# A Concurrency Problem with **Exponential DPLL**( $\mathcal{T}$ ) **Proofs**

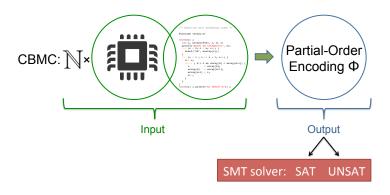
A Problem Harder Than Diamonds

Liana Hadarean<sup>1</sup> Alex Horn<sup>1</sup> Tim King<sup>2</sup>

<sup>1</sup>University of Oxford <sup>2</sup>Verimag

July 19, 2015

### **SAT/SMT-based Concurrency Verification**



[CAV '13] found bugs in:







### **A Concurrency Problem**

#### **Example**

Let the value at memory location x be initialized to 0, i.e. [x] = 0.

Thread $T_0$	Thread $T_1$		Thread $T_N$
	$   local v_1 := [x] $		local $v_N := [x]$
local $v_0 := [x]$ assert $(v_0 \le N)$	$[x] := v_1 + 1$	•••	$[x] := v_N + 1$

## **A Concurrency Problem**

#### **Example**

Let the value at memory location x be initialized to 0, i.e. [x] = 0.

Thread $T_0$	Thread $T_1$	Thread $T_N$
local $v_0 := [x]$ assert $(v_0 \le N)$	local $v_1 := [x]$   $[x] := v_1 + 1$	

#### Claim (Partial-Order DPLL( $\mathcal{T}$ ) Proof Complexity)

The size of  $DPLL(\mathcal{T})$  proofs<sup>\*</sup> for this problem is at least N! using partial-order encodings of concurrency<sup>\*\*</sup>.

#### **Our Contributions**

- The concept of non-interfering critical assignments;
- A proof complexity theorem for Fixed-Alphabet DPLL(T);
- A factorial-size lower bound for the concurrency problem;
- Experiments with multiple SMT solvers & theory combinations.

### Fixed-Alphabet DPLL( $\mathcal{T}$ ) Proofs

A simplified form of  $DPLL(\mathcal{T})$  with only two rules:

- Propositional resolution (Res);
- Learning  $\mathcal{T}$ -valid clauses over the literals of a fixed alphabet of  $\mathcal{T}$ -atoms ( $\mathcal{T}$ -LEARN).

### Fixed-Alphabet DPLL( $\mathcal{T}$ ) Proofs

A simplified form of  $DPLL(\mathcal{T})$  with only two rules:

- Propositional resolution (Res);
- Learning  $\mathcal{T}$ -valid clauses over the literals of a fixed alphabet of  $\mathcal{T}$ -atoms ( $\mathcal{T}$ -LEARN).

#### **Example (animation next)**

$$\phi \triangleq (x < y \lor x = y) \land y < x$$

 $\phi$  is a QF LIA-unsatisfiable CNF formula.

- Fix  $\mathcal{A} = \{x < y, x = y, y < x\}$  to be the alphabet of  $\mathcal{T}$ -atoms;
- let X = {A, B, C} be propositional variables;
- let  $-^{\mathbb{B}} : \mathcal{A} \to X$  be an injective function, e.g.  $(x < y)^{\mathbb{B}} = A$ .

Using  $\mathcal{A}$ , X and  $-^{\mathbb{B}}$ , we now illustrate Res and  $\mathcal{T}$ -LEARN.

$$\phi = (\mathbf{x} < \mathbf{y} \lor \mathbf{x} = \mathbf{y}) \land \mathbf{y} < \mathbf{x}$$

An execution of a lazy DPLL( $\mathcal{T}$ ) solver (with a fixed-alphabet):

$$\sim (\underbrace{A}_{A \leftarrow \top} \lor \underbrace{B}_{B \leftarrow \bot}) \land \underbrace{C}_{C \leftarrow \top}$$

(Bool Assignment)

$$\phi = (\mathbf{x} < \mathbf{y} \lor \mathbf{x} = \mathbf{y}) \land \mathbf{y} < \mathbf{x}$$

An execution of a lazy  $DPLL(\mathcal{T})$  solver (with a fixed-alphabet):

$$\phi = (\mathbf{x} < \mathbf{y} \lor \mathbf{x} = \mathbf{y}) \land \mathbf{y} < \mathbf{x}$$

An execution of a lazy DPLL( $\mathcal{T}$ ) solver (with a fixed-alphabet):

$$(A \lor B) \land C$$

$$A \leftarrow T \quad B \leftarrow \bot \quad C \leftarrow T$$

$$(Bool Assignment)$$

$$(T-LEARN)$$

$$(A \lor B) \land C \land (\neg A \lor \neg C)$$

$$(T-LEARN)$$

$$(A \lor B) \land C \land (\neg A \lor \neg C)$$

$$(Bool Assignment)$$

$$(Bool Assignment)$$

$$\phi = (\mathbf{x} < \mathbf{y} \lor \mathbf{x} = \mathbf{y}) \land \mathbf{y} < \mathbf{x}$$

An execution of a lazy DPLL( $\mathcal{T}$ ) solver (with a fixed-alphabet):

$$\phi = (\mathbf{x} < \mathbf{y} \lor \mathbf{x} = \mathbf{y}) \land \mathbf{y} < \mathbf{x}$$

An execution of a lazy  $DPLL(\mathcal{T})$  solver (with a fixed-alphabet):

$$(Bool Assignment)$$

$$A \leftarrow T \qquad B \leftarrow \bot \qquad C \leftarrow T$$

$$(A \lor B) \land C \land (\neg A \lor \neg C) \qquad (\mathcal{T}\text{-Lemma})$$

$$(\mathcal{T}\text{-Lemma})$$

$$A \leftarrow \bot \qquad B \leftarrow T \qquad C \leftarrow T \qquad T \qquad \bot$$

$$(A \lor B) \land C \land (\neg A \lor \neg C) \land (\neg B \lor \neg C) \qquad (\mathcal{T}\text{-Lemma})$$

$$(\mathcal{T}\text{-Lemma})$$

$$(\mathcal{T}\text{-Lemma})$$

 $\rightarrow \bot$  (i.e. formula is  $\mathcal{T}$ -unsatisfiable)

12

(Res, multiple steps)

$$\phi = (\mathbf{x} < \mathbf{y} \lor \mathbf{x} = \mathbf{y}) \land \mathbf{y} < \mathbf{x}$$

An execution of a lazy  $DPLL(\mathcal{T})$  solver (with a fixed-alphabet):

$$(Bool Assignment)$$

$$A \lor B \land C$$

$$A \lor B \land C \land (\neg A \lor \neg C)$$

$$\mathcal{T}\text{-Lemma}$$

$$A \lor B \land C \land (\neg A \lor \neg C)$$

$$A \lor B \land C \land (\neg A \lor \neg C)$$

$$A \lor B \land C \land (\neg A \lor \neg C)$$

$$A \lor B \land C \land (\neg A \lor \neg C) \land (\neg B \lor \neg C)$$

$$(\mathcal{T}\text{-Lemma})$$

$$(\mathcal{T}\text{-Lemma})$$

$$(\mathcal{T}\text{-Lemma})$$

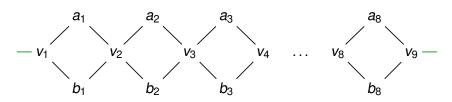
 $\rightarrow \bot$  (i.e. formula is  $\mathcal{T}$ -unsatisfiable)

13

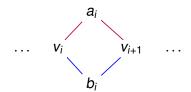
(Res, multiple steps)

## Known Challenges for DPLL( $\mathcal{T}$ ): Diamonds

Let 
$$\phi \diamond$$
 be  $\left( \bigwedge_{i=1}^{8} \left( \mathbf{v}_i = \mathbf{a}_i \wedge \mathbf{a}_i = \mathbf{v}_{i+1} \right) \vee \left( \mathbf{v}_i = \mathbf{b}_i \wedge \mathbf{b}_i = \mathbf{v}_{i+1} \right) \right) \wedge \mathbf{v}_1 \neq \mathbf{v}_9$ .



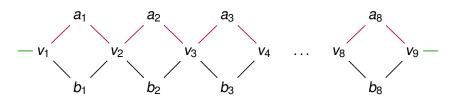
A diamond:



Let 
$$X = \{ /i, /i, /i, -i, -i : 1 \le i \le 8 \}$$
, e.g.  $/i$  denotes  $v_i = a_i$ .

# Known Challenges for DPLL( $\mathcal{T}$ ): Diamonds

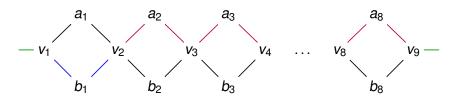
Let 
$$\phi_{\diamondsuit}$$
 be  $\left(\bigwedge_{i=1}^{8} \left(v_i = a_i \wedge a_i = v_{i+1}\right) \vee \left(v_i = b_i \wedge b_i = v_{i+1}\right)\right) \wedge v_1 \neq v_9$ .



- **1.** Let  $M_1 = \{ /i, /i : 1 \le i \le 8 \} \cup \{ -- \}.$
- **2.** Then  $M_1 \models \phi^{\mathbb{B}}_{\diamond}$ . But  $M_1$  leads to a (unique)  $\mathcal{T}$ -conflict.
- **3.** DPLL( $\mathcal{T}$ ) learns the  $\mathcal{T}$ -lemma  $\neg M_1$ .

# Known Challenges for DPLL( $\mathcal{T}$ ): Diamonds

Let 
$$\phi \diamond$$
 be  $\left( \bigwedge_{i=1}^{8} \left( v_i = a_i \wedge a_i = v_{i+1} \right) \vee \left( v_i = b_i \wedge b_i = v_{i+1} \right) \right) \wedge v_1 \neq v_9$ .



- **1.** Let  $M_2 = \{ \setminus_1, \setminus_1 \} \cup \{ \setminus_i, \setminus_i : 2 \le i \le 8 \} \cup \{ -- \}.$
- **2.** Then  $M_2 \models \phi_{\diamondsuit}^{\mathbb{B}}$ . But  $M_2$  leads to a (unique)  $\mathcal{T}$ -conflict.
- **3.** DPLL( $\mathcal{T}$ ) learns the  $\mathcal{T}$ -lemma  $\neg M_2$ .
- **4.** Note that  $\neg M_1$  is disjoint from  $\neg M_2$ .
- **5.** In general, DPLL( $\mathcal{T}$ ) enumerates  $2^8 \mathcal{T}$ -conflicts [LPAR '08].

### Our DPLL( $\mathcal{T}$ ) Proof Complexity Theorem

A strict generalization of the diamonds problem:

#### **Definition (Critical Assignments)**

An assignment M is *critical* if  $M \models \phi^{\mathbb{B}}$  and there is exactly one minimal  $\mathcal{T}$ -conflict  $\neg L$  such that  $\neg L^{\mathbb{B}} \subseteq M$ . Let Q be a set of critical assignments for  $\phi$ .

# Our DPLL( $\mathcal{T}$ ) Proof Complexity Theorem

A strict generalization of the diamonds problem:

#### **Definition (Critical Assignments)**

An assignment M is *critical* if  $M \models \phi^{\mathbb{B}}$  and there is exactly one minimal  $\mathcal{T}$ -conflict  $\neg L$  such that  $\neg L^{\mathbb{B}} \subseteq M$ . Let Q be a set of critical assignments for  $\phi$ .

#### **Definition (Non-interfering Critical Assignments)**

Q is non-interfering if, for all  $M_i \neq M_j$  in Q,  $\neg L_i^{\mathbb{B}}$  isn't a subset of  $M_j$ .

## Our DPLL( $\mathcal{T}$ ) Proof Complexity Theorem

A strict generalization of the diamonds problem:

#### **Definition (Critical Assignments)**

An assignment M is *critical* if  $M \models \phi^{\mathbb{B}}$  and there is exactly one minimal  $\mathcal{T}$ -conflict  $\neg L$  such that  $\neg L^{\mathbb{B}} \subseteq M$ . Let Q be a set of critical assignments for  $\phi$ .

#### **Definition (Non-interfering Critical Assignments)**

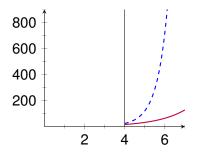
Q is non-interfering if, for all  $M_i \neq M_j$  in Q,  $\neg L_i^{\mathbb{B}}$  isn't a subset of  $M_j$ .

#### Theorem (DPLL( $\mathcal{T}$ ) Lower Bound Proof Complexity)

Let  $\phi$  be an unsatisfiable  $\mathcal T$ -formula, and Q be a non-interfering set of critical assignments for  $\phi$ . Every Fixed-Alphabet-DPLL( $\mathcal T$ ) proof that  $\phi$  is UNSAT contains at least |Q| applications of  $\mathcal T$ -LEARN.

# A New Challenge Problem for SMT Community

We use the previous theorem to establish the factorial-size lower bound proof complexity of our concurrency problem challenge.



 $---\Omega(N!)$  (Concurrency Problem)  $---\Omega(2^N)$  (Diamonds Problem)

## **SMT Encodings of Concurrency Problem**

Thread $T_0$	Thread $T_1$		Thread $T_N$
$r_0$	<i>r</i> <sub>1</sub>		$r_N$
	<i>W</i> <sub>1</sub>	•••	$w_N$

For N = 2, if restricted to  $T_1 \parallel T_2$ , we get the following interleavings:

- (1)  $r_1$ ;  $w_1$ ;  $r_2$ ;  $w_2$  (2)  $r_1$ ;  $r_2$ ;  $w_1$ ;  $w_2$  (3)  $r_1$ ;  $r_2$ ;  $w_2$ ;  $w_1$
- (4)  $r_2$ ;  $r_1$ ;  $w_1$ ;  $w_2$  (5)  $r_2$ ;  $r_1$ ;  $w_2$ ;  $w_1$  (6)  $r_2$ ;  $w_2$ ;  $r_1$ ;  $w_1$ .

Symbolically encode all interleavings, e.g. [CAV '13, FORTE '15].

### **SMT Encodings of Concurrency Problem**

Thread $T_0$	Thread $T_1$		Thread $T_N$
<i>r</i> <sub>0</sub>	r <sub>1</sub> w <sub>1</sub>		r <sub>N</sub> w <sub>N</sub>

Let  $R \triangleq \{r_0, \ldots, r_N\}$  and  $W \triangleq \{w_{init}, w_0, \ldots, w_N\}$ .

Our encodings are parameterized by three SMT theories:

•  $\mathcal{T}_C$ : clocks

• T<sub>S</sub>: selections

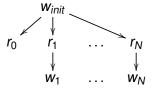
•  $\mathcal{T}_V$ : value

Using  $\mathcal{T}_C$ ,  $\mathcal{T}_S$  and  $\mathcal{T}_V$ , we encode partial-order axioms (see next).

#### **Clock Constraints**

#### Example

Preserved-program order (PPO) for  $T_0 \parallel T_1 \parallel ... \parallel T_N$ :

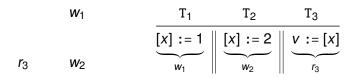


#### **Example**

By write consistency, writes in W are totally ordered in  $\mathcal{T}_C$ , e.g. either  $w_1 < w_2$  or  $w_2 < w_1$  in  $\mathcal{T}_C$ .

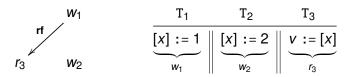
#### **Selection and Value Constraints**

#### **Example**



#### **Selection and Value Constraints**

#### Example



To encode that " $r_3$  reads from  $w_1$ ":

- $s_{r_3} = s_{w_1}$  in  $\mathcal{T}_S$
- $v_3 = 1$  in  $\mathcal{T}_V$

#### Intuition: Factorial Lower Bound for Proof Size

Shuffle threads, e.g. for  $T_1$ ,  $T_2$  and  $T_3$  we get:

$$T_1; T_2; T_3; T_0$$
  $(\pi_1)$ 

$$T_2; T_1; T_3; T_0$$
  $(\pi_2)$ 

$$T_2; T_3; T_1; T_0$$
  $(\pi_3)$ 

$$\dots$$
  $(\pi_k)$ 

Each shuffling is satisfiable in  $\mathcal{T}_C + \mathcal{T}_S$  but leads to a unique minimal  $\mathcal{T}_V$ -conflict:

$$v_1 = 0 \land v_2 = v_1 + 1 \land v_3 = v_2 + 1 \land v_{assert} = v_3 \land v_{assert} > N \quad (\pi_1)$$

$$v_2 = 0 \land v_1 = v_2 + 1 \land v_3 = v_1 + 1 \land v_{assert} = v_3 \land v_{assert} > N$$
 ( $\pi_2$ )

$$v_2 = 0 \land v_3 = v_2 + 1 \land v_1 = v_3 + 1 \land v_{assert} = v_3 \land v_{assert} > N$$
 ( $\pi_3$ )

$$\dots$$
  $(\pi_k)$ 

#### **Factorial Lower Bound for Proof Size**

Let  $\phi^3$  be a variant of the cubic-size encoding in [CAV '13] by our colleagues Alglave, Kroening and Tautschnig.

#### Theorem (Lower Bound for Cubic Partial-Order Encoding)

All Fixed-Alphabet-DPLL( $\mathcal{T}$ ) proofs for the problem challenge encoded with  $\phi^3$  contain at least N! applications of  $\mathcal{T}$ -LEARN.

We also studied a quadratic-size partial-order encoding [FORTE '15]. Here, we show that this asymptotically smaller encoding has also at least factorial-sized DPLL( $\mathcal{T}$ ) proofs!

# **Experiments with Two Partial-Order Encodings**

 $\mathcal{E}^3$  and  $\mathcal{E}^2$  are partial-order encodings of asymptotically different size, parameterized by three theories  $\mathcal{T}_C$ ,  $\mathcal{T}_S$  and  $\mathcal{T}_V$ .

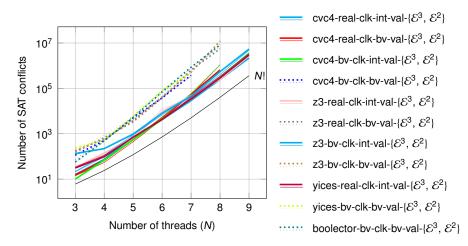
We instantiate  $\langle \mathcal{T}_C, \mathcal{T}_S, \mathcal{T}_V \rangle$  to four configurations:

- 1. "real-clk-int-val":  $\mathcal{T}_C = \mathcal{T}_S = \mathcal{T}_{\mathbb{R}}$  and  $\mathcal{T}_V = \mathcal{T}_{\mathbb{Z}}$
- 2. "bv-clk-int-val":  $\mathcal{T}_C = \mathcal{T}_S = \mathcal{T}_{\mathbb{BV}}$  and  $\mathcal{T}_V = \mathcal{T}_{\mathbb{Z}}$
- 3. "real-clk-bv-val":  $\mathcal{T}_C = \mathcal{T}_S = \mathcal{T}_\mathbb{R}$  and  $\mathcal{T}_V = \mathcal{T}_{\mathbb{BV}}$
- 4. "bv-clk-bv-val":  $\mathcal{T}_C = \mathcal{T}_S = \mathcal{T}_{\mathbb{BV}}$  and  $\mathcal{T}_V = \mathcal{T}_{\mathbb{BV}}$

We use the following SMT solvers: Boolector, CVC4, Yices2, Z3.

Example: "z3-bv-clk-int-val- $\mathcal{E}^2$ " denotes experiments with the  $O(N^2)$  encoding using Z3 where  $\mathcal{T}_C = \mathcal{T}_S = \mathcal{T}_{\mathbb{BV}}$  and  $\mathcal{T}_V = \mathcal{T}_{\mathbb{Z}}$ . We have a total of 56 SMT-LIB benchmarks. Timeout is 1 hour.

#### **Experimental Results**



Factorial growth of conflicts in fkp2013-unsat benchmark.

### **Concluding Remarks**

• A simple, yet challenging, SMT benchmark:

Thread $T_0$	Thread $T_1$	Thread $T_N$
$\overline{ \mathbf{ocal} \ v_0 := [x] }$		
$assert(v_0 \le N)$	$[x] := v_1 + 1$	$  [x]  := v_N + 1$

- A new diagnosis tool for SMT encodings:
  - **1.** Proof-size for  $DPLL(\mathcal{T})$  via non-interfering critical assignments
  - 2. N! lower bound for two state-of-the-art partial-order encodings
  - 3. Theory and experiments pinpoint value constraints as culprit



Morgan Deters

## **Cubic-size Encoding of Concurrency Problem**

Let  $\phi^3$  be the  $O(N^3)$  partial-order encoding of  $T_0 \parallel T_1 \parallel ... \parallel T_N$ :

$$C_{W_{init}} < C_{r_{assert}} \land \bigwedge_{i=1...N} C_{W_{init}} < C_{r_i} < C_{W_i} \land \bigwedge_{W_i,W' \in W,W \neq W'} (C_W < C_{W'} \lor C_{W'} < C_W) \land S_W \neq S_{W'} \land W_i \land$$

### **SC-relaxed Consistency Encoding**

Let *E* be the set of events,  $\ll$  be the PPO,  $val : E \rightarrow \mathcal{T}_V$ -terms,  $guard : E \rightarrow \mathcal{T}_V$ -formulas and *L* be the set of memory locations.

$$\begin{split} & \text{PPO} \triangleq \bigwedge \left\{ (guard(e) \land guard(e')) \Rightarrow (c_e < c_{e'}) \mid e, e' \in E \colon e \ll e' \right\} \\ & \text{WW}[x] \triangleq \bigwedge \left\{ (c_w < c_{w'} \lor c_{w'} < c_w) \land s_w \neq s_{w'} \mid w, w' \in W_x \land w \neq w' \right\} \\ & \text{RW}[x] \triangleq \bigwedge \left\{ (c_w < c_r \lor c_r < c_w \mid w \in W_x \land r \in R_x \right\} \\ & \text{RF}_{\text{TO}}[x] \triangleq \bigwedge \left\{ (guard(r) \Rightarrow \bigvee \left\{ s_w = s_r \mid w \in W_x \right\} \mid r \in R_x \right\} \\ & \text{RF}^3[x] \triangleq \bigwedge \left\{ (s_w = s_r) \Rightarrow (guard(w) \land val(w) = v_r \land c_w < c_r) \mid r \in R_x \land w \in W_x \right\} \\ & \text{FR}[x] \triangleq \bigwedge \left\{ (s_w = s_r \land c_w < c_{w'} \land guard(w')) \Rightarrow (c_r < c_{w'}) \mid w, w' \in W_x \land r \in R_x \right\} \\ & \mathcal{E}^3 \triangleq \bigwedge \left\{ \text{RF}_{\text{TO}}[x] \land \text{RF}^3[x] \land \text{FR}[x] \land \text{WW}[x] \land \text{RW}[x] \mid x \in L \right\} \land \text{PPO} \\ & \text{RF}^2[x] \triangleq \bigwedge \left\{ (s_w = s_r) \Rightarrow (c_w = \sup_r \land guard(w) \land val(w) = v_r \land c_w < c_r) \mid r \in R_x \land w \in W_x \right\} \\ & \mathcal{E}^2 \triangleq \bigwedge \left\{ \text{RF}_{\text{TO}}[x] \land \text{RF}^2[x] \land \text{SUP}[x] \land \text{WW}[x] \land \text{RW}[x] \mid x \in L \right\} \land \text{PPO} \end{split}$$