# $\begin{array}{c} \text{Matching } \mu\text{-Logic:} \\ \text{Foundation of A Unifying Programming} \\ \text{Language Framework} \end{array}$

Xiaohong Chen
PhD Final Exam
University of Illinois Urbana-Champaign
Department of Computer Science

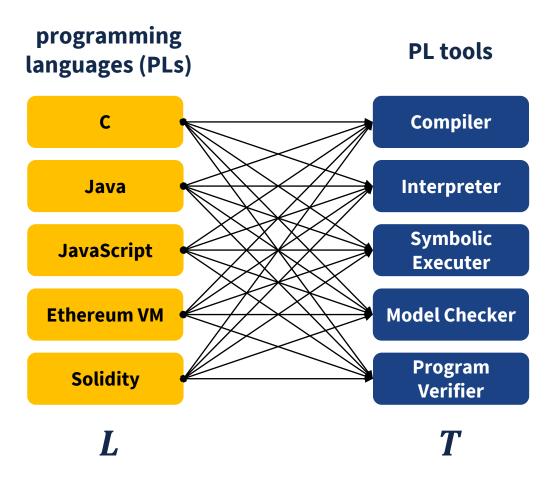
#### **Overview**

- Introduction to a Unifying Programming Language Framework
  - Motivating Example: The K Semantic Framework
  - Research Challenge: Proving the Correctness of K
- Main Contribution: Matching  $\mu$ -Logic
  - Basic Definitions
  - Expressive Power
  - Proof System and Proof Checker
  - Automatic Theorem Prover
- Using Matching  $\mu$ -Logic to Prove the Correctness of K
- Concluding Remarks

#### **Overview**

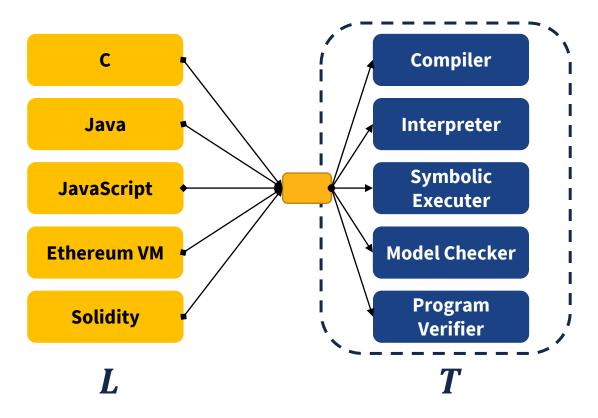
- Introduction to a Unifying Programming Language Framework
  - Motivating Example: The K Semantic Framework
  - Research Challenge: Proving the Correctness of K
- Main Contribution: Matching  $\mu$ -Logic
  - Basic Definitions
  - Expressive Power
  - Proof System and Proof Checker
  - Automatic Theorem Prover
- Using Matching  $\mu$ -Logic to Prove the Correctness of K
- Concluding Remarks

# Programming Language Design & Implementation: State-of-the-Art



 $L \times T$  systems to develop and maintain

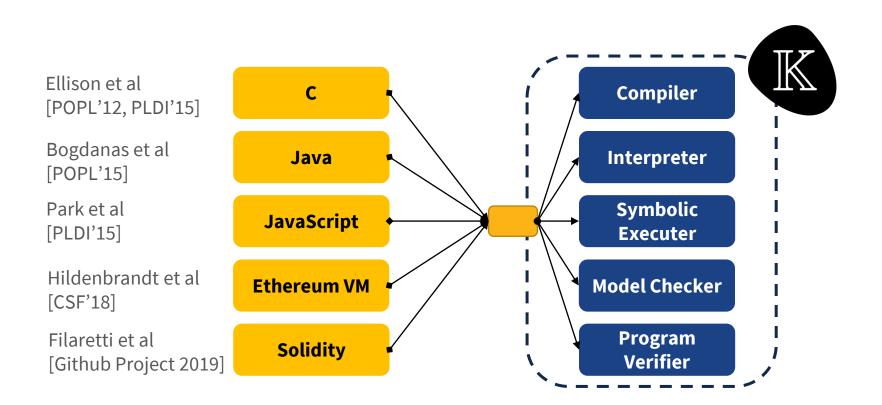
#### **Vision: A Unifying Programming Language Framework**



L + T systems to develop and maintain

#### **K Semantic Framework**

https://kframework.org/



#### K has wide applications



**RV-Match** 



**RV-Monitor** 







# Research Challenge: Proving the Correctness of K

- K has a large code base
  - >500k LOC in 4 programming languages
  - complex data structures, algorithms, and optimizations
- K is constantly evolving
  - latest release: 3 days ago



 It's not practical to fully verify K using traditional methods.

# Main Idea: Let's translate K to logic.

| K  | Logical Foundation of K  |  |  |  |  |
|--|--|--|--|--|--|
| A PL definition Ethereum VM  | A logical theory Γ <sup>EVM</sup>  |  |  |  |  |
| <ul> <li>Any PL task</li> <li>program execution Interpreter</li> <li>formal verification Program Verifier</li> </ul> | A logical theorem proved by a proof system $ \Gamma^{\text{EVM}} \vdash t_{\text{init}} \Rightarrow_{\text{exec}} t_{\text{final}} $ $ \Gamma^{\text{EVM}} \vdash \varphi_{\text{pre}} \Rightarrow_{\text{verify}} \varphi_{\text{post}} $ |  |  |  |  |
| Correctness of the task  | Generating the proof and checking it using a <i>proof checker</i>  |  |  |  |  |
|  |  |  |  |  |  |

correctness of any task done by any tool of any PL  $\Longrightarrow$  correctness of  $\underline{1}$  task done by  $\underline{1}$  program

# Which Logic?

#### We tried many logics/calculi/foundations

First-order logic; Second/higher-order logic; Least fixpoint logic; Modal logics; Temporal logics (LTL, CTL, CTL\*, ...),  $\lambda$ -calculus; Type systems (parametric, dependent, inductive, ...);  $\mu$ -calculus; Hoare logics; Separation logics; Dynamic logics; Rewriting logic; Reachability logic; Equational logic; Small-/big-step SOS; Evaluation contexts; Abstract machines (CC, CK, CEK, SECD, ...); Chemical abstract machine; Axiomatic; Continuations; Denotational; Initial Algebras; ...

#### ... but each of the above had limitations

- Some only handle certain aspects of K (e.g., only execution)
- Some are "design patterns" (e.g., Hoare logics)
- Modularity and heavy notation

#### • Matching $\mu$ -logic: keep advantages and avoid limitations

- PLs defined as theories; PL tools specified by theorems
- High expressive power (K and beyond)
- A 15-rule proof system and a 200-LOC proof checker: small trust base

#### **Overview**

- Introduction to a Unifying Programming Language Framework
  - Motivating Example: The K Semantic Framework
  - Research Challenge: Proving the Correctness of K
- Main Contribution: Matching  $\mu$ -Logic
  - Basic Definitions
  - Expressive Power
  - Proof System and Proof Checker
  - Automatic Theorem Prover
- Using Matching  $\mu$ -Logic to Prove the Correctness of K
- Concluding Remarks

#### Matching $\mu$ -Logic 101

Matching  $\mu$ -logic formulas, called *patterns*:

$$\varphi ::= x \mid \sigma(\varphi_1, ..., \varphi_n) \mid \varphi_1 \land \varphi_2 \mid \neg \varphi \mid \exists x. \varphi \mid X \mid \mu X. \varphi$$

$$\text{structures} \qquad \text{logical constraints} \qquad \text{first-order quantification} \qquad \text{fix points} \qquad \text{(in this talk)}$$

- *X* a *set variable*, ranging over sets
- $\mu X. \varphi$  the *least fixpoint* of  $\varphi$ , explained later
- $\nu X. \varphi \equiv \neg \mu X. \neg \varphi [\neg X/X]$  the greatest fixpoint of  $\varphi$
- $\mu X. \varphi$  and  $\nu X. \varphi$  require that X occurs positively in  $\varphi$

# Matching $\mu$ -Logic 101

A matching  $\mu$ -logic *model* has:

- a carrier set M
- a function  $\sigma_M: M \times \cdots \times M \to \mathcal{P}(M)$  for each symbol  $\sigma$

Given a model M and a variable valuation  $\rho$ :

$$oldsymbol{arphi}$$
 pattern matching  $|oldsymbol{arphi}|_{M,
ho}\subseteq M$ 

- $|x|_{M,\rho} = {\rho(x)}$
- $|\sigma(\varphi_1, ..., \varphi_n)|_{M,\rho} = \bigcup \{\sigma_M(a_1, ..., a_n) \mid a_i \in |\varphi_i|_{M,\rho} \}$
- $|\varphi_1 \wedge \varphi_2|_{M,\rho} = |\varphi_1|_{M,\rho} \cap |\varphi_2|_{M,\rho}$
- $|\neg \varphi|_{M,\rho} = M \setminus |\varphi|_{M,\rho}$
- $|\exists x. \varphi|_{M,\rho} = \bigcup \{|\varphi|_{M,\rho[a/x]} \mid a \in M\}$
- $|X|_{M,\rho} = \rho(X)$
- $|\mu X. \varphi|_{M,\rho} = \mathbf{lfp} \left( A \mapsto |\varphi|_{M,\rho[A/X]} \right)$

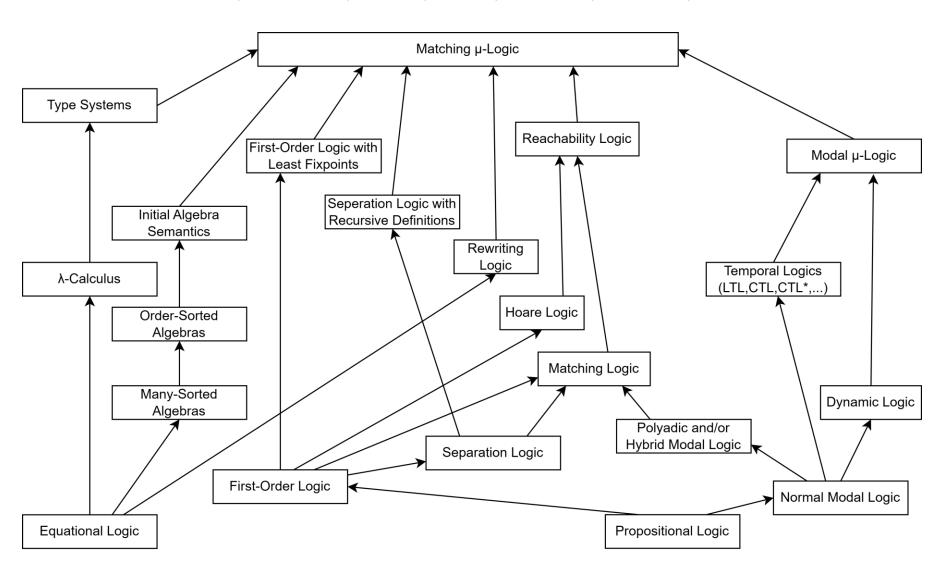
#### **Examples of Fixpoint Patterns**

- inductive datatypes [JLAMP'21]
  - type list = Nil | Cons of int \* list
  - $\mu L$ . Nil  $\vee \exists x$ . Cons(x, L)
  - $L =_{lfp} {Nil} \cup {Cons(x, l) \mid x \in int, l \in L}$
- program execution [LICS'19, CAV'21]
  - an execution trace from  $t_{\rm init}$  to  $t_{\rm final}$
  - $t_{\text{init}} \Rightarrow_{\text{exec}} t_{\text{final}} \equiv t_{\text{init}} \rightarrow \text{eventually } t_{\text{final}}$   $\mu S. t_{\text{final}} \lor (\text{next } S)$
- formal verification [LICS'19, OOPSLA'23]
  - if  $\varphi_{pre}$  holds when P starts, then  $\varphi_{post}$  holds when P terminates
  - $\varphi_{\text{pre}} \Rightarrow_{\text{verify}} \varphi_{\text{post}} \equiv \varphi_{\text{pre}} \rightarrow \text{weak-eventually } \varphi_{\text{post}}$   $\nu S. \varphi_{\text{post}} \lor (\text{next } S)$

Various forms/instances of fixpoints are definable by patterns.

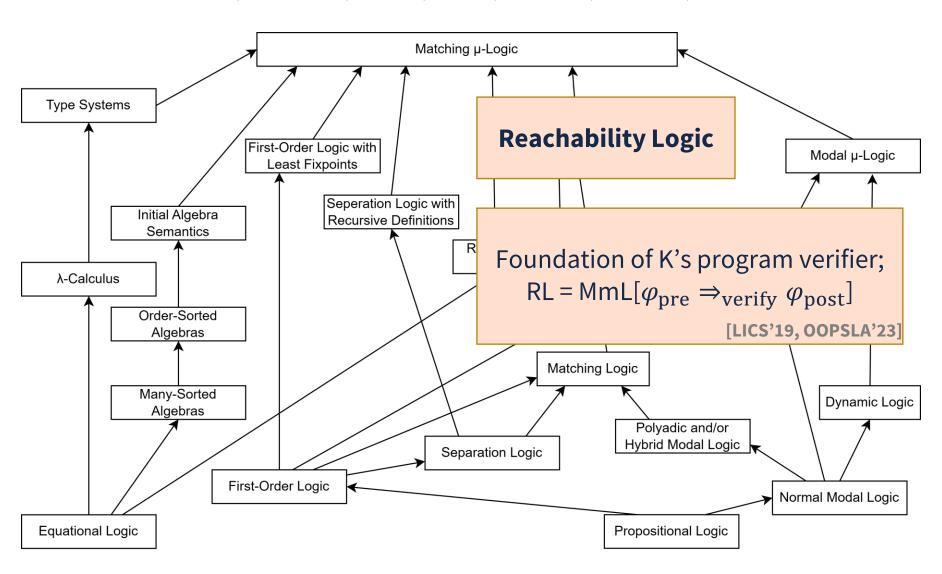
# Matching $\mu$ -Logic (MmL) Expressive Power

[LICS'19, OOPSLA'20, ICFP'20, CAV'21, JLAMP'21, JLAMP'22, OOPSLA'23]



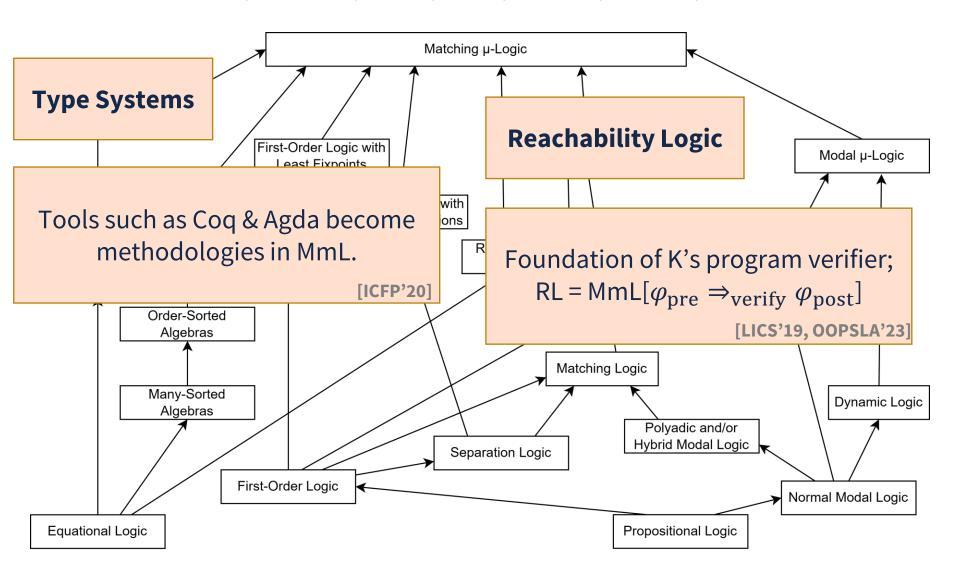
# Matching $\mu$ -Logic (MmL) Expressive Power

[LICS'19, OOPSLA'20, ICFP'20, CAV'21, JLAMP'21, JLAMP'22, OOPSLA'23]



# Matching $\mu$ -Logic (MmL) Expressive Power

[LICS'19, OOPSLA'20, ICFP'20, CAV'21, JLAMP'21, JLAMP'22, OOPSLA'23]



# Matching $\mu$ -Logic Proof System

(only 15 proof rules)

(Existence)

(Singleton)

Technical Rules  $\exists x. x$ 

 $\neg (C_1[x \land \varphi] \land C_2[x \land \neg \varphi])$ 

proof rules for fixpoints

# Deriving Mathematical Induction in Matching $\mu$ -Logic

**Mathematical Induction**: To show a property *P* holds for all naturals, prove:

(basis). The number 0 satisfies P

(**step**). If n satisfies P then n + 1 also satisfies P.

Step 1. Note that  $\mu N$ .  $0 \vee \mathbf{succ}(N)$  captures all natural numbers.

Step 2. Set the proof goal 
$$\vdash (\mu N. \ 0 \lor \mathbf{succ}(N)) \rightarrow \psi_P$$

Step 3. Apply (Knaster Tarski) and get two sub-goals:

Sub-Goal-1 
$$0 \rightarrow \psi_P$$
 (basis Sub-Goal-2  $\operatorname{succ}(\psi_P) \rightarrow \psi_P$  (step)

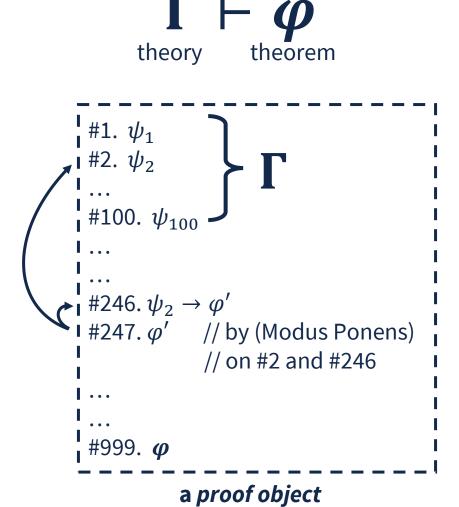
$$(\textbf{Knaster Tarski})$$

$$\frac{\varphi[\psi/X] \to \psi}{\mu X. \ \varphi \to \psi}$$

Various forms/instances of fixpoints reasoning are supported by (Knaster Tarski)

# Matching $\mu$ -Logic Proof System

# (Modus Ponens) $\frac{\varphi_1 \quad \varphi_1 \quad \varphi_1 \rightarrow \varphi_2}{\varphi_2}$



# Matching $\mu$ -Logic Proof Checker

- We use Metamath [Megill & Wheeler] <a href="http://metamath.org">http://metamath.org</a>
  - to encode proof objects &
  - check them automatically
- Very small trust base
  - Matching  $\mu$ -logic: 200 LOC
  - Metamath itself:
    - 350 LOC in Python
    - 400 LOC in Haskell
    - 550 LOC in C#
    - •

```
$c \imp ( ) #Pattern |- $.
      $v ph1 ph2 ph3 $.
      phl-is-pattern $f #Pattern phl $.
      ph2-is-pattern $f #Pattern ph2 $.
      ph3-is-pattern $f #Pattern ph3 $.
      imp-is-pattern
        $a #Pattern ( \imp ph1 ph2 ) $.
10
      axiom-1
11
       $a |- ( \imp ph1 ( \imp ph2 ph1 ) ) $.
12
13
      axiom-2
14
        $a |- ( \imp ( \imp ph1 ( \imp ph2 ph3 ) )
              ( \imp ( \imp ph1 ph2 )
15
16
                     ( \imp ph1 ph3 ) ) $.
17
18
19
        rule-mp.θ $e |- ( \imp ph1 ph2 ) $.
        rule-mp.1 $e |- ph1 $.
21
        rule-mp $a |- ph2 $.
22
```

Matching μ-logic syntax & proof rules; Defined in 200 LOC

```
imp-refl $p |- ( \imp phl phl )
24
25
        phl-is-pattern phl-is-pattern
        phl-is-pattern imp-is-pattern
        imp-is-pattern phl-is-pattern
27
        phl-is-pattern imp-is-pattern
28
        ph1-is-pattern ph1-is-pattern
30
        phl-is-pattern imp-is-pattern
31
        phl-is-pattern imp-is-pattern
32
        imp-is-pattern phl-is-pattern
33
        phl-is-pattern phl-is-pattern
34
        imp-is-pattern imp-is-pattern
35
        phl-is-pattern phl-is-pattern
        imp-is-pattern imp-is-pattern
37
        phl-is-pattern phl-is-pattern
        phl-is-pattern imp-is-pattern
39
        phl-is-pattern axiom-2
        phl-is-pattern phl-is-pattern
41
        phl-is-pattern imp-is-pattern
42
        axiom-1 rule-mp ph1-is-pattern
43
        phl-is-pattern axiom-1 rule-mp
44
```

Proof objects (checked by Metamath)

#### Checking proof objects is fast and trustworthy.

#### Where do Proof Objects Come From?

Q1: Is there always a proof object for a true statement?

• Completeness of matching  $\mu$ -logic (briefly)

Q2: Can we find proof objects automatically?

• Automatic theorem prover for matching  $\mu$ -logic (briefly)

Q3: Can we generate proof objects from K?

Proving the correctness of K

# Where do Proof Objects Come From?

**Q1**: Is there always a proof object for a true statement?

• Completeness of matching  $\mu$ -logic (briefly)

Q2: Can we find proof objects automatically?

• Automatic theorem prover for matching  $\mu$ -logic (briefly)

Q3: Can we generate proof objects from K?

Proving the correctness of K

# Completeness of Matching $\mu$ -Logic

- Matching  $\mu$ -logic is incomplete (because of  $\exists$  and  $\mu$ )
- What if there is no  $\mu$ ?
  - (Local Completeness)

$$\emptyset \vDash \varphi \implies \emptyset \vdash \varphi$$

• (Definedness Completeness)

$$\Gamma \vDash \varphi \implies \Gamma \vdash \varphi$$
, if  $\Gamma$  includes definedness/equality.

- The rest is open problem.
- What if there is no ∃?
  - Open problem.

# Where do Proof Objects Come From?

Q1: Is there always a proof object for a true statement?

• Completeness of matching  $\mu$ -logic (briefly)

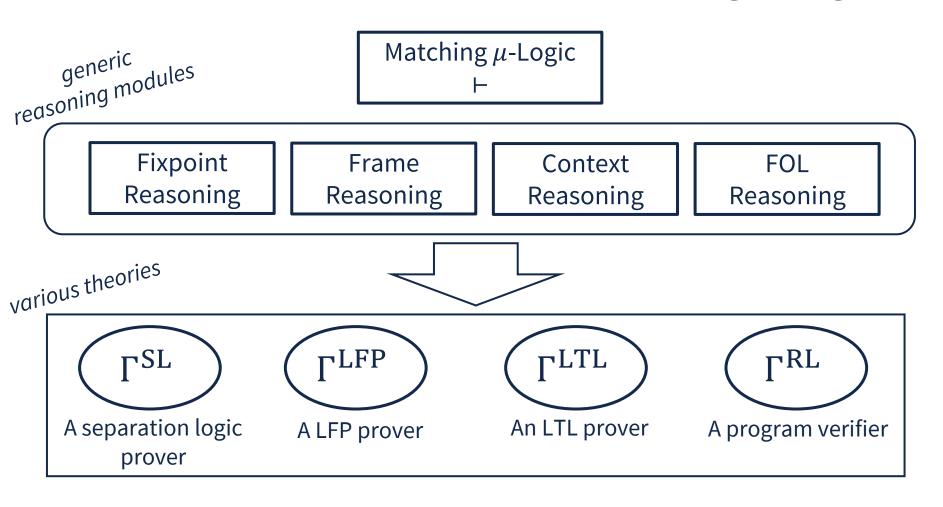
Q2: Can we find proof objects automatically?

• Automatic theorem prover for matching  $\mu$ -logic (briefly)

Q3: Can we generate proof objects from K?

Proving the correctness of K

#### Automatic Theorem Prover for Matching $\mu$ -Logic



- Separation logic: Proved 265/280 benchmark tests in SL-COMP'19
- LTL: Proved all the axioms in the complete LTL proof system
- LTP & RL: Proved the correctness of the SUM program

#### Where do Proof Objects Come From?

Q1: Is there always a proof object for a true statement?

• Completeness of matching  $\mu$ -logic (briefly)

Q2: Can we find proof objects automatically?

• Automatic theorem prover for matching  $\mu$ -logic (briefly)

Q3: Can we generate proof objects from K?

Proving the correctness of K

#### **Overview**

- Introduction to a Unifying Programming Language Framework
  - Motivating Example: The K Semantic Framework
  - Research Challenge: Proving the Correctness of K
- Main Contribution: Matching  $\mu$ -Logic
  - Basic Definitions
  - Expressive Power
  - Proof System and Proof Checker
  - Automatic Theorem Prover
- Using Matching  $\mu$ -Logic to Prove the Correctness of K
  - Translating K to Matching  $\mu$ -Logic
  - Proving the Correctness of the K Program Verifier
- Concluding Remarks

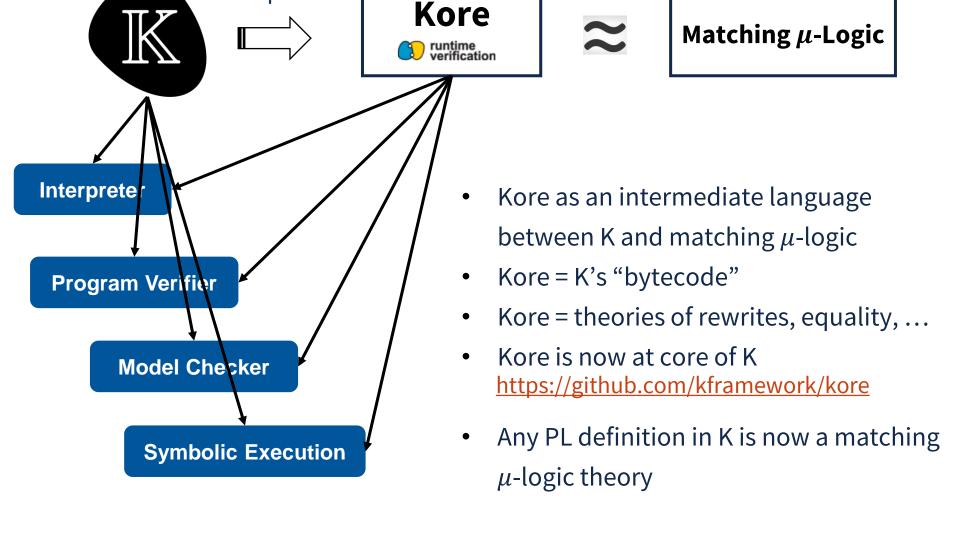
# Translating K to Matching $\mu$ -Logic

| K  | Matching $\mu$ -Logic  |  |  |  |  |
|--|--|--|--|--|--|
| A PL definition Ethereum VM  | A logical theory $\Gamma^{	ext{EVM}}$  |  |  |  |  |
| <ul> <li>Any PL task</li> <li>program execution Interpreter</li> <li>formal verification Program Verifier</li> </ul> | A theorem proved by the 15-rule proof system $ \cdot  \Gamma^{\text{EVM}} \vdash t_{\text{init}} \Rightarrow_{\text{exec}} t_{\text{final}} $ $ \cdot  \Gamma^{\text{EVM}} \vdash \varphi_{\text{pre}} \Rightarrow_{\text{verify}} \varphi_{\text{post}} $ |  |  |  |  |
| Correctness of the task  | Generating the proof and checking it using the 200-LOC <i>proof checker</i>  |  |  |  |  |

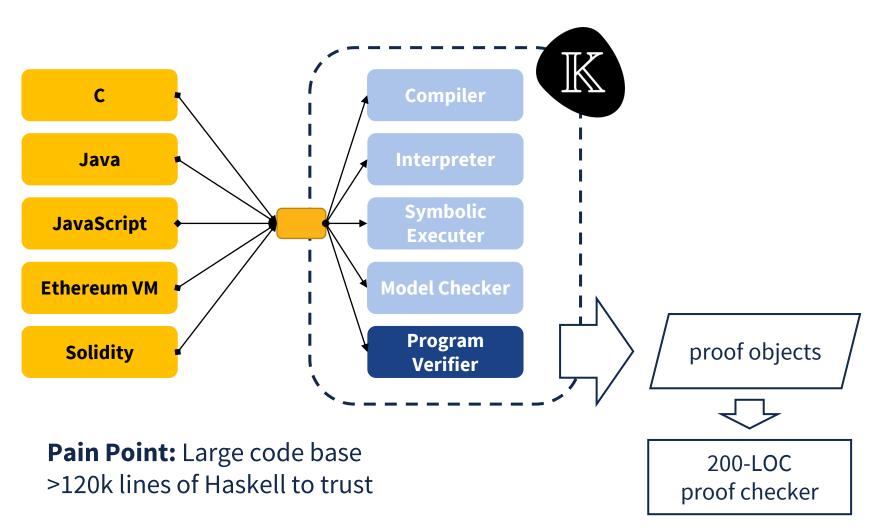
- **Task 1**: Generating the logical theory (e.g.,  $\Gamma^{\text{EVM}}$ )
- **Task 2**: Generating the proof for a given PL task (e.g., verifying a program)

#### Translating PL Semantics to Matching $\mu$ -Logic Theories

kompile

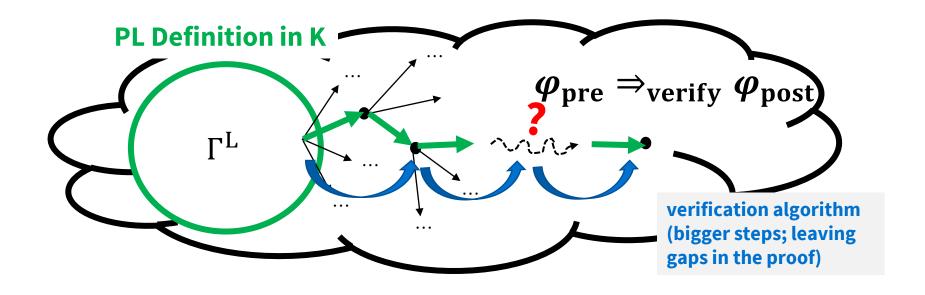


# **Proving the Correctness of the K Program Verifier**



**Solution**: Generating and checking proof objects

#### **Program Verification is Actually Proof Search**



A program verifier is a specialized, optimized, proof searcher.

# **Proof Generation for Program Verification**

The K program verifier checks that P satisfies the pre/post-conditions  $\varphi_{\rm pre}$  and  $\varphi_{\rm post}$  in L

#### proof generation

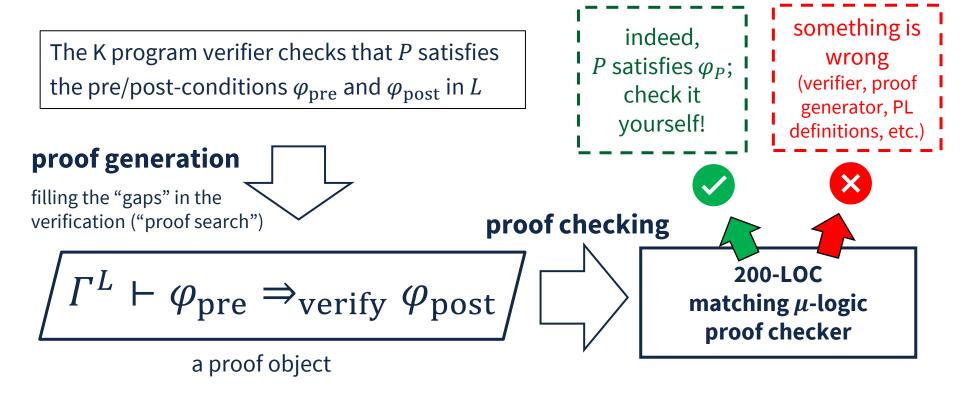
filling the "gaps" in the verification ("proof search")

$$/\Gamma^L \vdash \varphi_{\text{pre}} \Rightarrow_{\text{verify}} \varphi_{\text{post}}$$

a proof object

```
#1. \psi_1
#2. \psi_2
...
#100. \psi_{100}
...
#247. \psi_2 \to \varphi
#247. \varphi // by (Modus Ponens)
// on #2 and #246
...
...
#99999. \varphi_{\text{pre}} \Rightarrow_{\text{verify}} \varphi_{\text{post}}
```

# **Proof Generation for Program Verification**



# **Proof Generation: Complicated ...**

top-level proof goal 
$$\Gamma^L \vdash \varphi_{\mathrm{pre}} \Rightarrow_{\mathrm{verify}} \varphi_{\mathrm{post}}$$

... but none of the above needs to be trusted.

#### **Evaluation**

proof generation Time (seconds)

checking

We tested on 3 PL paradigms:

- imperative
- register-based
- functional

Reduced K trust base (~120k lines of Haskell)

Found issues in K (missing axioms etc.)

Future work

Apply it to more PLs

|               |           |       |                   |                      |            | THIE (Sec |  | Onus  |
|---------------|-----------|-------|-------------------|----------------------|------------|-----------|--|-------|
| Task          | Spec. LOC | Steps | Hint Size         | Proof Size           | K Verifier | Gen.      |  | Check |
| sum.imp       | 40        | 42    | $0.58\mathrm{MB}$ | 37/1.6 MB            | 4.2        | 105       |  | 1.8   |
| sum.reg       | 46        | 108   | $2.24\mathrm{MB}$ | 111/3.6 MB           | 9.1        | 259       |  | 5.4   |
| sum.pcf       | 18        | 22    | $0.29\mathrm{MB}$ | 38/1.5 MB            | 2.9        | 119       |  | 2.4   |
| exp.imp       | 27        | 31    | $0.5\mathrm{MB}$  | 37/1.5 MB            | 3.7        | 108       |  | 2.0   |
| exp.reg       | 27        | 43    | 0.96 MB           | 70/2.3 MB            | 4.7        | 177       |  | 3.1   |
| exp.pcf       | 20        | 29    | $0.5\mathrm{MB}$  | 65/2.3 MB            | 3.8        | 199       |  | 3.1   |
| collatz.imp   | 25        | 55    | $1.14\mathrm{MB}$ | 49/1.7 MB            | 4.8        | 138       |  | 2.6   |
| collatz.reg   | 37        | 100   | $3.66\mathrm{MB}$ | $209/4.7\mathrm{MB}$ | 9.3        | 414       |  | 5.5   |
| collatz.pcf   | 26        | 39    | 1.51 MB           | $110/2.2\mathrm{MB}$ | 5.3        | 247       |  | 5.2   |
| product.imp   | 44        | 42    | $0.62\mathrm{MB}$ | 44/1.8 MB            | 3.9        | 124       |  | 2.4   |
| product.reg   | 24        | 42    | 0.81 MB           | 65/2.3 MB            | 4.3        | 164       |  | 4.0   |
| product.pcf   | 21        | 48    | $0.82\mathrm{MB}$ | 80/2.8 MB            | 5.3        | 234       |  | 4.9   |
| gcd.imp       | 51        | 93    | 1.9 MB            | 74/2.3 MB            | 22.9       | 237       |  | 2.7   |
| gcd.reg       | 27        | 73    | 1.92 MB           | $124/3.3\mathrm{MB}$ | 18.6       | 306       |  | 3.6   |
| gcd.pcf       | 22        | 38    | 1.35 MB           | $150/3.2\mathrm{MB}$ | 12.8       | 367       |  | 5.2   |
| ln/count-by-1 | 44        | 25    | $0.24\mathrm{MB}$ | 28/1.3 MB            | 2.7        | 81        |  | 1.6   |
| ln/count-by-2 | 44        | 25    | $0.26\mathrm{MB}$ | 28/1.3 MB            | 9.0        | 88        |  | 1.4   |
| ln/gauss-sum  | 51        | 39    | $0.53\mathrm{MB}$ | 38/1.6 MB            | 4.6        | 107       |  | 2.0   |
| ln/half       | 62        | 65    | 1.3 MB            | 63/2.2 MB            | 13.1       | 173       |  | 3.0   |
| ln/nested-1   | 92        | 84    | 1.88 MB           | $104/3.4\mathrm{MB}$ | 7.5        | 231       |  | 5.9   |

#### **Overview**

- Introduction to a Unifying Programming Language Framework
  - Motivating Example: The K Semantic Framework
  - Research Challenge: Proving the Correctness of K
- Main Contribution: Matching  $\mu$ -Logic
  - Basic Definitions
  - Expressive Power
  - Proof System and Proof Checker
  - Automatic Theorem Prover
- Using Matching  $\mu$ -Logic to Prove the Correctness of K
  - Translating K to Matching  $\mu$ -Logic
  - Proving the Correctness of the K Program Verifier
- Concluding Remarks

# Matching $\mu$ -Logic: A Unifying Foundation for Programming

