Unit 1 - Activity 6 - Linear & Binary Search Reading

Binary Search

Think about how you look for a name in the phone book. You open it up somewhere in the middle and then see if you need to go forward or backward. We need to give the computer more precise instructions than this so we will tell it to divide the list exactly in half and compare the middle item to our target. If the middle item is smaller than our target, then we can look at the top half of the list. If it is bigger than our target, we will tell the computer to look in the bottom half of the list. We call this kind of search **Binary Searching** because we always divide the number of items we will search in half.

Here is the algorithm:

1. Set the list to be the whole list
2. Find the middle value of the list
3. If the middle value is equal to the target then we declare victory and stop.
4. If the middle item is less than the target, then we set the new list to be the upper half of the old list and we repeat from step 2 using the new list.
5. If the middle value is greater than the target, then we set the new list to be the bottom half of the list, and we repeat from step 2 with the new list.

A binary search is when you start with the middle of a sorted list, and see whether that's greater than or less than the value you're looking for, which determines whether the value is in the first or second half of the list. Jump to the half way through the sublist, and compare again etc. This is pretty much how humans typically look up a word in a dictionary (although we use better heuristics, obviously - if you're looking for "cat" you don't start off at "M"). In complexity terms this is an O(log n) search - the number of search operations grows more slowly than the list does, because you're halving the "search space" with each operation.

Linear Search

A linear search looks down a list, one item at a time, without jumping. In complexity terms this is an O(n) search - the time taken to search the list gets bigger at the same rate as the list does.

Suppose we want to write an algorithm to get the computer to search through a list of numbers (like a list of airline flight numbers). The list is ordered from the smallest to the biggest. The easiest way to find our number is to start at the beginning and compare our number (which we will call the **target**) to each number in the list. If we reach our target, then we are done. This method of searching is called **Linear Searching**.

Here is the algorithm:

1. Start with the first item in the list.
2. Compare the current item to the target
3. If the current value matches the target then we declare victory and stop.
4. If the current value is less than the target then set the current item to be the next item and repeat from 2.

Now, how do we tell the computer our algorithm? We need to give instructions to the computer in a language (code) it understands. When you understand the symbols in the language, the algorithm looks pretty similar to the code.