HW #4
CSE 378

(Please note that this homework needs to be a paper submit in class on Dec 5th.)

1. (Problem 7.9 in Hennessy and Patterson-3rd Edition)
   Here is a series of address references given as word addresses: 2, 3, 11, 16, 21, 13, 64, 48, 19, 11, 3, 22, 4, 27, 6 and 11.
   a. Assuming a Direct-mapped cache with 16 one-word blocks that is initially empty, label each reference in the list as a hit or a miss.
   b. Now assume that the cache is a 2-way set associative, one-word blocks, total 16-word capacity.

   If there is a cache-miss, indicate what type of miss it is (compulsory/conflict (with what)/capacity miss). For each of your answers to both parts a & b fill in the table shown below. Show the final contents of the cache in each case.
   Assume a Least Recently Used replacement policy where applicable.

   (a)

Address	Hit/Miss	Type
2		
3		
11		
16		
21		
13		
64		
48		
19		
11		
3		
22		
4		
27		
6		
11		

(b)

Address	Hit/Miss	Type
2		
3		
11		
16		
21		
13		
64		
48		
19		
11		
3		
22		
4		
27		
6		
11		

2. You've been tasked to upgrade a mission critical function in a visiting computational linguistics professor's code. The function (as defined in find_letter.h) searches for a string with a given letter somewhere in it. The current version runs too slowly. We think it may have something to do with locality, but we're not certain. You will find the code linked off from the schedule.

   Your tasks:
   1. Rewrite the function to run faster.
   2. Explain why the old function was slow.
   3. Explain why yours is faster.
   4. Explain in what circumstances your code would be slower than the provided code.

   These tasks should be printed out and attached to the other part of your assignment.

   Provided is some code to load the specially formatted files that the professor likes to use in his research. You are not expected to grok any code outside of the function you are to rewrite. Also provided is a simple main function for testing purposes.

   The file format is as follows:

   int (describes how long the longest string is)
   words
   words
   ...
   lastword

   The file words.txt is a simple example of this type of file.

   Due to the...odd string format that the professor uses, you may not use any of the standard c string library functions. The format he uses allows embedded nulls for some reason.

3. Do you see any problems with the indicated decoding of addresses for caches? If so, what is it?

Index	Tag	Offset
31		0

4. With respect to TLBs, indicate whether the three fields (Tag, Index and Offset) shown here should be Physical or Virtual addresses and also state the reason. For your answer, fill in the table provided below.

Tag	Index	Offset

Table:

Field	**Physical/Virtual**	**Reason**
Tag		
Index		
Offset		

5. As we saw in class, page tables require fairly large amounts of memory, even if most of the entries are invalid. One solution is to use a hierarchy of page tables. The virtual page number can be broken up into two pieces (described in Fig 7.20 on pg 513) , a “page table number” and a “page table offset”. The page table number can be used to index a first-level page table that provides a physical address for a second-level page table. The page table offset is used to index into the second-level page table to retrieve the physical page number. One way to arrange such a scheme is to have the second-level page tables occupy exactly one page of memory. (References to the book are for the 3rd edition).

   Assuming a 32-bit virtual address space with 4KB pages and 4 bytes per page table entry, how many bytes will each program need to use to store the first-level page table (which must always be in memory)? Provide information on how to decode the address (i.e. what bits for the page table number, pagetable offset and physical page offset). Explain your solution.