



#### A Verified SAT Solver with Watched Literals Using Imperative HOL

Peter

Lammich

Mathias Fleury

Jasmin C. Blanchette







## SAT Solving

Given a formula in conjunctive normal form

$$\varphi = \bigwedge_{i} \bigvee_{j} L_{i,j}$$

is there an assignment making the formula true?

Most used algorithm: CDCL, an improvement over DPLL







## How reliable are SAT solvers?

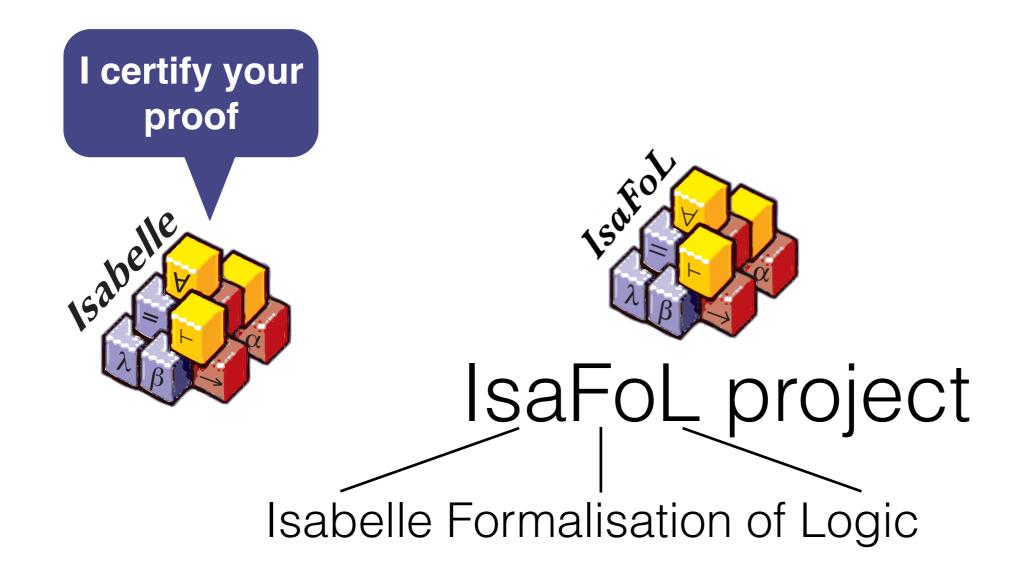
Two ways to ensure correctness:

- certify the certificate
  - certificates are huge
- verification of the code
  - code will not be competitive
  - allows to study metatheory















## IsaFoL

- FO resolution
   by Schlichtkrull (ITP 2016)
- CDCL with learn, forget, restart, and incrementality by Blanchette, Fleury, Weidenbach (IJCAR 2016)
- GRAT certificate checker
   by Lammich (CADE-26, 2017)
- A verified SAT solver with watched literals by Fleury, Blanchette, Lammich (now)







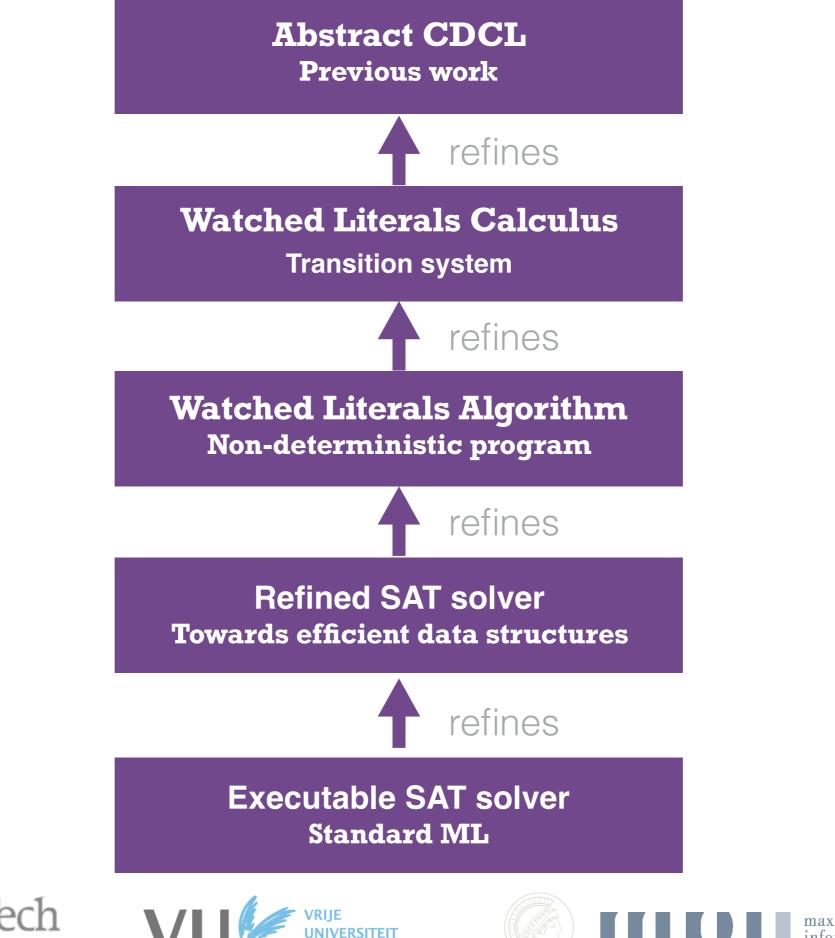
## IsaFoL

- FO resolution
   by Schlichtkrull (ITP 2016)
- CDCL with learn, forget, restart, and incrementality by Blanchette, Fleury, Weidenbach (IJCAR 2016)
- GRAT certificate checker
   by Lammich (CADE-26, 2017)
- A verified SAT solver with watched literals by Fleury, Blanchette, Lammich (now)

















#### Abstract CDCL Previous work







#### **Candidate model**

#### 1. Guess

- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict

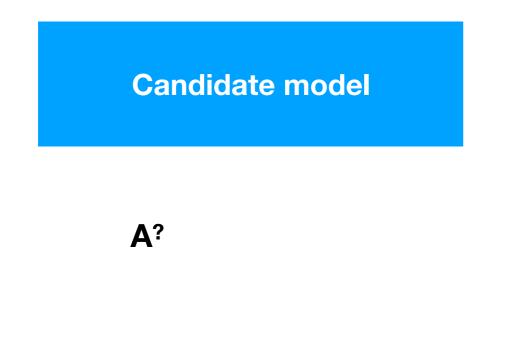
#### Clause

- **1.**  $\neg B \lor C \lor A$
- 2. ¬C ∨ ¬B ∨ ¬A
- 3. ¬A ∨ ¬B ∨ C
- 4. ¬A ∨ B



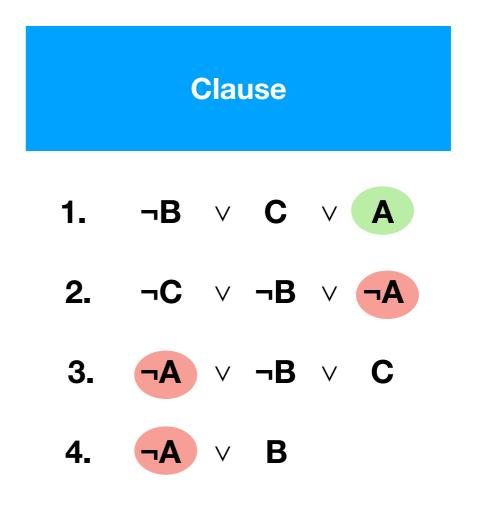








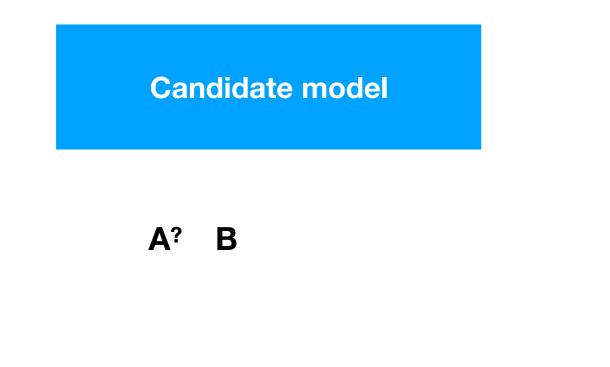
- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict





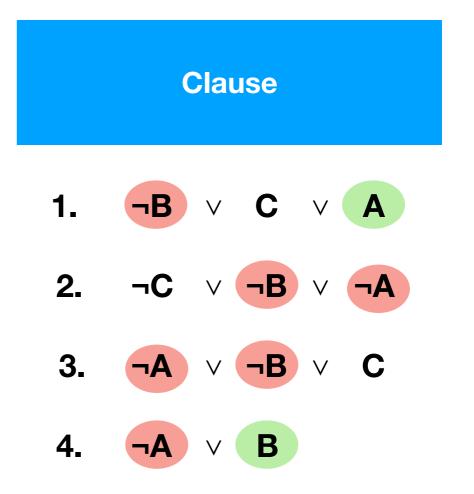








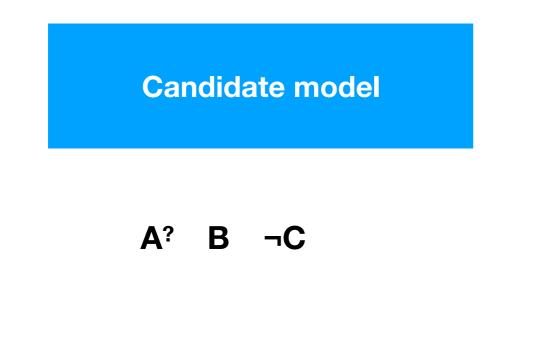
- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict





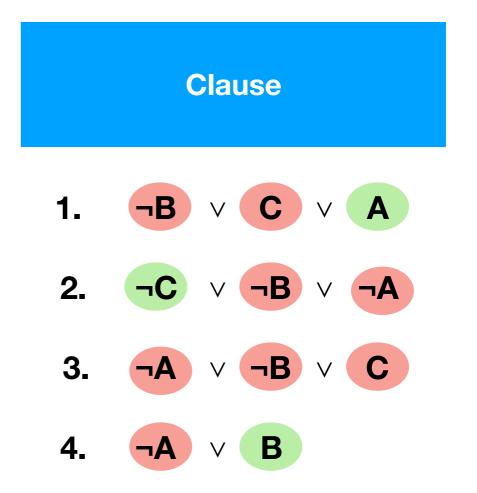








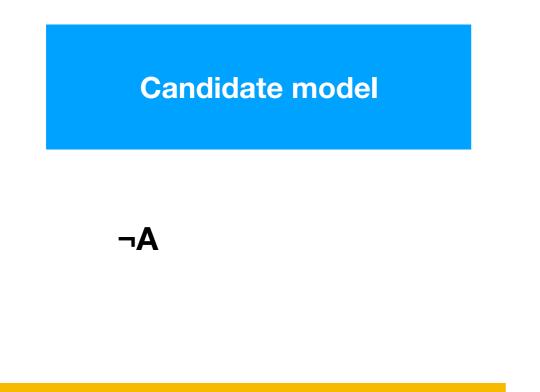
- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict





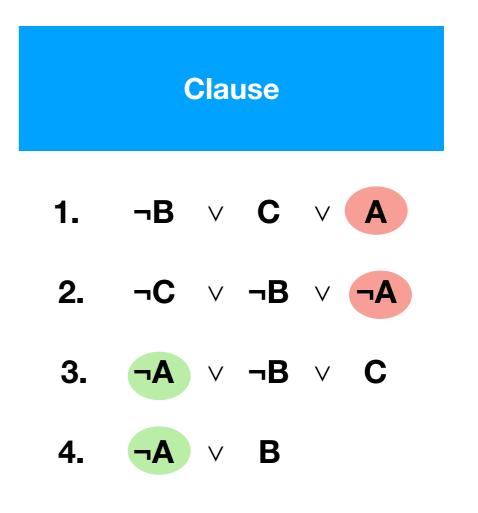








- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict







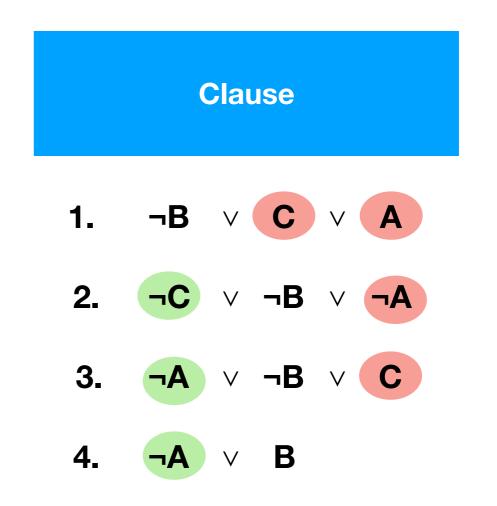




¬A ¬C?



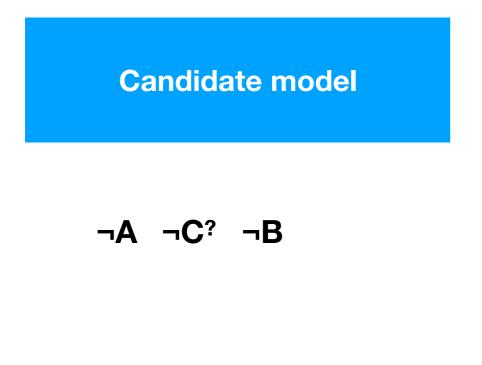
- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict





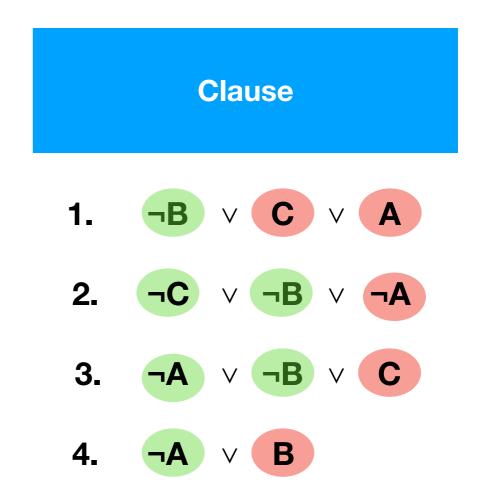








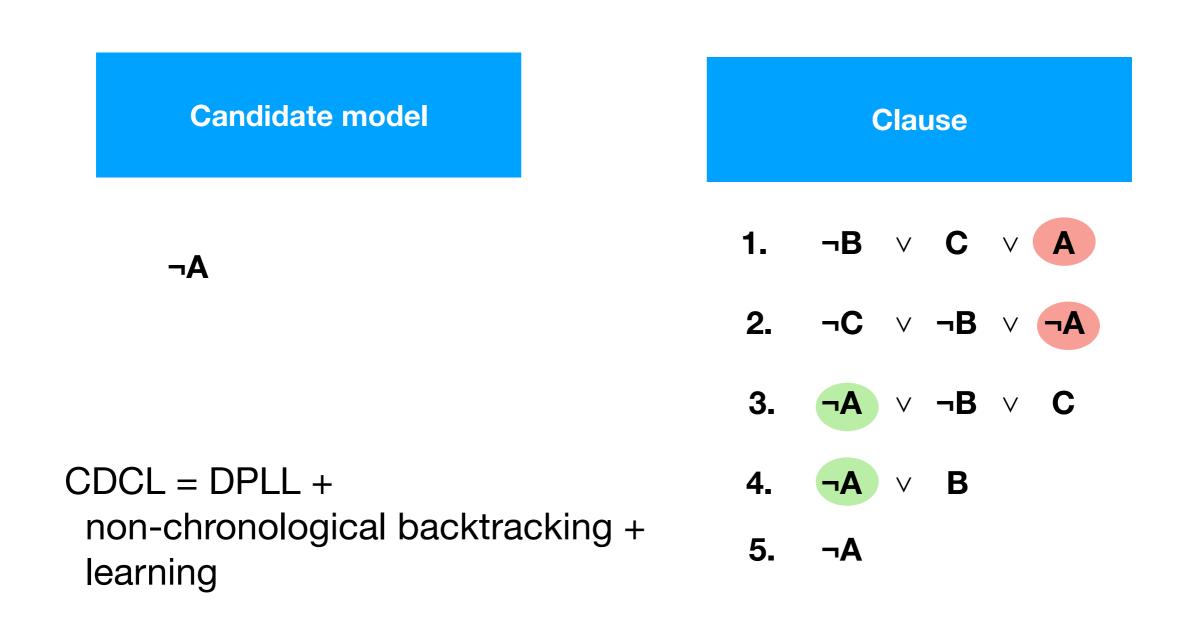
- 2. or propagate information
- 3. or take the opposite of the last guess if there is a conflict

















Propagate rule  

$$C \lor L \in N \implies M \models as \neg C \implies undefined_lit M L \implies$$
  
 $(M, N) \Rightarrow_{CDCL} (L \# M, N)$ 







Propagate rule  

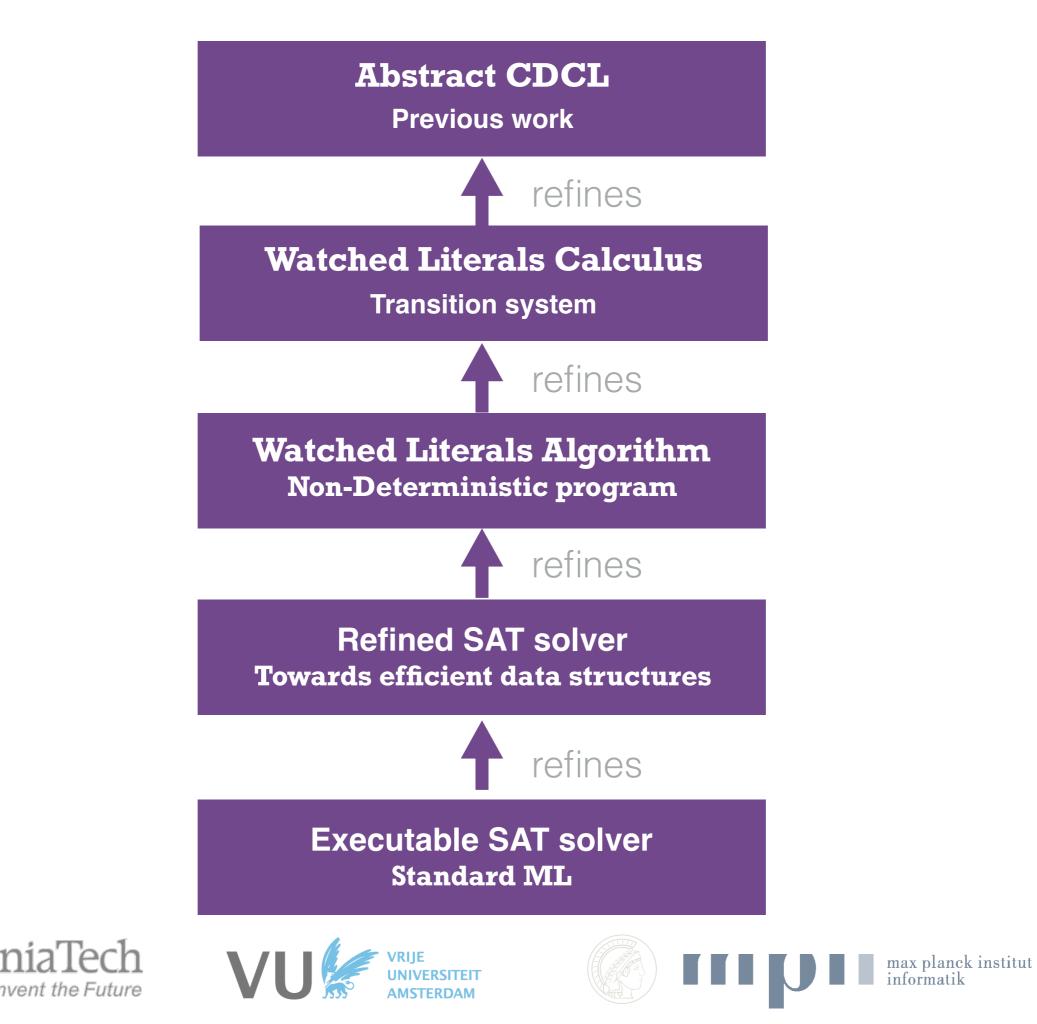
$$C \lor L \in N \implies M \models as \neg C \implies undefined_lit M L \implies$$
  
 $(M, N) \Rightarrow_{CDCL} (L \# M, N)$ 

Problem: Iterating over the clauses is inefficient









#### Watched Literals Calculus

#### **Transition system**







#### **Candidate model**

- 1. Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false

# 1. <mark>' ¬B ∨' C</mark> ∨ A

Clause

#### To update:



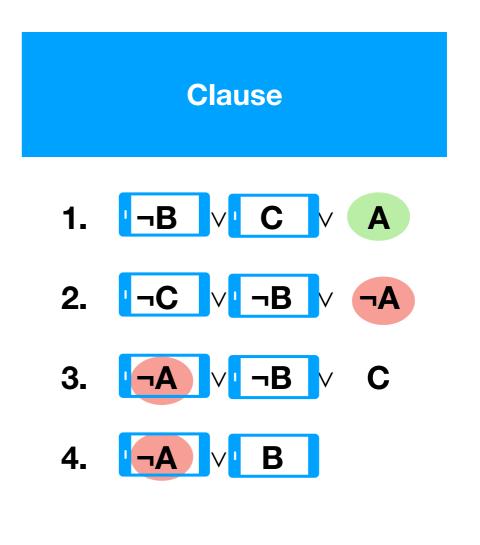






**A**?

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false



#### To update:



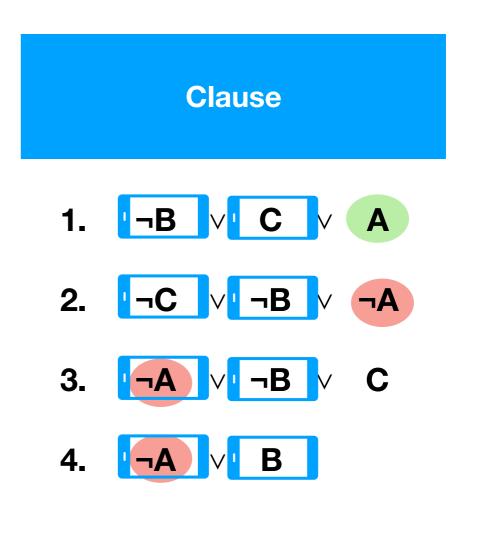






**A**?

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false



#### To update: 3. 4.



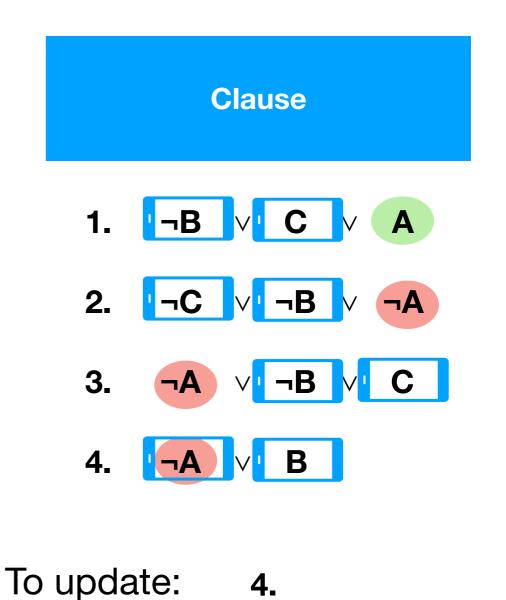






**A**?

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false





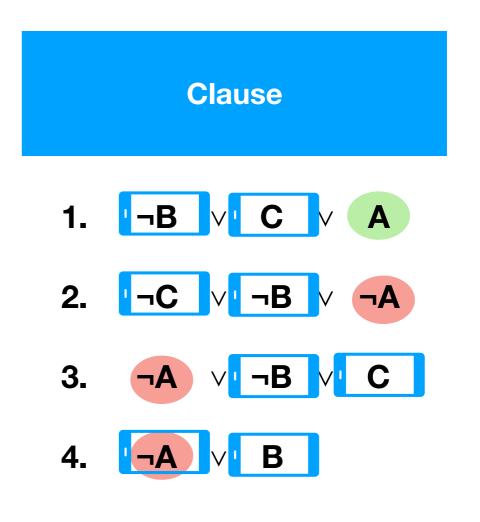






**A**?

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false



#### To update:



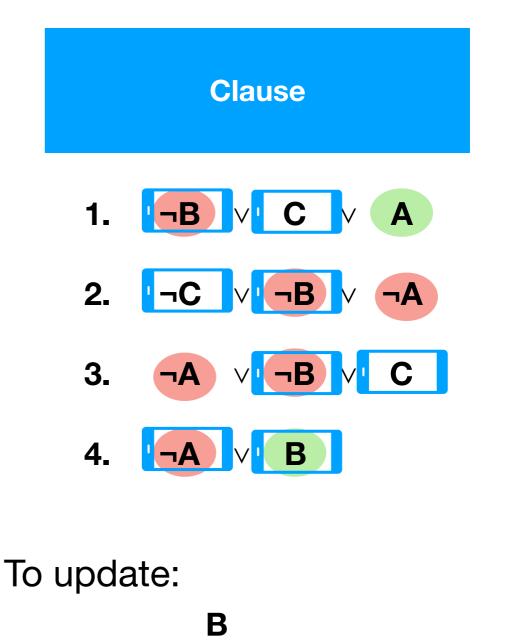






**A**? **B** 

- 1. Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false







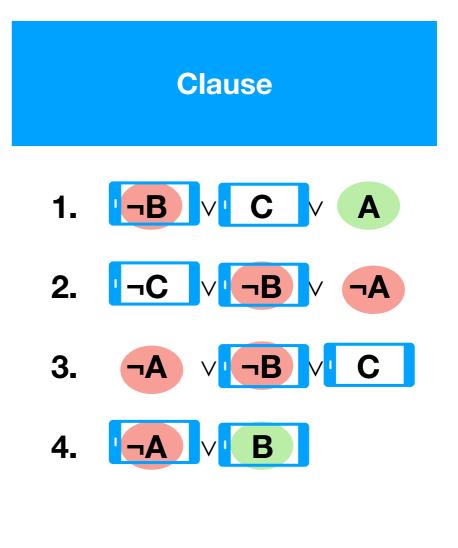
max planck institut

informatik



**A**? **B** 

- 1. Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false

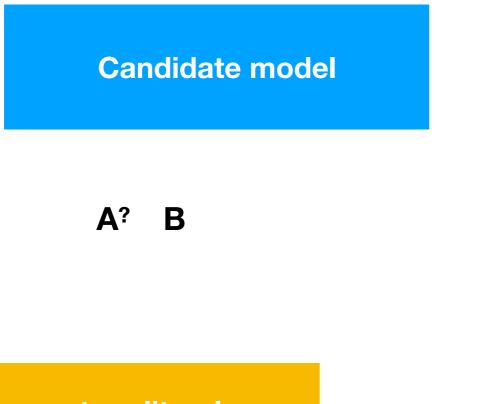


To update: 1. 2. 3.

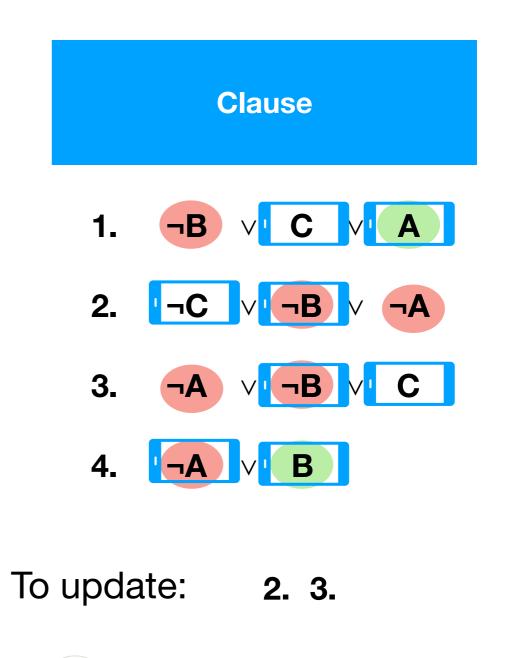








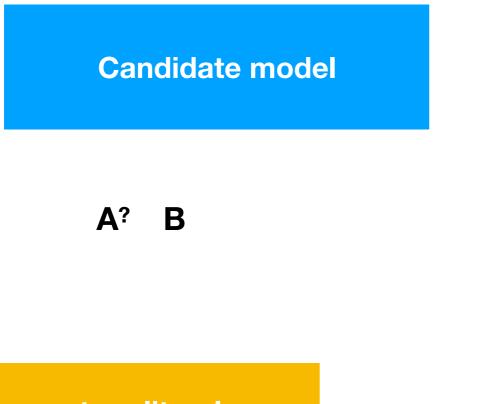
- **1. Watch one true literals**
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false



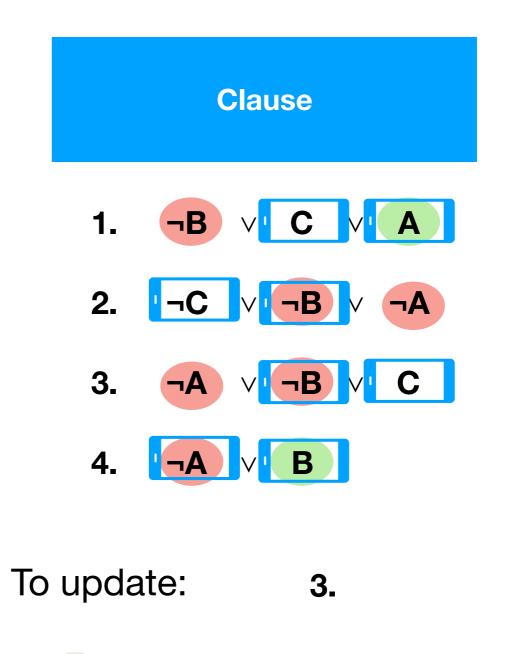








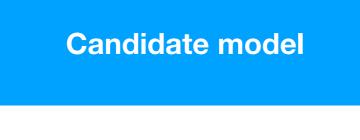
- **1. Watch one true literals**
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false









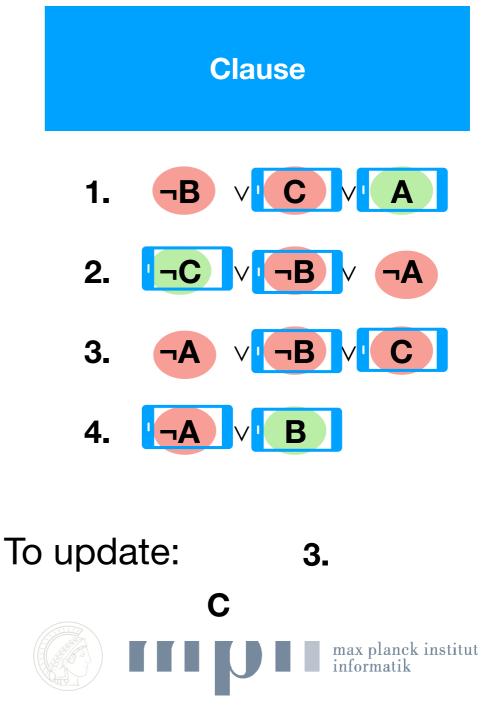


**A**? **B** ¬**C** 

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false



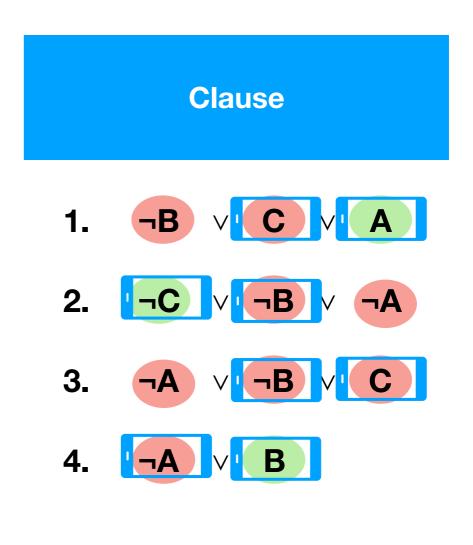






A? B ¬C

- 1. Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false

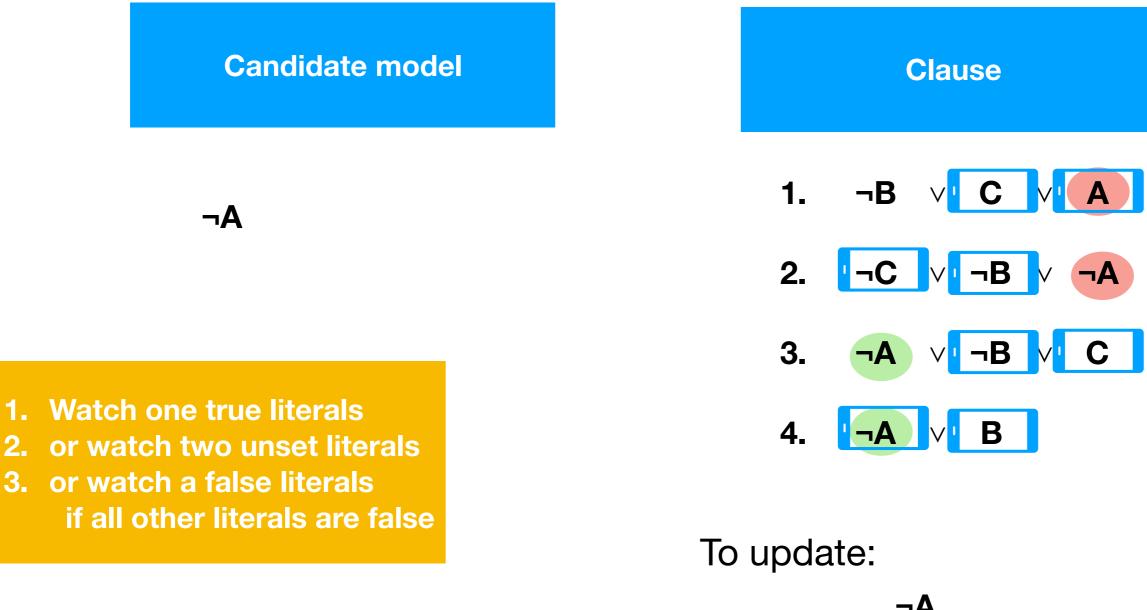


#### To update:







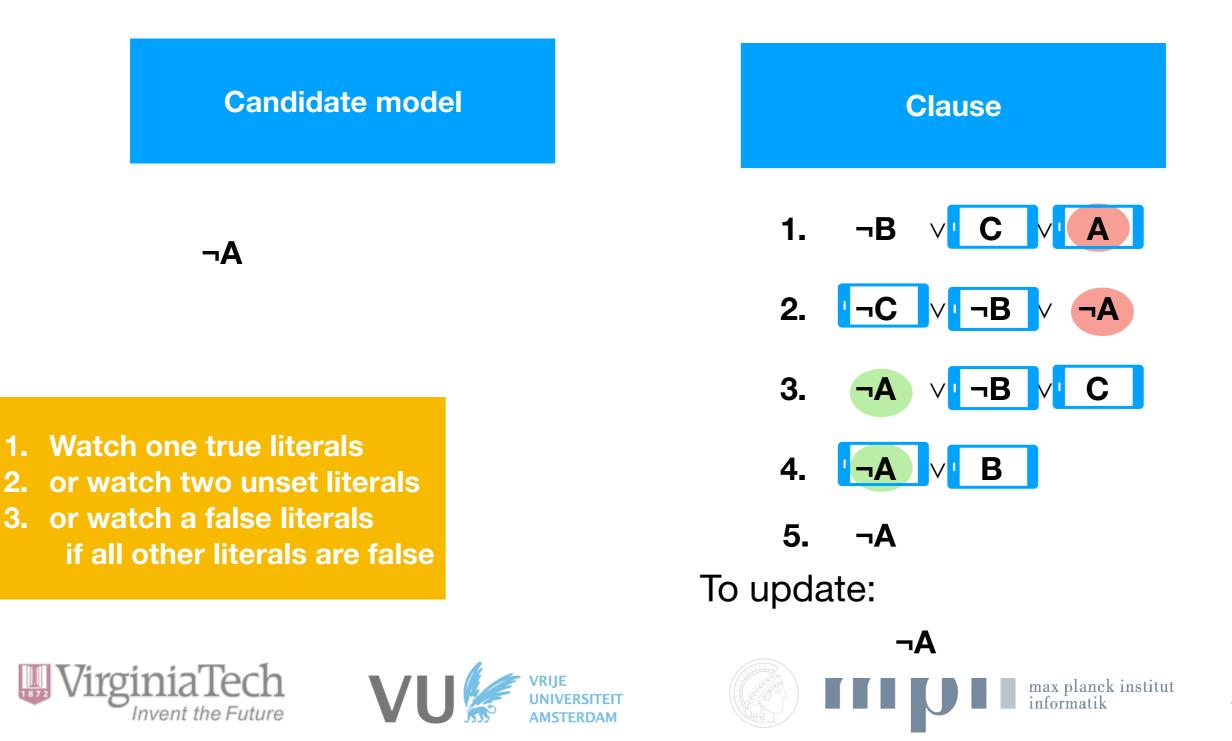




2.







2.

#### Watched literals invariant

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false







#### Watched literals invariant

- **1.** Watch one true literals
- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false

unless a conflict has been found







#### Watched literals invariant



- 2. or watch two unset literals
- 3. or watch a false literals if all other literals are false

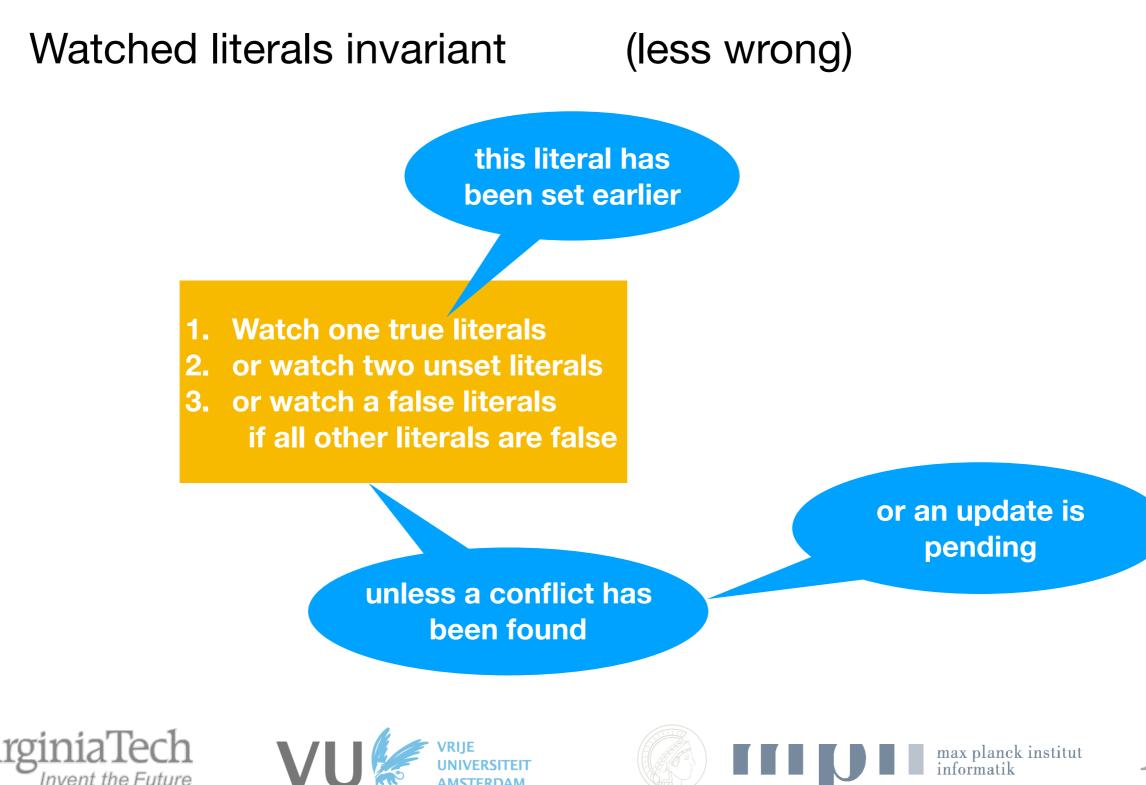
or an update is pending

unless a conflict has been found









## • Finding invariants (11 new ones)



No high-level description









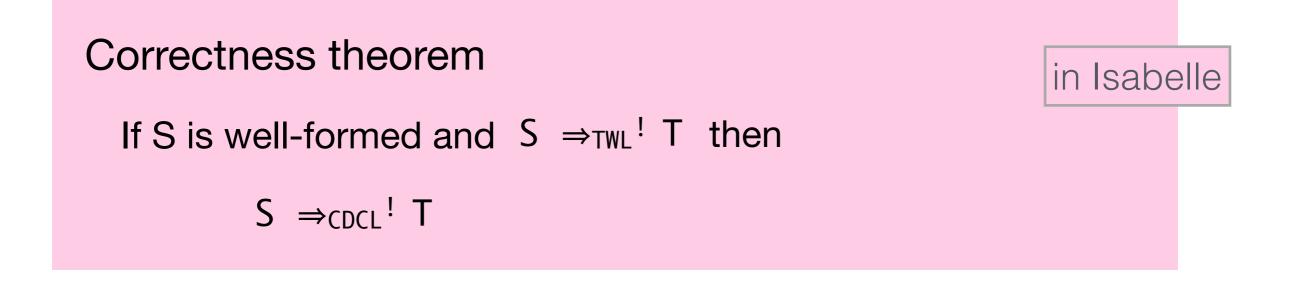
# Finding invariants (11 new ones)



No high-level description



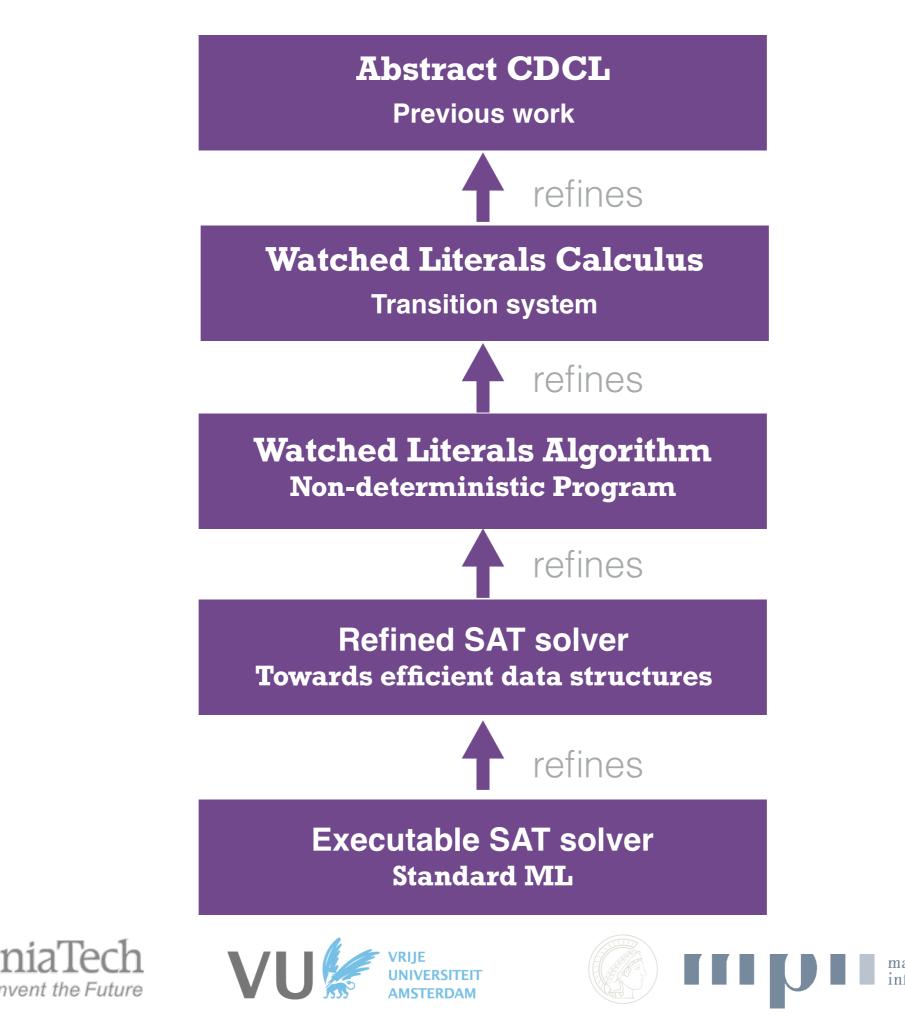
sledgehammer











# Watched Literals Calculus

## **Transition system**



# Watched Literals Algorithm Non-deterministic Program







# **Picking Next Clause**

```
propagate_conflict_literal L S :=

WHILE<sub>T</sub>

(\lambdaT. clauses_to_update T \neq {})

(\lambdaT. do {

ASSERT(clauses_to_update T \neq {})

C \leftarrow SPEC (\lambdaC. C \in clauses_to_update T);

U \leftarrow remove_from_clauses_to_update C T;

update_clause (L, C) U

}
```









```
propagate_conflict_literal L S :=

WHILE<sub>T</sub>

(\lambdaT. clauses_to_update T \neq {})

(\lambdaT. do {

ASSERT(clauses_to_update T \neq {})

C \leftarrow SPEC (\lambdaC. C \in clauses_to_update T);

U \leftarrow remove_from_clauses_to_update C T;

update_clause (L, C) U

}

)
```







```
propagate_conflict_literal L S :=

WHILE<sub>T</sub>

(\lambdaT. clauses_to_update T \neq {})

(\lambdaT. do {

ASSERT(clauses_to_update T \neq {})

C \leftarrow SPEC (\lambdaC. C \in clauses_to_update T);

U \leftarrow remove_from_clauses_to_update C T;

update_clause (L, C) U

}

)
```







```
propagate_conflict_literal L S :=

WHILE<sub>T</sub>

(\lambdaT. clauses_to_update T \neq {})

(\lambdaT. do {

ASSERT(clauses_to_update T \neq {})

C \leftarrow SPEC (\lambdaC. C \in clauses_to_update T);

U \leftarrow remove_from_clauses_to_update C T;

update_clause (L, C) U

}
```







```
propagate_conflict_literal L S :=

WHILE<sub>T</sub>

(\lambdaT. clauses_to_update T \neq {})

(\lambdaT. do {

ASSERT(clauses_to_update T \neq {})

C \leftarrow SPEC (\lambdaC. C \in clauses_to_update T);

U \leftarrow remove_from_clauses_to_update C T;

update_clause (L, C) U

}

)
```







- More deterministic (order of the rules)
- But still non deterministic (decisions)
- Goals of the form







- More deterministic (order of the rules)
- But still non deterministic (decisions)
- Goals of the form

propagate\_conflict\_literal L S 
$$\leq$$
 SPEC( $\lambda$ T. S  $\Rightarrow_{TWL}^*$  T)







in Isabelle



# VCG's goals hard to read



Very tempting to write fragile proofs

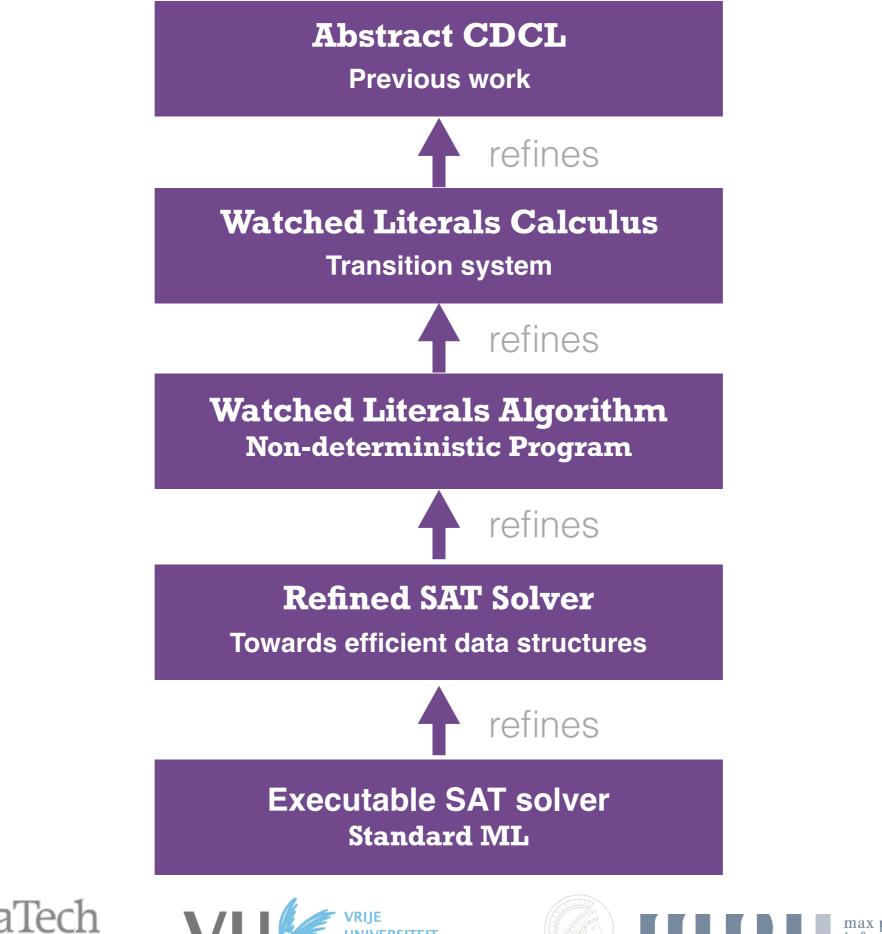


sledgehammer









nvent the Future

## Watched Literals Algorithm Non-deterministic Program



## **Refined SAT Solver**

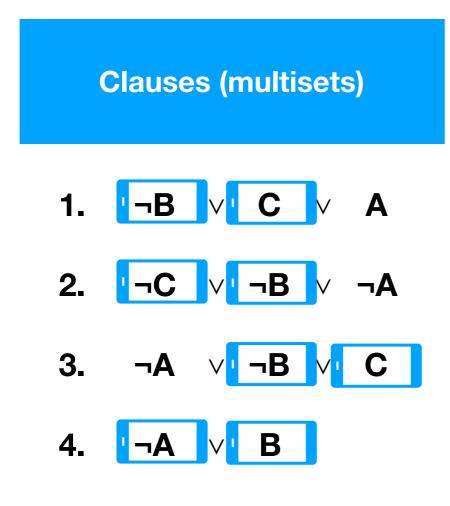
**Towards efficient data structures** 







# **DPLL with Watched Literals**

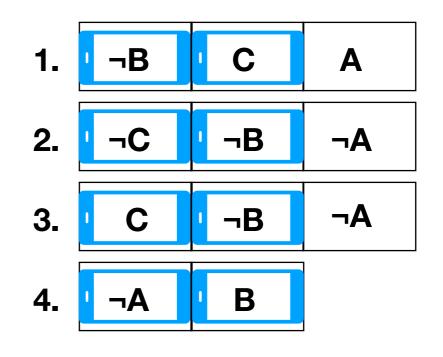


To update:





#### Clauses after refinement (lists)

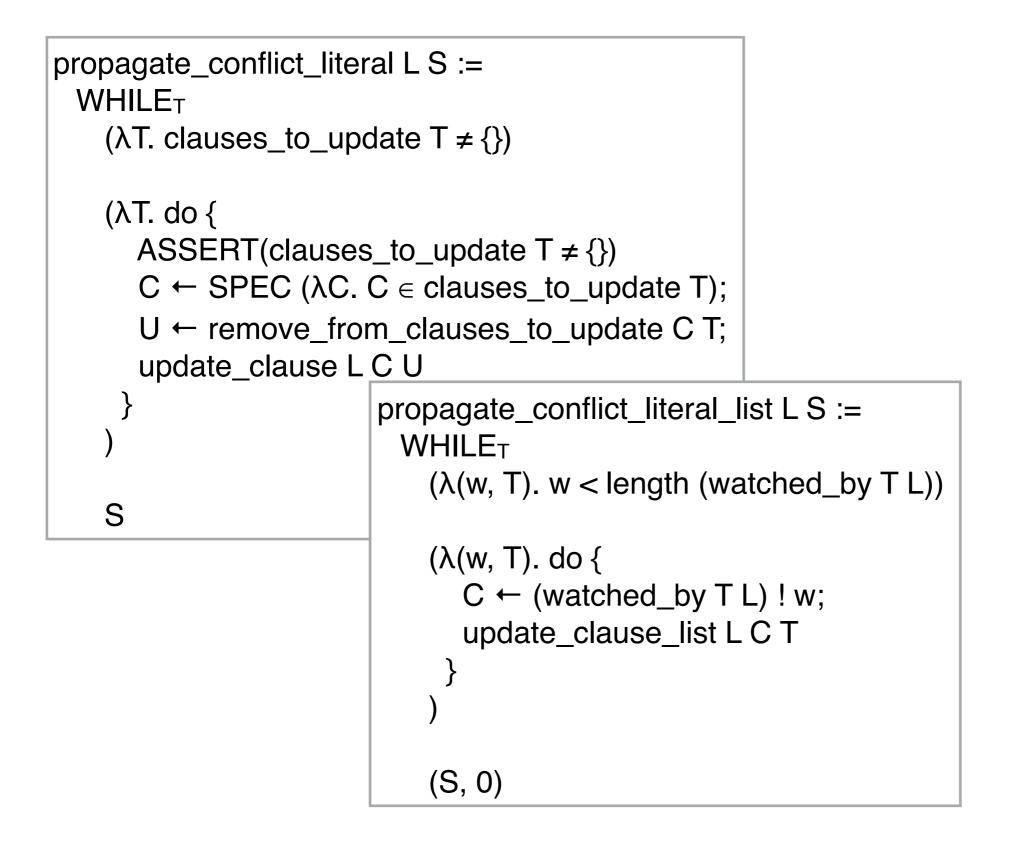


A: ¬A: 4 C: 1,3 ¬C: 2 B: 4 ¬B: 1,2,3



max planck institut

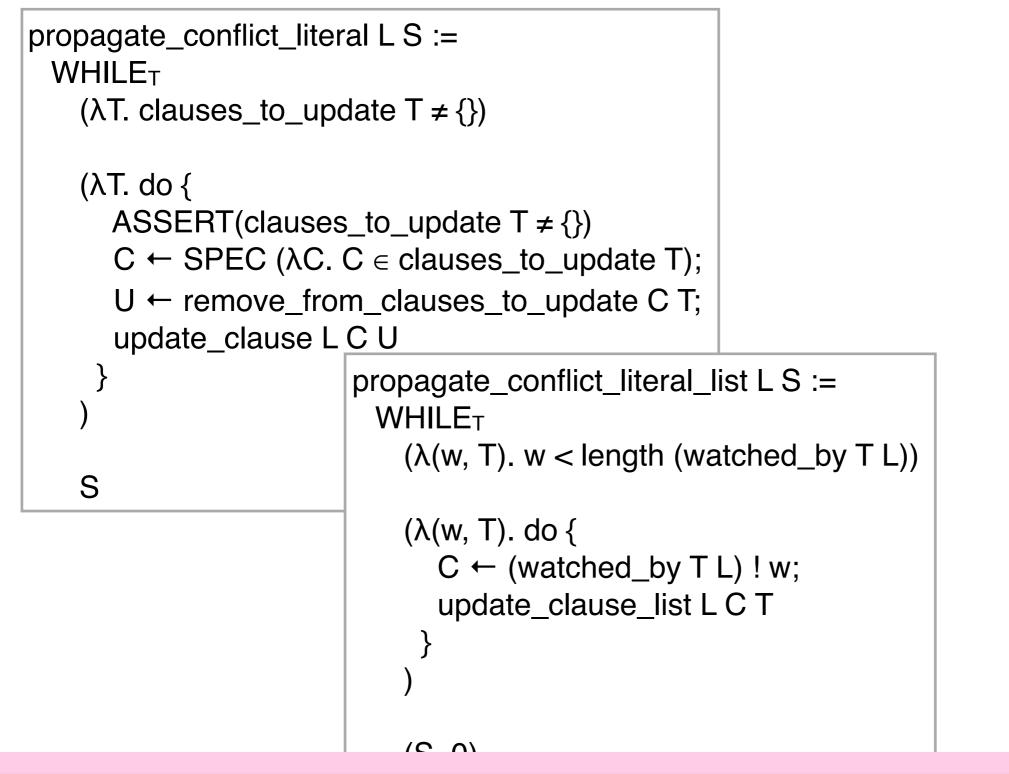
informatik











propagate\_conflict\_literal\_list  $L S \leq \Downarrow$  conversion\_between\_states (propagate\_conflict\_literal L T) in Isabelle







# More new invariants

- Aligning goals is hard...
- X
- Fast code uses many invariants
- For
  - Forgotten and new invariants



sledgehammer







- Choice on the heuristics
- Om Prepare code synthesis







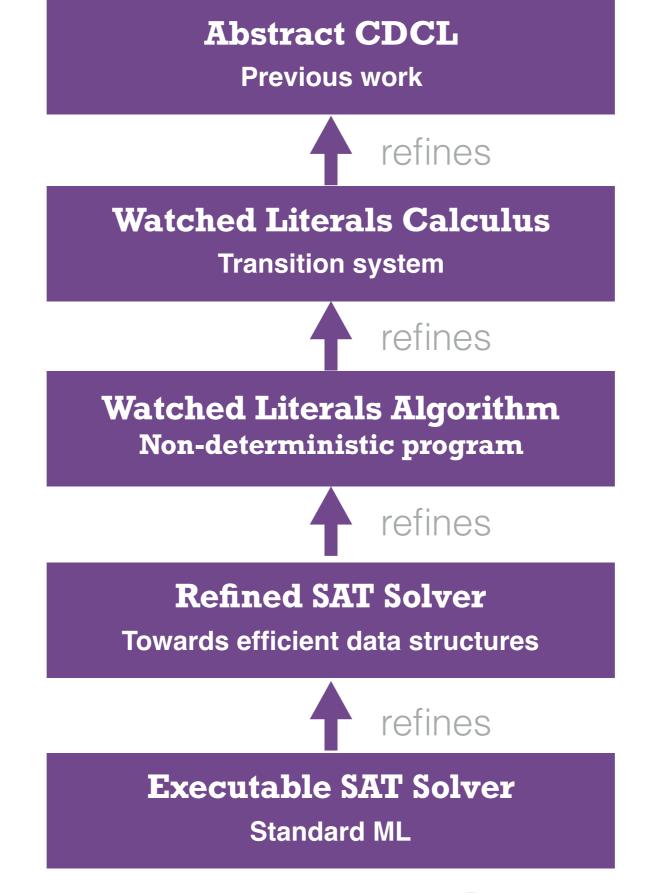
# **Decision heuristic**

- Variable-move-to-front heuristic
- No correctness w.r.t. a standard implementation
- Behaves correctly:
  - returns an unset literal if there is one
  - no exception (out-of-bound array accesses)















max planck institut informatik

# **Refined SAT Solver**

## **Towards efficient data structures**



# **Executable SAT Solver**

**Standard ML** 







### sepref\_definition executable\_version

- <propagate\_conflict\_literal\_heuristics></propagate\_conflict\_literal\_heuristics>
- ::  $\operatorname{unat_lit_assn^k *_a state_assn^d \rightarrow_a state_assn^}$
- by sepref

Synthesise imperative code and a refinement relation







### sepref\_definition executable\_version

- <propagate\_conflict\_literal\_heuristics></propagate\_conflict\_literal\_heuristics>
- ::  $\operatorname{unat_lit_assn^k *_a state_assn^d \rightarrow_a state_assn^}$
- by sepref

Synthesise imperative code and a refinement relation

```
main_loop S :=
    heap_WHILET
    (\(finished, _). return (¬ finished))
    (\(_, state).
        propagate state >=
            analyse_or_decide)
    (False, state) >=
    (\((_, final_state). return final_state))
```







```
sepref_definition executable_version
```

- <propagate\_conflict\_literal\_heuristics></propagate\_conflict\_literal\_heuristics>
- ::  $\operatorname{unat_lit_assn^k *_a state_assn^d \rightarrow_a state_assn^}$
- by sepref

Synthesise imperative code and a refinement relation

```
fun main_loop state =
 fn () =>
  let
   val (_, final_state) =
    heap_WHILET
      (fn (done, _) => (fn () => not done))
      (fn (\_, state) =>
        (analyse_or_decide (propagate state ()) ()))
      (false, xi)
      ();
  in final_state end;
```







### sepref\_definition executable\_version

- <propagate\_conflict\_literal\_heuristics></propagate\_conflict\_literal\_heuristics>
- ::  $\operatorname{unat_lit_assn^k *_a state_assn^d \rightarrow_a state_assn^}$
- by sepref

## Synthesise imperative code and a refinement relation

```
fun cdcl_twl_stgy_prog_wl_D_code x =
 (fn xi => fn () =>
  let
   val a =
    heap_WHILET (fn (a1, _) => (fn () => (not a1)))
      (fn (_, a2) =>
       (fn f_=> fn () => f_((unit_propagation_outer_loop_wl_D a2) ()) ())
        cdcl_twl_o_prog_wl_D_code)
      (false, xi) ();
  in
   let
    val (_, aa) = a;
   in
    (fn () => aa)
   end
     ()
```







24

Choice on the data structures

Clauses: resizable arrays of (fixed sized) arrays

However, no aliasing

- Indices instead of pointers
- N[C] makes a copy, so only use N[C][i]





No error messages



Transformations before generating code







Clauses of length 0 and 1

Once combined with an initialisation:

<(IsaSAT\_code, model\_if\_satisfiable)
 ∈ [λN. each\_clause\_is\_distinct N ∧
 literals\_fit\_in\_32\_bit\_integer N]a
 clauses\_as\_lists<sup>k</sup> → model>

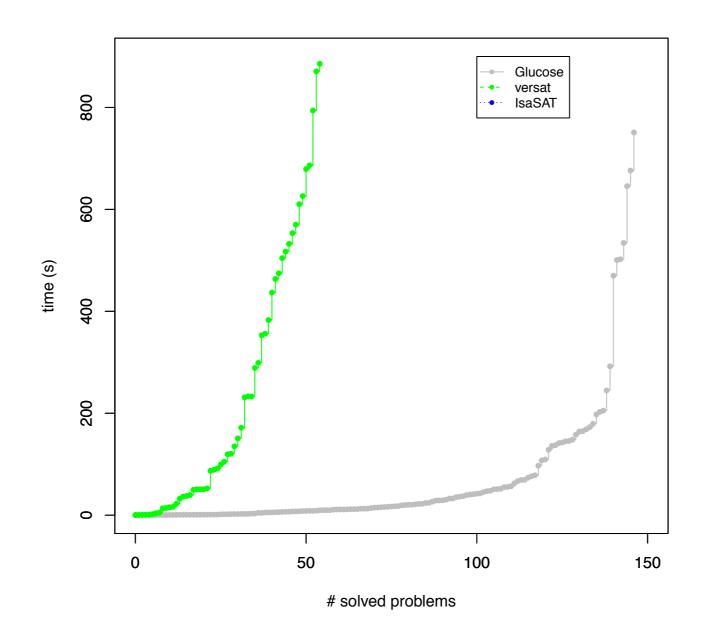
Exported code tested with an unchecked parser (easy and medium problems from the SAT competition 2009)







in Isabelle

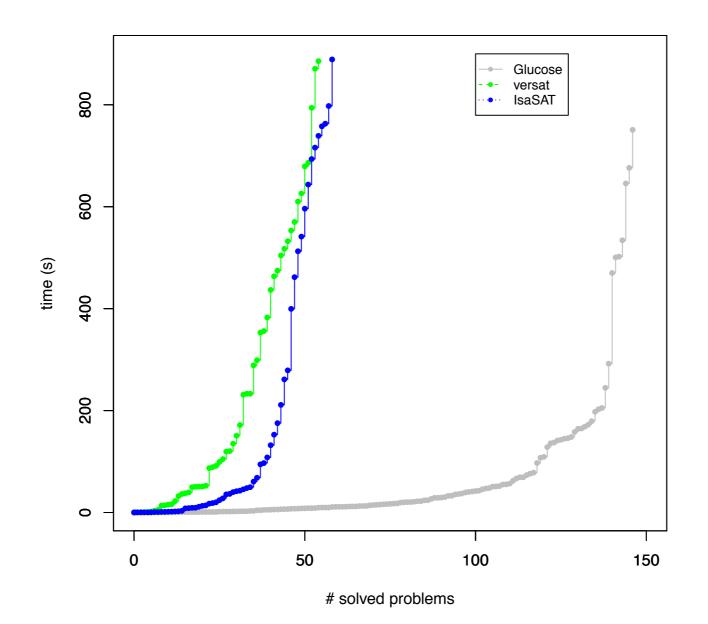


#### Performance of the first executable version







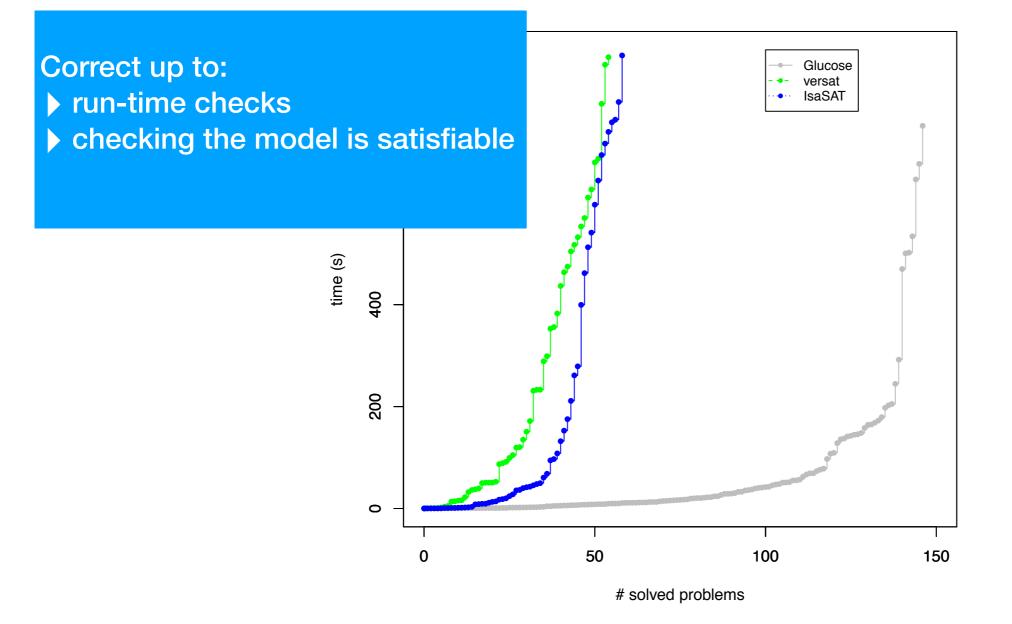


#### Performance of IsaSAT









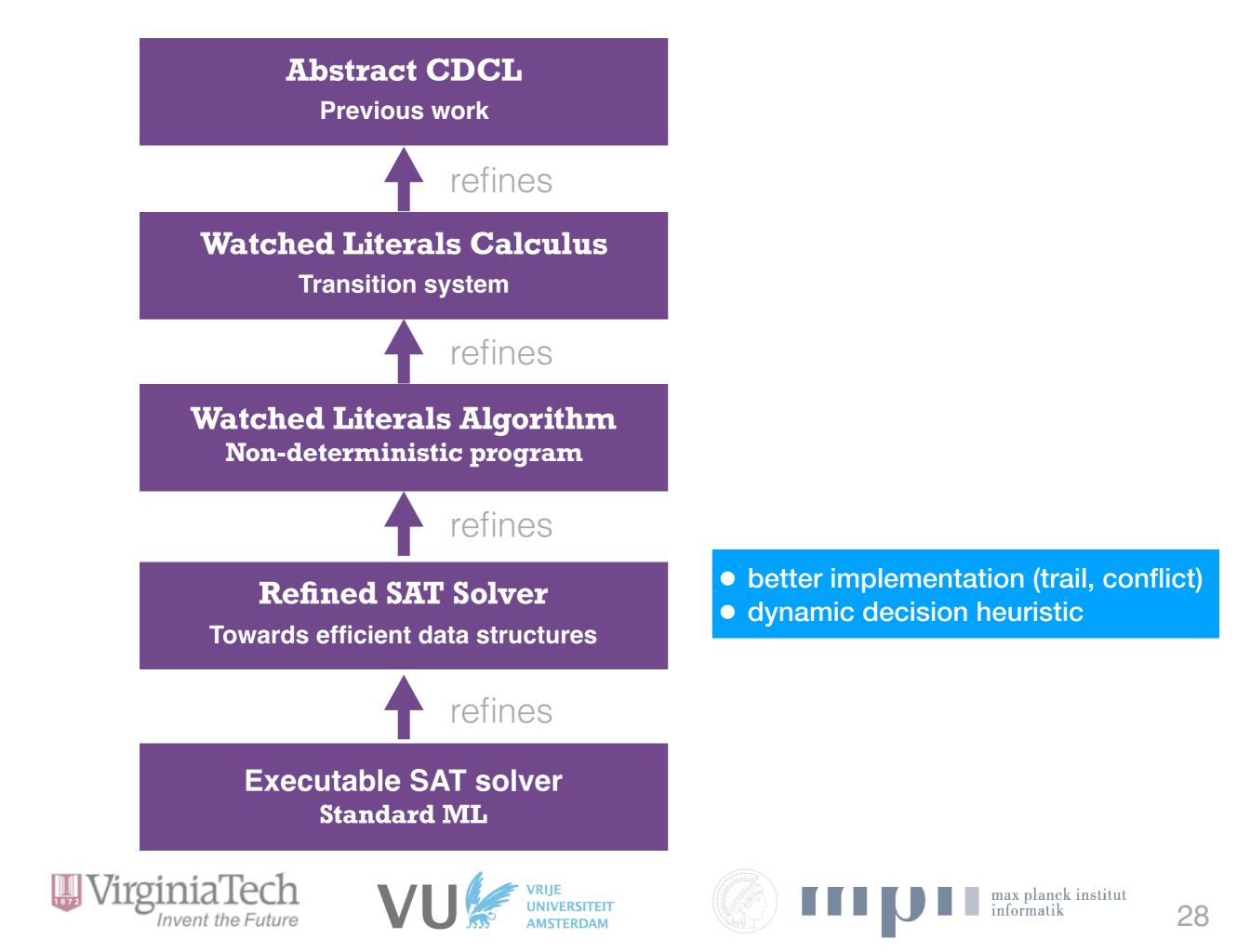
#### Performance of IsaSAT











#### **Abstract CDCL**

**Previous work** 



Watched Literals Calculus

**Transition system** 



Watched Literals Algorithm Non-deterministic program



**Refined SAT Solver** 

**Towards efficient data structures** 



Executable SAT solver Standard ML

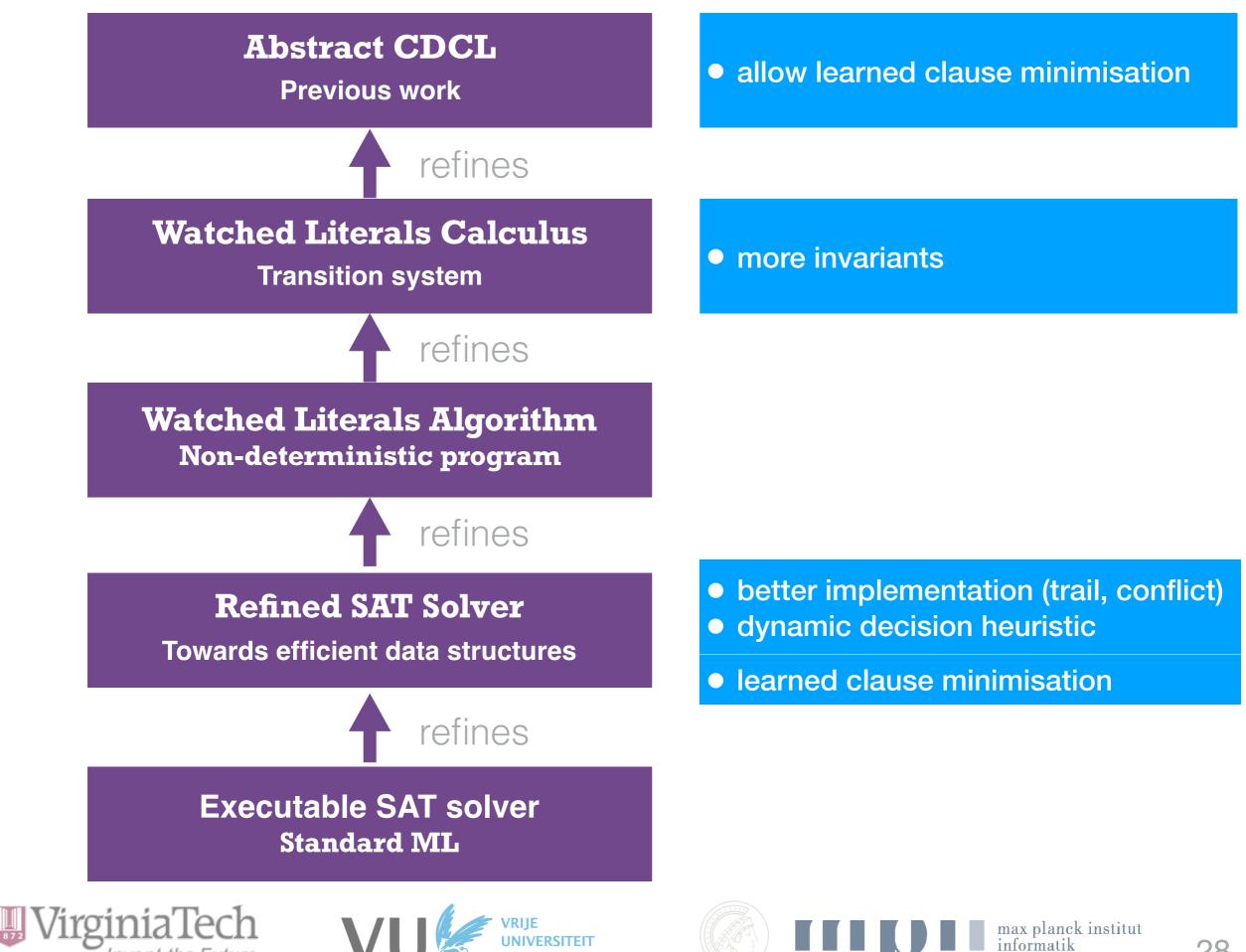




#### • allow learned clause minimisation

- better implementation (trail, conflict)
- dynamic decision heuristic
- learned clause minimisation





## How hard is it?

	Paper	Proof assistant
Very abstract	13 pages	50 pages
Abstract CDCL	9 pages (½ month)	90 pages (5 months)
Watched Literals	1 page	600 pages
	(C++ code of MiniSat)	(15 months)







# Conclusion

## Concrete outcome

- Watched literals optimisation
- Verified executable SAT solver

## Methodology

Refinement using the Refinement Framework

## Future work

- Restarts
- Use SAT solver in IsaFoR
- SAT Modulo Theories







(e.g., CVC or z3)

## Annex







```
for (i = j = 1; i < out_learnt.size(); i++)
if (reason(var(out_learnt[i])) == CRef_Undef II
    !litRedundant(out_learnt[i]))
    out_learnt[j++] = out_learnt[i];</pre>
```







```
fun minimize_and_extract_highest_lookup_conflict_code x =
 (fn ai => fn bid => fn bic => fn bib => fn bia => fn bi => fn () =>
  let
   val a =
    heap WHILET
     (fn (_, (a1a, (_, a2b))) =>
       (fn f_=> fn () => f_((length_arl_u_code heap_uint32 a2b) ()) ())
        (fn x_a => (fn () => (Word32.< (a1a, x_a)))))
     (fn (a1, (a1a, (a1b, a2b))) =>
       (fn f_ => fn () => f_
        (((fn () => Array.sub (fst a2b, Word32.toInt a1a))) ()) ())
        (fn x a =>
         (fn f_ => fn () => f_
          ((literal redundant wl lookup code ai bid a1 a1b x a bia) ())
           ())
           (fn (a1c, (_, a2d)) =>
            (if not a2d
             then (fn () =>
                 (a1, (Word32.+ (a1a, (Word32.fromInt 1)),
                      (a1c, a2b))))
             else (fn f => fn () => f
                  ((delete_from_lookup_conflict_code x_a a1) ()) ())
                 (fn x e =>
                   (fn f_=> fn () => f_((arl_last heap_uint32 a2b))
                    ())())
                    (fn xa =>
                     (fn f_ => fn () => f_
                      ((arl_set_u heap_uint32 a2b a1a xa) ()) ())
                      (fn xb =>
                        (fn f_{=} = fn () = f_{=}
((arl_butlast heap_uint32 xb) ()) ())
(bic, ((Word32.fromInt 1), (bib, bi))) ();
  in
   let
    val (a1, (_, (a1b, a2b))) = a;
```









# What is in IsaSAT?

## **Conflict Analysis**

- conflict as lookup table (Minisat)
- and as explicit array (Minisat's "outl", to simplify proofs)

## Decisions

Variable move to front (Splatz, cadical)

## Propagations

Mostly following MiniSAT (without BLIT)







Thank you, Norbert & Mate!

10

#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- learning additional units and binary clauses (multiple UIPs)
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15

#### **Slides by Armin Biere**

Features (II)

- stamping based VMTF instead of VSIDS
- subsumption for both irredundant and learned clauses
- inprocessing blocked clause decomposition (BCD) enabling ...
- ... inprocessing SAT sweeping for backbones and equivalences
- equivalent literal substitution (ELS)
- bounded variable elimination (BVE)
- blocked clause elimination (BCE)
- dynamic sticky clause reduction
- exponential moving average based restart scheduling
- delaying restarts
- trail reuse

Splatz @ POS'15







#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- learning additional units and binary clauses (multiple UIPs)
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15

#### Code only

Thank you, Norbert & Mate! Slides by Armin Biere

#### Features (II)

10

- stamping based VMTF instead of VSIDS
- subsumption for both irredundant and learned clauses
- inprocessing blocked clause decomposition (BCD) enabling ...
- ... inprocessing SAT sweeping for backbones and equivalences
- equivalent literal substitution (ELS)
- bounded variable elimination (BVE)
- blocked clause elimination (BCE)
- dynamic sticky clause reduction
- exponential moving average based restart scheduling
- delaying restarts
- trail reuse

Splatz @ POS'15







Thank you, Norbert & Mate!

10

#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- Iearning additional units and binary clauses (multiple UIPs
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15

#### Code only

Strengthening

#### **Slides by Armin Biere**

Features (II)

- stamping based VMTF instead of VSIDS
- subsumption for both irredundant and learned clauses
- inprocessing blocked clause decomposition (BCD) enabling ...
- ... inprocessing SAT sweeping for backbones and equivalences
- equivalent literal substitution (ELS)
- bounded variable elimination (BVE)
- blocked clause elimination (BCE)
- dynamic sticky clause reduction
- exponential moving average based restart scheduling
- delaying restarts
- trail reuse

Splatz @ POS'15







10

#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- learning additional units and binary clauses (multiple UIPs
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15

#### Code only

Strengthening





#### Thank you, Norbert & Mate!

#### **Slides by Armin Biere**

Features (II)

- stamping based VMTF instead of VSIDS
- subsumption for both irredundant and learned clauses
- inprocessing blocked clause decomposition (BCD) enabling ...
- ... inprocessing SAT sweeping for backbones and equivalences
- equivalent literal substitution (ELS)
- bounded variable elimination (BVE)
- blocked clause elimination (BCE)
- dynamic sticky clause reduction
- exponential moving average based restart scheduling
- delaying restarts
- trail reuse

Splatz @ POS'15



34

10

#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- learning additional units and binary clauses (multiple UIPs)
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15

# Code only

#### Strengthening





Thank you, Norbert & Mate!

#### **Slides by Armin Biere**

Features (II)

- stamping based VMTF instead of VSIDS
- subsumption for both irredundant and learned clauses
- inprocessing blocked clause decomposition (BCD) enabling ...
- ... inprocessing SAT sweeping for backbones and equivalences
- equivalent literal substitution (ELS)
- bounded variable elimination (BVE)
- blocked clause elimination (BCE)
- dynamic sticky clause reduction
- exponential moving average based restart scheduling
- delaying restarts
- trail reuse

Splatz @ POS'15



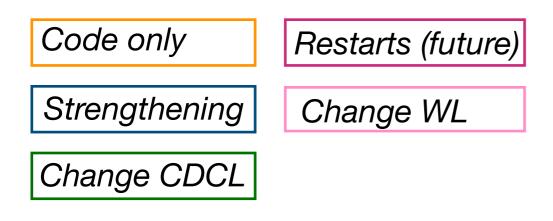




#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- Iearning additional units and binary clauses (multiple UIPs
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15



#### Thank you, Norbert & Mate!

10

#### **Slides by Armin Biere**

Features (II)

- stamping based VMTF instead of VSIDS
- subsumption for both irredundant and learned clauses
- inprocessing blocked clause decomposition (BCD) enabling ...
- ... inprocessing SAT sweeping for backbones and equivalences
- equivalent literal substitution (ELS)
- bounded variable elimination (BVE)
- blocked clause elimination (BCE)
- dynamic sticky clause reduction
- exponential moving average based restart scheduling
- delaying restarts
- trail reuse

Splatz @ POS'15







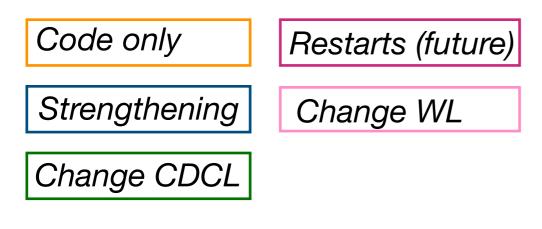
Thank you, Norbert & Mate!

10

#### Features (I)

- arena based memory allocation for clauses and watchers
- blocking literals (BLIT)
- special handling of binary clause watches
- literal-move-to-front watch replacement (LMTF)
- learned clause minimization with poison
- on-the-fly hyper-binary resolution (HBR)
- learning additional units and binary clauses (multiple UIPs)
- on-the-fly self-subsuming resolution (OTFS)
- decision only clauses (DECO)
- failed literal probing on binary implication graph roots
- eager recent learned clause subsumption

Splatz @ POS'15



#### **Slides by Armin Biere** Features (II) 11 stamping based VMTE instead of VSIDS Unchecked array accesses (Isabelle takes care) of it) • No unbounded integers (in theory, not complete anymore) Restarts exponential moving average based restart scheduling delaying restarts trail reuse

Splatz @ POS'15







## A first idea

#### A better strategy



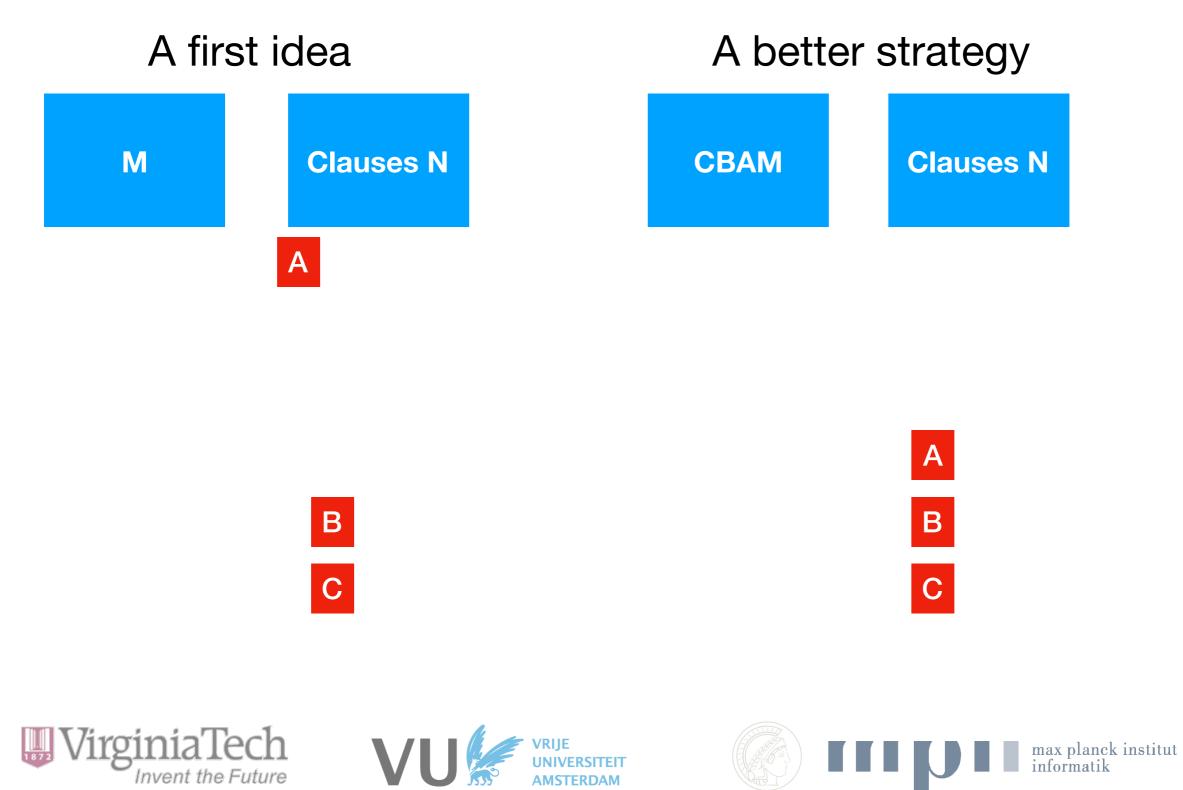


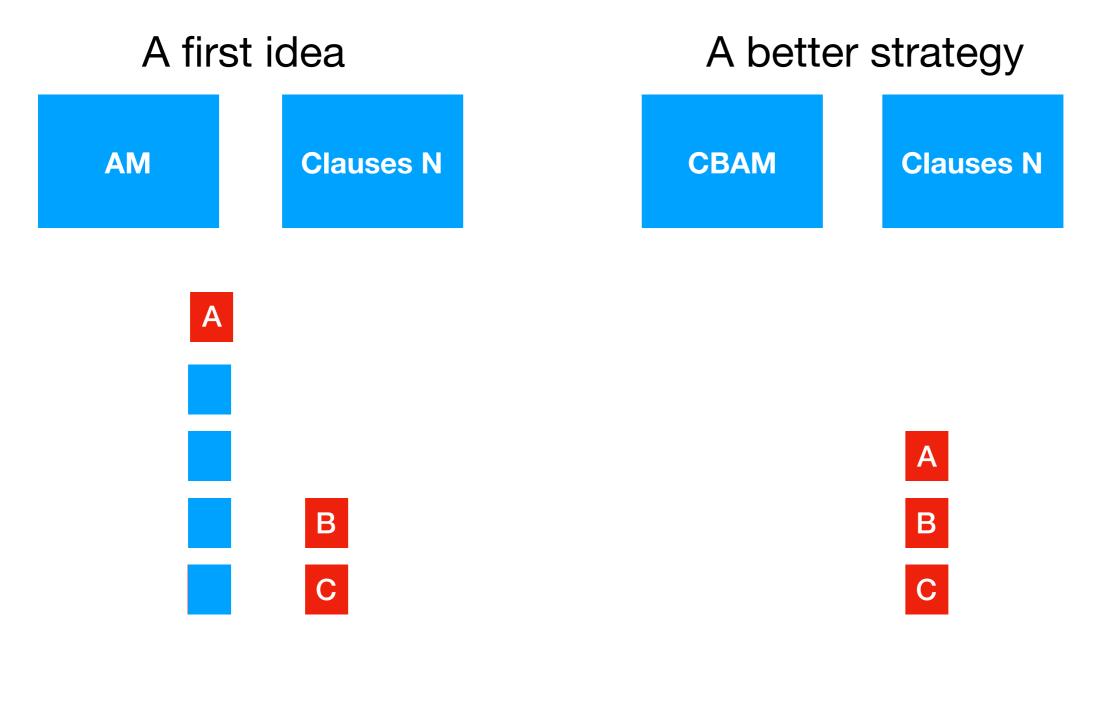








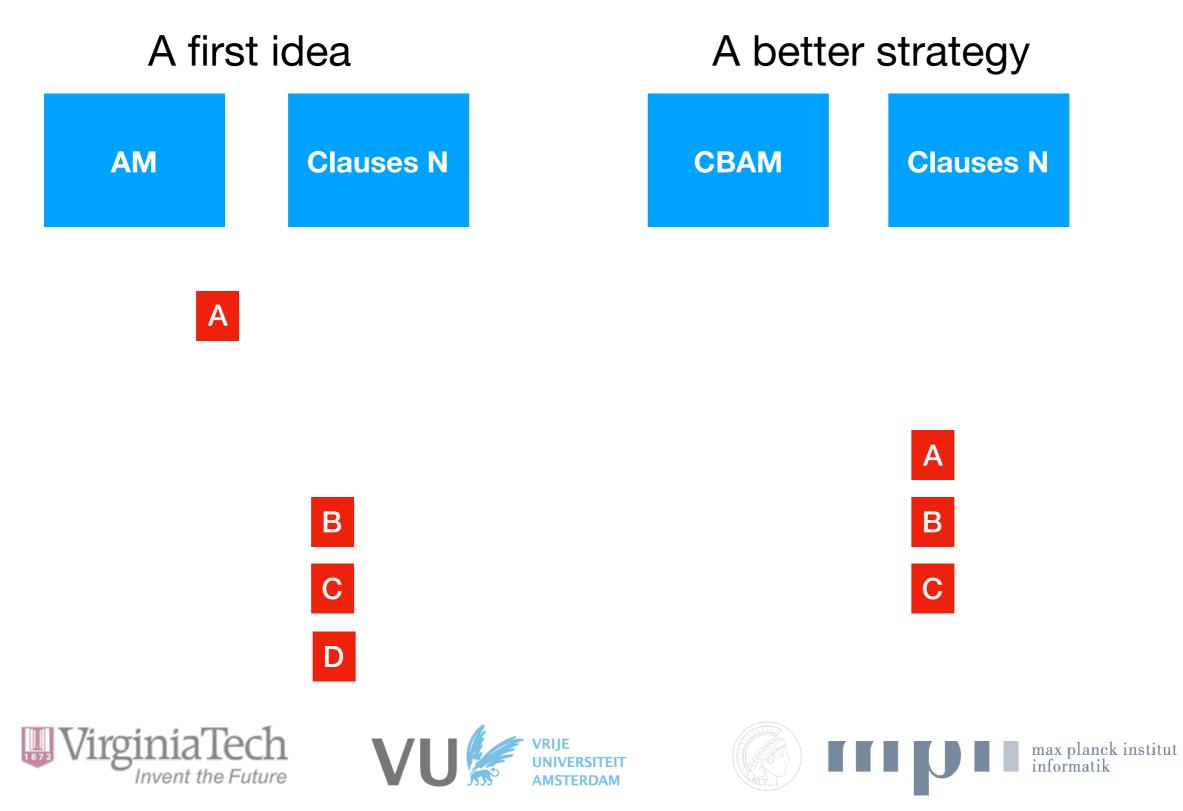


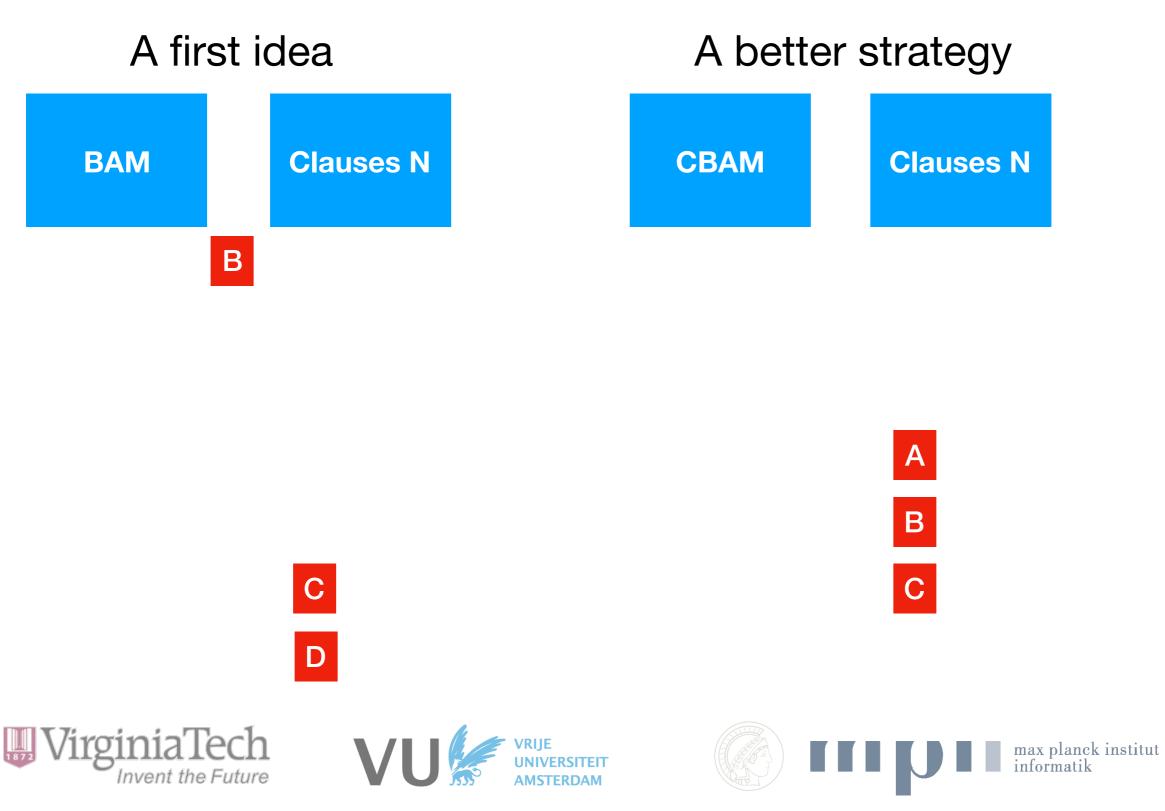


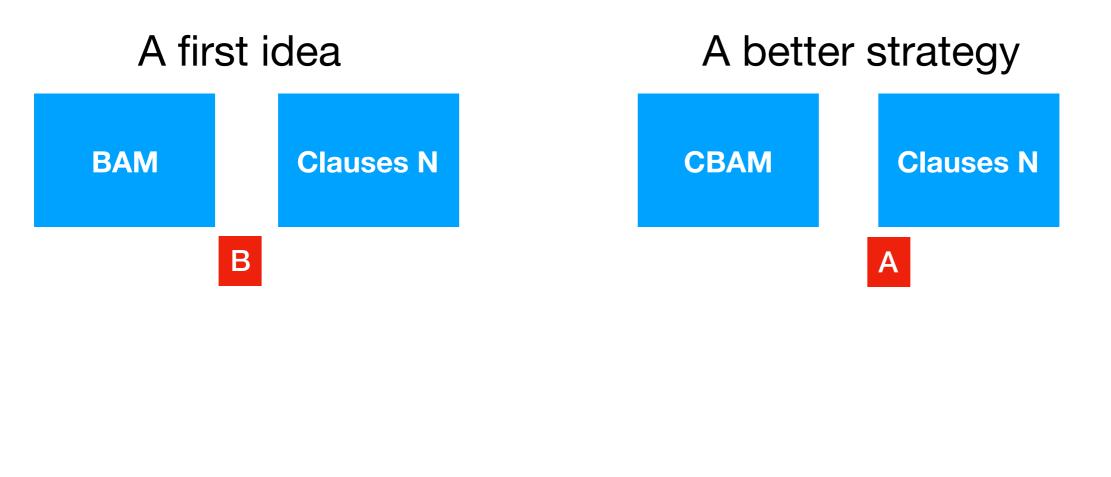




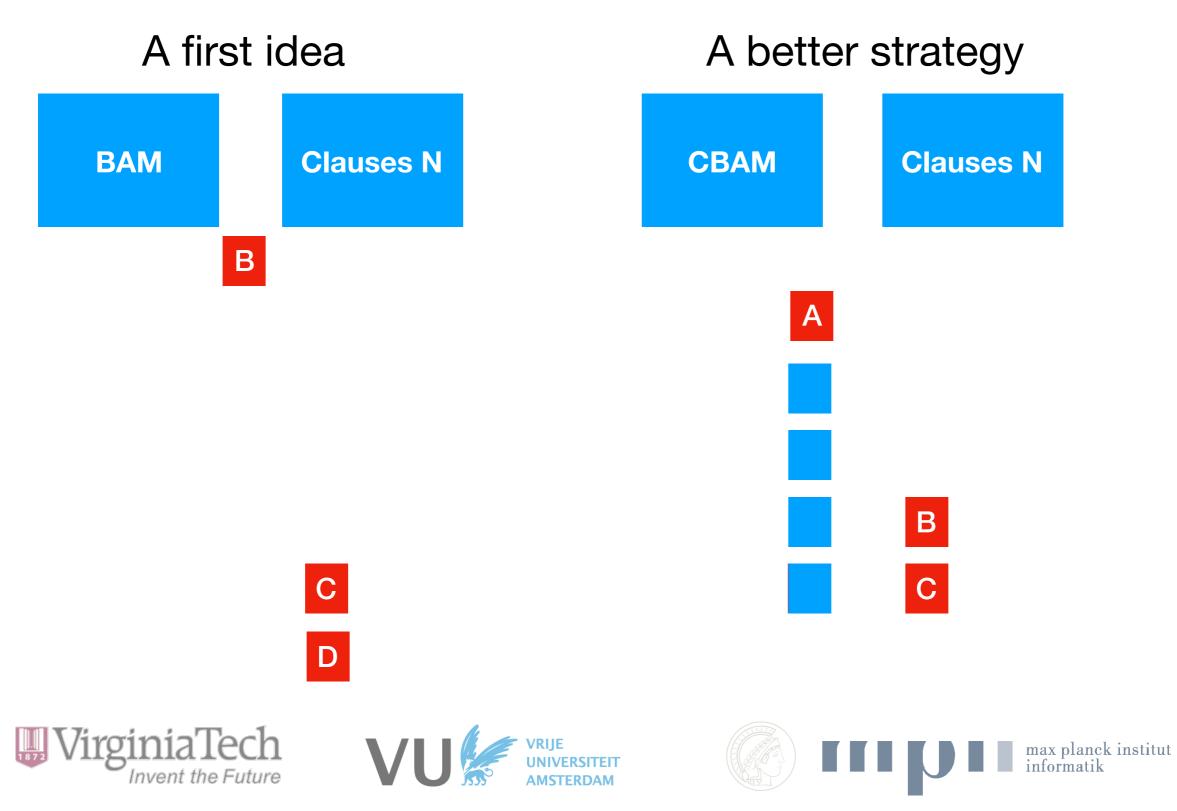












## A first idea

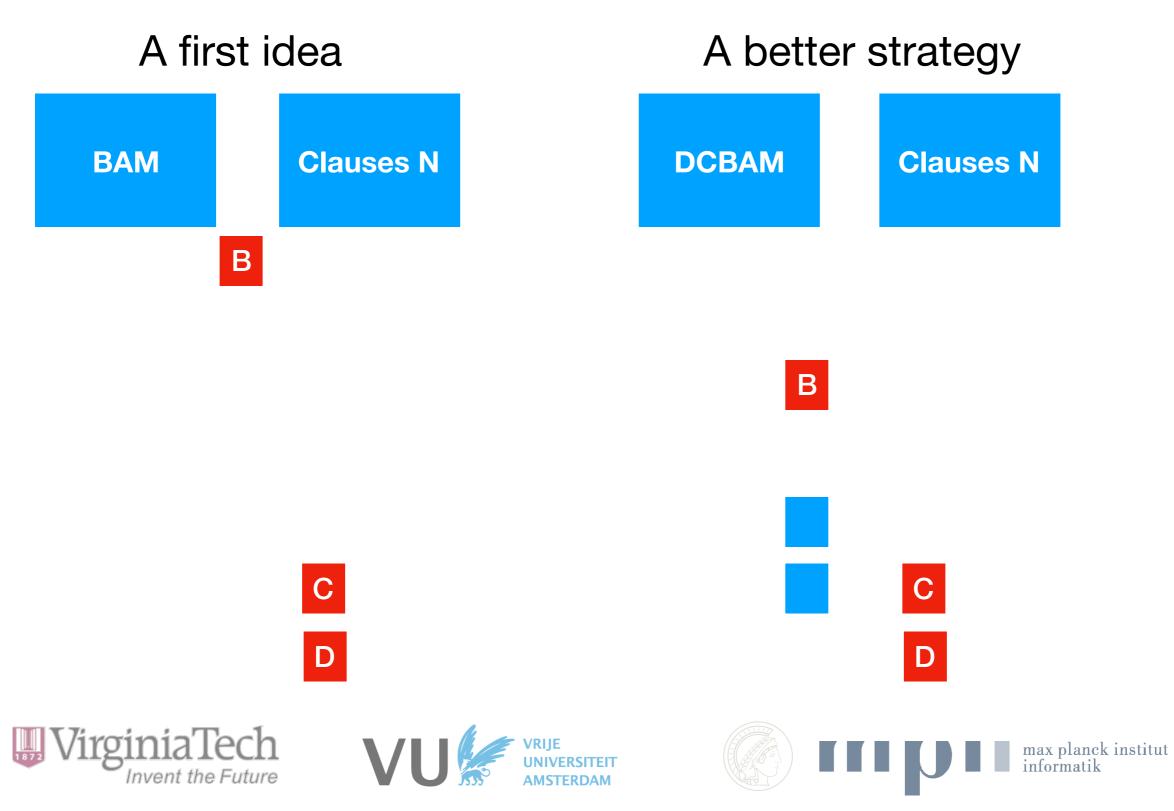
## A better strategy





# A first idea A better strategy BAM Clauses N DCBAM Clauses N





## A first idea

#### A better strategy









