Verification & Testing
Memory Debuggers

Anja Karl

V&T 3

Who has programmed in C?

Who had memory problems like invalid reads/writes or memory leaks?

Why are they so difficult to fix?

Who uses valgrind?

Memory Problems

Uninitialized read

```
int a, b;
a = b;
```

Unallocated read

```
int *p =
   (int*)malloc(4*sizeof(int));
printf("%d", p[4]);
```

Unallocated write

```
int *p =
   (int*)malloc(3*sizeof(int));
p[3] = 10;
```

Write after free

```
int *p =
   (int*)malloc(4*sizeof(int));
free(p);
p[2] = 10;
```

Memory Leak

```
int *p =
   (int*)malloc(4*sizeof(int));
end of program
```

Freeing unallocated memory

```
int *p;
free(p);
```

or

```
p = malloc(10 * sizeof(int));
free(p);
free(p);
```

Are these real problems?

None of These Errors Dump Core

- These errors do not always dump core. (Depending on compiler, OS)

- They sometimes produces expected results, sometimes unexpected results

- Uninitialized read: results depend on previous function call
  - `int a, b;`
  - `a = b;`

- Unallocated write may overwrite other data. May dump core if p points to the end of an allocated page,
  - `int *p = (int*)malloc(3*sizeof(int));`
  - `p[3] = 10;`

- Write after free: may overwrite other data if memory is reallocated before write. May dump core if memory is returned to OS

```
int *p = malloc(4*sizeof(int));
free(p);
p[2] = 10;
```

- Unallocated read. Returns data from different data structure.

```
int *p = malloc(4*sizeof(int));
int b;
b = p[4];
```

- Memory Leak. Slows program down and may dump core if in a loop.

```
int *p = malloc(4*sizeof(int));
end of program
```

Memory Errors

Memory Errors are

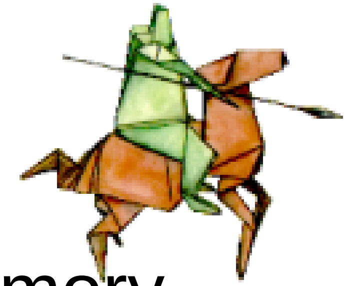
- hard to find
- often show themselves only occasionally
- often become apparent in different piece of code
- happen frequently!

Finding Memory Errors

List of tools that help with memory errors:

- IBM’s Purify (Rational)
- Valgrind (open source, Linux)
- electric fence (open source)
- dmalloc (open source)
- Clang & gcc sanitizer

Valgrind



Valgrind is a suite of tools, including a memory checker

- Translate to intermediate code
- Instrument intermediate code
- Execute on virtual CPU

Memcheck: increases code size 12x. Runs 25-50x slower.

Null: adds nothing, runs 4x slower

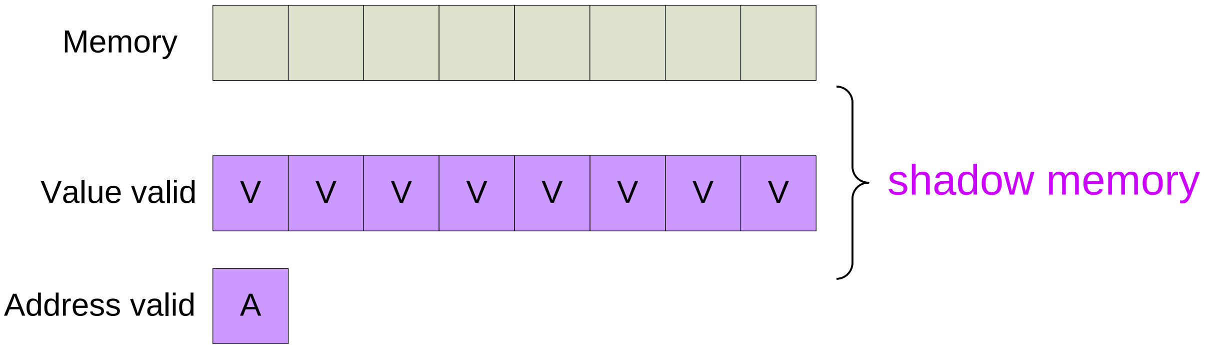
Valgrind Workings

Per byte of memory add

- Eight bits to store whether each of the bits has a valid value
- One bit to store whether the byte has been allocated

We want to find

- accesses where memory is not allocated
- decisions that depend on uninitialized values. (But: uninitialized copies are OK)



Valgrind Workings

- **Read or write:** Check A-bit
- **Load memory to CPU register:** also load value bits into shadow register
- **Writes:** set V-bits
- **Store register to memory:** store value bits into shadow memory
- **Value is used as address:** check V-bits
- **Branch depends on values:** check V-bits
- When value bits have been checked, they are set (prevents same error from being reported again)
- **Malloc/new:** address is valid, value is not. Keep “red zone” (address bits set to false) between memory chunks
- **free/delete:** check that memory has been allocated, prevent memory from being reallocated for as long as possible. Set A-bit to 0.

Examples

OK:

```
main(){
  int* a = malloc(sizeof(int));
  int *b = malloc(sizeof(int));
  *b = *a;
}

main(){
  int* a;
  *a = *a & 0xfffe;
  // bit 0 now initial'd
}
```

Wrong:

```
main(){
  int* a = malloc(sizeof(int));
  int* b = malloc(sizeof(int));
  *b = *a;
  printf("%d\n",*b);
}
```

Example

```
int *p;
int x = 1;
p = malloc(sizeof(int));
if(x){
   *p = 3;
   free(p);
   printf("%d",*p);
} else {
   printf("%d",*p);
}
```

More Details

Validity is kept on bit level. Need to properly handle

- Bit operations such as AND and OR
  - – $? \wedge 0 = 0$, but $? \wedge 1 = ?$
  - – $? \vee 0 = ?$, but $? \vee 1 = 1$
- Additions
- Shifts
- a XOR a
- Etc…

Example: Uninitialized Copy

```
int *p, *q;
max = user input, < 1024

p = (int*) malloc(1024*sizeof(int));
q = (int*) malloc(1024*sizeof(int));

for(i = 0; i < max; i++)
  p[i] = 0;

memcpy(q, p, 1024 * sizeof(int));

for(i = 0; i < max; i++)
  if(q[i])
    printf(“strange!\n”);

free(p); free(q);
```

This program is deemed correct by valgrind. Note that uninitialized values may be copied, as long as they are not visible.

Another example: a struct with four allocated bytes often takes up 8 bytes. Copying the struct copies uninitialized memory.

Bugs Valgrind Cannot Catch

```
void f(){
  int a[10];
  int b[10];

  printf(“%d\n”,b[0]);
  a[10] = 5;
  printf(“%d\n”,b[0]);
}
```

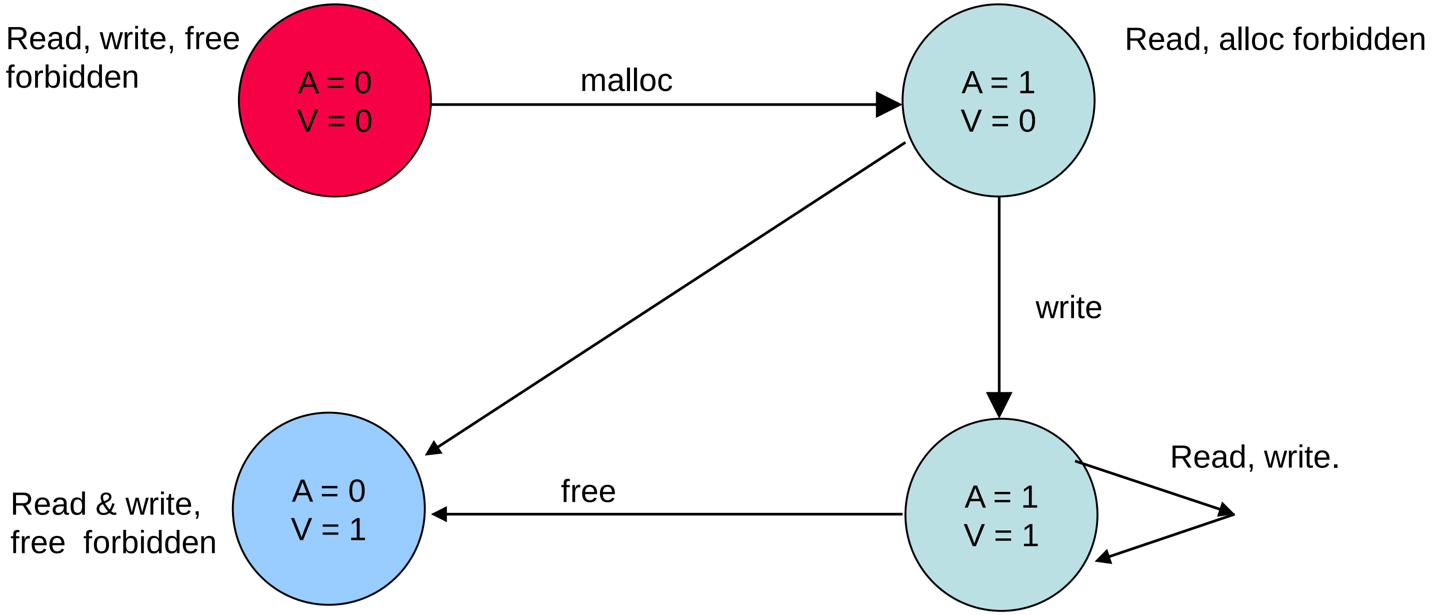
Valgrind cannot catch buffer overflows on static and local data. (only on malloc’ed data.) (*Why?*)

```
Valgrind --tool=memcheck --leak-check=yes
--suppressions=suppress.supp
```

Purify

Purify uses two bits of status per byte of memory

- Valid address?
- Valid data?



Purify

Less memory overhead: per byte, not per bit

No virtual CPU

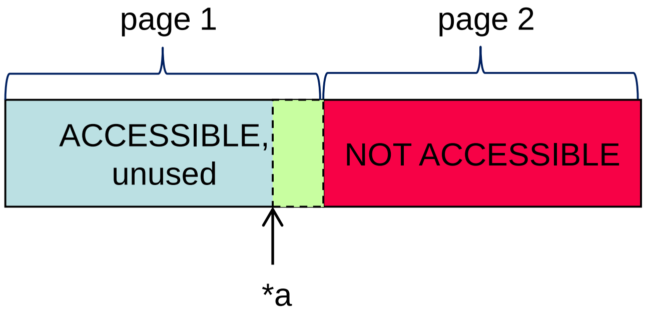
- Error flagged when uninitialized bytes read: uninitialized copies not allowed
- Faster, but more spurious warnings

Electric Fence

Memory is divided into *pages* (4096 bytes, usually)

- For every malloc, adjacent page of inaccessible memory is allocated
- MMU checks accesses to inaccessible pages without time overhead
- Memory overhead: every datastructure is at least 1 page
  - Big overhead if you have small datastructures!
  - The inaccessible page does not really count
- No virtual CPU, no annotation
- Only catches index too large accesses

```
a = malloc(128*sizeof(int))
```



More Valgrind Tools

Valgrind also includes

- Helgrind & Data Race Detector implement race condition detection (‘happens-before’)
- Massif is a heap profiler
- Callgrind is a profiler
- Cachegrind analyzes cache usage
- AddrCheck uses only A bits
- NullGrind