

Approximate Relational Reasoning for Higher-Order Probabilistic Programs

Philipp G. Haselwarter¹, Kwing Hei Li¹, Alejandro Aguirre¹, Simon Oddershede Gregersen², Joseph Tassarotti², and Lars Birkedal¹



Our Contribution: Approxis

- Relational higher-order separation logic for **approximate program refinement** for **Randomized ML**!

Our Contribution: Approxis

- Relational higher-order separation logic for **approximate program refinement** for **Randomized ML**!
 - **error probability** ϵ represented as a first-class separation logic **resource** $\text{!}(\epsilon)$
 - full power of non-probabilistic reasoning preserved (invariants, ghost state, ...)
 - expressive probabilistic reasoning rules manipulate $\text{!}(\epsilon)$ locally

Our Contribution: Approxis

- Relational higher-order separation logic for **approximate program refinement** for **Randomized ML**!
 - **error probability** ϵ represented as a first-class separation logic **resource** $\text{\textit{!}}(\epsilon)$
 - full power of non-probabilistic reasoning preserved (invariants, ghost state, ...)
 - expressive probabilistic reasoning rules manipulate $\text{\textit{!}}(\epsilon)$ locally
- New **logical relation** for program approximation & contextual equivalence

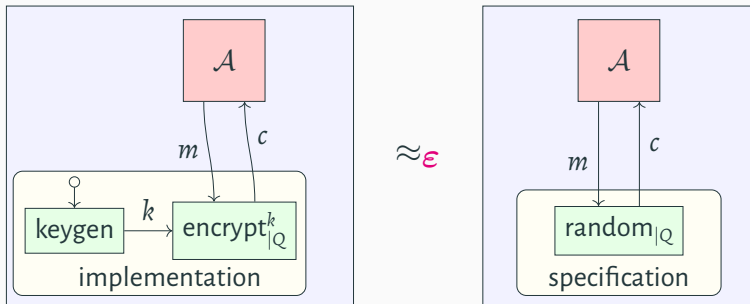
Our Contribution: Approxis

- Relational higher-order separation logic for **approximate program refinement** for **Randomized ML**!
 - **error probability** ϵ represented as a first-class separation logic **resource** $\text{\textit{!}}(\epsilon)$
 - full power of non-probabilistic reasoning preserved (invariants, ghost state, ...)
 - expressive probabilistic reasoning rules manipulate $\text{\textit{!}}(\epsilon)$ locally
- New **logical relation** for program approximation & contextual equivalence
- Case studies
 - cryptographic security: symmetric key encryption, ...
 - rejection samplers: sampling a node from a B+ tree, ...

Our Contribution: Approxis

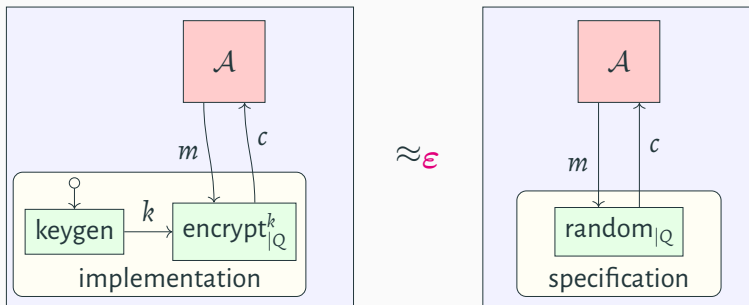
- Relational higher-order separation logic for **approximate program refinement** for **Randomized ML**!
 - **error probability** ϵ represented as a first-class separation logic **resource** $\text{!}(\epsilon)$
 - full power of non-probabilistic reasoning preserved (invariants, ghost state, ...)
 - expressive probabilistic reasoning rules manipulate $\text{!}(\epsilon)$ locally
- New **logical relation** for program approximation & contextual equivalence
- Case studies
 - cryptographic security: symmetric key encryption, ...
 - rejection samplers: sampling a node from a B+ tree, ...
- Fully mechanised in  and 

Example from Cryptography



Security: “Encrypted messages c appear random if the key k is unknown to $\mathcal{A}$.”

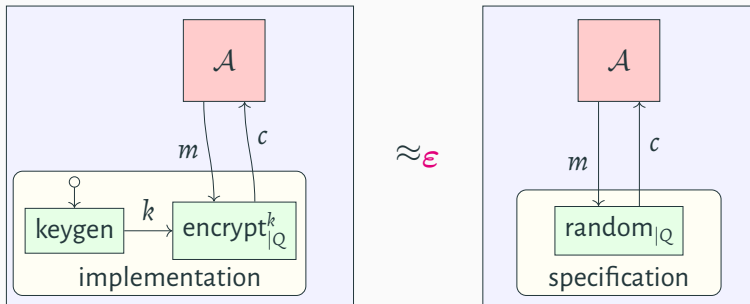
Example from Cryptography



Security: “Encrypted messages c appear random if the key k is unknown to $\mathcal{A}$.”

Game: Adversary $\mathcal{A}$ can query an abstract encryption “oracle” with Q messages m .

Example from Cryptography

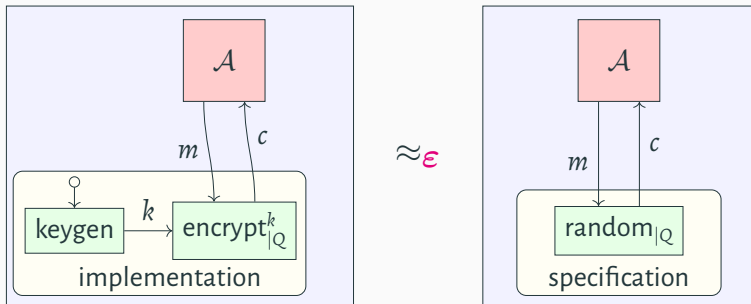


Security: “Encrypted messages c appear random if the key k is unknown to $\mathcal{A}$.”

Game: Adversary $\mathcal{A}$ can query an abstract encryption “oracle” with Q messages m .

Goal: Return **true** for the implementation and **false** for the specification.

Example from Cryptography



Security: “Encrypted messages c appear random if the key k is unknown to $\mathcal{A}$.”

Game: Adversary $\mathcal{A}$ can query an abstract encryption “oracle” with Q messages m .

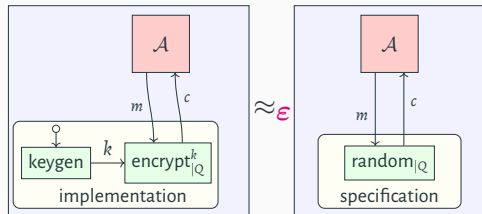
Goal: Return **true** for the implementation and **false** for the specification.

Claim: No adversary $\mathcal{A}$ can distinguish with probability better than ϵ !

Security in Approxis

Intuition:

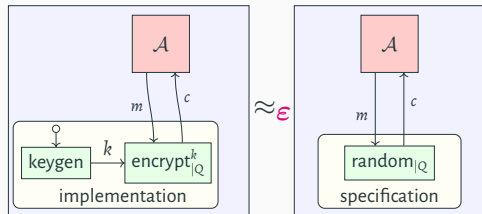
“encrypt behaves (almost) like the uniform distribution on ciphertexts”



Security in Approxis

Intuition:

“encrypt behaves (almost) like the uniform distribution on ciphertexts”

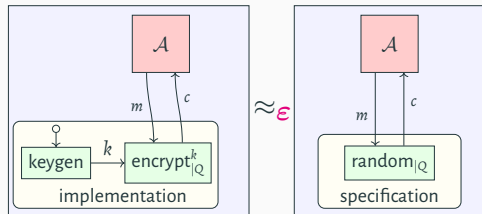


Mathematically:
$$\Pr[\mathcal{A}(\text{encrypt}_Q^k) \Downarrow \text{true}] = \Pr[\mathcal{A}(\text{random}_Q) \Downarrow \text{true}] \pm \epsilon$$

Security in Approxis

Intuition:

“encrypt behaves (almost) like the uniform distribution on ciphertexts”

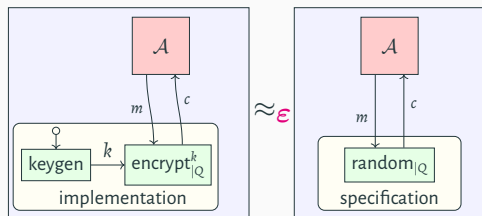


Mathematically: $\Pr[\mathcal{A}(\text{encrypt}_{|Q}^k) \Downarrow \text{true}] = \Pr[\mathcal{A}(\text{random}_{|Q}) \Downarrow \text{true}] \pm \epsilon$

Security in Approxis

Intuition:

“encrypt behaves (almost) like the uniform distribution on ciphertexts”



Mathematically: $\Pr[\mathcal{A}(\text{encrypt}_Q^k) \Downarrow \text{true}] = \Pr[\mathcal{A}(\text{random}_Q) \Downarrow \text{true}] \pm \epsilon$

In Approxis: $\text{!}(\epsilon) \multimap \text{rwp } \mathcal{A}(\text{encrypt}_Q^k) \lesssim \mathcal{A}(\text{random}_Q) \{b_1 = b_2\}$

Reasoning about Error Probability: Error Credits

- Error as first class **resource** [Eris, ICFP'24]: $\not\vdash(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\not\vdash(\epsilon) \vdash P$: “ P holds **up to error ϵ** ” *i.e.* $\Pr[\neg P] \leq \epsilon$

Reasoning about Error Probability: Error Credits

- Error as first class **resource** [Eris, ICFP'24]: $\zeta(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\zeta(\epsilon) \vdash P$: “ P holds **up to error ϵ** ” *i.e.* $\Pr[\neg P] \leq \epsilon$

- Approxis is flexible about when to “spend” this budget (*e.g.*, amortization)

Reasoning about Error Probability: Error Credits

- Error as first class **resource** [Eris, ICFP'24]: $\text{!}(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\text{!}(\epsilon) \vdash P$: “ P holds **up to error ϵ** ” *i.e.* $\Pr[\neg P] \leq \epsilon$

- Approxis is flexible about when to “spend” this budget (*e.g.*, amortization)
- Laws: $\text{!}(\epsilon_1 + \epsilon_2) \dashv\vdash \text{!}(\epsilon_1) * \text{!}(\epsilon_2)$ $\text{!}(1) \vdash \perp$

Reasoning about Error Probability: Error Credits

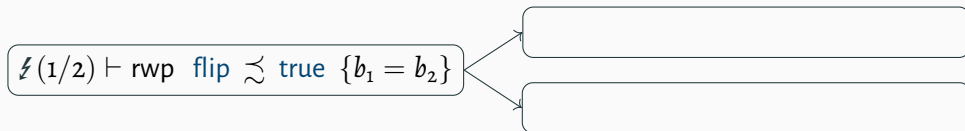
- Error as first class **resource** [Eris, ICFP'24]: $\text{!}(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\text{!}(\epsilon) \vdash P$: “ P holds **up to error** ϵ ” *i.e.* $\Pr[\neg P] \leq \epsilon$

- Approxis is flexible about when to “spend” this budget (*e.g.*, amortization)

- Laws: $\text{!}(\epsilon_1 + \epsilon_2) \dashv\vdash \text{!}(\epsilon_1) * \text{!}(\epsilon_2)$ $\text{!}(1) \vdash \perp$

- Error budget can increase during execution if expectation is preserved, *e.g.*:



Reasoning about Error Probability: Error Credits

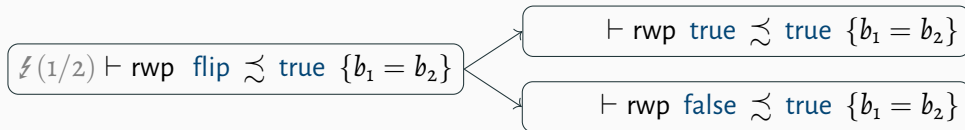
- Error as first class **resource** [Eris, ICFP'24]: $\text{!}(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\text{!}(\epsilon) \vdash P$: “ P holds **up to error** ϵ ” *i.e.* $\Pr[\neg P] \leq \epsilon$

- Approxis is flexible about when to “spend” this budget (*e.g.*, amortization)

- Laws: $\text{!}(\epsilon_1 + \epsilon_2) \dashv\vdash \text{!}(\epsilon_1) * \text{!}(\epsilon_2)$ $\text{!}(1) \vdash \perp$

- Error budget can increase during execution if expectation is preserved, *e.g.*:



Reasoning about Error Probability: Error Credits

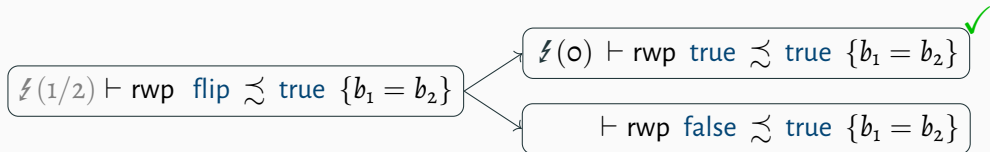
- Error as first class **resource** [Eris, ICFP'24]: $\text{!}(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\text{!}(\epsilon) \vdash P$: “ P holds **up to error** ϵ ” i.e. $\Pr[\neg P] \leq \epsilon$

- Approxis is flexible about when to “spend” this budget (e.g., amortization)

- Laws: $\text{!}(\epsilon_1 + \epsilon_2) \dashv\vdash \text{!}(\epsilon_1) * \text{!}(\epsilon_2)$ $\text{!}(1) \vdash \perp$

- Error budget can increase during execution if expectation is preserved, e.g.:



Reasoning about Error Probability: Error Credits

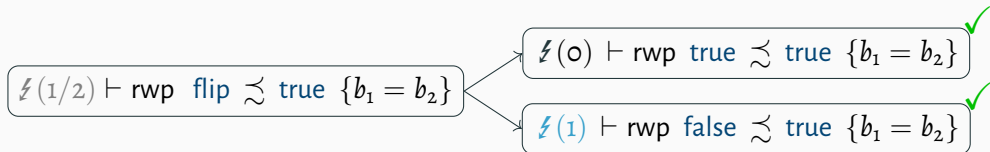
- Error as first class **resource** [Eris, ICFP'24]: $\text{!}(\epsilon)$ “up to ϵ ” $\epsilon \in [0, 1]$

Budget intuition for $\text{!}(\epsilon) \vdash P$: “ P holds **up to error** ϵ ” i.e. $\Pr[\neg P] \leq \epsilon$

- Approxis is flexible about when to “spend” this budget (e.g., amortization)

- Laws: $\text{!}(\epsilon_1 + \epsilon_2) \dashv\vdash \text{!}(\epsilon_1) * \text{!}(\epsilon_2)$ $\text{!}(1) \vdash \perp$

- Error budget can increase during execution if expectation is preserved, e.g.:



Reasoning about Known and Unknown Code

$$\nexists(\epsilon) \multimap \text{rwp } \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) \{b_1 = b_2\}$$

Reasoning about Known and Unknown Code

$$\neg(\epsilon) \multimap \text{rwp } \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) \{b_1 = b_2\}$$

- Reason **separately** about $\mathcal{A}$ and oracles

Reasoning about Known and Unknown Code

$$\not\approx(\epsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{ \phi \} \quad (1)$$

(2)

$$\not\approx(\epsilon) \multimap \text{rwp } \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) \{ b_1 = b_2 \}$$

- Reason **separately** about $\mathcal{A}$ and oracles
- For $\text{encrypt}_{|Q}^k$ and $\text{random}_{|Q}$: use **refinement logic rules** to prove (1) invariants, local state, probabilistic rules for errors, ...

Reasoning about Known and Unknown Code

$$\not\vdash(\epsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{ \phi \} \quad (1)$$

$$\forall f_1, f_2. \phi(f_1, f_2) \multimap \text{rwp } \mathcal{A}(f_1) \lesssim \mathcal{A}(f_2) \{ b_1 = b_2 \} \quad (2)$$

$$\not\vdash(\epsilon) \multimap \text{rwp } \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) \{ b_1 = b_2 \}$$

- Reason **separately** about $\mathcal{A}$ and oracles
- For $\text{encrypt}_{|Q}^k$ and $\text{random}_{|Q}$: use **refinement logic rules** to prove (1) invariants, local state, probabilistic rules for errors, ...
- But we don't know the code of $\mathcal{A}$

Reasoning about Known and Unknown Code

$$\not\vdash(\epsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{ \phi \} \quad (1)$$

$$\forall f_1, f_2. \phi(f_1, f_2) \multimap \text{rwp } \mathcal{A}(f_1) \lesssim \mathcal{A}(f_2) \{ b_1 = b_2 \} \quad (2)$$

$$\not\vdash(\epsilon) \multimap \text{rwp } \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) \{ b_1 = b_2 \}$$

- Reason **separately** about $\mathcal{A}$ and oracles
- For $\text{encrypt}_{|Q}^k$ and $\text{random}_{|Q}$: use **refinement logic rules** to prove (1)
invariants, local state, probabilistic rules for errors, ...
- But we don't know the code of $\mathcal{A} \implies$ use **type** of $\mathcal{A}$ to derive (2) !

Approximate Refinement: Errors & Invariants

$$\mathcal{I}(\varepsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{ \phi \}$$

Approximate Refinement: Errors & Invariants

$$\mathcal{E}(\epsilon) \multimap \text{rwp encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{\phi\}$$

- The oracles satisfy an invariant. Initially, $i = 0$ and $\epsilon = \sum_{k=0}^{Q-1} \frac{k}{2N} = \frac{(Q-1)Q}{2N}$.

$$\text{Inv} \triangleq \exists (i \leq Q) . \text{counter} \mapsto i * \text{counter}' \mapsto_s i * \mathcal{E}\left(\frac{(Q-1)Q}{2N} - \frac{(i-1)i}{2N}\right)$$

Approximate Refinement: Errors & Invariants

$$\text{!}(\epsilon) \multimap \text{rwp encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{\phi\}$$

- The oracles satisfy an invariant. Initially, $i = 0$ and $\epsilon = \sum_{k=0}^{Q-1} \frac{k}{2N} = \frac{(Q-1)Q}{2N}$.

$$\text{Inv} \triangleq \exists (i \leq Q) . \text{counter} \mapsto i * \text{counter}' \mapsto_s i * \text{!} \left(\frac{(Q-1)Q}{2N} - \frac{(i-1)i}{2N} \right)$$

- We transfer ownership of the initial error credits into the invariant

$$\boxed{\text{Inv}}^\gamma \multimap \text{rwp encrypt}_{|Q}^k \lesssim \text{random}_{|Q} \{\phi\}$$

Approximate Refinement: Error Resource Management

Every invocation of the oracles preserves the invariant.

- Increment counters: $\text{counter} \mapsto i+1 * \text{counter}' \mapsto_s i+1 * \text{?}(\varepsilon_i)$

Approximate Refinement: Error Resource Management

Every invocation of the oracles preserves the invariant.

- Increment counters: $\text{counter} \mapsto i+1 \ * \ \text{counter}' \mapsto_s i+1 \ * \ \text{!}(\epsilon_i)$

$$\frac{\text{counter} \mapsto i \quad \text{counter} \mapsto i \multimap \text{rwp } i \lesssim e \{ \Phi \}}{\text{rwp } !\text{counter} \lesssim e \{ \Phi \}}$$

Approximate Refinement: Error Resource Management

Every invocation of the oracles preserves the invariant.

- Increment counters: $\text{counter} \mapsto i+1 * \text{counter}' \mapsto_s i+1 * \text{!}(\epsilon_i)$
- Split credits: $\text{!}(\epsilon_i) \multimap \text{!}(\epsilon_{i+1}) * \text{!}(i/N)$

$$\frac{\text{counter} \mapsto i \quad \text{counter} \mapsto i \multimap \text{rwp } i \lesssim e \{ \Phi \}}{\text{rwp } !\text{counter} \lesssim e \{ \Phi \}} \quad \begin{array}{l} \text{!}(\epsilon + \epsilon') \dashv\vdash \text{!}(\epsilon) * \text{!}(\epsilon') \\ \epsilon_i = \epsilon_{i+1} + i/N \end{array}$$

Approximate Refinement: Error Resource Management

Every invocation of the oracles preserves the invariant.

- Increment counters: $\text{counter} \mapsto i+1 * \text{counter}' \mapsto_s i+1 * \mathcal{I}(\epsilon_i)$
- Split credits: $\mathcal{I}(\epsilon_i) \multimap \mathcal{I}(\epsilon_{i+1}) * \mathcal{I}(i/N)$
- Spend $\mathcal{I}(\epsilon_{i+1})$ credits to re-establish invariant with $i+1$

$$\frac{\text{counter} \mapsto i \quad \text{counter} \mapsto i \multimap \text{rwp } i \lesssim e \{ \Phi \}}{\text{rwp } !\text{counter} \lesssim e \{ \Phi \}} \quad \begin{array}{l} \mathcal{I}(\epsilon + \epsilon') \dashv\vdash \mathcal{I}(\epsilon) * \mathcal{I}(\epsilon') \\ \epsilon_i = \epsilon_{i+1} + i/N \end{array}$$

Approximate Refinement: Error Resource Management

Every invocation of the oracles preserves the invariant.

- Increment counters: $\text{counter} \mapsto i+1 * \text{counter}' \mapsto_s i+1 * \mathcal{L}(\epsilon_i)$
- Split credits: $\mathcal{L}(\epsilon_i) \multimap \mathcal{L}(\epsilon_{i+1}) * \mathcal{L}(i/N)$
- Spend $\mathcal{L}(\epsilon_{i+1})$ credits to re-establish invariant with $i+1$
- Spend $\mathcal{L}(i/N)$ credits to sample a fresh nonce $n \notin \text{nonces}$

$$\frac{\text{counter} \mapsto i \quad \text{counter} \mapsto i \multimap \text{rwp } i \lesssim e \{ \Phi \}}{\text{rwp } !\text{counter} \lesssim e \{ \Phi \}} \quad \begin{array}{l} \mathcal{L}(\epsilon + \epsilon') \dashv\vdash \mathcal{L}(\epsilon) * \mathcal{L}(\epsilon') \\ \epsilon_i = \epsilon_{i+1} + i/N \end{array}$$

Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\ell(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{\phi\}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{\phi\}}$$

Approximate Refinement: Errors Credits for Random Sampling

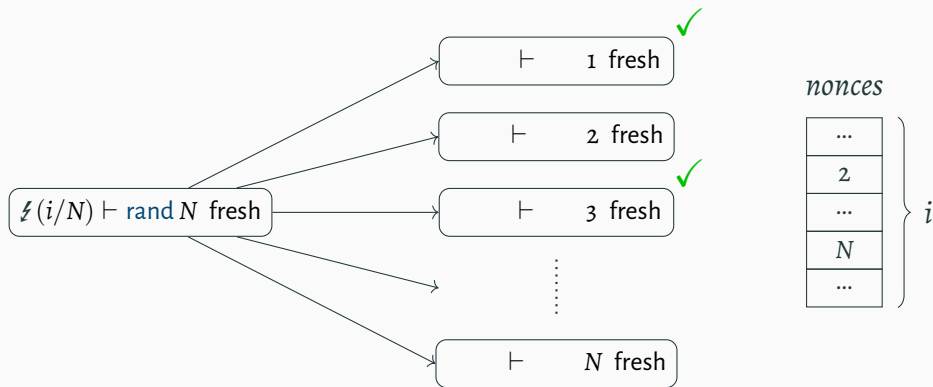
$$\frac{\ell(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{\phi\}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{\phi\}}$$

Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\ell(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{\phi\}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{\phi\}}$$

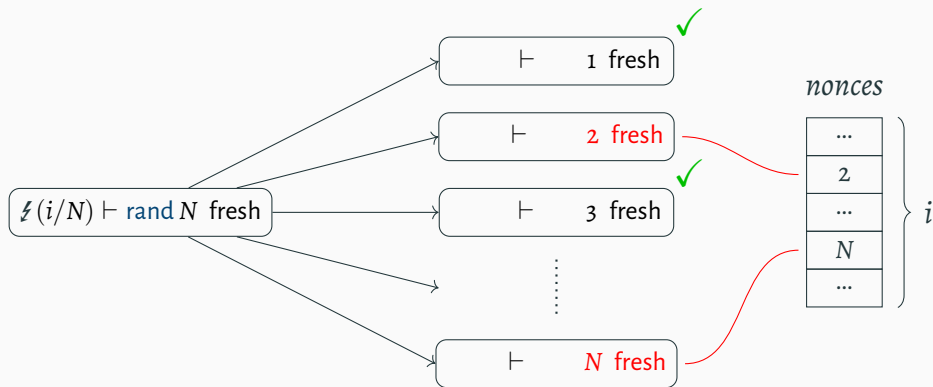
Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\not\vdash(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{\phi\}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{\phi\}}$$



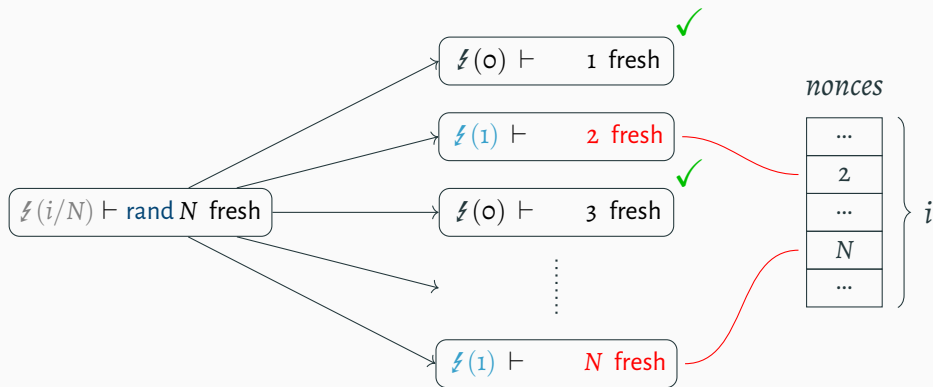
Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\not\vdash(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{\phi\}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{\phi\}}$$



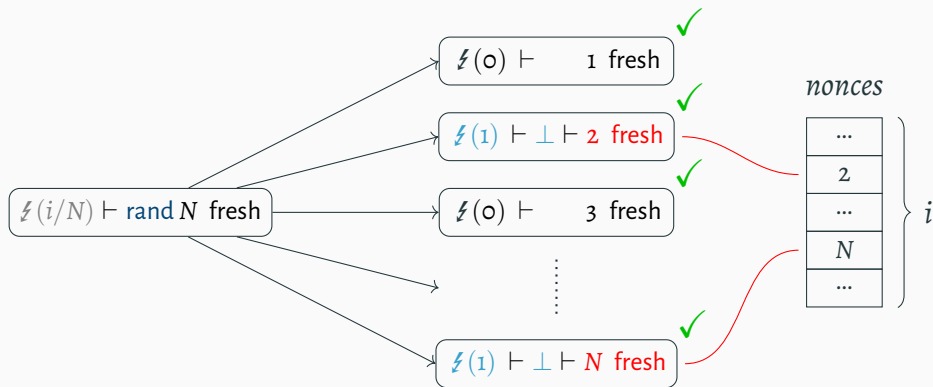
Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\not\vdash(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{\phi\}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{\phi\}}$$



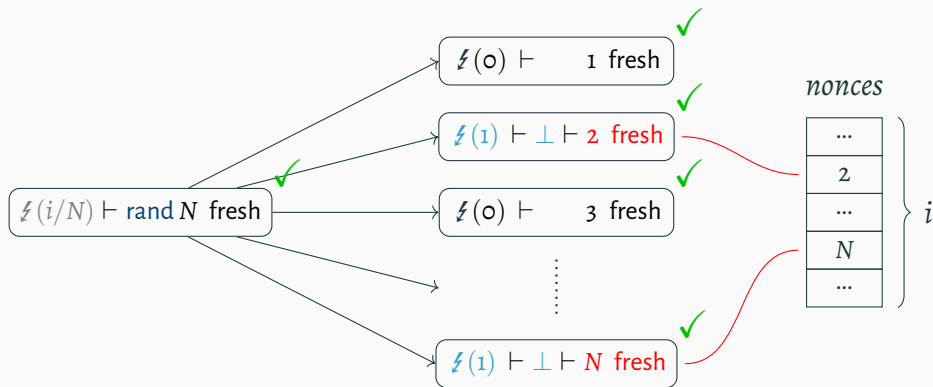
Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\text{!}(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{ \phi \}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{ \phi \}}$$



Approximate Refinement: Errors Credits for Random Sampling

$$\frac{\not\vdash(i/N) \quad i = \text{length}(\text{nonces}) \quad \forall n < N. n \notin \text{nonces} \multimap \text{rwp } n \lesssim n \{ \phi \}}{\text{rwp } \text{rand } N \lesssim \text{rand } N \{ \phi \}}$$



Refinement via Logical Relation

$$\text{rwp } \mathcal{A} \lesssim \mathcal{A} \{ \psi \}$$

- Code of $\mathcal{A}$ unknown

Refinement via Logical Relation

$$\text{rwp } \mathcal{A} \lesssim \mathcal{A} \{ \psi \}$$

- Code of $\mathcal{A}$ unknown
- Language property: all well-typed programs respect abstraction

Refinement via Logical Relation

$$\frac{\models \mathcal{A} \lesssim \mathcal{A} : (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool}}{\text{rwp } \mathcal{A} \lesssim \mathcal{A} \{ \psi \}}$$

- Code of $\mathcal{A}$ unknown
- Language property: all well-typed programs respect abstraction
- Captured via a **type-indexed logical relation**: $\models e_1 \lesssim e_2 : \tau$

Refinement via Logical Relation

$$\frac{\models \mathcal{A} \lesssim \mathcal{A} : (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool}}{\text{rwp } \mathcal{A} \lesssim \mathcal{A} \{ \psi \}}$$

- Code of $\mathcal{A}$ unknown
- Language property: all well-typed programs respect abstraction
- Captured via a **type-indexed logical relation**: $\models e_1 \lesssim e_2 : \tau$
- Logical relations model of the type system defined in terms of refinement

Refinement via Logical Relation

$$\frac{\models \mathcal{A} \lesssim \mathcal{A} : (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool}}{\text{rwp } \mathcal{A} \lesssim \mathcal{A} \{ \llbracket (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool} \rrbracket \}}$$

- Code of $\mathcal{A}$ unknown
- Language property: all well-typed programs respect abstraction
- Captured via a **type-indexed logical relation**: $\models e_1 \lesssim e_2 : \tau$
- Logical relations model of the type system defined in terms of refinement
- **Derive a rwp for $\mathcal{A}$ purely from its type signature**
 $\mathcal{A} : (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool}$

Summary: Approximate Refinement

- Approximate refinement of randomized programs

$$\mathbb{Z}(\epsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \precsim \text{random}_{|Q} \{ \phi \}$$

Summary: Approximate Refinement

- Approximate refinement of randomized programs

$$\mathbb{Z}(\epsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \precsim \text{random}_{|Q} \{ \phi \}$$

- Typed logical relation for unknown code

$$\models \mathcal{A} \precsim \mathcal{A} : (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool}$$

Summary: Approximate Refinement

- Approximate refinement of randomized programs

$$\mathcal{Z}(\epsilon) \multimap \text{rwp } \text{encrypt}_{|Q}^k \precsim \text{random}_{|Q} \{ \phi \}$$

- Typed logical relation for unknown code

$$\models \mathcal{A} \precsim \mathcal{A} : (\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool}$$

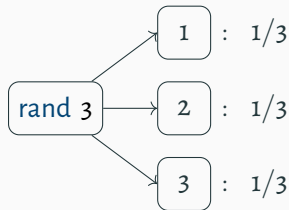
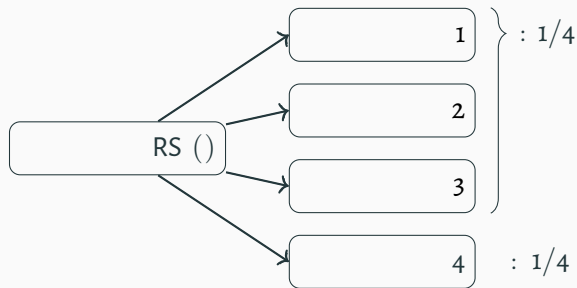
- By the adequacy theorem: encryption is secure

$$\Pr \left[\mathcal{A}(\text{encrypt}_{|Q}^k) \Downarrow \text{true} \right] = \Pr \left[\mathcal{A}(\text{random}_{|Q}) \Downarrow \text{true} \right] \pm \epsilon$$

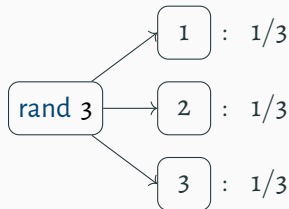
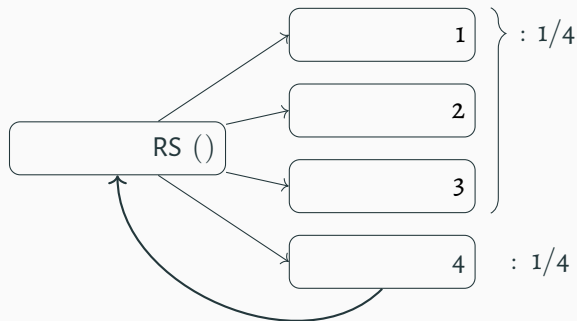
Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand}_4 \text{ in} \quad \approx \quad \text{rand}_3 : \text{int}$$

Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand}_4 \text{ in} \quad \approx \quad \text{rand}_3 : \text{int}$$


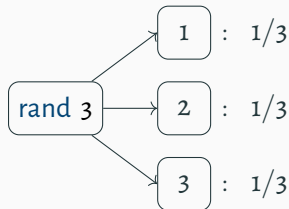
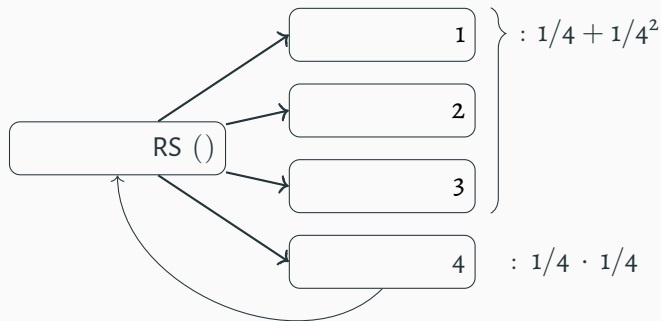
Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand}_4 \text{ in} \quad \approx \quad \text{rand}_3 : \text{int}$$


Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in}$$
$$\text{if } k < 4 \text{ then } k \text{ else RS } ()$$

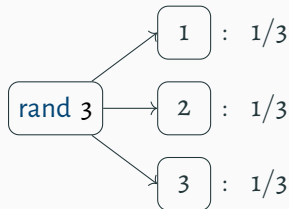
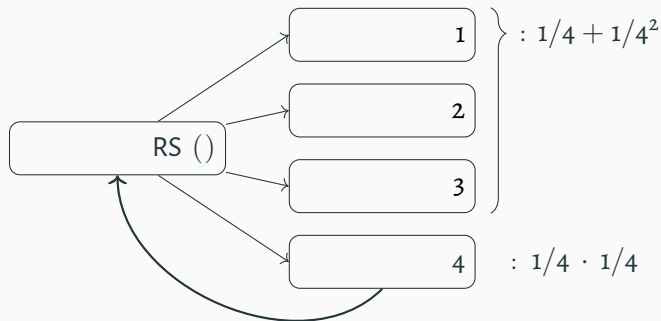
```
28     rand3 : int
```



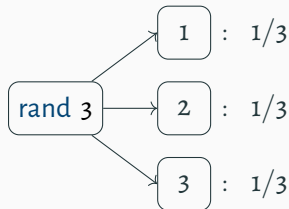
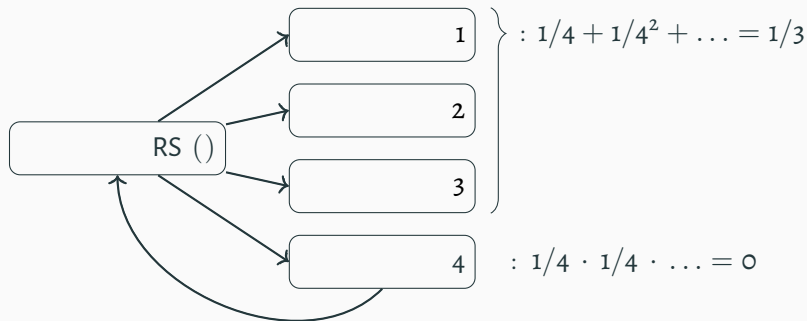
Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in}$$
$$\text{if } k < 4 \text{ then } k \text{ else RS } ()$$

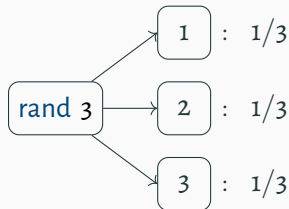
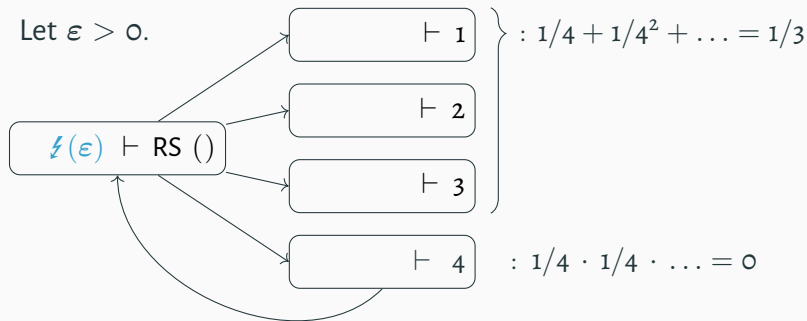
```
28     rand3 : int
```



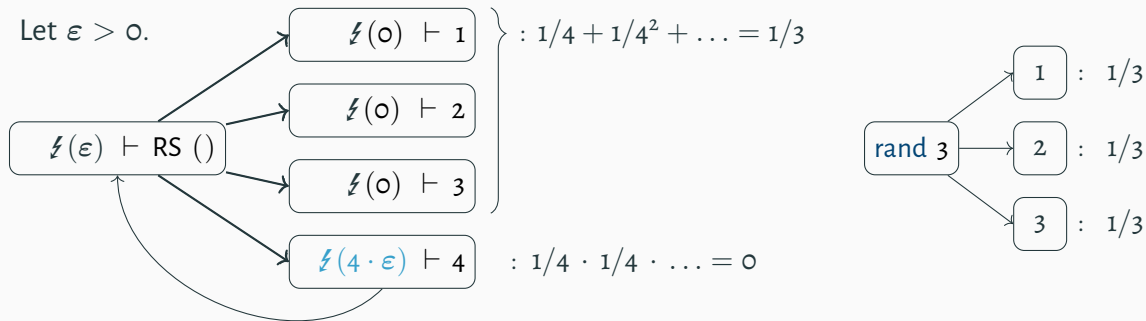
Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in} \quad \sim \quad \text{rand } 3 : \text{int}$$


Exact Contextual Refinement by Approximation in the Limit

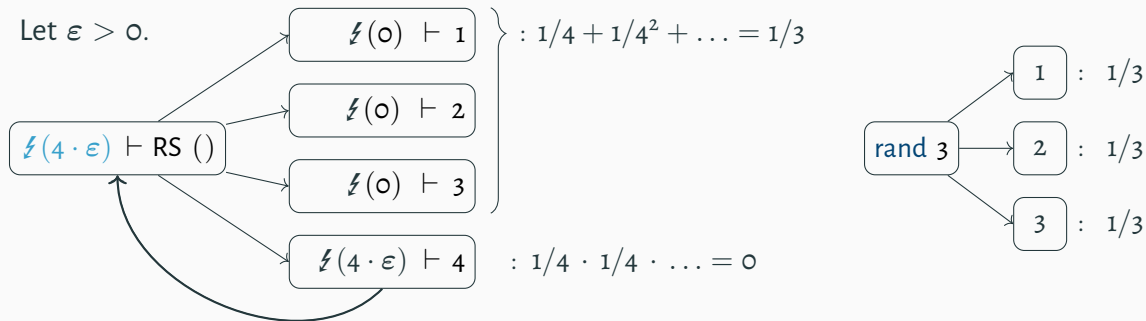
$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in} \quad \sim \quad \text{rand } 3 : \text{int}$$


Exact Contextual Refinement by Approximation in the Limit

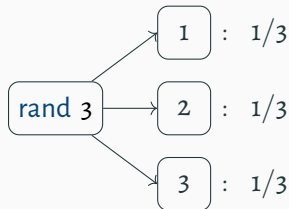
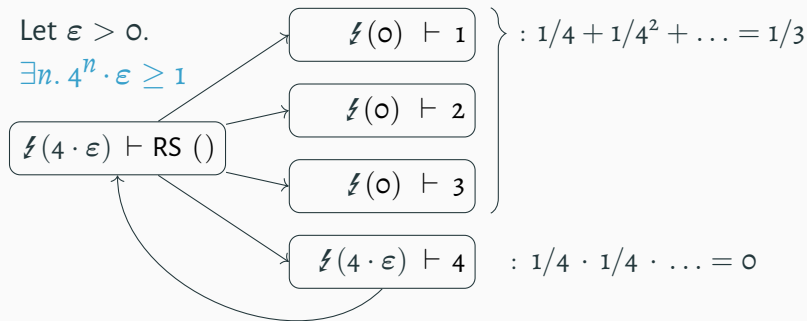
$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in} \quad \sim \quad \text{rand } 3 : \text{int}$$


Exact Contextual Refinement by Approximation in the Limit

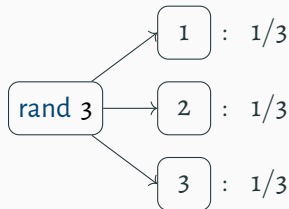
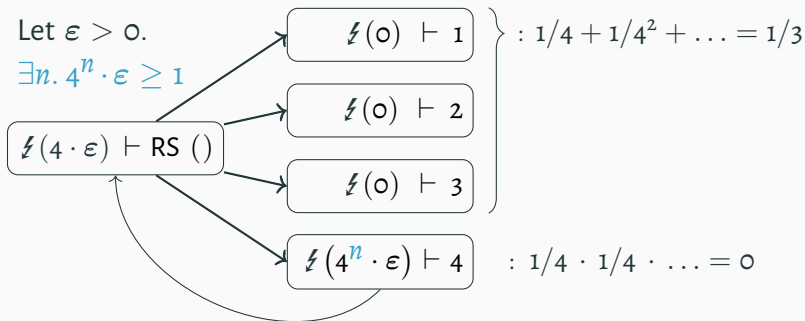
$\models$ $\text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in}$
 $\text{if } k < 4 \text{ then } k \text{ else RS } ()$ $\approx$ $\text{rand } 3 : \text{int}$



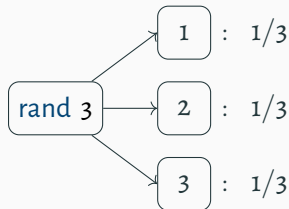
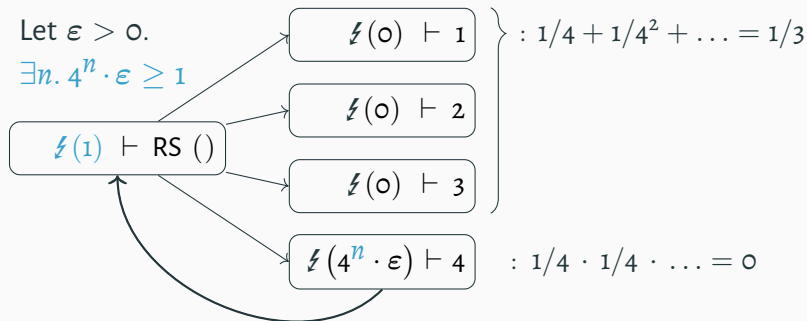
Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in} \quad \sim \quad \text{rand } 3 : \text{int}$$


Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in} \quad \rightsquigarrow \quad \text{rand } 3 \quad : \quad \text{int}$$


Exact Contextual Refinement by Approximation in the Limit

$$\models \text{rec RS } x = \text{let } k = \text{rand } 4 \text{ in} \quad \sim \quad \text{rand } 3 \quad : \quad \text{int}$$


The Bigger Picture

Related projects:

- Clutch: refinement, asynchronous sampling via tapes
- Eris: unary, bound probability of “bad events” via error credits
- Caliper: termination-preserving refinement
- Tachis: expected cost (e.g., time or entropy) via cost credits

The Bigger Picture

Related projects:

- Clutch: refinement, asynchronous sampling via tapes
- Eris: unary, bound probability of “bad events” via error credits
- Caliper: termination-preserving refinement
- Tachis: expected cost (e.g., time or entropy) via cost credits

Ongoing work:

- Concurrency
- Continuous distributions
- Differential privacy

The Bigger Picture

Related projects:

- Clutch: refinement, asynchronous sampling via tapes
- Eris: unary, bound probability of “bad events” via error credits
- Caliper: termination-preserving refinement
- Tachis: expected cost (e.g., time or entropy) via cost credits

Ongoing work:

- Concurrency
- Continuous distributions
- Differential privacy

Future work:

- Distributed systems
- Tail bounds for complexity
- Realistic languages & cross-language reasoning
- Unifying framework, more security applications,

Summary

- **Approxis**: higher-order approximate refinement separation logic
- **Typed logical relation** for **unknown code** & **limiting arguments**
- Fully **mechanized** in Rocq and Iris
- More in the paper:
 - applications, *e.g.*, rejection samplers for **data structures** (B+ trees from DBs/FSs)
 - semantic model (based on new **coupling laws**)

Thank you!

E-mail `pgh@cs.au.dk`

Website `github.com/logsem/clutch`

Encore

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

Program refinement: $e_1 \preceq e_2 \triangleq \Pr[\mathcal{C}[e_1] \Downarrow v] \leq \Pr[\mathcal{C}[e_2] \Downarrow v]$

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

Contextual refinement: $\mathcal{C}[e_1] \precsim \mathcal{C}[e_2] \triangleq \Pr[\mathcal{C}[e_1] \Downarrow \text{true}] \leq \Pr[\mathcal{C}[e_2] \Downarrow \text{true}]$

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

Refinement Logic: $\vdash \text{rwp } \mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] \{ \llbracket \text{bool} \rrbracket \}$

$\Downarrow$

Contextual refinement: $\mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] \triangleq \Pr[\mathcal{C}[e_1] \Downarrow \text{true}] \leq \Pr[\mathcal{C}[e_2] \Downarrow \text{true}]$

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

Logical relation: $\models \mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] : \text{bool}$

$\Downarrow$

Refinement Logic: $\vdash \text{rwp } \mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] \{ \llbracket \text{bool} \rrbracket \}$

$\Downarrow$

Contextual refinement: $\mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] \triangleq \Pr[\mathcal{C}[e_1] \Downarrow \text{true}] \leq \Pr[\mathcal{C}[e_2] \Downarrow \text{true}]$

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

$\Downarrow$ stable under typed ctx ($\mathcal{C} : \tau \Rightarrow \text{bool}$)

Logical relation: $\models \mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] : \text{bool}$

$\Downarrow$

Refinement Logic: $\vdash \text{rwp } \mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] \{ \llbracket \text{bool} \rrbracket \}$

$\Downarrow$

Contextual refinement: $\mathcal{C}[e_1] \lesssim \mathcal{C}[e_2] \triangleq \Pr[\mathcal{C}[e_1] \Downarrow \text{true}] \leq \Pr[\mathcal{C}[e_2] \Downarrow \text{true}]$

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

$$\models e_1 \lesssim e_2 : \tau$$

$\Downarrow$ stable under typed ctx ($C : \tau \Rightarrow \text{bool}$)

Logical relation: $\models C[e_1] \lesssim C[e_2] : \text{bool}$

$\Downarrow$

Refinement Logic: $\vdash \text{rwp } C[e_1] \lesssim C[e_2] \{ \llbracket \text{bool} \rrbracket \}$

$\Downarrow$

Contextual refinement: $C[e_1] \lesssim C[e_2] \triangleq \Pr[C[e_1] \Downarrow \text{true}] \leq \Pr[C[e_2] \Downarrow \text{true}]$

Methodology for Contextual Refinement

Consider the case $\varepsilon = 0$:

$\forall \varepsilon' > 0. \not\vdash (\varepsilon') \models e_1 \lesssim e_2 : \tau$ Can approximate via “error induction”!

$\Downarrow$ stable under typed ctx ($C : \tau \Rightarrow \text{bool}$)

Logical relation: $\models C[e_1] \lesssim C[e_2] : \text{bool}$

$\Downarrow$

Refinement Logic: $\vdash \text{rwp } C[e_1] \lesssim C[e_2] \{ \llbracket \text{bool} \rrbracket \}$

$\Downarrow$

Contextual refinement: $C[e_1] \lesssim C[e_2] \triangleq \Pr[C[e_1] \Downarrow \text{true}] \leq \Pr[C[e_2] \Downarrow \text{true}]$

IND\$-CPA Security in Approxis

Logical Relation: $\mathcal{I}(\epsilon) \models \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) : \text{bool}$

$\Downarrow$

Refinement Logic: $\mathcal{I}(\epsilon) \vdash \text{rwp } \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) \{b_1 \ b_2 . b_1 = b_2\}$

$\Downarrow$

Approximation: $\Pr[\text{red } \mathcal{A}(\text{encrypt}_{|Q}^k) \Downarrow \text{blue true}] \leq \Pr[\text{red } \mathcal{A}(\text{random}_{|Q}) \Downarrow \text{blue true}] + \epsilon$

IND\$-CPA proof I: Logical Relation

Let $\tau_{\mathcal{A}} \triangleq ((\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool})$ and assume $\vdash \mathcal{A} : \tau_{\mathcal{A}}$.

- To show: $\mathcal{Z}(\varepsilon) \models \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) : \text{bool}$

IND\$-CPA proof I: Logical Relation

Let $\tau_{\mathcal{A}} \triangleq ((\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool})$ and assume $\vdash \mathcal{A} : \tau_{\mathcal{A}}$.

- To show: $\mathcal{Z}(\varepsilon) \models \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) : \text{bool}$
- By Fundamental Lemma of the logical relation, we get $\vdash \mathcal{A} \lesssim \mathcal{A} : \tau_{\mathcal{A}}$

IND\$-CPA proof I: Logical Relation

Let $\tau_{\mathcal{A}} \triangleq ((\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool})$ and assume $\vdash \mathcal{A} : \tau_{\mathcal{A}}$.

- To show: $\mathcal{Z}(\epsilon) \models \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) : \text{bool}$
- By Fundamental Lemma of the logical relation, we get $\models \mathcal{A} \lesssim \mathcal{A} : \tau_{\mathcal{A}}$
- By the rule for application, we can “drop” $\mathcal{A}$ and focus on the oracles:

$$\mathcal{Z}(\epsilon) \models \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} : \text{msg} \rightarrow \text{cipher option}$$

IND\$-CPA proof I: Logical Relation

Let $\tau_{\mathcal{A}} \triangleq ((\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool})$ and assume $\vdash \mathcal{A} : \tau_{\mathcal{A}}$.

- To show: $\not\models (\epsilon) \models \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) : \text{bool}$
- By Fundamental Lemma of the logical relation, we get $\models \mathcal{A} \lesssim \mathcal{A} : \tau_{\mathcal{A}}$

- By the rule for application, we can “drop” $\mathcal{A}$ and focus on the oracles :

$$\not\models (\epsilon) \models \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} : \text{msg} \rightarrow \text{cipher option}$$

- The oracles satisfy an *invariant*. Initially, $i = 0$ and $\epsilon = \sum_{k=0}^{Q-1} \frac{k}{2N} = \frac{(Q-1)Q}{2N}$.

$$\text{Inv} \triangleq \exists i. \text{counter} \mapsto i * \text{counter}' \mapsto_s i * \not\models \left(\frac{(Q-1)Q}{2N} - \frac{(i-1)i}{2N} \right)$$

IND\$-CPA proof I: Logical Relation

Let $\tau_{\mathcal{A}} \triangleq ((\text{msg} \rightarrow \text{cipher option}) \rightarrow \text{bool})$ and assume $\vdash \mathcal{A} : \tau_{\mathcal{A}}$.

- To show: $\not\vdash (\varepsilon) \models \mathcal{A}(\text{encrypt}_{|Q}^k) \lesssim \mathcal{A}(\text{random}_{|Q}) : \text{bool}$
- By Fundamental Lemma of the logical relation, we get $\vdash \mathcal{A} \lesssim \mathcal{A} : \tau_{\mathcal{A}}$
- By the rule for application, we can “drop” $\mathcal{A}$ and focus on the oracles:

$$\not\vdash (\varepsilon) \models \text{encrypt}_{|Q}^k \lesssim \text{random}_{|Q} : \text{msg} \rightarrow \text{cipher option}$$

- The oracles satisfy an *invariant*. Initially, $i = 0$ and $\varepsilon = \sum_{k=0}^{Q-1} \frac{k}{2N} = \frac{(Q-1)Q}{2N}$.

$$\text{Inv} \triangleq \exists i. \text{counter} \mapsto i * \text{counter}' \mapsto_s i * \not\vdash \left(\frac{(Q-1)Q}{2N} - \frac{(i-1)i}{2N} \right)$$

- By the rule for functions, show refinement for an argument $(m : \text{msg})$

$$\boxed{\text{Inv}}^\gamma \models \text{encrypt}_{|Q}^k m \lesssim \text{random}_{|Q} m : \text{cipher option}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- No more unknown code, a program at ground type

$$\boxed{Inv}^\gamma \models \text{encrypt}_{|Q}^k m \lesssim \text{random}_{|Q} m : \text{cipher option}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- No more unknown code, a program at ground type

$$\boxed{Inv}^\gamma \models \text{encrypt}_{|Q}^k m \lesssim \text{random}_{|Q} m : \text{cipher option}$$

- Switch to rwp, open invariant, symbolically execute:

$$Inv_i \vdash \text{rwp} \left(\begin{array}{l} \text{if ! counter} < Q \text{ then} \\ \quad \text{incr counter;} \\ \quad \text{Some (encrypt } k \text{ m)} \\ \text{else None} \end{array} \right) \lesssim \left(\begin{array}{l} \text{if ! counter}' < Q \text{ then} \\ \quad \text{incr counter}'; \\ \quad \text{Some (random m)} \\ \text{else None} \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- Switch to rwp, open invariant, symbolically execute:

$$Inv_i \vdash \text{rwp} \left(\begin{array}{l} \text{if } \text{!counter} < Q \text{ then} \\ \quad \text{incr counter;} \\ \quad \text{Some (encrypt } k \text{ m)} \\ \text{else None} \end{array} \right) \approx \left(\begin{array}{l} \text{if } \text{!counter}' < Q \text{ then} \\ \quad \text{incr counter}'; \\ \quad \text{Some (random m)} \\ \text{else None} \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

- Open invariant, case on the if-statement, compute:

$$\dots * \mathcal{Z}(\epsilon_i) \vdash \text{rwp}_{T \setminus \gamma} \left(\begin{array}{l} \text{let } r = \text{rand } N \text{ in} \\ \text{let } \text{nonce} = \text{prf } k \text{ } r \text{ in} \\ \text{let } c = \text{xor } m \text{ nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } r = \text{rand } N \text{ in} \\ \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- Open invariant, case on the **if**-statement, compute:

$$\dots * \mathcal{I}(\varepsilon_i) \vdash \text{rwp}_{T \setminus \gamma} \left(\begin{array}{l} \text{let } r = \text{rand } N \text{ in} \\ \text{let nonce} = \text{prf } k \ r \text{ in} \\ \text{let } c = \text{xor } m \ \text{nonce} \text{ in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } r = \text{rand } N \text{ in} \\ \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

- Spend some error credits on sampling a fresh nonce

$$\frac{\mathcal{I}\left(\frac{\text{length}(l)}{N+1}\right) \quad \forall n \leq N. n \notin l \multimap \text{rwp } n \approx n \{ \Phi \}}{\text{rwp } \text{rand } N \approx \text{rand } N \{ \Phi \}}$$

where l is the list of previously generated nonces and $\text{length}(l) = i$.

IND\$-CPA proof II: refinement weakens-pre with error credits

- Open invariant, case on the **if**-statement, compute:

$$\dots * \mathcal{I}(\varepsilon_i) \vdash \text{rwp}_{T \setminus \gamma} \left(\begin{array}{l} \text{let } r = \text{rand } N \text{ in} \\ \text{let nonce} = \text{prf } k \ r \text{ in} \\ \text{let } c = \text{xor } m \ \text{nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } r = \text{rand } N \text{ in} \\ \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

- Spend some error credits on sampling a fresh nonce, close invariant:

$$\text{Inv}_{i+1} * r \text{ fresh} \vdash \text{rwp}_{T \setminus \gamma} \left(\begin{array}{l} \text{let nonce} = \text{prf } k \ r \text{ in} \\ \text{let } c = \text{xor } m \ \text{nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- Spend some error credits on sampling a fresh nonce, close invariant:

$$Inv_{i+1} * r \text{ fresh} \vdash \text{rwp}_{T \setminus \gamma} \left(\begin{array}{l} \text{let nonce} = \text{prf } k \text{ } r \text{ in} \\ \text{let } c = \text{xor } m \text{ nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

- PRF assumption: $\text{prf } k \text{ } r$ behaves randomly if r is fresh!

$$r \text{ fresh} \vdash \text{rwp} \left(\begin{array}{l} \text{let nonce} = \text{rand } N \text{ in} \\ \text{let } c = \text{xor } m \text{ nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- PRF assumption: $\text{prf } k \ r$ behaves randomly if r is fresh!

$$r \text{ fresh} \vdash \text{rwp} \left(\begin{array}{l} \text{let nonce} = \text{rand } N \text{ in} \\ \text{let } c = \text{xor } m \text{ nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

- xor acts as one-time pad because nonce is random:

$$\text{Inv}_{i+1} \vdash \text{rwp}_{\top \setminus \gamma} \left(\begin{array}{l} \text{let nonce} = \text{rand } N \text{ in} \\ \text{let } c = \text{xor } m \text{ nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{l} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

IND\$-CPA proof II: refinement weakens-pre with error credits

- xor acts as one-time pad because nonce is random:

$$\text{Inv}_{i+1} \vdash \text{rwp}_{\top \setminus \gamma} \left(\begin{array}{c} \text{let nonce} = \text{rand } N \text{ in} \\ \text{let } c = \text{xor } m \text{ nonce in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{c} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

- done:

$$\vdash \text{rwp} \left(\begin{array}{c} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \approx \left(\begin{array}{c} \text{let } c = \text{rand } N \text{ in} \\ (r, c) \end{array} \right) \{ \llbracket \text{cipher option} \rrbracket \}$$

Overview: Reasoning about Known and Unknown Code

“(Observable) behaviours of e_1 are *included* in those of e_2 .”

Program refinement: $e_1 \lesssim e_2$

Overview: Reasoning about Known and Unknown Code

“(Observable) behaviours of e_1 are *included* in those of e_2 .”

Program refinement: $e_1 \lesssim e_2 \triangleq \forall (\sigma : \text{State})(v : \text{Val}).$
 $\Pr[(e_1, \sigma) \Downarrow v] \leq \Pr[(e_2, \sigma) \Downarrow v]$

Overview: Reasoning about Known and Unknown Code

“(Observable) behaviours of e_1 are *included* in those of e_2 , up to probability ϵ .”

Program approximation: $e_1 \lesssim^{\epsilon} e_2 \triangleq \forall (\sigma : State)(v : Val).$
 $\Pr[(e_1, \sigma) \Downarrow v] \leq \Pr[(e_2, \sigma) \Downarrow v] + \epsilon$

Overview: Reasoning about Known and Unknown Code

Refinement Logic: $\text{rl}(\epsilon) \vdash \text{rwp } e_1 \lesssim e_2 \{ \phi \}$

$\Downarrow$ Adequacy

Program approximation: $e_1 \lesssim^\epsilon e_2 \triangleq \forall (\sigma : \text{State})(v : \text{Val}).$
 $\Pr[(e_1, \sigma) \Downarrow v] \leq \Pr[(e_2, \sigma) \Downarrow v] + \epsilon$

Overview: Reasoning about Known and Unknown Code

Typed Logical Relation: $\Downarrow(\epsilon) \models e_1 \lesssim e_2 : \tau$

$\Downarrow$ Congruence

Refinement Logic: $\Downarrow(\epsilon) \vdash \text{rwp } e_1 \lesssim e_2 \{ \phi \}$

$\Downarrow$ Adequacy

Program approximation: $e_1 \lesssim^\epsilon e_2 \triangleq \forall(\sigma : \text{State})(v : \text{Val}).$
 $\Pr[(e_1, \sigma) \Downarrow v] \leq \Pr[(e_2, \sigma) \Downarrow v] + \epsilon$