

# Simuliris

## A Separation Logic Framework for Verifying Concurrent Program Optimizations

Lennard Gäher, Michael Sammler, Simon Spies, Ralf Jung,  
Hoang-Hai Dang, Robbert Krebbers, Jeehoon Kang, Derek Dreyer



Clang optimizes



```
int mult(int *x, int *y) {  
  
    int i = 0; int sum = *y;  
    while (i != *x - 1) {  
        i += 1; sum += *y;  
    }  
    return sum;  
}
```

```
int opt(int *x, int *y) {  
    int n = *x;  
    int m = *y;  
    int i = 0; int sum = m;  
    while (i != n - 1) {  
        i += 1; sum += m;  
    }  
    return sum;  
}
```

implements multiplication of positive integers



```
int mult(int *x, int *y) {  
  
    int i = 0; int sum = *y;  
    while (i != *x - 1) {  
        i += 1; sum += *y;  
    }  
    return sum;  
}
```

```
int opt(int *x, int *y) {  
    int n = *x;  
    int m = *y;  
    int i = 0; int sum = m;  
    while (i != n - 1) {  
        i += 1; sum += m;  
    }  
    return sum;  
}
```

implements multiplication of positive integers

**How can we prove correctness of these program optimizations?**

# Can concurrent writes break the optimization?

unoptimized:

```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

optimized:

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```

# Can concurrent writes break the optimization?

unoptimized:

```
int i = 0; int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

optimized:

```
int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

||

```
*x = 2;  
*y = 42;
```

# Can concurrent writes break the optimization?

unoptimized:


```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

optimized:

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```

|| \*x = 2;  
|| \*y = 42;

can produce results not possible  
for the unoptimized program!



# Can concurrent writes break the optimization?

current state:        \*x        \*y  
                         1        99

optimized program:

```
→int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

||

```
→*x = 2;  
*y = 42;
```

# Can concurrent writes break the optimization?

current state:        \*x        \*y        n  
                         1        99        1

optimized program:

```
int n = *x;  
→int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

||

```
→*x = 2;  
*y = 42;
```



# Can concurrent writes break the optimization?

current state:      \*x      \*y      n  
                         2      99      1

optimized program:

```
int n = *x;  
→int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

||

```
*x = 2;  
→*y = 42;
```

# Can concurrent writes break the optimization?

current state:        \*x        \*y        n  
                         2        42        1

optimized program:

```
int n = *x;  
→int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

||

```
*x = 2;  
*y = 42;  
→
```

# Can concurrent writes break the optimization?

current state:        \*x        \*y        n        m  
                         2        42        1        42

optimized program:

```
int n = *x;  
int m = *y;  
→ int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

||

```
*x = 2;  
*y = 42;  
→
```

# Can concurrent writes break the optimization?

current state:      \*x      \*y      n      m      sum

                         2      42      1      42      42

optimized program:

```

int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
→return sum;

```

||

```

*x = 2;
*y = 42;
→

```

The optimized program can produce the result 42  
 with initial \*x = 1 and \*y = 99  
 (by using the old value 1 of x and the new value 42 of y)

# The unoptimized program cannot produce the result 42

current state:            \*x     \*y  
                             1     99

unoptimized program:

```
int i = 0;
int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

||

```
*x = 2;
*y = 42;
```

# The unoptimized program cannot produce the result 42

current state:            \*x     \*y  
                             1     99

unoptimized program:

```
int i = 0;
int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

||

```
*x = 2;
*y = 42;
```

# The unoptimized program cannot produce the result 42

current state:                    \*x        \*y  
                                     2        42

unoptimized program:

```
int i = 0;  
int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

||

```
*x = 2;  
*y = 42;
```



# The unoptimized program cannot produce the result 42

current state:                    \*x        \*y  
                                     2        42

unoptimized program:

```
int i = 0;  
int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

||            \*x = 2;  
              \*y = 42;  
→



# The unoptimized program cannot produce the result 42

current state:                    \*x        \*y  
                                       2        42

unoptimized program:

```

int i = 0;
int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;

```

||            \*x = 2;  
                                       \*y = 42;  
                                       →

The unoptimized program can **not** produce the result 42  
 with initial \*x = 1 and \*y = 99  
 (if the new value 42 of y is read, also the new value 2 of x is read)

# The unoptimized program cannot produce the result 42

current state:            \*x     \*y  
                             2     42

unoptimized program:

```
int i = 0;            // ...
```

The optimization seems to introduce new program behavior!

```
}  
return sum;
```

The unoptimized program can **not** produce the result 42  
with initial  $*x = 1$  and  $*y = 99$   
(if the new value 42 of  $y$  is read, also the new value 2 of  $x$  is read)

# Correctness under concurrency: relying on undefined behavior

unoptimized:

```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

optimized:

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```

||

```
*x = 2;
*y = 42;
```

# Correctness under concurrency: relying on undefined behavior

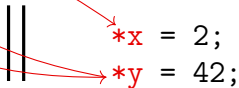
5

unoptimized:

```
int i = 0; int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

optimized:

```
int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```



**Data races** are **undefined behavior (UB)** in C/C++/unsafe Rust.

# Correctness under concurrency: relying on undefined behavior

unoptimized:

```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

optimized:

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```

||

~~\*x = 2;~~  
~~\*y = 42;~~

**Data races** are **undefined behavior (UB)** in C/C++/unsafe Rust.  
The compiler may assume their absence.

# Correctness under concurrency: relying on undefined behavior

unoptimized:

```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

o The optimization is correct. But how can we prove that?

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```

**Data races** are **undefined behavior (UB)** in C/C++/unsafe Rust.  
The compiler may assume their absence.

# The problem

How can we prove correctness of **concurrent** program optimizations relying on **data race UB** and involving **loops**?

	data race UB	concurrent	loops
[Ševčík, 2009], [Morisset et al., 2013]	✓	✓	~
[Vafeiadis et al., 2015]	✓	✓	~
CAS/Concurrent CompCert	~	~	✓
CompCertTSO [Ševčík et al., 2013]	✗	✓	✓
CCAL (CompCertX) [Gu et al., 2018]	✗	~	✓
[Liang and Feng, 2016]	✗	✓	✓
ReLoC [Frumin et al., 2018]	✗	✓	✓
[Tassarotti et al., 2017]	✗	✓	✓
Transfinite Iris [Spies et al., 2021]	✗	✗	✓
Stacked Borrows [Jung et al., 2020]	✗	✗	~

# The problem

How can we prove correctness of **concurrent** program optimizations relying on **data race UB** and involving **loops**?

	<b>data race UB</b>	<b>concurrent</b>	<b>loops</b>
[Ševčík, 2009], [Morisset et al., 2013]	✓	✓	~
[Vafeiadis et al., 2015]	✓	✓	~
CAS/Concurrent CompCert	~	~	✓
CompCertTSO [Ševčík et al., 2013]	✗	✓	✓
CCAL (CompCertX) [Gu et al., 2018]	✗	~	✓
[Liang and Feng, 2016]	✗	✓	✓
ReLoC [Frumin et al., 2018]	✗	✓	✓
[Tassarotti et al., 2017]	✗	✓	✓
Transfinite Iris [Spies et al., 2021]	✗	✗	✓
Stacked Borrows [Jung et al., 2020]	✗	✗	~



# The problem

How can we prove correctness of **concurrent** program optimizations relying on **data race UB** and involving **loops**?

	data race UB	concurrent	loops
[Ševčík, 2009], [Morisset et al., 2013]	✓	✓	~
[Vafeiadis et al., 2015]	✓	✓	~
CAS/Concurrent CompCert	~	~	✓
CompCertTSO [Ševčík et al., 2013]	✗	✓	✓

- can only handle finite traces
- cannot handle potentially unbounded loops

[Passarotti et al., 2017]	✗	✓	✓
Transfinite Iris [Spies et al., 2021]	✗	✗	✓
Stacked Borrows [Jung et al., 2020]	✗	✗	~

# The problem

How can we prove correctness of **concurrent** program optimizations relying on **data race UB** and involving **loops**?

	data race UB	concurrent	loops
[Ševčík, 2009], [Morisset et al., 2013]	✓	✓	~
[Vafeiadis et al., 2015]	✓	✓	~
CAS/Concurrent CompCert	~	~	✓
CompCertTSO [Ševčík et al., 2013]	✗	✓	✓

- no optimizations involving synchronizing operations (e.g., atomic reads)

[Tassarotti et al., 2017]	✗	✓	✓
Transfinite Iris [Spies et al., 2021]	✗	✗	✓
Stacked Borrows [Jung et al., 2020]	✗	✗	~

# The problem

How can we prove correctness of **concurrent** program optimizations relying on **data race UB** and involving **loops**?

	<b>data race UB</b>	<b>concurrent</b>	<b>loops</b>
<b>Our approach</b>	✓	✓	✓
[Ševčík, 2009], [Morisset et al., 2013]	✓	✓	~
[Vafeiadis et al., 2015]	✓	✓	~
CAS/Concurrent CompCert	~	~	✓
CompCertTSO [Ševčík et al., 2013]	✗	✓	✓
CCAL (CompCertX) [Gu et al., 2018]	✗	~	✓
[Liang and Feng, 2016]	✗	✓	✓
ReLoC [Frumin et al., 2018]	✗	✓	✓
[Tassarotti et al., 2017]	✗	✓	✓
Transfinite Iris [Spies et al., 2021]	✗	✗	✓
Stacked Borrows [Jung et al., 2020]	✗	✗	~

# Key idea: ownership acquisition on unsynchronized accesses

7

unoptimized:

```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

{True}

$\rightsquigarrow$

optimized:

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```

# Key idea: ownership acquisition on unsynchronized accesses

7

unoptimized:

reach unsynchronized  
access to y



```
int i = 0; int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

{True}

$\approx$

optimized:

```
int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

# Key idea: ownership acquisition on unsynchronized accesses

7

unoptimized:

reach unsynchronized  
access to y

```
int i = 0; int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

{True}

$\{y \mapsto^{\text{src}} z_y * y \mapsto^{\text{tgt}} z_y\}$

$\Vdash$

optimized:

obtain ownership  
with proof rule

```
int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

# Key idea: ownership acquisition on unsynchronized accesses

7

unoptimized:

reach unsynchronized  
access to x

```
int i = 0; int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

{True}

$\{y \mapsto^{\text{src}} z_y * y \mapsto^{\text{tgt}} z_y\}$

$\approx$

optimized:

obtain ownership  
with proof rule

```
int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

# Key idea: ownership acquisition on unsynchronized accesses

7

unoptimized:

optimized:

reach unsynchronized  
access to x

obtain ownership  
with proof rule

{True}

$\{y \mapsto^{\text{src}} z_y * y \mapsto^{\text{tgt}} z_y * x \mapsto^{\text{src}} z_x * x \mapsto^{\text{tgt}} z_x\}$

```
int i = 0; int sum = *y;
while (i != *x - 1) {
    i += 1; sum += *y;
}
return sum;
```

$\approx$

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n - 1) {
    i += 1; sum += m;
}
return sum;
```



# Key idea: ownership acquisition on unsynchronized accesses

unoptimized:

```
int i = 0; int sum = *y;  
while (i != *x - 1) {  
    i += 1; sum += *y;  
}  
return sum;
```

{True}

$\{y \mapsto^{\text{src}} z_y * y \mapsto^{\text{tgt}} z_y * x \mapsto^{\text{src}} z_x * x \mapsto^{\text{tgt}} z_x\}$  ←

⌋

optimized:

obtain ownership  
with proof rule

```
int n = *x;  
int m = *y;  
int i = 0; int sum = m;  
while (i != n - 1) {  
    i += 1; sum += m;  
}  
return sum;
```

$\{y \mapsto^{\text{src}} z_y * y \mapsto^{\text{tgt}} z_y * x \mapsto^{\text{src}} z_x * x \mapsto^{\text{tgt}} z_x\}$  ←

retain ownership  
throughout the loop

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...

# Simuliris: a separation logic-based simulation framework

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris**: separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...

# Simuliris: a separation logic-based simulation framework

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...



fully mechanized in the Coq proof assistant  
based on the Iris framework



**Iris**



# Simuliris: a separation logic-based simulation framework

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...



fully mechanized in the Coq proof assistant  
based on the Iris framework



# Simuliris: a separation logic-based simulation framework

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...



fully mechanized in the Coq proof assistant  
based on the Iris framework



Iris



Key ingredient: a powerful simulation relation

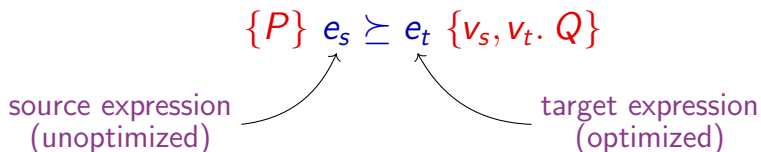
10

traditional coinductive simulation + modern separation logic

# Key ingredient: a powerful simulation relation

$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$

source expression (unoptimized)      target expression (optimized)



coinductive simulation

separation logic



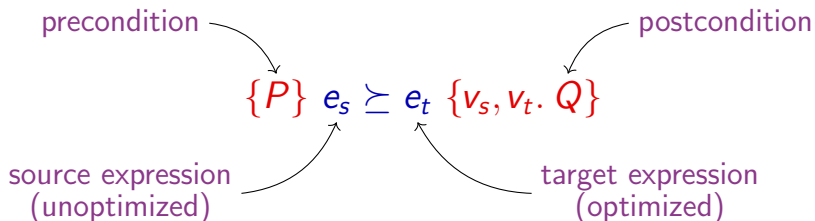
$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$

source expression (unoptimized)      target expression (optimized)

## coinductive simulation

- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

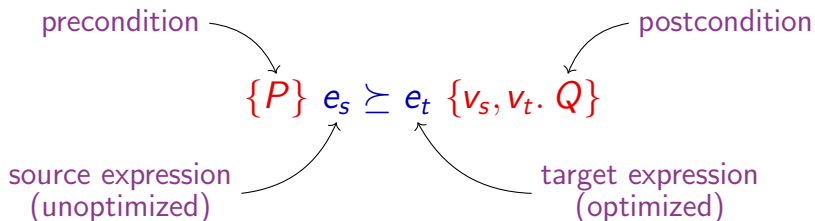


## coinductive simulation

- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

- compositional proof rules



## coinductive simulation

- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

- compositional proof rules
  - ownership reasoning
- with custom resources  $\boxed{\text{I}r/s^*}$

$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$

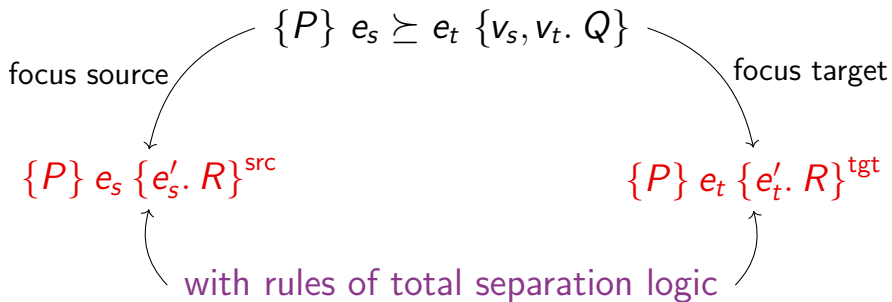
## coinductive simulation

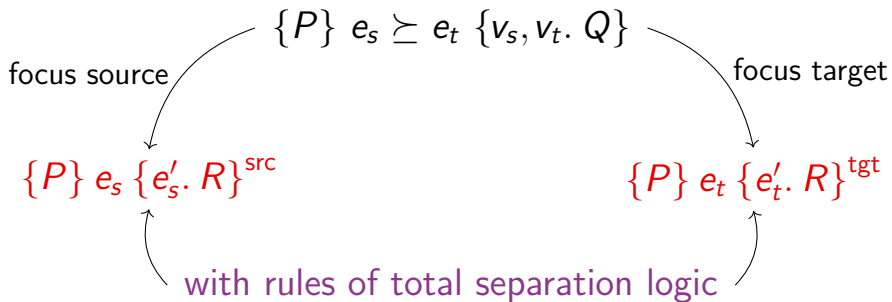
- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

- compositional proof rules
- ownership reasoning  
with custom resources

$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$





Enabled by a flexible **implicit stuttering** mechanism without explicit step counting!

$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$

## coinductive simulation

- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

- compositional proof rules
- ownership reasoning  
with custom resources  $\boxed{\text{I}r/s^*}$



# Ownership is useful for justifying optimizations

let x := new(42) in

call f ();

\*x

$\Downarrow$

let x := new(42) in

call f ();

42

# Ownership is useful for justifying optimizations

	$\{True\}$	
let x := new(42) in		let x := new(42) in
call f ();	$\Downarrow$	call f ();
$*x$		42
	$\{v_s, v_t. v_s = v_t = 42\}$	

# Ownership is useful for justifying optimizations

$$\begin{array}{ccc} & \{\text{True}\} & \\ \text{let } x := \text{new}(42) \text{ in} & & \text{let } x := \text{new}(42) \text{ in} \\ & \{x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42\} & \\ \text{call } f (); & \Downarrow & \text{call } f (); \\ *x & & 42 \\ & \{v_s, v_t. v_s = v_t = 42\} & \end{array}$$

# Ownership is useful for justifying optimizations

	$\{ \text{True} \}$	
let $x := \text{new}(42)$ in		let $x := \text{new}(42)$ in
	$\{ x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42 \}$	
call $f()$ ;	$\sqcap$	call $f()$ ;
	$\{ x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42 \}$	
$*_x$		42
	$\{ v_s, v_t. v_s = v_t = 42 \}$	

# Ownership is useful for justifying optimizations

	$\{ \text{True} \}$	
let $x := \text{new}(42)$ in		let $x := \text{new}(42)$ in
	$\{ x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42 \}$	
call $f()$ ;	$\sqsubseteq$	call $f()$ ;
	$\{ x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42 \}$	
$*_x$		42
	$\{ v_s, v_t. v_s = v_t = 42 \}$	

Unknown code must respect the ownership principles of our logic!

$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$

## coinductive simulation

- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

- compositional proof rules
- ownership reasoning  
with custom resources

	$\{True\}$	
let x := new(42) in		let x := new(42) in
while call f (*x) do	$\approx$	while call f (42) do
()		()
od		od
	$\{v_s, v_t \cdot v_s = v_t = ()\}$	

$$\begin{array}{ccc}
 & \{\text{True}\} & \\
 \text{let } x := \text{new}(42) \text{ in} & & \text{let } x := \text{new}(42) \text{ in} \\
 & \{x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42\} & \\
 \text{while call f } (*x) \text{ do} & \supseteq & \text{while call f } (42) \text{ do} \\
 \quad () & & \quad () \\
 \text{od} & & \text{od} \\
 & \{v_s, v_t \cdot v_s = v_t = ()\} &
 \end{array}$$



$W_s = \text{while } c_s \text{ do } e_s \text{ od}$        $W_t = \text{while } c_t \text{ do } e_t \text{ od}$

---

$\{I\} W_s \succeq W_t \{v_s, v_t. Q\}$

loop invariant

$$W_s = \text{while } c_s \text{ do } e_s \text{ od} \quad W_t = \text{while } c_t \text{ do } e_t \text{ od}$$

new proof goal

$$\{I\}$$

$$\text{if } c_s \text{ then } e_s; W_s \text{ else } () \succeq \text{if } c_t \text{ then } e_t; W_t \text{ else } ()$$

$$\{e'_s, e'_t. (\exists v_s, v_t. e'_s = v_s * e'_t = v_t * Q)\}$$


---


$$\{I\} W_s \succeq W_t \{v_s, v_t. Q\}$$

loop invariant

$$W_s = \text{while } c_s \text{ do } e_s \text{ od} \quad W_t = \text{while } c_t \text{ do } e_t \text{ od}$$

new proof goal

$$\{I\}$$

$$\text{if } c_s \text{ then } e_s; W_s \text{ else } () \succeq \text{if } c_t \text{ then } e_t; W_t \text{ else } ()$$

allows use of  
coinduction hypothesis

$$\{e'_s, e'_t. (\exists v_s, v_t. e'_s = v_s * e'_t = v_t * Q) \vee (e'_s = W_s * e'_t = W_t * I)\}$$


---


$$\{I\} W_s \succeq W_t \{v_s, v_t. Q\}$$

loop invariant

$$\begin{array}{ccc}
 & \{\text{True}\} & \\
 \text{let } x := \text{new}(42) \text{ in} & & \text{let } x := \text{new}(42) \text{ in} \\
 & \{x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42\} & \\
 \text{while call f } (*x) \text{ do} & \succeq & \text{while call f } (42) \text{ do} \\
 \quad () & & \quad () \\
 \text{od} & & \text{od} \\
 & \{v_s, v_t \cdot v_s = v_t = ()\} &
 \end{array}$$

Pick invariant  $I \triangleq x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42$

$\{ \text{True} \}$   
 $\text{let } x := \text{new}(42) \text{ in}$        $\text{let } x := \text{new}(42) \text{ in}$   
 $\{ x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42 \}$

Ownership reasoning is a powerful tool in combination  
with coinductive simulations!

$\{ v_s, v_t \cdot v_s = v_t = () \}$

Pick invariant  $I \triangleq x \mapsto^{\text{src}} 42 * x \mapsto^{\text{tgt}} 42$

$$\{P\} e_s \succeq e_t \{v_s, v_t. Q\}$$

## coinductive simulation

- coinduction
- reasoning about UB
- flexible stuttering

## separation logic

- compositional proof rules
- ownership reasoning  
with custom resources

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...



fully mechanized in the Coq proof assistant  
based on the Iris framework



# Optimizations with external locations: motivating example

```
fn foo(x) {  
  x ← 41;  
  x ← 42;  
  *x  
}  
  
≡  
  
fn foo_opt(x) {  
  x ← 42;  
  42  
}
```



# Optimizations with external locations: motivating example

in separation logic verification: assume ownership in precondition

<code>fn foo(x) {</code>	$\{x \mapsto^{\text{src}} z * x \mapsto^{\text{tgt}} z\}$	<code>fn foo_opt(x) {</code>
<code>  x ← 41;</code>		
<code>  x ← 42;</code>	$\Downarrow$	<code>  x ← 42;</code>
<code>  *x</code>		<code>  42</code>
<code>}</code>		<code>}</code>

in separation logic verification: assume ownership in precondition

<code>fn foo(x) {</code>	<code>{x ↦<sup>src</sup> z * x ↦<sup>tgt</sup> z}</code>	<code>fn foo_opt(x) {</code>
<code>  x ← 41;</code>		
<code>  x ← 42;</code>	<code>⊧</code>	<code>  x ← 42;</code>
<code>  *x</code>		<code>  42</code>
<code>}</code>		<code>}</code>

in compiler optimizations: surrounding code is **not cooperative!**

<code>fn foo(x) {</code>	<code>{??}</code>	<code>fn foo_opt(x) {</code>
<code>  x ← 41;</code>		
<code>  x ← 42;</code>	<code>⊧</code>	<code>  x ← 42;</code>
<code>  *x</code>		<code>  42</code>
<code>}</code>		<code>}</code>

# Optimizations with external locations: motivating example

in separation logic verification: assume ownership in precondition

<code>fn foo(x) {</code>	$\{x \mapsto^{\text{src}} z * x \mapsto^{\text{tgt}} z\}$	<code>fn foo_opt(x) {</code>
<code>  x ← 41;</code>		
<code>  x ← 42;</code>	$\supseteq$	<code>  x ← 42;</code>
<code>  *x</code>		<code>  42</code>
<code>}</code>		<code>}</code>

in compiler optimizations: surrounding code is **not cooperative!**

<code>fn foo(x) {</code>	$\{x_s \approx x_t\}$	<code>fn foo_opt(x) {</code>
<code>  x ← 41;</code>		
<code>  x ← 42;</code>	$\supseteq$	<code>  x ← 42;</code>
<code>  *x</code>		<code>  42</code>
<code>}</code>		<code>}</code>

**contract:** similar values  $v_s \approx v_t$  in source and target

for integers:  $z_s \approx z_t \triangleq z_s = z_t$

for memory locations  $l_s \approx l_t$ :

- **contract:** stored values are related by  $\approx$
- **accessible by anyone** as long as the contract is observed

**contract:** similar values  $v_s \approx v_t$  in source and target

for integers:  $z_s \approx z_t \triangleq z_s = z_t$

for memory locations  $l_s \approx l_t$ :

- **contract:** stored values are related by  $\approx$
- **accessible by anyone** as long as the contract is observed

How can we use this to justify optimizations?

**contract:** similar values  $v_s \approx v_t$  in source and target

Idea: we can break the contract as long as no other thread will notice

- **contract:** stored values are related by  $\approx$
- **accessible by anyone** as long as the contract is observed

How can we use this to justify optimizations?

When the source program does an unsynchronized access to  $l_s \approx l_t$ , we temporarily obtain ownership of  $l_s$  and  $l_t$ .

When the source program does an unsynchronized access to  $l_s \approx l_t$ , we temporarily obtain ownership of  $l_s$  and  $l_t$ .

An **unsynchronized write**  $l_s \leftarrow \_$  is reachable:  
 $\Rightarrow$  *all* concurrent accesses would be conflicting  
 $\Rightarrow$  obtain **exclusive** ownership  $l_s \mapsto^{\text{src}} v_s, l_t \mapsto^{\text{tgt}} v_t$



# Acquiring ownership on writes, formally: first attempt

---

$$\{l_s \approx l_t * P\} K[l_s \leftarrow v_0] \succeq e_t \{\Phi\}$$

public locations  $\curvearrowright$   $\{l_s \approx l_t * P\}$   $\curvearrowleft$   $K[l_s \leftarrow v_0]$  unsynchronized write  
in the source

# Acquiring ownership on writes, formally: first attempt

$$\frac{\forall v_s, v_t. \{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * P\} K[l_s \leftarrow v_0] \succeq e_t \{\Phi\}}{\{l_s \approx l_t * P\} K[l_s \leftarrow v_0] \succeq e_t \{\Phi\}}$$

obtain ownership

public locations

unsynchronized write  
in the source

# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$

$x_s \leftarrow 41;$

$x_s \leftarrow 42;$

$*x_s$

$\Vdash$

$x_t \leftarrow 42;$

42

# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$
$$\{x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 41;$

$x_s \leftarrow 42;$

$*x_s$

$\approx$

$x_t \leftarrow 42;$

42

# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$

$$\{x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 41;$

contract temporarily broken

$$\{x_s \mapsto^{\text{src}} 41 * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 42;$

$\Vdash$

$x_t \leftarrow 42;$

$*x_s$

42

# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$

$$\{x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 41;$

$$\{x_s \mapsto^{\text{src}} 41 * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 42;$

$\supseteq$

$x_t \leftarrow 42;$

$$\{x_s \mapsto^{\text{src}} 42 * x_t \mapsto^{\text{tgt}} 42\}$$

$*x_s$

42

# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$

$$\{x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 41;$

$$\{x_s \mapsto^{\text{src}} 41 * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 42;$

$\supseteq$

$x_t \leftarrow 42;$

$$\{x_s \mapsto^{\text{src}} 42 * x_t \mapsto^{\text{tgt}} 42\}$$

$*x_s$

42

$$\{v_s, v_t. v_s = v_t = 42 * x_s \approx x_t\}$$

# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$

$$\{x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$$\{x_s \mapsto^{\text{src}} z' * x_t \mapsto^{\text{tgt}} z' * x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 41;$

$$\{x_s \mapsto^{\text{src}} 41 * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 42;$

$\supseteq$

$x_t \leftarrow 42;$

$$\{x_s \mapsto^{\text{src}} 42 * x_t \mapsto^{\text{tgt}} 42\}$$

$*_{x_s}$

42

$$\{v_s, v_t. v_s = v_t = 42 * x_s \approx x_t\}$$

What prevents us from acquiring ownership multiple times?



# Proving the optimization with ownership

$$\{x_s \approx x_t\}$$

$$\{x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$$\{x_s \mapsto^{\text{src}} z' * x_t \mapsto^{\text{tgt}} z' * x_s \mapsto^{\text{src}} z * x_t \mapsto^{\text{tgt}} z\}$$

$x_s \leftarrow 41;$

$$\{x_s \mapsto^{\text{src}} 41 * x_t \mapsto^{\text{tgt}} z\}$$

The acquisition rule is unsound! ☹️

$$\{x_s \mapsto^{\text{src}} 42 * x_t \mapsto^{\text{tgt}} 42\}$$

$*x_s$

42

$$\{v_s, v_t. v_s = v_t = 42 * x_s \approx x_t\}$$

What prevents us from acquiring ownership multiple times?

# Solution: avoiding duplication of ownership

Track locations exploited by the current thread  $\pi$ :  $\text{exploit}_\pi C$

obtain ownership

$$\forall v_t, v_s. \{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t *$$

$$\{l_s \approx l_t *$$

public locations

$$P \} K[l_s \leftarrow v_0] \succeq_\pi e_t \{ \Phi \}$$

$$P \} K[l_s \leftarrow v_0] \succeq_\pi e_t \{ \Phi \}$$

unsynchronized write  
in the source

# Solution: avoiding duplication of ownership

Track locations exploited by the current thread  $\pi$ :  $\text{exploit}_\pi C$

$$\frac{\forall v_t, v_s. \{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * \text{exploit}_\pi(C, l_s \mapsto W) * P\} K[l_s \leftarrow v_0] \succeq_\pi e_t \{\Phi\}}{\{l_s \approx l_t * \text{exploit}_\pi C * P\} K[l_s \leftarrow v_0] \succeq_\pi e_t \{\Phi\}}$$

Diagram annotations:
 

- obtain ownership**: points to the source location  $l_s$  in the premise.
- exploit once**: points to the  $\text{exploit}_\pi$  operator in the premise.
- $l_s \notin C$ : a red annotation above the  $\text{exploit}_\pi$  operator in the premise.
- remember  $l_s$** : points to the source location  $l_s$  in the conclusion.
- public locations**: points to the source location  $l_s$  in the conclusion.
- track exploited locations for current thread**: points to the  $\text{exploit}_\pi$  operator in the conclusion.
- unsynchronized write in the source**: points to the  $W$  in the premise.

# Solution: avoiding duplication of ownership

Track locations exploited by the current thread  $\pi$ :  $\text{exploit}_\pi C$

defined with custom  
Iris ghost state

$$\frac{\forall v_t, v_s. \{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * \text{exploit}_\pi (C, l_s \mapsto W) * P\} K[l_s \leftarrow v_0] \succeq_\pi e_t \{\Phi\}}{\{l_s \approx l_t * \text{exploit}_\pi C * P\} K[l_s \leftarrow v_0] \succeq_\pi e_t \{\Phi\}}$$

obtain ownership  $\curvearrowright$   $\{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * \text{exploit}_\pi (C, l_s \mapsto W) * P\}$   
 exploit once  $\curvearrowright$   $l_s \notin C$   
 remember  $l_s$   $\curvearrowright$   $l_s \mapsto W$   
 public locations  $\curvearrowright$   $\{l_s \approx l_t * \text{exploit}_\pi C * P\}$   
 track exploited locations for current thread  $\uparrow$   $\text{exploit}_\pi C$   
 unsynchronized write in the source  $\curvearrowright$   $K[l_s \leftarrow v_0]$

# Solution: avoiding duplication of ownership

Track locations exploited by the current thread  $\pi$ :  $\text{exploit}_\pi C$

defined with custom

generalization: unsynchronized access just needs to be  
**reachable** (not the directly next instruction)

obtain ownership

$$\frac{\forall V_t, V_s. \{l_s \mapsto^{\text{src}} V_s * l_t \mapsto^{\text{tgt}} V_t * V_s \approx V_t * \text{exploit}_\pi (C, l_s \mapsto W) * P\} K[l_s \leftarrow v_0] \succeq_\pi e_t \{\Phi\}}{\{l_s \approx l_t * \text{exploit}_\pi C * P\} K[l_s \leftarrow v_0] \succeq_\pi e_t \{\Phi\}}$$

public locations

track exploited locations  
for current thread

unsynchronized write  
in the source

remember  $l_s$

$l_s \notin C$

# Solution: avoiding duplication of ownership

Track locations exploited by the current thread  $\pi$ :  $\text{exploit}_\pi C$

defined with custom

generalization: unsynchronized access just needs to be  
**reachable** (not the directly next instruction)

obtain ownership

$\forall v_t, v_s. \{$

rule for **reads**: obtain **fractional (read-only) ownership**

$e_t \{ \Phi \}$

public locations

track exploited locations  
for current thread

unsynchronized write  
in the source

# Maintaining & releasing ownership on synchronization

We can maintain ownership until the thread  
**observably** synchronizes.

action (potentially) visible  
by other threads 

We can maintain ownership until the thread  
**observably** synchronizes.

action (potentially) visible  
by other threads 

Example: rule for atomic writes

$$\{l_s \approx l_t * v_s \approx v_t * \text{exploit}_{\pi} \emptyset\}$$

$$l_s \leftarrow^{sc} v_s \succeq_{\pi} l_t \leftarrow^{sc} v_t$$

$$\{v'_s, v'_t. \text{exploit}_{\pi} \emptyset\}$$



# Verifying the motivating example

```
int i = 0; int sum = *y;
while (i != *x) {
    i += 1; sum += *y;
}
return sum;
```

```
int n = *x;
int m = *y;
int i = 0; int sum = m;
while (i != n) {
    i += 1; sum += m;
}
return sum;
```

# Verifying the motivating example

$$\{x_s \approx x_t * y_s \approx y_t\}$$

```
let (i, sum) := (new(0), new(*y)) in
while *i ≠ *x do
  i ← *i + 1;
  sum ← *sum + *y
od; *sum
```

$$\supseteq_{\pi}$$

```
let (m, n) := (*y, *x) in
let (i, sum) := (new(0), new(m)) in
while *i ≠ n do
  i ← *i + 1;
  sum ← *sum + m
od; *sum
```

$$\{v_s, v_t. v_s \approx v_t\}$$

1. Obtain ownership of  $x$  and  $y$  due to unsynchronized reads in the source
2. Initiate coinduction

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...



fully mechanized in the Coq proof assistant  
based on the Iris framework



Soundness: fair termination-preserving contextual refinement

30

fair termination-preserving contextual refinement

# Soundness: fair termination-preserving contextual refinement

30

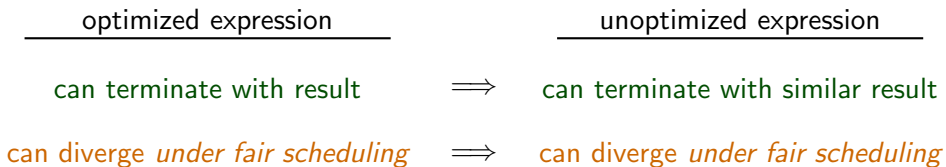
fair termination-preserving contextual refinement

optimized expression  $\Rightarrow$  unoptimized expression  
can terminate with result can terminate with similar result

# Soundness: fair termination-preserving contextual refinement

30

fair termination-preserving contextual refinement



## fair termination-preserving contextual refinement

for any surrounding program, assuming no UB in unoptimized program

optimized expression

unoptimized expression

can terminate with result

$\implies$

can terminate with similar result

can diverge *under fair scheduling*

$\implies$

can diverge *under fair scheduling*

## fair termination-preserving contextual refinement

for any surrounding program, assuming no UB in unoptimized program

optimized expression

unoptimized expression

can terminate with result

$\implies$

can terminate with similar result

can diverge *under fair scheduling*

$\implies$

can diverge *under fair scheduling*

Core soundness proof: proved once and for all!



# More in the paper . . .

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris**: separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, . . .



fully mechanized in the Coq proof assistant  
based on the Iris framework



Iris



# Simuliris: a separation logic-based simulation framework

logic for data race  
based optimizations

Stacked Borrows for Rust  
[Jung et al., 2020] + concurrency

**Simuliris:** separation logic-based simulation framework

- soundness: fair termination-preserving contextual refinement
- proof rules for verifying optimizations: coinduction, ...



fully mechanized in the Coq proof assistant  
based on the Iris framework



<https://gitlab.mpi-sws.org/iris/simuliris>



Thanks for listening!

# Generalization: reachability of an unsynchronized write

The argument even works if the write is just  
**(unconditionally) reachable!**

$$\frac{\forall v_t, v_s. \{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * \text{exploit}_{\pi} (C, l_s \mapsto W) * P\} \quad e_s \succeq_{\pi} e_t \{\Phi\}}{\{l_s \approx l_t * \text{exploit}_{\pi} C * P\} \quad e_s \succeq_{\pi} e_t \{\Phi\}}$$

Diagram annotations:

- exploit once (points to  $l_s \notin C$ )
- reach unsynchronized write in the source (points to  $e_s \rightarrow_{?}^* K[l_s \leftarrow v_0]$ )
- obtain ownership (points to  $l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s$ )
- public locations (points to  $l_s \approx l_t * \text{exploit}_{\pi} C * P$ )
- track exploited locations for current thread (points to  $\text{exploit}_{\pi} C * P$ )

# Generalization: reachability of an unsynchronized write

The argument even works if the write is just  
**(unconditionally) reachable!**

$$\frac{\begin{array}{l} \text{exploit once} \quad \text{reach unsynchronized write} \\ \text{obtain ownership} \quad \text{in the source} \\ \forall v_t, v_s. \{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * \text{exploit}_\pi (C, l_s \mapsto W) * P\} \quad e_s \succeq_\pi e_t \{\Phi\} \\ l_s \notin C \quad e_s \rightarrow^*_? K[l_s \leftarrow v_0] \end{array}}{\{l_s \approx l_t * \text{exploit}_\pi C * P\} \quad e_s \succeq_\pi e_t \{\Phi\}}$$

public locations      track exploited locations for current thread

A similar rule holds for reads!

When the source program does an unsynchronized access to  $l_s \approx l_t$ , we temporarily obtain ownership of  $l_s$  and  $l_t$ .

An unsynchronized read  $*l_s$  is reachable:

⇒ concurrent *write* accesses would be conflicting

⇒ obtain **fractional** ownership  $l_s \mapsto_q^{\text{src}} v_s, l_t \mapsto_q^{\text{tgt}} v_t$

When the source program does an unsynchronized access to  $l_s \approx l_t$ , we temporarily obtain ownership of  $l_s$  and  $l_t$ .

Do we ever have to give up ownership again?

⇒ concurrent *write* accesses would be conflicting

⇒ obtain **fractional** ownership  $l_s \mapsto_q^{\text{src}} V_s, l_t \mapsto_q^{\text{tgt}} V_t$

$$\frac{
 \begin{array}{l}
 \text{obtain ownership} \quad \text{exploit once} \\
 \forall v_t, v_s, q. \{ l_s \mapsto_q^{\text{src}} v_s * l_t \mapsto_q^{\text{tgt}} v_t * v_s \approx v_t * \text{exploit}_{\pi} (C, l_s \mapsto R(q)) * P \} e_s \succeq_{\pi} e_t \{ \Phi \} \\
 \text{reach unsynchronized read} \\
 \text{in the source} \\
 e_s \rightarrow^* K[*l_s]
 \end{array}
 }{
 \begin{array}{l}
 \text{public locations} \quad \text{track exploited locations} \\
 \{ l_s \approx l_t * \text{exploit}_{\pi} C * P \} e_s \succeq_{\pi} e_t \{ \Phi \} \\
 \text{for current thread}
 \end{array}
 }$$

$$\frac{C(l_s) = W \quad \{\text{exploit}_\pi (C \setminus l_s) * P\} e_s \succeq_\pi e_t \{\Phi\}}{\{l_s \mapsto^{\text{src}} v_s * l_t \mapsto^{\text{tgt}} v_t * v_s \approx v_t * l_s \approx l_t * \text{exploit}_\pi C * P\} e_s \succeq_\pi e_t \{\Phi\}}$$



The compiler should not be allowed to perform the following transformation:

<pre>while !lock(l) do   () od; unlock(l)</pre>	$\parallel$	<pre>if lock(l) then unlock(l) else ()</pre>	$\approx$	<pre>while true do   () od</pre>	$\parallel$	<pre>if lock(l) then unlock(l) else ()</pre>
---	-------------	--	-----------	----------------------------------	-------------	--

The source program only has diverging executions under an unfair scheduler, while the target program diverges even under a fair scheduler.

# In the paper: proofs of optimizations relying on data races

- optimization hoisting reads out of a while loop

```
let (i, sum) := (new(0), new(*ys)) in
while *i ≠ *xs do
  i ← *i + 1; sum ← *sum + *ys
od; *sum
```

 $\Downarrow_{\pi}$ 

```
let (n, m) := (*xt, *yt) in
let (i, sum) := (new(0), new(m)) in
while *i ≠ n do
  i ← *i + 1; sum ← *sum + m
od; *sum
```

Requires reasoning about potentially infinite loops!

# In the paper: proofs of optimizations relying on data races

- optimization hoisting reads out of a while loop
- optimization eliminating reads and writes over unknown read-only code

$$\begin{array}{ccc}
 x_s \leftarrow 42; & & x_t \leftarrow 42; \\
 e^{\text{RO}}; & \succeq_{\pi} & e^{\text{RO}}; \\
 *x_s & & 42
 \end{array}$$

# In the paper: proofs of optimizations relying on data races

- optimization hoisting reads out of a while loop
- optimization eliminating reads and writes over unknown read-only code

$$\begin{array}{ccc}
 x_s \leftarrow 42; & & x_t \leftarrow 42; \\
 *^{sc}y_s; & \succeq_{\pi} & *^{sc}y_t; \\
 *x_s & & 42
 \end{array}$$

Not supported by CAS/Concurrent CompCert!

# In the paper: proofs of optimizations relying on data races

- optimization hoisting reads out of a while loop
- optimization eliminating reads and writes over unknown read-only code
- eliminations and reorderings using data races by [Ševčík, 2009]

**Stacked Borrows:** an experimental aliasing model for Rust

- determines which kinds of memory accesses are allowed
- enables powerful optimizations



**Stacked Borrows:** an experimental aliasing model for Rust

- determines which kinds of memory accesses are allowed
- enables powerful optimizations



Using Simuliris, we have...

- extended the optimization proofs by [Jung et al., 2020] to concurrent environments
- developed a new proof of an optimization involving loops:

```
// x: &i32, g: &Fn() -> (),  
// f: &Fn(i32) -> bool  
while f(*x) {  
    g();  
}
```

 $\rightsquigarrow$ 

```
let r = *x;  
while f(r) {  
    g();  
}
```