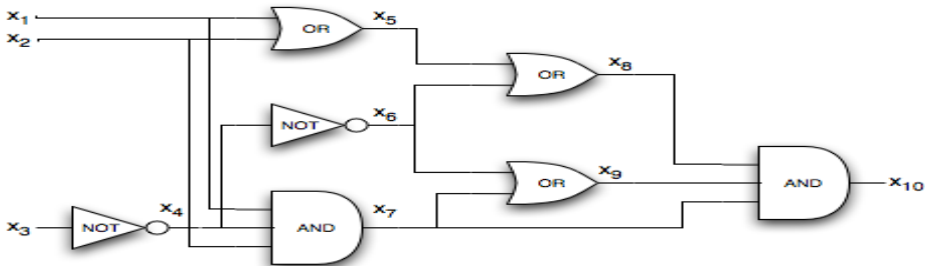


Practical SAT Solving

Lecture 6

Carsten Sinz, Tomáš Balyo, Markus Iser | May 30, 2016

INSTITUTE FOR THEORETICAL COMPUTER SCIENCE



- Details on implementing DPLL
 - Decision Heuristics
 - Restarts
 - Implementing Unit Propagation

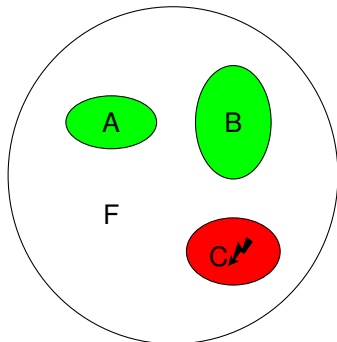
- More Details on implementing DPLL
 - Phase Saving
- Towards CDCL
 - Review: Resolution
 - Non-Chronological Backtracking (Backjumping)
 - Implication Graph
 - Clause Learning

Phase Saving

First implemented in RSAT (2006)

<http://reasoning.cs.ucla.edu/rsat/>

- Observation: Frequent *Restarts* decrease performance on some SAT instances
- Goal: Cache partial solutions to subsets of the formula and reuse them after a restart
- Idea: Remember last assignment of each variable and use it *first* in branching
- Result: *Phase Saving* stabilizes positive effect of restarts; best results in combination with *non-chronological backtracking* (later in this lecture)



Example: *A* and *B* are satisfied, search works on component *C*

- General deduction rule for clauses: $(A \vee I), (B \vee \bar{I}) \implies (A \vee B)$
- Can be used to extend DPLL Search by adding resolved clauses to the input formula
- Special Case: Unit Resolution (central part of DPLL Search)
- Resolution Operator: $C \circ_I D$ denotes the resolvent of clauses C and D on literal I

Trail

The partial assignment that represents the current path in the search tree

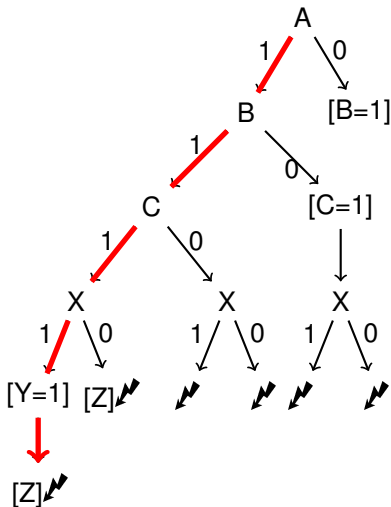
Decision Level

Starting with decision level 0 every branching step increases the decision level by 1. Note: Unit-propagation does not change the decision level.

Conflicting Clause

Clause with all literals assigned to 0 under the current trail (empty under unit resolution)

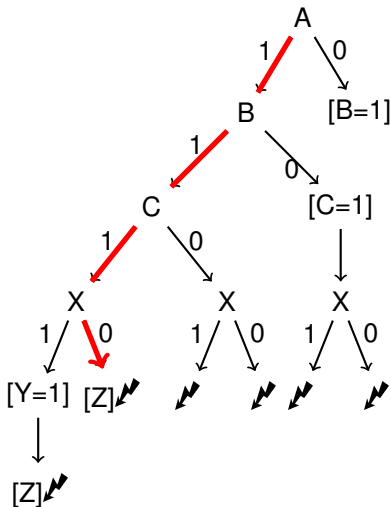
DPLL: Chronological Backtracking



$$F = \{ \{A, B\}, \\ \{B, C\}, \\ \{\neg A, \neg X, Y\}, \\ \{\neg A, X, Z\}, \\ \{\neg A, \neg Y, Z\}, \\ \{\neg A, X, \neg Z\}, \\ \{\neg A, \neg Y, \neg Z\} \}$$

- Trail: A, B, C, X, Y, Z
- Conflicting Clause:
 $\{\neg A, \neg Y, \neg Z\}$

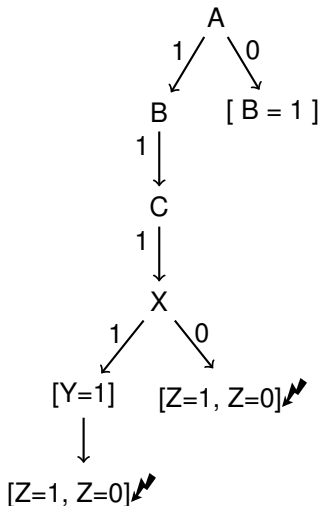
DPLL: Chronological Backtracking



$$F = \{ \{A, B\}, \\ \{B, C\}, \\ \{\neg A, \neg X, Y\}, \\ \{\neg A, X, Z\}, \\ \{\neg A, \neg Y, Z\}, \\ \{\neg A, X, \neg Z\}, \\ \{\neg A, \neg Y, \neg Z\} \}$$

- Trail: $A, B, C, \neg X, Z$
- Conflicting Clause:
 $\{\neg A, X, \neg Z\}$

Non-Chronological Backtracking



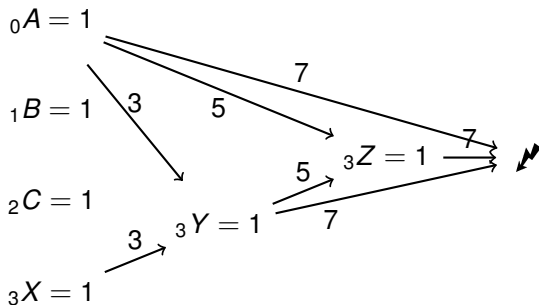
- The first two conflicting clauses $\{\neg A, \neg Y, \neg Z\}$, $\{\neg A, X, \neg Z\}$ contain only a fraction of the assignments on the trail
- Assignments to **B** and **C** obviously play no role in the present conflicting state and we could immediately backtrack to flip the assignment to **A**
- How to find out which assignments on the trail are relevant for the actual conflict?

Implication Graph

The Implication Graph $G_I = (V \cup \{\blacksquare\}, E)$ (for a given formula F and a trail T ending in a conflict) is a directed acyclic graph (DAG) where each vertex $v \in V$ is a tuple $(x = b, d)$ consisting of an assignment $x = b$ on the trail and its decision level d . There is an additional vertex \blacksquare (with no assignment and decision level) indicating a conflicting assignment. For each clause $C = (l_1, \dots, l_k, u)$ that became unit and caused a unit propagation of u there are edges $e \in E$ connecting the assignment of literal l_i to the assignment of unit u . Similarly, for a clause C , in which all literals are assigned false, there are edges from each literal in C to \blacksquare .

Example: Implication Graph

The graph shows the implication graph for the conflicting state under the trail A, B, C, X, Y, Z . The edge labels denote clauses, node labels indicate an assignment and its decision level.

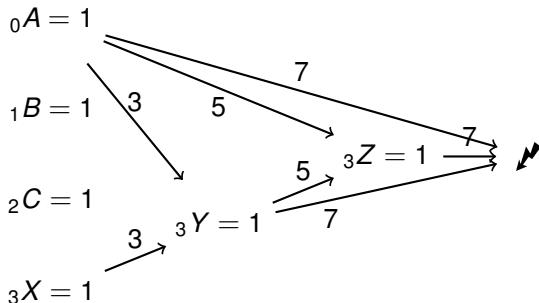


$$F = \left\{ \begin{array}{l} \{A, B\}, \quad =: 1 \\ \{B, C\}, \quad =: 2 \\ \{\neg A, \neg X, Y\}, \quad =: 3 \\ \{\neg A, X, Z\}, \quad =: 4 \\ \{\neg A, \neg Y, Z\}, \quad =: 5 \\ \{\neg A, X, \neg Z\}, \quad =: 6 \\ \{\neg A, \neg Y, \neg Z\} \quad =: 7 \end{array} \right.$$

Implication Graph

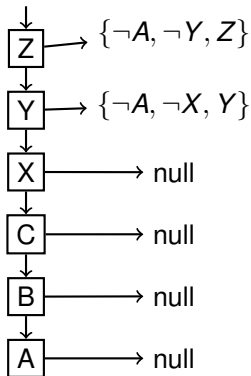
In the implication graph

- the sink is always the conflicting assignment
- the sources are the decision literals that take part in the conflict
- we can use it to detect the *reasons* for a conflict



Implication Graph: Implementation

Trail with conflicting clause $\{\neg A, \neg Y, \neg Z\}$:



- for each assignment store a pointer to the reason clause, the one that has become unit
- decision literals store a null pointer
- with the clause pointers and the trail we can trace all the implications back to their sources

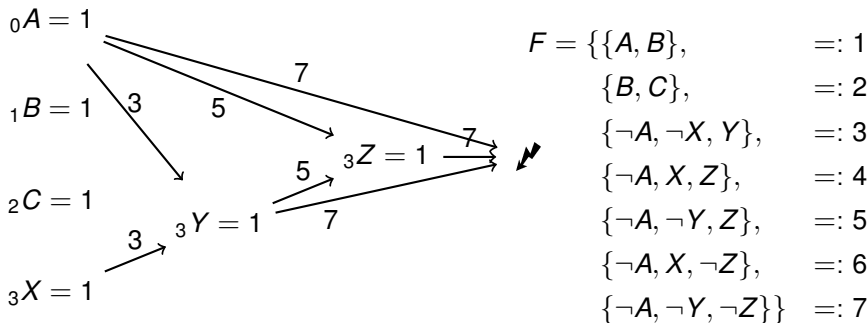
Back-jumping does not fully avoid the occurrence of the same conflict, the same (partial) assignments may generate the same conflict in another part of the search tree.

- Idea: store no-good constraint like in CSP (No-Good-Learning)
- Great in SAT: no-good is clause!
- generate conflict clauses and add them to CNF
- the literals contributing to a conflict form a partial assignment and its negation is a clause (implied by the original CNF)
- adding this clause avoids the conflicting partial assignment

Conflict Analysis: Example

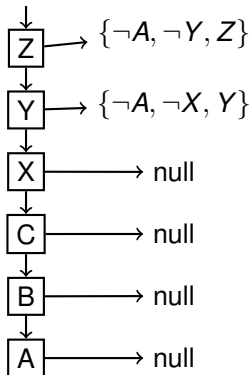
In the first conflict of the previous example we can learn the clause $\{\neg A, \neg X\}$ in order to prevent the solver to pick the same partial assignment again. This can also be expressed as the following sequence of resolution steps along the edges in the implication graph:

$$C = (7 \circ_Z 5) \circ_Y 3$$



Conflict Clause

Trail with conflicting clause
 $\{\neg A, \neg Y, \neg Z\}$:



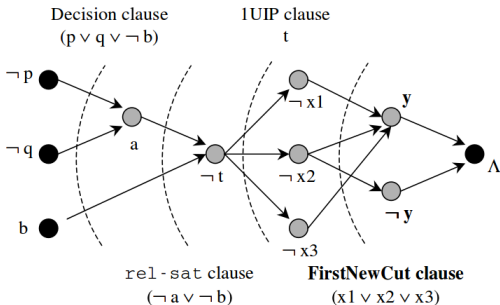
Given the trail the conflict clause (learned clause) can be generated by subsequent resolution, starting with the conflict clause $\{\neg A, \neg Y, \neg Z\}$ and the first reason clause on the trail $\{\neg A, \neg Y, Z\}$ we can resolve the intermediate clause $\{\neg A, \neg Y\}$. An additional resolution step results in the conflict clause $C = \{\neg A, \neg X\}$.

Properties of conflict clause C:

- $F \models C$
- $F \cup \neg C \vdash_{UP} \perp$
- $D \notin F, \forall D \subseteq C$

Unique Implication Points (UIP)

In general: several possibilities to learn a clause from an implication graph. A learned clause can be resolved for every cut in the the implication graph.



- UIP is a dominator in the implication graph
- A node v is a dominator (for \swarrow), if all paths from the last decision to Δ contain v
- FirstUIP: “first” dominator (seen from conflict side)

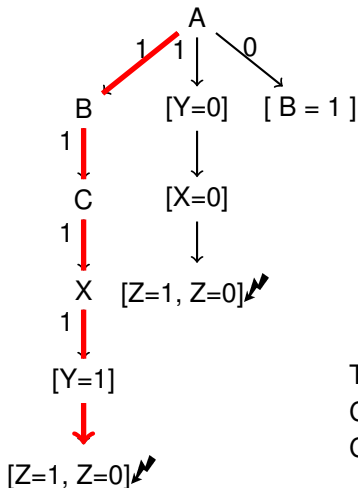
Graphic from Beame et al.: “Understanding the Power of Clause Learning”

Common Implementation (1UIP Learning Scheme)

- FirstUIP-clause: resolve conflicting clause and reason clauses until only a single literal of the current decision level remains
- Advantage: Stopping at a UIP always leads to an asserting clause. Algorithm becomes easier: backtrack until clause becomes asserting
- Undo all decisions until that level where the learned clause becomes asserting. The assertion level is the second highest level in a conflict clause.

Backtracking with Clause Learning

In our previous example the conflict clause (1UIP) with backtracking to assertion level changes the decision tree like this:



$$F = \{ \{A, B\}, \{B, C\}, \\ \{ \neg A, \neg X, Y \}, \\ \{ \neg A, X, Z \}, \\ \{ \neg A, \neg Y, Z \}, \\ \{ \neg A, X, \neg Z \}, \\ \{ \neg A, \neg Y, \neg Z \} \}$$

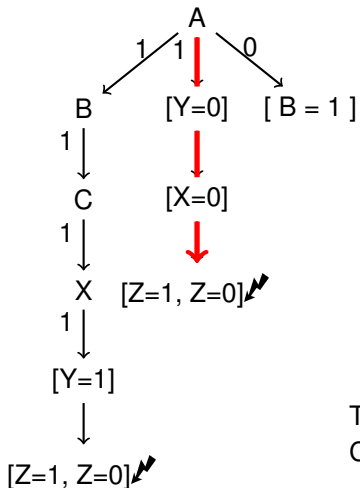
Trail: A, B, C, X, Y, Z

Conflicting Clause: $\{ \neg A, \neg Y, \neg Z \}$

Conflict Clause (1UIP): $\{ \neg A, \neg Y \}$

Backtracking with Clause Learning

In our previous example the conflict clause (1UIP) with backtracking to assertion level changes the decision tree like this:



$$F = \{ \{A, B\}, \{B, C\}, \\ \{ \neg A, \neg X, Y \}, \\ \{ \neg A, X, Z \}, \\ \{ \neg A, \neg Y, Z \}, \\ \{ \neg A, X, \neg Z \}, \\ \{ \neg A, \neg Y, \neg Z \} \\ \{ \neg A, \neg Y \} \}$$

Trail: $A, \neg Y, \neg X, Z$

Conflicting Clause: $\{ \neg A, X, \neg Z \}$

Without restrictions the clause database grows exponentially. Besides the risk of running out of memory this also slows down unit propagation. Clause database is regularly cleaned up. Quality measures for clauses (e.g. clause size, clause activity, LBD value)

Data: Formula F

Result: SAT / UNSAT

```
1  $dl \leftarrow 0$ 
2  $V \leftarrow \emptyset$ 
3 if unit propagation ( $F, V$ ) == CONFLICT then
4   return UNSAT
5 while not all variables assigned do
6    $(x, l) \leftarrow$  pick branching literal
7    $dl \leftarrow dl + 1$ 
8    $V \leftarrow \cup(x, l)$ 
9   if unit propagation ( $F, V$ ) == CONFLICT then
10     $bl \leftarrow$  analyze conflict
11    if  $bl < 0$  then
12      return UNSAT
13    else
14      backtrack to  $bl$ 
15       $dl \leftarrow bl$ 
16 return SAT
```

With Clause Learning a complete resolution refutation can be derived with CDCL

- proof can serve as a certificate for validating the correctness of the SAT solver
- resolution refutations based on clause learning find key practical applications (e.g. model checking)
- can help to determine minimally unsatisfiable subsets in an unsatisfiable formula

NEWS: Boolean Pythagorean triples problem was recently cracked using SAT technology

The problem asks whether it is possible to colour each positive integer either red or blue, so that no trio of integers a , b and c that satisfy Pythagoras famous equation $a^2 + b^2 = c^2$ are all the same colour.

- Solutions exist for numbers up to 7824
- Proof that from 7825 onwards there is no such coloring
- "Largest math proof" ever built (200 TB)
- Paper will be presented at the SAT conference 2016.

- Phase Saving: “A Lightweight Component Caching Scheme for Satisfiability Solvers” (Pipatsrisawat and Darwiche)
- Clause Learning: “Understanding the Power of Clause Learning” (Beame et al.)
- Two-Coloring of Pythagorean Triplets:
`http://www.nature.com/news/
two-hundred-terabyte-maths-proof-is-largest-ever-1.
19990`