On Propositional Model Counting and Enumeration and Automated Model Building in Predicate Logic

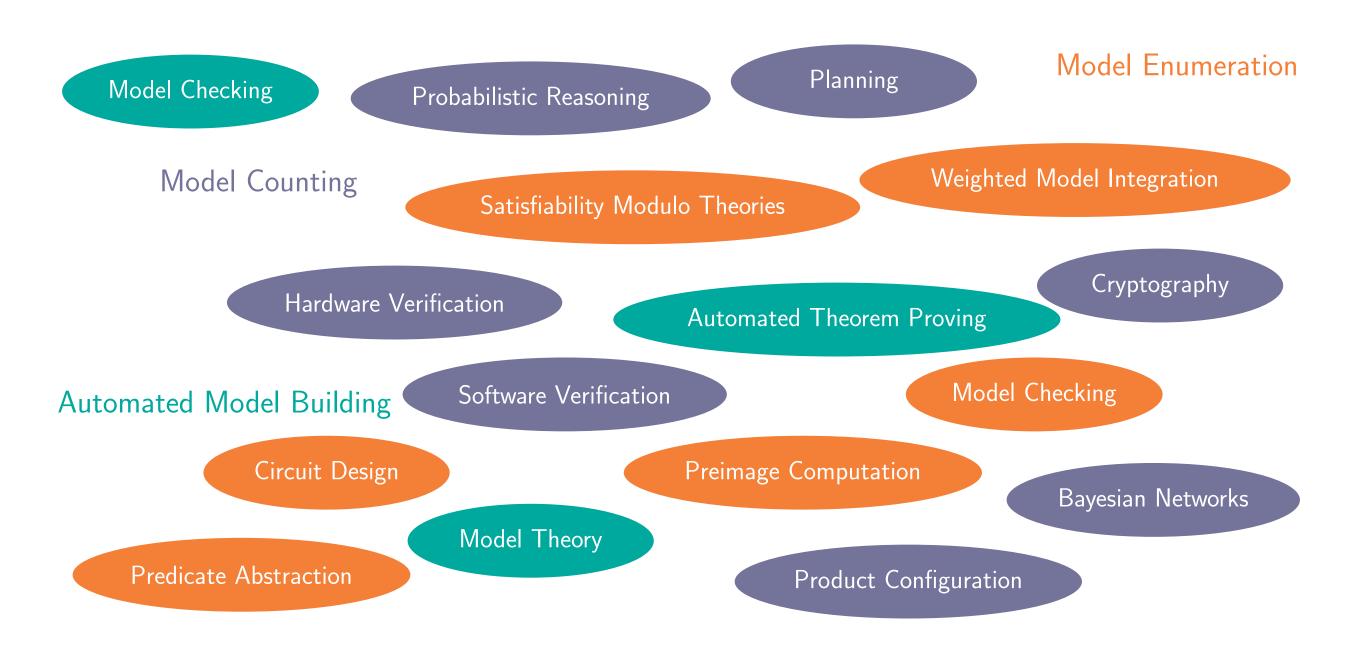
Sibylle Möhle

Theoretical Computer Science Group University of Regensburg

EuroProofNet Workshop on Program Verification

September 17, 2025

Where Do These Things Come Into Play?

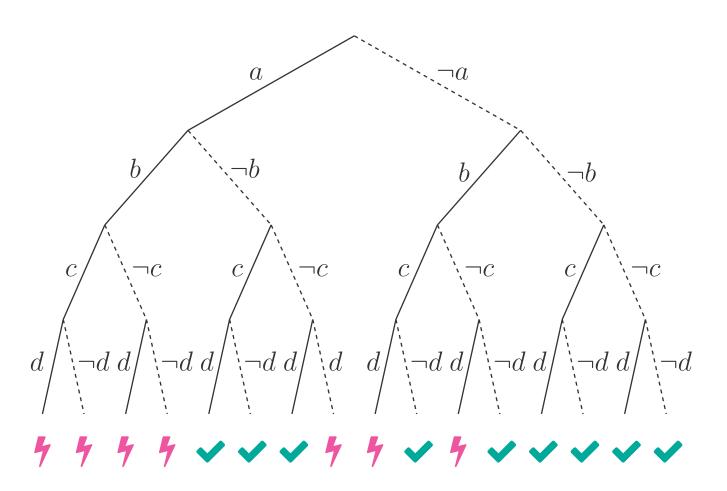


Counting and Enumerating Propositional Models

Challenge: Search Space Need be Explored Exhaustively

n	$2^{ n }$ assignments
1	2
2	4
10	1'024
100	$\approx 1.27 \cdot 10^{30}$
1000	$\approx 1.07 \cdot 10^{301}$

$$F = (\neg a \lor \neg b) \land (a \lor \neg b \lor \neg d) \land (\neg a \lor b \lor c \lor d)$$



State-Of-The-Art Solution for Model Counting: Component-Based Reasoning

$$F = \underbrace{(a \lor b)}_{\mathcal{C}_1} \land \underbrace{(c \lor d)}_{\mathcal{C}_2}$$

$$\Downarrow$$

$$\#F = 9 = \#\mathcal{C}_1 \cdot \#\mathcal{C}_2$$

R.J. Bayardo, J.D. Pehoushek. Counting Models Using Connected Components". AAAI'00. T. Sang et al. Combining Component Caching and Clause Learning for Effective Model Counting. SAT'04. M. Thurley. sharpSAT—Counting Models with Advanced Component Caching and Implicit BCP. SAT'06. J. Burchard, T. Schubert, B. Becker. Laissez-Faire Caching for Parallel #SAT Solving. SAT'15. J. Burchard, T. Schubert, B. Becker. Distributed Parallel #SAT Solving. CLUSTER'16.

State-Of-The-Art Solution for Model Counting: Component-Based Reasoning

$$F = \underbrace{(a \lor b)}_{\mathcal{C}_1} \land \underbrace{(c \lor d)}_{\mathcal{C}_2}$$

$$\Downarrow$$

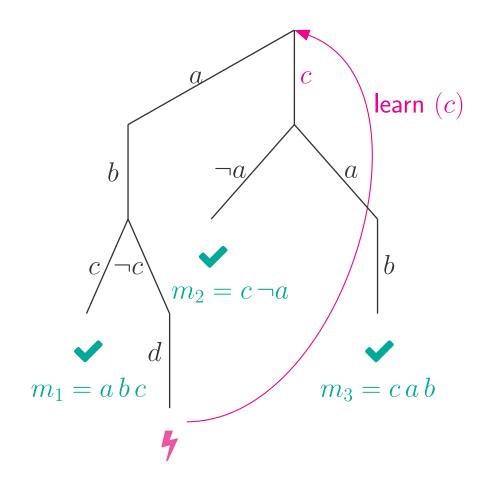
$$\#F = 9 = \#\mathcal{C}_1 \cdot \#\mathcal{C}_2$$

reduces work in individual computation

R.J. Bayardo, J.D. Pehoushek. Counting Models Using Connected Components". AAAI'00. T. Sang et al. Combining Component Caching and Clause Learning for Effective Model Counting. SAT'04. M. Thurley. sharpSAT—Counting Models with Advanced Component Caching and Implicit BCP. SAT'06. J. Burchard, T. Schubert, B. Becker. Laissez-Faire Caching for Parallel #SAT Solving. SAT'15. J. Burchard, T. Schubert, B. Becker. Distributed Parallel #SAT Solving. CLUSTER'16.

Challenge: Avoid Finding Models Multiple Times in CDCL With Backjumping

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d) \land (c)$$

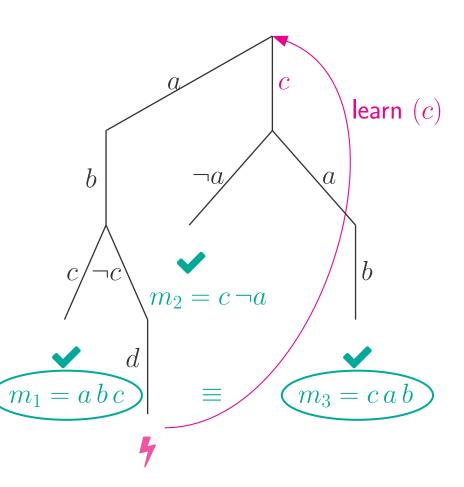


J.P. Marques-Silva, K.A. Sakallah. *GRASP—A New Search Algorithm for Satisfiability.* ICCAD'96. J.P. Marques-Silva, K.A. Sakallah. *GRASP: A Search Algorithm for Propositional Satisfiability.* IEEE Trans Comput, 1999. M.W. Moskewicz, C.F. Madigan, Y. Zhao, L. Zhang, S. Malik. *Chaff: Engineering an Efficient SAT Solver.* DAC'01.

Challenge: Avoid Finding Models Multiple Times in CDCL With Backjumping

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d) \land \textcolor{red}{(c)}$$

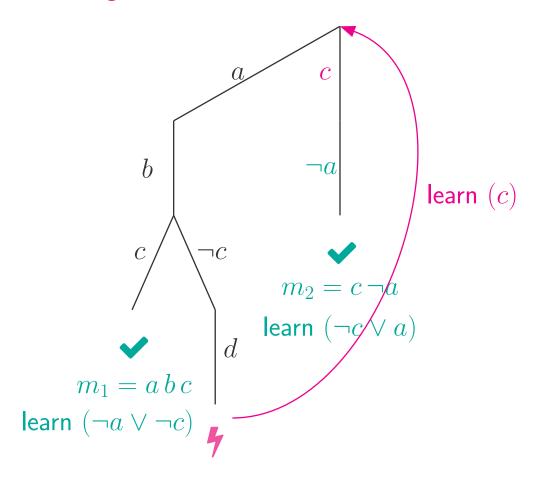
- learn from conflicts
- might find redundant models



J.P. Marques-Silva, K.A. Sakallah. *GRASP—A New Search Algorithm for Satisfiability.* ICCAD'96. J.P. Marques-Silva, K.A. Sakallah. *GRASP: A Search Algorithm for Propositional Satisfiability.* IEEE Trans Comput, 1999. M.W. Moskewicz, C.F. Madigan, Y. Zhao, L. Zhang, S. Malik. *Chaff: Engineering an Efficient SAT Solver.* DAC'01.

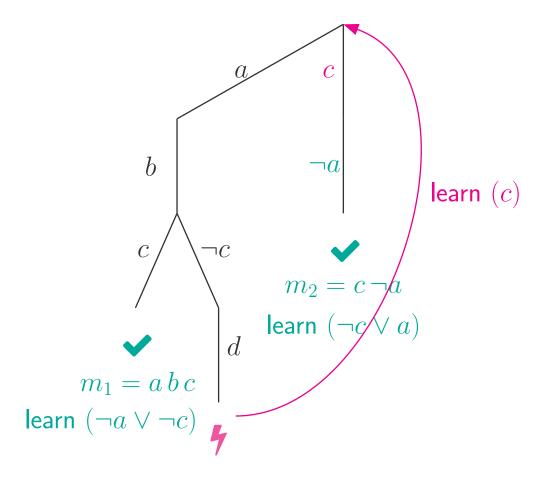
Blocking clauses

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d) \land (\neg a \lor \neg c) \land (c)$$



Blocking clauses

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d) \land (\neg a \lor \neg c) \land (c)$$

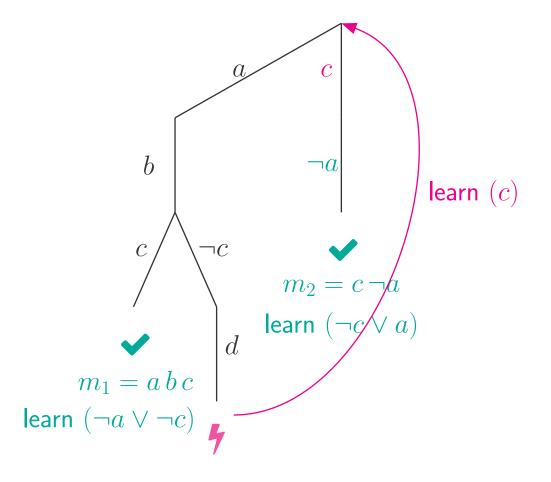


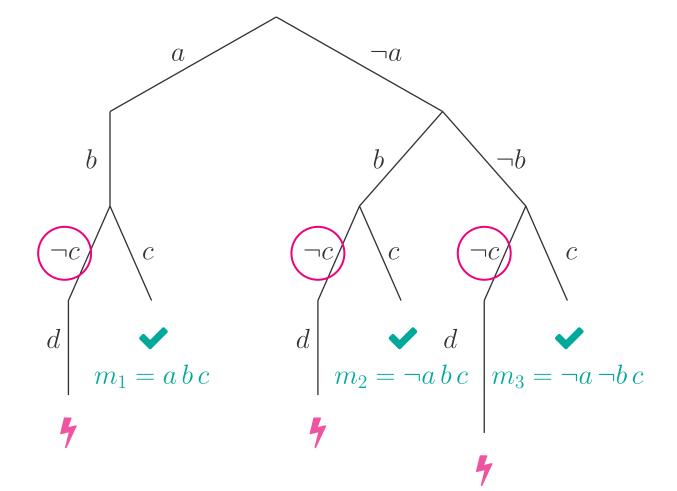
- if finds no redundant models
- learns from conflicts
- blocking clauses must be kept ⇒ formula blowup

Blocking clauses

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d)$$

Chronological backtracking



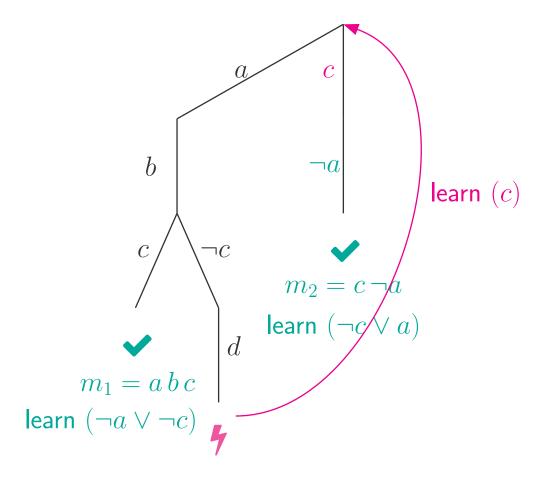


- finds no redundant models
- learns from cor
- blocking clause

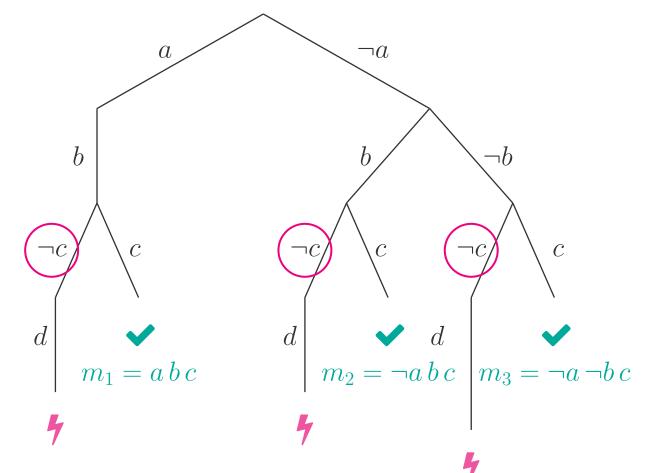
Blocking clauses

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d)$$

Chronological backtracking



- if inds no redundant models
- learns from conflicts
- blocking clauses must be kept ⇒ formula blowup

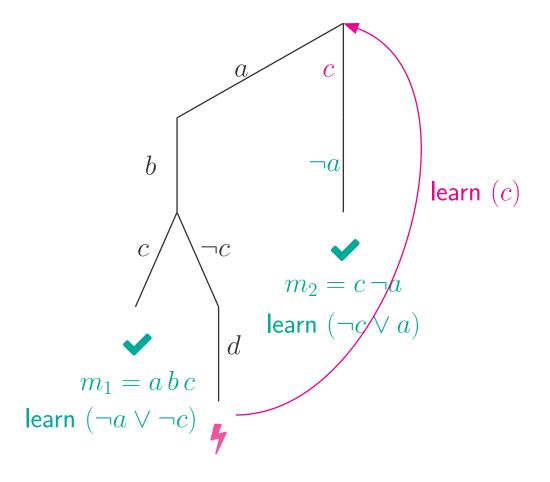


- if inds no redundant models
- no formula blowup
- no learning from conflicts

Blocking clauses

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d)$$

Chronological CDCL



- if finds no redundant models
- learns from cor
- blocking clause

A. Nadel, V. Ryvchin, *Chronological Backtracking*. SAT'18. S. Möhle, A. Biere, *Backing Backtracking*. SAT'19.

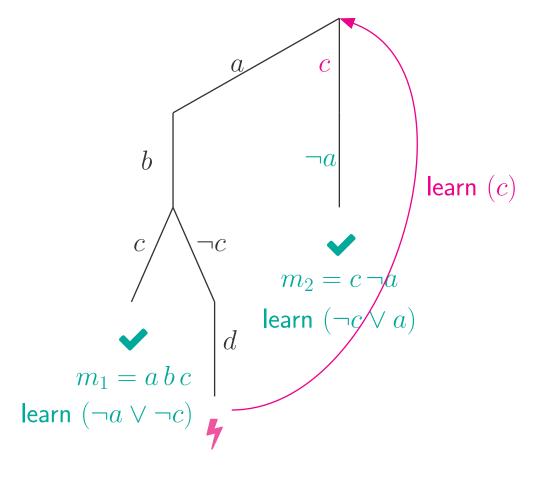
S. Möhle, A. Biere. Combining Conflict-Driven Clause Learning and Chronological Backtracking for Propositional Model Counting. GCAl'19.

S. Möhle, R. Sebastiani, A. Biere. Four Flavors of Entailment. SAT'20.

Blocking clauses

$$F = (\neg a \lor b) \land (c \lor d) \land (c \lor \neg d)$$

Chronological CDCL



- if finds no redundant models
- learns from conflicts
- \P blocking clauses must be kept \Rightarrow formula blowup

- if inds no redundant models
- no formula blowup
- learning from conflicts

Counting and Enumerating Propositional Models Using Chronological CDCL

Generalizing,

$$\#F = \sum_{C \in M} 2^{|V-C|}$$
 and $F \equiv \bigvee_{C \in M} C$

and

M is a Disjoint-Sum-of-Products (DSOP) representation of F

- lacksquare M is a disjunction of conjunctions of literals (cubes)
- lacktriangle The cubes in M are pairwise contradicting
- $lue{M}$ is logically equivalent to F
- lacksquare M is not unique

Counting and Enumerating Propositional Models Using Chronological CDCL

$$F = (\overline{p} \vee q) \wedge (p \vee q)$$

$$\longrightarrow$$

$$Rules$$

$$M = (r \wedge q) \vee (\overline{r} \wedge p \wedge q) \vee (\overline{r} \wedge \overline{p} \wedge q) = C_1 \vee C_2 \vee C_3$$

$$M \equiv F \quad \text{and} \quad \#M = \sum_{i=1}^3 2^{|V-C_i|} = 4 = \#F$$

Generalizing,

$$\#F = \sum_{C \in M} 2^{|V-C|}$$
 and $F \equiv \bigvee_{C \in M} C$

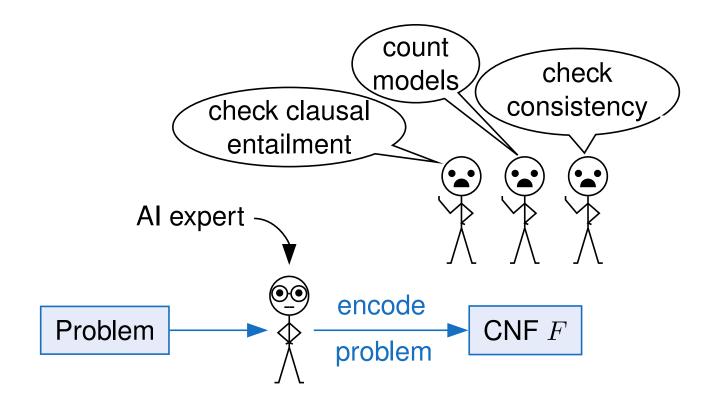
and

M is a Disjoint-Sum-of-Products (DSOP) representation of F \longleftarrow we did some kind of knowledge compilation, right?

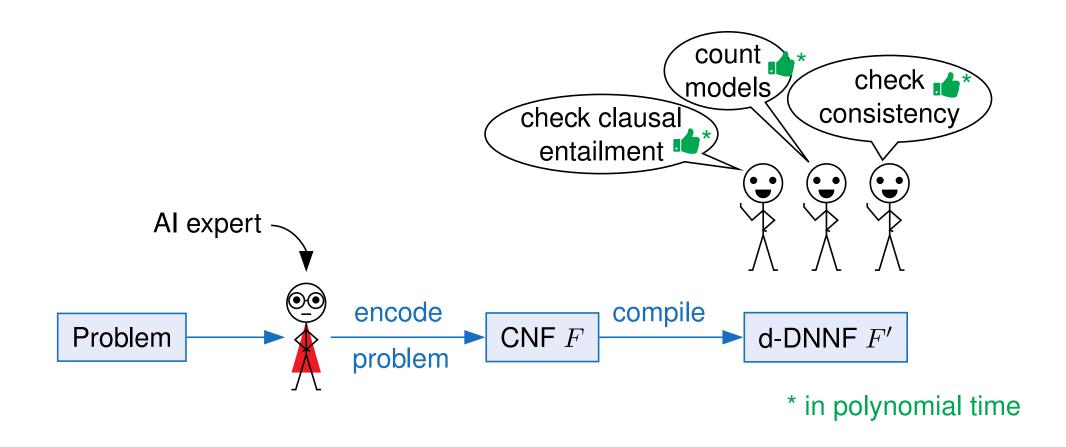
- \blacksquare M is a disjunction of conjunctions of literals (cubes)
- lacktriangle The cubes in M are pairwise contradicting
- $lue{M}$ is logically equivalent to F
- lacksquare M is not unique

Why Is Knowledge Compilation Interesting, Again?

What's the Problem with Conjunctive Normal Form (CNF)?



What's the Problem with Conjunctive Normal Form (CNF)?



Deterministic Decomposable Negation Normal Form (d-DNNF)

d-DNNF:
$$F=(a \wedge (b \vee (\neg b \wedge c)) \wedge d) \vee (\neg a \wedge b)$$

negations in front of variables

for all conjunctions: conjuncts do not share variables

for all disjunctions: disjuncts are pairwise contradicting

Model Counting in d-DNNF

$$\textbf{(1)} \ \ F = G \land H \ \implies \ \#F = \#G \cdot \#H \quad \text{ provided } \mathsf{var}(G) \cup \mathsf{var}(H) = \mathsf{var}(F) \text{ and } \mathsf{var}(G) \cap \mathsf{var}(H) = \emptyset$$

Model Counting in d-DNNF

(1)
$$F = G \land H \implies \#F = \#G \cdot \#H \quad \text{provided } \text{var}(G) \cup \text{var}(H) = \text{var}(F) \text{ and } \text{var}(G) \cap \text{var}(H) = \emptyset$$

(2)
$$F = C \lor D \implies \#F = 2^{|\mathsf{var}(F)| - |\mathsf{var}(C)|} + 2^{|\mathsf{var}(F)| - |\mathsf{var}(D)|} \quad \mathsf{provided} \ C \land D \equiv \bot$$

$$F = (a \land (b \lor (\neg b \land c)) \land d) \lor (\neg a \land b)$$
 (2)

$$\#F = \#(a \land (b \lor (\neg b \land c)) \land d) \cdot 2^0 + \#(\neg a \land b) \cdot 2^2$$

$$\tag{1}$$

$$= [\#(a) \cdot \#(\underline{b} \vee (\neg b \wedge c)) \cdot \#(d)] \cdot 2^{0} + [\#(\neg a) \cdot \#(b)] \cdot 2^{2}$$
 (2)

$$= [1 \cdot [\#(b) \cdot 2^1 + \#(\neg b \wedge c)] \cdot 2^0] + [1 \cdot 1] \cdot 2^2 \tag{1}$$

$$= [1 \cdot [1 \cdot 2^{1} + [\#(\neg b) \cdot \#(c)] \cdot 2^{0}] + 2^{2}$$

$$= [1 \cdot [1 \cdot 2^1 + [1 \cdot 1] \cdot 2^0] + 2^2 = 7$$

CNF vs. d-DNNF — the Model Counting Case

CNF:
$$F = (\neg a \lor b \lor c) \land (\neg a \lor d) \land (a \lor b)$$

d-DNNF:
$$F' = (a \land (b \lor (\neg b \land c)) \land d) \lor (\neg a \land b)$$

CNF vs. d-DNNF — the Model Counting Case

CNF:
$$F = (\neg a \lor b \lor c) \land (\neg a \lor d) \land (a \lor b)$$

$$\#F = ? \quad \text{not that easy}$$

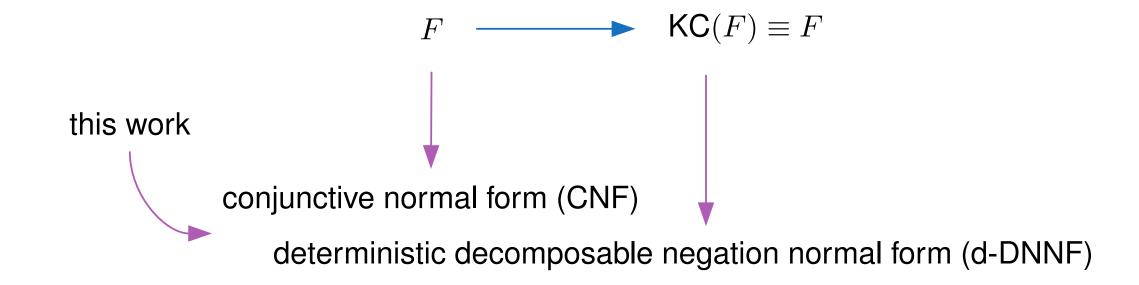
d-DNNF:
$$F' = (a \land (b \lor (\neg b \land c)) \land d) \lor (\neg a \land b)$$

$$\#F' = (1 \cdot (1 \cdot 2^1 + 1 \cdot 1 \cdot 2^0) \cdot 1) \cdot 2^0 + (1 \cdot 1) \cdot 2^2 = 7$$

The Task: Knowledge Compilation

$$F \longrightarrow \mathsf{KC}(F) \equiv F$$

The Task: CNF-to-d-DNNF Compilation



Automated Model Building

What do Knowledge Compilation and Automated Model Building Have in Common?

Knowledge compilation:

Translate a logical formula into another language in which some tasks of interest are executable efficiently.

Automated model building:

Find and finitely represent typically infinite models of a formula facilitating their computational processing.



The focus is on a convenient representation.

Automated Model Building Postulates

Let Σ be a vocabulary and F be a formula over Σ . Then a model representation formalism should ideally meet the following postulates:

- Uniqueness. Each model representation M specifies a single interpretation over Σ .
- Atom Test. There exists a fast procedure to evaluate arbitrary ground atoms over the signature Σ in M.
- Formula Evaluation. There exists an algorithm deciding the truth value of an arbitrary formula over Σ in M.
- **Equivalence Test.** There exists an algorithm deciding whether two representations M and M' over Σ describe the same interpretation.

The Satisfiability Problem of First-Order Logic

```
\begin{array}{ll} \textbf{Signature} & \Sigma = (\Omega, \Pi) \quad \text{where} \quad \Omega \quad \text{set of function symbols of arity} \geq 0 \\ & \Pi \quad \text{set of predicate symbols of arity} \geq 0 \\ & \mathcal{X} \quad \text{set of variables} \end{array}
```

Term $t \in T(\Sigma, \mathcal{X})$ where t is either a $x \in \mathcal{X}$ or $f(t_1, \ldots, t_n)$ with $f \in \Omega$ and $t_1, \ldots, t_n \in T(\Sigma, \mathcal{X})$

Ground term $t \in T(\Sigma)$ where t contains no variables

Atom $P(t_1, \ldots, t_n)$ where $P \in \Pi$ and $t_1, \ldots, t_n \in T(\Sigma, \mathcal{X})$

Literal an atom or its negation

Clause a disjunction of literals

Substitution $\sigma: \mathcal{X} \to T(\Sigma, \mathcal{X})$ such that $\sigma(x) \neq x$ for only finitely many variables

Unification terms s and t are unifiable if there exists a substitution σ such that $\sigma(s) = \sigma(t)$

Solving the Satisfiability Problem of First-Order Logic

$$\Sigma = (\{a/0, f/1\}, \{P/1\}) \qquad N = \{\underbrace{P(a)}_{C_1}, \underbrace{\neg P(x) \lor P(f(x))}_{C_2}\}$$

- 1. Propagate P(a) from C_1 : $\Gamma = P(a)$
- 2. Apply $\sigma = \{x \mapsto a\}$ to C_2 and propagate P(f(a)): $\Gamma = P(a) P(f(a))$
- 3. Apply $\sigma = \{x \mapsto f(a)\}$ to C_2 and propagate P(f(f(a))): $\Gamma = P(a) P(f(a)) P(f(f(a)))$
- 4. . . .

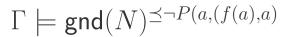
But the infinite model of N could be just represented by P(x)...

Representing Interpretations as Set of Literals

$$N = \{ P(a, f(x), a), P(x, f(x), x) \lor \neg P(x, g(x), x) \}$$

$$\Gamma = [P(a, f(a), a), \neg P(a, g(a), a)]$$

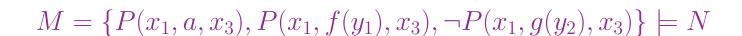
$$\Sigma = (\{a/0, f/1, g/1\}, \{P/3\})$$



Knuth-Bendix ordering with

$$\neg P \succ P \succ f \succ g \succ a$$

$$w(P) = w(f) = w(g) = w(a) = 1$$



Given $N = \{P(a, f(x), a), P(x, f(x), x) \lor P(x, g(x), x)\}$ and a trail $\Gamma = P(a, f(a), a) \neg P(a, g(a), a)$

Γ	Δ
$P(a, f(a), a)$ $\neg P(a, g(a), a)$	
$\mathcal{I} \models \Gamma$	$M \models \Delta$
$P(x_1, x_2, x_3)$	

Given $N = \{P(a, f(x), a), P(x, f(x), x) \lor P(x, g(x), x)\}$ and a trail $\Gamma = P(a, f(a), a) \neg P(a, g(a), a)$

Δ
$M \models \Delta$

Given $N = \{P(a, f(x), a), P(x, f(x), x) \lor P(x, g(x), x)\}$ and a trail $\Gamma = P(a, f(a), a) \neg P(a, g(a), a)$

Γ	Δ
$P(a, f(a), a)$ $\neg P(a, g(a), a)$	
$\mathcal{I} \models \Gamma$	$M \models \Delta$
$P(x_1, a, x_3)$ $P(x_1, f(y_1), x_3)$ $P(x_1, g(y_2), x_3)$	

Given $N = \{P(a, f(x), a), P(x, f(x), x) \lor P(x, g(x), x)\}$ and a trail $\Gamma = P(a, f(a), a) \neg P(a, g(a), a)$

Γ	Δ
$P(a, f(a), a)$ $\neg P(a, g(a), a)$	
$\mathcal{I} \models \Gamma$	$M \models \Delta$
$P(x_1, f(y_1), x_3)$ $P(x_1, g(y_2), x_3)$	$P(x_1, a, x_3)$

Given $N = \{P(a, f(x), a), P(x, f(x), x) \lor P(x, g(x), x)\}$ and a trail $\Gamma = P(a, f(a), a) \neg P(a, g(a), a)$

Wanted: representation of models of N= generalization of Γ

Γ	Δ
$\neg P(a, g(a), a)$	P(a, f(a), a)
$\mathcal{I} \models \Gamma$	$M \models \Delta$
$P(x_1, g(y_2), x_3)$	$P(x_1, a, x_3)$
	$P(x_1, f(y_1), x_3)$

Given $N = \{P(a, f(x), a), P(x, f(x), x) \lor P(x, g(x), x)\}$ and a trail $\Gamma = P(a, f(a), a) \neg P(a, g(a), a)$

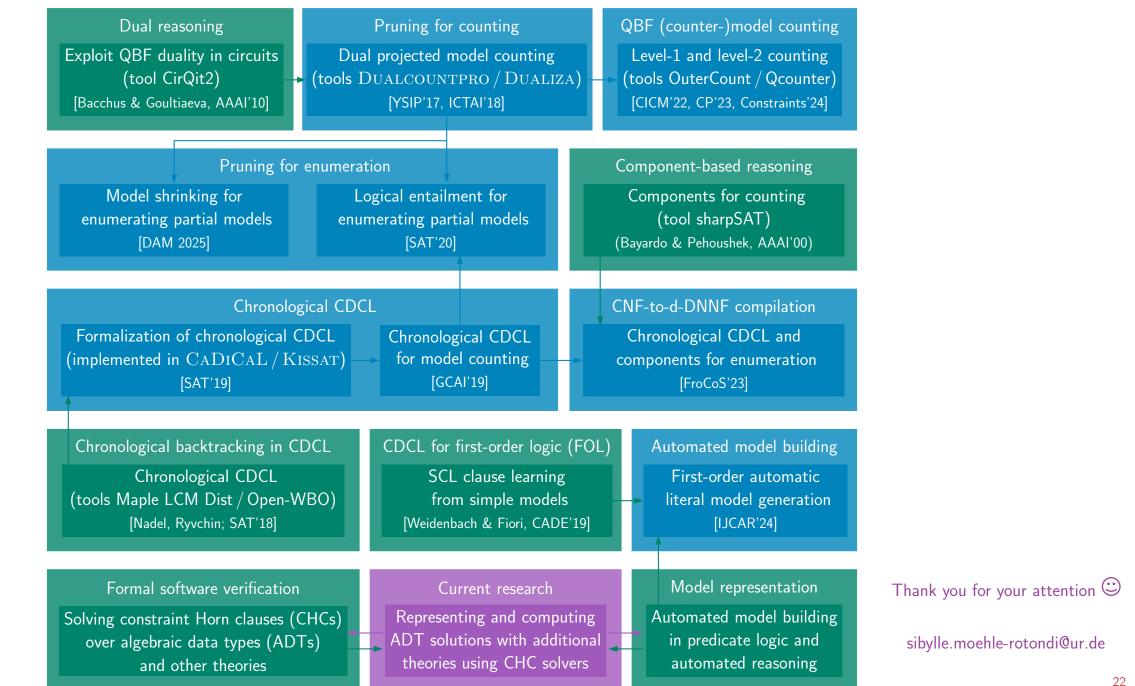
Wanted: representation of models of N= generalization of Γ

Γ	Δ
	$P(a, f(a), a)$ $\neg P(a, g(a), a)$
$\mathcal{I} \models \Gamma$	$M \models \Delta$
	$P(x_1, a, x_3)$ $P(x_1, f(y_1), x_3)$ $P(x_1, g(y_2), x_3)$

Some Context



Some Context



From DSOP to d-DNNF

$$F = (\neg a \lor b \lor c) \land (\neg a \lor d) \land (a \lor b)$$

$$\mathsf{DSOP}(F) = (\mathbf{a} \wedge b \wedge \mathbf{d}) \vee (\mathbf{a} \wedge \neg b \wedge c \wedge \mathbf{d}) \vee (\neg a \wedge b)$$

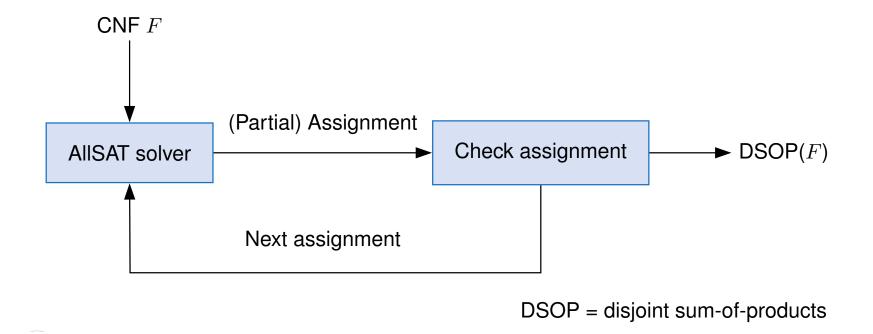
From DSOP to d-DNNF

$$F = (\neg a \lor b \lor c) \land (\neg a \lor d) \land (a \lor b)$$

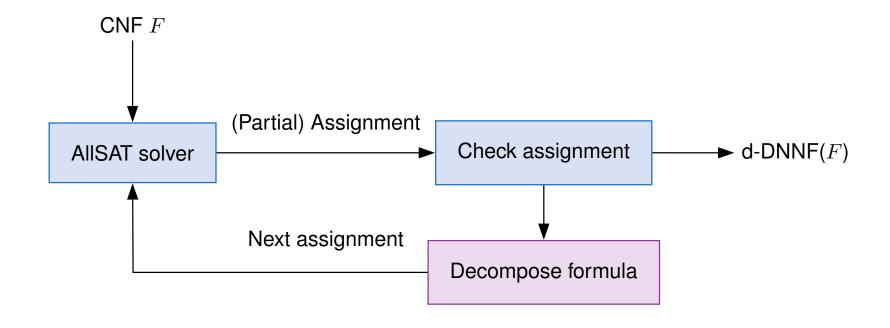
$$\mathsf{DSOP}(F) = (\mathbf{a} \wedge b \wedge \mathbf{d}) \vee (\mathbf{a} \wedge \neg b \wedge c \wedge \mathbf{d}) \vee (\neg a \wedge b)$$

$$d-\mathsf{DNNF}(F) = (a \land (b \lor (\neg b \land c)) \land d) \lor (\neg a \land b)$$

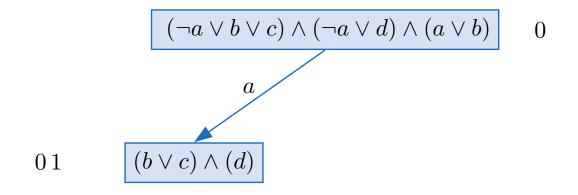
CNF-to-d-DNNF Compilation: Main Idea

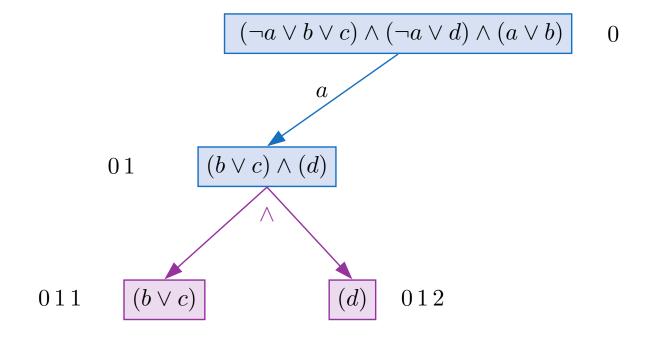


CNF-to-d-DNNF Compilation: Main Idea

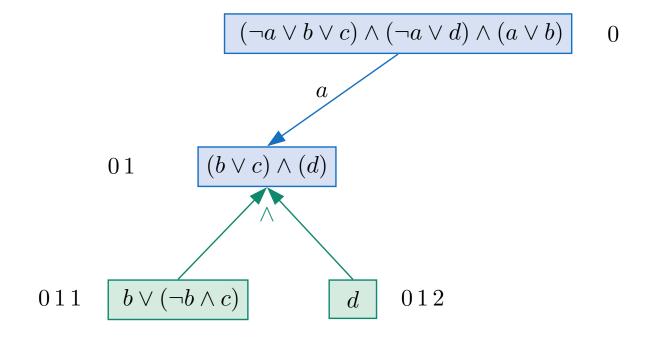


$$(\neg a \lor b \lor c) \land (\neg a \lor d) \land (a \lor b)$$

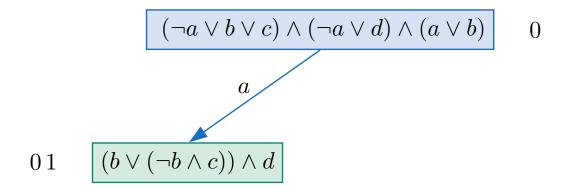


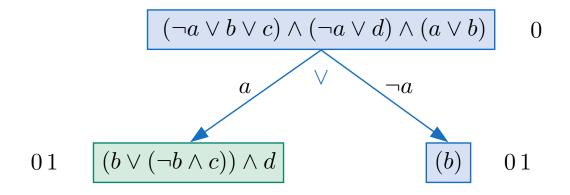


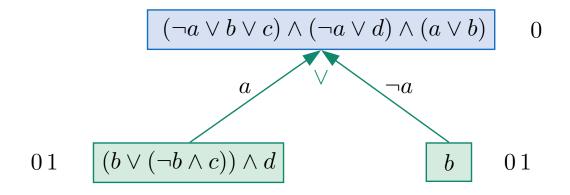
26



26







$$(a \wedge (b \vee (\neg b \wedge c)) \wedge d) \vee (\neg a \wedge b) \qquad 0$$

au	· · · 4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I	4	5	30	47	15	18	6	-7	-8	45	9	38	-23	17	44	-16
δ	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	6

au	· · · 4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I	4	5	30	47	15	18	6	-7	-8	45	9	38	-23	17	44	-16
δ	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	6

decision literal

au	· · · 4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I	4	5	30	47	15	18	6	-7	-8	45	9	38	-23	17	44	-16
δ	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	6

 $\mathsf{block}(I,4)$

au	· · · 4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I	4	5	30	47	15	18	6	-7	-8	45	9	38	-23	17	44	-16
δ	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	6

slice(I,4)

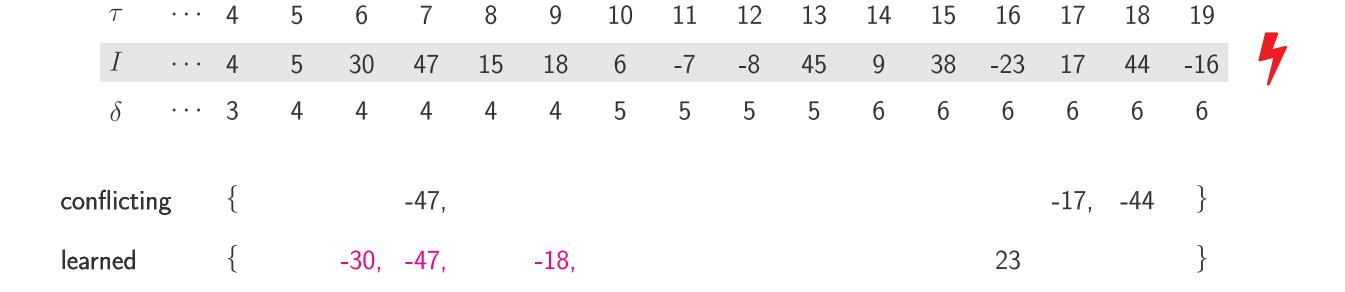
au	· · · 4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
I	· · · 4	5	30	47	15	18	6	-7	-8	45	9	38	-23	17	44	-16
δ	3	4	4	4	4	4	5	5	5	5	6	6	6	6	6	6

 $I_{\leqslant 4}$

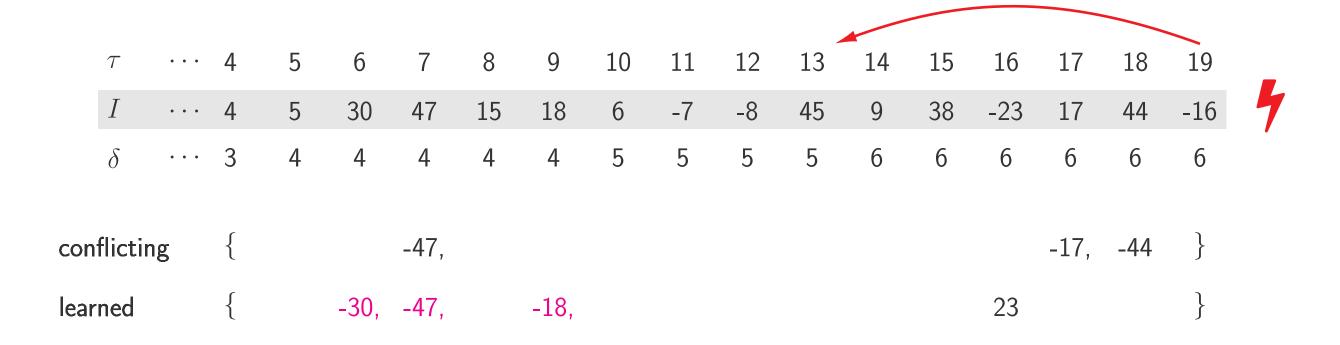
	• • •																	
I	• • •	4	5	30	47	15	18	6	-7	-8	45	9	38	-23	17	44	-16	7











backtrack level 5

au	• • •	4	5	6	7	8	9	10	11	12	13
I		4	5	30	47	15	18	6	-7	-8	45
δ		3	4	4	4	4	4	5	5	5	5

au	• • •	4	5	6	7	8	9	10	11	12	13
I	• • •	4	5	30	47	15	18	6	-7	-8	45
δ		3	4	4	4	4	4	5	5	5	5

$$au$$
 ... 4 5 6 7 8 9 10 11 12 13 14 I ... 4 5 30 47 15 18 6 -7 -8 45 23 δ ... 3 4 4 4 4 4 5 5 5 5 5 4

out of order

au	• • •	4	5	6	7	8	9	10	11	12	13
I		4	5	30	47	15	18	6	-7	-8	45
δ		3	4	4	4	4	4	5	5	5	5

$$au$$
 4 5 6 7 8 9 10 11 12 13 14 I 4 **5 30 47 15 18** 6 -7 -8 45 23 δ 3 4 4 4 4 4 5 5 5 5 5 4

 $\mathsf{block}(I,4)$

au	• • •	4	5	6	7	8	9	10	11	12	13
I		4	5	30	47	15	18	6	-7	-8	45
δ		3	4	4	4	4	4	5	5	5	5

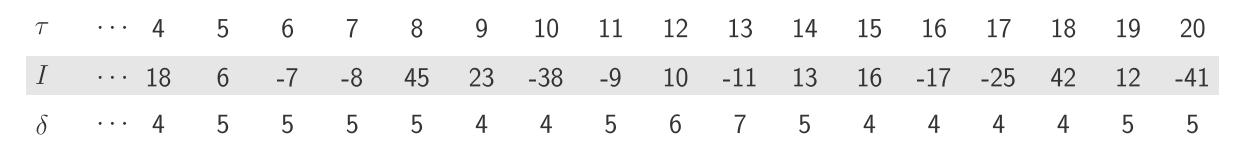
$$au$$
 ... 4 5 6 7 8 9 10 11 12 13 14 I ... 4 5 30 47 15 18 6 -7 -8 45 23 δ ... 3 4 4 4 4 4 5 5 5 5 5 4

slice(I,4)

au	• • •	4	5	6	7	8	9	10	11	12	13
I	• • •	4	5	30	47	15	18	6	-7	-8	45
δ		3	4	4	4	4	4	5	5	5	5

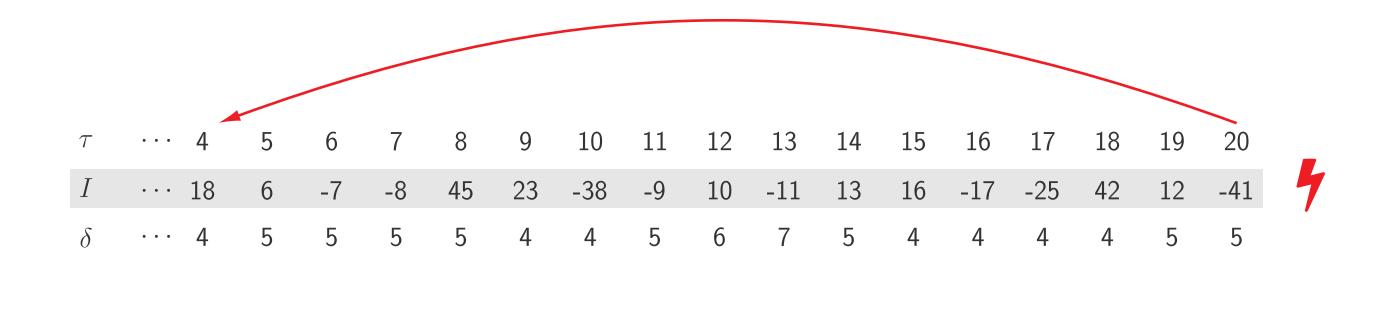
$$au$$
 4 5 6 7 8 9 10 11 12 13 14 I 4 5 30 47 15 18 6 -7 -8 45 23 δ 3 4 4 4 4 4 5 5 5 5 5 4

au	· · · 4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
I	· · · 18	6	-7	-8	45	23	-38	-9	10	-11	13	16	-17	-25	42	12	-41
δ	· · · 4	5	5	5	5	4	4	5	6	7	5	4	4	4	4	5	5





conflicting { 17, -42, -12 }



conflicting { 17, -42, -12 }

backtrack level 4

au	· · · 4	5	6	7	8	9	10	11
I	· · · 18	23	-38	16	-17	-25	42	-12
	· · · 4							