



软件分析

# 符号执行

熊英飞

北京大学



# 符号执行

- 抽象解释：每次分析完整程序，但在一个抽象域上进行
- 符号执行：每次分析一条路径，按某种顺序遍历路径
- 首先我们假设程序中
  - 没有数组
  - 没有指针
  - 没有函数调用
  - 没有系统调用



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

程序是否满足如下规约:

前条件:  $y > 0$

后条件:  $\text{main}(x,y) > 0$

即  $y > 0 \rightarrow \text{main}(x,y) > 0$



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

待探状态

x=a  
y=b  
z=0  
Next: 2  
Cond: b>0



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

当前状态

x=a  
y=b  
z=0  
Next: 2  
Cond: b>0

待探状态

x=a  
y=b+10  
z=0  
Next: 3  
Cond: b>0



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

当前状态

x=a  
y=b+10  
z=0  
Next: 3  
Cond: b>0

待探状态

x=a  
y=b+10  
z=0  
Next: 4  
Cond:  $a > 0 \wedge b > 0$

x=a  
y=b+10  
z=0  
Next: 8  
Cond:  $a \leq 0 \wedge b > 0$



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

当前状态

x=a  
y=b+10  
z=0  
Next: 4  
Cond:  $a > 0 \wedge b > 0$

待探状态

x=a+10  
y=b+10  
z=0  
Next: 5  
Cond:  $a > 0 \wedge b > 0$

x=a  
y=b+10  
z=0  
Next: 8  
Cond:  $a \leq 0 \wedge b > 0$



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

## 当前状态

x=a+10  
y=b+10  
z=0  
Next: 5  
Cond:  $a > 0 \wedge b > 0$

## 待探状态

x=a+10  
y=b+10  
z=(a+10)/5  
Next: 11  
Cond:  $a > 0 \wedge b > 0$

x=a  
y=b+10  
z=0  
Next: 8  
Cond:  $a \leq 0 \wedge b > 0$





# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

当前状态

x=a+10  
y=b+10  
z=(a+10)/5  
Next: 11  
Cond:  $a > 0 \wedge b > 0$

待探状态

x=a+10  
y=b+10  
z=(a+10)/5+(b+10)  
Next: 12  
Cond:  $a > 0 \wedge b > 0$

x=a  
y=b+10  
z=0  
Next: 8  
Cond:  $a \leq 0 \wedge b > 0$



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

## 当前状态

x=a+10  
y=b+10  
z=(a+10)/5+(b+10)  
Next: 12  
Cond:  $a > 0 \wedge b > 0$

- 求  $a > 0 \wedge b > 0 \rightarrow (a+10)/5+(b+10) > 0$  是否恒成立
- 取反后判可满足性
- 如果满足，则程序不满足规约

## 待探状态

x=a  
y=b+10  
z=0  
Next: 8  
Cond:  $a \leq 0 \wedge b > 0$



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

当前状态

x=a  
y=b+10  
z=0  
Next: 8  
Cond:  $a \leq 0 \wedge b > 0$

待探状态

x=a+10  
y=b+10  
z=a/5+2+(b+10)  
Next: 12  
Cond:  $a \leq 0 \wedge b > 0$



# 符号执行

```
1.  int main(x,y) {  
2.    y+=10;  
3.    if (x>0) {  
4.      x+=10;  
5.      z=x/5;  
6.    }  
7.    else {  
8.      z=x/5+2;  
9.      x+=10;  
10.   }  
11.   z+=y;  
12.   return z;  
13. }
```

当前状态

x=a+10  
y=b+10  
z=a/5+2+(b+10)  
Next: 12  
Cond:  $a \leq 0 \wedge b > 0$

- 求  $a \leq 0 \wedge b > 0 \rightarrow a/5+2+(b+10)>0$  是否恒成立
- 取反后判可满足性
- 如果满足，则程序不满足规约

待探状态



# 符号执行小结

- 状态：
  - 当前变量的符号值
  - 下一条待执行的语句
  - 当前状态的路径约束
- 不断遍历状态，遇上结束状态时用SMT求解器判断规约是否被满足
- 程序不满足规约=任意路径不满足规约
- 程序满足规约=所有路径满足规约
  - 路径无穷多时做不到
  - 通常遍历有限次循环来模拟



# 符号执行 vs 抽象解释

- 抽象解释是所有执行的上近似
- 符号执行是所有执行的下近似
- 抽象解释迭代过程中的任意状态表示了长度小于等于当前迭代次数的所有执行踪迹集合的超集
- 符号执行迭代过程中的任意状态表示了沿着某条路径执行特定步数之后的任意精确执行踪迹集合



# 符号执行检查程序内部错误

- 用符号执行发现缓冲区溢出
  - .....
  - `a[i] = 4;`
    - 判断后条件 $0 \leq i \&\& i < a.length$ 是否总是成立
- 用符号执行发现除0错误
  - .....
  - `x = 3 / i;`
    - 判断后条件 $i \neq 0$ 是否总是成立
- 用符号执行发现路径可行性
  - 判断给定路径上的路径约束是否可满足



# 调用约束求解的时机

1. if ( $x > 0$ ) {
  2.      $y = \dots$
  3. }
  4. if ( $x \leq 0$ ) {
  5.      $y = \dots$
  6. }
- 无需对不可达路径继续探索或验证
  - Eager evaluation: 在分支的时候就判断路径的可达性
  - Lazy evaluation: 只对完整路径判断
    - 和之前的方法等价
  - 在不同程序中两种方法各有优劣
    - eager evaluation对同一条路径可能调用更多次，但探索的路径总数会减少
    - 如果路径不可达，lazy evaluation的约束不会更简单，但有可能更容易产生冲突，导致解得更快





# 从冲突中学习

```
1.  if (x>0) {  
2.      y = ...  
3.  }  
4.  if (z > 0) {  
5.      ...  
6.  } else {  
7.      ...  
8.  }  
9.  if (x<=0) {  
10.     y = ...  
11. }
```

- 第1行的条件和第9行的条件互斥，但有多条路径都包括这两个条件
- 假设先选择路径1-2-4-5-9-10，则生成路径约束
  - 1:  $x > 0$
  - 4:  $z > 0$
  - 9:  $x \leq 0$
- SMT判断不能满足，返回矛盾集(1, 9)
- 假设又选择路径1-2-4-7-9-10，则路径约束包含(1,9)，不用调用SMT约束求解直接可知不可满足



# 处理数组

```
1. int main(int[] array, int index) {  
2.   array[index]=1;  
3.   assert(array[0]==1);  
4. }
```

通过SMT的Array Theory处理

array=a  
index=b  
Next: 2

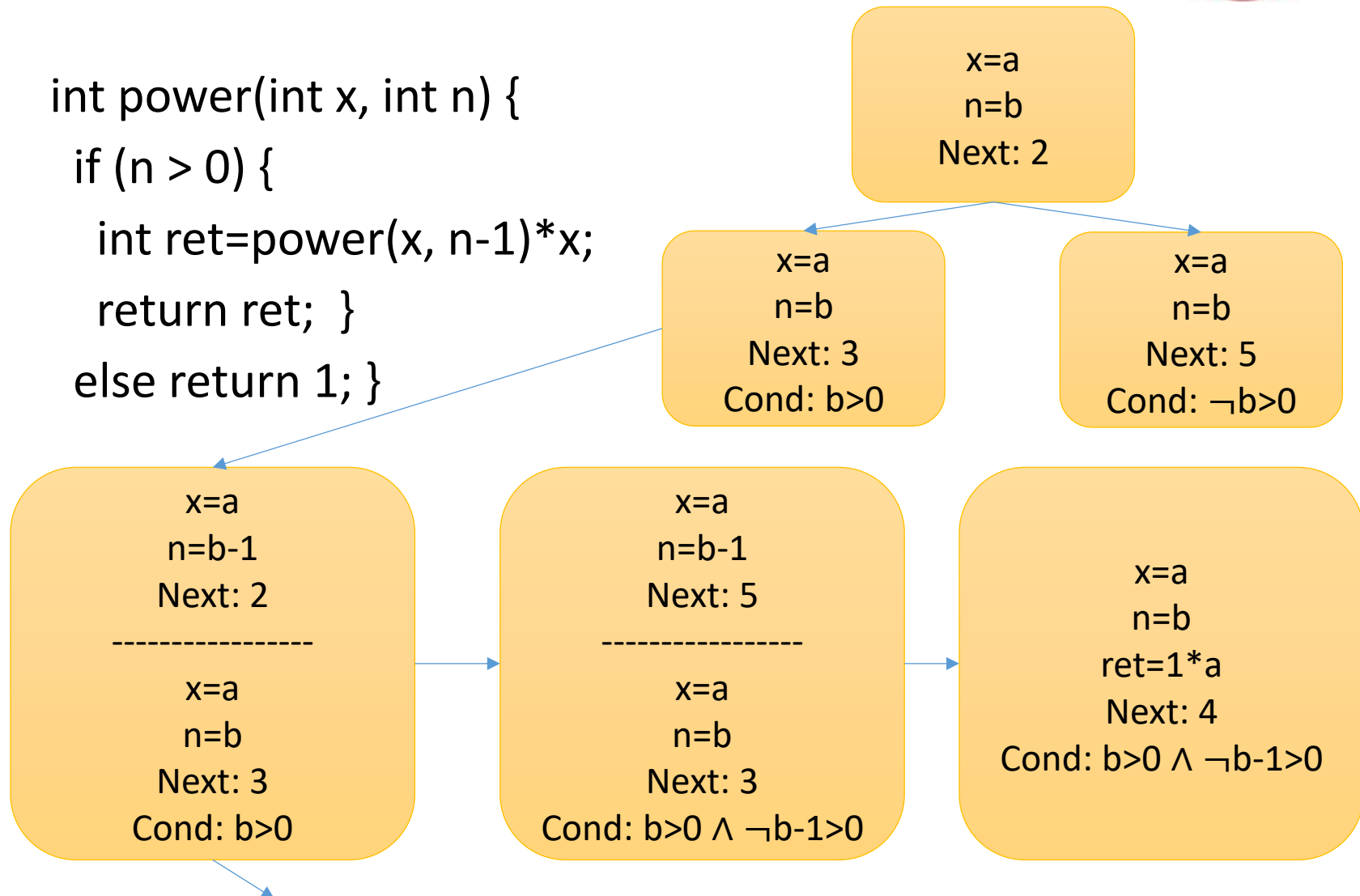
array=write(a, b, 1)  
index=b  
Next: 3

判断True->read(write(a, b,  
1), 0)=1是否恒成立



# 处理函数调用

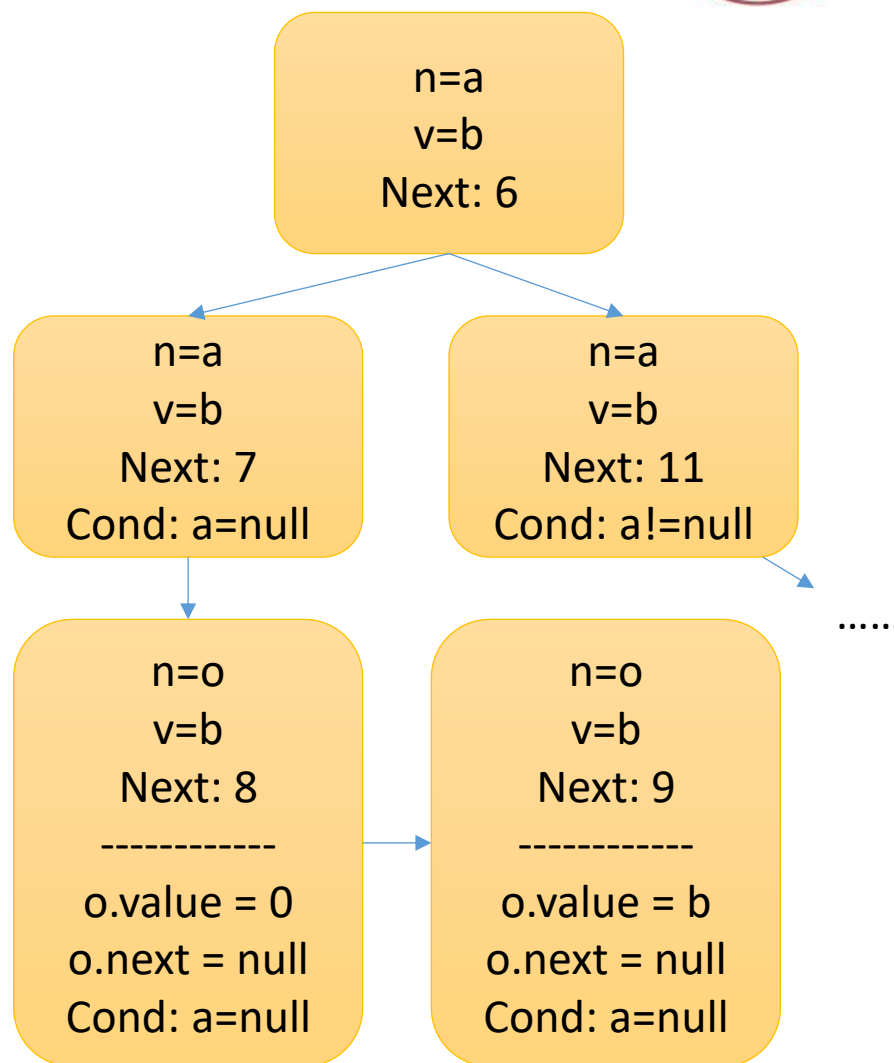
```
1. int power(int x, int n) {  
2.   if (n > 0) {  
3.     int ret=power(x, n-1)*x;  
4.     return ret; }  
5.   else return 1; }
```





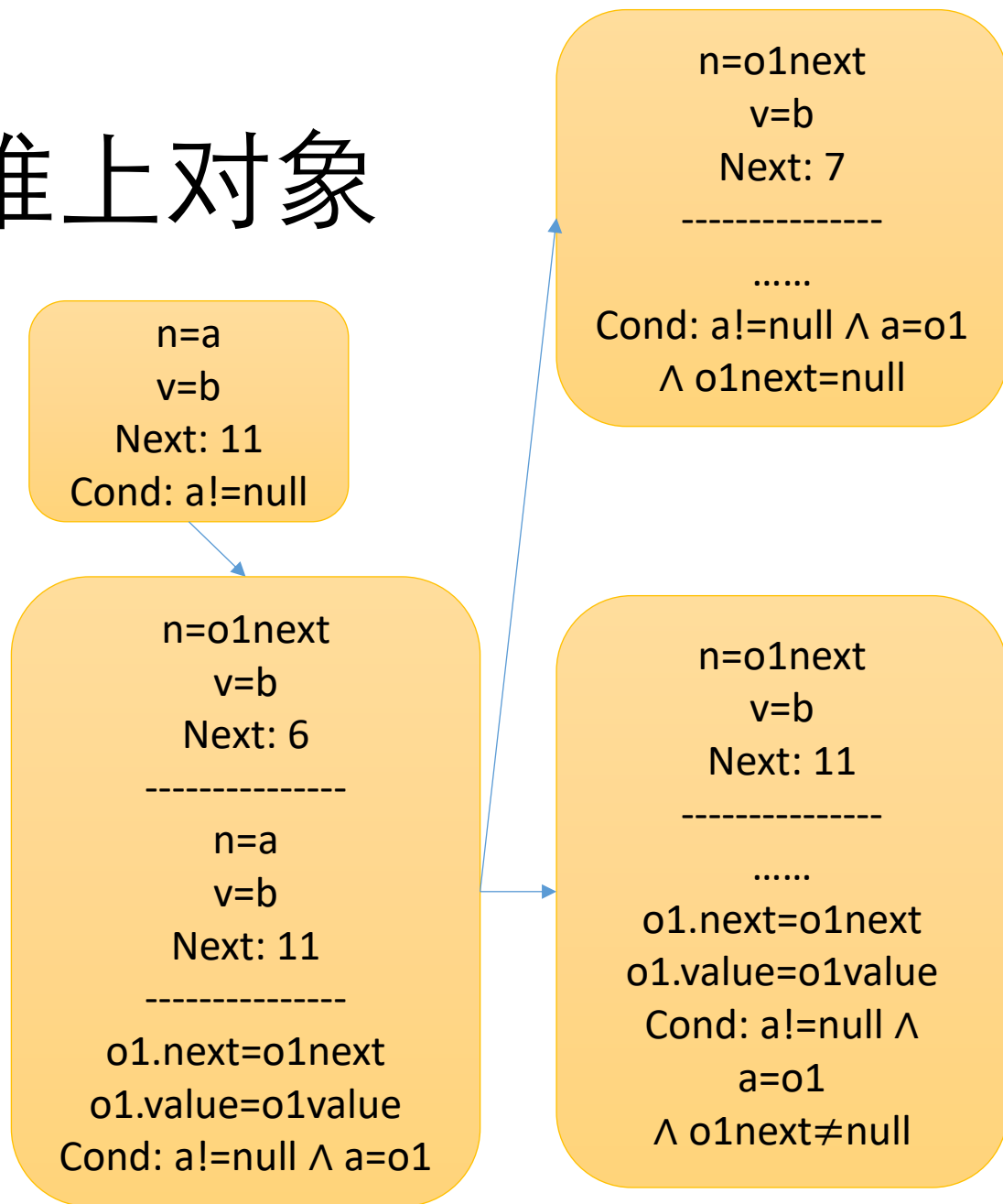
# 处理指针和堆上对象

```
1.  Class Node {  
2.      int value;  
3.      Node next;  
4.  }  
5.  int append(Node n, int v) {  
6.      if(n == null) {  
7.          n = new Node();  
8.          n.value = v;  
9.          return n;  
10.     }  
11.     n.next = append(n.next, v);  
12.     return n;  
13. }
```



# 处理指针和堆上对象

```
1.  Class Node {  
2.    int value;  
3.    Node next;  
4.  }  
5.  int append(Node n, int v) {  
6.    if(n == null) {  
7.      n = new Node();  
8.      n.value = v;  
9.      return n;  
10.   }  
11.   n.next = append(n.next, v);  
12.   return n;  
13. }
```



# 处理指针和堆上对象

```
1.  Class Node {
2.      int value;
3.      Node next;
4.  }
5.  int append(Node n, int v) {
6.      if(n == null) {
7.          n = new Node();
8.          n.value = v;
9.          return n;
10.     }
11.     n.next = append(n.next, v);
12.     return n;
13. }
```

n=o1next  
v=b  
Next: 11  
-----  
.....  
o1.next=o1next  
o1.value=o1value  
Cond: a!=null  $\wedge$   
a=o1  
 $\wedge$  o1next!=null

n=o1next  
v=b  
Next: 11  
-----  
.....  
o1.next=o1next  
o1.value=o1value  
Cond: a!=null  $\wedge$   
a=o1  
 $\wedge$  o1next!=null  
 $\wedge$  o1next = o1

n=o1next  
v=b  
Next: 11  
-----  
.....  
o1.next=o1next  
o1.value=o1value  
o2.next=o2next  
o2.value=o2value  
Cond: a!=null  $\wedge$  a=o1  
 $\wedge$  o1next!=null  
 $\wedge$  o1next = o2

需要访问o1next.next, 遍历o1next的所有可能性



# 处理指针和堆上对象

- 以上处理方法会形成非法对象
  - `o1.next==o1`，这在实际链表中不会出现
- 解决方法1：添加repOK函数

```
public static boolean repOK(Node head) {
    Node slow = head;
    Node fast = head;
    while (fast != null && fast.next != null) {
        slow = slow.next;
        fast = fast.next.next;
        if (slow == fast) return false; // found cycle
    }
    return true; // no cycle
}
```
- 对所有参数先执行repOK检查

```
int append(Node n, int v) {
    if (repOK(n)) {
        ...
    }
}
```
- 缺陷： repOK需要用户撰写，工作量较大



# 处理指针和堆上对象

- 解决方法2：只利用公共方法创建对象

```
class LinkedList<E> {  
    add(E element);  
    remove(int index);}
```

```
LinkedList<int> randomList() {  
    LinkedList<int> a = new LinkedList();  
    while(randomBool()) {  
        switch(randomInt(2)) {  
            case 0:  
                a.add(randomInt()); break;  
            case 1:  
                a.remove(randomInt()); break;  
        }  
    }  
}
```

```
int append(int v) {  
    n = randomList(); ....
```

每一次调用randomInt或者randomBool都创建一个新的符号值





# 处理指针和堆上对象

- 更复杂一点的例子，对象作为参数和返回值，需维护对象池

```
class MyString {  
    MyString(int); //根据传输参数选择一个预先定义好的字符串  
    MyString concat(MyString);  
}
```

```
MyString randomMyString() {  
    Set<MyString> as = new HashSet<MyString>();  
    as.add(new MyString(randomInt()));  
    while(true) {  
        MyString a;  
        switch(randomInt(2)) {  
            case 0:  
                a = new MyString(randomInt()); break;  
            case 1:  
                a = randomFrom(as).concat(randomFrom(as)); break;  
        }  
        if (randomBool()) return a; else as.add(a);  
    }  
}
```



# 处理指针和堆上对象

- 解决方案2的优化：
  - 在搜索过程中检查对象池中的对象的指向图
  - 如果两条路径到达的对象池，其对象指向图同构，则用合并两条路径到达的对象池，并将路径条件用“或者”连接
  - 比如：
    - `a = new LinkedList<int>();`
    - `a.add(1);`
    - `a.remove(0);`
  - 和
    - `a = new LinkedList<int>();`
- 详见：
  - Yunfan Zhang, Ruidong Zhu, Yingfei Xiong, Tao Xie. Efficient Synthesis of Method Call Sequences for Test Generation and Bounded Verification. ASE'22



# 约束求解失败的情况

- 形成了复杂条件
  - $x^5 + 3x^3 == y$
  - `p->next->value == x`
- 调用了系统调用
  - `If (file.read()==x)`
- 动态符号执行
  - 混合程序的真实执行和符号执行
  - 在约束求解无法进行的时候，用真实值代替符号值
    - 如果真实值 $x=10$ ，则 $x^5 + 3x^3 == y$ 变为 $103000==y$ ，可满足



# 动态符号执行-例子

```
1. int foo (int v) {  
2.     return (v*v) % 50; }  
3. void testme (int x, int y) {  
4.     z = foo (y);  
5.     if (x > y+10) {  
6.         if (z == x)  
7.             assert(false);  
8.     }}
```

x=a  
y=b  
z=b\*b % 50  
Next: 7  
Cond:  $a > b + 10$   
 $\wedge b*b \% 50 == a$

求解约束:

$a > b + 10 \wedge b*b \% 50 = a \rightarrow \text{false}$

会返回Unknown



# 动态符号执行-例子

```
1. int foo (int v) {  
2.   return (v*v) % 50; }  
3. void testme (int x, int y) {  
4.   z = foo (y);  
5.   if (x > y+10) {  
6.     if (z == x)  
7.       assert(false);  
8.   }}
```

x=1, a  
y=5, b  
z=0, 0  
Next: 4  
Cond:

x=1, a  
y=5, b  
z=25, (b\*b) % 50  
Next: 5  
Cond:

x=1, a  
y=5, b  
z=25, (b\*b) % 50  
Next: --  
Cond:  $\neg(a > b + 10)$

保存条件的  
列表



# 动态符号执行-例子

```
1. int foo (int v) {  
2.   return (v*v) % 50; }  
3. void testme (int x, int y) {  
4.   z = foo (y);  
5.   if (x > y+10) {  
6.     if (z == x)  
7.       assert(false);  
8.   }}
```

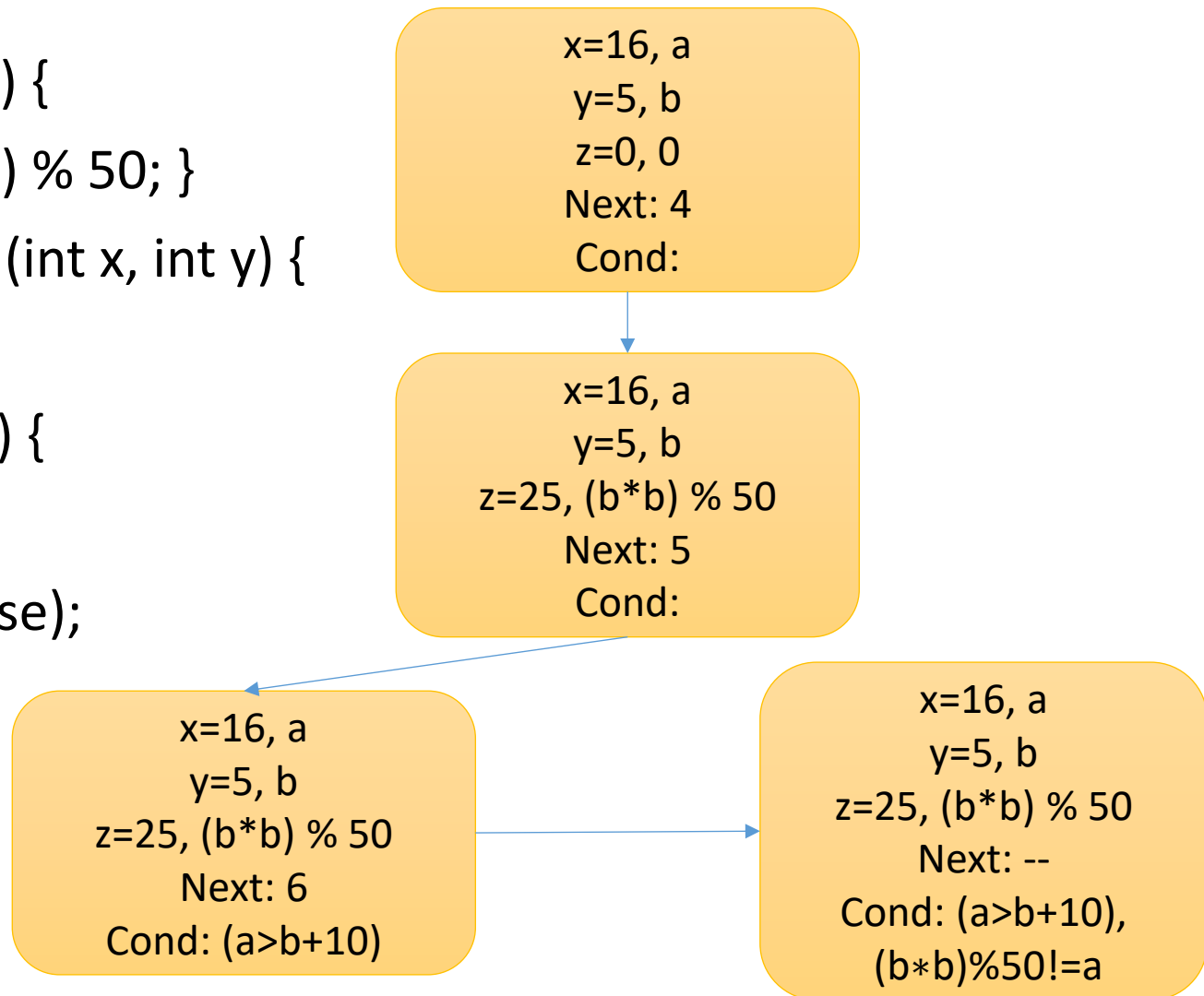
x=1, a  
y=5, b  
z=25, (b\*b) % 50  
Next: --  
Cond:  $\neg(a > b + 10)$

- 从列表中取最后一个没有被取反过的条件取反
  - $a > b + 10$
- 求解，得到  $a=16, b=5$



# 动态符号执行-例子

```
1. int foo (int v) {  
2.   return (v*v) % 50; }  
3. void testme (int x, int y) {  
4.   z = foo (y);  
5.   if (x > y+10) {  
6.     if (z == x)  
7.       assert(false);  
8.   }}
```





# 动态符号执行-例子

```
1. int foo (int v) {  
2.     return (v*v) % 50; }  
3. void testme (int x, int y) {  
4.     z = foo (y);  
5.     if (x > y+10) {  
6.         if (z == x)  
7.             assert(false);  
8.     }}
```

x=16, a  
y=5, b  
z=25, (b\*b) % 50  
Next: --  
Cond: (a>b+10),  
(b\*b)%50!=a

- 从列表中取最后一个没有被取反过的条件取反
  - $(a > b + 10) \wedge (b * b) \% 50 = a$
- 求解，发现%不被SMT支持
- 将%运算涉及的变量替换成具体值
  - $(a > 5 + 10) \wedge (5 * 5) \% 50 = a$
- 得到a=25





# 动态符号执行-例子

```
1. int foo (int v) {  
2.   return (v*v) % 50; }  
3. void testme (int x, int y) {  
4.   z = foo (y);  
5.   if (x > y+10) {  
6.     if (z == x)  
7.       assert(false);  
8.   }}
```

x=25, a  
y=5, b  
z=0, 0  
Next: 4  
Cond:

x=25, a  
y=5, b  
z=25, (b\*b) % 50  
Next: 7  
Cond: (a>b+10)  
 $\wedge$  (b\*b)%50=a

具体执行直接触发AssertionError



# 动态符号执行小结

- 替换变量为具体值的方法不保证完整性
  - 可满足的约束可能变得不可满足
  - $(a > b + 10) \wedge (b * b) \% 50 = a$  中，如果  $b = 0$ ，则不可满足
- 但效果一定优于静态符号执行
  - 替换只在原约束无法求解的情况下进行
- 为什么将具体值代入执行而不是在约束无法求解的时候替换为随机值？
  - 因为有可能拿不到完整的路径约束
  - 比如路径约束有可能是  $\text{foo}(x) == 0$ ，但  $\text{foo}$  的实现代码拿不到， $\text{foo}$  也不能脱离系统状态单独调用



# 常见符号执行工具

- C语言：KLEE
- Java语言：SymbolicPathFinder, JBSE
- C#语言：Pex
  - 北京大学谢涛老师和微软合作开发



# 基于霍尔逻辑的符号 执行



# 霍尔逻辑

- 霍尔三元组
  - $\{\text{前条件}\} \text{语句} \{\text{后条件}\}$
- 霍尔逻辑表示三者之间的推导关系
- 又称为公理语义



# IMP语言

Statement ::=

| skip

| while (Expr) Statement

| if (Expr) Statement else Statement

| Statement; Statement

| Var = Expr



# 霍尔逻辑规则

$$\text{SKIP} \frac{}{\{P\} \text{ skip } \{P\}}$$

$$\text{ASSIGN} \frac{}{\{P[a/x]\} x := a \{P\}}$$

$$\text{SEQ} \frac{\{P\} c_1 \{R\} \quad \{R\} c_2 \{Q\}}{\{P\} c_1; c_2 \{Q\}}$$

$$\text{IF} \frac{\{P \wedge b\} c_1 \{Q\} \quad \{P \wedge \neg b\} c_2 \{Q\}}{\{P\} \text{ if } b \text{ then } c_1 \text{ else } c_2 \{Q\}}$$

$$\text{CONSEQUENCE} \frac{\models (P \Rightarrow P') \quad \{P'\} c \{Q'\} \quad \models (Q' \Rightarrow Q)}{\{P\} c \{Q\}}$$

$$\text{WHILE} \frac{\{P \wedge b\} c \{P\}}{\{P\} \text{ while } b \text{ do } c \{P \wedge \neg b\}}$$



# 用霍尔逻辑证明举例

- $\text{if } (x > 0) \ x \ += \ 10; \text{ else } x = 20;$ 
  - 该程序执行结束后， $x$ 是否一定大于0?
- 根据Assign，可得
  - $\{x+10>0\} \ x+=10 \ \{x > 0\}$
  - $\{\text{True}\} \ x=20 \ \{x > 0\}$
- 因为 $x>0 \Rightarrow x+10 > 0$ 且 $\neg x>0 \Rightarrow \text{True}$ ，根据Consequence，可得
  - $\{x>0\} \ x+=10 \ \{x > 0\}$
  - $\{\neg x>0\} \ x=20 \ \{x > 0\}$
- 根据If，可得
  - $\{\text{True}\} \ \text{if } (x > 0) \ x \ += \ 10; \text{ else } x = 20; \{x>0\}$





# 用霍尔逻辑证明练习

- `while (x < 10) x += 1;`
  - 该程序执行结束后，`x`是否一定大于0?
- 根据Assign，可得
  - $\{True\} x+=1 \{True\}$
- 根据Consequence，可得
  - $\{x<10 \wedge True\} x+=10 \{True\}$
- 根据While，可得
  - $\{True\} \text{while } (x < 10) x += 1; \{x \geq 10\}$
- 根据Consequence，可得
  - $\{True\} \text{while } (x < 10) x += 1; \{x > 0\}$



# 谓词转换计算

- 最弱前条件计算：给定后条件和语句，求能形成霍尔三元组的最弱前条件
- 最强后条件计算：给定前条件和语句，求能形成霍尔三元组的最强后条件
- 基于谓词转换的符号执行
  - 给定输入需要满足的条件 $P$ ，代码 $c$ ，输出需要满足的条件 $Q$
  - 前向符号执行：基于 $P$ 和 $c$ 计算最强后条件 $Q'$ ，验证 $Q' \rightarrow Q$ 是否恒成立
  - 后向符号执行：基于 $Q$ 和 $c$ 计算最弱前条件 $P'$ ，验证 $P \rightarrow P'$ 是否恒成立



# 最弱前条件计算

- $wp(skip, Q) = Q$

$$\text{SKIP} \frac{}{\{P\} \text{ skip } \{P\}}$$

- $wp(x := a, Q) = Q[a/x]$

$$\text{ASSIGN} \frac{}{\{P[a/x]\} x := a \{P\}}$$

- $wp(c_1; c_2, Q) =$   
 $wp(c_1, wp(c_2, Q))$

$$\text{SEQ} \frac{\{P\} c_1 \{R\} \quad \{R\} c_2 \{Q\}}{\{P\} c_1; c_2 \{Q\}}$$

- $wp(\text{if } b \text{ then } c_1 \text{ else } c_2, Q) =$   
 $(b \rightarrow wp(c_1, Q))$   
 $\wedge (\neg b \rightarrow wp(c_2, Q))$

$$\text{IF} \frac{\{P \wedge b\} c_1 \{Q\} \quad \{P \wedge \neg b\} c_2 \{Q\}}{\{P\} \text{ if } b \text{ then } c_1 \text{ else } c_2 \{Q\}}$$



# 最弱前条件： 举例

- $\text{wp}(\text{if } (x > 0) \ x \ += \ 10; \text{ else } \ x = 20, \ x > 0)$ 
  - $= (x > 0 \rightarrow \text{wp}(x += 10, x > 0)) \wedge (x \leq 0 \rightarrow \text{wp}(x = 20, x > 0))$
  - $= (x > 0 \rightarrow x + 10 > 0) \wedge (x \leq 0 \rightarrow 20 > 0)$
  - $= \text{True}$



# 最弱前条件计算：循环

- $wp(\text{while } b \text{ do } c, Q) = \exists i \in \text{Nat}. L_i(Q)$ 
  - where
    - $L_0(Q) = \text{false}$
    - $L_{i+1}(Q) = (\neg b \Rightarrow Q) \wedge (b \Rightarrow wp(c, L_i(Q)))$
- $i$ 代表循环最多执行了 $i - 1$ 次
- 注意这个最弱前条件蕴含了循环必然终止



# 最强后条件计算

- $sp(P, skip) = P$
- $sp(P, x := a) = \exists n. x = a[n/x] \wedge P[n/x]$
- $sp(P, c_1; c_2) = sp(sp(P, c_1), c_2)$
- $sp(P, if\ b\ then\ c_1\ else\ c_2) = sp(b \wedge P, c_1) \vee sp(\neg b \wedge P, c_2)$
- $sp(P, while\ b\ do\ c) = \neg b \wedge \exists i. L_i(P)$ 
  - where
    - $L_0(P) = P$
    - $L_{i+1}(P) = sp(b \wedge L_i(P), c)$

因为约束更复杂，实际使用较少



# 符号执行和谓词转换

- 在没有循环时，谓词转换计算和符号执行等价
- 例：{ $x \leq 0$ } if ( $y > 0$ )  $x += 1$ ; else  $x += 2$ ; { $x < 3$ }
- 符号执行
  - 令  $x=a$ ,  $y=b$ , 计算得到
    - $a \leq 0 \rightarrow (b > 0 \wedge a + 1 < 3) \vee (\neg b > 0 \wedge a + 2 < 3)$
- 最弱前条件
  - $wp(x += 1, x < 3) = x + 1 < 3$
  - $wp(x += 2, x < 3) = x + 2 < 3$
  - $wp(\text{原程序}, x < 3) = (y > 0 \rightarrow x + 1 < 3) \wedge (\neg y > 0 \rightarrow x + 2 < 3) = (y > 0 \wedge x + 1 < 3) \vee (\neg y > 0 \wedge x + 2 < 3)$
  - 原三元组等价于
    - $x \leq 0 \rightarrow (y > 0 \wedge x + 1 < 3) \vee (\neg y > 0 \wedge x + 2 < 3)$



# 符号执行和谓词转换

- 在没有循环时，谓词转换计算和符号执行等价
- 例：  $\{x \leq 0\} \text{If } (y > 0) \ x += 1; \text{ else } x += 2; \{x < 3\}$
- 符号执行
  - 令  $x=a, y=b$ , 计算得到
    - $a \leq 0 \rightarrow (b > 0 \wedge a + 1 < 3) \vee (\neg b > 0 \wedge a + 2 < 3)$
- 最强后条件
  - $sp(x \leq 0 \wedge y > 0, x = x + 1) = \exists n. (x = n + 1 \wedge n \leq 0 \wedge y > 0)$
  - $sp(x \leq 0 \wedge \neg y > 0, x = x + 2) = \exists n. (x = n + 2 \wedge n \leq 0 \wedge \neg y > 0)$
  - $sp(x = 0, \text{原程序}) = (\exists n. (x = n + 1 \wedge n \leq 0 \wedge y > 0)) \vee (\exists n. (x = n + 2 \wedge n \leq 0 \wedge \neg y > 0))$
  - 原三元组等价于
    - $(\exists n. (x = n + 1 \wedge n \leq 0 \wedge y > 0)) \vee (\exists n. (x = n + 2 \wedge n \leq 0 \wedge \neg y > 0)) \rightarrow x < 3 =$
    - $(\exists n. (x = n + 1 \wedge n \leq 0 \wedge y > 0) \vee (x = n + 2 \wedge n \leq 0 \wedge \neg y > 0)) \rightarrow x < 3 =$
    - $\forall n. ((x = n + 1 \wedge n \leq 0 \wedge y > 0) \vee (x = n + 2 \wedge n \leq 0 \wedge \neg y > 0) \rightarrow x < 3) =$
    - $n \leq 0 \rightarrow ((y > 0 \rightarrow n + 1 < 3) \wedge (\neg y > 0 \rightarrow n + 2 < 3)) =$
    - $n \leq 0 \rightarrow ((y > 0 \wedge n + 1 < 3) \vee (\neg y > 0 \wedge n + 2 < 3))$





# 作业

- 符号执行是否具备部分收敛性？给定任意程序和任意前后条件，期望判断后条件是否总是成立，是否总是存在一个值 $k$ ，使得循环展开任意 $l \geq k$ 次后的分析结果都和理想结果相同。
  - 简单起见，只考虑IMP程序，不考虑数组、指针、函数调用等。



# 参考资料

- 《程序设计语言的形式语义》 Glynn Winskel著
- Baldoni R , Coppa E , D'elia, Daniele Cono, et al. A Survey of Symbolic Execution Techniques[J]. Acm Computing Surveys, 2016.