CS 3423 Operating Systems

Fall Semester 2019

Prof. Pai H. Chou

Assignment 10

(Scope: Ch. 9)

Due Date: Sunday, November 17 (24), 2019, 11:59pm

## 1. Definitions and Short Answers

1. In the base-and-limit registers scheme such as pictured below,


	1. What are the **addresses spaces** of the three processes? 256000~300040, 300040 ~ 420940, 420940~880000
	2. Who loads the values of the base and limit registers? loader/linker
	3. What happens when a process attempts to address memory location outside its address space? trap to OS
2. Does a compiler determine the absolute address of the memory references? not always If not, what kind of address does it generate? relocatable address
3. Which tool (compiler, linker, loader) determines the actual address to access memory for **compile-time binding**? linker What must be known in order to perform binding? base address
4. For **load-time binding**, what does the loader actually do? fill in address And what kind of code must the compiler-linker generate? relocatable code
5. If **execution-time binding** is used, what assumption can be made about the base address of the process? need hardware support Who calculates the base address + offset to form the actual memory address MMU, and at what time?move between memory
6. What is the meaning of a **logical address**? addresses generated by generated by the CPU
7. What is the meaning of a **physical address**? address seen by the memory unit How is logical address different or same as the physical address for
	1. compile-time binding same as the physical address
	2. load-time binding same as the physical address
	3. execution-time binding remap
8. What does **dynamic loading** mean? load code into memory on demand Why is it a good idea? Better memory-space utilization
9. What is a problem with the combination of **dynamic loading with static linking**? still could load in multiple copies
10. What is the mechanism for **dynamic linking**? stub, replaces itself with address of the routine, and executes the routine How does it address the problem of dynamic loading with static linking? check if referred library is in memory
11. What are two partitioning schemes of doing contiguous allocation? fixed/variable
12. What is a limitation imposed by **fixed partition** scheme of contiguous allocation? degree of multiprogramming is bounded by the #partitions
13. What is a problem caused by **variable partition** scheme of contiguous allocation? Holes of different sizes are scattered in memory
14. What are the three common schemes for variable-partition contiguous memory allocation? first fit, best fit, worst fit Which one is generally faster? first fit Which ones may need to search the entire list of holes? best fit, worst fit
15. What is the rationale for using **worst-fit**? Produces the largest leftover hole
16. What is **external fragmentation**? total memory space exists to satisfy a request, but not contiguous
17. What is **internal fragmentation**? free memory internal to a partition, but too small to be used
18. In the following two scenarios, identify what kind of fragmentation it is. Gray color means allocated, white color means free.
	1. **external fragmentation**
	
	2. **internal fragmentation**
	
19. **Compaction** may be a way to reduce fragmentation, but it is not always possible. What kind of **address binding** (compile time, load time, execution time) is required for compaction to work? execution time Even if it is supported, under what situation may it still not work correctly?backing store
20. Paging is a way of organizing noncontiguous allocation. What does **noncontiguous** mean in this case?Avoids problem of varying sized memory chunks
21. Does paging use variable partition or fixed partition? fixed partition
22. Does paging have internal fragmentation? yes external fragmentation? no
23. Why does paging tend to have less external fragmentation than contiguous allocation? fixed size
24. What is a **page** and what is a **frame**? page: logical block, frame: physical block
25. To support paging, how should a logical address be divided so that it can be mapped to a physical address? page number and page offset Which part of the address is mapped and which part is the same? page offset
26. What are the advantages and disadvantages of a
	1. smaller page size pro: smaller internal fragmentation, con: need more entries in page table
	2. larger page size pro: less entries in page table , con: bigger internal fragmentation

? Express in terms of page table size (i.e., number of entries) and fragmentation (say which kind fragmentation).

1. What does a **page table** map from and to? from logical memory space to frame What does a **frame table** map from and to? frame to page Do you need **one per process** or **one for the entire system**? frame table: **one for the entire system, page table: one per process**
2. What are two **registers** that identify a page table? Page-table base register, Page-table length register
3. What does TLB stand for? for fast access Is it hardware or software, hardware and what does it do?OS loads page-table entry into the TLB for faster access next time
4. What happens on a TLB miss? search page table and update TLB What happens if all TLB entries are occupied? replace other entry
5. What is ASID in a TLB entry? identifier for process Is it mandatory?no What are its benefits? TLB can contain entries for different processes
6. What is the purpose of bits for indicating **access rights** of a page? memory protection
7. Why would some pages be marked **invalid** in a page table entry? page not in process’s logical address
8. Why is the reason for using a PTLR (page table length register)? save memory when many entry are unused Isn't the size of a page table fixed? NO
9. Can **shared memory** between processes be supported? yes Do different process need to use the same virtual address? NO Do the virtual addresses of the different processes map to the same physical address? YES
10. How does a 2-level **hierarchical page table** scheme divide the logical address into different fields, and what are the steps in looking up the frame number? 1. highest order bit index to find page table T 2. next highest bit index into T' to find page of data
11. How does a **hashed page table** store its entries? page table contains linked list of elements hashed to same location
12. Is a **clustered page table** a form of a hashed page table? yes How is it more economical? entry refers to several pages

## 2. Programming Exercise

In this programming exercise, you are to implement algorithms for contiguous memory allocation, similar to malloc() and free() in the standard library (stdlib).

malloc(), for memory-allocate, is a stdlib function for dynamically allocating a contiguous block of memory. The parameter is the number of bytes to allocate. The return value is the pointer (here an int in Python) to the allocated memory block, or None if it cannot be allocated, possibly due to memory fragmentation.

free() will free a previously allocated memory (as returned by a previous call to malloc()). The textbook talks about three policies: **First-Fit**, **Best-Fit**, and **Worst-Fit**. You are to implement all three policies in Python. Use the following API:

class MemAlloc:

 \_POLICIES = {'FirstFit', 'BestFit', 'WorstFit'}

 def \_\_init\_\_(self, totalMemSize, policy = 'BestFit'):

 if not policy in MemAlloc.\_POLICIES:

 raise ValueError('policy must be in %s' % MemAlloc.\_POLICIES)

 self.allocation = { } # use this dictionary to map allocated

 # pointer to the allocated size

 # keep a list of holes, which are tuples with (pointer, size)

 self.holes = [(0, totalMemSize)] # sorting by pointer

 # your own code here …

 def malloc(self, reqSize):

 '''return the starting address of the block of memory, or None'''

 # your code here

 def free(self, pointer):

 '''free the previously allocated memory starting at pointer'''

 # your code here

You will find some test cases in the [template](https://drive.google.com/file/d/1CF86oT4I6TDim0eqqaIiEUjjZXrFB2CT/view?usp=sharing) file. Rename it memalloc.py

### 2.1 malloc(size\_t size)

malloc() and free() use of two data structures:

* list of holes (self.holes), which consists of tuples (*address*, *size*)
* mapping from allocated addresses to sizes (self.allocation)

malloc() will iterate over the list of holes, kept in sorted order by address.

* if the policy is **First-Fit**, then it uses the first hole that is big enough to serve the requested size
* if the policy is **Best-Fit**, then it continues looking for the smallest hole that is big enough to serve the requested size.
* if the policy is **Worst-Fit**, then it looks for the biggest hole that can serve the requested size.

If no holes are big enough, then malloc() returns None.

But if there is one hole that can work, then

* if the chosen hole is used up completely by this malloc() request, then it should be deleted from the list of holes.
* Otherwise, if there is still some remaining unused space in this hole, then update the hole’s address and size.

In any case, the new allocation should be recorded in the self.allocation dictionary. Use the address as the key and size as the value. Finally, return the address for the newly allocated memory chunk.

You should test this part thoroughly, possibly with your own test cases, before proceeding to the next part.

### 2.2 free(void\* p)

free() is the inverse operation of malloc(). It takes a previously allocated address as parameter, looks up the size from the allocation, and

* update the holes list
* delete the freed entry from the allocation dictionary

Updating the holes list is potentially the tricky part, because there are several possible cases. Let (*p*, *s*) denote the pointer to the freed block and the size of the block to be freed.

* if empty holes list: just add (*p*, *s*) to the holes list.
* if (*p*, *s*) goes before the first hole on the list:
	+ if (*p*, *s*) and first hole are disjoint, just prepend (*p*, *s*) to holes list
	+ if contiguous, then merge (*p*, *s*) into the first hole by updating the first hole’s starting address and size.
* if (*p*, *s*) goes after the last hole on the list:
	+ mirror image to the “before the first hole”
* if (*p*, *s*) goes between hole [i] and hole [i+1]:
	+ if all three are contiguous, merge all three (and delete hole [i+1])
	+ if (*p*, *s*) contiguous with [i], merge them
	+ if (*p*, *s*) contiguous with [i+1], merge them
	+ if all three are disjoint, insert (*p*, *s*) between [i] and [i+1] on the list

Here is sample output:

a=malloc(10):

 FirstFit symbols={'a': 0} holes=[(10, 10)] allocation={0: 10}

 BestFit symbols={'a': 0} holes=[(10, 10)] allocation={0: 10}

 WorstFit symbols={'a': 0} holes=[(10, 10)] allocation={0: 10}

b=malloc(1):

 FirstFit symbols={'a': 0, 'b': 10} holes=[(11, 9)] allocation={0: 10, 10: 1}

 BestFit symbols={'a': 0, 'b': 10} holes=[(11, 9)] allocation={0: 10, 10: 1}

 WorstFit symbols={'a': 0, 'b': 10} holes=[(11, 9)] allocation={0: 10, 10: 1}

c=malloc(4):

 FirstFit symbols={'a': 0, 'b': 10, 'c': 11} holes=[(15, 5)] allocation={0: 10, 10: 1, 11: 4}

 BestFit symbols={'a': 0, 'b': 10, 'c': 11} holes=[(15, 5)] allocation={0: 10, 10: 1, 11: 4}

 WorstFit symbols={'a': 0, 'b': 10, 'c': 11} holes=[(15, 5)] allocation={0: 10, 10: 1, 11: 4}

free(c)

 FirstFit symbols={'a': 0, 'b': 10} holes=[(11, 9)] allocation={0: 10, 10: 1}

 BestFit symbols={'a': 0, 'b': 10} holes=[(11, 9)] allocation={0: 10, 10: 1}

 WorstFit symbols={'a': 0, 'b': 10} holes=[(11, 9)] allocation={0: 10, 10: 1}

free(a)

 FirstFit symbols={'b': 10} holes=[(0, 10), (11, 9)] allocation={10: 1}

 BestFit symbols={'b': 10} holes=[(0, 10), (11, 9)] allocation={10: 1}

 WorstFit symbols={'b': 10} holes=[(0, 10), (11, 9)] allocation={10: 1}

d=malloc(9):

 FirstFit symbols={'b': 10, 'd': 0} holes=[(9, 1), (11, 9)] allocation={10: 1, 0: 9}

 BestFit symbols={'b': 10, 'd': 11} holes=[(0, 10)] allocation={10: 1, 11: 9}

 WorstFit symbols={'b': 10, 'd': 0} holes=[(9, 1), (11, 9)] allocation={10: 1, 0: 9}

e=malloc(10):

 FirstFit symbols={'b': 10, 'd': 0, 'e': None} holes=[(9, 1), (11, 9)] allocation={10: 1, 0: 9}

 BestFit symbols={'b': 10, 'd': 11, 'e': 0} holes=[] allocation={10: 1, 11: 9, 0: 10}

 WorstFit symbols={'b': 10, 'd': 0, 'e': None} holes=[(9, 1), (11, 9)] allocation={10: 1, 0: 9}

------------------------

a=malloc(3):

 FirstFit symbols={'a': 0} holes=[(3, 17)] allocation={0: 3}

 BestFit symbols={'a': 0} holes=[(3, 17)] allocation={0: 3}

 WorstFit symbols={'a': 0} holes=[(3, 17)] allocation={0: 3}

b=malloc(6):

 FirstFit symbols={'a': 0, 'b': 3} holes=[(9, 11)] allocation={0: 3, 3: 6}

 BestFit symbols={'a': 0, 'b': 3} holes=[(9, 11)] allocation={0: 3, 3: 6}

 WorstFit symbols={'a': 0, 'b': 3} holes=[(9, 11)] allocation={0: 3, 3: 6}

c=malloc(2):

 FirstFit symbols={'a': 0, 'b': 3, 'c': 9} holes=[(11, 9)] allocation={0: 3, 3: 6, 9: 2}

 BestFit symbols={'a': 0, 'b': 3, 'c': 9} holes=[(11, 9)] allocation={0: 3, 3: 6, 9: 2}

 WorstFit symbols={'a': 0, 'b': 3, 'c': 9} holes=[(11, 9)] allocation={0: 3, 3: 6, 9: 2}

d=malloc(5):

 FirstFit symbols={'a': 0, 'b': 3, 'c': 9, 'd': 11} holes=[(16, 4)] allocation={0: 3, 3: 6, 9: 2, 11: 5}

 BestFit symbols={'a': 0, 'b': 3, 'c': 9, 'd': 11} holes=[(16, 4)] allocation={0: 3, 3: 6, 9: 2, 11: 5}

 WorstFit symbols={'a': 0, 'b': 3, 'c': 9, 'd': 11} holes=[(16, 4)] allocation={0: 3, 3: 6, 9: 2, 11: 5}

free(a)

 FirstFit symbols={'b': 3, 'c': 9, 'd': 11} holes=[(0, 3), (16, 4)] allocation={3: 6, 9: 2, 11: 5}

 BestFit symbols={'b': 3, 'c': 9, 'd': 11} holes=[(0, 3), (16, 4)] allocation={3: 6, 9: 2, 11: 5}

 WorstFit symbols={'b': 3, 'c': 9, 'd': 11} holes=[(0, 3), (16, 4)] allocation={3: 6, 9: 2, 11: 5}

free(c)

 FirstFit symbols={'b': 3, 'd': 11} holes=[(0, 3), (9, 2), (16, 4)] allocation={3: 6, 11: 5}

 BestFit symbols={'b': 3, 'd': 11} holes=[(0, 3), (9, 2), (16, 4)] allocation={3: 6, 11: 5}

 WorstFit symbols={'b': 3, 'd': 11} holes=[(0, 3), (9, 2), (16, 4)] allocation={3: 6, 11: 5}

e=malloc(2):

 FirstFit symbols={'b': 3, 'd': 11, 'e': 0} holes=[(2, 1), (9, 2), (16, 4)] allocation={3: 6, 11: 5, 0: 2}

 BestFit symbols={'b': 3, 'd': 11, 'e': 9} holes=[(0, 3), (16, 4)] allocation={3: 6, 11: 5, 9: 2}

 WorstFit symbols={'b': 3, 'd': 11, 'e': 16} holes=[(0, 3), (9, 2), (18, 2)] allocation={3: 6, 11: 5, 16: 2}

free(b)

 FirstFit symbols={'d': 11, 'e': 0} holes=[(2, 9), (16, 4)] allocation={11: 5, 0: 2}

 BestFit symbols={'d': 11, 'e': 9} holes=[(0, 9), (16, 4)] allocation={11: 5, 9: 2}

 WorstFit symbols={'d': 11, 'e': 16} holes=[(0, 11), (18, 2)] allocation={11: 5, 16: 2}

f=malloc(11):

 FirstFit symbols={'d': 11, 'e': 0, 'f': None} holes=[(2, 9), (16, 4)] allocation={11: 5, 0: 2}

 BestFit symbols={'d': 11, 'e': 9, 'f': None} holes=[(0, 9), (16, 4)] allocation={11: 5, 9: 2}

 WorstFit symbols={'d': 11, 'e': 16, 'f': 0} holes=[(18, 2)] allocation={11: 5, 16: 2, 0: 11}

### 2.3 Test case showing advantage of First-Fit

In the provided test cases, we included one example that shows Best-Fit succeeding while the other two fail, and another example showing Worst-Fit succeeding. For this bonus problem, you are to generate a test case that shows First-Fit succeeding and Best-Fit and Worst-Fit fail. You must provide the test case in the same format as in the template. You must provide an explanation in the PDF file and a typescript. If multiple students submit identical test cases, then the bonus points will be divided evenly among them.