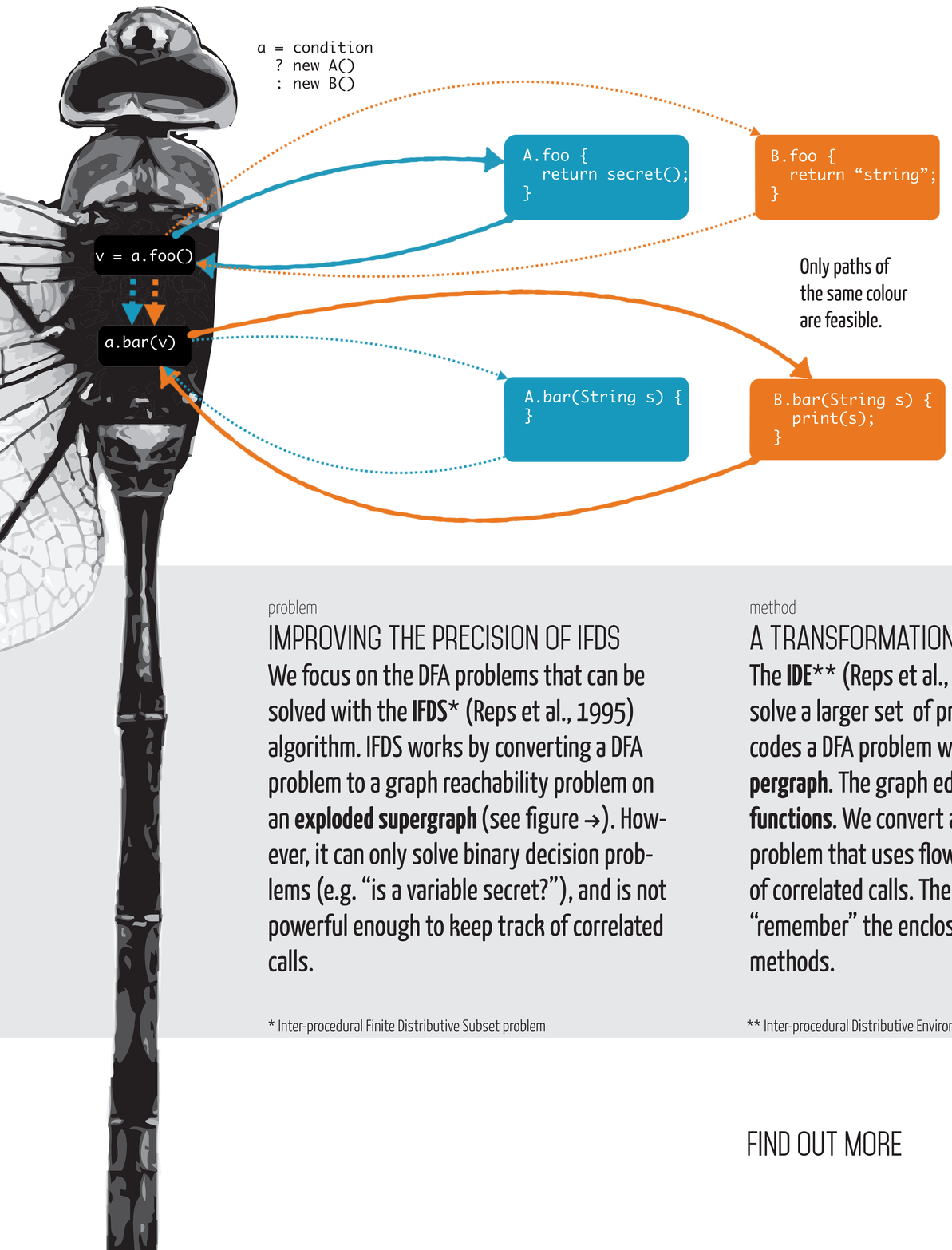
DATA FLOW ANALYSIS IN THE PRESENCE OF CORRELATED CALLS

Marianna Rapoport, Ondřej Lhoták, Frank Tip
University of Waterloo



summary

THE PRECISION OF DATA-FLOW ANALYSES CAN BE IMPROVED IN THE PRESENCE OF CORRELATED CALLS.

intro

IS YOUR DATA REALLY SECRET?

Data-flow analysis (DFA) approximates properties of programs without running them. For instance, in a **taint analysis**, we find out which variables are **secret**, e.g. to discover confidential information leaks. However, **infeasible paths** in a program's control-flow graph can affect the accuracy of an analysis.

goal

ELIMINATE INFEASIBLE PATHS

An infeasible path is one that cannot occur during program execution. In an object-oriented language, two method calls are **correlated** if they dispatch to multiple targets. The goal of this work is to **eliminate** the infeasible paths caused by correlated calls.

result

CORRELATED CALLS ANALYSIS

The correlated calls analysis improves the precision of IFDS results that contain correlated calls. Infeasible paths caused by correlated calls are removed by transforming an IFDS problem into a special type of IDE problem and solving the latter.

problem

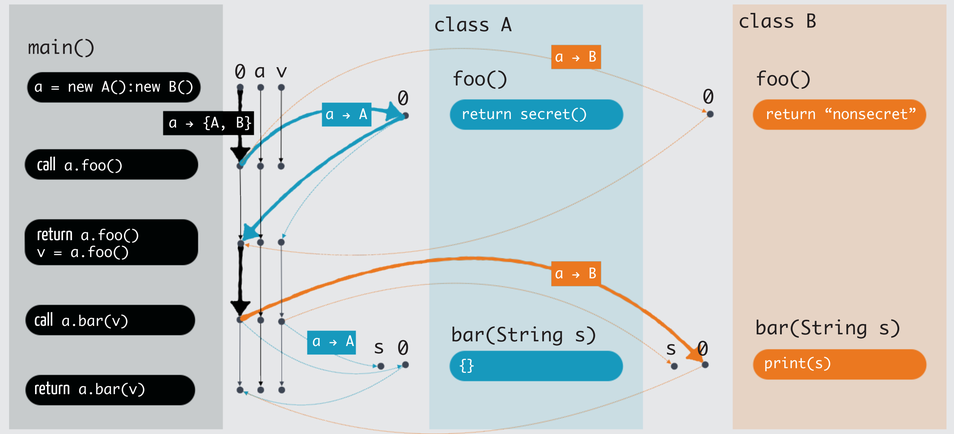
IMPROVING THE PRECISION OF IFDS

We focus on the DFA problems that can be solved with the **IFDS*** (Reps et al., 1995) algorithm. IFDS works by converting a DFA problem to a graph reachability problem on an **exploded supergraph** (see figure →). However, it can only solve binary decision problems (e.g. “is a variable secret?”), and is not powerful enough to keep track of correlated calls.

method

A TRANSFORMATION FROM IFDS TO IDE

The **IDE**** (Reps et al., 1996) algorithm can solve a larger set of problems than IFDS. IDE encodes a DFA problem with a **labeled exploded supergraph**. The graph edges are labeled with **flow functions**. We convert an IFDS problem to an IDE problem that uses flow functions to keep track of correlated calls. The flow functions serve to “remember” the enclosing classes of dispatched methods.



* Inter-procedural Finite Distributive Subset problem

** Inter-procedural Distributive Environment problem

FIND OUT MORE

- How do IFDS and IDE work?
- How are flow functions represented?
- How can we implement the correlated-calls analysis?
- How do we know the analysis is correct?

cs.uwaterloo.ca/~mrapopor