

PROFESSOR: Hi, have you found it particularly difficult to find a specific item in your house? Let's say you're looking for a pair of gloves, but you just can't find it, and you spend your entire afternoon looking for it. Well, you've just encountered the same problem that companies like Google or Microsoft encounter every single day. And that's the problem of search.

Just like a house which stores thousands of different items, Google stores 45 billion index pages of information. If every page was a sheet of paper and we stack them up real high, we'll create a tower 600 times taller than Mount Everest. Well, how can Google find my results so quickly when I find it so difficult to find a pair of gloves?

Well, searching on Google is kind of like looking for person in a big school. Let's say you're looking for James in a row of classrooms. One of the easiest method would be to go to every classroom nearest to you until you find James. There's a better method called binary search.

Now, let's say the students were arranged from a to z, one in each classroom. We would then go to the middle room first and see if James is there. And if James isn't there we will look at the first letter in the name, and if the letter is before j, we head to the right, if not, we head to the left. We would then approach the middle room in the newly sectioned area. Eventually we will repeat the process over and over again until we find James.

Now, this is just like the first method, but it's at a much, much faster rate. How much faster would that be? Well, that depends on the number of students in the school. Let's say there are 500 students and we're looking for one. It will take about 80 minutes in the first method, But one and a half minutes with binary search. But let's say there are 1,000 students in the school. It would take 160 minutes with the first method, but 1.6 minutes with the second method. Now that's a whole lot of difference.

So a name is just a word. But Google searches a combination of words, making it a little bit more complicated. So just like how we identified the first letter of each alphabet of the name, Google identifies 200 unique factors making your search terms faster.

Well, if you recall, the effectiveness of binary search depends on the prearrangement of data. And that's why computer scientists are actively looking for ways to sort, manage, and eventually retrieve data faster and better. In the same way, the TV remote goes near the TV, the shoes go to the shoe rack, the coats go into the cupboard, and the winter gloves go into

the winter jacket. Aha, so that's where my gloves are.