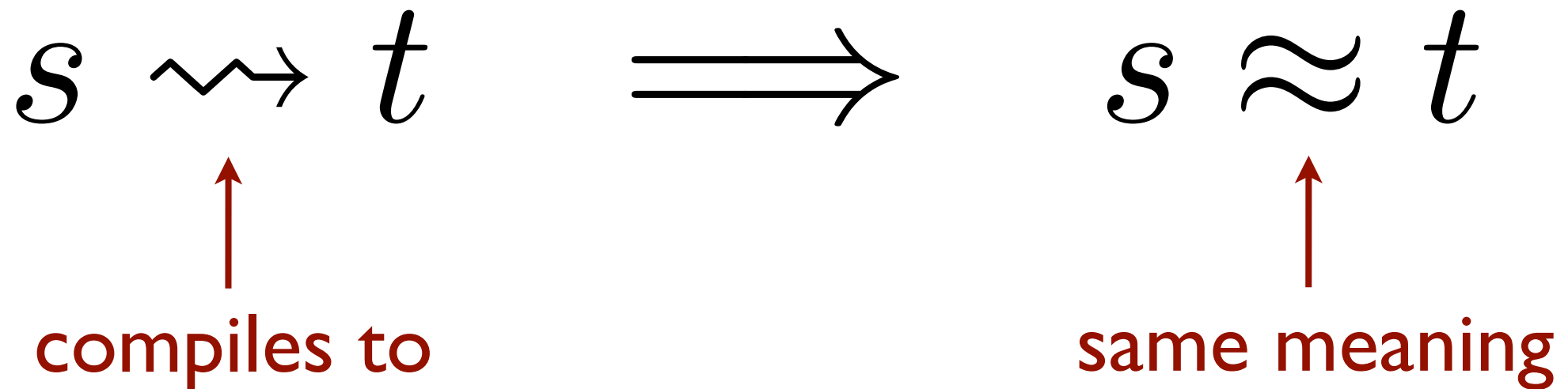


Verifying an Open Compiler from System F to Assembly

James T. Perconti & Amal Ahmed

Northeastern University

Compiler Correctness



Semantics-preserving compilation

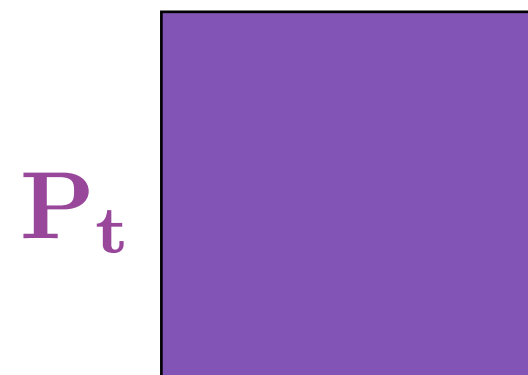
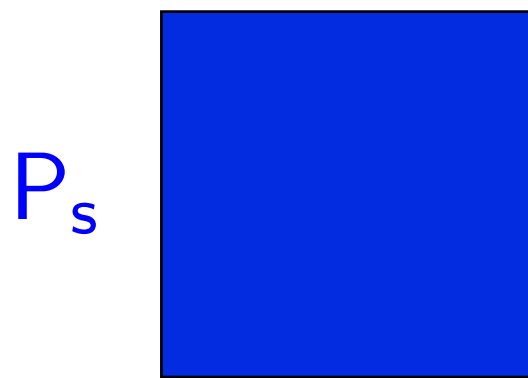
$s \rightsquigarrow t$
↑
compiles to



$s \approx t$
↑
same meaning

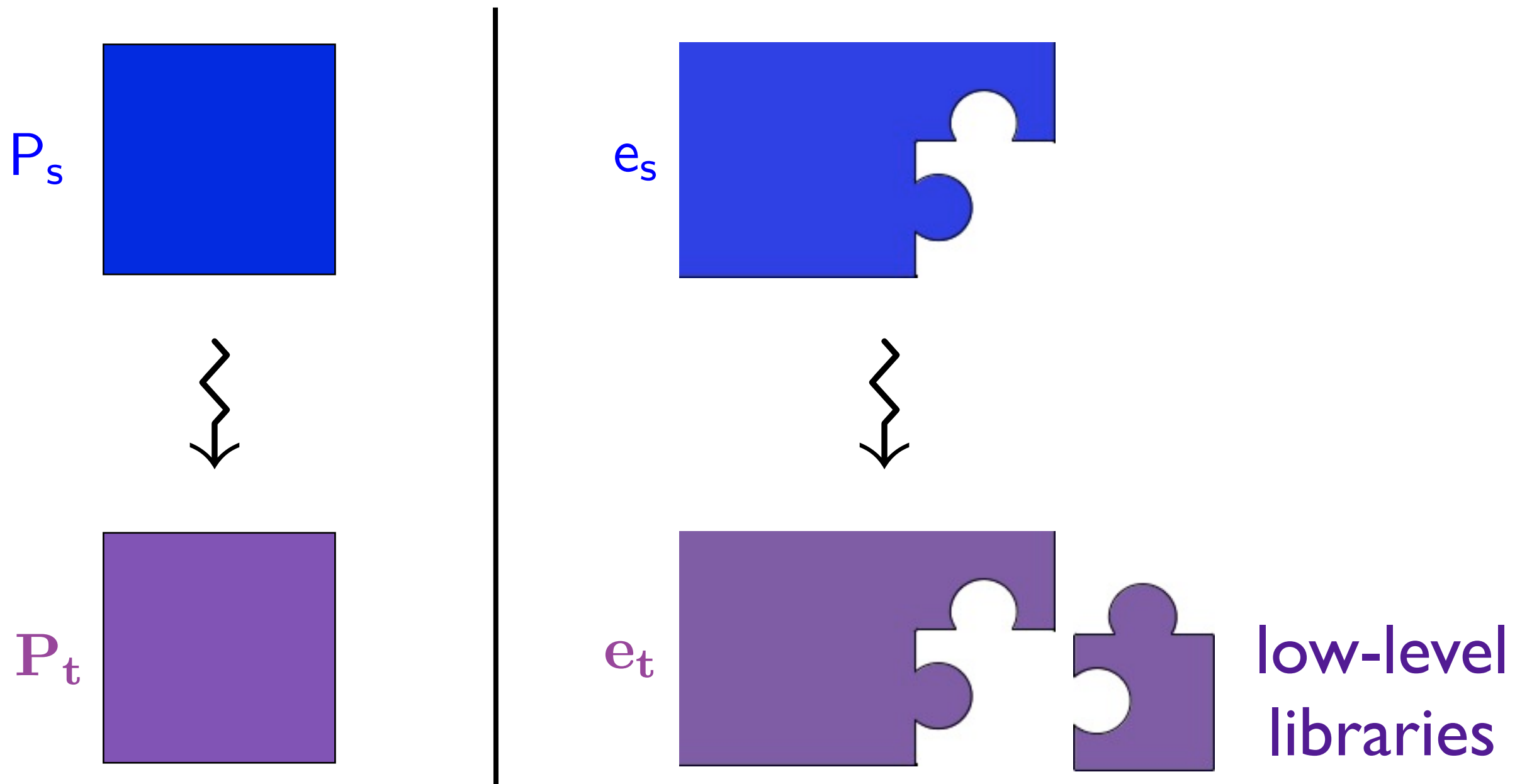
Problem: Closed-World Assumption

Correct compilation guarantee only applies to **whole** programs!



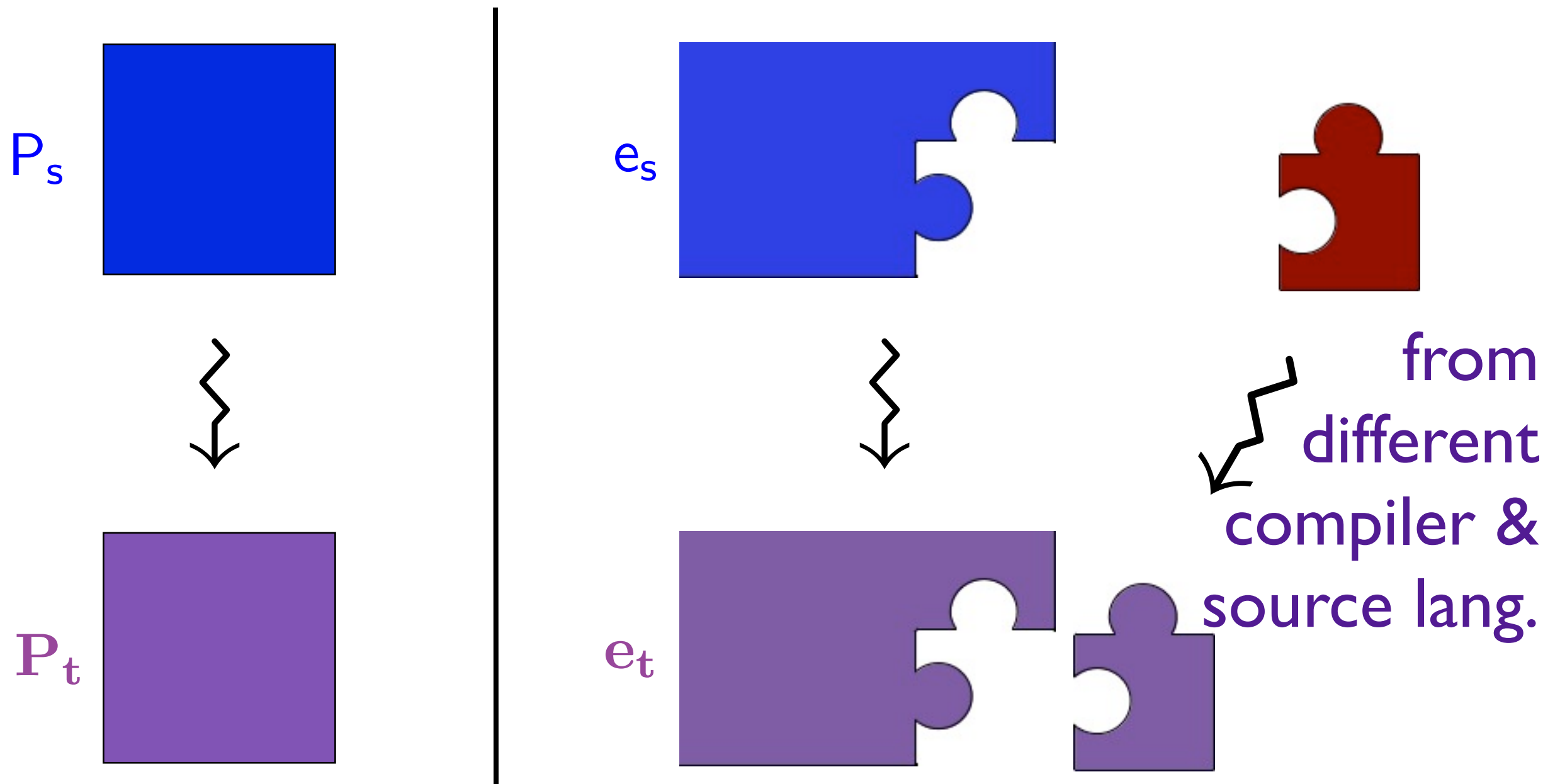
Problem: Closed-World Assumption

Correct compilation guarantee only applies to **whole** programs!



Problem: Closed-World Assumption

Correct compilation guarantee only applies to **whole** programs!



Why Whole Programs?

$$s \rightsquigarrow t \quad \Longrightarrow \quad s \approx t$$

↑
expressed how?

Why Whole Programs?

$$P_s \rightsquigarrow P_t \implies P_s \approx P_t$$

↑
expressed how?

CompCert

$$\begin{array}{ccccccc} P_s & \longmapsto & \dots & \longmapsto & P_s^i & \longmapsto & P_s^{i+1} & \longmapsto & \dots \\ \text{\color{red}!} & & & & \text{\color{red}!} & & \text{\color{red}!} & & \\ P_t & \longmapsto & \dots & \longmapsto & P_t^j & \longmapsto^* & P_t^{j+n} & \longmapsto & \dots \end{array}$$

Verifying Open Compilers: Benton-Hur

$$x : \tau' \vdash e_s : \tau \rightsquigarrow e_t \implies x : \tau' \vdash e_s \simeq e_t : \tau$$

[Benton-Hur, ICFP'09, MSR'10]

[Hur-Dreyer POPL'11]

Verifying Open Compilers: Benton-Hur

$$x : \tau' \vdash e_s : \tau \rightsquigarrow e_t \implies x : \tau' \vdash e_s \simeq e_t : \tau$$

cross-language logical relation

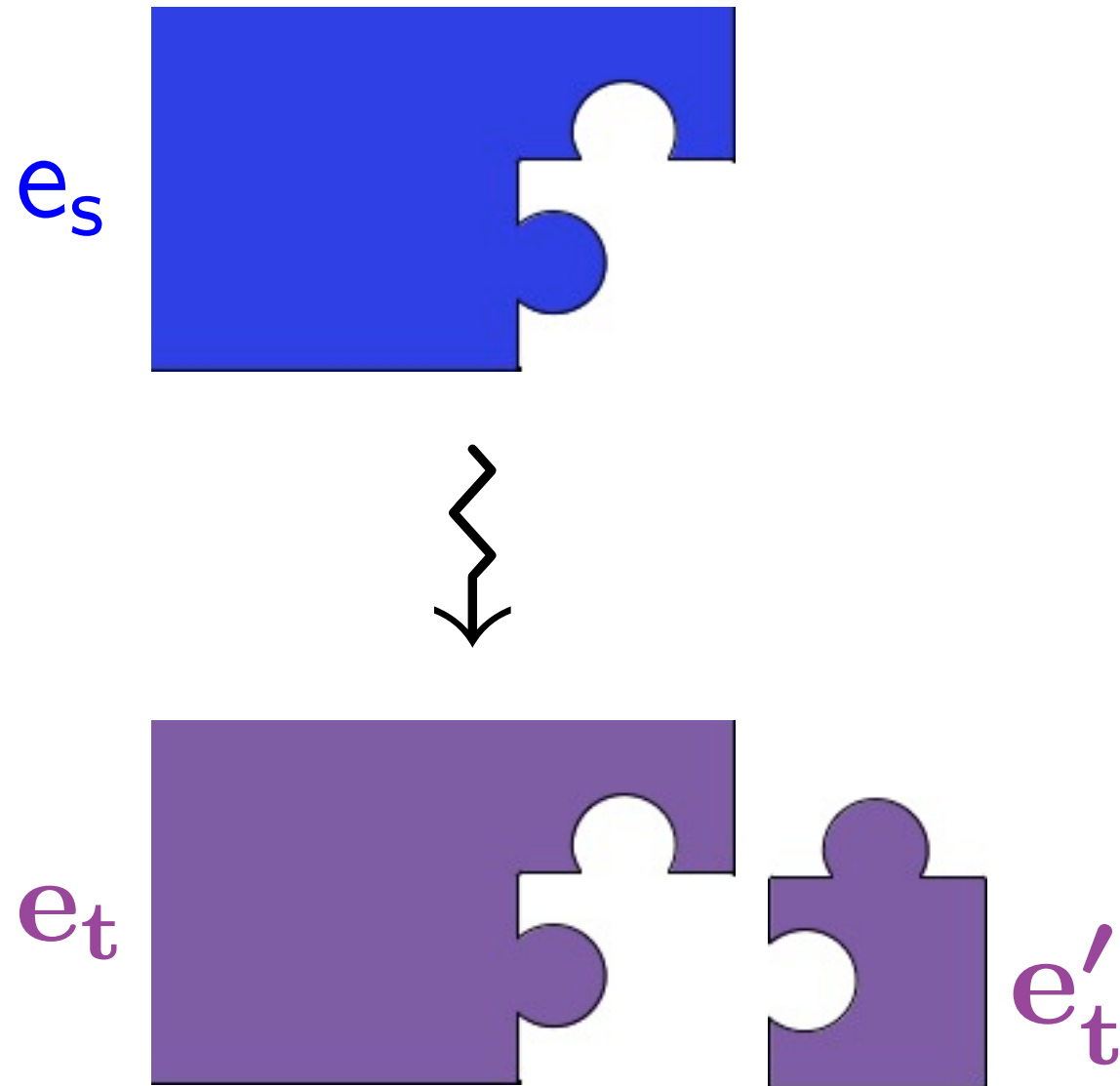
$$\forall v_s, v_t. \vdash v_s \simeq v_t : \tau' \implies \vdash e_s[v_s/x] \simeq e_t \dagger v_t : \tau$$

[Benton-Hur, ICFP'09, MSR'10]

[Hur-Dreyer POPL'11]

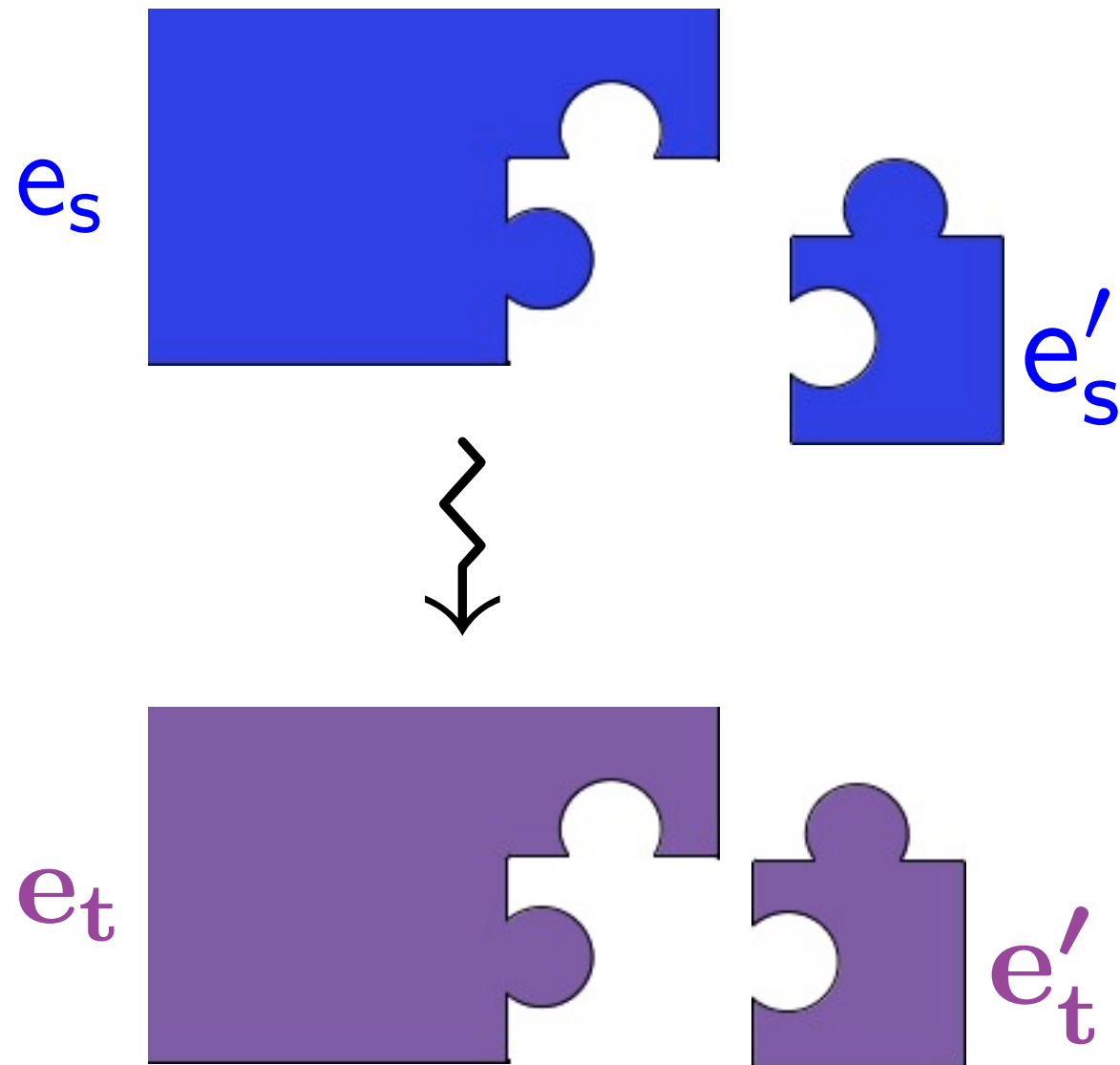
Benton-Hur: Problem 1

Have $x : \tau' \vdash e_s \approx e_t : \tau$

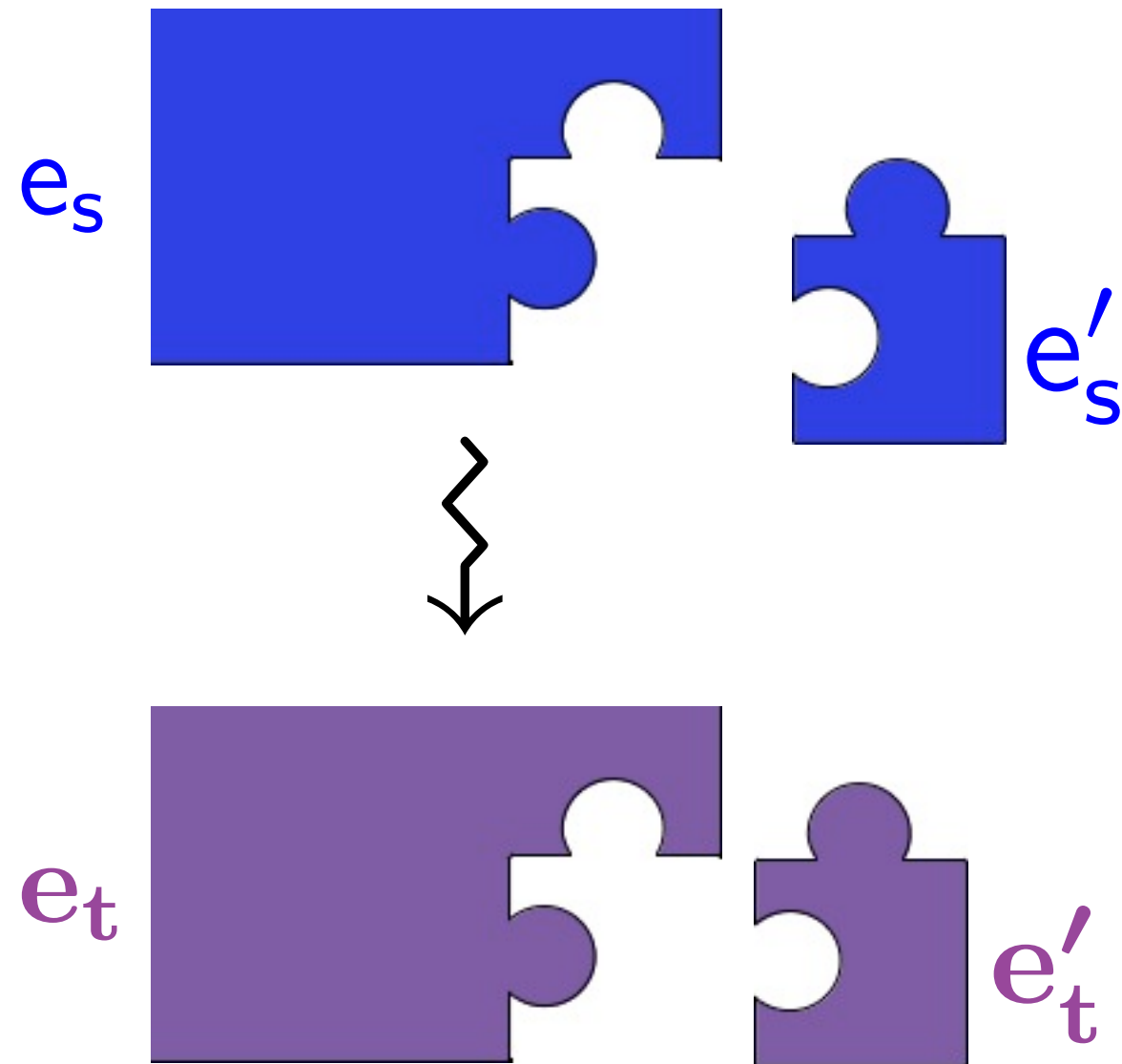


Benton-Hur: Problem 1

Have $x : \tau' \vdash e_s \approx e_t : \tau$



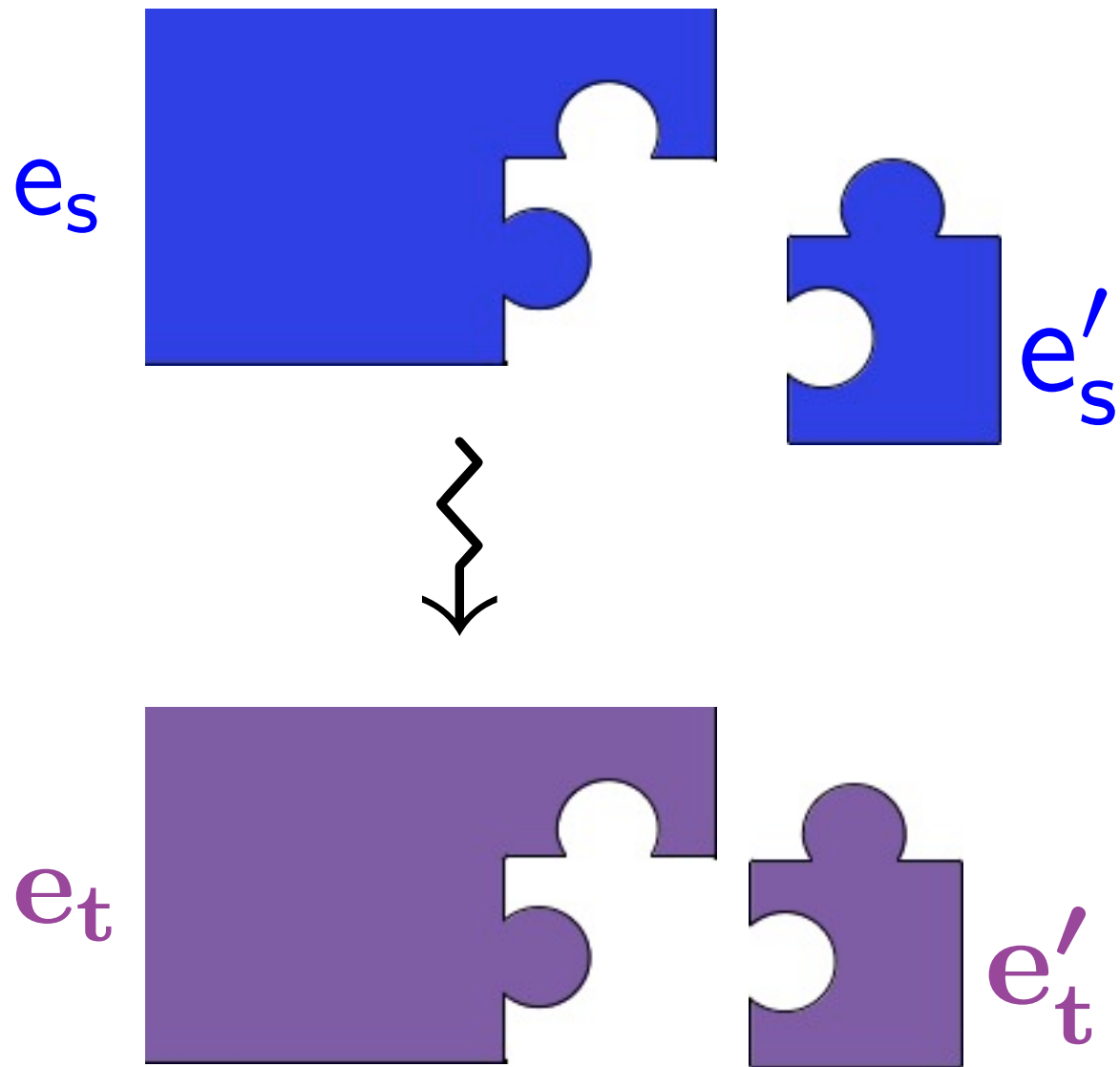
Benton-Hur: Problem 1



Have $x : \tau' \vdash e_s \approx e_t : \tau$

$\vdash e'_s \approx e'_t : \tau'$

Benton-Hur: Problem 1

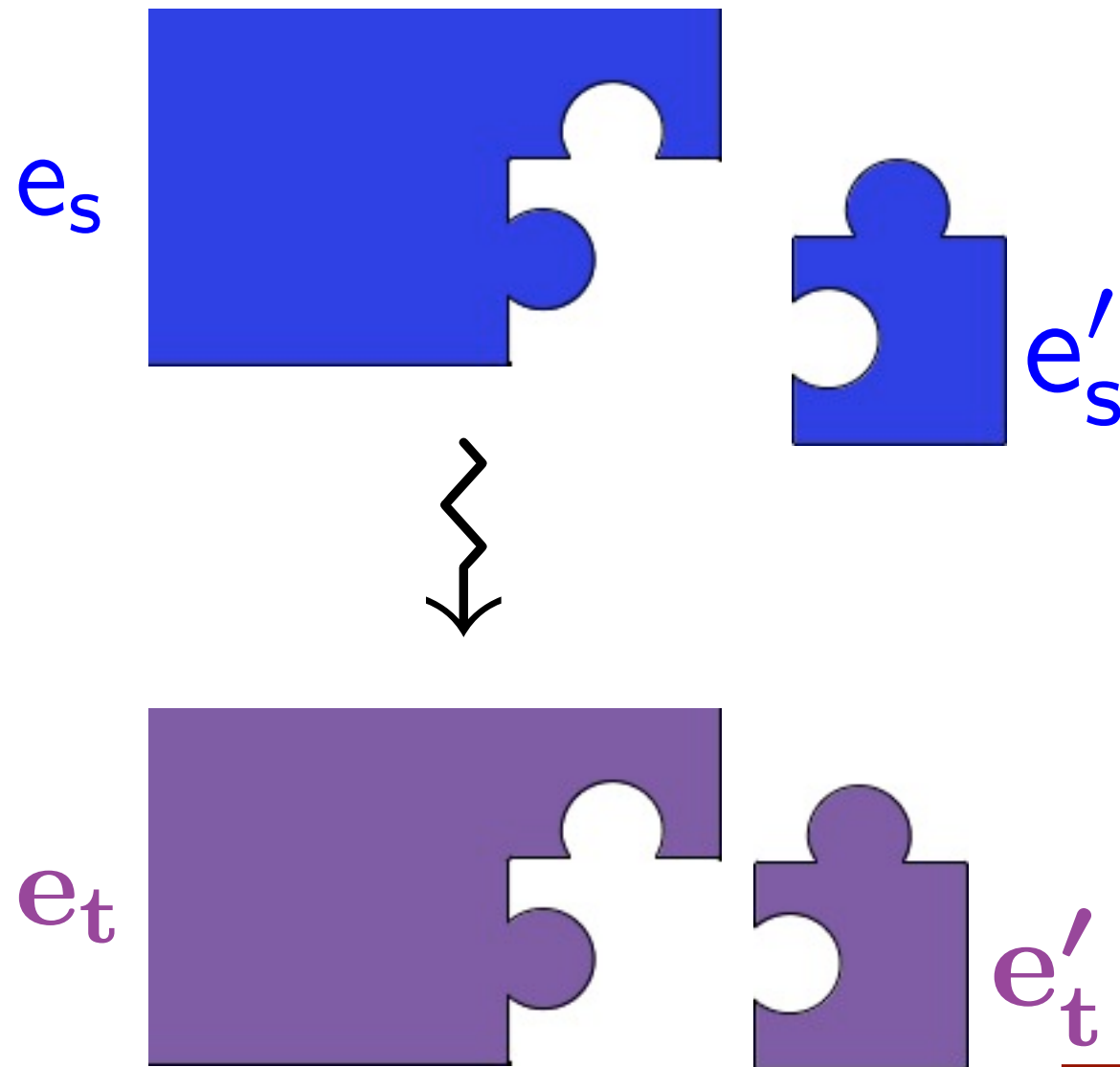


Have $x : \tau' \vdash e_s \approx e_t : \tau$

$\vdash e'_s \approx e'_t : \tau'$

$\therefore \vdash e_s[e'_s/x] \approx e_t \dagger e'_t : \tau$

Benton-Hur: Problem 1



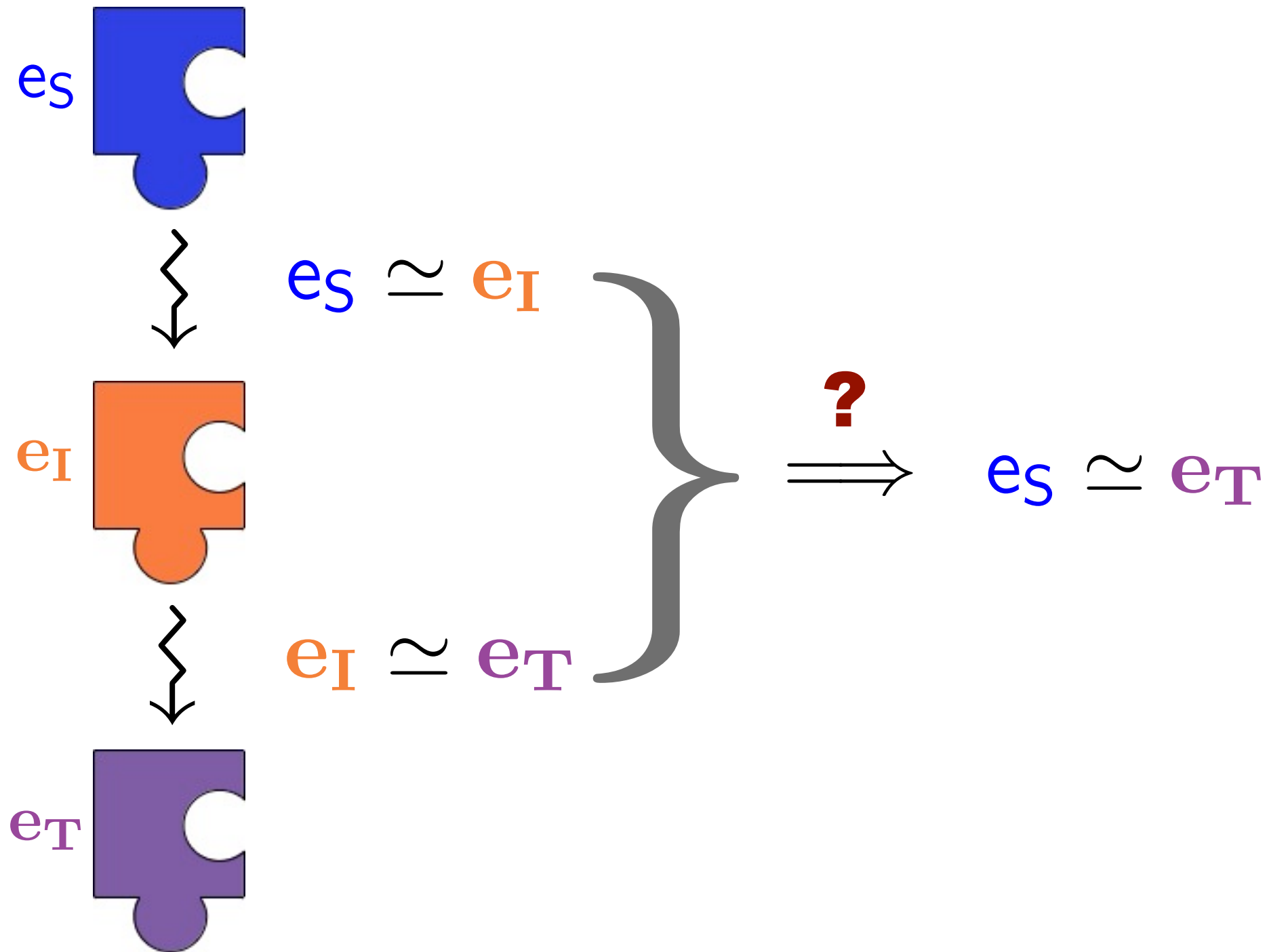
Have $x : \tau' \vdash e_s \approx e_t : \tau$

$\vdash e'_s \approx e'_t : \tau'$

$\therefore \vdash e_s[e'_s/x] \approx e_t \dagger e'_t : \tau$

Need to come up with e'_s
-- not feasible in practice!

Benton-Hur: Problem 2



Transitivity for single-lang. logical relation?

$$\left. \begin{array}{l} e_1 \approx e_2 \\ e_2 \approx e_3 \end{array} \right\} \stackrel{?}{\implies} e_1 \approx e_3$$

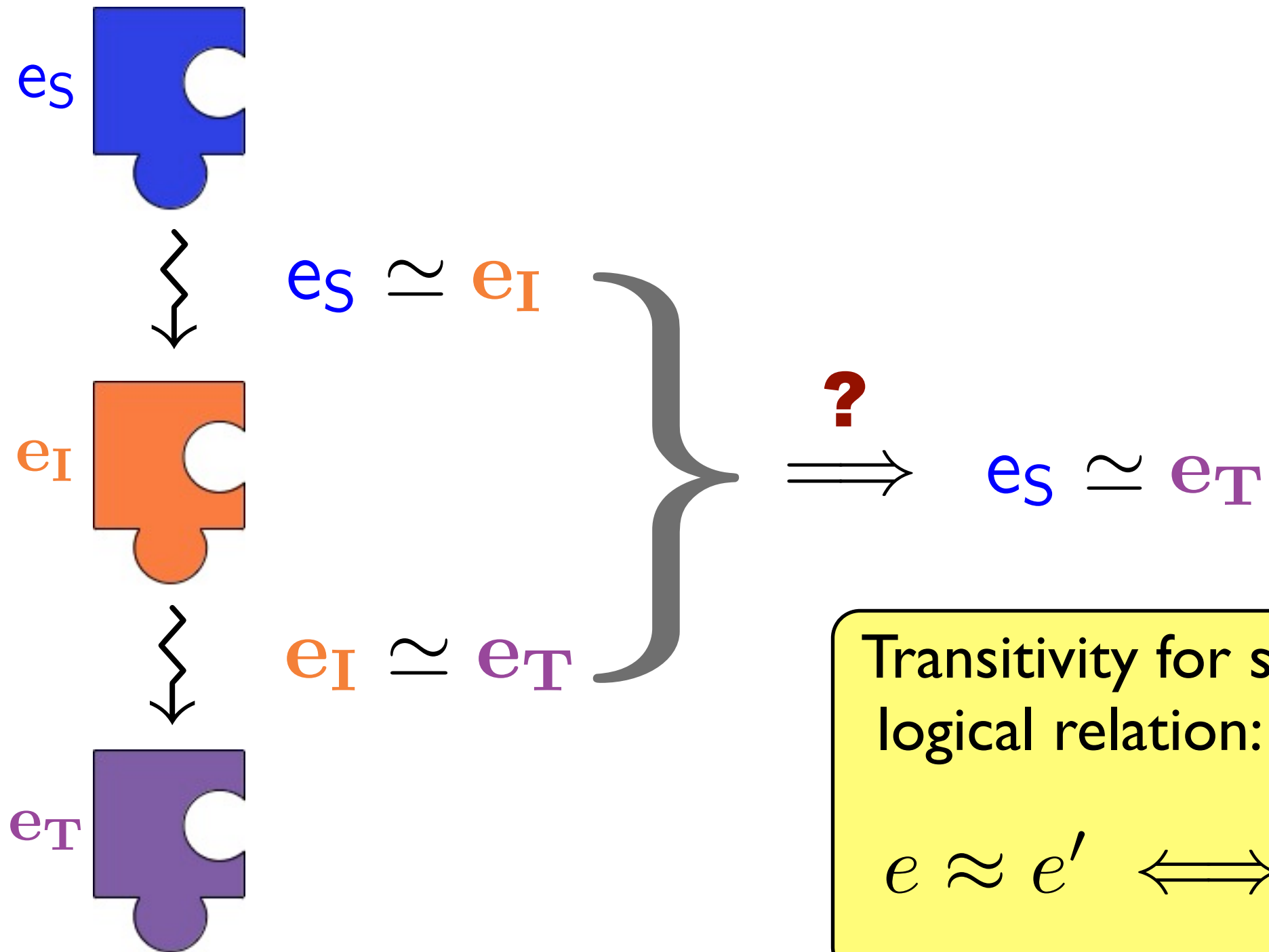
Transitivity for single-lang. logical relation?

$$\left. \begin{array}{l} e_1 \approx e_2 \\ e_2 \approx e_3 \end{array} \right\} \stackrel{?}{\implies} e_1 \approx e_3$$

- Prove: $e \approx e' \iff e \approx^{ctx} e'$

- \approx^{ctx} is transitive, $\therefore \approx$ is transitive

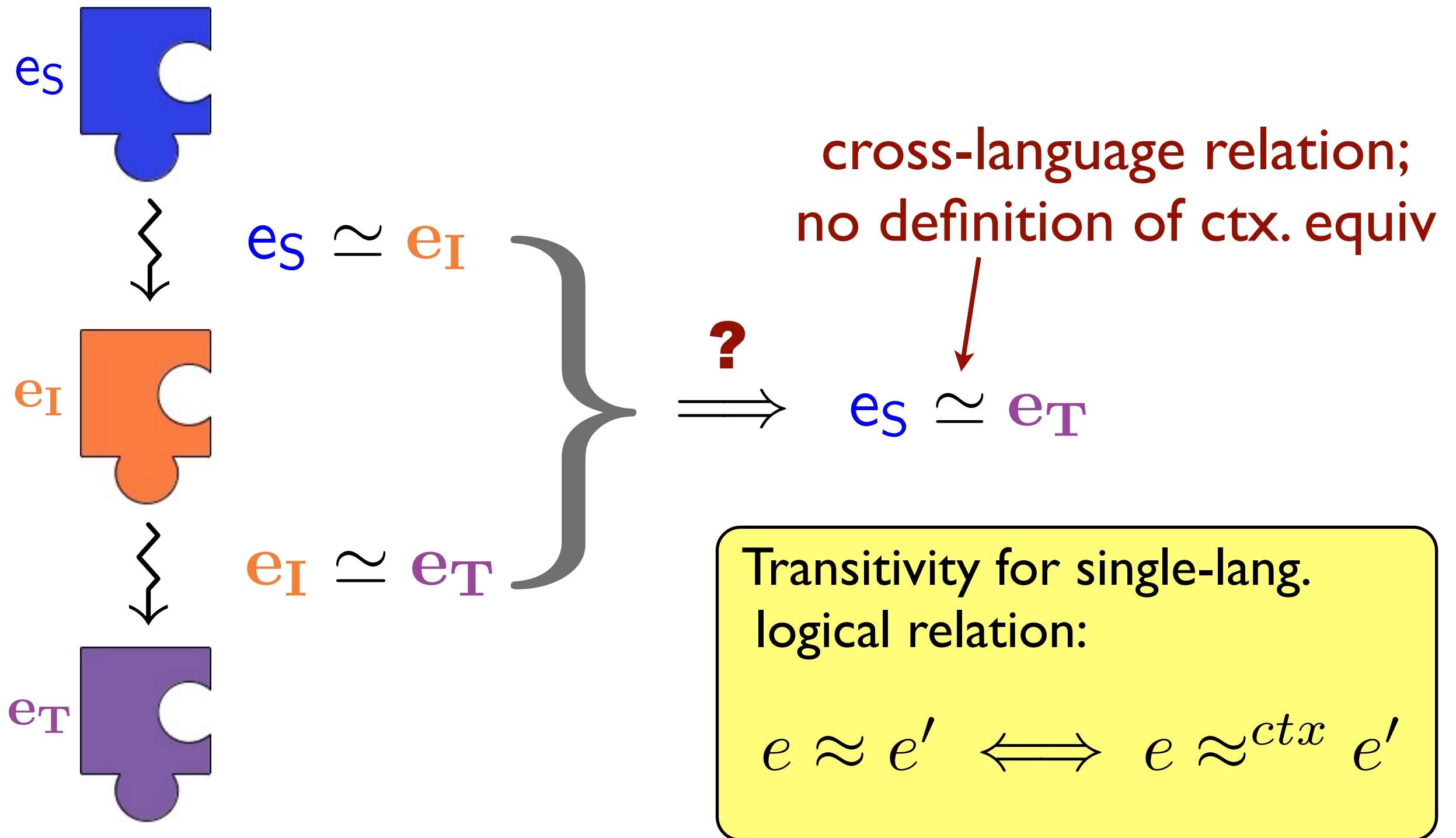
Benton-Hur: Problem 2



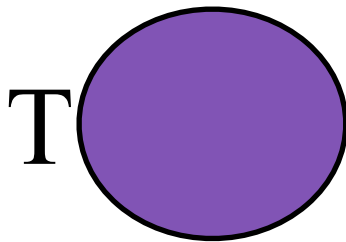
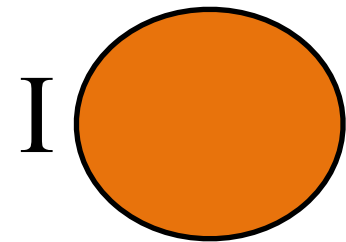
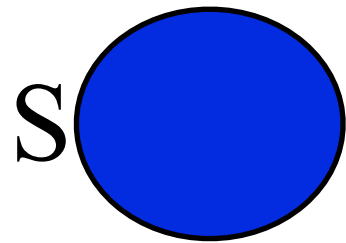
Transitivity for single-lang.
logical relation:

$$e \approx e' \iff e \approx^{ctx} e'$$

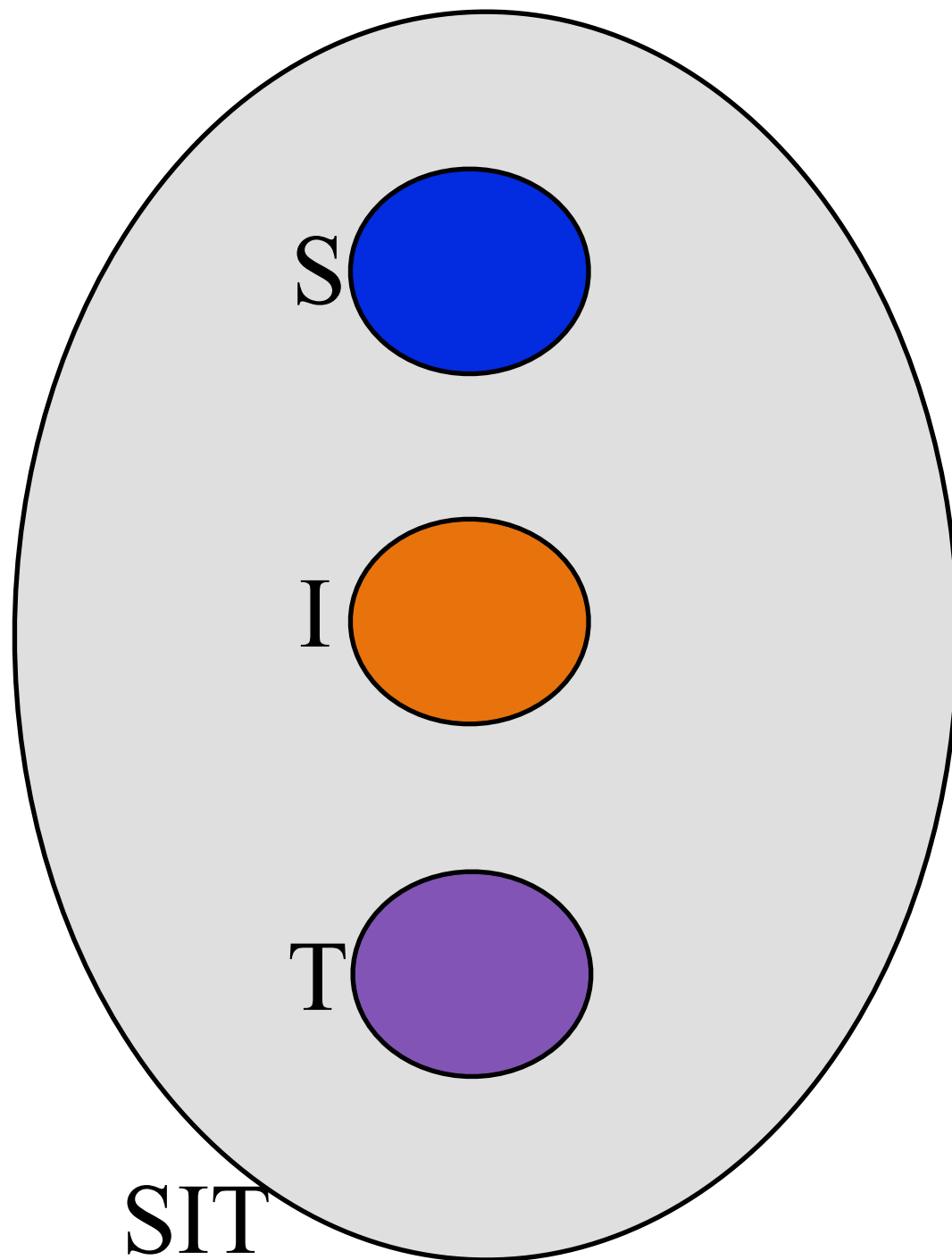
Benton-Hur: Problem 2



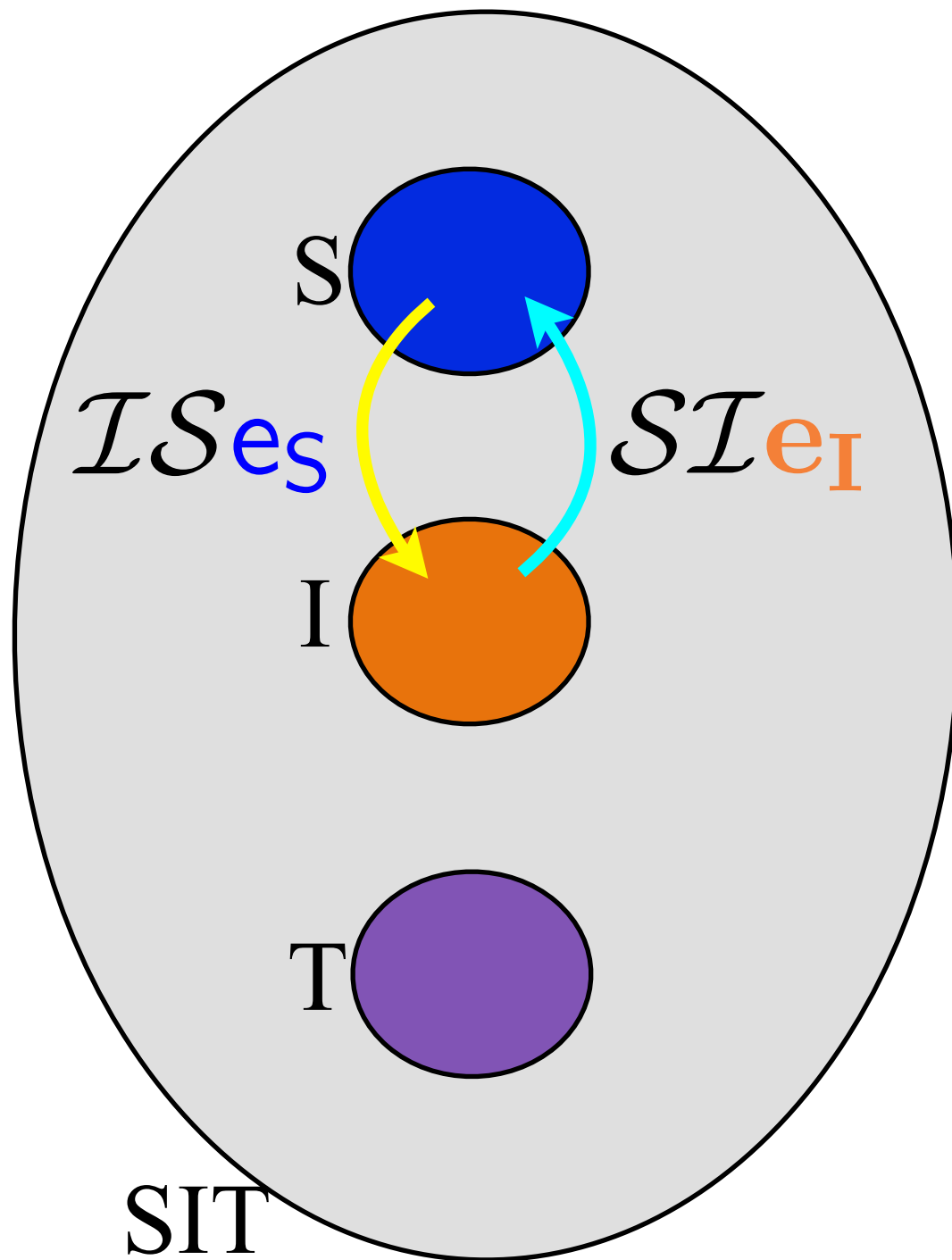
Our Approach



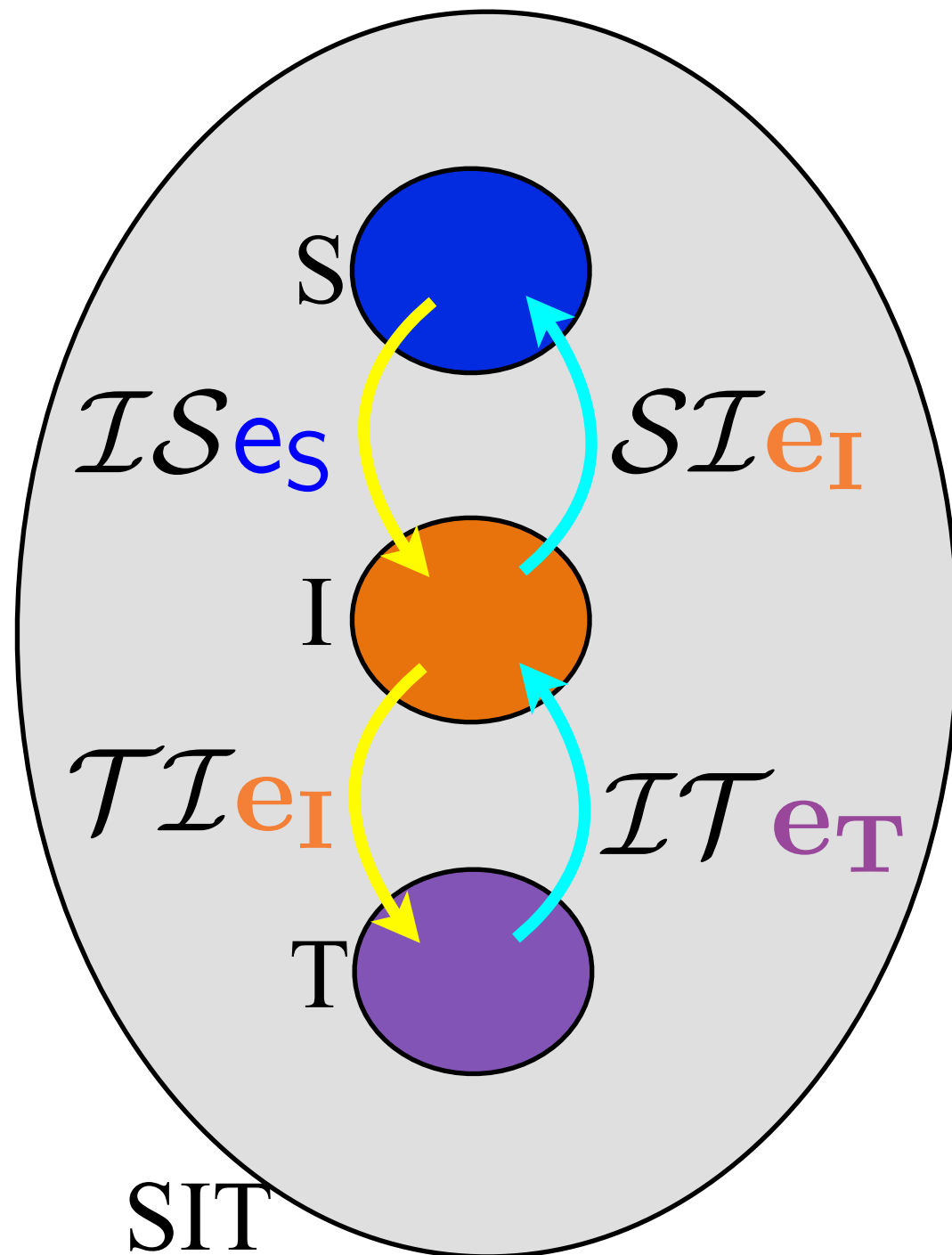
Our Approach



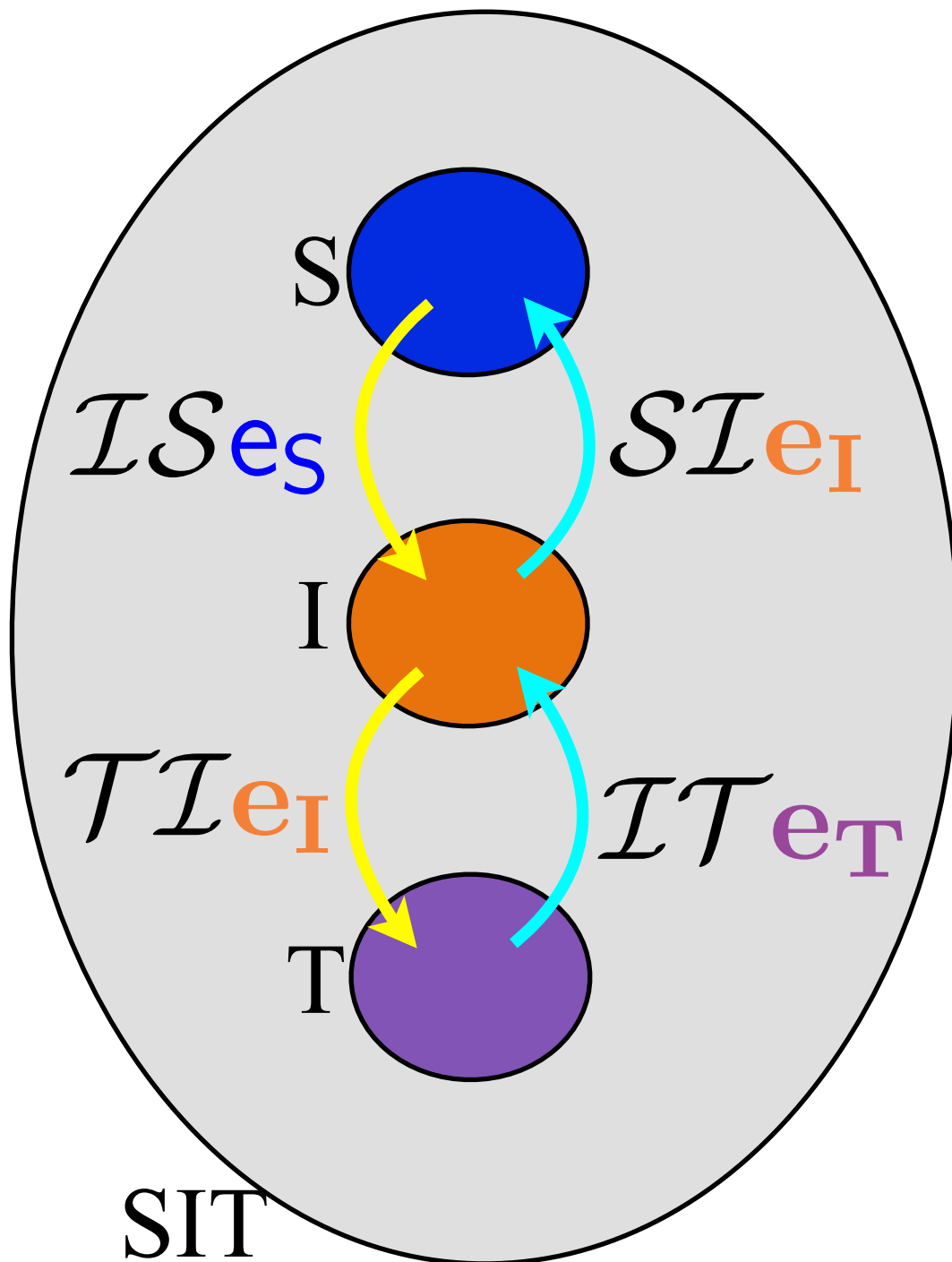
Our Approach



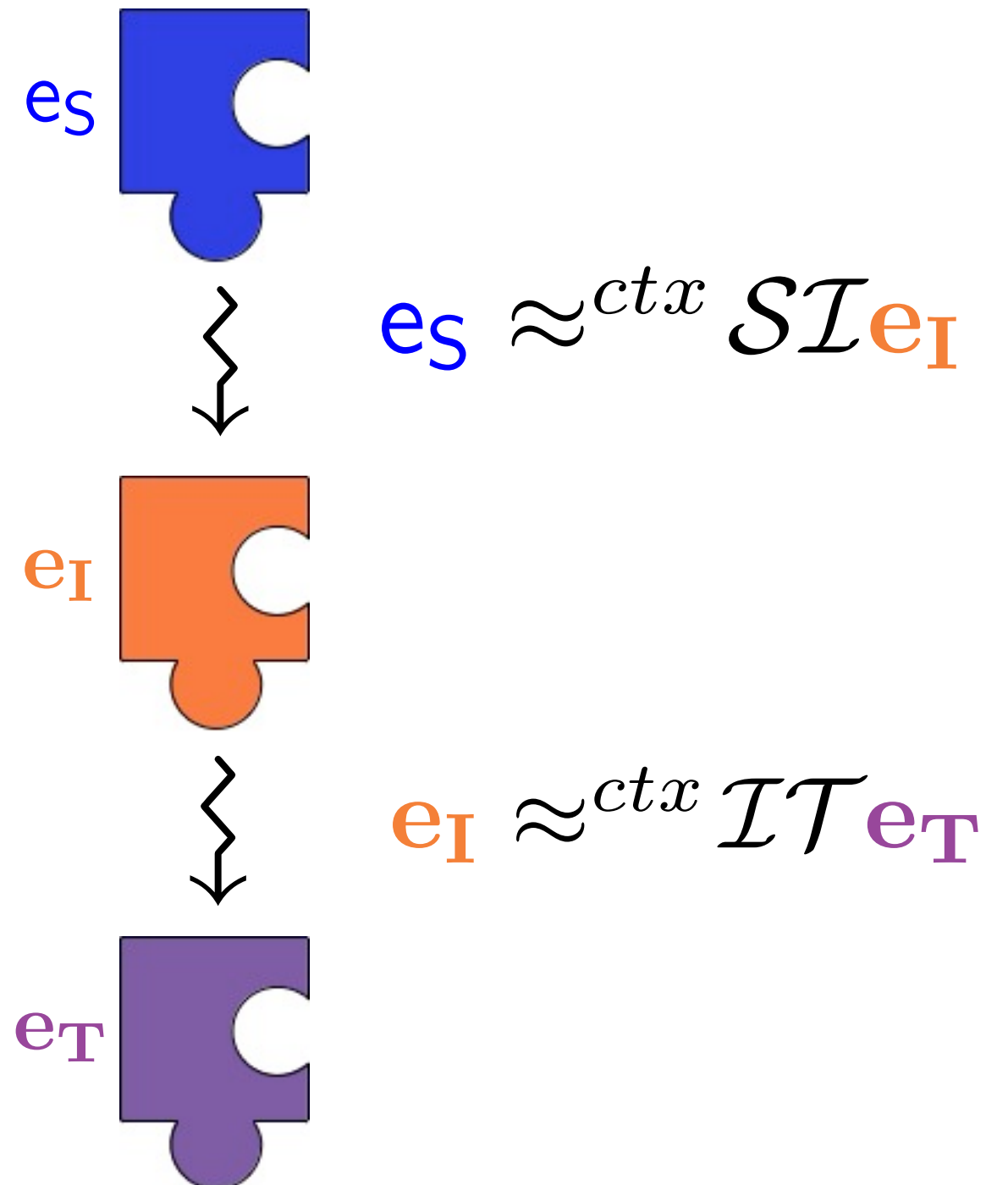
Our Approach



Our Approach

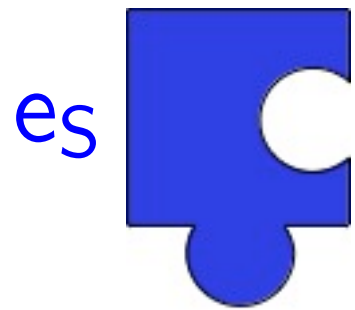


Compiler Correctness



Our Approach: Fixes Problem 2

Compiler Correctness



$$e_S \approx^{ctx} \mathcal{SI} e_I$$

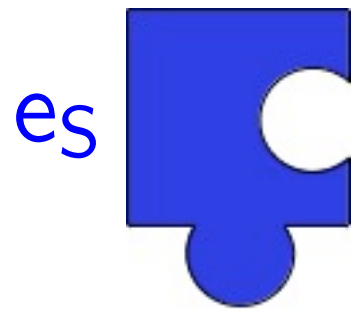


$$e_I \approx^{ctx} \mathcal{IT} e_T$$



Our Approach: Fixes Problem 2

Compiler Correctness



$$e_S \approx^{ctx} \mathcal{SI}e_I$$

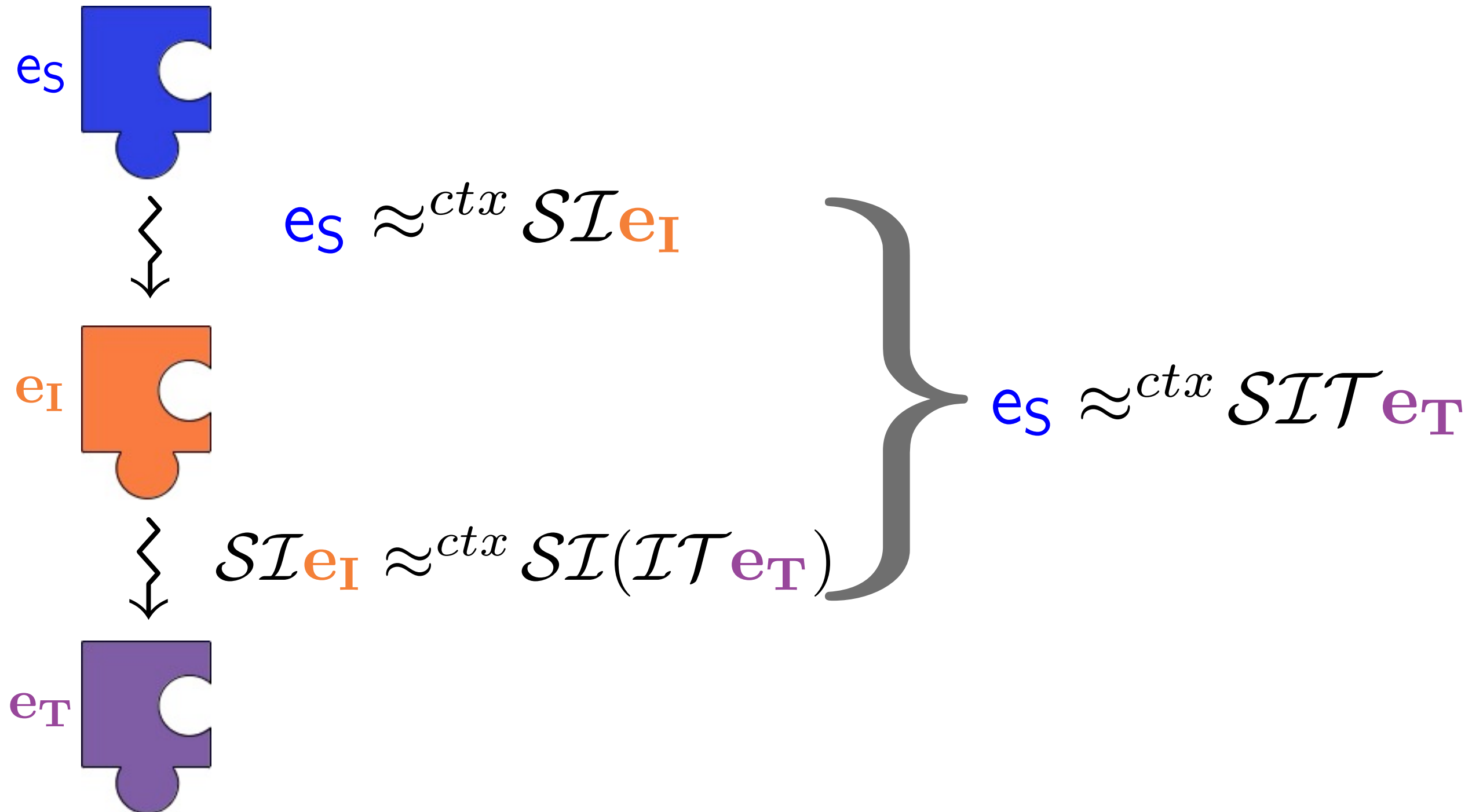


$$\mathcal{SI}e_I \approx^{ctx} \mathcal{SI}(\mathcal{IT}e_T)$$

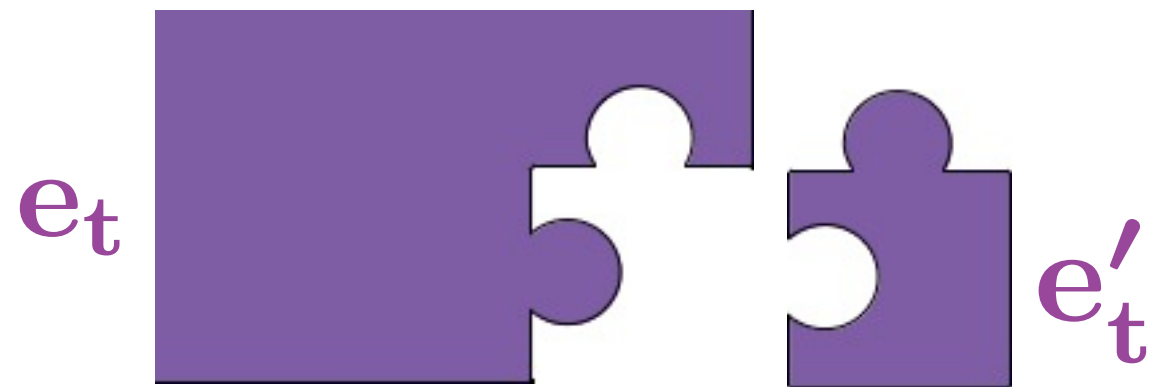
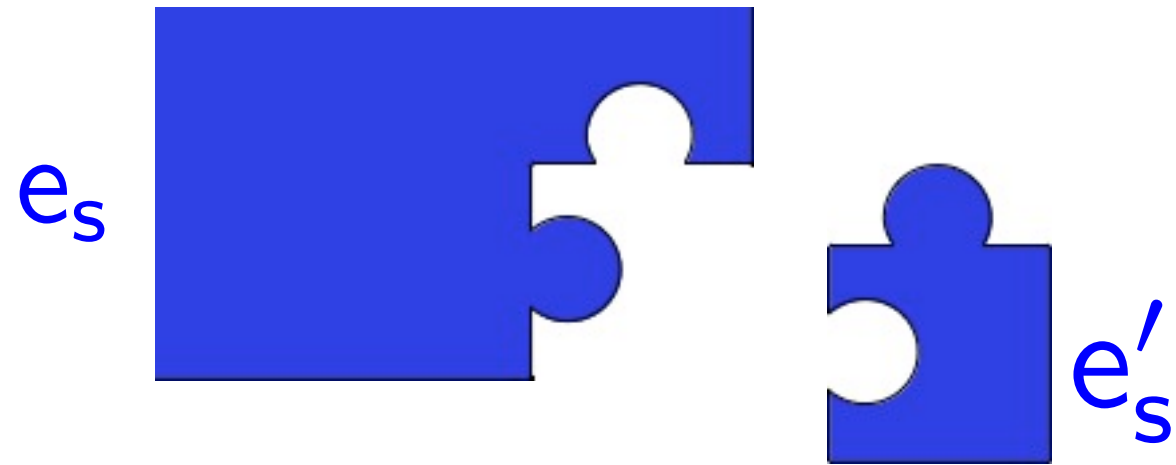


Our Approach: Fixes Problem 2

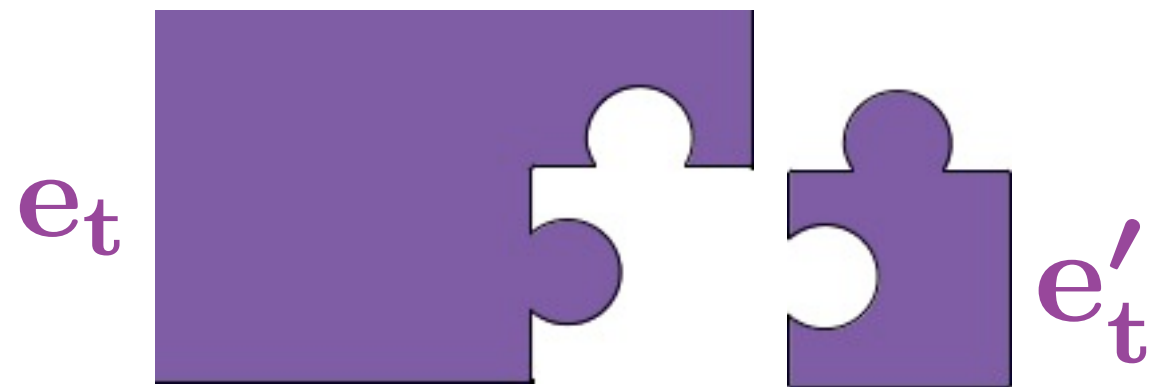
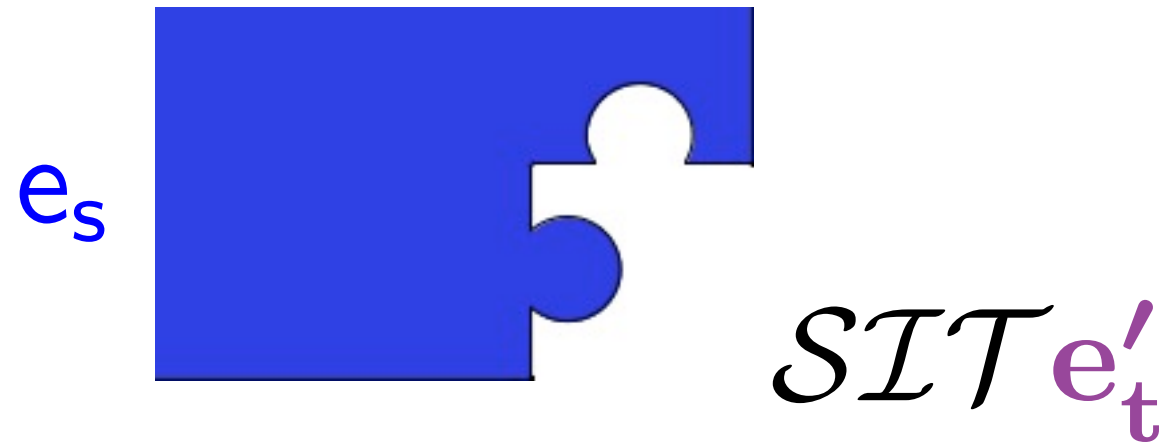
Compiler Correctness



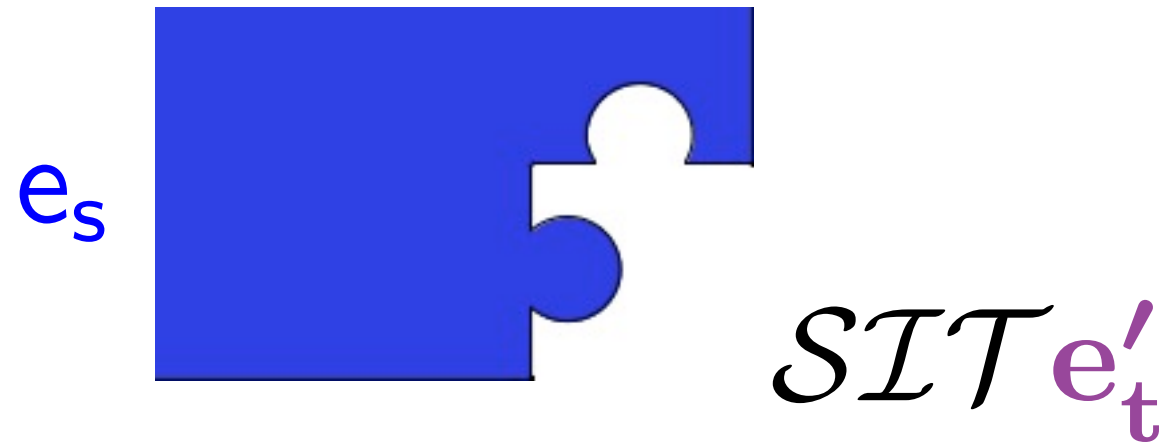
Our Approach: Fixes Problem 1



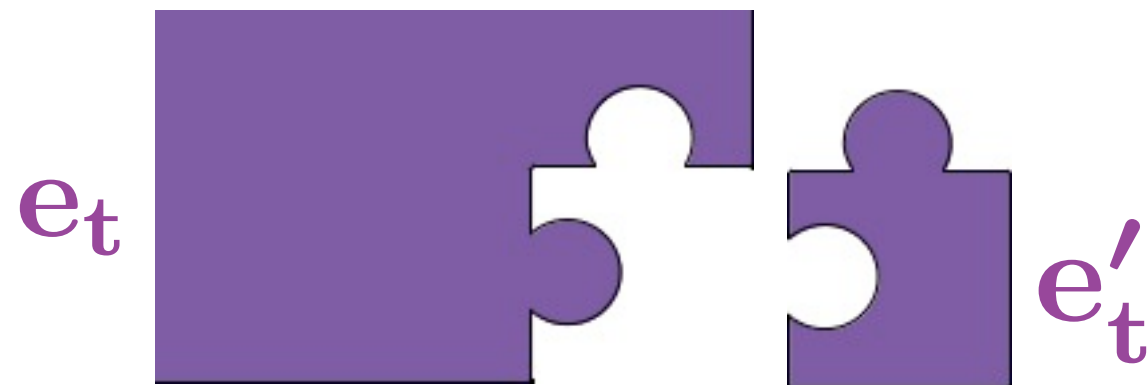
Our Approach: Fixes Problem 1



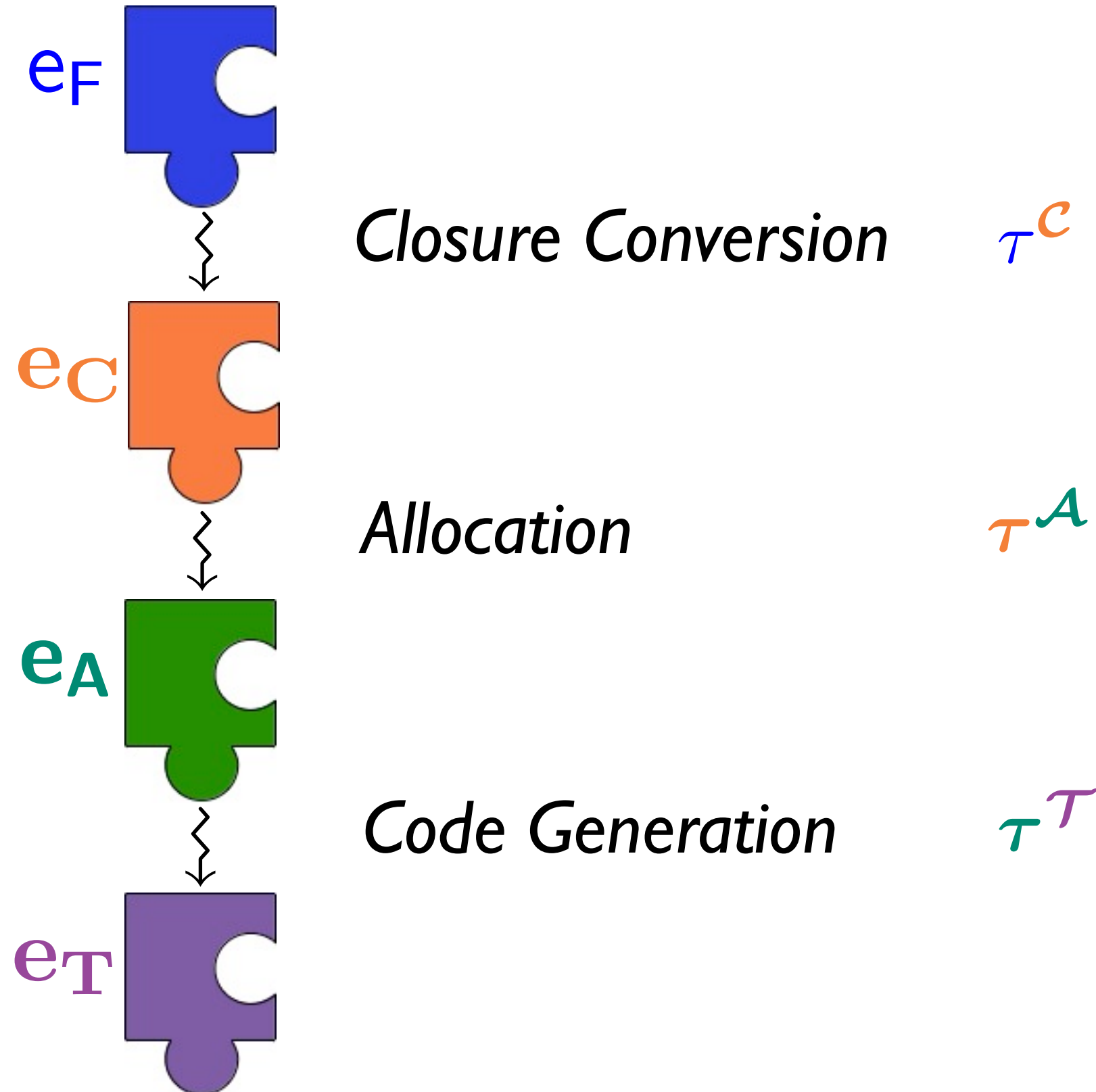
Our Approach: Fixes Problem 1



$$TIS(e_s (SIT e'_t)) \approx^{ctx} e_t e'_t$$

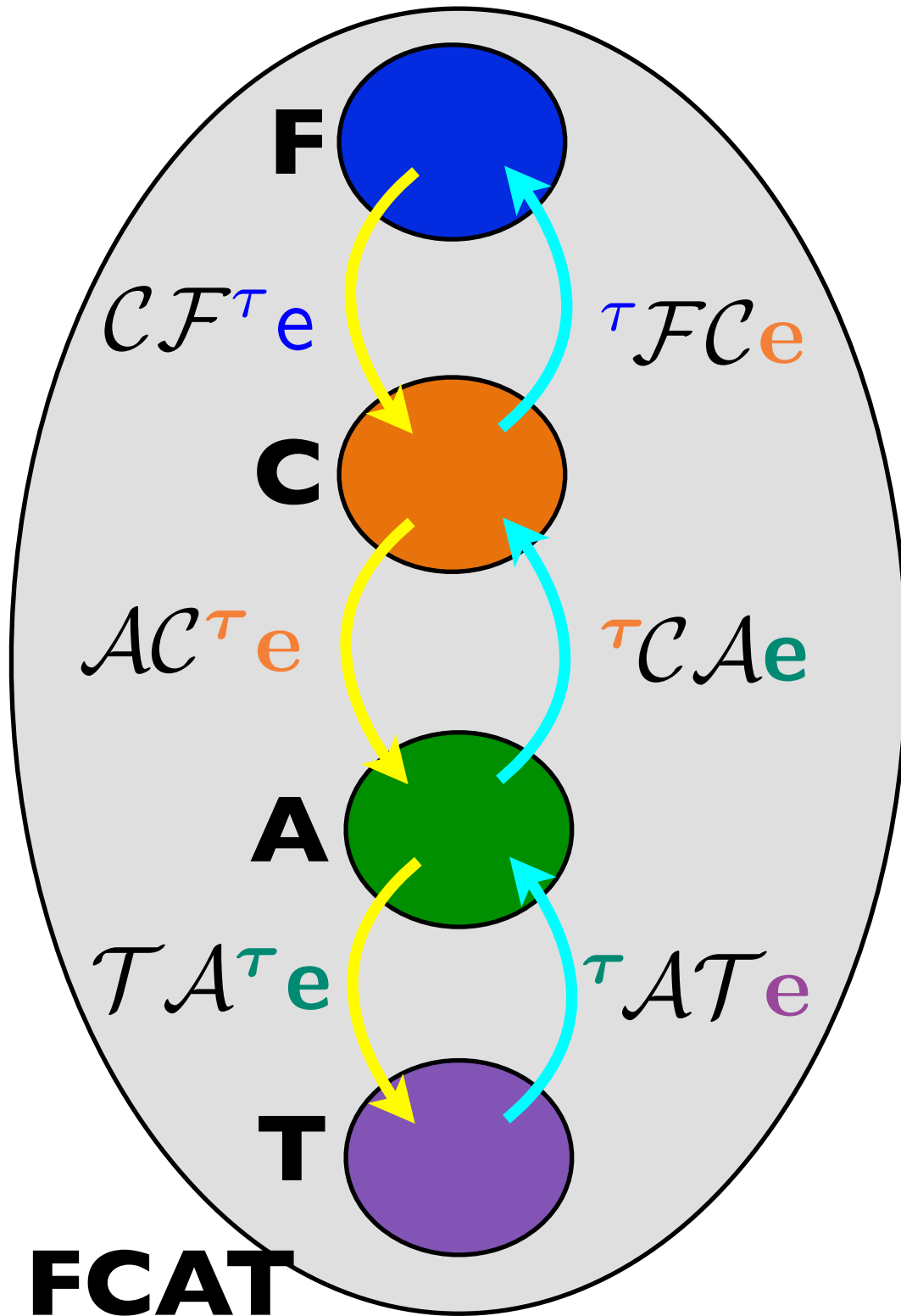


Our Compiler: System F to TAL

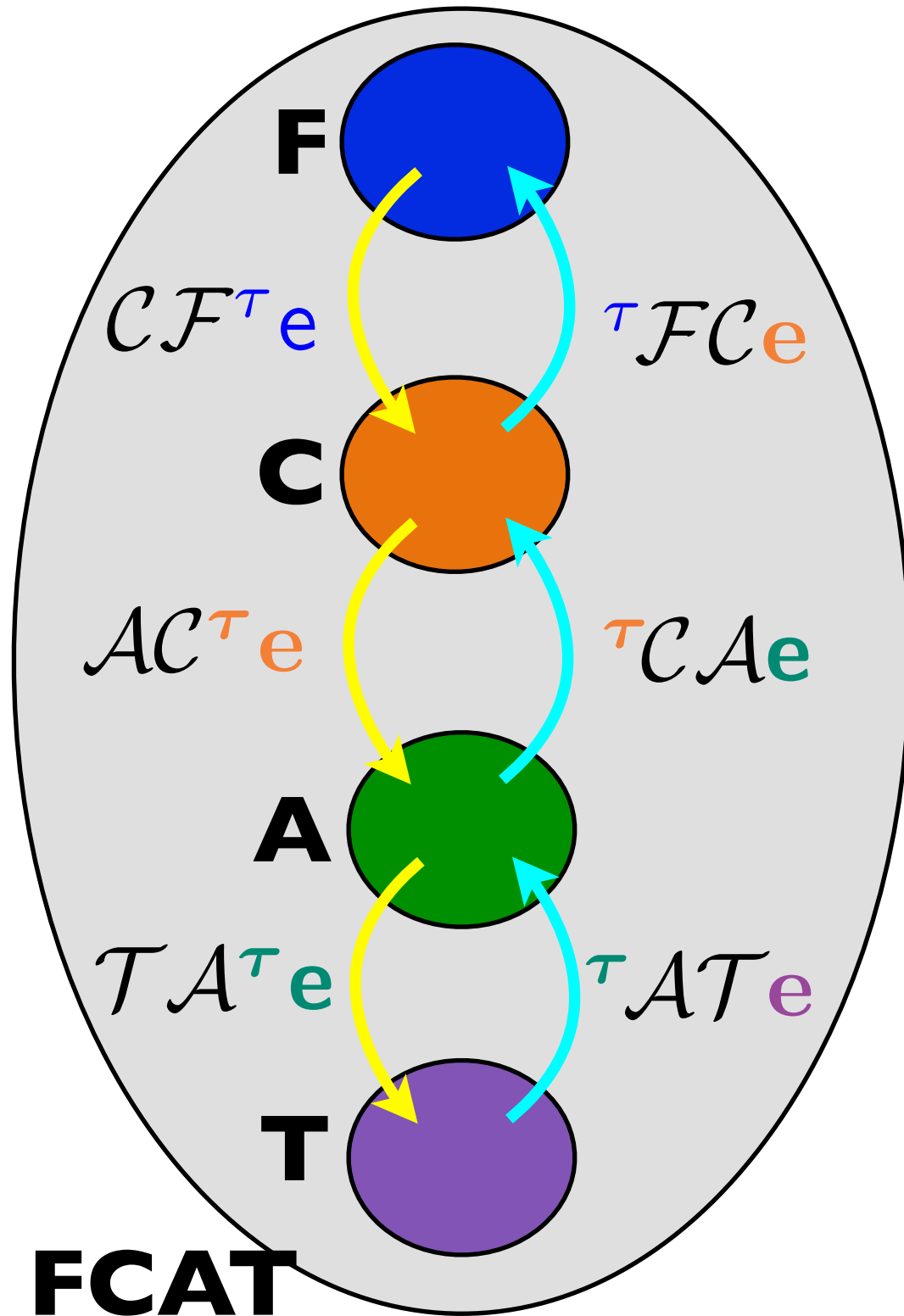


Combined language **FCAT**

- Boundaries mediate between
- τ & τ^C τ & τ^A τ & τ^T



Combined language **FCAT**



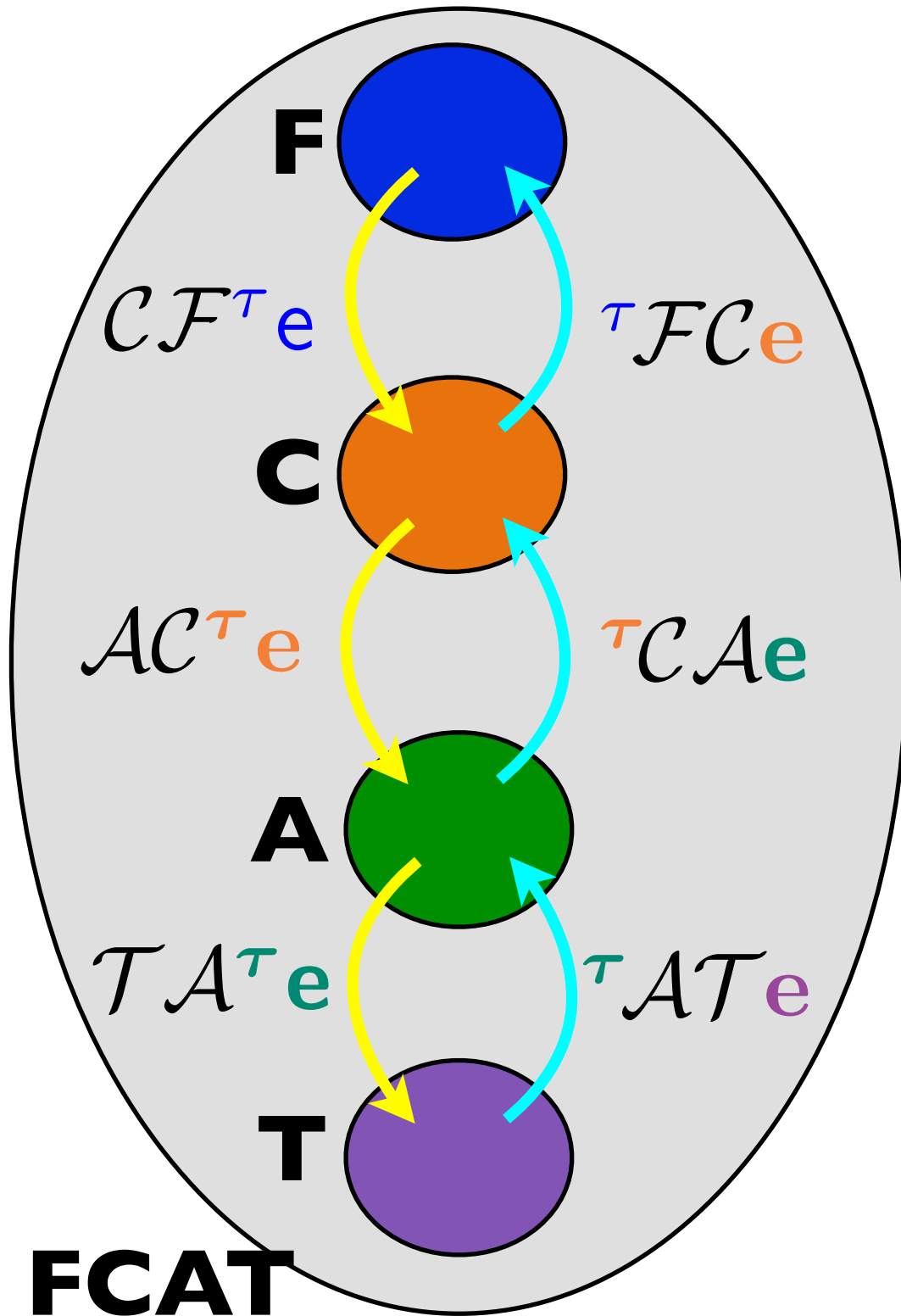
- Boundaries mediate between
 - τ & τ^C τ & τ^A τ & τ^T

- Operational semantics

$$CF^{\tau}e \mapsto^* CF^{\tau}v \mapsto v$$

$$\tau FC_e \mapsto^* \tau FC_v \mapsto v$$

Combined language **FCAT**



- Boundaries mediate between
 - $\tau \& \tau^C$ $\tau \& \tau^A$ $\tau \& \tau^T$

- Operational semantics

$$CF^\tau e \mapsto^* CF^\tau v \mapsto v$$

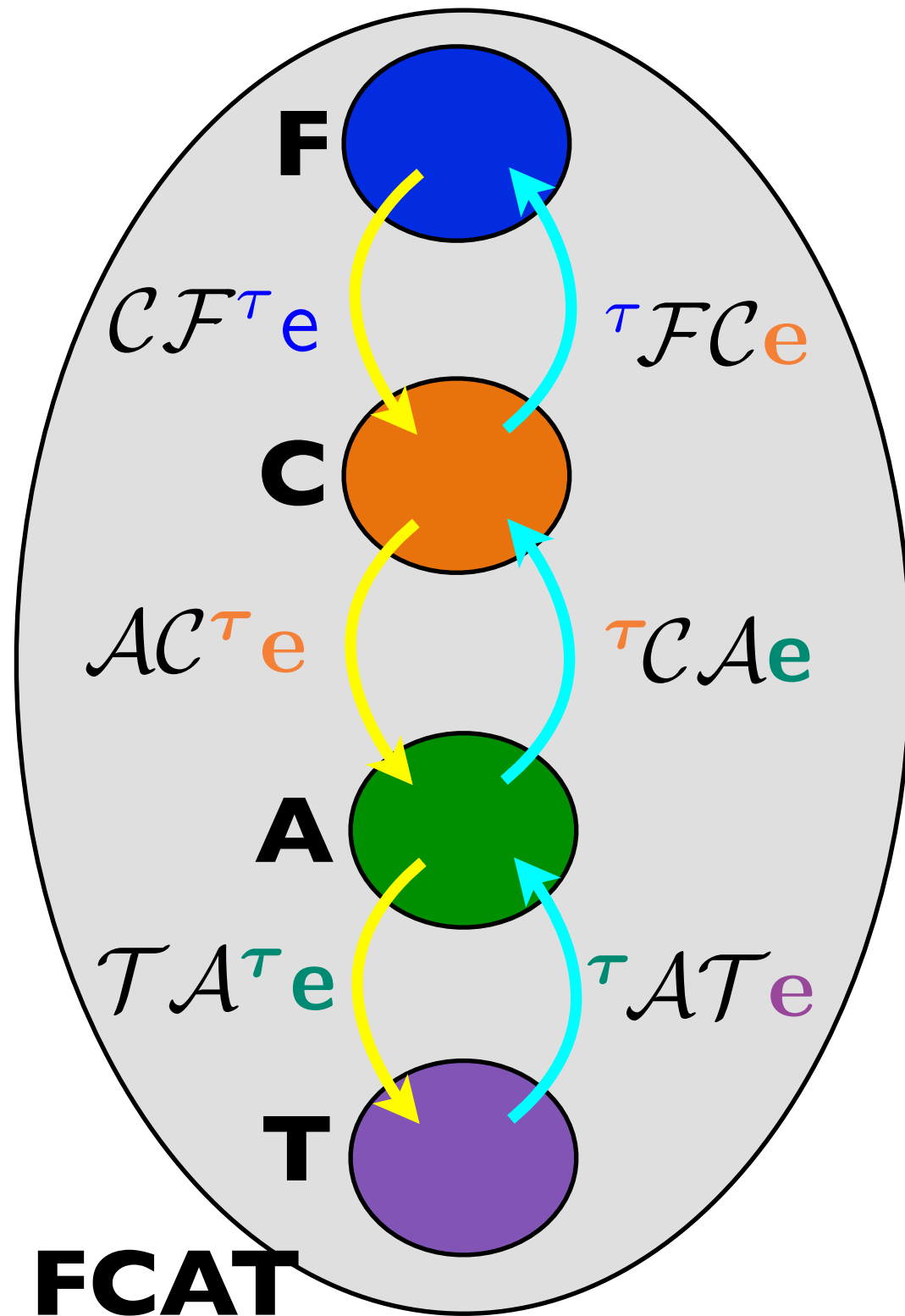
$$\tau FC e \mapsto^* \tau FC v \mapsto v$$

- Boundary cancellation

$$\tau FCCF^\tau e \approx^{ctx} e : \tau$$

$$CF^\tau \tau FC e \approx^{ctx} e : \tau^C$$

Challenges / Roadmap for rest of talk

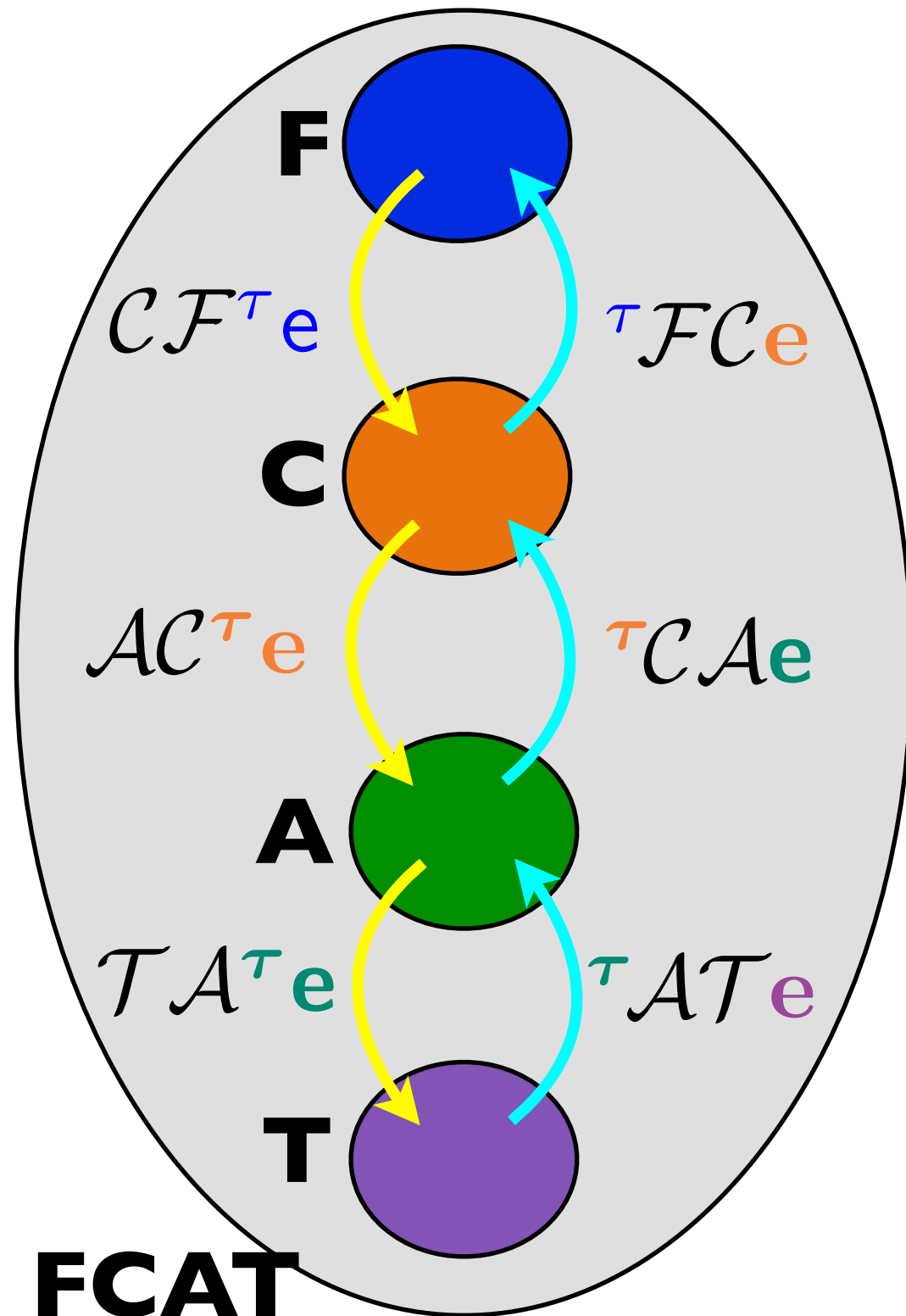


F+C: Interoperability semantics with type abstraction in both languages

C+A: Interoperability when compiler pass allocates code & tuples on heap

A+T: What is e ? What is v ?
How to define contextual equiv. for TAL *components*?
How to define logical relation?

Challenges / Roadmap for rest of talk



F+C: Interoperability semantics with type abstraction in both languages

C+A: Interoperability when compiler pass allocates code & tuples on heap

A+T: What is e ? What is v ?
How to define contextual equiv. for TAL *components*?
How to define logical relation?

F

$\tau ::= \alpha \mid \text{unit} \mid \text{int} \mid \forall[\bar{\alpha}].(\bar{\tau}) \rightarrow \tau \mid \exists\alpha.\tau \mid \mu\alpha.\tau \mid \langle\bar{\tau}\rangle$

$e ::= t$

$t ::= x \mid () \mid n \mid t \text{ p } t \mid \text{if0 } t \text{ t } t \mid \lambda[\bar{\alpha}](\overline{x:\tau}).t \mid t[\bar{\tau}]\bar{t}$
 $\mid \text{pack}\langle\tau,t\rangle \text{ as } \exists\alpha.\tau \mid \text{unpack}\langle\alpha,x\rangle = t \text{ in } t \mid \text{fold}_{\mu\alpha.\tau} t$
 $\mid \text{unfold } t \mid \langle\bar{t}\rangle \mid \pi_i(t)$

$p ::= + \mid - \mid *$

$v ::= () \mid n \mid \lambda[\bar{\alpha}](\overline{x:\tau}).t \mid \text{pack}\langle\tau,v\rangle \text{ as } \exists\alpha.\tau \mid \text{fold}_{\mu\alpha.\tau} v \mid \langle\bar{v}\rangle$

$\Delta; \Gamma \vdash e : \tau$

where $\Delta ::= \cdot \mid \Delta, \alpha$ and $\Gamma ::= \cdot \mid \Gamma, x : \tau$

C

$\tau ::= \alpha \mid \text{unit} \mid \text{int} \mid \forall[\bar{\alpha}].(\bar{\tau}) \rightarrow \tau \mid \exists\alpha.\tau \mid \mu\alpha.\tau \mid \langle\bar{\tau}\rangle$

$e ::= t$

$t ::= x \mid () \mid n \mid t \text{ p } t \mid \text{if0 } t \text{ t } t \mid \lambda[\bar{\alpha}] (\bar{x}:\bar{\tau}).t \mid t [] \bar{t}$
 $\mid t[\tau] \mid \text{pack}\langle\tau,t\rangle \text{ as } \exists\alpha.\tau \mid \text{unpack}\langle\alpha,x\rangle = t \text{ in } t$
 $\mid \text{fold}_{\mu\alpha.\tau} t \mid \text{unfold } t \mid \langle\bar{t}\rangle \mid \pi_i(t)$

$p ::= + \mid - \mid *$

$v ::= () \mid n \mid \lambda[\bar{\alpha}] (\bar{x}:\bar{\tau}).t \mid \text{pack}\langle\tau,v\rangle \text{ as } \exists\alpha.\tau$
 $\mid \text{fold}_{\mu\alpha.\tau} v \mid \langle\bar{v}\rangle \mid v[\tau]$

$\Delta; \Gamma \vdash e : \tau$

$\bar{\alpha}; \bar{x}:\bar{\tau} \vdash t : \tau'$

$\Delta; \Gamma \vdash \lambda[\bar{\alpha}] (\bar{x}:\bar{\tau}).t : \forall[\bar{\alpha}].(\bar{\tau}) \rightarrow \tau'$

Closure Conversion: **F** to **C**

τ^c Type Translation

$$\alpha^c = \alpha \quad \forall[\bar{\alpha}].(\bar{\tau}) \rightarrow \tau'^c = \exists\beta.\langle(\forall[\bar{\alpha}].(\beta, \bar{\tau}^c)) \rightarrow \tau'^c\rangle, \beta\rangle$$

$$\text{unit}^c = \text{unit} \quad \exists\alpha.\tau^c = \exists\alpha.\tau^c$$

$$\text{int}^c = \text{int} \quad \mu\alpha.\tau^c = \mu\alpha.\tau^c$$

$$\langle\tau_1, \dots, \tau_n\rangle^c = \langle\tau_1^c, \dots, \tau_n^c\rangle$$

$\Delta; \Gamma \vdash e : \tau \rightsquigarrow e$ where $\Delta^c; \Gamma^c \vdash e : \tau^c$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

Value Translation

$$\mathbf{CF}^{\text{int}}(\mathbf{n}) = \mathbf{n}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$\text{int}\mathbf{FC}(\mathbf{n}) = \mathbf{n}$$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\tau \rightarrow \tau')^{\mathbf{c}} = \exists\beta. \langle ((\beta, \tau^{\mathbf{c}}) \rightarrow \tau'^{\mathbf{c}}), \beta \rangle$$

$$\mathbf{CF}^{\tau \rightarrow \tau'}(\mathbf{v}) =$$

$$\text{pack} \langle \text{unit}, \langle \mathbf{v}, () \rangle \rangle \text{ as } \exists\beta. \langle ((\beta, \tau^{\mathbf{c}}) \rightarrow \tau'^{\mathbf{c}}), \beta \rangle$$

$$\mathbf{v} = \lambda(\mathbf{z} : \text{unit}, \mathbf{x} : \tau^{\mathbf{c}}). \mathbf{CF}^{\tau'}(\mathbf{v} \tau\mathbf{FC} \mathbf{x})$$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\tau \rightarrow \tau')^{\mathbf{c}} = \exists\beta. \langle ((\beta, \tau^{\mathbf{c}}) \rightarrow \tau'^{\mathbf{c}}), \beta \rangle$$

$$\mathbf{CF}^{\tau \rightarrow \tau'}(\mathbf{v}) =$$

$$\text{pack} \langle \text{unit}, \langle \mathbf{v}, () \rangle \rangle \text{ as } \exists\beta. \langle ((\beta, \tau^{\mathbf{c}}) \rightarrow \tau'^{\mathbf{c}}), \beta \rangle$$

$$\mathbf{v} = \lambda(\mathbf{z} : \text{unit}, \mathbf{x} : \tau^{\mathbf{c}}). \mathbf{CF}^{\tau'}(\mathbf{v} \tau\mathbf{FC} \mathbf{x})$$

$$\tau \rightarrow \tau' \mathbf{FC}(\mathbf{v}) = \lambda(\mathbf{x} : \tau). \tau' \mathbf{FC}(\text{unpack} \langle \beta, \mathbf{y} \rangle = \mathbf{v} \\ \text{in } \pi_1(\mathbf{y}) \pi_2(\mathbf{y}) \mathbf{CF}^{\tau} \mathbf{x})$$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\forall[\alpha].(\alpha) \rightarrow \alpha)^{\mathbf{c}} = \exists\beta. \langle (\forall[\alpha].(\beta, \alpha) \rightarrow \alpha), \beta \rangle$$

$$\alpha^{\mathbf{c}} = \alpha$$

$$\mathbf{CF}^{\forall[\alpha].(\alpha) \rightarrow \alpha}(\mathbf{v}) = \mathbf{pack} \langle \mathbf{unit}, \langle \mathbf{v}, () \rangle \rangle \mathbf{as} (\forall[\alpha].(\alpha) \rightarrow \alpha)^{\mathbf{c}}$$

$$\mathbf{v} = \lambda[\alpha](z : \mathbf{unit}, x : \alpha). \mathcal{CF}^{\alpha}(\mathbf{v} [\alpha] \alpha \mathcal{FC} \mathbf{x})$$

Interoperability: **F** and **C**

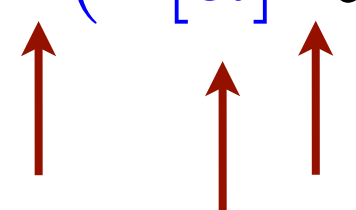
$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\forall[\alpha].(\alpha) \rightarrow \alpha)^{\mathbf{c}} = \exists\beta. \langle (\forall[\alpha].(\beta, \alpha) \rightarrow \alpha), \beta \rangle$$

$$\alpha^{\mathbf{c}} = \alpha$$

$$\mathbf{CF}^{\forall[\alpha].(\alpha) \rightarrow \alpha}(\mathbf{v}) = \mathbf{pack} \langle \mathbf{unit}, \langle \mathbf{v}, () \rangle \rangle \mathbf{as} (\forall[\alpha].(\alpha) \rightarrow \alpha)^{\mathbf{c}}$$

$$\mathbf{v} = \lambda[\alpha](z : \mathbf{unit}, x : \alpha). \mathcal{CF}^{\alpha}(\mathbf{v}[\alpha])^{\alpha} \mathcal{FC} \mathbf{x}$$


Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\forall[\alpha].(\alpha) \rightarrow \alpha)^{\mathbf{c}} = \exists\beta. \langle (\forall[\alpha].(\beta, \alpha) \rightarrow \alpha), \beta \rangle$$

$$\alpha^{\mathbf{c}} = \alpha$$

$$\mathbf{L}\langle\tau\rangle^{\mathbf{c}} = \tau$$

$$\mathbf{CF}^{\forall[\alpha].(\alpha) \rightarrow \alpha}(\mathbf{v}) = \mathbf{pack} \langle \mathbf{unit}, \langle \mathbf{v}, () \rangle \rangle \mathbf{as} (\forall[\alpha].(\alpha) \rightarrow \alpha)^{\mathbf{c}}$$

$$\mathbf{v} = \lambda[\alpha](z:\mathbf{unit}, x:\alpha). \mathcal{CF}^{\mathbf{L}\langle\alpha\rangle}(\mathbf{v} [\mathbf{L}\langle\alpha\rangle] \mathbf{L}\langle\alpha\rangle \mathcal{FC}_{\mathbf{x}})$$

Add new type $\mathbf{L}\langle\tau\rangle$ & new value form $\mathbf{L}\langle\tau\rangle \mathcal{FC}_{\mathbf{v}}$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\forall[\alpha].(\alpha) \rightarrow \alpha)^{\langle c \rangle} = \exists\beta. \langle (\forall[\alpha].(\beta, \alpha) \rightarrow \alpha), \beta \rangle$$

$$\alpha^{\langle c \rangle} = \alpha$$

$$L\langle \tau \rangle^{\langle c \rangle} = \tau$$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\forall[\alpha].(\alpha) \rightarrow \alpha)^{\langle \mathbf{c} \rangle} = \exists\beta. \langle (\forall[\alpha].(\beta, \alpha) \rightarrow \alpha), \beta \rangle$$

$$\alpha^{\langle \mathbf{c} \rangle} = \alpha \qquad \mathbf{L}\langle \tau \rangle^{\langle \mathbf{c} \rangle} = \tau$$

$$\forall[\alpha].(\alpha) \rightarrow \alpha \mathbf{FC}(\mathbf{v}) = \lambda[\alpha](x:\alpha). \alpha\mathcal{FC}(\text{unpack } \langle \beta, y \rangle = \mathbf{v} \\ \text{in } \pi_1(y) [\alpha^{\mathbf{c}}] \pi_2(y) \mathcal{CF}^{\alpha}x)$$

Interoperability: **F** and **C**

$$\mathbf{CF}^{\tau}(\mathbf{v}) = \mathbf{v}$$

$$\tau\mathbf{FC}(\mathbf{v}) = \mathbf{v}$$

$$(\forall[\alpha].(\alpha) \rightarrow \alpha)^{\langle \mathbf{c} \rangle} = \exists\beta. \langle (\forall[\alpha].(\beta, \alpha) \rightarrow \alpha), \beta \rangle$$

$$\alpha^{\mathbf{c}} = \lceil \alpha \rceil$$

$$L\langle \tau \rangle^{\langle \mathbf{c} \rangle} = \tau$$

$$\forall[\alpha].(\alpha) \rightarrow \alpha \mathbf{FC}(\mathbf{v}) = \lambda[\alpha](x:\alpha). \alpha\mathcal{FC}(\text{unpack } \langle \beta, y \rangle = \mathbf{v} \\ \text{in } \pi_1(y) \lceil \alpha \rceil \pi_2(y) \mathcal{CF}^{\alpha} x)$$

Add new type $\lceil \alpha \rceil$ & define $\lceil \alpha \rceil[\tau/\alpha] = \tau^{\langle \mathbf{c} \rangle}$

Interoperability: **F** and **C**

$\tau \langle \mathcal{C} \rangle$ Operational Type Translation

$$\forall [\bar{\alpha}]. (\bar{\tau}) \rightarrow \tau' \langle \mathcal{C} \rangle$$

$$= \exists \beta. \langle (\forall [\bar{\alpha}]. (\beta, \overline{\tau \langle \mathcal{C} \rangle [\alpha / [\alpha]]}) \rightarrow \tau' \langle \mathcal{C} \rangle [\alpha / [\alpha]]), \beta \rangle$$

$$\alpha \langle \mathcal{C} \rangle = [\alpha] \quad \exists \alpha. \tau \langle \mathcal{C} \rangle = \exists \alpha. (\tau \langle \mathcal{C} \rangle [\alpha / [\alpha]])$$

$$\text{unit} \langle \mathcal{C} \rangle = \text{unit} \quad \mu \alpha. \tau \langle \mathcal{C} \rangle = \mu \alpha. (\tau \langle \mathcal{C} \rangle [\alpha / [\alpha]])$$

$$\text{int} \langle \mathcal{C} \rangle = \text{int} \quad \langle \tau_1, \dots, \tau_n \rangle \langle \mathcal{C} \rangle = \langle \tau_1 \langle \mathcal{C} \rangle, \dots, \tau_n \langle \mathcal{C} \rangle \rangle$$

$$L \langle \tau \rangle \langle \mathcal{C} \rangle = \tau$$

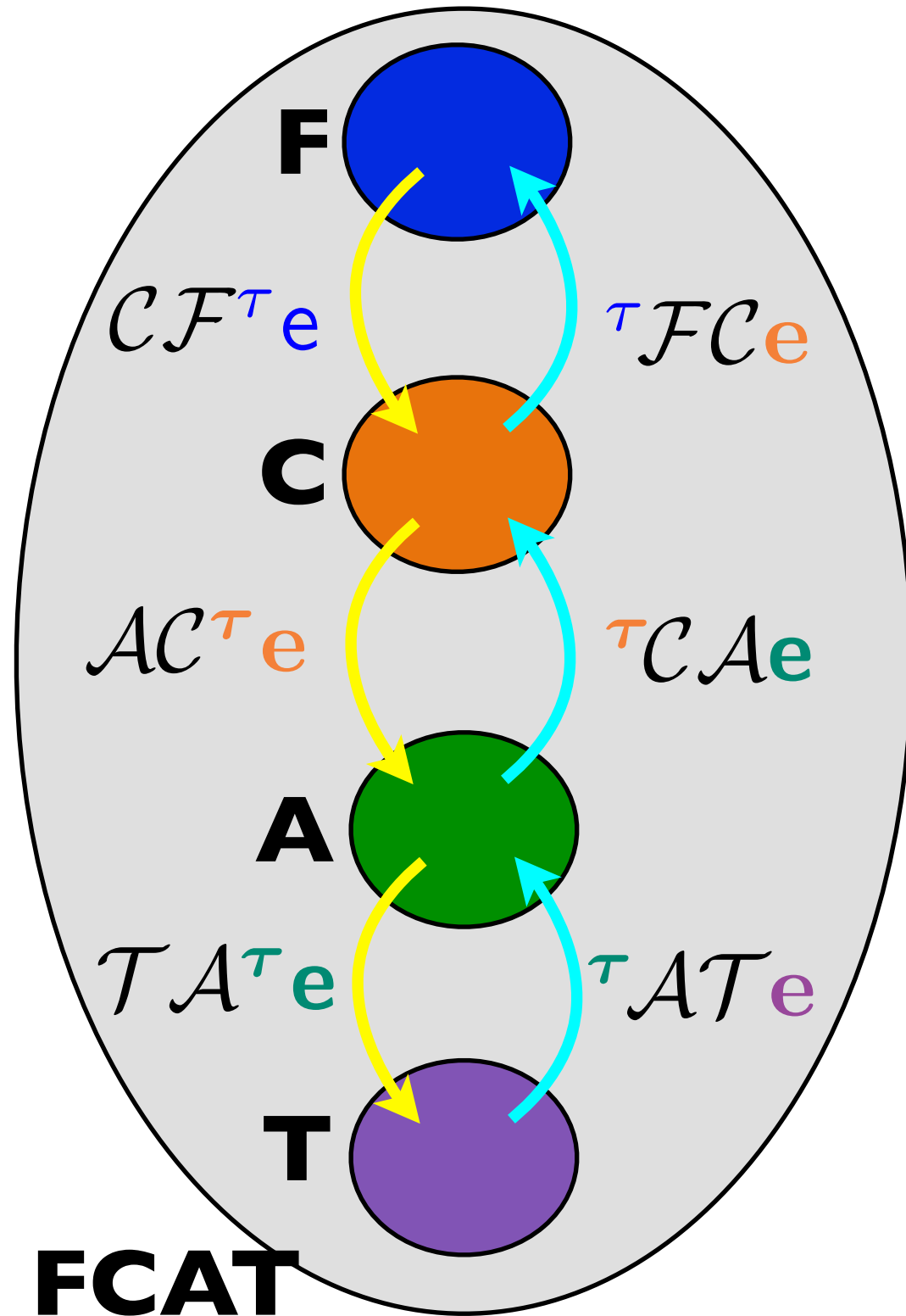
Type Substitution: $[\alpha] [\tau / \alpha] = \tau \langle \mathcal{C} \rangle$

$\Delta; \Gamma \vdash e : \tau$ Include **F** and **C** rules, with environments replaced by $\Delta; \Gamma$

$$\frac{\Delta; \Gamma \vdash e : \tau \langle \mathcal{C} \rangle}{\Delta; \Gamma \vdash \mathcal{F}^{\tau} e : \tau}$$

$$\frac{\Delta; \Gamma \vdash e : \tau}{\Delta; \Gamma \vdash \mathcal{C}^{\tau} e : \tau \langle \mathcal{C} \rangle}$$

Challenges / Roadmap



F+C: Interoperability semantics with type abstraction in both languages

C+A: Interoperability when compiler pass allocates code & tuples on heap

A+T: What is e ? What is v ?
How to define contextual equiv. for TAL *components*?
How to define logical relation?

A

$\tau ::= \alpha \mid \text{unit} \mid \text{int} \mid \exists\alpha.\tau \mid \mu\alpha.\tau \mid \text{box } \psi$

$\psi ::= \forall[\bar{\alpha}].(\bar{\tau}) \rightarrow \tau \mid \langle \tau, \dots, \tau \rangle$

$e ::= (t, \mathbf{H}) \mid t$

$t ::= x \mid () \mid n \mid t \text{ p } t \mid \text{if0 } t \ t \ t \mid \ell \mid t [] \bar{t} \mid t[\tau]$
 $\mid \text{pack } \langle \tau, t \rangle \text{ as } \exists\alpha.\tau \mid \text{unpack } \langle \alpha, x \rangle = t \text{ in } t \mid \text{fold}_{\mu\alpha.\tau} t$
 $\mid \text{unfold } t \mid \text{balloc } \langle \bar{t} \rangle \mid \text{read}[i] t$

$p ::= + \mid - \mid *$

$v ::= () \mid n \mid \text{pack } \langle \tau, v \rangle \text{ as } \exists\alpha.\tau \mid \text{fold}_{\mu\alpha.\tau} v \mid \ell \mid v[\tau]$

$\mathbf{H} ::= \cdot \mid \mathbf{H}, \ell \mapsto h$

$h ::= \lambda[\bar{\alpha}] (\bar{x} : \bar{\tau}).t \mid \langle v, \dots, v \rangle$

$\langle \mathbf{H} \mid e \rangle \longmapsto \langle \mathbf{H}' \mid e' \rangle$ Reduction Relation (selected cases)

$\langle \mathbf{H} \mid (t, \mathbf{H}') \rangle \longmapsto \langle (\mathbf{H}, \mathbf{H}') \mid t \rangle \quad \text{dom}(\mathbf{H}) \cap \text{dom}(\mathbf{H}') = \emptyset$

$\langle \mathbf{H} \mid \mathbf{E}[\ell [\bar{\tau}'] \bar{v}] \rangle \longmapsto \langle \mathbf{H} \mid \mathbf{E}[t[\bar{\tau}'/\bar{\alpha}][\bar{v}/\bar{x}]] \rangle \quad \mathbf{H}(\ell) = \lambda[\bar{\alpha}] (\bar{x} : \bar{\tau}).t$

Allocation: **C** to **A**

$\tau^{\mathcal{A}}$ Type Translation

$$\alpha^{\mathcal{A}} = \alpha \quad \forall[\bar{\alpha}].(\bar{\tau}) \rightarrow \tau'^{\mathcal{A}} = \mathbf{box} \forall[\bar{\alpha}].(\overline{\tau^{\mathcal{A}}}) \rightarrow \tau'^{\mathcal{A}}$$

$$\mathbf{unit}^{\mathcal{A}} = \mathbf{unit} \quad \exists\alpha.\tau^{\mathcal{A}} = \exists\alpha.\tau^{\mathcal{A}}$$

$$\mathbf{int}^{\mathcal{A}} = \mathbf{int} \quad \mu\alpha.\tau^{\mathcal{A}} = \mu\alpha.\tau^{\mathcal{A}}$$

$$\langle \tau_1, \dots, \tau_n \rangle^{\mathcal{A}} = \mathbf{box} \langle (\tau_1^{\mathcal{A}}), \dots, (\tau_n^{\mathcal{A}}) \rangle$$

$\Delta; \Gamma; \vdash e: \tau \rightsquigarrow (t, H: \Psi)$

where $\Delta; \Gamma \vdash e: \tau$, $\cdot \vdash H: \Psi$, and
 $\cdot; \Delta^{\mathcal{A}}; \Gamma^{\mathcal{A}} \vdash (t, H): \tau^{\mathcal{A}}$

Interoperability: **C** and **A**

$$\tau ::= \dots \mid \mathbf{L}\langle\tau\rangle$$

$$\tau ::= \dots \mid \lceil\alpha\rceil \mid \lceil\alpha\rceil$$

$$\lceil\alpha\rceil[\tau/\alpha] = (\tau\langle\mathbf{C}\rangle)\langle\mathcal{A}\rangle$$

$$\lceil\alpha\rceil[\tau/\alpha] = \tau\langle\mathcal{A}\rangle$$

$$\frac{\Psi; \Delta; \Gamma \vdash \mathbf{e} : \tau\langle\mathcal{A}\rangle}{\Psi; \Delta; \Gamma \vdash \mathcal{C}\mathcal{A}\mathbf{e} : \tau}$$

$$\frac{\Psi; \Delta; \Gamma \vdash \mathbf{e} : \tau}{\Psi; \Delta; \Gamma \vdash \mathcal{A}\mathcal{C}\tau\mathbf{e} : \tau\langle\mathcal{A}\rangle}$$

Interoperability: **C** and **A**

$$\mathbf{AC}^{\tau}(\mathbf{v}, M) = (\mathbf{v}, M')$$

$$\mathbf{AC}^{\langle \bar{\tau} \rangle}(\langle \bar{\mathbf{v}} \rangle, M_1) = (\ell, (M_{n+1}, \ell \mapsto \langle \bar{\mathbf{v}} \rangle))$$

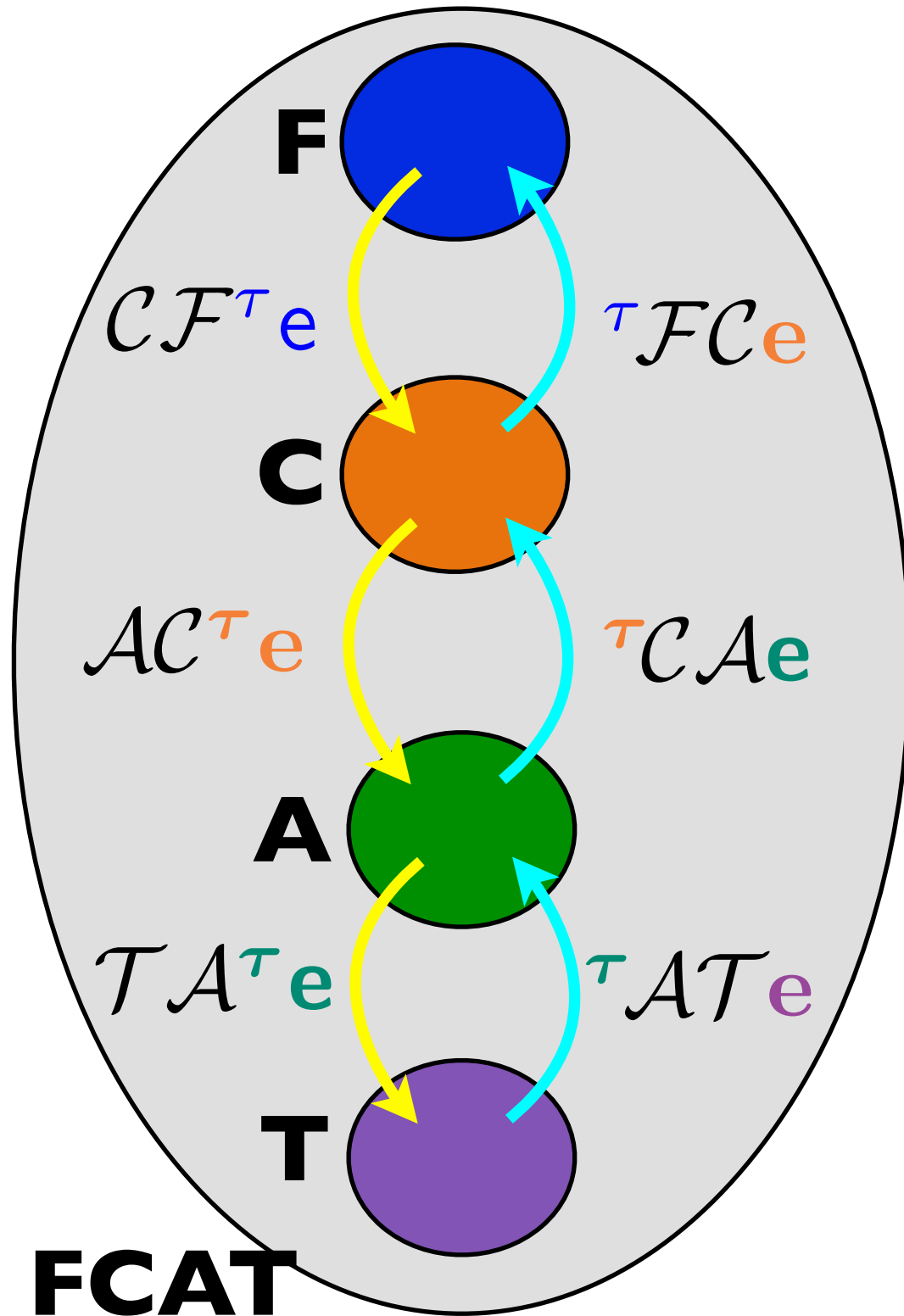
$$\text{where } \mathbf{AC}^{\tau_i}(\mathbf{v}_i, M_i) = (\mathbf{v}_i, M_{i+1})$$

$$\tau \mathbf{CA}(\mathbf{v}, M) = (\mathbf{v}, M')$$

$$\langle \bar{\tau} \rangle \mathbf{CA}(\ell, M_1) = (\langle \bar{\mathbf{v}} \rangle, M_{n+1})$$

$$\text{where } M_1(\ell) = \langle \bar{\mathbf{v}} \rangle \text{ and } \tau_i \mathbf{CA}(\mathbf{v}_i, M_i) = (\mathbf{v}_i, M_{i+1})$$

Challenges / Roadmap



F+C: Interoperability semantics with type abstraction in both languages

C+A: Interoperability when compiler pass allocates code & tuples on heap

A+T: What is e ? What is v ?
How to define contextual equiv. for TAL *components*?
How to define logical relation?

T

τ	$::= \alpha \mid \text{unit} \mid \text{int} \mid \exists\alpha.\tau \mid \mu\alpha.\tau$	<i>Type</i>
	$\mid \text{ref} \langle \tau, \dots, \tau \rangle \mid \text{box } \psi$	
ψ	$::= \forall[\Delta].\{\chi; \sigma\}^q \mid \langle \tau, \dots, \tau \rangle$	<i>Heap value type</i>
χ	$::= \cdot \mid \chi, r : \tau$	<i>Register file type</i>
σ	$::= \zeta \mid \bullet \mid \tau :: \sigma$	<i>Stack type</i>
q	$::= \epsilon \mid r \mid i \mid \text{end}[\tau; \sigma]$	<i>Return marker</i>
Δ	$::= \cdot \mid \Delta, \alpha \mid \Delta, \zeta \mid \Delta, \epsilon$	<i>Type variable environment</i>
ω	$::= \tau \mid \sigma \mid q$	<i>Instantiation of type variable</i>
r	$::= r1 \mid r2 \mid \dots \mid r7 \mid ra$	<i>Register</i>
h	$::= \text{code}[\Delta]\{\chi; \sigma\}^q.I \mid \langle w, \dots, w \rangle$	<i>Heap value</i>
w	$::= () \mid n \mid \ell \mid \text{pack} \langle \tau, w \rangle \text{ as } \exists\alpha.\tau$	<i>Word value</i>
	$\mid \text{fold}_{\mu\alpha.\tau} w \mid w[\omega]$	
u	$::= w \mid r \mid \text{pack} \langle \tau, u \rangle \text{ as } \exists\alpha.\tau$	<i>Small value</i>
	$\mid \text{fold}_{\mu\alpha.\tau} u \mid u[\omega]$	
I	$::= \iota; I \mid \text{jmp } u \mid \text{ret } q, r$	<i>Instruction sequence</i>

T

ℓ	$::=$ <code>aop r_d, r_s, u $bnz\ r, u$ $mv\ r_d, u$</code>	<i>Instruction</i>
	<code> $ralloc\ r_d, n$ $balloc\ r_d, n$ $ld\ r_d, r_s[i]$ $st\ r_d[i], r_s$</code>	
	<code> $unpack\ \langle \alpha, r_d \rangle\ u$ $unfold\ r_d, u$ $salloc\ n$ $sfree\ n$</code>	
	<code> $sld\ r_d, i$ $sst\ i, r_s$</code>	
<code>aop</code>	$::=$ <code>add sub mult</code>	<i>Arithmetic operation</i>
<code>e</code>	$::=$ <code>(I, H) I</code>	<i>Component</i>
<code>v</code>	$::=$ <code>ret q, r</code>	<i>Term value</i>
<code>E</code>	$::=$ <code>(E_I, \cdot)</code>	<i>Evaluation context</i>
<code>E_I</code>	$::=$ <code>[\cdot]</code>	<i>Instruction evaluation context</i>
<code>H</code>	$::=$ <code>\cdot $H, \ell \mapsto h$</code>	<i>Heap or Heap fragment</i>
<code>R</code>	$::=$ <code>\cdot $R, r \mapsto w$</code>	<i>Register file</i>
<code>S</code>	$::=$ <code>nil $w :: S$</code>	<i>Stack</i>
<code>M</code>	$::=$ <code>($H, R, S : \sigma$)</code>	<i>Memory</i>

$$\langle M \mid e \rangle \longmapsto \langle M' \mid e' \rangle$$

Well-typed Components in \mathbf{T}

$$\Psi; \Delta; \chi; \sigma; \mathbf{q} \vdash e : \tau; \sigma'$$

$$\frac{\Psi \vdash \mathbf{H} : \Psi_e \quad \text{boxheap}(\Psi_e) \quad \text{ret-type}(\mathbf{q}, \chi, \sigma) = \tau; \sigma' \quad (\Psi, \Psi_e); \Delta; \chi; \sigma; \mathbf{q} \vdash \mathbf{I}}{\Psi; \Delta; \chi; \sigma; \mathbf{q} \vdash (\mathbf{I}, \mathbf{H}) : \tau; \sigma'}$$

Equivalence of \mathbf{T} Components: Tricky!

Logical relations: related inputs to related outputs

$$\mathcal{V}[\tau_1 \rightarrow \tau_2] = \{(W, \lambda x.e_1, \lambda x.e_1) \mid \dots\}$$

$$\mathcal{HV}[\forall[\Delta].\{\chi; \sigma\}^q] = \{(W, \text{code}[\Delta]\{\chi; \sigma\}^q.I_1, \text{code}[\Delta]\{\chi; \sigma\}^q.I_2) \mid \dots\}$$

Equivalence of \mathbf{T} Components: Tricky!

Logical relations: related inputs to related outputs

$$\mathcal{V}[\tau_1 \rightarrow \tau_2] = \{(W, \lambda x.e_1, \lambda x.e_1) \mid \dots\}$$

$$\mathcal{HV}[\forall[\Delta].\{\chi; \sigma\}^q] = \{(W, \text{code}[\Delta]\{\chi; \sigma\}^q.I_1, \text{code}[\Delta]\{\chi; \sigma\}^q.I_2) \mid \dots\}$$

Equivalence of \mathbf{T} Components: Tricky!

Logical relations: related inputs to related outputs

$$\mathcal{V}[\tau_1 \rightarrow \tau_2] = \{(W, \lambda x.e_1, \lambda x.e_1) \mid \dots\}$$

$$\mathcal{HV}[\forall[\Delta].\{\chi; \sigma\}^q] = \{(W, \text{code}[\Delta]\{\chi; \sigma\}^q.I_1, \text{code}[\Delta]\{\chi; \sigma\}^q.I_2) \mid \dots\}$$



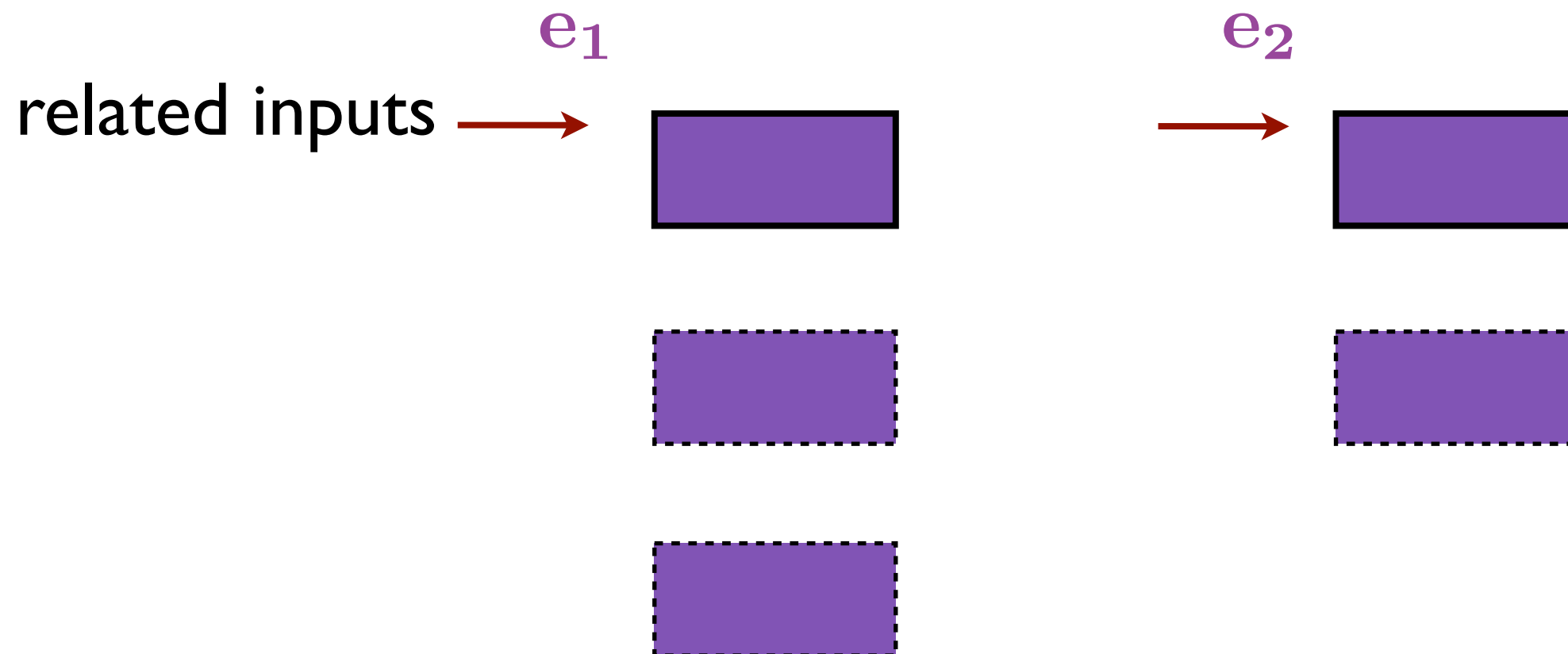
 = $\text{code}[\Delta]\{\chi; \sigma\}^q.I$

Equivalence of **T** Components: Tricky!

Logical relations: related inputs to related outputs

$$\mathcal{V}[\tau_1 \rightarrow \tau_2] = \{(W, \lambda x.e_1, \lambda x.e_1) \mid \dots\}$$

$$\mathcal{HV}[\forall[\Delta].\{\chi; \sigma\}^q] = \{(W, \text{code}[\Delta]\{\chi; \sigma\}^q.I_1, \text{code}[\Delta]\{\chi; \sigma\}^q.I_2) \mid \dots\}$$

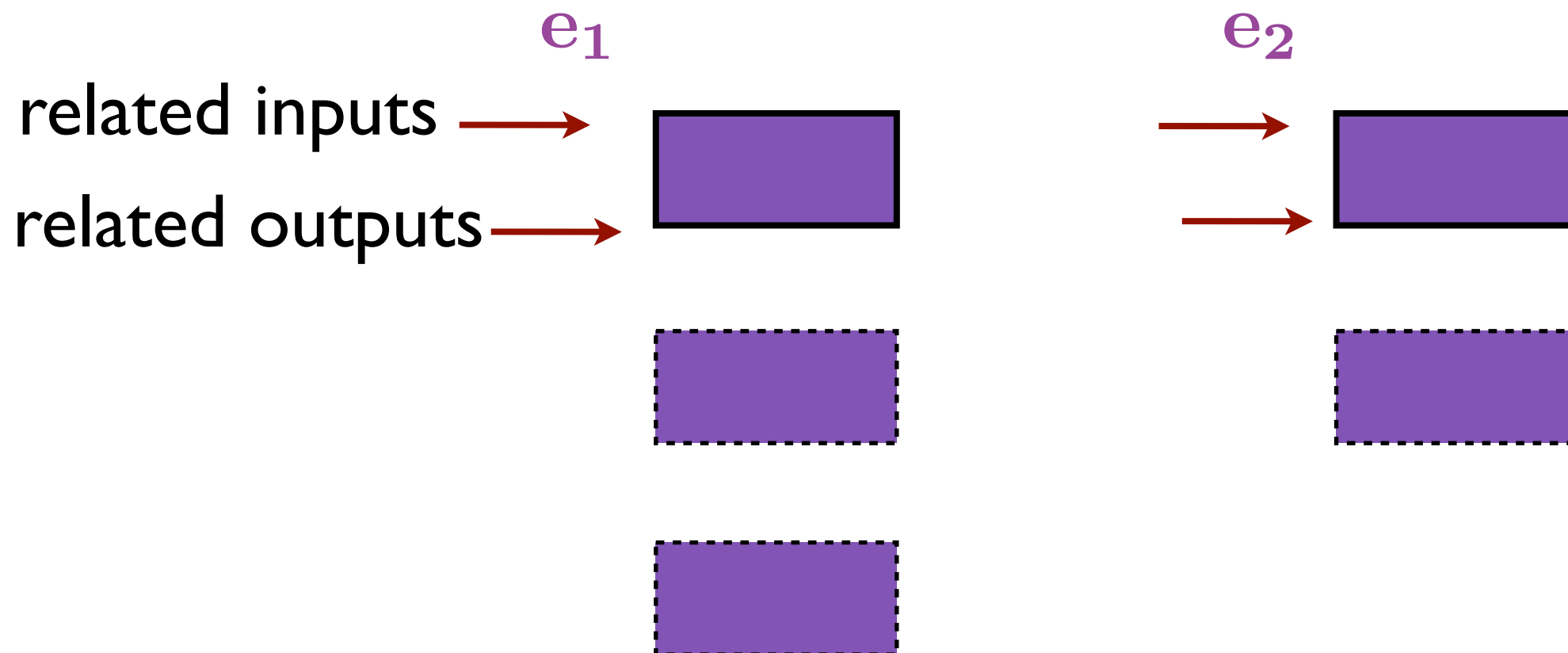


Equivalence of **T** Components: Tricky!

Logical relations: related inputs to related outputs

$$\mathcal{V}[\tau_1 \rightarrow \tau_2] = \{(W, \lambda x.e_1, \lambda x.e_1) \mid \dots\}$$

$$\mathcal{HV}[\forall[\Delta].\{\chi; \sigma\}^q] = \{(W, \text{code}[\Delta]\{\chi; \sigma\}^q.I_1, \text{code}[\Delta]\{\chi; \sigma\}^q.I_2) \mid \dots\}$$

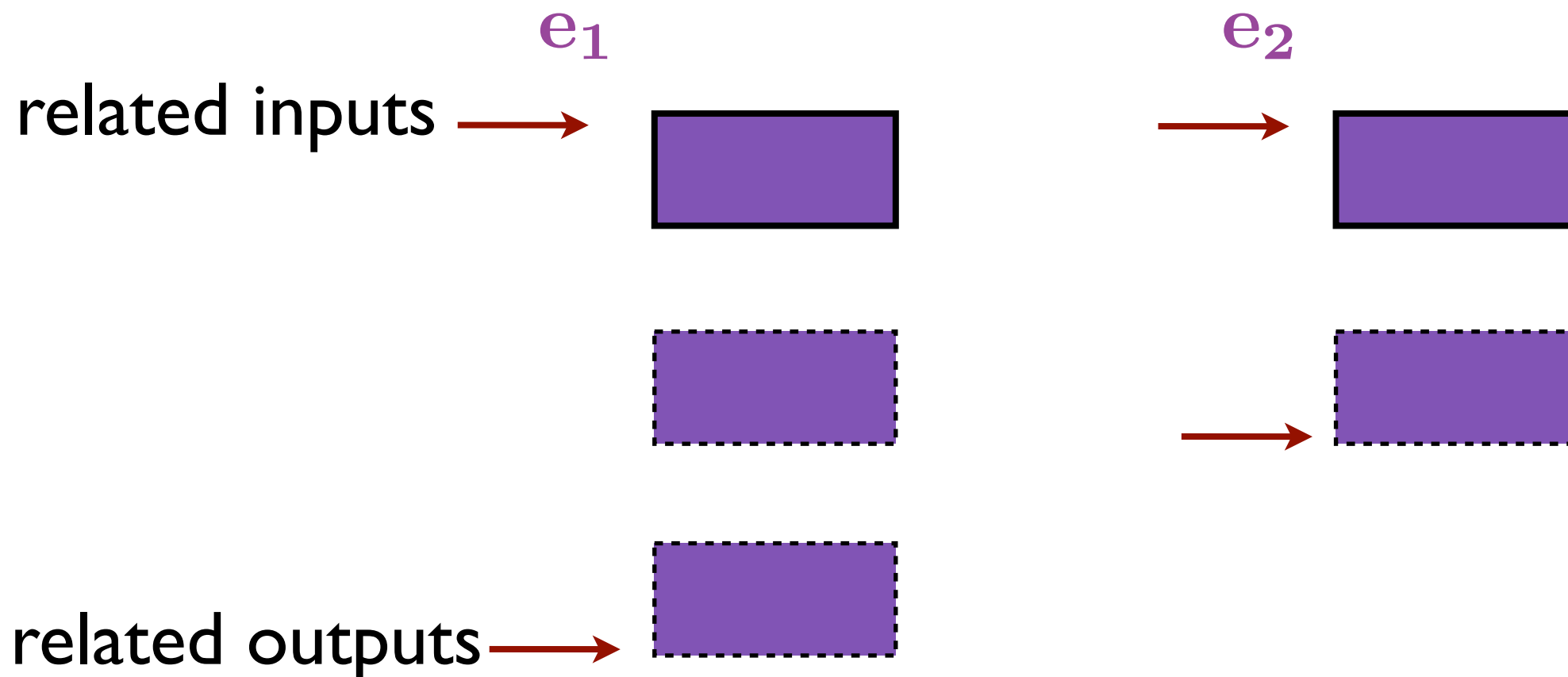


Equivalence of **T** Components: Tricky!

Logical relations: related inputs to related outputs

$$\mathcal{V}[\tau_1 \rightarrow \tau_2] = \{(W, \lambda x.e_1, \lambda x.e_1) \mid \dots\}$$

$$\mathcal{HV}[\forall[\Delta].\{\chi; \sigma\}^q] = \{(W, \text{code}[\Delta]\{\chi; \sigma\}^q.I_1, \text{code}[\Delta]\{\chi; \sigma\}^q.I_2) \mid \dots\}$$



Code Generation: **A** to **T**



Type translation

$$\begin{aligned} \text{box } \forall[\bar{\alpha}].(\tau_1, \dots, \tau_n) &\rightarrow \tau' T \\ &= \text{box } \forall[\bar{\alpha}, \zeta, \epsilon]. \\ &\quad \{ra : \text{box } \forall[].\{r1 : \tau' T ; \zeta\}^\epsilon ; \\ &\quad \tau_n T :: \dots :: \tau_1 T :: \zeta\}^{ra} \end{aligned}$$

Code Generation: **A** to **T**



Type translation

$$\begin{aligned} \text{box } \forall[\bar{\alpha}].(\tau_1, \dots, \tau_n) &\rightarrow \tau' \mathcal{T} \\ &= \text{box } \forall[\bar{\alpha}, \zeta, \epsilon]. \\ &\quad \{ra : \text{box } \forall[].\{r1 : \tau' \mathcal{T}; \zeta\}^\epsilon; \\ &\quad \tau_n \mathcal{T} :: \dots :: \tau_1 \mathcal{T} :: \zeta\}^{ra} \end{aligned}$$

$$\boxed{\Psi; \Delta; \Gamma \vdash e : \tau \rightsquigarrow e} \quad \text{where } \Psi \mathcal{T}; (\Delta \mathcal{T}, \zeta, \epsilon); \chi; \sigma; ra \vdash e : \tau \mathcal{T}; \sigma$$

for $\chi = ra : \forall[].\{r1 : \tau \mathcal{T}; \sigma\}^\epsilon$ and $\sigma = \text{order}(\Gamma, \zeta) \mathcal{T}$

Interoperability: **A** and **T**

$$\frac{\Psi; \Delta; \Gamma; \cdot; \sigma; \text{end}[\tau \langle \mathcal{T} \rangle; \sigma'] \vdash e : \tau \langle \mathcal{T} \rangle; \sigma'}{\Psi; \Delta; \Gamma; \chi; \sigma; \text{out} \vdash {}^{\tau} \mathcal{A} \mathcal{T} e : \tau; \sigma'}$$

$$\frac{{}^{\tau} \mathbf{AT}(M.\mathbf{M.R}(\mathbf{r}), M) = (\mathbf{v}, M')}{\langle M \mid E[{}^{\tau} \mathcal{A} \mathcal{T} \text{ret end}[\tau \langle \mathcal{T} \rangle; \sigma], \mathbf{r}] \rangle \longmapsto \langle M' \mid E[\mathbf{v}] \rangle}$$

Interoperability: **A** and **T**

$\iota ::= \dots \mid \text{import } r_d, \sigma \mathcal{T} \mathcal{A}^\tau e$

$$\mathbf{TA}^\tau(\mathbf{v}, M) = (\mathbf{w}, M')$$

$$\langle M \mid E[\text{import } r_d, \sigma' \mathcal{T} \mathcal{A}^\tau \mathbf{v}; \mathbf{I}] \rangle \longmapsto \langle M' \mid E[\text{mv } r_d, \mathbf{w}; \mathbf{I}] \rangle$$

Conclusions

- Compiler verification methodology that
 - guarantees correct compilation of components, not just whole programs
 - works for multi-pass compilers
 - supports reasoning about whole programs produced by linking with arbitrary target code
- **Interoperability semantics** provides specification of when source and target code are related
 - easier to understand compiler correctness theorem
 - but, have to get all the languages to fit together!

Questions?
