CS 489 / 698: Software and Systems Security

Module: Bug Finding Tools and Practices
Lecture: symbolic execution

Meng Xu *(University of Waterloo)*

Fall 2024

Outline

Motivation

Q: Why research on symbolic execution when we have unit testing or even fuzzing?

Motivation

Q: Why research on symbolic execution when we have unit testing or even fuzzing?

A: A more complete exploration of program states.

Illustration

```
fn foo(x: u64): u64 {
    if (x * 3 == 42) {
        some_hidden_bug();
    }
    if (x * 5 == 42) {
        some_hidden_bug();
    }
    return 2 * x;
}
```

Illustration

Unit Test

```
foo(0);
foo(1);
```

```
fn foo(x: u64): u64 {
    if (x * 3 == 42) {
        some_hidden_bug();
    }
    if (x * 5 == 42) {
        some_hidden_bug();
    }
    return 2 * x;
}
```

Illustration

```rust
fn foo(x: u64): u64 {
    if (x * 3 == 42) {
        some_hidden_bug();
    }
    if (x * 5 == 42) {
        some_hidden_bug();
    }
    return 2 * x;
}
```

Unit Test

```
foo(0);
foo(1);
```

Fuzzing

```
foo(0);
foo(1);
foo(12);
foo(78);
......
foo(9,223,372,036,854,775,808);
```

Illustration

```
fn foo(x: u64): u64 {
    if (x * 3 == 42) {
        some_hidden_bug();
    }
    if (x * 5 == 42) {
        some_hidden_bug();
    }
    return 2 * x;
}
```

Unit Test

foo(0);

foo(1);

Fuzzing

foo(0);

foo(1);

foo(12);

foo(78);

......

foo(9,223,372,036,854,775,808);

Symbolic execution

foo(x)

aborts when $x = 14$

returns $2x$ otherwise

Satisfiability Modulo Theories (SMT)

Definition: A procedure that decides whether a mathematical formula is satisfiable.

Example:

- $3x = 42$
- $2x \geq 2^{64}$
- $5x = 42$

Satisfiability Modulo Theories (SMT)

Definition: A procedure that decides whether a mathematical formula is satisfiable.

Example:

- $3x = 42 \longrightarrow$ satisfiable with $x = 14$
- $2x \geq 2^{64} \longrightarrow$ satisfiable with $x \geq 2^{63}$
- $5x = 42 \longrightarrow$ unsatisfiable, cannot find an x

Ask two question whenever you see a symbolic execution work:

- How does it convert code into mathematical formula?
- What does it try to solve for?

Program Modeling Desiderata

- Control-flow graph exploration
- Loop handling
- Memory modeling
- Concurrency

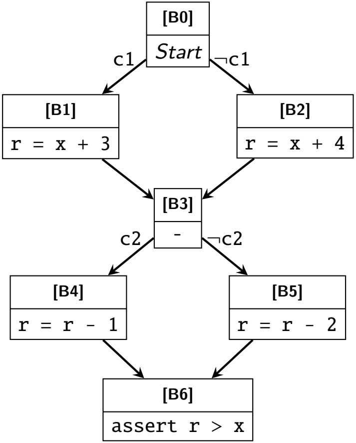
Outline

An example of a pure function

```
fn foo(
    c1: bool, c2: bool,
    x: u64
) -> u64 {
    let r = if (c1) {
        x + 3
    } else {
        x + 4
    };

    let r = if (c2) {
        r - 1
    } else {
        r - 2
    };

    r
}
spec foo {
    ensures r > x;
}
```

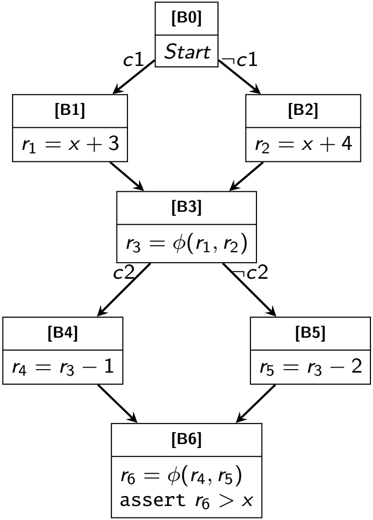
An example of a pure function

```
fn foo(
    c1: bool, c2: bool,
    x: u64
) -> u64 {
    let r = if (c1) {
        x + 3
    } else {
        x + 4
    };

    let r = if (c2) {
        r - 1
    } else {
        r - 2
    };

    r
}
spec foo {
    ensures r > x;
}
```



The example in SSA form

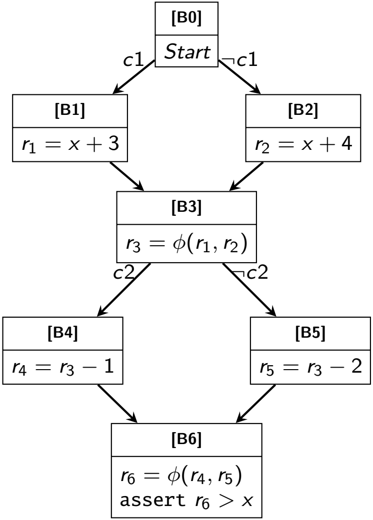
```
fn foo(
    c1: bool, c2: bool,
    x: u64
) -> u64 {
    let r = if (c1) {
        x + 3
    } else {
        x + 4
    };

    let r = if (c2) {
        r - 1
    } else {
        r - 2
    };

    r
}
spec foo {
    ensures r > x;
}
```



Path-based exploration

Vars: $c1$, $c2$, x, r_{1-6}

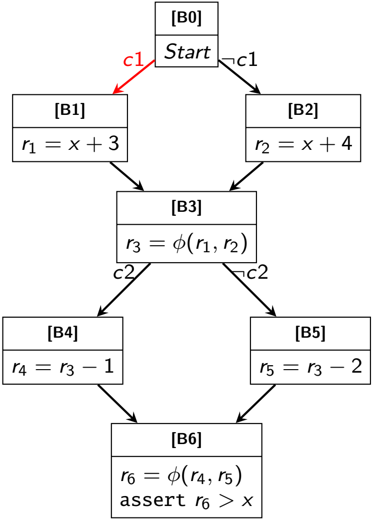
B0	Sym. repr. Path cond.	$\emptyset$ True



Path-based exploration

Vars: $c1$, $c2$, x, r_{1-6}

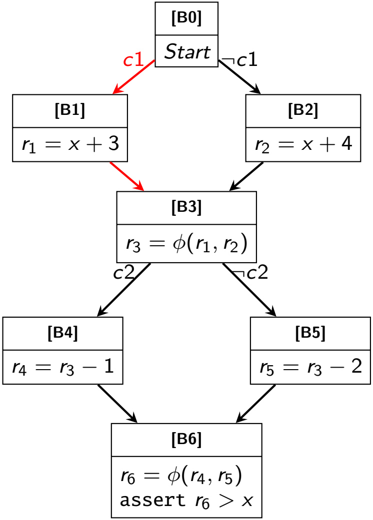
B0	Sym. repr. Path cond.	$\emptyset$ True
B1	Sym. repr. Path cond.	$r_1 = x + 3$ $c1$



Path-based exploration

Vars: $c1$, $c2$, x, r_{1-6}

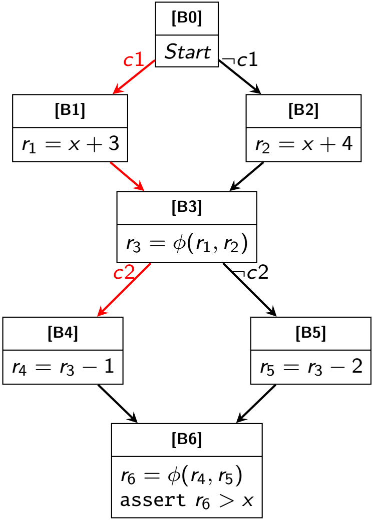
B0	Sym. repr. Path cond.	$\emptyset$ True
B1	Sym. repr. Path cond.	$r_1 = x + 3$ $c1$
B3	Sym. repr. Path cond.	$r_1 = x + 3$ $r_3 = r_1$ $c1$



Path-based exploration

Vars: $c1$, $c2$, x, r_{1-6}

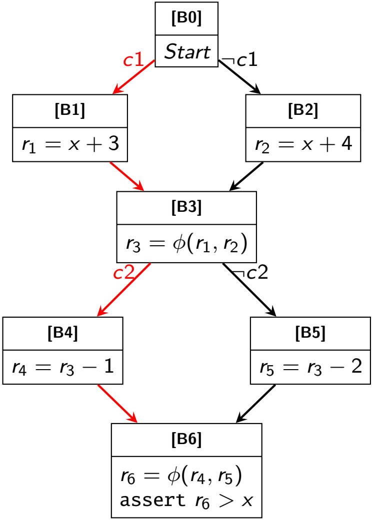
B0	Sym. repr. Path cond.	$\emptyset$ True
B1	Sym. repr. Path cond.	$r_1 = x + 3$ $c1$
B3	Sym. repr. Path cond.	$r_1 = x + 3$ $r_3 = r_1$ $c1$
B4	Sym. repr. Path cond.	$r_1 = x + 3$ $r_3 = r_1$ $r_4 = r_3 - 1$ $c_1 \wedge c_2$



Path-based exploration

Vars: $c1$, $c2$, x, r_{1-6}

B0	Sym. repr. Path cond.	$\emptyset$ True
B1	Sym. repr. Path cond.	$r_1 = x + 3$ $c1$
B3	Sym. repr. Path cond.	$r_1 = x + 3$ $r_3 = r_1$ $c1$
B4	Sym. repr. Path cond.	$r_1 = x + 3$ $r_3 = r_1$ $r_4 = r_3 - 1$ $c_1 \wedge c_2$
B6	Sym. repr. Path cond.	$r_1 = x + 3$ $r_3 = r_1$ $r_4 = r_3 - 1$ $r_6 = r_4$ $c_1 \wedge c_2$

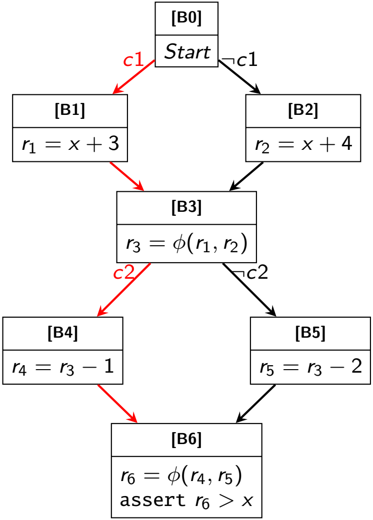


Proving procedure (per path)

Vars: $c1$, $c2$, x, r_{1-6}

B6	Sym. repr.	$r_1 = x + 3$ $r_3 = r_1$ $r_4 = r_3 - 1$ $r_6 = r_4$
	Path cond.	$c_1 \wedge c_2$

$\rightsquigarrow$



Proving procedure (per path)

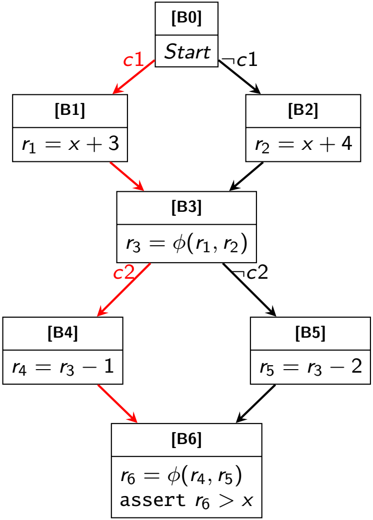
Vars: $c1,\ c2,\ x,\ r_{1-6}$

B6	Sym. repr.	$r_1 = x + 3$ $r_3 = r_1$ $r_4 = r_3 - 1$ $r_6 = r_4$
	Path cond.	$c_1 \wedge c_2$

$\rightsquigarrow$

Prove that $\forall\ c1, c2, x, r_{1-6}$:

$$((c1 \wedge c2) \wedge (\\ (r_1 = x + 3) \\ (r_3 = r_1) \\ (r_4 = r_3 - 1) \\ (r_6 = r_4) \\ )) \Rightarrow (r_6 > x)$$

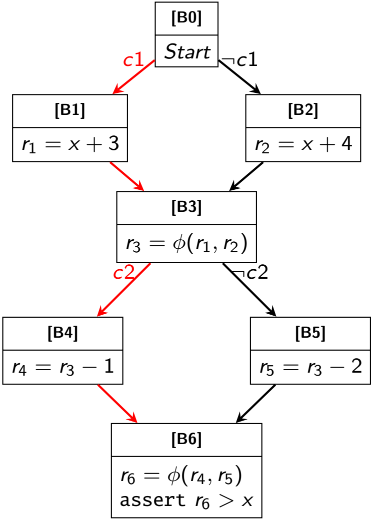


Proving procedure (all paths)

Prove that

$\forall\, c1, c2, x, r_{1-6}$:

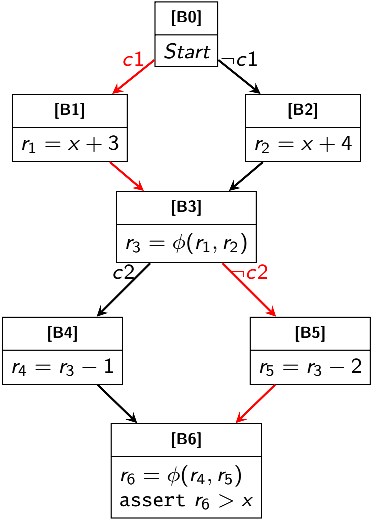
$((c1 \wedge c2) \wedge ($
$(r_1 = x + 3)$
$(r_3 = r_1)$
$(r_4 = r_3 - 1)$
$(r_6 = r_4)$
$)) \Rightarrow (r_6 > x)$



Proving procedure (all paths)

Prove that
$\forall\ c1, c2, x, r_{1-6}$:

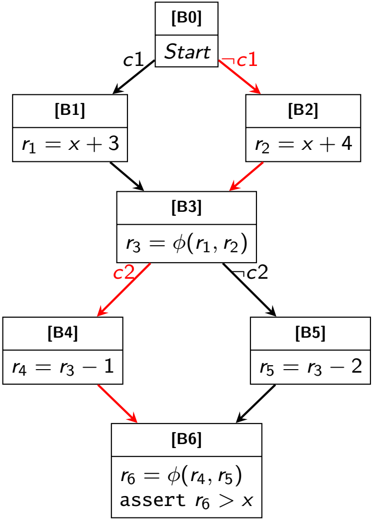
$$((c1 \land \neg c2) \land ($$
$$(r_1 = x + 3)$$
$$(r_3 = r_1)$$
$$(r_5 = r_3 - 2)$$
$$(r_6 = r_5)$$
$$)) \Rightarrow (r_6 > x)$$



Proving procedure (all paths)

Prove that
$\forall\ c1, c2, x, r_{1-6}$:

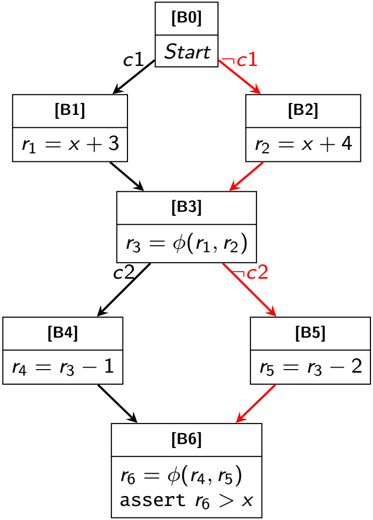
$$((\neg c1 \land c2) \land ($$
$$(r_2 = x + 4)$$
$$(r_3 = r_2)$$
$$(r_4 = r_3 - 1)$$
$$(r_6 = r_4)$$
$$)) \Rightarrow (r_6 > x)$$



Proving procedure (all paths)

Prove that
$\forall\ c1, c2, x, r_{1-6}$:

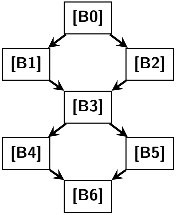
$$((\neg c1 \wedge \neg c2) \wedge (\\ (r_2 = x + 4) \\ (r_3 = r_2) \\ (r_5 = r_3 - 2) \\ (r_6 = r_5) \\ )) \Rightarrow (r_6 > x)$$



Path explosion

Path explosion

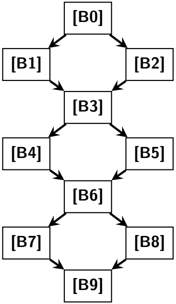
2^{2} paths



Path explosion

2^{2} paths

2^{3} paths



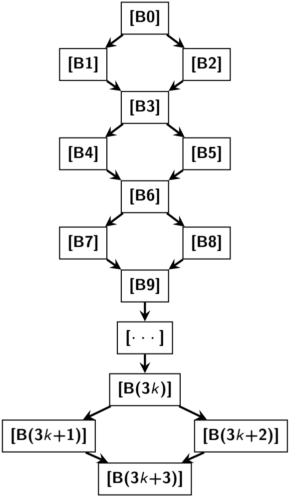
Path explosion

2^{2} paths

2^{3} paths

...

2^{k} paths

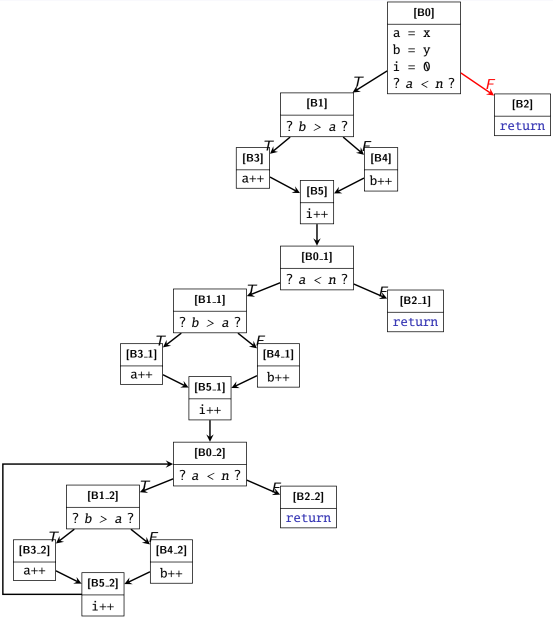


How about loops?

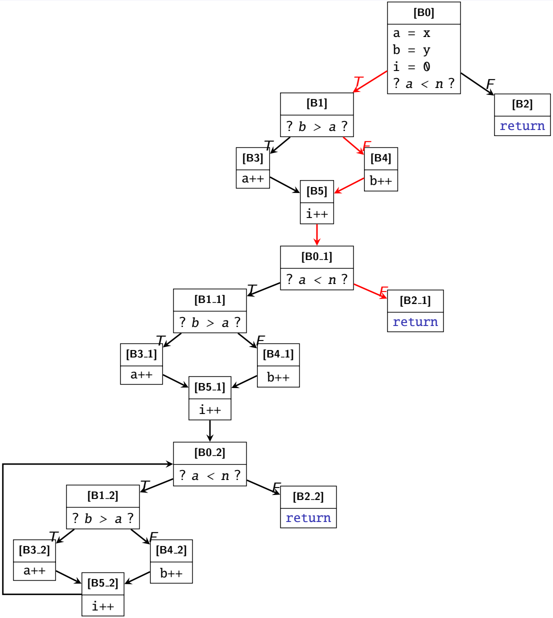
```
// a library function
fn sync(
  x: u64, y: u64, n: u64
) -> (u64, u64, u64) {
  let a = x, b = y, i = 0;
  while (a < n) {
    if (b > a) {
      a++;
    } else {
      b++;
    }
    i++;
  }
  return (a, b, i);
}
```

```
// core application logic
pub fn main() {
  let (x, y, n) = input();
  let (a, b, i) = sync(x, y, n);
  assert!(i == 0 || i < 2*n);
  \\\\\\\\\\\\\\\\\\\\\\\\\\\
}
```

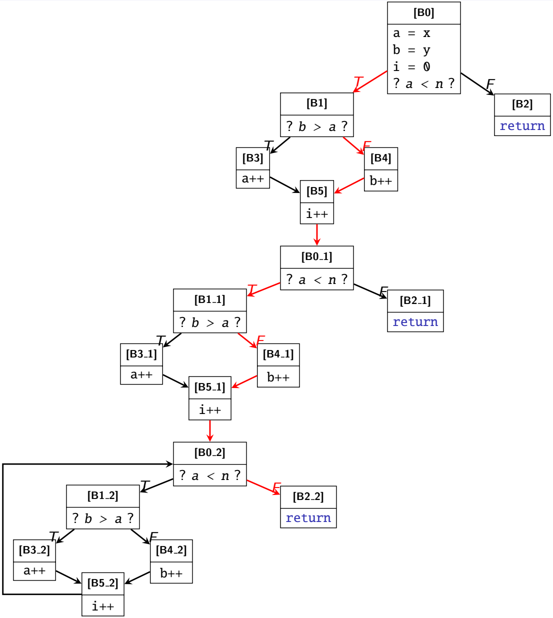
Conventional symbolic execution



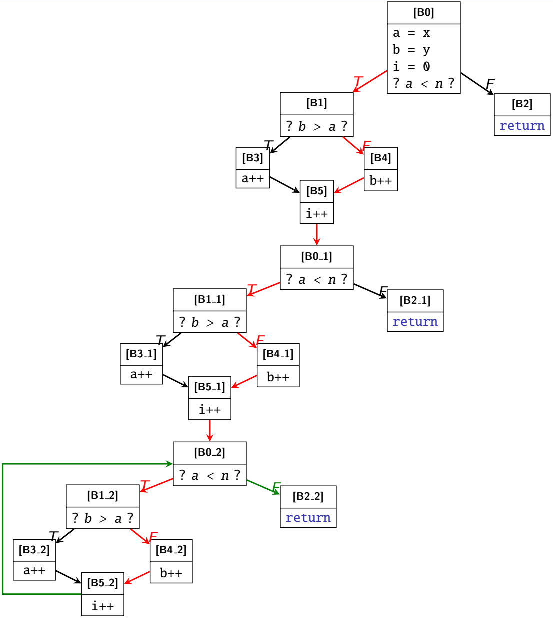
Conventional symbolic execution



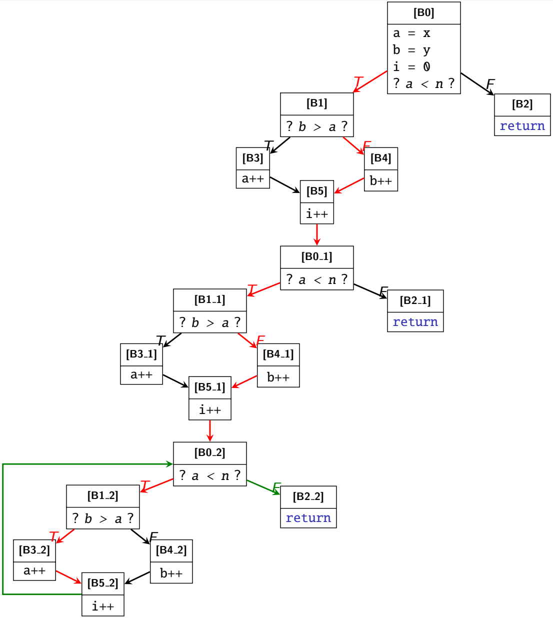
Conventional symbolic execution



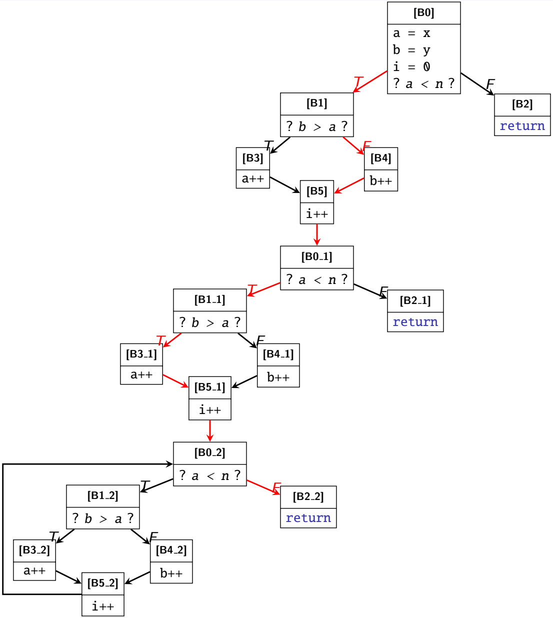
Conventional symbolic execution



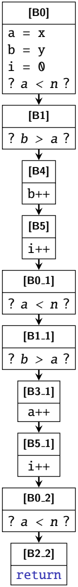
Conventional symbolic execution



Conventional symbolic execution



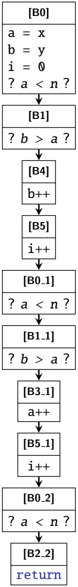
Encoding the path conditions



Find x, y, n such that

- $x < n$ (from [B0])
- $y \leq x$ (from [B1])
- $x < n$ (from [B0_1])
- $y + 1 > x$ (from [B1_1])
- $x + 1 \geq n$ (from [B0_2])
- $n \neq 0 \wedge i \geq 2n$ (from `assert!`)

Encoding the path conditions



Find x, y, n such that

- $x < n$ (from [B0])
- $y \leq x$ (from [B1])
- $x < n$ (from [B0_1])
- $y + 1 > x$ (from [B1_1])
- $x + 1 \geq n$ (from [B0_2])
- $n \neq 0 \wedge i \geq 2n$ (from `assert!`)

Solving the predicates yield:
$\{\ x = 0,\ y = 0,\ n = 1\ \}$

Symbolic execution for bug finding

```
// a library function
fn sync(
  x: u64, y: u64, n: u64
) -> (u64, u64, u64) {
  let a = x, b = y, i = 0;
  while (a < n) {
    if (b > a) {
      a++;
    } else {
      b++;
    }
    i++;
  }
  return (a, b, i);
}
```

```
// core application logic
pub fn main() {
  let (x, y, n) = input();
  let (a, b, i) = sync(x, y, n);
  assert!(i == 0 || i < 2*n);
  \\\\\\\\\\\\\\\\\\\\\\\\\\\
}
```

- x=0, y=0, n=1 → a=1, b=1, i=2
- x=0, y=0, n=2 → a=2, b=2, i=4
- ......
- x=0, y=0, n=k → a=k, b=k, i=$2k$

Path explosion in symbolic execution

```
// a library function
fn sync(
  x: u64, y: u64, n: u64
) -> (u64, u64, u64) {
  let a = x, b = y, i = 0;
  while (a < n) {
    if (b > a) {
      a++;
    } else {
      b++;
    }
    i++;
  }
  return (a, b, i);
}
```

```
// core application logic
pub fn main() {
  let (x, y, n) = input();
  let (a, b, i) = sync(x, y, n);
  assert!(n-a-b+i != 42);
  \\\\\\\\\\\\\\\\\\\\\\\\
}
```

Q: What if a bug can only be triggered after exploring k branches?

Path explosion in symbolic execution

```
// a library function
fn sync(
  x: u64, y: u64, n: u64
) -> (u64, u64, u64) {
  let a = x, b = y, i = 0;
  while (a < n) {
    if (b > a) {
      a++;
    } else {
      b++;
    }
    i++;
  }
  return (a, b, i);
}
```

```
// core application logic
pub fn main() {
  let (x, y, n) = input();
  let (a, b, i) = sync(x, y, n);
  assert!(n-a-b+i != 42);
  \\\\\\\\\\\\\\\\\\\\\\\
}
```

Q: What if a bug can only be triggered after exploring *k* branches?

In fact, this bug can only be triggered after at least 42 levels of loop unrolling.

- x=0, y=0, n=42 → a=42, b=42, i=84
- x=9, y=5, n=56 → a=56, b=56, i=98

In the conventional way of symbolic execution, finding this bug requires an exhaustive search of 2^{42} paths.

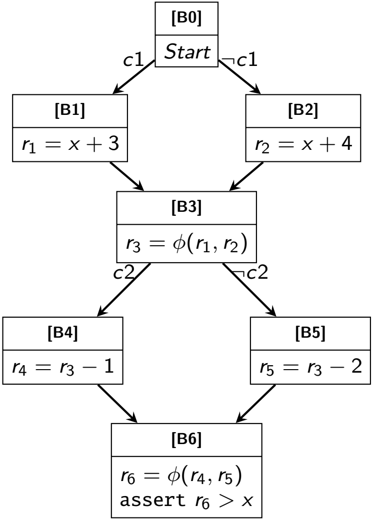
Outline

Weakest precondition calculus

Move prover (Boogie actually) adopts a backward state exploration process, following the weakest precondition calculus.

The running example, once again

```
fn foo(
    c1: bool, c2: bool,
    x: u64
) -> u64 {
    let r = if (c1) {
        x + 3
    } else {
        x + 4
    };

    let r = if (c2) {
        r - 1
    } else {
        r - 2
    };

    r
}
spec foo {
    ensures r > x;
}
```



The passification process

Convert the program into a dynamic single assignment (DSA) form.

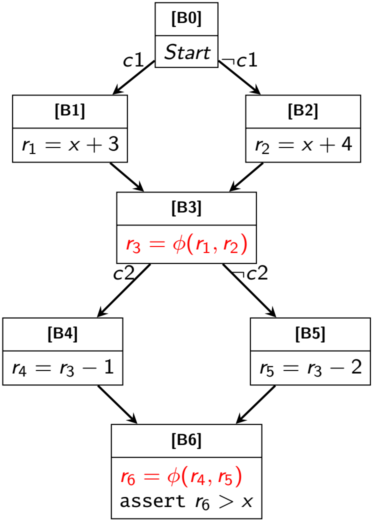
The passification process

Convert the program into a dynamic single assignment (DSA) form.

DSA is extremely similar to static single assignment (SSA) with the ϕ-node eagerly uplifted.

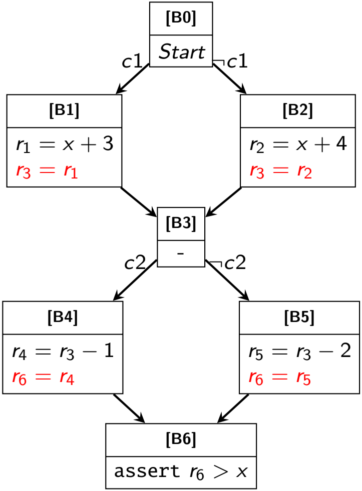
The passification process

```
fn foo(
    c1: bool, c2: bool,
    x: u64
) -> u64 {
    let r = if (c1) {
        x + 3
    } else {
        x + 4
    };

    let r = if (c2) {
        r - 1
    } else {
        r - 2
    };

    r
}
spec foo {
    ensures r > x;
}
```



The passification process

```
fn foo(
    c1: bool, c2: bool,
    x: u64
) -> u64 {
    let r = if (c1) {
        x + 3
    } else {
        x + 4
    };

    let r = if (c2) {
        r - 1
    } else {
        r - 2
    };

    r
}
spec foo {
    ensures r > x;
}
```



The walk-up process

Do a topological sort on the CFG and traverse backward.

The walk-up process

Do a topological sort on the CFG and traverse backward.

This ensures that for each block in the CFG, we visit it *once and only once* (assuming no loops).

The walk-up algorithm

Follow these rules for the intra-block walk-up process:

- $wp(\text{assert } c) = c$
- $wp(\text{assert } c, Q) = c \land Q$
- $wp(\text{assign } e, Q) = e \implies Q$
- $wp(s_1; s_2, Q) = wp(s_1, wp(s_2, Q))$

The walk-up algorithm

Follow these rules for the intra-block walk-up process:

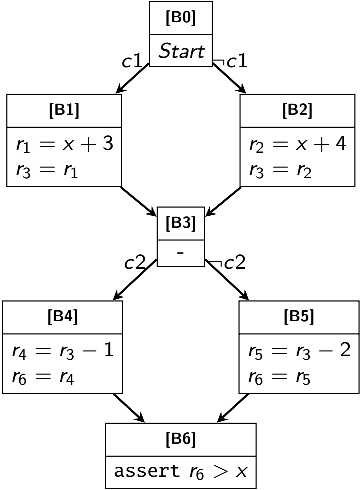
- $wp(\text{assert } c) = c$
- $wp(\text{assert } c, Q) = c \wedge Q$
- $wp(\text{assign } e, Q) = e \implies Q$
- $wp(s_1; s_2, Q) = wp(s_1, wp(s_2, Q))$

The rule for inter-block walk-up is:

$$A \leftarrow wp(s_1; s_2; \ldots; s_n, \bigwedge_{B \in \text{Succ}(A)} B)$$

The walk-up process with an example

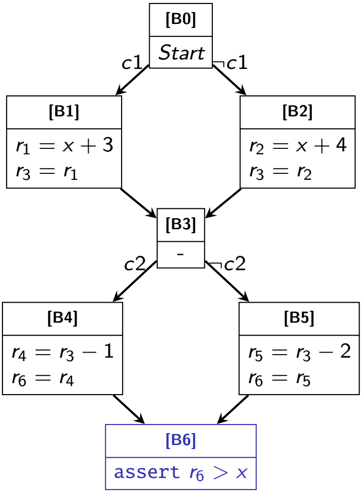
Vars: $c1$, $c2$, x, r_{1-6}, B_{0-6}



The walk-up process with an example

Vars: $c1$, $c2$, x, r_{1-6}, B_{0-6}

$B_6 \leftarrow r_6 > x$

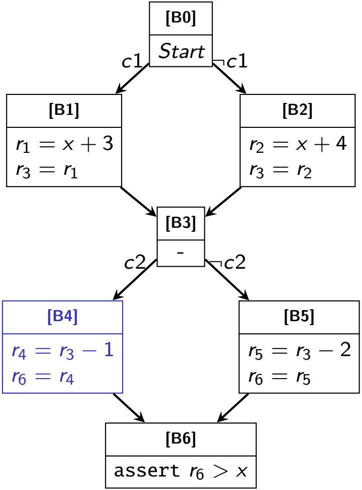


The walk-up process with an example

Vars: $c1$, $c2$, x, r_{1-6}, B_{0-6}

$B_6 \leftarrow r_6 > x$

$B_4 \leftarrow (c2) \Rightarrow ($
$\quad (r_4 = r_3 - 1) \Rightarrow ($
$\quad\quad (r_6 = r_4) \Rightarrow B_6))$



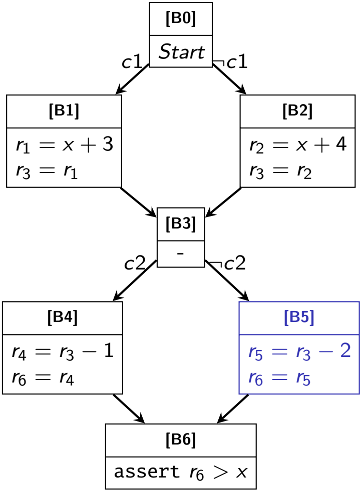
The walk-up process with an example

Vars: $c1,\ c2,\ x,\ r_{1-6},\ B_{0-6}$

$B_6 \leftarrow r_6 > x$

$$B_4 \leftarrow (c2) \Rightarrow ((r_4 = r_3 - 1) \Rightarrow ((r_6 = r_4) \Rightarrow B_6))$$

$$B_5 \leftarrow (\neg c2) \Rightarrow ((r_5 = r_3 - 2) \Rightarrow ((r_6 = r_5) \Rightarrow B_6))$$



The walk-up process with an example

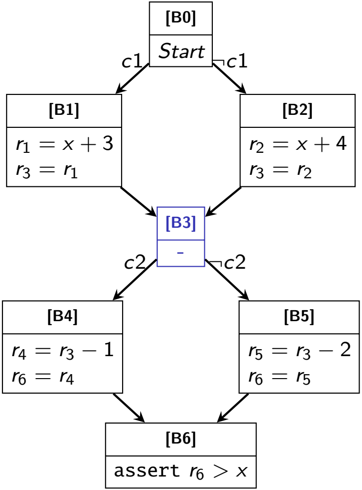
Vars: $c1,\ c2,\ x,\ r_{1-6},\ B_{0-6}$

$B_6 \leftarrow r_6 > x$

$$B_4 \leftarrow (c2) \Rightarrow ((r_4 = r_3 - 1) \Rightarrow ((r_6 = r_4) \Rightarrow B_6))$$

$$B_5 \leftarrow (\neg c2) \Rightarrow ((r_5 = r_3 - 2) \Rightarrow ((r_6 = r_5) \Rightarrow B_6))$$

$B_3 \leftarrow B_4 \wedge B_5$



The walk-up process with an example

Vars: $c1,\ c2,\ x,\ r_{1-6},\ B_{0-6}$

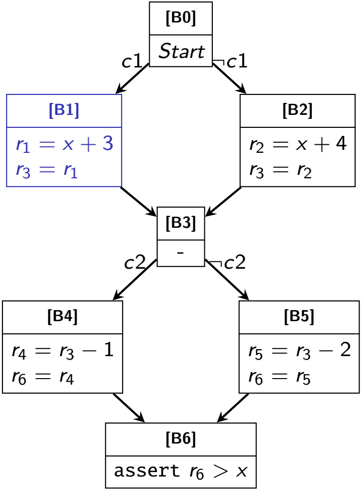
$B_6 \leftarrow r_6 > x$

$$B_4 \leftarrow (c2) \Rightarrow ((r_4 = r_3 - 1) \Rightarrow ((r_6 = r_4) \Rightarrow B_6))$$

$$B_5 \leftarrow (\neg c2) \Rightarrow ((r_5 = r_3 - 2) \Rightarrow ((r_6 = r_5) \Rightarrow B_6))$$

$B_3 \leftarrow B_4 \wedge B_5$

$$B_1 \leftarrow (c1) \Rightarrow ((r_1 = x + 3) \Rightarrow ((r_3 = r_1) \Rightarrow B_3))$$



The walk-up process with an example

Vars: $c1,\ c2,\ x,\ r_{1-6},\ B_{0-6}$

$B_6 \leftarrow r_6 > x$

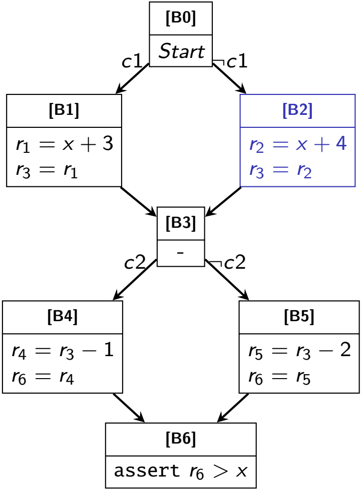
$$B_4 \leftarrow (c2) \Rightarrow ((r_4 = r_3 - 1) \Rightarrow ((r_6 = r_4) \Rightarrow B_6))$$

$$B_5 \leftarrow (\neg c2) \Rightarrow ((r_5 = r_3 - 2) \Rightarrow ((r_6 = r_5) \Rightarrow B_6))$$

$B_3 \leftarrow B_4 \wedge B_5$

$$B_1 \leftarrow (c1) \Rightarrow ((r_1 = x + 3) \Rightarrow ((r_3 = r_1) \Rightarrow B_3))$$

$$B_2 \leftarrow (\neg c1) \Rightarrow ((r_2 = x + 4) \Rightarrow ((r_3 = r_2) \Rightarrow B_3))$$



The walk-up process with an example

Vars: $c1, c2, x, r_{1-6}, B_{0-6}$

$B_6 \leftarrow r_6 > x$

$B_4 \leftarrow (c2) \Rightarrow ($
$(r_4 = r_3 - 1) \Rightarrow ($
$(r_6 = r_4) \Rightarrow B_6))$

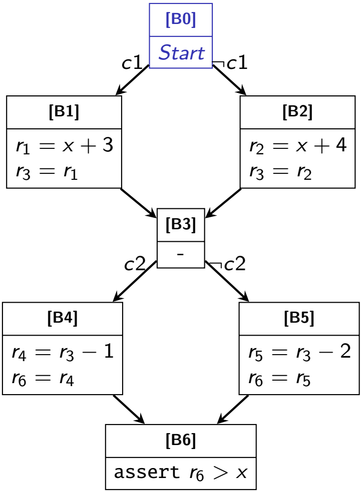
$B_5 \leftarrow (\neg c2) \Rightarrow ($
$(r_5 = r_3 - 2) \Rightarrow ($
$(r_6 = r_5) \Rightarrow B_6))$

$B_3 \leftarrow B_4 \wedge B_5$

$B_1 \leftarrow (c1) \Rightarrow ($
$(r_1 = x + 3) \Rightarrow ($
$(r_3 = r_1) \Rightarrow B_3))$

$B_2 \leftarrow (\neg c1) \Rightarrow ($
$(r_2 = x + 4) \Rightarrow ($
$(r_3 = r_2) \Rightarrow B_3))$

$B_0 \leftarrow B_1 \wedge B_2$

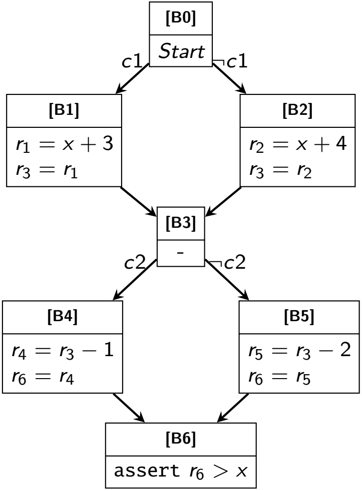


Proving procedure

Prove that
$\forall\ c1, c2, x, r_{1-6}, B_{0-6}$:

$$
\begin{aligned}
&B_6 \leftarrow r_6 > x\\
&B_4 \leftarrow (c2) \Rightarrow (\\
&\quad (r_4 = r_3 - 1) \Rightarrow (\\
&\qquad (r_6 = r_4) \Rightarrow B_6))\\
&B_5 \leftarrow (\neg c2) \Rightarrow (\\
&\quad (r_5 = r_3 - 2) \Rightarrow (\\
&\qquad (r_6 = r_5) \Rightarrow B_6))\\
&B_3 \leftarrow B_4 \wedge B_5\\
&B_1 \leftarrow (c1) \Rightarrow (\\
&\quad (r_1 = x + 3) \Rightarrow (\\
&\qquad (r_3 = r_1) \Rightarrow B_3))\\
&B_2 \leftarrow (\neg c1) \Rightarrow (\\
&\quad (r_2 = x + 4) \Rightarrow (\\
&\qquad (r_3 = r_2) \Rightarrow B_3))\\
&B_0 \leftarrow B_1 \wedge B_2
\end{aligned}
$$

$B_0 = \text{True}$



Comparison of forward and backward symbolic execution

Prove that $\forall\ c1, c2, x, r_{1-6}$:

$$
\begin{aligned}
&((c1 \wedge c2) \wedge (\\
&\quad (r_1 = x + 3) \\
&\quad (r_3 = r_1) \\
&\quad (r_4 = r_3 - 1) \\
&\quad (r_6 = r_4) \\
&)) \Rightarrow (r_6 > x)
\end{aligned}
$$

However, need to repeat this process multiple (worst case exponential) times.

Prove that
$\forall\ c1, c2, x, r_{1-6}, B_{0-6}$:

$$
\begin{aligned}
&B_6 \leftarrow r_6 > x \\
&B_4 \leftarrow (c2) \Rightarrow (\\
&\qquad (r_4 = r_3 - 1) \Rightarrow (\\
&\qquad\quad (r_6 = r_4) \Rightarrow B_6)) \\
&B_5 \leftarrow (\neg c2) \Rightarrow (\\
&\qquad (r_5 = r_3 - 2) \Rightarrow (\\
&\qquad\quad (r_6 = r_5) \Rightarrow B_6)) \\
&B_3 \leftarrow B_4 \wedge B_5 \\
&B_1 \leftarrow (c1) \Rightarrow (\\
&\qquad (r_1 = x + 3) \Rightarrow (\\
&\qquad\quad (r_3 = r_1) \Rightarrow B_3)) \\
&B_2 \leftarrow (\neg c1) \Rightarrow (\\
&\qquad (r_2 = x + 4) \Rightarrow (\\
&\qquad\quad (r_3 = r_2) \Rightarrow B_3)) \\
&B_0 \leftarrow B_1 \wedge B_2
\end{aligned}
$$

$B_0 = \text{True}$

Outline

Breaking cycles in the CFG

Loop invariants are keys to break cycles in the CFG

Breaking cycles in the CFG

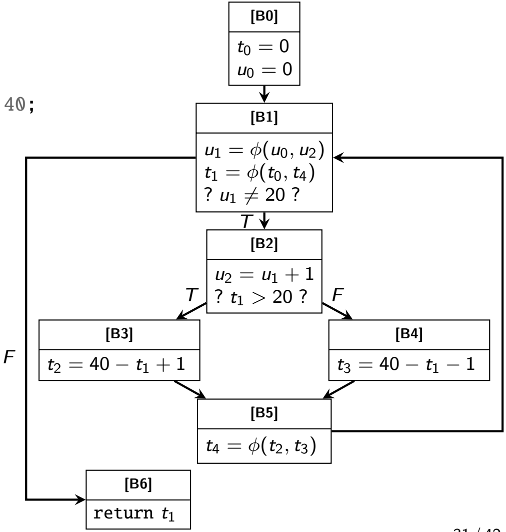
Loop invariants are keys to break cycles in the CFG

A loop invariant is transformed into statements that:

- Assert the invariant at the beginning of the loop
- Havoc (i.e., re-symbolize) the loop induction variables
- Assume the invariant to re-establish relations among the induction variables being havoc-ed
- Assert the invariant at the end of the loop body

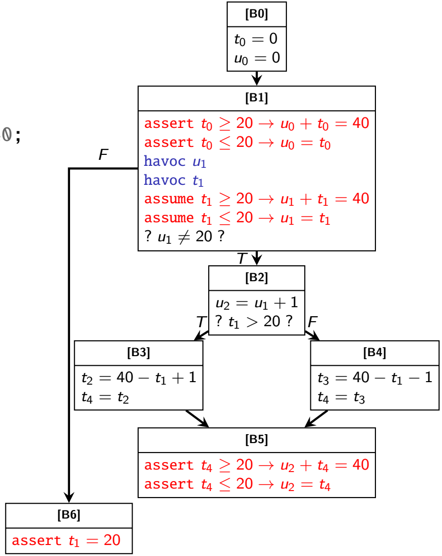
A running example

```
fn bar(): u64 {
    t: u64 = 0;
    u: u64 = 0;
    while ({
spec {
    invariant t >= 20 ==> u + t == 40;
    invariant t <= 20 ==> u == t;
}
        (u != 20)
    }) {
        u = u + 1;
        if (t > 20) {
            t = 40 - t + 1;
        } else {
            t = 40 - t - 1;
        }
    }
    t
}
spec bar {
    ensures result == 20;
}
```



A running example

```
fn bar(): u64 {
    t: u64 = 0;
    u: u64 = 0;
    while ({
spec {
    invariant t >= 20 ==> u + t == 40;
    invariant t <= 20 ==> u == t;
}
        (u != 20)
    }) {
        u = u + 1;
        if (t > 20) {
            t = 40 - t + 1;
        } else {
            t = 40 - t - 1;
        }
    }
    t
}
spec bar {
    ensures result == 20;
}
```



A running example

$$B_6 \leftarrow (u_1 = 20) \Rightarrow ((t_1 = 20))$$

$$B_5 \leftarrow ((t_4 \leq 20 \rightarrow u_2 = t_4) \wedge (t_4 \geq 20 \rightarrow u_2 + t_4 = 40))$$

$$B_4 \leftarrow (t1 \leq 20) \Rightarrow ((t_3 = 40 - t_1 - 1) \Rightarrow ((t_4 = t_3) \Rightarrow B_5))$$

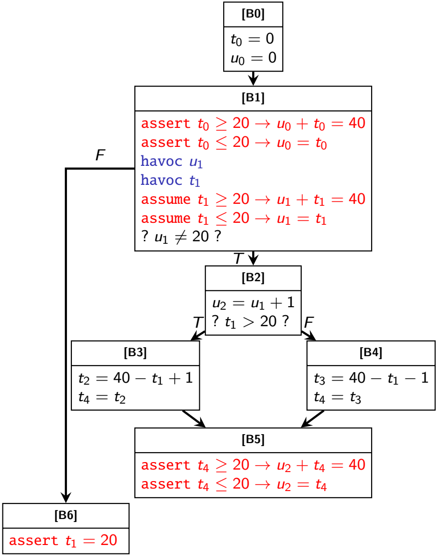
$$B_3 \leftarrow (t1 > 20) \Rightarrow ((t_2 = 40 - t_1 + 1) \Rightarrow ((t_4 = t_2) \Rightarrow B_5))$$

$$B_2 \leftarrow (u_1 \neq 20) \Rightarrow ((u_2 = u_1 + 1) \Rightarrow (B_3 \wedge B4))$$

$$B_1 \leftarrow ((t_0 \geq 20 \rightarrow u_0 + t_0 = 40) \wedge ((t_0 \leq 20 \rightarrow u_0 = t_0) \wedge ((t_1 \geq 20 \rightarrow u_1 + t_1 = 40) \Rightarrow ((t_1 \leq 20 \rightarrow u_1 = t_1) \Rightarrow (B_2 \wedge B6)))))$$

$$B_0 \leftarrow ((t_0 = 0) \Rightarrow ((u_0 = 0) \Rightarrow B_1))$$

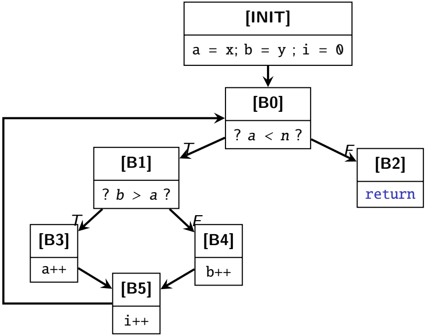
Prove that: $B_0 = \text{True}$



Back to the same symbolic loop unrolling example

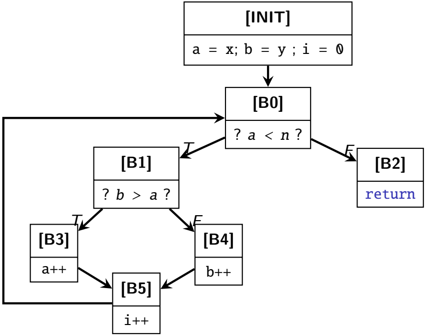
```
// a library function
fn sync(
  x: u64, y: u64, n: u64
) -> (u64, u64, u64) {
  let a = x, b = y, i = 0;
  while (a < n) {
    if (b > a) {
      a++;
    } else {
      b++;
    }
    i++;
  }
  return (a, b, i);
}
```

```
// core application logic
pub fn main() {
  let (x, y, n) = input();
  let (a, b, i) = sync(x, y, n);
  assert!(n-a-b+i != 42);

}
```



Constrained Horn Clause (CHC)

$$a = x \land b = y \land i = 0 \implies A(a, b, i)$$

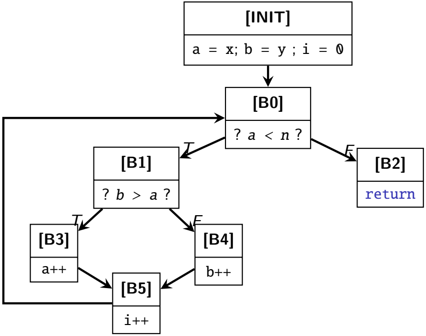


Constrained Horn Clause (CHC)

$$a = x \land b = y \land i = 0 \implies A(a, b, i)$$

$$A(a, b, i) \;\land\; a < n \land \\ b > a \;\land\; a' = a + 1 \;\land\; b' = b \land \\ i' = i + 1 \implies A(a', b', i')$$

$$A(a, b, i) \;\land\; a < n \land \\ b \leq a \;\land\; a' = a \;\land\; b' = b + 1 \land \\ i' = i + 1 \implies A(a', b', i')$$



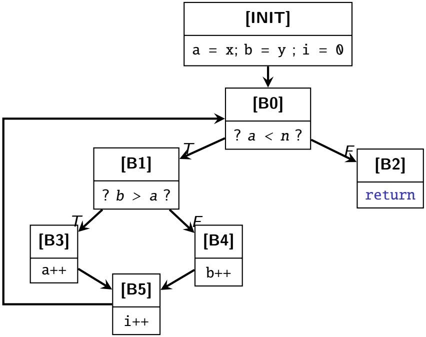
Constrained Horn Clause (CHC)

$$a = x \wedge b = y \wedge i = 0 \implies A(a, b, i)$$

$$A(a, b, i) \wedge a < n \wedge b > a \wedge a' = a + 1 \wedge b' = b \wedge i' = i + 1 \implies A(a', b', i')$$

$$A(a, b, i) \wedge a < n \wedge b \leq a \wedge a' = a \wedge b' = b + 1 \wedge i' = i + 1 \implies A(a', b', i')$$

$$A(a, b, i) \wedge a \geq n \implies B(a, b, i)$$



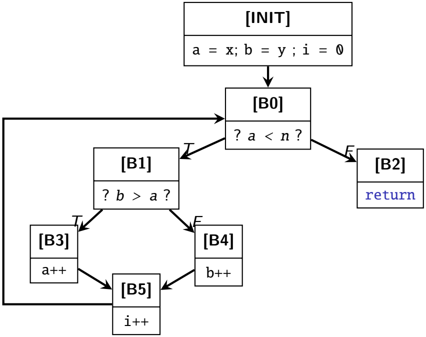
Constrained Horn Clause (CHC)

$$a = x \land b = y \land i = 0 \implies A(a, b, i)$$

$$\begin{aligned} &A(a, b, i) \ \land \ a < n \land \\ &\quad b > a \ \land \ a' = a + 1 \ \land \ b' = b \land \\ &\quad i' = i + 1 \implies A(a', b', i') \\ &A(a, b, i) \ \land \ a < n \land \\ &\quad b \leq a \ \land \ a' = a \ \land \ b' = b + 1 \land \\ &\quad i' = i + 1 \implies A(a', b', i') \end{aligned}$$

$$A(a, b, i) \ \land \ a \geq n \implies B(a, b, i)$$

solve $B(a, b, i) \ \land \ n - a - b + i = 42$



Benefits of the CHC encoding

Encoding the program in CHC allows solvers to infer invariants about the loop. In this particular case, an invariant can be inferred:

$$a + b - i = x + y$$

Benefits of the CHC encoding

Encoding the program in CHC allows solvers to infer invariants about the loop. In this particular case, an invariant can be inferred:

$$a + b - i = x + y$$

This helps to solve for a concrete test case that there exists x, y, and n such that $n = x + y + 42$.

Outline

Simple borrow

```
struct S {
    f1: u64,
    f2: u64,
}

fn foo(x: &mut S) {
    let p = &mut x.f1;
    *p = 1;
}
```

Simple borrow

```
struct S {
    f1: u64,
    f2: u64,
}

fn foo(x: &mut S) {
    let p = &mut x.f1;
    *p = 1;
}
```

```
fn _foo_(x: Mutation<S>) -> Mutation<S> {
    // p := borrow_field<S>.f1(x);
    let p = Mutation<u64> {
        root: x.root, // Param(0),
        paths: concat!(x.paths, Field(0)),
        value: x.value.f1,
    };

    // p2 := write_ref(p, 1);
    let p2 = update!(p, @value = 1);

    // x2 := write_back[x.f1](p2);
    let v = update!(x.value, @f1 = p2.value)
    let x2 = update!(x, @value = v)

    // return x2;
    x2
}
```

Mutations under borrow semantics

```rust
enum Root {
    Param(usize),
    Local(usize),
}

enum Path {
    Field(usize),
    Index(usize),
}

struct Mutation<T> {
    root: Root,
    paths: Vec<Path>,
    value: T,
}
```

Mutations under borrow semantics

```rust
enum Root {
    Param(usize),
    Local(usize),
}

enum Path {
    Field(usize),
    Index(usize),
}

struct Mutation<T> {
    root: Root,
    paths: Vec<Path>,
    value: T,
}
```

```rust
struct S {
    f1: u64,
    f2: u64,
}

fn foo(x: &mut S) {
    let p = &mut x.f1;
    *p = 1;
}
```

```rust
fn _foo_(x: Mutation<S>) -> Mutation<S> {
    Mutation<S> {
        root: x.root, // Root::Param(0)
        paths: x.paths, // vec[]
        value: S {
            f1: 1,
            f2: x.value.f2,
        }
    }
}
```

Outline

Definition of concolic execution

Background: **concolic**, as the name suggests, is the combination of two English words: *concrete* and *symbolic*, and the order matters!

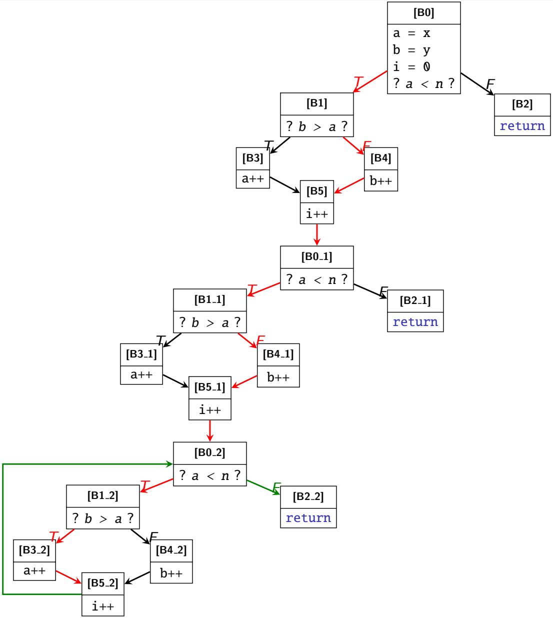
Definition of concolic execution

Background: **concolic**, as the name suggests, is the combination of two English words: *concrete* and *symbolic*, and the order matters!

The basic idea of concolic execution is:

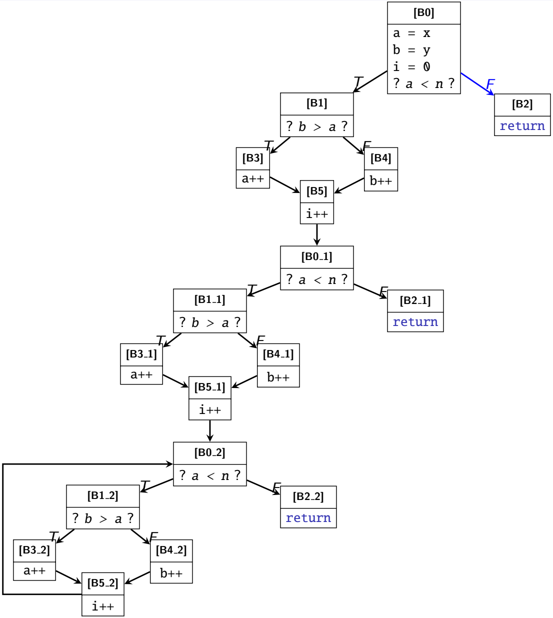
1. Execute a test case concretely
2. For each branch encountered in the test case, find another test case that toggles this branch symbolically.

Concolic execution with the running example



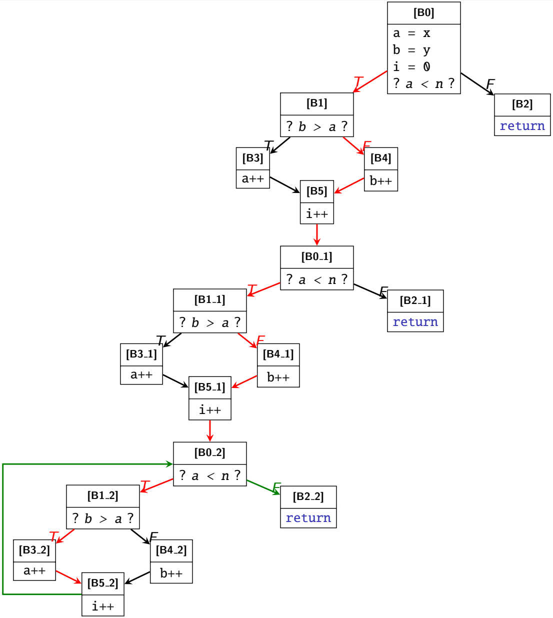
{x=1, y=0, n=2}

Concolic execution with the running example



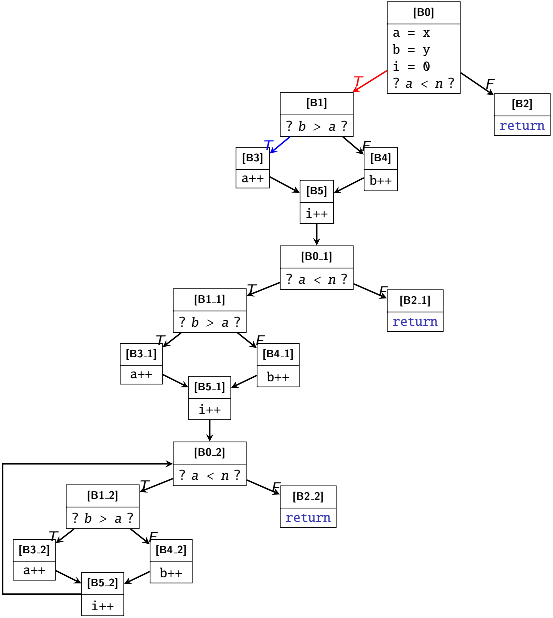
{x=9, y=2, n=6}

Concolic execution with the running example



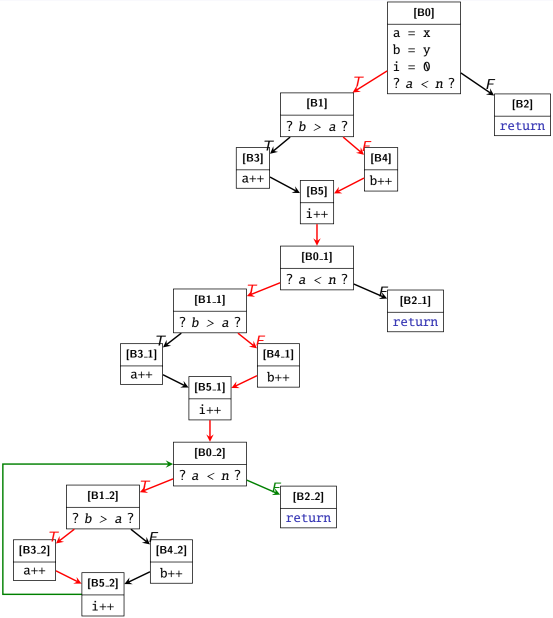
{x=1, y=0, n=2}

Concolic execution with the running example



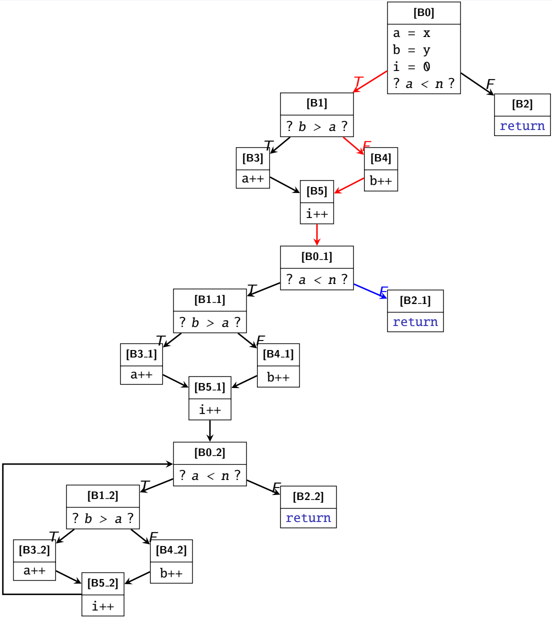
{x=3, y=4, n=5}

Concolic execution with the running example



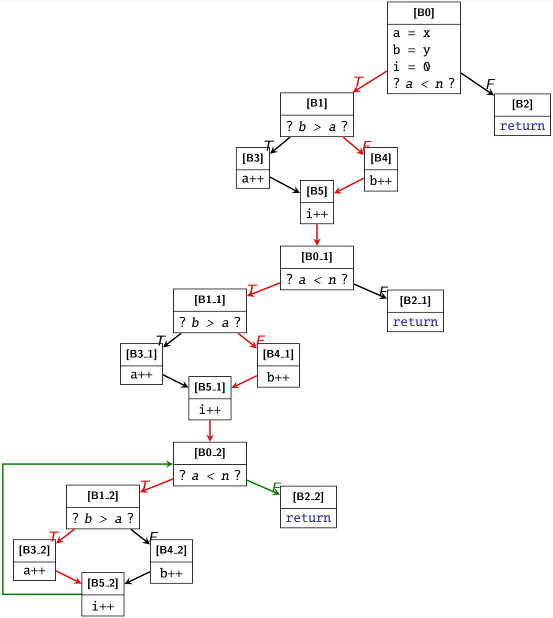
{x=1, y=0, n=2}

Concolic execution with the running example



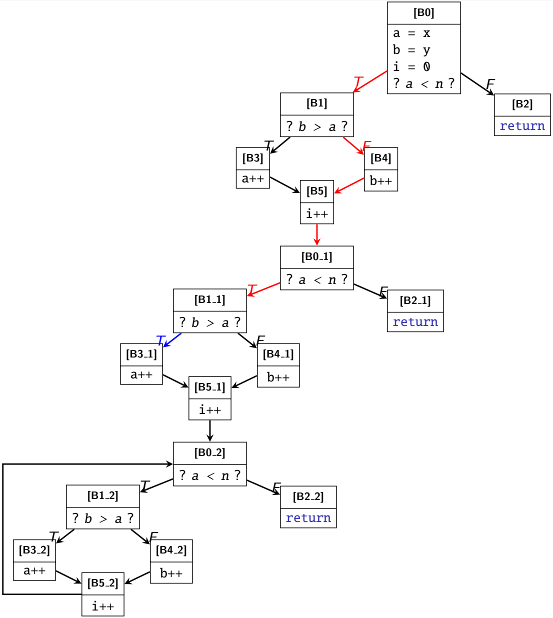
⟨*infeasible*⟩

Concolic execution with the running example



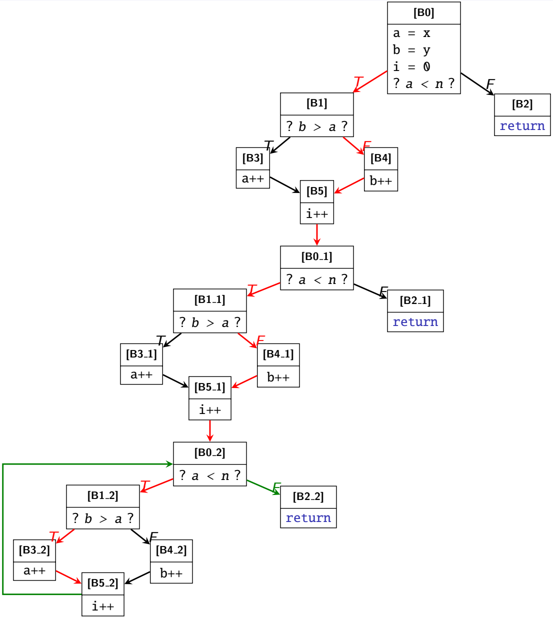
{x=1, y=0, n=2}

Concolic execution with the running example



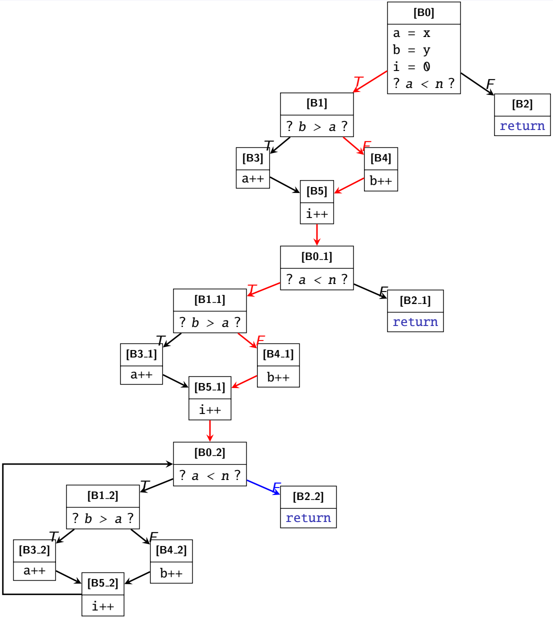
{x=5, y=5, n=8}

Concolic execution with the running example



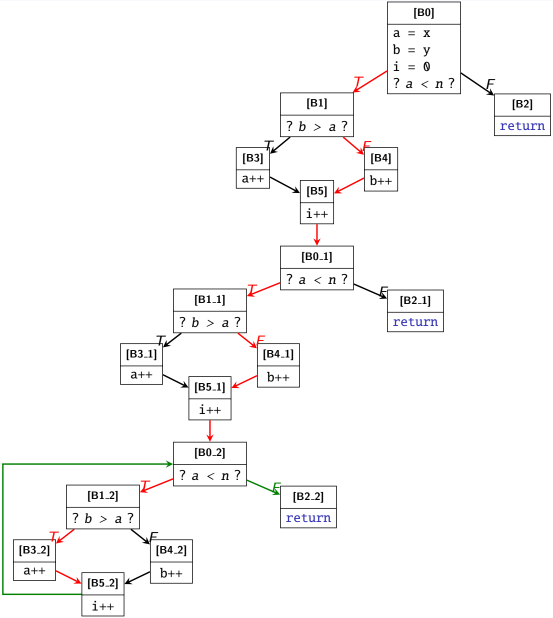
{x=1, y=0, n=2}

Concolic execution with the running example



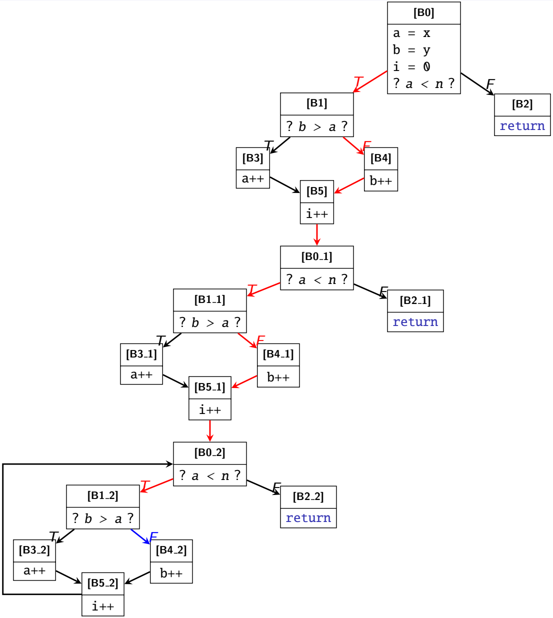
⟨*infeasible*⟩

Concolic execution with the running example



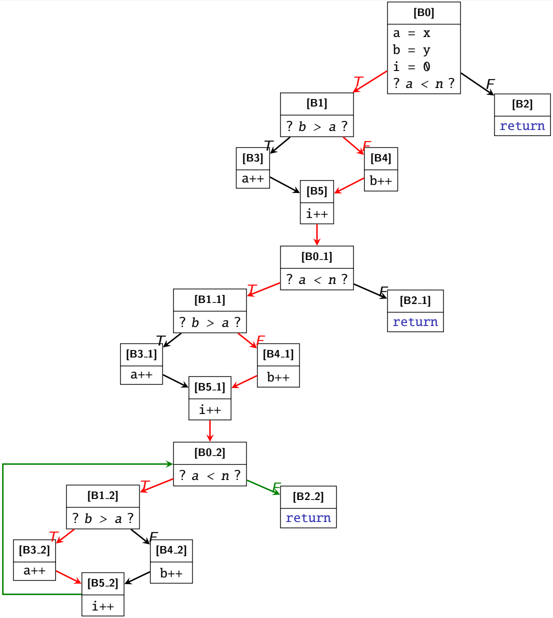
{x=1, y=0, n=2}

Concolic execution with the running example



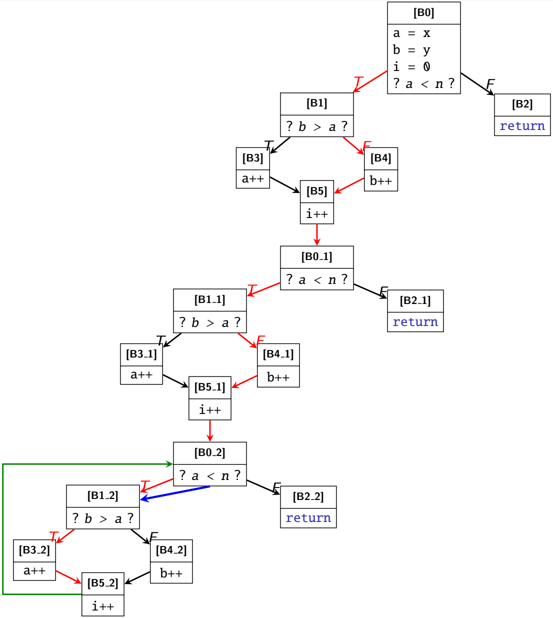
{x=7, y=3, n=9}

Concolic execution with the running example



{x=1, y=0, n=2}

Concolic execution with the running example



. . . endless loop . . .

$\langle$ **End** $\rangle$