ECS 150 - GDB tutorial

Prof. Joël Porquet-Lupine

UC Davis - 2020/2021

UC DAVIS

COMPUTER SCIENCE

GDB

GNU project

- Started by Richard Stallman in 1983
- Free software, mass collaboration project in response to proprietary UNIX
  - Copyleft license: GNU GPL
  - User programs: text editor (Emacs), compiler (GCC toolchain), debugger (GDB), and various utilities (ls, grep, awk, make, etc.)
  - Kernel: GNU Hurd

GDB

- GNU DeBugger
- Supports many languages
  - Including C and C++
- Inspection of program during execution
  - Execution flow
  - Data
- Helps finding errors like *segmentation fault*
- Read the fully-detailed manual: https://sourceware.org/gdb/current/onlinedocs/gdb/

GDB usage

Compilation flags

- Canonical compilation command line:

```
$ gcc [cflags] -o <output> <input>
```

- Optimize for speed (`-O2`)

```
$ gcc -Wall -Werror -O2 -o myprogram main.c
```


- Enable debugging support (`-g`)


```
$ gcc -Wall -Werror -g myprogram main.c
```


  - Not recommended to use debugging along with optimizations
    - No optimization option is equivalent to `-O0`

GDB usage

Makefile digression

- During development, very useful to be able to debug your program
- For production, probably better to disable the debug support and activate all possible optimization support
  - Reduce size of the executable (can easily be by 50%!)
  - Increase performance (can also be by 50%!)

Makefile automatism


```
ifeq ($(D),1)
CFLAGS    += -g   # Enable debugging
else
CFLAGS    += -O2  # Enable optimization
endif
```

Building mode

```
$ make D=1    # compile with debug support and no optimization
```

```
$ make        # compile with optimizations (production)
```

- Probably want to use `make clean` when changing building mode

GDB usage

Starting GDB

- Start GDB, specify the program to debug

```
$ gdb
...
(gdb) file myprogram
Reading symbols from myprogram...done.
(gdb)
```

- Or, start GDB with the program to debug as argument

```
$ gdb myprogram
...
Reading symbols from myprogram...done.
(gdb)
```


Running the program

- Without any argument:

```
(gdb) run
```

- With arguments:

```
(gdb) run argv1 argv2...
```

GDB usage

Interactive help

- GDB offers an interactive shell
  - History management
  - Auto-complete (with TAB)

In order to discover what you can do, just ask:

```
(gdb) help
List of classes of commands:
aliases -- Aliases of other commands
breakpoints -- Making program stop at certain points
...

(gdb) help breakpoints
Making program stop at certain points.
List of commands:
awatch -- Set a watchpoint for an expression
break -- Set breakpoint at specified location
...

(gdb) help break
Set breakpoint at specified location.
break [PROBE_MODIFIER] [LOCATION] [thread THREADNUM] [if CONDITION]
...
```

GDB usage

Possible scenarios

1. Program doesn't have bugs:
   - It will run fine until completion

```
$ ./myprogram
I worked, hurray!
```

2. *Best*-case scenario, regarding bugs:
   - *Segmentation fault*

```
$ ./myprogram
segmentation fault (core dumped) ./myprogram
```

3. *Worst*-case scenario:
   - Doesn't crash but wrong result

```
$ ./myprogram
I work□□, □□rray!
```

   - Bugs that don't trigger any segmentation fault
   - In this case, you'll probably have to spend more time...

Segmentation faults

Example #1

```c
#include <stdio.h>
#include <stdlib.h>
#include <string.h>

size_t foo_len (const char *s)
{
    return strlen(s);
}

int main (int argc, char *argv[])
{
    char *a = NULL;

    printf ("size of a = %d\n", foo_len(a));

    return 0;
}
```


Execution

```
$ ./strlen-test
zsh: segmentation fault (core dumped)  ./strlen-test
```

Segmentation faults

Run with GDB

- *(After compiling the code with `-g`)*

```
$ gdb ./strlen-test
(gdb) run
Starting program: /home/joel/tmp/test/strlen-test

Program received signal SIGSEGV, Segmentation fault.
0x00007ffff7abc446 in strlen () from /usr/lib/libc.so.6
(gdb)
```

Backtrace

- First thing to do when getting a *segfault*:
  - Understand what is the sequence of calls that brought us there


```
(gdb) backtrace     # use just 'bt'
#0  0x00007ffff7abc446 in strlen () from /usr/lib/libc.so.6
#1  0x000000000040055e in foo_len (s=0x0) at strlen-test.c:7
#2  0x0000000000400583 in main (argc=1, argv=0x7fffffffd7a8) at strlen-test.c:14
```


Investigate

- `foo_len()` is supposed to receive a pointer
- Here it receives `0` (aka `NULL`)
- Looks like this `NULL` pointer probably gets dereferenced in `strlen()`...

Segmentation faults

Fix...

- Here, the problem is fairly obvious

```
size_t foo_len (const char *s)
{
    return strlen(s);
}

int main (int argc, char *argv[])
{
    char *a = "This is a valid string";

    printf ("size of a = %d\n", foo_len(a));

    return 0;
}
```


And, celebrate!

```
$ ./strlen-test
size of a = 22
```

Segmentation faults

Better fix

- Prevent the same bug from happening again

```
size_t foo_len (const char *s)
{
    assert(s && "String cannot be NULL here!");
    return strlen(s);
}

int main (int argc, char *argv[])
{
    char *a = NULL;

    printf ("size of a = %d\n", foo_len(a));

    return 0;
}
```



```
$ ./strlen-test
strlen-test: strlen-test.c:8: foo_len:
    Assertion `s && "String cannot be NULL here!"' failed.
```

Segmentation faults

Example #2

```c
#include<stdio.h>
#include<stdlib.h>

const static int len = 10;

int main(void)
{
    int *tab;
    unsigned int i;

    tab = malloc(len * sizeof(int));

    for (i = len - 1; i >= 0; i--)
        tab[i] = i;

    free(tab);
    return 0;
}
```


Execution

```
$ ./tablen-test
segmentation fault (core dumped)  ./tablen-test
```

Segmentation faults

Run GDB

```
$ gdb ./tablen-test
(gdb) run
Starting program: /home/joel/tmp/test/tablen-test

Program received signal SIGSEGV, Segmentation fault.
0x0000000000400535 in main () at tablen-test.c:14
14              tab[i] = i;
```

Backtrace


```
(gdb) bt
#0  0x0000000000400535 in main () at tablen-test.c:14
```


- Except that here, it's not much of help...

Inspect variables

- Display index `i` so that we know which index in the array was being accessed:

```
(gdb) print i
$1 = 4294967295
```

Segmentation faults

Fix...

- Problem is a case of overflow
  - An `unsigned int` type automatically wraps from 0 to 4294967295

```c
#include<stdio.h>
#include<stdlib.h>

const static int len = 10;

int main(void)
{
    int *tab;
    int i;

    tab = malloc(len * sizeof(int));

    for (i = len - 1; i >= 0; i--)
        tab[i] = i;

    free(tab);
    return 0;
}
```

Tracking bugs

Behavior bugs

- Behavioral bugs more complicated to find because program doesn't crash
- It's just that the output is wrong

```c
#include <ctype.h>
#include <stdio.h>
#include <string.h>

int main(void)
{
    int i;
    char str[] = "Tracking bugs is my passion";

    printf("Before: %s\n", str);

    for (i = 0; i < strlen(str) - 1; i++)
        str[i] = toupper(str[i]);

    printf("After: %s\n", str);

    return 0;
}
```


Execution

```
Before: Tracking bugs is my passion
After: TRACKING BUGS IS MY PASSIOn
```

Tracking bugs

Setting breakpoints

- Stop the program during the execution at a designated point
- Set as many breakpoints as necessary
- GDB will always stop the execution when reaching them

Breaking at exact location in code

```
(gdb) break string-test.c:13
Breakpoint 1 at 0x4005ef: file string-test.c, line 13.

(gdb) r
Starting program: /home/joel/tmp/test/string_test
Before: Tracking bugs is my passion

Breakpoint 1, main () at string-test.c:13
13              str[i] = toupper(str[i]);
```

Tracking bugs

Breaking at a particular function

```
(gdb) b main
Breakpoint 1 at 0x40059f: file string-test.c, line 8.

(gdb) r
Starting program: /home/joel/tmp/test/string-test

Breakpoint 1, main () at string-test.c:8
8           char str[] = "Tracking bugs is my passion";
```


Breaking only if condition is satisfied


```
(gdb) b string-test.c:13 if i == 5
Breakpoint 1 at 0x4005ef: file string-test.c, line 13.

(gdb) r
Starting program: /home/joel/tmp/test/string-test
Before: Tracking bugs is my passion

Breakpoint 1, main () at string-test.c:13
13              str[i] = toupper(str[i]);

(gdb) print i
$1 = 5
```

Tracking bugs

Dealing with breakpoints

- Set at least one breakpoint before running the program
  - Otherwise the program will run until completion
- Once the program stops and the gdb shell is available, a few options:
  1. Continue the execution until hitting the same or another breakpoint

     ```
     (gdb) continue      # or just 'c'
     ```

  2. Execute only the next line of code and break again

     ```
     (gdb) step          # or just 's'
     ```

     Careful, `step` enters function calls
  3. Jump over function calls

     ```
     (gdb) next          # or just 'n'
     ```

Tip: typing `<enter>` in the interactive GDB shell repeats the last command

Tracking bugs

Printing variables

```
int a = 2;
char b = 'x';
int *c = &a;
char *s = "A string";
                        // <= breaking here
...
```

- Inspect the value of all your variables with command `print`

Default

- By default, prints variables according to their type

```
(gdb) print a
$1 = 2
(gdb) p b
$2 = 120 'x'
(gdb) p c
$3 = (int *) 0x7fffffffd65c
(gdb) p s
$4 = 0x40070b "A string"
```

Tweak

- Can tweak both the way `print` prints and what it prints

```
(gdb) print /x a
$1 = 0x2
(gdb) p /c b+2
$2 = 122 'z'
(gdb) p *c
$3 = 2
(gdb) p s[0]
$4 = 65 'A'
```

Tracking bugs

Printing data structures

```
struct entry {
    int     key;
    char    *name;
} obj = {
    .key    = 2,
    .name   = "toto",
};

struct entry *e = &obj;
```


- With `print`, you can access the pointer and the object it's pointing to:


```
(gdb) print e
$1 = (struct entry *) 0x7fffffffd640
(gdb) print &obj
$2 = (struct entry *) 0x7fffffffd640
(gdb) p *e
$3 = {key = 2, name = 0x400734 "toto"}
(gdb) p e->key
$4 = 2
(gdb) p obj.name
$5 = 0x400734 "toto"
```

Misc

Setting watchpoint

- Breakpoints are for interrupting the execution flow at a specific location
- Watchpoints are for interrupting the program when a variable is modified

```
(gdb) watch i
Hardware watchpoint 2: i

(gdb) c
Continuing.
Before: Tracking bugs is my passion

Hardware watchpoint 2: i

Old value = 0
New value = 1
0x0000000000400612 in main () at string_test.c:12
12          for (i = 0; i < strlen(str) - 1; i++)
```

Misc

Other useful commands

- `finish`
  - Runs until the current function is finished
- `until`
  - When executed in a loop, continues the execution until the loop ends
- `info breakpoints`
  - Shows informations about all declared breakpoints

```
(gdb) info b
Num     Type           Disp Enb Address            What
1       breakpoint     keep y   0x000000000040059f in main at string-test.c:8
breakpoint already hit 1 time
2       hw watchpoint  keep y                      i
breakpoint already hit 2 times
```


- `delete`
  - Deletes a breakpoint

Valgrind

Example

```c
#include <stdlib.h>

void f(void)
{
        int *x = malloc(10 * sizeof(int));
        x[10] = 0;
}

int main(void)
{
        f();
        return 0;
}
```

Valgrind

Run

```
$ valgrind --leak-check=full ./valgrind_example
...
==31134== Invalid write of size 4
==31134==    at 0x108668: f (in /home/joel/work/ecs150/slides/tuto_gdb/code/valgrind_example)
==31134==    by 0x108679: main (in /home/joel/work/ecs150/slides/tuto_gdb/code/valgrind_example)
==31134==  Address 0x51f0068 is 0 bytes after a block of size 40 alloc'd
==31134==    at 0x4C2CEDF: malloc (vg_replace_malloc.c:299)
==31134==    by 0x10865B: f (in /home/joel/work/ecs150/slides/tuto_gdb/code/valgrind_example)
==31134==    by 0x108679: main (in /home/joel/work/ecs150/slides/tuto_gdb/code/valgrind_example)
==31134==
==31134==
==31134== HEAP SUMMARY:
==31134==     in use at exit: 40 bytes in 1 blocks
==31134==   total heap usage: 1 allocs, 0 frees, 40 bytes allocated
==31134==
==31134== 40 bytes in 1 blocks are definitely lost in loss record 1 of 1
==31134==    at 0x4C2CEDF: malloc (vg_replace_malloc.c:299)
==31134==    by 0x10865B: f (in /home/joel/work/ecs150/slides/tuto_gdb/code/valgrind_example)
==31134==    by 0x108679: main (in /home/joel/work/ecs150/slides/tuto_gdb/code/valgrind_example)
==31134==
==31134== LEAK SUMMARY:
==31134==    definitely lost: 40 bytes in 1 blocks
==31134==    indirectly lost: 0 bytes in 0 blocks
==31134==      possibly lost: 0 bytes in 0 blocks
==31134==    still reachable: 0 bytes in 0 blocks
==31134==         suppressed: 0 bytes in 0 blocks
==31134==
==31134== For counts of detected and suppressed errors, rerun with: -v
==31134== ERROR SUMMARY: 2 errors from 2 contexts (suppressed: 0 from 0)
```