

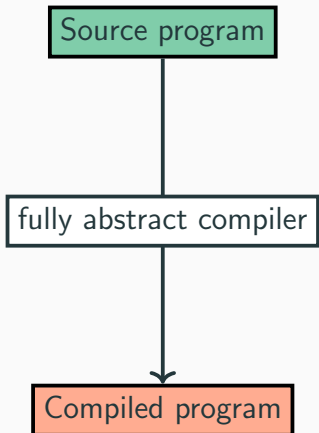
Linking Types: Secure compilation of multi-language programs

Daniel Patterson and Amal Ahmed

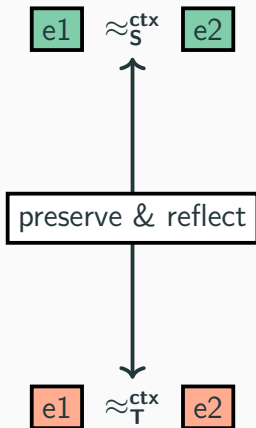
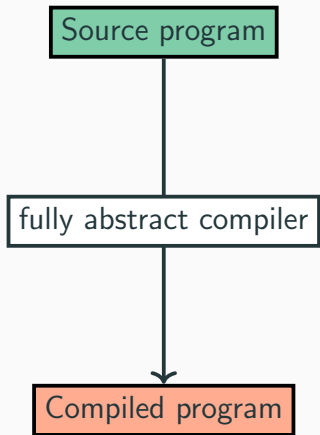
January 15, 2017

Northeastern University

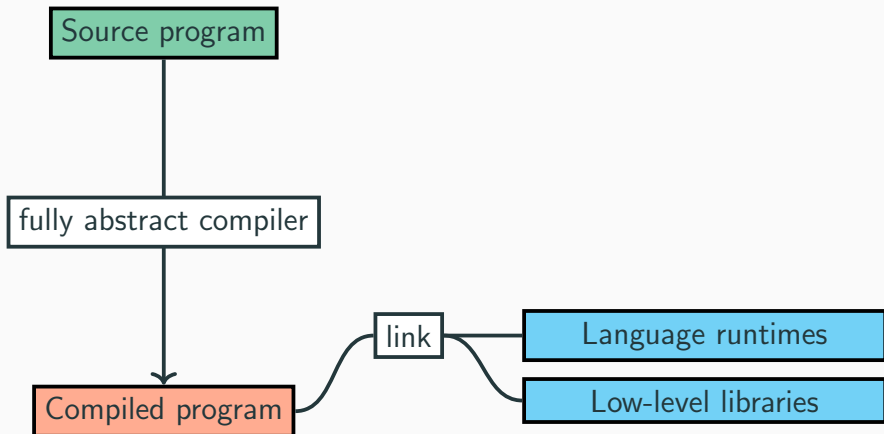
Fully abstract compilers



Fully abstract compilers

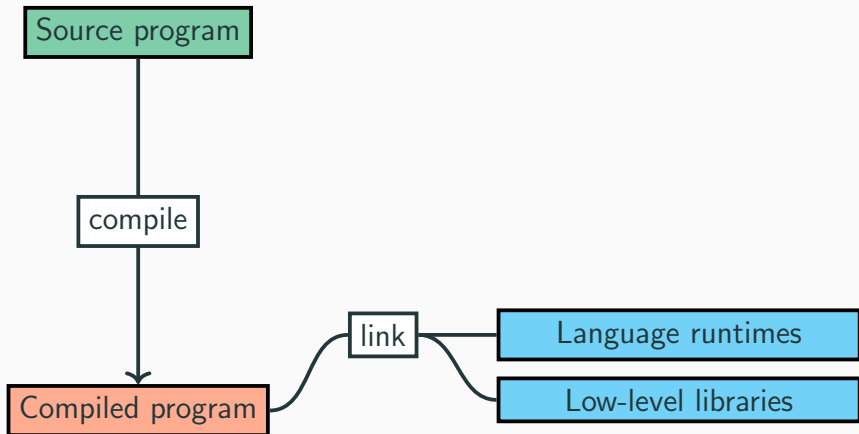


Fully abstract compilers in the real world

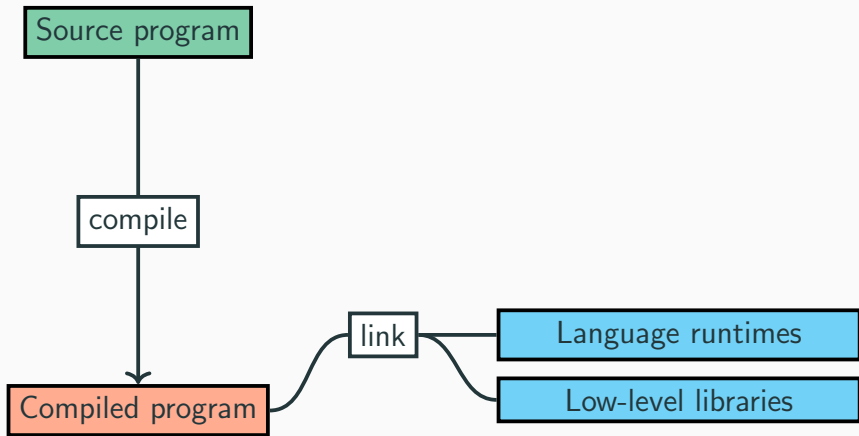


How do we implement fully
abstract compilers for
multi-language programs?

Fully abstract compilers in the real world

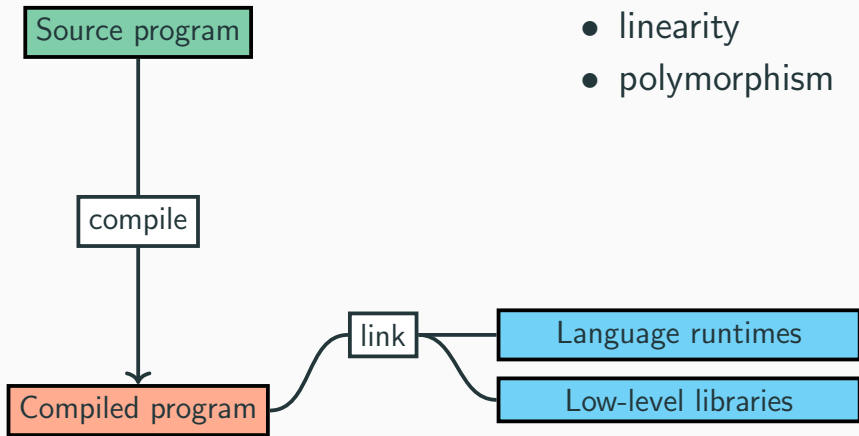


Fully abstract compilers in the real world



- references
- first-class control

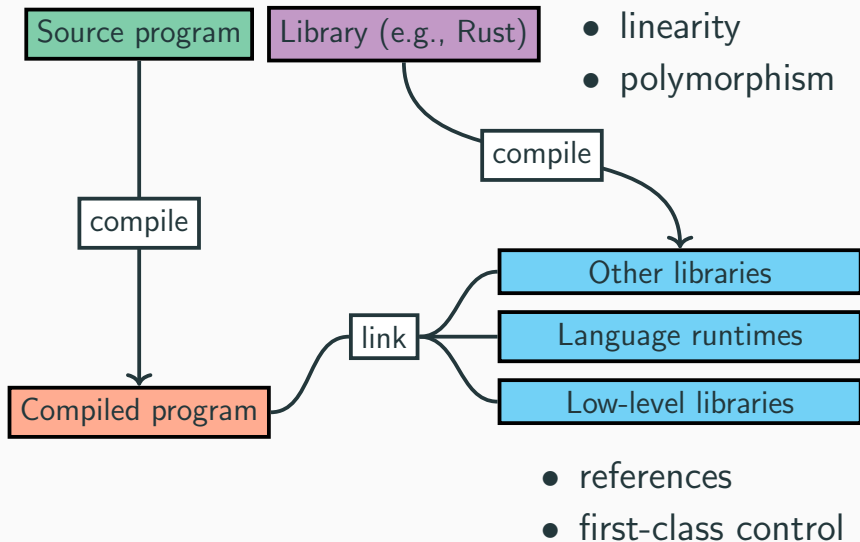
Fully abstract compilers in the real world



- linearity
- polymorphism

- references
- first-class control

Fully abstract compilers in the real world



Without a full accounting for **linking**, full abstraction—and secure compilation—is doomed.

Full abstraction: by example

e.g. a pure language λ linking with a counter library

$\lambda c. c()$

Full abstraction: by example

e.g. a pure language λ linking with a counter library

$$\lambda c. c() \approx_{\lambda}^{ctx} \lambda c. c(); c(): (\text{unit} \rightarrow \text{int}) \rightarrow \text{int}$$

Full abstraction: by example

e.g. a pure language λ linking with a counter library

$$\lambda c. c() \approx_{\lambda}^{ctx} \lambda c. c(); c(): (\text{unit} \rightarrow \text{int}) \rightarrow \text{int}$$

```
Ccounter = let v = ref 0 in  
           let c' () = v := !v + 1; !v in  
           [.] c'
```

Full abstraction: by example

e.g. a pure language λ linking with a counter library

$$\lambda c. c() \approx_{\lambda}^{ctx} \lambda c. c(); c(): (\text{unit} \rightarrow \text{int}) \rightarrow \text{int}$$

```
Ccounter = let v = ref 0 in
             let c' () = v := !v + 1; !v in
             [·] c'
```

$$C^{\text{counter}}[\lambda c. c()] \Downarrow 1 \quad C^{\text{counter}}[\lambda c. c(); c()] \Downarrow 2$$

Full abstraction: by example

e.g. a pure language λ linking with a counter library

$$\lambda c. c() \approx_{\lambda}^{ctx} \lambda c. c(); c(): (\text{unit} \rightarrow \text{int}) \rightarrow \text{int}$$

```
Ccounter = let v = ref 0 in
           let c' () = v := !v + 1; !v in
           [·] c'
```

$$C^{\text{counter}}[\lambda c. c()] \Downarrow 1 \quad C^{\text{counter}}[\lambda c. c(); c()] \Downarrow 2$$

$$c': \text{unit} \rightarrow \text{int} \quad \neq \quad c: \text{unit} \rightarrow \text{int}$$

Full abstraction: by example

e.g. a pure language λ linking with a counter library

$\lambda c. c() \not\approx_{\lambda}^{ctx} \lambda c. c(); c(): (\text{unit} \rightarrow \text{int}) \rightarrow \text{int}$

$C^{\text{counter}} = \text{let } v = \text{ref } 0 \text{ in}$
 $\quad \text{let } c' () = v := !v + 1; !v \text{ in}$
 $\quad [\cdot] c'$

$C^{\text{counter}}[\lambda c. c()] \Downarrow 1 \quad C^{\text{counter}}[\lambda c. c(); c()] \Downarrow 2$

$c': \text{unit} \rightarrow \text{int} \neq c: \text{unit} \rightarrow \text{int}$

Writing a type which
corresponds to behavior
inexpressible in your language is
the essence of **linking types**.

Linking types: in more detail

Two source languages: λ and λ^{ref} .

$$\begin{array}{ll} \tau ::= \text{unit} \mid \text{int} \mid \tau \rightarrow \tau & \tau ::= \dots \mid \text{ref } \tau \\ e ::= () \mid n \mid x \mid \lambda x : \tau. e & e ::= \dots \mid \text{ref } e \mid e := e \mid !e \\ & \mid ee \mid e + e \mid e * e \\ v ::= () \mid n \mid \lambda x : \tau. e & v ::= \dots \mid \ell \end{array}$$

Compiled to a shared, typed, target (not shown).

Linking types: extended language λ^{κ}

Extend λ to λ^{κ} :

$$\begin{aligned}\tau &::= \text{unit} \mid \text{int} \mid \text{ref } \tau \mid \tau \rightarrow \mathbb{R}^{\epsilon} \tau \\ e &::= () \mid \mathbf{n} \mid \mathbf{x} \mid \lambda \mathbf{x} : \tau. e \mid e e \mid e + e \\ &\quad \mid e * e \\ v &::= () \mid \mathbf{n} \mid \lambda \mathbf{x} : \tau. e \\ \epsilon &::= \bullet \mid \circ\end{aligned}$$

Linking types: extended language λ^{κ}

Extend λ **types** to λ^{κ} :

$$\begin{aligned}\tau &::= \text{unit} \mid \text{int} \mid \text{ref } \tau \mid \tau \rightarrow \mathbf{R}^{\epsilon} \tau \\ e &::= () \mid \mathbf{n} \mid \mathbf{x} \mid \lambda \mathbf{x} : \tau. e \mid e e \mid e + e \\ &\quad \mid e * e \\ v &::= () \mid \mathbf{n} \mid \lambda \mathbf{x} : \tau. e \\ \epsilon &::= \bullet \mid \circ\end{aligned}$$

- $\mathbf{R}^{\circ} \tau$ is a pure computation.
- $\mathbf{R}^{\bullet} \tau$ may allocate, read, or write references.

Linking types: extended language λ^{κ}

Extend λ **terms** to λ^{κ} :

$$\begin{aligned}\tau &::= \text{unit} \mid \text{int} \mid \text{ref } \tau \mid \tau \rightarrow \mathbb{R}^{\epsilon} \tau \\ e &::= () \mid \text{n} \mid \text{x} \mid \lambda \text{x} : \tau. e \mid e e \mid e + e \\ &\quad \mid e * e \mid \text{ref } e \mid e := e \mid !e \\ v &::= () \mid \text{n} \mid \lambda \text{x} : \tau. e \mid \ell \\ \epsilon &::= \bullet \mid \circ\end{aligned}$$

- Add *representative* terms for new behavior.
- Programmers use these to *reason* about inhabitants of type $\mathbb{R}^{\bullet} \tau$.

Linking types: using them

Consider the following λ programs

program 1 $\lambda f : \text{int} \rightarrow \text{int}. f 0$

program 2 $\lambda f : \text{int} \rightarrow \text{int}. f 0; f 0$

In λ^{κ} , by “default” f has type: $\text{int} \rightarrow \mathbb{R}^{\circ} \text{int}$

Linking types: using them

Consider the following λ programs

```
program 1  $\lambda f : \text{int} \rightarrow \text{int}. f 0$ 
```

```
program 2  $\lambda f : \text{int} \rightarrow \text{int}. f 0; f 0$ 
```

In λ^{κ} , by “default” f has type: $\text{int} \rightarrow R^{\circ} \text{int}$

But a programmer can annotate:

```
program 1  $\lambda f : \text{int} \rightarrow R^{\bullet} \text{int}. f 0$ 
```

```
program 2  $\lambda f : \text{int} \rightarrow R^{\bullet} \text{int}. f 0; f 0$ 
```

Linking types: default embedding κ^+

Property 1: κ^+ is Equivalence Preserving & Reflecting

Linking types: default embedding κ^+

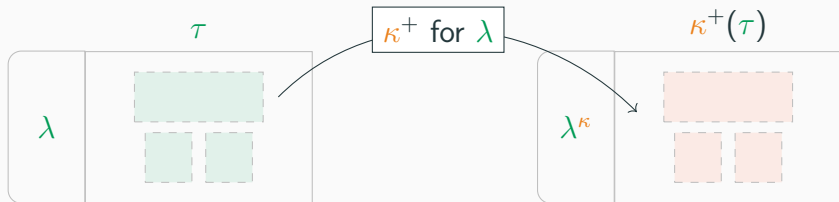
Property 1: κ^+ is Equivalence Preserving & Reflecting

$$\forall e_1, e_2. e_1 \approx_{\lambda}^{ctx} e_2 : \tau \implies e_1 \approx_{\lambda \kappa}^{ctx} e_2 : \kappa^+(\tau).$$

Linking types: default embedding κ^+

Property 1: κ^+ is Equivalence Preserving & Reflecting

$$\forall e_1, e_2. e_1 \approx_{\lambda}^{ctx} e_2 : \tau \implies e_1 \approx_{\lambda \kappa}^{ctx} e_2 : \kappa^+(\tau).$$



Linking types: default embedding κ^+

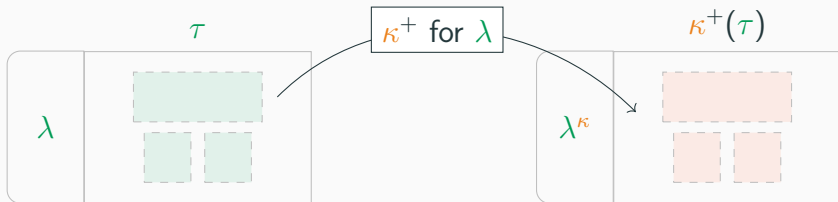
Property 1: κ^+ is Equivalence Preserving & Reflecting

$$\forall e_1, e_2. e_1 \approx_{\lambda}^{ctx} e_2 : \tau \implies e_1 \approx_{\lambda^{\kappa}}^{ctx} e_2 : \kappa^+(\tau).$$

$$\kappa^+(\text{unit}) = \text{unit}$$

$$\kappa^+(\text{int}) = \text{int}$$

$$\kappa^+(\tau_1 \rightarrow \tau_2) = \kappa^+(\tau_1) \rightarrow \mathbb{R}^{\circ} \kappa^+(\tau_2)$$



Linking types: project κ^-

Property 2: No New Programs.

Linking types: project κ^-

Property 2: No New Programs.

For e made of λ terms: $\forall \tau. e : \tau \implies e : \kappa^-(\tau)$

Linking types: project κ^-

Property 2: No New Programs.

For e made of λ terms: $\forall \tau. e : \tau \implies e : \kappa^-(\tau)$

$$\kappa^-(\text{unit}) = \text{unit}$$

$$\kappa^-(\text{int}) = \text{int}$$

$$\kappa^-(\text{ref } \tau) = \kappa^-(\tau)$$

$$\kappa^-(\tau_1 \rightarrow \text{R}^e \tau_2) = \kappa^-(\tau_1) \rightarrow \kappa^-(\tau_2)$$

Linking types: project κ^-

Property 2: No New Programs.

For e made of λ terms: $\forall \tau. e : \tau \implies e : \kappa^-(\tau)$

$$\kappa^-(\text{unit}) = \text{unit}$$

$$\kappa^-(\text{int}) = \text{int}$$

$$\kappa^-(\text{ref } \tau) = \kappa^-(\tau)$$

$$\kappa^-(\tau_1 \rightarrow \mathbb{R}^e \tau_2) = \kappa^-(\tau_1) \rightarrow \kappa^-(\tau_2)$$

e.g.,

$$\lambda(r : \text{ref int})(f : \text{ref int} \rightarrow \mathbb{R}^e \text{int}). f r$$

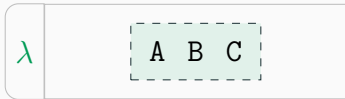
$$\xRightarrow{\kappa^-} \lambda(r : \text{int})(f : \text{int} \rightarrow \text{int}). f r$$

More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$



$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$

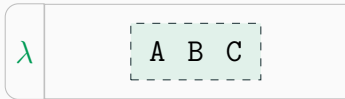
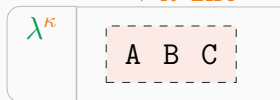
More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

$(\text{int} \rightarrow \mathbb{R}^\circ \text{int})$
 $\rightarrow \mathbb{R}^\circ \text{int}$



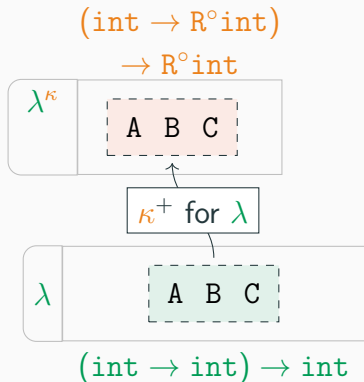
$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$

More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

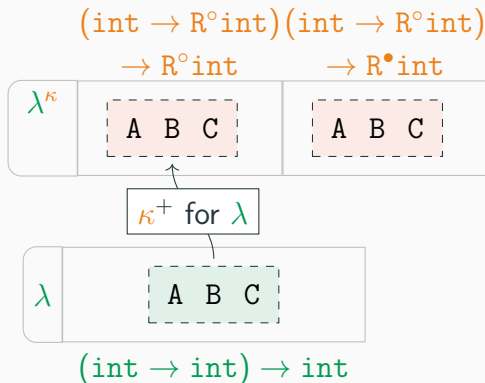


More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$



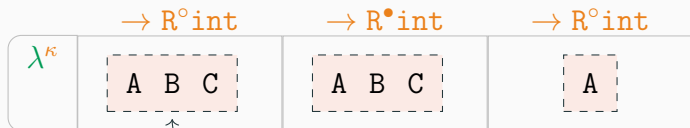
More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

$(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\circ \text{int})$



$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$

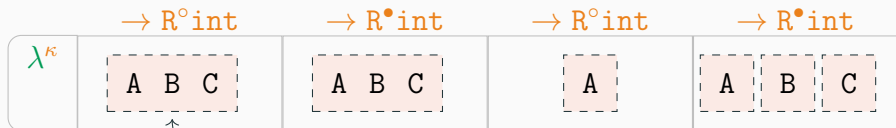
More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

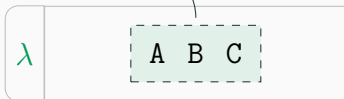
program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

$(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\bullet \text{int})(\text{int} \rightarrow R^\bullet \text{int})$



κ^+ for λ



$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$

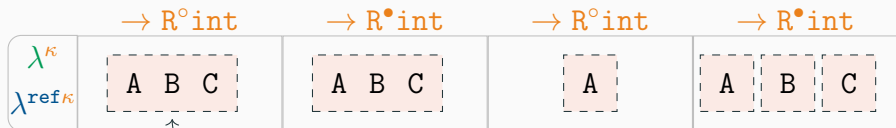
More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

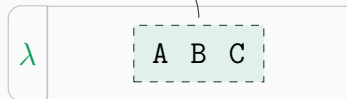
program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

$(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\bullet \text{int})(\text{int} \rightarrow R^\bullet \text{int})$



κ^+ for λ



$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$

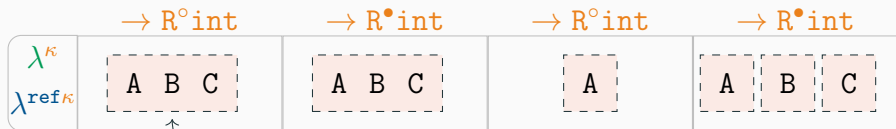
More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

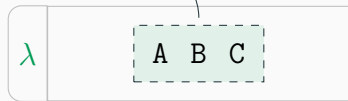
program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

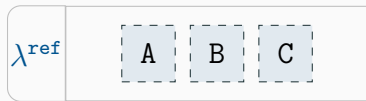
$(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\bullet \text{int})(\text{int} \rightarrow R^\bullet \text{int})$



κ^+ for λ



$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$



$(\text{int} \rightarrow \text{int}) \rightarrow \text{int}$

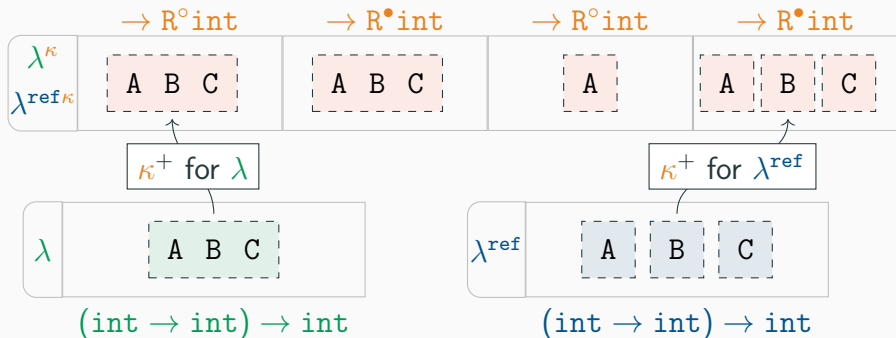
More about equivalences

program A – $\lambda f : \text{int} \rightarrow \text{int}. 1$

program B – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; 1$

program C – $\lambda f : \text{int} \rightarrow \text{int}. f\ 0; f\ 0; 1$

$(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\circ \text{int})(\text{int} \rightarrow R^\bullet \text{int})(\text{int} \rightarrow R^\bullet \text{int})$



Linking types: back to example

$$\lambda c: \text{unit} \rightarrow \text{int}. c() \approx_{\lambda^{\kappa}}^{\text{ctx}} \lambda c: \text{unit} \rightarrow \text{int}. c(); c()$$

Linking types: back to example

$$\lambda c: \text{unit} \rightarrow \mathbb{R}^\circ \text{int}. c() \approx_{\lambda_{\kappa}^{\text{ctx}}} \lambda c: \text{unit} \rightarrow \mathbb{R}^\circ \text{int}. c(); c()$$

Linking types: back to example

$\lambda c: \text{unit} \rightarrow \mathbb{R}^\circ \text{int}. c() \approx_{\lambda^{\text{ctx}}} \lambda c: \text{unit} \rightarrow \mathbb{R}^\circ \text{int}. c(); c()$

$c^{\text{counter}} = \text{let } v = \text{ref } 0 \text{ in}$
 $\text{let } c' () = v := !v + 1; !v \text{ in}$
 $[\cdot] c'$

Linking types: back to example

$\lambda c: \text{unit} \rightarrow \mathbb{R}^\circ \text{int}. c() \approx_{\lambda^{\text{ctx}}} \lambda c: \text{unit} \rightarrow \mathbb{R}^\circ \text{int}. c(); c()$

```
ccounter = let v = ref 0 in
           let c' () = v := !v + 1; !v in
           [·] c'
```

$c': \text{unit} \rightarrow \mathbb{R}^\circ \text{int}$

Linking types: back to example

$\lambda c: \text{unit} \rightarrow R^{\bullet} \text{int}. c() \not\approx_{\lambda^{\text{ctx}}} \lambda c: \text{unit} \rightarrow R^{\bullet} \text{int}. c(); c()$

```
ccounter = let v = ref 0 in
           let c' () = v := !v + 1; !v in
           [·] c'
```

$c': \text{unit} \rightarrow R^{\bullet} \text{int}$

Linking types: back to example

$\lambda c: \text{unit} \rightarrow R^{\bullet} \text{int}. c() \not\approx_{\lambda^{\kappa} \text{ctx}} \lambda c: \text{unit} \rightarrow R^{\bullet} \text{int}. c(); c()$

```
ccounter = let v = ref 0 in
           let c' () = v := !v + 1; !v in
           [·] c'
```

$c': \text{unit} \rightarrow R^{\bullet} \text{int}$

$c^{\text{counter}}[\lambda c. c()] \Downarrow 1$ $c^{\text{counter}}[\lambda c. c(); c()] \Downarrow 2$

With **linking types**, the programmer specifies the equivalences she wants, in a language she can understand.

Linking types: next steps

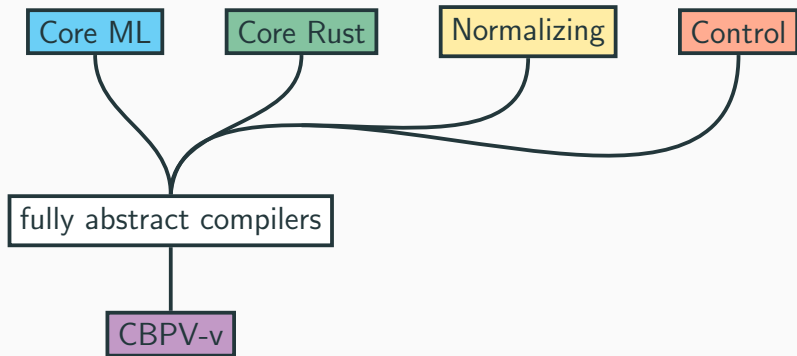
Core ML

Core Rust

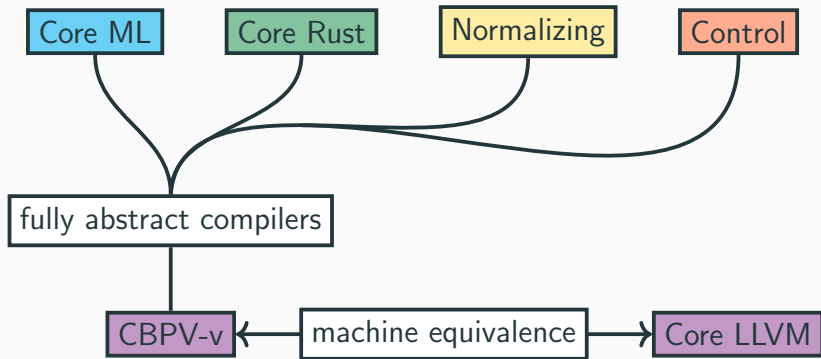
Normalizing

Control

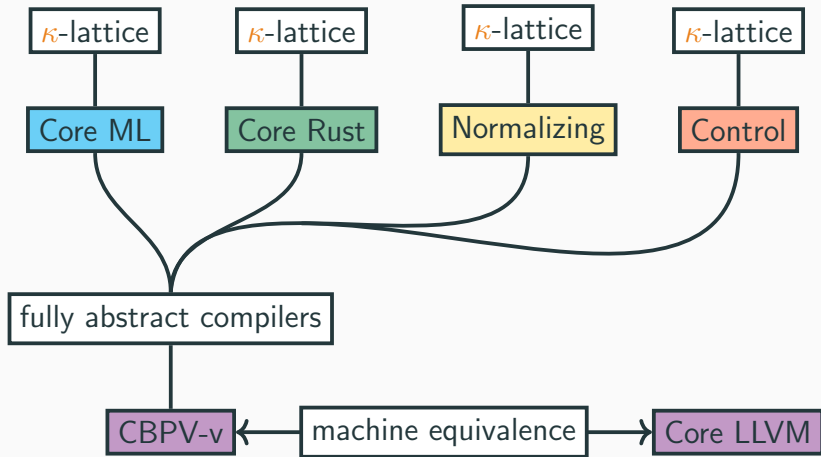
Linking types: next steps



Linking types: next steps



Linking types: next steps



Linking types make realizing fully abstract compilation possible in multi-language systems

Learn more:

<https://dbp.io/pubs/2017/linking-types-snapl-submission.pdf>

Linking types make realizing fully abstract compilation possible in multi-language systems (which are the only systems that really exist).

Learn more:

<https://dbp.io/pubs/2017/linking-types-snapl-submission.pdf>

Linking types: extra content

Backwards Compatibility

- Annotations are optional for programmers.
- To change target, only need to fully abstractly compile old target to the new.

Linking types: extra content

Backwards Compatibility

- Annotations are optional for programmers.
- To change target, only need to fully abstractly compile old target to the new.

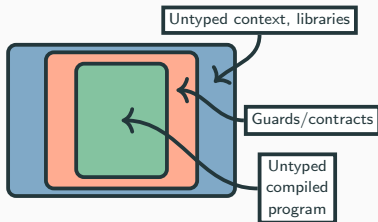
Developer Tooling

- With typed target, type translations can guide if two components can be linked together.
- Type translate to target, attempt reverse type translation to other language. Enables *cross-language type errors*.

Full abstraction: different approaches

Dynamic enforcement

- Devriese *et al.*, POPL16
- Patrignani *et al.*, TOPLAS15
- Patrignani, 2015 Dissertation
- *etc.*



Typed target languages

- New *et al.*, ICFP16
- Ahmed, Blume, ICFP11
- *etc.*

