Presentation of

Solving SAT and SAT Modulo Theories: From an Abstract Davis–Putnam–Logemann–Loveland Procedure to DPLL(*T*)

ROBERT NIEUWENHUIS AND ALBERT OLIVERAS

Technical University of Catalonia, Barcelona, Spain

AND

CESARE TINELLI

The University of Iowa, Iowa City, Iowa

Abstract. We first introduce *Abstract DPLL*, a rule-based formulation of the Davis–Putnam– Logemann–Loveland (DPLL) procedure for propositional satisfiability. This abstract framework allows one to cleanly express practical DPLL algorithms and to formally reason about them in a simple way. Its properties, such as soundness, completeness or termination, immediately carry over to the modern DPLL implementations with features such as backjumping or clause learning.

We then extend the framework to Satisfiability Modulo background Theories (SMT) and use it to model several variants of the so-called *lazy approach* for SMT. In particular, we use it to introduce a few variants of a new, efficient and modular approach for SMT based on a general DPLL(X) engine, whose parameter X can be instantiated with a specialized solver *Solver*_T for a given theory T, thus producing a DPLL(T) system. We describe the high-level design of DPLL(X) and its cooperation with *Solver*_T, discuss the role of *theory propagation*, and describe different DPLL(T) strategies for some theories arising in industrial applications.

Our extensive experimental evidence, summarized in this article, shows that DPLL(T) systems can significantly outperform the other state-of-the-art tools, frequently even in orders of magnitude, and have better scaling properties.

Categories and Subject Descriptors: B.6.3 [Logic Design]: Design Aids—Verification; F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic—Computational logic; verification; I.2.3 [Artificial Intelligence]: Deduction and Theorem Proving—Deduction (e.g., natural, rule-based)

General Terms: Theory, Verification

Additional Key Words and Phrases: SAT solvers, Satisfiability Modulo Theories

by Ramūnas Gutkovas

SMT Reading Group Uppsala 2012-10-31

Background for Abstract DPLL

Modern **SAT** (~2006) solvers based on **DPLL** use many **extensions** and combinations thereof

Original **DPLL** [Davis et al. '62]

Problem: these extensions lack formal treatment

Background for Abstract DPLL

Modern **SAT** (~2006) solvers based on **DPLL** use many **extensions** and combinations thereof



Problem: these extensions lack formal treatment

Abstract DPLL

Solution: Abstract DPLL

[NOT '06]

66 ... a uniform, declarative framework for describing DPLLbased solvers, both for propositional satisfiability and for satisfiability modulo theories.

(emphasis is mine)

Abstract DPLL Notions and Classical DPLL

Formal Preliminaries (1/3) Propositional Case

(in examples this set will be identified with natural numbers) l is a literal whenever l is p or $\neg p$

 $\neg l = \mathrm{if}\ l = p\ \mathrm{then}\ \neg p\ \mathrm{else}\ p$ negation op. on literal

 $C = l_1 \lor \cdots \lor l_n$ clause is a set of literals

 $F=C_1,\ldots,C_m$ formula is a set of clauses

F is in ${f CNF}$

()/3) Formal Preliminaries Propositional Case

 $M = l_1 l_2 \dots l_n$ assignment is a consistent sequence of literals M is consistent if there is no p s.t. $eg p \in M$ and $p \in M$ l is defined in M if $l \in M$ or $eg l \in M$ $M \vDash C$ if there is $l \in C$ s.t. $l \in M$ conflict $M \models \neg C$ if for every $l \in C$ implies $\neg l \in M$ $M \vDash F$ if for every $C \in F$ implies $M \vDash C$ $\left\{\begin{array}{l} M \text{ is a model of } F \\ F \text{ is satisfiable} \end{array}\right\} \text{whenever for some } M \quad M \vDash F$ $F \vDash F'$ if for every $M \vDash F$ implies $M \vDash F'$

p atom

l literal

 $C = l_1 \vee \cdots \vee l_n$

 $F = C_1, \ldots, C_m$

$$\begin{array}{l} \textbf{Abstract DPLL System} \\ \textbf{Transition System} & \left(\begin{array}{c} p \text{ atom} \\ l \text{ literal} \\ G = l_1 \lor \cdots \lor l_n \text{ clause} \\ F = C_1, \ldots, C_m \text{ care} \\ M = l_1 l_2 \ldots l_n \text{ model} \end{array} \right) \\ \textbf{Transition system} & \left(\Gamma, \Longrightarrow\right) \text{ where } \Longrightarrow \in \Gamma \times \Gamma \\ \textbf{State } S \in \Gamma \text{ in case of ADPLL is one of} \\ 1 \cdot FailState \\ 2 \cdot M \parallel F \\ \textbf{Transition } S \Longrightarrow S' \\ \textbf{Final state } S \text{ wrt } \Longrightarrow \text{ whenever } S \nleftrightarrow \\ \textbf{Reflexive-transitive closure } \Rightarrow^* \text{ of } \Longrightarrow \\ \textbf{Decision literal } l^d \text{ in } M = N l^d N' \end{array}$$

Classical DPLL system

$$i f = 0$$

$$i f = 0$$

$$M \parallel F, C \lor I \implies M I \parallel F, C \lor I \text{ if } \begin{cases} M \models \neg C \\ I \text{ is undefined in } M. \end{cases}$$

$$P = 0$$

$$M \parallel F \implies M I \parallel F \qquad \text{if } \begin{cases} I \text{ occurs in some clause of } F \\ \neg I \text{ occurs in no clause of } F \\ I \text{ is undefined in } M. \end{cases}$$

$$P = 0$$

$$M \parallel F \implies M I \parallel F \qquad \text{if } \begin{cases} I \text{ or } \neg I \text{ occurs in a clause of } F \\ I \text{ is undefined in } M. \end{cases}$$

$$P = 0$$

$$M \parallel F \implies M I^{d} \parallel F \qquad \text{if } \begin{cases} I \text{ or } \neg I \text{ occurs in a clause of } F \\ I \text{ is undefined in } M. \end{cases}$$

$$P = 0$$

$$M \parallel F \implies M I^{d} \parallel F \qquad \text{if } \begin{cases} I \text{ or } \neg I \text{ occurs in a clause of } F \\ I \text{ is undefined in } M. \end{cases}$$

$$P = 0$$

$$M \parallel F, C \implies FailState \qquad \text{if } \begin{cases} M \models \neg C \\ M \text{ contains no decision literals.} \end{cases}$$

$$P = 0$$

$$M \parallel F, C \implies M \neg I \parallel F, C \qquad \text{if } \begin{cases} M I^{d} N \models \neg C \\ N \text{ contains no decision literals.} \end{cases}$$



















$$\begin{array}{c} M \\ & \downarrow \\ &$$

Extensions TO DPLLAbstract DPLL + Backjump + Learning + Restart

(Informal) *Example* 2.3.

- \Longrightarrow_{B} (Decide)
- \Longrightarrow_B (UnitPropagate)
- $\Longrightarrow_B (\text{Decide})$
- \Longrightarrow_B (UnitPropagate)
- \Longrightarrow_B (Decide)
- \Longrightarrow_{B} (UnitPropagate)

(Informal) *Example* 2.3.

- \Longrightarrow_{B} (Decide)
- \Longrightarrow_B (UnitPropagate)
- $\Longrightarrow_B (\text{Decide})$
- \Longrightarrow_{B} (UnitPropagate)
- \Longrightarrow_B (Decide)
- \Longrightarrow_{B} (UnitPropagate)









(Informal) *Example* 2.3.



- \Longrightarrow_{B} (Decide)
- \implies_{B} (UnitPropagate)
- \implies_{B} (Decide)

 $1^{d}2\overline{5} \parallel \overline{1} \lor 2, \quad \overline{3} \lor 4, \quad \overline{5} \lor \overline{6}, \quad 6 \lor \overline{5} \lor \overline{2}$

(Informal) *Example* 2.3.



5 and 1 are incompatible or consequently 5 and 2 are incompatible we learn that $\overline{5 \land 2} = \overline{5} \lor \overline{2}$ \longrightarrow_{B} (Decide) \implies_{B} (UnitPropagate)

 $1^{d}2\overline{5} \parallel \overline{1} \lor 2, \quad \overline{3} \lor 4, \quad \overline{5} \lor \overline{6}, \quad 6 \lor \overline{5} \lor \overline{2}$
$$\begin{array}{c} \textbf{Basic DPLL system} \Longrightarrow_{B} \\ \textbf{Conflict Driven Backtracking} \\ \textbf{Definition 2.4.} \\ \textbf{Definition 2.4.} \\ \textbf{Date Propagate + Decide + Fail} \\ \textbf{Backjump:} \\ \textbf{M} \ l^{d} \ N \parallel F, \ C \ \Longrightarrow \ M \ l' \parallel F, \ C \ \textbf{if} \\ \begin{cases} M \ l^{d} \ N \models \neg C, \ \text{and there is} \\ \text{some clause } C' \lor l' \ \text{such that:} \\ F, \ C \models C' \lor l' \ \text{and } M \models \neg C', \\ l' \ \text{is undefined in } M, \ \text{and} \\ l' \ or \neg l' \ occurs \ in \ F \ or \ in \ M \ l^{d} \ N. \end{cases} \\ \end{array}$$

 ${\cal C}$ conflict clause

 $C' \lor l'$ backjump clause unit clause wrt Msatisfiable under the same models as F,C

Backtracks further than backtrack





M $\overbrace{M_0 l_1^{\mathrm{d}} M_1 \dots l_n^{\mathrm{d}}}^{I} \overbrace{M_n}^{I} || F, C$

$$\begin{array}{l} \textbf{Basic DPLL system} \Longrightarrow_{B} \\ \textbf{P atom} \\ \textbf{l iteral} \\ C = l_{1} \lor \cdots \lor l_{n} \text{ clause} \\ F = C_{1}, \ldots, C_{m} \text{ cNF} \\ M = l_{1}l_{2} \ldots l_{n} \text{ model} \end{array}$$
$$M \ l^{d} \ N \parallel F, \ C \implies M \ l' \parallel F, \ C \ \text{if} \qquad \begin{cases} M \ l^{d} \ N \models \neg C, \text{ and there is} \\ \text{some clause} \ C' \lor l' \text{ such that:} \\ F, \ C \models C' \lor l' \text{ and } M \models \neg C', \\ l' \text{ is undefined in } M, \text{ and} \\ l' \text{ or } \neg l' \text{ occurs in } F \text{ or in } M \ l^{d} \ N. \end{cases}$$

$$\overbrace{M_0 l_1^{\mathrm{d}} M_1 \dots l_n^{\mathrm{d}} M_n}^{M} || F, C$$

Take
$$C' \vee l' = \neg l_1 \vee \cdots \vee \neg l_n$$

$$\begin{array}{l} \textbf{Basic DPLL system} \Longrightarrow_{B} \\ \textbf{P atom} \\ \textbf{l iteral Conflict Driven Backtracking} \\ \textbf{C} = l_{1} \lor \cdots \lor l_{n} \text{ clause} \\ F = C_{1}, \ldots, C_{m} \text{ cNF} \\ M = l_{1}l_{2} \ldots l_{n} \text{ model} \end{array}$$
$$M \ l^{d} N \parallel F, C \implies M \ l' \parallel F, C \ \text{if} \begin{cases} M \ l^{d} N \models \neg C, \text{ and there is} \\ \text{some clause } C' \lor l' \text{ such that:} \\ F, C \models C' \lor l' \text{ and } M \models \neg C', \\ l' \text{ is undefined in } M, \text{ and} \\ l' \text{ or } \neg l' \text{ occurs in } F \text{ or in } M \ l^{d} N. \end{cases}$$

$$\underbrace{M}_{M_0 l_1^{\mathrm{d}} M_1 \dots l_n^{\mathrm{d}}} \bigwedge_{M_n}^N || F, C$$

Take $C' \lor l' = \neg l_1 \lor \cdots \lor \neg l_n$ Have $M \vDash \neg (\neg l_1 \lor \cdots \lor \neg l_{n-1}) \iff M \vDash l_1, \ldots, l_{n-1}$

$$\begin{array}{c} \textbf{Pasic DPLL system} \Longrightarrow_{B} \\ \textbf{P atom} \\ \textbf{l literal Conflict Driven Backtracking} \\ C = l_{1} \lor \cdots \lor l_{n} \text{ clause} \\ F = C_{1}, \ldots, C_{m \text{ CNF}} \\ M = l_{1}l_{2} \ldots l_{n} \text{ model} \end{array}$$
$$M \ l^{d} N \parallel F, C \implies M \ l' \parallel F, C \text{ if } \begin{cases} M \ l^{d} N \models \neg C, \text{ and there is} \\ \text{some clause } C' \lor l' \text{ such that:} \\ F, C \models C' \lor l' \text{ and } M \models \neg C', \\ l' \text{ is undefined in } M, \text{ and} \\ l' \text{ or } \neg l' \text{ occurs in } F \text{ or in } M \ l^{d} N. \end{cases}$$

$$\begin{split} & \overbrace{M_0 l_1^{\mathrm{d}} M_1 \dots l_n^{\mathrm{d}}}^M \overbrace{M_n}^N || F, C \\ & \text{Take } C' \lor l' = \neg l_1 \lor \dots \lor \neg l_n \\ & \text{Have } M \vDash \neg (\neg l_1 \lor \dots \lor \neg l_{n-1}) \iff M \vDash l_1, \dots, l_{n-1} \\ & \text{And } F, C, l_1, \dots, l_n \text{ is unsat } \iff F, C \vDash \neg l_1 \lor \dots \lor \neg l_n \end{split}$$





Backjump is a backtracking mechanism.

LEMMA 2.8. Assume that $\emptyset \parallel F \Longrightarrow_{L}^{*} M \parallel F'$ and that $M \models \neg C$ for some clause *C* in *F'*. Then either Fail or Backjump applies to $M \parallel F'$.

This follows by a bit more generalized construction of the presented one.

 $M_0 l_1^{\mathrm{d}} M_1 \dots l_n^{\mathrm{d}} M_n \parallel F, C \implies M_0 l_1^{\mathrm{d}} M_1 \dots \neg l_n \parallel F, C$ Take $C' \lor l' = \neg l_1 \lor \dots \lor \neg l_n$ Have $M \vDash \neg (\neg l_1 \lor \dots \lor \neg l_{n-1}) \iff M \vDash l_1, \dots, l_{n-1}$ And F, C, l_1, \dots, l_n is **unsat** $\iff F, C \vDash \neg l_1 \lor \dots \lor \neg l_n$

S



THEOREM 2.13. If $\emptyset \parallel F \Longrightarrow_{B}^{*} S$ where S is final with respect to Basic DPLL, then

(1) S is FailState if, and only if, F is unsatisfiable.

(2) If S is of the form $M \parallel F'$, then M is a model of F. As F' = F



THEOREM 2.10. There are no infinite derivations of the form $\emptyset \parallel F \implies_{B} S_1 \implies_{B} \cdots$.



THEOREM 2.10. There are no infinite derivations of the form $\emptyset \parallel F \implies_{B} S_1 \implies_{B} \cdots$.





 $l \mathcal{I}$

S'

final



THEOREM 2.10. There are no infinite derivations of the form $\emptyset \parallel F \implies_{B} S_{1} \implies_{B} \cdots$.

Search Progress

progress at decision level i = $M_0 l_1^d M_1 l_2 \dots l_n^d M_n || F$ number of literals at dl i

For $M||F \Longrightarrow M'||F'$ M'||F' is more **progressed** than M||Fif there is a **least** more progressed decision level, or the **model** is more progressed.



level, or the model is more progressed.

Conflict Driven Clause Learning Example (Conflict Graph)



Conflict Driven Clause Learning Example (Conflict Graph)



Conflict Driven Clause Learning Example (Conflict Graph)



We can learn $\overline{6} \lor 8 \lor 7$

Conflict Driven Clause Learning Example (Backward Conflict Resolution)



Conflict Driven Clause Learning Example (Backward Conflict Resolution)



We can learn $\overline{6} \lor 8 \lor 7$

DPLL System With Learning \Longrightarrow_L

Definition 2.5.

Basic DPLL +

Learn:

 $M \parallel F \implies M \parallel F, C \text{ if } \begin{cases} \text{each atom of } C \text{ occurs in } F \text{ or in } M \\ F \models C. \end{cases}$

Forget:

 $M \parallel F, C \implies M \parallel F \quad \text{if } \{ F \models C.$

DPLL System With Learning \Longrightarrow_L

Learn:

 $M \parallel F \implies M \parallel F, C \text{ if } \begin{cases} \text{each atom of } C \text{ occurs in } F \text{ or in } M \\ F \models C. \end{cases}$

Forget:

 $M \parallel F, C \implies M \parallel F \quad \text{if } \{ F \models C.$

Sound and Complete

THEOREM 2.12. If $\emptyset \parallel F \Longrightarrow_{L}^{*} S$ where S is final with respect to Basic DPLL, then

(1) *S* is FailState if, and only if, *F* is unsatisfiable. (2) If *S* is of the form $M \parallel F'$ then *M* is a model of *F*.

DPLL System With Learning \Longrightarrow_L

Learn:

$$M \parallel F \implies M \parallel F, C \text{ if } \begin{cases} \text{each atom of } C \text{ occurs in } F \text{ or in } M \\ F \models C. \end{cases}$$

Forget:

 $M \parallel F, C \implies M \parallel F \quad \text{if } \{ F \models C.$

Sound and Complete

THEOREM 2.12. If $\emptyset \parallel F \Longrightarrow_{L}^{*} S$ where S is final with respect to Basic DPLL, then

(1) S is FailState if, and only if, F is unsatisfiable.

(2) If S is of the form $M \parallel F'$ then M is a model of F.

Decidable

THEOREM 2.11. Every derivation $\emptyset \parallel F \Longrightarrow_L S_1 \Longrightarrow_L \cdots$ by the DPLL system with Learning is finite if it contains no infinite subderivations consisting of only Learn and Forget steps.



Definition 2.14. The Restart rule is:

$$M \parallel F \implies \emptyset \parallel F.$$



Definition 2.14. The Restart rule is:

$$M \parallel F \implies \emptyset \parallel F.$$

Definition 2.15.



increased periodicity if $\ m < n$

Restarts

Definition 2.14. The Restart rule is:

$$M \parallel F \implies \emptyset \parallel F.$$

Definition 2.15.



where m, n number of applications of Basic DPLL

increased periodicity if $\ m < n$

Abstract DPLL + Restart terminates when considering paths with increasing periodicity of Restart

THEOREM 2.16. Any derivation $\emptyset \parallel F \implies S_1 \implies \cdots$ by the transition system *L* extended with the Restart rule is finite if it contains no infinite subderivations consisting of only Learn and Forget steps, and Restart has increasing periodicity in it.

Abstract DPLL Modulo Theories

First-Order Case

F ground quantifier free formula in CNF

T theory is a set of closed first-order formulas

F T-satisfiable (T-consistent) if $F \wedge T$ is first-order satisfiable

$$M$$
 is T-model of F if $M \models F$ propositional

entails $F \models_T F'$ if $F \wedge \neg F'$ T-inconsistent











Lazy SMT based on DPLL

Incremental T-solver

Checks Theory Consistency every time a literal is added to the assignment.

On-line SAT Solver

Whenever an **assignment** is **Theory Inconsistent, backtracks** to the point where the theory is still **consistent**.

(Exhaustive) Theory propagation

The assignment may be extended with the literals which are entailed by the theory.

Abstract DPLL modulo Theories



Modeling Lazy SMT with Abstract DPLL Naive Lazy Approach

 $\begin{array}{ll} M \mid\mid F & \text{is final wrt Decide, Fail, UnitPropagate, T-Backjump} \\ \text{If } M & \text{is T-incosistent} \\ & \text{There is } \{l_1, \ldots, l_n\} \subseteq M \text{ s.t. } \emptyset \vDash_T \neg l_1 \lor \cdots \lor l_n \\ & M \mid\mid F \implies_{T-\text{learn}} M \mid\mid F, \neg l_1 \lor \cdots \lor \neg l_n \\ & \implies_{\text{restart}} \emptyset \mid\mid F, \neg l_1 \lor \cdots \lor \neg l_n \end{array}$

Incremental

Same as above, except it applies for any **T-inconsistent** state.

Incremental + Online

Does not restart, but uses the fact that learned lemma clause is conflicting to apply T-backjump.

Basic and Full DPLL Modulo Theories system

Definition 3.4. The *Basic DPLL Modulo Theories system* consists of the rules Decide, Fail, UnitPropagate, TheoryPropagate, and *T*-Backjump.

Definition 3.5. The *Full DPLL Modulo Theories system*, denoted by FT, consists of the rules of Basic DPLL Modulo Theories and the rules *T*-Learn, *T*-Forget, and Restart.

Sound and Complete

THEOREM 3.10. Let Der be a derivation $\emptyset \parallel F \implies_{FT} S$, where (i) S is final with respect to Basic DPLL Modulo Theories, and (ii) if S is of the form $M \parallel F'$ then M is T-consistent. Then

(1) S is FailState if, and only if, F is T-unsatisfiable.

(2) If S is of the form $M \parallel F'$, then M is a T-model of F.



THEOREM 3.7 (TERMINATION). Let Der be a derivation of the form: $\emptyset \parallel F = S_0 \Longrightarrow_{FT} S_1 \Longrightarrow_{FT} \cdots$ Then Der is finite if the following two conditions hold:

 Der has no infinite subderivations consisting of only T-Learn and T-Forget steps.
For every subderivation of Der of the form: S_{i-1} ⇒_{FT} S_i ⇒_{FT} ··· ⇒_{FT} S_j ⇒_{FT} ··· ⇒_{FT} S_k where the only three Restart steps are the ones producing S_i, S_j, and S_k, either: -there are more Basic DPLL Modulo Theories steps in S_j ⇒_{FT} ··· ⇒_{FT} S_k than in S_i ⇒_{FT} ··· ⇒_{FT} S_j, or -a clause is learned² in S_j ⇒_{FT} ··· ⇒_{FT} S_k that is not forgotten in Der.

Increasing Periodicity

Do not revisit failed search
DPLL (T)

Abstract DPLL Engine

DPLL(T) Solver

DPLL(X) engine parameterized by a theory solver

 ${\small Solver_T}$ theory solver for conjunction of formulas

DPLL(T) Solver

DPLL(X) engine parameterized by a theory solver

Solver_T theory solver for conjunction of formulas

 $Solver_T + DPLL(X) = DPLL(T)$

DPLL(T) SMT solver

DPLL(T) Architecture

Interface of Solver_T

-]: Literal markLitTrue: Literal ---- unit M: Assignment unmarkLastLits: $Int \longrightarrow$ unit **isConsistent:** Assignment \longrightarrow Strength \longrightarrow **bool explainInCons:** Assignment \longrightarrow Literal set Finds and returns $\{l_1, \ldots, l_n\} \subseteq M$ s.t. $\emptyset \models_T \neg l_1 \lor \cdots \lor \neg l_n$ **explainTProp:** Assignment \longrightarrow Literal \longrightarrow Literal **set** Finds and returns $\{l_1,\ldots,l_n\} \subset M$ for a given literal s.t. $l_1,\ldots,l_n \models_T l$
 - entails: Assignment \longrightarrow Literal set \longrightarrow Literal set Returns { $l \mid M \vDash_T l \ \& \ l \in L$ }

Quest

lons?