ECS 260 Software Engineering

Winter 2023

Homework #2

Due 5:00pm Thursday February 16th

Student Name: Yu-Cheng Hwang

- To complete the assignment (i) modify `hw2.tex`, (ii) create the corresponding PDF document (using pdflatex, for example), and (iii) submit the pdf electronically via Canvas by the due date.
- This homework has to be worked on individually.
- Remember to replace "Your Name" with your name above.
- Replace the `TODO` comments with your answers.
- Do not comment out the `\newpage` commands, and make sure each question (and answer) retain their original page numbers.
- For this assignment, we will use a Docker image that has KLEE installed. Visit the following link for instructions on how to install Docker: https://docs.docker.com/get-docker/
- Pull the Docker image:

  ```
  docker pull ucdavisplse/ecs260-hw2
  ```

- Create and run the Docker container:

  ```
  docker run --rm -ti --ulimit="stack=-1:-1" ucdavisplse/ecs260-hw2
  ```

1 Testing a Small Function (15 pts)

- In this part of the homework, we will follow KLEE tutorial #1:
  http://klee.github.io/tutorials/testing-function/
- Once you are inside the container, change your directory:
  `cd klee_src/examples/get_sign`
- Follow the instructions in the tutorial, and answer the questions below.

(a) What test cases does KLEE generate for this function?

According to the program, it has three constrains: $x = 0$, $x < 0$, and $x > 0$
So the KLEE generates three symbolic inputs for the function.

1. `input = 0`
2. `input = 16843009`
3. `input = (int)-2147483648 or input = (uint)2147483648`

(b) What outputs are produced when replaying the test cases?

We first set the link path to the directory in docker.
Then we run program with our 3 test cases, we get:

1. 0 for `input = 0`
2. 1 for `input = 16843009`
3. 255 for `input = (int)-2147483648 or input = (uint)2147483648`

(c) What kinds of files does KLEE produce for this function, and what kind of information do they contain?

Klee generates the following files:

- test00000x.ktest: test files used to generate symbolic input to test program.
- run.stats: various statistics emitted by Klee.
- run.isatss: binary file containing global statistics.
- message: message other than warning or info.
- info: record the exact command line with which Klee was run.
- assembly.ll: LLVM bitcode.
- warnings.txt: contain warning message emitted by Klee.

2 Testing the GNU Coreutils (40 pts)

- Your task for this part of the homework is to use KLEE to test GNU Coreutils, and to reproduce the problems reported in Figure 7 on page 10 of the OSDI'08 paper discussed in lecture: https://www.doc.ic.ac.uk/~cristic/papers/klee-osdi-08.pdf/

- Change the directory to ~/coreutils-6.10-bitcodes.

  `cd ~/coreutils-6.10-bitcodes`

  This directory contains LLVM bitecode files for all of the Coreutils version 6.10.

- As an example, we can use KLEE as a pure interpreter on `cat`:

  `klee --optimize --libc=uclibc --posix-runtime ./cat.bc -version`

- You may want to use the following flags (you can find most of this list by running a coreutils program under KLEE with –help).

  -sym-arg N
  Replace by a symbolic argument with length N

  -sym-args MIN MAX N
  Replace by at least MIN arguments and at most MAX arguments, each with maximum length N

  -sym-files NUM N
  Make NUM symbolic files ('A', 'B', 'C', etc.), each with size N

- **EXAMPLE 1.** We can reproduce bugs from Figure 7 as follows (use the inputs provided in Figure 7):

  `klee --optimize --libc=uclibc --posix-runtime ./mkdir.bc -Z a b`

```
KLEE: WorkaroundLLVMPR39177: replaced alias @fputc with clone of function @__fputc_unlocked
KLEE: WorkaroundLLVMPR39177: replaced alias @fputc_unlocked with clone of function @__fputc_unlocked
KLEE: WorkaroundLLVMPR39177: replaced alias @fputs with clone of function @fputs_unlocked
KLEE: WorkaroundLLVMPR39177: replaced alias @fwrite with clone of function @fwrite_unlocked
KLEE: WorkaroundLLVMPR39177: replaced alias @fgetc_unlocked with clone of function @__fgetc_unlocked
KLEE: WARNING: undefined reference to function: __ctype_get_mb_cur_max
KLEE: WARNING: undefined reference to function: bindtextdomain
KLEE: WARNING: undefined reference to function: dcgettext
KLEE: WARNING: undefined reference to function: textdomain
KLEE: WARNING: executable has module level assembly (ignoring)
KLEE: WARNING ONCE: calling external: syscall(16, 0, 21505, 94164462961936) at /tmp/klee_src/runtime/POSIX/fd.c:991 10
KLEE: WARNING ONCE: Alignment of memory from call "malloc" is not modelled. Using alignment of 8.
KLEE: WARNING ONCE: calling external: bindtextdomain(94164459072432, 94164460370240) at mkdir.c:158 3
KLEE: WARNING ONCE: calling external: textdomain(94164459072432) at mkdir.c:160 3
KLEE: WARNING ONCE: Alignment of memory from call "realloc" is not modelled. Using alignment of 8.
KLEE: WARNING ONCE: calling external: dcgettext(0, 94164460297600, 5) at mkdir.c:194 12
KLEE: WARNING ONCE: calling external: __ctype_get_mb_cur_max() at quotearg.c:181 36
KLEE: ERROR: quotearg.c:248: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location

KLEE: done: total instructions = 55094
KLEE: done: completed paths = 1
KLEE: done: generated tests = 1
```

- **EXAMPLE 2.** We can use KLEE to search for bugs:

  ```
  klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./mkdir.bc -sym-args
  3 3 10
  ```

- Answer the following questions:

(a) Which bugs from Figure 7 can you *reproduce* (from the 10 test cases provided)? List the command you ran, and the ERROR you observe.

I can reproduce all these errors listed in the paper.
The errors are **ALL** *memory error: out of bound pointer*.
Here are the commands I ran:

```
klee --optimize --libc=uclibc --posix-runtime ./paste.bc -d\\ abcdefghijklmnopqrstuvwxyz
klee --optimize --libc=uclibc --posix-runtime ./pr.bc -e t2.txt
klee --optimize --libc=uclibc --posix-runtime ./tac.bc -r t3.txt t3.txt
klee --optimize --libc=uclibc --posix-runtime ./mkdir.bc -Z a b
klee --optimize --libc=uclibc --posix-runtime ./mkfifo.bc -Z a b
klee --optimize --libc=uclibc --posix-runtime ./mknod.bc -Z a b p
klee --optimize --libc=uclibc --posix-runtime ./md5sum.bc -c t1.txt
klee --optimize --libc=uclibc --posix-runtime ./ptx.bc -F\\ abcdefghijklmnopqrstuvwxyz
klee --optimize --libc=uclibc --posix-runtime ./ptx.bc x t4.txt
klee --optimize --libc=uclibc --posix-runtime ./seq.bc -f \%0 1
```

(b) Run KLEE to *search* for bugs on the 9 GNU Coreutils for which the paper found bugs (listed in Figure 7). For each program, list the command line used to test it, the number of error-inducing inputs generated by KLEE, and an example of an error-inducing input (if any is found).

```
klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./paste.bc -sym-args 3 3 10
KLEE: ERROR: ../../src/paste.c:93: memory error: out of bound pointer
ktest file : 'test000001.ktest'
args       : ['./paste.bc', '-sym-args', '3', '3', '10']
num objects: 4
object 0: name: 'arg00'
object 0: size: 11
object 0: data: b'-d\x00\xff\xff\xff\xff\xff\xff\xff\xff'
object 0: hex : 0x2d6400ffffffffffffff
object 0: text: -d.........
object 1: name: 'arg01'
object 1: size: 11
object 1: data: b'\\\x00\xff\xff\xff\xff\xff\xff\xff\xff'
object 1: hex : 0x5c00ffffffffffffffff
object 1: text: \.........
object 2: name: 'arg02'
object 2: size: 11
object 2: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 2: hex : 0x00000000000000000000
object 2: text: ..........
object 3: name: 'model_version'
object 3: size: 4
object 3: data: b'\x01\x00\x00\x00'
object 3: hex : 0x01000000
object 3: int : 1
object 3: uint: 1
object 3: text: ....


klee --optimize --libc=uclibc --posix-runtime -max-time=300s ./pr.bc -sym-args 3 3 10
KLEE: done: total instructions = 1180295


klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./tac.bc -sym-args 3 3 10
KLEE: HaltTimer invoked
KLEE: halting execution, dumping remaining states


klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./mkdir.bc -sym-args 3 3 10
KLEE: ERROR: ../../lib/quotearg.c:248: memory error: out of bound pointer
ktest file : 'test000009.ktest'
args       : ['./mkdir.bc', '-sym-args', '3', '3', '10']
num objects: 4
object 0: name: 'arg00'
object 0: size: 11
object 0: data: b'--c\x00\xff\xff\xff\xff\xff\xff\xff'
object 0: hex : 0x2d2d6300ffffffffffff
object 0: text: --c.......
object 1: name: 'arg01'
object 1: size: 11
object 1: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 1: hex : 0x00000000000000000000
object 1: text: ..........
object 2: name: 'arg02'
object 2: size: 11
object 2: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 2: hex : 0x00000000000000000000
object 2: text: ..........
object 3: name: 'model_version'
object 3: size: 4
object 3: data: b'\x01\x00\x00\x00'
object 3: hex : 0x01000000
object 3: int : 1
object 3: uint: 1
object 3: text: ....


klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./mkfifo.bc -sym-args 3 3 10
KLEE: ERROR: ../../lib/quotearg.c:248: memory error: out of bound pointer
ktest file : 'test000006.ktest'
args       : ['./mkfifo.bc', '-sym-args', '3', '3', '10']
num objects: 4
object 0: name: 'arg00'
object 0: size: 11
object 0: data: b'-Z\x00\xff\xff\xff\xff\xff\xff\xff\xff'
object 0: hex : 0x2d5a00ffffffffffffff
object 0: text: -Z.........
object 1: name: 'arg01'
object 1: size: 11
object 1: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 1: hex : 0x00000000000000000000
```

```
object 1: text: ..........
object 2: name: 'arg02'
object 2: size: 11
object 2: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 2: hex : 0x00000000000000000000000
object 2: text: ..........
object 3: name: 'model_version'
object 3: size: 4
object 3: data: b'\x01\x00\x00\x00'
object 3: hex : 0x01000000
object 3: int : 1
object 3: uint: 1
object 3: text: ....
```

```
klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./mknod.bc -sym-args 3 3 10
KLEE: ERROR: ../../lib/quotearg.c:248: memory error: out of bound pointer
klee@ebc7b1ec2860:~/coreutils-6.10-bitcodes/klee-out-15$ ktest-tool test000015.ktest
ktest file : 'test000015.ktest'
args       : ['./mknod.bc', '-sym-args', '3', '3', '10']
num objects: 4
object 0: name: 'arg00'
object 0: size: 11
object 0: data: b'-Z\x01\xff\xff\xff\xff\xff\xff\xff\xff'
object 0: hex : 0x2d5a01ffffffffffffff
object 0: text: -Z.........
object 1: name: 'arg01'
object 1: size: 11
object 1: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 1: hex : 0x00000000000000000000000
object 1: text: ..........
object 2: name: 'arg02'
object 2: size: 11
object 2: data: b'p\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff'
object 2: hex : 0x70ffffffffffffffffff
object 2: text: p.........
object 3: name: 'model_version'
object 3: size: 4
object 3: data: b'\x01\x00\x00\x00'
object 3: hex : 0x01000000
object 3: int : 1
object 3: uint: 1
object 3: text: ....
```

```
klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./md5sum.bc -sym-args 3 3 10
KLEE: HaltTimer invoked
KLEE: halting execution, dumping remaining states
```

```
klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./ptx.bc -sym-args 3 3 10
KLEE: ERROR: ../../src/ptx.c:399: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:397: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:377: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:344: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:391: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:352: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:383: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:362: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:367: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:372: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:326: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:339: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:330: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:392: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: ERROR: ../../src/ptx.c:327: memory error: out of bound pointer
KLEE: NOTE: now ignoring this error at this location
KLEE: HaltTimer invoked
KLEE: halting execution, dumping remaining states
ktest file : 'test000010.ktest'
args       : ['./ptx.bc', '-sym-args', '3', '3', '10']
num objects: 4
object 0: name: 'arg00'
object 0: size: 11
object 0: data: b'-F\\\x00\\r\xff\xff\xff\xff\xff'
object 0: hex : 0x2d465c005c72ffffffffff
object 0: text: -F\.\r.....
```

```
object 1: name: 'arg01'
object 1: size: 11
object 1: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 1: hex : 0x00000000000000000000000
object 1: text: ...........
object 2: name: 'arg02'
object 2: size: 11
object 2: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 2: hex : 0x00000000000000000000000
object 2: text: ...........
object 3: name: 'model_version'
object 3: size: 4
object 3: data: b'\x01\x00\x00\x00'
object 3: hex : 0x01000000
object 3: int : 1
object 3: uint: 1
object 3: text: ....


klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./seq.bc -sym-args 3 3 10
KLEE: HaltTimer invoked
KLEE: halting execution, dumping remaining states
```

(c) Run KLEE on *three* GNU Coreutils *not* listed in Figure 7. For each program, list the command line used to test it, the number of error-inducing inputs generated by KLEE, and an example of an error-inducing input (if any is found).

```
klee --optimize --libc=uclibc --posix-runtime -max-time=60s ./cat.bc -sym-args 1 3 5
KLEE: HaltTimer invoked
KLEE: halting execution, dumping remaining states
KLEE: done: total instructions = 114683
KLEE: done: completed paths = 50
KLEE: done: generated tests = 50


klee --optimize --libc=uclibc --posix-runtime -max-time=120s ./dir.bc -sym-args 1 5 10
KLEE: HaltTimer invoked
KLEE: halting execution, dumping remaining states
KLEE: done: total instructions = 2082934
KLEE: done: completed paths = 18617
KLEE: done: generated tests = 18617


klee --optimize --libc=uclibc --posix-runtime --max-time=60 ./sha256sum.bc --sym-files 1 10 \
--sym-args 1 2 8

ktest file : 'test000007.ktest'
args       : ['./sha256sum.bc', '--sym-files', '1', '10', '--sym-args', '1', '2', '8']
num objects: 5
object 0: name: 'n_args'
object 0: size: 4
object 0: data: b'\x01\x00\x00\x00'
object 0: hex : 0x01000000
object 0: int : 1
object 0: uint: 1
object 0: text: ....
object 1: name: 'arg00'
object 1: size: 9
object 1: data: b'A\x00\x00\x00\x00\x00\x00\x00\x00'
object 1: hex : 0x4100000000000000
object 1: text: A........
object 2: name: 'A-data'
object 2: size: 10
object 2: data: b'\x00\x00\x00\x00\x00\x00\x00\x00\x00\x00'
object 2: hex : 0x00000000000000000000
object 2: text: ..........
object 3: name: 'A-data-stat'
object 3: size: 144
object 3: data: b'-\x00\x00\x00\x00\x00\x00\x00\x01\x00\x00\x00\x00\x00
\x00\x00\x01\x00\x00\x00\x00\x00\x00\x00\xa4\x81\x00\x00\xe8
\x03\x00\x00\xe8\x03\x00\x00\xff\xff\xff\xff\x00\x00\x00\x00\x00
\x00\x00\x00\xff\xff\xff\xff\xff\xff\xff\xff\x00\x10\x00\x00\x00\x00
\x00\x00\xff\xff\xff\xff\xff\xff\xff\xff\xaa\x82\xeec\x00\x00\x00\x00
\xff\xff\xff\xff\xff\xff\xff\xff\x16\x83\xeec\x00\x00\x00\x00\xff
\xff\xff\xff\xff\xff\xff\xff\x16\x83\xeec\x00\x00\x00\x00\xff\xff\xff
\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff
\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff\xff'
object 3: hex : 0x2d000000000000001000000000000001000000000000a4810000e8030000e8030000ffffffff
0000000000000000ffffffffffffffff0010000000000000ffffffffffffffff
aa82ee6300000000ffffffffffffffff1683ee6300000000ffffffffffffffff
1683ee6300000000ffffffffffffffffffffffffffffffffffffffffffffffff
object 3: text:  -...............................................................................
c...............c...............c................................
object 4: name: 'model_version'
object 4: size: 4
object 4: data: b'\x01\x00\x00\x00'
object 4: hex : 0x01000000
object 4: int : 1
object 4: uint: 1
object 4: text: ....
```

3 Checking Equivalence (30 pts)

- Test the equivalence of two *small* functions as shown in Figure 11 on page 13. Pick an example other than *mod*, and inject a bug in one of the implementations if needed. Please include the following: (1) the source code you tested using KLEE, (2) an explanation of what the bug is, (3) the command line you used to run KLEE, and (4) the test result (did KLEE find the bug?).

Function description

`fun1` and `fun2` add two input varialbes `(x,y)` and return the answer. `fun2` modify variable `y` to let it always be positive integer. Then we compare the return value of `fun1` and `fun2`.

Test code

```
#include<iostream>
#include<cassert>
#include<klee/klee.h>

using namespace std;

int fun1(int x, int y){
    return x + y;
}

int fun2(int x, int y){
    // insert error here
    if(y > 0) y = -1 * y;
    else y = 0;

    return x + y;
}

int main(){
    int x, y;
    klee_make_symbolic(&x, sizeof(x), "x");
    klee_make_symbolic(&y, sizeof(y), "y");
    int f1 = fun1(x, y);
    int f2 = fun2(x, y);
    assert(f1 == f2);
    return 0;
}
```

Command line

```
clang -I ../klee_src/include -emit-llvm -c -g -O0 -Xclang -disable-O0-optnone p3b.cpp
klee p3b.bc
```

Error

```
KLEE: ERROR: p3b.cpp:24: ASSERTION FAIL: f1 == f2
KLEE: NOTE: now ignoring this error at this location

KLEE: done: total instructions = 75
KLEE: done: completed paths = 3
KLEE: done: generated tests = 2
```

Error test case

```
ktest file : 'test000001.ktest'
args       : ['p3b.bc']
num objects: 2
object 0: name: 'x'
object 0: size: 4
object 0: data: b'\xff\xff\xff\xff'
object 0: hex : 0xffffffff
object 0: int : -1
object 0: uint: 4294967295
object 0: text: ....
object 1: name: 'y'
object 1: size: 4
object 1: data: b'\x00\x00\x00\x80'
object 1: hex : 0x00000080
object 1: int : -2147483648
object 1: uint: 2147483648
object 1: text: ....
```