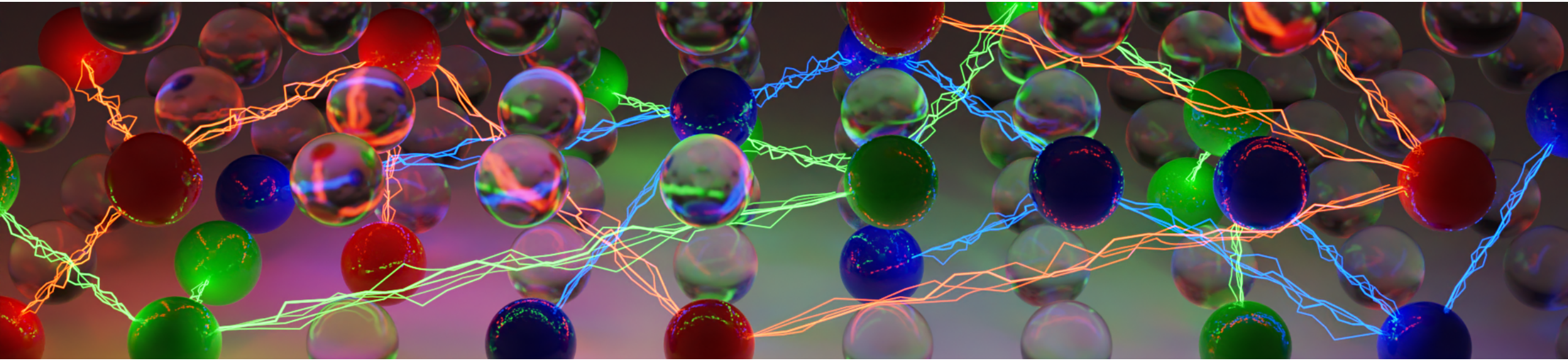
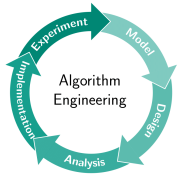


Practical SAT Solving



Lecture 4 – Application Highlights

Ashlin Iser, Dominik Schreiber | May 11, 2026

SAtRes
Scalable Automated Reasoning

Overview

Recap Lecture 3

- Stochastic local search
- Resolution, saturation algorithm
- From Davis Putnam to DPLL

Today: Applications I

- Why consider applications?
- Automated planning: Foundations and SAT-based methods
- From planning to checking: Bounded Model Checking basics
- Combinational Equivalence Checking

Why consider the Application View? (1/2)

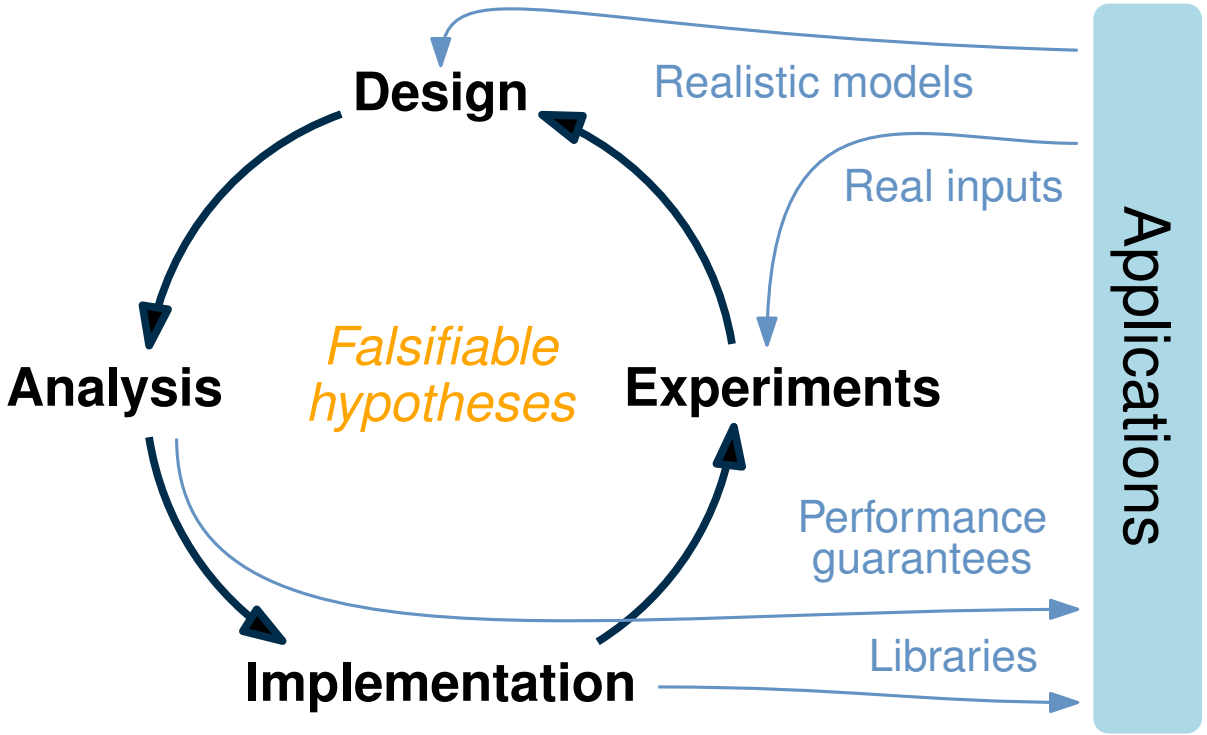
Result, instance family		1st author	Speedups of MallobSat over Kissat-MAB_HyWalk
SAT	Hypertree decomposition	Schidler	3, 5, 5, 5, 5, 7, 8, 9, 12, 13
SAT	Hamilton circle	Heule	4, 4, 7, 11, 17, 20, 21, 22, 24, 31, 33, 36, 42
SAT	Tree decomposition	Ehlers	5, 7, 87
UNSAT	Cellular automatons	Chowdhury	5, 8, 8, 9, 9, 10, 22, 22, 66
			⋮
UNSAT	Relativized pidgeon hole	Elffers	277, 542, 638
UNSAT	Bioinformatics	Bonet	292, 717
UNSAT	Balanced random	Spence	321, 388
SAT	Sum of three cubes	Riveros	384, 509, 1018, 3345
SAT	Circuit multiplication	Shunyang	17, 31, 32, 62, 105, 119, 213, 254, 393, 741, 746, 1401, 2650
UNSAT	Perfect matchings	Reeves	119, 413, 12 439, 108 593, 217 099

3072 cores (128 machines) of SuperMUC-NG · 400 problems from Int. SAT Competition 2021

Only instances with seq. time $\geq$ 60 s · Only families with $\geq$ 2 instances

Why consider the Application View? (2/2)

- Perform sound **algorithm engineering** with realistic applications in the loop
- Understand application-specific solver behavior, needs, shortcomings
- Advance the positive feedback loop between solvers and applications



SAT Competition Benchmarks – “Real” Applications?

5355 problems,
138 families + 72 unknown

Font size $\propto$ # problems

via `wordclouds.com`,
GBD
(`benchmark-database.de`)

Automated Planning: Introduction

Classical Automated Planning in a Nutshell

Given a set of world features, an according initial world state, and a set of conditional rules to modify the world state ("actions"), find a way to successively transform the initial world state until it satisfies some goal condition.

Properties of (classical) automated planning:

- Deterministic, full-knowledge, closed-world [23] ("*everything not explicitly true is assumed false*")
- Extremely generic "state transition system" model
  – essentially shortest path search but on a prohibitively large graph that must be generated on the go
- **PSPACE**-complete [4] – solutions may need to be exponentially long
- Applications of classical planning models mostly limited to highly idealized/simplified settings
  – But: Ideal "fruitfly problem" for considering interactions of applications with SAT solving
- Tons of extensions, enhancements, variants

Automated Planning: Definitions

Planning Problem (semi formal, simplified PDDL-style)

A **planning problem** is a tuple $P := (s_I, A, g)$, where s_I and g are sets of consistent propositional world state features (each true or false) and A is a set of actions. Each action has the shape $a = (pre_a, eff_a)$, where pre_a and eff_a are also sets of consistent propositional world state features. The **objective** is to find a plan $\Pi = \langle a_1, \ldots, a_n \rangle$ such that each pre_{a_k} is satisfied by $s_k :=$ (s_I updated by $eff_{a_1}, \ldots, eff_{a_{k-1}}$) and g is satisfied by the final arising state s_{n+1}.

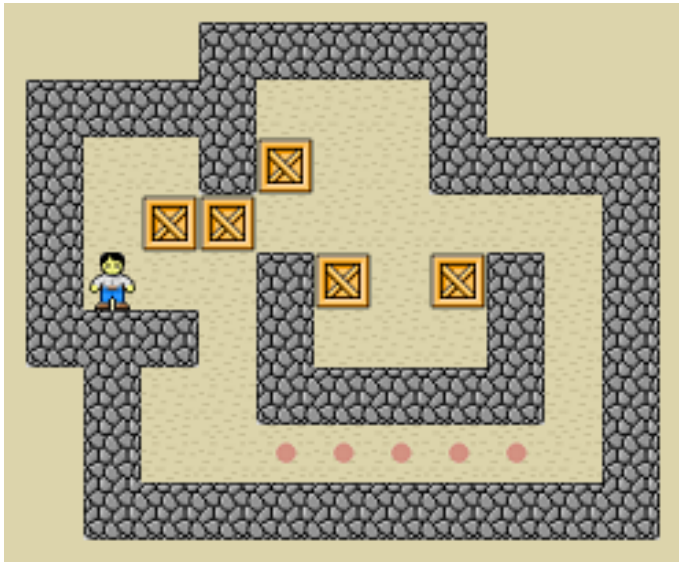
Automated Planning: Definitions

Planning Problem (semi formal, simplified PDDL-style)

A **planning problem** is a tuple $P := (s_I, A, g)$, where s_I and g are sets of consistent propositional world state features (each true or false) and A is a set of actions. Each action has the shape $a = (pre_a, eff_a)$, where pre_a and eff_a are also sets of consistent propositional world state features. The **objective** is to find a plan $\Pi = \langle a_1, \ldots, a_n \rangle$ such that each pre_{a_k} is satisfied by $s_k :=$ (s_I updated by $eff_{a_1}, \ldots, eff_{a_{k-1}}$) and g is satisfied by the final arising state s_{n+1}.

Example: Sokoban [6]

- **World state features:** Positions of walls, boxes, goals, player
- **Actions:** Walk and push. E.g., `push-right(x, y)`:
  - Preconditions: `player-at(x,y)`, `box-at(x+1,y)`, `¬wall-at(x+2,y)`, `¬box-at(x+2,y)`
  - Effects: `player-at(x+1,y)`, `¬box-at(x+1,y)`, `box-at(x+2,y)`
- **Goal:** `box-at(`x^*`,`y^*`)` for each marked location (x^*, y^*)
- **Plan:** Series of move/push actions leading to a goal state



Automated Planning

How do we solve planning problems?

- Most common: Forward state space search [12] – explore state space while heuristically closing in on a goal state
- More exotic: Backward search, plan-space search (not discussed here)
- Old and robust (since 1992 [14]): SAT-based planning

Automated Planning

How do we solve planning problems?

- Most common: Forward state space search [12] – explore state space while heuristically closing in on a goal state
- More exotic: Backward search, plan-space search (not discussed here)
- Old and robust (since 1992 [14]): SAT-based planning

(How) Can we encode planning problems into SAT?

Automated Planning

How do we solve planning problems?

- Most common: Forward state space search [12] – explore state space while heuristically closing in on a goal state
- More exotic: Backward search, plan-space search (not discussed here)
- Old and robust (since 1992 [14]): SAT-based planning

(How) Can we encode planning problems into SAT?

- We cannot translate planning to a single polynomially sized SAT encoding (NP vs. PSPACE?)
  – Unclear how many steps to encode!
- We can translate planning to a compact SAT encoding for a fixed number of K steps
  $\Rightarrow$ Incrementally increase K until plan is found!
- Main aspects of design space: Used SAT encoding, Scheduling strategies for values of K

Automated Planning: A SAT Encoding [13] (1/2)

1. The initial state has to hold at time 0: (Assume s_I explicitly defines *all relevant/reachable propositions*)

$$\forall \ell \in s_I : \quad \ell^{(0)}$$

Automated Planning: A SAT Encoding [13] (1/2)

1. The initial state has to hold at time 0: (Assume s_I explicitly defines *all relevant/reachable propositions*)

$$\forall \ell \in s_I : \quad \ell^{(0)}$$

2. Apply at most one action at each time: (Optionally also add "*At least one* action")

$$\forall t \in \{0, \ldots, K\} \;\; \forall a, a' \in A : \quad \neg a^{(t)} \vee \neg a'^{(t)}$$

Automated Planning: A SAT Encoding [13] (1/2)

1. The initial state has to hold at time 0: (Assume s_I explicitly defines *all relevant/reachable propositions*)

$$\forall \ell \in s_I : \quad \ell^{(0)}$$

2. Apply at most one action at each time: (Optionally also add "*At least one* action")

$$\forall t \in \{0, \ldots, K\} \;\; \forall a, a' \in A : \quad \neg a^{(t)} \vee \neg a'^{(t)}$$

3. Applying an action at time t implies its preconditions at time t:

$$\forall t \in \{0, \ldots, K\} \;\; \forall a \in A \;\; \forall \ell \in pre_a : \quad a^{(t)} \rightarrow \ell^{(t)}$$

Automated Planning: A SAT Encoding [13] (1/2)

1. The initial state has to hold at time 0: (Assume s_I explicitly defines *all relevant/reachable propositions*)

$$\forall \ell \in s_I : \quad \ell^{(0)}$$

2. Apply at most one action at each time: (Optionally also add "*At least one* action")

$$\forall t \in \{0, \ldots, K\} \;\; \forall a, a' \in A : \quad \neg a^{(t)} \vee \neg a'^{(t)}$$

3. Applying an action at time t implies its preconditions at time t:

$$\forall t \in \{0, \ldots, K\} \;\; \forall a \in A \;\; \forall \ell \in pre_a : \quad a^{(t)} \rightarrow \ell^{(t)}$$

4. Applying an action at time t implies its effects at time $t+1$:

$$\forall t \in \{0, \ldots, K\} \;\; \forall a \in A \;\; \forall \ell \in eff_a : \quad a^{(t)} \rightarrow \ell^{(t+1)}$$

Automated Planning: A SAT Encoding [13] (1/2)

1. The initial state has to hold at time 0: (Assume s_I explicitly defines *all relevant/reachable propositions*)

$$\forall \ell \in s_I : \quad \ell^{(0)}$$

2. Apply at most one action at each time: (Optionally also add "*At least one* action")

$$\forall t \in \{0, \ldots, K\} \;\; \forall a, a' \in A : \quad \neg a^{(t)} \vee \neg a'^{(t)}$$

3. Applying an action at time t implies its preconditions at time t:

$$\forall t \in \{0, \ldots, K\} \;\; \forall a \in A \;\; \forall \ell \in pre_a : \quad a^{(t)} \rightarrow \ell^{(t)}$$

4. Applying an action at time t implies its effects at time $t+1$:

$$\forall t \in \{0, \ldots, K\} \;\; \forall a \in A \;\; \forall \ell \in eff_a : \quad a^{(t)} \rightarrow \ell^{(t+1)}$$

5. Each goal has to hold at time K:

$$\forall \ell \in g : \quad \ell^{(K)}$$

Automated Planning: A SAT Encoding (2/2)

We are done, right?

Automated Planning: A SAT Encoding (2/2)

We are done, right?
$\Rightarrow$ Sokoban example: SAT Solver finds solution at $K = 1$, where boxes magically materialize at all goal spots.
What did we forget?

Automated Planning: A SAT Encoding (2/2)

We are done, right?
$\Rightarrow$ Sokoban example: SAT Solver finds solution at $K = 1$, where boxes magically materialize at all goal spots.
What did we forget?

6. **Frame axioms:** A state feature only changes if an action supports this change:

$$\forall t \in \{0, \ldots, K-1\} \;\; \forall \ell \in s_I \cup \bar{s}_I : \quad \bar{\ell}^{(t)} \wedge \ell^{(t+1)} \rightarrow \bigvee_{a \in A \mid \ell \in \mathit{eff}_a} a^{(t)}$$

Automated Planning: A SAT Encoding (2/2)

We are done, right?
$\Rightarrow$ Sokoban example: SAT Solver finds solution at $K = 1$, where boxes magically materialize at all goal spots.
What did we forget?

6. **Frame axioms:** A state feature only changes if an action supports this change:

$$\forall t \in \{0, \ldots, K-1\} \;\; \forall \ell \in s_I \cup \bar{s}_I : \quad \bar{\ell}^{(t)} \wedge \ell^{(t+1)} \rightarrow \bigvee_{a \in A \mid \ell \in \mathit{eff}_a} a^{(t)}$$

Proposition

Using clause rules 1–6, encoding a planning problem P to SAT for some fixed $K \geq 0$ yields a satisfiable CNF formula if and only if there exists a plan Π for P with at most K steps.
Such a Π can be decoded from a model by reading the satisfying assignment to the variables $a^{(t)}$.

SAT Planning

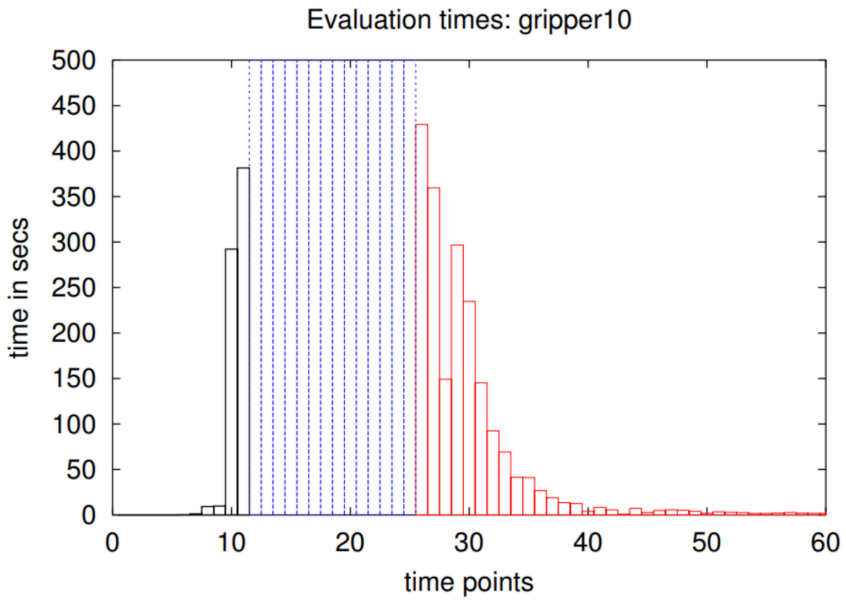
Improvements to SAT-based planning

- More efficient encodings, in particular for 2. (“at most one action”)
- Relaxed semantics: Execute several actions at once (e.g., [1])
- Exploit incremental SAT solving! [11]
  - Encode 1. initially
  - Encode 2.,3.,4.,6. at every new increment
  - Enforce 5. (“goal reached”) as temporary assumptions at each SAT call

SAT Planning

Improvements to SAT-based planning

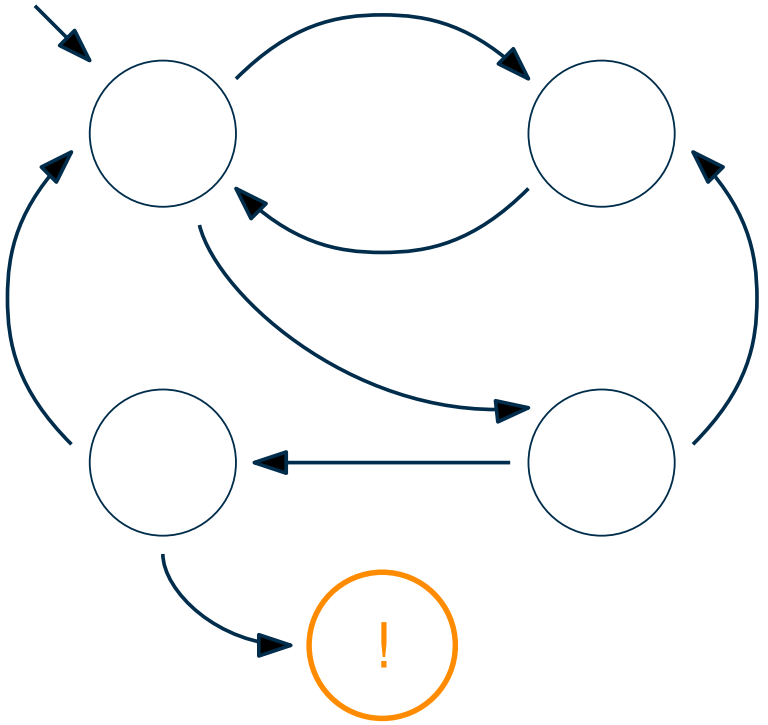
- More efficient encodings, in particular for 2. (“at most one action”)
- Relaxed semantics: Execute several actions at once (e.g., [1])
- Exploit incremental SAT solving! [11]
  - Encode 1. initially
  - Encode 2.,3.,4.,6. at every new increment
  - Enforce 5. (“goal reached”) as temporary assumptions at each SAT call
- Find more effective sequences of values for K to test (“makespan scheduling”) [24]
  - Try to bypass unsurmountable wall of exceedingly difficult K (esp. UNSAT)
  - Increase K linearly, exponentially, . . .
  - Test several K in parallel



From Planning to Verification?

Let's use our SAT-based planning to check the correctness of a system!

- World state features $\equiv$ State features of our system
- Actions $\equiv$ Valid transitions between states
- Goals $\equiv$

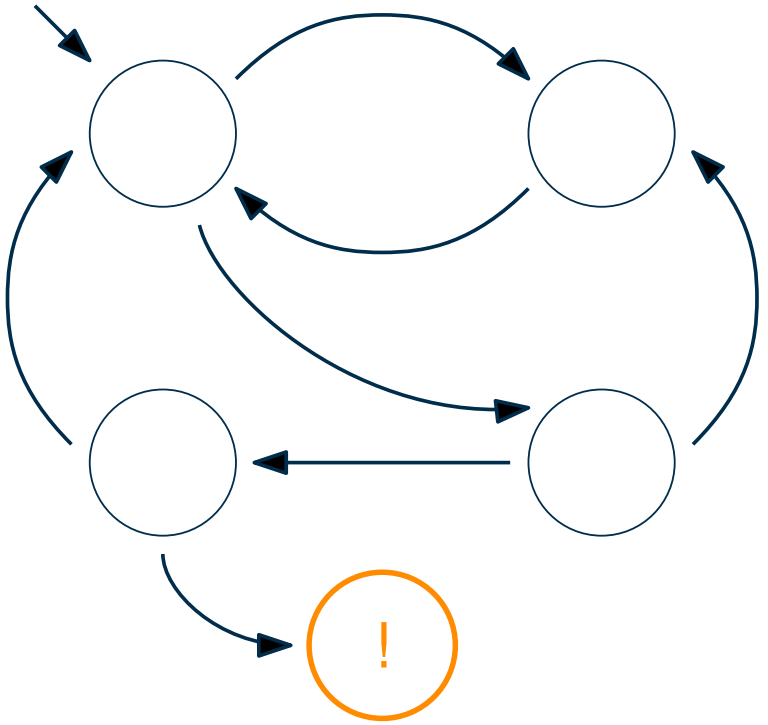


A snack machine or an electronic component or a C program or ...

From Planning to Verification?

Let's use our SAT-based planning to check the correctness of a system!

- World state features ≡ State features of our system
- Actions ≡ Valid transitions between states
- Goals ≡ incorrect state, violating some constraint
- Plan ≡

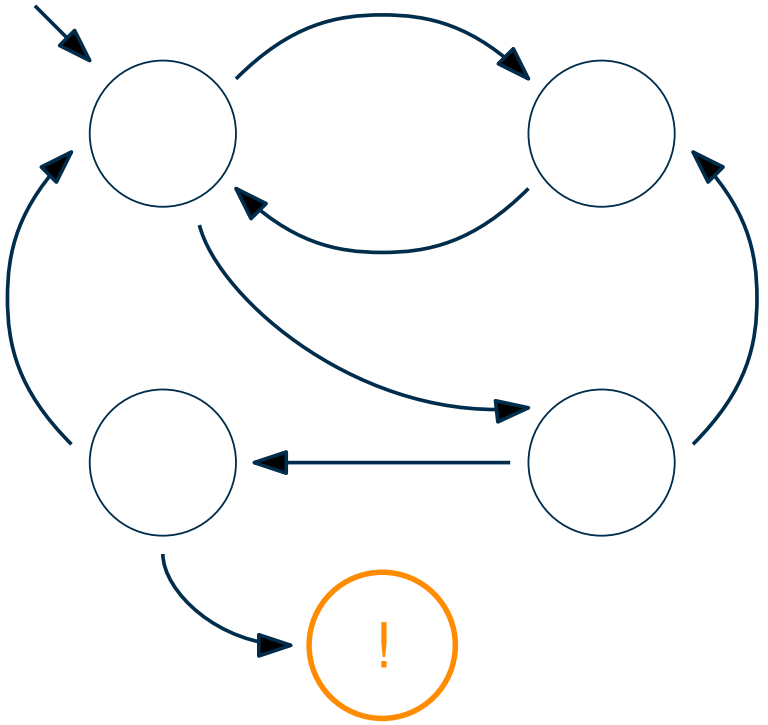


A snack machine or an electronic component or a C program or . . .

From Planning to Verification?

Let's use our SAT-based planning to check the correctness of a system!

- World state features ≡ State features of our system
- Actions ≡ Valid transitions between states
- Goals ≡ incorrect state, violating some constraint
- Plan ≡ reachable incorrectness
- Unsatisfiability ≡ system is always correct?

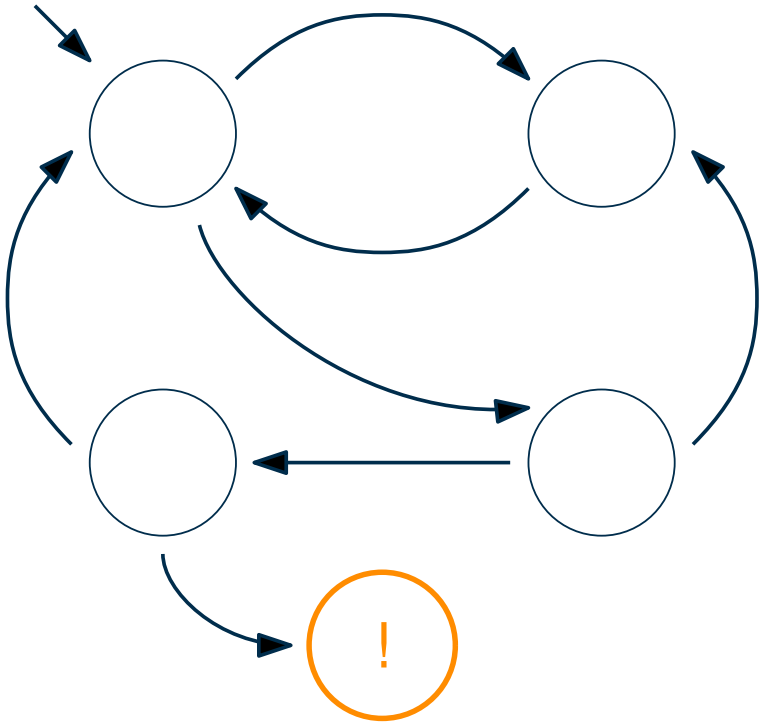


A snack machine or an electronic component or a C program or . . .

From Planning to Verification?

Let's use our SAT-based planning to check the correctness of a system!

- World state features ≡ State features of our system
- Actions ≡ Valid transitions between states
- Goals ≡ incorrect state, violating some constraint
- Plan ≡ reachable incorrectness
- Unsatisfiability at k steps ≡ system is correct within k steps



A snack machine or an electronic component or a C program or . . .

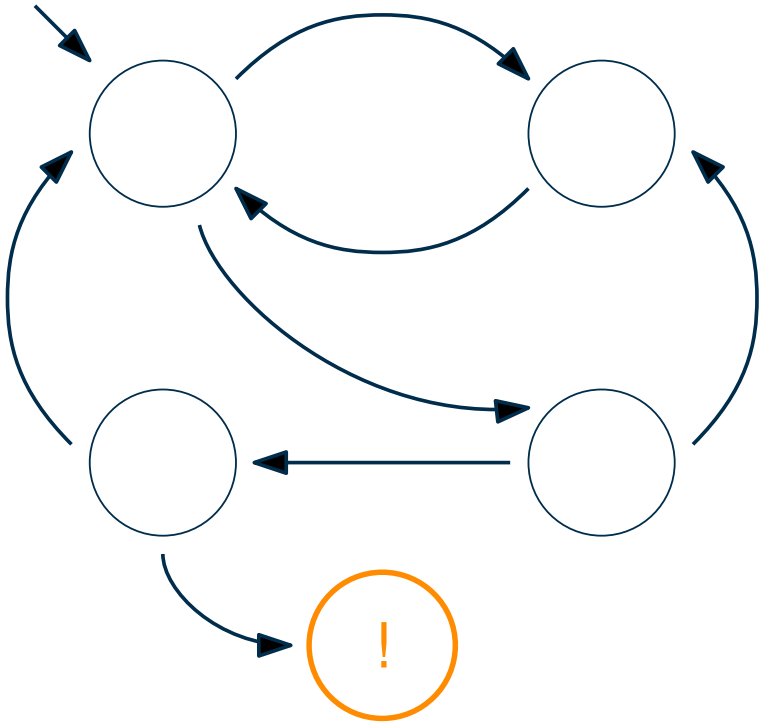
From Planning to Verification?

Let's use our SAT-based planning to check the correctness of a system!

- World state features $\equiv$ State features of our system
- Actions $\equiv$ Valid transitions between states
- Goals $\equiv$ incorrect state, violating some constraint
- Plan $\equiv$ reachable incorrectness
- Unsatisfiability at k steps $\equiv$ system is correct within k steps

Bounded Model Checking

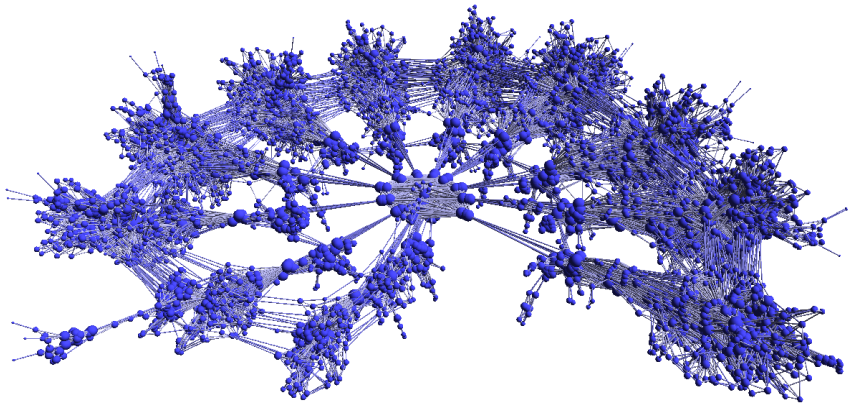
- Encode and check transition system for $k = 1, 2, \ldots$
- Satisfying assignment $\equiv$ counter example!
- Crucial tool for hardware and software verification [28]



A snack machine or an electronic component or a C program or ...

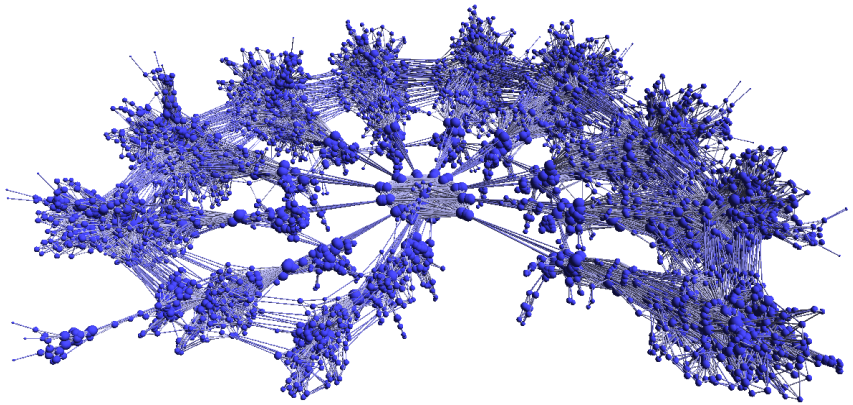
Bounded Model Checking

- Invented by Clarke & Biere in ∼2000 [5], mostly replacing BDD-based model checking
- State transition system based on temporal logic (LTL, CTL, . . .); solving via SAT or SMT
- Applications: Computer-aided design (CAD), software verification, invariant checking, bug detection, . . . [28]



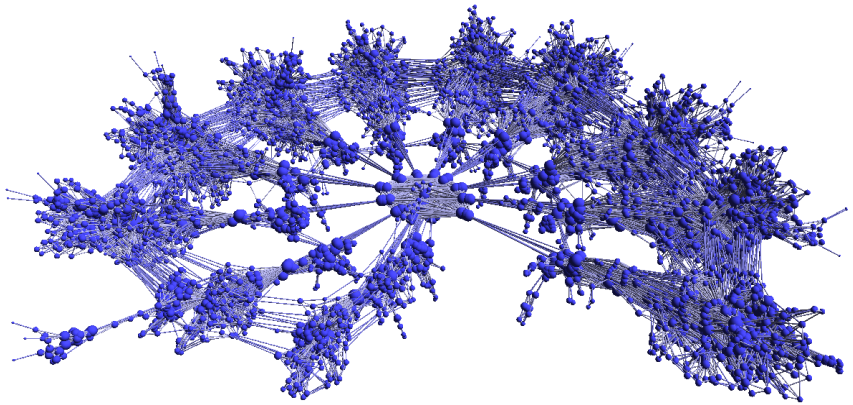
Bounded Model Checking

- Invented by Clarke & Biere in ∼2000 [5], mostly replacing BDD-based model checking
- State transition system based on temporal logic (LTL, CTL, . . .); solving via SAT or SMT
- Applications: Computer-aided design (CAD), software verification, invariant checking, bug detection, . . . [28]
- One of the most essential real-world applications of SAT
  - Pushed industrial interest in SAT solvers in 2000s
  - Actively influenced solver design and algorithms
  - Some of the largest, structurally most distinct benchmarks



Bounded Model Checking

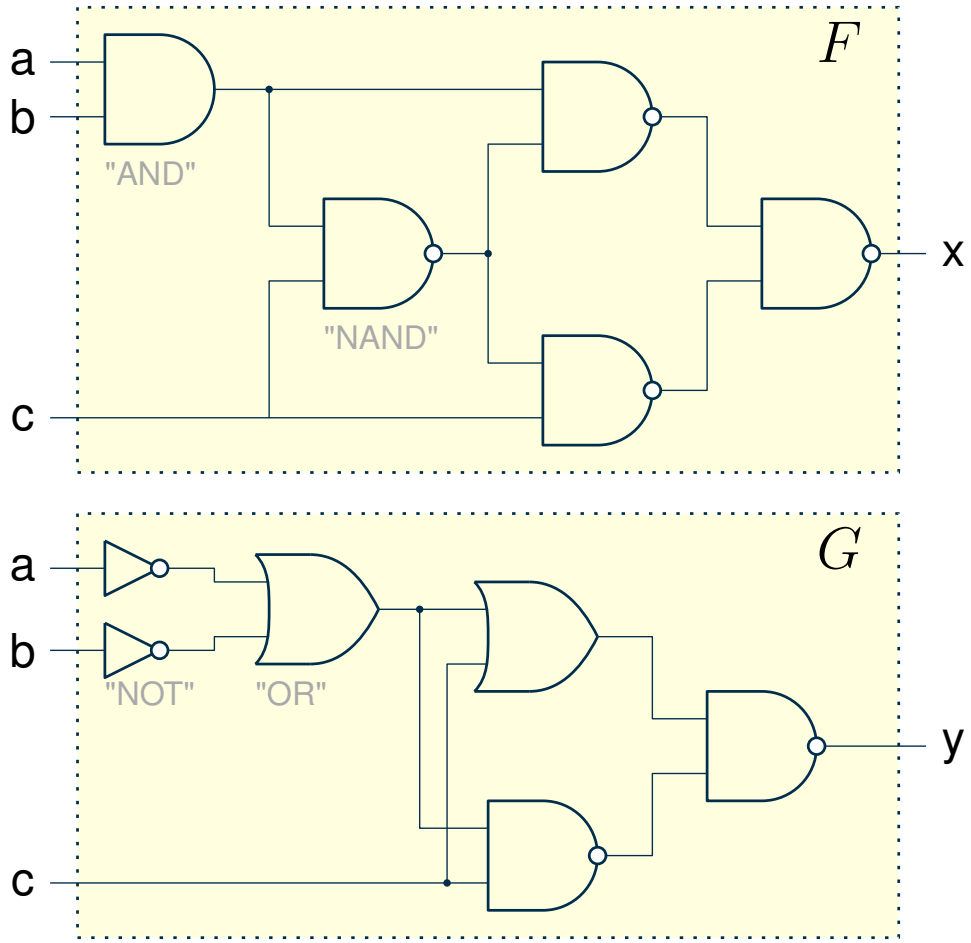
- Invented by Clarke & Biere in ∼2000 [5], mostly replacing BDD-based model checking
- State transition system based on temporal logic (LTL, CTL, . . .); solving via SAT or SMT
- Applications: Computer-aided design (CAD), software verification, invariant checking, bug detection, . . . [28]
- One of the most essential real-world applications of SAT
  - Pushed industrial interest in SAT solvers in 2000s
  - Actively influenced solver design and algorithms
  - Some of the largest, structurally most distinct benchmarks
- Examples for BMC @ KIT:
  - Low-Level Bounded Model Checker (LLBMC) [8] (C program verification)
  - Verification of Java contracts [2] (see right)
  - Cryptography [15]



```
/*@ requires 0 <= x1;
  @ ensures \result == x1 * x2;
  @ assignable \nothing;
  @*/
public int mult(int x1, int x2) {
  int res = 0;
  /*@ loop_invariant 0 <= i && i <= x1 && res == i * x2;
    @ decreases x1 - i;
    @ assignable \nothing;
    @*/
  for (int i = 0; i < x1; ++i) res += x2;
  return res;
}
```

Combinational Equivalence Checking

- Given: two combinational circuits
  stateless "input→output" circuit, no feedback
- Question: are the circuits logically equivalent?
- Right example: Is $F(a, b, c) \equiv G(a, b, c)$?
- How to solve with SAT?

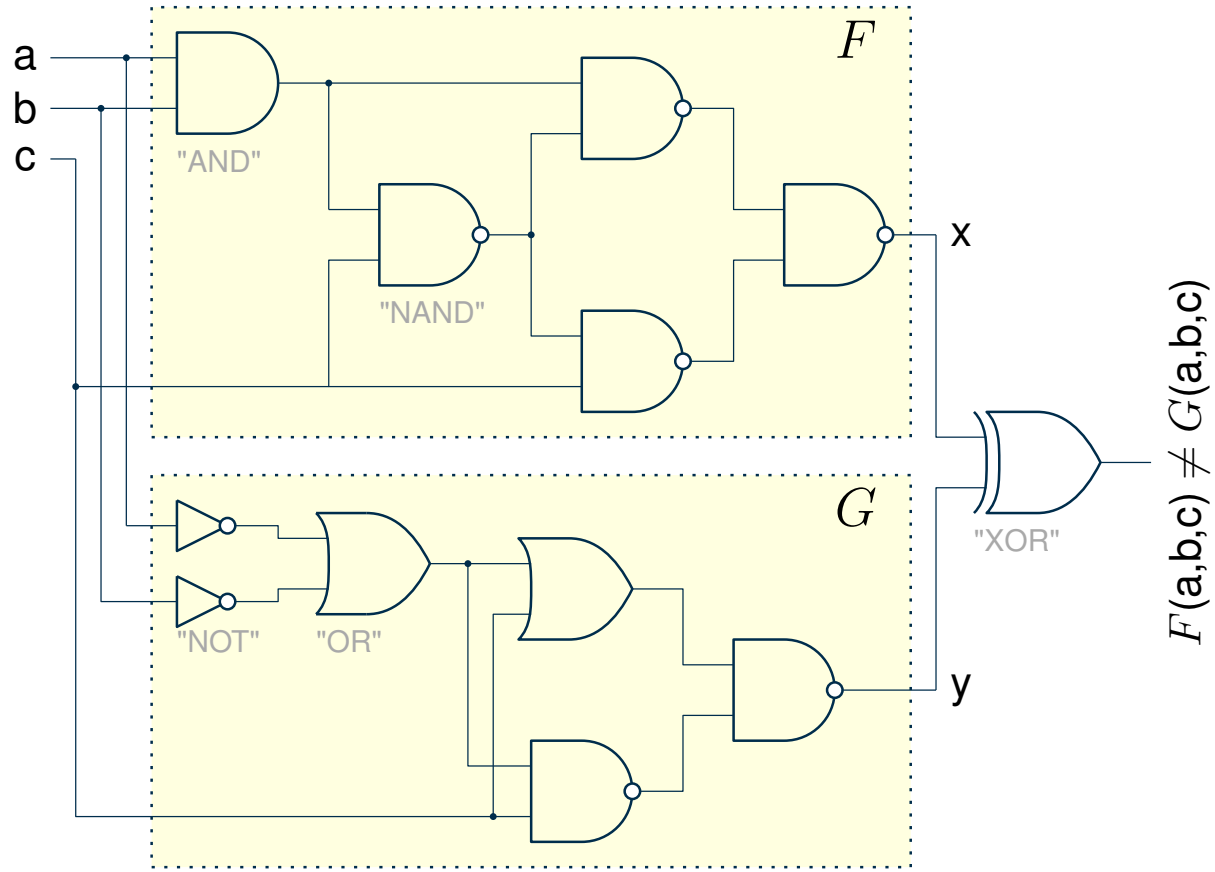


Combinational Equivalence Checking

- Given: two combinational circuits
  stateless "input→output" circuit, no feedback
- Question: are the circuits logically equivalent?
- Right example: Is $F(a, b, c) \equiv G(a, b, c)$?
- How to solve with SAT?

Miter Formula

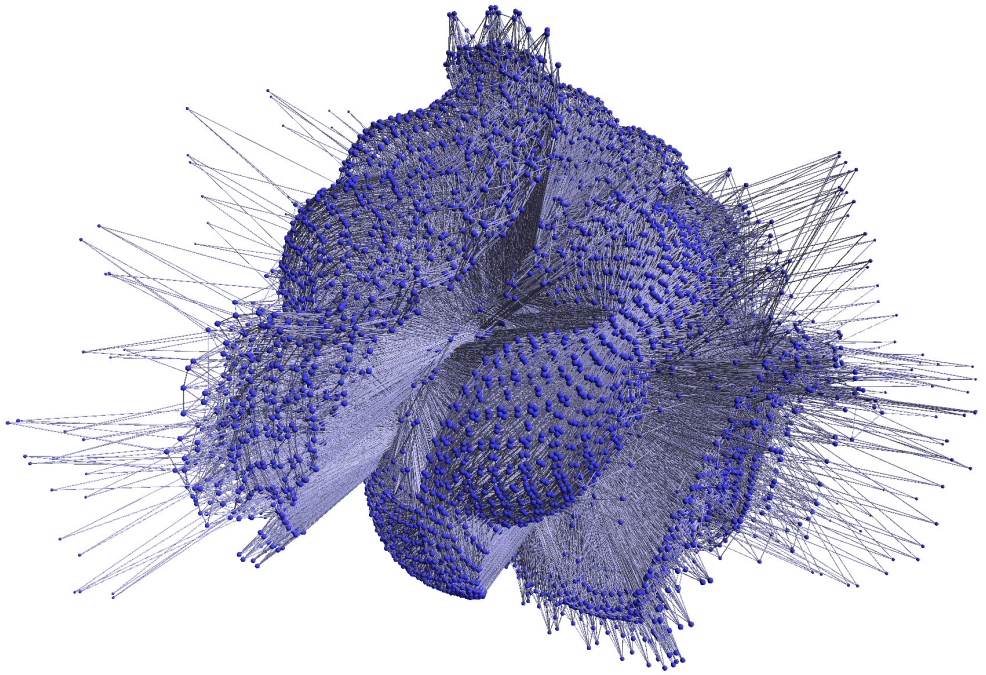
- Encode F, G relative to shared input bits
- Assert $x \neq y$ (multi-bit output: $\bigvee_{x_i, y_i} x_i \neq y_i$)
- Satisfiable $\Leftrightarrow$ ($F \not\equiv G$) (why this way?)



CEC: How Hard Can It Be?

Two Miter examples from SAT Competition 2023

- Instance A: 260k variables, 850k clauses
  - Circuits are not equivalent (“buggy miter”)
  - Solved in 1.33 s by KISSAT_MAB_PROP-NO_SYM [10]

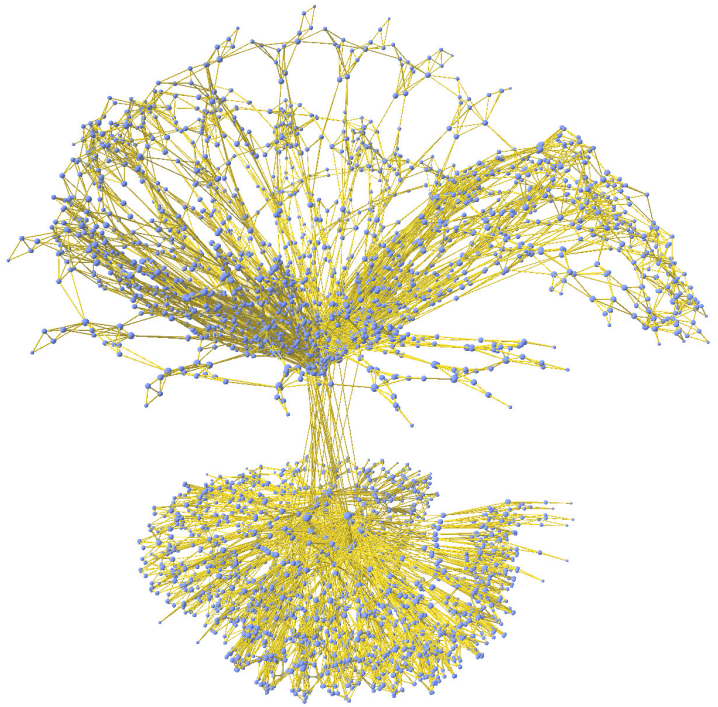


Nodes = variables;
Edges = common clause(s);
variables contracted by factor ≈ 16

CEC: How Hard Can It Be?

Two Miter examples from SAT Competition 2023

- Instance A: 260k variables, 850k clauses
  - Circuits are not equivalent (“buggy miter”)
  - Solved in 1.33 s by KISSAT_MAB_PROP-NO_SYM [10]
- Instance B: 4k variables, 13k clauses
  - Circuits are equivalent
  - Unsolved by sequential solvers within 5000 s
  - Solved by some parallel solvers :)



Nodes = variables;

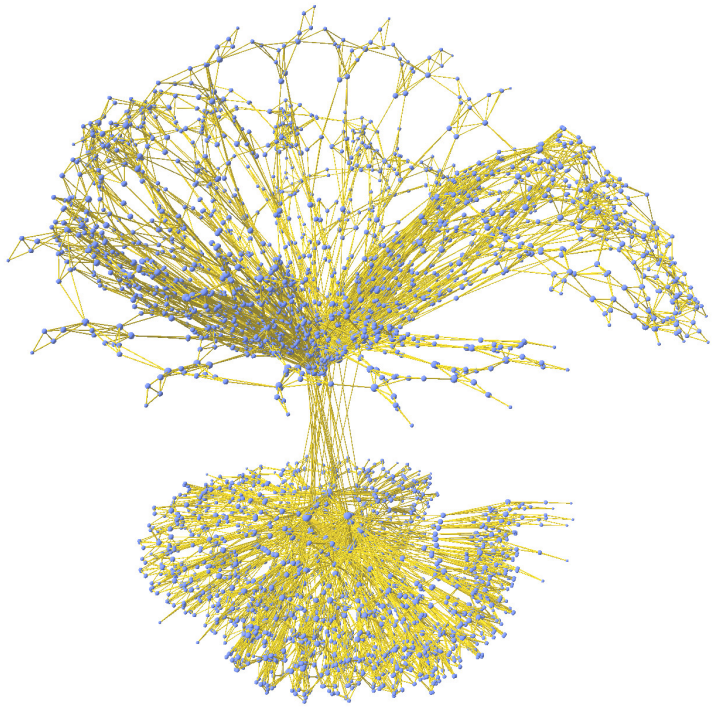
Edges = common clause(s)

CEC: How Hard Can It Be?

Two Miter examples from SAT Competition 2023

- Instance A: 260k variables, 850k clauses
  - Circuits are not equivalent (“buggy miter”)
  - Solved in 1.33 s by KISSAT_MAB_PROP-NO_SYM [10]
- Instance B: 4k variables, 13k clauses
  - Circuits are equivalent
  - Unsolved by sequential solvers within 5000 s
  - Solved by some parallel solvers :)

Generally: co-NP-complete, can require very large proofs

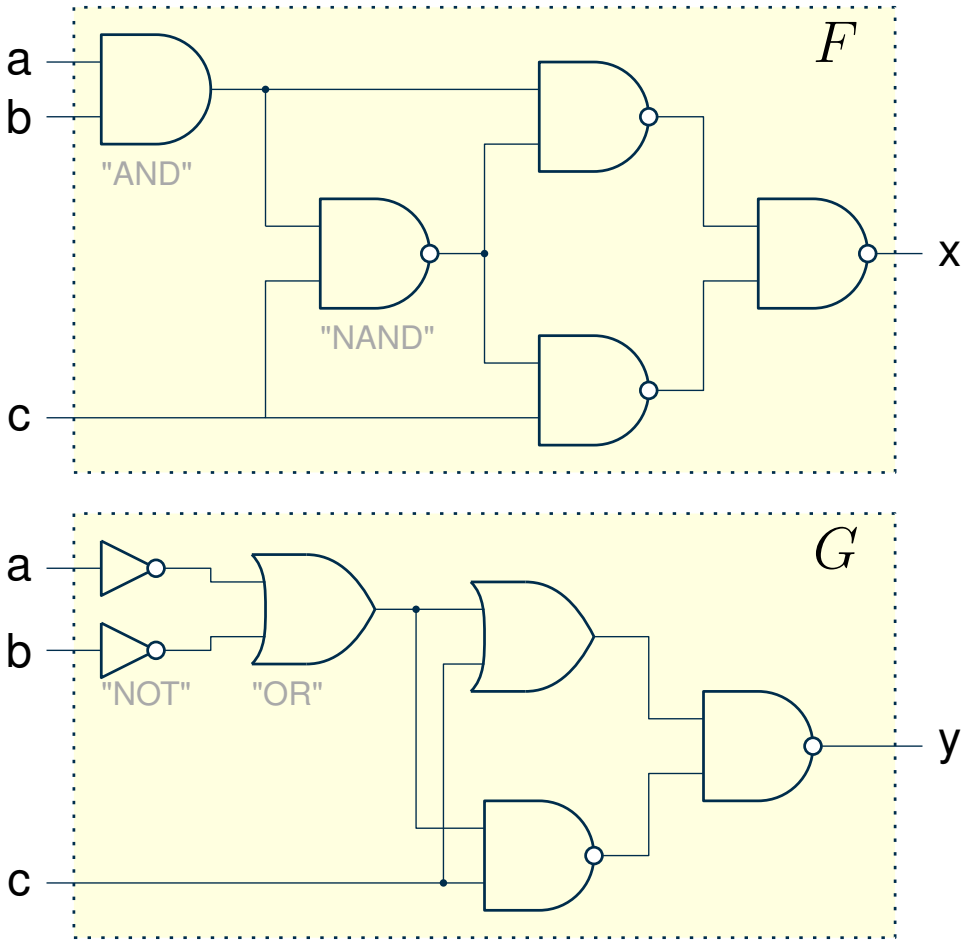


Nodes = variables;
Edges = common clause(s)

CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

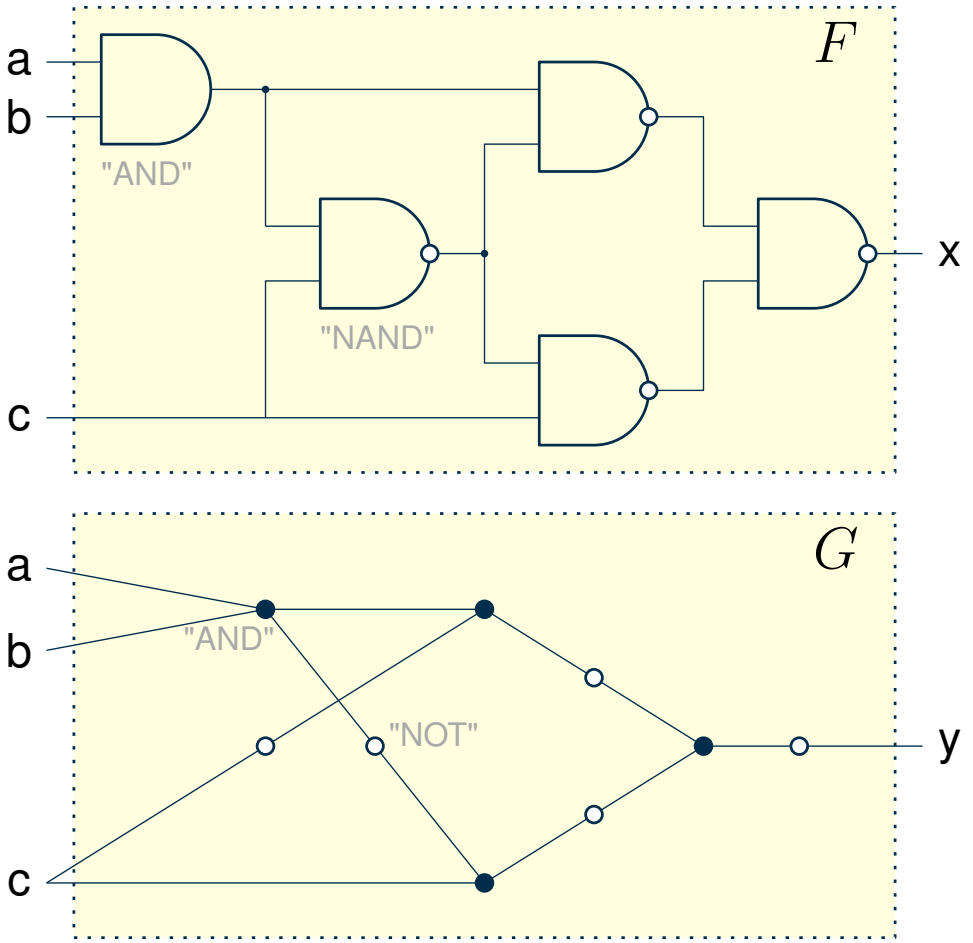
- Randomly test different inputs, collecting pairs of potentially equivalent nodes



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

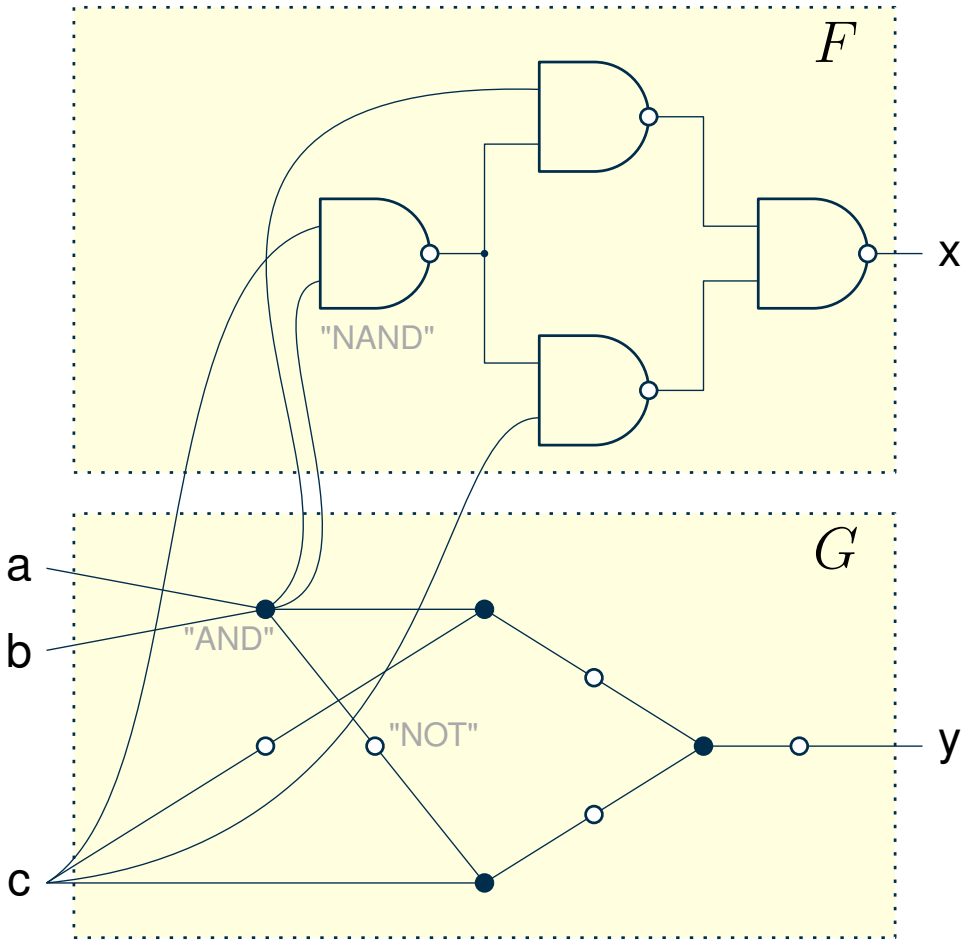
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

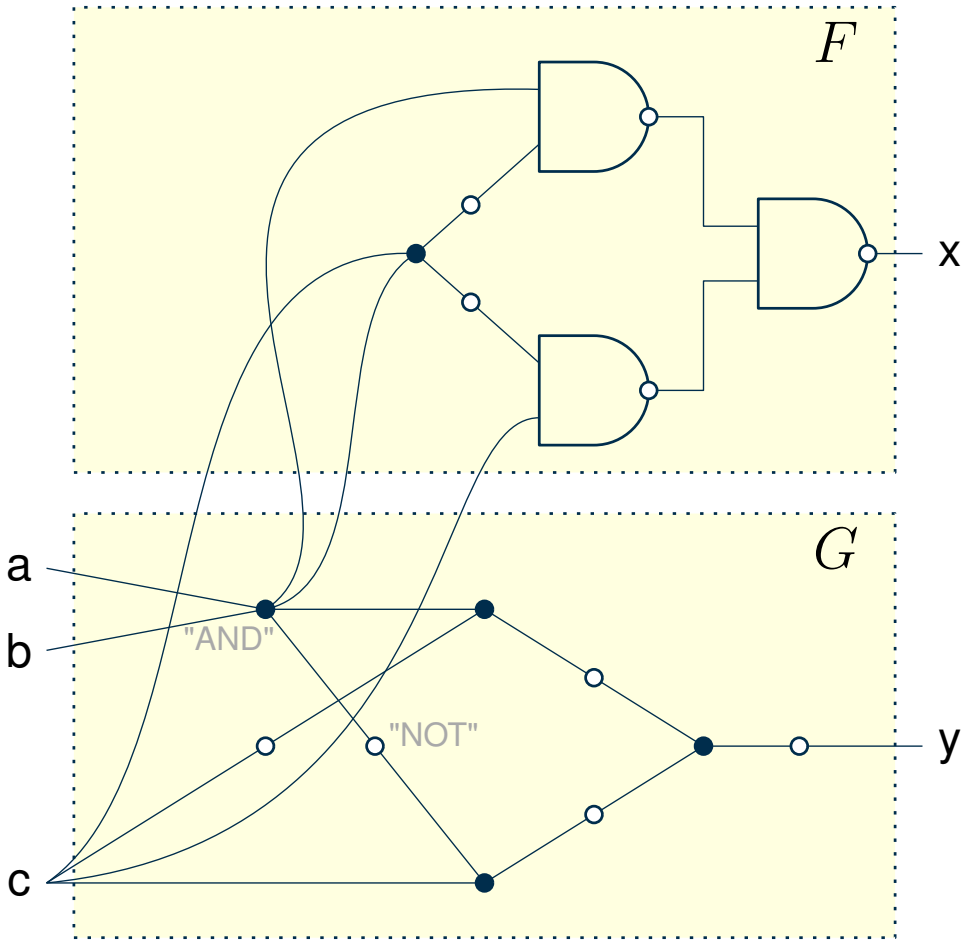
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

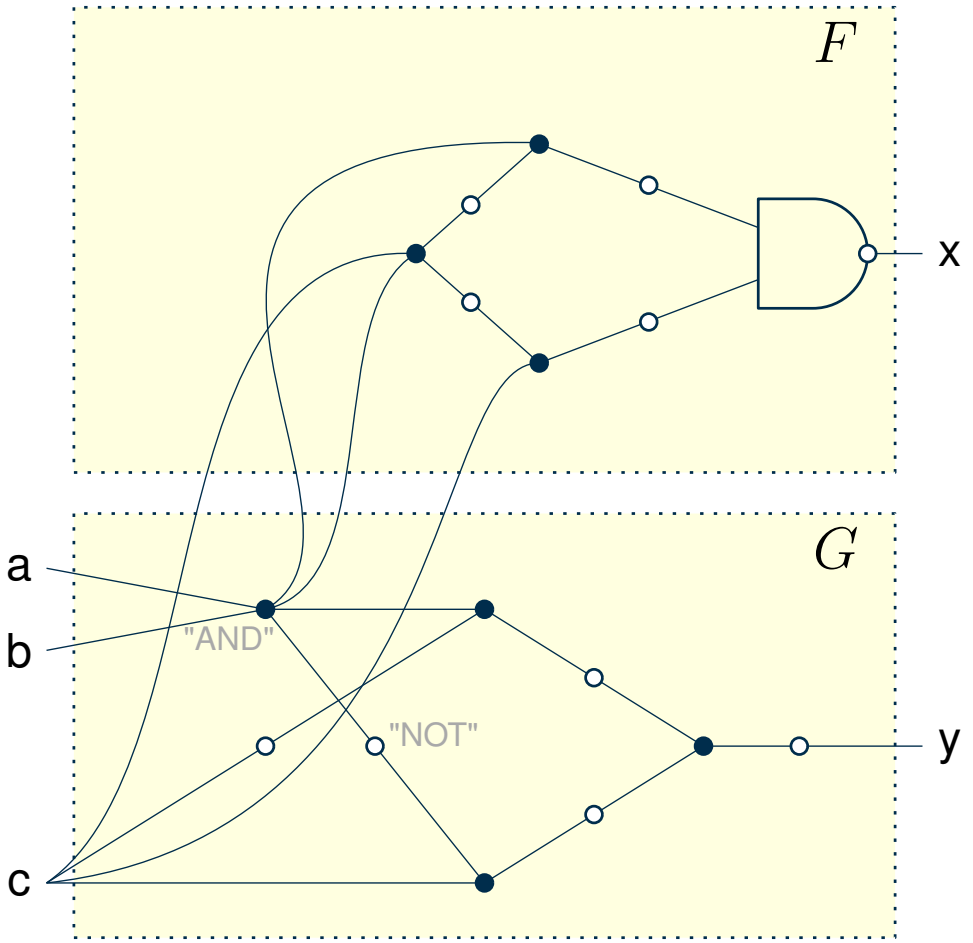
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

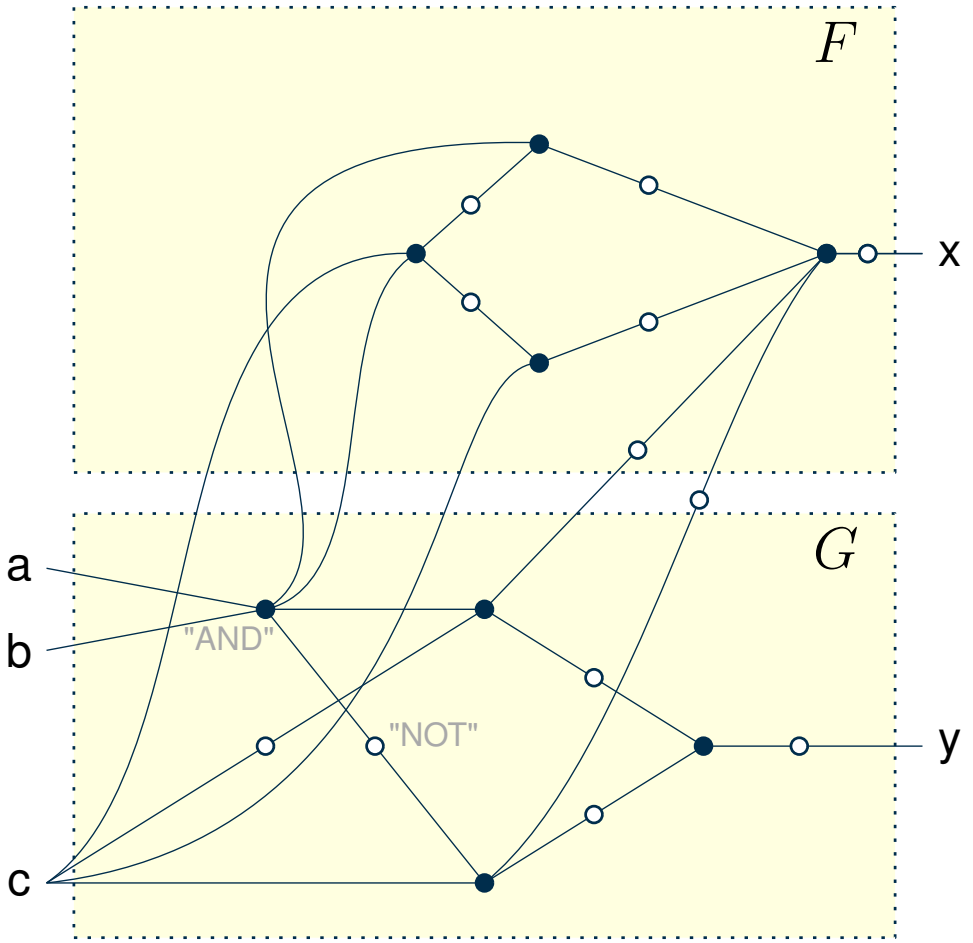
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

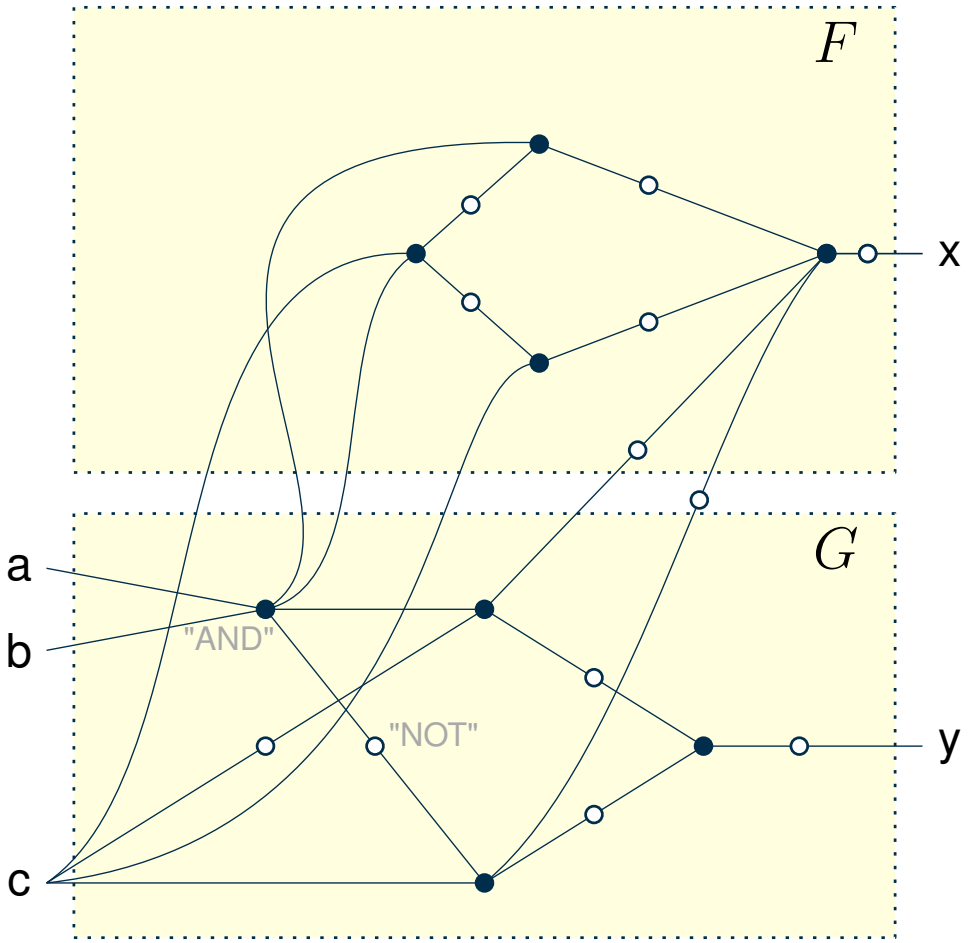
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

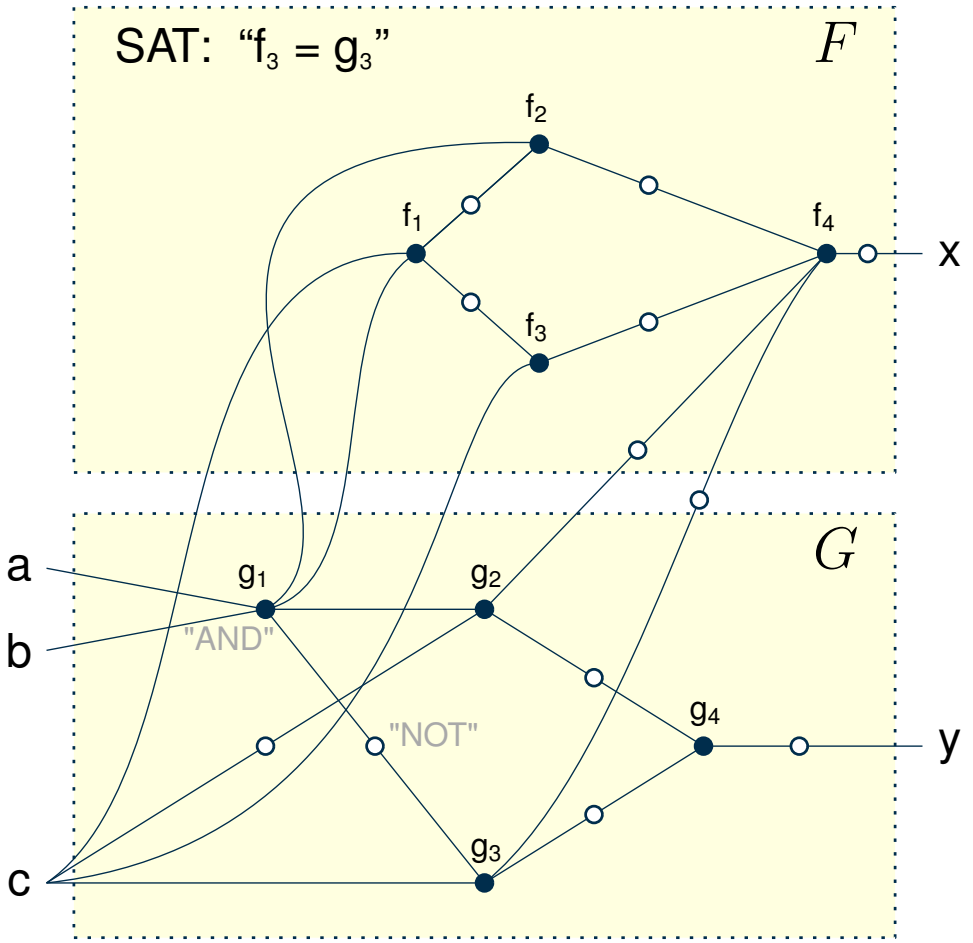
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging
- Structural hashing: Ensure that each functionally distinct sub-circuit is encoded only once



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

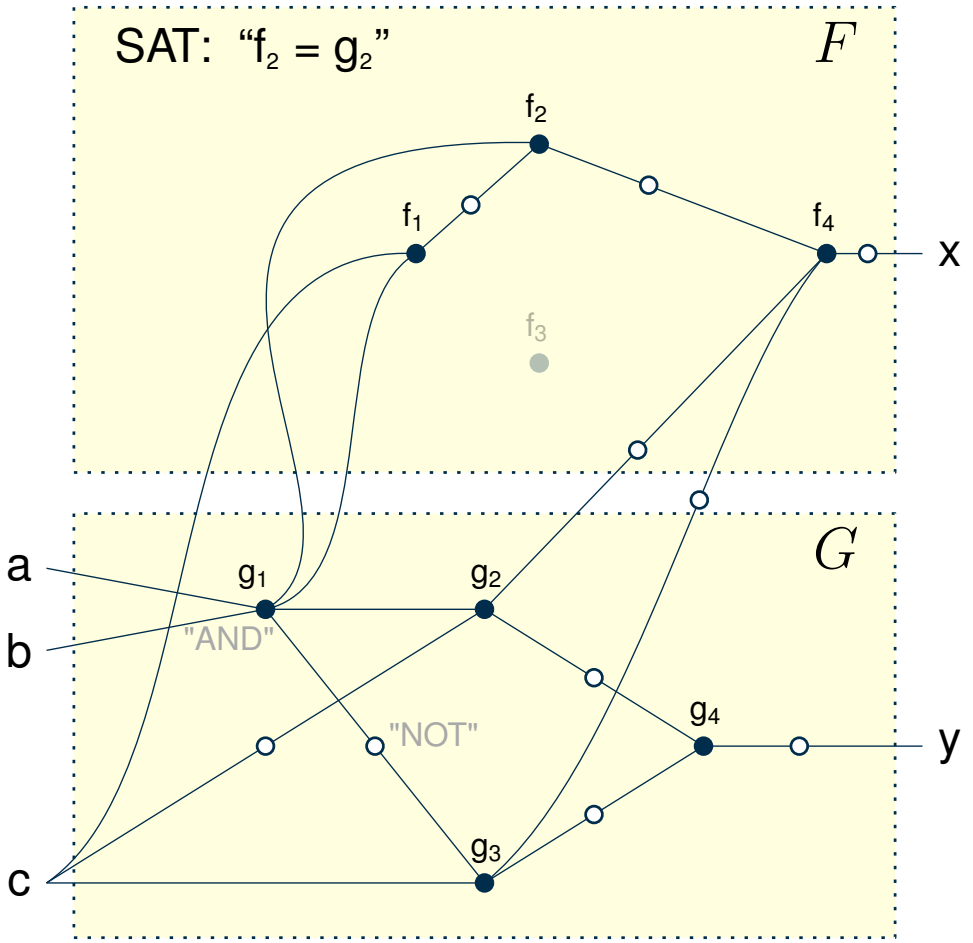
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging
- Structural hashing: Ensure that each functionally distinct sub-circuit is encoded only once
- SAT sweeping: Use SAT sub-program to test whether potentially equivalent nodes are actually equivalent



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

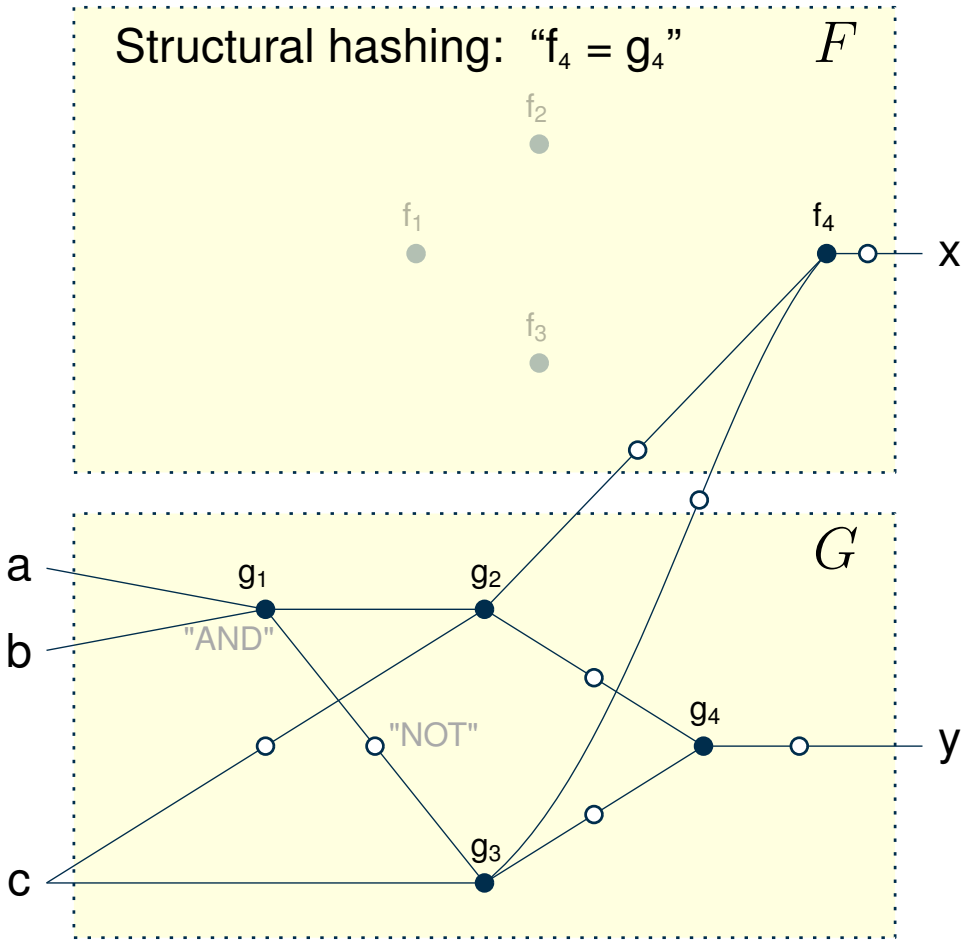
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging
- Structural hashing: Ensure that each functionally distinct sub-circuit is encoded only once
- SAT sweeping: Use SAT sub-program to test whether potentially equivalent nodes are actually equivalent



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

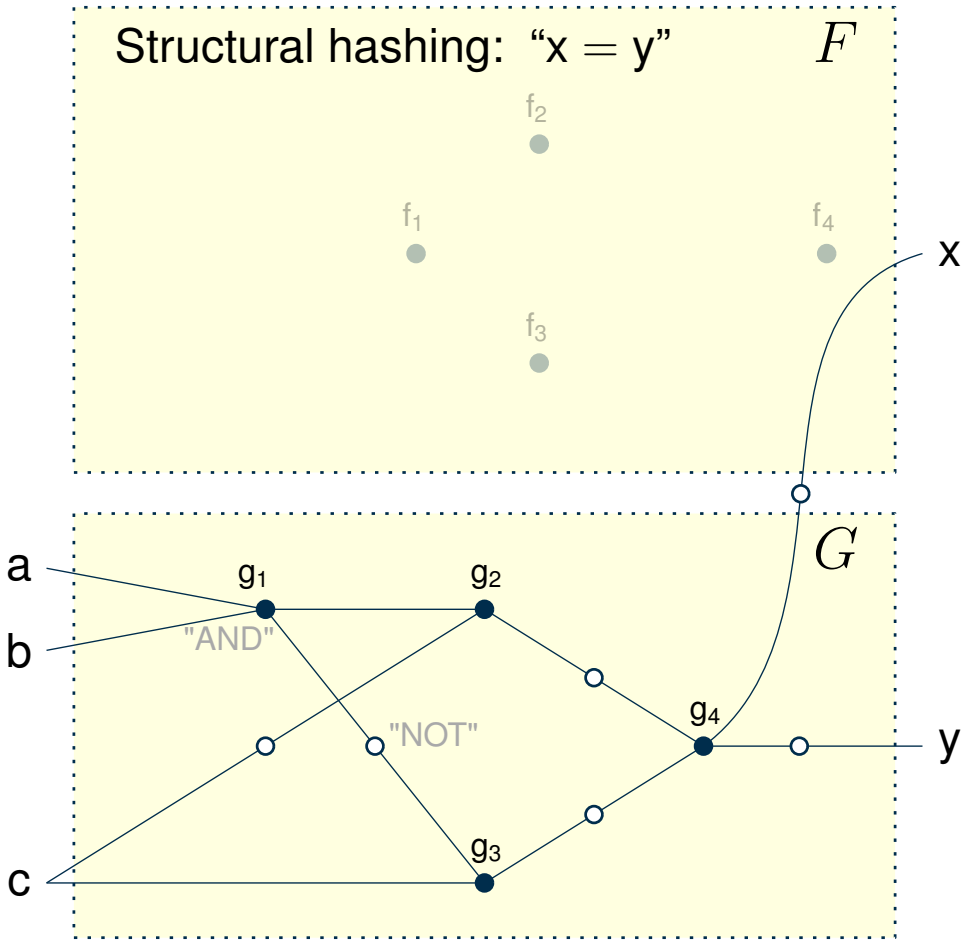
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging
- Structural hashing: Ensure that each functionally distinct sub-circuit is encoded only once
- SAT sweeping: Use SAT sub-program to test whether potentially equivalent nodes are actually equivalent



CEC Techniques [29]

Improving CEC: Try to merge equivalent sub-circuits

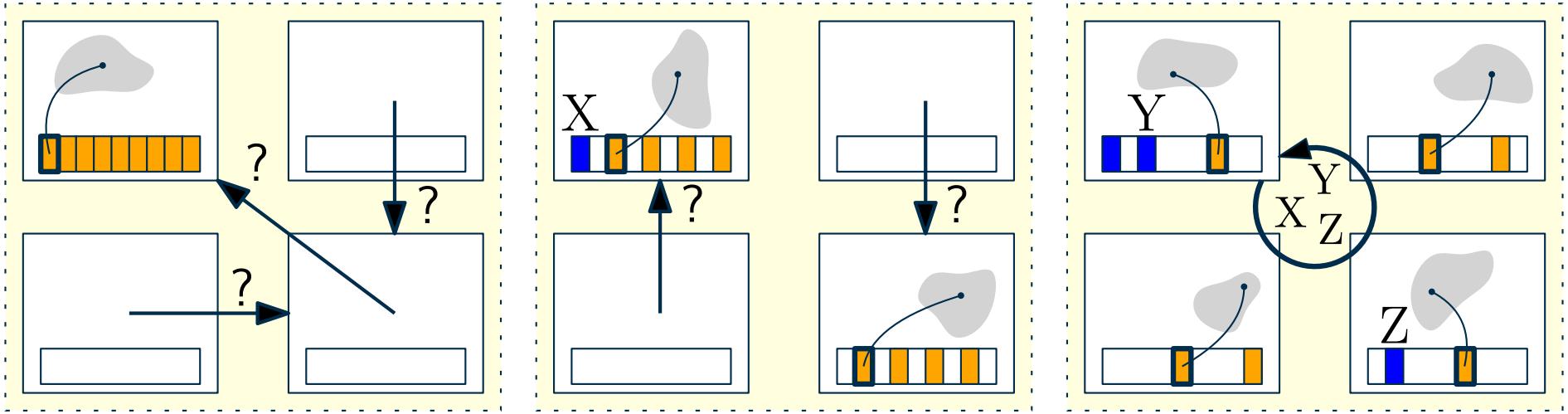
- Randomly test different inputs, collecting pairs of potentially equivalent nodes
- Use And/Inverter-Graph (AIG) for simple circuit manipulation and merging
- Structural hashing: Ensure that each functionally distinct sub-circuit is encoded only once
- SAT sweeping: Use SAT sub-program to test whether potentially equivalent nodes are actually equivalent



Our Latest Work on CEC

Parallel SAT Sweeping (Rigi-Luperti, Biere, Schreiber 2026; to be published)

- No information on circuits – operates purely on CNF!
- Parallelization of Kissat's sweeping:
  Sweeps limited environments in the Variable Incidence Graph starting from selected variables

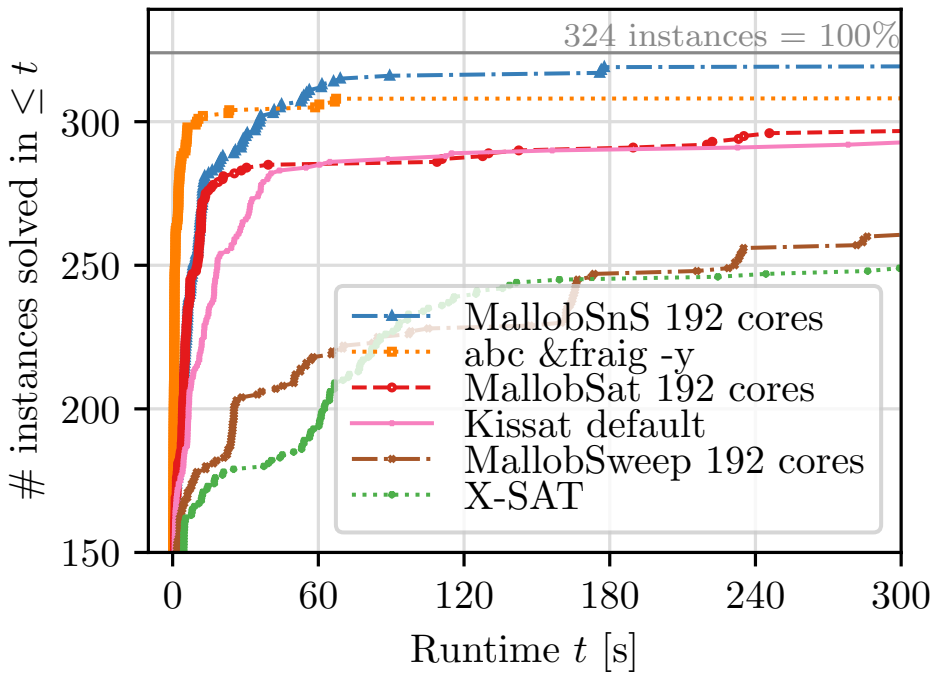


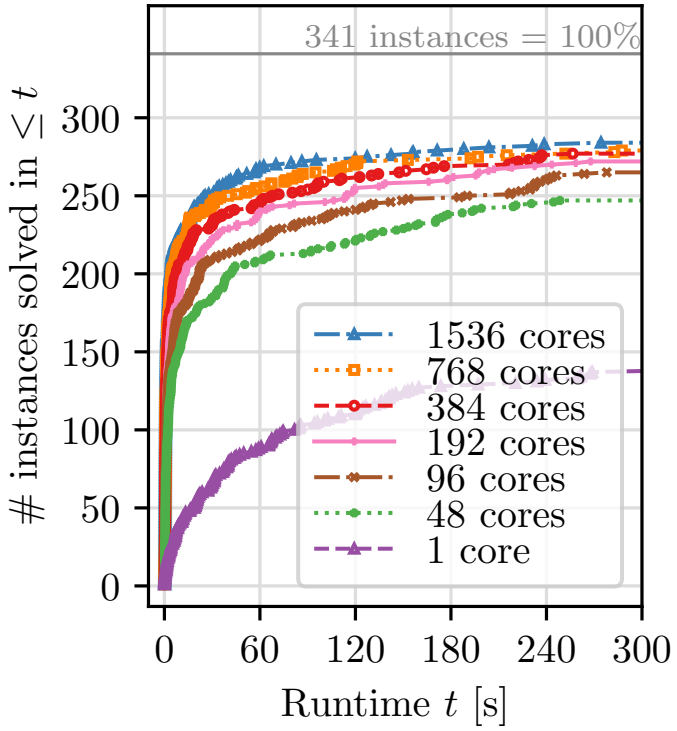
Four workers

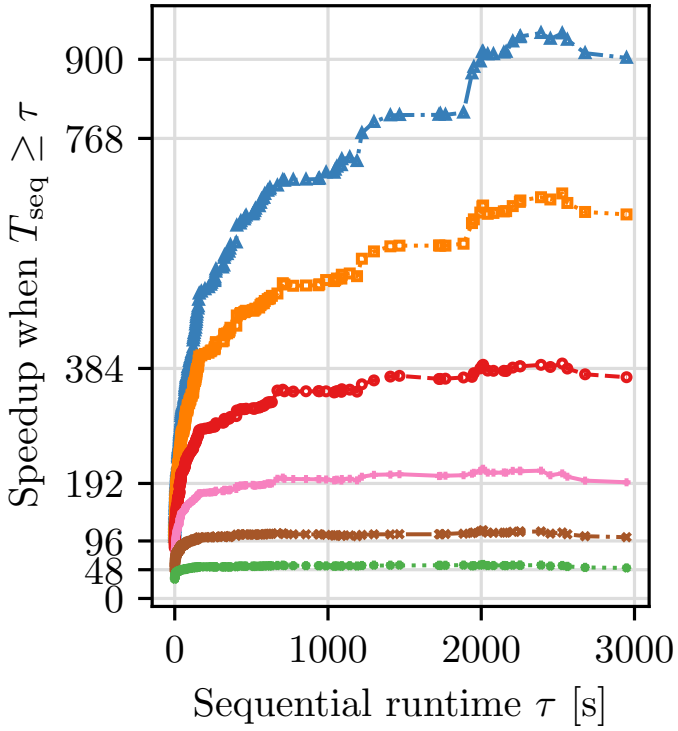
Our Latest Work on CEC

Parallel SAT Sweeping (Rigi-Luperti, Biere, Schreiber 2026; to be published)

- No information on circuits – operates purely on CNF!
- Parallelization of Kissat's sweeping:
  Sweeps limited environments in the Variable Incidence Graph starting from selected variables







Benchmarks from Hardware Model Checking Competition (HWMCC) 2012 and 2020

CEC: Remarks

- Important cornerstone of **Electronic Design Automation** [20]
  – can be used to validate implementation based on specification
  – other EDA techniques: model checking, Automated Test Pattern Generation (ATPG)
- Crucial “intrinsically Boolean” benchmark problem throughout history of SAT solving
  – Every SAT competition features miters
- SAT sweeping originally proposed for bounded model checking [16]
- Gate recognition and merging now a form of general inprocessing for any formula, connected to variable elimination [3]

Analyzing Cryptographic Building Blocks

Cryptanalysis = analyze, attempt to “break” cryptographic building blocks to test, advance them

- Building blocks: Stream ciphers ((msg,key) $\mapsto$ encrypted msg) [26], hash functions [7], . . .
- Algebraic cryptanalysis: try to build equations relating output to input [26]
  - SAT solver should support XOR clauses
  - SAT solver should use Gaussian Elimination as a sub-program

Analyzing Cryptographic Building Blocks

Cryptanalysis = analyze, attempt to “break” cryptographic building blocks to test, advance them

- Building blocks: Stream ciphers ((msg,key) $\mapsto$ encrypted msg) [26], hash functions [7], . . .
- Algebraic cryptanalysis: try to build equations relating output to input [26]
  - SAT solver should support XOR clauses
  - SAT solver should use Gaussian Elimination as a sub-program
- Established SAT-based approaches:
  - Prove mathematical properties of internal states [7]
  - Find weak keys and preimages [17]
  - Find collisions of hash functions [21]

Analyzing Cryptographic Building Blocks

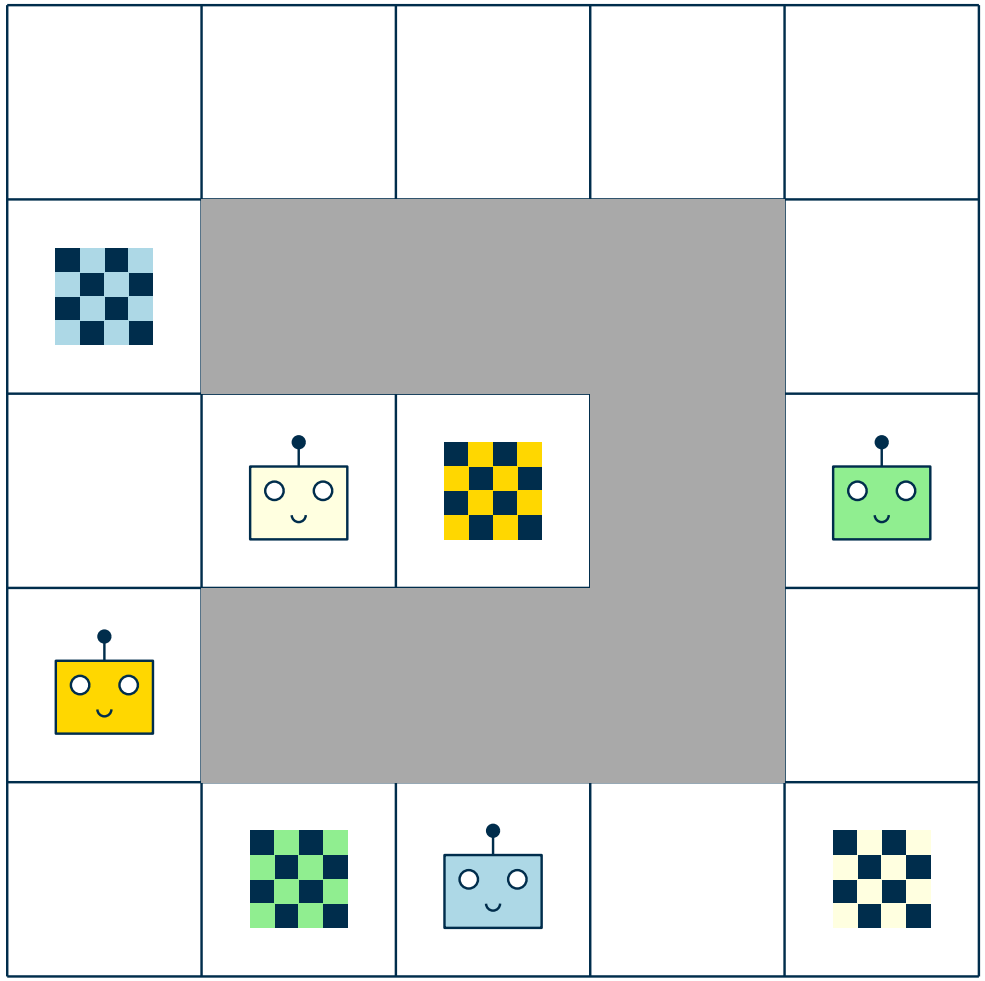
Cryptanalysis = analyze, attempt to "break" cryptographic building blocks to test, advance them

- Building blocks: Stream ciphers ((msg,key) $\mapsto$ encrypted msg) [26], hash functions [7], ...
- Algebraic cryptanalysis: try to build equations relating output to input [26]
  - SAT solver should support XOR clauses
  - SAT solver should use Gaussian Elimination as a sub-program
- Established SAT-based approaches:
  - Prove mathematical properties of internal states [7]
  - Find weak keys and preimages [17]
  - Find collisions of hash functions [21]
- Cross-application use of SAT techniques:
  - Cryptanalysis via SMT solving [31]
  - Cryptanalysis via bounded model checking [19]
  - Hash function analysis also used in algorithm design [30]

"*it turns out that the highly combinatorial nature of the problem is not well suited for linear solvers, and that SAT solvers are a better fit for this type of problem*" —Dobraunig et al. after trying MILP for ASCON [7]

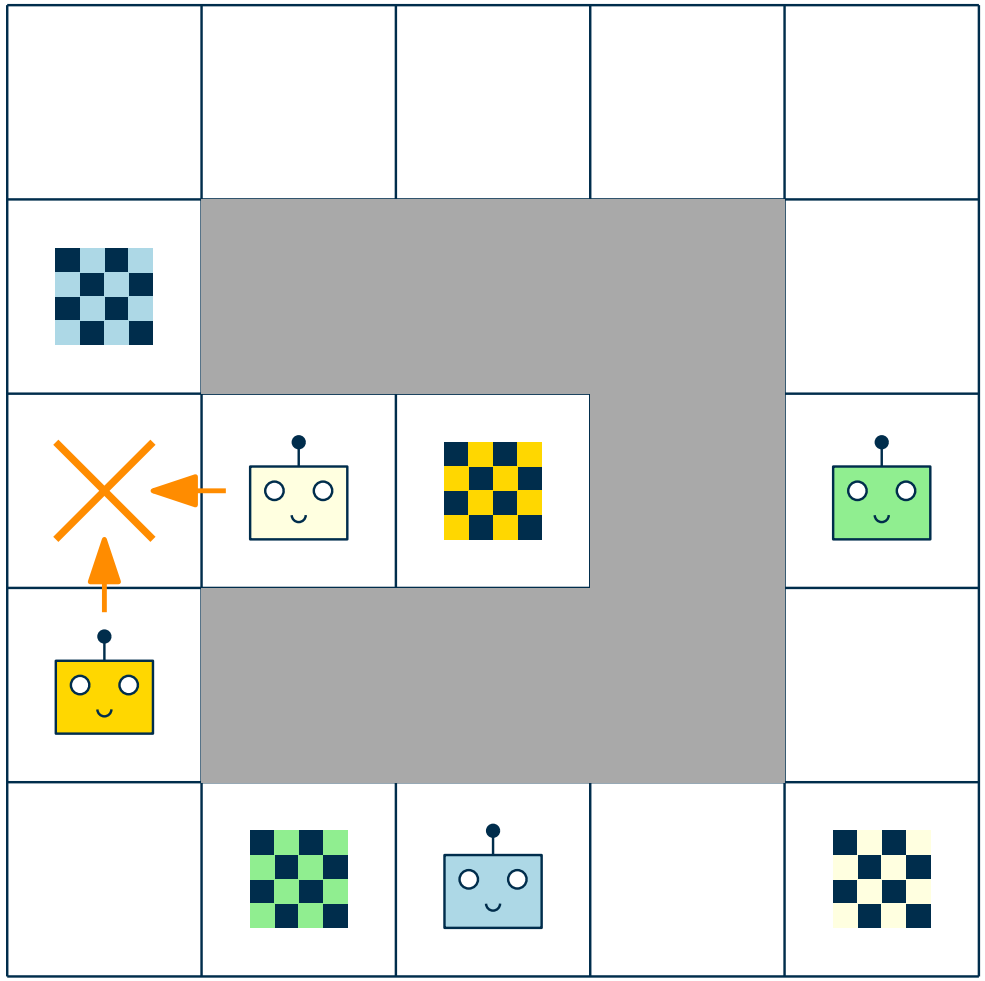
Multi Agent Path Finding [27]

- Discretized 2D grid of positions, n cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position



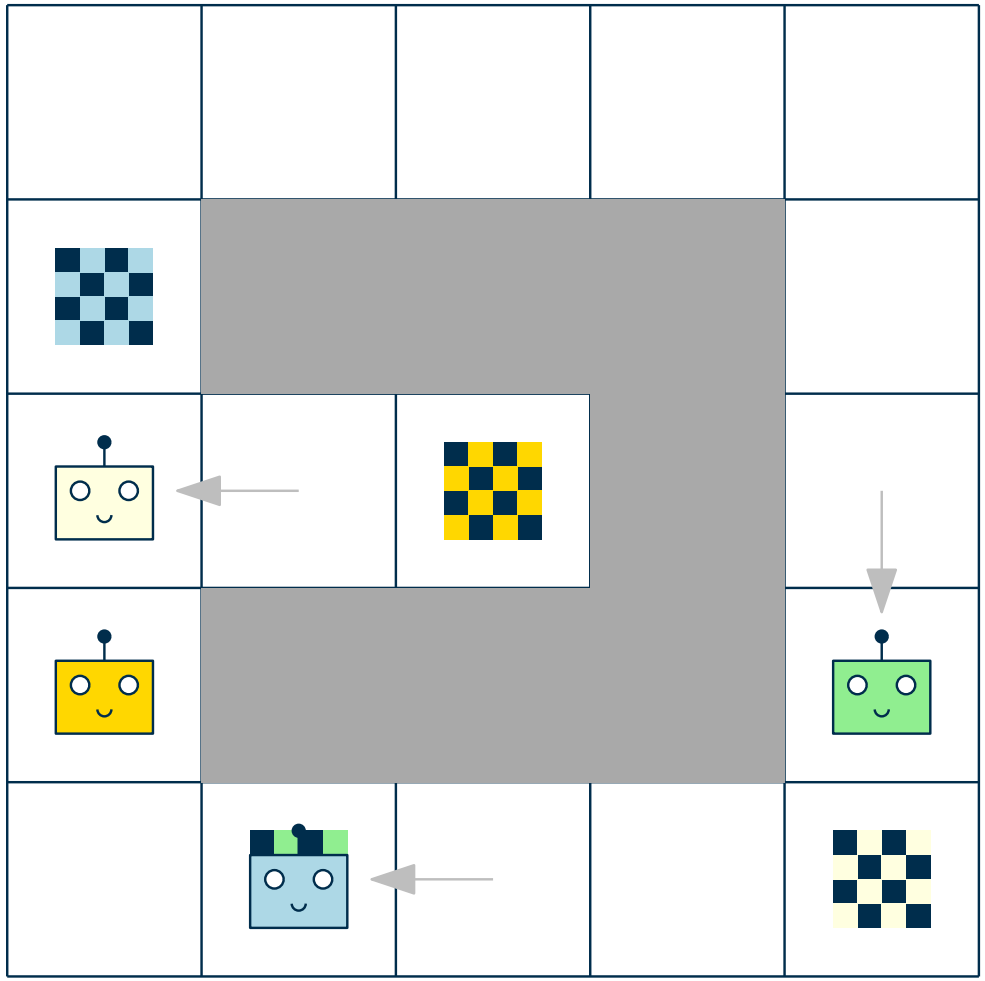
Multi Agent Path Finding [27]

- Discretized 2D grid of positions, n cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position
- Collisions disallowed
- Optimize makespan (= steps until all goals reached) or Sum of Costs (= total number of actions performed)



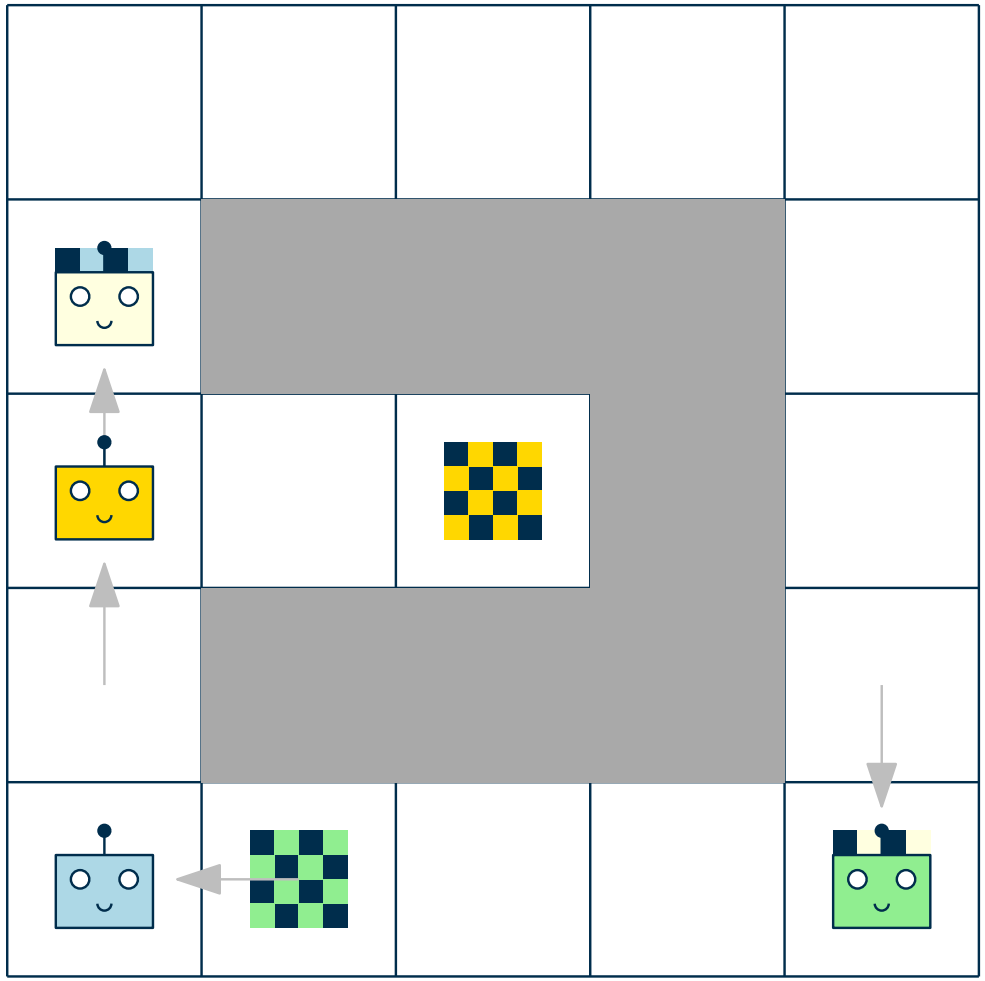
Multi Agent Path Finding [27]

- Discretized 2D grid of positions, n cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position
- Collisions disallowed
- Optimize makespan (= steps until all goals reached) or Sum of Costs (= total number of actions performed)



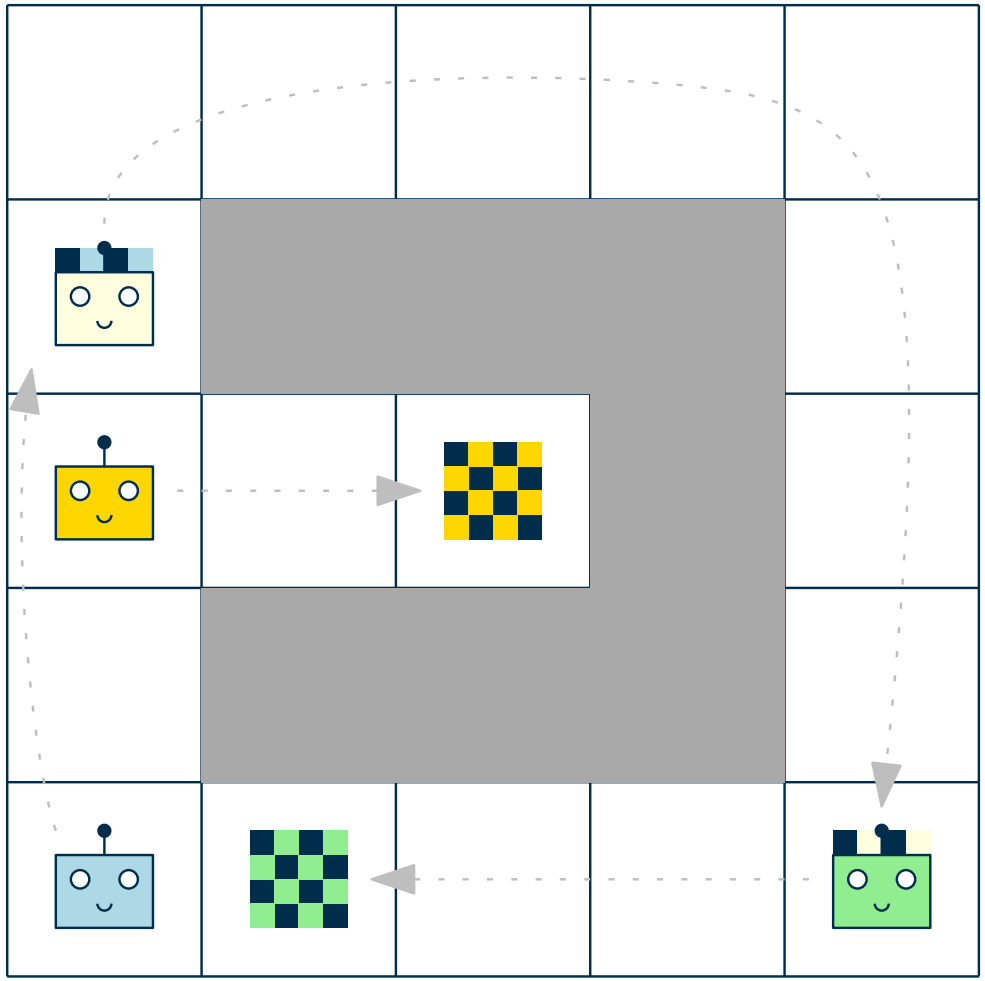
Multi Agent Path Finding [27]

- Discretized 2D grid of positions, *n* cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position
- Collisions disallowed
- Optimize makespan (= steps until all goals reached) or Sum of Costs (= total number of actions performed)



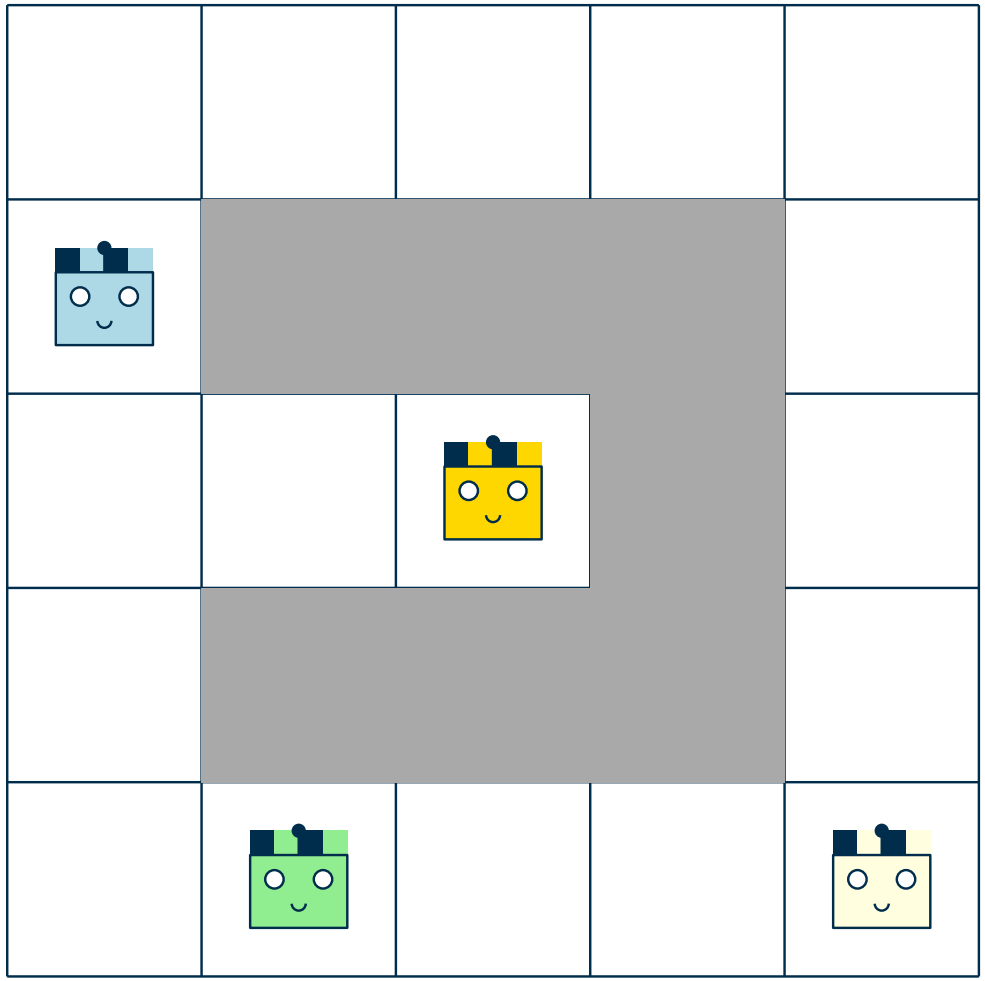
Multi Agent Path Finding [27]

- Discretized 2D grid of positions, n cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position
- Collisions disallowed
- Optimize makespan (= steps until all goals reached) or Sum of Costs (= total number of actions performed)



Multi Agent Path Finding [27]

- Discretized 2D grid of positions, n cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position
- Collisions disallowed
- Optimize makespan (= steps until all goals reached) or Sum of Costs (= total number of actions performed)

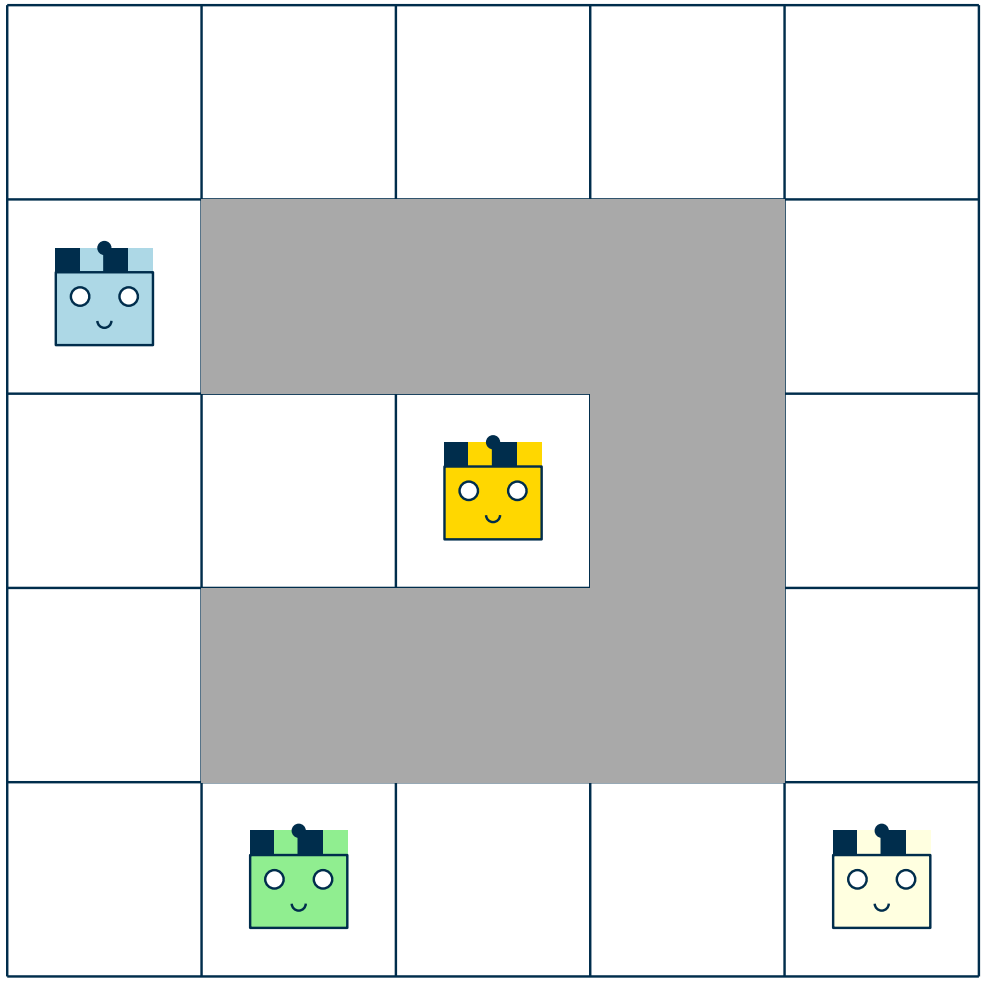


Multi Agent Path Finding [27]

- Discretized 2D grid of positions, n cooperative agents
- Discretized time steps: move 0-1 cells per time step
- Per agent: Initial position and goal position
- Collisions disallowed
- Optimize makespan (= steps until all goals reached) or Sum of Costs (= total number of actions performed)

Optimal Approaches to MAPF

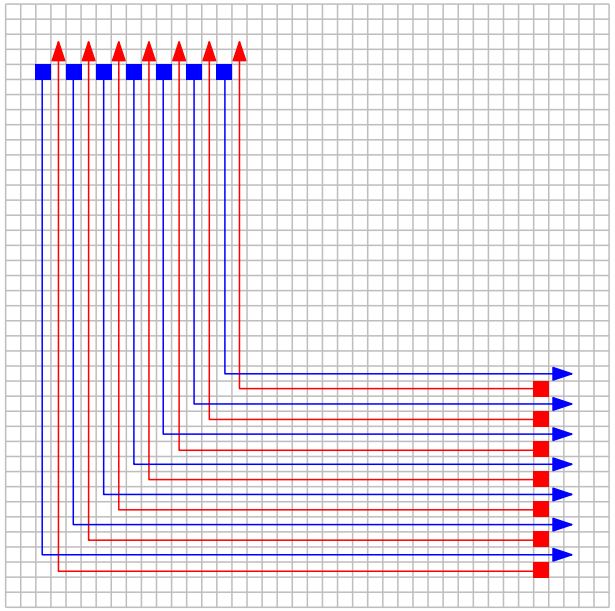
- M* algorithm: adjusted A^* with collision handling and backtracking
- Conflict Based Search (CBS): route individually; at collision, add constraint to a colliding agent and re-route
- Reduction-based approaches: SAT, MaxSAT, ASP, CSP



Is SAT-based MAPF worthwhile?

Observation on MAPF [27] (and planning, and scheduling, and probably many other problems . . .):

Direct search-based approaches
perform especially well on large,
lightly constrained instances.



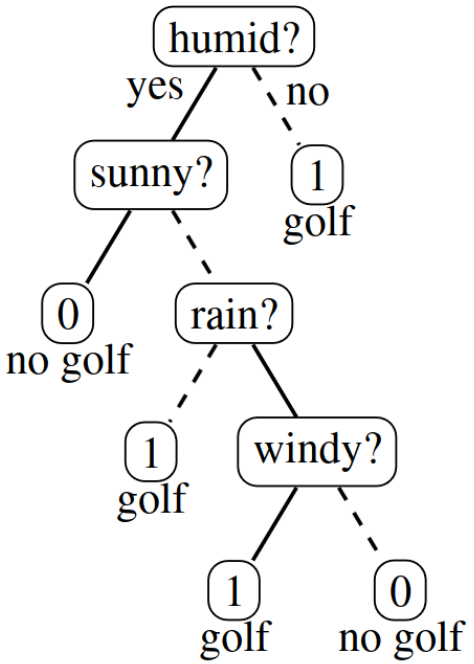
SAT-based approaches
perform especially well on small-sized,
highly constrained instances.

13	7	2	4
3	8	15	5
9	12	1	6
14	11	10	

Explainable AI: Learning Decision Trees

- Given: n d-dimensional sample vectors (d features) mapped to a (binary) class
- Task: Learn decision tree classifying all samples
- Explainable classifier (the more shallow the better)

sample	sunny	rain	overcast	temp:mild	temp:hot	temp:cool	humid	windy	play golf
e_1	0	1	0	0	0	1	0	1	1
e_2	0	1	0	1	0	0	0	1	1
e_3	1	0	0	0	1	0	1	1	0
e_4	0	0	1	0	0	1	0	0	1
e_5	1	0	0	1	0	0	0	0	1
e_6	0	1	0	1	0	0	1	0	0
e_7	0	1	0	1	0	0	1	1	1
e_8	0	0	1	0	1	0	1	1	1
e_9	1	0	0	1	0	0	1	1	0
e_{10}	1	0	0	0	0	1	0	1	1
e_{11}	0	0	1	1	0	0	1	0	1
e_{12}	1	0	0	0	1	0	1	0	0



taken from [25]

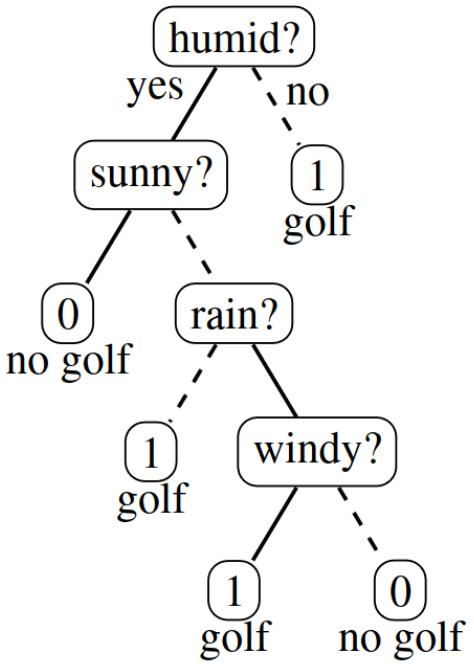
Explainable AI: Learning Decision Trees

- Given: n d-dimensional sample vectors (d features) mapped to a (binary) class
- Task: Learn decision tree classifying all samples
- Explainable classifier (the more shallow the better)

SAT-based approach [22]

- Encode complete binary tree of depth k
- Encode recursively for each node which samples are excluded along its path
- Constrain that 0-leaves exclude all 1-labeled samples and vice versa
- Solver picks a sub-tree and each node's feature

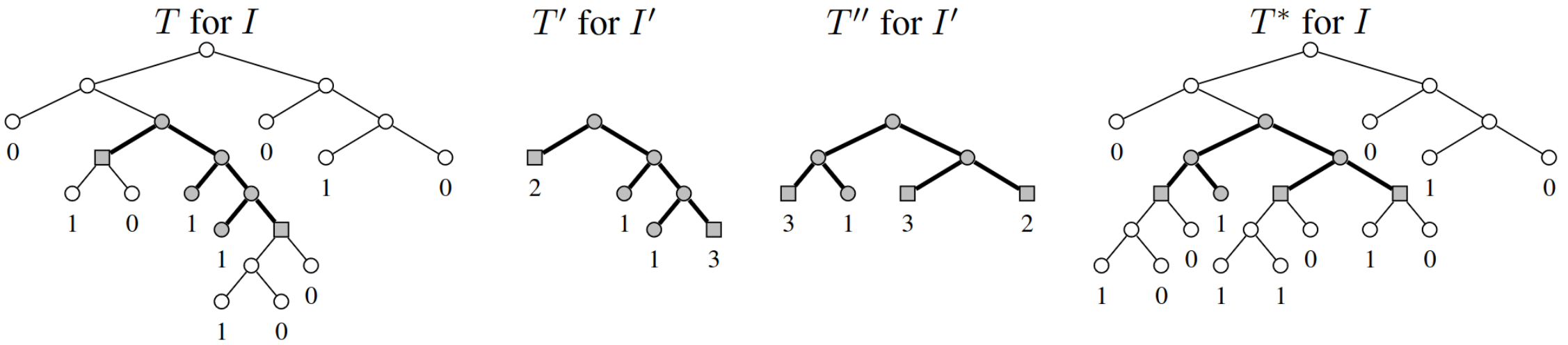
sample	sunny	rain	overcast	temp:mild	temp:hot	temp:cool	humid	windy	play golf
e_1	0	1	0	0	0	1	0	1	1
e_2	0	1	0	1	0	0	0	1	1
e_3	1	0	0	0	1	0	1	1	0
e_4	0	0	1	0	0	1	0	0	1
e_5	1	0	0	1	0	0	0	0	1
e_6	0	1	0	1	0	0	1	0	0
e_7	0	1	0	1	0	0	1	1	1
e_8	0	0	1	0	1	0	1	1	1
e_9	1	0	0	1	0	0	1	1	0
e_{10}	1	0	0	0	0	1	0	1	1
e_{11}	0	0	1	1	0	0	1	0	1
e_{12}	1	0	0	0	1	0	1	0	0



taken from [25]

SAT-based Improvement of Decision Trees [25]

- SAT-based approach slow/infeasible for large data sets
- Better: Construct initial decision tree heuristically, then locally improve sub-trees via SAT
  - **Hybrid approach**, also beneficial in other contexts, e.g., CEC, planning [9]
- Enables to scale up merits of SAT to arbitrarily large data sets



taken from [25]

More on SAT Solving × Machine Learning

- Algorithm selection

  Xu, Lin, et al. “SATzilla: portfolio-based algorithm selection for SAT.” JAIR 2008.

  Eggensperger, Katharina, Marius Lindauer, and Frank Hutter. “Neural networks for predicting algorithm runtime distributions.” IJCAI 2018.

- SAT Solving featuring ML techniques

  Liang, Jia Hui, et al. “Learning rate based branching heuristic for SAT solvers.” SAT 2016.

  Guo, Wenxuan, et al. “Machine learning methods in solving the boolean satisfiability problem.” Machine Intelligence Research (2023).

- Verify Neural Networks via SAT/SMT solving

  Huang, Xiaowei, et al. “Safety verification of deep neural networks.” CAV 2017.

  Ehlers, Ruediger. “Formal verification of piece-wise linear feed-forward neural networks.” ATVA 2017.

- Analyze and understand behavior of SAT solvers and instances

  Soos, Mate, Raghav Kulkarni, and Kuldeep S. Meel. “CrystalBall: gazing in the black box of SAT solving.” SAT 2019.

  Fuchs, Tobias, Jakob Bach, and Ashlin Iser. “Active Learning for SAT Solver Benchmarking.” TACAS 2023.

Train Scheduling with Disruptions via MaxSAT [18]

Can SAT Solving fix the Deutsche Bahn?

Train Scheduling with Disruptions via MaxSAT [18]

Can SAT Solving fix the Deutsche Bahn?

No.

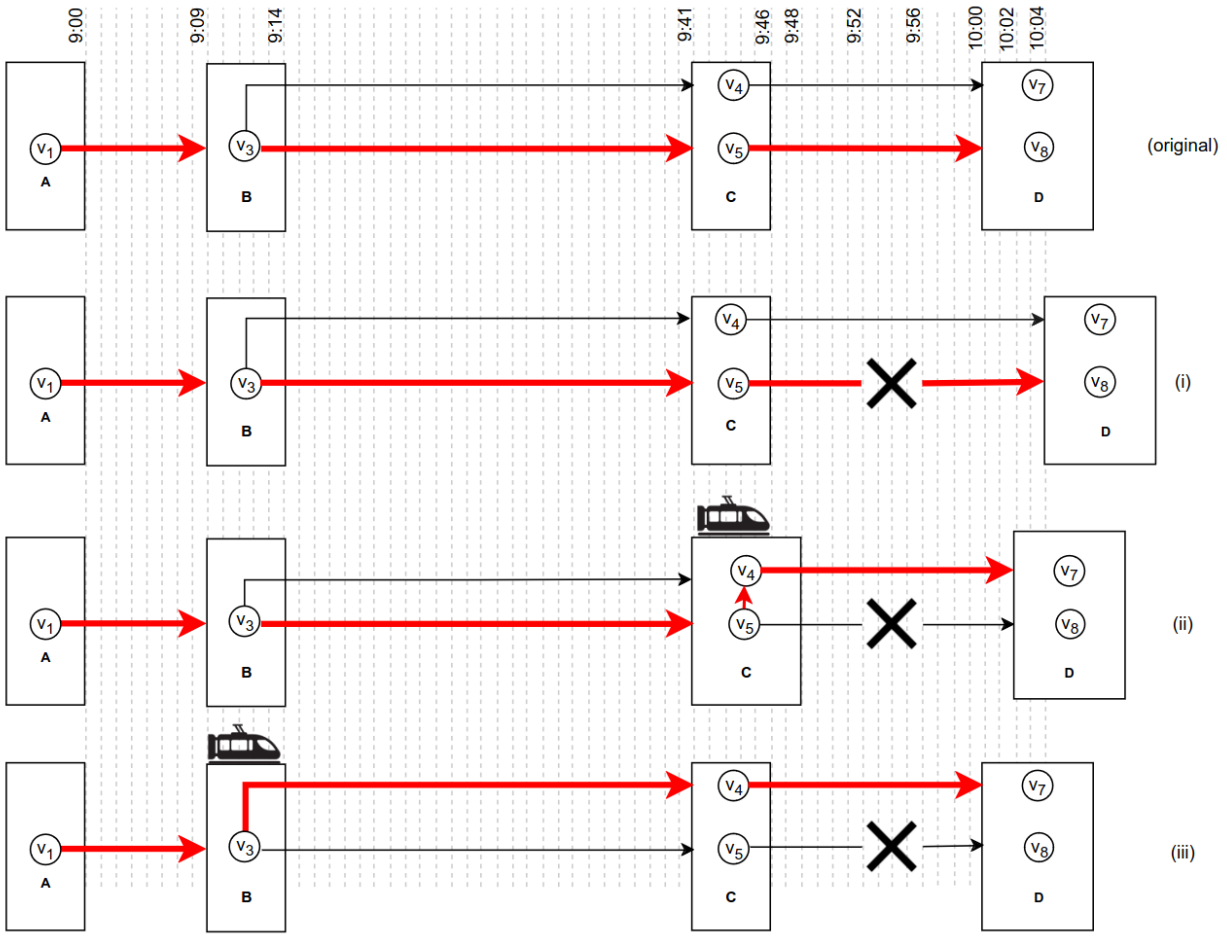
Train Scheduling with Disruptions via MaxSAT [18]

Can SAT Solving fix the Deutsche Bahn?
No.
But it might make the Swiss trains run even better!

Train Scheduling with Disruptions via MaxSAT [18]

Train Scheduling Optimization Problem (TSOP)

- Route trains through predefined stations
- Schedule train timetable subject to time and resource constraints



taken from [18]

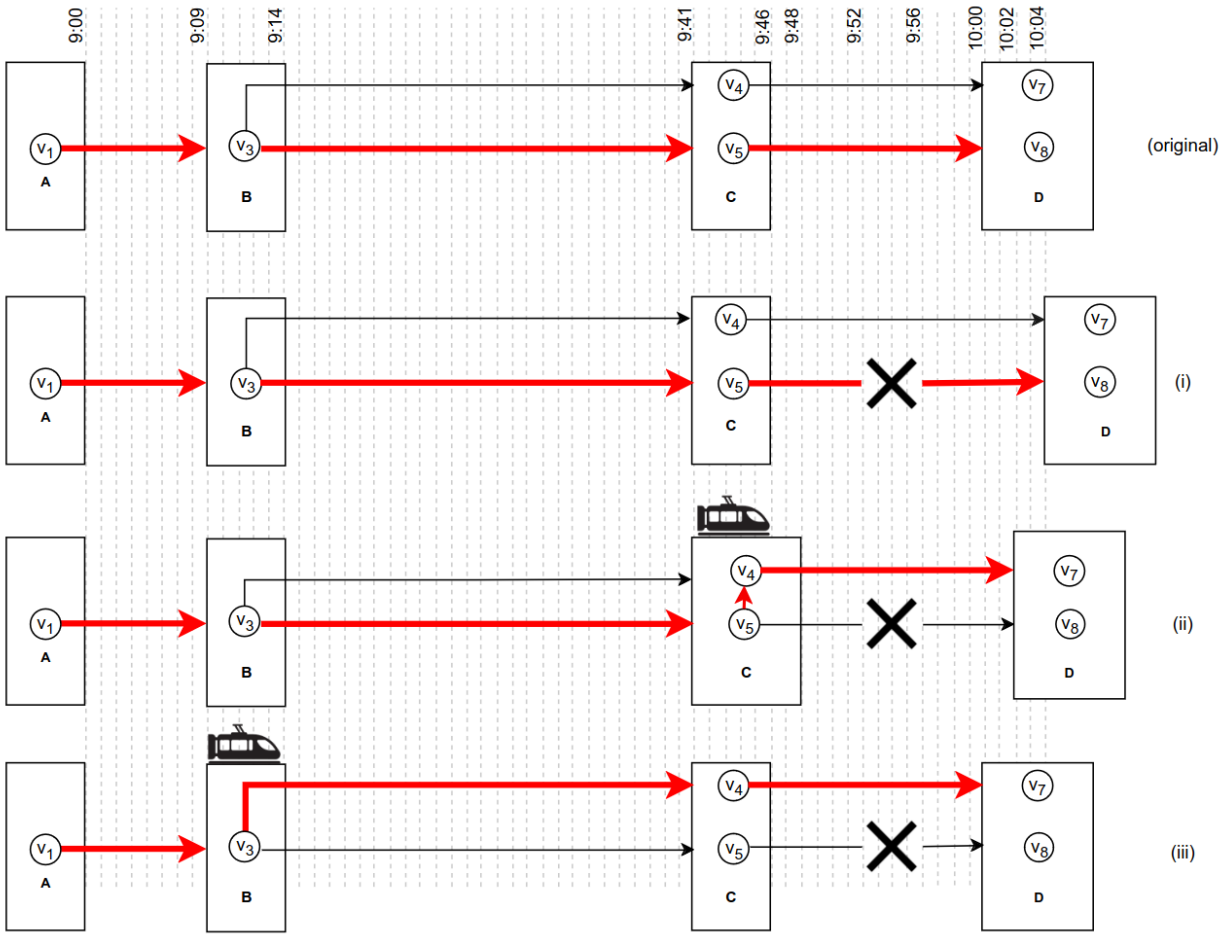
Train Scheduling with Disruptions via MaxSAT [18]

Train Scheduling Optimization Problem (TSOP)

- Route trains through predefined stations
- Schedule train timetable subject to time and resource constraints

TSOP Under Disruption (TSOPUD)

- Numerous disruptions: slowdown, train blocked, track blocked, staffing / rolling stock
- Reroute, reschedule to minimize delays



taken from [18]

Train Scheduling with Disruptions via MaxSAT [18]

Train Scheduling Optimization Problem (TSOP)

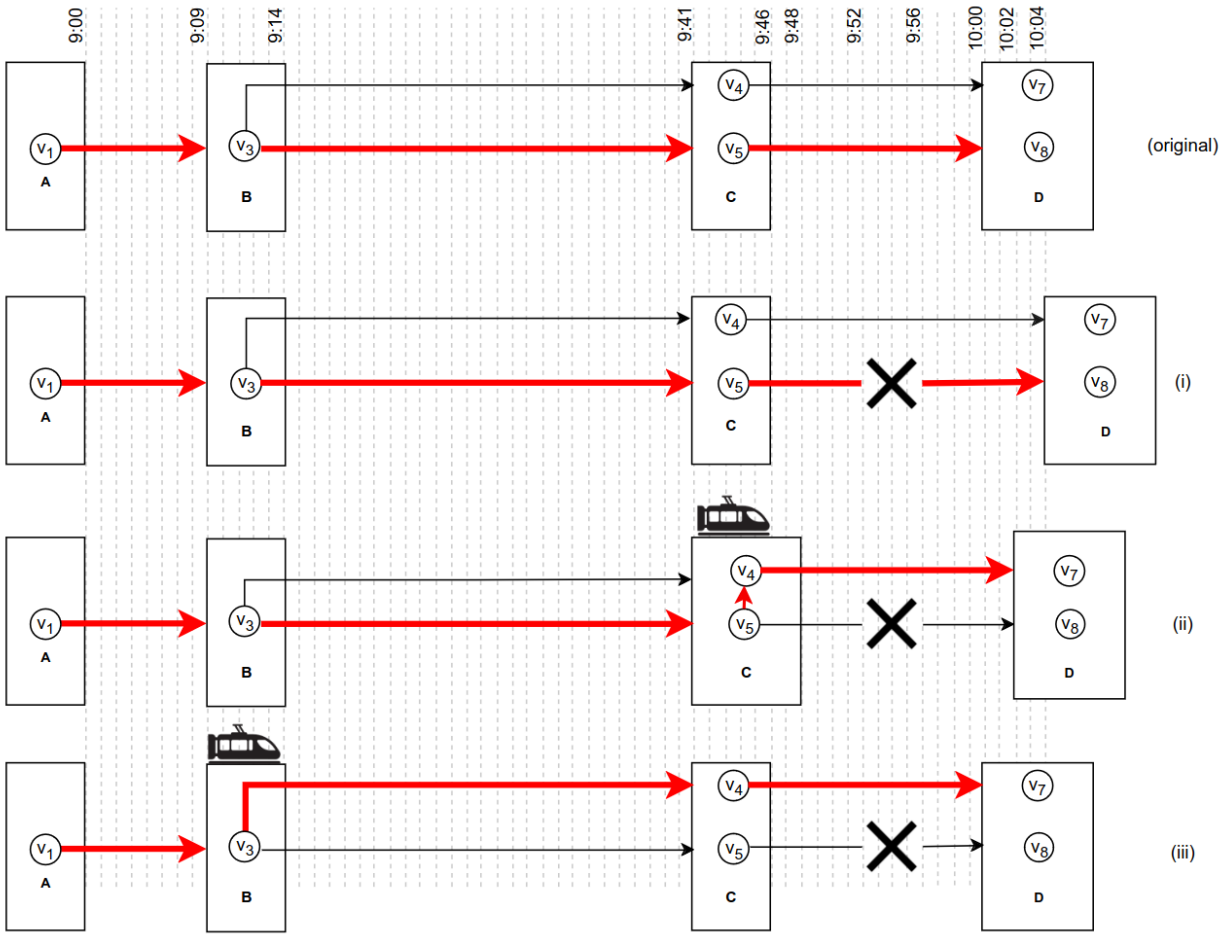
- Route trains through predefined stations
- Schedule train timetable subject to time and resource constraints

TSOP Under Disruption (TSOPUD)

- Numerous disruptions: slowdown, train blocked, track blocked, staffing / rolling stock
- Reroute, reschedule to minimize delays

MaxSAT-based approach

- Encode time requirements as hard clauses, route/train cost as soft clauses
- Relax timings incrementally until feasible
- Add disruption(s), relax timings as needed



taken from [18]

Application Highlights: Takeaways

- SAT is an **essential and well-established tool** particularly for
  - software & hardware verification
  - electronic design automation
  - security and cryotography
- Indicators for a problem to best use SAT for?
  - large portion of “intrinsically Boolean” constraints
  - combinatorial search space / set of decisions, NP-hard (or harder)
  - problem description not too large
- Cross-application techniques and insights:
  - Often promising to hybridize SAT with direct (search) methods: use SAT to resolve difficult cores
  - Positive feedback loop between solver techniques and applications
  - Cross-fertilization between different applications (e.g., BMC, SMT)
  - More often than not, incremental and iterative approaches are needed

References I

[1] Toma Balyo. „Relaxing the relaxed exist-step parallel planning semantics“. In: *2013 IEEE 25th International Conference on Tools with Artificial Intelligence*. IEEE. 2013, S. 865–871.

[2] Bernhard Beckert u. a. „Modular verification of JML contracts using bounded model checking“. In: *Int. Symposium on Leveraging Applications of Formal Methods (ISoLA)*. Springer. 2020, S. 60–80.

[3] Armin Biere und Mathias Fleury. „Gimsatul, IsaSAT, Kissat Entering the SAT Competition 2022“. In: *SAT Competition*. http://hdl.handle.net/10138/359079. 2022, S. 10–11.

[4] Tom Bylander. „The computational complexity of propositional STRIPS planning“. In: *Artificial Intelligence* 69.1-2 (1994), S. 165–204. DOI: 10.1016/0004-3702(94)90081-7.

[5] Edmund Clarke u. a. „Bounded model checking using satisfiability solving“. In: *Formal methods in system design* 19 (2001), S. 7–34. DOI: 10.1023/A:1011276507260.

[6] Joseph Culberson. „Sokoban is PSPACE-complete“. In: *Technical report, Department of Computing Science, University of Alberta* (1997).

[7] Christoph Dobraunig u. a. „Cryptanalysis of ascon“. In: *Topics in Cryptology—CT-RSA 2015: The Cryptographer's Track at the RSA Conference 2015, San Francisco, CA, USA, April 20-24, 2015. Proceedings*. Springer. 2015, S. 371–387.

[8] Stephan Falke, Florian Merz und Carsten Sinz. „The bounded model checker LLBMC“. In: *2013 28th IEEE/ACM International Conference on Automated Software Engineering (ASE)*. IEEE. 2013, S. 706–709.

[9] Nils Froleyks, Tomáš Balyo und Dominik Schreiber. „PASAR—Planning as Satisfiability with Abstraction Refinement“. In: *Proc. SoCS*. Bd. 10. 1. 2019, S. 70–78. URL: https://ojs.aaai.org/index.php/SOCS/article/download/18504/18295.

[10] Yu Gao. „Kissat_MAB_prop in SAT Competition 2023“. In: *SAT Competition*. 2023, S. 16.

References II

[11] Stephan Gocht und Tomas Balyo. „Accelerating SAT Based Planning with Incremental SAT Solving“. In: *Proc. ICAPS*. 2017, S. 135–139. DOI: 10.1609/icaps.v27i1.13798.

[12] Malte Helmert. „The Fast Downward planning system“. In: *Journal of Artificial Intelligence Research* 26 (2006), S. 191–246.

[13] Henry Kautz und Bart Selman. „Pushing the envelope: Planning, propositional logic, and stochastic search“. In: *AAAI Conference on Artificial Intelligence*. 1996, S. 1194–1201.

[14] Henry A. Kautz und Bart Selman. „Planning as Satisfiability“. In: *Proc. ECAI*. Bd. 92. Citeseer. 1992, S. 359–363.

[15] Alexander Koch, Michael Schrempp und Michael Kirsten. „Card-based cryptography meets formal verification“. In: *New Generation Computing* 39.1 (2021), S. 115–158. DOI: 10.1007/s00354-020-00120-0.

[16] Andreas Kuehlmann. „Dynamic transition relation simplification for bounded property checking“. In: *IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004.* IEEE. 2004, S. 50–57.

[17] Frédéric Lafitte, Jorge Nakahara Jr und Dirk Van Heule. „Applications of SAT solvers in cryptanalysis: finding weak keys and preimages“. In: *Journal on Satisfiability, Boolean Modeling and Computation* 9.1 (2014), S. 1–25. DOI: 10.3233/sat190099.

[18] Alexandre Lemos u. a. „Iterative Train Scheduling under Disruption with Maximum Satisfiability“. In: *Journal of Artificial Intelligence Research* 79 (2024), S. 1047–1090.

[19] Norbert Manthey. „Testing the ASCON Hash Function“. In: *SAT Competition*. 2023, S. 63.

[20] João P. Marques-Silva und Karem A. Sakallah. „Boolean satisfiability in electronic design automation“. In: *Proc. Design Automation Conference*. 2000, S. 675–680. DOI: 10.1145/337292.337611.

References III

[21] Ilya Mironov und Lintao Zhang. „Applications of SAT solvers to cryptanalysis of hash functions“. In: *Theory and Applications of Satisfiability Testing-SAT 2006: 9th International Conference, Seattle, WA, USA, August 12-15, 2006. Proceedings 9*. Springer. 2006, S. 102–115.

[22] Nina Narodytska u. a. „Learning optimal decision trees with SAT“. In: *Int. Joint Conf.s on AI (IJCAI)*. 2018, S. 1362–1368. DOI: 10.24963/ijcai.2018/189.

[23] Raymond Reiter. „On closed world data bases“. In: *Readings in artificial intelligence*. Elsevier, 1981, S. 119–140. DOI: 10.1016/b978-0-934613-03-3.50014-3.

[24] Jussi Rintanen. „Evaluation strategies for planning as satisfiability“. In: *Proc. ECAI*. Bd. 16. 2004, S. 682.

[25] André Schidler und Stefan Szeider. „SAT-based decision tree learning for large data sets“. In: *AAAI Conference on Artificial Intelligence*. Bd. 35. 5. 2021, S. 3904–3912. DOI: 10.1609/aaai.v35i5.16509.

[26] Mate Soos, Karsten Nohl und Claude Castelluccia. „Extending SAT Solvers to Cryptographic Problems“. In: *Theory and Applications of Satisfiability Testing (SAT)*. Springer. 2009, S. 244–257. DOI: 10.1007/978-3-642-02777-2_24.

[27] Pavel Surynek u. a. „Migrating Techniques from Search-Based Multi-Agent Path Finding Solvers to SAT-Based Approach“. In: *JAIR* 73 (2022), S. 553–618. DOI: 10.1613/jair.1.13318.

[28] Yakir Vizel, Georg Weissenbacher und Sharad Malik. „Boolean Satisfiability Solvers and Their Applications in Model Checking“. In: *IEEE*. Bd. 103. 11. 2015, S. 2021–2035. DOI: 10.1109/JPROC.2015.2455034.

[29] Sean Weaver. *Equivalence Checking*. https://www21.in.tum.de/~lammich/2015_SS_Seminar_SAT/resources/Equivalence_Checking_11_30_08.pdf. 2015.

[30] Sean Weaver und Marijn J. H. Heule. „Constructing minimal perfect hash functions using SAT technology“. In: *AAAI Conference on Artificial Intelligence*. Bd. 34. 02. 2020, S. 1668–1675. DOI: 10.1609/aaai.v34i02.5529.

[31] Wenqian Xin u. a. „Improved cryptanalysis on SipHash“. In: *International Conference on Cryptology and Network Security*. Springer. 2019, S. 61–79.