Mini Research Projects

These mini-research projects cover the more accessible topics we have covered, and that we will cover, in this Computer Science Summer School. There is plenty of information in textbooks and online, and I encourage you to use both: online material is great for getting started quickly but textbooks have a lot more detail and will allow you to show that you have understood the topics in depth.

Pick any of these topics (noting that item 6 could be 5 separate projects and item 15 could be 4 separate projects!). I hope you enjoy researching these areas of Computer Science – they will certainly set you up very well for further study of Computer Science at university.

John Fawcett, July 2023

1. The lectures showed the "anatomy of a simple computer" with separate memory buses and Peripheral I/O buses. Why separate these? Is there ever a reason to merge them into one? Is there ever a reason to split them further (e.g. two/more memory buses or two/more I/O buses)?
2. I left the representation of floating point numbers as an interesting exercise for a mini research project. IEEE754 is an international standard and specifies (among many things) one way to represent floating point numbers. Find out how it works and, besides numbers, what other "mathematical objects" it can represent. What happens in IEEE754 if you ask the CPU to do something that is not mathematically defined, such as to calculate the square root of -2 or to find the arcsin of 14?
3. We did not consider a data cache in the lectures.
    a. Why might it be more difficult to build a data cache (D-cache) than an I-cache?
    b. Under what circumstances would a D-cache speed up a program?
    c. How could you decide when to replace an entry in a D-cache?
4. In the days of MS DOS, Microsoft's Disk Operating System, the extended memory manager (emm386.exe) was used to implement 'solution 1' on slide 3. Research how this worked and what shortcomings it had compared to the paging approach described in lectures. Does the DOS approach have any advantages?
5. We said that the operating system keeps a translation table for each process but we did not explain how it remembers where it put each translation table! Research the unix operating system's concept of *process control block (PCB)* and write up what is stored there and what it's used for.
6. We talked about keeping commands for the hard disk in a *queue* data structure, which implements first-come-first-served (FCFS). Each of the following would be an interesting research project – you could do any one of these in detail for your project, or more briefly look at all 5 for your mini-research project:
    a. Compare FCFS to the SCAN algorithm.
    b. Compare FCFS to the C-SCAN algorithm.
    c. Compare FCFS to the LOOK algorithm.
    d. Compare FCFS to the CLOOK algorithm.
    e. Compare FCFS to the FSCAN algorithm.
7. Research how the Windows NT (4.0) CPU scheduling algorithm works.
8. How does the Unix dynamic priority scheduler work?
9. Research *Named Pipes*. How are they different from pipes? What problem do they solve? Show how to use them with pseudo-code or a real program.

10. For two programs you use regularly, figure out how they use interprocess communication. Explain using pseudo-code (or a real program) how you would implement them.
11. Research and write-up how the CSMA/CD Datalink protocol manages medium access control and collisions. Do any common network technologies use CSMA/CD?
12. Research and write-up how the CSMA/CA Datalink protocol manages medium access control and collisions. Do any common network technologies use CSMA/CA?
13. Research and write-up how the Token Ring Datalink layer works.
14. How is the Address Resolution Protocol (ARP) used to translate Layer-3 IPv4 addresses into Layer-2 Ethernet station addresses (ESA/MAC addresses)?
15. We can "use Physics to invent random bits". There are many different ways to do this. Four possible mini-research projects might investigate how to do it based on:
    a. A radioactive sample and a geiger counter
    b. A very sensitive thermometer
    c. Quantum physics
    d. The timings of packets of data flowing through the Internet