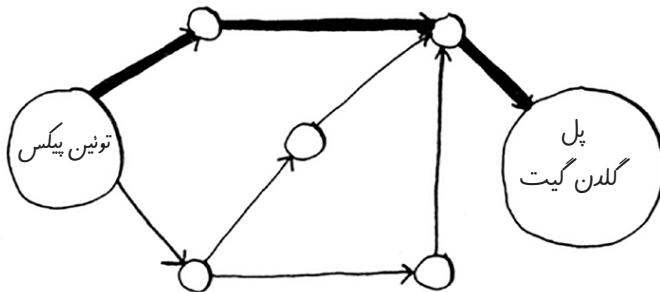
- مبحث گراف‌ها را ادامه می‌دهیم و شما با گراف‌های وزن‌دار آشنا می‌شوید؛ روشی برای اختصاص وزن کم‌تر یا بیشتر به برخی از یال‌ها.

- الگوریتم دایجسترا را یاد می‌گیرید. این الگوریتم به شما امکان می‌دهد برای گراف‌های وزن‌دار به پرسش «کوتاه‌ترین مسیر به X چیست؟» پاسخ دهید.

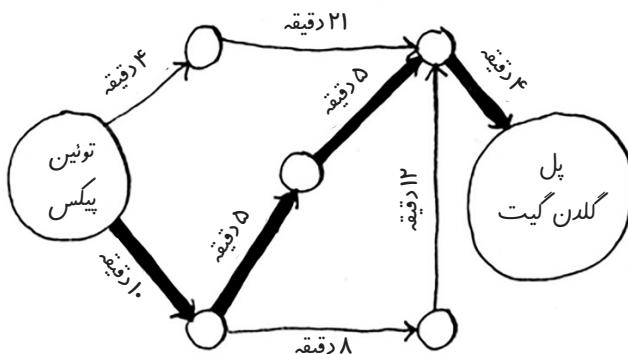
- در مورد چرخه‌های در گراف، جایی که الگوریتم دایجسترا جواب‌گو نیست، یاد می‌گیرید.

-
1. Dijkstra's algorithm
 2. weighted graphs
 3. cycles

در فصل قبل، مسیری برای رفتن از نقطه‌ی A به نقطه‌ی B پیدا کردید.



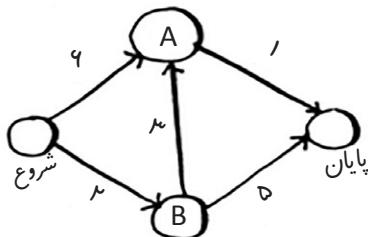
این لزوماً سریع‌ترین مسیر نیست بلکه کوتاه‌ترین مسیر است، زیرا کمترین تعداد قطعه (سه قطعه) را دارد. اما فرض کنید زمان سفر را به آن قطعه‌ها اضافه کنید. اکنون می‌بینید که مسیر سریع‌تری وجود دارد.



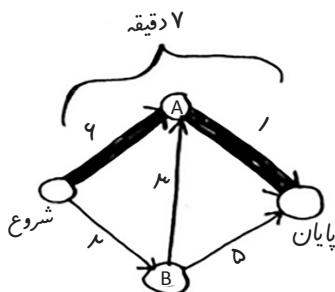
در فصل پیش از جست‌وجوی سطح اول استفاده کردید. جست‌وجوی سطح اول مسیری را که کمترین قطعه را دارد (گراف اول در اینجا نشان داده شده) پیدا می‌کند. اگر به جای آن سریع‌ترین مسیر را بخواهید (گراف دوم) چه می‌شود؟ شما می‌توانید این کار را به سریع‌ترین حالت با یک الگوریتم متفاوت به نام الگوریتم دایجسترا انجام بدھید.

کار با الگوریتم دایجسترا

عملکرد آن را با یک گراف بررسی کنیم:



هر قطعه یک زمان سفر بر حسب دقیقه دارد. شما از الگوریتم دایجسترا استفاده می‌کنید تا در کوتاه‌ترین زمان ممکن از نقطه‌ی شروع تا نقطه‌ی پایان بروید. اگر در این گراف جست‌وجوی سطح اول انجام دهید، کوتاه‌ترین مسیر را به این شکل خواهید داشت:



این مسیر ۷ دقیقه طول می‌کشد. بیایید ببینیم می‌توان مسیری یافت که زمان کمتری طول بکشد! چهار گام برای الگوریتم دایجسترا وجود دارد:

۱. «ارزان‌ترین» گره را پیدا کنید. این همان گره‌ای است که می‌توانید در کمترین زمان به آن برسید.
۲. هزینه‌های همسایگان این گره را به روز کنید. جلوتر توضیح می‌دهم منظورم

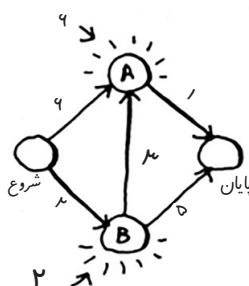
چیست.

۳. این کار را تا انجام آن برای تمامی گره‌های گراف تکرار کنید.

۴. مسیر نهایی را محاسبه کنید.

گام ۱: ارزان‌ترین گره را پیدا کنید. شما در نقطه‌ی شروع ایستاده‌اید و نمی‌دانید که آیا

باید به گره‌ی A بروید یا گره‌ی B. چقدر طول می‌کشد تا به هر گره برسید؟



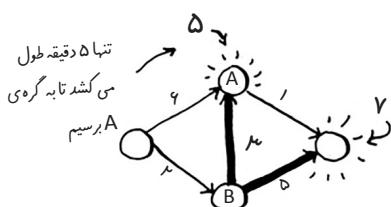
۶ دقیقه تا گره‌ی A و ۲ دقیقه برای رسیدن به گره‌ی B طول می‌کشد. در مورد بقیه‌ی گره‌ها، هنوز چیزی نمی‌دانید.

از آنجاکه هنوز نمی‌دانید چقدر طول می‌کشد تا به نقطه‌ی پایان برسید، بی‌نهایت را کنار می‌گذارد (به زودی دلیلش را می‌فهمید). گره‌ی B نزدیک‌ترین گره است ... ۲ دقیقه با آن فاصله دارد.

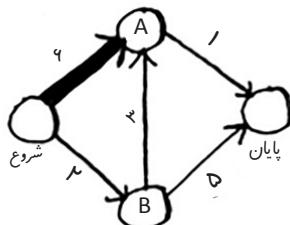
گره	زمان رسیدن به گره
A	۶
B	۲
پایان	∞

گام ۲: محاسبه کنید که چقدر طول می‌کشد تا از طریق یک یال از B به همه‌ی همسایگان گره‌ی B برسید.

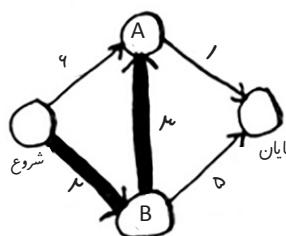
گره	زمان
A	۵
B	۲
پایان	۷



بسیار خب. شما یک مسیر کوتاه‌تر برای گرهی A پیدا کردید! قبل‌تر ۶ دقیقه طول می‌کشید تا به گرهی A برسید.



اما اگر از گرهی B عبور کنید، مسیری وجود دارد که فقط ۵ دقیقه طول می‌کشد!



وقتی مسیر کوتاه‌تری برای همسایه‌ی B پیدا کردید، هزینه‌ی آن را به روز کنید. در این مورد، یافته‌ها عبارتند از:

- یک مسیر کوتاه‌تر به A (از ۶ دقیقه به ۵ دقیقه کاهش می‌یابد)

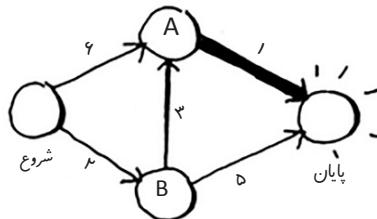
- یک مسیر کوتاه‌تر تا پایان (از بی‌نهایت به ۷ دقیقه کاهش می‌یابد)

گام ۳: تکرار کنید!

گره	زمان
A	۵
B	۲
پایان	۷

دوباره گام ۱: گره‌ای را پیدا کنید که کم‌ترین زمان را برای رسیدن نیاز دارد. کار با گرهی B تمام شده است، بنابراین گرهی A کوچک‌ترین تخمین زمان بعدی را شامل می‌شود.

دوباره گام ۲: هزینه‌ها را برای همسایگان گرهی A به روز کنید.



چه خوب، اکنون ۶ دقیقه طول می‌کشد تا به نقطه‌ی پایان برسیم!
شما الگوریتم دایجسترا را برای هر گره اجرا کرده‌اید (نیازی به اجرای آن برای گرهی پایانی ندارید). در این مرحله، شما می‌دانید

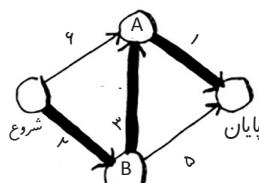
۲۰ دقیقه طول می‌کشد تا به گرهی B برسید.

۵ دقیقه طول می‌کشد تا به گرهی A برسید.

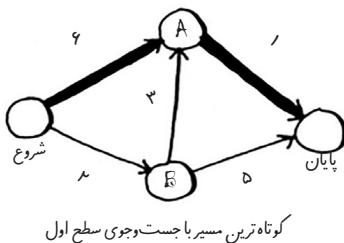
۶ دقیقه طول می‌کشد تا به پایان برسید.

گره	زمان
A	۵
B	۲
پایان	۶

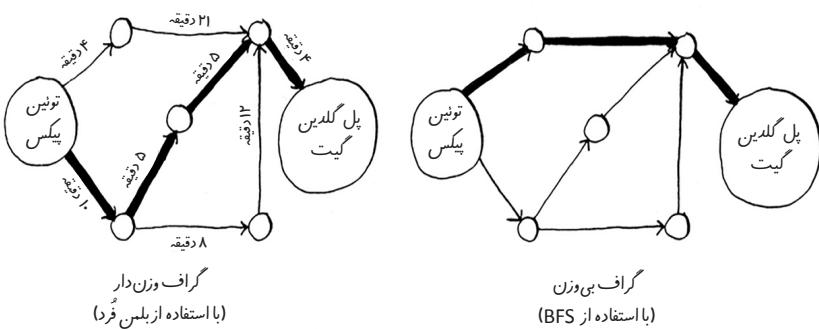
آخرین مرحله یعنی محاسبه‌ی مسیر نهایی را برای بخش بعدی نگه می‌دارم. فعلاً فقط نشان می‌دهم که مسیر نهایی چیست.



جستجوی سطح اول این را به عنوان کوتاه‌ترین مسیر پیدا نمی‌کند، چون سه قطعه است. و راهی دو قطعه‌ای برای شروع تا پایان وجود دارد.



در فصل پیش، شما از جستجوی سطح اول برای یافتن کوتاه‌ترین مسیر بین دو نقطه استفاده کردید. در آن زمان، «کوتاه‌ترین مسیر» به معنای مسیری با کمترین قطعات بود. اما در الگوریتم دایجسترا، شما به هر بخش یک عدد یا وزن اختصاص می‌دهید. سپس الگوریتم دایجسترا مسیری با کمترین وزن کل را پیدا می‌کند.



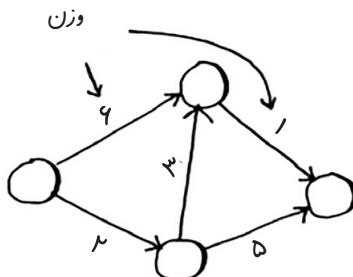
به طور خلاصه، الگوریتم دایجسترا دارای چهار گام است:

۱. ارزان‌ترین گره را پیدا کنید. این گره‌ای است که می‌توانید در کوتاه‌ترین زمان به آن برسید.
۲. بررسی کنید که آیا مسیر ارزان‌تری به همسایگان این گره وجود دارد یا خیر. اگر چنین است، هزینه‌های آن‌ها را به روز کنید.

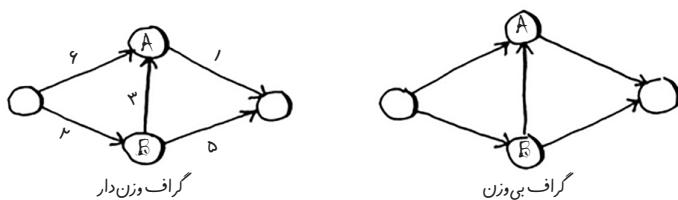
۳. این روند را تکرار کنید تا زمانی که این کار را برای هر گره در گراف انجام داده باشد.
۴. مسیر نهایی را محاسبه کنید. (در بخش بعدی آمده است!)

اصطلاح‌شناسی

قرار است چند مثال دیگر از الگوریتم دایجسترا را به شکل عملی به شما نشان بدهم. اما ابتدا اجازه دهید به تشریح چند اصطلاح بپردازم. هنگامی که با الگوریتم دایجسترا کار می‌کنید، هر یال در گراف همراه با یک عدد است. به این اعداد وزن^۱ گفته می‌شود.

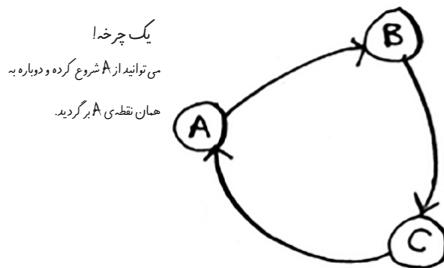


به گراف دارای وزن، گراف وزن دار^۲ می‌گویند. گراف بدون وزن را گراف بی وزن^۳ می‌نامند.

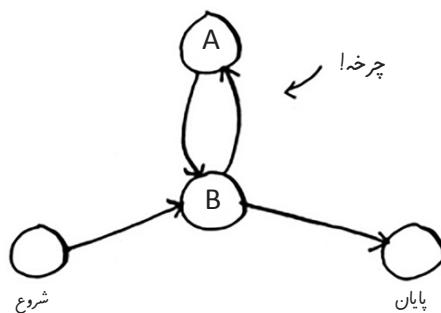


برای محاسبه‌ی کوتاه‌ترین مسیر در یک گراف بی وزن، از جست‌وجوی سطح اول استفاده کنید. برای محاسبه‌ی کوتاه‌ترین مسیر در یک گراف وزن دار، از الگوریتم دایجسترا استفاده کنید. گراف‌ها همچنین می‌توانند چرخه داشته باشند. یک چرخه به این شکل است.

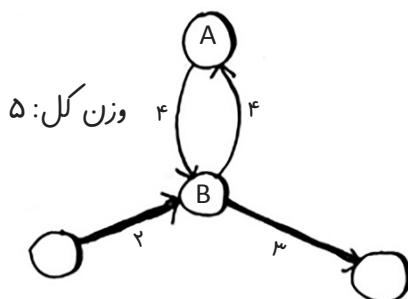
1. weights
2. Weighted graph
3. Unweighted graph



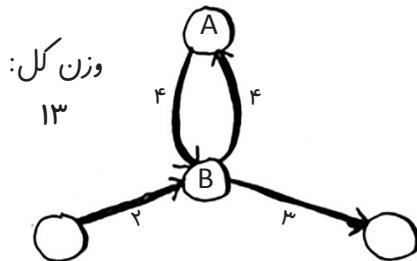
به این معنی که می توانید از یک گره شروع کرده، به اطراف حرکت کنید و در نهایت به همان گره برسید. فرض کنید می خواهید کوتاهترین مسیر در این گراف را که دارای یک چرخه است پیدا کنید.



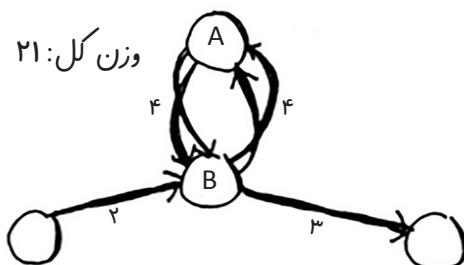
دنبال کردن چرخه منطقی است؟ خوب، شما می توانید از مسیری استفاده کنید که از چرخه اجتناب می کند.



یا می‌توانید چرخه را طی کنید.

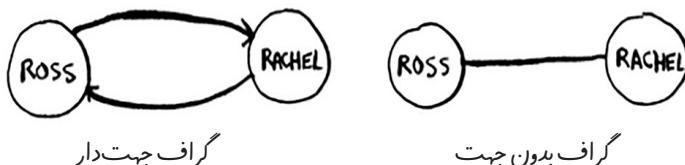


در هر صورت به گرهی A ختم می‌شود، اما این چرخه وزن بیشتری دارد. حتی اگر بخواهید می‌توانید چرخه را دو بار طی کنید.

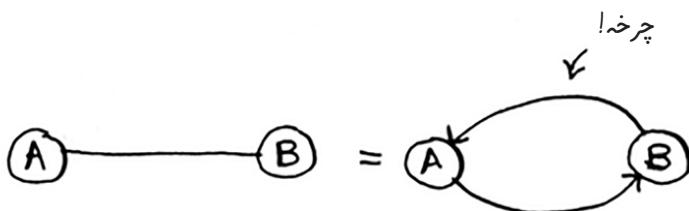


اما هر بار که چرخه را طی می‌کنید، مقداری معادل ۸ به وزن کل اضافه می‌کنید. در نتیجه با طی کردن چرخه هرگز به کوتاه‌ترین مسیر دست پیدا نمی‌کنید.

در آخر، به خاطر دارید در فصل ۶ درباره‌ی گراف‌های جهت‌دار و غیر جهت‌دار چه گفتیم؟



گراف بدون جهت به این معنی است که هر دو گره به یکدیگر اشاره می‌کنند. در واقع همان چرخه است!



با یک گراف بدون جهت، هر یال چرخه‌ی دیگری اضافه می‌کند. الگوریتم دایجسترا فقط با گراف‌های جهت‌دار غیر مدور^۱ که به اختصار DAG نامیده می‌شوند کار می‌کند.

معامله بر سر یک پیانو



صحبت از اصطلاحات کافی است، بباید به مثال دیگری نگاه کنیم! این تصویر راما است. راما در تلاش است تا یک کتاب موسیقی را با پیانو عوض کند.

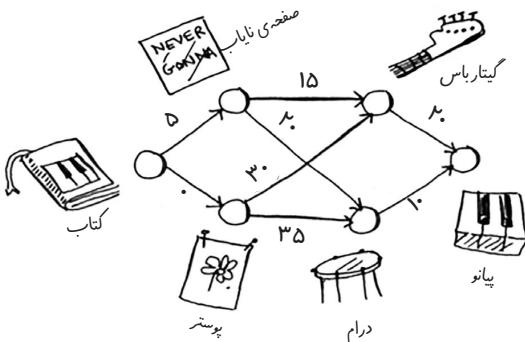
الکس می‌گوید: «این پوستر را با کتابت تاخت می‌زنم. این پوستر گروه محبوبم دیسترویره. کتابت رو با این ال پی کمیاب ریک استلی عوض می‌کنم و پنج دلار هم میدارم روش.» امی می‌گوید: «او، شنیده‌ام که صفحه‌ی آهنگ واقعاً محشری دارد. من گیتار یا درام را با پوستر یا صفحه‌ی تو عوض می‌کنم.»



بتههون فریاد می‌زنند: «گیتارو من می‌خواستم. پیانو رو در عوض یکی از چیزهای امی به تو می‌دهم.»

بسیار عالی! راما با کمی پول می‌تواند یک کتاب پیانو را به یک پیانوی واقعی تبدیل کند. فقط باید بفهمد که چگونه برای انجام این معامله‌ها کمترین پول را خرج کند. بباید آنچه را که به او پیشنهاد شده است، به گراف تبدیل کنیم.

1. Directed acyclic graph



در این گراف، گره‌ها تمام آیتم‌هایی را شامل می‌شوند که راما می‌تواند سرآن‌ها معامله کند. وزن روی یال‌ها مبلغی است که او باید برای انجام معامله بپردازد. بنابراین می‌تواند پوستر را با گیتار به قیمت ۳۰ دلار عوض کند، یا صفحه را با گیتار با ۱۵ دلار عوضه کند. راما چگونه می‌تواند مسیر کتاب تا پیانو را که کمترین پول را در آن خرج می‌کند، بفهمد؟ الگوریتم دایجسترا نجات‌بخش است! به یاد داشته باشید، الگوریتم دایجسترا چهار مرحله دارد. در این مثال، شما هر چهار مرحله را انجام خواهید داد، بنابراین مسیر نهایی را در پایان نیز محاسبه خواهید کرد.

گره	هزینه
صفحه	۵
پوستر	.
گیتار	∞
درام	∞
پیانو	∞

} هنوز به این گره‌ها رسیده‌ایم

ابتدا نیاز به مقداری مقدمه‌چینی است.

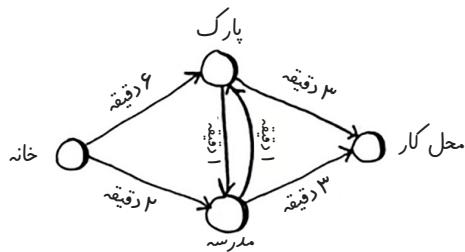
جدولی از هزینه‌ی هر گره تهیه کنید. هزینه‌ی یک گره به معنای آن است که رسیدن به آن چقدر هزینه‌بردار است.

گره	والد
صفحه	کتاب
پوستر	کتاب
گیتار	—
درام	—
پیانو	—

با ادامه الگوریتم، این جدول را به تناوب به روز خواهید کرد. برای محاسبه‌ی مسیر نهایی، همچنین به یک ستون والد در این جدول نیاز دارید.

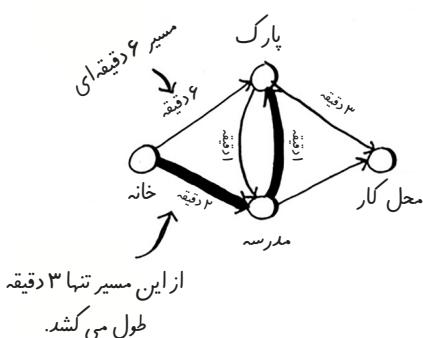
به زودی نحوه‌ی کار این ستون را به شما نشان خواهم داد. بیایید الگوریتم آن را بنویسیم.

گام ۱: ارزان‌ترین گره را پیدا کنید. در این مورد، پوستر ارزان‌ترین معامله است، صفر دلار. آیا راه ارزان‌تری برای معامله‌ی پوستر وجود دارد؟ این نکته‌ی بسیار مهمی است، پس راجع به آن فکر کنید. آیا می‌توانید یک سری معاملات را ببینید که پوستر راما را با کمتر از صفر دلار دریافت می‌کند؟ اگر به جوابی در ذهن خود رسیدید، ادامه‌ی مطلب را بخوانید. پاسخ: خیر. چون پوستر ارزان‌ترین گره‌ای است که راما می‌تواند به آن دسترسی پیدا کند، راهی برای ارزان‌تر کردن آن وجود ندارد. در اینجا یک راه متفاوت برای نگاه کردن به آن وجود دارد. فرض کنید از خانه به محل کار سفر می‌کنید.



اگر مسیر را به سمت مدرسه طی کنید، ۲ دقیقه طول می‌کشد. اگر مسیر را به سمت پارک طی کنید، ۴ دقیقه طول می‌کشد. آیا راهی وجود دارد که بتوانید مسیر پارک را طی کنید و در کمتر از ۲ دقیقه به مدرسه برسید؟ غیرممکن است، زیرا فقط تا خود پارک

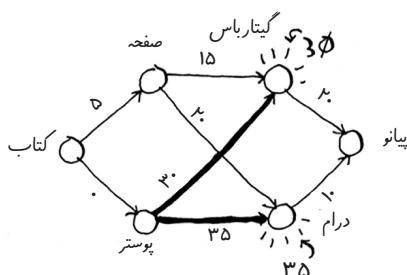
بیشتر از ۲ دقیقه طول می‌کشد. از طرف دیگر آیا می‌توانید مسیر سریع‌تری به پارک پیدا کنید؟ بله.



این ایده‌ی کلیدی الگوریتم دایجسṭra است: به ارزان‌ترین گره در گراف خود نگاه کنید. هیچ راه ارزان‌تری برای رسیدن به این گره وجود ندارد!

به مثال موسیقی برگردیم. پوستر ارزان‌ترین معامله است.

گام ۲: مشخص کنید که چقدر طول می‌کشد تا به همسایگانش برسید (هزینه).



هزینه	گره	والد	صفحه	کتاب
5				کتاب
0				پوستر
30				گیتار
35				درام
50				پیانو

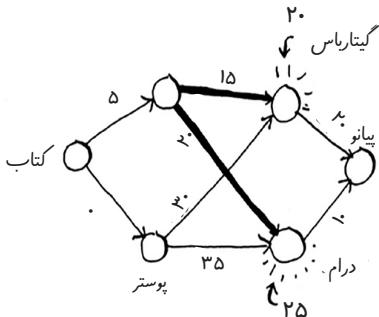
قیمت گیتار باس و درام را در جدول دارید. وقتی به سراغ پوستر رفتید ارزش آن‌ها مشخص شده بود، بنابراین پوستر به عنوان والد آن‌ها در نظر گرفته می‌شود. یعنی برای رسیدن به گیتار بیس، یال پوستر را دنبال می‌کنید، همین را برای درام‌ها هم می‌شود گفت.

هزینه	گره	والد	صفحه	کتاب
5				کتاب
0				پوستر
30				گیتار
35				درام
∞				پیانو

از «پوستر» که شروع کنید به این گره‌ها می‌رسیم {

دوباره گام ۱: صفحه ارزان‌ترین گرهی بعدی با قیمت ۵ دلار است.

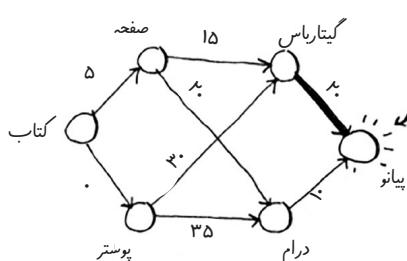
دوباره گام ۲: مقادیر همه‌ی همسایگانش را به روز کنید.



کتاب	صفحه	۵
کتاب	پوستر	۰
صفحه	گیتار	۲۰
صفحه	درام	۲۵
—	پیانو	∞

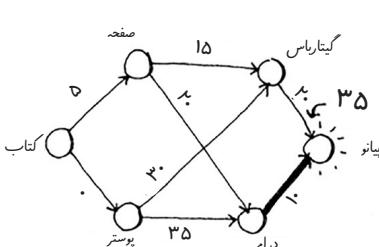
خب، قیمت درام و گیتار را به روز کردی! این به این معناست که با دنبال کردن یال های صفحه، رسیدن به درام و گیتار ارزان تر است. بنابراین صفحه را به عنوان والد جدید برای هر دو ساز در نظر می گیرید.

گیتار باس ارزان ترین کالای بعدی است. همسایگان آن را به روز کنید.



کتاب	صفحه	۵
کتاب	پوستر	۰
صفحه	گیتار	۲۰
صفحه	درام	۲۵
گیتار	پیانو	۴۰

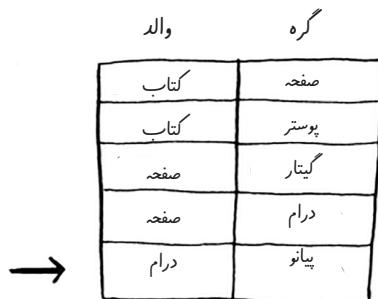
خوب، بالاخره با مبادله ی گیتار با پیانو، به یک قیمت برای پیانو دست پیدا کردیم. بنابراین شما گیتار را به عنوان والد در نظر می گیرید. در نهایت، آخرین گره، درام است.



کتاب	صفحه	۵
کتاب	پوستر	۰
صفحه	گیتار	۲۰
صفحه	درام	۲۵
درام	پیانو	۳۵

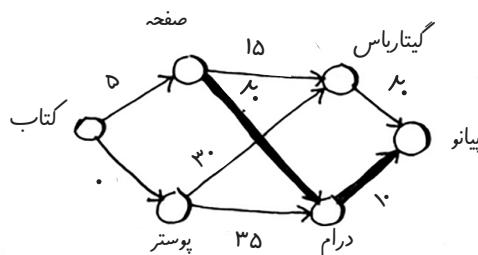
راما می تواند با معامله ی درام با پیانو، پیانو را حتی ارزان تر دریافت کند. بنابراین ارزان ترین مجموعه معاملات ۳۵ دلار برای راما هزینه در بر دارد.

حالا، همان طور که قول داده بودم، باید مسیر را مشخص کنید. تا اینجا، می‌دانید که کوتاهترین مسیر ۳۵ دلار هزینه دارد، اما چگونه می‌توانید مسیر را مشخص کنید؟ برای شروع، به والد پیانو نگاه کنید.

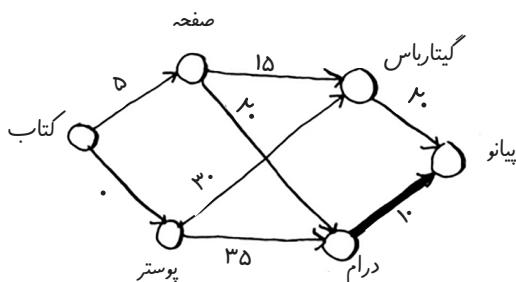


درام والد پیانو است. یعنی راما درام‌ها را با پیانو عوض می‌کند. پس شما این یال را دنبال می‌کنید.

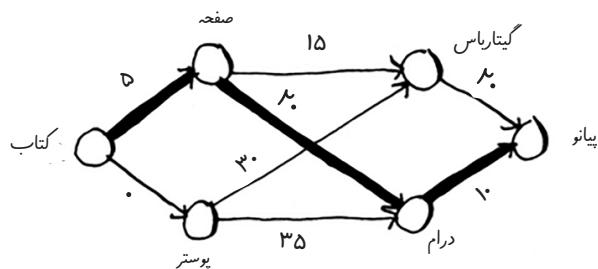
باید ببینیم یال‌ها را چگونه دنبال می‌کنید. پیانو درام را به عنوان والد خود دارد.



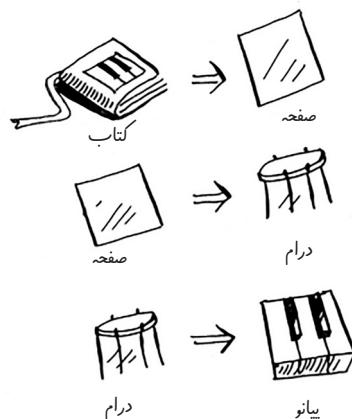
و صفحه والد درام‌ها است.



بنابراین راما صفحه را با درام‌ها معاوضه می‌کند. و البته، این کتاب را با صفحه عوض خواهد کرد. با عقب‌گرد از سمت والدین، اکنون مسیر کامل را در اختیار دارید.



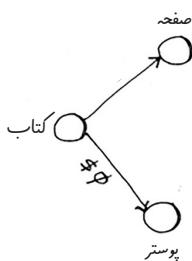
مجموعه معاملاتی که راما باید انجام بدهد در زیر آمده است.



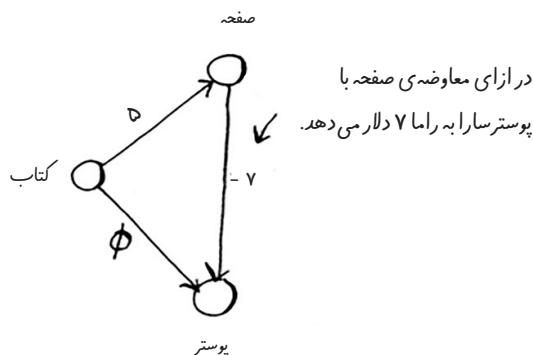
تا اینجا، از عبارت کوتاهترین مسیر به معنای واقعی کلمه‌ی آن استفاده کرده‌ام: محاسبه‌ی کوتاهترین مسیر بین دو مکان یا بین دو نفر. امیدوارم این مثال به شما نشان داده باشد که کوتاهترین مسیر نباید درباره‌ی فاصله‌ی فیزیکی باشد. این می‌تواند در مورد به حداقل رساندن^۱ چیزی باشد. راما در این مورد می‌خواست مقدار پولی را که خرج می‌کند به حداقل برساند. با تشکر از دایجسترا!

یال‌های با وزن منفی^۱

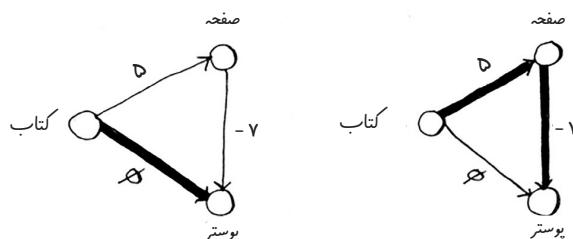
در این مثال، الکس پیشنهاد داد که کتاب را با دو آیتم معامله کند.



فرض کنید سارا پیشنهاد می‌دهد که صفحه را با پوستر معوضه کند، و او به راما ۷ دلار دیگر می‌دهد. انجام این معامله برای راما هیچ هزینه‌ای ندارد. در عوض، او ۷ دلار به دست می‌آورد. در گراف چگونه باید نشان داد؟

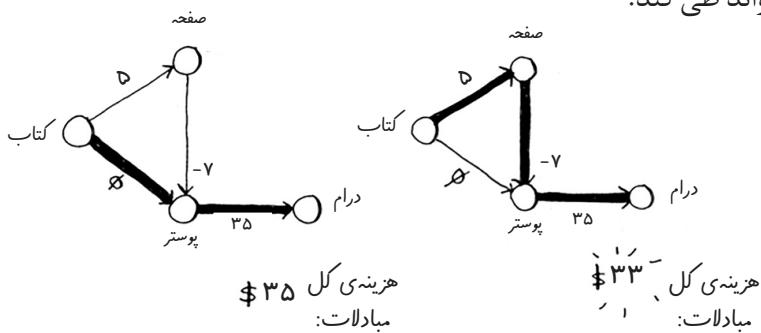


یال میان صفحه و پوستر وزن منفی دارد! اگر راما این معامله را انجام بددهد، ۷ دلار دریافت می‌کند. حالا راما دو راه برای رسیدن به پوستر دارد.



از این مسیر راما دلار به دست نمی‌آورد.

بنابراین انجام معامله‌ی دوم منطقی است - راما از این طریق ۲ دلار پس می‌گیرد! حالا اگر یادتان باشد راما می‌تواند پوستر را با طبل عوض کند. دو مسیر هست که می‌تواند طی کند.

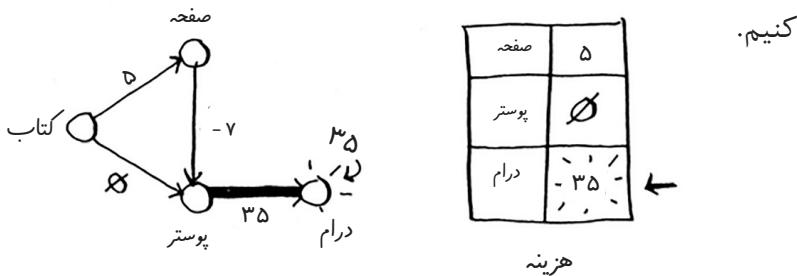


مسیر دوم برای او ۲ دلار کمتر هزینه دارد، بنابراین او باید آن مسیر را طی کند، درست است؟ خب حدس بزنید چه اتفاقی افتاده؟ اگر الگوریتم دایجسترا را روی این گراف اجرا کنید، راما مسیر اشتباهی را در پیش خواهد گرفت. مسیرش طولانی تر می‌شود. اگر یال‌هایی با وزن منفی دارید، نمی‌توانید از الگوریتم دایجسترا استفاده کنید. یال‌های با وزن منفی الگوریتم را خراب می‌کنند. بیایید ببینیم وقتی الگوریتم دایجسترا را در این مورد اجرا می‌کنید چه اتفاقی می‌افتد. ابتدا جدولی از هزینه‌ها ایجاد کنید.

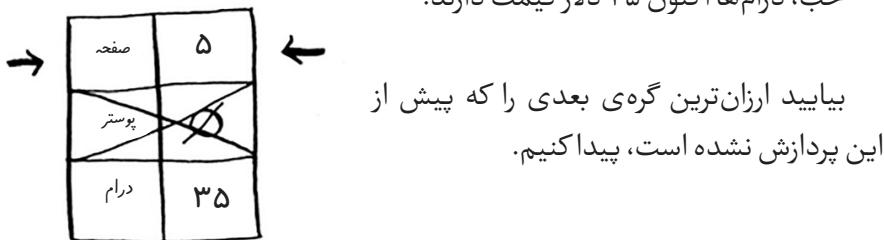
صفحه	۵
پوستر	.
درام	∞

هزینه

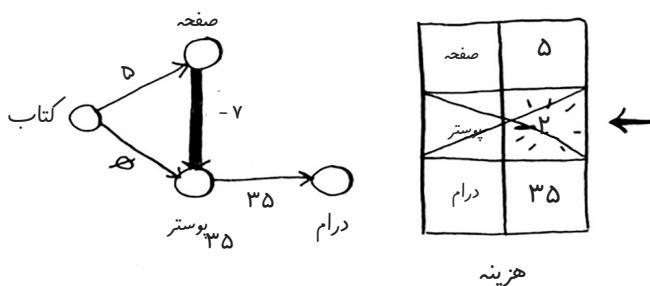
در مرحله‌ی بعد، گرهی کم هزینه را پیدا کنید و هزینه‌ها را برای همسایگانش به روز کنید. در این مورد، پوستر کم هزینه‌ترین گره است. بنابراین، طبق الگوریتم دایجسترا، هیچ راه ارزان‌تری از پرداخت صفر دلار برای رسیدن به پوستر وجود ندارد (می‌دانید که چنین چیزی اشتباه است!). به هر حال، بیایید هزینه‌ها را برای همسایگانش به روز



خب، درام‌ها اکنون ۳۵ دلار قیمت دارند.



هزینه‌ها را برای همسایگان خود به روز کنید.



على رغم اینکه پیش از این گرهی پوستر را پردازش کرده‌اید، هزینه‌ی آن را به روز می‌کنید. این یک هشدار جدی است. هنگامی که یک گره را پردازش می‌کنید، به این معنی است که هیچ راه ارزان‌تری برای رسیدن به آن گره وجود ندارد. با این حال

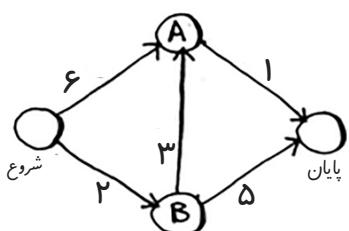
الان شما یک راه ارزان‌تر برای پوستر پیدا کردید! درام‌ها هیچ همسایه‌ای ندارند، بنابراین این نقطه پایان الگوریتم است. در تصویر زیر هزینه‌های نهایی آمده است.

صفحه	۵
پوستر	-۲
درام	۳۵

هزینه‌ی نهایی

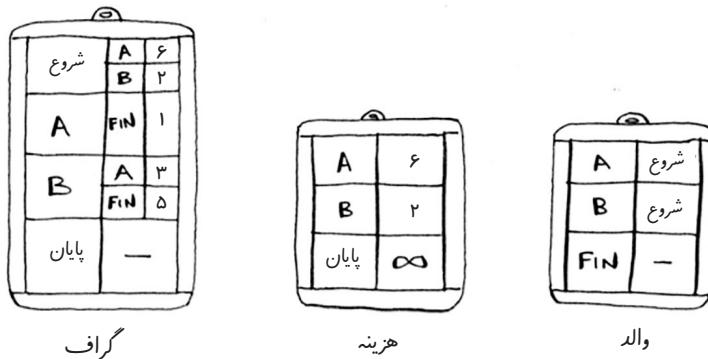
درام‌ها ۳۵ دلار هستند. مسیری را می‌شناشید که فقط ۳۳ دلار هزینه دارد، اما الگوریتم دایجسترا آن را پیدا نکرد. الگوریتم دایجسترا فرض می‌کند که چون شما گرهی پوستر را پردازش می‌کنید، راه سریع‌تری برای رسیدن به آن گره وجود ندارد. این فرض تنها در صورتی کارساز است که یال‌هایی با وزن منفی نداشته باشید. بنابراین نمی‌توانید از یال‌های وزن منفی با الگوریتم دایجسترا استفاده کنید. اگر می‌خواهید کوتاه‌ترین مسیر را در گرافی پیدا کنید که دارای یال‌هایی با وزن منفی است، چنین الگوریتمی برای آن وجود دارد! نام این الگوریتم بلمن-فورد^۱ نام دارد. الگوریتم بلمن-فورد خارج از محدوده‌ی این کتاب است، اما می‌توانید توضیحات خیلی خوبی درباره‌ی این الگوریتم در اینترنت پیدا کنید.

پیاده‌سازی



بیایید ببینیم که چگونه الگوریتم دایجسترا را با کد پیاده‌سازی کنیم. از این گراف برای این مثال استفاده خواهم کرد.

برای کدنویسی این مثال، به سه جدول هش نیاز دارید.



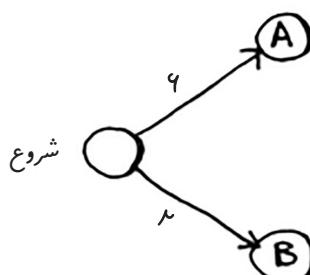
با پیشرفت الگوریتم، هزینه‌ها و جدول‌های والدها را به روز می‌کنید. ابتدا باید گراف را پیاده‌سازی کنید. از جدول هش مانند آن‌چه در فصل ۶ انجام دادید استفاده می‌کنید:

```
graph = {}
```

در فصل پیش، تمام همسایگان یک گره را مانند کد زیر در جدول هش ذخیره کردید:

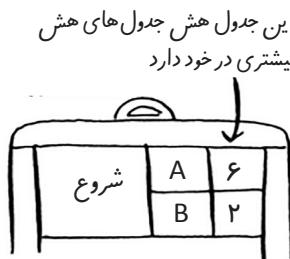
```
graph["you"] = ["alice", "bob", "claire"]
```

اما این بار باید همسایه‌ها و هزینه‌ی رسیدن به آن همسایه را ذخیره کنید. برای مثال، دو همسایه دارد: A و B و Start.



وزن آن یال‌ها را چگونه نشان می‌دهید؟ چرا صرفاً از یک جدول هش دیگر استفاده نکنید؟

```
graph["start"]={}
graph["start"]["a"] = 6
graph["start"]["b"] = 2
```



بنابراین `graph["start"]` یک جدول هش است. می‌توانید همه‌ی همسایگان را به این شکل دریافت کنید:

```
>>> print graph["start"].keys()
["a", "b"]
```

یک یال از Start به A و یک یال از Start به B وجود دارد. اگر بخواهید وزن آن یال‌ها را بیابید چه؟

```
>>> print graph["start"]["a"]
2
>>> print graph["start"]["b"]
6
```

بیاباید گره‌های دیگر و همسایگان آن‌ها را به گراف اضافه کنیم:

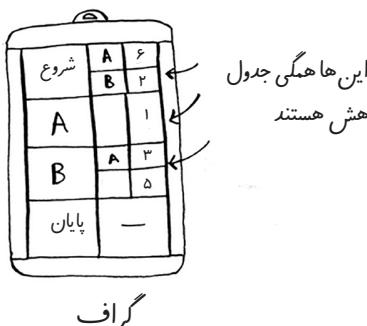
```
graph["a"] = {}
```

```

graph["a"]["fin"] = 1
graph["b"] = {}
graph["b"]["a"] = 3
graph["b"]["fin"] = 5
graph["fin"] = {} ← گرهی پایان هیچ حسابی ای ندارد

```

جدول هش گراف کامل به این شکل است.



در مرحله‌ی بعد به یک جدول هش برای ذخیره‌ی هزینه‌های هر گره نیاز دارد.

A	6
B	2
پایان	∞

هزینه

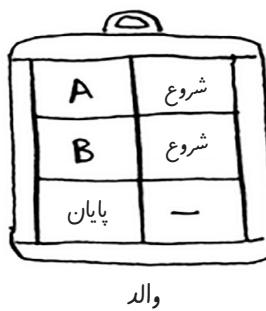
هزینه‌ی یک گره مدت زمانی است که طول می‌کشد تا از نقطه‌ی شروع به آن گره برسید. شما می‌دانید که از شروع تا گره‌ی ۲ دقیقه و تا گره‌ی A ۶ دقیقه طول می‌کشد. (اگرچه ممکن است مسیری را پیدا کنید که زمان کمتری می‌برد). از زمان مورد نیاز برای رسیدن به نقطه‌ی پایان اطلاعی ندارید. اگر هنوز هزینه را نمی‌دانید، بی‌نهایت را کنار می‌گذارید. آیا می‌توانید بی‌نهایت را به زبان پایتون بنویسید؟ بله می‌توانید:

```
infinity = float("inf")
```

در اینجا کدی برای ایجاد جدول هزینه‌ها آمده است:

```
infinity = float("inf")
costs = {}
costs["a"] = 6
costs["b"] = 2
costs["fin"] = infinity
```

همچنین به جدول هش دیگری برای والدها نیاز دارد:



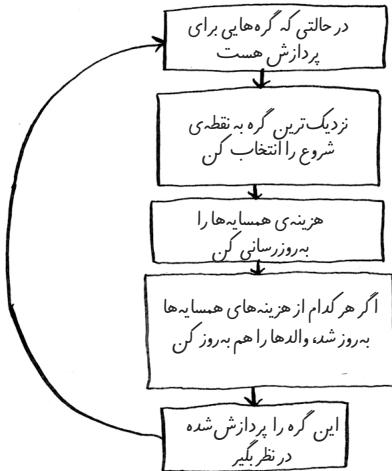
در اینجا کد برای ایجاد جدول هش والدها آمده است:

```
parents = {}
parents["a"] = "start"
parents["b"] = "start"
parents["fin"] = None
```

در نهایت، شما به یک آرایه نیاز دارید تا تمام گره‌هایی را که قبل‌پردازش کرده‌اید را در نظر داشته باشد، چراکه نیازی به پردازش بیش از یک‌بار یک گره ندارید:

```
processed = []
```

تمام تنظیمات همین است. حالا باید به الگوریتم نگاه کنیم.



ابتدا کد را به شما نشان می‌دهم و سپس آن را مرور می‌کنم.

```

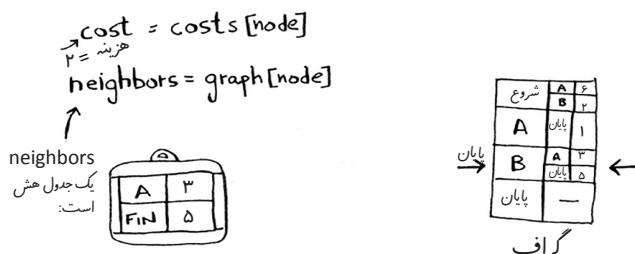
node = find_lowest_cost_node(costs) ← گره‌ای با کم‌ترین هزینه که پردازش نشده را پیدا می‌کند
while node is not None: ← اگر تمام گره‌ها پردازش شده باشند، کار این حلقه به پایان می‌رسد.
    cost = costs[node]
    neighbors = graph[node]
    for n in neighbors.keys(): ← از تمام همسایه‌های این گره عبور می‌کند.
        new_cost = cost + neighbors[n] ← اگر رسیدن به این همسایه کم‌هزینه‌تر است
        if costs[n] > new_cost: ← با انتخاب این گره...
            costs[n] = new_cost ← هزینه این گره بروز می‌شود.
            parents[n] = node ← این گره والد جدید برای این همسایه می‌شود.
            processed.append(node) ← گره به عنوان پردازش شده علامت‌گذاری می‌شود.
            node = find_lowest_cost_node(costs) ← یال بعدی را برای پردازش و... ایجاد حلقه پیدا می‌کند.
  
```

این همان الگوریتم دایجسترا در پایتون است! من کد تابع را بعداً به شما نشان خواهم داد. ابتدا، بیایید کد الگوریتم `find_lowest_cost_node(costs)` را در عمل ببینیم.

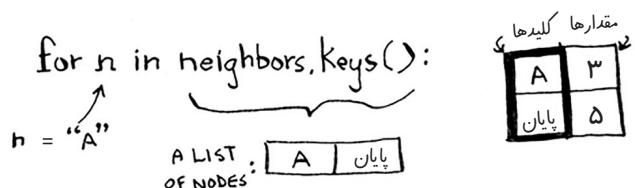
اپن گرہ را با کمترین هزینہ پیدا کنید.



هزینه و همسایه‌های آن گره را پیدا کنید.



د، همسایه‌ها حلقه نند.



هر گره هزینه‌ای دارد. هزینه به این معنی است که از ابتدا چقدر طول می‌کشد تا به آن گره برسید. در اینجا محاسبه می‌کنید که چقدر طول می‌کشد تا به گرهی A برسید به شرط آن‌که مسیر شما به جای Start > node A باشد:

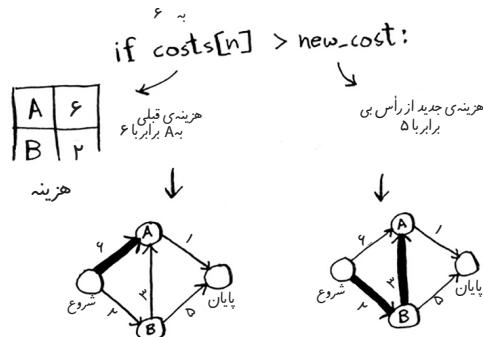
Start > node B > node A

$$\text{new_cost} = \text{cost} + \text{neighbors}[n]$$

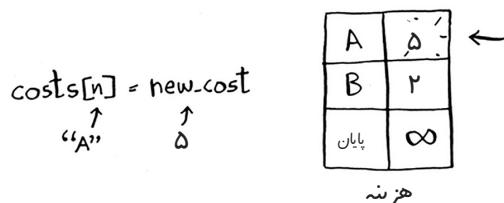
↑ ↓
 COST OF DISTANCE FROM
 "B", i.e. γ B TO A: γ'

$\left. \begin{array}{l} \text{new_cost} = \gamma + \gamma' \\ = \Delta \end{array} \right\}$

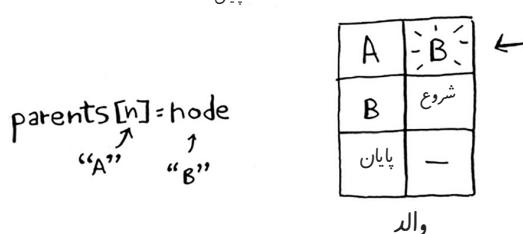
بیایید هزینه‌ها را با هم مقایسه کنیم.



مسیر کوتاه‌تری برای گرهی A پیدا کردید! هزینه را به روز کنید.



مسیر جدید از گرهی B می‌گذرد، بنابراین B را به عنوان والد جدید در نظر بگیرید.



خوب، شما به بالای حلقه بازگشته‌اید. همسایه‌ی بعدی `for` گرهی `Finish` است.

`for n in neighbors.keys():`

\uparrow $\brace{}$

$n =$ "پایان"

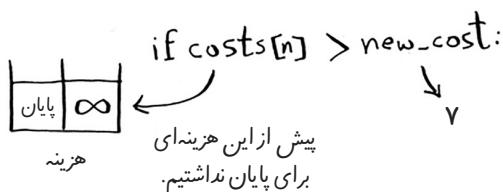
اگر از گرهی B عبور کنید چقدر طول می کشد تا به گرهی Finish برسید؟

$$\text{new_cost} = \text{cost} + \text{neighbors}[n]$$

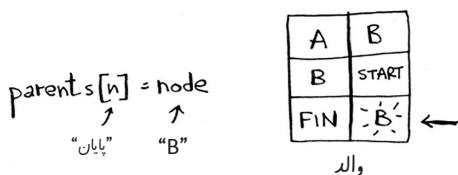
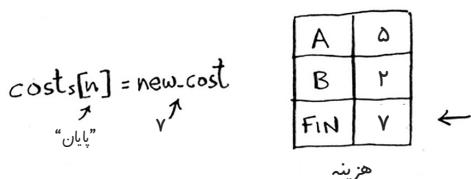
↓ ↓
 ۲ فاصله ب
 پایان: Δ

$2 + \Delta = V$

۷ دقیقه طول می‌کشد. هزینه‌ی قبلی بی‌نهایت دقیقه بود و ۷ دقیقه زمان کمتری از آن است.



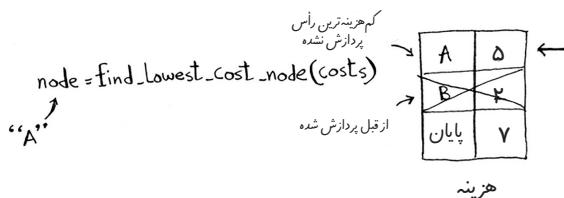
هزینه‌ی جدید و والد جدید را برای گرهی Finish در نظر بگیرید.



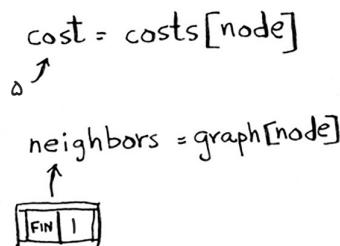
بسیار خوب، شما هزینه‌های همه‌ی همسایگان گرهی B را به روز کردید. این گره را به عنوان پردازش شده علامت بزنید.

processed.append(node)
“B”[↗]

گرهی بعدی را برای پردازش پیدا کنید.



هزینه و همسایگان گرهی A را به دست بیاورید.



گرهی A فقط یک همسایه دارد: گرهی Finish.

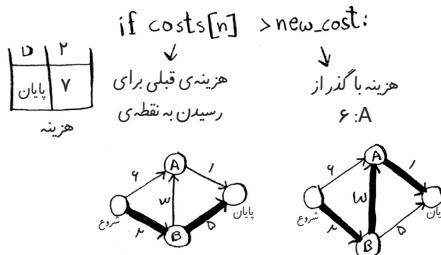
for n in neighbors.keys():
 ↗ {
 "پایان" } پایان

در حال حاضر ۷ دقیقه طول می‌کشد تا به گرهی Finish برسید. اگر از گرهی A عبور کنید چقدر طول می‌کشد تا به آنجا برسید؟

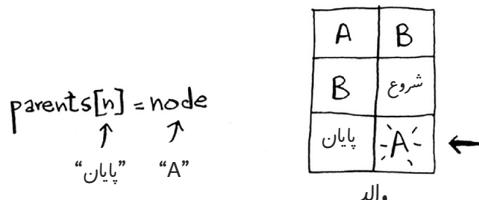
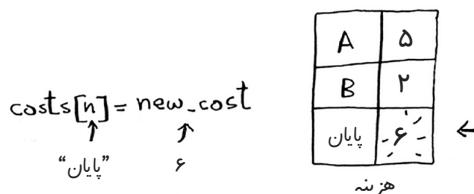
$$\text{new_cost} = \text{cost} + \text{neighbors}[n]$$

هزینه‌ی رسیدن فاصله‌ی A تا
 به از نقطه‌ی پایان: 1
 شروع: 5

} $\Delta + 1 = 6$



از گرهی A سریع‌تر می‌توان به Finish رسید! بیایید هزینه و والدها را به روز کنیم.



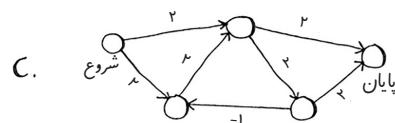
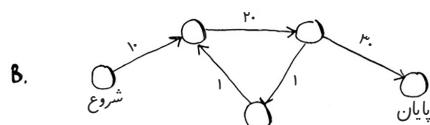
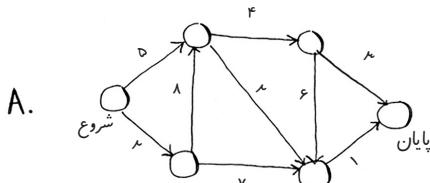
هنگامی که تمام گره‌ها را پردازش کردیم، الگوریتم به پایان می‌رسد. امیدوارم این راهنمایی به شما کمک کند تا الگوریتم را کمی بهتر درک کرده باشید. یافتن کم‌هزینه‌ترین گره با تابع `find_lowest_cost_node` بسیار آسان است. کد آن در ادامه آمده است:

```
def find_lowest_cost_node(costs):
    lowest_cost = float("inf")
    lowest_cost_node = None
    for node in costs: node را بررسی می‌کند
        cost = costs[node]
        if cost < lowest_cost and node not in processed: اگر تا اینجا کل کشته هزینه باشد و هنوز پروسس نشده باشد
            lowest_cost = cost آن را کم‌ترین هزینه‌ی جدید در نظر می‌گیرد
            lowest_cost_node = node
```

```
return lowest_cost_node
```

تمرین

۷.۱ در هر یک از این گراف‌ها، وزن کوتاه‌ترین مسیر ابتدا تا انتهای چقدر است؟



جمع‌بندی

- جست‌وجوی سطح اول برای محاسبه‌ی کوتاه‌ترین مسیر برای یک گراف بدون وزن استفاده می‌شود.
- الگوریتم دایجسترا برای محاسبه‌ی کوتاه‌ترین مسیر برای یک گراف وزن‌دار استفاده می‌شود.
- الگوریتم دایجسترا زمانی مفید است که همه‌ی وزن‌ها مثبت باشند.
- اگر وزن‌های منفی دارید، از الگوریتم بلمن-فورد استفاده کنید.



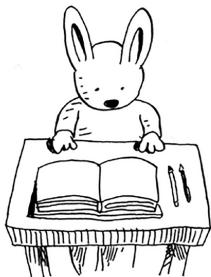
۸ | الگوریتم حریصانه^۱

در این فصل

- یاد می‌گیرید چگونه از پس غیرممکن‌ها بریایید: مسئله‌هایی که راه حل الگوریتمی سریعی ندارند (مسئله‌های ان‌پی کامل^۲).
- یاد می‌گیرید در مواجهه با این مسئله‌ها چگونه آن‌ها را شناسایی کنید و وقت خود را برای یافتن یک الگوریتم سریع برای آنها تلف نکنید.
- شما در مورد الگوریتم‌های تقریبی یاد می‌گیرید که می‌توانید از آن‌ها برای یافتن سریع یک راه حل تقریبی برای یک مسئله NP-complete استفاده کنید.
- در مورد استراتژی حریصانه که یک استراتژی ساده‌ای برای حل مسئله‌ی است، یاد می‌گیرید.

-
1. greedy algorithm
 2. NP-complete

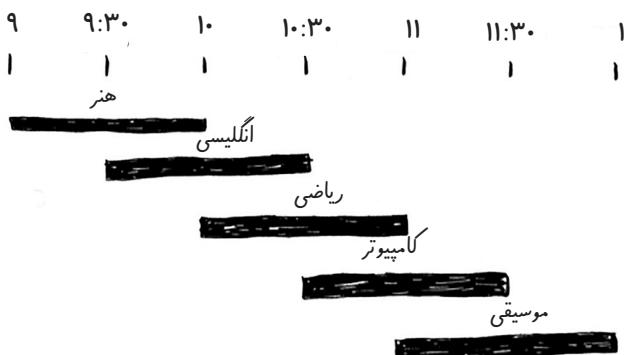
مسئله‌ی برنامه‌ریزی برای کلاس درس



فرض کنید تنها یک اتاق برای کلاس درس دارید و می‌خواهید بیشترین تعداد ممکن کلاس را در آنجا برگزار کنید. لیستی از کلاس‌ها در اختیار دارید:

	کلاس	پایان شروع
هر	۹ AM	10 AM
انگلیسی	9:۳۰ AM	10:۳۰ AM
ریاضی	10 AM	11 AM
کامپیوتر	10:۳۰ AM	11:۳۰ AM
موسیقی	11 AM	12 PM

شما نمی‌توانید همه‌ی این کلاس‌ها را در آنجا برگزار کنید، چون برخی از آن‌ها با یکدیگر هم‌پوشانی زمانی دارند.



شما می‌خواهید تا جایی که ممکن است کلاس‌های بیشتری را در این اتاق برگزار کنید. چگونه مجموعه‌ای از کلاس‌ها را انتخاب می‌کنید به شرط آنکه بزرگ‌ترین تعداد کلاس‌های ممکن را داشته باشد؟

مسئله‌ی دشواری به نظر می‌رسد، درست است؟ در واقع، آن قدر الگوریتم آسانی دارد که ممکن است شما را شگفت‌زده کند. نحوه‌ی عملکرد آن در اینجا آمده است:

۱. کلاسی را انتخاب کنید که زودتر به اتمام می‌رسد. این اولین کلاسی است که در این اتاق برگزار می‌کنید.
۲. حالا باید کلاسی را انتخاب کنید که بعد از اولین کلاس شروع شود. دوباره کلاسی را انتخاب کنید که زودتر تمام می‌شود. این دومین کلاسی است که برگزار می‌کنید.

به این کار ادامه داده تا در نهایت به پاسخ برسید! بیایید آن را امتحان کنیم. کلاس هنر قبل از همه، در ساعت ۱۰ صبح به پایان می‌رسد، بنابراین این یکی از کلاس‌هایی است که شما انتخاب می‌کنید.

هر	۹ AM	۱۰ AM	✓
انگلیسی	۹:۳۰ AM	۱۰:۳۰ AM	
ریاضی	۱۰ AM	۱۱ AM	
کامپیوتر	۱۰:۳۰ AM	۱۱:۳۰ AM	
موسیقی	۱۱ AM	۱۲ PM	

اکنون به کلاسی نیاز دارید که بعد از ساعت ۱۰ صبح شروع شده و زودتر از باقی کلاس‌ها به پایان می‌رسد.

هر	۹ AM	۱۰ AM	✓
انگلیسی	۹:۳۰ AM	۱۰:۳۰ AM	✗
ریاضی	۱۰ AM	۱۱ AM	✓
کامپیوتر	۱۰:۳۰ AM	۱۱:۳۰ AM	
موسیقی	۱۱ AM	۱۲ PM	

زبان انگلیسی به دلیل تداخل با هنر از فهرست مورد نظر ما حذف می‌شود، اما کلاس ریاضی مناسب است. در نهایت، علوم کامپیوتر با ریاضی تداخل زمانی دارد، اما کلاس موسیقی مشکلی ندارد.

هنر	9 AM	10 AM	✓
انگلیسی	9:۳۰ AM	10:۳۰ AM	✗
ریاضی	10 AM	11 AM	✓
کامپیوتر	10:۳۰ AM	11:۳۰ AM	✗
موسیقی	11 AM	12 PM	✓

بنابراین این سه کلاس را در این کلاس درس برگزار خواهید کرد.

۹	۹:۳۰	۱۰	۱۰:۳۰	۱۱	۱۱:۳۰	۱۲
۱	۱	۱	۱	۱	۱	۱
هنر	ریاضی	موسیقی				

بسیاری به من می‌گویند که از آنجایی که این الگوریتم آسان است و بیش از اندازه واضح است، پس لابد اشتباه است. اما همین ویژگی زیبایی الگوریتم‌های حریصانه است: آن‌ها آسان هستند! یک الگوریتم حریصانه ساده است: در هر مرحله، حرکت بهینه را انتخاب کنید. در این حالت، هر بار که کلاسی را انتخاب می‌کنید، کلاسی را انتخاب می‌کنید که زودتر از همه به پایان می‌رسد. از نظر فنی: در هر مرحله شمارا را حل بهینه محلی^۱ را انتخاب می‌کنید و در پایان با راه حل بهینه‌ی سراسری^۲ روبرو می‌شوید. باور کنید یا نه، این الگوریتم ساده راه حل بهینه را برای این مسئله‌ی زمان‌بندی پیدا می‌کند!

بدیهی است که الگوریتم‌های حریصانه همیشه کارساز نیستند. اما نوشتن آن‌ها ساده است! بیایید به مثال دیگری نگاه کنیم.

- Locally optimal solution
- Globally optimal solution



مسئله‌ی کوله‌پشتی

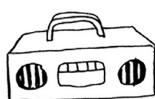
فرض کنید شما یک دزد طمع‌کار هستید. با یک کوله‌پشتی در یک فروشگاه هستید، و همه‌ی اقلام زیر برای دزدی موجود است اما تنها می‌توانید چیزهایی را که در کوله‌پشتی جا می‌شوند بددید. کوله‌پشتی می‌تواند ۳۵ پوند را در خود جای بدهد.

شما سعی می‌کنید ارزش اقلامی را که در کوله‌پشتی خود قرار می‌دهید به حداقل برسانید. از چه الگوریتمی استفاده می‌کنید؟

با استراتژی حریصانه باز هم بسیار ساده خواهد بود:

۱. گران‌ترین چیزی را که در کوله‌پشتی شما جا می‌شود انتخاب کنید.
۲. گران‌ترین چیز بعدی را که در کوله‌پشتی شما جا می‌شود انتخاب کنید. و به همین ترتیب.

تنها مشکل این است که این الگوریتم اینجا به کار نمی‌آید! به عنوان مثال، فرض کنید سه چیز برای سرقت موجود است.



ضبط صوت
۳۰۰ دلار
۳۰ پوند

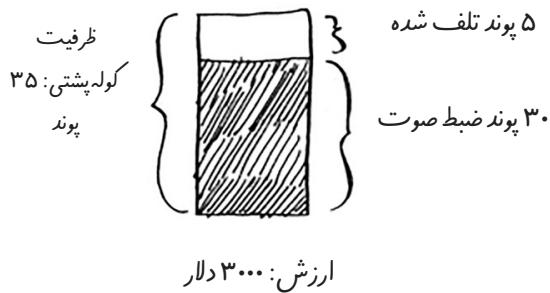


لپ‌تاپ
۲۰۰ دلار
۲۰ پوند



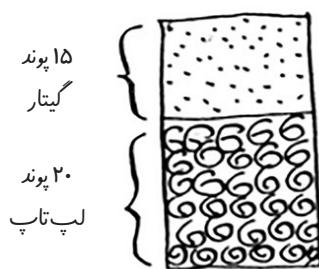
گیتار
۱۵۰ دلار
۱۵ پوند

در کوله‌پشتی ۳۵ پوند آیتم را می‌توان قرار داد. سیستم ضبط صوت از تمامی وسایل بالارزش‌تر است، بنابراین شما آن را می‌دزدید. اکنون برای هیچ چیز دیگری جای خالی باقی نمانده است.



ارزش: ۳۰۰۰ دلار

شما ۳۰۰۰ دلار کالا گیرتان آمده‌است. اما صبر کنید! اگر به جای آن لپ‌تاپ و گیتار را انتخاب می‌کردید، می‌توانستید معادل ۳۵۰۰ دلار سرقت کنید!



ارزش: ۳۵۰۰ دلار

واضح است که در اینجا استراتژی حریصانه راه حل بهینه را به شما نمی‌دهد. با این حال بسیار نزدیک به آن است. در فصل بعدی نحوه‌ی محاسبه‌ی راه حل صحیح را توضیح می‌دهم. اما اگر سارقی در یک مرکز خرید باشد، وضعیت ایده‌آل برای شما اهمیت چندانی ندارد. در چنین شرایطی «به نسبت خوب» کفایت می‌کند.

نتیجه‌ای که از مثال دوم می‌توان گرفت: گاهی، بی‌نقص بودن دشمن خوب بودن است. گاهی اوقات تنها چیزی که نیاز دارید یک الگوریتم است که مشکل را به خوبی حل کند. و ایجاست که الگوریتم‌های حریصانه می‌درخشنند، زیرا نوشتمن آن‌ها ساده است و

معمولًاً با نتایج ایده‌آل چندان تفاوتی ندارند.

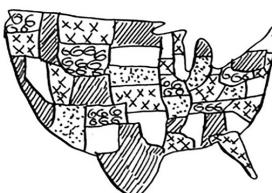
تمرین

۸.۱ شما برای یک شرکت مبلمان کار می‌کنید و باید تولیدات این شرکت را به سراسر کشور ارسال کنید. شما باید بسته‌ها را در کامیون بار بزنید. بسته‌ها اندازه‌های مختلفی دارند و شما سعی می‌کنید حداقل استفاده را از فضای کامیون داشته باشید. بسته‌ها را چگونه انتخاب می‌کنید تا از فضای داخل کامیون به شکل بهینه استفاده کنید؟ از استراتژی حریصانه استفاده کنید. آیا این الگوریتم شما را به راه حل بهینه می‌رساند؟

۸.۲ شما به اروپا می‌روید و هفت روز فرصت دارید تا جایی که زمان اجازه بدهد از مکان‌های گوناگون بازدید کنید. شما یک ارزش نقطه‌ای^۱ به هر مورد اختصاص می‌دهید (میزان اشتیاق به دیدن آن) و تخمین می‌زنید که بازدید هر کدام چقدر طول می‌کشد. چگونه می‌توانید مجموع ارزش را به حداقل برسانید (بازدید از تمام مکان‌هایی که واقعاً دوست دارید از آن‌ها بازدید کنید)؟ با یک استراتژی حریصانه. آیا این راه حلی بهینه به شما می‌دهد؟

باید به آخرین نمونه نگاه کنیم. این مثالی است که در آن استفاده از الگوریتم‌های حریصانه کاملاً ضروری است.

مسئله‌ی پوشش مجموعه^۲



فرض کنید می‌خواهید برنامه‌ای رادیویی داشته باشید و این برنامه در هر ۵۰ ایالت برای شنوندگان در دسترس باشد. باید بررسی کنید که باید این برنامه را از کدام ایستگاه‌ها مخابره کرد تا به گوش همه‌ی مخاطب‌ها برسد. پخش در هر ایستگاه هزینه دارد، بنابراین سعی می‌کنید تعداد ایستگاه‌هایی را که از آن مخابره می‌کنید به حداقل برسانید. شما لیستی از ایستگاه‌ها دارید.

1. Point value
2. The set-covering problem

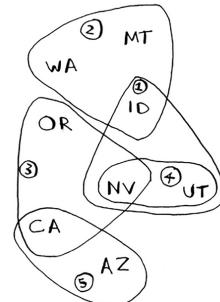
دسترسی در ایستگاه رادیو

KONE	ID,NV,UT
KTWO	WA, ID, MT
KTHREE	OR, NV, CA
KFOUR	NV, UT
KFIVE	CA, AZ

... وغیره

هر ایستگاه یک منطقه را پوشش می‌دهد و البته هم‌پوشانی هم وجود دارد.

چگونه می‌توانید کوچک‌ترین مجموعه ایستگاه را انتخاب کنید به شرط آنکه برنامه در تمامی 5^n ایالت مخابره شود؟ آسان به نظر می‌رسد، این طور نیست؟ در عمل کاری بسیار



دشوار است. در اینجا شیوه‌ای انجام آن آمده است:

۱. تمام زیرمجموعه‌های ممکن از ایستگاه‌ها را فهرست کنید. به این زیرمجموعه‌ها مجموعه‌ی توانی^۱ گفته می‌شود و تعداد آن‌ها 2^{8n} است.

SET #1 ... SET #8 ... SET #500



۲. از میان آن‌ها، مجموعه‌ای را انتخاب کنید که کمترین تعداد ایستگاه را با فرض پوشش تمامی 5^n ایالت شامل می‌شود.

مشکل این است که محاسبه‌ی هر زیرمجموعه‌ی ممکن از ایستگاه‌ها زمان بر است. زمان $O(2^{8n})$ طول می‌کشد، چون 2^{8n} ایستگاه وجود دارد. اگر مجموعه‌ی کوچکی از ۵ تا ۱۰ ایستگاه دارید، این کار ممکن است. اما با تمام مثال‌های اینجا، به این فکر کنید که اگر آیتم‌های زیادی داشته باشید چه پیش می‌آید. اگر ایستگاه‌های بیشتری داشته باشید خیلی بیشتر طول می‌کشد. فرض کنید بتوانید در هر ثانیه ۱۰ زیرمجموعه حساب کنید. هیچ الگوریتمی نیست که بتواند آن را با سرعتی مناسب حل کند! چاره چیست؟

تعداد ایستگاه	زمان لازم
۵	۳.۲ ثانیه
۱۰	۱۰۲.۴ ثانیه
۳۲	۱۳.۶ سال
۱۰۰	4×10^{11} سال

الگوریتم‌های تقریبی^۱

الگوریتم‌های حریصانه نجات‌بخش هستند! در اینجا یک الگوریتم حریصانه معرفی می‌شود که تقریباً به همین صورت است:

۱. ایستگاهی را انتخاب کنید که برای ایالت‌هایی که هنوز امکان دریافت برنامه‌ی رادیویی را ندارند پوشش حداکثری داشته باشد. مشکلی نیست اگر ایستگاهی برای برخی از ایالت‌ها با ایستگاه‌های قبلی هم‌پوشانی داشته باشد.
۲. این کار را تا زمانی که تمام ایالت‌ها تحت پوشش قرار بگیرند، ادامه بدهید.

به این الگوریتم تقریبی می‌گویند. در شرایطی که محاسبه‌ی راه حل دقیق زمان زیادی می‌برد، یک الگوریتم تقریبی کارساز خواهد بود. الگوریتم‌های تقریبی بر اساس معیارهای زیر داوری می‌شوند:

- چقدر سریع هستند
- چقدر به راه حل بهینه نزدیک هستند

الگوریتم‌های حریصانه انتخاب خوبی هستند زیرا نه تنها دستیابی به آن‌ها ساده است، بلکه به این معنی است که معمولاً سریع هم اجرا می‌شوند. در این حالت، الگوریتم

حریصانه در زمان $O(n^2)$ اجرا می‌شود. در اینجا n تعداد ایستگاه‌های رادیویی است. بیایید کد این مسئله را بررسی کنیم.

کد

در این مثال، برای سادگی مسئله تنها از زیر مجموعه‌ی ایالت‌ها و ایستگاه‌ها استفاده می‌کنم.

ابتدا فهرستی تهیه کنید از ایالت‌هایی که قصد پوشش آن‌ها را دارید:

`states_needed = set(["mt", "wa", "or", "id", "nv", "ut", "ca", "az"])` یک آرایه وارد کرده و به شما `set` بررسی گرداند.

برای این کار از یک `set` استفاده کردم. یک `set` مانند یک لیست است، با این تفاوت که در `set` هر آیتم تنها یک بار در مجموعه ظاهر می‌شود. `Set` ها نمی‌توانند تکراری باشند. برای مثال، فرض کنید این لیست را دارید:

```
>>> arr = [1, 2, 2, 3, 3, 3]
```

و شما آن را به یک `set` تبدیل کردید:

```
>>> set(arr)
set([1, 2, 3])
```

۱، ۲، و ۳ در `set` فقط یک بار نمایش داده می‌شوند.

[۱،۲،۲،۳،۳،۳] → تبديل به
set → (۱،۲،۳)
SET

شما همچنین به لیست ایستگاه‌هایی که از بین آن‌ها انتخاب می‌کنید نیاز دارید. برای این منظور از هش استفاده کردم:

```

stations = { }
stations["kone"] = set(["id", "nv", "ut"])
stations["ktwo"] = set(["wa", "id", "mt"])
stations["kthree"] = set(["or", "nv", "ca"])
stations["kfour"] = set(["nv", "ut"])
stations["kfive"] = set(["ca", "az"])

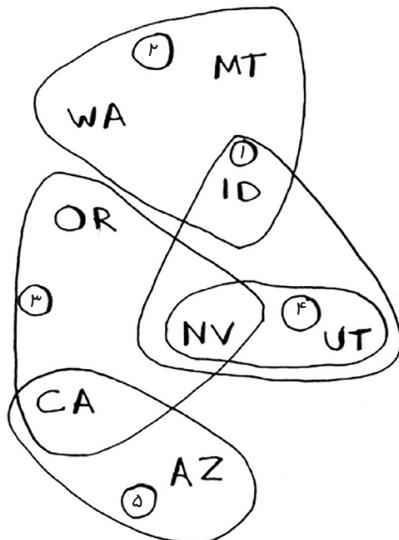
```

کلیدها نام ایستگاه‌ها هستند و مقدارها ایالت‌هایی هستند که آن ایستگاه‌ها پوشش می‌دهند.

بنابراین در این مثال، ایستگاه `kone` آیداهو، نوادا و یوتا را پوشش می‌دهد. تمام مقادیر نیز `set` هستند. به زودی می‌بینید که در همه‌ی شرایط استفاده از `set` زندگی شما را آسان‌تر می‌کند.

در نهایت، نیاز است تا `set` نهایی ایستگاه‌هایی که استفاده می‌کنید در جایی نگهداری بشود:

```
final_stations = set()
```



محاسبه‌ی پاسخ

اکنون باید حساب کنید که از چه ایستگاه‌هایی استفاده می‌کنید. به تصویر سمت چپ نگاه کرده و ببینید آیا می‌توانید پیش‌بینی کنید که از چه ایستگاه‌هایی باید استفاده کرد.

ممکن است پیش از یک راه حل صحیح وجود داشته باشد. شما باید به هر یک از ایستگاه‌ها نگاه کرده و ایستگاهی را انتخاب کنید که بیشتر ایالت‌های پوشش نیافرته را شامل شود. نام آن را `best_station`

در نظر می‌گیرم:

```
best_station = None
states_covered = set()

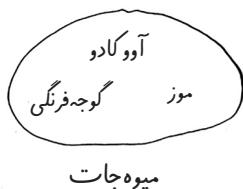
for station, states_for_station in stations.items():
```

ستی است از تمام ایالت‌های هنوز پوشش داده‌نشده‌ای که ایستگاه مورد اشاره آن‌ها را پوشش می‌دهد. حلقه‌ی `for` به شما این امکان را می‌دهد که بر روی هر ایستگاهی حلقه اعمال کنید تا ببینید کدام یک مناسب‌ترین ایستگاه است. بباید به کد حلقه‌ی `for` نگاه کنیم:

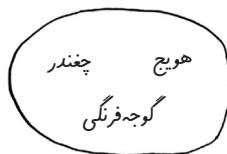
```
covered = states_needed & states_for_station <----  
if len(covered) > len(states_covered):    سیتکش جدیداً آن  
    best_station = station                    set intersection  
    states_covered = covered                  می‌گویند.
```

در اینجا یک خط کد عجیب و غریب می‌بینید:

```
covered = states_needed & states_for_station  
قضیه چیست؟
```



فرض کنید یک سیت از میوه‌ها دارد.



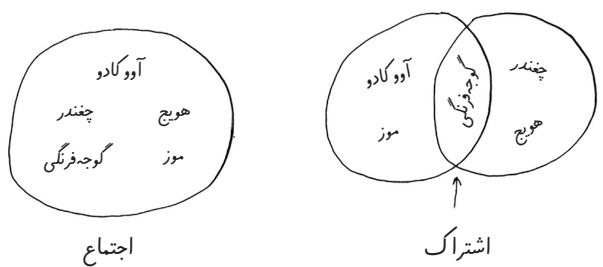
همچنین یک سیت از سبزیجات دارد.

وقتی دو دسته سِت داشته باشید، می‌توانید کارهای جالبی با آن‌ها انجام دهید.

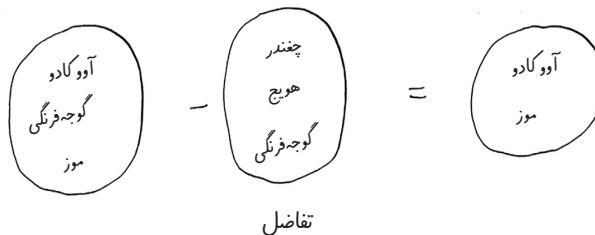
در ادامه کارهایی که می‌توانید با ستها انجام دهید، آمده است.

یا سبزیجات یا میوه‌جات:

هم سبزیجات هم میوه‌جات:



چیزهایی که میوه هستند اما
سیر یا جات نیستند



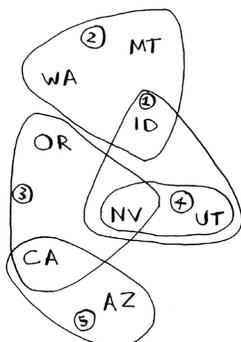
- اجتماع مجموعه^۱ به معنای «ترکیب هر دو مجموعه» است.
 - اشتراک مجموعه^۲ به معنی «یافتن آیتم‌هایی است که در هر دو مجموعه ظاهر می‌شوند» (در این مورد، فقط گوجه فرنگی).
 - تفضیل مجموعه^۳ به این معنی است که «اقلام یک مجموعه را از آیتم‌های مجموعه‌ی دیگر کم کنید.»

1. t
2. Set intersection
3. Set difference

به عنوان مثال:

```
>>> fruits = set(["avocado", "tomato", "banana"])
>>> vegetables = set(["beets", "carrots", "tomato"])
>>> fruits | vegetables <----- اجتماع مجموعه
set(["avocado", "beets", "carrots", "tomato", "banana"])
>>> fruits & vegetables <----- اشتراک مجموعه
set(["tomato"])
>>> fruits - vegetables <----- تفاضل مجموعه
set(["avocado", "banana"])
>>> vegetables - fruits <----- حاصل این عبارت چیست؟
```

جمع‌بندی:



- سِت‌ها مانند لیست‌ها هستند، با این تفاوت که سِت‌ها نمی‌توانند عضو تکراری داشته باشند.
- می‌توانید چند عملیات جالب مانند اجتماع، اشتراک، تفاضل را روی سِت‌ها اعمال کنید.

بازگشت به کد

بیایید به مثال اصلی برگردیم.
این کد یک اشتراک مجموعه است:

```
covered = states_needed & states_for_station
```

مجموعه‌ای از ایالت‌ها است که در هر دوی covered و states_needed از ایالت‌ها است. covered حضور دارد. بنابراین set covered ایالت‌های بدون states_for_station

پوشش است که این ایستگاه پوشش می‌دهد! بعد بررسی می‌کنید که آیا این ایستگاه ایالت‌های بیشتری را نسبت به `best_station` فعلى پوشش می‌دهد یا خیر:

```
if len(covered) > len(states_covered):
    best_station = station
    states_covered = covered
```

اگر چنین باشد، این ایستگاه `best_station` جدید است. در نهایت، پس از اتمام حلقه‌ی `for` را به لیست نهایی ایستگاه‌ها اضافه می‌کنید:

```
final_stations.add(best_station)
```

همچنین باید `states_needed` را به روز کنید. از آنجایی که این ایستگاه برخی از ایالت‌ها را پوشش می‌دهد، دیگر به این ایالت‌ها نیازی نیست:

```
states_needed -= states_covered
```

و شما حلقه می‌زنید تا `states_needed` خالی شود. در اینجا کد کامل این حلقه آمده است:

```
while states_needed:
    best_station = None
    states_covered = set()

    for station, states in stations.items():
        covered = states_needed & states

        if len(covered) > len(states_covered):
            best_station = station
            states_covered = covered

    states_needed -= states_covered
    final_stations.add(best_station)
```

در نهایت، می‌توانید `final_stations` را پرینت کنید، باید این نوشه را مشاهده کنید:

```
>>> print final_stations
set(['ktwo', 'kthree', 'kone', 'kfive'])
```

این همان چیزی است که انتظار داشتید؟ به جای ایستگاه‌های ۱، ۲، ۳ و ۵، می‌توانید ایستگاه‌های ۴، ۳، ۲ و ۵ را انتخاب کنید. بیایید زمان اجرای الگوریتم حریصانه را با الگوریتم دقیق مقایسه کنیم.

تعداد ایستگاه‌ها	$O(n!)$ الگوریتم دقیق	$O(n^2)$ الگوریتم حریصانه
۵	۳.۲ ثانیه	۲.۵ ثانیه
۱۰	۱۰۲.۴ ثانیه	۱۰ ثانیه
۳۲	۱۳.۶ سال	۱۰۲.۴ ثانیه
۱۰۰	4×10^{۲۹} سال	۱۶.۶۷ دقیقه

تمرین

هر یک از الگوریتم‌های زیر را بررسی کنید که آیا الگوریتم حریصانه هستند یا خیر.

۸.۳ مرتب‌سازی سریع

۸.۴ جست‌وجوی سطح اول

۸.۵ الگوریتم دایجسترا



مسئله‌ی ان‌پی‌کامل

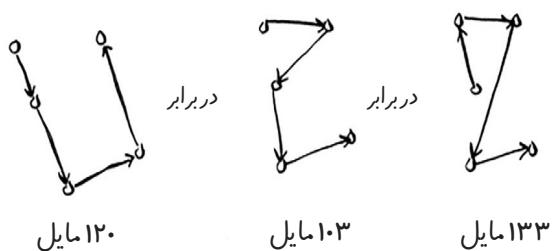
برای حل مسئله‌ی پوشش-مجموعه، باید تمامی مجموعه‌های ممکن را محاسبه می‌کردید.

1. Exact algorithm

شاید به یاد مسئله‌ی فروشنده‌ی دوره‌گرد فصل یک افتاده باشید. در این مسئله، یک فروشنده باید به پنج شهر مختلف سفر کند.



و او در تلاش است کوتاه‌ترین مسیری بیابد که از هر پنج شهر بگذرد. برای یافتن کوتاه‌ترین مسیر، ابتدا باید هر مسیر ممکن را محاسبه کنید.

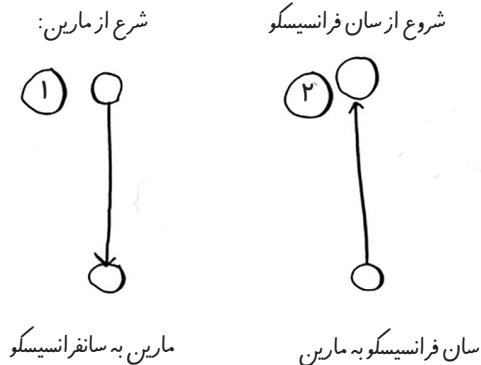


... 9

چند مسیر برای پنج شهر پاید محاسبه کنید؟

فروشنده‌ی دوره‌گرد، گام به گام

از جزء شروع می‌کنیم. تنها دو شهر را فرض کنید. دو مسیر برای انتخاب وجود دارد.



مسیر یکسان یا متفاوت؟

ممکن است فکر کنید این باید همان مسیر باشد. در نهایت مگر مسیر $SF > M$ با $M > SF$ برابر نیست؟ نه لزوماً. برخی از شهرها (مانند سانفرانسیسکو) خیابان‌های یک طرفه‌ی زیادی دارند، بنابراین نمی‌توانید از راهی که آمده‌اید بازگردید. همچنین ممکن است مجبور شوید ۱ یا ۲ مایل از مسیر خود دور شوید تا یک ورودی به بزرگراه پیدا کنید. بنابراین این دو مسیر یکسان نیستند.

ممکن است از خود بپرسید، «در مسئله‌ی فروشنده‌ی دوره‌گرد، آیا شهر خاصی وجود دارد که باید از آن شروع کنید؟» برای مثال، فرض کنید من یک فروشنده‌ی دوره‌گرد هستم و در سانفرانسیسکو زندگی می‌کنم و باید به چهار شهر دیگر سفر کنم. سانفرانسیسکو مبداء سفر من است.

اما گاهی اوقات شهر مبداء تعیین نمی‌شود. فرض کنید شما در FedEx کار می‌کنید و

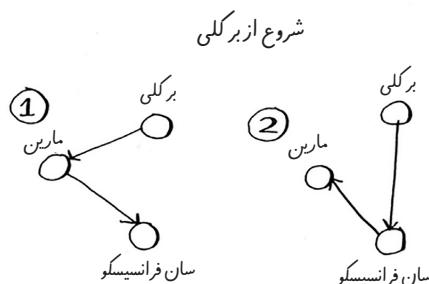
می‌خواهید بسته‌ای را در Bay Area تحویل دهید. این بسته از شیکاگو به یکی از ۵۰ مرکز FedEx در منطقه Bay Area منتقل می‌شود. سپس آن بسته بار کامیونی می‌شود تا برای تحویل بسته‌ها به محل‌های مختلف برود. به کدام مکان باید رفت؟ در اینجا مبدأ ناشناخته است. محاسبه‌ی مسیر بهینه و مبدأ فروشنده‌ی دوره‌گرد به عهده‌ی شماست.

زمان اجرای هر دو نسخه یکسان است. اما اگر شهری به عنوان مبدأ مشخص نشود، مثال ساده‌تری می‌شود و برای همین با همین مثال کار را پیش می‌برم.

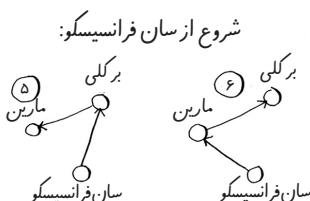
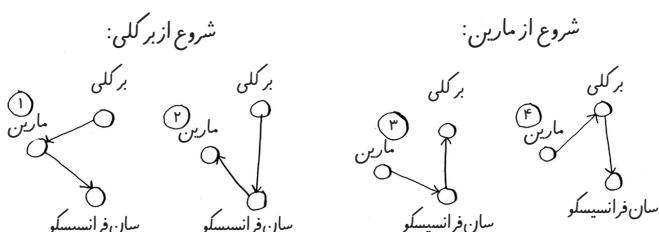
دو شهر = دو مسیر ممکن.

۳ شهر

حالا فرض کنید یک شهر دیگر اضافه کنید. چند مسیر ممکن وجود دارد؟ اگر از برکلی شروع کنید، دو شهر دیگر هم برای بازدید دارید.



در مجموع شش مسیر وجود دارد، دو مسیر برای سفر از هر شهر.



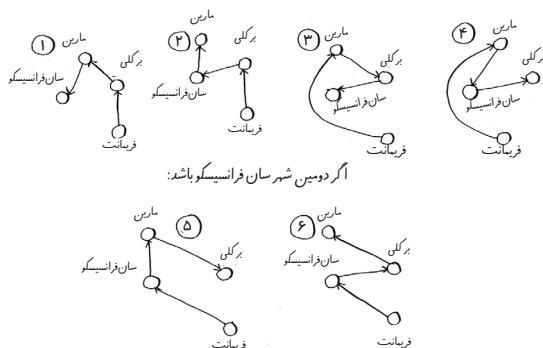
بنابراین سه شهر = شش مسیر ممکن.

۴ شهر

بیایید شهر فرمونت را اضافه کنیم. حالا فرض کنید سفر را از این شهر شروع می‌کنید.

شروع از

اگر دومین شهر مارین باشد: فرمیانت: اگر دومین شهر برکلی باشد:



شش مسیر ممکن از مبداء فرمیانت هست. نگاه کنید! آن‌ها بسیار شبیه به شش مسیری هستند که در زمانی که تنها سه شهر داشتید محاسبه کردید. با این تفاوت که در حال حاضر همه‌ی مسیرها یک شهر دیگر را هم شامل می‌شوند، فرمیانت! یک الگو وجود دارد. فرض کنید چهار شهر هست، و فرمیانت را شهر مبداء انتخاب می‌کنید. سه شهر باقی می‌ماند. و می‌دانید که اگر سه شهر باشد، شش مسیر مختلف برای رفت‌وآمد میان آن شهرها وجود دارد. اگر از فرمیانت شروع کنید، شش مسیر ممکن است. همچنین می‌توانید از شهر دیگری شروع کنید.

شروع از مارین:

شروع از سان فرانسیسکو:

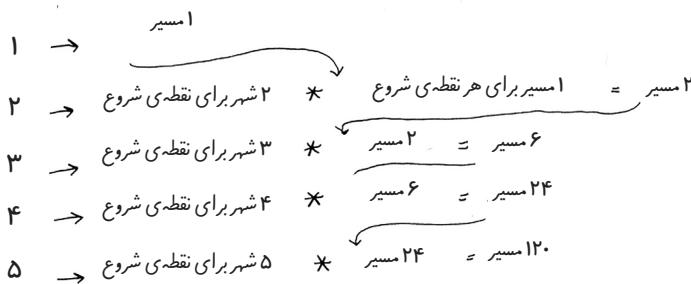
$$= 6 \text{ مسیر ممکن} = 6 \text{ مسیر ممکن} =$$

شروع از برکلی:

$$= 6 \text{ مسیر ممکن} =$$

چهارگزینه برای شهر مبداء، با شش مسیر ممکن برای هر شهر مبداء $4 * 6 = 24$ مسیر ممکن می‌شود.
الگوی مشخصی مشاهده می‌کنید؟ هر بار که شهر جدیدی اضافه می‌کنید، تعداد مسیرهایی را که باید حساب کنید افزایش می‌دهید.

تعداد شهرها



چند مسیر ممکن برای شش شهر وجود دارد؟ 720 مسیر. 5040 مسیر برای 7 شهر، 40320 مسیر برای 8 شهر.

این همان تابع فاکتوریل است (به خاطر دارید در مورد آن در فصل ۳ خواندید?). پس $120 = 5!$ فرض کیم 10 شهر باشد. چند مسیر ممکن وجود دارد؟ $3628800 = 10!$. برای 10 شهر باید بیش از 3 میلیون مسیر ممکن را حساب کنید. همان‌طور که می‌بینید تعداد مسیرهای ممکن به سرعت زیاد می‌شود!

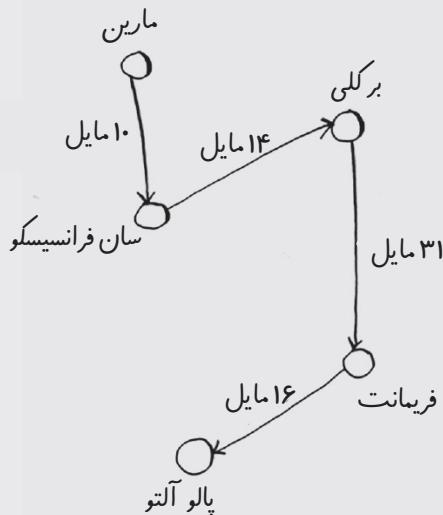
به همین دلیل در صورت داشتن تعداد زیادی شهر، محاسبه‌ی راه حل «صحیح» برای مسئله‌ی فروشنده‌ی دوره‌گرد غیرممکن است.

مسئله‌ی فروشنده‌ی دوره‌گرد و مسئله‌ی پوشش مجموعه در یک مورد مشترک هستند: شما هر راه حل ممکن را محاسبه می‌کنید و کوچک‌ترین/کوتاه‌ترین راه حل را انتخاب می‌کنید. هر دوی این مسئله‌ها ان‌پی کامل هستند.

تقریب

یک الگوریتم تقریبی مناسب برای مسئله‌ی فروشنده‌ی دوره‌گرد دارای چه ویژگی‌هایی است؟ باید الگوریتم ساده‌ای باشد که مسیر کوتاهی پیدا کند. قبل از آنکه ادامه دهیم، ببینید پاسخی می‌توانید بیابید.

روش کار من به این شکل است: به دلخواه یک شهر را به عنوان مبدأ انتخاب می‌کنم. سپس، هر بار که فروشنده باید شهر بعدی را برای سفر انتخاب کند، نزدیک‌ترین شهر بازدید نشده را انتخاب می‌کنم. فرض کنید مارین نقطه‌ی آغاز باشد.



مسافت کل: ۷۱ مایل. شاید این کوتاه‌ترین مسیر نباشد، با این حال مسیر بسیار کوتاهی است.

در ادامه توضیح کوتاهی درباره‌ی ان‌پی کامل آمده است: حل بعضی از مسئله‌ها حسابی سخت هستند. به عنوان مثال از مسئله‌های فروشنده‌ی دوره‌گرد و پوشش می‌توان نام برد. بر اساس نظر بسیاری از افراد باهوش و کاردست الگوریتمی برای حل سریع این مسئله‌ها نمی‌توان نوشت.

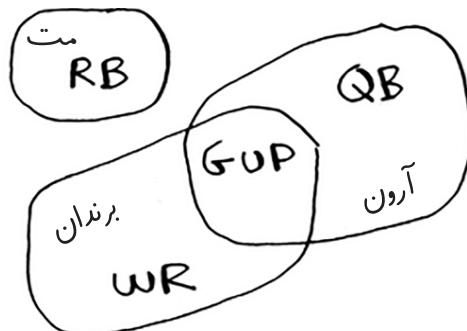
تشخیص مسئله‌ی ان‌پی کامل



یونا در حال انتخاب بازیکن برای تیم فوتبال خیالی خود است. او توانایی‌هایی را که برای بازیکن‌ها در نظر دارد، لیست کرده است: مهاجم نوک حمله‌ی کاردست، مهاجم توپ بر خوب، با توانایی بازی خوب در هوای بارانی، زیرفشار و نظایر آن. او لیستی از بازیکنان دارد که در آن توانایی هر بازیکن قید شده است.

بازیکن	ویژگی
مت فرت	RB
برندان مارشال	تحت‌فشار علکرده‌خوبی / WR دارد.
آرون راجرز	تحت‌فشار علکرده‌خوبی / QB دارد.
...	...

جونا به تیمی نیاز دارد که تمام توانایی‌های مورد نظر او را برأورده کند. همچنین تعداد نفرات تیم محدود است. یونا متوجه ماجرا می‌شود: «یک لحظه صبر کنید. این یک مسئله‌ی پوشش مجموعه است!»

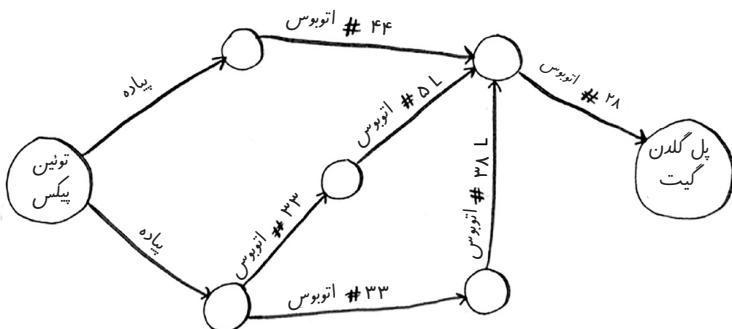


جونا می‌تواند از همان الگوریتم تقریبی برای ایجاد تیم خود استفاده کند:

۱. بازیکنی را انتخاب کنید که از میان توانایی‌های مورد نیاز بتواند بیشترین ویژگی‌ها را شامل بشود.

۲. این کار را تا زمانی تکرار کنید که بازیکنان انتخابی تیم تمام توانایی‌های مورد نظر را داشته باشند (یا فضای تیم شما تمام شود).

مسئله‌های ان‌پی کامل همه جا حضور پیدا می‌کنند! خوب است که تشخیص بدھید مسئله‌ای که مشغول حل آن هستید ان‌پی کامل است یا خیر. در این مرحله، می‌توانید از تلاش برای حل کامل آن دست کشیده و به جای آن با استفاده از یک الگوریتم تقریبی به جواب برسید. اما تشخیص اینکه آیا مسئله‌ای که روی آن کار می‌کنید ان‌پی کامل است یا خیر، دشوار است. عموماً بین مسئله‌ای که به راحتی قابل حل است و مسئله‌ای که جواب برسید. اما تفاوت بسیار کمی وجود دارد. به عنوان مثال، در فصل‌های قبل در مورد کوتاه‌ترین مسیرها بسیار صحبت کردم. شما می‌دانید که چگونه باید کوتاه‌ترین راه را برای رسیدن از نقطه‌ی A به نقطه‌ی B محاسبه کنید.



اما اگر بخواهید کوتاه‌ترین مسیری را بیابید که چندین نقطه را به هم متصل کند، مثل مسئله‌ی فروشنده‌ی دوره‌گرد و ان‌پی کامل خواهد بود. پاسخ کوتاه: هیچ راه آسانی برای تشخیص اینکه آیا مسئله‌ای که روی آن کار می‌کنید ان‌پی کامل است وجود ندارد. چند نکته را با هم مرور کنیم:

- الگوریتم شما اگر تعداد آیتم‌ها انگشت‌شمار باشد به سرعت اجرا می‌شود، اما با آیتم‌های بیشتر اساسی کند می‌شود.

- «همه‌ی ترکیبات X» عموماً به یک مسئله‌ی ان‌پی کامل اشاره می‌کنند.

- آیا باید «هر نسخه‌ی ممکن» X را محاسبه کنید چون نمی‌توانید آن را به مسئله‌های فرعی کوچک‌تر تقسیم کنید؟ ان پی کامل را امتحان کنید.
- اگر مسئله‌ی شما شامل یک توالی باشد (مانند دنباله‌ای از شهرها، مانند فروشنده‌ی دوره‌گرد)، و حل آن دشوار باشد، ان پی کامل را امتحان کنید.
- اگر مسئله‌ی شما شامل یک مجموعه (مانند مجموعه‌ای از ایستگاه‌های رادیویی) است و حل آن سخت است، ان پی کامل را امتحان کنید.
- می‌توانید مسئله‌ی خود را به عنوان مسئله‌ی پوشش مجموعه یا مسئله‌ی فروشنده‌ی دوره‌گرد بیان کنید؟ پس مسئله‌ی شما قطعاً ان پی کامل است.

تمرین

- ۸.۶** یک پستچی باید به ۲۰ خانه بسته تحویل بدهد. او باید کوتاه‌ترین مسیری را که به تمام ۲۰ خانه می‌رود، پیدا کند. آیا این یک مسئله‌ی ان پی کامل است؟
- ۸.۷** پیدا کردن بزرگ‌ترین دسته در یک مجموعه از افراد (یک دسته به مجموعه افرادی گفته می‌شود که همه‌ی افراد آن مجموعه هم‌دیگر را می‌شناسند). آیا این یک مسئله‌ی ان پی کامل است؟
- ۸.۸** شما مشغول رسم کردن نقشه‌ی ایالات متحده هستید و باید ایالت‌های مجاور را با رنگ‌های مختلف رنگ‌آمیزی کنید. باید حداقل تعداد رنگ مورد نیاز خود را پیدا کنید تا هیچ دو ایالت مجاور هم رنگ نباشند. آیا این یک مسئله‌ی ان پی کامل است؟

جمع‌بندی

- الگوریتم‌های حریصانه به صورت محلی بهینه‌سازی می‌شوند، به این امید که در نهایت به یک بهینه‌ی سراسری برسند.
- مسئله‌های ان‌پی کامل راه حل سریع شناخته شده‌ای ندارند.
- اگر مسئله‌ای از ان‌پی کامل دارید، بهترین گزینه استفاده از یک الگوریتم تقریبی است.
- نوشتن الگوریتم‌های حریصانه آسان و سرعت اجرای آن زیاد است، بنابراین برای الگوریتم‌های تقریبی مناسب هستند.

۹ | برنامه‌نویسی پویا^۱



در این فصل

- درباره‌ی برنامه‌نویسی پویا یاد می‌گیرید. برنامه‌نویسی پویا تکنیکی برای حل یک مسئله‌ی دشوار با تقسیم کردن آن به زیرمسئله^۲ و شروع کار با حل آن زیرمسئله‌ها است.

- با حل مثال، یاد می‌گیرید که یک راه حل برنامه‌نویسی پویا برای یک مسئله‌ی جدید ارائه کنید.



مسئله‌ی کوله‌پشتی

بیایید دوباره مسئله‌ی کوله‌پشتی فصل ۸ را مرور کنیم. شما یک سارق هستید با یک کوله‌پشتی که ظرفیت ۴ پوندی دارد.

1. Dynamic programming
2. subproblem

این سه آیتم را می‌توان در کوله‌پشتی قرار داد:



ضبط صوت

۳۰۰۰ دلار

۴ پوند

لپ تاپ

۲۰۰۰ دلار

۳ پوند

گیتار

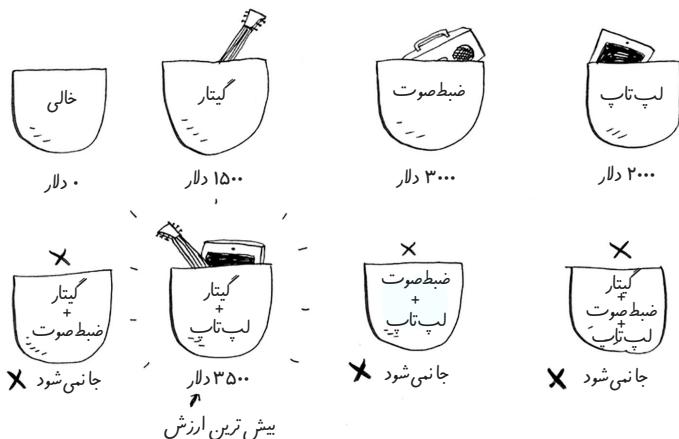
۱۵۰۰ دلار

۱ پوند

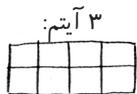
کدام یک از این‌ها را باید بدزدید تا ارزش کالاهای سوخت شده حداکثر باشد؟

راه حل ساده

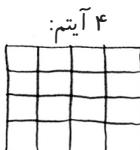
ساده‌ترین الگوریتم این است: تمام مجموعه کالاهای ممکن را امتحان می‌کنید و مجموعه‌ای را پیدا می‌کنید که بیشترین ارزش را داشته باشد.



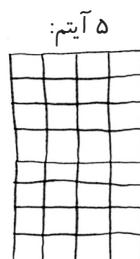
این روش هرچند نتیجه‌بخش است، اما بسیار کند است. برای ۳ آیتم باید ۸ مجموعه‌ی ممکن را حساب کنید. برای ۴ آیتم باید ۱۶ مجموعه حساب کنید. با هر آیتمی که اضافه می‌کنید، تعداد مجموعه‌هایی که باید حساب کنید دو برابر می‌شود! این الگوریتم به زمان $O(2^n)$ نیاز دارد که خیلی خیلی کند است.



۸ مجموعه مکن



۱۶ مجموعه مکن



۳۲ مجموعه مکن

۴ آیتم =
۴ بیلیارد
مجموعه مکن !!

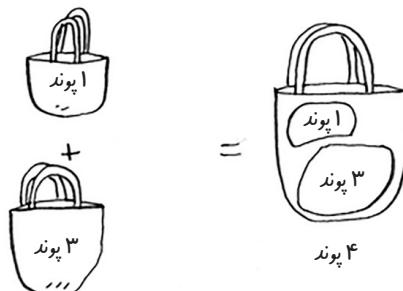
این روش برای تعداد متوسطی از کالا غیر عملی است. در فصل ۸ نحوه محاسبه‌ی یک جواب تقریبی را مشاهده کردید. آن راه حل اگر چه نزدیک به راه حل بهینه است، اما لزوماً می‌تواند راه حل بهینه نباشد.

پس چگونه جواب بهینه را محاسبه می‌کنید؟

برنامه‌نویسی پویا

پاسخ: با برنامه نویسی پویا! بیایید ببینیم در اینجا الگوریتم برنامه‌نویسی پویا به چه شکل کار می‌کند. در برنامه‌نویسی پویا با حل مسئله‌های فرعی شروع کرده و آن را به حل مسئله‌ی بزرگ بسط می‌دهید.

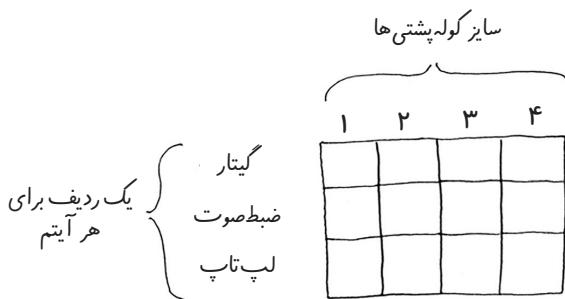
برای مسئله‌ی کوله پشتی، با حل مسئله‌ی کوله‌پشتی‌های کوچک‌تر (یا «کوله‌پشتی‌های فرعی») شروع می‌کنید و سپس تا حل مسئله‌ی اصلی پیش می‌روید.



برنامه‌نویسی پویا مفهوم سختی است، بنابراین نگران نشوید اگر فوراً آن را متوجه نشدید. مثال‌های زیادی را با هم بررسی می‌کنیم.

ابتدا با نشان دادن این الگوریتم و به شکل عملی شروع می‌کنم. بعد از آن که یک بار آن را در عمل دیدید، سؤال‌های زیادی در ذهن شما شکل می‌گیرد! تمام تلاشم را می‌کنم تا به تمامی این موارد بپردازم.

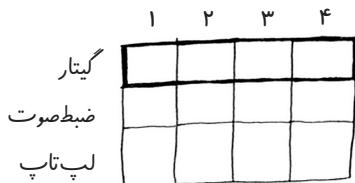
هر الگوریتم برنامه‌نویسی پویا با یک شبکه^۱ شروع می‌شود. در ادامه شبکه‌ای برای مسئله‌ی کوله پشتی آمده است.



ردیف‌های شبکه ایتم‌ها و ستون‌ها وزن‌های کوله‌پشتی از ۱ پوند تا ۴ پوند هستند. شما به تمام آن ستون‌ها نیاز دارید زیرا به شما در محاسبه‌ی مقادیر کوله‌های فرعی کمک می‌کنند. شبکه در ابتدا تهی است. شما باید تمام سلول‌های شبکه را پر کنید. با پرشدن شبکه، به پاسخ این مسئله دست می‌باید! لطفاً به همین شکل پیش بروید. شبکه‌ی خود را بسازید، تا با هم آن را پر کنیم.

ردیف گیتار

فرمول دقیق محاسبه‌ی این شبکه را بعدتر به شما می‌گوییم. ابتدا بیایید نگاهی به آن بیندازیم. از ردیف اول شروع کنید.



1. grid

ردیف اول به گیتار اختصاص دارد، به این معنی که شما سعی می‌کنید گیتار را در کوله‌پشتی جای بدهید. در هر سلول، یک تصمیم‌گیری ساده وجود دارد: گیتار را بذدید یا نه؟ به یاد داشته باشید، شما در تلاش برای یافتن مجموعه‌ای از اقلام هستید که بیشترین ارزش را از سرقت نصیب شما بکند.

سلول اول دارای یک کوله‌پشتی با ظرفیت ۱ پوند است. گیتار نیز ۱ پوند است، یعنی در کوله‌پشتی جا می‌شود! بنابراین ارزش این سلول ۱۵۰۰ دلار و حاوی یک گیتار است. بیایید شبکه را پر کنیم.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار ج			
ضبط صوت				
لپ تاپ				

مانند این، هر سلول در شبکه شامل لیستی از تمام مواردی است که در آن مقطع زمانی در کوله‌پشتی جا می‌شوند. بیایید به سلول بعدی نگاه کنیم. در اینجا شما یک کوله‌پشتی با ظرفیت ۲ پوند دارید. خوب، گیتار قطعاً در آن جا می‌شود!

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار ج	۱۵۰۰ دلار ج		
ضبط صوت				
لپ تاپ				

برای بقیه‌ی سلول‌های این ردیف هم به همین شکل عمل می‌کنیم. به یاد داشته باشید، این ردیف اول است، بنابراین شما فقط گزینه‌ی گیتار را برای انتخاب دارید. این طور فرض می‌کنید که دو آیتم دیگر در حال حاضر برای سرقت در دسترس نیستند.

	۱	۲	۳	۴
گیتار	۱۵۰.. G	۱۵۰.. G	۱۵۰.. G	۱۵۰.. G
ضبط صوت				
لپ تاپ				

در این مرحله، احتمالاً گیج شده‌اید. چرا این کار را برای کوله‌پشتی‌هایی با ظرفیت ۱ پوند، ۲ پوند و غیره انجام می‌دهید، در حالی‌که مسئله از یک کوله‌پشتی ۴ پوندی صحبت می‌کند؟ به یاد دارید که چگونه به شما گفتم که برنامه‌نویسی پویا با یک مسئله‌ی کوچک شروع می‌شود و به مسئله‌ای بزرگ تبدیل می‌شود؟ شما در اینجا مسئله‌هایی فرعی را حل می‌کنید که به شما در حل مسئله‌ی بزرگ کمک می‌کند. ادامه دهید، در ادامه همه چیز واضح‌تر می‌شود.

در این مرحله، شبکه‌ی شما باید به این شکل باشد.

	۱	۲	۳	۴
گیتار	۱۵۰.. G	۱۵۰.. G	۱۵۰.. G	۱۵۰.. G
ضبط صوت				
لپ تاپ				

به یاد داشته باشید که هدف شما به حداکثر رساندن ارزش کوله پشتی است. این ردیف بهترین حدس فعلی برای ارزش حداکثری را نشان می‌دهد. بنابراین در حال حاضر، طبق این ردیف، اگر یک کوله‌پشتی با ظرفیت ۴ پوند داشته باشید، حداکثر ارزش کالا ۱۵۰۰ دلار خواهد بود.

۱	۲	۳	۴
گیتار 1500	1500	1500	1500
ضبط صوت			
لپ تاپ			

بهترین گزینه‌ی فعلی
برای سرقت:
گیتار ۱۵۰۰ دلاری

می‌دانید که این راه حل نهایی نیست. همانطور که الگوریتم را مرور می‌کنیم، تخمین خود را بهبود می‌بخشید.

ردیفِ ضبط صوت

بیایید برویم سراغ ردیف بعدی. این ردیف برای ضبط صوت است. اکنون که در ردیف دوم هستیم، می‌توانید ضبط صوت یا گیتار را سرقت کنید. در هر ردیف، می‌توانید آیتم آن ردیف یا آیتم‌های ردیف‌های بالاتر را بدزدید. بنابراین در حال حاضر امکان سرقت لپ‌تاپ وجود ندارد، اما می‌توانید ضبط صوت و/یا گیتار را بدزدید. بیایید با سلول اول شروع کنیم، یک کوله‌پشتی با ظرفیت ۱ پوند. حداکثر مقداری که در این مقطع می‌توانید در یک کوله‌پشتی ۱ پوندی قرار بدهید، ۱۵۰۰ دلار است.

۱	۲	۳	۴
گیتار 1500	1500	1500	1500
ضبط صوت	1500	1500	1500
لپ تاپ		1500	1500

ماکریسم فعلی برای
کوله‌پشتی یک پوندی

ماکریسم جدید برای یک
کوله‌پشتی ۱ پوندی

آیا باید این ضبط صوت را بذدید یا نه؟

شما یک کوله‌پشتی با ظرفیت ۱ پوند دارید. آیا ضبط صوت در آن جا می‌شود؟ نه، خیلی بزرگ است! از آنجایی که ضبط صوت در کوله‌پشتی جانمی‌شود، ۱۵۰۰ دلار حداکثر مقدار برای یک کوله‌پشتی ۱ پوندی است.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G			
لپ تاپ				

برای دو سلوی بعدی هم به همین شکل عمل می‌کیم. این کوله‌پشتی‌ها دارای ظرفیت ۲ پوند و ۳ پوند هستند. حداکثر مقدار قبلی برای هر دو ۱۵۰۰ دلار بود.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	
لپ تاپ				

همچنان ضبط صوت جانمی‌شود، بنابراین فرض شما بدون تغییر باقی می‌ماند. اگر یک کوله‌پشتی با ظرفیت ۴ پوند داشته باشد چه؟ خب: بالاخره ضبط صوت جا می‌شود! حداکثر مقدار قبلی ۱۵۰۰ دلار بود، اما اگر در عوض ضبط صوت را در آن قرار بدهید، ارزش آن ۳۰۰۰ دلار می‌شود! باید ضبط صوت را انتخاب کنیم.

	۱	۲	۳	۴
گیتار	\$1500 G	\$1500 G	\$1500 G	\$1500 G
ضبط صوت	\$1500 G	\$1500 G	\$1500 G	\$3000 115
لپ تاپ				

برآورد خود را به روز کردید! اگر یک کوله پشتی ۴ پوندی دارید، می‌توانید حداقل ۳۰۰۰ دلار کالا در آن قرار بدهید. با نگاه به شبکه متوجه می‌شوید که در حال به روزرسانی تدریجی تخمین خود هستید.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار 115
لپ تاپ				

← برآورد قبلی
← برآورد جدید
← راه حل نهایی

ردیف لپ تاپ

بیایید همین کار را با لپ تاپ انجام بدهیم! وزن این لپ تاپ ۳ پوند است، بنابراین در یک کوله پشتی ۱ پوندی یا ۲ پوندی جای نمی‌گیرد. تخمین دو سلوول اول ۱۵۰۰ دلار است.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار 115
لپ تاپ				

برای ۳ پوند، تخمین قدیمی ۱۵۰۰ دلار بود. اما می‌توانید به جای آن لپتاپ را انتخاب کنید، و ارزش آن ۲۰۰۰ دلار است. بنابراین حداکثر برآورد جدید ۲۰۰۰ دلار است!

۱	۲	۳	۴
کیتار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
لبتاب L	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار L

برای ۴ پوند، اوضاع جالب می‌شود. این بخش مهم است. برآورد فعلی ۳۰۰۰ دلار است. شما می‌توانید لپتاپ را در کوله‌پشتی قرار بدهید، اما ارزش آن فقط ۲۰۰۰ دلار می‌شود.

لیتاتیپ	ضبط صوت	یا	۲۰۰۰ دلار
---------	---------	----	-----------

خب، نتیجه به خوبی تخمین قدیمی نیست. اما صبر کنید! وزن لپ‌تاپ فقط ۳ پوند است، بنابراین همچنان ۱ پوند فضای خالی دارید! می‌توانید چیزی در این ۱ پوند قرار گیرد.

یک پوند از فضای
خالی، $\frac{?}{300\dots} + \frac{?}{200\dots}$ دلار لپ تاپ یا ضبط صوت

حداکثر مقداری که می‌توانید در ۱ پوند فضای قرار بدهید چقدر است؟ خب، در طی این مدت همواره آن را محاسبه کرده‌اید.

۱	۲	۳	۴
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار G
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۲۰۰۰ دلار L	

→ مأکریم ارزش برای ۱ پوند

طبق بهترین تخمین آخر، می‌توانید گیتار ۱۵۰۰ دلاری را در فضای ۱ پوندی قرار بدهید. بنابراین مقایسه‌ی واقعی به شرح زیر است.

$$\begin{matrix} 1500 \text{ دلار} \\ \text{گیتار} \\ \text{ضبط صوت} \end{matrix} + \begin{matrix} 2000 \text{ دلار} \\ \text{لپ تاپ} \end{matrix} \quad \text{یا} \quad \left(\begin{matrix} 3000 \text{ دلار} \end{matrix} \right)$$

شاید با خودتان فکر می‌کردید که چرا حداکثر مقدار را برای کوله پشتی‌های کوچک‌تر محاسبه می‌کنید. حالا دیگر امیدوارم قابل فهم باشد! هنگامی که فضای خالی دارید، می‌توانید از پاسخ به آن مسئله‌های فرعی استفاده کنید تا بفهمید چه چیزی در آن فضا قرار می‌گیرد. بهتر است لپ تاپ + گیتار را با ارزش ۳۵۰۰ دلار انتخاب کنید.

شبکه‌ی نهایی به این صورت است.

۱	۲	۳	۴
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار G
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۲۰۰۰ دلار L	۳۵۰۰ دلار L G

↑ پاسخ!

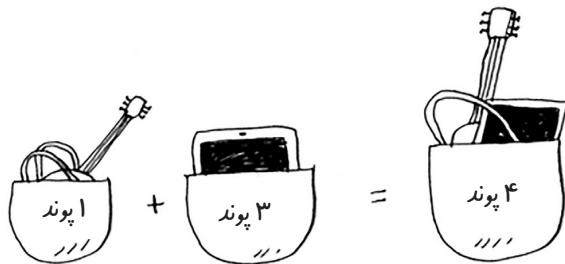
گیتار
ضبط صوت
لپ تاپ

پاسخ: حداکثر ارزشی که در کوله پشتی قرار می‌گیرد ۳۵۰۰ دلار است شامل یک گیتار و یک لپتاپ!

شاید فکر کنید که از فرمول دیگری برای محاسبهٔ مقدار آخرین سلول استفاده کردم. به این دلیل که هنگام پرکردن مقادیر سلول‌های قبلی از پیچیدگی‌های غیرضروری صرف نظر کردم. مقدار هر سلول با فرمولی یکسان محاسبه می‌شود:

$$\text{وزن آیتم} - \left(\begin{array}{l} \text{ارزش آیتم فعلی} \\ + \text{ارزش باقیمانده} \\ \hline \end{array} \right) = \text{ماکریسم} = \left\{ \begin{array}{l} \text{ماکریسم قبلی. ۱.} \\ \text{یا} \\ \text{ستون ردیف} \end{array} \right.$$

شما می‌توانید از این فرمول برای هر سلول این شبکه استفاده کنید، و باید در نهایت شبکه‌ی شما برابر باشد با همان شبکه‌ای که من به آن رسیدم. به یاد دارید که چگونه در مورد حل زیر مسئله‌ها صحبت کردم؟ شما راه حل‌های دو زیرمسئله را برای حل مشکل بزرگ‌تر ترکیب کردید.



پرسش‌های متناول در مسئله‌ی کوله پشتی

چون ممکن است همچنان ابهاماتی در مسئله‌ی کوله‌پشتی باقی مانده باشد، در این بخش به برخی از سؤالات رایج پاسخ داده می‌شود.



با اضافه‌کردن یک آیتم دیگر چه اتفاقی می‌افتد؟

فرض کنید متوجهه آیتم قابل سرقت چهارمی بشوید که قبل‌ب‌ه آن توجه نکردید! مثلاً می‌توانید یک آیفون هم بذدید.

آیا برای حساب این آیتم جدید باید همه چیز را از اول حساب کنید؟ جواب منفی است. به یاد داشته باشید، برنامه‌نویسی پویا به تدریج و مستمر بر اساس برآورد شما ساخته می‌شود. تا اینجا، این آیتم‌ها حداقل ارزش را دارند:

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
لپ‌تاپ	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۲۳۰۰ دلار L	۳۵۰۰ دلار LG

يعنى برای یک کوله‌پشتی ۴ پوندی، می‌توانید کالاهایی به ارزش ۳۵۰۰ دلار را سرقت کنید. فرض بر این بود که این بیشترین ارزش نهایی است. حالا بیایید یک ردیف برای آیفون اضافه کنید.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
لپ‌تاپ	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۲۳۰۰ دلار L	۳۵۰۰ دلار LG
آیفون				

پاسخ جدید

شما یک حداکثر ارزش جدید دارید! سعی کنید سلول‌های این ردیف جدید را پیش از خواندن باقی کتاب پر کنید.

باید با سلول اول شروع کنیم. آیفون در یک کوله پشتی ۱ پوندی قرار می‌گیرد. پیش از این حداکثر قیمت ۱۵۰۰ دلار بود، اما آیفون ۲۰۰۰ دلار ارزش دارد. در عوض آیفون را انتخاب می‌کنیم.

	۱	۲	۳	۴
گیتار	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G
ضبط صوت	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار S
لپ تاپ	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۲۰۰۰ دلار L	۳۵۰۰ دلار LG
آیفون	۲۰۰۰ دلار I			

در سلول بعدی می‌توانید آیفون و گیتار را جای بدھید.

	۱	۲	۳	۴
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۱۵۰۰ دلار G	۳۰۰۰ دلار S	
۱۵۰۰ دلار G	۱۵۰۰ دلار G	۲۰۰۰ دلار L	۳۵۰۰ دلار LG	
۲۰۰۰ دلار I	۳۵۰۰ دلار IG			

برای سلول ۳، گزینه‌ی بهتری از انتخاب دوباره‌ی آیفون و گیتار ندارید، پس در این مرحله کاری نمی‌کنید.

برای سلول آخر، وضعیت جالب می‌شود. حداکثر مقدار فعلی ۳۵۰۰ دلار است. به جای آن می‌توانید آیفون را بذرزدید و ۳ پوند فضای باقیمانده در اختیار داشته باشید.

$$(\text{لپتاپ + گیتار آیفون} \rightarrow \text{آیفون} + \text{لپتاپ}) \rightarrow \text{آیفون} \rightarrow \text{آیفون}$$

آن ۳ پوند ارزش دارد! ۲۰۰۰ دلار از آیفون + ۲۰۰۰ دلار از زیرمسئلهٔ قبلی: ۴۰۰۰ دلار می‌شود. یک حداکثر جدید!

نسخهٔ جدید شبکهٔ نهایی به شکل زیر است:

G	G	G	G
1۵۰۰ دلار	1۵۰۰ دلار	1۵۰۰ دلار	1۵۰۰ دلار
G	G	G	S
1۵۰۰ دلار	1۵۰۰ دلار	2۰۰۰ دلار	3۵۰۰ دلار
I	IG	IG	IL
3۵۰۰ دلار	3۵۰۰ دلار	3۵۰۰ دلار	4۰۰۰ دلار

↑
پاسخ جدید

پرسش: هیچ‌گاه ارزش یک ستون پایین هم می‌آید؟

ماکریم ارزش
نزولی است.

	1۵۰۰ دلار	1۵۰۰ دلار	1۵۰۰ دلار	1۵۰۰ دلار
Ø	Ø	Ø	Ø	3۰۰۰ دلار

لحظه‌ای تأمل کنید و به پاسخ این پرسش فکر کنید.

پاسخ: خیر. در هر تکرار، حداکثر برآورد فعلی را ذخیره می‌کنید. هرگز تخمین بدتر از قبل نمی‌شود!

تمرین

۹.۱ فرض کنید می‌توانید یک آیتم دیگر را سرقت کنید: پخش‌کننده‌ی MP^۳. وزن آن ۱ پوند است و ۱۰۰۰ دلار ارزش دارد. آیا باید آن را بدزدید؟

چه اتفاقی می‌افتد اگر ترتیب ردیف‌ها را تغییر بدھید؟

آیا تغییری در پاسخ ایجاد می‌شود؟ فرض کنید ردیف‌ها را به این ترتیب پر می‌کنید: ضبط صوت، لپ‌تاپ، گیتار. شبکه‌ی آن به چه شکل است؟ قبل از خواندن ادامه‌ی کتاب، شبکه را برای خود پر کنید.
شبکه به این شکل است:

	۱	۲	۳	۴
ضبط صوت	∅	∅	∅	۳۰۰... دلار
لپ‌تاپ	∅	∅	۲۰۰... دلار	۳۰۰... دلار
گیتار	۱۵۰۰... دلار	۱۵۰۰... دلار	۲۰۰... دلار	۳۵۰۰... دلار

پاسخ تغییر نمی‌کند. ترتیب ردیف‌ها اهمیتی ندارد.

آیا می‌توان شبکه را به جای ردیفی، ستونی پر کرد؟

خودتان آن را امتحان کنید! در این مسئله فرقی نمی‌کند. اما می‌تواند در مسئله‌های دیگر تفاوت ایجاد کند.

چه اتفاقی می‌افتد اگر یک آیتم کوچک‌تر اضافه کنید؟

فرض کنید بشود یک گردن بند دزدید. وزن آن ۵.۰ پوند است و ارزش آن ۱۰۰۰ دلار است. تا اینجا، شبکه‌ی شما فرض می‌کند که همه‌ی وزن‌ها اعداد صحیح هستند. حالا تصمیم می‌گیرید گردن بند را بدزدید. ۳.۵ پوند برای شما باقی مانده است. حداکثر مقداری که می‌توانید در ۳.۵ پوند قرار بدهید چقدر است؟ نمی‌دانید! شما فقط مقادیر را برای کوله‌های ۱ پوندی، ۲ پوندی، ۳ پوندی و ۴ پوندی محاسبه کرده‌اید. شما باید ارزش یک کوله پشتی ۳.۵ پوندی را بدانید.

برای گردن بند، شما باید دانه‌بندی ریزتری را در نظر بگیرید. بنابراین شبکه باید تغییر کند.

	۰.۵	۱	۱.۵	۲	۲.۵	۳	۳.۵	۴
گیتار								
ضبط صوت								
لپ تاپ								
جوهرات								

آیا می‌توان کسری از یک کالا را سرقت کرد؟

فرض کنید شما یک سارق هستید و در یک فروشگاه مواد غذایی ایستاده‌اید. می‌توانید کیسه‌های عدس و برنج را بدزدید. اگر یک کیسه‌ی کامل جا نشود، می‌توانید آن را باز کنید و تا جایی که می‌توانید بردارید. بنابراین دیگر همه‌ی یا هیچ نیست - می‌توانید کسری از یک آیتم را بدارید. چگونه با استفاده از برنامه‌نویسی پویا این موضوع را مدیریت می‌کنید؟

پاسخ: نمی‌توانید. در راه حل برنامه‌نویسی پویا، یا آیتم را برمی‌دارید یا خیر. انتخاب نیمی از آیتم جزو گزینه‌های برنامه‌نویسی پویا نیست.

اما این مورد نیز به راحتی با استفاده از یک الگوریتم حریصانه حل می‌شود! اول، تا جایی که می‌توانید از با ارزش‌ترین چیز بدارید. آن آیتم که تمام شد، تا آنجا که می‌توانید از با ارزش‌ترین کالای بعدی بدارید و غیره.

برای مثال، فرض کنید این آیتم‌ها را برای انتخاب دارید.

	کینوا	\$6/۱۶
	دال	\$3/۱۶
	برنج	\$2/۱۶



هر پوند کینوا با ارزش‌تر از چیزهای دیگر است. بنابراین، تمام کینوا‌ای را که امکان حمل آن را دارید، بردارید! چه بهتر اگر تمام کوله‌پشتی شما با همین کینوا پر شود.

اگر کینوا تمام شد و کوله‌پشتی پرنشد، کالای با ارزش بعدی را بدارید و غیره.

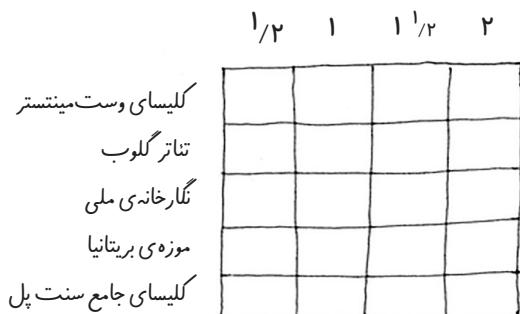
بهینه‌سازی برنامه‌ی سفر

فرض کنید برای تعطیلات فرح‌بخش به لندن می‌روید. دو روز فرصت دارید و کارهای زیادی برای انجام هست. اما فرصت انجام همه‌ی آن کارها را ندارید، بنابراین از آن‌ها یک لیست تهیه می‌کنید.

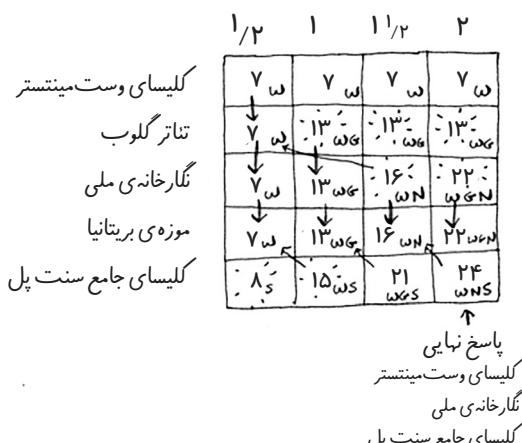
امتیاز	زمان	جاده
۷	۱/۲ روز	کلیسای وست‌مینستر
۶	۱/۲ روز	تناتر گلوب
۹	۱ روز	گلارخانه‌ی ملی
۹	۲ روز	گلارخانه‌ی ملی
۸	۱/۲ روز	کلیسای جامع سنت پل

برای هر یک از آن‌ها، زمان مورد نیاز برای بازدید از آن محل و میزان اشتیاقتان برای دیدن آن را یادداشت کنید. می‌توانید بر اساس این لیست درباره‌ی برنامه‌ی خود تصمیم بگیرید؟

همان مسئله‌ی کوله‌پشتی است! با این تفاوت که به جای کوله‌پشتی، زمان محدود است و عوض ضبط صوت ولپ تاپ، فهرست مکان‌هایی که می‌خواهید سر بزنید وجود دارد. قبل از حرکت، شبکه‌ی برنامه‌نویسی پویا را برای این لیست ترسیم کنید. شبکه به این شکل خواهد بود.



درست متوجه شدید؟ شبکه را پر کنید. در نهایت چه مکان‌هایی را باید ببینید؟ پاسخ آن به این شکل است:



بررسی آیتم‌های وابسته به یکدیگر

فرض کنید می‌خواهید به پاریس بروید، بنابراین چند آیتم را به لیست اضافه می‌کنید.

برج ایفل	۱ ۱/۲ روز	۸
لور	۱ ۱/۲ روز	۹
نوتدام	۱ ۱/۲ روز	۷

برای بازدید از این مکان‌ها به زمان زیادی نیاز است، زیرا ابتدا باید از لندن به پاریس سفر کنید که نصف روز طول می‌کشد. اگر بخواهید هر سه آیتم را در برنامه‌ی خود داشته باشید، چهار روز و نیم طول می‌کشد.

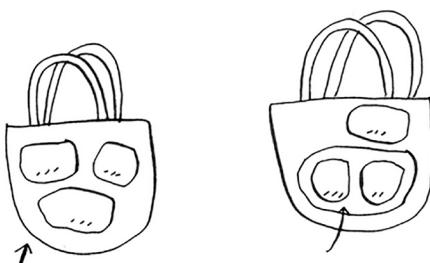
صبر کنید، به این صورت نیست. لازم نیست برای هر مورد به پاریس سفر کنید. هنگامی که در پاریس هستید، برای هر کدام تنها یک روز نیاز است. پس باید یک روز برای هر آیتم + نصف روز سفر = سه و نیم روز باشد نه چهار و نیم روز.

اگر برج ایفل را در کوله‌پشتی خود قرار بدهید، مورد «ارزان‌تر» موزه لور خواهد بود - به جای یک و نیم روز فقط یک روز برای شما هزینه خواهد داشت. چگونه این را در برنامه نویسی پویا مدل‌سازی می‌کنید؟

چنین چیزی امکان‌پذیر نیست. برنامه نویسی پویا قدرتمند است زیرا می‌تواند زیرمسئله‌ها را حل کند و از پاسخ آن‌ها برای حل مسئله‌ی بزرگ استفاده کند. برنامه نویسی پویا فقط زمانی به کار می‌آید که هر زیرمسئله ناپیوسته باشد - زمانی که به سایر زیرمسئله‌ها وابسته نباشد. به این معنا که هیچ راهی برای محاسبه‌ی مسئله‌ی پاریس با استفاده از الگوریتم برنامه نویسی پویا وجود ندارد.

آیا ممکن است که این راه حل نیاز به بیش از دو زیر-کوله‌پشتی داشته باشد؟

بهترین راه حل می‌تواند سرقت بیش از دو آیتم باشد. طبق این الگوریتم، حداکثر دو کوله‌پشتی را با هم ترکیب می‌کنید - هرگز بیشتر از دو کوله‌پشتی نخواهید داشت. اما این زیر-کوله‌پشتی‌ها می‌توانند زیر-کوله‌پشتی‌های خودشان را داشته باشند.



داشتن سه زیر-کولپشن‌های امکان‌پذیر نیست.

اما می‌توان زیر-کولپشن‌هایی داشت که زیر-کولپشن‌های خود را داشته باشند.

آیا ممکن است با بهترین راه حل، کوله‌پشتی به طور کامل پر نشود؟



بله. فرض کنید می‌توانید یک الماس را هم بذدید.

این الماس بزرگ است: وزن آن 5.3 پوند است و یک میلیون دلار قیمت دارد، بسیار باارزش‌تر از هر چیز دیگری. حتماً باید آن را بذدید! با سرقت آن نیم پوند فضای باقی می‌ماند و هیچ چیز در آن فضای باقی‌مانده جانمی‌شود.

تمرین

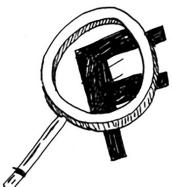
۹.۲ فرض کنید می‌خواهید به پیک‌نیک بروید. یک کوله‌پشتی دارید که 6 پوند را در خود جای می‌دهد و می‌توانید آیتم‌های زیر را بردارید. به هر کدام یک عدد اختصاص داده شده و هر چه مقدار آن بیشتر باشد، آیتم اهمیت بیشتری دارد:

- آب، 3 پوند، 15
- کتاب، 1 پوند، 3
- غذا، 2 پوند، 9
- ژاکت، 2 پوند، 5
- دوربین، 1 پوند، 6

مجموعه‌ی بهینه‌ای از آیتم‌هایی که باید در پیک‌نیک همراه خود داشته باشید شامل چه چیزهایی می‌شود؟

طولانی‌ترین زیرشته‌ی مشترک

تا به اینجا، با یک مسئله‌ی برنامه‌نویسی پویا مواجه شده‌اید. از آن چه نتیجه‌های می‌توان گرفت؟



- برنامه‌نویسی پویا زمانی مفید است که بخواهید چیزی را با توجه به یک محدودیت بهینه کنید. در مسئله‌ی کوله‌پشتی، شما باید ارزش کالاهایی را که دزدیده‌اید، با توجه به اندازه‌ی کوله‌پشتی به حداقل برسانید.
- کاربرد برنامه‌نویسی پویا در مواردی است که مسئله به زیرمسئله‌های مجزا تقسیم بشود و به یکدیگر وابسته نباشند.

دستیابی به یک راه حل از نوع برنامه‌نویسی پویا می‌تواند دشوار باشد. در این بخش بر روی این موضوع تمرکز می‌کنیم. تعدادی از نکات کلی آن به این شرح است:

- هر راه حل برنامه‌نویسی پویا شامل یک شبکه است.
- مقدارهای موجود در سلول‌ها معمولاً همان چیزی است که می‌خواهید بهینه‌سازی کنید. برای مسئله‌ی کوله‌پشتی، مقدارها ارزش اجناس بود.
- هر سلول یک مسئله‌ی فرعی است، پس به این فکر کنید که چگونه می‌توانید مسئله‌ی خود را به چند زیرمسئله تقسیم کنید. این روش به شما کمک می‌کند تا محورها را انتخاب کنید.



بیایید به مثال دیگری نگاه کنیم. فرض کنید dictionary.com را اداره می‌کنید. کاربری یک کلمه تایپ می‌کند و شما معنی آن کلمه را می‌نویسید. اما می‌خواهید که اگر کاربر کلمه را اشتباه نوشت، بتوانید

کلمه‌ی مورد نظر او را حدس بزنید. الکس در حال جست‌وجوی *fish* است، اما به طور تصادفی *hish* را تایپ کرده است. هرچند این کلمه در دیکشنری شما لیست نشده است، اما لیستی از کلمات مشابه دارید.

"HISH" مشابه

• FISH

• VISTA

(این یک مثال بسیار ساده است، بنابراین شما لیست خود را به دو کلمه محدود می‌کنید. در واقعیت، این لیست احتمالاً هزاران کلمه خواهد بود.)
کلمه *hish* را تایپ کرد. منظور الکس تایپ کدام کلمه بود: *fish* یا *vista*؟

ساخت شبکه

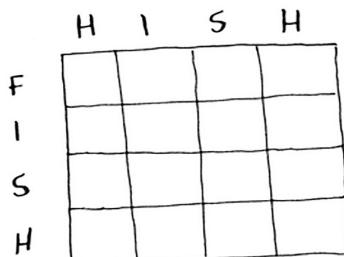
شبکه برای این مسئله چگونه است؟ شما باید به این پرسش‌ها پاسخ بدهید:

- مقدار هر سلول چقدر است؟
- چگونه این مسئله را به مسئله‌های فرعی تقسیم می‌کنید؟
- محورهای شبکه کدام هستند؟

در برنامه‌نویسی پویا، تلاش شما در به حداقل رساندن چیزی است. در این مورد، شما در تلاش برای یافتن طولانی‌ترین زیرشته‌ای هستید که دو کلمه‌ی مشترک دارند. *fish* و *vista* چه زیرشته‌ی مشترکی دارند؟ *hish* و *fish* چطور؟ این همان چیزی است که شما می‌خواهید آن را حساب کنید.

به یاد داشته باشید، مقدار سلول‌ها معمولاً همان چیزی است که می‌خواهید بهینه‌سازی کنید. در این مورد، مقادیر احتمالاً یک عدد خواهد بود: طول طولانی‌ترین زیرشته که در این دو رشته مشترک هستند.

چگونه این مسئله را به چند زیرمسئله تقسیم می‌کنید؟ می‌توانید زیرشته‌ها را با هم مقایسه کنید. به جای مقایسه‌ی *fish* و *hish* ابتدا می‌توانید *his* و *fis* را مقایسه کنید. هر سلول دارای طول طولانی‌ترین زیرشته‌ای است که دو زیرشته‌ی مشترک دارند. به شما این سرنخ را می‌دهد که احتمالاً محورها دو کلمه هستند. بنابراین احتمالاً شبکه چیزی شبیه به این باشد:



اگر از نظر شما این یک جادوچنبل است، نگران نباشید. موضوع دشواری است - به همین دلیل است که آموزش این موضوع را در بخش‌های آخر این کتاب در نظر گرفته‌ام! بعداً یک تمرین به شما می‌دهم تا برنامه‌نویسی پویا را تمرین کنید.

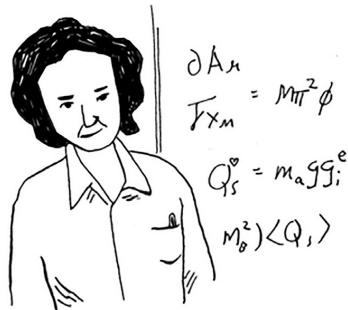
پر کردن شبکه

حالا درک درستی دارید از اینکه یک شبکه باید چگونه باشد. فرمول پر کردن هر سلول از شبکه چیست؟ می‌توانید کمی تقلب کنید، زیرا راه حل آن را می‌دانید - *hish* و *fish* یک زیرشته به طول مشترک ۳ دارند: *.ish*.

اما همچنان چیزی از فرمول قابل استفاده در اختیار شما نمی‌گذارد. دانشمندان علوم کامپیوتر گاهی درباره‌ی استفاده از الگوریتم فاینمن^۱ شوخی می‌کنند. نام الگوریتم فاینمن از فیزیک‌دان معروف ریچارد فاینمن گرفته شده و به این صورت است:

۱. مسئله را یادداشت کنید.
۲. به سختی فکر کنید.
۳. راه حل را یادداشت کنید.

1. Feynman algorithm



دانشمندان علوم کامپیوتر یک گوله نمک هستند!

حقیقت این است که در اینجا هیچ راه آسانی برای محاسبه فرمول وجود ندارد. شما باید آزمایش کنید و سعی کنید چیزی را پیدا کنید که کارساز باشد. گاهی اوقات الگوریتم‌ها دستورالعمل‌هایی دقیق نیستند. آن‌ها چارچوبی هستند که ایده‌ی خود را روی آن می‌سازید.

سعی کنید خودتان راه حلی برای این مسئله پیدا کنید. به شما سرنخی می‌دهم - بخشی از شبکه به این شکل است.

	H	I	S	H
F	•	•		
I				
S			2	•
H				3

مقادیر دیگر چیست؟ به یاد داشته باشید که هر سلول مقدار یک زیرمسئله است. چرا سلول (۳، ۳) دارای مقدار ۲ است؟ چرا سلول (۳، ۴) مقدار ۰ دارد؟ بعد از اینکه سعی کردید به فرمولی دست پیدا کنید، ادامه‌ی مطلب را بخوانید. حتی اگر به جواب درست هم نرسیدید، تلاش شما باعث می‌شود تا توضیح من برای شما بسیار قابل فهم‌تر باشد.

راه حل

شبکه‌ی نهایی به این صورت است.

	H	I	S	H
F
I	.	1	.	.
S	.	.	2	.
H	1	.	.	3

فرمول مورد نظر من برای پر کردن هر سلول به این شکل است.

۱. اگر حروف باهم منطبق باشند، مقدار صفر خواهد شد.

	H	I	S	H
F
I	.	1	.	.
S	.	.	2	.
H	1	.	.	3

۲. اگر حروف منطبق باشند، ارزش آن برابر است با ارزش همسایه + 1 + گوشی بالای چپ

شبه‌کد این فرمول به این صورت است:

```

if word_a[i] == word_b[j]: <----- حروف باهم سازگارند.
    cell[i][j] = cell[i-1][j-1] + 1
else: <----- حروف باهم سازگار نیستند.
    cell[i][j] = 0
  
```

در این جدول *hish* در برابر *vista* درج شده است.

	V	I	S	T	A
H
I	.	1	.	.	.
S	.	.	2	.	.
H

↑
پاسخ
نهاي
نهاي
نيست

توجه داشته باشید که در این مسئله، راه حل نهايی ممکن است در آخرین سلول نباشد! برای مسئله‌ی کوله پشتی، همیشه راه حل نهايی در سلول آخر بود. اما برای طولانی‌ترین زیررشه‌ی مشترک، راه حل بزرگترین عدد در شبکه است - و ممکن است آخرین سلول نباشد.

بیایید به سؤال اصلی برگردیم: کدام رشته با *hish* اشتراک بیشتری دارد؟ و *fish* زیررشه‌ای از سه حرف مشترک و *hish* و *vista* زیررشه‌ای از دو حرف مشترک دارند.
احتمالاً منظور الکس *fish* بوده است.

بزرگ‌ترین زیردنباله‌ی مشترک^۱

فرض کنید الکس اشتباهی *fosh* را جست و جو کرده است. منظورش کدام کلمه است: *fish* یا *?fort*

F	O	S	H	
F	1	0	.	.
O	.	2	0	.
R	0	1	0	.
T	0	0	0	.

F	O	S	H	
F	1	0	.	.
I	.	0	0	.
S	0	0	1	0
H	0	0	0	2

بیایید آنها را با استفاده از فرمول بزرگ‌ترین زیردنباله‌ی مشترک مقایسه کنیم.
هر دو یکسان هستند: دو حرف!

1. Longest common subsequence

اما *fish* به *fosh* نزدیکتر است.

$$\begin{matrix} F & O & S & H \\ \downarrow & \downarrow & \downarrow & \\ F & I & S & H \end{matrix} = ۳$$

$$\begin{matrix} F & O & S & H \\ \downarrow & \downarrow & & \\ F & O & R & T \end{matrix} = ۲$$

شما در حال مقایسه‌ی بزرگ‌ترین زیرشته‌ی مشترک هستید، اما در واقع باید بزرگ‌ترین زیردنباله‌ی مشترک را مقایسه کنید: تعداد حروف در یک دنباله که در این دو کلمه مشترک است. چگونه بزرگ‌ترین زیردنباله‌ی مشترک را محاسبه می‌کنید؟

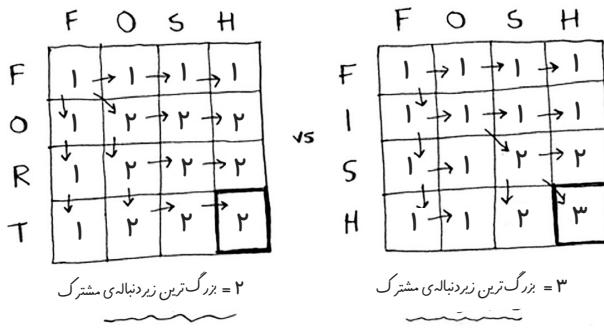
در اینجا بخشی از شبکه‌ی *fosh* و *fish* آمده است.

	F	O	S	H
F	1	1		
I	1			
S		1	۲	۲
H				

آیا می‌توانید فرمول این شبکه را پیدا کنید؟ طولانی‌ترین زیردنباله‌ی مشترک بسیار شبیه به طولانی‌ترین زیرشته‌ی مشترک است و فرمول آن‌ها نیز تقریباً مشابه هستند. پیش از آنکه به راه حل این مسئله بپردازم تلاش کنید آن را پیدا کنید.

بزرگ‌ترین زیردنباله‌ی مشترک - راه حل

شبکه‌ی نهايی به اين شكل است.



فرمول من برای پرکردن هر سلول :

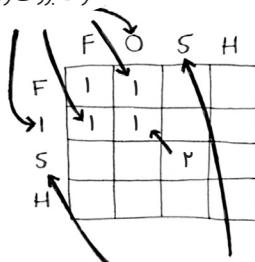
از همسایگان گوشی بالایی چپ

(ستواوت با بزرگ‌ترین)

۱- اگر حروف باهم برابر نبودند،

زیررشته‌ی مشترک)

حروف بزرگ‌تر را انتخاب کنید



۲- اگر باهم متنطبق باشند، مقدار آن برابر با همسایه‌ی

گوشی بالایی چپ به علاوه‌ی یک خواهد بود.

و اين شبه کد آن است:

```

if word_a[i] == word_b[j]: ←----- حروف برابر باشند.
    cell[i][j] = cell[i-1][j-1] + 1
else: ←----- حروف باهم برابر نباشد.
    cell[i][j] = max(cell[i-1][j], cell[i][j-1])
  
```

خب - موفق شدید! این قطعاً یکی از سخت‌ترین فصل‌های کتاب است. حالا واقعاً برنامه‌نویسی پویا کاربرد دارد؟ بله:

- زیست‌شناسان از بزرگ‌ترین زیردانبه‌ی مشترک برای یافتن تشابهات رشته‌های DNA استفاده می‌کنند. از این روش می‌توان برای تشخیص شباهت دو حیوان یا دو بیماری استفاده کرد. بزرگ‌ترین زیردانبه‌ی مشترک برای یافتن درمان بیماری ام اس استفاده می‌شود.
- آیا تا به حال از `diff` (مانند `git diff`) استفاده کرده‌اید؟ `Diff` تفاوت‌های میان دو فایل را به شما می‌گوید و برای انجام این کار از برنامه‌نویسی پویا استفاده می‌کند.
- در مورد تشابهات رشته‌ها صحبت کردیم. فاصله‌ی لون‌اشتاین^۱ میزان تشابهات دو رشته را اندازه‌گیری می‌کند و از برنامه‌نویسی پویا استفاده می‌کند. فاصله‌ی لون‌اشتاین در موارد متعددی کاربرد دارد، از بررسی درستی املاء گرفته تا فهمیدن اینکه آیا کاربران با داده‌هایی که بازگزاری می‌کنند کپی‌رایت را نقض می‌کنند یا خیر.
- آیا تا به حال از `word wrap` برنامه‌ای مانند `word` مایکروسافت استفاده کرده‌اید؟ چگونه می‌توان فهمید که کجا باید `wrap` بشود تا طول خط ثابت بماند؟ برنامه‌نویسی پویا!

تمرین

- ۹.۳** برای محاسبه‌ی بزرگ‌ترین زیردانبه‌ی مشترک، میان `blue` و `clues` شبکه را ترسیم و آن را پر کنید.

جمع‌بندی

- برنامه‌نویسی پویا زمانی مفید است که بخواهید با توجه به یک محدودیت چیزی را بهینه کنید.
- زمانی می‌توان از برنامه‌نویسی پویا استفاده کرد که مسئله به زیرمسئله‌های ناپیوسته^۱ تقسیم شود.
- برای هر راه حل برنامه‌نویسی پویا یک شبکه وجود دارد.
- مقادیر موجود در سلول‌ها معمولاً همان چیزی هستند که می‌خواهید بهینه‌سازی کنید.
- هر سلول یک مسئله‌ی فرعی است، پس به این فکر کنید که چگونه می‌توانید مسئله‌ی خود را به مسئله‌های فرعی تقسیم کنید.
- هیچ فرمول واحدی برای محاسبه‌ی راه حل یک برنامه‌نویسی پویا وجود ندارد.

۱۵ | الگوریتم K - نزدیک‌ترین همسایه^۱



در این فصل

• یاد می‌گیرید با استفاده از الگوریتم k نزدیک‌ترین همسایه، یک سیستم طبقه‌بندی بسازید.

• در مورد استخراج ویژگی^۲ یاد می‌گیرید.

• در مورد رگرسیون^۳ یاد می‌گیرید: پیش‌بینی یک عدد، مانند ارزش سهام در روز بعد، یا میزان رضایت کاربر از تماشای یک فیلم.

• موارد استفاده و محدودیت‌های الگوریتم K - نزدیک‌ترین همسایه را یاد می‌گیرید.

طبقه‌بندی پرتوال در مقابل طبقه‌بندی گریپ فروت



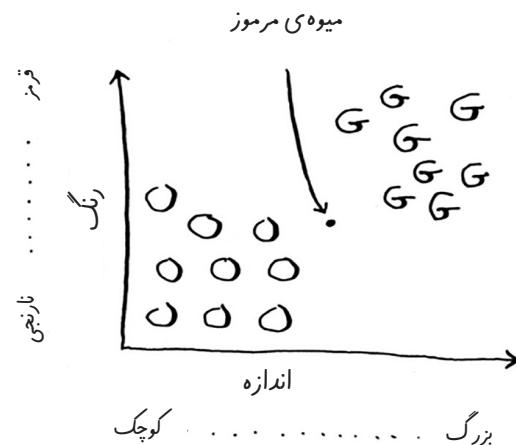
به این میوه نگاه کنید. پرتوال است یا گریپ فروت؟ خب، می‌دانیم که عموماً گریپ فروت‌ها بزرگ‌تر و قرمزتر هستند.

1. k-nearest neighbors
2. feature extraction
3. regression

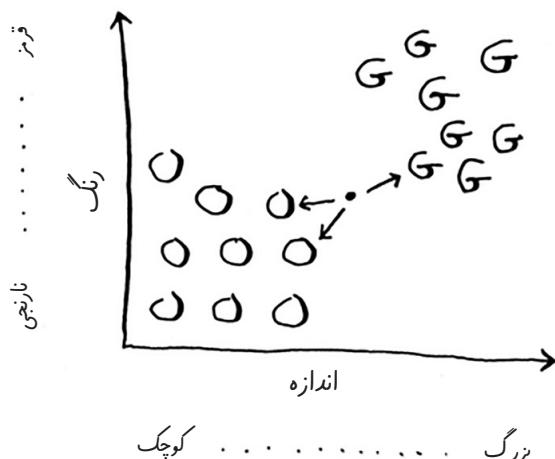
روند فکر کردن من چیزی شبیه به این است: یک گراف در ذهنم دارم.



عموماً، میوه‌های بزرگ‌تر و قرمزتر گریپ فروت هستند. این میوه هم بزرگ است و هم قرمز. بنابراین احتمالاً گریپ فروت باشد. اما اگر میوه‌ای مشابه داشته باشد چه؟

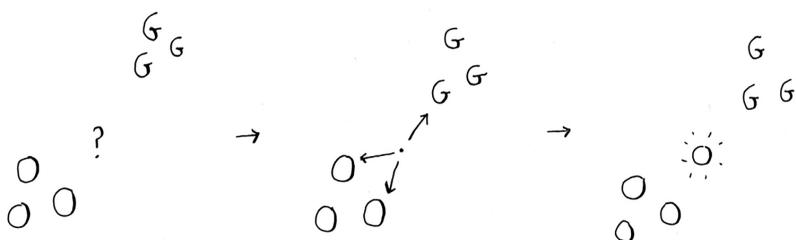


این میوه را چگونه طبقه‌بندی می‌کنید؟ یک راه این است که به همسایگان این نقطه نگاه کنید. به سه همسایه‌ی نزدیک‌تر این نقطه نگاهی بیندازید.



بزرگ کوچک

بیشتر همسایه‌ها پرتفال هستند. در نتیجه این میوه احتمالاً یک پرتفال است.
تبریک: شما الان از الگوریتم K - نزدیک‌ترین همسایه (KNN) برای طبقه‌بندی استفاده کردید! این الگوریتم بسیار ساده است.

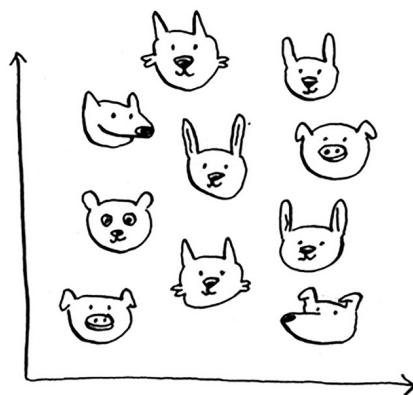


1. میوه‌ای جدید برای طبقه‌بندی در اختیار دارد.
2. به سه همسایه‌ی نزدیک آن نگاه می‌کنید
3. پرتفال‌های بیشتر در همسایگی، پس احتمالاً پرتفال باشد.

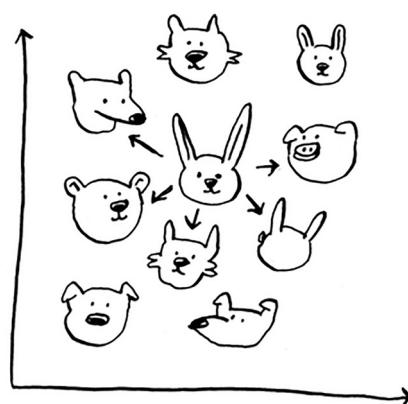
الگوریتم KNN ساده ولی مفید است! اگر چیزی را بخواهید طبقه‌بندی کنید، بد نیست ابتدا KNN را امتحان کنید. بیایید به یک مثال واقعی‌تر نگاه کنیم.

ایجاد سیستم پیشنهاد

فرض کنید نتفلیکس هستید و می‌خواهید یک سیستم پیشنهاد فیلم برای کاربران بسازید. در ظاهر، چیزی شبیه مسئله‌ی گریپ فروت است! شما می‌توانید هر کاربر را روی یک گراف رسم کنید.



ترسیم این کاربران بر اساس تشابهات است، بنابراین کاربرهایی با سلیقه‌ی مشابه در فاصله‌ی نزدیک تری به هم رسم می‌شوند. فرض کنید می‌خواهید فیلم‌هایی به پریانکا پیشنهاد کنید. پنج کاربر نزدیک به او را پیدا می‌کنید.



جاستین، جی‌سی، جوی، لنس و کریس همگی در فیلم‌ها سلیقه‌ای مشابه دارند. بنابراین اگر آن‌ها فیلمی را دوست داشته باشند، احتمالاً پریانکا هم از آن فیلم خوش خواهد آمد!

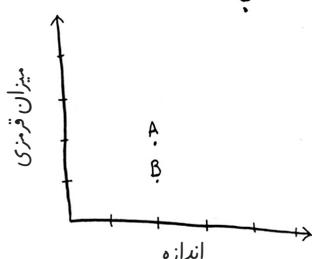
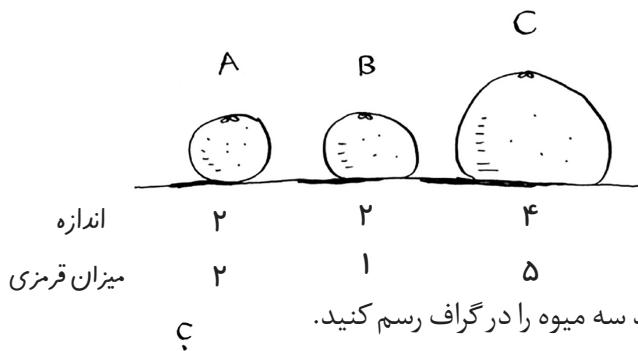
هنگامی که این گراف را دارید، ساختن یک سیستم پیشنهادات آسان است. اگر فیلمی هست که جاستین دوست داشته باشد، آن را به پریانکا توصیه می‌کنید.



اما هنوز یک بخش مهم از این روش جا مانده است. شما کاربران را بر اساس تشابهات ترسیم کردید. چگونه متوجه می‌شوید که دو کاربر چقدر شبیه هم هستند؟

استخراج ویژگی

در مثال گریپ‌فروت، میوه‌ها را بر اساس اندازه و میزان سرخی مقایسه کردید. اندازه و رنگ ویژگی‌هایی هستند که با هم مقایسه می‌کنید. حالا فرض کنید سه میوه دارید. می‌توانید ویژگی‌ها را استخراج کنید.



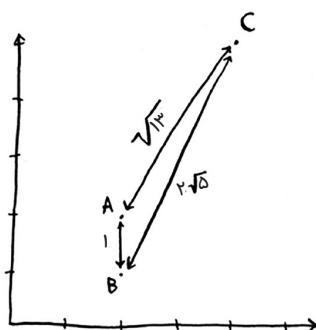
از گراف و به شکل تصویری می‌توان دریافت که میوه‌های A و B مشابه هستند. بیایید اندازه‌گیری کنیم که آن‌ها چقدر به هم نزدیک هستند. برای یافتن فاصله‌ی بین دو نقطه از فرمول فیثاغورث استفاده می‌کنیم.

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

برای مثال، فاصله‌ی بین A و B به صورت زیر حساب می‌شود:

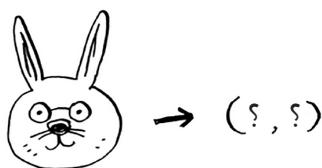
$$\begin{aligned} & \sqrt{(2-2)^2 + (2-1)^2} \\ & = \sqrt{0+1} \\ & = \sqrt{1} \\ & = 1 \end{aligned}$$

فاصله‌ی بین A و B یک است. شما می‌توانید بقیه‌ی فواصل را هم پیدا کنید.



فرمول فاصله تأییدی است برآنچه پیش از این و به شکل تصویری دیدید: میوه‌های A و B مشابه یکدیگر هستند.

فرض کنید در عوض مقایسه را میان کاربران نتفلیکس انجام بدھید. شما به راهی برای ترسیم گراف کاربران نیاز دارید. بنابراین، باید هر کاربر را مطابق آن چه برای میوه‌ها انجام دادید به مجموعه‌ای از مختصات تبدیل کنید.



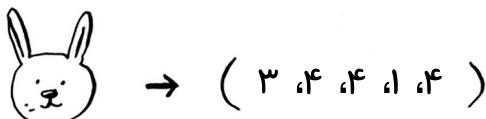
زمانی که بتوانید کاربران را در گراف در نظر بگیرید. می‌توانید فاصله‌ی میان آن‌ها را اندازه‌گیری کنید.

در اینجا شیوه‌ی تبدیل کاربران به مجموعه‌ای از اعداد آمده است. هنگامی که کاربران برای نتفلیکس ثبت‌نام می‌کنند، از آن‌ها خواسته می‌شود بر اساس میزان علاقه‌مندی به دسته‌بندی فیلم‌ها امتیاز بدهند. برای هر کاربر، مجموعه‌ای از رتبه‌بندی در اختیار دارید!

	پریانکا	جاستین	مرفوس
کمدی	۳	۴	۲
آشن	۴	۳	۵
درام	۴	۵	۱
ترسناک	۱	۱	۳
عاشقانه	۴	۵	۱

پریانکا و جاستین عاشق‌های عاشقانه هستند و از فیلم‌های ترسناک متنفرند. مورفوس فیلم اکشن دوست دارد و از فیلم‌های عاشقانه متنفر است (بیزار است از اینکه ببیند یک فیلم اکشن خوب با یک صحنه‌ی عاشقانه‌ی مسخره خراب می‌شود). به یاد

دارید در مسئله‌ی پرقال و گریپ فروت، هر میوه با مجموعه‌ای از دو عدد نشان داده می‌شد؟ در اینجا، هر کاربر با مجموعه‌ای از پنج عدد نشان داده می‌شود.



یک ریاضی‌دان خواهد گفت، به جای اینکه فاصله را در دو بعد محاسبه کنید، اکنون فاصله را در پنج بعد محاسبه می‌کنید. اما فرمول فاصله ثابت باقی می‌ماند.

$$\sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2 + (d_1 - d_2)^2 + (e_1 - e_2)^2}$$

تنها کافی است مجموعه‌ای از پنج عدد را به جای مجموعه‌ای از دو عدد در نظر بگیرید.

فرمول فاصله انعطاف‌پذیر است: شما می‌توانید مجموعه‌ای از یک میلیون عدد داشته باشید و همچنان از همان فرمول فاصله‌ی قدیمی برای پیداکردن فاصله استفاده کنید. شاید از خود بپرسید: «با داشتن پنج عدد فاصله به چه معناست؟» فاصله به شما می‌گوید که این مجموعه اعداد چقدر شبیه هم هستند.

$$\begin{aligned} & \sqrt{(3-4)^2 + (4-3)^2 + (4-5)^2 + (1-1)^2 + (4-5)^2} \\ &= \sqrt{1+1+1+0+1} \\ &= \sqrt{4} \\ &= 2 \end{aligned}$$

این عدد، فاصله‌ی پریانکا و جاستین است.

پریانکا و جاستین خیلی شبیه هم هستند. تفاوت پریانکا و مورفیوس چیست؟ قبل از ادامه‌ی خواندن کتاب مسافت را محاسبه کنید.
درست متوجه شدید؟ پریانکا و مورفیوس فاصله‌ای معادل ۲۴ دارند. فاصله به شما می‌گوید که سلیقه‌ی پریانکا بیشتر شبیه سلیقه‌ی جاستین است تا مورفیوس.

عالی! حالا پیشنهاد دادن فیلم به پریانکا آسان است: اگر جاستین فیلمی را دوست دارد، آن را به پریانکا توصیه کنید و برعکس. شما الان یک سیستم پیشنهاد فیلم ساختید!

اگر کاربر نتفلیکس هستید، نتفلیکس مدام از شما می‌خواهد: «لطفاً به فیلم‌های بیشتری امتیاز بدهید. هرچه به فیلم‌های بیشتری امتیاز بدهید، پیشنهادهایی که به شما داده می‌شود بهتر خواهد بود.» حالا دیگر می‌دانید چرا. هر چه تعداد فیلم‌های بیشتری را رتبه‌بندی کنید، نتفلیکس با دقت بیشتری می‌تواند مشاهده کند که شما به چه کاربران دیگری شبیه هستید.

تمرین

- ۱۰.۱** در مثال نتفلیکس، فاصله‌ی بین دو کاربر مختلف را با استفاده از فرمول فاصله محاسبه کردید. اما همه‌ی کاربران با یک معیار ثابت به فیلم‌ها امتیاز نمی‌دهند. فرض کنید دو کاربر با نام‌های یوگی و پینکی سلیقه‌ی یکسانی در فیلم دارند. یوگی به هر فیلمی که دوست داشته باشد ۵ امتیاز می‌دهد، در حالی که پینکی گزیده‌تر عمل می‌کند و امتیاز ۵ را فقط برای بهترین فیلم‌ها نگه می‌دارد. آن‌ها سلیقه‌ی مشابهی دارند، اما طبق الگوریتم فاصله، همسایه نیستند. چه راهکاری برای استراتژی متفاوت آن‌ها در رتبه‌بندی در نظر می‌گیرید؟

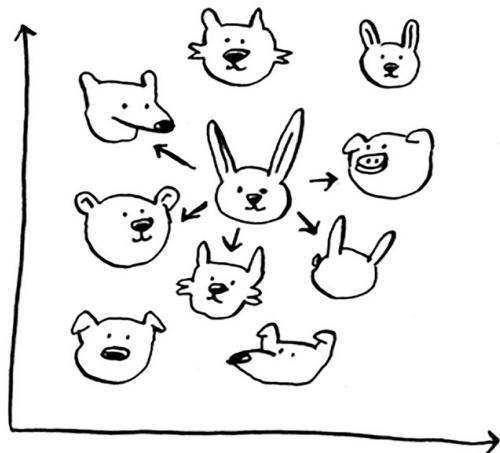
- ۱۰.۲** فرض کنید نتفلیکس گروهی از «اینفلوئنسرها» را گزینش کرده است. به عنوان

مثال، کوئنتین تارانتینو و وس اندرسون اینفلوئنسرهای نتفلیکس هستند، بنابراین رتبه‌بندی این افراد بیشتر از یک کاربر عادی حساب می‌شود. چه تغییراتی برای سیستم پیشنهادات نیاز است تا نسبت به رتبه‌بندی اینفلوئنسرها جانبدارانه عمل کند؟

رگرسیون

فرض کنید به دنبای چیزی بیش از صرفاً پیشنهاد فیلم باشید: می‌خواهید حدس بزنید که پریانکا به این فیلم چه امتیازی می‌دهد. پنج نفری که به هم نزدیک‌تر هستند را در نظر بگیرید.

راستی من مدام عدد ۵ را مثال می‌زنم. هیچ چیز خاصی در مورد عدد ۵ وجود ندارد:



شما می‌توانید ۲، یا ۱۰، یا ۱۰۰۰۰ نفر نزدیک‌تر را در نظر بگیرید. به همین دلیل است که الگوریتم k -نزدیک‌ترین همسایه نامیده می‌شود و نه پنج همسایه‌ی نزدیک‌تر! فرض کنید می‌خواهید رتبه‌بندی فیلم آواخوان حرفه‌ای را حدس بزنید. خوب، جاستین، جی‌سی، جوی، لنس و کریس چه امتیازی به آن دادند؟

جاستین	:	۵
جی‌سی	:	۴
جونی	:	۴
لنس	:	۵
کریس	:	۳

می‌توانید میانگین رتبه‌بندی آن‌ها را حساب کنید و ۰.۴ ستاره بدهید. به این کار رگرسیون می‌گویند. طبقه‌بندی و رگرسیون دو کار اساسی است که با KNN انجام می‌دهید:

• طبقه‌بندی = دسته‌بندی در یک گروه

• رگرسیون = پیش‌بینی پاسخ (مانند یک عدد)

رگرسیون بسیار کاربردی است. فرض کنید یک نانوایی کوچک در برکلی دارید و هر روز نان تازه پخت می‌کنید. شما می‌خواهید تخمین بزنید که برای امروز باید چه تعداد نان درست کنید. شما مجموعه‌ای از ویژگی‌ها را در اختیار دارید:

- آب و هوا در مقیاس ۱ تا ۵ (۱ = بد، ۵ = عالی).
- آخر هفته یا تعطیلات؟ (۱ = آخر هفته یا تعطیلات است، صفر = روزهای دیگر هفته).
- آیا مسابقه‌ای برگزار می‌شود؟ (۱ = بله، صفر = خیر).



واز میزان فروش نان در مجموعه ویژگی مختلف آگاه هستید.

$$\begin{array}{l}
 \boxed{A} (5, 1, 0) = 300 \quad \boxed{B} (3, 1, 1) = 225 \\
 \text{قرص نان} \qquad \qquad \qquad \text{قرص نان} \\
 \boxed{C} (1, 1, 0) = 75 \quad \boxed{D} (4, 0, 1) = 200 \\
 \text{قرص نان} \qquad \qquad \qquad \text{قرص نان} \\
 \boxed{E} (4, 0, 0) = 150 \quad \boxed{F} (2, 0, 0) = 50 \\
 \text{قرص نان} \qquad \qquad \qquad \text{قرص نان}
 \end{array}$$

امروز یک روز آخر هفته با هوای خوب است. بر اساس داده‌هایی که مشاهده کردید، چند قرص نان خواهید فروخت؟ بیایید از KNN استفاده کنیم، در شرایطی که $K = 4$. ابتدا، چهار همسایه‌ی نزدیکتر را برای این نقطه مشخص کنید.

$$\mathcal{C}_4 = \{A, B, C, D\}$$

D, B, A و E نزدیکترین همسایه‌ها هستند.

- A. ۱ ←
- B. ۲ ←
- C. ۹
- D. ۲ ←
- E. ۱ ←
- F. ۵

میانگین نان‌های فروخته شده در آن روزها ۲۱۸.۷۵ است. این تعداد نانی است که باید برای امروز پخت کنید!

شباخت کسینوسی^۱

تاکنون از فرمول فاصله برای مقایسه‌ی فاصله‌ی بین دو کاربر استفاده کرده‌اید. آیا این بهترین فرمول برای استفاده است؟ یکی از موارد رایج که در عمل مورد استفاده قرار می‌گیرد، شباخت کسینوسی است. فرض کنید دو کاربر شبیه هم هستند، اما یکی از آن‌ها در رتبه‌بندی محافظه‌کارتر است. هر دو عاشق فیلم آمر اکبر آنتونی از منموهان دسای هستند. پل به آن ۵ ستاره و روآن ۴ ستاره داده است. اگر کماکان از فرمول فاصله استفاده کنید، این دو کاربر حتی اگر سلیقه‌ی مشابهی داشته باشند، می‌توانند همسایه‌ی یکدیگر نباشند، شباخت کسینوسی فاصله‌ی بین دو بُردار را اندازه نمی‌گیرد. در عوض، زوایای دو بُردار را با هم مقایسه می‌کند. در برخورد با چنین مواردی بهتر است چنین رویکردی را داشته باشیم. موضوع شباخت کسینوسی خارج از محدوده‌ی این کتاب است، اما اگر از KNN استفاده می‌کنید، درباره‌ی آن جست‌وجو کنید!

1. Cosine similarity
2. vectors



انتخاب ویژگی‌های خوب

برای تعیین پیشنهادات، امتیاز کاربران به دسته‌بندی فیلم‌ها را در اختیار داشتید. در عوض اگر از کاربران می‌خواستید به عکس گربه‌ها امتیاز بدهند چه؟ اطلاعات کاربران را بر اساس تشابه امتیازدهی این عکس‌ها خواهید داشت.

در این شرایط، قاعده‌تاً موتور پیشنهادات عملکرد بدتری خواهد داشت، چون این «ویژگی‌ها» ارتباط چندانی با سلیقه در فیلم‌ها ندارند! یا فرض کنید از کاربران می‌خواهید به فیلم‌ها امتیاز بدهند تا بر اساس آن بتوانید به این افراد پیشنهاداتی ارائه کنید - اما تنها از آن‌ها می‌خواهید به اینیمیشن‌های داستان اسباب‌بازی ۲ و داستان اسباب‌بازی ۳ امتیاز بدهند. این به شما اطلاعات چندانی درباره‌ی سلیقه‌ی فیلمی کاربران نمی‌دهد.

هنگامی که با KNN کار می‌کنید، انتخاب ویژگی‌های مناسب برای مقایسه بسیار مهم است. انتخاب ویژگی‌های مناسب یعنی:

- ویژگی‌هایی که مستقیماً با فیلم‌هایی که می‌خواهید توصیه کنید مرتبط باشد.
- ویژگی‌هایی که صرفاً یک سوگیری مشخص ندارند (برای مثال، اگر از کاربران بخواهید فقط به فیلم‌هایی کمی امتیاز بدهند، به شما نمی‌گوید که آیا آن‌ها فیلم‌های اکشن را دوست دارند یا خیر)

به نظر شما رتبه‌بندی روش خوبی برای توصیه‌ی فیلم است؟ شاید من به سریال شنود بیشتر از برنامه‌ی تلویزیونی جویندگان خانه امتیاز دادم، اما در واقع زمان بیشتری را صرف تماشای جویندگان خانه کنم. چه پیشنهادی برای بهبود سیستم پیشنهادات نتفلیکس دارید؟

برگردیم به مثال نانوایی: آیا می‌توانید دو ویژگی خوب و دو ویژگی بد برای نانوایی انتخاب کنید؟ شاید بعد از آگهی در روزنامه یا دوشبته‌ها لازم باشد نان‌های بیشتری پخت کنید.

وقتی نوبت به انتخاب ویژگی‌های خوب می‌رسد، هیچ پاسخ درستی وجود ندارد. شما باید تمام جوانب را در نظر بگیرید.

تمرین

۱۰.۳ نتفلیکس میلیون‌ها کاربر دارد. در مثال قبلی به پنج همسایه‌ی نزدیک برای ساختن سیستم پیشنهادات نگاه کردیم. این تعداد خیلی کم است؟ یا خیلی زیاد است؟



مقدمه‌ای بر یادگیری ماشینی

KNN الگوریتمی واقعاً مفید و مقدمه‌ای برای ورود به دنیای جادویی یادگیری ماشین است! یادگیری ماشین چیزی نیست جز هوشمندتر کردن کامپیوتر. قبل از یک نمونه از یادگیری ماشین را مرور کردیم: ایجاد یک سیستم پیشنهادات. بیایید چند نمونه‌ی دیگر را بررسی کنیم.

OCR

OCR مخفف تشخیص نوری حرف^۱ است. به این معنا که می‌توانید از یک صفحه متن عکس بگیرید و کامپیوتر به طور خودکار متن را برای شما بخواند. گوگل از OCR برای دیجیتالی کردن کتاب‌ها استفاده می‌کند. OCR چگونه کار می‌کند؟ برای مثال این عدد را در نظر بگیرید.

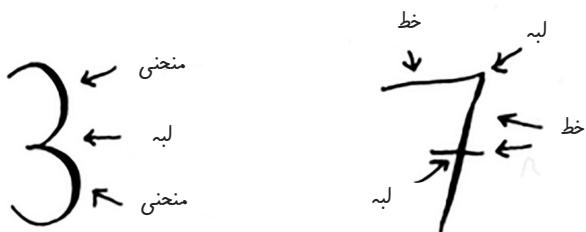
۷

چگونه به طور خودکار می‌توان فهمید که این عدد چیست؟ برای این کار می‌توانید از KNN استفاده کنید:

۱. تصاویر زیادی از اعداد را مرور کرده و ویژگی‌های آن اعداد را استخراج کنید.

1. optical character recognition

۲. هنگامی که یک تصویر جدید دریافت کردید، ویژگی‌های آن تصویر را استخراج می‌کند، و نزدیک‌ترین همسایه‌های آن را پیدا می‌کند! این همان مسئله‌ی پرتفال و گریپ فروت است. به طور کلی، الگوریتم‌های OCR خطوط، لبه‌ها و منحنی‌ها را اندازه‌گیری می‌کنند.



بعد وقتی یک کاراکتر جدید دریافت کردید، می‌توانید همان ویژگی‌ها را از آن استخراج کنید.

استخراج ویژگی در OCR بسیار پیچیده‌تر از مثال میوه است. اما باید بدانید که فناوری‌های پیچیده هم بر اساس ایده‌های ساده مانند KNN ساخته می‌شوند. می‌توانید از همین ایده‌ها برای تشخیص گفتار یا تشخیص چهره استفاده کنید. وقتی عکسی را در فیسبوک بارگزاری می‌کنید، این پلتفرم گاهی آن قدر هوشمند است که افراد به تصویر درآمده در عکس را به طور خودکار تگ کند. این یک کاربرد واقعی یادگیری ماشینی است!

اولین مرحله‌ی OCR که در آن تصاویر اعداد را مرور می‌کنید و ویژگی‌ها را استخراج می‌کنید، آموزش^۱ نامیده می‌شود. اکثر الگوریتم‌های یادگیری ماشینی یک مرحله آموزش دارند: به کامپیوتر پیش از آنکه بتواند کاری انجام دهد، باید آموزش داده بشود. مثال بعدی شامل فیلترهای اسپم است و یک مرحله‌ی آموزش دارد.

ساخت فیلتر اسپم

فیلترهای اسپم از الگوریتم ساده‌ی دیگری به نام دسته‌بندی‌کننده‌ی بیز ساده^۱ استفاده می‌کنند. در ابتدا، دسته‌بندی‌کننده‌ی بیز ساده را با تعدادی داده آموزش می‌دهید.

موضوع	اسپم؟
«پسورد خود را بست کنید»	اسپم نیست
«شایرندۀ امیلیون دلار شده‌اید»	اسپم
«پسوردت را برای من بفرست.»	اسپم
«شاهزادۀ نیجریه می‌خواهد ۱۰ میلیون دلار به شما منتقل کند»	اسپم
«تولدت مبارک»	اسپم نیست

فرض کنید ایمیلی با این مضمون دریافت می‌کنید که «یک میلیون دلار خود را دریافت کنید!» آیا این یک ایمیل اسپم است؟ می‌توانید این جمله را به کلمات تقسیم کنید. سپس، برای هر کلمه، ببینید احتمال نمایش آن کلمه در ایمیل اسپم چقدر است. برای مثال در این مدل بسیار ساده کلمه‌ی میلیون فقط در ایمیل‌های اسپم ظاهر می‌شود. بیز ساده احتمال اینکه چیزی اسپم باشد را تشخیص می‌دهد. عملکردی مشابه KNN دارد.

برای مثال، می‌توانید از بیز ساده برای دسته‌بندی میوه‌ها استفاده کنید: یک میوه‌ی بزرگ و قرمز دارید. احتمال اینکه این میوه گریپ‌فروت باشد، چقدر است؟ این یک الگوریتم ساده‌ی دیگر است که تا حد زیادی مؤثر است. ما این الگوریتم‌های ساده را دوست داریم!



پیش‌بینی بازار سهام

یک مثال از مسئله‌ای که پیاده‌سازی یادگیری ماشین در آن دشوار است: پیش‌بینی تغییرات ارزش بازار سهام. چگونه ویژگی‌های خوبی را در بازار سهام انتخاب می‌کنید؟ فرض کنید شما می‌گویید اگر دیروز ارزش سهام افزایش پیدا کرده، امروز هم افزایش می‌یابد. آیا این ویژگی خوبی است؟ یا فرض کنید می‌گویید که ارزش سهام همیشه در ماه می‌کاهش می‌یابد. آیا این رویکرد کارساز است؟ هیچ راه تضمینی برای استفاده از ارقام گذشته برای پیش‌بینی عملکرد آینده وجود ندارد. پیش‌بینی آینده زمانی که متغیرهای زیادی در آن دخیل هستند سخت و تقریباً غیرممکن است.

جمع‌بندی

امیدوارم با خواندن این بخش دیدی به دست آورده باشید از کارهایی که با KNN و یادگیری ماشین می‌توان انجام داد. یادگیری ماشین حوزه‌ی جالبی است که در صورت علاقه‌ی توانید دانش خود را در آن گسترش بدهید:

- KNN برای طبقه‌بندی و رگرسیون استفاده می‌شود و شامل نگاه‌کردن به k-نزدیک‌ترین همسایه است.
 - طبقه‌بندی = دسته‌بندی¹ در یک گروه.
 - رگرسیون = پیش‌بینی پاسخ (مانند عدد).
 - استخراج ویژگی به معنای تبدیل یک آیتم(مانند میوه یا کاربر) در لیستی از اعداد قابل مقایسه است.
 - انتخاب ویژگی‌های خوب بخش مهمی از یک الگوریتم موفق KNN است.
1. categorization

۱۱| گام‌های بعدی

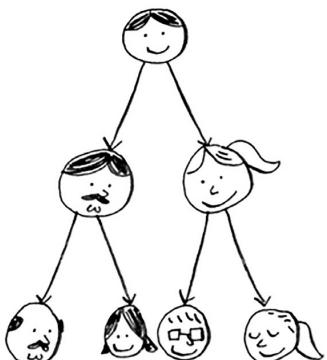


در این فصل

• مرور مختصری خواهیم کرد به ۱۰ الگوریتمی که در این کتاب به آن‌ها پرداخته نشده است و به دلایل مفید بودن آن‌ها اشاره می‌کنیم.

• به شما سرنخ‌هایی داده خواهد شد تا با توجه به علائق خود درباره آن‌ها مطالعات بیشتری داشته باشید.

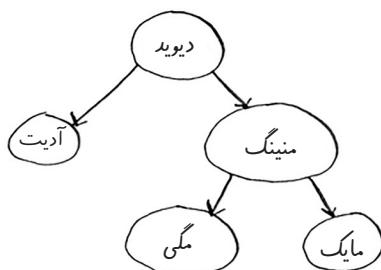
درخت



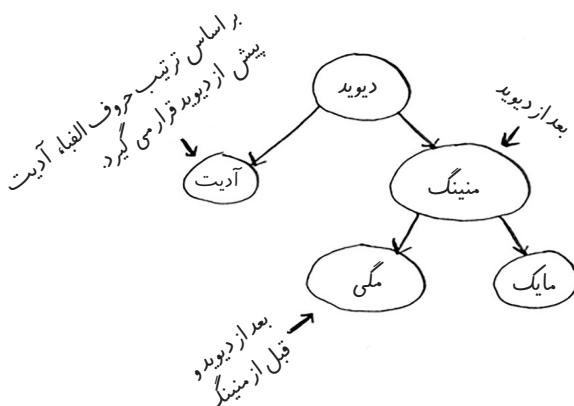
بیایید به مثال جستجوی دودویی برگردیم. هنگامی که کاربر وارد فیس‌بوک می‌شود، فیس‌بوک باید در یک آرایه‌ی بزرگ نگاه کند و ببیند آیا نام کاربری مورد نظر وجود دارد یا خیر. گفتیم سریع‌ترین راه برای جستجو در این آرایه، اجرای جستجوی دودویی است. اما یک مشکل وجود دارد: هر بار که یک کاربر جدید ثبت‌نام می‌کند، نام کاربری او را در آرایه

وارد می‌کنید. سپس باید آرایه را دوباره مرتب کنید، زیرا جستجوی دودویی فقط با آرایه‌های مرتب شده کار می‌کند. بهتر نخواهد بود اگر بتوانید نام کاربری را فوراً در روزنه‌ی سمت راست آرایه وارد کنید، تا بعد از آن مجبور نشوید آرایه را مرتب کنید؟ این ایده‌ی اصلی ساختمان داده‌ی درخت جستجوی دودویی است.

درخت جستجوی دودویی به این شکل است:

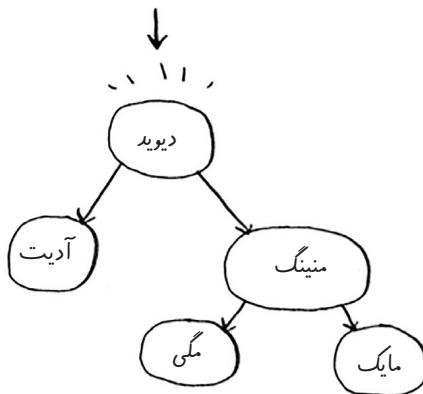


برای هر گره‌ی، گره‌های سمت چپ آن از نظر ارزش کوچک‌تر و گره‌های سمت راست از نظر ارزش بزرگ‌تر هستند.

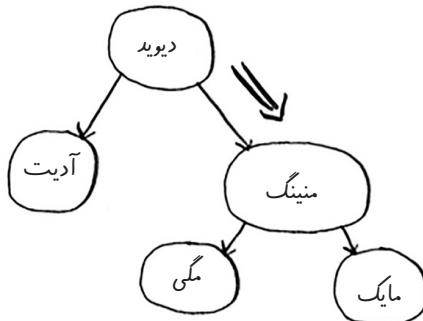


فرض کنید به دنبال نام کاربری مگی هستیم. شما از گرهی ریشه شروع می‌کنید.

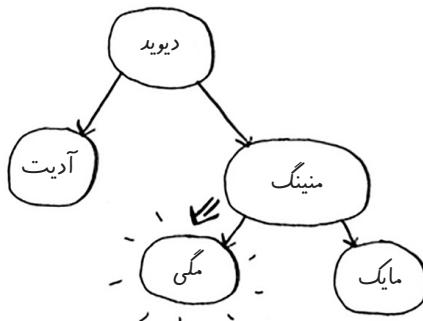
1. node



نام مگی بعد از نام دیوارد می‌آید، پس به سمت راست می‌روید.



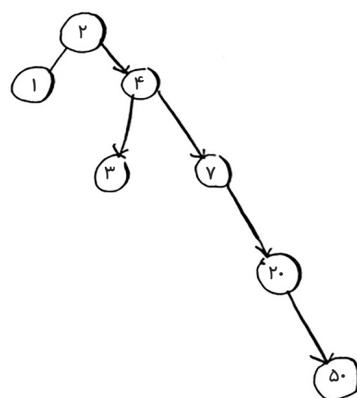
نام مگی قبل از نام منینگ می‌آید، پس به سمت چپ می‌روید.



مگی پیدا شد! تقریباً مانند اجرای یک جستجوی دودویی است! جستجوی یک عنصر در درخت جستجوی دودویی به طور متوسط زمان $O(\log n)$ و در بدترین حالت زمان $O(n)$ دارد. جستجوی یک آرایه‌ی مرتب شده در بدترین حالت به زمان $O(\log n)$ نیاز دارد، ممکن است تصور کنید آرایه‌ی مرتب شده بهتر است. اما درخت جستجوی دودویی به طور متوسط برای درج و حذف سرعت بسیار بیشتری دارد.

آرایه	درخت جستجوی دودویی
جستجو	$O(\log n)$
درج	$O(n)$
حذف	$O(n)$

درختان جستجوی دودویی عیب‌هایی هم دارند: اول ازهمه، شما دسترسی تصادفی دریافت نمی‌کنید. شما نمی‌توانید بگویید «عنصر پنجم این درخت را به من بدهید». این زمان‌های عملکرد نیز به طور متوسط و متغیر به توازن درخت هستند. فرض کنید یک درخت نامتوازن مانند آنچه در ادامه آمده، دارید.



همان‌طور که مشاهده می‌کنید درخت به سمت راست متمایل است. این درخت عملکرد چندان خوبی ندارد، زیرا متوازن نیست. درخت‌های جست‌وجوی دودویی خاصی وجود دارند که خود را متوازن می‌کنند. یک مثال درخت قرمز - سیاه است.

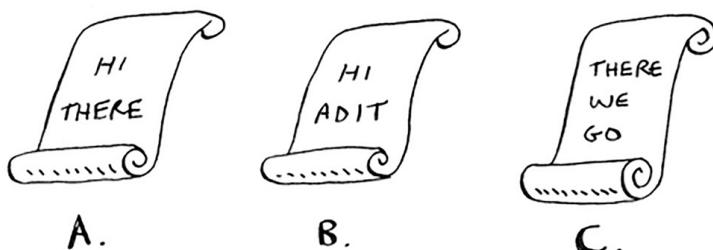
پس در چه شرایطی از درخت‌های جست‌وجوی دودویی استفاده می‌شود؟ درختان نوع خاصی از درخت دودویی، معمولاً برای ذخیره‌ی داده‌ها در پایگاه‌های داده^۱ استفاده می‌شوند.

اگر به پایگاه‌های داده یا ساختارهای داده‌ی پیشرفته‌تر علاقه دارید، این موارد را بررسی کنید:

- درختان B
- درختان قرمز-سیاه
- هیپ‌ها^۲
- درخت گسترده^۳

ایندکس‌های معکوس^۴

در اینجا یک نسخه‌ی بسیار ساده از نحوه‌ی کار یک موتور جست‌وجو ذکر شده است. فرض کنید سه صفحه‌ی وب با این محتوای ساده دارید.



1. Red-black tree
2. databases
3. Heaps
4. Splay trees
5. Inverted indexes

بایایید با این محتوا یک جدول هش بسازیم.

HI	A, B
THERE	A, C
ADIT	B
WE	C
GO	C

کلیدهای جدول هش کلمات هستند و مقادیر به شما می‌گوید که هر کلمه در چه صفحاتی ظاهر می‌شود. حال فرض کنید کاربری به دنبال hi است. بایایید ببینیم hi در چه صفحاتی نمایش داده می‌شود.



درست است! در صفحات A و B ظاهر می‌شود. اجازه بدھید آن صفحات را به عنوان نتیجه به کاربر نشان بدهیم. یا فرض کنید کاربر there را جستجو کرده است. می‌دانیید که در صفحات A و C نشان داده می‌شود. خیلی آسان است، مگر نه؟ این یک ساختمان داده‌ی کاربردی و مفید است: یک هش که کلمات را به مکان‌هایی که در آن ظاهر می‌شوند نگاشت می‌کند. به این ساختمان داده، ایندکس معکوس می‌گویند و معمولاً برای ساخت موتورهای جستجو استفاده می‌شود. اگر به جستجو علاقه‌مند هستید، این نقطه‌ی خوبی برای شروع است.

تبديل فوريه^۱

تبديل فوريه يکی از آن الگوریتم‌های نادر است: درخشان، زیبا و با هزاران کاربرد. بتر اکسپلینند^۲ (وب‌سایتی عالی که ریاضیات را به سادگی توضیح می‌دهد) بهترین تشبيه را برای تبديل فوريه می‌کند: اگر یک اسموتوی به تبديل فوريه بدهيد، آين الگوریتم، مواد تشکيل دهنده‌ی آن را به شما می‌گويد. يا اگر یک آهنگ تحويل بگيرد، فرانس‌های آن را جدا می‌کند.

1. The Fourier transform
2. <https://betterexplained.com/>

به نظر می‌رسد که این ایده‌ی ساده کاربردهای بسیاری دارد. برای مثال، اگر بتوانید یک آهنگ را بر اساس فرکانس‌ها جدا کنید، می‌توانید فرکانس‌هایی را که برایتان مهم هستند تقویت کنید. می‌توانید باس را تقویت کنید و تریبل^۱ را مخفی کنید. تبدیل فوریه برای پردازش سیگنال‌ها محشر است. می‌توانید از آن برای فشرده‌سازی موسیقی استفاده کنید. ابتدا یک فایل صوتی را به یادداشت‌های تشکیل دهنده‌ی آن تقسیم می‌کنید. تبدیل فوریه دقیقاً به شما می‌گوید که هر نُت در کل آهنگ چندبار استفاده شده است.

بنابراین می‌توانید فقط از نت‌هایی که مهم نیستند خلاص شوید. فرمت MP3 این‌گونه کار می‌کند!

موسیقی تنها نوع سیگنال دیجیتال نیست. فرمت JPEG یکی دیگر از فرمتهای فشرده است و به همین صورت عمل می‌کند. از تبدیل فوریه برای پیش‌بینی زلزله‌های آینده و تجزیه و تحلیل DNA استفاده می‌شود. می‌توانید از آن برای ساخت اپلیکیشنی مانند Shazam استفاده کنید که حدس می‌زند چه آهنگی در حال پخش است. تبدیل فوریه کاربردهای زیادی دارد. احتمال اینکه با آن برخورد کنید زیاد است!

الگوریتم‌های موازی^۲

سه مورد بعدی در مورد مقیاس‌پذیری و کار با داده‌های عظیم است. در گذشته، کامپیوترها سریع‌تر و سریع‌تر می‌شدند. اگر می‌خواستید سرعت الگوریتم خود را بهتر کنید، می‌توانستید چند ماه صبر کنید تا خود کامپیوترها سریع‌تر شوند. اما ما به پایان آن دوره نزدیک شده‌ایم. در عوض، لپ‌تاپ‌ها و رایانه‌ها با چندین هسته عرضه می‌شوند. برای اینکه الگوریتم‌های خود را سریع‌تر کنید، باید آن‌ها را طوری تغییر بدهید که بی‌درنگ و به صورت موازی در تمام هسته‌ها اجرا بشوند!

اگر بخواهیم با یک مثال ساده تشریح کنم در بهترین حالت یک الگوریتم مرتب‌سازی تقریباً زمان $O(n \log n)$ دارد. روشن است که نمی‌توانید یک آرایه را در زمان

1. treble
2. Parallel algorithms

مرتب کنید، مگر اینکه از یک الگوریتم موازی استفاده کنید! یک نسخه‌ی موازی از مرتب سازی سریع وجود دارد که یک آرایه را در زمان $O(n)$ مرتب می‌کند.

طراحی الگوریتم‌های موازی، اطمینان از کارکرد صحیح آن‌ها و اطلاع از میزان افزایش سرعت سخت است. یک مورد را می‌توان با قطعیت گفت - افزایش سرعت خطی نیست. بنابراین اگر به جای یک هسته، دو هسته در لپ‌تاپ خود دارید، به این معنا نخواهد بود که الگوریتم شما به شکلی جادویی دو برابر سریع‌تر اجرا می‌شود. چند دلیل برای این مسئله وجود دارد:

- سربار مدیریت موازی^۱ - فرض کنید باید یک آرایه از ۱۰۰۰ آیتم را مرتب کنید. چگونه این وظیفه را بین دو هسته تقسیم می‌کنید؟ آیا به هر هسته ۵۰۰ آیتم می‌دهید تا مرتب بشوند و سپس دو آرایه‌ی مرتب شده را در یک آرایه‌ی مرتب شده‌ی بزرگ ادغام می‌کنید؟ ادغام این دو آرایه زمان بر است.

- متعادل‌سازی بار^۲ - فرض کنید ۱۰ کار برای انجام باشد، بنابراین به هر هسته ۵ کار می‌دهید. اما هسته‌ی A تمام کارهای آسان را انجام می‌دهد، بنابراین در ۱۰ ثانیه کار خود را به پایان می‌رساند، در حالی که هسته‌ی B همه‌ی کارهای سخت را انجام می‌دهد، بنابراین یک دقیقه طول می‌کشد. یعنی هسته‌ی A به مدت ۵۰ ثانیه بی‌کار مانده بود در صورتی که هسته‌ی B مشغول انجام تمام کارها بود! چگونه کار را به طور مساوی توزیع می‌کنید تا هر دو هسته به یک اندازه کار کنند؟

اگر به جنبه‌ی نظری عملکرد و مقیاس‌پذیری علاقه دارید، مطالعه‌ی الگوریتم‌های موازی می‌تواند برای شما جالب باشند!

1. Overhead of managing the parallelism
2. Load balancing

نگاشت‌کاهش^۱

نوع خاصی از الگوریتم موازی هست که روزبه روز محبوب‌تر می‌شود: الگوریتم توزیع شده^۲. اگر به دو تا چهار هسته نیاز داشته باشید، اجرای یک الگوریتم موازی روی لپ‌تاپ کفایت می‌کند، اما اگر به صدها هسته نیاز داشته باشید چه؟ پس می‌توانید الگوریتم خود را برای اجرا در چندین ماشین بنویسید. الگوریتم نگاشت‌کاهش یک الگوریتم توزیع شده‌ی محبوب است. می‌توانید به کمک ابزار منبع باز و محبوب Apache Hadoop از آن استفاده کنید.

چرا الگوریتم‌های توزیع شده مفید هستند؟

فرض کنید جدولی با میلیارد‌ها یا تریلیون‌ها ردیف دارید و می‌خواهید یک کوئری SQL پیچیده را بر روی آن اجرا کنید. شما نمی‌توانید آن را در MySQL اجرا کنید، چراکه پس از چند میلیارد ردیف عملکرد خوبی نخواهد داشت. از نگاشت‌کاهش در Hadoop استفاده کنید!

یا فرض کنید باید فهرست طولانی از مشاغل را پردازش کنید. پردازش هر کار ۱۰ ثانیه طول می‌کشد و شما باید ۱ میلیون شغل را مانند این پردازش کنید. اگر این کار را روی یک دستگاه انجام بدھید، ماه‌ها طول می‌کشد! اگر بتوانید آن را روی ۱۰۰ کامپیوتر اجرا کنید، تنها در چند روز کار شما به انجام می‌رسد.

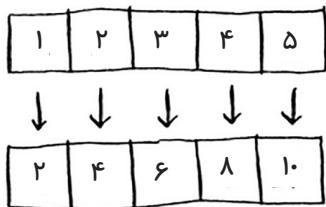
استفاده از الگوریتم‌های توزیع شده زمانی ایده‌آل است که کار زیادی برای انجام دادن دارید و می‌خواهید زمان لازم برای انجام آن را تسربیع کنید. نگاشت‌کاهش به طور خاص از دو ایده‌ی ساده ساخته شده است: تابع map و تابع reduce

1. MapReduce
2. distributed algorithm

تابع map

تابع مَپ ساده است: یک آرایه دریافت می‌کند و همان تابع را برای هر آیتم در آرایه اعمال می‌کند. به عنوان مثال، در اینجا ما هر آیتم در آرایه را دو برابر می‌کنیم:

```
>>> arr1 = [1, 2, 3, 4, 5]
>>> arr2 = map(lambda x: 2 * x, arr1)
[2, 4, 6, 8, 10]
```



شامل $[2, 4, 6, 8, 10]$ است - هر عنصر در arr1 دو برابر شد! arr2 دو برابر کردن یک عنصر بسیار سریع است. حال فرض کنید تابعی را اعمال می‌کنید که پردازش آن به زمان بیشتری نیاز دارد. به این شبه کد نگاه کنید:

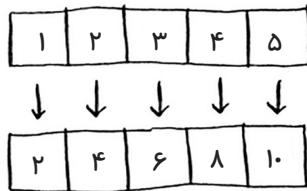
```
>>> arr1 = # A list of URLs
>>> arr2 = map(download_page, arr1)
```

در اینجا لیستی از URL ها دارید و می‌خواهید هر یک از این صفحات را دانلود کنید و محتویات آن را در arr2 ذخیره کنید. این فرآیند برای هر یک از صفحات می‌تواند چند ثانیه طول بکشد. اگر هزار URL داشته باشید، ممکن است چند ساعت طول بکشد!

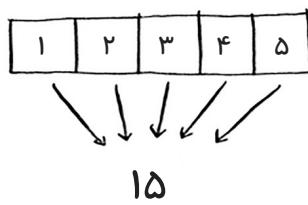
خوب نمی‌شد اگر ۱۰۰ کامپیوتر داشتید و مَپ کار را به طور خودکار میان همه آن‌ها پخش می‌کرد؟ در نتیجه شما در هر نوبت ۱۰۰ صفحه را دانلود می‌کردید و کار بسیار سریع‌تر پیش می‌رفت! این ایده‌ی اصلی پشت «map» در MapReduce است.

تابع reduce

تابع `reduce` گاهی می‌تواند گیج‌کننده باشد. طرح کلی این است که شما یک لیست کامل از آیتم‌ها را به یک مورد «کاہش» بدهید. با `map`، از یک آرایه به آرایه دیگر می‌روید.



با `reduce`، یک آرایه را به تنها یک آیتم تبدیل می‌کنید.



مثال:

```
>>> arr1 = [1, 2, 3, 4, 5]
>>> reduce(lambda x,y: x+y, arr1)
15
```

در این حالت، تمام عناصر آرایه را جمع می‌کنید: $1 + 2 + 3 + 4 + 5 = 15!$ من در اینجا جزئیات بیشتری از `reduce` ارائه نمی‌کنم، زیرا در این مورد آموزش‌های آنلاین زیادی در دسترس است.

نگاشت‌کاہش از این دو مفهوم ساده برای اجرای کوئی در مورد داده‌ها در چندین ماشین استفاده می‌کند. وقتی مجموعه داده‌ی بزرگی دارید (میلیارد‌ها ردیف)، نگاشت‌کاہش می‌تواند در عرض چند دقیقه به شما پاسخ بدهد، جایی که پایگاه داده‌ی سنتی ممکن است ساعت‌ها طول بکشد.

فیلترهای بلوم^۱ و هایپر لَگَ لَگَ^۲

فرض کنید اپلیکیشن Reddit را مدیریت می‌کنید. می‌خواهید ببینید لینکی که کاربر پست می‌کند، آیا قبلًا در سایت فرستاده شده است یا خیر. پیام‌هایی که تکراری نباشد با ارزش‌تر در نظر گرفته می‌شوند. بنابراین باید بفهمید که آیا این لینک قبل ارسال شده است یا خیر.

یا فرض کنید شما گوگل هستید و در حال خزیدن^۳ یا به عبارتی پیمایش در صفحات وب هستید. تنها در صورتی می‌خواهید یک صفحه‌ی وب را پیمایش کنید که پیش از این پیمایش نشده باشند. بنابراین باید بفهمید که آیا این صفحه قبلًا بررسی شده است یا خیر.

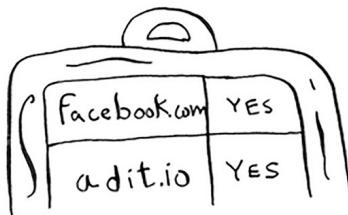
یا فرض کنید وب‌سایت bit.ly را اداره می‌کنید که یک کوتاه‌کننده‌ی URL است. نمی‌خواهید اجازه بددهید کاربران از طریق سایت شما به وب‌سایت‌های مخرب هدایت بشوند. شما مجموعه‌ای از URL‌ها دارید که مخرب در نظر گرفته می‌شوند. اکنون باید بفهمید که آیا کاربر را به یک URL در آن مجموعه هدایت می‌کنید یا خیر.

همه‌ی این نمونه‌ها مشکل یکسانی دارند. شما یک مجموعه‌ی بسیار بزرگ دارید.



1. Bloom filters
2. HyperLogLog
3. crawling

شما یک آیتم جدید دارید و می‌خواهید ببینید که آیا به آن مجموعه تعلق دارد یا خیر. شما می‌توانید این کار را به سرعت با هش انجام دهید. برای مثال، فرض کنید گوگل یک هش بزرگ دارد که کلیدهای آن تمام صفحاتی را شامل می‌شود که در آن‌ها پیمایش انجام شده است.



می‌خواهید ببینید آیا قبلًا adit.io را پیمایش کرده‌اید یا خیر. آن را در هش جست‌وجو کنید.

adit.io → YES

adit.io کلیدی است که در هش وجود دارد، بنابراین شما قبلًا آن را پیمایش کرده‌اید. میانگین زمان جست‌وجو برای جدول‌های هش O(1) است. در این هش است، بنابراین شما قبلًا به آن سر زده‌اید. شما آن را با زمان ثابت پیدا کردید. بسیار خوب!

با این تفاوت که این هش باید بزرگ باشد. گوگل تریلیون‌ها صفحه‌ی وب را ایندکس می‌کند. اگر این هش شامل تمامی URL‌هایی باشد که گوگل ایندکس کرده است، فضای زیادی را اشغال می‌کند. Reddit و bit.ly مشکل فضای یکسانی دارند. وقتی با داده‌های زیادی سروکار دارید، باید خلاق باشید!

فیلترهای بلوم

فیلترهای بلوم یک راه حل ارائه می‌کنند. فیلترهای بلوم ساختمان داده‌های احتمالی^۱ هستند و پاسخی به شما می‌دهند که هرچند امکان دارد اشتباه باشد اما احتمالاً صحیح باشد. به

جای هش، می‌توانید از فیلتر بلوم بپرسید که آیا قبلًا این URL را پیمایش کرده یا خیر. یک جدول هش به شما پاسخ دقیقی می‌دهد. فیلتر بلوم به شما پاسخی می‌دهد که احتمالاً درست است:

- موارد مثبت کاذب ممکن است. احتمال دارد گوگل اعلام کند: «شما قبلًا این سایت را پیمایش کرده‌اید»، حتی اگر این کار را نکرده باشید.
- منفی کاذب امکان‌پذیر نیست. اگر فیلتر بلوم اعلام کند: «شما این سایت را پیمایش نکرده‌اید»، پس حتماً همین طور است.

فیلترهای بلوم بسیار عالی هستند زیرا فضای بسیار کمی را اشغال می‌کنند. یک جدول هش باید هر آدرسی را که گوگل پیمایش کرده، ذخیره کند، اما فیلتر بلوم نیازی به انجام این کار ندارد. زمانی که به پاسخ دقیقی نیاز ندارید، بسیار خوب هستند، مانند تمام این مثال‌ها. برای [.ly](#) اشکالی ندارد که اعلام کند: «این سایت ممکن است مخرب باشد، بنابراین بیشتر مراقب باشید.»

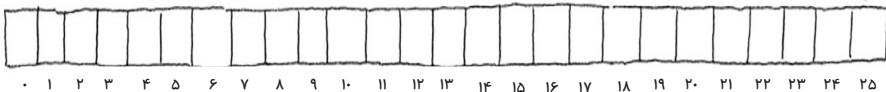
هایپرلاگ لاغ

در همین رابطه الگوریتم دیگری با نام HyperLogLog وجود دارد. فرض کنید گوگل می‌خواهد تعداد جستجوهای منحصر به فرد انجام شده به دست کاربرانش را بشمارد. یا فرض کنید آمازون می‌خواهد تعداد آیتم‌های منحصر به فردی را که امروز کاربران به آن‌ها نگاه می‌کنند، بشمارد. برای پاسخ به این پرسش‌ها فضای زیادی نیاز است! در مورد گوگل باید گزارشی از تمام جستجوهای منحصر به فرد داشته باشید. وقتی کاربر چیزی را جستجو می‌کند، باید بینید که آیا از قبل در گزارش وجود دارد یا خیر. اگر نه، باید آن را به گزارش اضافه کنید. حتی برای یک روز هم لاغ بسیار بزرگی می‌شود! هایپرلاگ لاغ تعداد عناصر منحصر به فرد در یک مجموعه را تخمین می‌زند. درست مانند فیلترهای بلوم، پاسخ دقیقی به شما نمی‌دهد، اما بسیار به پاسخ دقیق نزدیک است و برای عملیاتی مشخص به نسبت شرایط دیگر تنها کسری از حافظه را استفاده می‌کند.

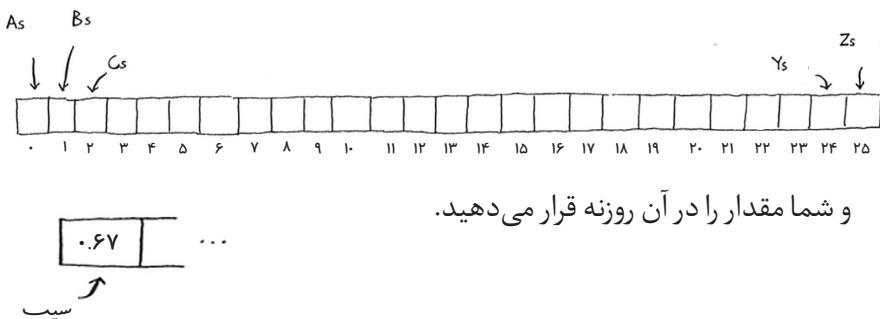
اگر داده‌های زیادی دارید و با پاسخ‌های تقریبی راضی هستید، الگوریتم‌های احتمالی را بررسی کنید!

الگوریتم‌های SHA

هش را از فصل ۵ به خاطر دارید؟ محض یادآوری، فرض کنید یک کلید دارید و می‌خواهید مقدار مرتبط را در یک آرایه قرار بدهید.



شما از یک تابع هش استفاده می‌کنید تا به شما بگوید که مقدار را در چه روزنه‌ای قرار بدهید.



و شما مقدار را در آن روزنه قرار می‌دهید.

این روش به شما امکان می‌دهد جستجوهایی با زمان-ثابت انجام بدهید. هنگامی که می‌خواهید مقدار یک کلید را بدانید، می‌توانید دوباره از تابع هش استفاده کنید و در زمان $O(1)$ به شما می‌گوید کدام روزنه را باید بررسی کنید.

در این حالت، شما می‌خواهید که تابع هش توزیع خوبی در اختیار شما بگذارد. بنابراین یک تابع هش یک رشته را می‌گیرد و شماره‌ی روزنه‌ی آن رشته را به شما بازمی‌گرداند.

مقایسه‌ی فایل‌ها

یکی دیگر از توابع هش یک تابع الگوریتم هش ایمن^۱ (SHA) است. با توجه به یک رشته، SHA یک هش برای آن رشته به شما می‌دهد.

شاید این اصطلاح برای شما کمی گیج‌کننده باشد. SHA یک تابع هش است. یک هش تولید می‌کند که فقط یک رشته‌ی کوتاه است. تابع هش برای جدول‌های هش از یک رشته به ایندکس آرایه می‌رود، در حالی که SHA از رشته‌ای به رشته‌ی دیگر می‌رود. یک هش متفاوت برای هر رشته تولید می‌کند.

“hello” \Rightarrow 2cf24db...

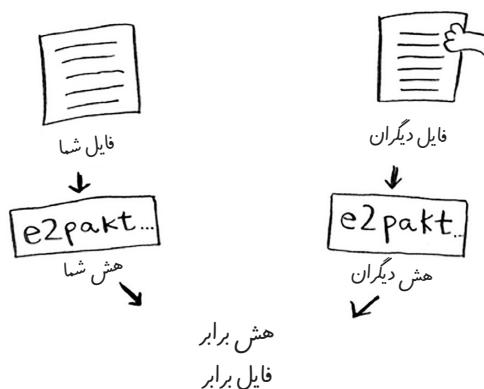
“algorithm” \Rightarrow b1eb2ec...

“password” \Rightarrow 5e88489...

نکته

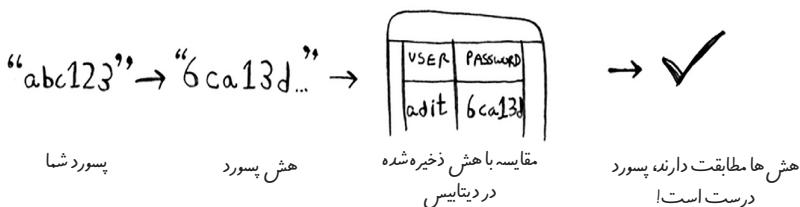
هش‌های SHA طولانی هستند و در اینجا کوتاه شده‌اند.

شما می‌توانید از SHA برای تشخیص یکسان بودن دو فایل استفاده کنید. این زمانی مفید است که فایل‌های بسیار بزرگی دارید. فرض کنید یک فایل ۴ گیگابایتی دارد. می‌خواهید بررسی کنید که آیا دوست شما همان فایل بزرگ را دارد یا خیر. لازم نیست فایل پرچم خود را ایمیل کنید. در عوض، می‌توانید هش SHA را محاسبه کرده و آن را مقایسه کنید.



چک کردن رمزهای عبور

همچنین زمانی مفید است که بخواهید رشته‌ها را بدون اینکه نشان دهید رشته‌ی اصلی چیست مقایسه کنید. به عنوان مثال، فرض کنید سرویس جیمیل هک می‌شود و مهاجم تمام رمزهای عبور را سرقت می‌کند! آیا رمز عبور شما فاش شده است؟ نه، اینطور نیست. گوگل رمز عبور اصلی را ذخیره نمی‌کند، فقط هش SHA رمز عبور را ذخیره می‌کند! وقتی رمز عبور خود را تایپ می‌کنید، گوگل آن را هش می‌کند و آن را با هش موجود در پایگاهداده‌ی خود بررسی می‌کند.



بنابراین فقط هش‌ها را مقایسه می‌کند - لازم نیست رمز عبور شما را ذخیره کند! معمولاً برای هش رمزهای عبور مانند این استفاده می‌شود. این یک هش یک‌طرفه است. می‌توانید هش یک رشته را دریافت کنید.

$\text{abc123} \rightarrow 6ca13d$

اما با در اختیار داشتن هش نمی‌توانید رشته‌ی اصلی را به دست بیاورید.

? ← 6ca13d

در نتیجه اگر هکری هش‌های SHA را از جیمیل به دست بیاورد، نمی‌تواند آن هش‌ها را به رمزهای عبور اصلی تبدیل کند! شما می‌توانید رمز عبور را به هش تبدیل کنید، اما برعکس آن امکان پذیر نیست.

SHA در واقع خانواده‌ای از الگوریتم‌ها است: SHA-0، SHA-1، SHA-2 و SHA-3. در زمان نگارش این کتاب، SHA-0 و SHA-1 دارای نقاط ضعفی هستند. اگر از الگوریتم SHA برای هش رمز عبور استفاده می‌کنید، از SHA-2 یا SHA-3 استفاده کنید. استاندارد طلایی برای توابع هش رمز عبور در حال حاضر bcrypt است (اگرچه امنیت هیچ‌کدام کاملاً تضمین شده نیست).

هشینگ حساس به مکان^۱

SHA یک ویژگی مهم دیگر نیز دارد: به محل حساس نیست. فرض کنید یک رشته دارید و یک هش برای آن ایجاد می‌کنید.

`dog → cd6357`

اگر فقط یک کاراکتر از رشته را تغییر دهید و هش را دوباره ایجاد کنید، کاملاً متفاوت می‌شود!

`dot → e392da`

این نکته‌ی مثبتی است چون مهاجم نمی‌تواند هش‌ها را مقایسه کند و ببیند به شکستن رمز عبور نزدیک شده است یا خیر. گاهی اوقات، برعکس است: یک تابع هش حساس به مکان می‌خواهد. اینجاست که Simhash وارد می‌شود. اگر تغییر کوچکی در یک رشته ایجاد کنید، Simhash هش ایجاد می‌کند که تنها تفاوت اندکی دارد. این به شما امکان می‌دهد هش‌ها را مقایسه کنید و ببینید دو رشته چقدر شبیه هم هستند که ویژگی بسیار مفیدی است!

- گوگل از Simhash برای شناسایی موارد تکراری هنگام پیمایش وب استفاده می‌کند.
- یک معلم می‌تواند از Simhash استفاده کند تا ببیند آیا دانش‌آموzan مقاله‌های خود را از وب کپی کرده‌اند یا خیر.

- Scribd به کاربران امکان می‌دهد اسناد یا کتاب‌هایی را برای اشتراک‌گذاری با

1. Locality-sensitive hashing

دیگران بارگزاری کنند. اما Scribd نمی‌خواهد کاربران محتوای دارای حق چاپ را آپلود کنند! این سایت می‌تواند از Simhash استفاده کند تا بیند که آیا یک فایل بارگزاری شده به طور مثال مشابه کتاب هری پاتر است یا خیر و اگر چنین است، به طور خودکار جلوی آن را بگیرد.

Simhash زمانی مفید است که می‌خواهید ایتم‌های مشابه را بررسی کنید.

تبادل کلید دیفی-هلمن^۱

در اینجا خوب است به الگوریتم دیفی-هلمن اشاره کنیم، چرا که یک مسئله‌ی قدیمی را به روشی زیبا حل می‌کند. شیوه‌ی رمزگذاری یک پیام به‌ نحوی که فقط در مقصد قابل خواندن باشد چیست؟

ساده‌ترین راه این است که یک رمز مانند $a = 2$ و غیره ایجاد کنید. سپس اگر پیام «۴، ۱۵، ۷» را برای شما ارسال کنم، می‌توانید آن را به « $d, 0, g$ » ترجمه کنید. اما برای چنین کاری، هر دو طرف باید بر سر یک رمز به توافق برسیم. مانعی توانیم در خود ایمیل به نتیجه برسیم، زیرا ممکن است شخصی ایمیل شما را هک کند، رمز را پیدا کرده، و پیام‌های ما را رمزگشایی کند. عجب، حتی اگر هم‌دیگر را حضوری ببینیم، باز هم ممکن است کسی رمز را حدس بزند - کار چندان پیچیده‌ای نیست. پس رمز را باید هر روز تغییر بدھیم. اما برای این کار لازم است هر روز هم‌دیگر را ملاقات کنیم!

حتی اگر هر روز موفق به تغییر آن شویم، رمز ساده‌ای مانند این به راحتی با یک حمله‌ی brute-force شکسته می‌شود. فرض کنید من پیام «۹، ۶، ۱۳، ۱۶، ۲۴، ۱۶، ۱۹، ۱۳، ۱۳، ۵» را می‌بینم. حدس من این است که در این مورد از $a = 1, b = 2$ و غیره استفاده می‌شود.

۹	۶	۱۳	۱۳	۱۶	۲۴	۱۶	۱۹	۱۳	۵
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
i	f	m	m	P	X	P	S	m	e

1. Diffie-Hellman key exchange

این کلمات معنی خاصی ندارند. بیایید $a = 3$ ، $b = 2$ وغیره را امتحان کنیم.

٩	٦	١٣	١٣	١٦	٢٤	١٦	١٩	٥
↓	↓	↓	↓	↓	↓	↓	↓	↓
h	e	l	l	o	w	o	r	d

نتیجه داد! یک رمز ساده مانند این به راحتی شکسته می‌شود. آلمانی‌ها از رمز بسیار پیچیده‌تری در جنگ جهانی دوم استفاده کردند، اما آن هم شکسته شد. دیفی هلمن این دو مشکل را حل می‌کند:

- نیازی نیست هر دو از رمز اطلاع داشته باشند. بنابراین نیازی به ملاقات برای توافق بر سر رمز نیست..
- رمزگشایی پیام‌های رمزنگاری شده بسیار دشوار است.

دیفی هلمن دو کلید دارد: یک کلید عمومی و یک کلید خصوصی. کلید عمومی دقیقاً همین است: عمومی. می‌توانید آن را در وب‌سایت خود ارسال کنید، به دوستان خود ایمیل کنید، یا هر کار دیگری که می‌خواهید با آن انجام بدھید. نیازی نیست آن را پنهان کنید. وقتی شخصی می‌خواهد برای شما پیامی ارسال کند، آن را با استفاده از کلید عمومی رمزنگاری می‌کند. یک پیام رمزنگاری شده فقط با استفاده از کلید خصوصی قابل رمزگشایی است. تا زمانی که شما تنها فردی باشید که کلید خصوصی را در اختیار دارید، فقط شما می‌توانید این پیام را رمزگشایی کنید!

الگوریتم دیفی هلمن همچنان به همراه جایگزین آن RSA استفاده می‌شود. اگر به رمزنگاری علاقه دارید، دیفی هلمن نقطه شروع خوبی است: ساده و ظریف است و دنبال‌کردن آن چندان سخت نیست.

برنامه‌ریزی خطی^۱

بهترین الگوریتم را برای آخر نگه داشتم. برنامه‌ریزی خطی یکی از جالب‌ترین چیزهایی است که می‌شناسم.

برنامه‌ریزی خطی برای به حداکثرسازدن یک چیز مشخص با در نظر داشتن به برخی محدودیت‌ها استفاده می‌شود. برای مثال، فرض کنید شرکت شما دو محصول، پیراهن و ساک پارچه‌ای و شلوار تولید می‌کند. پیراهن به ۱ متر پارچه و ۵ دکمه نیاز دارد. ساک‌ها به ۲ متر پارچه و ۲ دکمه نیاز دارند. ۱۱ متر پارچه و ۲۰ دکمه دارید. شما به ازای هر پیراهن ۲ دلار و برای هر ساک ۳ دلار به دست می‌آورید. برای به حداکثرسازدن سود باید چه تعداد پیراهن و ساک تولید کنید؟

در اینجا ضمن آن که به اقلامی که در اختیار دارید محدود شده‌اید تلاش دارید سود خود را به حداکثر برسانید.

مثال دیگر: فرض کنید یک سیاستمدار هستید و می‌خواهید تعداد آرای خود را به حداکثر برسانید. تحقیقات شما نشان داده است که به طور متوسط یک ساعت کار (بازاریابی، تحقیق و غیره) برای هر رأی از ساکنان سانفرانسیسکو، یا یک و نیم ساعت / رأی از یک ساکن شیکاگو طول می‌کشد. به حداقل ۵۰۰ ساکن سانفرانسیسکایی و ۳۰۰ ساکن شیکاگو نیاز دارید. ۵ روز فرصت دارید. همچنین برای شما ۲ دلار / ساکن سانفرانسیسکو در مقابل ۱ دلار / ساکن شیکاگو هزینه دارد. کل بودجه‌ی شما ۱۵۰۰ دلار است. حداکثر تعداد کل آرایی که می‌توانید کسب کنید (سانفرانسیسکو + شیکاگو) چقدر است؟

در اینجا شما سعی می‌کنید رأی‌ها را به حداکثر برسانید، و زمان و پول شما محدود است.

ممکن است تصور کنید، «در مورد بسیاری از موضوعات بهینه سازی در این کتاب صحبت شده است. ارتباط آن‌ها با برنامه‌ریزی خطی چیست؟ همه‌ی الگوریتم‌های گراف را می‌توان از طریق برنامه‌ریزی خطی انجام داد. برنامه‌ریزی خطی یک چارچوب بسیار کلی‌تر است و

مسئله‌های گراف زیرمجموعه‌ای از آن هستند. امیدوارم دود از کله‌ی شما بلند شده باشد! برنامه نویسی خطی از الگوریتم غیر مرکب^۱ استفاده می‌کند. این یک الگوریتم پیچیده است، به همین دلیل است که در این کتاب به آن پرداخته نشده است. اگر به بهینه‌سازی علاقه دارید، برنامه‌ریزی خطی را جستجو کنید!

سخن آخر

امیدوارم این مرور سریع ۱۰ الگوریتم به شما نشان داده باشد که چه موارد بیشتری برای کشف باقی مانده است. بهترین راه برای یادگیری این است که موضوع مورد علاقه‌ی خود را پیدا کنید و در آن شیرجه بزنید. این کتاب شالوده‌ای مستحکم برای چنین کاری فراهم کرده است.

پاسخ تمرین‌ها



فصل ۱

۱.۱ فرض کنید یک لیست مرتب شده از ۱۲۸ نام دارید، و با جست‌وجوی دودویی در میان آن لیست به جست‌وجوی پردازید. حداکثر تعداد مراحل چه مقدار خواهد بود؟

پاسخ: ۷.

۱.۲ تعداد اعضای لیست را دوباره مقدار قبلی در نظر بگیرید. این بار حداکثر تعداد مراحل چقدر خواهد بود؟

پاسخ: ۸.

۱.۳ نام یک نفر را در نظر دارید و می‌خواهید شماره تلفن آن شخص را در دفترچه تلفن پیدا کنید.

پاسخ: $O(\log n)$.

۱.۴ شما یک شماره تلفن دارید و می‌خواهید نام شخص را در دفترچه تلفن پیدا کنید.
(نکته: باید در تمام صفحات دفترچه بگردید!)

پاسخ: $O(n)$.

۱.۵ شما می‌خواهید شماره‌ی هر فرد نوشته شده در دفترچه تلفن را بخوانید.
پاسخ: $O(n)$.

۱.۶ شما می‌خواهید شماره‌ی افرادی را بخوانید که نام آن‌ها با حرف A شروع شده است.
پاسخ: $O(n)$. ممکن است فکر کنید، «من فقط برای ۱/۲۶ حروف این کار را انجام می‌دهم بنابراین زمان اجرا باید $O(n/2)$ باشد. یک قانون ساده که باید به خاطر بسیاری‌د است که از اعدادی که جمع، تفربیق، ضرب یا تقسیم می‌شوند چشم‌پوشی کنید. هیچکدام از این زمان‌های اجرای Big O صحیح نیستند:
 $O(n / 26)$ ، $O(n * 26)$ ، $O(n - 26)$ ، $O(n + 26)$ همه‌ی آن‌ها مانند! $O(n)$ هستند.
 چرا؟ اگر کنجدکاو هستید، به «بازبینی نماد O بزرگ» در فصل ۴ مراجعه کنید و درباره‌ی ثابت‌ها در نماد O مطالعه کنید (یک ثابت فقط یک عدد است؛ ۲۶ عدد ثابت این پرسش بود).

فصل ۲

۲.۱ فرض کنید در حال ساخت یک برنامه برای پیگیری امور مالی خود هستید.

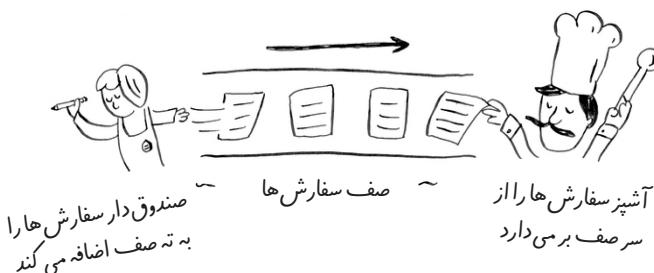
۱. خورد و خوراک روزمره
۲. سینما
۳. حق عضویت SFBC

هرروز خرچ‌های خود را یادداشت می‌کنید. در پایان ماه، هزینه‌ها را بررسی کرده و میزان هزینه‌های خود را جمع می‌زنید. بنابراین، شما تعداد زیادی درج کردن و مقدار کمی خواندن دارید. برای این برنامه باید از میان آرایه و لیست کدام را استفاده کنید؟

پاسخ: هر روز هزینه‌ها را به لیست اضافه می‌کنید و ماهی یک بار همه‌ی هزینه‌ها را مطالعه می‌کنید. آرایه‌ها دارای سرعت خواندن بالا و درج کند هستند. لیست‌های پیوندی دارای

سرعت خواندن کند و درج سریع هستند. از آن جایی که بیشتر از اینکه قرار باشد بخوانید درج می‌کنید، استفاده از یک لیست پیوندی منطقی است. همچنین، تنها در صورتی که بخواهید به عناصر تصادفی در لیست دسترسی داشته باشید، خواندن فهرست‌های پیوندی کند است. چون شما در حال خواندن تمام عنصر لیست هستید، لیست‌های پیوندی هم در خواندن به خوبی عمل می‌کنند. بنابراین یک لیست پیوندی راه حل خوبی برای این مشکل است.

۲.۲ فرض کنید در حال ساخت یک برنامه برای رستوران‌ها هستید تا سفارش مشتری‌ها در آن ثبت بشود. برنامه‌ی شما باید فهرستی از سفارشات را ذخیره کند. گارسون‌ها مدام به این لیست سفارش‌هایی را اضافه می‌کنند، و آشپزها سفارشات را از لیست می‌بینند و آن‌ها را آماده می‌کنند. این یک صف سفارش است: گارسون‌ها سفارشات را به ته صف اضافه می‌کنند، و آشپزها اولین سفارش را از سر صف بر می‌دارند و آن را حاضر می‌کنند.



برای پیاده‌سازی این صف از میان آرایه و لیست پیوندی کدام را استفاده می‌کنید؟ (راهنمایی: لیست‌های پیوندی برای درج/حذف و آرایه‌ها برای دسترسی تصادفی مناسب هستند. کدام یک را در اینجا به کار می‌برید؟)

پاسخ: لیست پیوندی. درجهای زیادی انجام می‌شود. (گارسون‌هایی که سفارش‌ها را اضافه می‌کنند)، و لیست‌های پیوندی برای آن‌ها مناسب‌تر هستند. شما نیازی به جست‌وجو یا دسترسی تصادفی ندارید (موردهی که آرایه‌ها در آن برتری دارند)، چون سرآشپزها همیشه اولین سفارش را از صف خارج می‌کنند.

۲.۳ بیایید یک آزمایش نظری انجام بدھیم. فرض کنید **فیس بوک** لیستی از نام‌های کاربری را نگه می‌دارد. هنگامی که شخصی سعی می‌کند به **فیس بوک** وارد بشود، نام کاربری او جست‌وجو می‌شود. اگر نام آن‌ها در لیست نام‌های کاربری باشد، می‌توانند وارد بشوند. به دفعات کاربران وارد **فیس بوک** می‌شوند، بنابراین جست‌وجوهای زیادی از طریق این لیست از نام‌های کاربری انجام می‌شود. فرض کنید **فیس بوک** از جست‌وجوی دودویی برای جست‌وجوی لیست استفاده می‌کند. جست‌وجوی دودویی نیاز به دسترسی تصادفی دارد. شما باید بتوانید فوراً به وسط لیست نام‌های کاربری برسید. با دانستن این موضوع، آیا لیست را به شکل یک آرایه یا یک لیست پیوندی پیاده‌سازی می‌کنید؟

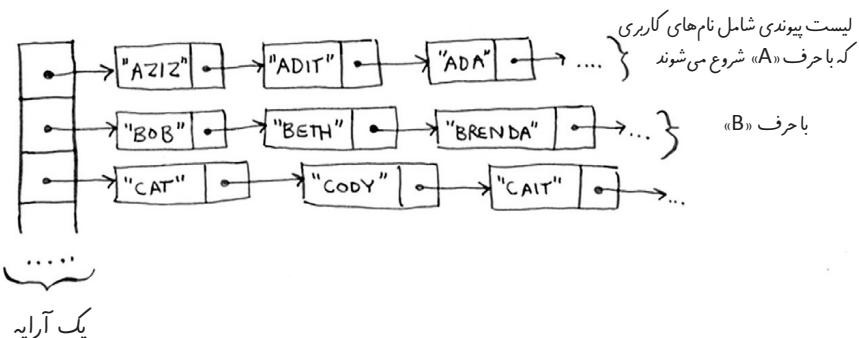
پاسخ: یک آرایه‌ی مرتب شده. آرایه‌ها به شما دسترسی تصادفی می‌دهند – شما می‌توانید یک عنصر را فوراً از وسط آرایه دریافت کنید. چنین کاری بالیست‌های پیوندی امکان‌پذیر نیست. برای رسیدن به عنصر میانی در یک لیست پیوندی، باید از عنصر اول شروع کنید و همه‌ی پیوندها را تا عنصر میانی دنبال کنید.

۲.۴ همچنین کاربرها بیشتر اوقات در **فیس بوک** ثبت‌نام می‌کنند. فرض کنید تصمیم گرفتید از یک آرایه برای ذخیره‌ی لیست کاربران استفاده کنید. استفاده از آرایه برای درج چه جنبه‌های منفی به دنبال دارد؟ به طور خاص، فرض کنید از جست‌وجوی دودویی برای لایگین استفاده می‌کنید. وقتی کاربران جدیدی به آرایه اضافه می‌کنید چه اتفاقی می‌افتد؟

پاسخ: درج در آرایه‌ها کند است. همچنین، اگر از جست‌وجوی دودویی برای جست‌وجوی نام‌های کاربری استفاده می‌کنید، آرایه باید مرتب شود. فرض کنید شخصی به نام Adit B در **فیس بوک** ثبت نام کرده است. نام‌ها در انتهای آرایه درج خواهد شد. بنابراین باید هر بار که یک نام در آن درج می‌شود آرایه را مرتب کنید!

۲.۵ در واقعیت، فیس‌بوک از آرایه و لیست پیوندی برای ذخیره‌ی اطلاعات کاربر استفاده نمی‌کند.

بیایید یک ساختمان داده‌ی ترکیبی را در نظر بگیریم: آرایه‌ای از لیست‌های پیوندی. شما یک آرایه با ۲۶ اسلات دارید. هر اسلات به یک لیست پیوندی اشاره می‌کند. به عنوان مثال، اولین روزنہ در آرایه به یک لیست پیوندی اشاره می‌کند که شامل همه‌ی نام‌های کاربری است که با حرف *a* شروع می‌شوند. اسلات دوم به یک لیست پیوندی اشاره می‌کند که شامل همه‌ی نام‌های کاربری است که با حرف *b* شروع می‌شوند وغیره.



فرض کنید Adit B در فیس بوک ثبت نام کرده است، و شما می‌خواهید او را به لیست اضافه کنید. شما به اسلات ۱ در آرایه می‌روید، به لیست پیوندی برای اسلات ۱ رفته و B را به پایان لیست اضافه می‌کنید. حال، فرض کنید می‌خواهید H را جست‌وجو کنید. به اسلات ۲۶ می‌روید، که به لیست پیوندی از همه‌ی نام‌های Z اشاره می‌کند. سپس در آن لیست جست‌وجو می‌کنید تا H Zakhir را پیدا کنید.

این ساختمن داده‌ی ترکیبی را با آرایه‌ها و لیست‌های پیوندی مقایسه کنید. به نسبت آرایه و فهرست پیوندی برای جست‌وجو و درج کنتر است یا سریع‌تر؟ نیازی نیست زمان اجرای ۰ بزرگ را محاسبه کنید، فقط اینکه ساختمن داده‌ی جدید سریع‌تر است یا کنتر؟

پاسخ: جست‌وجو: کنتر از آرایه‌ها، سریع‌تر از لیست‌های پیوندی. درج: سریع‌تر از آرایه‌ها، مدت زمان مشابه لیست‌های پیوندی. بنابراین برای جست‌وجو از یک آرایه کنتر است، اما سریع‌تر یا یکسان با لیست‌های پیوندی در تمامی موارد است.

ما در مورد ساختمان داده‌ی هیبریدی دیگری به نام جدول هش در این کتاب صحبت می‌کنیم. این موضوع به شما دید می‌دهد تا بتوانید ساختمان داده‌های پیچیده‌تر را به کمک ساختارهای ساده‌تر بسازید.

بالاخره فیس بوک از چه روشی استفاده می‌کند؟ احتمالاً از ده‌ها پایگاه داده‌ی مختلف، با ساختمان داده‌های متفاوت استفاده می‌کند: جدول‌های هش، درخت‌های -B، وغیره. آرایه‌ها و لیست‌های پیوندی، بلوک‌های سازنده‌ی این ساختمان داده‌های پیچیده‌تر هستند.

فصل ۳

۳.۱ فرض کنید من یک پشته فراخوانی مانند این را به شما نشان می‌دهم.



تنها بر اساس این پشته فراخوانی، چه اطلاعاتی می‌توان داد؟

پاسخ:

- ابتدا تابع greet با name = maggie فراخوانی می‌شود.

- سپس تابع greet2 تابع greet را با name = maggie فراخوانی می‌کند.

- در این مرحله، تابع greet در حالت ناتمام و تعلیق شده است،

- فراخوانی تابع فعلی، تابع greet2 است.

- پس از اتمام فراخوانی این تابع، فعالیت تابع greet از سرگرفته می‌شود.

۳.۲ فرض کنید به طور تصادفی یک تابع بازگشتی نوشته شد که تا ابد اجرا می‌شود. همان‌طور که دیدید، کامپیووتر شما برای هر فراخوانی تابع، حافظه را در پشته تخصیص می‌دهد. وقتی تابع بازگشتی شما تا ابد اجرا بشود چه اتفاقی برای پشته می‌افتد؟

پاسخ: پشته برای همیشه رشد می‌کند. هر برنامه فضای محدودی در پشته فراخوانی دارد. وقتی فضای برنامه‌ی شما تمام شود (که در نهایت اتفاق می‌افتد)، با یک خطای سرریز پشته^۱ خارج می‌شود.

فصل ۴

۴.۱ کد را برای تابع `sum` قبلی بنویسید.

پاسخ:

```
def sum(list):
    if list == []:
        return 0
    return list[0] + sum(list[1:])
```

۴.۲ یک تابع بازگشتی برای شمارش تعداد آیتم‌های یک لیست بنویسید.

پاسخ:

```
def count(list):
    if list == []:
        return 0
    return 1 + count(list[1:])
```

٤.٣ حداکثر عدد را در یک لیست بیابید.

پاسخ:

```
def max(list):
    if len(list) == 2:
        return list[0] if list[0] > list[1] else list[1]
    sub_max = max(list[1:])
    return list[0] if list[0] > sub_max else sub_max
```

٤.٤ جستجوی دودویی را باز فصل ۱ به خاطر دارید؟ این الگوریتم، این یک الگوریتم تقسیم و حل نیز هست. آیا می‌توانید برای جستجوی دودویی، صورت پایه و صورت بازگشته را پیدا کنید؟

پاسخ: صورت پایه برای جستجوی دودویی، آرایه‌ای با یک آیتم است. اگر آیتمی که به دنبال آن هستید با آیتم موجود در آرایه مطابقت دارد، آن را پیدا کرده‌اید! در غیر این صورت، در آرایه نیست.

در حالت بازگشته برای جستجوی دودویی، آرایه را به نصف تقسیم می‌کنید، نیمی را کنار می‌گذارد و نیمی دیگر را با جستجوی دودویی فرا می‌خوانید.

هر یک از این عملیات بر اساس نماد O بزرگ چقدر طول می‌کشد؟

٤.٥ چاپ مقدار هر عنصر در یک آرایه.

پاسخ: $O(n)$

٤.٦ دو برابر کردن مقدار هر عنصر در یک آرایه.

پاسخ: $O(n)$

٤.٧ دو برابر کردن مقدار فقط اولین عنصر در یک آرایه.

پاسخ: $O(1)$

۴.۸ ایجاد جدول ضرب با تمام عناصر موجود در آرایه. بنابراین اگر آرایه‌ی شما $[2, 3, 7, 8, 10]$ باشد، ابتدا هر عنصر را در ۲ ضرب کنید، سپس هر عنصر را در ۳، سپس در ۷ و غیره ضرب کنید.

پاسخ: $O(n^2)$

فصل ۵

کدام یک از این توابع هش معتبر هستند؟

۵.۱

$f(x) = 1 \leftarrow$ تابع در ازای هر ورودی «۱» را برمی‌گرداند

پاسخ: بله

۵.۲

$f(x) = \text{rand}() \leftarrow$ در نوبت یک عدد تصادفی را برمی‌گرداند

خیر

۵.۳

$f(x) = \text{next_empty_slot}() \leftarrow$ ایندکس اسلات خالی بعدی در جدول هش را برمی‌گرداند

پاسخ: خیر

۵.۴

$f(x) = \text{len}(x) \leftarrow$ از طول رشته برای ایندکس استفاده می‌کند

پاسخ: بله

فرض کنید این چهار تابع هش را دارید که با رشته‌ها کار می‌کنند:
آ. «۱» را برای همه‌ی ورودی‌ها برگردانید.

ب. از طول رشته به عنوان ایندکس استفاده کنید.

پ. از اولین کاراکتر رشته به عنوان ایندکس استفاده کنید. بنابراین، تمام رشته‌هایی که با a شروع می‌شوند با هم هش می‌شوند و به غیره.

ت. هر حرف را با یک عدد اول نگاشت کنید: $a = ۱, b = ۲, c = ۳, d = ۴, e = ۵, f = ۶, g = ۷, h = ۸, i = ۹, j = ۱۰$ و غیره. برای یک رشته، تابع هش مجموع تمام کاراکترهای باقی‌مانده‌ی اندازه‌ی هش است. به عنوان مثال، اگر اندازه‌ی هش شما ۱۰ است و رشته‌ی آن «bag» است، ایندکس $۳ + ۲ + ۱ + ۷ = ۱۷\% \quad ۱۰\% \quad ۲۲\% = ۱۰$ است.

برای هر یک از این مثال‌ها، کدام تابع‌های هش توزیع خوبی ارائه می‌کنند؟ اندازه‌ی جدول هش را ۱۰ اسلات فرض کنید.

۵.۵ دفترچه تلفنی که در آن کلیدها اسمی و مقادیر شماره‌ی تلفن هستند. اسمی عبارتند از: Dan، Bob، Ben، Esther و.

پاسخ: توابع هش C و D توزیع خوبی ارائه می‌کنند.

۵.۶ نگاشتی از اندازه‌ی باتری به توان باتری. اندازه‌های AA، AAA و AAAA هستند.

پاسخ: توابع هش B و D توزیع خوبی ارائه می‌کنند.

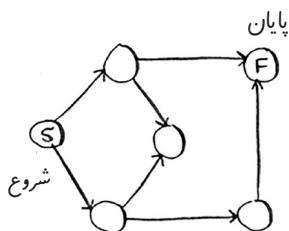
۵.۷ نگاشتی از عنوان کتاب به نویستده‌ها. عنوان‌های Maus، Fun Home و Watchmen هستند.

پاسخ: توابع هش B، C و D توزیع خوبی ارائه می‌کنند

فصل ۶

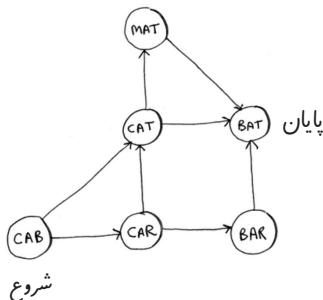
برای یافتن راه حل، الگوریتم جستجوی سطح اول را روی هر یک از این گراف‌ها اجرا کنید.

۶.۱ طول کوتاه‌ترین مسیر را از ابتدا تا انتهای پیدا کنید.



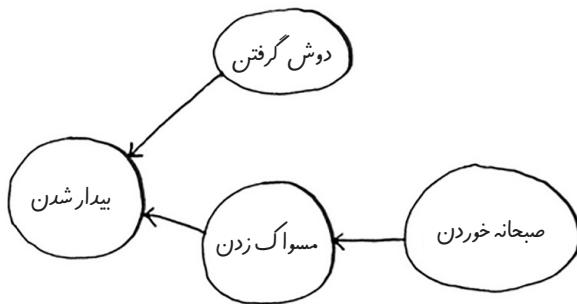
پاسخ: طول کوتاه‌ترین مسیر ۲ است.

۶.۲ طول کوتاه‌ترین مسیر را از «bat» تا «cab» بیابید.



پاسخ: طول کوتاه‌ترین مسیر ۲ است.

٦.٣ گراف کوچکی از برنامه‌ی صبحگاهی من به این شکل است:



درستی یا نادرستی هر یک از این سه لیست را مشخص کنید.

A.

۱. بیدارشدن
۲. دوش گرفتن
۳. صبحانه خوردن
۴. مسوآک زدن

B.

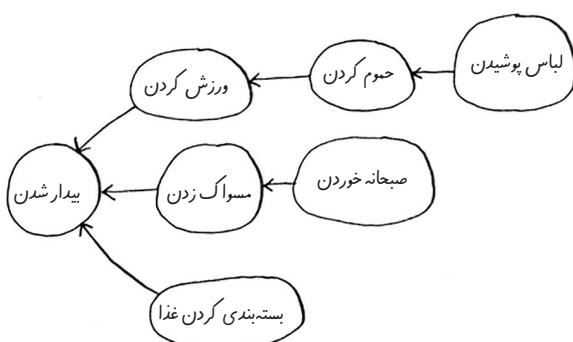
۱. بیدارشدن
۲. مسوآک زدن
۳. صبحانه خوردن
۴. دوش گرفتن

C.

۱. دوش گرفتن
۲. بیدارشدن
۳. مسوآک زدن
۴. صبحانه خوردن

پاسخ: A—نادرست. B—درست C—نادرست.

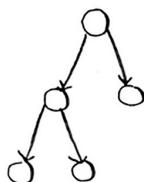
٦.٤ در اینجا یک گراف بزرگ‌تر وجود دارد. یک لیست صحیح برای این گراف تهیه کنید.



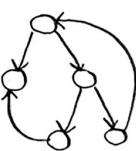
پاسخ: ۱-بیدار شدن. ۲-ورزش کردن ۳-دوش گرفتن ۴-مسواک زدن ۵-لباس
پوشیدن ۶-ناهار را بسته‌بندی کدن ۷-صبحانه خوردن

۶.۵ کدام یک از تصاویری که در ادامه آمده علاوه بر گراف درخت هم هستند؟

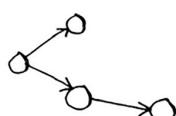
A.



B.



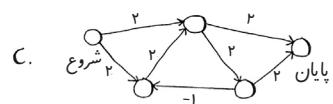
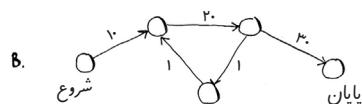
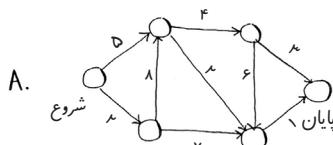
C.



پاسخ: A - درخت است، B - درخت نیست C - درخت است. آخرین مثال فقط یک درخت کناری است. درختان زیرمجموعه‌ای از گراف‌ها هستند. بنابراین یک درخت همیشه یک گراف ممکن است یک درخت باشد یا نباشد.

فصل ۷

۷.۱ در هر یک از این گراف‌ها وزن کوتاه‌ترین مسیر ابتدا تا انتهای چقدر است؟



پاسخ‌ها: A—60; B—8; C—A. سوال انحرافی. هیچ کوتاه‌ترین مسیری ممکن نیست (چرخه‌ی وزن منفی^(۱)).

فصل ۸

۸.۱ شما برای یک شرکت مبلمان کار می‌کنید و باید تولیدات این شرکت را به سراسر کشور ارسال کنید. شما باید بسته‌ها را در کامیون بار بزنید. بسته‌ها اندازه‌های مختلفی دارند و شما سعی می‌کنید حداکثر استفاده را از فضای کامیون داشته باشید. بسته‌ها را چگونه انتخاب می‌کنید تا از فضای داخل کامیون به شکل بهینه استفاده کنید؟ از استراتژی حریصانه استفاده کنید. آیا این الگوریتم شما را به راه حل بهینه می‌رساند؟

پاسخ: استراتژی حریصانه به این صورت است که بزرگ‌ترین جعبه‌ای را که در فضای باقی‌مانده جا می‌شود، انتخاب کنید و آن قدر این عمل را تکرار کنید که دیگر امکان جادaden جعبه‌های بیشتری وجود نداشته باشد. خیر، این راه حل بهینه را در اختیار شما قرار نمی‌دهد.

۸.۲ شما به اروپا می‌روید و هفت روز فرصت دارید تا جایی که زمان اجازه بدهد از مکان‌های گوناگون بازدید کنید. شما یک ارزش نقطه‌ای به هر مورد اختصاص می‌دهید (میزان اشتیاق به دیدن آن) و تخمین می‌زنید که بازدید هر کدام چقدر طول می‌کشد. چگونه می‌توانید مجموع ارزش را به حداکثر برسانید (بازدید از تمام مکان‌هایی که واقعاً دوست دارید از آن‌ها بازدید کنید)؟ با یک استراتژی حریصانه. آیا این راه حلی بهینه به شما می‌دهد؟

پاسخ: به انتخاب گزینه‌ای با بالاترین ارزش نقطه‌ای که هنوز در زمان باقی‌مانده می‌توانید انجام دهید، ادامه دهید. وقتی امکان آن نبود که کار بیشتری انجام دهید،

1. negative-weight cycle

متوقف شوید. خیر، این راه حل بهینه را به شما نمی‌دهد.

هر یک از الگوریتم‌های زیر را بررسی کنی که آیا یک الگوریتم حریصانه هستند یا خیر.
۸.۳ مرتب‌سازی سریع

پاسخ: خیر.

۸.۴ جست‌وجوی سطح اول

پاسخ: خیر.

۸.۵ الگوریتم دایجسترا

پاسخ: بله.

۸.۶ یک پستچی باید به ۲۰ خانه بسته تحویل بدهد. او باید کوتاه‌ترین مسیری را که به تمام ۲۰ خانه می‌رود، پیدا کند. آیا این یک مسئله‌ی ان‌پی کامل است؟

پاسخ: بله.

۸.۷ پیدا کردن بزرگ‌ترین دسته در یک مجموعه از افراد (یک دسته به مجموعه افرادی گفته می‌شود که همه‌ی افراد آن مجموعه یکدیگر را می‌شناسند). آیا این یک مسئله‌ی ان‌پی کامل است؟

پاسخ: بله.

۸.۸ شما مشغول رسم کردن نقشه‌ی ایالات متحده هستید و باید ایالت‌های مجاور را با

رنگ‌های مختلف رنگ‌آمیزی کنید. باید حداقل تعداد رنگ مورد نیاز خود را پیدا کنید تا هیچ دو ایالت مجاور هم رنگ نباشند. آیا این یک مسئله‌ی ان‌پی کامل است؟

پاسخ: بله.

فصل ۹

۹.۱ فرض کنید می‌توانید یک آیتم دیگر را سرقت کنید: پخش‌کننده‌ی MP3. وزن آن ۱ پوند است و ۱۰۰۰ دلار ارزش دارد. آیا باید آن را بدزدید؟

پاسخ: بله. سپس می‌توانید MP3 پلیر، آیفون و گیتار را به ارزش کل ۴۵۰۰ دلار بدزدید.

۹.۲ فرض کنید می‌خواهید به پیک‌نیک بروید. یک کوله پشتی دارید که ۶ پوند را در خود جای می‌دهد و می‌توانید آیتم‌های زیر را بردارید. به هر کدام یک عدد اختصاص داده شده و هر چه مقدار آن بیشتر باشد، آیتم اهمیت بیشتری دارد:

- آب، ۳ پوند، ۱۵
- کتاب، ۱ پوند، ۳
- غذا، ۲ پوند، ۹
- ژاکت، ۲ پوند، ۵
- دوربین، ۱ پوند، ۶

مجموعه‌ی بهینه‌ای از آیتم‌هایی که باید در سفر کمپینگ همراه خود داشته باشید چیست؟

پاسخ: آب و غذا و دوربین را انتخاب کنید.

۹.۳: برای محاسبه‌ی بزرگ‌ترین زیردنباله‌ی مشترک، میان blue و clues شبکه را ترسیم

و آن را پر کنید.

	C	L	U	E	S
B
L	.	۱	۰	.	.
U	.	.	۲	.	.
E	.	.	.	۳	.

پاسخ:

فصل ۱۰

۱۰.۱ در مثال نتفلیکس، فاصله‌ی بین دو کاربر مختلف را با استفاده از فرمول فاصله محاسبه کردید. اما همه‌ی کاربران با یک معیار ثابت به فیلم‌ها امتیاز نمی‌دهند. فرض کنید دو کاربر با نام‌های یوگی و پینکی سلیقه‌ی یکسانی در فیلم دارند. یوگی به هر فیلمی که دوست داشته باشد ۵ امتیاز می‌دهد، در حالی که پینکی گزیده‌تر عمل می‌کند و امتیاز ۵ را فقط برای بهترین فیلم‌ها نگه می‌دارد. آن‌ها سلیقه‌ی مشابهی دارند، اما طبق الگوریتم فاصله، همسایه نیستند. چه راهکاری برای استراتژی متفاوت آن‌ها در رتبه‌بندی در نظر می‌گیرید؟

پاسخ: می‌توانید از چیزی به نام نرمال‌سازی^۱ استفاده کنید. به میانگین رتبه‌بندی هر فرد نگاه کنید و از آن برای مقیاس‌بندی رتبه‌بندی آن‌ها استفاده کنید. برای مثال، ممکن است متوجه شوید که میانگین امتیاز پینکی ۳ است، در حالی که میانگین امتیاز یوگی ۵.۳ است. بنابراین رتبه‌بندی پینکی را کمی افزایش می‌دهید، تا زمانی که میانگین امتیاز او نیز به ۵.۳ برسد. سپس می‌توانید رتبه‌بندی آن‌ها را در همان مقیاس مقایسه کنید.

۱۰.۲ فرض کنید نتفلیکس گروهی از «اینفلوئنسرها» را گزینش کرده است. به عنوان مثال، کوئنتین تارانتینو و وس اندرسون اینفلوئنسرهای نتفلیکس هستند، بنابراین رتبه‌بندی این افراد بیشتر از یک کاربر عادی حساب می‌شود. چه تغییراتی برای سیستم پیشنهادات نیاز است تا نسبت به رتبه‌بندی اینفلوئنسرها جانبدارانه عمل کند؟

پاسخ: هنگام استفاده از KNN می‌توانید به رتبه‌بندی افراد تأثیرگذار اهمیت بیشتری بدهید. فرض کنید سه همسایه دارید: جو، دیو و وس اندرسون (یک اینفلوئنسر). آن‌ها به به فیلم پادوی گلف به ترتیب $3, 4, 5$ امتیاز داده‌اند. به جای اینکه فقط میانگین رتبه‌بندی آن‌ها را بگیرید ($\frac{3+4+5}{3} = 4$ ستاره)، می‌توانید به رتبه‌بندی وس اندرسون وزن بیشتری بدهید: $\frac{3+4+5}{5} = 4,4$ ستاره.

۱۰.۳ نتفلیکس میلیون‌ها کاربر دارد. در مثال قبلی به پنج همسایه‌ی نزدیک برای ساختن سیستم پیشنهادات نگاه کردیم. این تعداد خیلی کم است؟ یا خیلی زیاد است؟

پاسخ: خیلی کم است. اگر به همسایه‌های کمتری نگاه کنید، احتمال بیشتری وجود دارد که نتایج منحرف شوند. یک قانون خوب این است که اگر N کاربر دارید، باید به $\text{sqrt}(N)$ همسایه نگاه کنید.



