

~ FUNNY HALF-HALF ~

Written by:
Sridhar Iyer,
IIT Bombay

It was close to midnight when there was a loud knock on the door. "Come in!" said Sherlock Holmes, the famous private detective. Inspector Lestrade entered. "Holmes, we need your help," he said, in great agitation. "Pray take a seat," said Holmes, "Watson, be good enough to make some room for him."

Lestrade flung himself into the sofa beside Watson and started speaking at once. "You remember the arrest of the criminal Moriarty. His trial is tomorrow and an important document of evidence has somehow disappeared. The document had the details of all his associates and bank accounts. We know that there is a copy of the document in his house. We have got a search warrant and have been searching the house since yesterday without any luck so far. We need to find the document before morning, otherwise he will go scot-free."

"And what seems to be the problem?" murmured Holmes, settling himself down in his armchair and closing his eyes. Lestrade glanced at the lounging figure and continued ruefully, "We have looked everywhere. Inside the tables, under the mattresses, behind the pictures and all the usual places. We even broke into Moriarty's safe but only found a piece of paper with the words 'funny half-half' on it."

"A man of Moriarty's intelligence will not conceal something in any of the usual places. What are the places you haven't searched yet?" asked Holmes. "Only the library is left and it has thousands and thousands of books. Moriarty himself is a prolific writer" said Lestrade. "H'm. the library; document; funny half-half", mused Holmes falling into a reverie. Lestrade and Watson were

discussing the problem in hushed tones when Holmes suddenly sprang to his feet and cried "Let's go."

They reached Moriarty's house and went straight to the huge library. Holmes picked up the library catalog which had a list of all the books and quickly glanced through it. He then went to one section of the library and started prowling up and down the aisles of book shelves. Watson suggested, "If the document is likely to be hidden in a book, why don't we search them?" "Taking down the books one by one and searching all of them will take many days. Unless we are extremely lucky, we will not find the document before morning" objected Lestrade. "These books also seem to be filed randomly. Why cant Moriarty at least keep them in a sorted manner." he complained.

Meanwhile Holmes had stopped in front of one of the shelves and after a close scrutiny of the books, pulled out three books from the different rows of the shelf. "Here, Watson. Be kind enough to glance through this book." said Holmes handing one of the books to Watson. Lestrade took the other and Holmes the third. Lestrade was skeptically turning the pages when Watson suddenly shouted, "Here it is! We have found the document!".

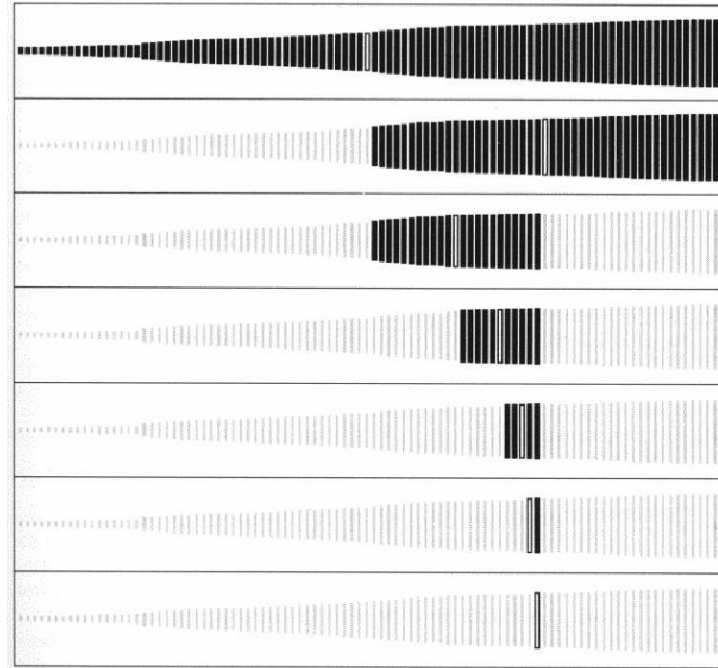


"Holmes! This is brilliant. How did you do it?" cried Watson. Holmes smilingly replied, "You have the clue in 'funny half-half'. Moriarty would not have kept such a paper in his safe unless it meant something. Also, if you notice in the catalog, the books in the library are indeed sorted. The only difference is that instead of sorting them alphabetically on the author name or the title, Moriarty has sorted them according to their subject." "Ah! I get it", said Lestrade, "The clue 'funny' means that we have to search among the books in the humour section".

"But there are many books in the humour section. How did you locate the book so quickly?", asked Watson. "That's from the clue 'half-half'", replied Holmes. "It could have meant one of two things. The first 'half' might refer to the position of the book while the second 'half' indicates the position of the document within the book. Hence I picked out the book in the exact middle of the section and gave it to you. Alternately, the entire 'half-half' might refer to only the position of the book, so I picked out the other two books." he stopped enigmatically.

"Holmes! Don't keep me guessing. How did you identify which were the other two books?" cried Watson despairingly. "OK, suppose there are 12 books." said Holmes, "The first 'half' corresponds to the 6th book. The second 'half' might indicate the 3rd book (which is half-way from the 1st till the 6th) or the 9th book (which half-way from the 6th till the 12th). Elementary, my dear Watson", he concluded.

Interestingly, similar principles called 'search techniques' are used in many areas of computer applications, wherever there is a need to locate information quickly. The simplest of these is called 'sequential search' wherein you go through the entire list of information one by one till the desired item is found. This works well when the list of items is small but it may take a long time when the list is very long.



One common improvement to sequential search is called 'binary search'. This works well when the list of items is sorted in some order, such as alphabetical or numerically ascending or descending. The basic idea of binary search is to compare the middle item of the list with the item to be found. If this is itself the item, then the search is successful. If the middle item is smaller than the item to be found, then we repeat the search among the items to the right (between the middle and the last items). If the middle item is greater than the item to be found, then we repeat the search among the items to the left (between the first and middle items). This process is similar to the half-half in the above story and it continues till the desired item is found.

Another example of binary search in action is a simple guessing game in which a player selects a number between 1 and N. Another player has to guess the number using only questions answered with yes or no. Supposing N is 20 and the number 8 is selected, the game might proceed as follows: Is the number greater than 10? (No); Is the number greater than 5? (Yes); Is the number greater than 8? (No); Is the number greater than 7? (Yes); Therefore, the number must be 8. At each step, we

choose a number right in the middle of the range of possible values. For example, once we know the number is greater than 5, but less than or equal to 10, we choose 8 as the next number since it is in the middle of the range [6, 10].

Binary search works nicely when the list of items is sorted in some order. How about when we are searching for a particular word in a page? How about when we are using Google to search for information on the Internet? The Google search engine rose to prominence around 2001. A search engine is basically a program designed to help find information stored on computer systems such as the World Wide Web (WWW). A search engine uses a more advanced principle called 'indexing'. First the engine gathers and stores the information from a large number of web pages. Then the contents of each page are analyzed to determine how it should be indexed (for example, words are extracted from the titles, headings). The main idea of indexing is similar to classifying and sorting books in a library according to their subject. This classification data about the web pages is called the index. So the index is something similar to the library catalog in the above story or the Index found at the back of a book. This index can be later used to find the pages relevant to a query. Now the search engine is ready. Later when a user comes to the search engine and makes a query, typically by giving keywords, the engine looks up the keywords in the index and provides a listing of best-matching web pages. While searching may sound simple, there are a lot of intricacies. For example, what happens to the size of the index when there are millions of Web pages? There may be many pages that include a particular keyword and match a given query. How does a search engine decide which of these pages constitute the "best" results for the given query?

Searching: The act of trying to find something or someone. A search engine is a system dedicated to the search and retrieval of information. Most commonly, it is a program designed to help users of the Internet locate information on the World Wide Web.

Indexing: The process of converting a collection of data into a form suitable for easy search and retrieval. It is often used to refer to the automatic selection of meaningful words (or keywords) from a website and storing them in a list or database that can be used by a search engine.
Keyword: A word or phrase that you use to search for information about a topic. The keyword describes a specific area of information.

Some interesting related websites are:

For Sherlock Holmes stories -

<http://www.gutenberg.org/etext/1661>

For Searching techniques -

http://en.wikipedia.org/wiki/Binary_search

http://en.wikipedia.org/wiki/Search_engine

<http://www.animal.ahrgr.de/en/Animation4.html>