

```
$Id: lab5c-gdb-valgrind.mm,v 1.50 2014-04-23 21:27:47-07 - - $
PWD: /afs/cats.ucsc.edu/courses/cms012b-wm/Labs-cms012m/lab5c-gdb-valgrind
URL: http://www2.ucsc.edu/courses/cms012b-wm/:/Labs-cms012m/lab5c-gdb-valgrind/
```

This lab will introduce you to **gdb**, the Gnu debugger associated with **gcc**, and **valgrind**, which can be used to track uninitialized variables, memory leak, and dangling pointers. Uninitialized variables are variables which are declared but whose value is used before being assigned to. Memory leak occurs when memory is not freed when no longer needed; C does not have a garbage collector. A dangling pointer points at storage that has been freed and should no longer be accessed.

Before beginning this lab, study some of the tutorials in **Tutorials/gdb-tutorials**. There are links in this directory to **gdb-tutorial-handout.pdf** and **gdb-tutorial-ohio.html**, which are fairly short, but also a link to **gdb-tutorial-rms**, which is much longer and more detailed.

```
http://www2.ucsc.edu/courses/cms012b-wm/:/Labs-cms012m/
lab5c-gdb-valgrind/gdb-tutorials/gdb-tutorial-ohio.html
```

1. Use of script

In this lab, you will follow some detailed steps. For each step, submit the files listed. After you have submitted the necessary files, verify that they are all in the submit directory by using **ls**. As in a previous lab, the command

```
grep Submit: *.tt
```

will summarize the files you need to submit. Most of the commands will be interactive, so make use of the **script** command to capture command line output. A terminal session can be captured with

```
script filename
```

where *filename* is the file into which you want your session captured. Be sure not to use anything other than line mode commands in this file, and examine it after to verify this. Specifically, never use an editor inside a terminal running **script**.

2. Detailed steps

Following are the items for this lab. Capture the output in the file specified at the end of each part.

Part (01).

A couple of uninitialized variables. For convenience, a script **mk** (Figure 1) has been provided to avoid the need for a **Makefile** in this lab. It contains compilation instructions. Start with the program **uninit.c**. (Figure 2). Use the following commands.

Submit: **part01.typescript**

mk uninit.c	Run the script to compile.
valgrind uninit	Check for uninitialized variables.
echo \$?	What is the exit status ? bash will capture the crash.
pstatus 139	Print the meaning of the crash.
exit	Get out of script .

Part (02).

Now look into your program with **gdb**, capturing your session into **part02.script**

Submit: **part02.typescript**

gdb uninit	Start gdb .
run	Run the program.
where	Ask where in the program it crashed.
list	Look at a few lines around where it crashed.
print foo	Print the values of some variables.
print pointer	
print *pointer	
print argv[0]	
quit	Quit gdb . Answer yes to the subprocess question.

Part (03).

Now step through the program a line at a time.

Submit: **part03.typescript**

gdb uninit	
break main	Set a breakpoint at the beginning of the main function.
run	Note that it stops at the breakpoint.
print foo	Note that the value is some number, but there is no way to figure out why it has this value.
next	Step one statement, stepping over, not into, functions. The command step would step into the function instead of over it.
print pointer	
next	Step one more statement. Note that it crashes at this point.
quit	

Part (04).

Now let look at **malloc** (similar to Java's **new**) and **free**, which releases storage. C does not have a garbage collector. Start with the program **malloc.c** (Figure 3).

Submit: **part04.typescript**

valgrind malloc	Note that there are a couple of malloc 's but only one free later. So one block was leaked.
gdb malloc	
break main	Set a breakpoint in the main function.
run	
print ptr	
print *ptr	Bad memory access because ptr is not initialized.
next	
print ptr	Now ptr has a value obtained from the heap.
print *ptr	But the value it points it is uninitialized. If it is 0, that is just a coincidence.
next	
next	
next	
print ptr	
print *ptr	Now it points at initialized storage.
next	The storage is freed by the call to free .
next	
next	
next	The reference to __libc_start_main is the startup function called by the operating system to set up the environment and call main.
quit	

Part (05).

Examine **list1.c** (Figure 4). Compile it with **mk list1.c** and look at the errors and warnings you see. Capture the output from this compilation and submit it. Read the man page **malloc(3)** to see what header file was not included.

Submit: **part05.typescript**

Part (06).

Copy `list1.c` to `list2.c` and fix the missing header file. Capture the output.

Submit: `part06.typescript`

Submit: `list2.c`

<code>mk list2.c</code>	
<code>valgrind list2 foo bar</code>	Note the complaints from valgrind . It complains about memory leak, but also about invalid access to memory.
<code>gdb list2</code>	
<code>break n</code>	Set a breakpoint at line <i>n</i> , where <i>n</i> is the line number of the return statement.
<code>run foo bar</code>	Note how arguments are given to a program, on the run not on the invocation of gdb .
<code>where</code>	
<code>list</code>	Does not list the lines around the point of the crash.
<code>list list2.c:23</code>	We can select the particular set of lines to list.
<code>print head</code>	Not in the current stack frame. Note that we have called several library functions, as shown by where .
<code>bt</code>	A backtrace is another way of looking at the stack.
<code>up</code>	
<code>up</code>	
<code>up</code>	We unwind the stack three levels here.
<code>print head</code>	Now we are in the correct frame.
<code>print *head</code>	
<code>print *(head->word)</code>	We can use more complicated C expressions.
<code>print head->link->link->link</code>	
<code>print *(head->link->link->link)</code>	

Part (07).

Run `list2` again, showing values in `argv`.

Submit: `part07.typescript`

<code>gdb list2</code>	
<code>break main</code>	
<code>run foo bar</code>	Run the program with two command line arguments, namely <code>foo</code> and <code>bar</code> .
<code>print argc</code>	
<code>print argv</code>	
<code>print argv[0]</code>	<code>argv[0]</code> is always the name of the program being run.
<code>print argv[1]</code>	
<code>print argv[2]</code>	
<code>print argv[3]</code>	<code>argv[argc]</code> is always the null pointer, represented as <code>0x0</code> in C.
<code>print argv[4]</code>	After <code>argv</code> is the default environment which you can display using the <code>env</code> or <code>printenv</code> command.
<code>print argv[5]</code>	
<code>print argv[6]</code>	

Part (08).

Copy `list2.c` to `list3.c` and use `valgrind` and `gdb` as appropriate so that you can track down and fix all of the memory faults. Ignore memory leak for now. The backslash in the command below is just for typographical reasons, to get the command within the box's width. You may type the entire command on one line.

Submit: `part08.typescript`

Submit: `list3.c`

<code>valgrind --leak-check=full \</code> <code>list3 foo bar baz qux</code>	Run <code>valgrind</code> with the option <code>--leak-check=full</code> to verify that your program in fact has no problems except leaks.
---	--

Part (09).

Copy `list3.c` to `list4.c`. Eliminate memory leak by using `free` to release all allocated storage.

Submit: `part09.typescript`

Submit: `list4.c`

<code>valgrind --leak-check=full \</code> <code>list4 foo bar baz qux</code>	Verify that your program now works with no memory faults and no memory leak.
<code>echo \$?</code>	Make sure the exit status is <code>EXIT_SUCCESS</code> .

Part (10).

Write a program called `environ.c` which will declare the external variable

```
extern char **environ;
```

and write a loop iterating over that array, printing each element per line. The stopping condition is meeting a null pointer, as there is no variable indicating how large the array is.

Submit: `part10.typescript`

Submit: `environ.c`

<code>./environ</code>	Print out all your environment variables.
------------------------	---

3. What to submit

Submit the 14 files mentioned above. If you are doing pair programming, submit the required files as well.

4. Debugging with ddd

An alternative to `gdb` is `ddd`, which is a GUI wrapper around `gdb`. It is not part of this lab and there is nothing to submit from using `ddd`, but you might want to explore it. For example:

- (1) Start with: `ddd unittest &`. The ampersand (&) at the end of the line causes the program to be run in the background. If you are using a terminal without X11 forwarding this will not work.
- (2) In the `gdb` window, type: `break main`. Note that a stop sign appears in the code.
- (3) Then type: `run`. An arrow shows the breakpoint.
- (4) Click on **Step** several times.
- (5) You may also use print statements in the `gdb` window to examine variables.

5. Debugging with gdbtui

There is also a program `gdbtui` which is the text user interface to `gdb`, which works in a terminal window in full screen mode. You might like to try that as well.

```
1  #!/bin/sh
2  # $Id: mk,v 1.8 2014-04-23 21:26:13-07 - - $
3  #
4  # This script takes the names of C source files and compiles them
5  # into executable images.  Each source must be a complete program.
6  #
7
8  export PATH=$PATH:/afs/cats.ucsc.edu/courses/cmcs012b-wm/bin
9  GCC='gcc -g -O0 -Wall -Wextra -std=gnu99'
10
11 function label() {
12     echo ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
13     echo "$@"
14     echo ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
15 }
16
17 function catnv() {
18     for arg in $*
19     do
20         if [ "$arg" = "-f" ]
21         then
22             echo -e '\f'
23         else
24             label $arg
25             cat -nv $arg
26         fi
27     done
28 }
29
30 for CSOURCE in *.c
31 do
32     EXECBIN=${CSOURCE%.c}
33     if [ "$CSOURCE" == "$EXECBIN" ]
34     then
35         echo $0: $CSOURCE: is not a C source file
36     else
37         echo CSOURCE=$CSOURCE EXECBIN=$EXECBIN
38         set -x
39         cid + $CSOURCE
40         checksource $CSOURCE
41         catnv $CSOURCE >$CSOURCE.log
42         label "Errors: $GCC $CSOURCE -o $EXECBIN" >>$CSOURCE.log
43         $GCC $CSOURCE -o $EXECBIN -lm >>$CSOURCE.log 2>&1
44         cat $CSOURCE.log
45         set +x
46         echo =====
47     fi
48 done
```

```
49
50 mkpspdf Listing.ps $0 *.log
51
```

Figure 1. code/mk

```
1 // $Id: uninit.c,v 1.3 2014-02-18 15:08:14-08 - - $
2
3 #include <stdio.h>
4 #include <stdlib.h>
5
6 int main (int argc, char **argv) {
7     int *pointer;
8     printf ("pointer=%p\n", pointer);
9     printf ("*pointer=%d\n", *pointer);
10    int foo;
11    printf ("foo=%d\n", foo);
12 }
```

Figure 2. code/uninit.c


```
1 // $Id: malloc.c,v 1.2 2013-02-08 18:01:12-08 - - $
2
3 #include <stdio.h>
4 #include <stdlib.h>
5
6 typedef struct node node;
7 struct node {
8     int value;
9     node *link;
10 };
11
12 int main (int argc, char **argv) {
13     node *ptr = malloc (sizeof (struct node));
14     ptr = malloc (sizeof (node));
15     ptr->value = 6;
16     ptr->link = NULL;
17     printf ("%p-> {%d, %p}\n", ptr, ptr->value, ptr->link);
18     free (ptr);
19     return EXIT_SUCCESS;
20 }
```

Figure 3. code/malloc.c

```
1 // $Id: list1.c,v 1.3 2013-02-08 18:00:41-08 - - $
2
3 #include <assert.h>
4 #include <stdio.h>
5
6 typedef struct node node;
7 struct node {
8     char *word;
9     node *link;
10 };
11
12 int main (int argc, char **argv) {
13     node *head;
14     for (int argi = 0; argi < 5; ++argi) {
15         node *node = malloc (sizeof (struct node));
16         assert (node != NULL);
17         node->word = argv[argi];
18         node->link = head;
19         head = node;
20     }
21     for (node *curr = head; curr->link != NULL; curr = curr->link) {
22         printf ("%p->node {word=%p->[%s], link=%p}\n",
23             curr, curr->word, curr->word, curr->link);
24     }
25     return 9;
26 }
```

Figure 4. code/list1.c