



© 1997–2009, Millennium Mathematics Project, University of Cambridge.

Permission is granted to print and copy this page on paper for non-commercial use. For other uses, including electronic redistribution, please contact us.

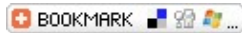
March 2010

Features



Of pancakes, mice and men

by Colva Roney-Dougal and Vince Vatter

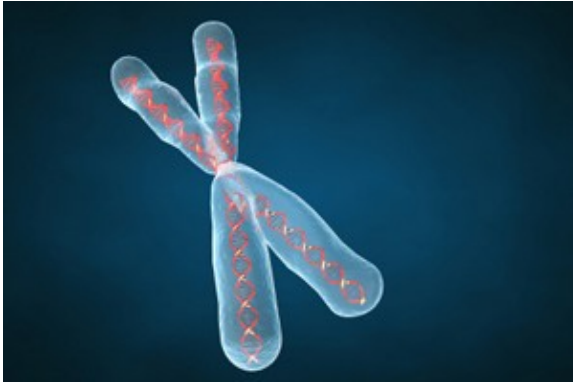


DNA normally evolves by tiny mutations, but every now and then something more radical occurs, and the order of entire genes along a chromosome is rearranged. Understanding these rearrangements is an important task, shedding light on the process of evolution. It boils down to solving a problem from pure mathematics: an instance of the many connections between maths and biology that have become increasingly apparent in recent years. In this article we'll explore this problem, embarking on a journey from waiters sorting pancakes, via one of the richest men in the world, to the genetic similarities of mice and humans.

Pancake sorting

In 1975 Jacob E. Goodman (under the pseudonym Harry Dweighter) posed the following problem:

The chef in our place is sloppy, and when he prepares a stack of pancakes they come out all different sizes. Therefore, when I deliver them to a customer, on the way to the table I rearrange them (so that the smallest winds up on top, and so on, down to the largest at the bottom) by grabbing several from the top and flipping them over, repeating this (varying the number I flip) as many times as necessary. If there are n pancakes, what is the maximum number of flips (in terms of n) that I will ever have to use to rearrange them?



Pancakes hold the secret to gene flipping along chromosomes.

We can model this as a *permutation sorting* problem. A *permutation* π of $1, 2, 3, \dots, n$ is just a list of the numbers $\{1, 2, 3, \dots, n\}$ that contains each number exactly once. For example, $3\ 1\ 4\ 2$ is a permutation of $1, 2, 3, 4$. If π is a permutation, we write $\pi(i)$ for the number in position i of the list π , so that if π is $3\ 1\ 2$ we can write $\pi(3) = 2$.

From a stack of pancakes we can make a permutation by numbering them from 1 up to n , by size, and then reading the stack from top to bottom. Our waiter's task is then to sort the stack by *prefix reversals*: he is given a permutation π , and wants to sort π to $1\ 2\ 3\ \dots\ n$. He does so by repeatedly changing the first part of the permutation. Suppose he has a stack of pancakes with the biggest on the top, the smallest in the middle, and the medium-sized one on the bottom. Then $\pi = 3\ 1\ 2$, and he should start by flipping the whole stack over, to get $2\ 1\ 3$. He can then flip just the first two pancakes to get $1\ 2\ 3$, which is correctly sorted.

Expressing this in general terms, the waiter first chooses a position i and then reverses the first i entries of π . Written out in full, this transforms π to

$$\pi(i)\ \pi(i-1)\ \dots\ \pi(1)\ \pi(i+1)\ \pi(i+2)\ \dots\ \pi(n).$$

The waiter will now carry on by choosing a new position, j say, and reversing the first j entries of this new permutation.

Harry Dweighter's question is therefore:

Given a whole number n , what is the maximum number of prefix reversals required to transform any permutation of length n into $1\ 2\ 3\ \dots\ n$?

It's never going to be harder than...

One easy sorting method proceeds as follows:

1. Find the biggest entry in the current permutation that's in the wrong place: let's say it's number k in position j .
2. Reverse the first j entries, so that k is now in position 1 .
3. Reverse the first k entries, so that the number k is now correctly in position k .
4. Go back to step 1 and repeat, until the whole permutation is sorted.



Want a neater stack?

As an example, suppose that we have four pancakes with $\pi = 1\ 3\ 2\ 4$. Now the biggest entry that's in the wrong place is 3, which is in position 2. Reversing the first two entries gives $3\ 1\ 2\ 4$. Then reversing the first three gives $2\ 1\ 3\ 4$, with 3 now in the correct place. Now the biggest entry in the wrong position is 2 in position 1. Reversing just the first entry amounts to doing nothing, then reversing the first two entries gives the required ordering $1\ 2\ 3\ 4$.

You can convince yourself that this method always works no matter how many and what ordering of pancakes you started with. So the answer to Harry Dweighthor's question is either equal to the number of reversals required to complete this method, or, if there is a better method, it is smaller. In other words, the number of reversals needed to complete this method is an *upper bound* on the difficulty of the problem.

With this method, each number takes at most two flips to end up in the right place, except that 1 will automatically be put in position 1 when we finally put 2 into position 2. So in total this method takes at most $2(n - 1)$ flips: this gives our initial upper bound on the difficulty of the problem.

And it's never going to be easier than...

What about a lower bound? Answering Harry Dweighthor's question is equivalent to finding the permutation that is *hardest* to sort and counting the number of flips required to sort it. We don't know what this hardest permutation is, but if we can find one that is reasonably hard, then this will give us a reasonably good lower bound on the difficulty of the problem: we know that sorting the hardest permutation is probably going to be harder than this, but hopefully not that much harder.

As a measure of "hardness" we look for consecutive entries that differ by 1. If this is the case for a pair of consecutive entries $\pi(i)$ and $\pi(i + 1)$ of a permutation π , we say that the entries are *adjacent*. For example, if π is $2\ 1\ 3\ 4$ then the entries $\pi(1) = 2$ and $\pi(2) = 1$ are adjacent, and so are $\pi(3) = 3$ and $\pi(4) = 4$, but $\pi(2) = 1$ and $\pi(3) = 3$ are not adjacent.

In the target permutation $1\ 2\ 3\ \dots\ n$ all entries are adjacent, so sorting a permutation is all about creating adjacent entries. It seems reasonable to suggest that the fewer adjacencies in a permutation, the harder it is to

And it's never going to be easier than...

Of pancakes, mice and men

sort. So let's suppose that the initial permutation π has no adjacent entries at all, say $3\ 1\ 4\ 2$, and work out the minimum number of flips it takes to sort it.

Each prefix reversal creates at most one pair of adjacent entries: either the beginning of the segment we flipped ends up next to an adjacent entry or it does not. So we need at least $n - 1$ prefix reversals to sort a permutation which starts with no adjacent entries. Notice also that if our permutation does not *end* with n , then we will need to bring n to the front at some point, and then reverse the whole permutation. This second flip certainly creates no new adjacencies so in fact at least n flips are required to sort a permutation without adjacent entries. Therefore, n is a lower bound on the maximum number of flips required to sort a permutation of length n .

This can be improved to a lower bound of $15n/14$, using a more complicated argument.



Bill Gates, an unlikely pancake flipper. Image: [World Economic Forum](#).

The first significant improvement on the *upper* bound of $2(n - 1)$ flips was proved by Bill Gates (of Microsoft fame) and Christos Papadimitriou, when Gates was an undergraduate at Harvard in the mid 70s. By classifying permutations into different types, they proved that every stack of pancakes (permutation) can be sorted by at most $(5n + 5)/3$ flips. This bound was not improved until 2009 by a team of seven researchers from the University of Texas at Dallas, who divided the problem into 2220 different cases to show that every permutation can be sorted by at most $18n/11$ prefix reversals.

So we have some grim news for our poor harried waiter: in the worst case, it will take him between $15n/14$ and $18n/11$ flips to sort n pancakes. For $n = 10$ this means between 11 and 16 flips.

Burnt Pancakes

Unfortunately, if the chef is even more careless and *burns* one side of the pancake, and our waiter needs to serve the pancakes with their unburnt side up, then his situation gets even bleaker. This can be modelled as a *signed permutation*, where we write over- and underlines on the entries to indicate their burnt sides.

Of pancakes, mice and men

For example, if the pancakes are given to the waiter as $\overline{4} \ \overline{1} \ \overline{2} \ \overline{3}$ (reading from top to bottom), with the burnt side of the first and last pancakes up and the burnt side of the other two down, we write this as the signed permutation $\overline{4} \ \underline{1} \ \underline{2} \ \overline{3}$. Our waiter's goal is to sort $\overline{4} \ \underline{1} \ \underline{2} \ \overline{3}$ into $\underline{1} \ \underline{2} \ \underline{3} \ \underline{4}$, which he can do with the following sequence of flips:

$$\begin{array}{l} \overline{4} \ \underline{1} \ \underline{2} \ \overline{3} \text{ flip all four pancakes} \\ \underline{3} \ \overline{2} \ \overline{1} \ \underline{4} \text{ flip the first pancake} \\ \overline{3} \ \overline{2} \ \overline{1} \ \underline{4} \text{ flip the first three pancakes} \\ \underline{1} \ \underline{2} \ \underline{3} \ \underline{4} \end{array}$$

Notice that when we flip a prefix of a signed permutation, the overlines become underlines (and *vice versa*), as well as the entries moving. In 1995, Daniel S. Cohen and Manuel Blum proved that every signed permutation can be sorted with at most $\frac{23n}{14}$ flips.

An extra plate

In response to the increasingly wayward chef, the waiter has a brilliant idea. He puts the plate of pancakes down, and uses a second plate. He first slides a batch of pancakes, still the same way up, onto the spare plate. Now he can flip the remaining pancakes as before – “flipping some initial part of this bottom stack” before returning the original pancakes back to the top of the pile.

So for example, given $\underline{1} \ \overline{2} \ \overline{3}$, he could slide off the top pancake onto the spare plate, then flip the second pancake so the burnt side is down, then return the top pancake to get a properly sorted stack of pancakes: $\underline{1} \ \underline{2} \ \underline{3}$. Mathematically, this is *sorting by reversals*, rather than sorting by prefix reversals. How much faster can he sort the pancakes this way?

Unlike in the prefix reversal situation, here we know the answer. It was shown in the mid-1990s that we can sort any permutation (stack of unburnt pancakes) in $n - 1$ reversals, and furthermore that there exist permutations, called *increasing oscillations* or *Gollan permutations*, which require exactly this number. These permutations have the form

$$\begin{array}{l} 3 \ 1 \ 5 \ 2 \ 7 \ 4 \ 9 \ 6 \ \dots \ n-6 \ n-1 \ n-4 \ n \ n-2 \ \text{if } n \text{ is even,} \\ 3 \ 1 \ 5 \ 2 \ 7 \ 4 \ 9 \ 6 \ \dots \ n-2 \ n-5 \ n \ n-3 \ n-1 \ \text{if } n \text{ is odd.} \end{array}$$

and consist of two subsequences (here coloured red and blue), both increasing.

If the chef burns one side of each pancake, the two plate system fares only slightly worse: every signed permutation can be sorted with at most $n + 1$ flips (only two more than if the pancakes aren't burnt), and again there are permutations which require precisely this many flips:

$$\begin{array}{l} \underline{n} \ \underline{n-1} \ \dots \ \underline{3} \ \underline{2} \ \underline{1} \ \text{if } n \text{ is even,} \\ \underline{2} \ \underline{1} \ \underline{3} \ \underline{n} \ \underline{n-1} \ \dots \ \underline{4} \ \text{if } n \text{ is odd and greater than } 3. \end{array}$$

Of pancakes, mice and men

(Have a go at sorting $\overline{3} \underline{2} \underline{1}$ with only 3 flips.)

Mice in the kitchen

It turns out that evolutionary processes flip genes in the same way as our waiter flips pancakes: soon we'll see a sequence of flips that takes us from mice to men.

Rather than a stack of pancakes, a chromosome is an ordered list of genes, a *sequence*. (Genes themselves are a subsequence of the long DNA molecule, the [Science Education Foundation](#) has a good illustration on how these relate to each other.) Normally genes evolve by tiny mutations, but every now and then something more radical occurs and entire genes get flipped: we're talking the two plate flipping technique here, as this can occur anywhere in the chromosome. Understanding gene flipping gives important clues about how an organism evolved. For example, two versions of a virus that are in fact quite closely related can look extremely different, if we look at the details within each gene. However, their similarities become more obvious if we concentrate instead on the broader picture: the sequence of genes along the chromosome. Furthermore, genes have beginnings and ends, so we can denote one direction of a gene by an underline and the reverse direction by an overline, enabling us to model chromosomes as signed permutations.



Close relatives.

Sticking with the food theme for now, let's look at a pair of vegetables. Despite outward appearances, cabbages and turnips are more similar than you might think. In fact, many of their genes are 99% identical in content, but they come in a different order in the two different vegetables. For example, if we look at one sequence of five genes shared between turnips and cabbages, and label them as $\underline{1} \underline{2} \underline{3} \underline{4} \underline{5}$ in the order they occur in turnips, then the cabbage genes are $\underline{1} \overline{5} \overline{4} \overline{3} \underline{2}$. By looking at the difference between these two permutations, we can get an estimate of how many reversals have occurred since cabbages and turnips evolved from their common ancestor, and this in turn gives us a rough idea of how long ago this division occurred. Our notion of the difference between these sequences of genes is precisely the waiter's notion of how hard it is to sort pancakes. In the cabbage/turnip case, this difference is 3:

$$\begin{array}{l} \underline{1} \overline{5} \overline{4} \overline{3} \underline{2} \text{ slide } \underline{1} \text{ on to plate; flip } \overline{5} \overline{4} \overline{3} \underline{2} \\ \underline{1} \overline{2} \underline{3} \overline{4} \underline{5} \text{ slide } \underline{1} \text{ onto plate; flip } \overline{2} \\ \underline{1} \underline{2} \underline{3} \overline{4} \underline{5} \text{ slide } \underline{1} \underline{2} \underline{3} \text{ onto plate; flip } \overline{4} \\ \underline{1} \underline{2} \underline{3} \underline{4} \underline{5} \end{array}$$

A second example occurs between humans and mice, in a case of eight shared genes in the X chromosome. If we label the genes of the mouse as

Of pancakes, mice and men

1 2 3 4 5 6 7 8

then these genes occur in human DNA as the signed permutation

$\bar{4}$ $\bar{6}$ 1 $\bar{7}$ $\bar{2}$ $\bar{3}$ $\bar{5}$ 8.

If we find the quickest way to sort the human genes, then it's likely that the midpoint of the sequence represents a common ancestor, from which humans and mice have diverged. In this example, transforming humans to mice takes a mere 6 flips:

$$\begin{array}{l}
\bar{4} \bar{6} \underline{1} \underline{7} \underline{2} \underline{3} \underline{5} \underline{8} \quad \text{flip } \bar{4} \bar{6} \underline{1} \underline{7} \underline{2} \underline{3} \\
\underline{3} \underline{2} \underline{7} \underline{1} \underline{6} \underline{4} \underline{5} \underline{8} \quad \text{flip } \underline{3} \underline{2} \underline{7} \underline{1} \\
\underline{1} \underline{7} \underline{2} \underline{3} \underline{6} \underline{4} \underline{5} \underline{8} \quad \text{slide } \underline{1} \text{ onto plate; flip } \underline{7} \underline{2} \\
\underline{1} \underline{2} \underline{7} \underline{3} \underline{6} \underline{4} \underline{5} \underline{8} \quad \text{slide } \underline{1} \underline{2} \underline{7} \text{ onto plate; flip } \underline{3} \underline{6} \\
\underline{1} \underline{2} \underline{7} \underline{6} \underline{3} \underline{4} \underline{5} \underline{8} \quad \text{slide } \underline{1} \underline{2} \text{ onto plate; flip } \underline{7} \underline{6} \underline{3} \underline{4} \underline{5} \\
\underline{1} \underline{2} \underline{5} \underline{4} \underline{3} \underline{6} \underline{7} \underline{8} \quad \text{slide } \underline{1} \underline{2} \text{ onto plate; flip } \underline{5} \underline{4} \underline{3} \\
\underline{1} \underline{2} \underline{3} \underline{4} \underline{5} \underline{6} \underline{7} \underline{8}
\end{array}$$

Thus it's likely that our common ancestor had gene order something similar to the third step in this sequence. Evolution performs the same process on our genes as a waiter with a stack of pancakes made by a sloppy cook, and the study of sorting signed permutations by reversals allows us to measure the genetic variation between different organisms. Let's give the last word to Mark Twain:

Training is everything. The peach was once a bitter almond; cauliflower is nothing but cabbage with a college education.

About this article

Colva Roney-Dougal is a lecturer in Pure Mathematics at the [University of St Andrews](#). She initially went to university to study English and Moral Philosophy, but got distracted along the way and wound up with a maths degree. Since then she has worked at Queen Mary, University of London and the University of Sydney. She is interested in symmetry and computation.

Vince Vatter is currently a John Wesley Young Research Instructor at Dartmouth College in New Hampshire. Born in Michigan, Vince received his PhD in mathematics from Rutgers University, and then spent two years across the pond as a research fellow at the University of St Andrews.

Colva and Vince have yet to make pancakes together.



Of pancakes, mice and men

Plus is part of the family of activities in the Millennium Mathematics Project, which also includes the [NRICH](#) and [MOTIVATE](#) sites.