

Mini-tutorial on conflict-driven clause learning solvers

Marijn J. H. Heule

The University of Texas at Austin

20/01/2014 @ BIRS

The Satisfiability (SAT) problem

$$\begin{aligned} & (x_5 \vee x_8 \vee \bar{x}_2) \wedge (x_2 \vee \bar{x}_1 \vee \bar{x}_3) \wedge (\bar{x}_8 \vee \bar{x}_3 \vee \bar{x}_7) \wedge (\bar{x}_5 \vee x_3 \vee x_8) \wedge \\ & (\bar{x}_6 \vee \bar{x}_1 \vee \bar{x}_5) \wedge (x_8 \vee \bar{x}_9 \vee x_3) \wedge (x_2 \vee x_1 \vee x_3) \wedge (\bar{x}_1 \vee x_8 \vee x_4) \wedge \\ & (\bar{x}_9 \vee \bar{x}_6 \vee x_8) \wedge (x_8 \vee x_3 \vee \bar{x}_9) \wedge (x_9 \vee \bar{x}_3 \vee x_8) \wedge (x_6 \vee \bar{x}_9 \vee x_5) \wedge \\ & (x_2 \vee \bar{x}_3 \vee \bar{x}_8) \wedge (x_8 \vee \bar{x}_6 \vee \bar{x}_3) \wedge (x_8 \vee \bar{x}_3 \vee \bar{x}_1) \wedge (\bar{x}_8 \vee x_6 \vee \bar{x}_2) \wedge \\ & (x_7 \vee x_9 \vee \bar{x}_2) \wedge (x_8 \vee \bar{x}_9 \vee x_2) \wedge (\bar{x}_1 \vee \bar{x}_9 \vee x_4) \wedge (x_8 \vee x_1 \vee \bar{x}_2) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_6) \wedge (\bar{x}_1 \vee \bar{x}_7 \vee x_5) \wedge (\bar{x}_7 \vee x_1 \vee x_6) \wedge (\bar{x}_5 \vee x_4 \vee \bar{x}_6) \wedge \\ & (\bar{x}_4 \vee x_9 \vee \bar{x}_8) \wedge (x_2 \vee x_9 \vee x_1) \wedge (x_5 \vee \bar{x}_7 \vee x_1) \wedge (\bar{x}_7 \vee \bar{x}_9 \vee \bar{x}_6) \wedge \\ & (x_2 \vee x_5 \vee x_4) \wedge (x_8 \vee \bar{x}_4 \vee x_5) \wedge (x_5 \vee x_9 \vee x_3) \wedge (\bar{x}_5 \vee \bar{x}_7 \vee x_9) \wedge \\ & (x_2 \vee \bar{x}_8 \vee x_1) \wedge (\bar{x}_7 \vee x_1 \vee x_5) \wedge (x_1 \vee x_4 \vee x_3) \wedge (x_1 \vee \bar{x}_9 \vee \bar{x}_4) \wedge \\ & (x_3 \vee x_5 \vee x_6) \wedge (\bar{x}_6 \vee x_3 \vee \bar{x}_9) \wedge (\bar{x}_7 \vee x_5 \vee x_9) \wedge (x_7 \vee \bar{x}_5 \vee \bar{x}_2) \wedge \\ & (x_4 \vee x_7 \vee x_3) \wedge (x_4 \vee \bar{x}_9 \vee \bar{x}_7) \wedge (x_5 \vee \bar{x}_1 \vee x_7) \wedge (x_5 \vee \bar{x}_1 \vee x_7) \wedge \\ & (x_6 \vee x_7 \vee \bar{x}_3) \wedge (\bar{x}_8 \vee \bar{x}_6 \vee \bar{x}_7) \wedge (x_6 \vee x_2 \vee x_3) \wedge (\bar{x}_8 \vee x_2 \vee x_5) \end{aligned}$$

Does there exist an assignment satisfying all clauses?

Search for a satisfying assignment (or proof none exists)

$$\begin{aligned} & (x_5 \vee x_8 \vee \bar{x}_2) \wedge (x_2 \vee \bar{x}_1 \vee \bar{x}_3) \wedge (\bar{x}_8 \vee \bar{x}_3 \vee \bar{x}_7) \wedge (\bar{x}_5 \vee x_3 \vee x_8) \wedge \\ & (\bar{x}_6 \vee \bar{x}_1 \vee \bar{x}_5) \wedge (x_8 \vee \bar{x}_9 \vee x_3) \wedge (x_2 \vee x_1 \vee x_3) \wedge (\bar{x}_1 \vee x_8 \vee x_4) \wedge \\ & (\bar{x}_9 \vee \bar{x}_6 \vee x_8) \wedge (x_8 \vee x_3 \vee \bar{x}_9) \wedge (x_9 \vee \bar{x}_3 \vee x_8) \wedge (x_6 \vee \bar{x}_9 \vee x_5) \wedge \\ & (x_2 \vee \bar{x}_3 \vee \bar{x}_8) \wedge (x_8 \vee \bar{x}_6 \vee \bar{x}_3) \wedge (x_8 \vee \bar{x}_3 \vee \bar{x}_1) \wedge (\bar{x}_8 \vee x_6 \vee \bar{x}_2) \wedge \\ & (x_7 \vee x_9 \vee \bar{x}_2) \wedge (x_8 \vee \bar{x}_9 \vee x_2) \wedge (\bar{x}_1 \vee \bar{x}_9 \vee x_4) \wedge (x_8 \vee x_1 \vee \bar{x}_2) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_6) \wedge (\bar{x}_1 \vee \bar{x}_7 \vee x_5) \wedge (\bar{x}_7 \vee x_1 \vee x_6) \wedge (\bar{x}_5 \vee x_4 \vee \bar{x}_6) \wedge \\ & (\bar{x}_4 \vee x_9 \vee \bar{x}_8) \wedge (x_2 \vee x_9 \vee x_1) \wedge (x_5 \vee \bar{x}_7 \vee x_1) \wedge (\bar{x}_7 \vee \bar{x}_9 \vee \bar{x}_6) \wedge \\ & (x_2 \vee x_5 \vee x_4) \wedge (x_8 \vee \bar{x}_4 \vee x_5) \wedge (x_5 \vee x_9 \vee x_3) \wedge (\bar{x}_5 \vee \bar{x}_7 \vee x_9) \wedge \\ & (x_2 \vee \bar{x}_8 \vee x_1) \wedge (\bar{x}_7 \vee x_1 \vee x_5) \wedge (x_1 \vee x_4 \vee x_3) \wedge (x_1 \vee \bar{x}_9 \vee \bar{x}_4) \wedge \\ & (x_3 \vee x_5 \vee x_6) \wedge (\bar{x}_6 \vee x_3 \vee \bar{x}_9) \wedge (\bar{x}_7 \vee x_5 \vee x_9) \wedge (x_7 \vee \bar{x}_5 \vee \bar{x}_2) \wedge \\ & (x_4 \vee x_7 \vee x_3) \wedge (x_4 \vee \bar{x}_9 \vee \bar{x}_7) \wedge (x_5 \vee \bar{x}_1 \vee x_7) \wedge (x_5 \vee \bar{x}_1 \vee x_7) \wedge \\ & (x_6 \vee x_7 \vee \bar{x}_3) \wedge (\bar{x}_8 \vee \bar{x}_6 \vee \bar{x}_7) \wedge (x_6 \vee x_2 \vee x_3) \wedge (\bar{x}_8 \vee x_2 \vee x_5) \end{aligned}$$

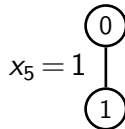
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} & (x_1 \vee x_4) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ & \mathcal{F}_{\text{extra}} \end{aligned}$$

0

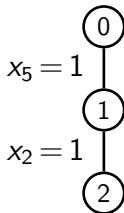
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} & (x_1 \vee x_4) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ & \mathcal{F}_{\text{extra}} \end{aligned}$$



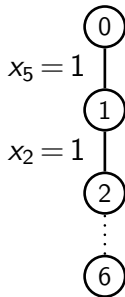
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} & (x_1 \vee x_4) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ & \mathcal{F}_{\text{extra}} \end{aligned}$$



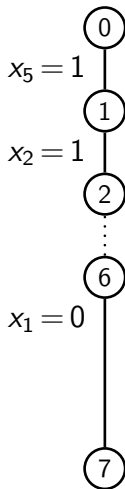
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} & (x_1 \vee x_4) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ & \mathcal{F}_{\text{extra}} \end{aligned}$$



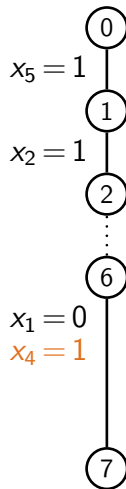
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} &(\bar{x}_1 \vee x_4) \wedge \\ &(x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ &(\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ &\mathcal{F}_{\text{extra}} \end{aligned}$$



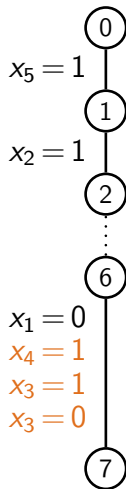
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} &(\bar{x}_1 \vee x_4) \wedge \\ &(x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ &(\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ &\mathcal{F}_{\text{extra}} \end{aligned}$$



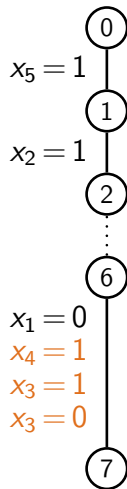
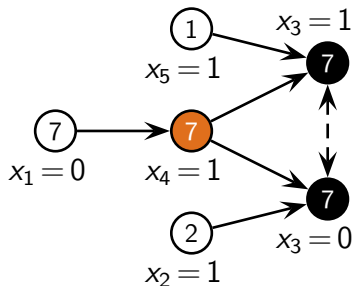
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} &(\bar{x}_1 \vee x_4) \wedge \\ &(x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ &(\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ &\mathcal{F}_{\text{extra}} \end{aligned}$$



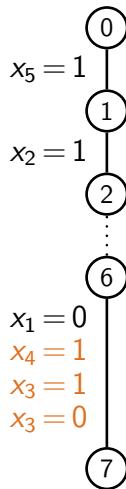
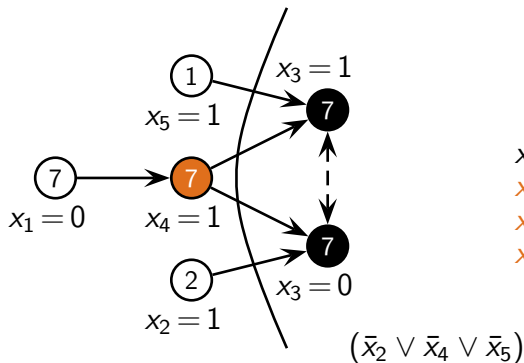
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} & (x_1 \vee x_4) \wedge \\ & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\ & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\ & \mathcal{F}_{\text{extra}} \end{aligned}$$



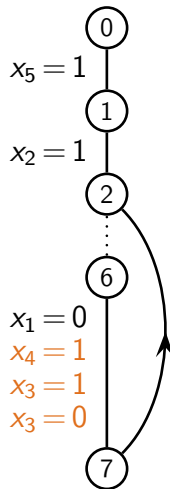
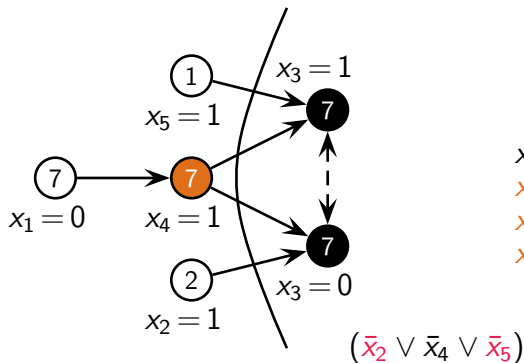
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned} & (\mathbf{x}_1 \vee \mathbf{x}_4) \wedge \\ & (\mathbf{x}_3 \vee \bar{\mathbf{x}}_4 \vee \bar{\mathbf{x}}_5) \wedge \\ & (\bar{\mathbf{x}}_3 \vee \bar{\mathbf{x}}_2 \vee \bar{\mathbf{x}}_4) \wedge \\ & \mathcal{F}_{\text{extra}} \end{aligned}$$



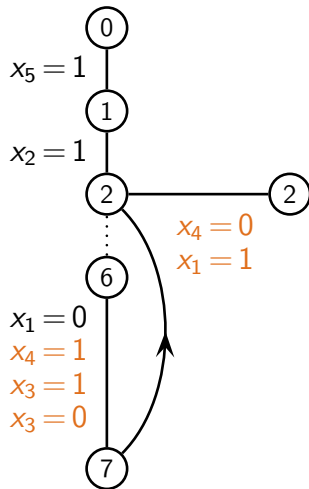
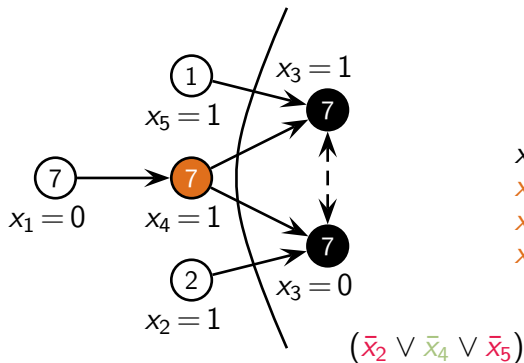
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned}
 &(x_1 \vee x_4) \wedge \\
 &(x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\
 &(\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\
 &\mathcal{F}_{\text{extra}}
 \end{aligned}$$



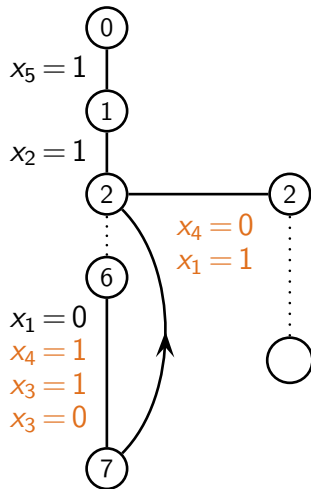
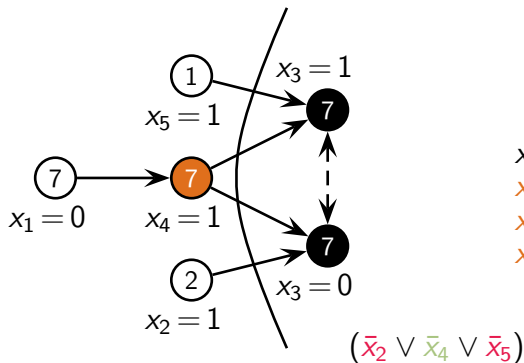
Conflict-driven SAT solvers: Search and Analysis

$$\begin{aligned}
 & (x_1 \vee x_4) \wedge \\
 & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\
 & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\
 & \mathcal{F}_{\text{extra}}
 \end{aligned}$$



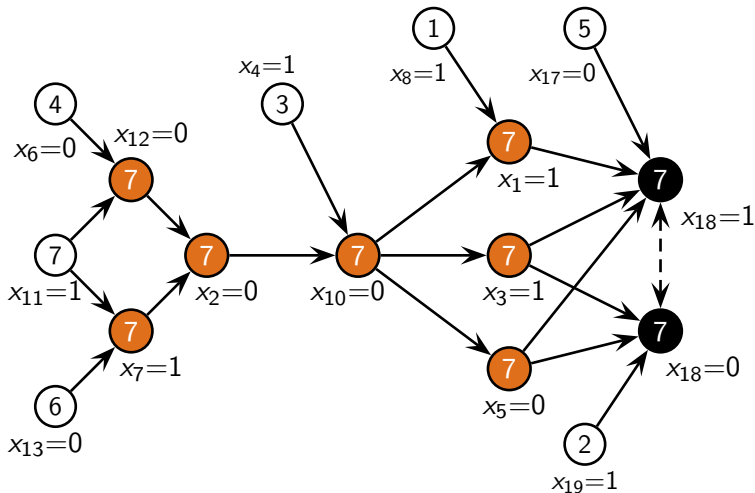
Conflict-driven SAT solvers: Search and Analysis

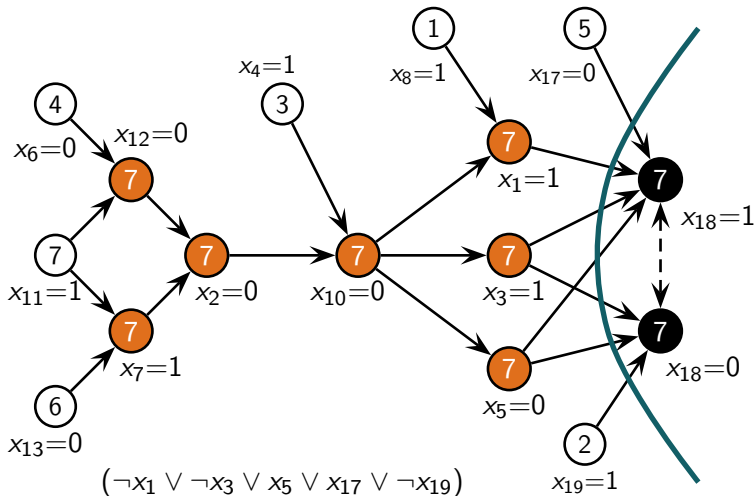
$$\begin{aligned}
 & (x_1 \vee x_4) \wedge \\
 & (x_3 \vee \bar{x}_4 \vee \bar{x}_5) \wedge \\
 & (\bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_4) \wedge \\
 & \mathcal{F}_{\text{extra}}
 \end{aligned}$$

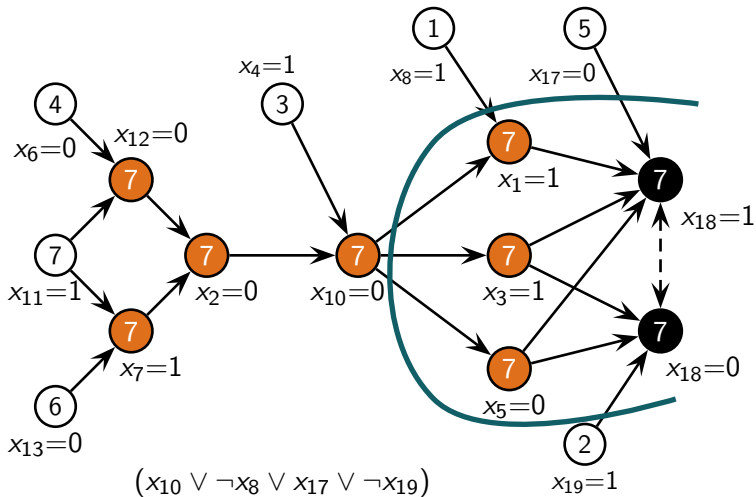


Conflict-driven SAT solvers: Pseudo-code

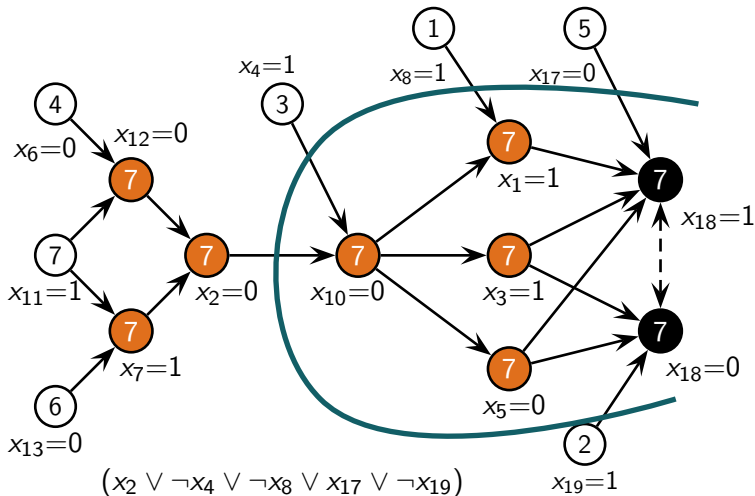
```
1: while TRUE do
2:    $l_{\text{decision}} := \text{GETDECISIONLITERAL}()$ 
3:   If no  $l_{\text{decision}}$  then return satisfiable
4:    $\mathcal{F} := \text{SIMPLIFY}(\mathcal{F}(l_{\text{decision}} \leftarrow 1))$ 
5:   while  $\mathcal{F}$  contains  $C_{\text{falsified}}$  do
6:      $C_{\text{conflict}} := \text{ANALYZECONFLICT}(C_{\text{falsified}})$ 
7:     If  $C_{\text{conflict}} = \emptyset$  then return unsatisfiable
8:      $\text{BACKTRACK}(C_{\text{conflict}})$ 
9:      $\mathcal{F} := \text{SIMPLIFY}(\mathcal{F} \cup \{C_{\text{conflict}}\})$ 
10:  end while
11: end while
```

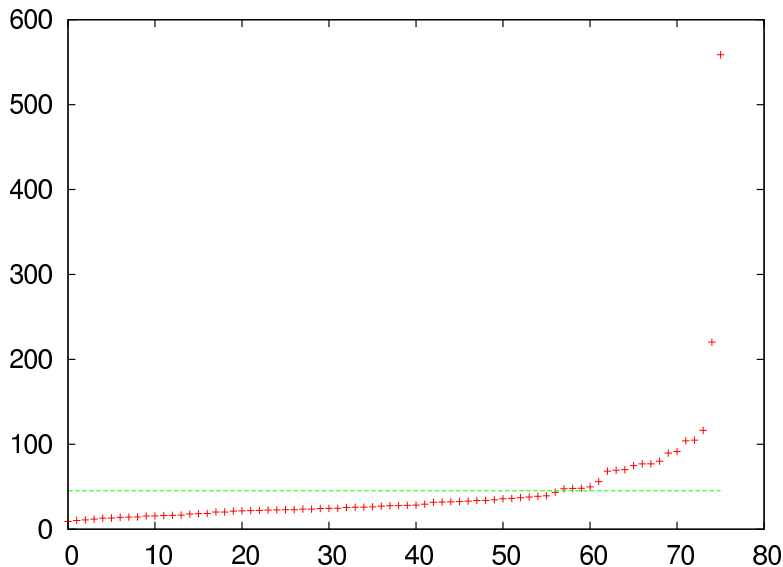




$(x_{10} \vee \neg x_8 \vee x_{17} \vee \neg x_{19})$
first unique implication point



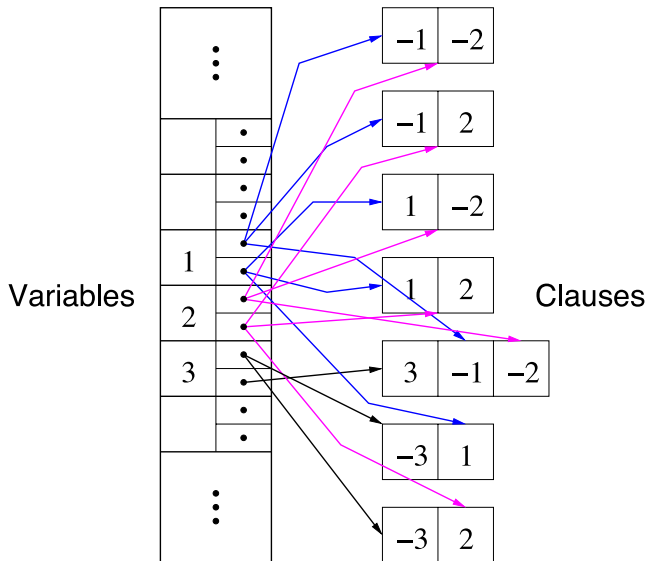
Average Learned Clause Length



Data-structures

Watch pointers

Simple data structure for unit propagation



Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = *, x_2 = *, x_3 = *, x_4 = *, x_5 = *, x_6 = *\}$$

$\neg x_1$	x_2	$\neg x_3$	$\neg x_5$	x_6
------------	-------	------------	------------	-------

x_1	$\neg x_3$	x_4	$\neg x_5$	$\neg x_6$
-------	------------	-------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

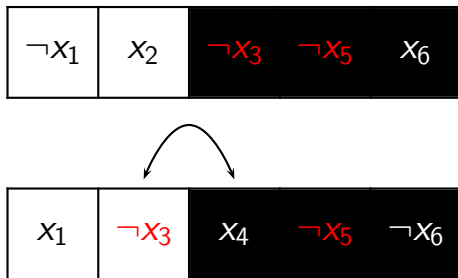
$$\varphi = \{x_1 = *, x_2 = *, x_3 = *, x_4 = *, x_5 = \mathbf{1}, x_6 = *\}$$

$\neg x_1$	x_2	$\neg x_3$	$\neg x_5$	x_6
------------	-------	------------	------------	-------

x_1	$\neg x_3$	x_4	$\neg x_5$	$\neg x_6$
-------	------------	-------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = *, x_2 = *, x_3 = \mathbf{1}, x_4 = *, x_5 = 1, x_6 = *\}$$



Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

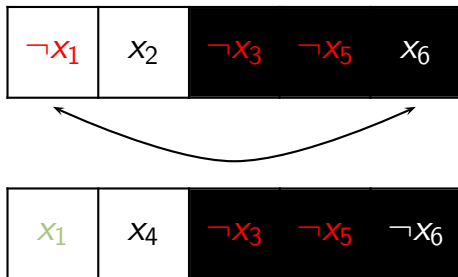
$$\varphi = \{x_1 = *, x_2 = *, x_3 = 1, x_4 = *, x_5 = 1, x_6 = *\}$$

$\neg x_1$	x_2	$\neg x_3$	$\neg x_5$	x_6
------------	-------	------------	------------	-------

x_1	x_4	$\neg x_3$	$\neg x_5$	$\neg x_6$
-------	-------	------------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = \mathbf{1}, x_2 = *, x_3 = 1, x_4 = *, x_5 = 1, x_6 = *\}$$



Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = 1, x_2 = *, x_3 = 1, x_4 = *, x_5 = 1, x_6 = *\}$$

x_6	x_2	$\neg x_3$	$\neg x_5$	$\neg x_1$
-------	-------	------------	------------	------------

x_1	x_4	$\neg x_3$	$\neg x_5$	$\neg x_6$
-------	-------	------------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = 1, x_2 = *, x_3 = 1, x_4 = \mathbf{0}, x_5 = 1, x_6 = *\}$$

x_6	x_2	$\neg x_3$	$\neg x_5$	$\neg x_1$
-------	-------	------------	------------	------------

x_1	x_4	$\neg x_3$	$\neg x_5$	$\neg x_6$
-------	-------	------------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = 1, x_2 = \mathbf{0}, x_3 = 1, x_4 = 0, x_5 = 1, x_6 = * \}$$

x_6	x_2	$\neg x_3$	$\neg x_5$	$\neg x_1$
-------	-------	------------	------------	------------

x_1	x_4	$\neg x_3$	$\neg x_5$	$\neg x_6$
-------	-------	------------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 0, x_5 = 1, x_6 = \mathbf{1}\}$$

x_6	x_2	$\neg x_3$	$\neg x_5$	$\neg x_1$
-------	-------	------------	------------	------------

x_1	x_4	$\neg x_3$	$\neg x_5$	$\neg x_6$
-------	-------	------------	------------	------------

Conflict-driven: Watch pointers (1) [MoskewiczMZZM'01]

$$\varphi = \{x_1 = 1, x_2 = 0, x_3 = 1, x_4 = 0, x_5 = 1, x_6 = 1\}$$

x_6	x_2	$\neg x_3$	$\neg x_5$	$\neg x_1$
-------	-------	------------	------------	------------

x_1	x_4	$\neg x_3$	$\neg x_5$	$\neg x_6$
-------	-------	------------	------------	------------

Conflict-driven: Watch pointers (2) [MoskewiczMZZM'01]

Only examine (get in the cache) a clause when both

- a watch pointer gets falsified
- the other one is not satisfied

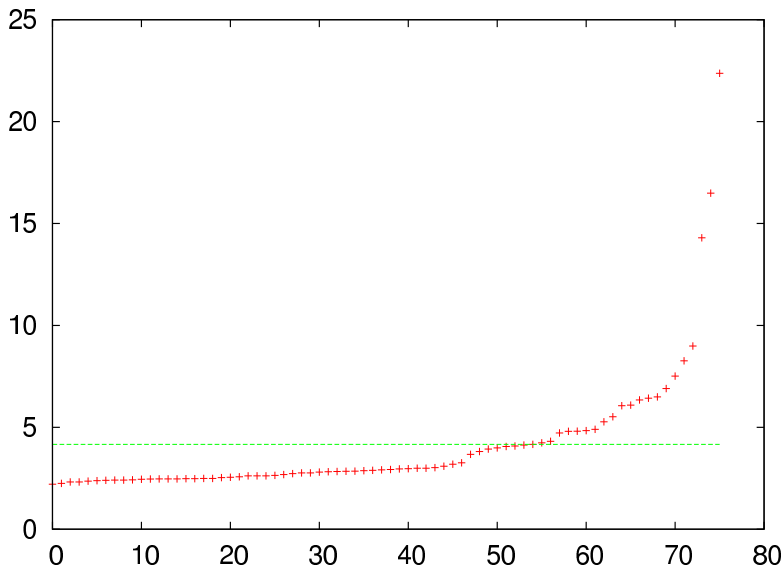
While backjumping, just unassign variables

Conflict clauses \rightarrow watch pointers

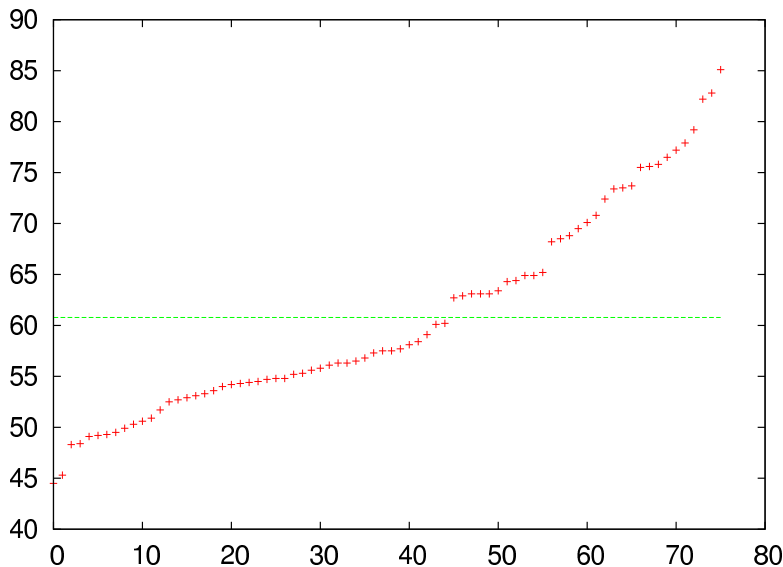
No detailed information available

Not used for binary clauses

Average Number Clauses Visited Per Propagation



Percentage visited clauses with other watched literal true



Heuristics

Most important CDCL heuristics

Variable selection heuristics

- aim: minimize the search space
- plus: could compensate a bad value selection

Most important CDCL heuristics

Variable selection heuristics

- aim: minimize the search space
- plus: could compensate a bad value selection

Value selection heuristics

- aim: guide search towards a solution (or conflict)
- plus: could compensate a bad variable selection, cache solutions of subproblems [PipatsrisawatDarwiche'07]

Most important CDCL heuristics

Variable selection heuristics

- aim: minimize the search space
- plus: could compensate a bad value selection

Value selection heuristics

- aim: guide search towards a solution (or conflict)
- plus: could compensate a bad variable selection, cache solutions of subproblems [PipatsrisawatDarwiche'07]

Restart strategies

- aim: avoid heavy-tail behavior [GomesSelmanCrato'97]
- plus: focus search on recent conflicts when combined with dynamic heuristics

Variable selection heuristics

Based on the occurrences in the (reduced) formula

- examples: Jeroslow-Wang, Maximal Occurrence in clauses of Minimal Size (MOMS), look-aheads
- not practical for CDCL solver due to watch pointers

Variable selection heuristics

Based on the occurrences in the (reduced) formula

- examples: Jeroslow-Wang, Maximal Occurrence in clauses of Minimal Size (MOMS), look-aheads
- not practical for CDCL solver due to watch pointers

Variable State Independent Decaying Sum (VSIDS)

- original idea (zChaff): for each conflict, increase the score of involved variables by 1, half all scores each 256 conflicts
[MoskewiczMZZM'01]
- improvement (MiniSAT): for each conflict, increase the score of involved variables by δ and increase $\delta := 1.05\delta$
[EenSörensson'03]

Visualization of VSIDS in PicoSAT

<http://www.youtube.com/watch?v=M0jhFywLre8>

Value selection heuristics

Based on the occurrences in the (reduced) formula

- examples: Jeroslow-Wang, Maximal Occurrence in clauses of Minimal Size (MOMS), look-aheads
- not practical for CDCL solver due to watch pointers

Value selection heuristics

Based on the occurrences in the (reduced) formula

- examples: Jeroslow-Wang, Maximal Occurrence in clauses of Minimal Size (MOMS), look-aheads
- not practical for CDCL solver due to watch pointers

Based on the encoding / consequently

- negative branching (early MiniSAT) [EenSörensson'03]

Value selection heuristics

Based on the occurrences in the (reduced) formula

- examples: Jeroslow-Wang, Maximal Occurrence in clauses of Minimal Size (MOMS), look-aheads
- not practical for CDCL solver due to watch pointers

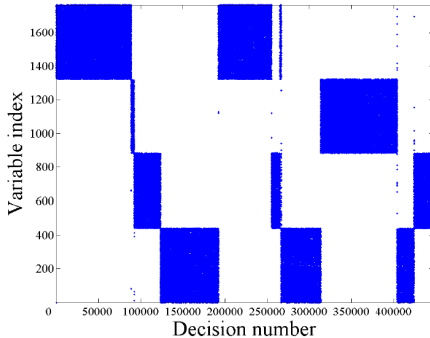
Based on the encoding / consequently

- negative branching (early MiniSAT) [EenSörensson'03]

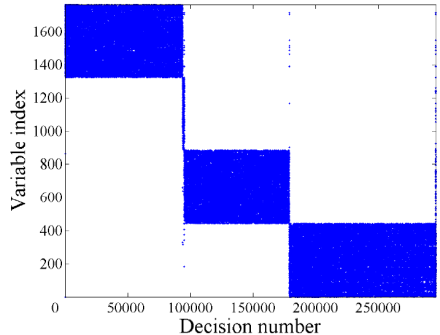
Based on the last implied value (phase-saving)

- introduced to CDCL [PipatsrisawatDarwiche'07]
- already used in local search [HirschKojevnikov'01]

Selecting the last implied value remembers solved components



negative branching



phase-saving

Restarts

Restarts in CDCL solvers:

- Counter heavy-tail behavior [GomesSelmanCrato'97]
- Unassign all variables but keep the (dynamic) heuristics

Restarts

Restarts in CDCL solvers:

- Counter heavy-tail behavior [GomesSelmanCrato'97]
- Unassign all variables but keep the (dynamic) heuristics

Restart strategies: [Walsh'99, LubySinclairZuckerman'93]

- Geometrical restart: e.g. 100, 150, 225, 333, 500, 750, ...
- Luby sequence: e.g. 100, 100, 200, 100, 100, 200, 400, ...

Restarts

Restarts in CDCL solvers:

- Counter heavy-tail behavior [GomesSelmanCrato'97]
- Unassign all variables but keep the (dynamic) heuristics

Restart strategies: [Walsh'99, LubySinclairZuckerman'93]

- Geometrical restart: e.g. 100, 150, 225, 333, 500, 750, ...
- Luby sequence: e.g. 100, 100, 200, 100, 100, 200, 400, ...

Rapid restarts by reusing trail: [vanderTakHeuleRamos'11]

- Partial restart same effect as full restart
- Optimal strategy Luby-1: 1, 1, 2, 1, 1, 2, 4, ...

Conflict-Clause Minimization

Self-Subsumption

Use self-subsumption to shorten conflict clauses

$$\frac{C \vee l \quad D \vee \bar{l}}{D} \quad C \subseteq D \quad \frac{(a \vee b \vee l) \quad (a \vee b \vee c \vee \bar{l})}{(a \vee b \vee c)}$$

Conflict clause minimization is an important optimization.

Self-Subsumption

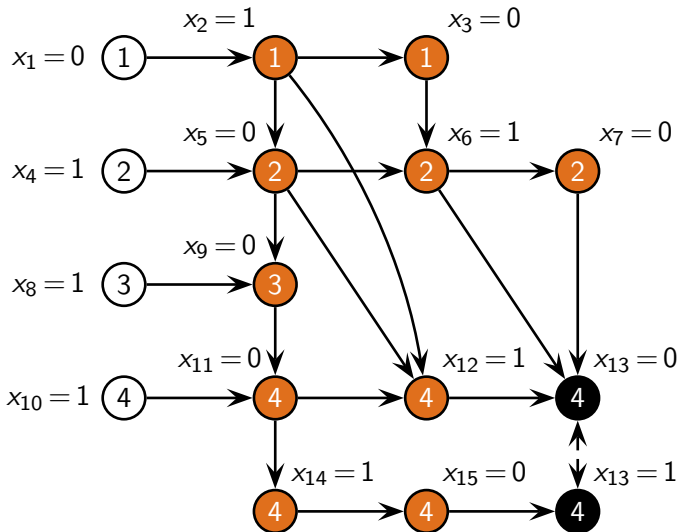
Use self-subsumption to shorten conflict clauses

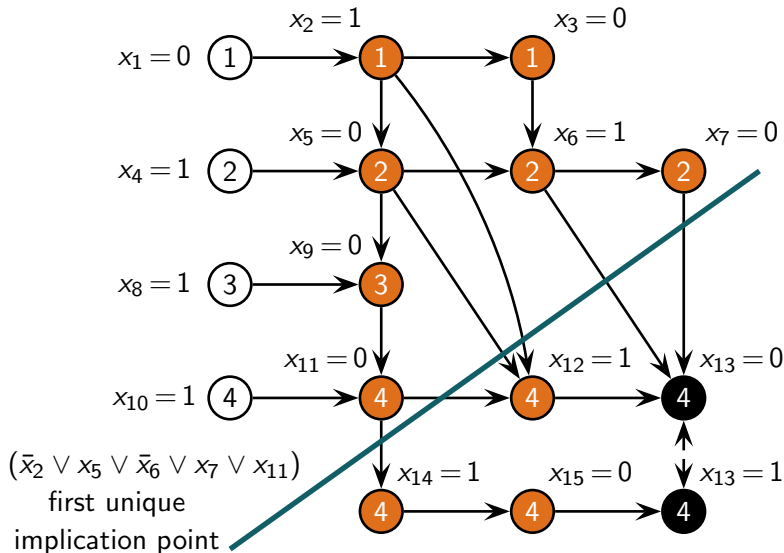
$$\frac{C \vee I \quad D \vee \bar{I}}{D} \quad C \subseteq D \quad \frac{(a \vee b \vee I) \quad (a \vee b \vee c \vee \bar{I})}{(a \vee b \vee c)}$$

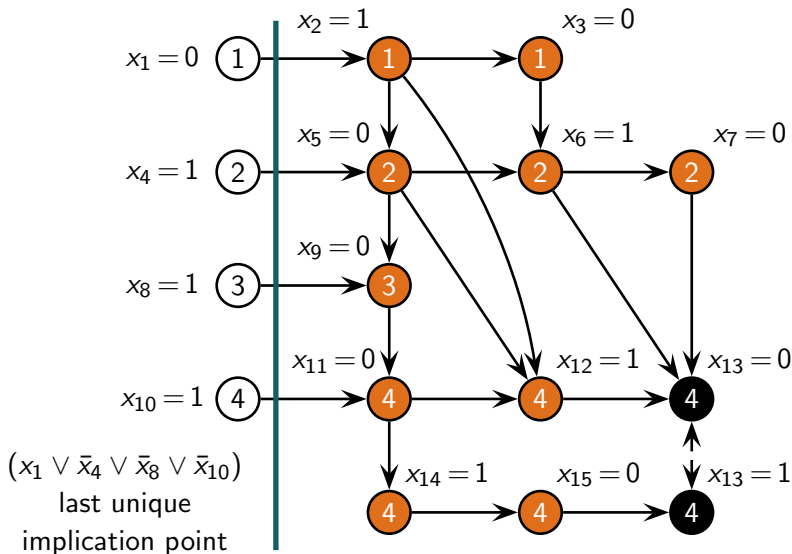
Conflict clause minimization is an important optimization.

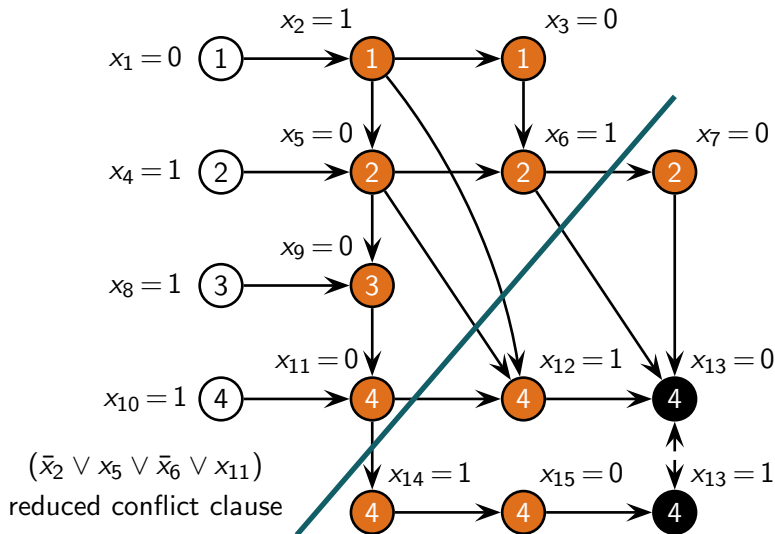
Use implication chains to further minimization:

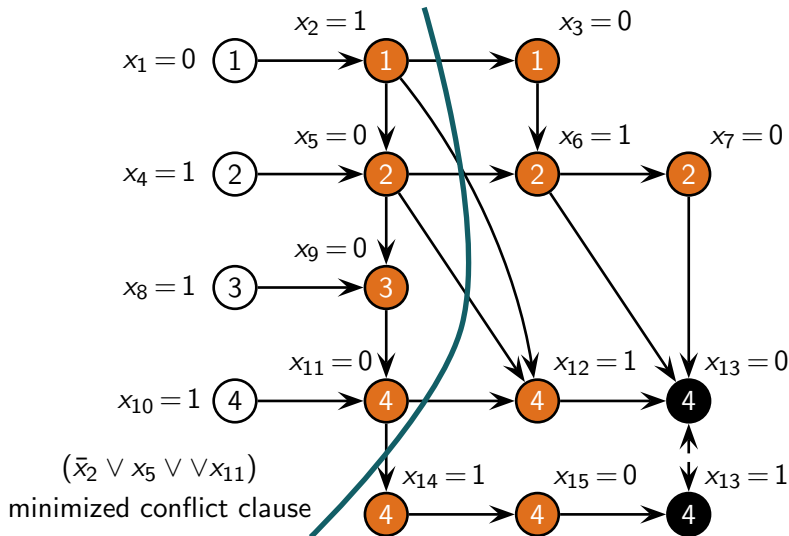
$$\begin{aligned} \dots (\bar{a} \vee b)(\bar{b} \vee c)(\textcolor{red}{a} \vee c \vee d) \dots &\Rightarrow \\ \dots (\bar{a} \vee b)(\bar{b} \vee c)(c \vee d) \dots \end{aligned}$$











Conclusions: state-of-the-art CDCL solver

Key contributions to CDCL solvers:

- concept of conflict clauses (grasp) [Marques-SilvaSakallah'96]
- restart strategies [GomesSC'97,LubySZ'93]
- 2-watch pointers and VSIDS (zChaff) [MoskewiczMZZM'01]
- efficient implementation (Minisat) [EenSörensson'03]
- phase-saving (Rsat) [PipatsrisawatDarwiche'07]
- conflict-clause minimization [SörenssonBiere'09]

+ Pre- and in-processing techniques