SAT-based Approaches for Test & Verification of Integrated Circuits

Albert-Ludwigs-Universität Freiburg

Dr. Tobias Schubert

Chair of Computer Architecture
Institute of Computer Science
Faculty of Engineering
schubert@informatik.uni-freiburg.de

Summer School on Verification Technology, Systems & Applications 2015

UNI FREIBURG

About Me

Just a very short CV

- Studied computer science & microsystems engineering at the University of Freiburg
- Made my PhD working on efficient parallel SAT solving at the University of Freiburg
- Member of the *Transregional Collaborative Research Center 14 AVACS – Automatic Verification and Analysis of Complex Systems*
- Principal investigator within the cluster of excellence *BrainLinks-BrainTools*
- Member of the part-time distance learning program *Intelligent Embedded Microsystems*

About Me

My research interests include

- Efficient (parallel) algorithms for SAT and related domains
- Real-world applications using
  - SAT,
  - #SAT,
  - MaxSAT,
  - QBF, and
  - SMT solvers

  as the underlying backend
- Embedded & cyber-physical systems
- Industrial internet & internet of things
- E-learning, blended learning, distance teaching

Collaborators

University of Freiburg

- Bernd Becker
- Jan Burchard
- Alejandro Czutro
- Linus Feiten
- Karina Gitina
- Paolo Marin
- Sven Reimer
- Matthias Sauer
- Karsten Scheibler
- Christoph Scholl
- Ralf Wimmer

University of Bremen

- Rolf Drechsler

University of Oldenburg

- Martin Fränzle

University of Passau

- Ilia Polian

University of Potsdam

- Torsten Schaub

MPI Saarbrücken

- Christoph Weidenbach

Motivation: Embedded Systems

Embedded Systems

- Information processing systems embedded into a “larger” product

Without Embedded Systems

- No cars would drive today
- No planes would fly today
- No factory would work today
- No mobile communication would be possible

Motivation: Embedded Systems

Embedded Systems

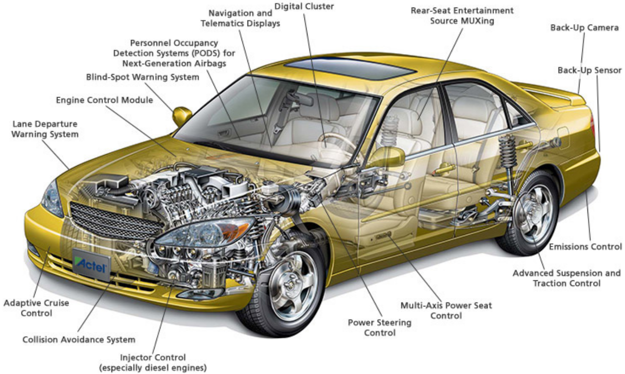
- Information processing systems embedded into a “larger” product

Without Embedded Systems

- No cars would drive today
- No planes would fly today
- No factory would work today
- No mobile communication would be possible

Verifying designs and testing produced chips are mandatory tasks, in particular for safety-critical applications!

Motivation: Automotive Area



- Many functions controlled by embedded systems
- Multiple networks / system busses
- Up to 70 different processors within one car

Motivation: Automotive Area

Consequences

- Increasing system complexity
- Increasing number of dependencies between different subsystems
- Up to 40% of the total costs are caused by electronics & software
- Up to 90% of the innovations are driven by electronics & software
- 40–50% of all car breakdowns are caused by electronics & software
- Errors related to electronics or software are responsible for more than 40% of all call-backs
- Reliable function is of outmost importance, because otherwise human lives can be endangered!

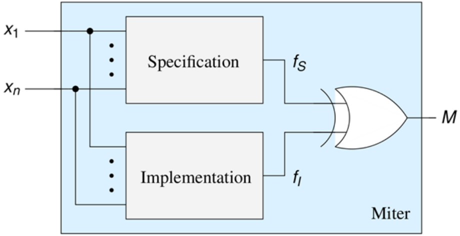
⇒ Safety-critical application of embedded systems!

Verifying Integrated Circuit Designs

Focus is on detecting design errors

- Errors which occur during the translation of a specification into the final integrated circuit ($\leadsto$ implementation)
- Errors in the design make all produced chips erroneous

$\Rightarrow$ Formal methods to avoid design errors before producing any chip

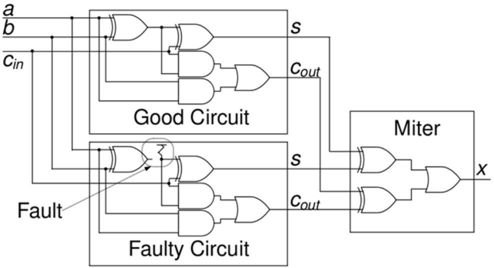


Testing Integrated Circuits

Focus is on production errors

- Defects which are caused during the production of single chips and which change their functionality
- Causes are contaminations, shifted exposure masks, wrong doping, ...

⇒ Formal methods to ensure that all production errors can be found



UNI FREIBURG

But why using SAT Solvers?

- Tremendous performance improvements within the last 15 years
- Nowadays SAT solvers (and their extensions) are able to ...
  - solve problems coming from real-world applications (e.g., large industrial circuits)
  - handle optimization & enumeration problems, multi-valued domains, hybrid systems

Typical SAT-based Flow

"real problem"

"real solution"

$$(a+b+\overline{c})\cdot(\overline{a}+c)\cdot(\overline{b}+c)\cdot(c+\overline{d})\cdot(\overline{c}+d)\cdot(\overline{d}+\overline{e}+f)\cdot(d+\overline{f})\cdot(e+\overline{f})\cdot f$$

SAT instance

SAT solver

$$a=0, b=0, c=0, d=1, e=1, f=1$$

SAT solution

Outline

Applications

- Bounded Model / Property Checking
- Path Compaction
- Security Issues
- Test Pattern Relaxation
- Automatic Test Pattern Generation
- Hybrid System Verification
- Black Box Verification
- Combinational Equivalence Checking
- The End

- SAT
- MaxSAT
- #SAT
- QBF
- DQBF
- SMT

Core Algorithms

Outline

Applications

- Bounded Model / Property Checking
- Path Compaction
- Security Issues
- Test Pattern Relaxation
- Automatic Test Pattern Generation
- Hybrid System Verification
- Black Box Verification
- Combinational Equivalence Checking
- The End

Core Algorithms

- **SAT**
- MaxSAT
- #SAT
- QBF
- DQBF
- SMT

Boolean Satisfiability Problem (SAT)

- Given
  - A Boolean formula φ in Conjunctive Normal Form (CNF)
    - A CNF is a conjunction of clauses: $C_1 \wedge \ldots \wedge C_m$
    - A clause is a disjunction of literals: $(l_1 \vee \ldots \vee l_k)$
    - A literal l is a Boolean variable or its negation: l or $\neg l$
- Question
  - Is there a valuation of the variables that satisfies φ?
- Example
  - $x_1 = x_2 = 0, x_3 = 1$ satisfies
    $\varphi = (\neg x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee \neg x_2 \vee \neg x_3)$
- Techniques for solving instances of the SAT problem are called SAT algorithms or SAT solvers
- Complexity of the “general” SAT problem: NP-complete (S.A. Cook, 1971)

Overview of SAT Algorithms

Focus here is on complete methods

- Due to a systematic procedure complete solvers are able to prove the unsatisfiability of a CNF formula
- DP algorithm
  - M. Davis, H. Putnam, 1960
  - Based on resolution
- DLL algorithm
  - M. Davis, G. Logemann, D. Loveland, 1962
  - Based on depth-first search
- Modern SAT algorithms
  - Based on the DLL algorithm, but enriched with efficient data structures and several acceleration & optimization techniques
  - zChaff, MiniSat, MiraXT, lingeling, antom, Glucose

Preliminaries

Definition (Empty Clause)

The empty clause, denoted with $\square$, describes the empty set of literals, and it is unsatisfiable by definition.

Definition (Empty Formula)

The empty formula describes an empty set of clauses and it is satisfiable by definition.

Preliminaries

Definition (Pure Literal)

Let F be a CNF formula and L be a literal contained in F. L is called a pure literal iff L occurs in F only positive or only negative.

Steps in order to simplify a CNF formula F

- Delete from F all clauses in which a pure literal L occurs, because these ones will be satisfied by an appropriate assignment to L

Remark

- As it is rather time consuming, pure literal detection is applied by modern SAT solvers during pre-/inprocessing only

UNI FREIBURG

Preliminaries

Definition (Unit Clause)

A clause consisting of a single literal L is called a unit clause with L being the corresponding unit literal.

Steps in order to simplify a CNF formula F

- Assign a unit literal L to 1
- Delete from F all clauses containing L
- Delete all occurrences of $\neg L$

Preliminaries

Definition (Subsumption)

Let C_1 and C_2 be two clauses. C_1 subsumes C_2 iff all literals occurring in C_1 also occur in C_2: $C_1 \subseteq C_2$.

Steps in order to simplify a CNF formula F

- Delete all clauses from F that are subsumed by at least one other clause of F

Remark

- Typically, modern SAT solvers apply subsumption checks during pre-/inprocessing only

Preliminaries

Definition (Resolution)

Let C_1 and C_2 be two clauses and L be a literal with the following property: $L \in C_1$ and $\neg L \in C_2$. Then one can compute the clause R

$$R = \left(C_1 - \{L\}\right) \cup \left(C_2 - \{\neg L\}\right)$$

that is denoted as the resolvent of the clauses C_1 and C_2 over L. Typically, the notation $R = C_1 \otimes_L C_2$ is used.

Lemma (Resolution Lemma)

Let F be a CNF formula and R be the resolvent of two clauses C_1 and C_2 from F. Then F and $F \cup \{R\}$ are equivalent: $F \equiv F \cup \{R\}$.

Preliminaries

Definition

Let F be a CNF formula. Then $Res(F)$ is defined as

$$Res(F) = F \cup \{R \mid R \text{ is the resolvent of two clauses in } F\}.$$

Moreover, let us define:

$$Res^0(F) = F$$

$$Res^{t+1}(F) = Res(Res^t(F)) \text{ for } t \geq 0$$

$$Res^*(F) = \lim_{t \geq 0} Res^t(F)$$

Theorem (Resolution Theorem)

A CNF formula F is unsatisfiable iff $\square \in Res^*(F)$.

UNI FREIBURG

Preliminaries

Definition

Let F be a CNF formula and x_i a variable occurring in F with $L = x_i$ and $\neg L = \neg x_i$. The we define P, N and W as follows:

- P is the set of clauses in F which contain L:

$$P = \{C \in F \mid L \in C\}$$

- N is the set of clauses in F which contain $\neg L$:

$$N = \{C \in F \mid \neg L \in C\}$$

- W is the set of clauses in F which contain neither L nor $\neg L$:

$$W = \{C \in F \mid L \notin C \wedge \neg L \notin C\}$$

Obviously, we have $F = P \cup N \cup W$.

Preliminaries

Definition (Pairwise Resolution)

Using this partitioning of the clauses we define $P \otimes_{x_i} N$ as the set of clauses, which can be constructed by resolution of all pairs $(p,n) \in P \times N$:

$$P \otimes_{x_i} N = \{R \,|\, (R = C_1 \otimes_{x_i} C_2) \wedge (C_1 \in P) \wedge (C_2 \in N)\}.$$

Theorem (Variable Elimination)

Let F be a formula in CNF and x_i a variable which appears both positive and negative in F. Further let the sets P, N, and W be the partition of F as defined before.
Then $F = P \cup N \cup W$ and $F' = (P \otimes_{x_i} N) \wedge W$ are satisfiability equivalent.

DLL Algorithm

- Main idea: If a CNF formula F is satisfiable, then for an arbitrary variable x_i occuring in F either $x_i = 1$ or $x_i = 0$ must hold

  $\Rightarrow$ Try both cases one after the other

  $\Rightarrow$ Depth-first search

- Applying unit clause & pure literal rule to accelerate the search
- Recursive algorithm, in particular the given formula gets modified when going from recursion level r to $r + 1$
- In the literature both "DLL" and "DPLL" can be found

DLL Algorithm

bool DLL(CNF F)
{
  if ($F = \emptyset$) { **return** *SATISFIABLE*; } // Empty set of clauses
  if ($\square \in F$) { **return** *UNSATISFIABLE*; } // Empty Clause

  if (F contains a unit clause (L)) // Unit Clause
    {
      // Unit Subsumption.
      $F' = F - \{C \mid (L \in C) \wedge (C \in F) \wedge (C \neq (L))\}$;

      // Unit Resolution.
      $P = \{(L)\}$;
      $N = \{C \mid (\neg L \in C) \wedge (C \in F')\}$;
      $W = F' - P - N$;
      return DLL($[P \otimes_L N] \wedge W$);
    }

  if (F contains a pure literal L) // Pure Literal
    {
      // Delete from F every clause containing L.
      $F' = F - \{C \mid (L \in C) \wedge (C \in F)\}$;
      return DLL(F');
    }

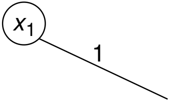
  L = SelectLiteral(F); // Choose a Literal
  if (DLL($F \cup \{(L)\}$) $==$ *SATISFIABLE*) // Case distinction
    { **return** *SATISFIABLE*; }
  else
    { **return** DLL($F \cup \{(\neg L)\}$); }
}

DLL Algorithm

$$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$$

UNI FREIBURG

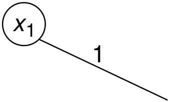
DLL Algorithm



$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$

Case distinction $x_1 = 1$

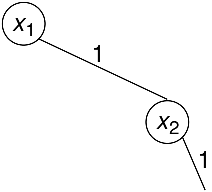
DLL Algorithm



$(\quad , \neg x_2, \neg x_3) \wedge (\quad , \neg x_2, x_3) \wedge (\quad , x_2, \neg x_3) \wedge (\quad , x_2, x_3) \wedge$

Case distinction $x_1 = 1$

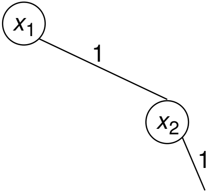
DLL Algorithm



$(\quad , \neg x_2, \neg x_3) \wedge (\quad , \neg x_2, x_3) \wedge (\quad , x_2, \neg x_3) \wedge (\quad , x_2, x_3) \wedge$

Case distinction $x_2 = 1$

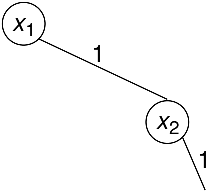
DLL Algorithm



$(\quad , \quad , \neg x_3) \wedge (\quad , \quad , x_3) \wedge \qquad \wedge \qquad \wedge$

Case distinction $x_2 = 1$

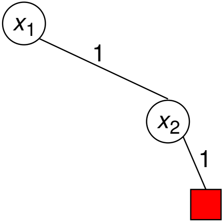
DLL Algorithm



$(\quad , \quad , \neg x_3) \wedge (\quad , \quad , x_3) \wedge \qquad \wedge \qquad$

Unit clauses $x_3 = 0$ and $x_3 = 1$

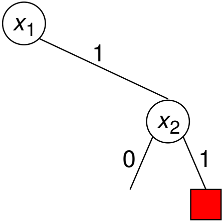
DLL Algorithm



$(\quad , \quad , \neg x_3) \wedge (\quad , \quad , x_3) \wedge \quad \wedge \quad \wedge$

Contradiction/conflict

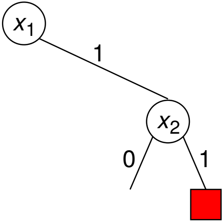
DLL Algorithm



$(\quad, \neg x_2, \neg x_3) \wedge (\quad, \neg x_2, x_3) \wedge (\quad, x_2, \neg x_3) \wedge (\quad, x_2, x_3) \wedge$

Case distinction $x_2 = 0$

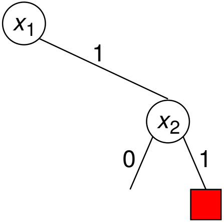
DLL Algorithm



$\wedge \quad\quad \wedge(\quad , \quad ,\neg x_3)\wedge(\quad , \quad ,x_3)\wedge$

Case distinction $x_2 = 0$

DLL Algorithm

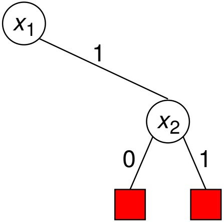


$\wedge \qquad \wedge(\quad , \quad , \neg x_3) \wedge (\quad , \quad , x_3) \wedge$

Unit clauses $x_3 = 0$ and $x_3 = 1$

UNI FREIBURG

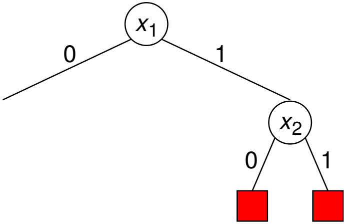
DLL Algorithm



$\wedge \qquad \wedge(\quad , \quad , \neg x_3) \wedge (\quad , \quad , x_3) \wedge$

Contradiction/conflict

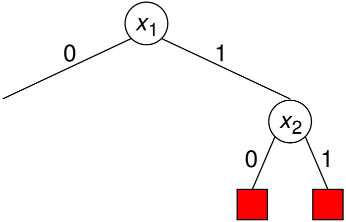
DLL Algorithm



$$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$$

Case distinction $x_1 = 0$

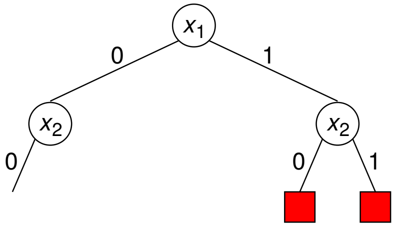
DLL Algorithm



$\wedge \qquad \wedge \qquad \wedge \qquad \wedge (\quad , \neg x_2, \neg x_3)$

Case distinction $x_1 = 0$

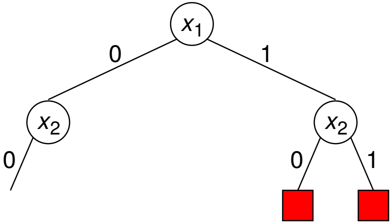
DLL Algorithm



$\wedge$ $\wedge$ $\wedge$ $\wedge (\quad , \neg x_2, \neg x_3)$

Pure literal $x_2 = 0$

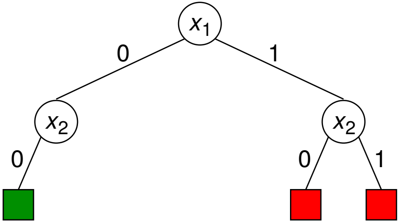
DLL Algorithm



$\wedge$ $\wedge$ $\wedge$ $\wedge$

Pure literal $x_2 = 0$

DLL Algorithm



$\wedge$ $\wedge$ $\wedge$ $\wedge$

Formula satisfiable

From DLL to modern SAT Algorithms

Overall

- DLL algorithm
  - Recursive procedure
  - For the transition from recursion level r to level $r+1$ the given formula gets modified
  - For backtracking from level $r+1$ to r the original (sub)formula at level r has to be restored
- Modern SAT algorithms
  - Non-recursive implementation
  - Apart from special cases (preprocessing), the CNF remains unmodified
  - Typically, the pure literal rule is not applied

From DLL to modern SAT Algorithms

Unit clause

- DLL algorithm
  - A clause consisting exactly one literal
- Modern SAT algorithms
  - In addition to the rule above, clauses where all literals but one are assigned with negated polarity are also referred to as unit clauses
  - Example: Assignment $x_1 = 0, x_2 = 1$ turns $(x_1, \neg x_2, x_3)$ into a unit clause
  - In the example, the evaluation $x_1 = 0, x_2 = 1$ forces the assignment $x_3 = 1$ in order to satisfy the clause $(x_1, \neg x_2, x_3)$ $\Rightarrow$ implication

From DLL to modern SAT Algorithms

Unit propagation to determine all implications forced by a variable assignment

- DLL algorithm
  - Repeated application of the unit clause rule on successsive recursion levels until the rule cannot be applied anymore
- Modern SAT algorithms
  - Done non-recursively, also called Boolean Constraint Propagation (BCP)
  - Example: For the CNF $F = (x_1, \neg x_2) \wedge (x_1, x_2, x_3) \wedge (\neg x_3, x_4)$, $x_1 = 0$ leads to the implications $x_2 = 0, x_3 = 1, x_4 = 1$

From DLL to modern SAT Algorithms

Contradiction/conflict

- DLL algorithm
  - Empty clause
- Modern SAT algorithms
  - Unsatisfied clause
  - Example: Valuation $x_1 = 0, x_2 = 1, x_3 = 0$ makes $(x_1, \neg x_2, x_3)$ unsatisfied, and so the whole CNF formula containing it cannot be satisfied anymore

From DLL to modern SAT Algorithms

Conflict analysis & backtracking

- DLL algorithm
  - The combination of the decisions done before will always be considered as the origin of a conflict
  - Backtracking to the recursion level of the last “branching” in which one case for a variable assignment has not been explored yet
  - If such a recursion level does not exist, the given CNF formula is unsatisfiable

From DLL to modern SAT Algorithms

Conflict analysis & backtracking

- Modern SAT algorithms
  - Complex analysis of the conflict setting, because not all "branchings" done before have to be involved in the current conflict
  - Learning of a conflict clause via resolution to avoid running into the same conflict again
  - (Non-)chronological backtracking according to the derived conflict clause
  - If a conflict occurs on decision level 0, the given CNF formula is unsatisfiable

Modern SAT Algorithms

Main techniques of today's SAT solvers

- Preprocessing
- In turn…
  - Choose the next decision variable
  - Boolean constraint propagation / unit propagation
  - If necessary, conflict analysis & backtracking
- At some fixed points during the search process
  - Unlearning (of some conflict clauses)
  - Restarts
  - Inprocessing
- In case of a satisfiable CNF formula
  - Output the satisfying variable assignment $\Rightarrow$ model

Modern SAT Algorithms

```
bool SequentialSatEngine(CNF F)
{
  if (PreprocessCNF(F) == CONFLICT)                      // Preprocessing the CNF formula
    { return UNSATISFIABLE; }                            // Problem unsatisfiable
  while (true)
    {
      if (DecideNextBranch())                            // Choice of the next unassigned variable
        {
          while (BCP() == CONFLICT)                      // Boolean Constraint Propagation
            {
              BLevel = AnalyzeConflict();                // Conflict analysis
              if (BLevel > 0)
                { Backtrack(BLevel); }                   // Cancel the „incorrect“ assignment
              else
                { return UNSATISFIABLE; }                // Problem unsatisfiable
            }
        }
      else
        { return SATISFIABLE; }                          // All variables assigned, problem satisfiable
    }
}
```

Not explicitly stated: Inprocessing, unlearning, restarts, model output

UNI FREIBURG

Modern SAT Algorithms

```
bool SequentialSatEngine(CNF F)
{
  if (PreprocessCNF(F) == CONFLICT)                          // Preprocessing the CNF formula
    { return UNSATISFIABLE; }                                 // Problem unsatisfiable
  while (true)
    {
      if (DecideNextBranch())                     // Choice of the next unassigned variable
        {
          while (BCP() == CONFLICT)                        // Boolean Constraint Propagation
            {
              BLevel = AnalyzeConflict();                               // Conflict analysis
              if (BLevel > 0)
                { Backtrack(BLevel); }                  // Cancel the „incorrect“ assignment
              else
                { return UNSATISFIABLE; }                          // Problem unsatisfiable
            }
        }
      else
        { return SATISFIABLE; }            // All variables assigned, problem satisfiable
    }
}
```

Not explicitly stated: Inprocessing, unlearning, restarts, model output

Preprocessing

- Goal
  - Reduce the formula's size in terms of clauses and literals to speed up the search process
- Observation from the experience
  - As a rule of thumb, the size of a formula is related to the time necessary for the SAT algorithm to solve it
- Identification & preprocessing of unit clauses within the original set of clauses belong to the common operations done in modern SAT algorithms
- It is very important to find a good compromise between the additional effort required by preprocessing and the expected saving during the search process

Preprocessing

Unit Propagation Lookahead (UPLA)

- Fix a variable x_i to 0, check implications; then change its value to $x_i = 1$, check implications. Simplify the formula exploiting the following consequences:
  - $(x_i = 0 \rightarrow \mathit{conflict}) \wedge (x_i = 1 \rightarrow \mathit{conflict}) \Rightarrow \mathit{UNSAT}$
  - $(x_i = 0 \rightarrow \mathit{conflict}) \Rightarrow x_i = 1$
  - $(x_i = 1 \rightarrow \mathit{conflict}) \Rightarrow x_i = 0$
  - $(x_i = 0 \rightarrow x_j = 1) \wedge (x_i = 1 \rightarrow x_j = 1) \Rightarrow x_j = 1$
  - $(x_i = 0 \rightarrow x_j = 0) \wedge (x_i = 1 \rightarrow x_j = 0) \Rightarrow x_j = 0$
  - $(x_i = 0 \rightarrow x_j = 0) \wedge (x_i = 1 \rightarrow x_j = 1) \Rightarrow x_i \equiv x_j$

Preprocessing

Unit Propagation Lookahead (UPLA)

- Advantage
  - Built on top of the components already available in the solver
- Disadvantages
  - Requires binary clauses in the original formula
  - Necessary to extend the model when e. g. $x_i \equiv x_j$ is detected and all the occurrences of x_i are substituted with x_j
  - In general quite time consuming, in particular if all the variables are tested

Preprocessing

Application of resolution

- Advantages
  - No particular kind of clauses necessary in the original formula
  - Usually, simplifies effectively within a manageable time
- Disadvantages
  - In case of a satisfiable CNF formula, model extension required
- Techniques (SatELite)
  - Self-subsuming resolution
  - Elimination by clause distribution
  - Variable elimination by substitution
  - Forward subsumption
  - Backward subsumption

UNI FREIBURG

Preprocessing

Self-subsuming resolution

- Original formula
  - $F = (x_1 \vee \neg x_3) \wedge (x_1 \vee x_2 \vee x_3) \wedge \ldots$
- Resolution applied to the first two clauses
  - $(x_1 \vee \neg x_3) \otimes_{x_3} (x_1 \vee x_2 \vee x_3) = (x_1 \vee x_2)$

  $\Rightarrow$ $(x_1 \vee x_2)$ subsumes $(x_1 \vee x_2 \vee x_3)$

  $\Rightarrow$ Replace $(x_1 \vee x_2 \vee x_3)$ with $(x_1 \vee x_2)$
- Simplified formula
  - $F' = (x_1 \vee \neg x_3) \wedge (x_1 \vee x_2) \wedge \ldots$
- Saving
  - 1 literal

Preprocessing

Elimination by clause distribution

- Sometimes also called variable elimination
- Original formula
  - $F = (x_1 \vee x_2) \wedge (x_1 \vee \neg x_3) \wedge (\neg x_1 \vee x_3) \wedge (\neg x_1 \vee \neg x_2)$
- Variable elimination applied to x_1 leads to
  - $F' = (x_2 \vee x_3) \wedge (\neg x_3 \vee \neg x_2)$
- Saving
  - 1 variable, 2 clauses, 4 literals
- Applied only if it leads to a reduction of the formula's size

Preprocessing

Variable elimination by substitution

- Original formula
  - $F = (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge$
    $(x_4 \vee \neg x_5) \wedge (\neg x_4 \vee x_5 \vee x_6)$
- The first three clauses represent an AND gate ($\rightsquigarrow$ Tseitin transformation)
  - $[(\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2)] \leftrightarrow [x_5 \equiv x_1 \wedge x_2]$
- Removing the first three clauses, and replacing the occurrences of x_5 by $x_1 \wedge x_2$ in the other clauses leads to
  - $F' = (x_4 \vee \neg(x_1 \wedge x_2)) \wedge (\neg x_4 \vee (x_1 \wedge x_2) \vee x_6)$
- Transformation into CNF
  - $F'' = (x_4 \vee \neg x_1 \vee \neg x_2) \wedge (\neg x_4 \vee x_1 \vee x_6) \wedge (\neg x_4 \vee x_2 \vee x_6)$
- Saving: 1 variable, 2 clauses, 3 literals
- Applied only if it leads to a reduction of the formula's size
- Procedure for OR, NAND, other “basic gates” quite similar

Preprocessing

Forward subsumption

- Test if a clause generated during one of the preprocessing techniques described before is already subsumed by one clause of the current CNF formula

Backward subsumption

- Test if a clause generated during one of the preprocessing techniques described before subsumes one (or more) clauses of the current CNF formula

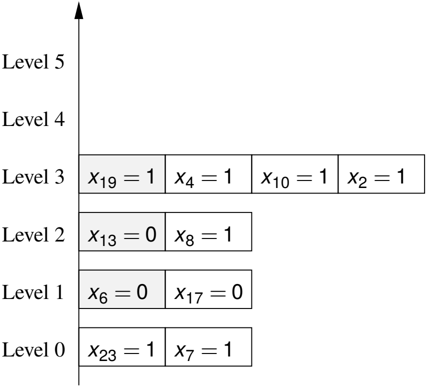
$\Rightarrow$ Remove all the clauses subsumed

Modern SAT Algorithms

```
bool SequentialSatEngine(CNF F)
{
  if (PreprocessCNF(F) == CONFLICT)                    // Preprocessing the CNF formula
    { return UNSATISFIABLE; }                          // Problem unsatisfiable
  while (true)
    {
      if (DecideNextBranch())                          // Choice of the next unassigned variable
        {
          while (BCP() == CONFLICT)                    // Boolean Constraint Propagation
            {
              BLevel = AnalyzeConflict();              // Conflict analysis
              if (BLevel > 0)
                { Backtrack(BLevel); }                 // Cancel the „incorrect“ assignment
              else
                { return UNSATISFIABLE; }              // Problem unsatisfiable
            }
        }
      else
        { return SATISFIABLE; }                        // All variables assigned, problem satisfiable
    }
}
```

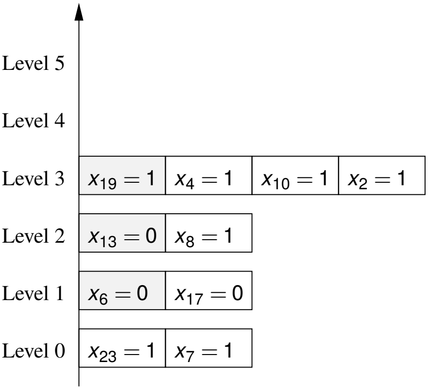
Not explicitly stated: Inprocessing, unlearning, restarts, model output

Decision Stack



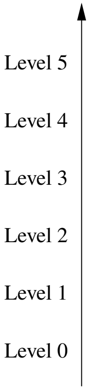
- Central data structure of modern SAT algorithms
- Decision stack stores the order of the executed assignments
- If a model for a CNF formula could be found, the decision stack stores the satisfying assignment

Decision Stack



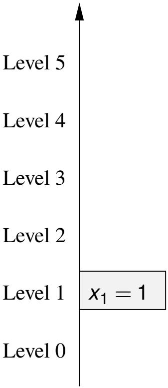
- Each variable assignment has an associated decision level
- Decision level gets initialized with 0; before a decision is made, it is incremented by one; backtracking decrements the decision level appropriately
- Decision level 0 plays a special role: It stores only implications from unit clauses in the original formula, but no decisions
- A conflict on decision level 0 means that the CNF is unsatisfiable

Decision Stack – First Example



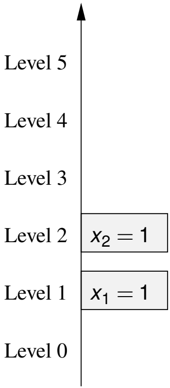
$$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$$

Decision Stack – First Example



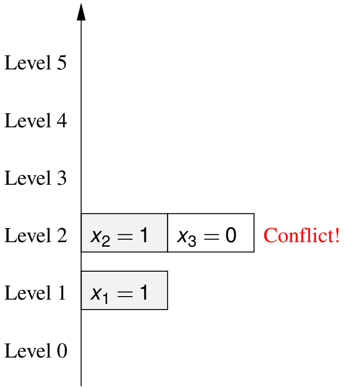
$$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$$

Decision Stack – First Example



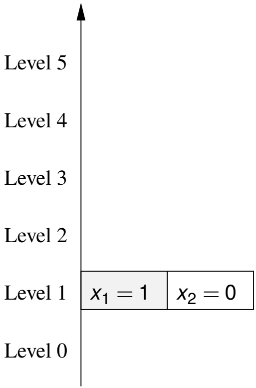
$$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$$

Decision Stack – First Example



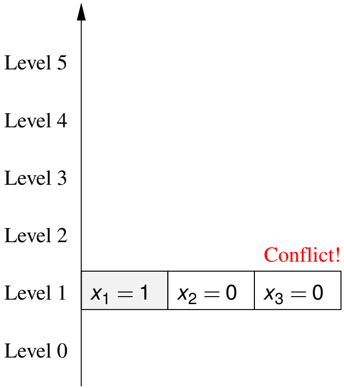
$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$

Decision Stack – First Example



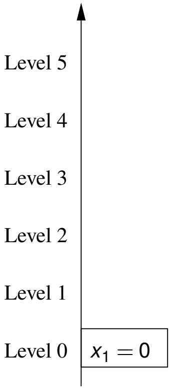
$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$

Decision Stack – First Example



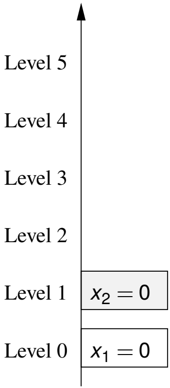
$$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$$

Decision Stack – First Example



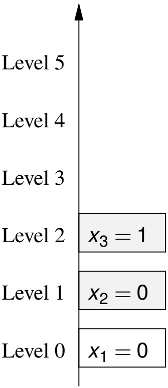
$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$

Decision Stack – First Example



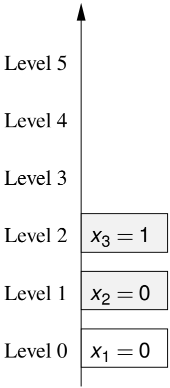
$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$

Decision Stack – First Example



$(\neg x_1, \neg x_2, \neg x_3) \wedge (\neg x_1, \neg x_2, x_3) \wedge (\neg x_1, x_2, \neg x_3) \wedge (\neg x_1, x_2, x_3) \wedge (x_1, \neg x_2, \neg x_3)$

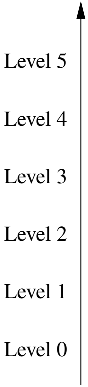
Decision Stack – First Example



$\Rightarrow$ Formula satisfiable with, e. g., $x_1 = 0, x_2 = 0, x_3 = 1$

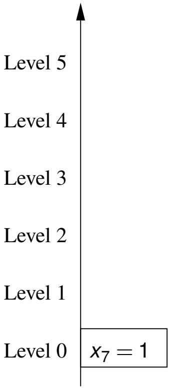
UNI FREIBURG

Decision Stack – Second Example



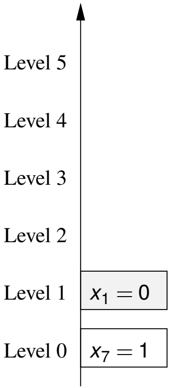
$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



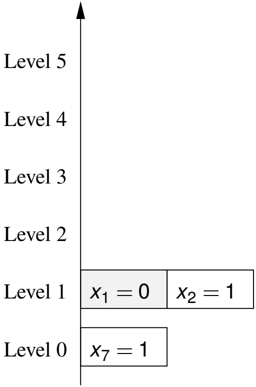
$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



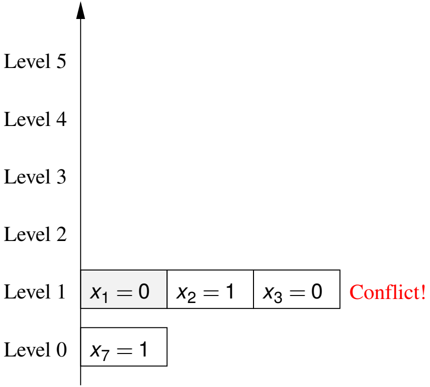
$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



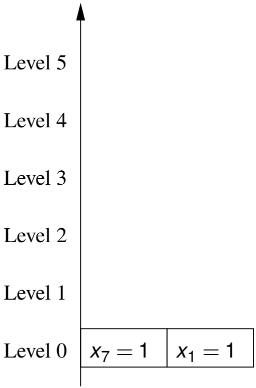
$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



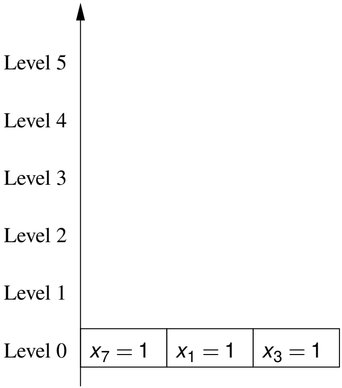
$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



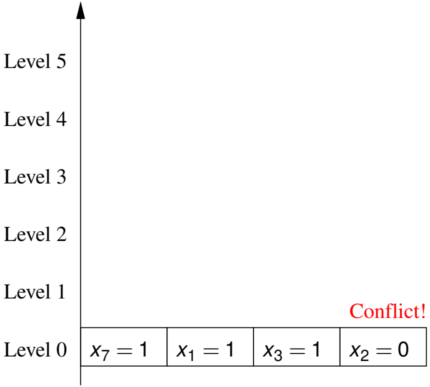
$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



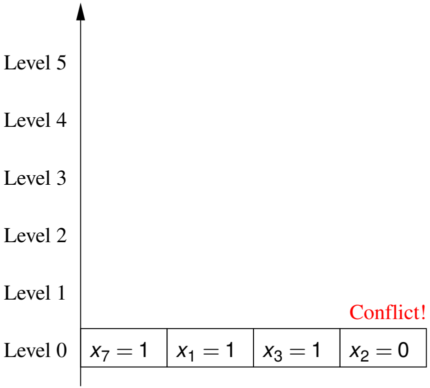
$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$

Decision Stack – Second Example



$$(x_1, x_2) \wedge (x_1, \neg x_3) \wedge (\neg x_1, x_3) \wedge (\neg x_1, \neg x_2) \wedge (x_3, \neg x_2) \wedge (\neg x_3, x_2) \wedge (x_7)$$

Decision Stack – Second Example



$\Rightarrow$ Formula unsatisfiable due to a conflict on decision level 0

Modern SAT Algorithms

```
bool SequentialSatEngine(CNF F)
{
  if (PreprocessCNF(F) == CONFLICT)                     // Preprocessing the CNF formula
    { return UNSATISFIABLE; }                           // Problem unsatisfiable
  while (true)
    {
      if (DecideNextBranch())                           // Choice of the next unassigned variable
        {
          while (BCP() == CONFLICT)                     // Boolean Constraint Propagation
            {
              BLevel = AnalyzeConflict();               // Conflict analysis
              if (BLevel > 0)
                { Backtrack(BLevel); }                  // Cancel the „incorrect“ assignment
              else
                { return UNSATISFIABLE; }               // Problem unsatisfiable
            }
        }
      else
        { return SATISFIABLE; }                         // All variables assigned, problem satisfiable
    }
}
```

Not explicitly stated: Inprocessing, unlearning, restarts, model output

Decision Heuristics

- Have the role of choosing the next decision variable
- Comparable with “case distinction” in the DLL algorithm
- Affects the search process significantly
- Modern SAT algorithms do not test whether the CNF formula is already satisfied during the search, rather it is indirectly guaranteed from assigning all variables without running into a conflict
  - Example: $F = (x_1, x_2, x_3) \wedge (\neg x_1, x_4)$

  $\Rightarrow$ A satisfying assignment is for example $x_1 = 1, x_4 = 1$

  $\Rightarrow$ Today’s solvers do no test whether $x_1 = x_4 = 1$ already satisfies all the clauses, but assign the remaining variables without generating a conflict (e. g., $x_2 = x_3 = 0$) before they conclude that the CNF is satisfiable

Decision Heuristics

Classical decision heuristics

- Several flavors
  - Dynamic Largest Individual/Combined Sum
  - Maximum Occurrences on Clauses of Minimal Size
- Choice criteria
  - "How often does a still unassigned variable occur in currently unresolved clauses?"
  - Among the unassigned variables, choose the one that occurs most frequently in unresolved clauses
  - In most cases also weighted with the length of those clauses
- These heuristics are quite time consuming, because both the status of each clause and the distribution of the variables within the set of clauses have to be computed and kept up to date

  $\Rightarrow$ Computation complexity defined over #clauses

Decision Heuristics

Variable State Independent Decaying Sum (VSIDS)

- Today's standard method used by almost every SAT solver
- Computation complexity defined over #variables
- No update is mandatory during the backtrack phase
- Each variable x_i has two activity counters P_{x_i} and N_{x_i}
- For each literal L in a learned clause C the activity is incremented as follows:

$$\begin{aligned} P_{x_i} &= P_{x_i} + 1, \text{ if } L = x_i \\ N_{x_i} &= N_{x_i} + 1, \text{ if } L = \neg x_i \end{aligned}$$

- The unassigned variable x_i with the highest activity (P_{x_i} or N_{x_i}) is chosen as the next decision variable
- Polarity depends on whether $P_{x_i} > N_{x_i}$ holds or not

Decision Heuristics

Variable State Independent Decaying Sum (VSIDS)

- Periodically, the activities are "normalized", i. e., divided by a constant factor
  - ⇒ After the normalization, the recently learned clauses have a higher weight in comparison to the clauses learned before the last normalization process
  - ⇒ Takes into account the "history" of the search process
- Several optimizations possible
  - By which amount should the activities be incremented?
  - How often should the normalization take place?
  - By which factor should the activity scores be divided?

Modern SAT Algorithms

```
bool SequentialSatEngine(CNF F)
{
  if (PreprocessCNF(F) == CONFLICT)                        // Preprocessing the CNF formula
    { return UNSATISFIABLE; }                              // Problem unsatisfiable
  while (true)
    {
      if (DecideNextBranch())                              // Choice of the next unassigned variable
        {
          while (BCP() == CONFLICT)                        // Boolean Constraint Propagation
            {
              BLevel = AnalyzeConflict();                  // Conflict analysis
              if (BLevel > 0)
                { Backtrack(BLevel); }                     // Cancel the „incorrect“ assignment
              else
                { return UNSATISFIABLE; }                  // Problem unsatisfiable
            }
        }
      else
        { return SATISFIABLE; }                            // All variables assigned, problem satisfiable
    }
}
```

Not explicitly stated: Inprocessing, unlearning, restarts, model output

Boolean Constraint Propagation

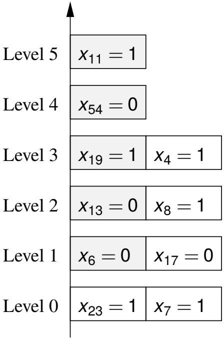
- Tasks
  - Detect all implications forced by a variable assignment
  - Detect conflicts
- Comparable to the repeated application of the unit clause rule of the DLL algorithm
- Efficient implementation mandatory, because roughly 80% of the runtime of a SAT algorithm is spent by the BCP routine

Boolean Constraint Propagation

General flow

- After every variable assignment, identify the implications that have arisen, and push them into the implication queue
- As long as there are items in the implication queue...
  1. Remove the first element from the queue
  2. Assign to each implied variable its forced truth value
  3. Check which consecutive implications arise, and push them into the implication queue
  4. Check for conflicts

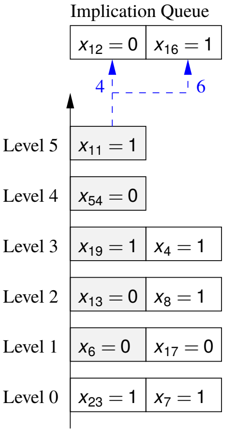
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

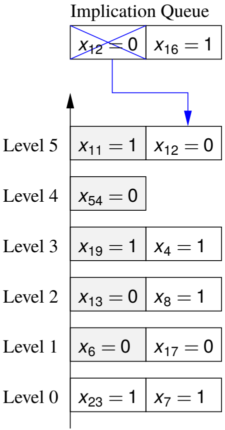
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

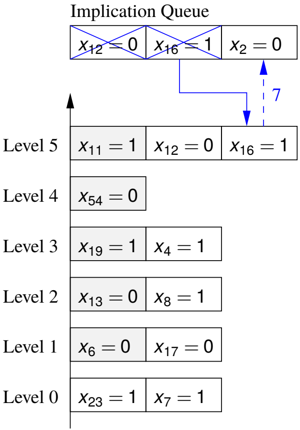
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

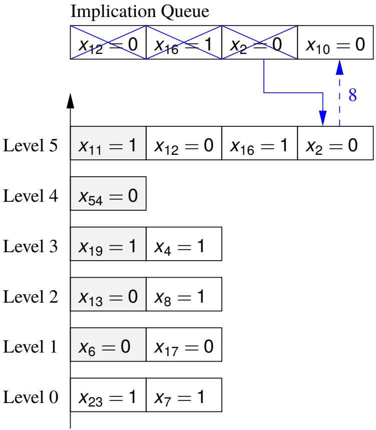
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

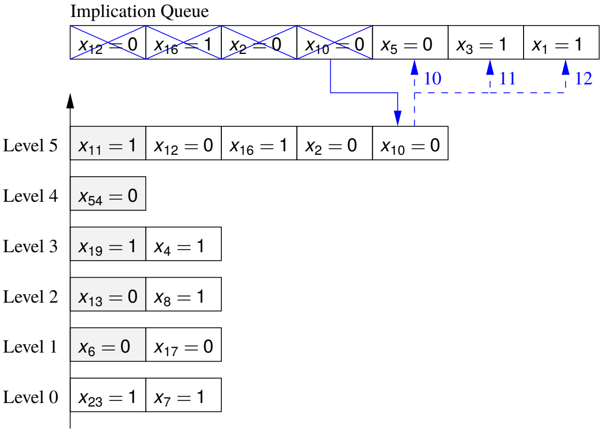
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

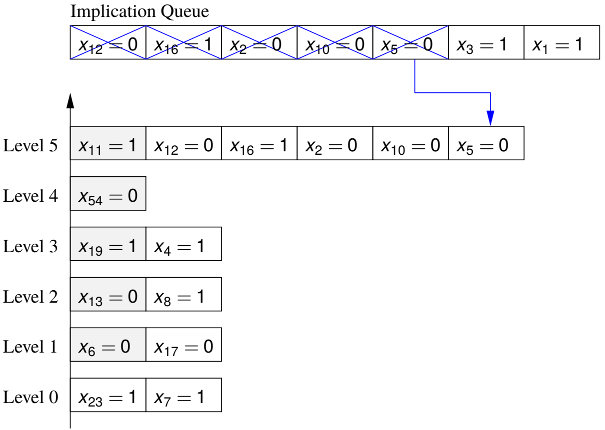
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

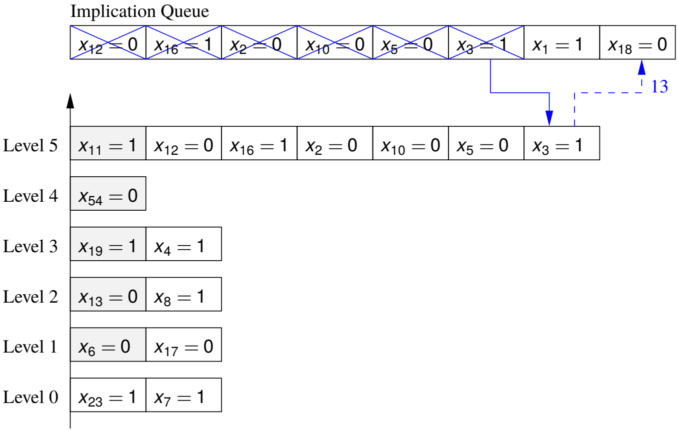
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

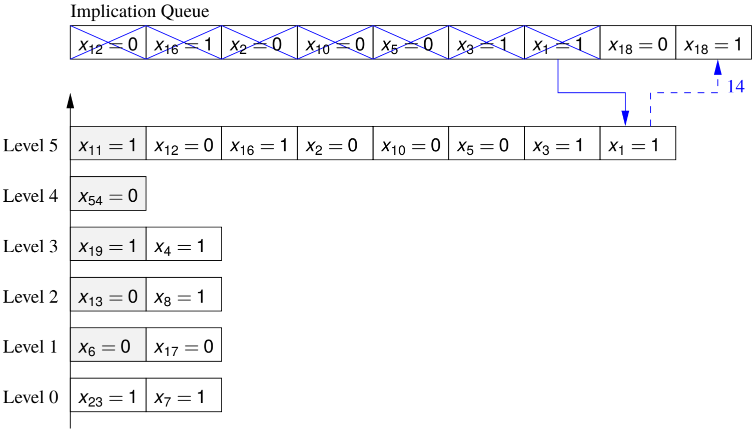
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

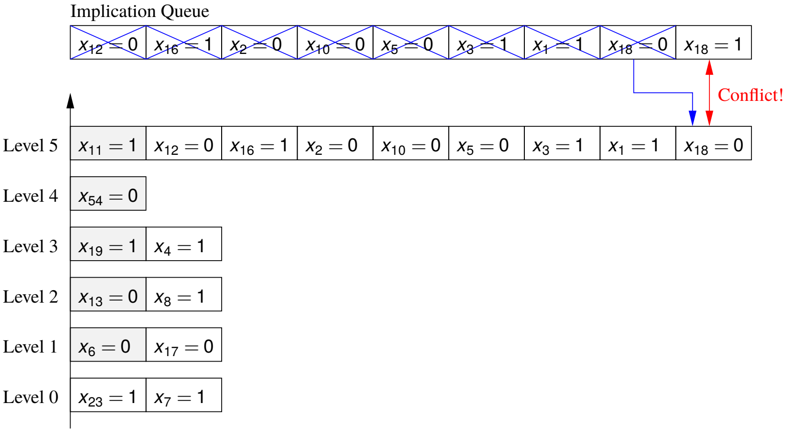
Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

Boolean Constraint Propagation – Example



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

Boolean Constraint Propagation

Approaches for the implementation of a BCP routine

- Counter-Based Schemes
- Watched Literals / 2-Literal Watching Scheme

Boolean Constraint Propagation

Counter-Based Schemes

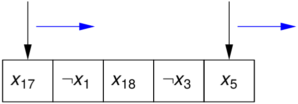
- 2-Counter Scheme
  - Two counters for each clause
    - One counter for the literals which satisfy the clause
    - One counter for the unassigned literals
- 1-Counter Scheme
  - One counter for each clause to count the number of not falsifying literals
- Disadvantages
  - “Unnecessary” counter updates
  - Adjustment of the counter values during backtrack
  - Requires a list for each variable and polarity to store all the clauses where the “related literal” (variable having that polarity) occurs

Boolean Constraint Propagation

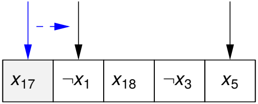
Watched Literals

- For each clause mark two different literals
- Invariant
  - Watched literals of a clause are either unassigned or satisfy the clause
- Advantages in comparison to counter-based schemes
  - Update operations only when necessary, i. e., when an assignment “breaks” the invariant
  - One list for each variable and polarity (like before), but containing only the clauses currently watched by that literal
- Disadvantage
  - Literals of a clause are checked several times

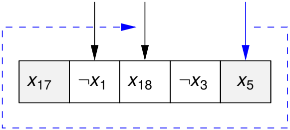
Watched Literals



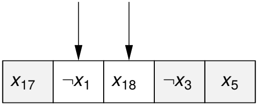
(a) Initial state



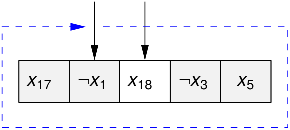
(b) $x_{17} = 0$



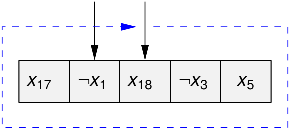
(c) $x_5 = 0$



(d) $x_3 = 1$



(e) $x_1 = 1 \Rightarrow x_{18} = 1$



(f) $x_{18} = 0 \Rightarrow$ Conflict!

Watched Literals

Possible optimizations

- Always store the watched literals in the first two positions of a clause
  - Allows for a fast access to the “second” watched literal of a clause
  - If the second watched literal satisfies the clause, it is not necessary to find a replacement for the first one (in case the status of the first one switches from unresolved to false)

Nowadays, the BCP procedures of almost all modern SAT solvers are based on watched literals!

Modern SAT Algorithms

```
bool SequentialSatEngine(CNF F)
{
  if (PreprocessCNF(F) == CONFLICT)                          // Preprocessing the CNF formula
    { return UNSATISFIABLE; }                                // Problem unsatisfiable
  while (true)
    {
      if (DecideNextBranch())                                // Choice of the next unassigned variable
        {
          while (BCP() == CONFLICT)                          // Boolean Constraint Propagation
            {
              BLevel = AnalyzeConflict();                    // Conflict analysis
              if (BLevel > 0)
                { Backtrack(BLevel); }                       // Cancel the „incorrect“ assignment
              else
                { return UNSATISFIABLE; }                    // Problem unsatisfiable
            }
        }
      else
        { return SATISFIABLE; }                              // All variables assigned, problem satisfiable
    }
}
```

Not explicitly stated: Inprocessing, unlearning, restarts, model output

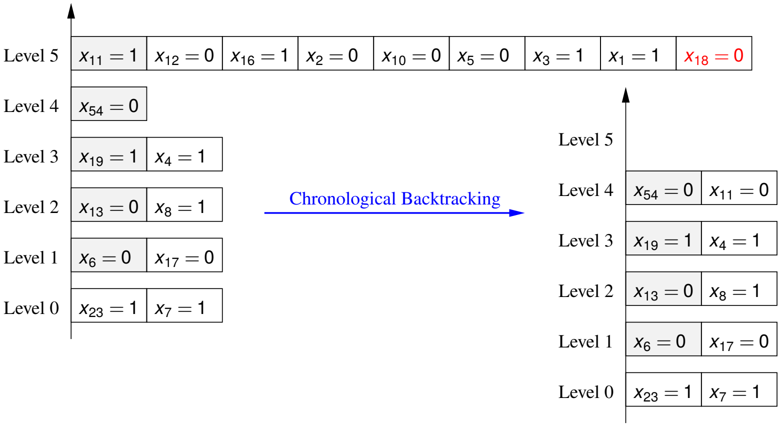
Conflict Analysis & Backtracking

DLL algorithm

- The combination of the decisions done before will always be considered as the origin of a conflict
- Backtracking to the recursion level of the last “branching” in which one case for a variable assignment has not been explored yet (chronological backtracking)
- If such a recursion level does not exist, the given CNF formula is unsatisfiable

UNI FREIBURG

Conflict Analysis & Backtracking



$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

Conflict Analysis & Backtracking

Modern SAT algorithms

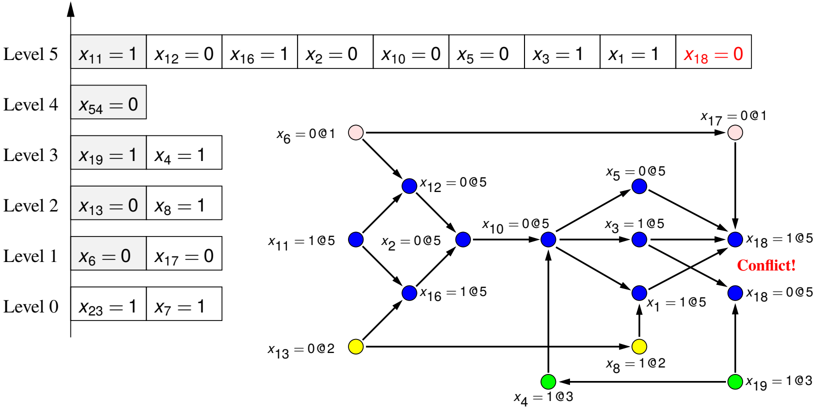
- Complex analysis of the conflict setting, because not all "branchings" done before have to be involved in the current conflict
- Learning of a conflict clause via resolution to avoid running into the same conflict again
- (Non-)chronological backtracking according to the derived conflict clause
- If a conflict occurs on decision level 0, the given CNF formula is unsatisfiable

Conflict Analysis & Backtracking

Implication graph

- Data structure for performing the conflict analysis in today's SAT solvers
- Directed, acyclic graph
- Nodes represent assignments to variables
- Edges represent which set of assignments have caused an implication
- Implication graph gets updated after every variable assignment and after every backtrack operation

Conflict Analysis & Backtracking



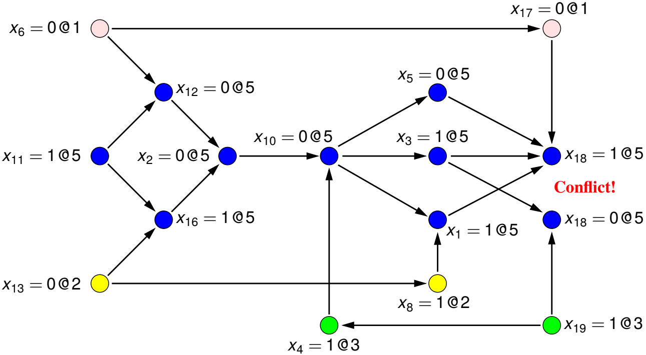
$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

Conflict Analysis & Backtracking

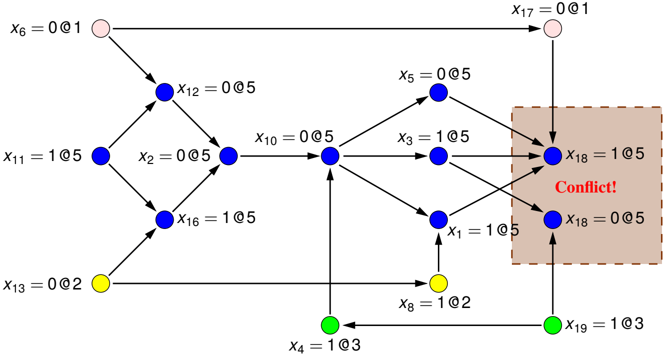
- During the conflict analysis the implication graph gets traversed backwards (in reverse order of the assignments stored by the decision stack) starting from the conflicting point, to allow to compute the succession of resolution steps which finally lead to the conflict clause
- Different termination criteria for interrupting the resolution steps lead to different conflict clauses
- Implementations
  - 1UIP (standard technique explained in the following)
  - RelSat
  - Grasp
  - ...

Conflict Analysis & Backtracking



$$F = (x_{23}) \wedge (x_7, \neg x_{23}) \wedge (x_6, \neg x_{17}) \wedge (x_6, \neg x_{11}, \neg x_{12}) \wedge (x_{13}, x_8) \wedge (\neg x_{11}, x_{13}, x_{16}) \wedge (x_{12}, \neg x_{16}, \neg x_2) \wedge (x_2, \neg x_4, \neg x_{10}) \wedge$$
$$(\neg x_{19}, x_4) \wedge (x_{10}, \neg x_5) \wedge (x_{10}, x_3) \wedge (x_{10}, \neg x_8, x_1) \wedge (\neg x_{19}, \neg x_{18}, \neg x_3) \wedge (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \wedge \ldots$$

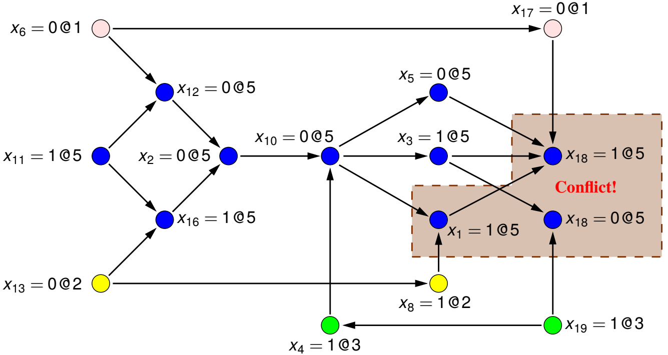
Conflict Analysis & Backtracking



$F = (x_{23}) \wedge (x_7, \neg x_{23}) \wedge (x_6, \neg x_{17}) \wedge (x_6, \neg x_{11}, \neg x_{12}) \wedge (x_{13}, x_8) \wedge (\neg x_{11}, x_{13}, x_{16}) \wedge (x_{12}, \neg x_{16}, \neg x_2) \wedge (x_2, \neg x_4, \neg x_{10}) \wedge$
$(\neg x_{19}, x_4) \wedge (x_{10}, \neg x_5) \wedge (x_{10}, x_3) \wedge (x_{10}, \neg x_8, x_1) \wedge (\neg x_{19}, \neg x_{18}, \neg x_3) \wedge (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \wedge \ldots$

$R_1 = (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \otimes_{x_{18}} (\neg x_{19}, \neg x_{18}, \neg x_3) = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19})$

Conflict Analysis & Backtracking

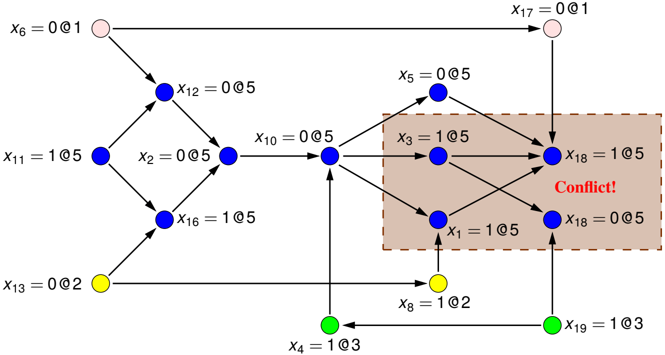


$F = (x_{23}) \wedge (x_7, \neg x_{23}) \wedge (x_6, \neg x_{17}) \wedge (x_6, \neg x_{11}, \neg x_{12}) \wedge (x_{13}, x_8) \wedge (\neg x_{11}, x_{13}, x_{16}) \wedge (x_{12}, \neg x_{16}, \neg x_2) \wedge (x_2, \neg x_4, \neg x_{10}) \wedge$
$(\neg x_{19}, x_4) \wedge (x_{10}, \neg x_5) \wedge (x_{10}, x_3) \wedge (x_{10}, \neg x_8, x_1) \wedge (\neg x_{19}, \neg x_{18}, \neg x_3) \wedge (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \wedge \ldots$

$R_1 = (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \otimes_{x_{18}} (\neg x_{19}, \neg x_{18}, \neg x_3) = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19})$

$R_2 = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19}) \otimes_{x_1} (x_1, x_{10}, \neg x_8) = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8)$

Conflict Analysis & Backtracking



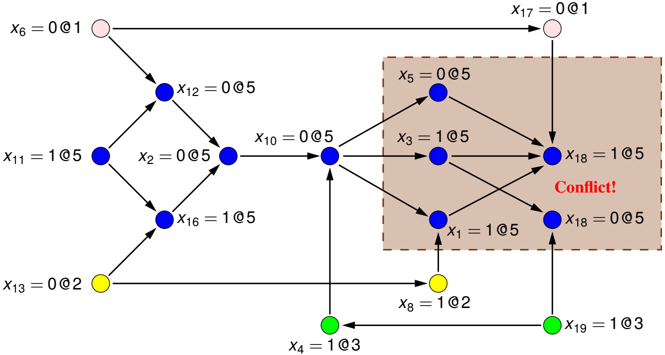
$F = (x_{23}) \wedge (x_7, \neg x_{23}) \wedge (x_6, \neg x_{17}) \wedge (x_6, \neg x_{11}, \neg x_{12}) \wedge (x_{13}, x_8) \wedge (\neg x_{11}, x_{13}, x_{16}) \wedge (x_{12}, \neg x_{16}, \neg x_2) \wedge (x_2, \neg x_4, \neg x_{10}) \wedge$
$(\neg x_{19}, x_4) \wedge (x_{10}, \neg x_5) \wedge (x_{10}, x_3) \wedge (x_{10}, \neg x_8, x_1) \wedge (\neg x_{19}, \neg x_{18}, \neg x_3) \wedge (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \wedge \ldots$

$R_1 = (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \otimes_{x_{18}} (\neg x_{19}, \neg x_{18}, \neg x_3) = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19})$

$R_2 = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19}) \otimes_{x_1} (x_1, x_{10}, \neg x_8) = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8)$

$R_3 = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8) \otimes_{x_3} (x_{10}, x_3) = (x_{17}, x_5, \neg x_{19}, x_{10}, \neg x_8)$

Conflict Analysis & Backtracking



$F = (x_{23}) \wedge (x_7, \neg x_{23}) \wedge (x_6, \neg x_{17}) \wedge (x_6, \neg x_{11}, \neg x_{12}) \wedge (x_{13}, x_8) \wedge (\neg x_{11}, x_{13}, x_{16}) \wedge (x_{12}, \neg x_{16}, \neg x_2) \wedge (x_2, \neg x_4, \neg x_{10}) \wedge$
$(\neg x_{19}, x_4) \wedge (x_{10}, \neg x_5) \wedge (x_{10}, x_3) \wedge (x_{10}, \neg x_8, x_1) \wedge (\neg x_{19}, \neg x_{18}, \neg x_3) \wedge (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \wedge \ldots$

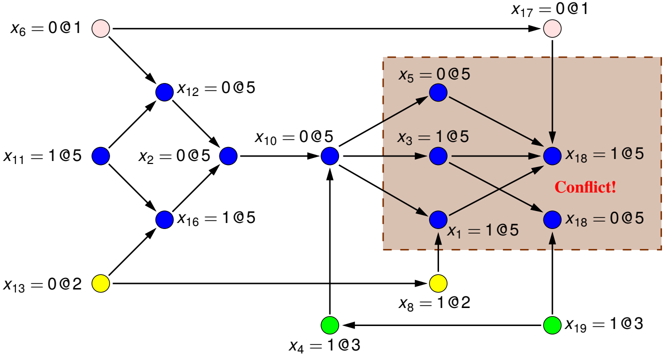
$R_1 = (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \otimes_{x_{18}} (\neg x_{19}, \neg x_{18}, \neg x_3) = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19})$

$R_2 = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19}) \otimes_{x_1} (x_1, x_{10}, \neg x_8) = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8)$

$R_3 = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8) \otimes_{x_3} (x_{10}, x_3) = (x_{17}, x_5, \neg x_{19}, x_{10}, \neg x_8)$

$R_4 = (x_{17}, x_5, \neg x_{19}, x_{10}, \neg x_8) \otimes_{x_5} (x_{10}, \neg x_5) = (x_{17}, \neg x_{19}, x_{10}, \neg x_8)$

Conflict Analysis & Backtracking



$F = (x_{23}) \wedge (x_7, \neg x_{23}) \wedge (x_6, \neg x_{17}) \wedge (x_6, \neg x_{11}, \neg x_{12}) \wedge (x_{13}, x_8) \wedge (\neg x_{11}, x_{13}, x_{16}) \wedge (x_{12}, \neg x_{16}, \neg x_2) \wedge (x_2, \neg x_4, \neg x_{10}) \wedge$
$(\neg x_{19}, x_4) \wedge (x_{10}, \neg x_5) \wedge (x_{10}, x_3) \wedge (x_{10}, \neg x_8, x_1) \wedge (\neg x_{19}, \neg x_{18}, \neg x_3) \wedge (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \wedge \ldots$

$R_1 = (x_{17}, \neg x_1, x_{18}, \neg x_3, x_5) \otimes_{x_{18}} (\neg x_{19}, \neg x_{18}, \neg x_3) = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19})$

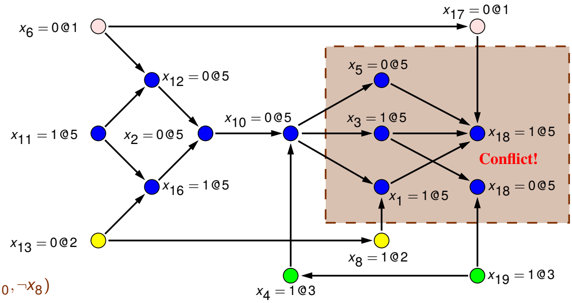
$R_2 = (x_{17}, \neg x_1, \neg x_3, x_5, \neg x_{19}) \otimes_{x_1} (x_1, x_{10}, \neg x_8) = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8)$

$R_3 = (x_{17}, \neg x_3, x_5, \neg x_{19}, x_{10}, \neg x_8) \otimes_{x_3} (x_{10}, x_3) = (x_{17}, x_5, \neg x_{19}, x_{10}, \neg x_8)$

$R_4 = (x_{17}, x_5, \neg x_{19}, x_{10}, \neg x_8) \otimes_{x_5} (x_{10}, \neg x_5) = (x_{17}, \neg x_{19}, x_{10}, \neg x_8) \Leftarrow$ Final conflict clause

Conflict Analysis & Backtracking

Level									
Level 5	$x_{11}=1$	$x_{12}=0$	$x_{16}=1$	$x_2=0$	$x_{10}=0$	$x_5=0$	$x_3=1$	$x_1=1$	$x_{18}=0$
Level 4	$x_{54}=0$								
Level 3	$x_{19}=1$	$x_4=1$							
Level 2	$x_{13}=0$	$x_8=1$							
Level 1	$x_6=0$	$x_{17}=0$							
Level 0	$x_{23}=1$	$x_7=1$							



Conflict clause: $(x_{17}, \neg x_{19}, x_{10}, \neg x_8)$

$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

UNI FREIBURG

Conflict Analysis & Backtracking

Observations

- Conflict analysis according to the 1UIP scheme (First Unique Implication Point) terminates as soon as the computed resolvent contains exactly one literal at the current decision level (the so-called UIP), whereas all other literals were assigned at lower decision levels
- Conflict clauses represent combinations of variables that will inevitably lead to a conflict
- Resolution Lemma allows to insert a conflict clause into the CNF formula, and consequently to “prune” the whole search tree by preventing the solver from running into the same conflict again
- Compared to others, the 1UIP scheme turned out to be the most powerful one (shorter conflict clauses, more effective pruning, faster runtime)

Conflict Analysis & Backtracking

(Non)-chronological backtracking

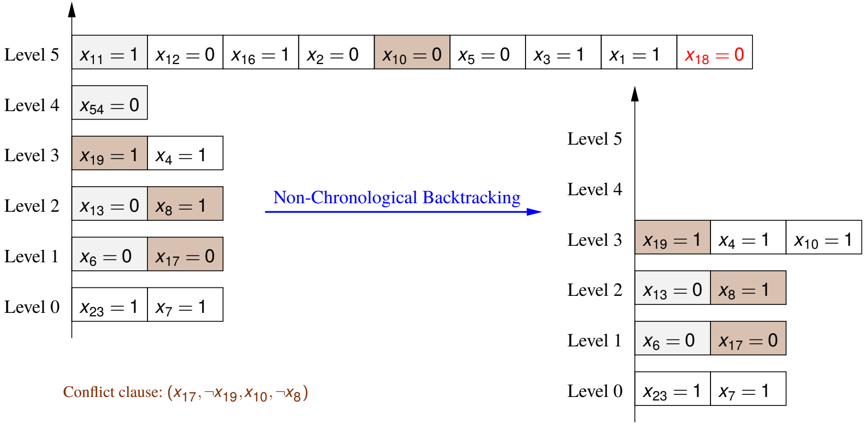
- In today's SAT algorithms the backtrack level is determined by the derived conflict clause only
- The backtrack level matches the maximum decision level among all the literals in the conflict clause except the UIP, which becomes an implication after backtracking
- Idea: "What would have happened if the conflict clause had already been contained into the original CNF formula?"

Conflict Analysis & Backtracking

(Non-)chronological backtracking

- Procedure
  1. Backtrack down to the given backtrack level
  2. Assign the truth value implied by the UIP (after backtracking, the conflict clause will be automatically a unit clause)
  3. Proceed with the search process
- If a conflict appears at decision level 0, the CNF formula is unsatisfiable

Conflict Analysis & Backtracking



Conflict clause: $(x_{17}, \neg x_{19}, x_{10}, \neg x_8)$

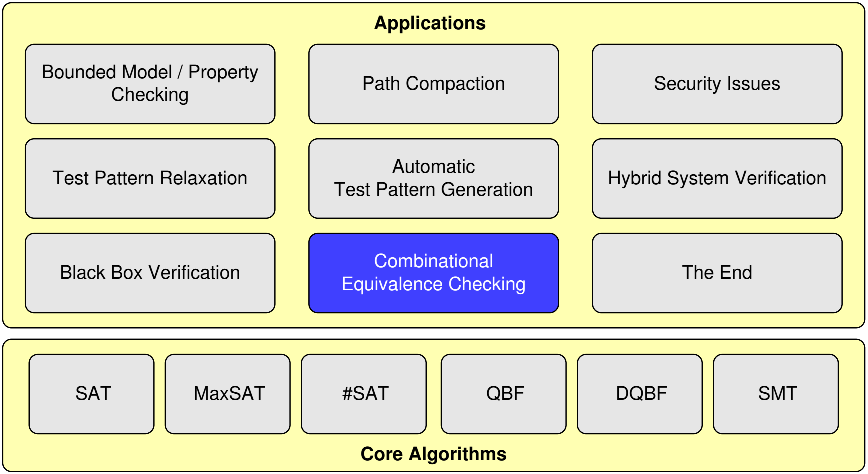
$$F = \underbrace{(x_{23})}_{1} \wedge \underbrace{(x_7, \neg x_{23})}_{2} \wedge \underbrace{(x_6, \neg x_{17})}_{3} \wedge \underbrace{(x_6, \neg x_{11}, \neg x_{12})}_{4} \wedge \underbrace{(x_{13}, x_8)}_{5} \wedge \underbrace{(\neg x_{11}, x_{13}, x_{16})}_{6} \wedge \underbrace{(x_{12}, \neg x_{16}, \neg x_2)}_{7} \wedge \underbrace{(x_2, \neg x_4, \neg x_{10})}_{8} \wedge$$

$$\underbrace{(\neg x_{19}, x_4)}_{9} \wedge \underbrace{(x_{10}, \neg x_5)}_{10} \wedge \underbrace{(x_{10}, x_3)}_{11} \wedge \underbrace{(x_{10}, \neg x_8, x_1)}_{12} \wedge \underbrace{(\neg x_{19}, \neg x_{18}, \neg x_3)}_{13} \wedge \underbrace{(x_{17}, \neg x_1, x_{18}, \neg x_3, x_5)}_{14} \wedge \ldots$$

Other Features of modern SAT Solvers

- Unlearning of conflict clauses
- Inprocessing
- Restarts
- Termination guarantees
- Unsatisfiability certificates
- Assumptions
- Incremental SAT solving
- Parallel SAT algorithms
- Incomplete SAT algorithms

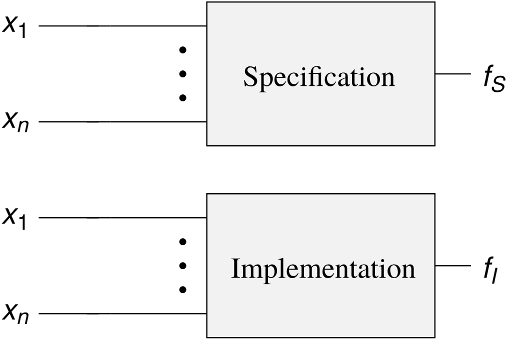
Outline


Applications
Bounded Model / Property Checking
Path Compaction
Security Issues
Test Pattern Relaxation
Automatic Test Pattern Generation
Hybrid System Verification
Black Box Verification
Combinational Equivalence Checking
The End
SAT
MaxSAT
#SAT
QBF
DQBF
SMT
Core Algorithms

Combinational Equivalence Checking

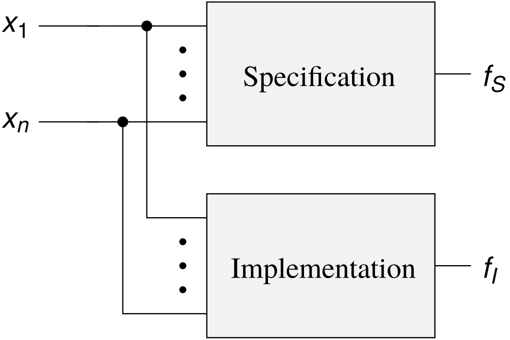
- Given
  - Specification and implementation of a combinatorial circuit
- Question
  - Are specification and implementation equivalent?
- Approach for SAT-based equivalence checking
  - Generate a so-called Miter from specification and implementation
  - Build a CNF formula from the Miter representation
  - Solve the formula with a SAT algorithm
  - Specification and implementation of a combinatorial circuit are equivalent iff the CNF formula generated from the Miter is unsatisfiable

Miter



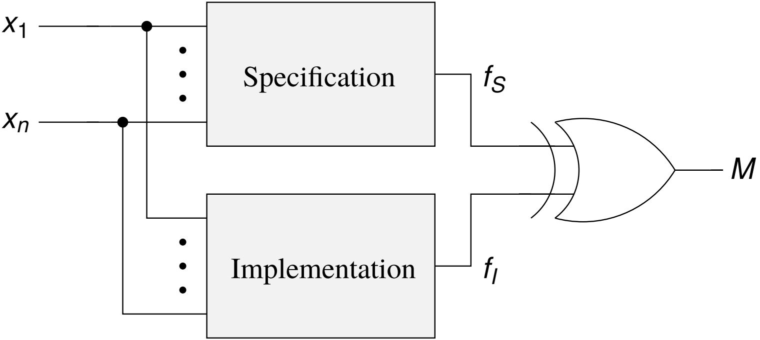
⇒ Connect corresponding inputs

Miter



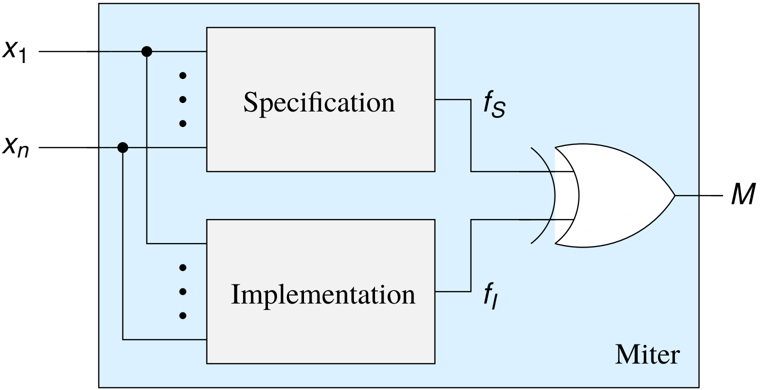
⇒ Link corresponding outputs by EXOR gates

Miter



⇒ Miter circuit

Miter

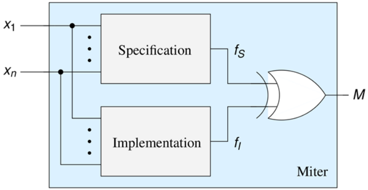


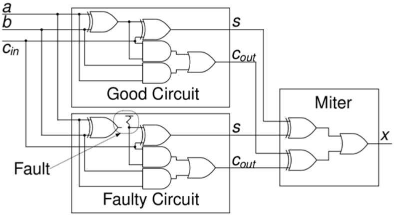
$\Rightarrow M = 1 \Leftrightarrow$ Specification & implementation not equivalent

Miter

Remarks

- Drafted method can be extended to combinatorial circuits having multiple outputs





- Usually, SAT-algorithms take as input only CNF formulas, that means the Boolean function of the Miter circuit must be translated into a CNF representation ⇝ Tseitin transformation

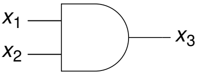
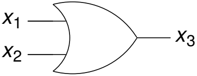
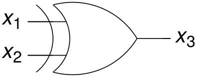
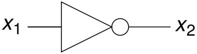
Tseitin Transformation

In order to avoid the exponential size of the CNF form obtained from the formula created from the function F of the circuit, some alternative techniques can be applied:

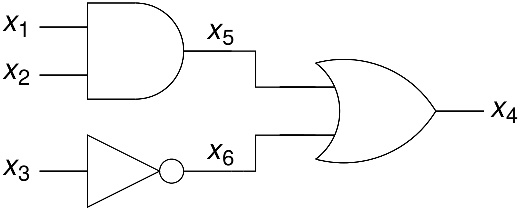
- Define a satisfiability equivalent CNF F' equivalent to F that is satisfiable iff F is satisfiable
- For each gate output insert an additional variable $\Rightarrow$ in general the CNF F' will have variables which do not occur in F
- For each gate realize a “characteristic function” in CNF which evaluates to 1 for every possible consistent signal configuration
- Put together the individual gates using an AND connection to obtain the final CNF formula

$\Rightarrow$ Tseitin transformation

Tseitin Transformation

Gates	Function	CNF formula
	$x_3 \equiv x_1 \wedge x_2$	$(\neg x_3 \vee x_1) \wedge (\neg x_3 \vee x_2) \wedge (x_3 \vee \neg x_1 \vee \neg x_2)$
	$x_3 \equiv x_1 \vee x_2$	$(x_3 \vee \neg x_1) \wedge (x_3 \vee \neg x_2) \wedge (\neg x_3 \vee x_1 \vee x_2)$
	$x_3 \equiv x_1 \oplus x_2$	$(\neg x_3 \vee x_1 \vee x_2) \wedge (\neg x_3 \vee \neg x_1 \vee \neg x_2) \wedge (x_3 \vee \neg x_1 \vee x_2) \wedge (x_3 \vee x_1 \vee \neg x_2)$
	$x_2 \equiv \neg x_1$	$(x_2 \vee x_1) \wedge (\neg x_2 \vee \neg x_1)$

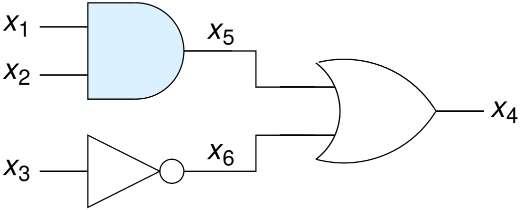
Tseitin Transformation – Example



$$F_{SK} = (x_1 \wedge x_2) \vee \neg x_3$$

$$\begin{aligned} F_{SK}^{CNF} = \; & (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge \\ & (x_6 \vee x_3) \wedge (\neg x_6 \vee \neg x_3) \wedge \\ & (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \end{aligned}$$

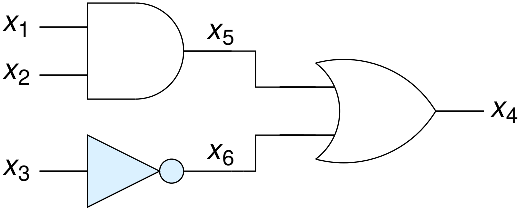
Tseitin Transformation – Example



$$F_{SK} = (x_1 \wedge x_2) \vee \neg x_3$$

$$\begin{aligned} F_{SK}^{CNF} = & (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge \\ & (x_6 \vee x_3) \wedge (\neg x_6 \vee \neg x_3) \wedge \\ & (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \end{aligned}$$

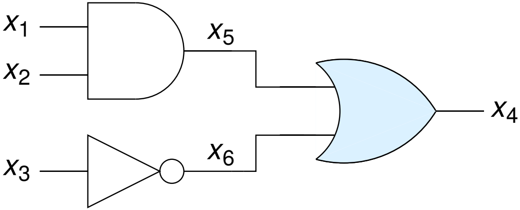
Tseitin Transformation – Example



$$F_{SK} = (x_1 \land x_2) \lor \neg x_3$$

$$\begin{aligned} F_{SK}^{CNF} = & (\neg x_5 \lor x_1) \land (\neg x_5 \lor x_2) \land (x_5 \lor \neg x_1 \lor \neg x_2) \land \\ & (x_6 \lor x_3) \land (\neg x_6 \lor \neg x_3) \land \\ & (x_4 \lor \neg x_5) \land (x_4 \lor \neg x_6) \land (\neg x_4 \lor x_5 \lor x_6) \end{aligned}$$

Tseitin Transformation – Example



$$F_{SK} = (x_1 \wedge x_2) \vee \neg x_3$$

$$\begin{aligned} F_{SK}^{CNF} = & (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge \\ & (x_6 \vee x_3) \wedge (\neg x_6 \vee \neg x_3) \wedge \\ & (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \end{aligned}$$

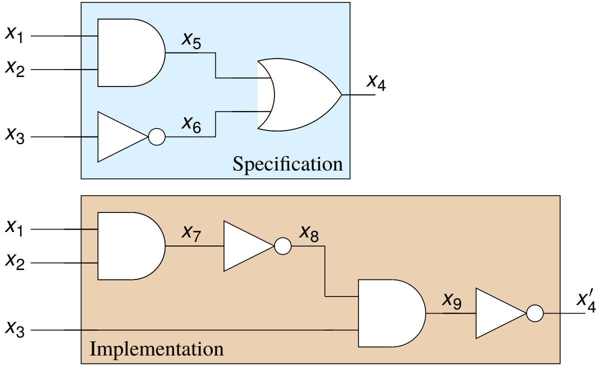
Tseitin Transformation

Important property

- As long as for the CNF representation of each single gate only a constant number of clauses is required, the number of clauses in the final CNF will be linear in the number of gates in the circuit

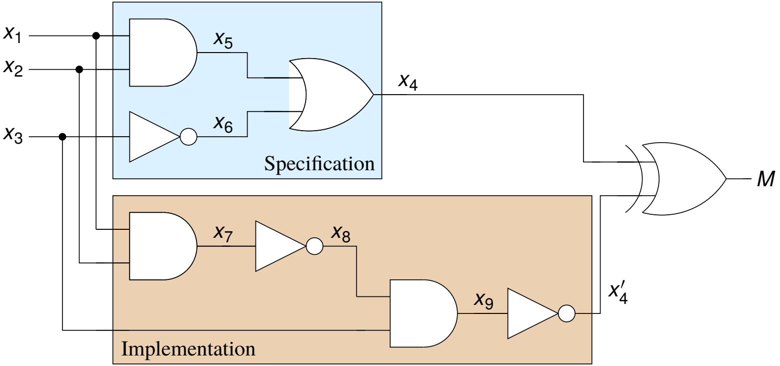
Combinational Equivalence Checking – Example

Let the specification and the implementation of a combinatorial circuit be defined as follows:



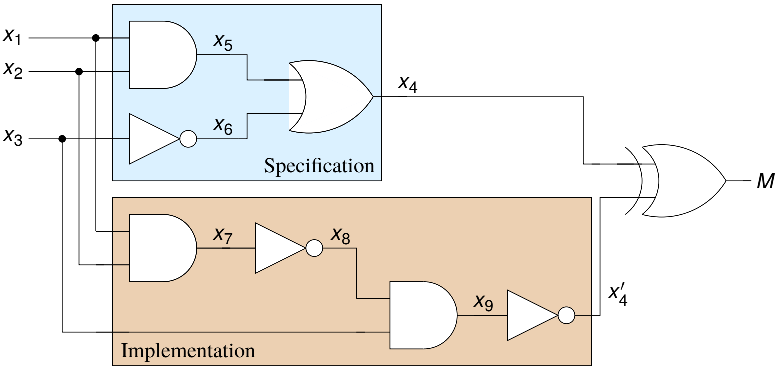
Question: Are the specification and the implementation equivalent?

Combinational Equivalence Checking – Example



$$
\begin{aligned}
F_M = {} & (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge (x_6 \vee x_3) \wedge (\neg x_6 \vee \neg x_3) \wedge \\
& (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \wedge (\neg x_7 \vee x_1) \wedge (\neg x_7 \vee x_2) \wedge \\
& (x_7 \vee \neg x_1 \vee \neg x_2) \wedge (x_7 \vee x_8) \wedge (\neg x_7 \vee \neg x_8) \wedge (\neg x_9 \vee x_3) \wedge (\neg x_9 \vee x_8) \wedge \\
& (x_9 \vee \neg x_3 \vee \neg x_8) \wedge (x_9 \vee x_4') \wedge (\neg x_9 \vee \neg x_4') \wedge (\neg M \vee \neg x_4 \vee \neg x_4') \wedge \\
& (\neg M \vee x_4 \vee x_4') \wedge (M \vee \neg x_4 \vee x_4') \wedge (M \vee x_4 \vee \neg x_4') \wedge (M)
\end{aligned}
$$

Combinational Equivalence Checking – Example



$$
\begin{aligned}
F_M = {} & (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge (x_6 \vee x_3) \wedge (\neg x_6 \vee \neg x_3) \wedge \\
& (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \wedge (\neg x_7 \vee x_1) \wedge (\neg x_7 \vee x_2) \wedge \\
& (x_7 \vee \neg x_1 \vee \neg x_2) \wedge (x_7 \vee x_8) \wedge (\neg x_7 \vee \neg x_8) \wedge (\neg x_9 \vee x_3) \wedge (\neg x_9 \vee x_8) \wedge \\
& (x_9 \vee \neg x_3 \vee \neg x_8) \wedge (x_9 \vee x_4') \wedge (\neg x_9 \vee \neg x_4') \wedge (\neg M \vee \neg x_4 \vee \neg x_4') \wedge \\
& (\neg M \vee x_4 \vee x_4') \wedge (M \vee \neg x_4 \vee x_4') \wedge (M \vee x_4 \vee \neg x_4') \wedge (M)
\end{aligned}
$$

F_M is unsatisfiable ⇒ Implementation and specification are equivalent!

Structural Methods

Nowadays SAT solvers can handle problems with millions of clauses. But how to compare (large) combinatorial circuits for which SAT methods still fail? ⇒ Structural methods

- Solve several “small” problems instead of one “large” problem
- Various options
  - Compute equivalent gates inside the miter circuit
  - And-Inverter-Graphs (AIGs)
  - ...

UNI FREIBURG

Structural Methods

Observation from real-world instances

- In most cases circuits which have to be compared show structural similarities
  - Example: Only small changes in later design phases
  - In many cases logic optimizations respect hierarchy boundaries
  - Thus, changes are not fundamental in most cases

Structural Methods

Observation from real-world instances

- In most cases circuits which have to be compared show structural similarities
  - Example: Only small changes in later design phases
  - In many cases logic optimizations respect hierarchy boundaries
  - Thus, changes are not fundamental in most cases

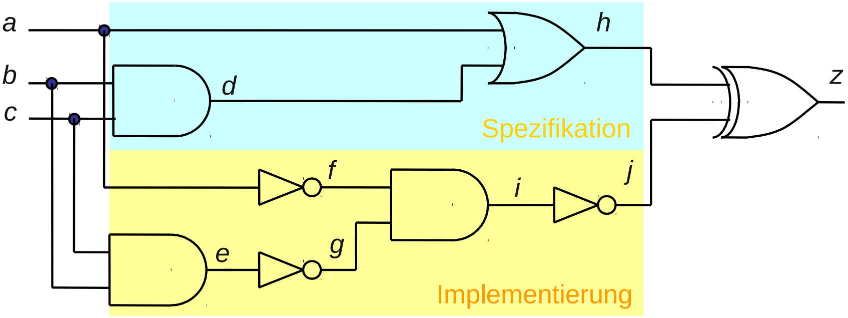
How can we exploit structural similarities?

Structural Methods

Approach

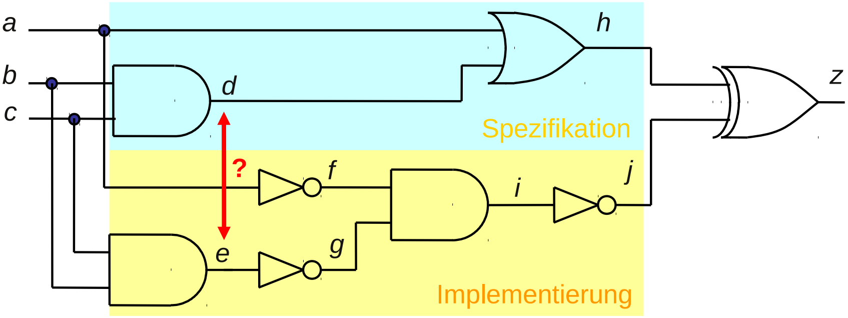
1. Traverse the circuits which have to be compared from inputs to outputs
   - Identify equivalences at the internal signals of the miter
   - If there are any equivalences, replace equivalent nodes by one (shared) representative
2. Check satisfiability of the simplified miter circuit

Structural Methods – Simple Example



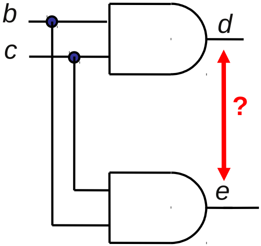
Starting point

Structural Methods – Simple Example



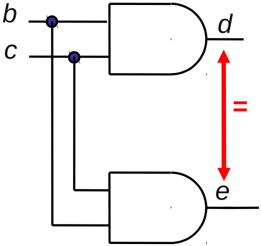
Are the internal signals *d* and *e* equivalent?

Structural Methods – Simple Example



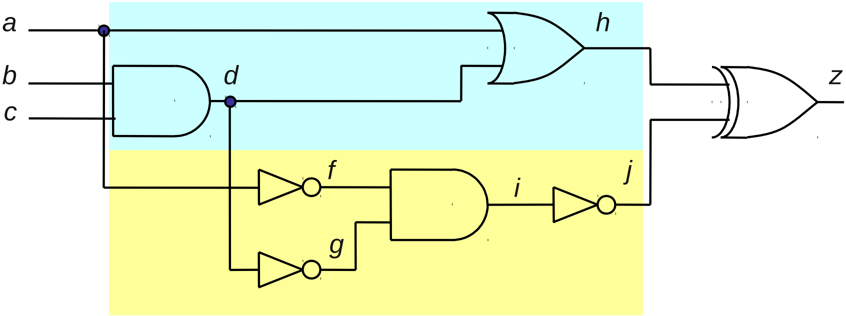
Parts of the miter which are relevant for the proof of $d \equiv e$

Structural Methods – Simple Example



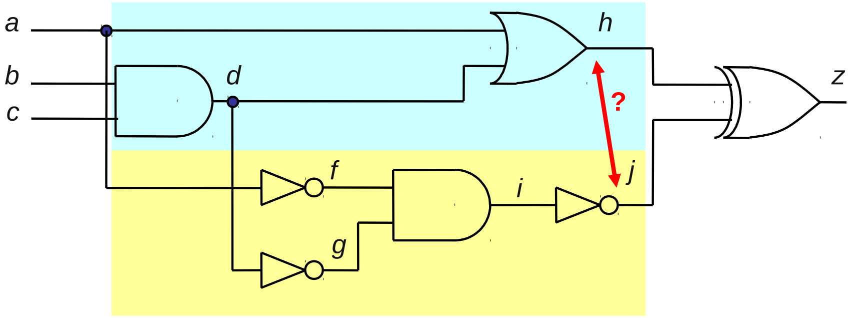
Local analysis is sufficient to show that $d \equiv e$

Structural Methods – Simple Example



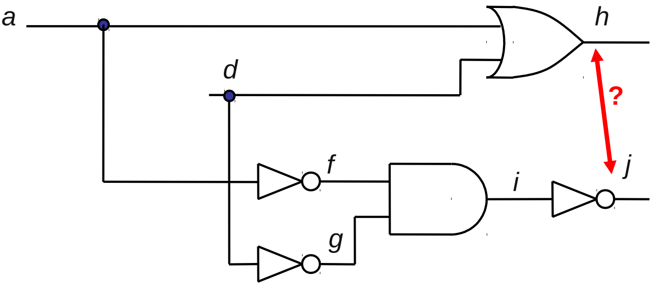
Simplified miter

Structural Methods – Simple Example



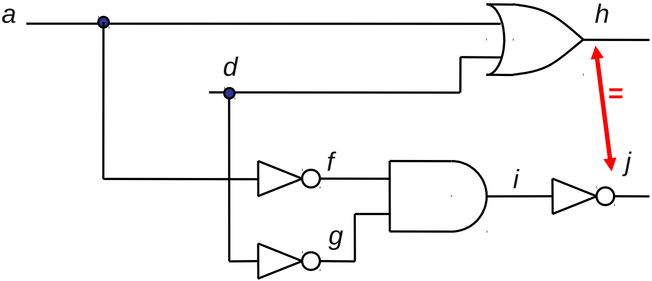
Are the internal signals *h* and *j* equivalent?

Structural Methods – Simple Example



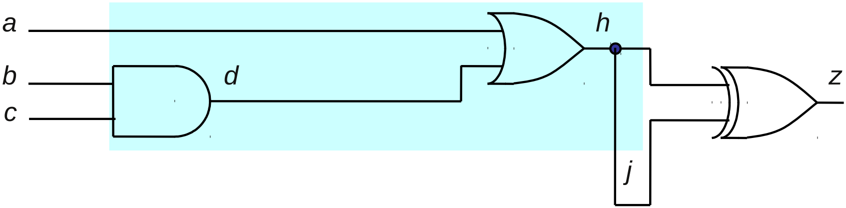
Parts of the miter which are relevant for the proof of $h \equiv j$

Structural Methods – Simple Example



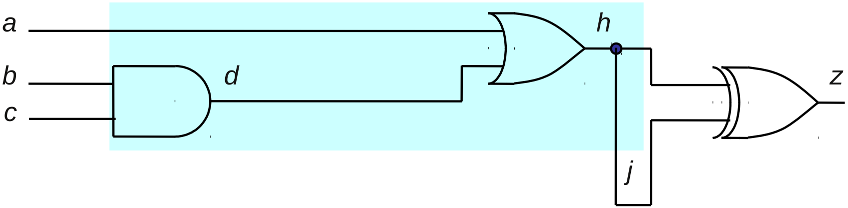
Local analysis is sufficient to show that $h \equiv j$

Structural Methods – Simple Example



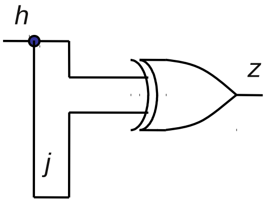
More simplified miter

Structural Methods – Simple Example



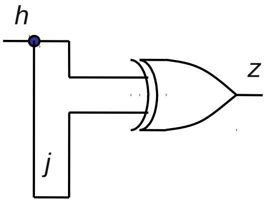
Does $z = 0$ hold? Are specification and implementation equivalent?

Structural Methods – Simple Example



Parts of the miter which are relevant for the proof of $z = 0$

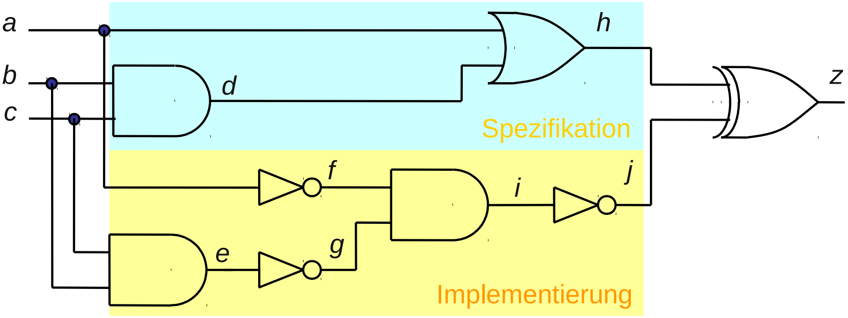
Structural Methods – Simple Example



Local analysis is sufficient to show that $z = 0$

UNI FREIBURG

Structural Methods – Simple Example



⇒ Specification and implementation are equivalent!

Structural Methods – Detection of Equivalences

Detect potential candidates for pairs of equivalent nodes by simulation with random patterns

- By an (incomplete) simulation of a restricted number of patterns we can only show “non-equivalence”
- Use simulation to partition the nodes into equivalence classes which consist of the nodes with identical simulation results
- Use a complete method (e.g. SAT) to detect equivalent nodes within the computed equivalence classes

Structural Methods – Detection of Equivalences

Using SAT to prove equivalences

- In order to keep the miter circuit “small”, the inputs of the SAT problem are not necessarily primary inputs, but rather equivalent internal nodes which have already been detected to be equivalent
- Two nodes are equivalent, if the SAT instance representing the corresponding miter is unsatisfiable
- If two nodes are proved to be equivalent, then one of the nodes may be replaced by its equivalent counterpart
- Be careful: If the SAT instance is satisfiable, then this does not necessarily mean that the corresponding nodes are not equivalent!

Structural Methods – Detection of Equivalences

Equivalent nodes can be used as so-called cut points after they have been replaced by a common representative

- Cut points will be new input variables during miter construction and thus keep the miter “small”
- If the resulting circuits are equivalent, then the original circuits have already been equivalent

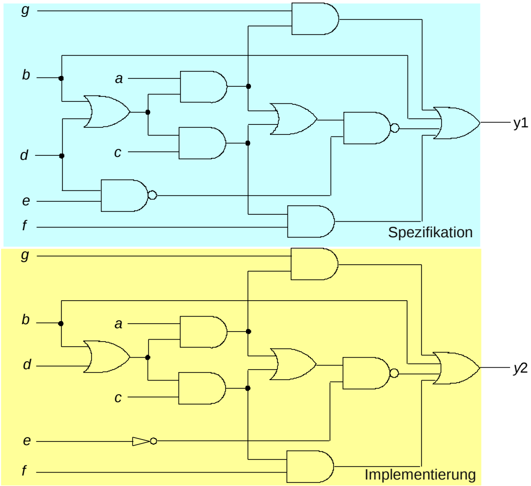
Structural Methods – Detection of Equivalences

Equivalent nodes can be used as so-called cut points after they have been replaced by a common representative

- Cut points will be new input variables during miter construction and thus keep the miter “small”
- If the resulting circuits are equivalent, then the original circuits have already been equivalent

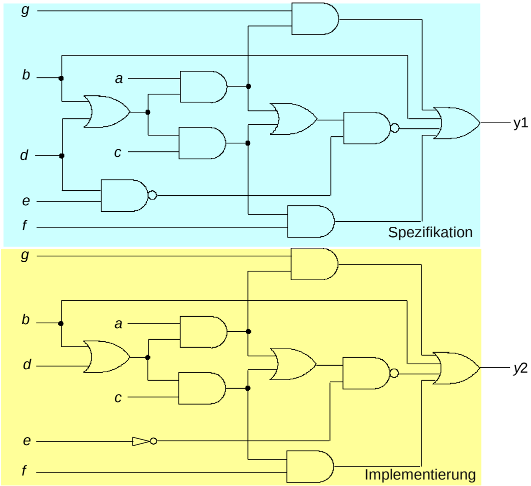
  Problem: Using cut points may lead to so-called “false negatives”, i.e., two equivalent nodes are not classified to be equivalent!

Structural Methods – Example



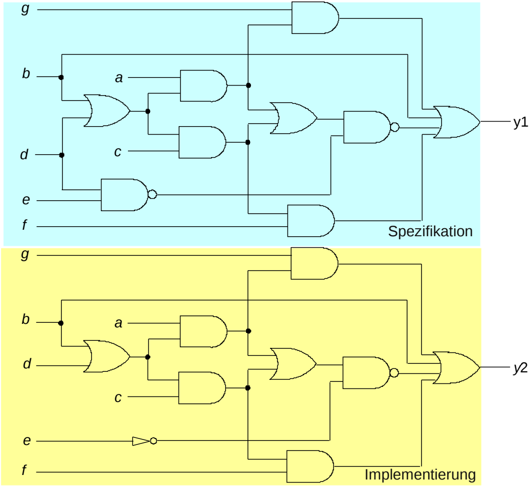
Starting point

Structural Methods – Example



Note: Specification and implementation are equivalent

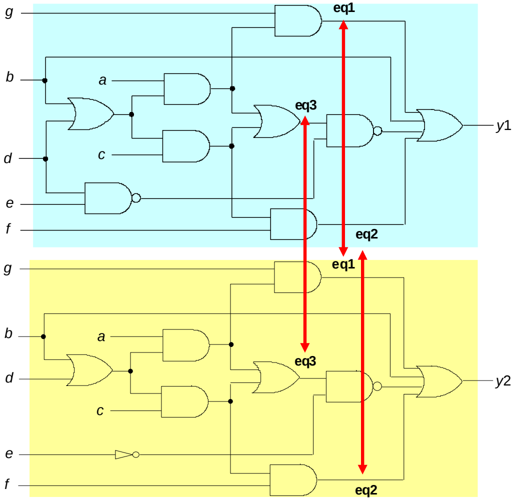
Structural Methods – Example



Try to show equivalence of y_1 and y_2 using cut points

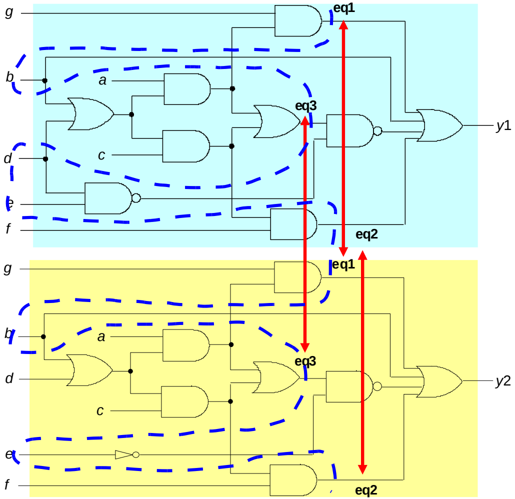
UNI FREIBURG

Structural Methods – Example



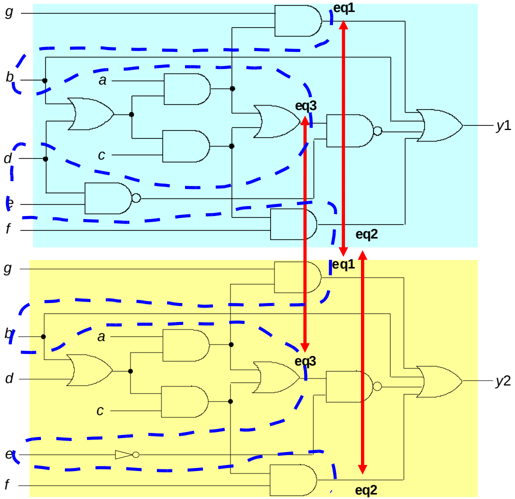
Assumption: Equivalences eq_1, eq_2, and eq_3 already shown

Structural Methods – Example



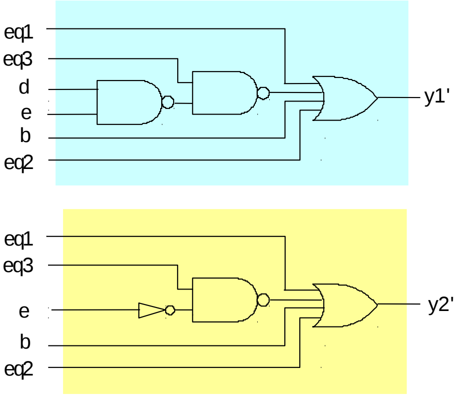
Cut the circuits at the internal equivalent signals

Structural Methods – Example

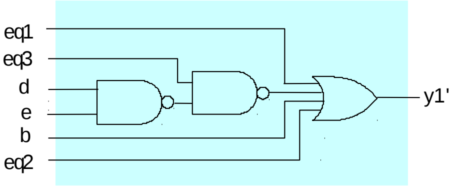


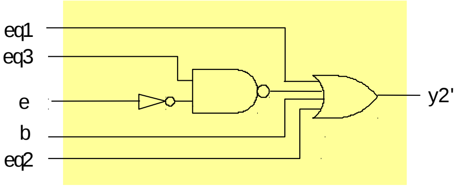
Compute the miter depending on "cut variables"

Structural Methods – Example



Structural Methods – Example





Corresponding CNF formula satisfiable

$\Rightarrow$ y_1 and y_2 not equivalent

$\Rightarrow$ Specification and implementation not equivalent

$\Rightarrow$ But it is a False Negative!

Structural Methods – False Negatives

Problem

- New variables at cut points may be assigned to arbitrary values

But…

- The “rightmost” parts of the circuit need only to be equivalent for values at the cut points which can be produced by the “leftmost” parts

Structural Methods – Avoiding False Negatives

- Do not use cut points
  - Makes proofs of equivalence for two nodes much more difficult in many cases, since the corresponding SAT problems become significantly “larger”
- SAT sweeping
  - In a first step stop at cut points when constructing the miter
  - If necessary (satisfiable CNF) include more parts of the circuit into the SAT problem to check for false negative results

Outline

Applications

Bounded Model / Property Checking

Path Compaction

Security Issues

Test Pattern Relaxation

Automatic Test Pattern Generation

Hybrid System Verification

Black Box Verification

Combinational Equivalence Checking

The End

SAT

MaxSAT

#SAT

QBF

DQBF

SMT

Core Algorithms

Automatic Test Pattern Generation

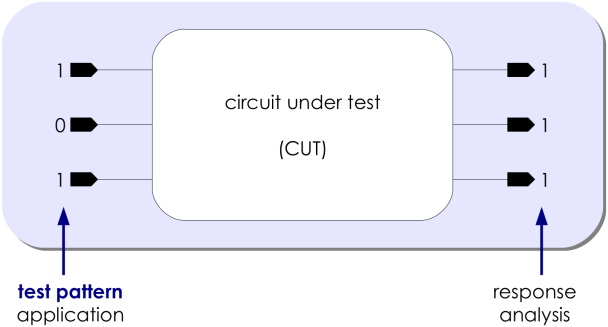
Motivation

- Post-production test is a crucial step
  - Have there been problems during production?
  - Does the circuit contain faults?
- In particular when used in safety-critical applications, every produced chip has to be tested
- Testing comprises more than 40% of costs in semiconductor industry

Automatic Test Pattern Generation

Testing: Experiment on real manufactored chips

- Goal is to check whether the chip behaves correctly
- 1. step: Apply an appropriate test pattern
- 2. step: Analyse the response of the circuit under test



UNI FREIBURG

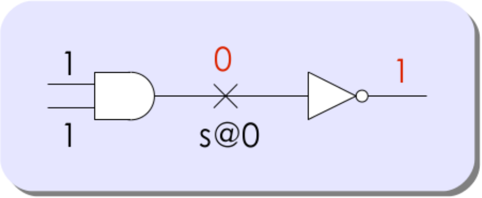
Automatic Test Pattern Generation

- Physical defects are modeled on the Boolean level according to a fault model
- Fault models are an abstract representation of real defects
  - Single stuck-at
  - Bridging faults
  - Interconnect opens
  - Path delay faults
  - ...
- Automatic Test Pattern Generation (ATPG)
  - Given: Circuit *CUT* and fault model *FM*
  - Goal: Determine test patterns for (all) faults in *CUT* wrt. *FM*

Automatic Test Pattern Generation

Single stuck-at fault model (s@)

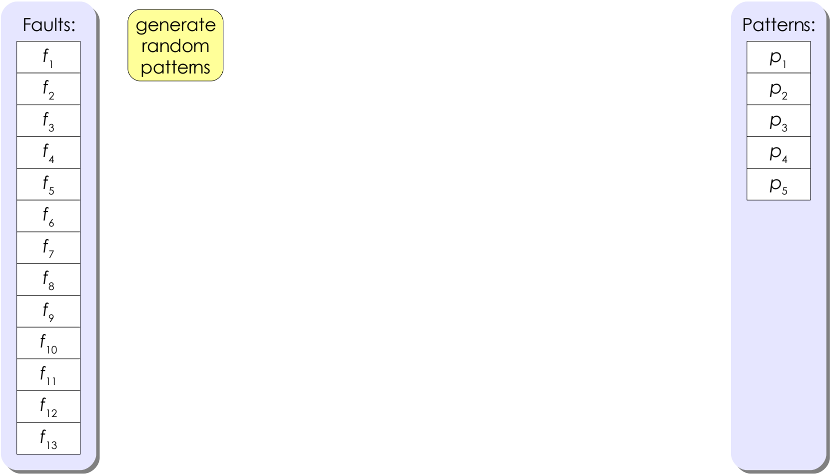
- s@0: One line is always at logic 0
- s@1: One line is always at logic 1
- In total only (2 $\times$ number_of_signals_CUT) faults to be checked
- High amount of real defects detected by the s@ fault model!



Automatic Test Pattern Generation – Typical Flow

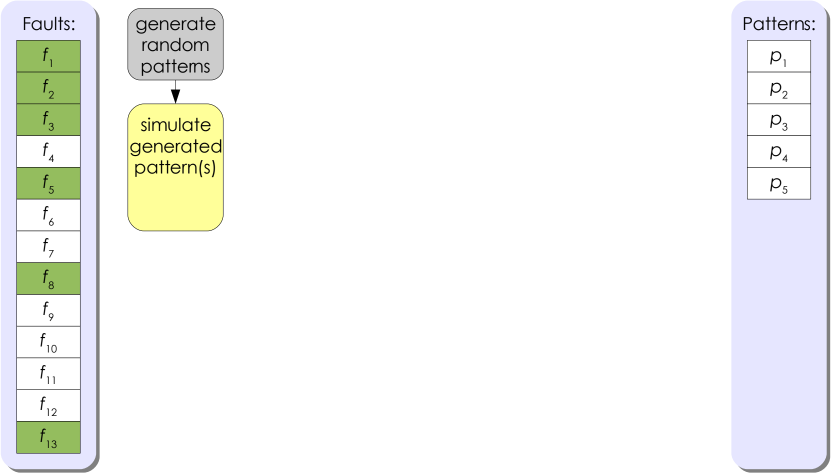
Faults:
f_1
f_2
f_3
f_4
f_5
f_6
f_7
f_8
f_9
f_{10}
f_{11}
f_{12}
f_{13}

Automatic Test Pattern Generation – Typical Flow

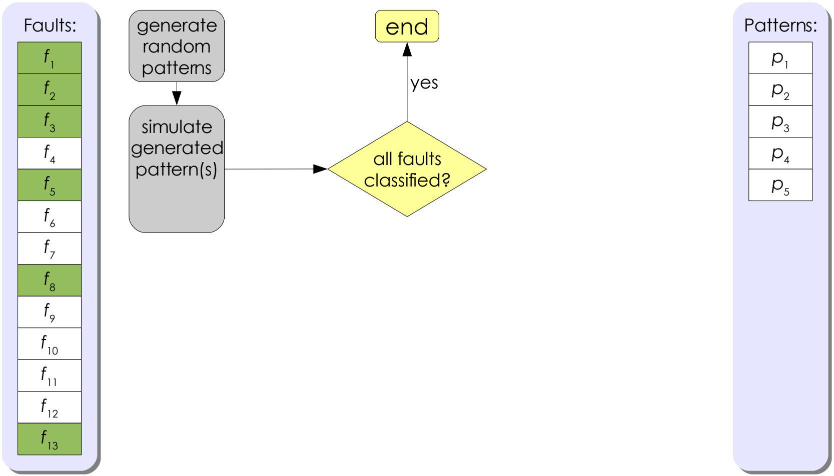


UNI FREIBURG

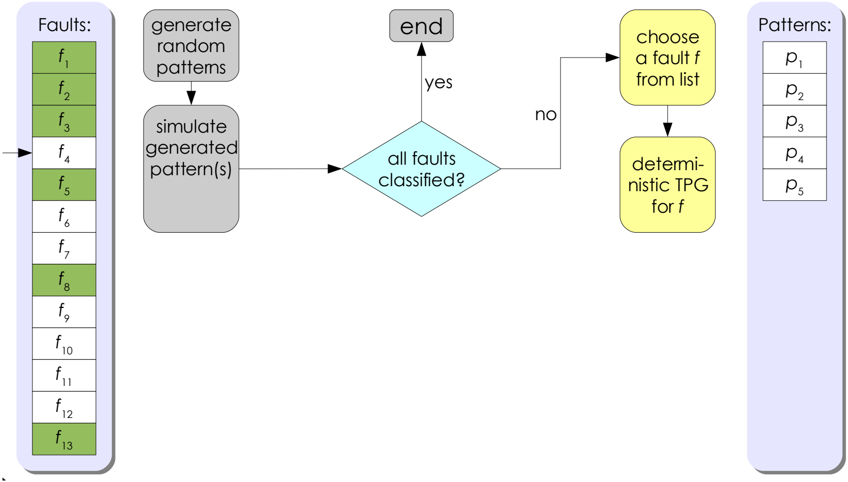
Automatic Test Pattern Generation – Typical Flow



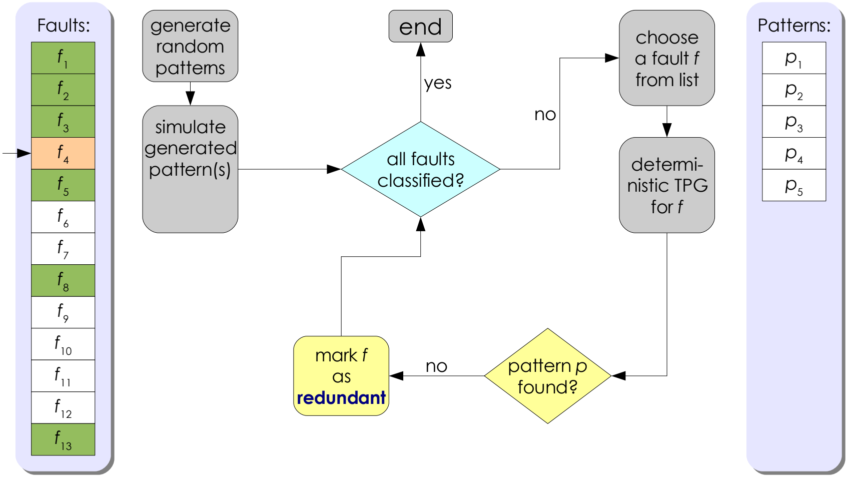
Automatic Test Pattern Generation – Typical Flow



Automatic Test Pattern Generation – Typical Flow

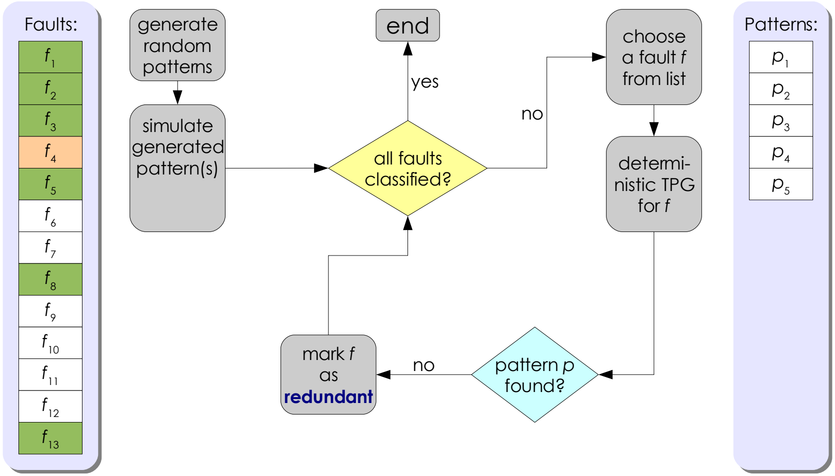


Automatic Test Pattern Generation – Typical Flow

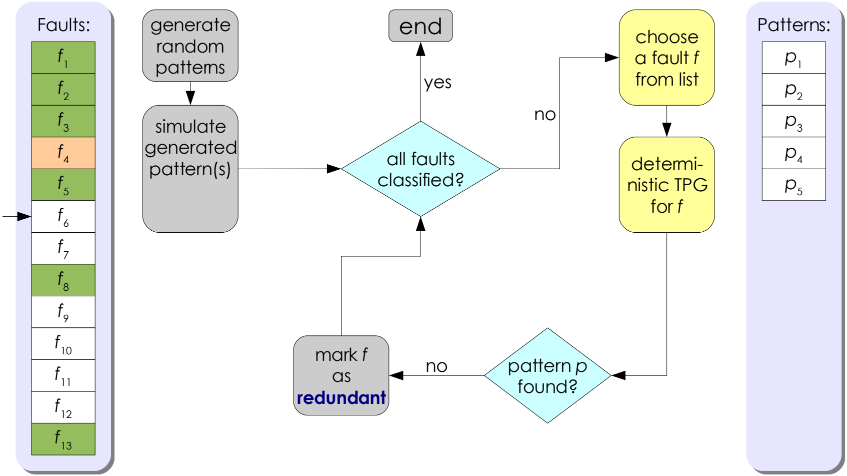


UNI FREIBURG

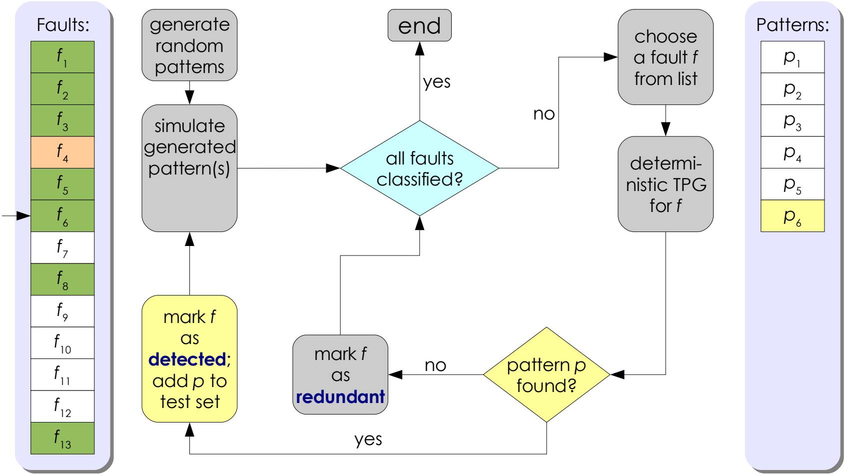
Automatic Test Pattern Generation – Typical Flow



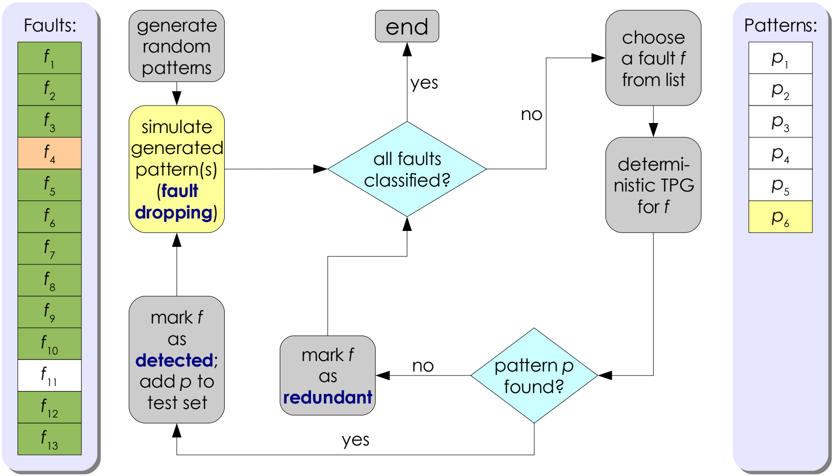
Automatic Test Pattern Generation – Typical Flow



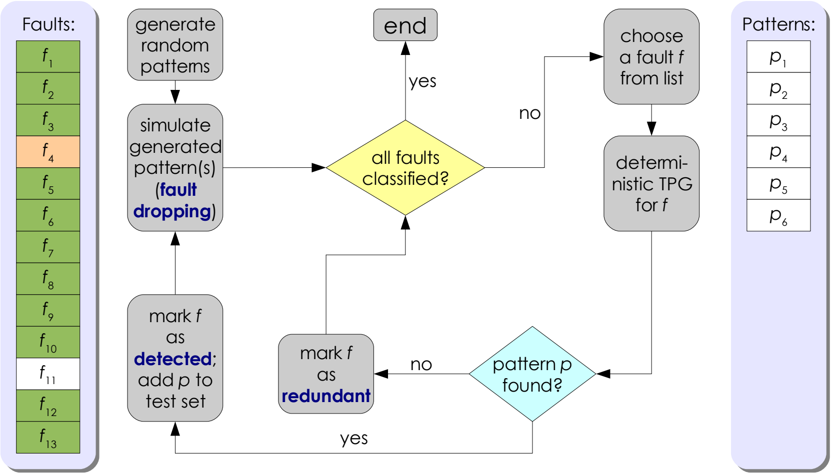
Automatic Test Pattern Generation – Typical Flow



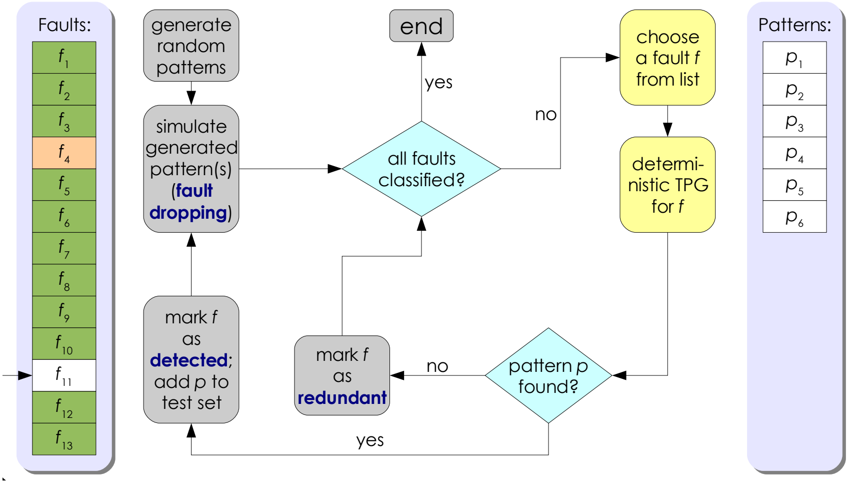
Automatic Test Pattern Generation – Typical Flow



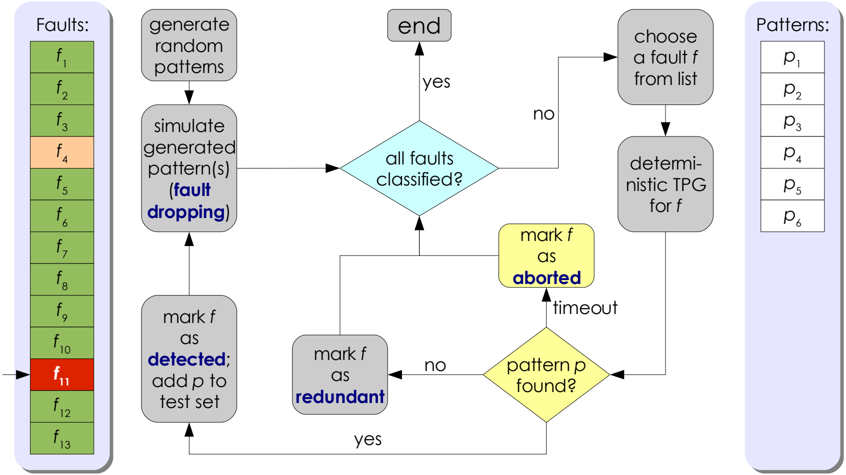
Automatic Test Pattern Generation – Typical Flow



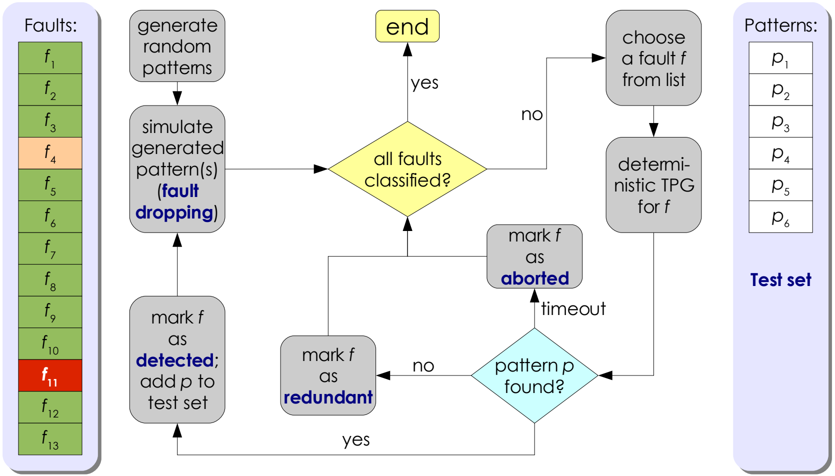
Automatic Test Pattern Generation – Typical Flow



Automatic Test Pattern Generation – Typical Flow



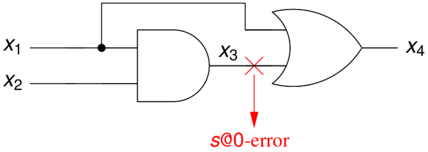
Automatic Test Pattern Generation – Typical Flow



Automatic Test Pattern Generation

Redundant faults: s@0 at x_3 is redundant

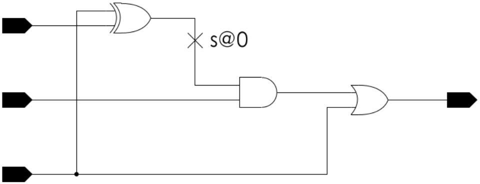
- Justifying the error requires $x_1 = 1$ and $x_2 = 1$
- But propagating the error to output x_4 requires $x_1 = 0$



Automatic Test Pattern Generation

Main concept of automatic test pattern generation

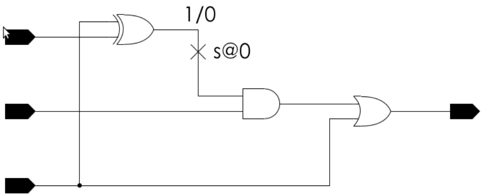
- Justify the fault and find a propagation path



Automatic Test Pattern Generation

Main concept of automatic test pattern generation

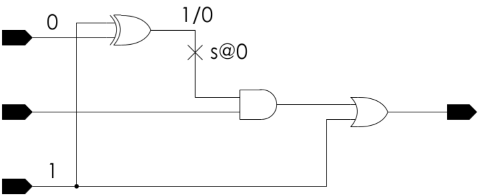
- Justify the fault and find a propagation path



Automatic Test Pattern Generation

Main concept of automatic test pattern generation

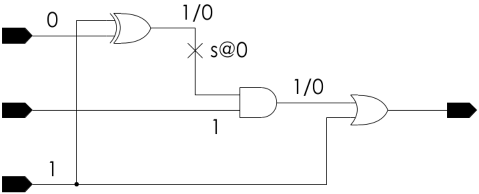
- Justify the fault and find a propagation path



Automatic Test Pattern Generation

Main concept of automatic test pattern generation

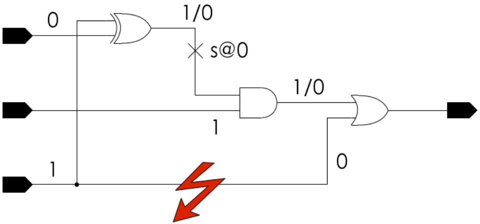
- Justify the fault and find a propagation path



Automatic Test Pattern Generation

Main concept of automatic test pattern generation

- Justify the fault and find a propagation path

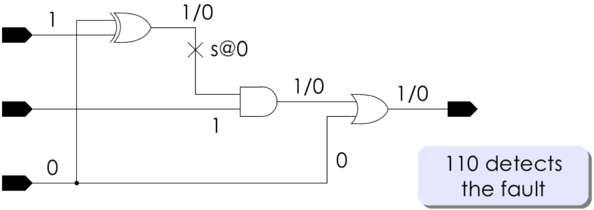


UNI FREIBURG

Automatic Test Pattern Generation

Main concept of automatic test pattern generation

- Justify the fault and find a propagation path



Automatic Test Pattern Generation

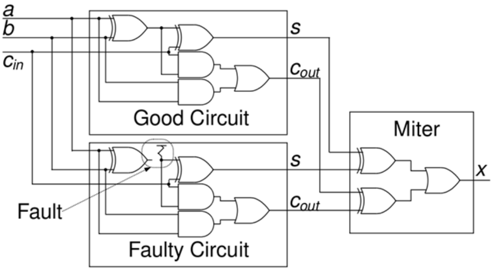
Several ATPG-Approaches

- Structural methods
  - D-algorithm
  - PODEM
  - FAN
- SAT-based methods

SAT-based ATPG

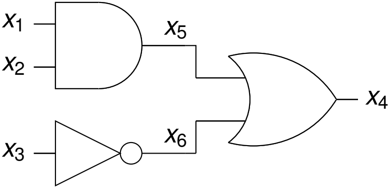
Main flow

- Construct the miter containing the correct and the faulty circuit
- Encode the miter as CNF & solve the SAT problem
- If the SAT formula is satisfiable we have found a test pattern for the particular fault under consideration
- Otherwise, the fault is redundant

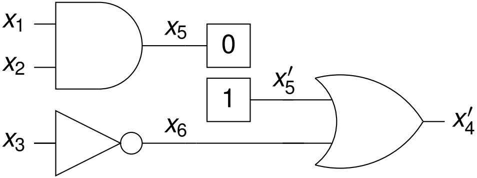


UNI FREIBURG

SAT-based ATPG – Example

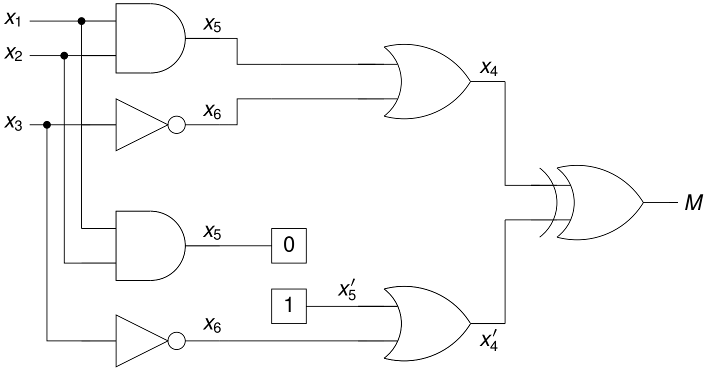


(a) Correct circuit

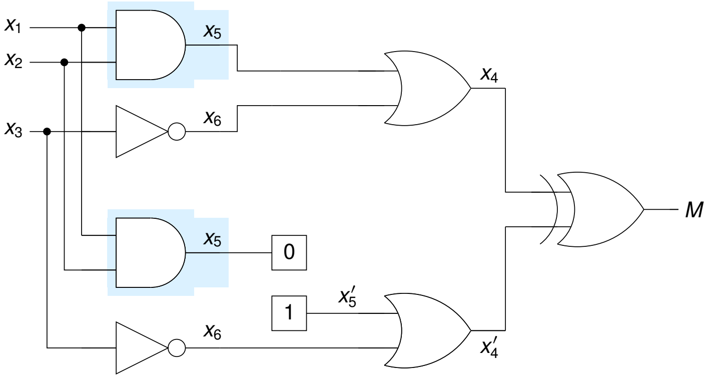


(b) Faulty circuit, *s*@1-error at x_5

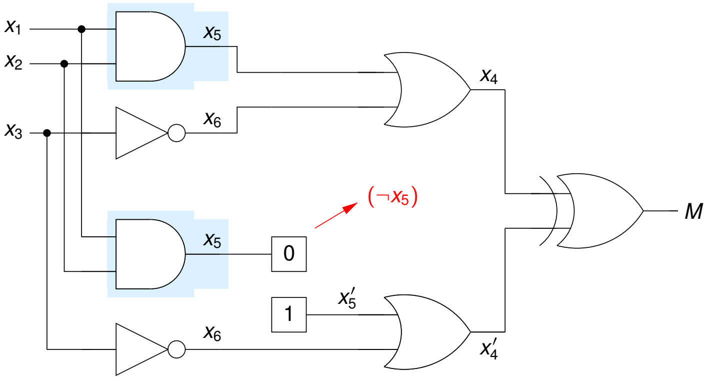
SAT-based ATPG – Example



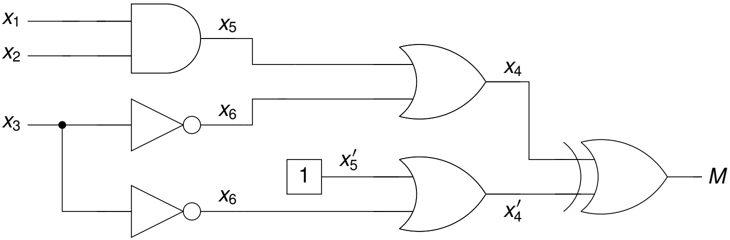
SAT-based ATPG – Example



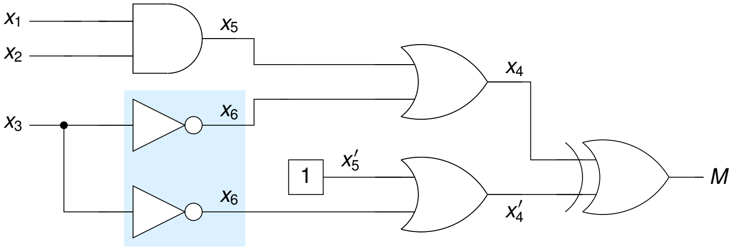
SAT-based ATPG – Example



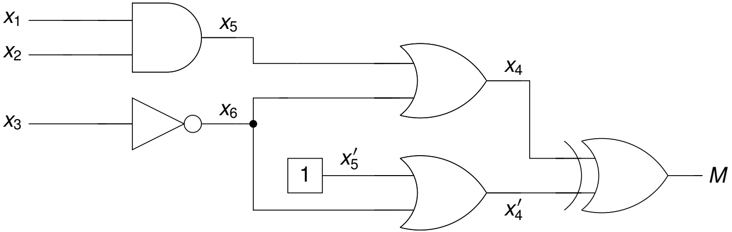
SAT-based ATPG – Example



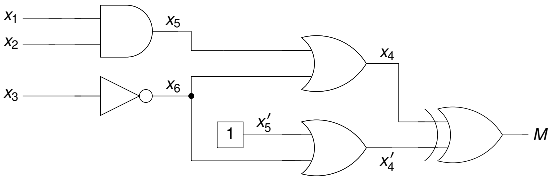
SAT-based ATPG – Example



SAT-based ATPG – Example

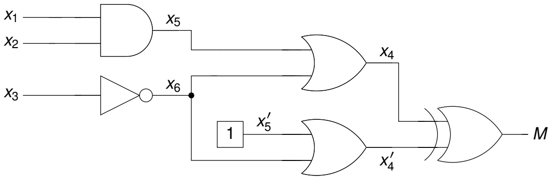


SAT-based ATPG – Example



$$
\begin{aligned}
F_M = {} & (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge (x_6 \vee x_3) \wedge \\
& (\neg x_6 \vee \neg x_3) \wedge (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \wedge \\
& (x_4' \vee \neg x_5') \wedge (x_4' \vee \neg x_6) \wedge (\neg x_4' \vee x_5' \vee x_6) \wedge (\neg M \vee x_4 \vee x_4') \wedge \\
& (\neg M \vee \neg x_4 \vee \neg x_4') \wedge (M \vee \neg x_4 \vee x_4') \wedge (M \vee x_4 \vee \neg x_4') \wedge \\
& (M) \wedge (\neg x_5) \wedge (x_5')
\end{aligned}
$$

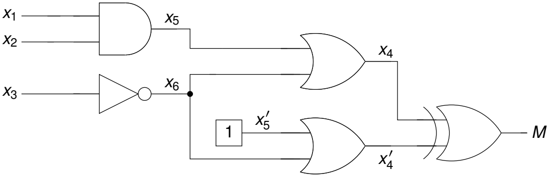
SAT-based ATPG – Example



$$
\begin{aligned}
F_M = {} & (\neg x_5 \lor x_1) \land (\neg x_5 \lor x_2) \land (x_5 \lor \neg x_1 \lor \neg x_2) \land (x_6 \lor x_3) \land \\
& (\neg x_6 \lor \neg x_3) \land (x_4 \lor \neg x_5) \land (x_4 \lor \neg x_6) \land (\neg x_4 \lor x_5 \lor x_6) \land \\
& (x_4' \lor \neg x_5') \land (x_4' \lor \neg x_6) \land (\neg x_4' \lor x_5' \lor x_6) \land (\neg M \lor x_4 \lor x_4') \land \\
& (\neg M \lor \neg x_4 \lor \neg x_4') \land (M \lor \neg x_4 \lor x_4') \land (M \lor x_4 \lor \neg x_4') \land \\
& (M) \land (\neg x_5) \land (x_5')
\end{aligned}
$$

$$
F_M' = (\neg x_1 \lor \neg x_2) \land (x_3) \land (\neg x_6) \land (x_4') \land (\neg x_4) \land (M) \land (\neg x_5) \land (x_5')
$$

SAT-based ATPG – Example



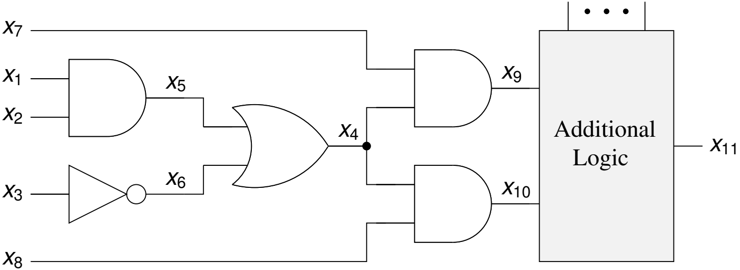
$$F_M = (\neg x_5 \vee x_1) \wedge (\neg x_5 \vee x_2) \wedge (x_5 \vee \neg x_1 \vee \neg x_2) \wedge (x_6 \vee x_3) \wedge$$
$$(\neg x_6 \vee \neg x_3) \wedge (x_4 \vee \neg x_5) \wedge (x_4 \vee \neg x_6) \wedge (\neg x_4 \vee x_5 \vee x_6) \wedge$$
$$(x'_4 \vee \neg x'_5) \wedge (x'_4 \vee \neg x_6) \wedge (\neg x'_4 \vee x'_5 \vee x_6) \wedge (\neg M \vee x_4 \vee x'_4) \wedge$$
$$(\neg M \vee \neg x_4 \vee \neg x'_4) \wedge (M \vee \neg x_4 \vee x'_4) \wedge (M \vee x_4 \vee \neg x'_4) \wedge$$
$$(M) \wedge (\neg x_5) \wedge (x'_5)$$

$$F'_M = (\neg x_1 \vee \neg x_2) \wedge (x_3) \wedge (\neg x_6) \wedge (x'_4) \wedge (\neg x_4) \wedge (M) \wedge (\neg x_5) \wedge (x'_5)$$

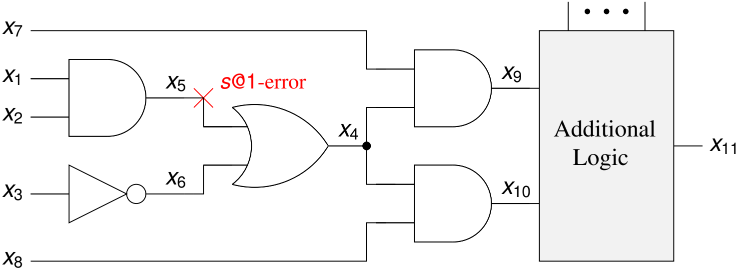
Test set: $(x_1, x_2, x_3) = \{(0,0,1), (1,0,1), (0,1,1)\}$

UNI FREIBURG

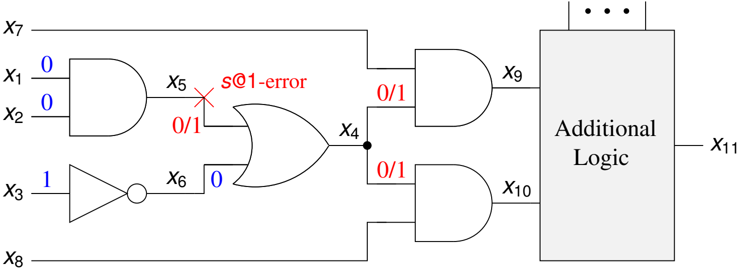
SAT-based ATPG – Adding Structural Information



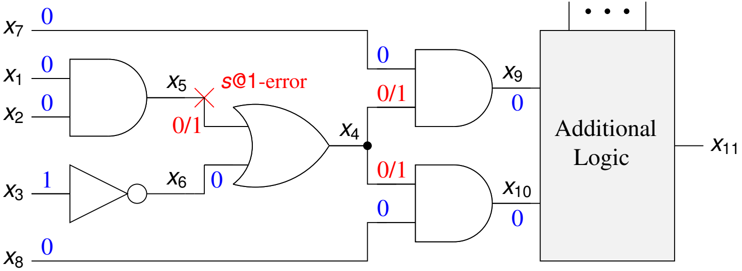
SAT-based ATPG – Adding Structural Information



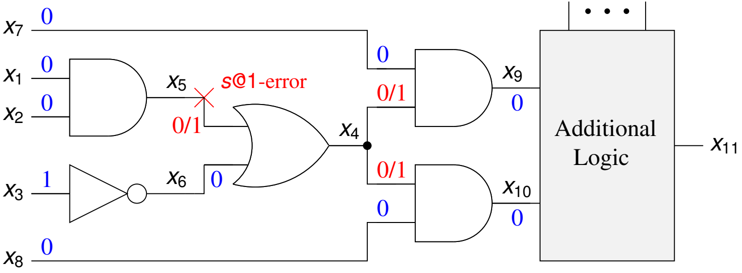
SAT-based ATPG – Adding Structural Information



SAT-based ATPG – Adding Structural Information

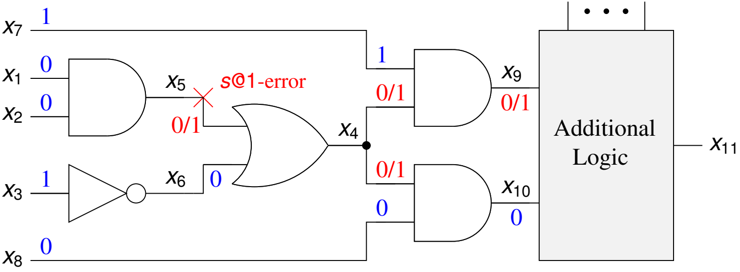


SAT-based ATPG – Adding Structural Information



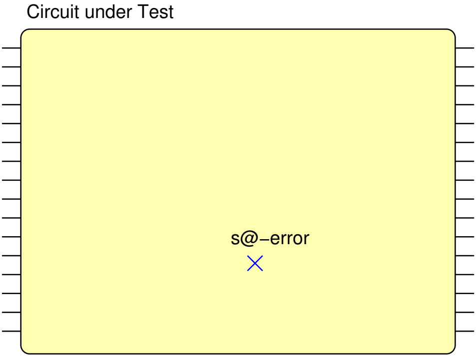
Add (x_7, x_8) to the CNF

SAT-based ATPG – Adding Structural Information

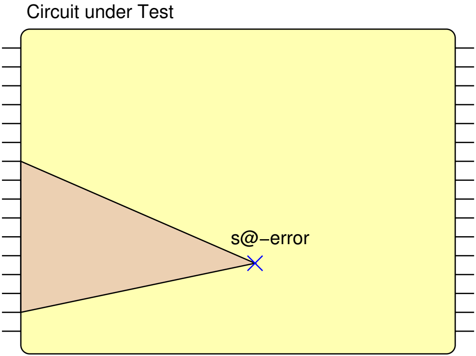


Add (x_7, x_8) to the CNF

SAT-based ATPG – Cone-of-Influence Reduction

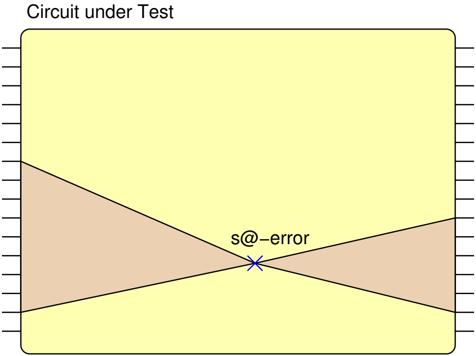


SAT-based ATPG – Cone-of-Influence Reduction



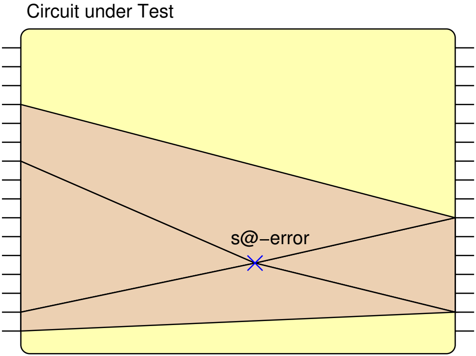
Which inputs might be relevant for justifying the fault?

SAT-based ATPG – Cone-of-Influence Reduction



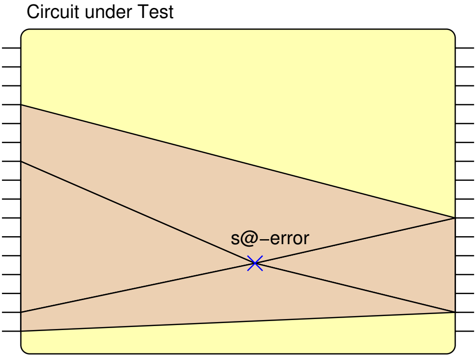
Which outputs might be on the propagation path?

SAT-based ATPG – Cone-of-Influence Reduction



What about side-effects?

SAT-based ATPG – Cone-of-Influence Reduction



⇒ Only the “brown” parts have to be transformed into CNF!

SAT-based ATPG – Testing of Sequential Circuits