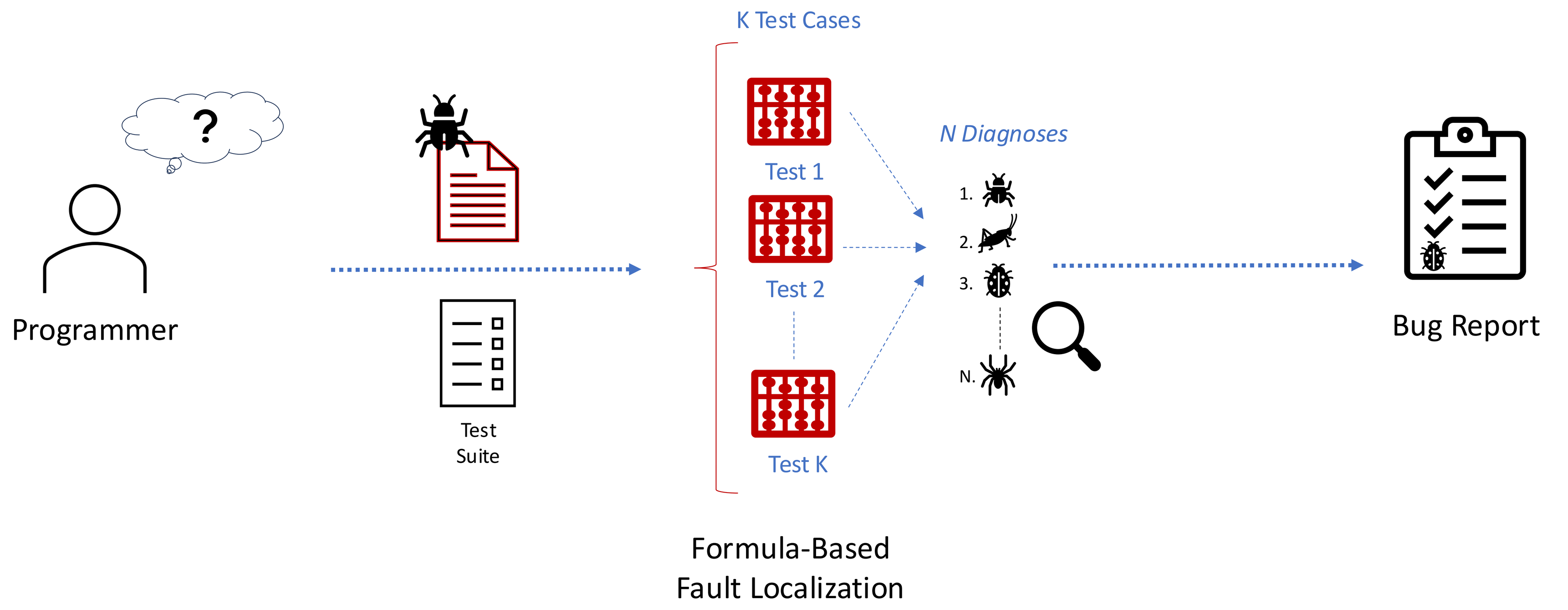
CFaults: Model-Based Diagnosis for Fault Localization in C with Multiple Test Cases

Pedro Orvalho, Mikoláš Janota, Vasco Manquinho

pmorvalho@inesc-id.pt

Motivation: Formula-Based Fault Localization (FBFL)

- Debugging is one of the most time-consuming and expensive tasks in software development.
  - In 2000, the total cost of the work done in preparation for Year 2000 Problem likely surpassed 400 Billion US$ [The Guardian 2019];
  - In 2024, the estimated global cost of Crowdstrike's error that hit Microsoft systems, is 24 Billion US$ [The Sun UK].



State-of-the-art FBFL tools especially for programs with multiple faults:

- **do not ensure a minimal diagnosis** across all failing tests (e.g., BugAssist);
- may produce an overwhelming number of **redundant sets of diagnoses** (e.g., SNIPER).

Contributions

- We formulate the FL problem as a **single optimization problem**;
- We leverage MaxSAT and the theory of *Model-Based Diagnosis (MBD)*, **integrating all failing test cases simultaneously**;
- We present CFaults, **a fault localization tool for ANSI-C programs**, that:
  - **refines localized faults** to pinpoint the bug's location more precisely;
  - is **fast and only produces subset-minimal diagnoses**, unlike other FBFL tools.

Model-Based Diagnosis with Multiple Test Cases

Model-Based Diagnosis Theory:

- A system description $\mathcal{P}$ **is composed of a set of components** $\mathcal{C} = \{c_1, \ldots, c_n\}$.
- Each component in $\mathcal{C}$ can be declared **healthy** or **unhealthy**.
- For each component $c \in \mathcal{C}$, $h(c) = 0$ **if** c **is unhealthy, otherwise,** $h(c) = 1$.
- $\mathcal{P}$ is described by a CNF formula, where $\mathcal{F}_c$ denotes the encoding of component c:

$$\mathcal{P} \triangleq \bigwedge_{c \in \mathcal{C}} (\neg h(c) \vee \mathcal{F}_c) \quad (1)$$

We **integrate all failing test cases** in a single MaxSAT formula:

- We **generate only minimal diagnoses** capable of identifying all faulty components within the system, in our case, a C program;
- Given m observations (failing test cases), $\mathcal{O} = \{o_1, \ldots, o_m\}$, a distinct replica of the system, denoted as $\mathcal{P}_i$, is required for each observation o_i;
- The hard clauses, ϕ_h, in our MaxSAT formulation correspond to:

$$\phi_h = \bigwedge_{o_i \in \mathcal{O}} (\mathcal{P}_i \wedge o_i);$$

- The soft clauses are formulated as:

$$\phi_s = \bigwedge_{c \in \mathcal{C}} h(c).$$

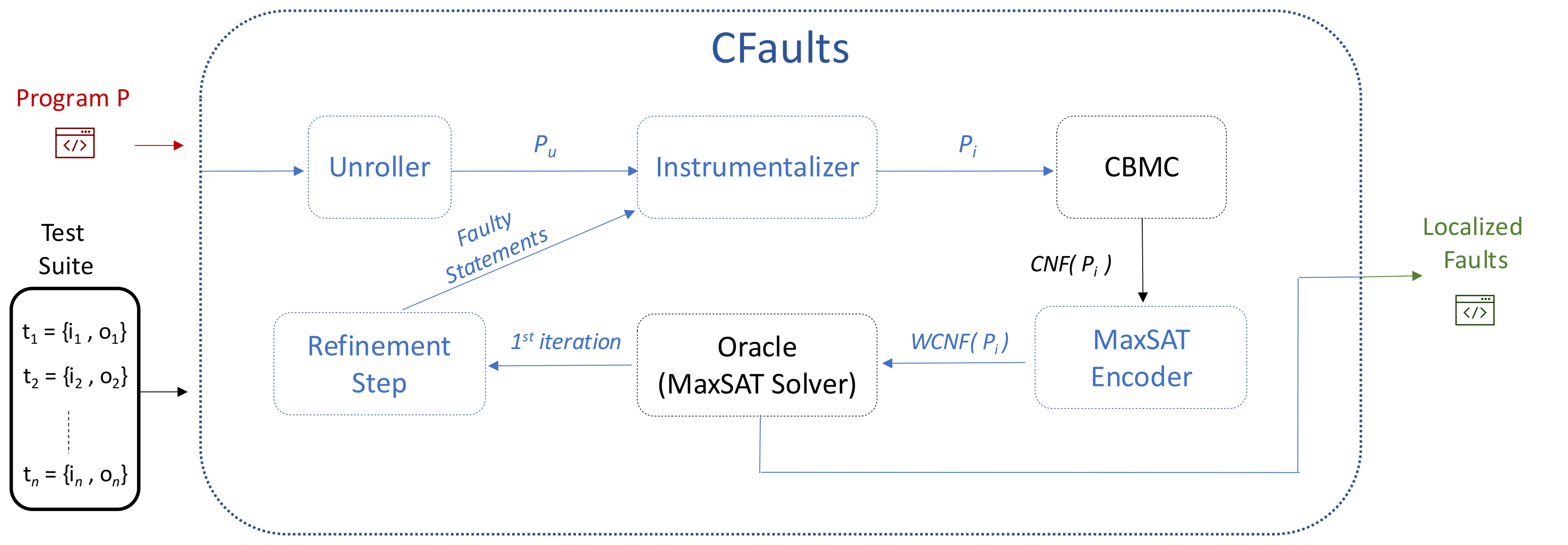
- The complement of the MaxSAT solution, i.e., **the set of unhealthy components** ($h(c) = 0$), **corresponds to a subset-minimal aggregated diagnosis**.
- This **diagnosis is a subset-minimal of components** that, when declared unhealthy (deactivated), make the system consistent with all observations, as follows:

$$\bigwedge_{o_i \in \mathcal{O}} (\mathcal{P}_i \wedge o_i) \wedge \bigwedge_{c \in \mathcal{C} \setminus \Delta} h(c) \wedge \bigwedge_{c \in \Delta} \neg h(c) \nvDash \bot \quad (2)$$

- A diagnosis Δ is minimal iff no subset of Δ, $\Delta' \subsetneq \Delta$, is a diagnosis, and Δ is of minimal cardinality if there is no other diagnosis $\Delta'' \subseteq \mathcal{C}$ with $|\Delta''| < |\Delta|$.

inesc id lisboa
TÉCNICO LISBOA
CZECH INSTITUTE OF INFORMATICS ROBOTICS AND CYBERNETICS CTU IN PRAGUE

CFaults



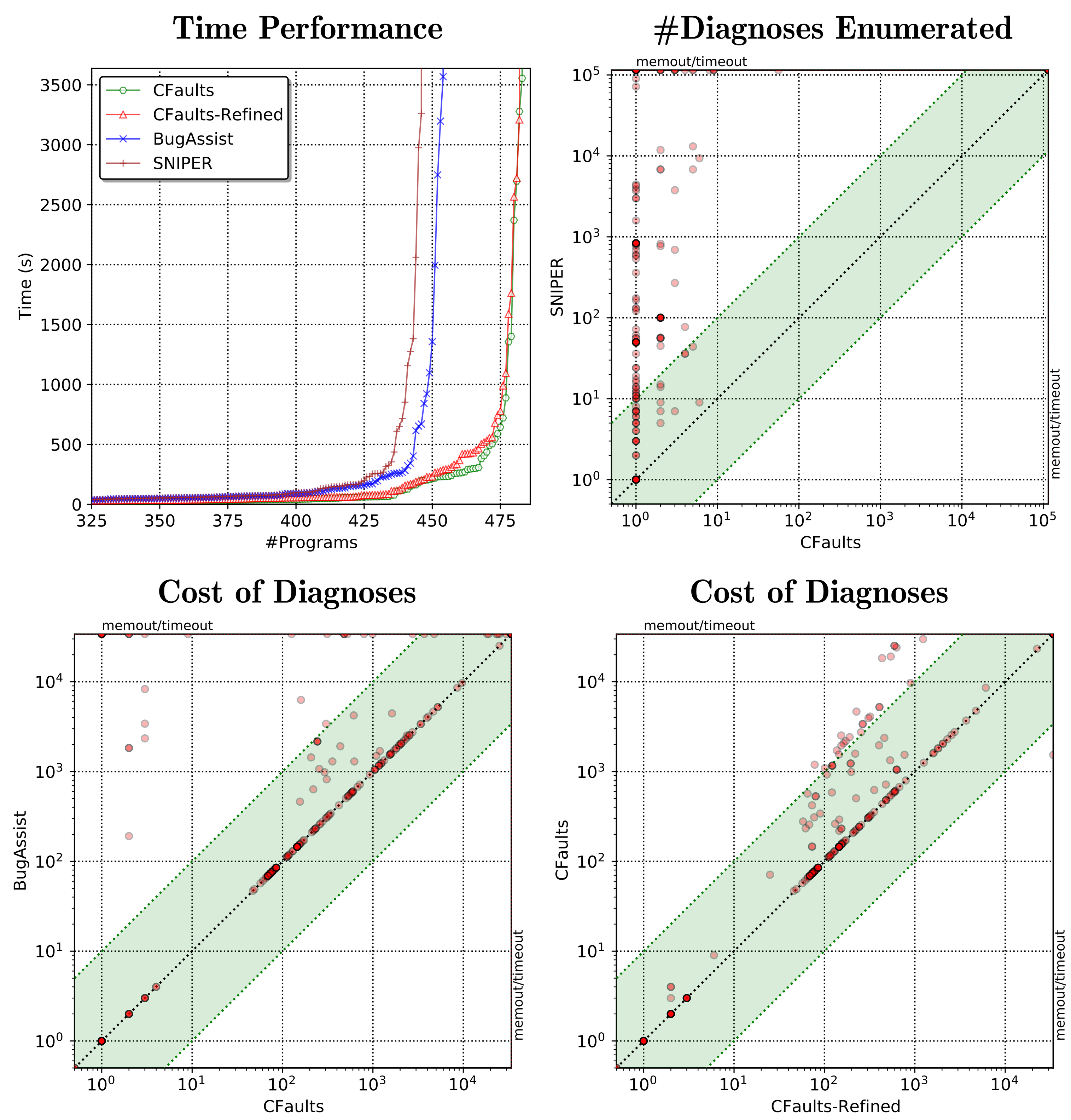
Experimental Evaluation

Benchmark: **TCAS**

	Valid Diagnosis	Memouts	Timeouts
BugAssist	41 (100.0%)	0 (0.0%)	0 (0.0%)
SNIPER	7 (17.07%)	34 (82.93%)	0 (0.0%)
CFaults	41 (100.0%)	0 (0.0%)	0 (0.0%)
CFaults-Refined	41 (100.0%)	0 (0.0%)	0 (0.0%)

Benchmark: **C-Pack-IPAs**

	Valid Diagnosis	Memouts	Timeouts
BugAssist	454 (93.42%)	0 (0.0%)	32 (6.58%)
SNIPER	446 (91.77%)	4 (0.82%)	36 (7.41%)
CFaults	483 (99.38%)	1 (0.21%)	2 (0.41%)
CFaults-Refined	482 (99.18%)	1 (0.21%)	3 (0.62%)



- **SNIPER generates significantly more diagnoses**, most of them redundant;
- **BugAssist yields a non-optimal diagnosis in 10% of TCAS**;
- **BugAssist fails to provide an optimal diagnosis in almost 6%** of C-Pack-IPAs;
- The refinement enables CFaults to identify smaller **diagnoses at a reduced cost in approximately 16%** of C-Pack-IPAs.

LaTeX TikZposter