

# Automated Software Analysis Techniques For High Reliability: A Concolic Testing Approach

Moonzoo Kim

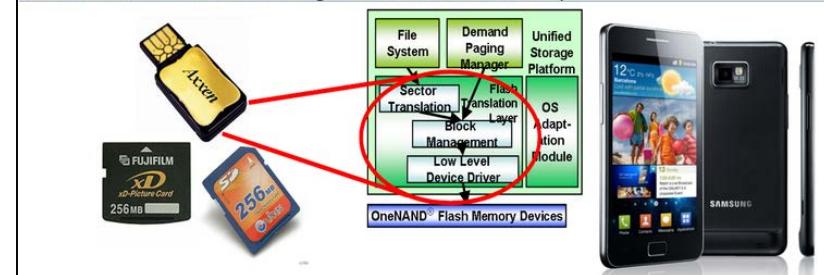


## SW Testing & Verification Group

Home Members Research Projects Publications Courses Lab Seminar Tools Data Link

You are here: Home

### Welcome to Software Testing and Verification Group

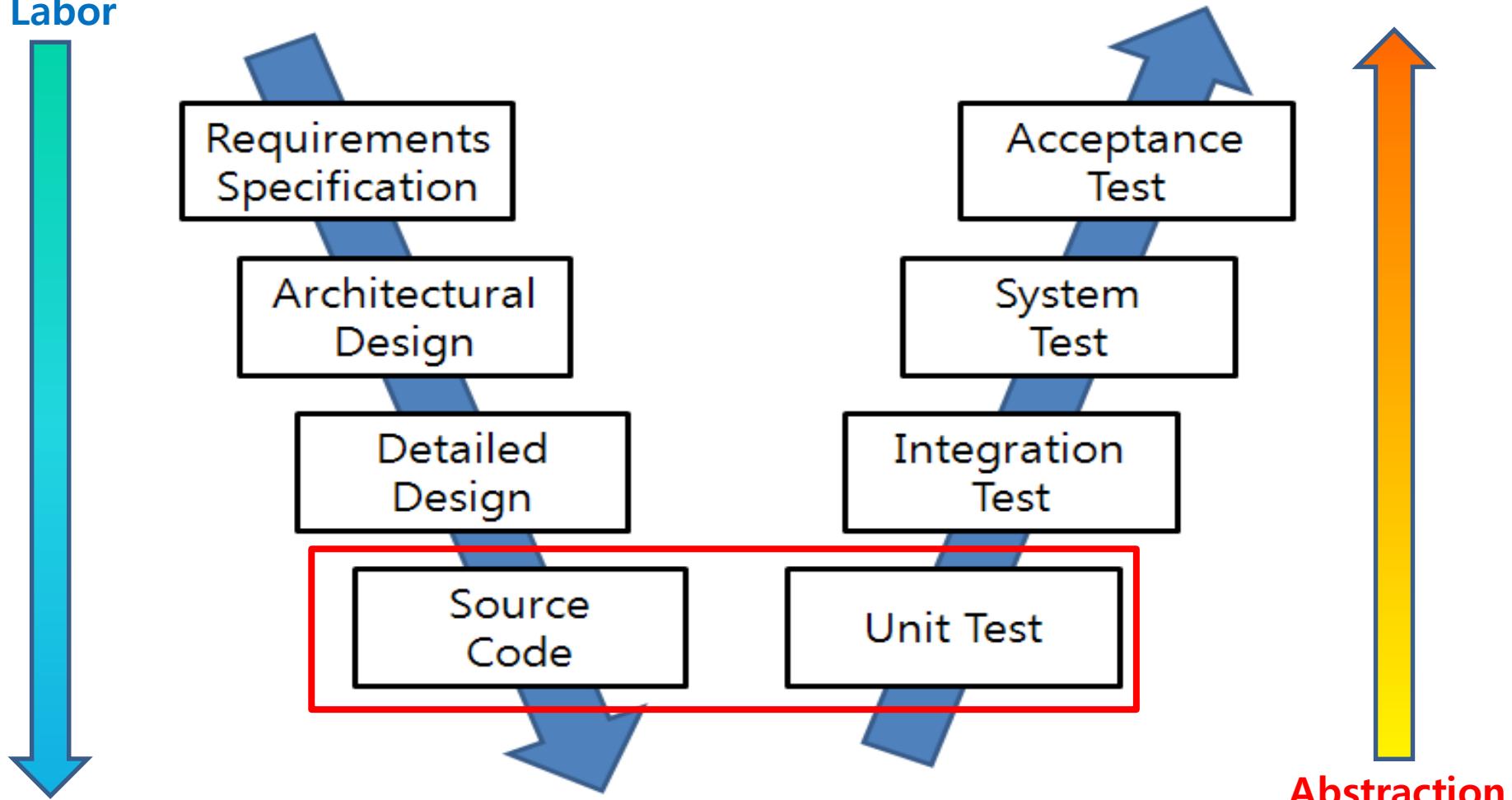


# Contents

- Automated Software Analysis Techniques
  - Background
  - Concolic testing process
  - Example of concolic testing
- Case Study: Busybox utility
- Future Direction and Conclusion

# Main Target of Automated SW Analysis

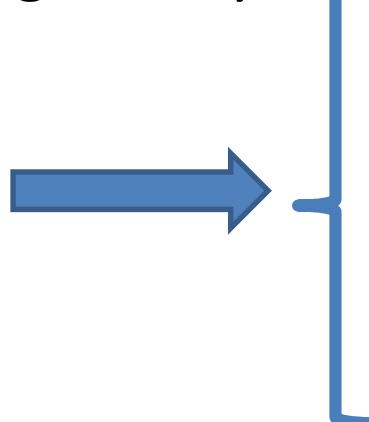
Manual  
Labor



# Automated Software Analysis Techniques

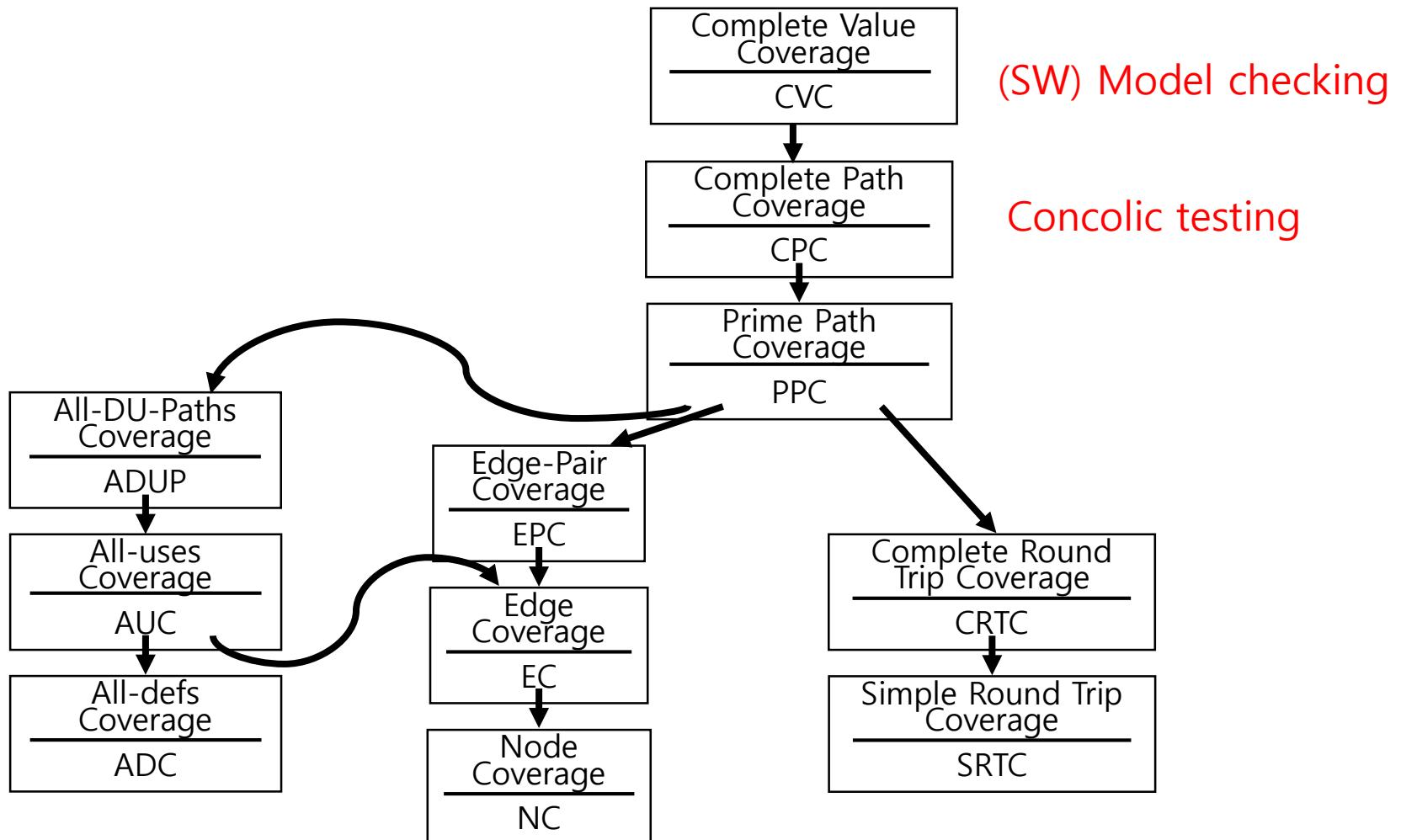
- Aims to explore possible behaviors of target systems in an exhaustive manner
- Key methods:
  - Represents a target program/or executions as a “logical formula”
  - Then, analyze the logical formula (a target program) by using logic analysis techniques

*Weakness of  
conventional  
testing*



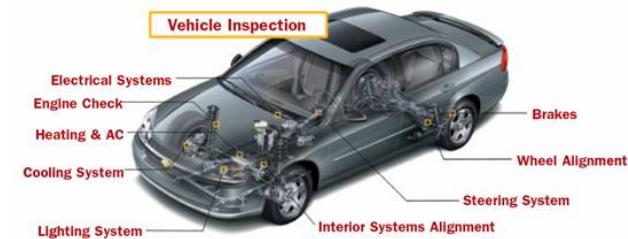
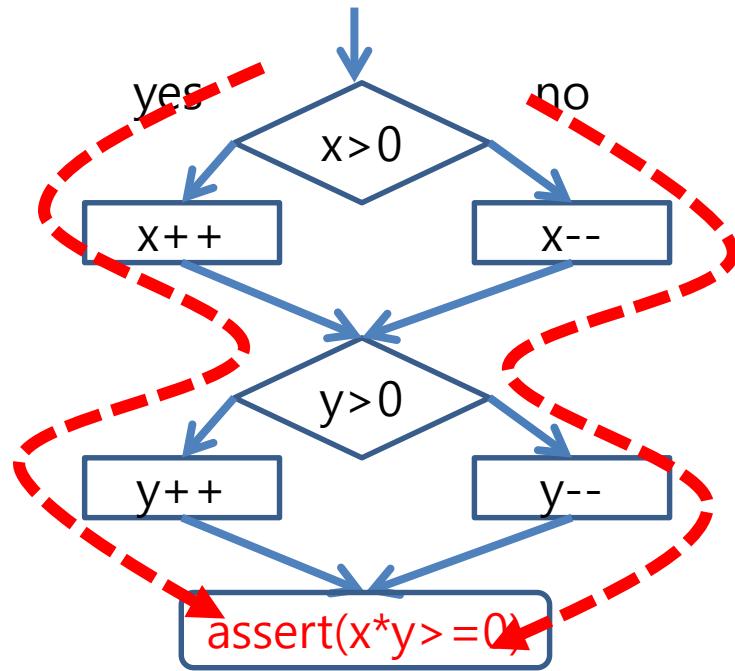
- Symbolic execution (1970)
- Model checking (1980)
- SW model checking (2000)
- Concolic testing (2005 ~)

# Hierarchy of SW Coverages



# Weaknesses of the Branch Coverage

```
/* TC1: x= 1, y= 1;  
TC2: x=-1, y=-1;*/  
void foo(int x, int y) {  
    if ( x > 0)  
        x++;  
    else  
        x--;  
    if(y >0)  
        y++;  
    else  
        y--;  
    assert (x * y >= 0);  
}
```



Systematic testing techniques are necessary for quality software!  
-> Integration testing is not enough  
-> Unit testing with automated test case generation is desirable  
for both productivity and quality

# Dynamic v.s. Static Analysis

	<b>Dynamic Analysis (i.e., testing)</b>	<b>Static Analysis (i.e. model checking)</b>
Pros	<ul style="list-style-type: none"><li>•Real result</li><li>•No environmental limitation</li><li>•Binary library is ok</li></ul>	<ul style="list-style-type: none"><li>•Complete analysis result</li><li>•Fully automatic</li><li>•Concrete counter example</li></ul>
Cons	<ul style="list-style-type: none"><li>•Incomplete analysis result</li><li>•Test case selection</li></ul>	<ul style="list-style-type: none"><li>•Consumed huge memory space</li><li>•Takes huge time for verification</li><li>•False alarms</li></ul>

**-> Concolic testing**

# Concolic Approach

- Combine concrete and symbolic execution
  - **Concrete + Symbolic = Concolic**
- In a nutshell, concrete execution over a concrete input guides symbolic execution
  - No false positives
- **Automated** testing of real-world C Programs
  - Execute target program on **automatically** generated test inputs
  - **All possible execution paths** are to be explored
  - Higher branch coverage than random testing

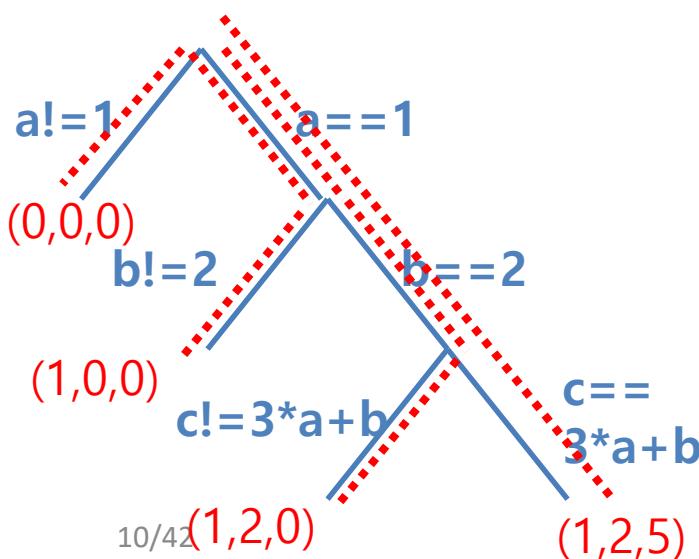
# Overview of Concolic Testing Process

1. Select input variables to be handled symbolically
2. A target C program is statically instrumented with probes, which record symbolic path conditions
3. The instrumented C program is executed with given input values
  - Initial input values are assigned randomly
4. Obtain a symbolic path formula  $\varphi_i$  from a concrete execution over a concrete input
5. One branch condition of  $\varphi_i$  is **negated** to generate the next symbolic path formula  $\psi_i$
6. A constraint solver solves  $\psi_i$  to get next concrete input values
  - Ex.  $\varphi_i: (x < 2) \&\& (x + y \geq 2)$  and  $\psi_i: (x < 2) \&\& (x + y < 2)$ .  
One solution is  $x=1$  and  $y=0$
7. Repeat step 3 until all feasible execution paths are explored

Iterations

# Concolic Testing Example

```
// Test input a, b, c
void f(int a, int b, int c) {
    if (a == 1) {
        if (b == 2) {
            if (c == 3*a + b) {
                Error();
            }
        }
    }
}
```



- Random testing
  - Probability of reaching `Error()` is extremely low
- Concolic testing generates the following 4 test cases
  - $(0,0,0)$ : initial random input
    - Obtained symbolic path formula (SPF)  $\phi: a \neq 1$
    - Next SPF  $\psi$  generated from  $\phi: !(a \neq 1)$
  - $(1,0,0)$ : a solution of  $\psi$  (i.e.  $!(a \neq 1)$ )
    - SPF  $\phi: a == 1 \&& b \neq 2$
    - Next SPF  $\psi: a == 1 \&& !(b \neq 2)$
  - $(1,2,0)$ 
    - SPF  $\phi: a == 1 \&& (b == 2) \&& (c \neq 3a + b)$
    - Next SPF  $\psi: a == 1 \&& (b == 2) \&& !(c \neq 3a + b)$
  - $(1,2,5)$ 
    - Covered all paths and



# Example

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

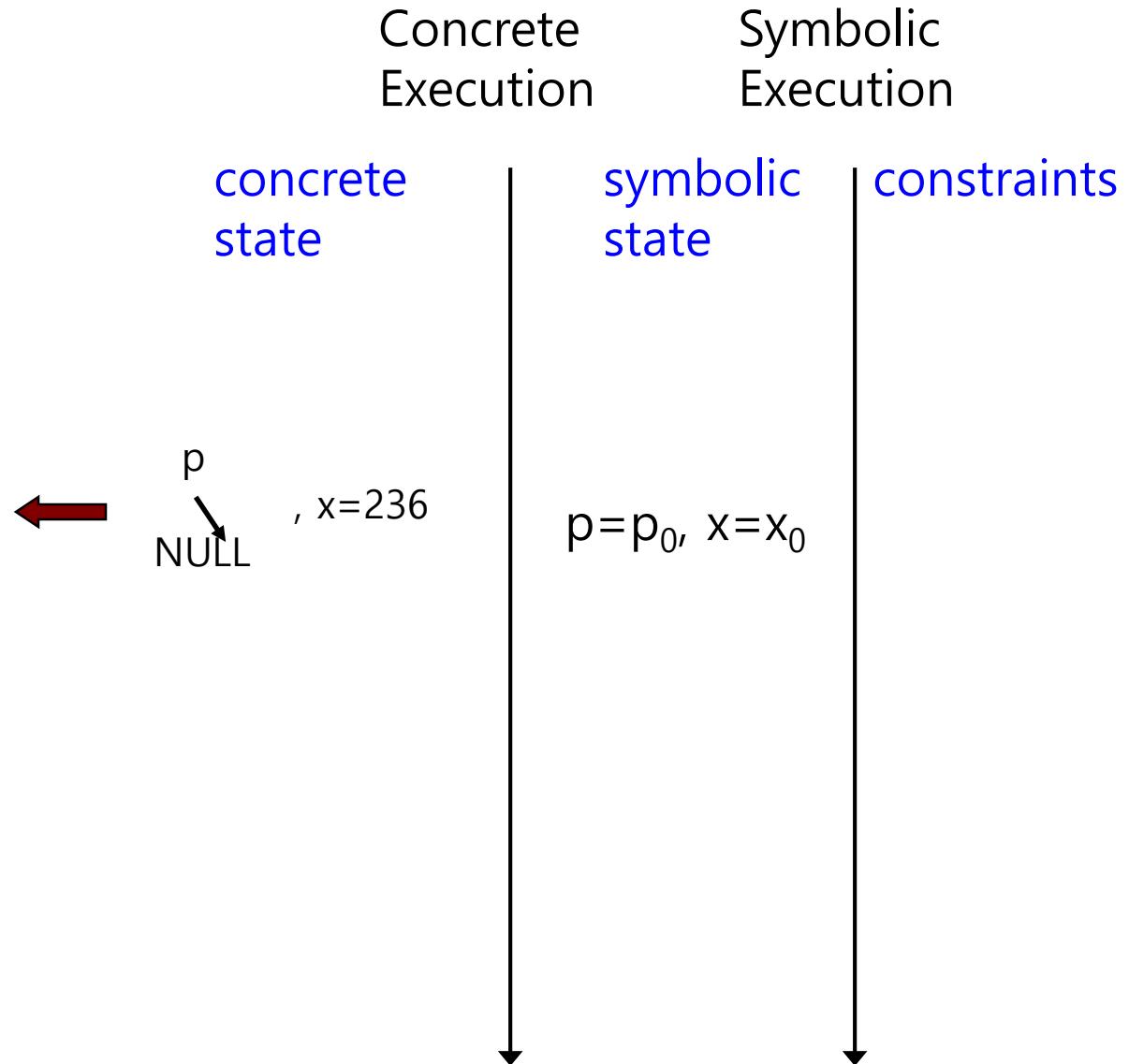
- Random Test Driver:
  - random memory graph reachable from p
  - random value for x
- Probability of reaching **Error()** is extremely low

# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

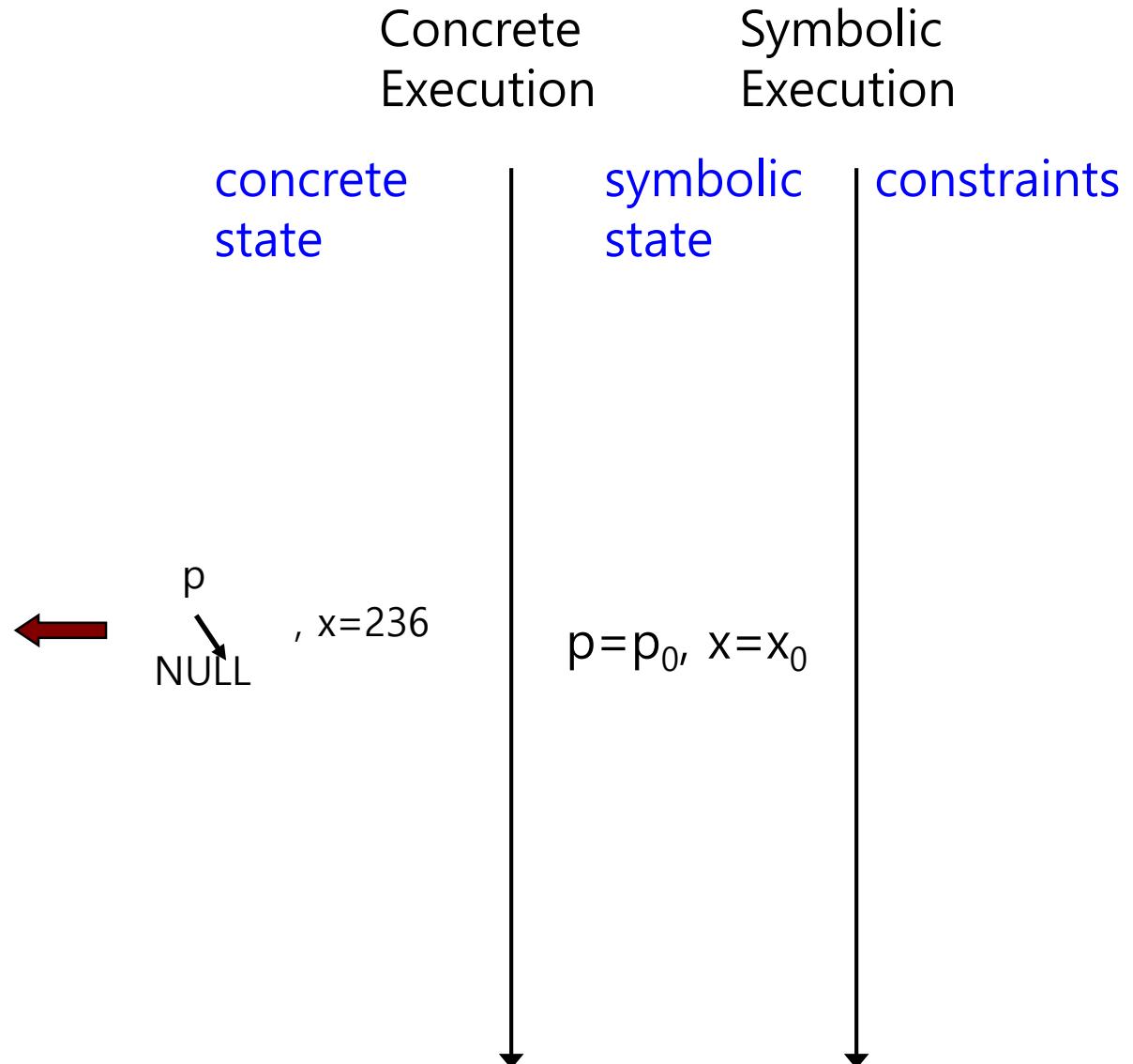


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

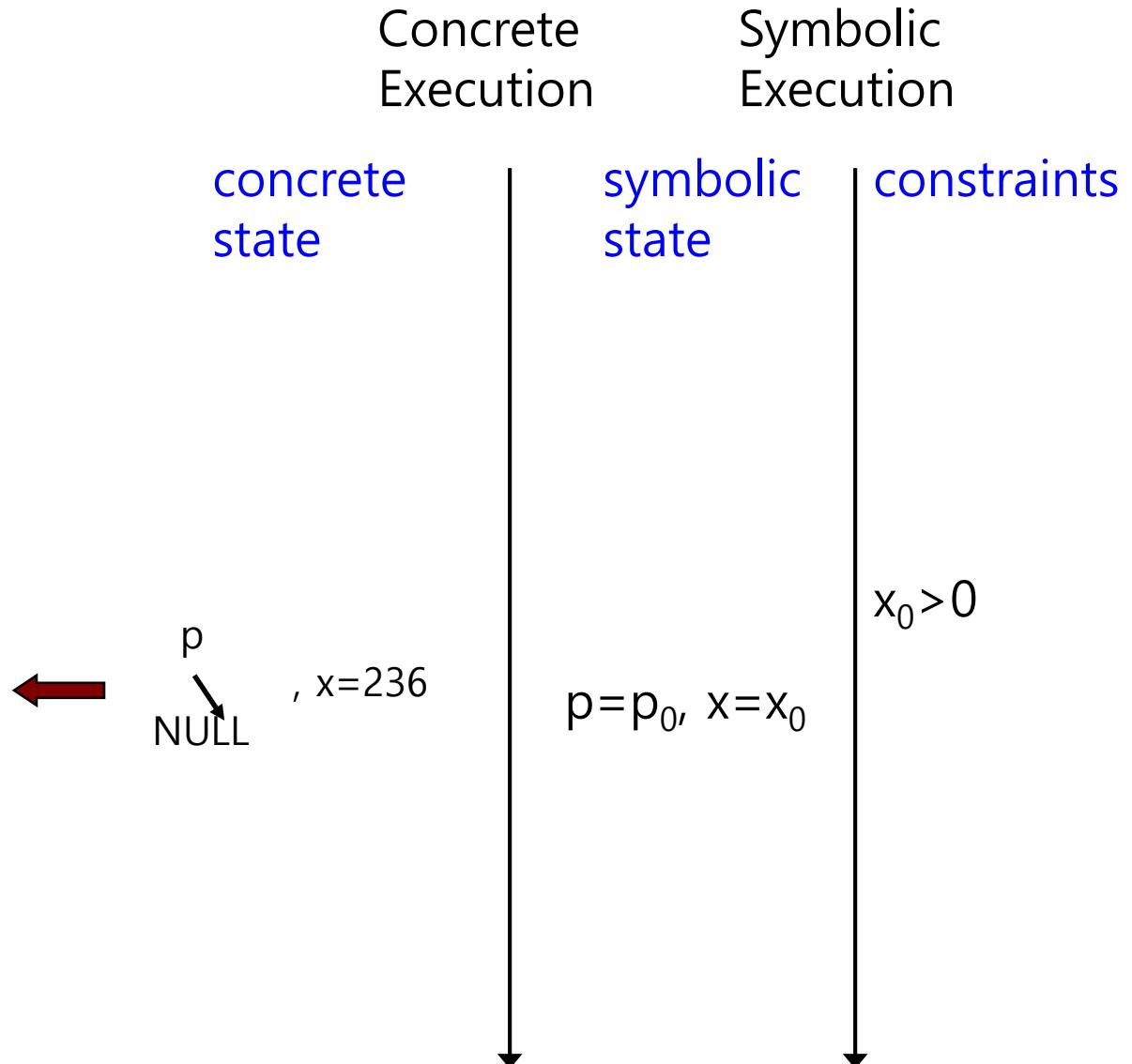


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

Concrete  
Execution

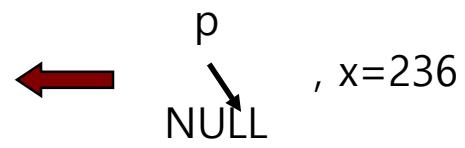
concrete  
state

Symbolic  
Execution

symbolic  
state

constraints

$x_0 > 0$   
 $!(p_0 \neq \text{NULL})$



# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

Concrete  
Execution

Symbolic  
Execution

concrete

symbolic

constraints

solve:  $x_0 > 0$  and  $p_0 \neq \text{NULL}$

$x_0 > 0$   
 $p_0 = \text{NULL}$

$\leftarrow$   
p  
NULL ,  $x=236$

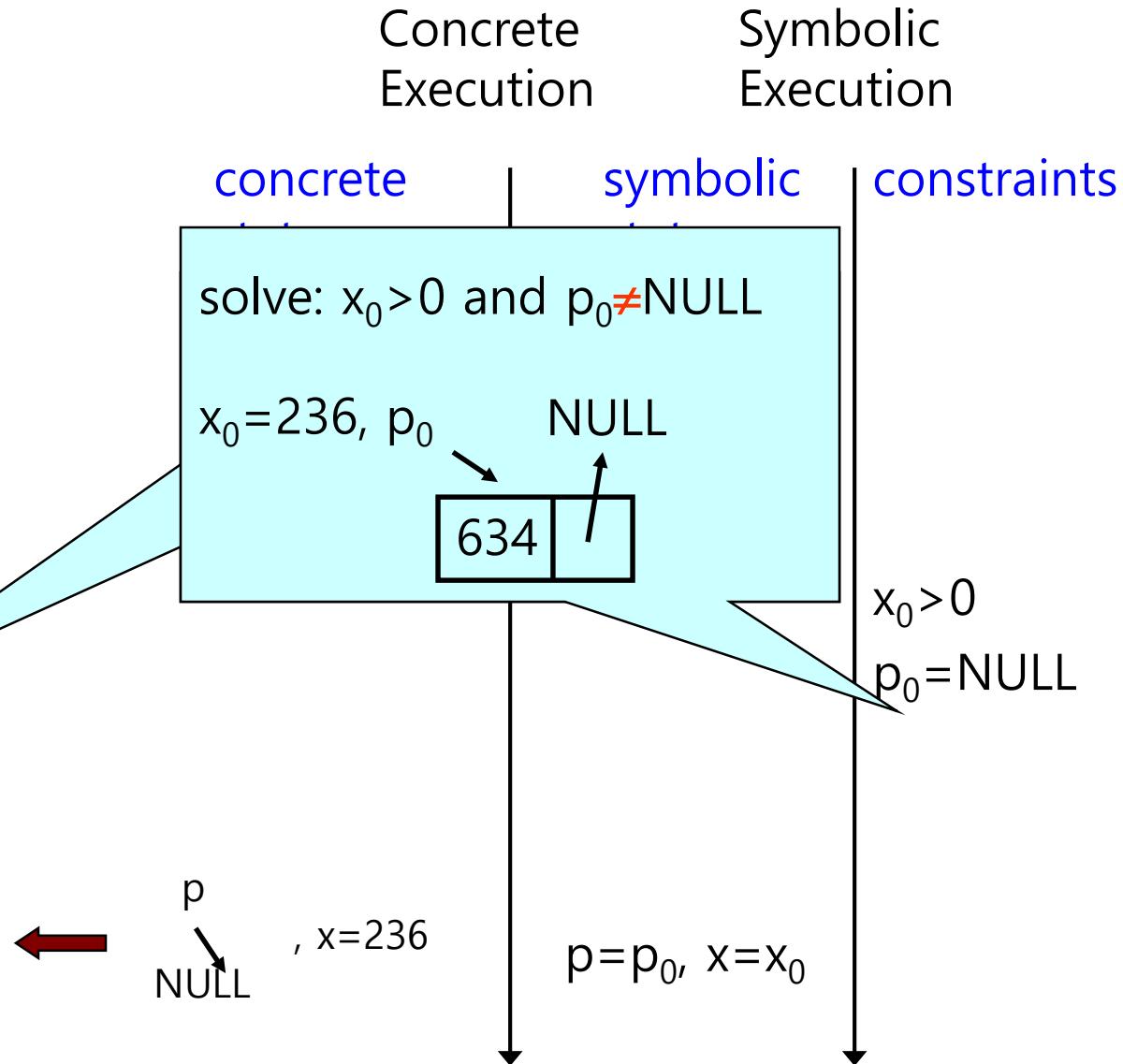
$p=p_0, x=x_0$

# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

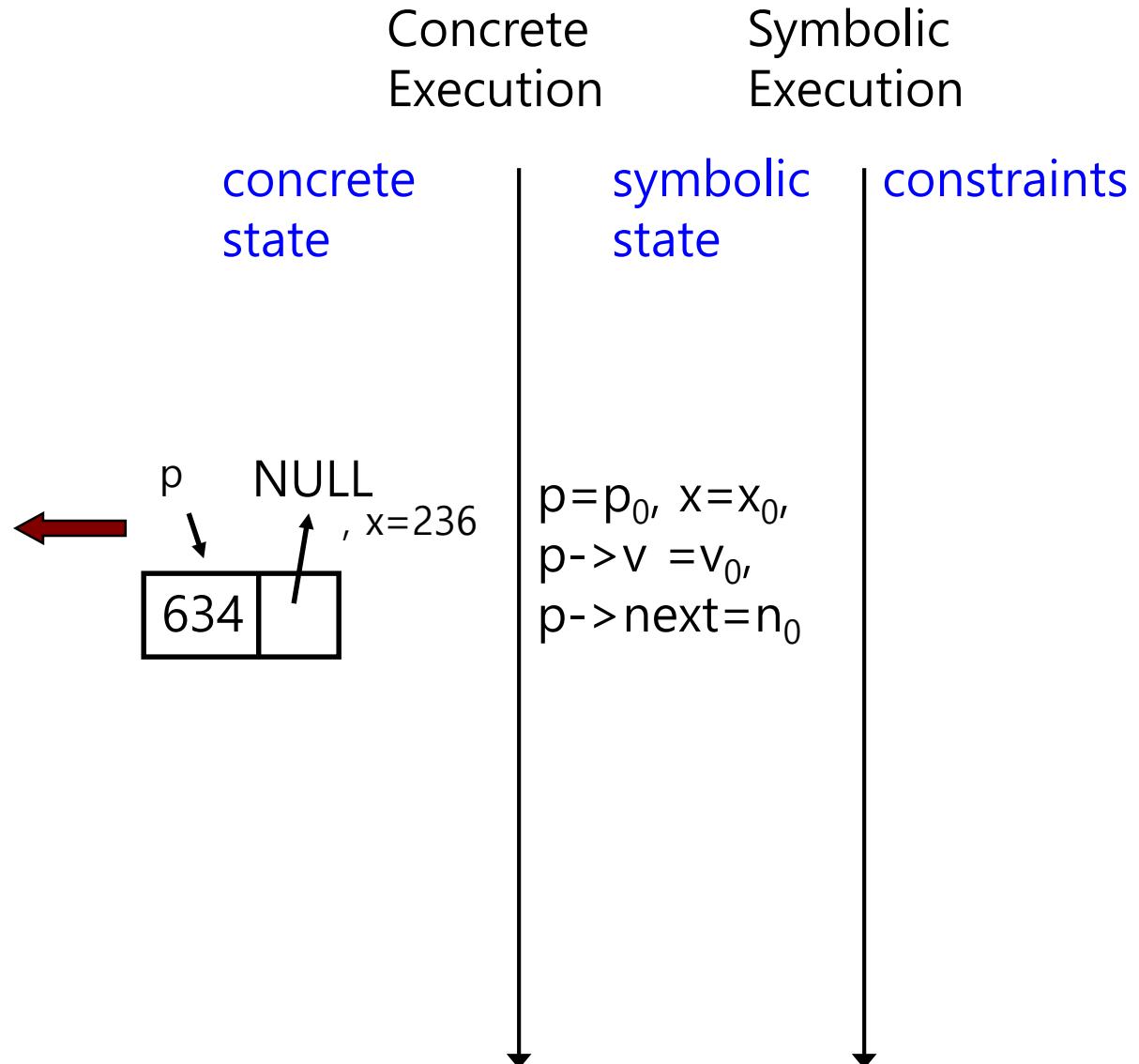


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

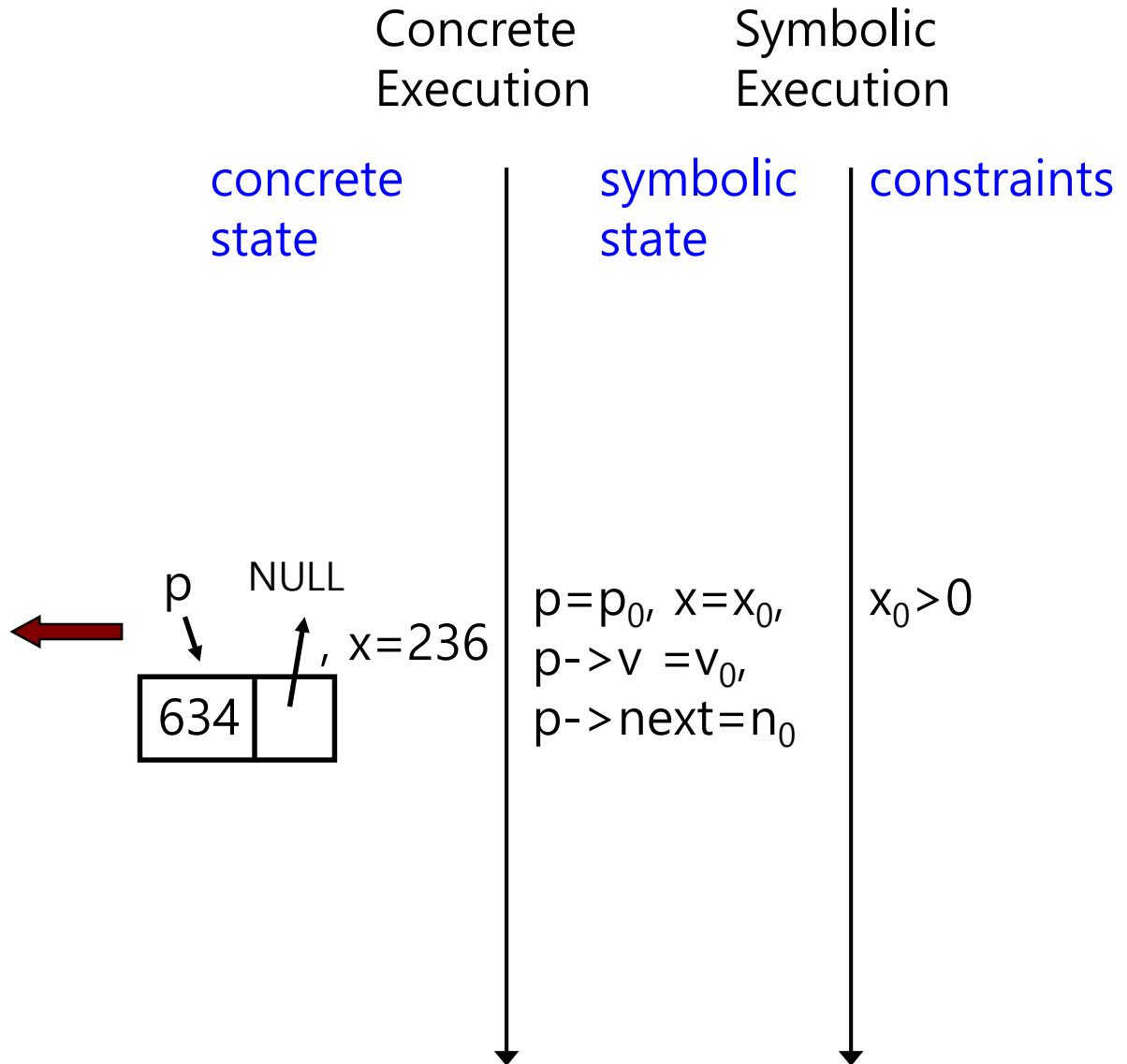


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

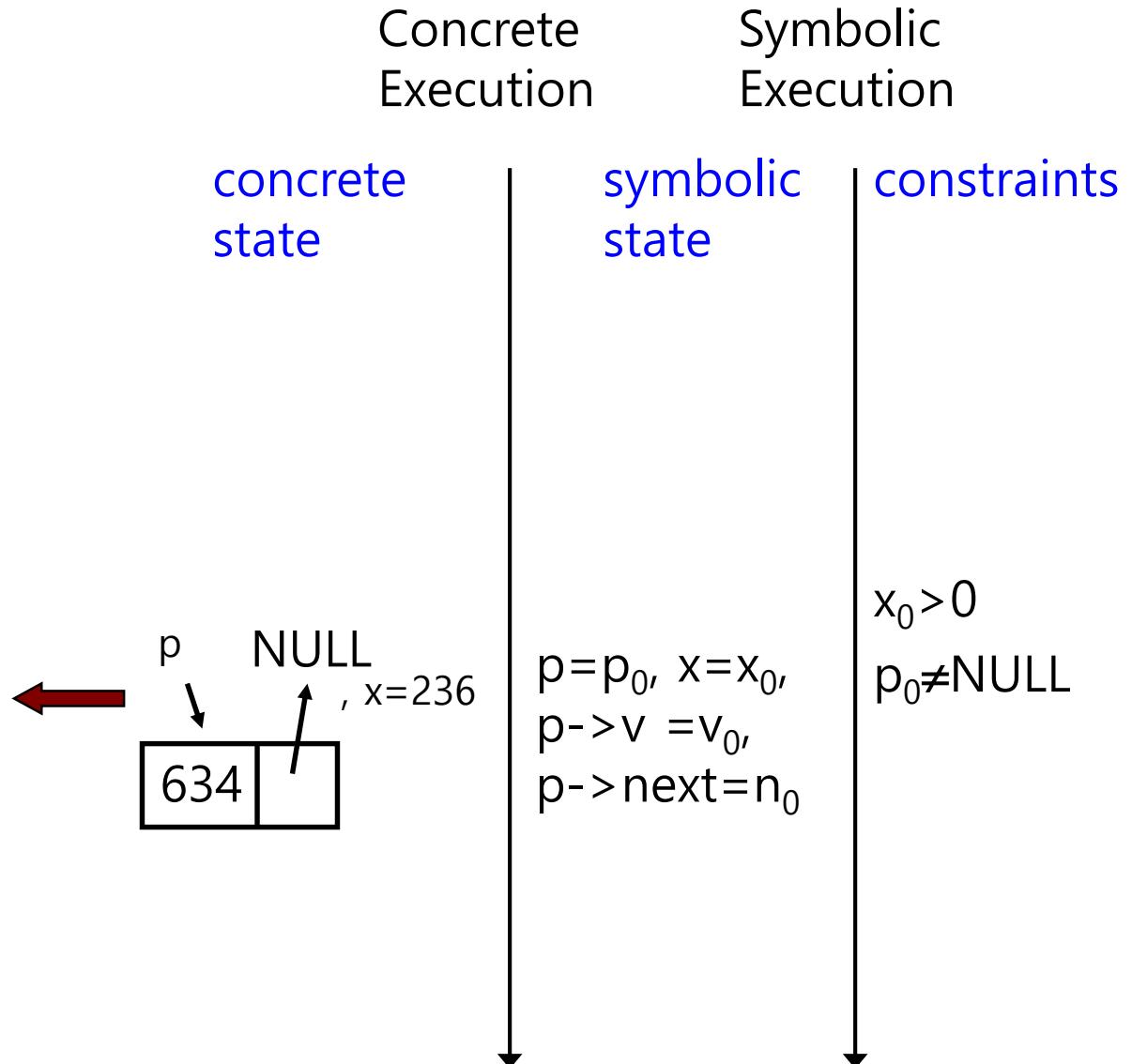


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

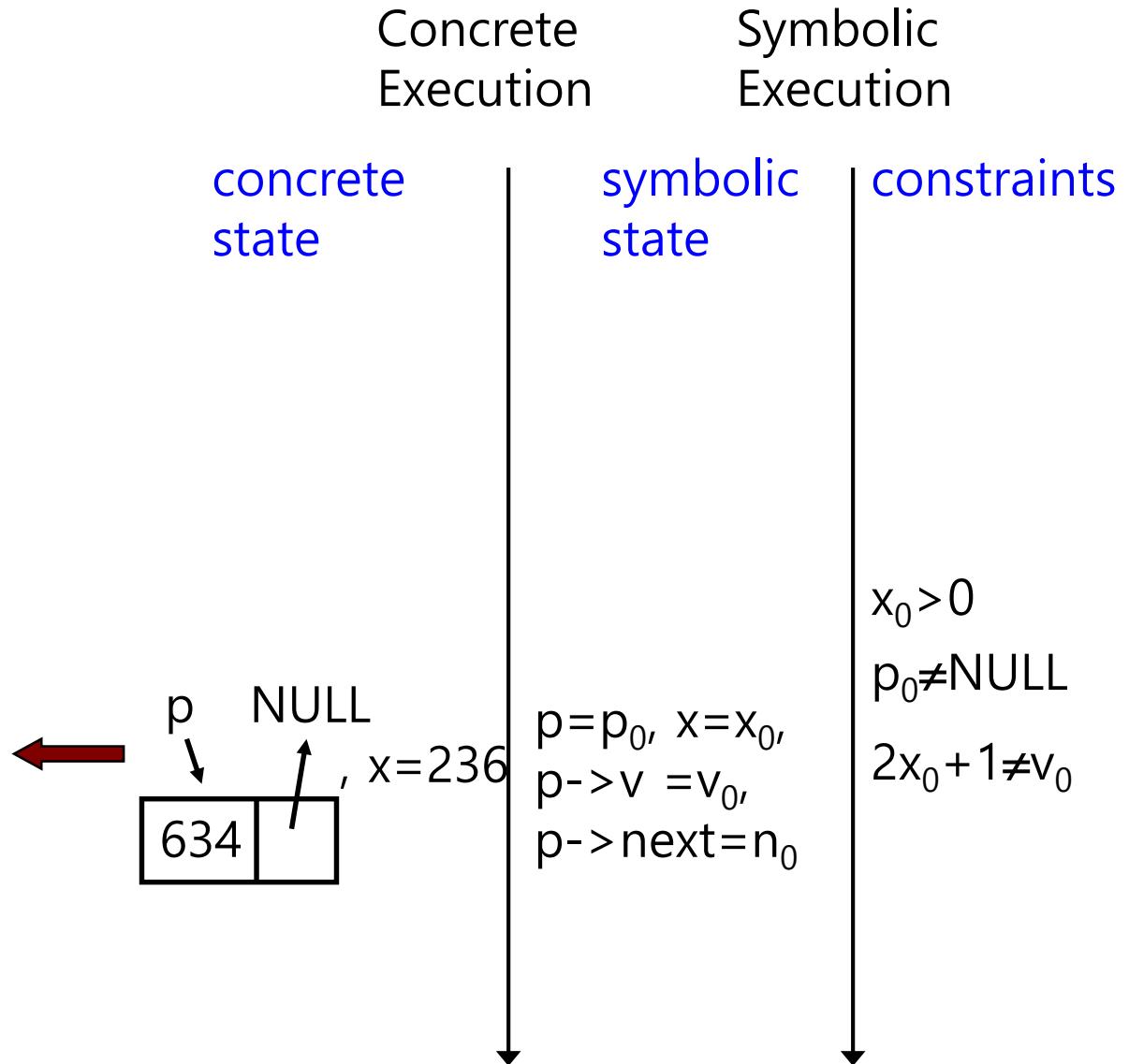


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

Concrete  
Execution

concrete  
state

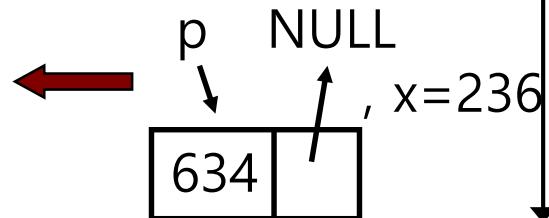
Symbolic  
Execution

symbolic  
state

constraints

$x_0 > 0$   
 $p_0 \neq \text{NULL}$   
 $2x_0 + 1 \neq v_0$

$p = p_0, x = x_0,$   
 $p->v = v_0,$   
 $p->next = n_0$



# Concolic Testing

```
typedef struct cell {
    int v;
    struct cell *next;
} cell;
```

```
int f(int v) {
    return 2*v + 1;
}
```

```
int testme(cell *p, int x) {
    if (x > 0)
        if (p != NULL)
            if (f(x) == p->v)
                if (p->next == p)
                    Error();
    return 0;
}
```

Concrete  
Execution

Symbolic  
Execution

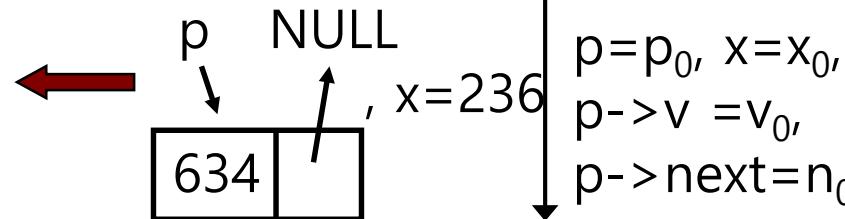
concrete

symbolic

constraints

solve:  $x_0 > 0$  and  $p_0 \neq \text{NULL}$   
and  $2x_0 + 1 = v_0$

$x_0 > 0$   
 $p_0 \neq \text{NULL}$   
 $2x_0 + 1 \neq v_0$

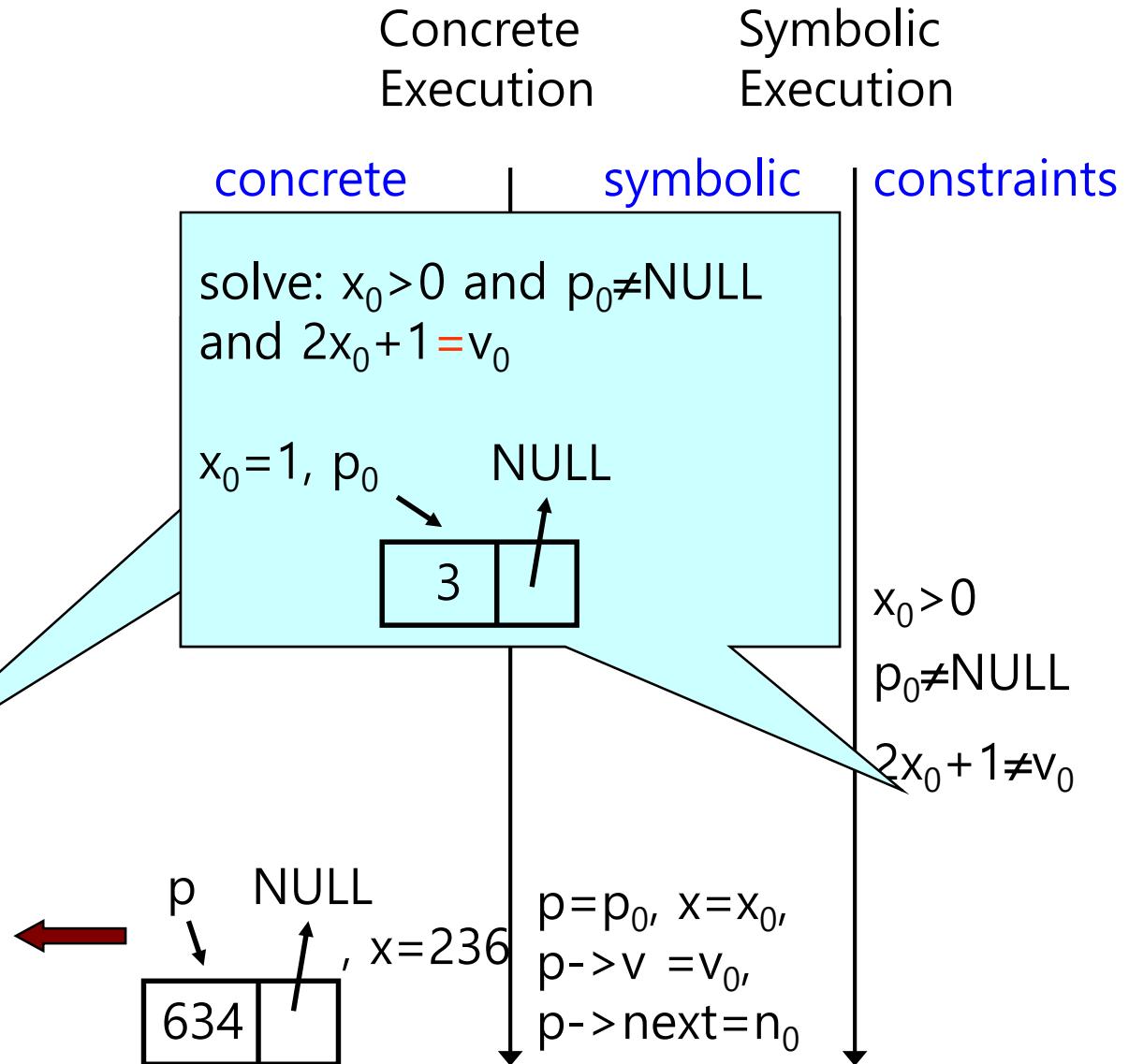


# Concolic Testing

```
typedef struct cell {
    int v;
    struct cell *next;
} cell;
```

```
int f(int v) {
    return 2*v + 1;
}
```

```
int testme(cell *p, int x) {
    if (x > 0)
        if (p != NULL)
            if (f(x) == p->v)
                if (p->next == p)
                    Error();
    return 0;
}
```

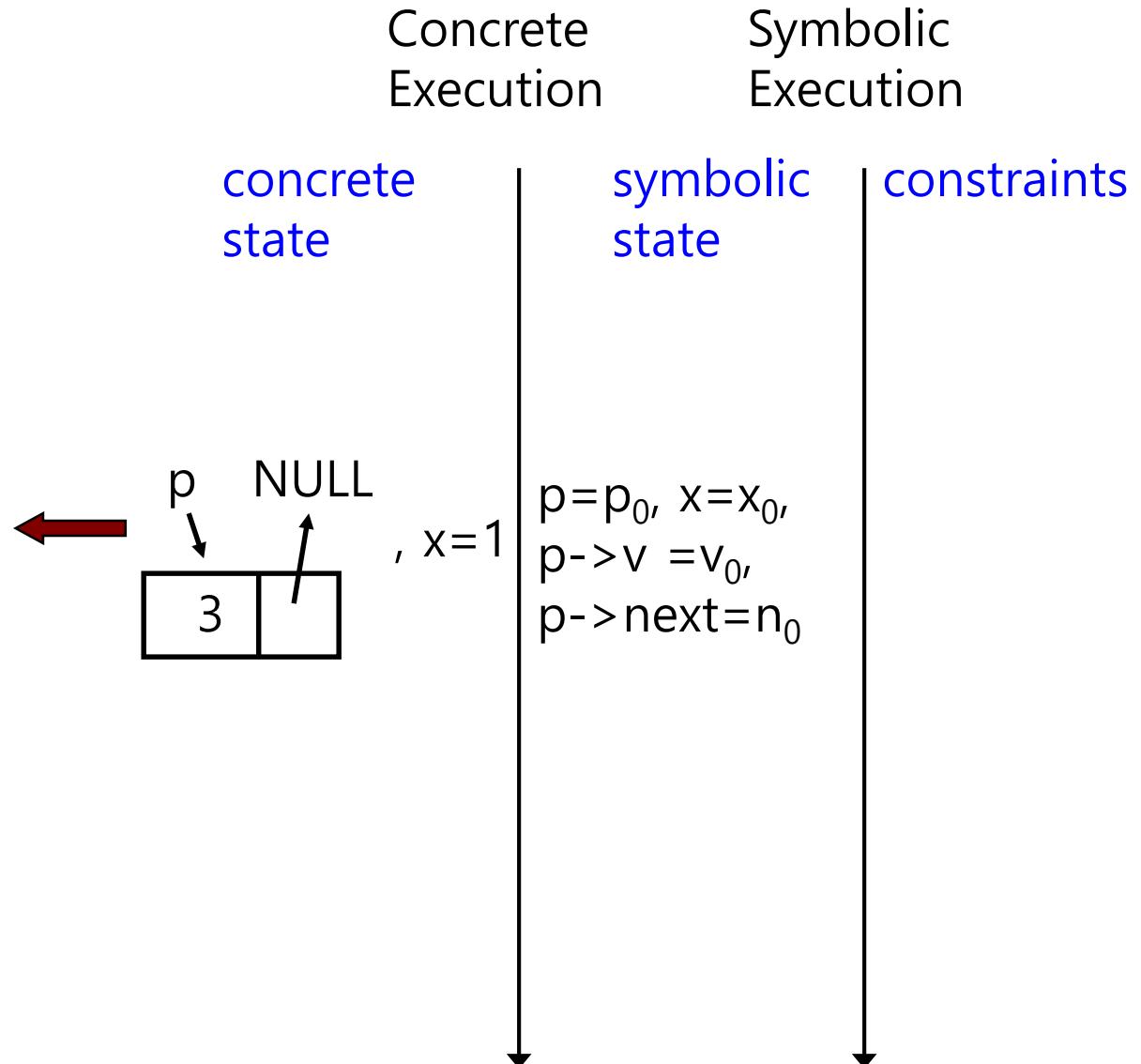


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

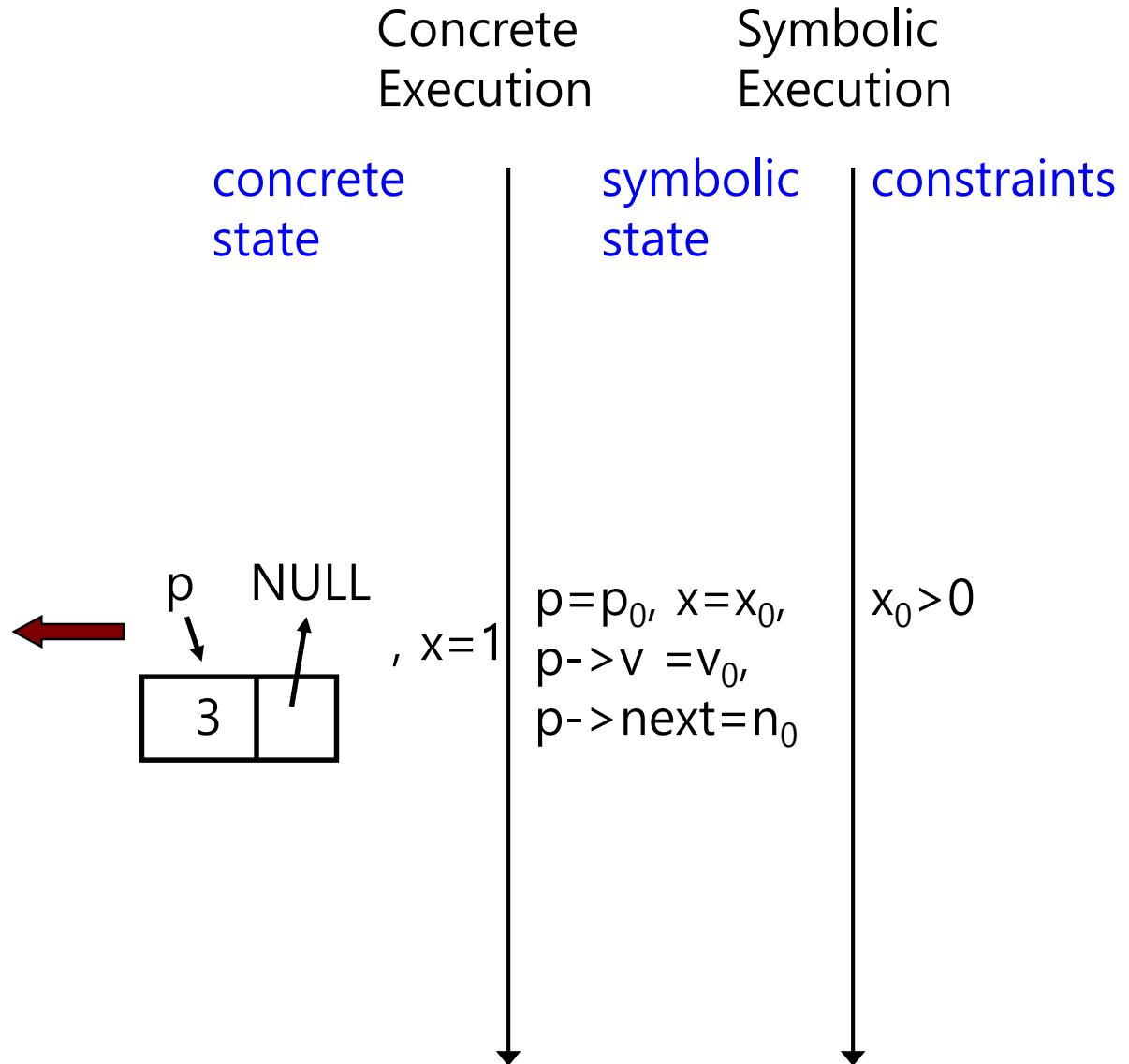


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

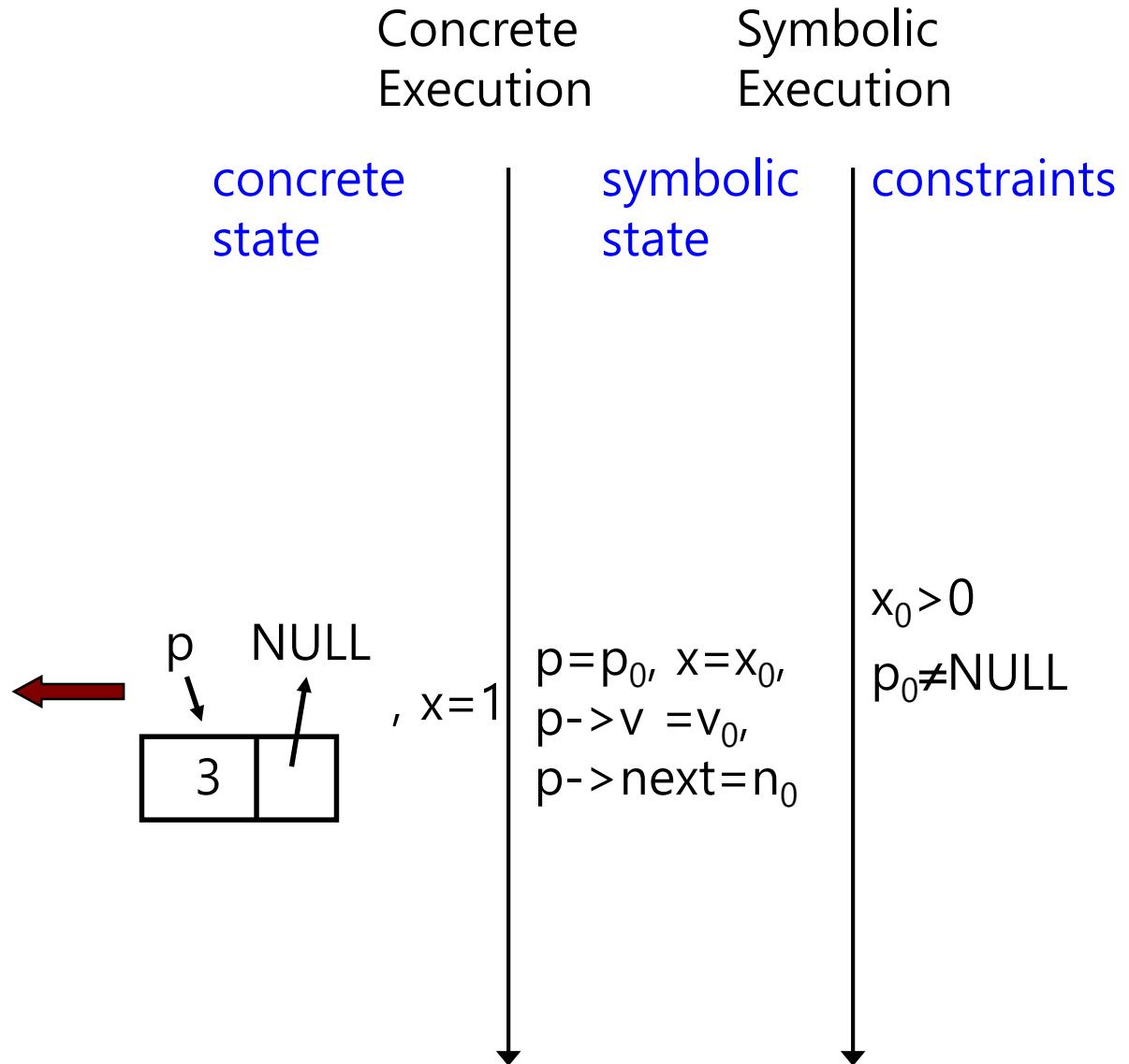


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

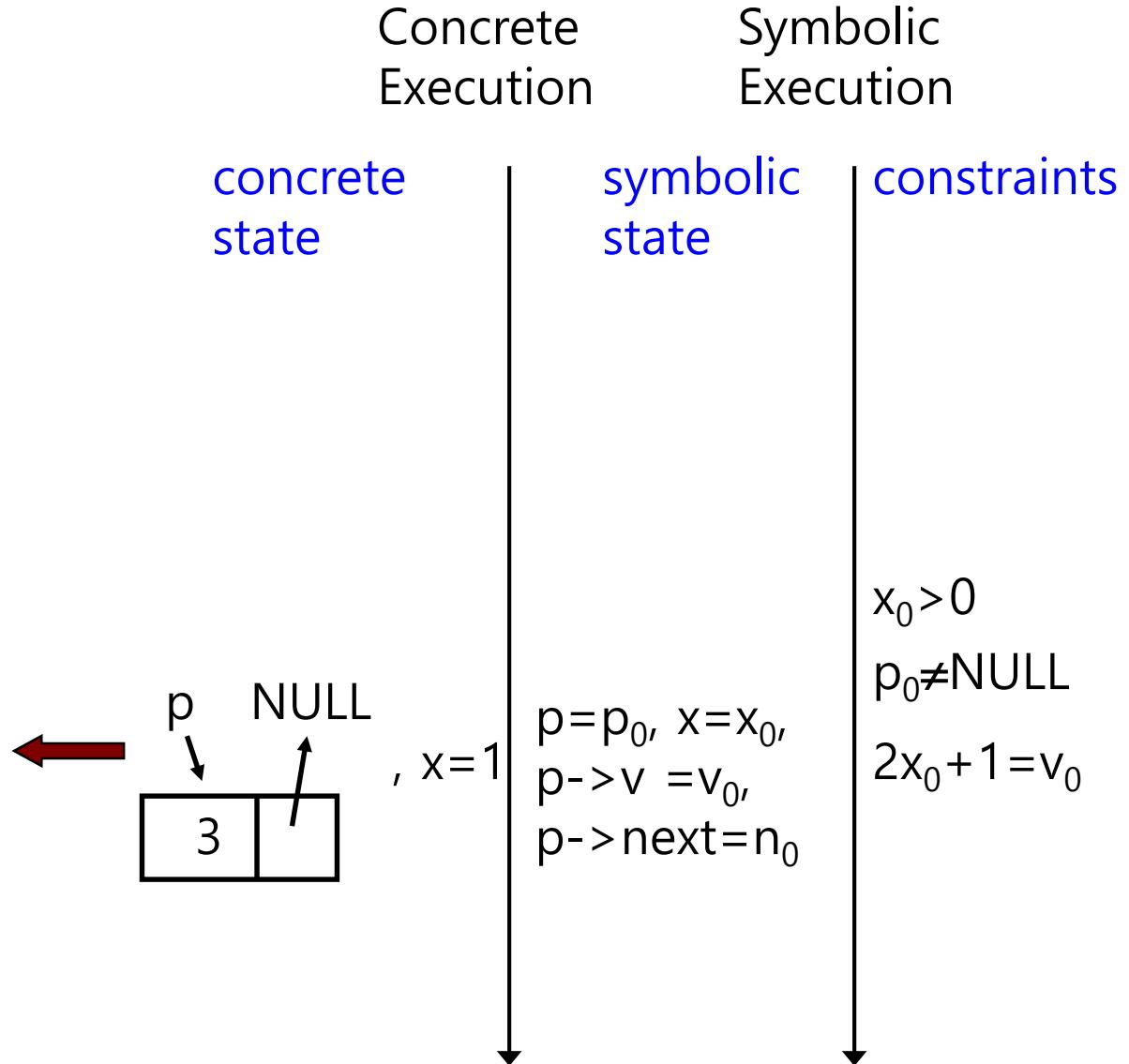


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

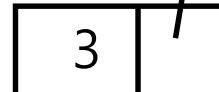


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



Concrete  
Execution

concrete  
state

Symbolic  
Execution

symbolic  
state

constraints

, x=1

$p=p_0, x=x_0,$   
 $p->v=v_0,$   
 $p->next=n_0$

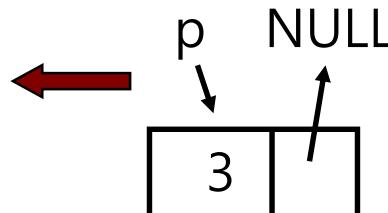
$x_0 > 0$   
 $p_0 \neq \text{NULL}$   
 $2x_0 + 1 = v_0$   
 $n_0 \neq p_0$

# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



Concrete  
Execution

concrete  
state

Symbolic  
Execution

symbolic  
state

constraints

$x_0 > 0$   
 $p_0 \neq \text{NULL}$   
 $2x_0 + 1 = v_0$   
 $n_0 \neq p_0$

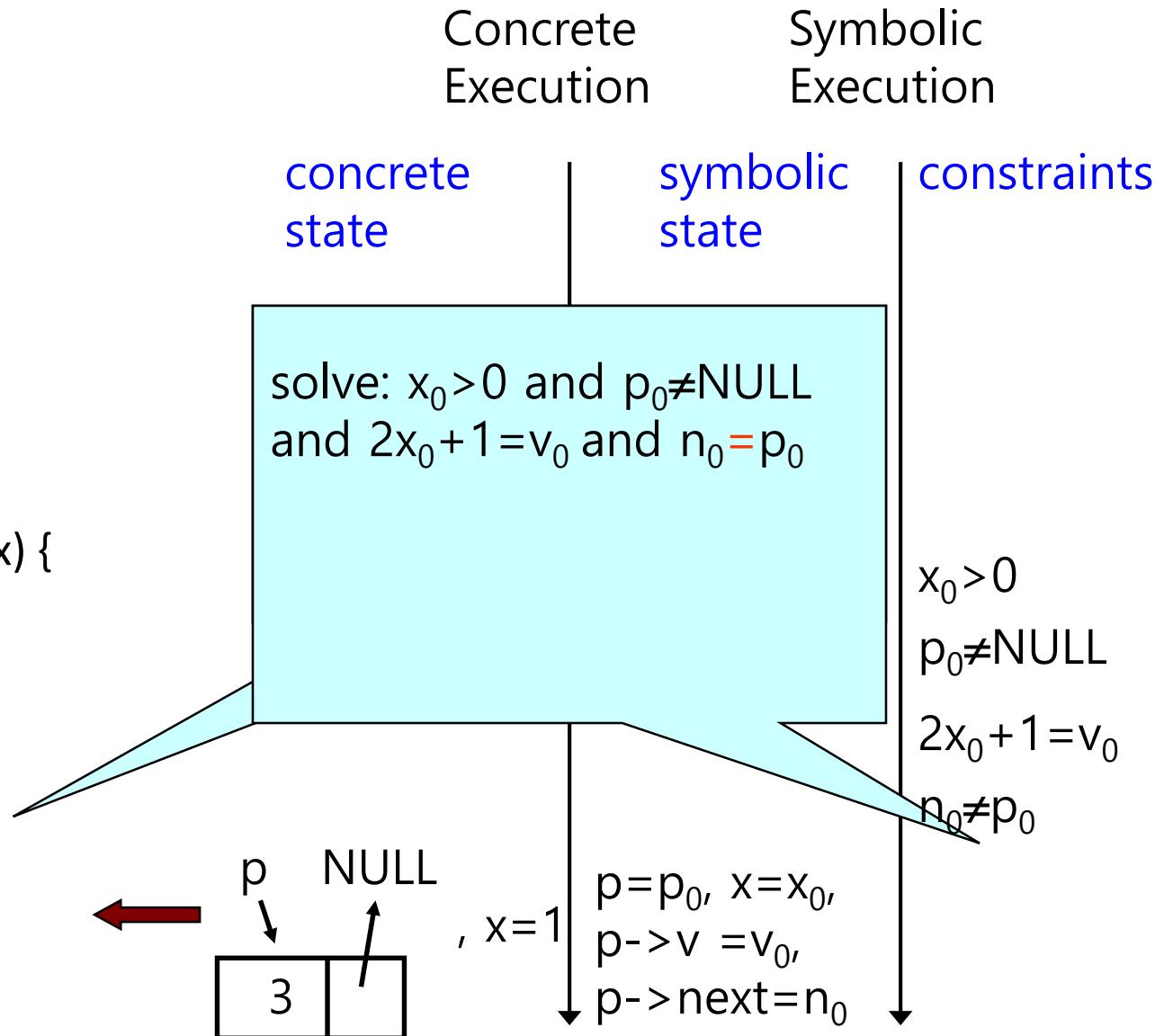
$p = p_0, x = x_0,$   
 $p->v = v_0,$   
 $p->next = n_0$

# Concolic Testing

```
typedef struct cell {
    int v;
    struct cell *next;
} cell;
```

```
int f(int v) {
    return 2*v + 1;
}
```

```
int testme(cell *p, int x) {
    if (x > 0)
        if (p != NULL)
            if (f(x) == p->v)
                if (p->next == p)
                    Error();
    return 0;
}
```

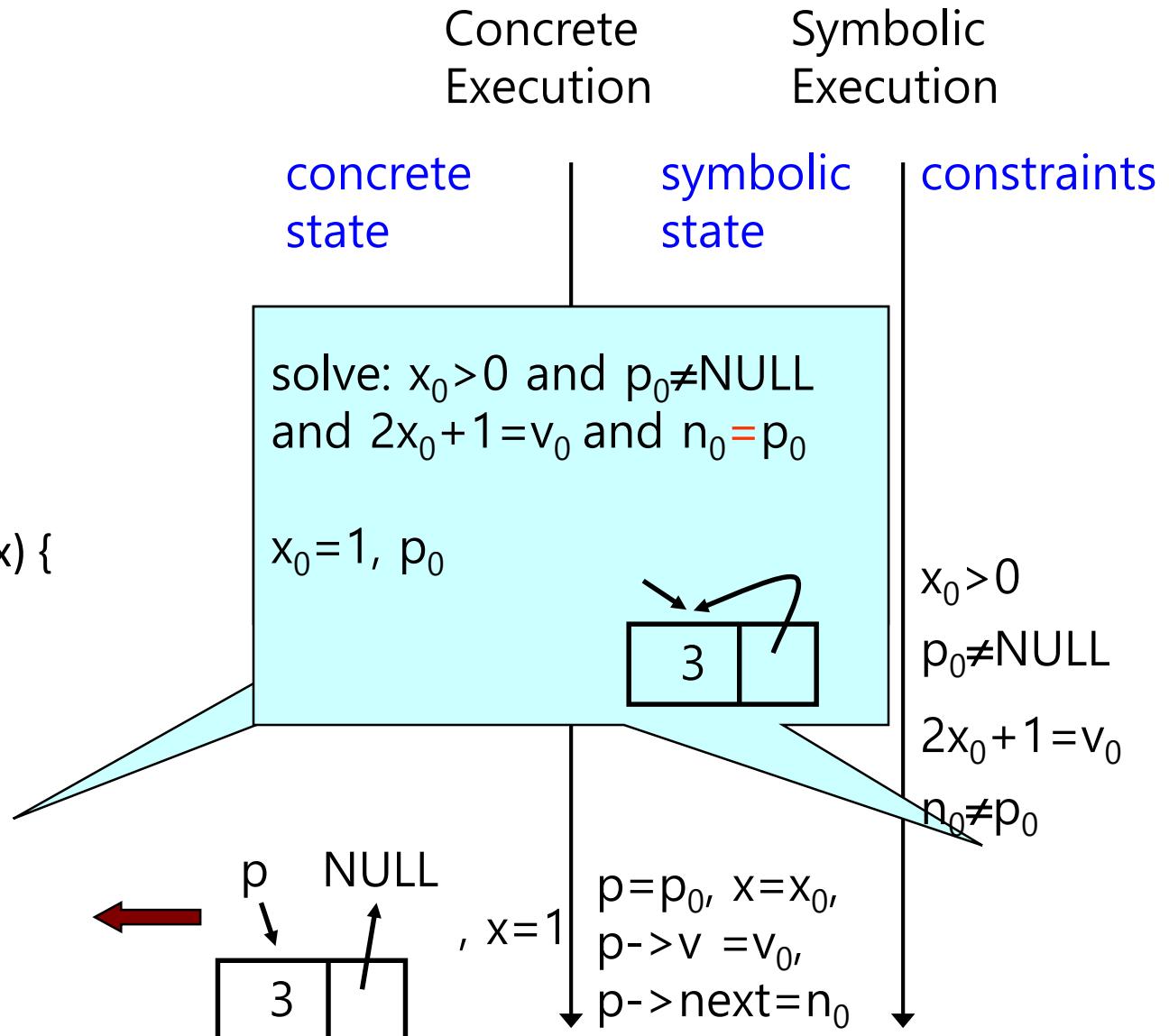


# Concolic Testing

```
typedef struct cell {
    int v;
    struct cell *next;
} cell;
```

```
int f(int v) {
    return 2*v + 1;
}
```

```
int testme(cell *p, int x) {
    if (x > 0)
        if (p != NULL)
            if (f(x) == p->v)
                if (p->next == p)
                    Error();
    return 0;
}
```

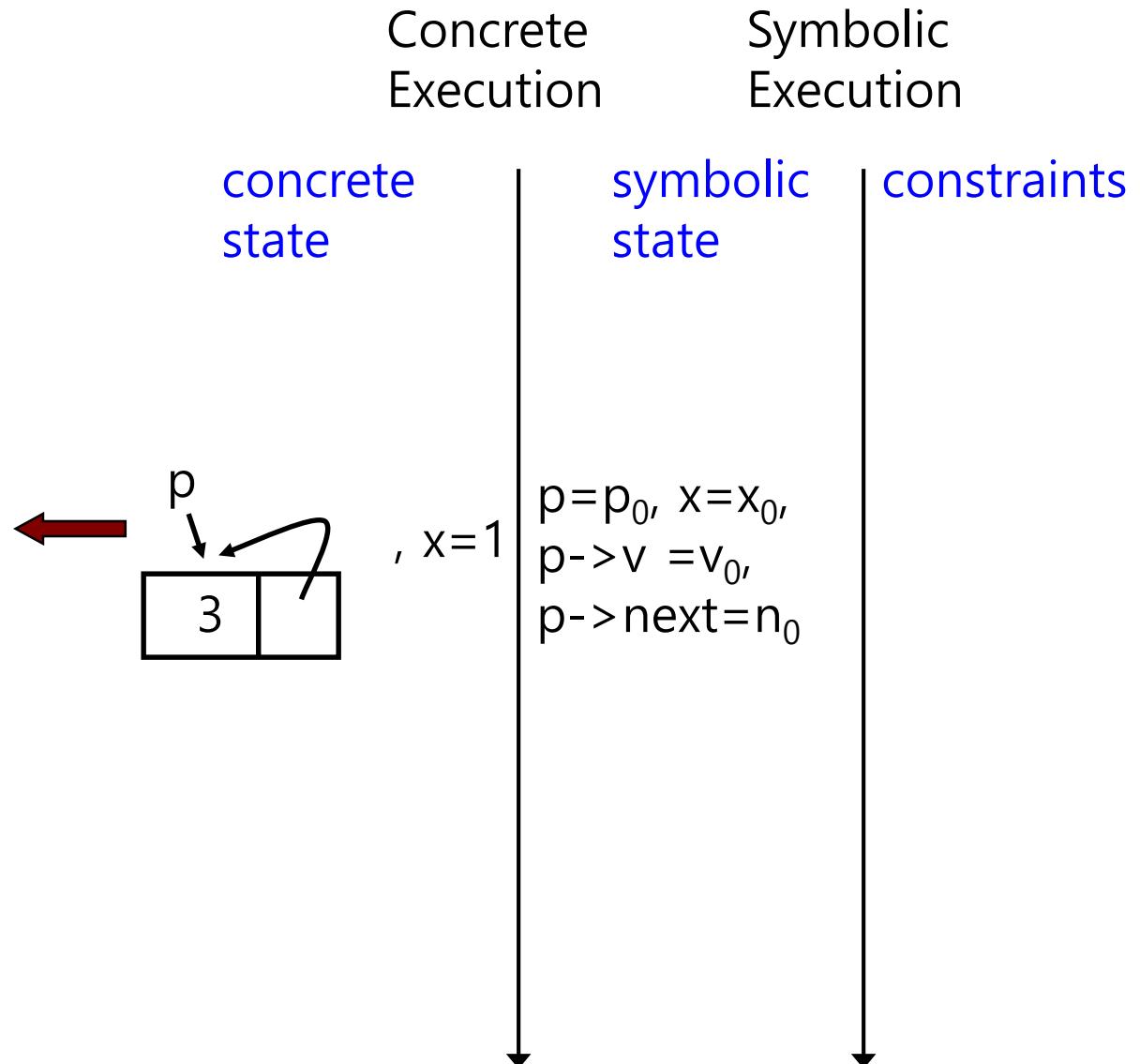


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```

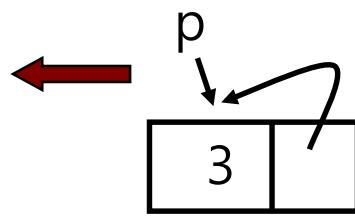


# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



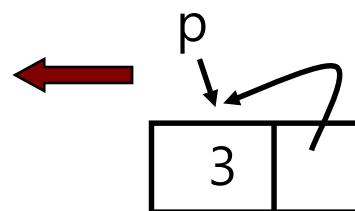
Concrete Execution	Symbolic Execution
concrete state	symbolic state
, $x=1$	$p=p_0, x=x_0,$ $p->v=v_0,$ $p->next=n_0$

# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



,  $x=1$

Concrete  
Execution

concrete  
state

Symbolic  
Execution

symbolic  
state

constraints

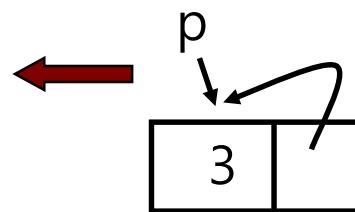
$x_0 > 0$   
 $p_0 \neq \text{NULL}$

# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



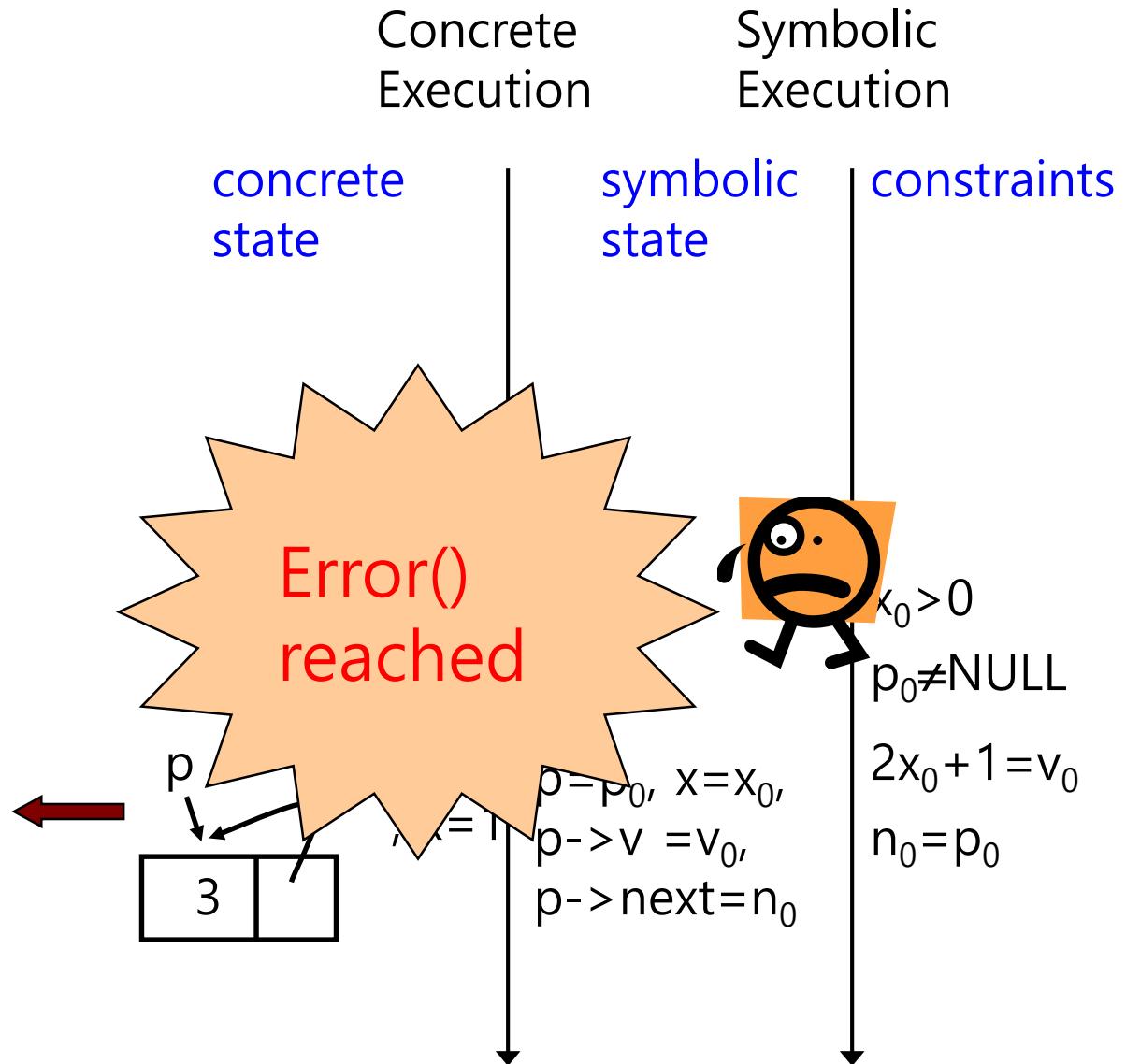
Concrete Execution	Symbolic Execution
concrete state	symbolic state
, $x=1$	$p=p_0, x=x_0,$ $p->v=v_0,$ $p->next=n_0$
	$x_0 > 0$
	$p_0 \neq \text{NULL}$
	$2x_0 + 1 = v_0$

# Concolic Testing

```
typedef struct cell {  
    int v;  
    struct cell *next;  
} cell;
```

```
int f(int v) {  
    return 2*v + 1;  
}
```

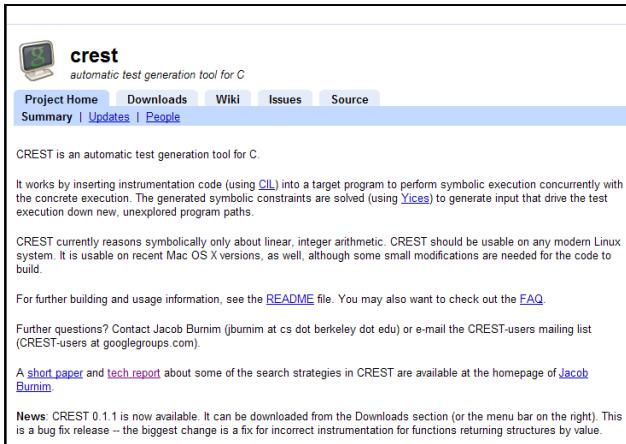
```
int testme(cell *p, int x) {  
    if (x > 0)  
        if (p != NULL)  
            if (f(x) == p->v)  
                if (p->next == p)  
                    Error();  
    return 0;  
}
```



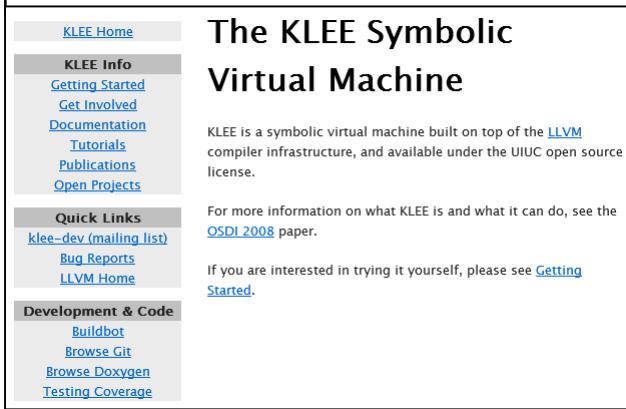
# Summary: Concolic Testing

- Pros
  - Automated test case generation
  - High coverage
  - High applicability (no restriction on target programs)
- Cons
  - If a target program has external binary function calls, coverage might not be complete
    - Ex. if( sin(x) + cos(x) == 0.3 ) { error(); }
  - Current limitation on pointer and array
  - Slow analysis speed due to a large # of TCs

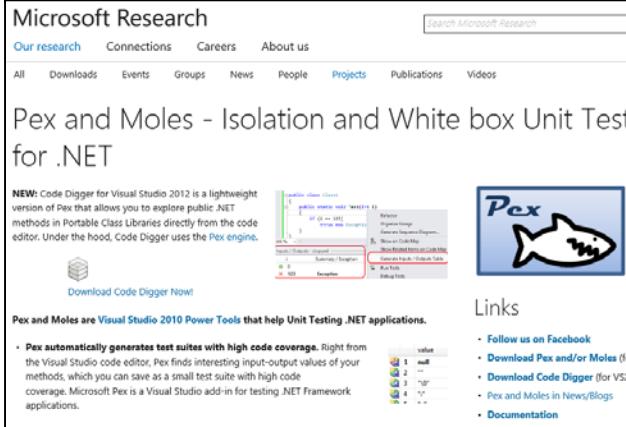
# Concolic Testing Tools



The screenshot shows the CREST project page on SourceForge. It includes a logo, navigation links (Project Home, Downloads, Wiki, Issues, Source), and sections for Summary, Updates, and People. The main content discusses CREST as an automatic test generation tool for C, mentioning its use of instrumentation code (using CIL) and symbolic execution. It also notes that CREST currently reasons symbolically only about linear, integer arithmetic.



The screenshot shows the KLEE Symbolic Virtual Machine page. It features a sidebar with links like KLEE Home, KLEE Info (Getting Started, Get Involved, Documentation, Tutorials, Publications, Open Projects), Quick Links (klee-dev mailing list, Bug Reports, LLVM Home), and Development & Code (Buildbot, Browse Git, Browse Doxygen, Testing Coverage). The main content area is titled "The KLEE Symbolic Virtual Machine" and explains that KLEE is a symbolic virtual machine built on top of the LLVM compiler infrastructure.



The screenshot shows the Microsoft Research Pex and Moles page. It highlights "Pex and Moles - Isolation and White box Unit Test for .NET". A section for "Code Digger" is shown, stating it's a lightweight version of Pex that allows you to explore public .NET methods in Portable Class Libraries directly from the code editor. Below this is a screenshot of the Visual Studio interface showing Pex annotations. The page also mentions "Pex and Moles are Visual Studio 2010 Power Tools that help Unit Testing .NET applications." and provides links for "Follow us on Facebook", "Download Pex and/or Moles (VS2010)", "Download Code Digger (for VS2010)", "Pex and Moles in News/Blogs", and "Documentation".

- **CROWN (open source)**
  - Target: C
  - Instrumentation based extraction
  - BV supported
  - <https://github.com/swtv-kaist/CROWN>
- **KLEE (open source)**
  - Target: LLVM
  - VM based symbolic formula extraction
  - BV supported
  - Symbolic POSIX library supported
  - <http://ccadar.github.io/klee/>
- **PEX (IntelliTest incorporated in Visual Studio 2015)**
  - Target: C#
  - VM based symbolic formula extraction
  - BV supported
  - Integrated with Visual Studio
  - <http://research.microsoft.com/en-us/projects/pex/>
- **CATG (open source)**
  - Target: Java
  - Trace/log based symbolic formula extraction
  - LIA supported

Moonzoo

KAIST

# Case Study: Busybox

- We test a busybox by using CREST.
  - BusyBox is a one-in-all command-line utilities providing a fairly complete programming/debugging environment
  - It combines tiny versions of ~300 UNIX utilities into a single small executable program suite.
  - Among those 300 utilities, we focused to test the following 10 utilities
    - `grep`, `vi`, `cut`, `expr`, `od` , `printf`, `tr`, `cp`, `ls`, `mv`.
    - We selected these 10 utilities, because their behavior is easy to understand so that it is clear what variables should be declared as symbolic
    - Each utility generated 40,000 test cases for 4 different search strategies

# Busybox Testing Result

<u>Utility</u>	<u>LOC</u>	<u># of branches</u>	<u>DFS #of covered branch/time</u>	<u>CFG #of covered branch/time</u>	<u>Random #of covered branch/time</u>	<u>Random input #of covered branch/time</u>	<u>Merge of all 4 strategies #of covered branch/time</u>
grep	914	178	105(59.0%)/278s	85(47.8%)/56s	136(76.4%)/85s	50(28.1%)/45s	136(76.4%)/
vi	4000	1498	855(57.1%)/1495s	965(64.4%)/1036s	1142(76.2%)/723s	1019(68.0%)/463s	1238(82.6%)/
cut	209	112	67(59.8%)/42s	60(53.6%)/45s	84(75.0%)/53s	48(42.9%)/45s	90(80.4%)/
expr	501	154	104(67.5%)/58s	101(65.6%)/44s	105(68.1%)/50s	48(31.2%)/31s	108(70.1%)/
od	222	74	59(79.7%)/35s	72(97.3%)/41s	66(89.2%)/42s	44(59.5%)/30s	72(97.3%)/
printf	406	144	93(64.6%)/84s	109(75.7%)/41s	102(70.8%)/40s	77(53.5%)/30s	115(79.9%)/
tr	328	140	67(47.9%)/58s	72(51.4%)/50s	72(51.4%)/50s	63(45%)/42s	73(52.1%)/
cp	191	32	20(62.5%)/38s	20(62.5%)/38s	20(62.5%)/38s	17(53.1%)/30s	20(62.5%)/
ls	1123	270	179(71.6%)/87s	162(64.8%)/111s	191(76.4%)/86s	131(52.4%)/105s	191(76.4%)/
mv	135	56	24(42.9%)/0s	24(42.9%)/0s	24(42.9%)/0s	17(30.3%)/0s	24(47.9%)/
<b>AVG</b>	<b>803</b>	<b>264</b>	<b>157.3(59.6%)/809s</b>	<b>167(63.3%)/146s</b>	<b>194.2(73.5%)/117s</b>	<b>151.4(57.4%)/83s</b>	<b>206.7(78.4%)/1155s</b>

# Result of grep

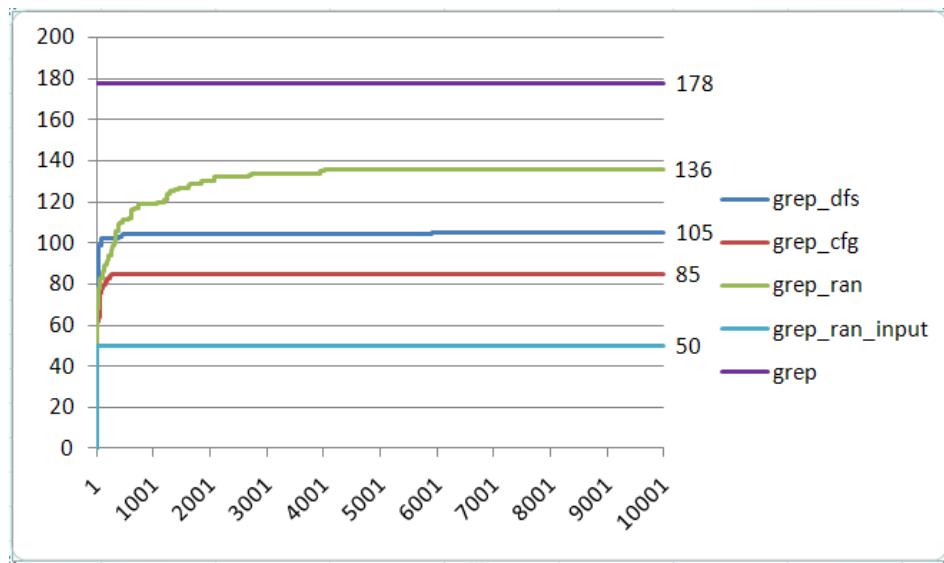
Experiment 1:

Iterations: 10, 000

branches in grep.c : 178

Execution Command:

```
run_crest './busybox grep "define" test_grep.dat' 10000 -dfs  
run_crest './busybox grep "define" test_grep.dat' 10000 -cfg  
run_crest './busybox grep "define" test_grep.dat' 10000 -random  
run_crest './busybox grep "define" test_grep.dat' 10000 -random_input
```



Strategy	Time cost (s)
Dfs	2758
Cfg	56
Random	85
Pure_random	45

# Test Oracles

- In the busybox testing, we do not use any explicit test oracles
  - Test oracle is an orthogonal issue to test case generation
  - However, still violation of runtime conformance (i.e., no segmentation fault, no divide-by-zero, etc) can be checked
- Segmentation fault due to integer overflow detected at grep 2.0
  - This bug was detected by test cases generated using DFS
  - The bug causes segmentation fault when
    - -B 1073741824 (i.e.  $2^{32}/4$ )
    - PATTERN should match line(s) after the 1<sup>st</sup> line
    - Text file should contain at least two lines
  - Bug scenario
    - Grep tries to dynamically allocate memory for buffering matched lines (-B option).
    - But due to integer overflow (# of line to buffer \* sizeof(pointer)), memory is allocated in much less amount
    - Finally grep finally accesses illegal memory area

## Bug 2653 - busybox grep with option -B can cause segmentation fault

<u>Status</u> : RESOLVED FIXED	<u>Reported</u> : 2010-10-02 06:35 UTC by Yunho Kim
<u>Product</u> : Busybox	<u>Modified</u> : 2010-10-03 21:50 UTC ( <a href="#">History</a> )
<u>Component</u> : Other	<u>CC List</u> : 1 user ( <a href="#">show</a> )
<u>Version</u> : 1.17.x	
<u>Platform</u> : PC Linux	
<u>Importance</u> : P5 major	<u>Host</u> :
<u>Target Milestone</u> : ---	<u>Target</u> :
<u>Assigned To</u> : unassigned	<u>Build</u> :
<u>URL</u> :	
<u>Keywords</u> :	
<u>Depends on</u> :	
<u>Blocks</u> :	
	Show dependency tree / <a href="#">graph</a>

### Attachments

[Add an attachment](#) (proposed patch, testcase, etc.)

### Note

You need to [log in](#) before you can comment on or make changes to this bug.

Yunho Kim 2010-10-02 06:35:09 UTC

I report an integer overflow bug in a busybox grep applet, which causes an memory corruption.

```
**** findutils/grep.c ****
634     if (option_mask32 & OPT_C) {
635         /* -C unsets prev -A and -B, but following -A or -B
636          may override it */
637         if (!(option_mask32 & OPT_A)) /* not overridden */
638             lines_after = Copt;
639         if (!(option_mask32 & OPT_B)) /* not overridden */
640             lines_before = Copt;
```

- Bug patch was immediately made in 1 day, since this bug is critical one
  - Importance: P5 major
    - major loss of function
  - Busybox 1.18.x will have fix for this bug

# SAGE: Whitebox Fuzzing for Security Testing @ Microsoft

- X86 binary concolic testing tool to generate millions of test files automatically targeting large applications
  - used daily in Windows, Office, etc.
- Mainly targets crash bugs in various windows file parsers (>hundreds)
- Impact: since 2007 to 2013
  - 500+ machine years
  - 3.4 Billion+ constraints
  - 100s of apps, 100s of bugs
    - 1/3 of all security bugs detected by Win7 WEX team were found by SAGE
  - Millions of dollars saved

*This slide quotes PLDI 2013 MSR Open House Event poster  
“SAGE: WhiteboxFuzzing for Security Testing” by  
E.Bounimova, P.Godefroid, and D.Molnar*

# Microsoft Project Springfield

- Azure-based cloud service to find security bugs in x86 windows binary
- Based on concolic testing techniques of SAGE



The landing page for Project Springfield features a blue header with the text "Project Springfield" and "Fuzz your code before hackers do". A "Sign up" button is located in the top-left corner. To the right is a large graphic of a shield with a lock, surrounded by arrows and a gear, symbolizing security and protection. Below the header, the text "What is Project Springfield?" is displayed, followed by a description: "Project Springfield is Microsoft's unique fuzz testing service for finding security critical bugs in software. Project Springfield helps customers quickly adopt practices and technology battle-tested over the last 15 years at Microsoft." The page then highlights four key features with icons: a dollar sign inside wavy lines for "Million Dollar" Bugs, a binary code grid for Battle tested tech, a cluster of squares for Fast, Consistent Roll-out, and two people with a magnifying glass for Available now.

What is Project Springfield?

Project Springfield is Microsoft's unique fuzz testing service for finding security critical bugs in software. Project Springfield helps customers quickly adopt practices and technology battle-tested over the last 15 years at Microsoft.

"Million Dollar" Bugs

Project Springfield uses "Whitebox Fuzzing" technology which discovered 1/3rd of the "million dollar" security bugs during Windows 7 development.

Battle tested tech

The same state-of-the-art tools and practices used inside Microsoft to fuzz Windows and Office applications.

Fast, Consistent Roll-out

Project Springfield provides the platform to ensure systematic security risk assessment and testing consistency.

Available now

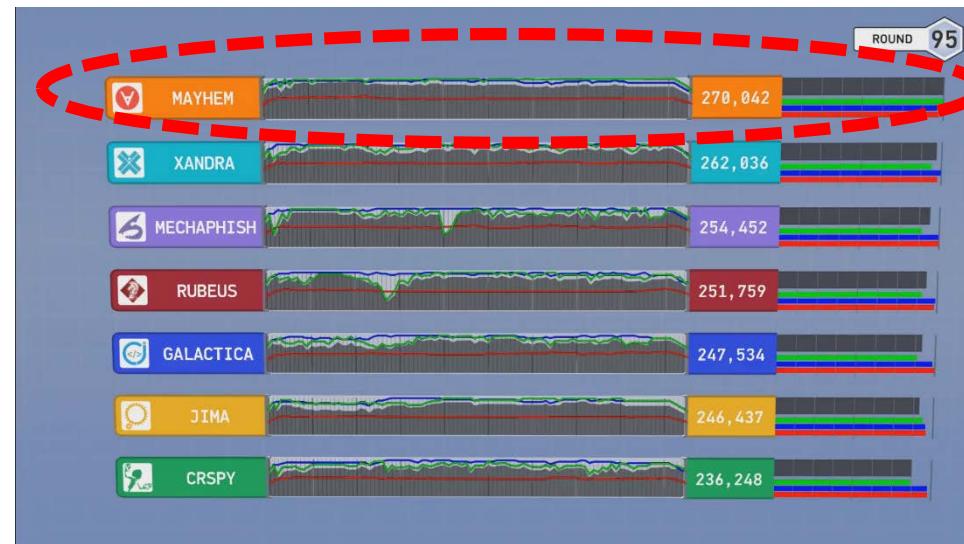
Enterprise customers are currently using Project Springfield to find and fix critical security issues.

[See examples >](#)

# 2016 Aug DARPA Cyber Grand Challenge

-the world's 1<sup>st</sup> all-machine hacking tournament

- Each team's Cyber Reasoning System **automatically** identifies security flaws and applies patches to its own system in a hack-and-defend style contest targeting a new Linux-based OS DECREE
- Mayhem won the best score, which is CMU's concolic testing based tool



# 현대모비스, AI 기반 소프트웨어 검증시스템 도입..."효율 2배로"

2018-07-22 10:00

댓글 f TALK ...

가- 가+

'마이스트' 적용...대화형 검색 로봇 '마이봇'도 도입

(서울=연합뉴스) 윤보람 기자 = 현대모비스[012330]가 인공지능(AI)을 활용해 자율주행, 커넥티비티(연결성) 등 미래 자동차 소프트웨어(SW) 개발에 속도를 낸다.

현대모비스는 AI를 기반으로 하는 소프트웨어 검증시스템 '마이스트'(MAIST: Mobis Artificial Intelligence Software Testing)를 최근 도입했다고 22일 밝혔다.

Hyundai Mobis and a research team lead by Prof. Moonzoo Kim at KAIST jointly developed MAIST for automated testing

MAIST automates unit coverage testing performed by human engineers by applying concolic unit testing

Google에 의해 종료된 광고입니다.

이 광고 그만 보기

이 광고가 표시된 이유 ⓘ

실제 현대모비스가 통합형 차체제어시스템(IBU)과 써라

MAIST can reduce 53% and 70% of manual testing effort for IBU(Integrated Body Unit) and SVM(Surround View Monitoring)

현대모비스는 하반기부터 소프트웨어가 탑재되는 제동, 조향 등 모든 전장부품으로 마이스트를 확대 적용할 계획이다. 글로벌 소프트웨어 연구기지인 인도연구소에도 적용한다.

## ■ 현대모비스 인공지능 도입 사례

AI 시스템	목적	도입 효과
MAIST	소프트웨어 검증 자동화	IBU 53%
마이봇 (Mobis AI Robot)	소프트웨어 개발문서 검색	SVM 70%

<http://m.yna.co.kr/kr/contents/?cid=AKR2018072015880003&mobile>

Moonzoo Kim

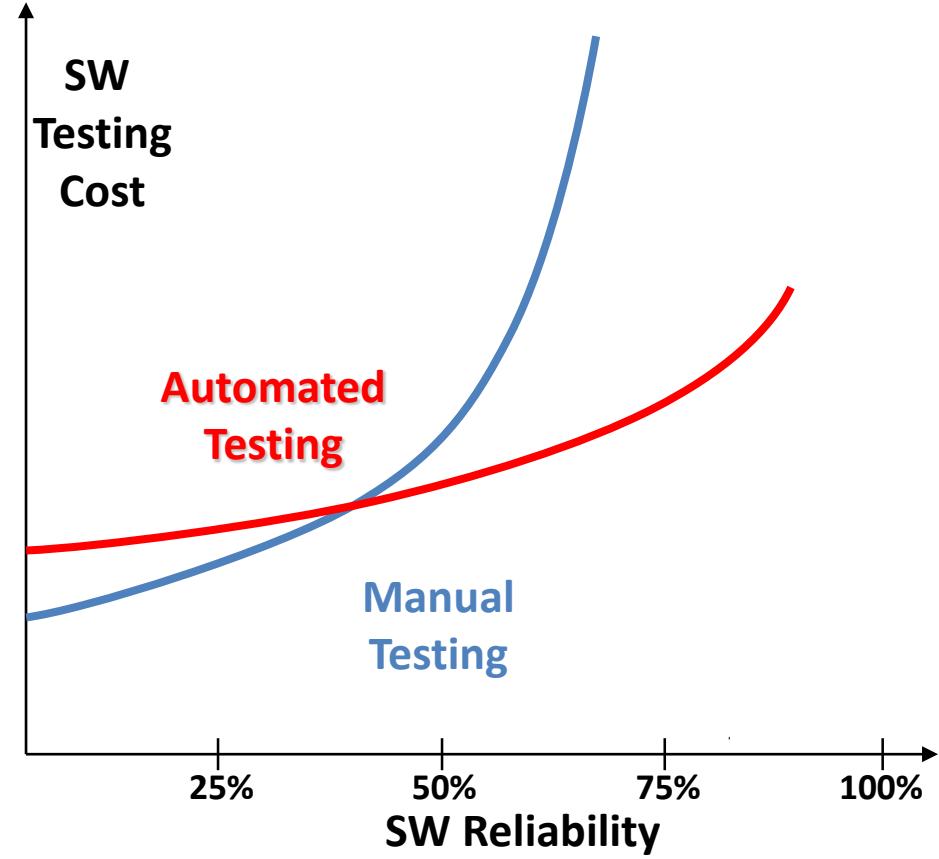
KAIST

# Future Direction

- Tool support will be strengthened for automated SW analysis
  - Ex. CBMC, BLAST, CREST, KLEE, and Microsoft PEX
  - Automated SW analysis will be performed routinely like GCC
  - Labor-intensive SW analysis => automated SW analysis by few experts
- Supports for concurrency analysis
  - Deadlock/livelock detection
  - Data-race detection
- Less user input, more analysis result and less false alarm
  - Fully automatic C++ syntax & type check (1980s)
  - (semi) automatic null-pointer dereference check (2000s)
  - (semi) automatic user-given assertion check (2020s)
  - (semi) automatic debugging (2030s)

# Conclusion

- Automated concolic testing is effective and efficient for testing industrial embedded software including vehicle domain as well as consumer electronics domain
  - S사에서의 성공사례에 힘입어, M사에서도 2017년부터 산학과제를 통해 자체 개발 sw에 도입
- Successful application of automated testing techniques requires expertise of human engineers



## Traditional testing

- Manual TC gen
- Testing main scenarios
- System-level testing
- Small # of TCs



## Concolic testing

- Automated TC gen
- Testing exceptional scenarios
- Unit-level testing
- Large # of TCs