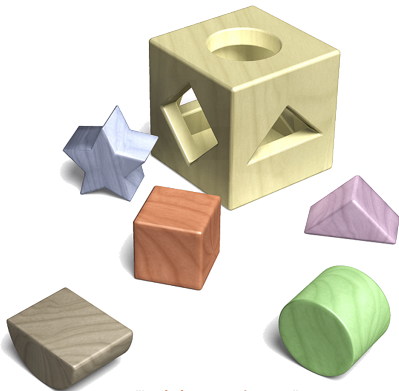


# SAT Solving – Systematic Search

**Steffen Hölldobler**

International Center for Computational Logic  
Technische Universität Dresden  
Germany

- ▶ **Truth Tables**
- ▶ **Semantic Trees**
- ▶ **DPLL**
- ▶ **DPLL-NB**
- ▶ **DPLL-CDBL**
- ▶ **GenericCDCL**



*"Logic is everywhere ..."*



## Truth Tables

- ▶ How can we compute the value of a formula  $F$  under all possible interpretations?
- ▶ Computing a truth table
  - 1 Let  $m = |\mathcal{S}(F)|$  be the number of subformulae of  $F$
  - 2 Let  $\mathcal{R}_F = \{A \mid A \in \mathcal{R} \text{ and } A \in \mathcal{S}(F)\}$  and  $n = |\mathcal{R}_F|$  be the number of propositional variables occurring in  $F$
  - 3 Form a table  $\text{TT}(F)$  with  $2^n$  rows and  $m$  columns, where the first  $n$  columns are marked by the elements of  $\mathcal{R}_F$ , the last column is marked by  $F$ , and the remaining columns are marked by the other subformulas of  $F$
  - 4 Fill in the first  $n$  columns with  $\top$  and  $\perp$  as follows:  
 In the first column fill in alternating downwards  $\top \perp \top \perp \dots$ ,  
 in the second column  $\top \top \perp \perp \dots$ ,  
 in the third column  $\top \top \top \perp \perp \perp \dots$ , etc.
  - 5 Calculate the values in the remaining columns using the known functions on the set of truth values



## Some Details

- ▶ For a row  $\zeta$  in  $\text{TT}(F)$  we denote by  $\zeta(G)$  the truth value in the column marked by the formula  $G \in \mathcal{S}(F)$
- ▶ With this, step 5 can be reformulated as follows

5 For each row  $\zeta$  in  $\text{TT}(F)$  and for all  $F \circ G, \neg F \in \mathcal{S}(F) \setminus \mathcal{R}_F$  calculate:

$$\zeta(F \circ G) = \zeta(F) \circ^* \zeta(G) \text{ and } \zeta(\neg F) = \neg^* \zeta(F)$$



## Some Observations

- ▶ Let  $I$  be an interpretation,  $F$  a formula,  
 $\mathcal{R}_F$  the set of variables occurring in  $F$  and  $n = |\mathcal{R}_F|$
- ▶ If we fix  $A^I$  for all  $A \in \mathcal{R}_F$ , then  $F^I$  is uniquely determined
- ▶ There are exactly  $2^n$  different possibilities of assigning truth values to  $\mathcal{R}_F$
- ▶ Each row in  $\text{TT}(F)$  corresponds precisely to one of these possibilities
- ▶ For each interpretation  $I$  of the language  $\mathcal{L}(\mathcal{R})$   
exists exactly one row  $\zeta_I$  in  $\text{TT}(F)$  with  $G^I = \zeta_I(G)$  for all  $G \in \mathcal{S}(F)$
- ▶ For every row  $\zeta$  in  $\text{TT}(F)$  exists an interpretation  $I$  of the language  $\mathcal{L}(\mathcal{R})$   
with  $G^I = \zeta(G)$  for all  $G \in \mathcal{S}(F)$ 
  - ▷  $I$  is not uniquely determined



## Determining Satisfiability using Truth Tables

- ▶  $F$  is satisfiable iff  $\text{TT}(F)$  contains a row  $\zeta$  with  $\zeta(F) = \top$
- ▶  $F$  is unsatisfiable iff for all rows  $\zeta$  in  $\text{TT}(F)$  we find  $\zeta(F) = \perp$
- ▶  $F$  is valid iff for all rows  $\zeta$  in  $\text{TT}(F)$  we find  $\zeta(F) = \top$
- ▶  $F$  is falsifiable iff  $\text{TT}(F)$  contains a row  $\zeta$  with  $\zeta(F) = \perp$



## Semantic Trees – Main Characteristics

- ▶ Optimization of the truth table method
- ▶ Stepwise partitioning of interpretations (through branching)
- ▶ Usually conceived for formulas in clausal form
- ▶ **Notation**

In the sequel nodes are (labeled by) expressions of the form  $F :: J$ , where

- ▷  $F$  is a formula and
- ▷  $J$  is either a partial interpretation for  $F$ , SAT, or CONFLICT

with the following informal meaning:

- ▷ If  $J$  is a partial interpretation, then  $F|_J$  is the “remaining” SAT-problem
- ▷ If  $J$  is SAT or CONFLICT, then  $F|_J$  is undefined
- ▷  $F :: J$  has successor  $F :: \text{SAT}$  iff  $J \models F$  iff  $F|_J = \langle \rangle$
- ▷  $F :: J$  has successor  $F :: \text{CONFLICT}$  iff  $J \not\models F$  iff  $[\ ] \in F|_J$



## Semantic Trees

- ▶ A **semantic tree** for a CNF-formula  $F$  is a binary tree satisfying the following conditions
  - ▷ The root node is  $F :: ()$
  - ▷ If  $F :: J$  is a node and  $F|_J = \langle \rangle$  then it has a successor node  $F :: \text{SAT}$
  - ▷ If  $F :: J$  is a node and  $[] \in F|_J$  then it has a successor node  $F :: \text{CONFLICT}$
  - ▷ If  $F :: J$  is a node,  $[] \notin F|_J$  and  $A \in \text{atoms}(F|_J)$  then  $F :: J$  has two successor nodes  $F :: J, \dot{A}$  and  $F :: J, \bar{A}$
- ▶ **Note**
  - ▷ The conditions in the three if-statements are mutually exclusive; the corresponding rules (see next slide) do not **overlap**
  - ▷ Instead of  $F :: J$  we could label the nodes with  $F|_J$
  - ▷ From an implementational point of view it is beneficial to separate  $F$  and  $J$   
**Why?**



## The ST Calculus

- ▶ Given a CNF-formula  $F$
- ▶ Computations are initialized by  $F :: ()$
- ▶ The rules of the calculus are

$$\begin{array}{llll}
 F :: J & \rightsquigarrow_{SAT} & F :: SAT & \text{iff } F|_J = \langle \rangle \\
 F :: J & \rightsquigarrow_{CONF} & F :: CONFLICT & \text{iff } [] \in F|_J \\
 F :: J & \rightsquigarrow_{SPLIT} & F :: J, \dot{A} \mid F :: J, \bar{\dot{A}} & \text{iff } A \in \text{atoms}(F|_J) \text{ and } [] \notin F|_J
 \end{array}$$

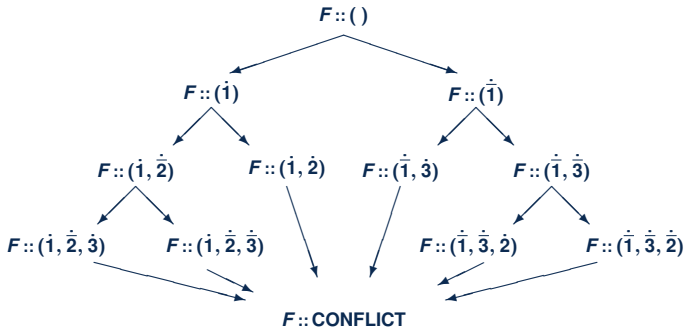
- ▶ SPLIT leads to branching
- ▶ Computation terminates if
  - ▷ a node  $F :: SAT$  is reached in which case  $F$  is satisfiable or
  - ▷ all leaf nodes are of the form  $F :: CONFLICT$  in which case  $F$  is unsatisfiable
- ▶  $F :: J \rightsquigarrow F' :: J'$   
 iff  $F :: J \rightsquigarrow_{SAT} F' :: J'$  or  $F :: J \rightsquigarrow_{CONF} F' :: J'$  or  $F :: J \rightsquigarrow_{SPLIT} F' :: J'$
- ▶  $\rightsquigarrow^*$  is the reflexive and transitive closure of  $\rightsquigarrow$





## Example

- Let  $F = \langle [2, \bar{3}], [2, 3], [\bar{1}, \bar{2}], [1, \bar{3}], [1, \bar{2}, 3] \rangle$  in

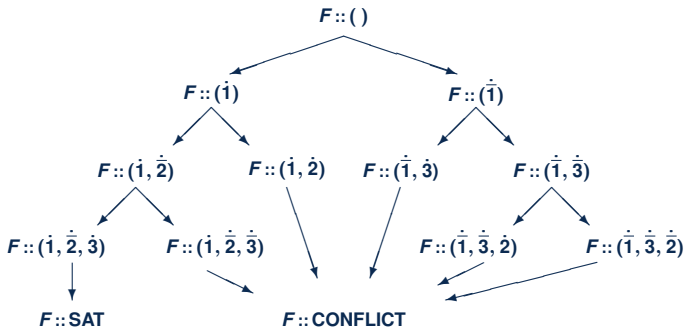


- $F$  is unsatisfiable



## Another Example

- Let  $F = \langle [2, 3], [\bar{1}, \bar{2}], [1, \bar{3}], [1, \bar{2}, 3] \rangle$  in



- $(\dot{1}, \dot{\bar{2}}, \dot{3}) \models F$



## Abstract Reduction Systems

▶ **The ST calculus is an abstract reduction system**

Baader, Nipkow: Term Rewriting and All That. Cambridge University Press: 1998

▶ **An abstract reduction system  $(\mathcal{R}, \rightarrow)$  is said to be**

**terminating** iff there is no infinite descending chain  $t_0 \rightarrow t_1 \rightarrow \dots$

**confluent** iff  $t_1 \leftarrow^* t \rightarrow^* t_2$  implies  $(\exists t') t_1 \rightarrow^* t' \leftarrow^* t_2$

**locally confluent** iff  $t_1 \leftarrow t \rightarrow t_2$  implies  $(\exists t') t_1 \rightarrow^* t' \leftarrow^* t_2$

**canonical** iff is terminating and confluent

▶ **Newman's Lemma** A terminating relation is confluent if it is locally confluent

Newman: On theories with a combinatorial definition of 'equivalence'

Annals of Mathematics 43(2), 223-243: 1942



## ST Termination

- ▶ **Theorem** ST is terminating
- ▶ **Proof (sketch)**
  - ▷ SAT, CONF and SPLIT do not overlap,  
i.e., at most one of these rules is applicable to a node  $F :: J$
  - ▷ If SAT or CONF are applied then their only successor nodes  $F :: \text{SAT}$  and  $F :: \text{CONFLICT}$  are irreducible
  - ▷ We turn to SPLIT
    - ▶▶  $\text{atoms}(F)$  is finite
    - ▶▶ SPLIT is applied to  $F :: J$  yielding two successor nodes  $F :: J, \dot{A}$  and  $F :: J, \bar{\dot{A}}$  if  $A \in \text{atoms}(F|_J)$
    - ▶▶  $A \notin \text{atoms}(F|_{(J, \dot{A})}) \cup \text{atoms}(F|_{(J, \bar{\dot{A}})})$
    - ▶▶ There are no infinite sequences of SPLIT

qed



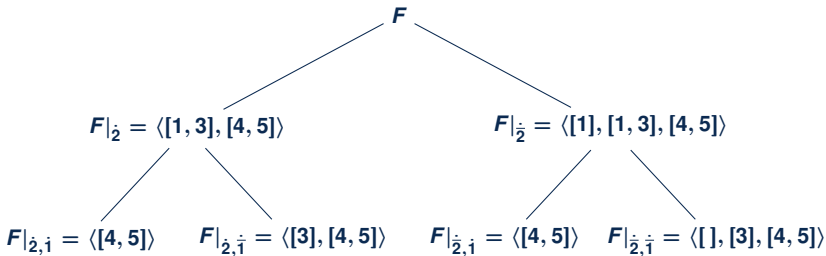
## ST Confluency – Preliminaries

- ▶ We assume that nodes are labeled by  $F|_J$  instead of  $F :: J$
- ▶ **Observations**
  - ▷  $F|_{J,L_1,L_2} = F|_{J,L_2,L_1}$
  - ▷  $F|_J = F|_{J,A} = F|_{J,\bar{A}}$  if  $A \notin \text{atoms}(F|_J)$  and neither  $A \in J$  nor  $\bar{A} \in J$



## Example

- Let  $F = \langle [1, 2], [1, 3], [4, 5] \rangle$

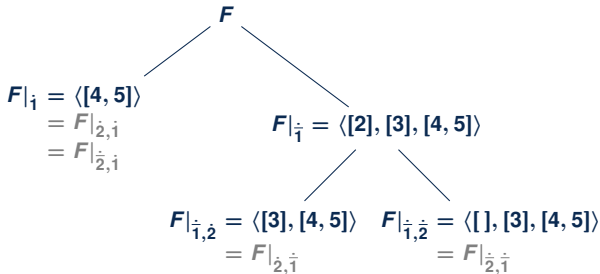


- The set of leaves is  $\{ \langle [4, 5] \rangle, \langle [3], [4, 5] \rangle, \langle [], [3], [4, 5] \rangle \}$



## Example – Continued

- ▶ Let  $F = \langle [1, 2], [1, 3], [4, 5] \rangle$



- ▶ The set of leaves is  $\{ \langle [4, 5] \rangle, \langle [3], [4, 5] \rangle, \langle [], [3], [4, 5] \rangle \}$ , which is identical to the set of leaves of the previous tree
- ▶ Two trees with identical root are **similar** if they have identical sets of leaves



## ST Confluency

- ▶ **Proposition** ST is confluent (modulo similarity of trees)
- ▶ **Proof (sketch)** Because ST is terminating and, thus, Newman's Lemma is applicable, it suffices to show that ST is locally confluent
  - ▷ Because SAT, CONF, and SPLIT do not overlap, the only possible overlap is between two different instance of SPLIT applicable to some node  $F|_J$
  - ▷ Let  $F|_{J,\dot{A}}$  and  $F|_{J,\dot{\bar{A}}}$  as well as  $F|_{J,\dot{B}}$  and  $F|_{J,\dot{\bar{B}}}$  be the respective extensions of  $F|_J$ , where  $A \neq B$
  - ▷ If  $B \in \text{atoms}(F|_{J,\dot{A}})$  then SPLIT can be applied to  $F|_{J,\dot{A}}$  yielding  $F|_{J,\dot{A},\dot{B}}$  and  $F|_{J,\dot{A},\dot{\bar{B}}}$ ; otherwise,  $F|_{J,\dot{A}} = F|_{J,\dot{A},\dot{B}} = F|_{J,\dot{A},\dot{\bar{B}}}$
  - ▷ Similar arguments can be made for the remaining three cases
  - ▷ Because literals can be swapped in an interpretation, the two trees rooted in  $F|_J$  and generated by the two different initial applications of SPLIT are similar

qed





## A Comment

- ▶  $[\ ] \notin F|_J$  may be omitted
  - ▷ in the last condition of the definition of a semantic tree and, consequently,
  - ▷ in the definition of SPLIT
- ▶ Hence,
  - ▷ CONF and SPLIT overlap
  - ▷ However, ST is still confluent
  - ▷ One can show that CONF is a simplification rule
  - ▷ Thus, CONF should always be applied first and no alternatives need to be considered



## ST Soundness

- ▶ **Lemma** Suppose  $F :: J \rightsquigarrow_{SPLIT} F :: J, \dot{A} \mid F :: J, \dot{\bar{A}}$ . Then,  $F|_J$  is satisfiable **iff** either  $F|_{(J, \dot{A})}$  or  $F|_{(J, \dot{\bar{A}})}$  is satisfiable
- ▶ **Proof**  $\rightsquigarrow$  Exercise
- ▶ **Theorem** ST is sound
- ▶ **Proof (sketch)**
  - ▷ To show if  $F :: () \rightsquigarrow^* F :: SAT$  then the CNF-formula  $F$  is satisfiable
  - ▷ Suppose  $F :: () \rightsquigarrow^* F :: SAT$
  - ▷  $F :: SAT$  is generated iff its parent node is  $F :: J$  and  $F|_J = \langle \rangle$
  - ▷  $\langle \rangle$  is satisfiable
  - ▷ By induction on the length of the given derivation and using the above mentioned lemma we can show that  $F$  is satisfiable qed
- ▶ **Exercise** Complete the proof of the Theorem



## ST Completeness

- ▶ **Corollary** Suppose  $F :: J \rightsquigarrow_{SPLIT} F :: J, \dot{A} \mid F :: J, \dot{\bar{A}}$ . Then,  $F|_J$  is unsatisfiable iff  $F|_{(J, \dot{A})}$  and  $F|_{(J, \dot{\bar{A}})}$  are unsatisfiable
- ▶ **Proof** Follows from the previous lemma by negating both sides of the equivalence qed
- ▶ **Theorem** ST is complete
- ▶ **Proof (sketch)**
  - ▷ To show if a CNF-formula  $F$  is satisfiable, then  $F :: () \rightsquigarrow^* F :: SAT$
  - ▷ Suppose  $F$  is satisfiable, but  $F :: () \not\rightsquigarrow^* F :: SAT$
  - ▷ Because ST is terminating, all leaf nodes are of the form  $F :: CONFLICT$
  - ▷ For all leaf nodes we find  $J$  such that  $F :: J \rightsquigarrow_{CONF} F :: CONFLICT$  and  $[] \in F|_J$
  - ▷ By induction on the length of the given derivation and using the above mentioned corollary we learn that  $F$  is unsatisfiable qed



## Relationship to the Truth Table Method

- ▶ Each branch of a semantic tree corresponds to rows in the truth table
- ▶ Different branches correspond to different rows in the truth table
- ▶ Advantage over the truth table method
  - ▷ Leaf nodes  $F :: \text{CONFLICT}$  and  $F :: \text{SAT}$  may be reached even if not all members of atoms( $F$ ) have been assigned to truth values



## Controlling the Generation of Semantic Trees

- ▶ Which leaf node  $F :: J$  shall be selected?
- ▶ Which atom  $A$  shall be selected in an application of SPLIT?
- ▶ Which branch shall be investigated first after an application of SPLIT?
- ▶ What about redundancies?



## DPLL

- ▶ **This section is based on**
  - ▷ Davis, Putnam: A Computing Procedure for Quantification Theorem  
Journal of the ACM 7, 201-215: 1960
  - ▷ Davis, Logemann, Loveland: A Machine Program for Theorem Provin.  
Communications of the ACM 5, 394-397: 1962
- ▶ **DPLL is an acronym for the authors**
- ▶ **The DPLL method was originally specified to show unsatisfiability**
- ▶ **Here, we present a version for showing satisfiability  
leading to an improved algorithm for the generation of semantic trees**
- ▶ **We consider clauses as sets**
  - ▷ **There are no multiple occurrences of literals in a clause**
  - ▷ **An implementation has to guarantee this!**



## Simplification Rules

- ▶ Consider a CNF-formula  $F$
- ▶ Consider rules which yield  $F'$  such that  $F' \equiv F$  and  $F'$  is “simpler” than  $F$
- ▶ Such rules can be applied at any time
- ▶ They are often called **simplification rules**
- ▶ **Here** TAUT and SUBS



## TAUT: Tautological Clauses

### ▶ Definition

A clause is a **tautology** iff it contains a complementary pair of literals

### ▶ Proposition

Tautologies can be deleted while preserving semantic equivalence  
i.e.,  $F, C \equiv F$  if  $C$  is a tautology

$$\langle [1, \bar{2}, 2, 3, 4, \bar{7}], [5, \bar{6}] \rangle \equiv \langle [5, \bar{6}] \rangle$$

▶  $F, C :: J \rightsquigarrow_{\text{TAUT}} F :: J$  iff  $C$  is a tautology

### ▶ Applicable

- ▷ in the initialization phase
- ▷ whenever a new clause is generated, e.g., by resolution

▶ TAUT reduces the number of clauses in a formula





## SUBS: Subsumption

▶ **Definition**  $C_1$  **subsumes**  $C_2$  **iff**  $C_1 \subseteq C_2$

▶ **Proposition**

Subsumed clauses can be deleted while preserving semantic equivalence  
i.e.,  $F, C \equiv F$  if there exists  $C' \in F$  with  $C' \subseteq C$

$$\langle [2, \bar{3}], [\bar{2}], [1, 2, \bar{3}, \bar{4}] \rangle \equiv \langle [2, \bar{3}], [\bar{2}] \rangle$$

▶  $F, C :: J \rightsquigarrow_{SUBS} F :: J$  **iff** there exists  $C' \in F$  such that  $C' \subseteq C$

▶ **Applicable**

▷ in the initialization phase

▷ whenever a new clause is generated, e.g., by resolution

▶ **SUBS reduces the number of clauses in a formula**

▷ **How complex is the removal of subsumed clauses?**

▷ **Are there forms of subsumption which are less costly?**



## Remaining Rules

- ▶ **SAT, SPLIT and CONFLICT** as in the ST calculus
- ▶ **UNIT** as a special variant of SPLIT
- ▶ **PURE**



## UNIT

- ▶ Let  $F :: J$  be a node in the computation of a semantic tree for  $F$
- ▶ Suppose SPLIT was applied yielding the new nodes  $F :: J, \dot{L}$  and  $F :: J, \dot{\bar{L}}$ 
  - ▶ If  $[L] \in F|_J$  then  $[\ ] \in F|_{(J, \dot{L})}$  and, thus,  $F :: J, \dot{L} \rightsquigarrow_{CONF} F :: CONFLICT$
  - ▶ If  $[\bar{L}] \in F|_J$  then  $[\ ] \in F|_{(J, \dot{\bar{L}})}$  and, thus,  $F :: J, \dot{\bar{L}} \rightsquigarrow_{CONF} F :: CONFLICT$
- ▶ Hence, unit clauses should eagerly trigger SPLITS
- ▶  $F :: J \rightsquigarrow_{UNIT} F :: J, L$  iff  $[L] \in F|_J$
- ▶  $L$  is a propagation variable
- ▶ Applicable
  - ▶ in the initialization phase
  - ▶ whenever a new clause is generated, e.g., by resolution
  - ▶ whenever literals are deleted from a clause, e.g., by UNIT
- ▶ How complex is the application of UNIT?

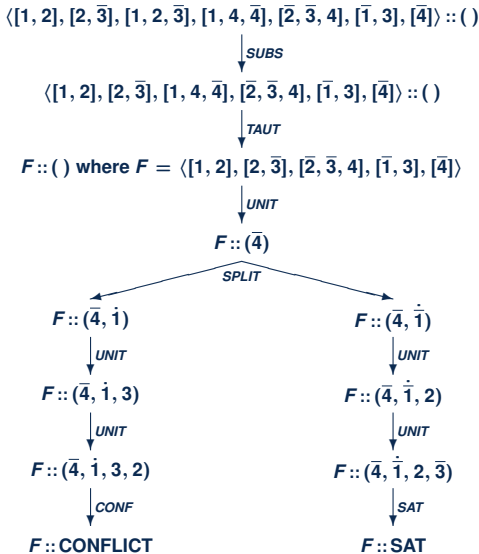


# PURE

- ▶ **Definition** A literal  $L \in \text{lits}(F)$  is **pure** iff  $\bar{L} \notin \text{lits}(F)$
- ▶ Clauses containing a pure  $L$  are satisfied by any interpretation containing  $L$
- ▶ Interpretations containing  $\bar{L}$  need not be considered
- ▶  $F :: J \rightsquigarrow_{\text{PURE}} F :: J, L$  iff there exists  $L \in \text{lits}(F|_J)$  which is pure in  $F|_J$ .
- ▶  $L$  is a propagation variable
- ▶ **Applicable**
  - ▷ in the initialization phase
  - ▷ whenever clauses have been deleted, e.g., by SUBS or TAUT



## An Example



## The DPLL Calculus

- ▶ Given a CNF-formula  $F$
- ▶ The computation is initialized by  $F :: ()$
- ▶ The rules of the calculus are  
SAT, CONFLICT, SPLIT, TAUT, SUBS, UNIT and PURE
- ▶ Computation terminates if
  - ▷ a node  $F' :: \text{SAT}$  is reached in which case  $F$  is satisfiable or
  - ▷ all leaf nodes are of the form  $F' :: \text{CONFLICT}$  in which case  $F$  is unsatisfiable
- ▶ **Note**
  - ▷ TAUT and SUBS may still be applicable to  $F' :: \text{CONFLICT}$
  - ▷ However, TAUT and SUBS can only be applied finitely many times because in each application a clause from  $F'$  is removed and  $F'$  is finite



## UNIT and PURE Revisited

- ▶ **Proposition** UNIT and PURE are satisfiability preserving, i.e.,
  - ▷ Suppose  $F :: J \rightsquigarrow_{UNIT} F :: J, L$ . Then,  
 $F|_J$  is satisfiable **iff**  $F|_{J,L}$  is satisfiable
  - ▷ Suppose  $F :: J \rightsquigarrow_{PURE} F :: J, L$ . Then,  
 $F|_J$  is satisfiable **iff**  $F|_{J,L}$  is satisfiable
- ▶ **Exercise** Prove the proposition
- ▶ **Note** Whenever UNIT or PURE is applied to  $F :: J$  yielding  $F :: J, L$  then
  - ▷  $L \notin \text{lits}(F|_{J,L})$
  - ▷  $\text{atoms}(L) \not\subseteq \text{atoms}(F|_{J,L})$
  - ▷  $\text{lits}(F|_{J,L}) \cup \{L\} \subseteq \text{lits}(F|_J)$
- ▶ **Exercise** Give examples for the application of UNIT and PURE where the subset relation is proper



## DPLL Termination

- ▶ **Theorem** DPLL is terminating
- ▶ **Proof (sketch)** The theorem follows from the following observations:
  - ▷ **SAT**: the node cannot be further extended
  - ▷ **CONFLICT**: the node will not be further extended
  - ▷ **TAUT, SUBS, UNIT, PURE**: the number of clauses occurring in  $F$  decreases
  - ▷ **SPLIT**:
    - ▶▶ the number of clauses does not increase
    - ▶▶ the number of atoms occurring in  $F|_J$  decreases
- ▶ **Exercise** Complete the proof

qed





## DPLL Confluency

- ▶ **Claim** DPLL is confluent
- ▶ **Some Consequences**
  - ▷ The inference rules can be applied in any order
  - ▷ Starting with  $F :: ()$  we can first apply TAUT and SUBS as often as possible
  - ▷ Thereafter, TAUT and SUBS will not be applicable anymore
  - ▷ We delay applications of SPLIT as long as possible, i.e., if TAUT and SUBS are no longer applicable, we apply PURE and UNIT eagerly



## DPLL Soundness

- ▶ **Theorem** DPLL is sound
- ▶ **Proof** Follows from the corresponding result of the GenericCDCL calculus which will be presented later in the lecture



## DPLL Completeness

▶ **Theorem** DPLL is complete

▶ **Proof (sketch)**

The proof is in analogy to the proof of the completeness of the ST calculus

▷ Recall that

▶ TAUT and SUBS are simplification rules and

▶ PURE and UNIT are satisfiability preserving

▷ Hence, if  $F :: J \rightsquigarrow F' :: J'$ , where  $\rightsquigarrow \in \{\rightsquigarrow_{TAUT}, \rightsquigarrow_{UNIT}, \rightsquigarrow_{SUBS}, \rightsquigarrow_{PURE}\}$ , then the following holds: If  $F'|_{J'}$  is unsatisfiable, so is  $F|_J$ .

▷ The remaining steps are similar to those in the proof of the completeness of the ST calculus except that between two splits TAUT, SUBS, PURE, and UNIT may be applied

qed

▶ **Exercise** Complete the proof



## Naive Backtracking

- ▶ The ST and the DPLL calculus are branching due to SPLIT
- ▶ We would like to linearize DPLL and, thereby, ST
- ▶ TAUT and SUBS simplify the formula, SAT is a termination rule
- ▶ UNIT and PURE are satisfiability preserving and add propagation literals
- ▶ SPLIT is replaced by
  - ▷  $F :: J \rightsquigarrow_{DECIDE} F :: J, \dot{L} \text{ iff } [] \notin F|_J \text{ and } L \in \text{atoms}(F|_J) \cup \overline{\text{atoms}(F|_J)}$
- ▶ If  $[] \in F|_J$  then  $J$  may or may not contain decision literals
  - ▷  $F :: J \rightsquigarrow_{UNSAT} F :: UNSAT$   
iff  $[] \in F|_J$  and  $J$  does not contain a decision literal
  - ▷  $F :: J, \dot{L}, P \rightsquigarrow_{NB} F :: J, \bar{L} \text{ iff } [] \in F|_{J, \dot{L}, P}$ , where
    - ▶▶  $P$  is a sequence of propagation literals
    - ▶▶  $\dot{L}$  is the decision literal with the highest level in  $J, \dot{L}, P$
    - ▶▶  $\bar{L}$  is a propagation literal
    - ▶▶ NB is called **naive backtracking**



## A Note on PURE

- ▶ Each application of PURE can be replaced by DECIDE, in which case the pure literal  $L$  used by PURE becomes a decision literal
- ▶ In most of the literature and almost all systems PURE is not considered
- ▶ We would like to keep it in the moment as we do not fully understand why an implementation of PURE is so costly or why a particular result is affected by PURE
- ▶ If, however, one of the methods and techniques presented in the sequel causes a problem due to the fact that PURE adds  $L$  as a propagation literal to the current partial interpretation then PURE shall be replaced by DECIDE



## The Previous Example Revisited

$\langle [1, 2], [2, \bar{3}], [1, 2, \bar{3}], [1, 4, \bar{4}], [\bar{2}, \bar{3}, 4], [\bar{1}, 3], [\bar{4}] \rangle :: ()$

$\rightsquigarrow_{SUBS} \langle [1, 2], [2, \bar{3}], [1, 4, \bar{4}], [\bar{2}, \bar{3}, 4], [\bar{1}, 3], [\bar{4}] \rangle :: ()$

$\rightsquigarrow_{TAUT} F :: ()$  where  $F = \langle [1, 2], [2, \bar{3}], [\bar{2}, \bar{3}, 4], [\bar{1}, 3], [\bar{4}] \rangle$

$\rightsquigarrow_{UNIT} F :: (\bar{4})$  ( $F|_{(\bar{4})} = \langle [1, 2], [2, \bar{3}], [\bar{2}, \bar{3}], [\bar{1}, 3] \rangle$ )

$\rightsquigarrow_{DECIDE} F :: (\bar{4}, \dot{1})$  ( $F|_{(\bar{4}, \dot{1})} = \langle [2, \bar{3}], [\bar{2}, \bar{3}], [3] \rangle$ )

$\rightsquigarrow_{UNIT} F :: (\bar{4}, \dot{1}, 3)$  ( $F|_{(\bar{4}, \dot{1}, 3)} = \langle [2], [\bar{2}] \rangle$ )

$\rightsquigarrow_{UNIT} F :: (\bar{4}, \dot{1}, 3, 2)$  ( $F|_{(\bar{4}, \dot{1}, 3, 2)} = \langle [] \rangle$ )

$\rightsquigarrow_{NB} F :: (\bar{4}, \bar{1})$  ( $F|_{(\bar{4}, \bar{1})} = \langle [2], [2, \bar{3}], [\bar{2}, \bar{3}] \rangle$ )

$\rightsquigarrow_{UNIT} F :: (\bar{4}, \bar{1}, 2)$  ( $F|_{(\bar{4}, \bar{1}, 2)} = \langle [\bar{3}] \rangle$ )

$\rightsquigarrow_{UNIT} F :: (\bar{4}, \bar{1}, 2, \bar{3})$  ( $F|_{(\bar{4}, \bar{1}, 2, \bar{3})} = \langle \rangle$ )

$\rightsquigarrow_{SAT} F :: SAT$



## Another Example Revisited

$F :: ()$

$\rightsquigarrow_{DECIDE}$

$F :: (1)$

$\rightsquigarrow_{UNIT}$

$F :: (1, 2)$

$\rightsquigarrow_{UNIT}$

$F :: (1, 2, 3)$

$\rightsquigarrow_{NB}$

$F :: (1)$

$\rightsquigarrow_{UNIT}$

$F :: (1, 3)$

$\rightsquigarrow_{UNIT}$

$F :: (1, 3, 2)$

$\rightsquigarrow_{UNSAT}$

$F :: UNSAT$

where  $F = \langle [2, \bar{3}], [2, 3], [\bar{1}, \bar{2}], [1, \bar{3}], [1, \bar{2}, 3] \rangle$

$(F|_{(1)} = \langle [2, \bar{3}], [2, 3], [\bar{2}] \rangle)$

$F|_{(1,2)} = \langle [\bar{3}], [3] \rangle)$

$(F|_{(1,2,3)} = \langle [] \rangle)$

$(F|_{(\bar{1})} = \langle [2, \bar{3}], [2, 3], [\bar{3}], [\bar{2}, 3] \rangle)$

$(F|_{(\bar{1},\bar{3})} = \langle [2], [\bar{2}] \rangle)$

$(F|_{(\bar{1},\bar{3},2)} = \langle [] \rangle)$



## The DPLL-NB Calculus

- ▶ Given a CNF-formula  $F$
- ▶ The computation is initialized by  $F :: ()$
- ▶ The rules of the calculus are SAT, UNSAT, DECIDE, NB, TAUT, SUBS, UNIT and PURE
- ▶ Computation terminates if
  - ▷ a node  $F' :: \text{SAT}$  is reached in which case  $F$  is satisfiable or
  - ▷ a node  $F' :: \text{UNSAT}$  is reached in which case  $F$  is unsatisfiable
- ▶ **Note** In DPLL-NB and, in particular, in  $F :: J$ ,  $J$  may be a partial interpretation, SAT or UNSAT





## DPLL-NB – Results

- ▶ **Theorem** DPLL-NB is terminating, sound, and complete
  - ▶ **Proof (sketch)**
    - ▷ **Termination and soundness follow from corresponding results for the GenericCDCL calculus, which will be presented later in the lecture**
    - ▷ **Completeness**
      - ▶▶ DPLL is complete
      - ▶▶ The search space is finite
      - ▶▶ NB specifies just a specific order of traversing this space
- qed



## Heuristics

- ▶ Whenever **DECIDE** is applied to  $F :: J$  in the following examples, then
  - ▷ the smallest atom  $A$  occurring in  $F|_J$  is selected
- ▶ Whenever **UNIT** is applied to  $F :: J$  in the following examples, then
  - ▷ the leftmost unit clause occurring in  $F|_J$  is selected



## Backtracking and Redundancies (1)

$F :: ()$	where $F = \langle [\bar{1}, \bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [\bar{5}, 6], [\bar{1}, \bar{5}, \bar{6}], [5, 7], [\bar{1}, 5, \bar{7}], [1, 3] \rangle$
$\rightsquigarrow$ DECIDE	$F :: (\bar{1})$ $(F _{(\bar{1})} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ DECIDE	$F :: (\bar{1}, \bar{2})$ $(F _{(\bar{1}, \bar{2})} = \langle [\bar{3}], [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, \bar{3})$ $(F _{(\bar{1}, \bar{2}, \bar{3})} = \langle [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, \bar{3}, 4)$ $(F _{(\bar{1}, \bar{2}, \bar{3}, 4)} = \langle [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ DECIDE	$F :: (\bar{1}, \bar{2}, \bar{3}, 4, \bar{5})$ $(F _{(\bar{1}, \bar{2}, \bar{3}, 4, \bar{5})} = \langle [6], [\bar{6}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, \bar{3}, 4, \bar{5}, 6)$ $(F _{(\bar{1}, \bar{2}, \bar{3}, 4, \bar{5}, 6)} = \langle [] \rangle)$
$\rightsquigarrow$ NB	$F :: (\bar{1}, \bar{2}, \bar{3}, 4, \bar{5})$ $(F _{(\bar{1}, \bar{2}, \bar{3}, 4, \bar{5})} = \langle [7], [\bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, \bar{3}, 4, \bar{5}, 7)$ $(F _{(\bar{1}, \bar{2}, \bar{3}, 4, \bar{5}, 7)} = \langle [] \rangle)$
$\rightsquigarrow$ NB	$F :: (\bar{1}, \bar{2})$ $(F _{(\bar{1}, \bar{2})} = \langle [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, 4)$ $(F _{(\bar{1}, \bar{2}, 4)} = \langle [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ DECIDE	$F :: (\bar{1}, \bar{2}, 4, \bar{5})$ $(F _{(\bar{1}, \bar{2}, 4, \bar{5})} = \langle [6], [\bar{6}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, 4, \bar{5}, 6)$ $(F _{(\bar{1}, \bar{2}, 4, \bar{5}, 6)} = \langle [] \rangle)$
$\rightsquigarrow$ NB	$F :: (\bar{1}, \bar{2}, 4, \bar{5})$ $(F _{(\bar{1}, \bar{2}, 4, \bar{5})} = \langle [7], [\bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, \bar{2}, 4, \bar{5}, 7)$ $(F _{(\bar{1}, \bar{2}, 4, \bar{5}, 7)} = \langle [] \rangle)$
$\rightsquigarrow$ NB	$F :: (\bar{1})$ $(F _{(\bar{1})} = \langle [\bar{2}, 4], [2, 4], [\bar{5}, 6], [5, 7], [3] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\bar{1}, 3)$ $(F _{(\bar{1}, 3)} = \langle [\bar{2}, 4], [2, 4], [\bar{5}, 6], [5, 7] \rangle)$
$\rightsquigarrow$ PURE	$F :: (\bar{1}, 3, 7)$ $(F _{(\bar{1}, 3, 7)} = \langle [\bar{2}, 4], [2, 4], [\bar{5}, 6] \rangle)$
$\rightsquigarrow$ PURE	$F :: (\bar{1}, 3, 7, \bar{5})$ $(F _{(\bar{1}, 3, 7, \bar{5})} = \langle [\bar{2}, 4], [2, 4] \rangle)$
$\rightsquigarrow$ PURE	$F :: (\bar{1}, 3, 7, \bar{5}, 4)$ $(F _{(\bar{1}, 3, 7, \bar{5}, 4)} = \langle \rangle)$
$\rightsquigarrow$ SAT	$F :: \text{SAT}$



## Conflict Analysis

- ▶ In the previous example the following clauses triggered UNIT propagation

$$\begin{aligned}
 C_1 &= [\bar{1}, \bar{2}, \bar{3}] & (C_1|_{(1,2)} &= [\bar{3}]) \\
 C_2 &= [\bar{2}, 4] & (C_2|_{(1,2)} &= [4]) \\
 C_3 &= [\bar{5}, 6] & (C_3|_{(1,2,\bar{3},4,5)} &= [6])
 \end{aligned}$$

Subsequently, the clause  $C = [\bar{1}, \bar{5}, \bar{6}]$  became empty and caused a conflict

- ▶ We can find the following (linear) resolution derivation from  $C$  wrt  $\{C_1, C_2, C_2\}$

$$C_4 = [\bar{1}, \bar{5}] \quad (\text{res}(C, C_3))$$

- ▶ **Note**

- ▶ Resolvents can be added while preserving semantic equivalence
- ▶  $[\bar{1}, \bar{5}]|_{(1)} = [\bar{5}]$



## Backtracking and Redundancies (2)

$F :: ()$  where  $F = \langle [\bar{1}, \bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [\bar{5}, 6], [\bar{1}, \bar{5}, \bar{6}], [5, 7], [\bar{1}, 5, \bar{7}], [1, 3] \rangle$

$\rightsquigarrow_{DECIDE} F :: (\dot{1}) \quad (F|_{(1)} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$

$\rightsquigarrow_{DECIDE} F :: (\dot{1}, \dot{2}) \quad (F|_{(1,2)} = \langle [\bar{3}], [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$

$\rightsquigarrow_{UNIT} F :: (\dot{1}, \dot{2}, \bar{3}) \quad (F|_{(1,2,\bar{3})} = \langle [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$

$\rightsquigarrow_{UNIT} F :: (\dot{1}, \dot{2}, \bar{3}, 4) \quad (F|_{(1,2,\bar{3},4)} = \langle [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$

$\rightsquigarrow_{DECIDE} F :: (\dot{1}, \dot{2}, \bar{3}, 4, \dot{5}) \quad (F|_{(1,2,\bar{3},4,5)} = \langle [6], [\bar{6}] \rangle)$

$\rightsquigarrow_{UNIT} F :: (\dot{1}, \dot{2}, \bar{3}, 4, \dot{5}, 6) \quad (F|_{(1,2,\bar{3},4,5,6)} = \langle [] \rangle)$

$\rightsquigarrow_{LEARN} F, [\bar{1}, \bar{5}] :: (\dot{1}, \dot{2}, \bar{3}, 4, \dot{5}, 6)$

$\rightsquigarrow_{BACK} F, [\bar{1}, \bar{5}] :: (\dot{1})$

$\rightsquigarrow_{UNIT} F, [\bar{1}, \bar{5}] :: (\dot{1}, \bar{5}) \quad ((F, [\bar{1}, \bar{5}])|_{(1,\bar{5})} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [7], [\bar{7}] \rangle)$

$\rightsquigarrow_{CDBL} F, [\bar{1}, \bar{5}] :: (\dot{1}, \bar{5}) \quad ((F, [\bar{1}, \bar{5}])|_{(1,\bar{5})} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [7], [\bar{7}] \rangle)$

$\rightsquigarrow_{UNIT} F, [\bar{1}, \bar{5}] :: (\dot{1}, \bar{5}, 7) \quad ((F, [\bar{1}, \bar{5}])|_{(1,\bar{5},7)} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [] \rangle)$



## Conflict Analysis (2)

- ▶ The following clauses triggered UNIT propagation:

$$C_1 = [5, 7] \quad (C_1|_{(1, \bar{5})} = [7])$$

$$C_2 = [\bar{1}, \bar{5}] \quad (C_2|_{(1)} = [\bar{5}])$$

- ▶ The new conflict was caused by  $C = [\bar{1}, 5, \bar{7}]$
- ▶ We can find the following (linear) resolution derivation from  $C$  wrt  $\{C_1, C_3\}$ :

$$C_3 = [\bar{1}, 5] \quad (\text{res}(C, C_1))$$

$$C_4 = [\bar{1}] \quad (\text{res}(C_3, C_2))$$

- ▶ **Note**

- ▶ A unit clause can be added
- ▶ This clause should be considered at the start
- ▶  $[\bar{1}]$  subsumes  $[\bar{1}, \bar{5}]$



## Backtracking and Redundancies (3)

	$F :: ()$	where $F = \langle [\bar{1}, \bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [\bar{5}, 6], [\bar{1}, \bar{5}, \bar{6}], [5, 7], [\bar{1}, 5, \bar{7}], [1, 3] \rangle$
$\rightsquigarrow$ DECIDE	$F :: (\dot{1})$	$(F _{(1)} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ DECIDE	$F :: (\dot{1}, \dot{2})$	$(F _{(1,2)} = \langle [\bar{3}], [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\dot{1}, \dot{2}, \bar{3})$	$(F _{(1,2,\bar{3})} = \langle [4], [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\dot{1}, \dot{2}, \bar{3}, 4)$	$(F _{(1,2,\bar{3},4)} = \langle [\bar{5}, 6], [\bar{5}, \bar{6}], [5, 7], [5, \bar{7}] \rangle)$
$\rightsquigarrow$ DECIDE	$F :: (\dot{1}, \dot{2}, \bar{3}, 4, \dot{5})$	$(F _{(1,2,\bar{3},4,5)} = \langle [6], [\bar{6}] \rangle)$
$\rightsquigarrow$ UNIT	$F :: (\dot{1}, \dot{2}, \bar{3}, 4, \dot{5}, 6)$	$(F _{(1,2,\bar{3},4,5,6)} = \langle [] \rangle)$
$\rightsquigarrow$ CDBL	$F, [\bar{1}, \bar{5}] :: (\dot{1}, \bar{5})$	$((F, [\bar{1}, \bar{5}]) _{(1,\bar{5})} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [7], [\bar{7}] \rangle)$
$\rightsquigarrow$ UNIT	$F, [\bar{1}, \bar{5}] :: (\dot{1}, \bar{5}, 7)$	$((F, [\bar{1}, \bar{5}]) _{(1,\bar{5},7)} = \langle [\bar{2}, \bar{3}], [\bar{2}, 4], [2, 4], [] \rangle)$
$\rightsquigarrow$ CDBL	$F, [\bar{1}, \bar{5}], [\bar{1}] :: (\bar{1})$	$(F, [\bar{1}, \bar{5}], [\bar{1}]) _{(\bar{1})} = \langle [\bar{2}, 4], [2, 4], [\bar{5}, 6], [5, 7], [3] \rangle)$
$\rightsquigarrow$ UNIT	$F, [\bar{1}, \bar{5}], [\bar{1}] :: (\bar{1}, 3)$	$(F, [\bar{1}, \bar{5}], [\bar{1}]) _{(\bar{1},3)} = \langle [\bar{2}, 4], [2, 4], [\bar{5}, 6], [5, 7] \rangle)$
$\rightsquigarrow$ PURE	$F, [\bar{1}, \bar{5}], [\bar{1}] :: (\bar{1}, 3, 4, 6, 7)$	$(F, [\bar{1}, \bar{5}], [\bar{1}]) _{(\bar{1},3,4,6,7)} = \langle \rangle)$
$\rightsquigarrow$ SAT	$F, [\bar{1}, \bar{5}], [\bar{1}] :: SAT$	



## Relevant Clauses

- ▶ **Definition** A clause  $C$  is **relevant** in  $F :: J$   
iff  $C \in F$  and there exist  $I, L, I'$  such that  $J = I, L, I'$  and  $C|_I = [L]$
- ▶  $\text{relevant}(F :: J) = \{C \in F \mid C \text{ is relevant in } F :: J\}$





## Conflict-Directed Backtracking and Learning

- ▶ Replace naive backtracking by conflict-directed backtracking and learning
- ▶  $F :: J, \dot{L}, J' \rightsquigarrow_{CDBL} F, D :: J, L'$  **iff**
  - ▷ there exists  $C \in F$  such that  $C|_{J, \dot{L}, J'} = []$
  - ▷ there is a linear resolution derivation from  $C$  to  $D$  wrt  $\text{relevant}(F :: J, \dot{L}, J')$
  - ▷  $D|_J = [L']$



## The DPLL-CDBL Calculus

- ▶ Given a CNF-formula  $F$
- ▶ The computation is initialized by  $F :: ()$
- ▶ The rules of the calculus are  
SAT, UNSAT, DECIDE, CDBL, TAUT, SUBS, UNIT and PURE
- ▶ Computation terminates if
  - ▷ a node  $F' :: \text{SAT}$  is reached in which case  $F$  is satisfiable or
  - ▷ a node  $F' :: \text{UNSAT}$  is reached in which case  $F$  is unsatisfiable



## Another Example

	$F :: ()$	where $F = \langle [\bar{1}, 2], [\bar{2}, 3], [\bar{4}, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{3}, \bar{8}, \bar{9}] \rangle$
$\rightsquigarrow_{DECIDE}$	$F :: (\dot{1})$	$(F _{(1)} = \langle [2], [\bar{2}, 3], [4, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{3}, \bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2)$	$(F _{(1,2)} = \langle [3], [4, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{3}, \bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3)$	$(F _{(1,2,3)} = \langle [\bar{4}, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{DECIDE}$	$F :: (\dot{1}, 2, 3, \dot{4})$	$(F _{(1,2,3,4)} = \langle [5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5)$	$(F _{(1,2,3,4,5)} = \langle [6], [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6)$	$(F _{(1,2,3,4,5,6)} = \langle [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{DECIDE}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \dot{7})$	$(F _{(1,2,3,4,5,6,7)} = \langle [8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \dot{7}, 8)$	$(F _{(1,2,3,4,5,6,7,8)} = \langle [9], [\bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \dot{7}, 8, 9)$	$(F _{(1,2,3,4,5,6,7,8,9)} = \langle [] \rangle)$
$\rightsquigarrow_{CDBL}$	$F, [\bar{3}, \bar{8}] :: (\dot{1}, 2, 3, \bar{8})$	$((F, [\bar{3}, \bar{8}]) _{(1,2,3,\bar{8})} = \langle [\bar{4}, 5], [\bar{5}, 6], [\bar{7}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F, [\bar{3}, \bar{8}] :: (\dot{1}, 2, 3, \bar{8}, \bar{7})$	$((F, [\bar{3}, \bar{8}]) _{(1,2,3,\bar{8},\bar{7})} = \langle [\bar{4}, 5], [\bar{5}, 6] \rangle)$
$\rightsquigarrow_{PURE}$	$F, [\bar{3}, \bar{8}] :: (\dot{1}, 2, 3, \bar{8}, \bar{7}, \bar{4})$	$((F, [\bar{3}, \bar{8}]) _{(1,2,3,\bar{8},\bar{7},\bar{4})} = \langle [\bar{5}, 6] \rangle)$
$\rightsquigarrow_{PURE}$	$F, [\bar{3}, \bar{8}] :: (\dot{1}, 2, 3, \bar{8}, \bar{7}, \bar{4}, \bar{5})$	$((F, [\bar{3}, \bar{8}]) _{(1,2,3,\bar{8},\bar{7},\bar{4},\bar{5})} = \langle \rangle)$



## Another Example – Conflict Analysis

- ▶  $\text{relevant}(F :: (\bar{1}, 2, 3, \bar{4}, 5, 6, \bar{7}, 8, 9))$  contains of the following clauses

$$\begin{array}{ll}
 C_1 & = \quad [\bar{1}, 2] & ([\bar{1}, 2]|_{(1)} = [2]) \\
 C_2 & = \quad [\bar{2}, 3] & ([\bar{2}, 3]|_{(1,2)} = [3]) \\
 C_3 & = \quad [\bar{4}, 5] & ([\bar{4}, 5]|_{(1,2,3,4)} = [5]) \\
 C_4 & = \quad [\bar{5}, 6] & ([\bar{5}, 6]|_{(1,2,3,4,5)} = [6]) \\
 C_5 & = \quad [\bar{7}, 8] & ([\bar{7}, 8]|_{(1,2,3,4,5,6,7)} = [8]) \\
 C_6 & = \quad [\bar{8}, 9] & ([\bar{8}, 9]|_{(1,2,3,4,5,6,7,8)} = [9])
 \end{array}$$

- ▶ The conflict clause is  $C = [\bar{3}, \bar{8}, \bar{9}]$
- ▶ We obtain the following linear derivation from  $C$  wrt  $\{C_i \mid 1 \leq i \leq 6\}$

$$\begin{array}{ll}
 C_7 & = \quad [\bar{8}, \bar{3}] & (\text{res}(C, C_6)) \\
 C_8 & = \quad [\bar{7}, \bar{3}] & (\text{res}(C_7, C_5)) \\
 C_9 & = \quad [\bar{7}, \bar{2}] & (\text{res}(C_8, C_2)) \\
 C_{10} & = \quad [\bar{7}, \bar{1}] & (\text{res}(C_9, C_1))
 \end{array}$$

- ▶ In principle, all derived clauses could have been added!



## Another Example – Continued

► There is an alternative application of CDBL

	$F :: ()$	where $F = \langle [\bar{1}, 2], [\bar{2}, 3], [\bar{4}, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{3}, \bar{8}, \bar{9}] \rangle$
$\rightsquigarrow_{DECIDE}$	$F :: (\dot{1})$	$(F _{(1)} = \langle [2], [2, 3], [\bar{4}, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{3}, \bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2)$	$(F _{(1,2)} = \langle [3], [\bar{4}, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{3}, \bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3)$	$(F _{(1,2,3)} = \langle [\bar{4}, 5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{DECIDE}$	$F :: (\dot{1}, 2, 3, \dot{4})$	$(F _{(1,2,3,4)} = \langle [5], [\bar{5}, 6], [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5)$	$(F _{(1,2,3,4,5)} = \langle [6], [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6)$	$(F _{(1,2,3,4,5,6)} = \langle [\bar{7}, 8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{DECIDE}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \dot{7})$	$(F _{(1,2,3,4,5,6,7)} = \langle [8], [\bar{8}, 9], [\bar{8}, \bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \dot{7}, 8)$	$(F _{(1,2,3,4,5,6,7,8)} = \langle [9], [\bar{9}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \dot{7}, 8, 9)$	$(F _{(1,2,3,4,5,6,7,8,9)} = \langle [] \rangle)$
$\rightsquigarrow_{CDBL}$	$F, [\bar{3}, \bar{8}] :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \bar{8})$	$((F, [\bar{3}, \bar{8}]) _{(1,2,3,4,5,6,\bar{8})} = \langle [\bar{7}] \rangle)$
$\rightsquigarrow_{UNIT}$	$F, [\bar{3}, \bar{8}] :: (\dot{1}, 2, 3, \dot{4}, 5, 6, \bar{8}, \bar{7})$	$((F, [\bar{3}, \bar{8}]) _{(1,2,3,4,5,6,\bar{8},\bar{7})} = \langle \rangle)$



## DPLL-CDBL – Results

- ▶ **Theorem** DPLL-CDBL is terminating, sound and complete
- ▶ **Proof**
  - ▷ **Termination and soundness follow from corresponding results for the GenericCDCL calculus, which will be presented later in the lecture**
  - ▷ **Completeness: to do**



## GenericCDCL

- ▶ H., Manthey, Philipp, Steinke: GenericCDCL – A Formalization of Modern Propositional Satisfiability Solvers. In: Proc. POS-14, Le Berre (ed.), EPIc Series 27, 89-102: 2014, EasyChair, <http://www.easychair.org>
- ▶  $F$  and  $F'$  are **equisatisfiable**, in symbols  $F \equiv_{SAT} F'$ ,  
**iff** either both are satisfiable or both are unsatisfiable



## The Rules of the GenericCDCL Calculus

$F :: J$	$\rightsquigarrow_{SAT}$	<b>SAT</b>	iff	$F _J = \langle \rangle$
$F :: J$	$\rightsquigarrow_{UNSAT}$	<b>UNSAT</b>	iff	$[\ ] \in F _J$ and $J$ contains only propagation variables
$F :: J$	$\rightsquigarrow_{DEC}$	$F :: J, \bar{L}$	iff	$L \in \text{atoms}(F) \cup \overline{\text{atoms}(F)}$ and $\{L, \bar{L}\} \cap J = \emptyset$
$F :: J$	$\rightsquigarrow_{INF}$	$F :: J, L$	iff	$F _J \equiv_{SAT} F _{J,L}$ , $L \in \text{atoms}(F) \cup \overline{\text{atoms}(F)}$ and $\{L, \bar{L}\} \cap J = \emptyset$
$F :: J$	$\rightsquigarrow_{LEARN}$	$F, C :: J$	iff	$F \models C$
$F :: J$	$\rightsquigarrow_{REMOVE}$	$F \setminus \{C\} :: J$	iff	$F \setminus \{C\} \models C$
$F :: J, J'$	$\rightsquigarrow_{BACK}$	$F :: J$		
$F :: ()$	$\rightsquigarrow_{INP}$	$F' :: ()$	iff	$F \equiv_{SAT} F'$





## Some Comments

- ▶ **The Rules**
  - ▶ **SAT UNSAT DEC** minor changes
  - ▶ **INF** covers **UNIT** and **PURE** as well as many other techniques
  - ▶ **LEARN** covers all learning techniques in SAT solvers
  - ▶ **REMOVE** covers **SUBS** as well as **TAUT**, but also allows to remove previously learned clauses if they are not effective
  - ▶ **BACK** covers naive backtracking, backjumping as well as restarts
  - ▶ **INP** allows the application of all simplification techniques
- ▶ **GenericCDCL** covers all systematic SAT solvers
- ▶ **'All'** is to be understood as **'to the best of our knowledge, all'**



## The GenericCDCL Calculus

- ▶ Given a CNF-formula  $F$
- ▶ The computation is initialized by  $F :: ()$
- ▶ The rules of the calculus are  
SAT, UNSAT, DEC, INF, LEARN, REMOVE, BACK and INP
- ▶ Computation terminates if
  - ▷ the node SAT is reached in which case  $F$  is satisfiable or
  - ▷ the node UNSAT is reached in which case  $F$  is unsatisfiable



## Invariants

- ▶ **Proposition** If  $F :: () \xrightarrow{n} G :: J$ , then
  - ▷  $F \equiv_{SAT} G$
  - ▷  $G|_{J_1} \equiv_{SAT} G|_{J_1, L}$ , for all  $J_1, J_2$  and propagation literal  $L$  with  $J = J_1, L, J_2$
- ▶ **Proof** by induction on  $n$
- ▶ **Exercise** complete the proof



## Soundness

- ▶ **Theorem** If  $F :: () \rightsquigarrow^* G :: J \rightsquigarrow_{SAT} SAT$ , then  $F$  is satisfiable  
 If  $F :: () \rightsquigarrow^* G :: J \rightsquigarrow_{UNSAT} UNSAT$ , then  $F$  is unsatisfiable
- ▶ **Proof** follows immediately from the previous proposition
- ▶ **Exercise** complete the proof



# Completeness

- ▶ **Theorem** If  $F$  is satisfiable, then  $F :: () \rightsquigarrow^* \text{SAT}$   
 If  $F$  is unsatisfiable, then  $F :: () \rightsquigarrow^* \text{UNSAT}$

- ▶ **Proof**

- ▷ **Suppose  $F$  is satisfiable**

Then we find a model  $J = (L_1, \dots, L_n)$  for  $F$

Then  $F :: () \rightsquigarrow_{DEC}^n F :: (\dot{L}_1, \dots, \dot{L}_n) \rightsquigarrow_{SAT} \text{SAT}$

- ▷ **Suppose  $F$  is unsatisfiable**

Then  $F \models []$

Then  $F :: () \rightsquigarrow_{LEARN} F, [] :: () \rightsquigarrow_{UNSAT} \text{UNSAT}$



## Confluence for Reachable States

- ▶ **Theorem** If  $F::() \xrightarrow{*} T_1$  and  $F::() \xrightarrow{*} T_2$ ,  
then there exists  $T$  with  $T_1 \xrightarrow{*} T$  and  $T_2 \xrightarrow{*} T$
- ▶ **Proof** follows from the completeness of GenericCDCL  
and its ability to perform restarts with the help of the BACK rule
- ▶ **Exercise** complete the proof



## Termination Analysis

- ▶ **GenericCDCL does not terminate due to**
  - ▷ possibly infinite sequences of LEARN and REMOVE
  - ▷ possibly infinite sequences of restarts
  - ▷ possibly infinite sequences of INP
- ▶ **Fairness Criteria**
  - ▷ Each clause  $C \subseteq \text{lits}(F)$  is learned at most finitely many times
    - ▶▶ Eventually, LEARN is no longer applicable
  - ▷ The number of restarts and the number of applications of INP is bounded
    - ▶▶ Eventually, restarts and INP are no longer applicable
- ▶ **Alternative fairness criteria are possible**



## Conflict-Driven Backtracking and Learning Revisited

- ▶ Let  $F$  be a formula in CNF,  $L$  a literal, and  $J$  a partial interpretation.
- ▶  $C \in F$  is called **conflict clause** under  $J$  iff  $C|_J = []$ .
- ▶ A clause  $C$  is **relevant for  $L$**  (or is a **reason for  $L$** ) in  $F :: J$   
iff  $C \in F$  and there exist  $I'$  and  $I$  such that  $J = I', L, I$  and  $C|_I = [L]$ .
- ▶  $\text{relevant}_L(L, F :: J) = \{C \in F \mid C \text{ is relevant for } L \text{ in } F :: J\}$ .
- ▶ A clause  $C$  is **relevant** (or is a **reason clause**) in  $F :: J$   
iff there exists an  $L$  such that  $C$  is relevant for  $L$  in  $F :: J$ .
- ▶  $\text{relevant}(F :: J) = \{C \in F \mid C \text{ is relevant in } F :: J\}$ .





## An Old Example Revisited

▶ Let  $F = \langle [1, 2], [2, \bar{3}], [\bar{2}, \bar{3}, 4], [\bar{1}, 3], [\bar{4}] \rangle$

▶ We may obtain

$$\begin{array}{llll}
 F :: () & \rightsquigarrow_{UNIT} & F :: (\bar{4}) & (F|_{(\bar{4})} = \langle [1, 2], [2, \bar{3}], [\bar{2}, \bar{3}], [\bar{1}, 3] \rangle) \\
 & \rightsquigarrow_{DECIDE} & F :: (\bar{1}, \bar{4}) & (F|_{(\bar{1}, \bar{4})} = \langle [2, \bar{3}], [\bar{2}, \bar{3}], [3] \rangle) \\
 & \rightsquigarrow_{UNIT} & F :: (3, \bar{1}, \bar{4}) & (F|_{(3, \bar{1}, \bar{4})} = \langle [2], [\bar{2}] \rangle) \\
 & \rightsquigarrow_{UNIT} & F :: (2, 3, \bar{1}, \bar{4}) & (F|_{(2, 3, \bar{1}, \bar{4})} = \langle [] \rangle)
 \end{array}$$

▶ We find

$$\begin{array}{ll}
 \text{relevantL}(\bar{4}, F :: (2, 3, \bar{1}, \bar{4})) & = \{[\bar{4}]\} \\
 \text{relevantL}(\bar{1}, F :: (2, 3, \bar{1}, \bar{4})) & = \emptyset \\
 \text{relevantL}(3, F :: (2, 3, \bar{1}, \bar{4})) & = \{[\bar{1}, 3]\} \\
 \text{relevantL}(2, F :: (2, 3, \bar{1}, \bar{4})) & = \{[2, \bar{3}]\} \\
 \text{relevant}(F :: (2, 3, \bar{1}, \bar{4})) & = \{[\bar{4}], [\bar{1}, 3], [2, \bar{3}]\}
 \end{array}$$



## Another Example

▶ Consider  $F = \langle [\bar{1}, \bar{3}, 4], [\bar{2}, \bar{3}, 4] \rangle$ .

▶ We may obtain

$$\begin{array}{lll}
 F :: () & \rightsquigarrow_{DECIDE} & F :: (\dot{1}) & (F|_{(1)} = \langle [\bar{3}, 4], [\bar{2}, \bar{3}, 4] \rangle) \\
 & \rightsquigarrow_{DECIDE} & F :: (\dot{2}, \dot{1}) & (F|_{(2,1)} = \langle [\bar{3}, 4], [\bar{3}, 4] \rangle) \\
 & \rightsquigarrow_{DECIDE} & F :: (\dot{3}, \dot{2}, \dot{1}) & (F|_{(3,2,1)} = \langle [4], [4] \rangle) \\
 & \rightsquigarrow_{UNIT} & F :: (4, \dot{3}, \dot{2}, \dot{1}) & (F|_{(4,3,2,1)} = \langle \rangle) \\
 & \rightsquigarrow_{SAT} & SAT &
 \end{array}$$

▶ We find

$$\text{relevantL}(4, F :: (4, \dot{3}, \dot{2}, \dot{1})) = \{[\bar{1}, \bar{3}, 4], [\bar{2}, \bar{3}, 4]\}$$



## Multiple Conflict Clauses

- ▶ In the sequel
  - ▷ we will apply UNIT whenever possible,
  - ▷ we will prefer literals wrt their absolut value,
  - ▷ we prefer positive over negative literals.
- ▶ Consider  $F = \langle [\bar{1}, 2], [\bar{1}, 3], [\bar{1}, 4], [\bar{2}, \bar{4}], [\bar{3}, \bar{4}] \rangle$ .
- ▶ We obtain

$$\begin{array}{llll}
 F :: () & \rightsquigarrow_{DECIDE} & F :: (\dot{1}) & (F|_{(1)} = \langle [2], [3], [4], [\bar{2}, \bar{4}], [\bar{3}, \bar{4}] \rangle) \\
 & \rightsquigarrow_{UNIT} & F :: (2, \dot{1}) & (F|_{(2,1)} = \langle [3], [4], [\bar{4}], [\bar{3}, \bar{4}] \rangle) \\
 & \rightsquigarrow_{UNIT} & F :: (3, 2, \dot{1}) & (F|_{(3,2,1)} = \langle [4], [\bar{4}], [\bar{4}] \rangle) \\
 & \rightsquigarrow_{UNIT} & F :: (4, 3, 2, \dot{1}) & (F|_{(4,3,2,1)} = \langle [], [] \rangle)
 \end{array}$$

- ▶ There are two conflict clauses, viz.  $[\bar{2}, \bar{4}]$  and  $[\bar{3}, \bar{4}]$ .



## Multiple Conflicts

- ▶ **Recall** We will apply UNIT whenever possible.
- ▶ **Observation** Conflicts can only arise
  - ▷ after an application of UNIT to some  $F :: J$  and
  - ▷ if  $F|_J$  contains two unit clauses  $[A]$  and  $[\bar{A}]$ .
- ▶ Consider  $F = \langle [\bar{1}, \bar{2}, 3], [\bar{2}, \bar{3}] \rangle$ .
  - ▷ We obtain

$$\begin{array}{l}
 F :: () \quad \rightsquigarrow_{DECIDE} \quad F :: (\dot{1}) \quad (F|_{(1)} = \langle [\bar{2}, 3], [\bar{2}, \bar{3}] \rangle) \\
 \rightsquigarrow_{DECIDE} \quad F :: (\dot{2}, \dot{1}) \quad (F|_{(1)} = \langle [3], [\bar{3}] \rangle)
 \end{array}$$

- ▷ Two different conflicts are obtained
  - ▶▶ by applying unit wrt  $[3]$  or  $[\bar{3}]$
  - ▶▶ yielding the conflict clauses  $[\bar{2}, \bar{3}]$  or  $[\bar{1}, \bar{2}, 3]$ , respectively.
- ▶ We assume that one particular conflict is selected.



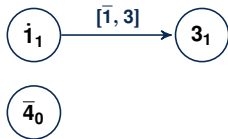
## Implication Graphs

- ▶ In the sequel we assume that  $|\text{relevantL}(L, F :: J)| \leq 1$  for all  $L$ .
  - ▷ If  $|\text{relevantL}(L, F :: J)| > 1$  for some  $L$ , then
    - ▶▶ all relevant clauses wrt  $L$  except one are deleted and
    - ▶▶ we write  $\text{relevantL}(L, F :: J) = C$  instead of  $\text{relevantL}(L, F :: J) = \{C\}$ .
  - ▷ Different selections lead to different implication graphs.
- ▶ Let  $F$  be a formula in CNF and  $J$  a partial interpretation.
- ▶ An **implication graph for  $F :: J$**  is a graph  $(\mathcal{V}, \mathcal{E})$ , where
  - ▷  $\mathcal{V} = J$ ,
  - ▷  $\mathcal{E} = \{(L, L') \in \mathcal{V} \times \mathcal{V} \mid \text{there exists } C = \text{relevantL}(L', F :: J) \text{ and } \bar{L} \in C\}$ ,
  - ▷ each  $(L, L') \in \mathcal{E}$  is directed from  $L$  to  $L'$ .
- ▶ It is sometimes convenient
  - ▷ to label a vertex by its decision level in  $J$  (depicted as subscript)
  - ▷ to label an edge by the clause who caused it.

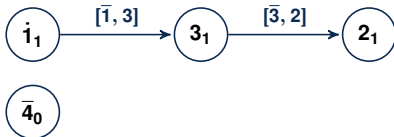


## Implication Graphs for an Old Example

- ▶ Let  $F = \langle [1, 2], [2, \bar{3}], [\bar{2}, \bar{3}, 4], [\bar{1}, 3], [\bar{4}] \rangle$
- ▶ Consider  $J_1 = (3, \dot{1}, \bar{4})$ 
  - ▶ relevant( $L, F :: J_1$ ) =  $\{[\bar{4}], [\bar{1}, 3]\}$

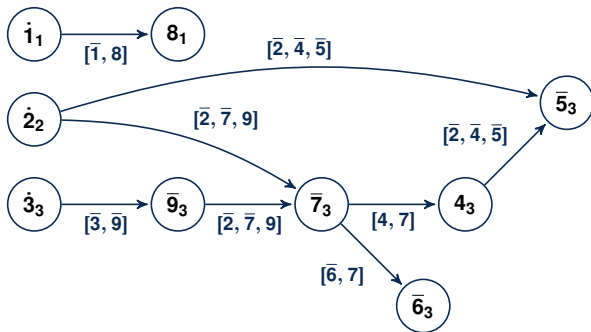


- ▶ Consider  $J_2 = (2, 3, \dot{1}, \bar{4})$ 
  - ▶ relevant( $L, F :: J_2$ ) =  $\{[\bar{4}], [\bar{1}, 3], [\bar{3}, 2]\}$



## Implication Graphs – Another Example

- ▶ Let  $F = \langle [\bar{1}, 8], [\bar{2}, \bar{4}, \bar{5}], [\bar{3}, \bar{9}], [\bar{2}, \bar{7}, 9], [4, 7], [\bar{6}, 7] \rangle$
- ▶ Let  $J = (\bar{5}, \bar{6}, 4, \bar{7}, 9, \dot{3}, \dot{2}, 8, \dot{1})$ .



## Conflict Graphs

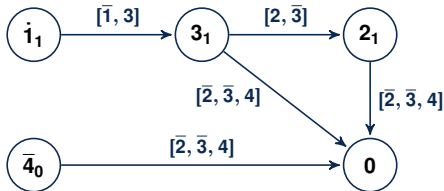
- ▶ Consider  $F :: J$  and let  $(\mathcal{V}, \mathcal{E})$  be the implication graph for  $F :: J$ .
- ▶ Let  $C \in F$  and suppose  $C|_J = []$ .
- ▶  $(\mathcal{V}', \mathcal{E}')$  is the **conflict graph for  $F :: J$  and  $C$**  if
  - ▶  $\mathcal{V}' = \mathcal{V} \cup \{0\}$  and
  - ▶  $\mathcal{E}' = \mathcal{E} \cup \{(L, 0) \mid L \in \mathcal{V} \text{ and } \bar{L} \in C\}$ .
- ▶  $0$  is called **conflict node**.
- ▶ Conflict graphs are acyclic.





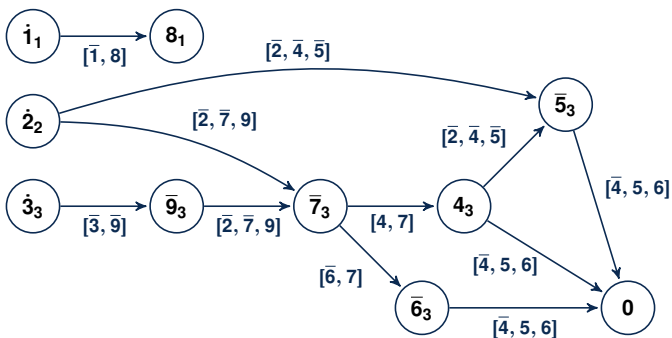
## The Conflict Graph of an Old Example

- ▶ Let  $F = \langle [1, 2], [2, \bar{3}], [\bar{2}, \bar{3}, 4], [\bar{1}, 3], [\bar{4}] \rangle$ .
- ▶ Consider  $J = (2, 3, \bar{1}, \bar{4})$ .
- ▶ **Note**  $[\bar{2}, \bar{3}, 4] \mid J = []$ .



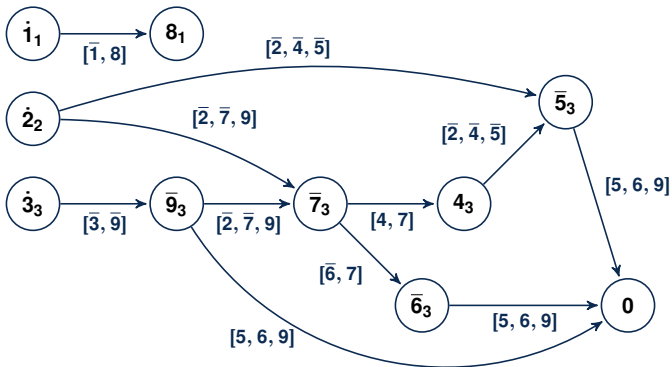
## Conflict Graphs – Another Example (1)

- ▶ Let  $F = \langle [\bar{1}, 8], [\bar{2}, \bar{4}, \bar{5}], [\bar{3}, \bar{9}], [\bar{2}, \bar{7}, 9], [4, 7], [\bar{6}, 7], [\bar{4}, 5, 6], [5, 6, 9] \rangle$ .
- ▶ Let  $J = (\bar{5}, \bar{6}, 4, \bar{7}, \bar{9}, \dot{3}, \dot{2}, 8, \dot{1})$ .
- ▶ **Note**  $[\bar{4}, 5, 6] \upharpoonright_J = []$ .



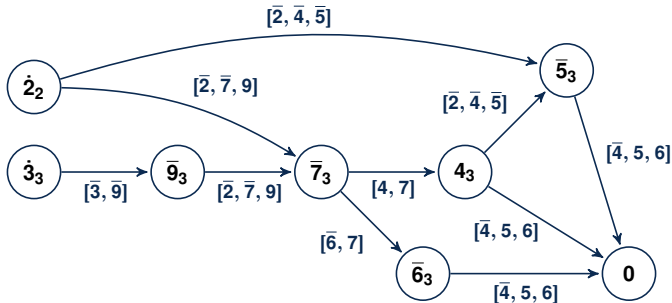
## Conflict Graphs – Another Example (2)

- ▶ Let  $F = \langle [\bar{1}, 8], [\bar{2}, \bar{4}, \bar{5}], [\bar{3}, \bar{9}], [\bar{2}, \bar{7}, 9], [4, 7], [\bar{6}, 7], [\bar{4}, 5, 6], [5, 6, 9] \rangle$ .
- ▶ Let  $J = (\bar{5}, \bar{6}, 4, \bar{7}, \bar{9}, \dot{3}, \dot{2}, 8, \dot{1})$ .
- ▶ **Note** There is another conflict clause  $[5, 6, 9] \mid_J = []$ .



## Reduced Conflict Graph

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a conflict graph.
- ▶ The **reduced conflict graph** of  $(\mathcal{V}, \mathcal{E})$  is the subgraph of  $(\mathcal{V}, \mathcal{E})$  containing all vertices which are connected to 0 and all edges between these vertices.

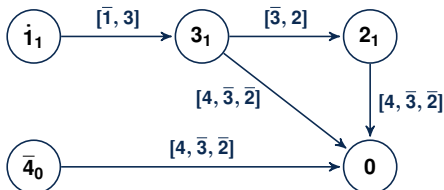


- ▶ In the sequel, we consider only reduced conflict graphs.



## Paths

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a reduced conflict graph.
- ▶ Let  $\mathcal{V}_S = \{L \in \mathcal{V} \mid L \text{ is a decision literal or its decision level is } 0\}$ .
- ▶ Let  $paths(\mathcal{V}, \mathcal{E})$  be the set of all directed paths from nodes in  $\mathcal{V}_S$  to 0 in  $(\mathcal{V}, \mathcal{E})$ .
- ▶ Consider

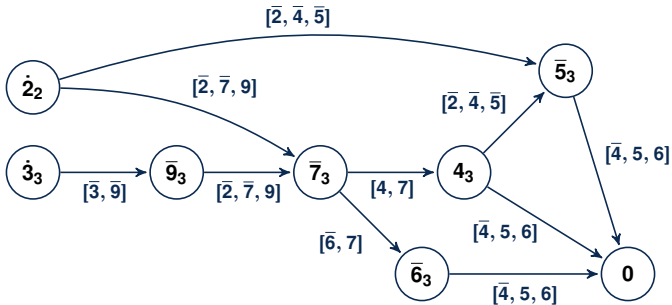


- ▶  $\mathcal{V}_S = \{1, \bar{4}\}$ .
- ▶  $paths(\mathcal{V}, \mathcal{E}) = \{(1, 3, 2, 0), (1, 3, 0), (\bar{4}, 0)\}$ .



## Paths – Another Example

► Consider



►  $\mathcal{V}_S = \{2, 3\}$

►  $paths(\mathcal{V}, \mathcal{E}) = \{ (2, \bar{5}, 0), (2, \bar{7}, 4, \bar{5}, 0), (2, \bar{7}, 4, 0), (2, \bar{7}, \bar{6}, 0), (3, \bar{9}, \bar{7}, 4, \bar{5}, 0), (3, \bar{9}, \bar{7}, 4, 0), (3, \bar{9}, \bar{7}, \bar{6}, 0) \}$



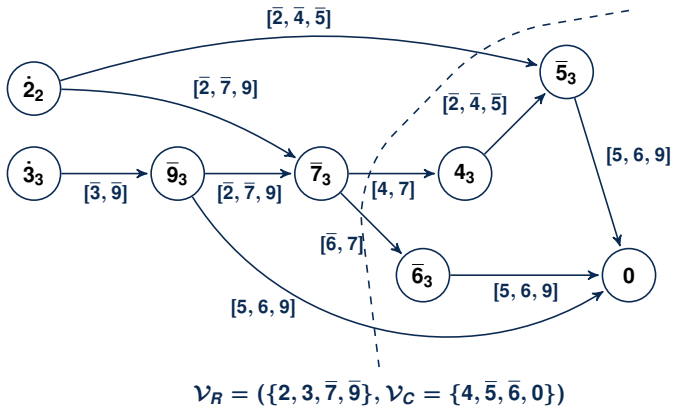
## Cuts

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a reduced conflict graph.
- ▶ A **cut**  $(\mathcal{V}_R, \mathcal{V}_C)$  through  $(\mathcal{V}, \mathcal{E})$  is a partition of  $\mathcal{V}$  into  $\mathcal{V}_R$  and  $\mathcal{V}_C$  such that
  - ▷  $\mathcal{V}_R \cap \mathcal{V}_C = \emptyset$ ,
  - ▷  $\mathcal{V}_R \cup \mathcal{V}_C = \mathcal{V}$ ,
  - ▷  $\mathcal{V}_S \subseteq \mathcal{V}_R$ ,
  - ▷  $\{0\} \subseteq \mathcal{V}_C$ , and
  - ▷ each  $p \in \text{paths}(\mathcal{V}, \mathcal{E})$  is partitioned into two subpaths  $p_R$  and  $p_C$  such that
    - ▶▶  $p_R$  and  $p_C$  have no vertex in common,
    - ▶▶  $p$  is obtained by adding an edge from the end of  $p_R$  to the start of  $p_C$ ,
    - ▶▶  $\mathcal{V}_R$  contains all vertices occurring in  $p_R$ , and
    - ▶▶  $\mathcal{V}_C$  contains all vertices occurring in  $p_C$ .



## Cuts – Example

► Consider





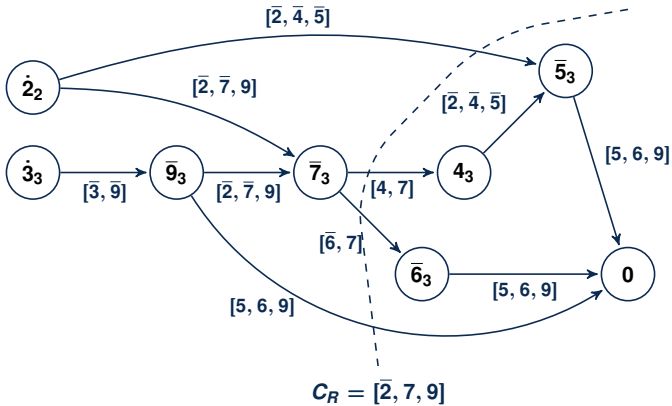
## Cut Clauses

- ▶ Let  $(\mathcal{V}_R, \mathcal{V}_C)$  be a cut through a reduced conflict graph  $(\mathcal{V}, \mathcal{E})$ .
- ▶ Let  $(\mathcal{V}_R, \mathcal{E}_R)$  be the subgraph of  $(\mathcal{V}, \mathcal{E})$  which consists only of the vertices  $\mathcal{V}_R$  and edges between elements of  $\mathcal{V}_R$ .
- ▶ Let  $\mathcal{V}'_R$  be the subset of  $\mathcal{V}_R$  containing all vertices, where an outgoing edge was cut by  $(\mathcal{V}_R, \mathcal{V}_C)$ .
- ▶ The **cut clause**  $C_R$  determined by  $(\mathcal{V}_R, \mathcal{V}_C)$  is the clause  $\overline{\mathcal{V}'_R}$ .



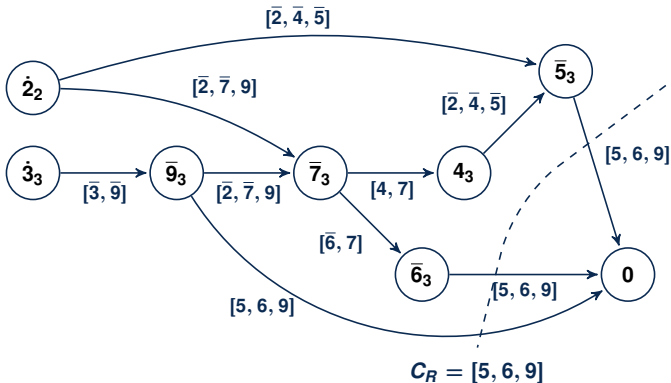
## Cut Clauses – Example

► Consider



## Initial Cuts

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a reduced conflict graph.
- ▶ Let  $(\mathcal{V}_R, \mathcal{V}_C)$  be a cut through  $(\mathcal{V}, \mathcal{E})$ .
- ▶  $(\mathcal{V}_R, \mathcal{V}_C)$  is an **initial cut** if  $\mathcal{V}_C = \{0\}$ .



- ▶ **Note** The cut clause is the conflict clause.

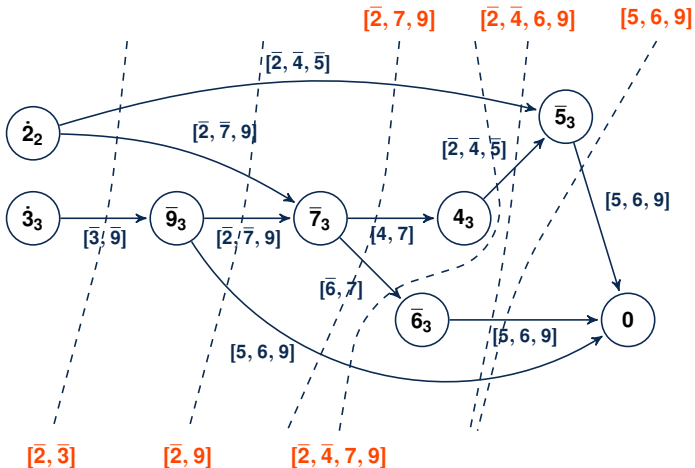


## Shifted Cuts

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a reduced conflict graph.
- ▶ **Recall**  $\mathcal{V}_S = \{L \in \mathcal{V} \mid L \text{ is a decision literal or its decision level is } 0\}$ .
- ▶ Let  $(\mathcal{V}_R, \mathcal{V}_C)$  be a cut through  $(\mathcal{V}, \mathcal{E})$ .
- ▶ Let  $(\mathcal{V}_R, \mathcal{E}_R)$  be the subgraph of  $(\mathcal{V}, \mathcal{E})$  which consists only of the vertices  $\mathcal{V}_R$  and edges between elements of  $\mathcal{V}_R$ .
- ▶ Let  $L \in \mathcal{V}_R \setminus \mathcal{V}_S$  such that the outdegree of  $L$  in  $(\mathcal{V}_R, \mathcal{E}_R)$  is 0.
- ▶  $(\mathcal{V}_R \setminus \{L\}, \mathcal{V}_C \cup \{L\})$  is the cut obtained from  $(\mathcal{V}_R, \mathcal{V}_C)$  by **shifting**  $L$ .



## Shifted Cuts – Examples



## Corresponding Linear Resolution Derivation

► We obtain

1	$[\bar{3}, \bar{9}]$	relevant clause
2	$[\bar{2}, \bar{7}, 9]$	relevant clause
3	$[4, \bar{7}]$	relevant clause
4	$[\bar{6}, \bar{7}]$	relevant clause
5	$[\bar{2}, \bar{4}, \bar{5}]$	relevant clause
6	$[5, \bar{6}, 9]$	conflict clause
7	$[\bar{2}, \bar{4}, 6, 9]$	res(6,5)
8	$[\bar{2}, \bar{4}, 7, 9]$	res(7,4)
9	$[\bar{2}, \bar{7}, 9]$	res(8,3)
10	$[\bar{2}, \bar{9}]$	res(9,2)
11	$[\bar{2}, \bar{3}]$	res(10,1)

- **Observation** Cut-clauses corresponding to a cut, which was generated by a sequence of shifts from the initial cut, can also be generated by linear resolution derivations from the conflict clause using relevant clauses.



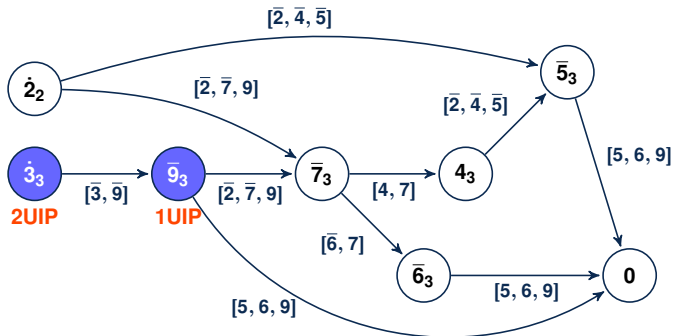
## Unique Implication Points

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a (reduced) conflict graph.
- ▶ Let  $L^* \in \mathcal{V}$  be the decision literal with the highest decision level in  $\mathcal{V}$ .
  - ▷  $L^*$  is unique.
- ▶ Let  $paths^*(\mathcal{V}, \mathcal{E})$  be the set of all directed paths from  $L^*$  to 0 in  $(\mathcal{V}, \mathcal{E})$ .
  - ▷  $paths^*(\mathcal{V}, \mathcal{E}) \subseteq paths(\mathcal{V}, \mathcal{E})$ .
- ▶  $L \in \mathcal{V}$  is a **unique implication point (UIP)** if it occurs in all paths in  $paths^*(\mathcal{V}, \mathcal{E})$ .
  - ▷  $L^*$  is always a UIP.
  - ▷ The decision level of an UIP is equal to the decision level of  $L^*$ .
- ▶ The UIPs of a reduced conflict graph can be ordered:
  - ▷ The **1UIP** is the UIP which is closest to 0.
  - ▷ The **nUIP** is the UIP which is  $n$ -closest to 0.



## Unique Implication Points – Example

► Consider





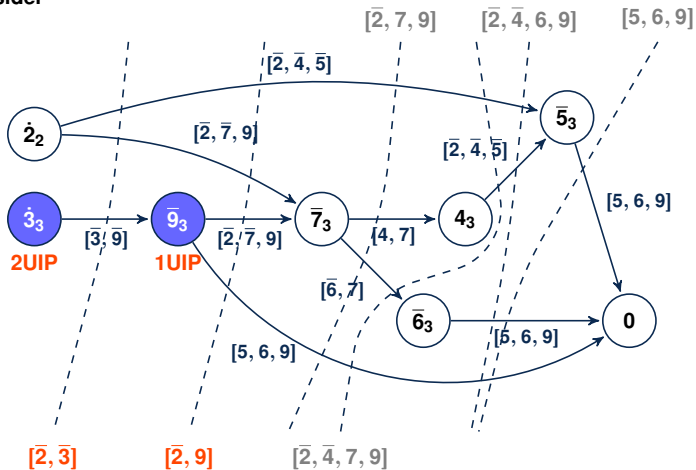
## UIP Clauses

- ▶ Let  $(\mathcal{V}, \mathcal{E})$  be a (reduced) conflict graph.
- ▶ Let  $(\mathcal{V}_R, \mathcal{V}_C)$  be a cut through  $(\mathcal{V}, \mathcal{E})$ .
- ▶ Let  $C$  be the cut clause determined by  $(\mathcal{V}_R, \mathcal{V}_C)$ .
- ▶  $C$  is a **UIP clause** if
  - ▷  $C$  contains a literal  $L$  such that  $\bar{L}$  is a UIP in  $(\mathcal{V}, \mathcal{E})$  and
  - ▷ there is no literal in  $C$  whose level is identical to the level of  $L$ .
- ▶  $C$  is a **nUIP clause** if
  - ▷ if  $C$  is a UIP clause and
  - ▷ if  $L \in C$  such that  $\bar{L}$  is a UIP in  $(\mathcal{V}, \mathcal{E})$  then  $\bar{L}$  is the nUIP in  $(\mathcal{V}, \mathcal{E})$ .



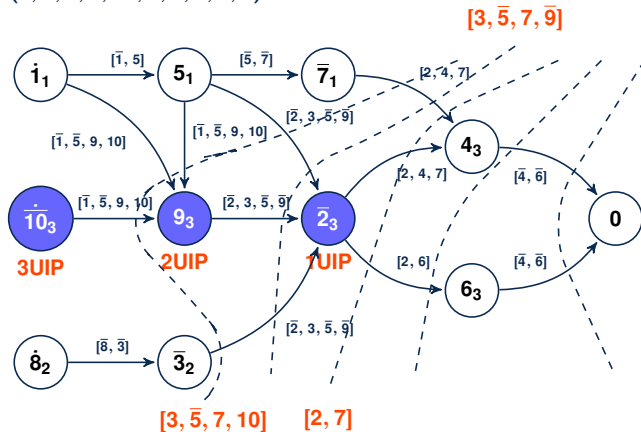
## UIP Clauses – Examples

► Consider



## Yet Another Example

- ▶ Let  $F = \langle [2, 6], [\bar{8}, \bar{3}], [\bar{4}, \bar{6}], [\bar{1}, \bar{5}, 9, 10], [2, 4, 7], [\bar{1}, 5], [2, 3, \bar{5}, \bar{9}], [\bar{5}, \bar{7}] \rangle$ .
- ▶ Let  $J = (6, 4, \bar{2}, 9, \bar{10}, \bar{3}, \bar{8}, \bar{7}, 5, \bar{1})$ .



## UIP Clauses – Remarks

- ▶ **UIP clauses are not unique.**
  - ▷ In the previous example,  $[2, 7]$  and  $[2, \bar{5}]$  are 1UIP clauses.
- ▶ **Let**
  - ▷  $C$  be a UIP clause,
  - ▷  $m$  the level of the literal whose complement is UIP, and
  - ▷  $n$  the next highest level assigned to some literal in  $C$ .

An application of CDBL to  $F :: J$  using  $C$  removes all literals from  $J$  whose level is higher than  $n$ .

- ▶ **Most modern SAT-solvers learn 1UIP clauses.**
- ▶ **If several 1UIP clauses exist, then often the shortest one is preferred.**



## Conflict Graphs vs Resolution Derivations

- Consider the following linear resolution derivation:

1	$[\bar{1}, 5]$	relevant clause
2	$[\bar{5}, \bar{7}]$	relevant clause
3	$[\bar{8}, \bar{3}]$	relevant clause
4	$[\bar{1}, \bar{5}, 9, 10]$	relevant clause
5	$[\bar{2}, 3, \bar{5}, \bar{9}]$	relevant clause
6	$[2, 4, 7]$	relevant clause
7	$[2, 6]$	relevant clause
8	$[\bar{4}, \bar{6}]$	conflict clause
9	$[2, \bar{4}]$	res(8,7)
10	$[2, 7]$	res(9,6)
11	$[2, 5]$	res(10,2)
12	$[\bar{1}, 2]$	res(12,1)

- Note  $[\bar{1}, 2]$  is not a cut-clause!
- Pipatsriawat, Darwiche: On Modern Clause-Learning Satisfiability Solvers. Journal Automated Reasoning 44, 277-301:2010.

