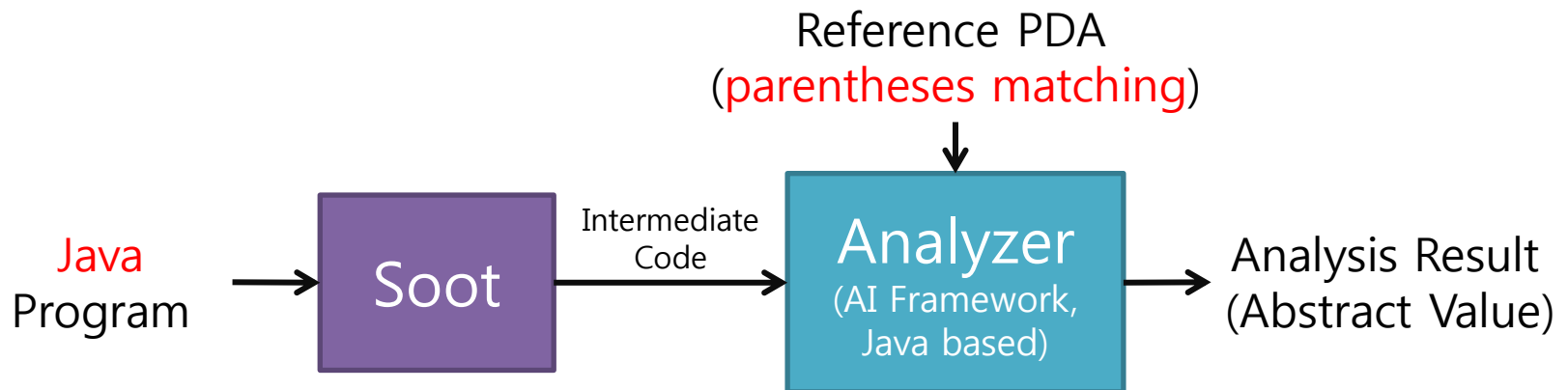# Implementing String Analysis for PHP Programs

Jaejun Choi

Programming Language Laboratory,

Computer Science Dept., KAIST

Previous Work

- ## String analysis using AI framework
  - ### Abstract value from string
    - Se-won Kim and Kwang-moo Choe, "String analysis as an Abstract interpretation"
  - ### To verify it really works

Reference PDA
(parentheses matching)

Java
Program

Soot

Intermediate
Code

Analyzer
(AI Framework,
Java based)

Analysis Result
(Abstract Value)

Current Work

- String analysis using AI framework
  - To apply it to real-world programs
    - PHP programs
    - Support *replace* operation
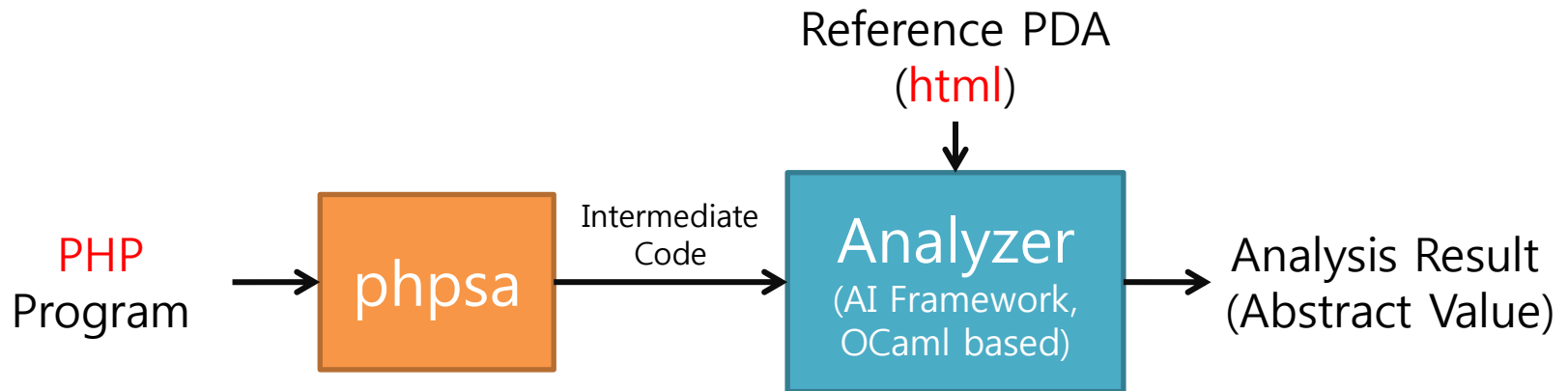
replace operation

- $C_{s'} := replace(C_1, C_2, C_S)$
  - Is possible when
    - $C_1 \equiv C_2$
    - $C_1 \sqsubseteq C_2$

  - Otherwise
    - $C_{s'} := \bot$

  - C_1, C_2, C_S are abstract values

replace operation

- Problem

  – Does this approach really work well for real-world programs?

    - Applicability

    - Accuracy

Work to Do

- Continue the implementation
  - Modify phpsa
  - Implement analyzer

- Experiments
  - Dose it work well?
  - If so, how much?
  - It not, why?