Resolution-Based Certificate Extraction for QBF (Tool Presentation)

Aina Niemetz, Mathias Preiner,
Florian Lonsing, Martina Seidl, and Armin Biere

Institute for Formal Models and Verification (FMV)
Johannes Kepler University, Linz, Austria
http://fmv.jku.at/

SAT'12
June 17 - 20, 2012
Trento, Italy

Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$

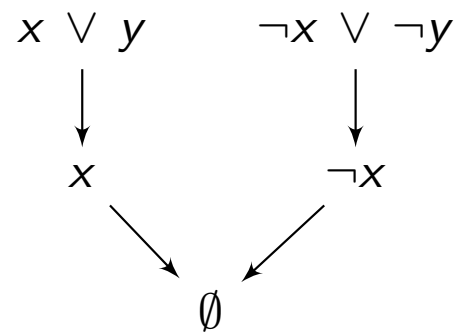
Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$

Truth Table

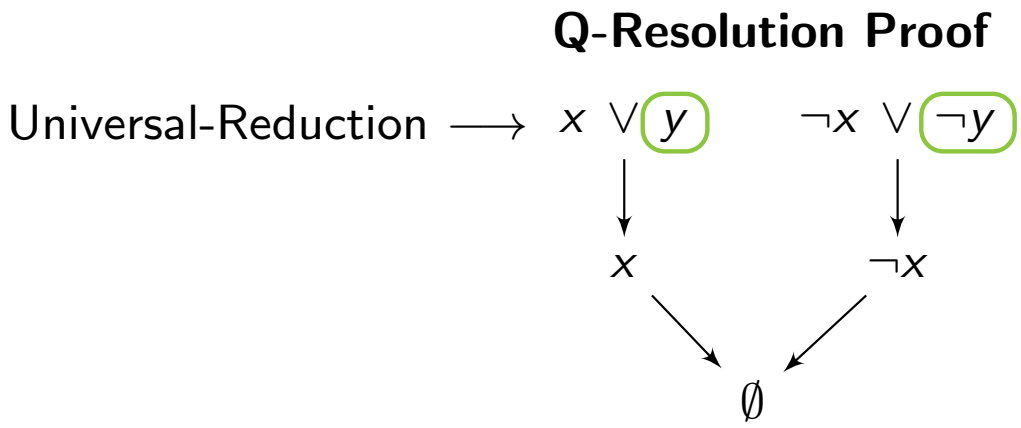
x	y	ψ
0	0	0
0	1	1
1	0	1
1	1	0

Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$

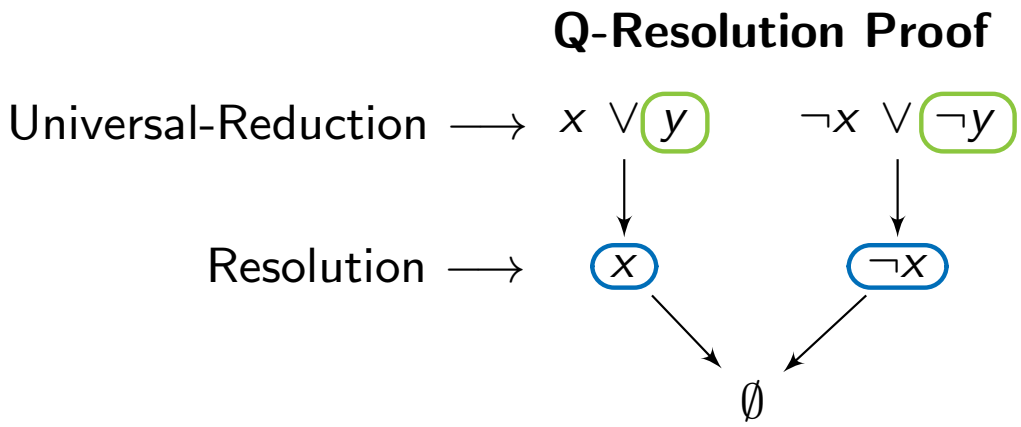
Q-Resolution Proof



Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$



Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$



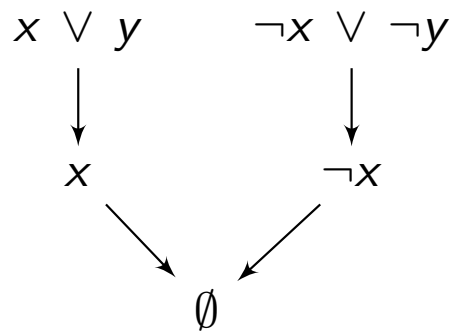
Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$

Truth Table

x	y	ψ
0	0	0
0	1	1
1	0	1
1	1	0

unsat

Q-Resolution Proof



$$\longrightarrow \quad y = x \quad \Rightarrow \quad \psi = 0$$

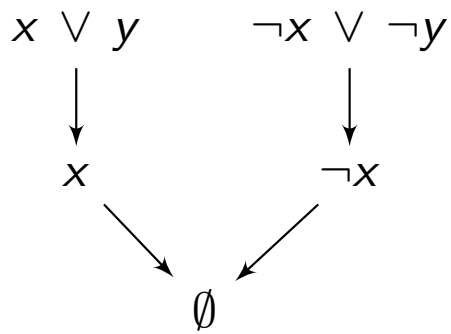
Exclusive OR (XOR): QBF $\psi = \exists x \forall y.(x \vee y) \wedge (\neg x \vee \neg y)$

Truth Table

x	y	ψ
0	0	0
0	1	1
1	0	1
1	1	0

unsat

Q-Resolution Proof



$\longrightarrow \quad y = x \quad \Rightarrow \quad \psi = 0$

$\longrightarrow \quad f_y(x) = x \quad$ (counter model)

Our Goal

- verify correctness of a QBF solver's result
- concrete solutions (certificates), e.g. counter examples, strategies
  $\longrightarrow$ Skolem/Herbrand function-based certificates

Our Goal

- verify correctness of a QBF solver's result
- concrete solutions (certificates), e.g. counter examples, strategies
  $\longrightarrow$ Skolem/Herbrand function-based certificates

QBF Certificates

- as set of Skolem/Herbrand functions (e.g. $f_y(x) = x$ in prev. example)
- representation of model/counter model
- novel approach presented at CAV'11 [BJ11] for true and false QBF
  $\longrightarrow$ extraction of Skolem/Herbrand functions from Q-resolution proofs

Our Goal

- verify correctness of a QBF solver's result
- concrete solutions (certificates), e.g. counter examples, strategies
  $\longrightarrow$ Skolem/Herbrand function-based certificates

QBF Certificates

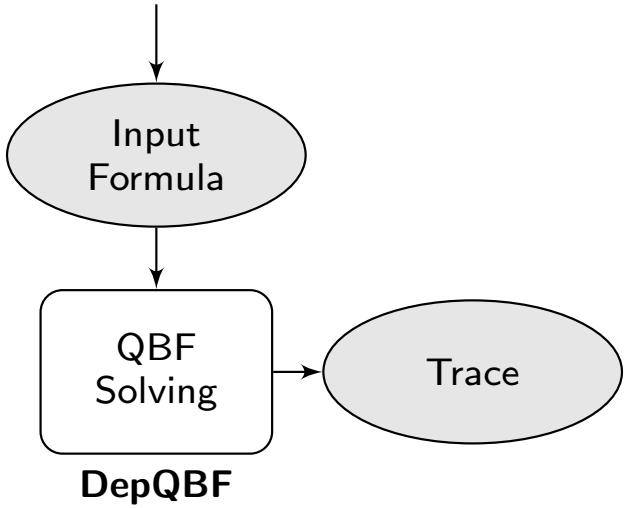
- as set of Skolem/Herbrand functions (e.g. $f_y(x) = x$ in prev. example)
- representation of model/counter model
- novel approach presented at CAV'11 [BJ11] for true and false QBF
  $\longrightarrow$ extraction of Skolem/Herbrand functions from Q-resolution proofs

Our Work

- solver-independent framework for
- resolution-based certificate extraction and validation
- for true and false QBF

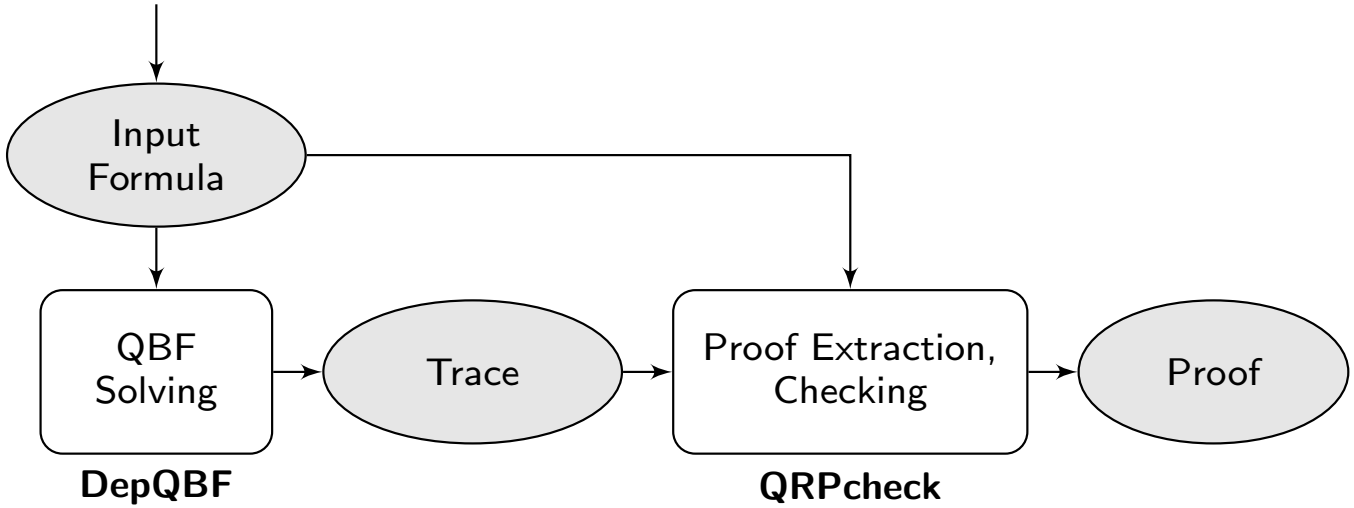
Certificaton Workflow

Overview



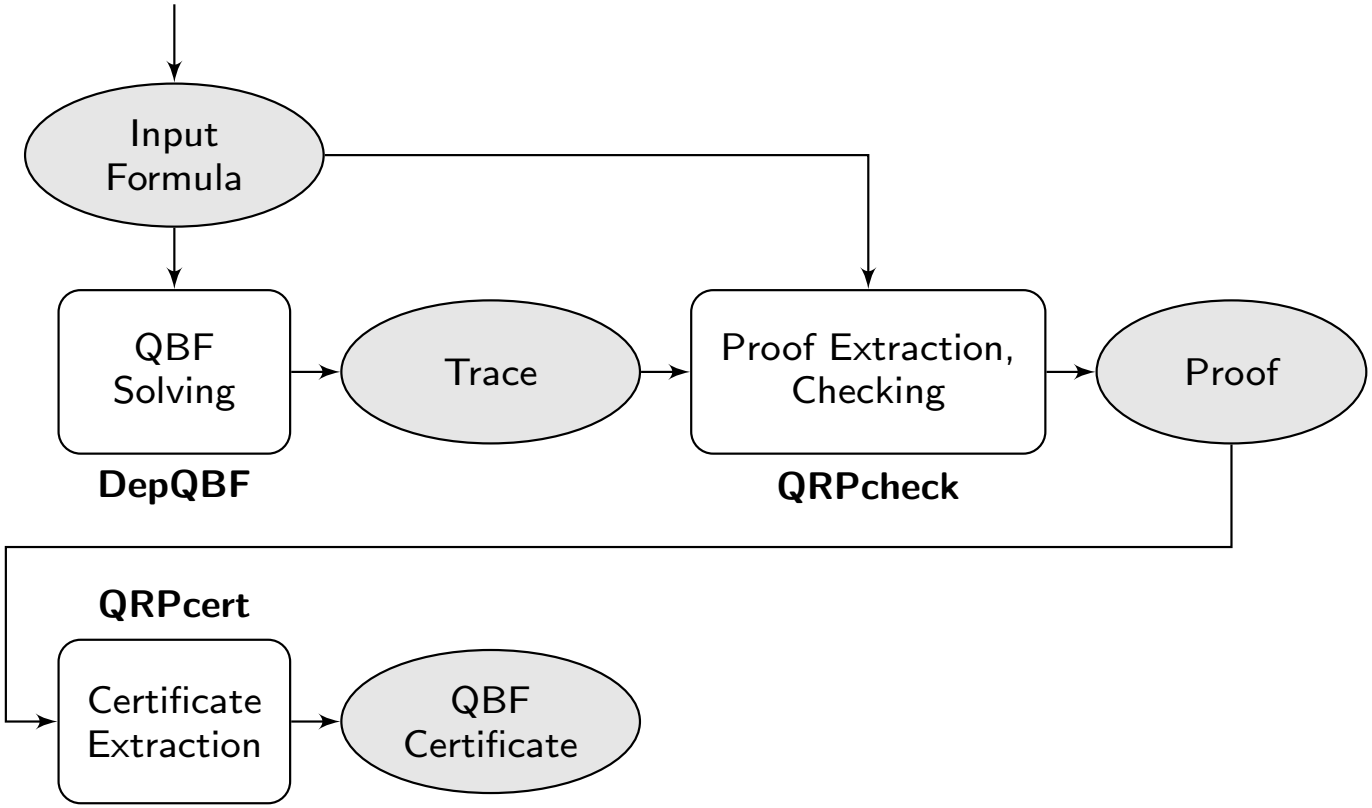
Certificaton Workflow

Overview



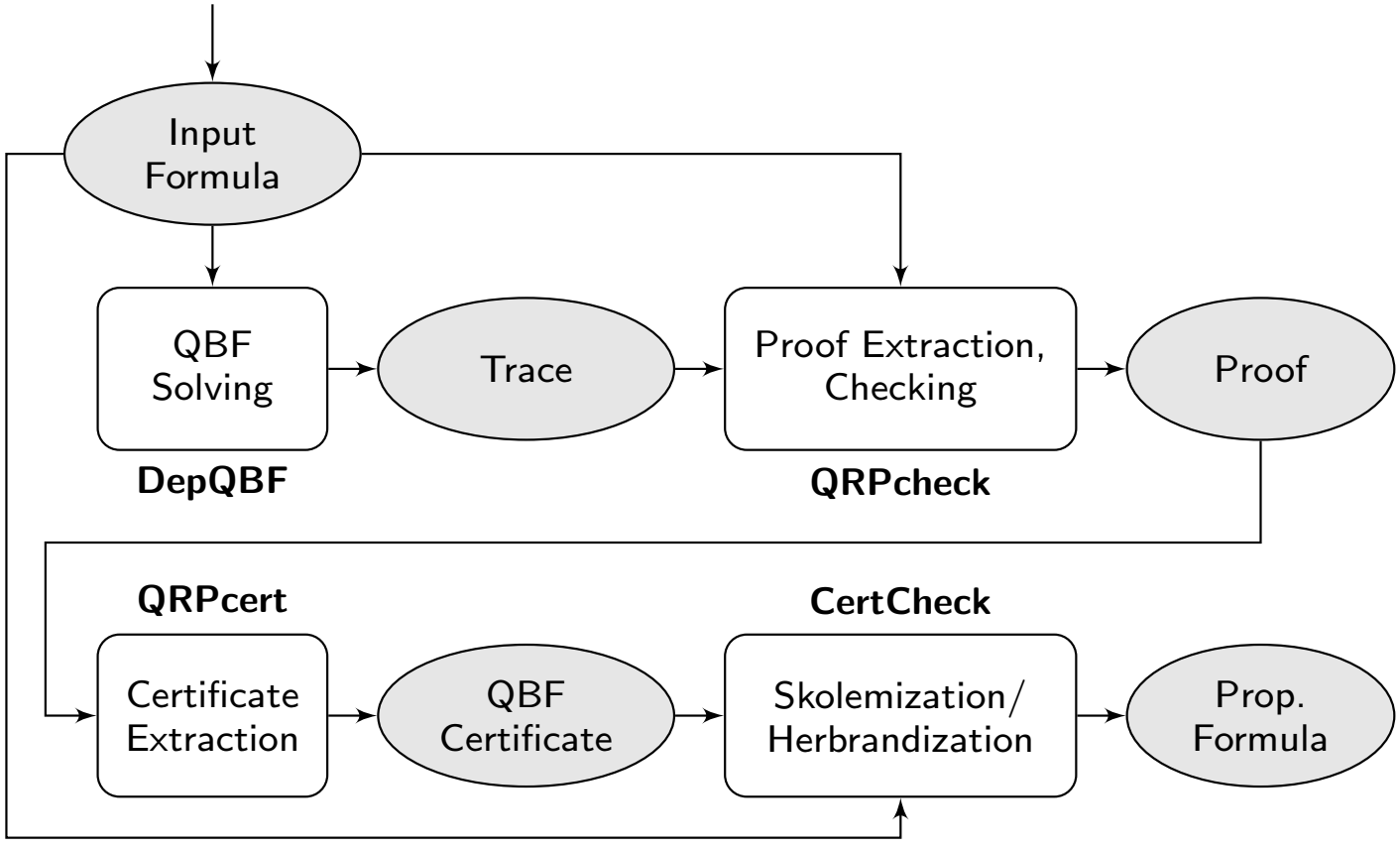
Certificaton Workflow

Overview



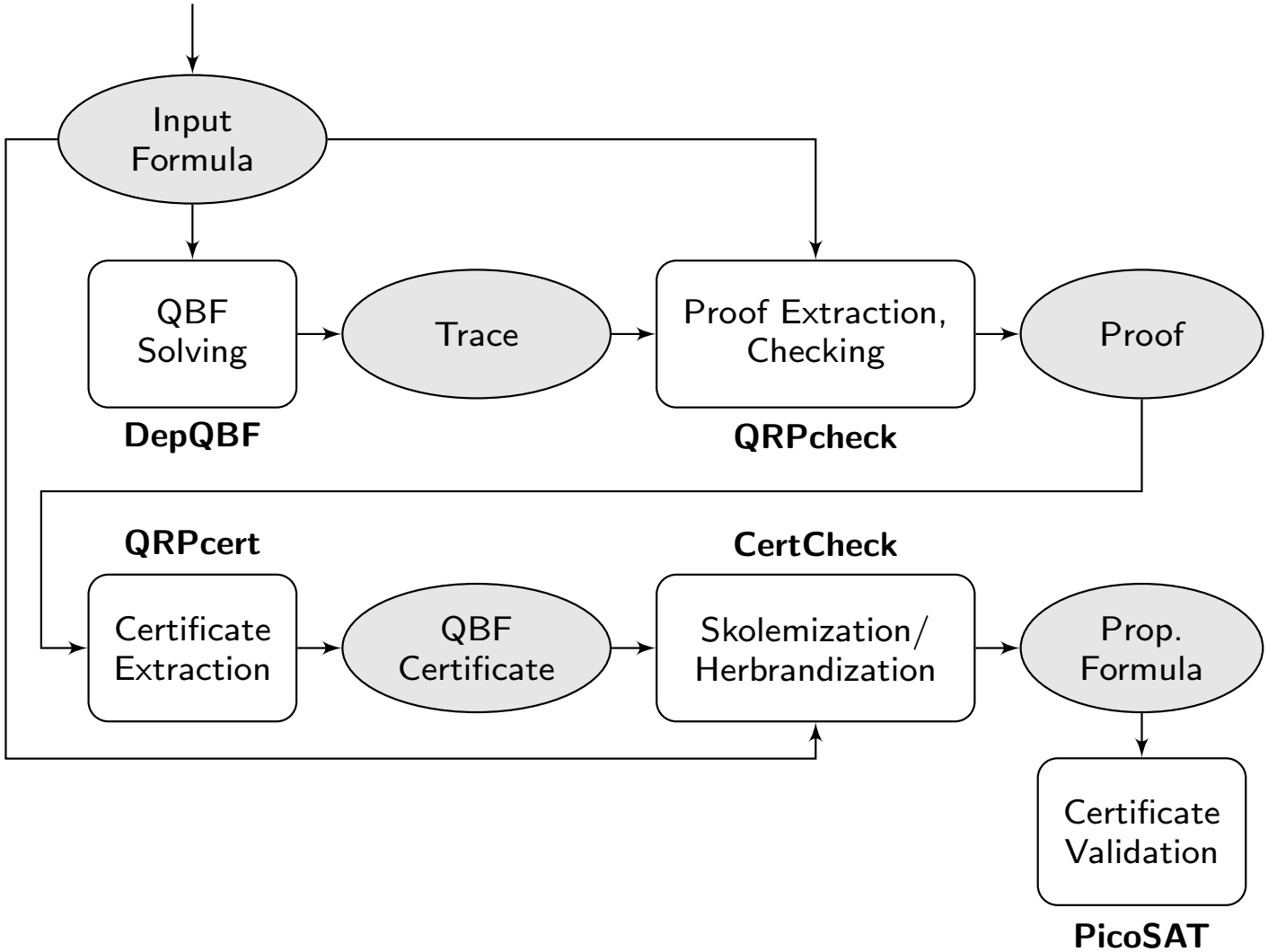
Certificaton Workflow

Overview



Certificaton Workflow

Overview



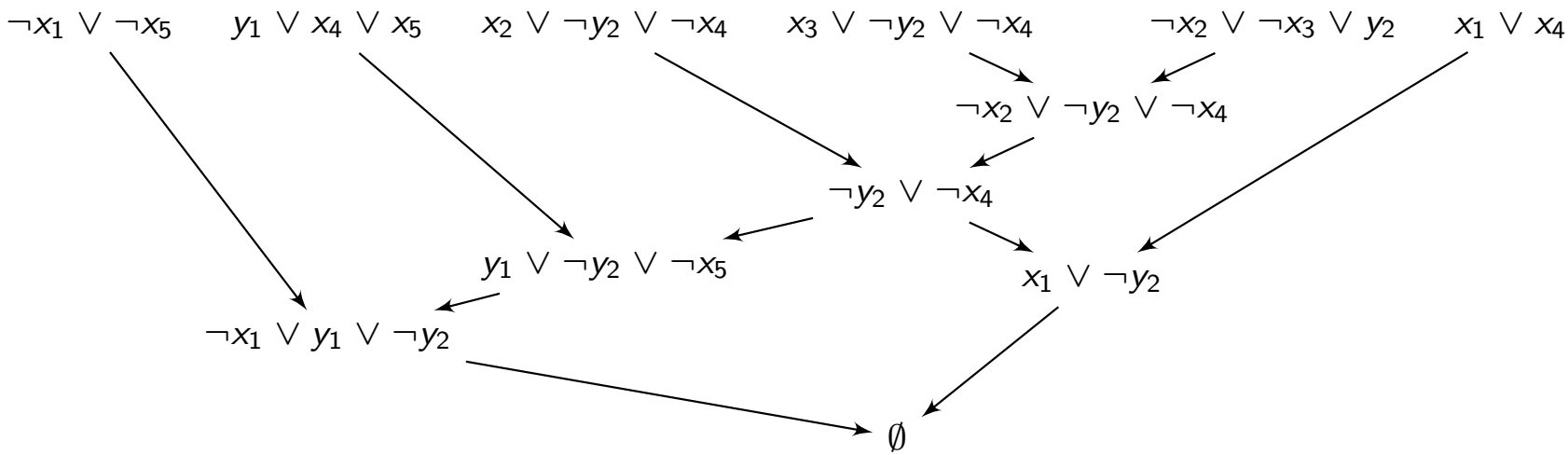
Input Formula

$$\exists x_1 \forall y_1 \exists x_2 x_3 \forall y_2 \exists x_4 x_5 . (\neg x_1 \vee \neg x_5) \wedge (y_1 \vee x_4 \vee x_5) \wedge (x_2 \vee \neg y_2 \vee \neg x_4) \wedge (x_3 \vee \neg y_2 \vee \neg x_4) \wedge (\neg x_2 \vee \neg x_3 \vee y_2) \wedge (x_1 \vee x_4)$$

Input Formula

$$\exists x_1 \forall y_1 \exists x_2 x_3 \forall y_2 \exists x_4 x_5 . (\neg x_1 \vee \neg x_5) \wedge (y_1 \vee x_4 \vee x_5) \wedge (x_2 \vee \neg y_2 \vee \neg x_4) \wedge (x_3 \vee \neg y_2 \vee \neg x_4) \wedge (\neg x_2 \vee \neg x_3 \vee y_2) \wedge (x_1 \vee x_4)$$

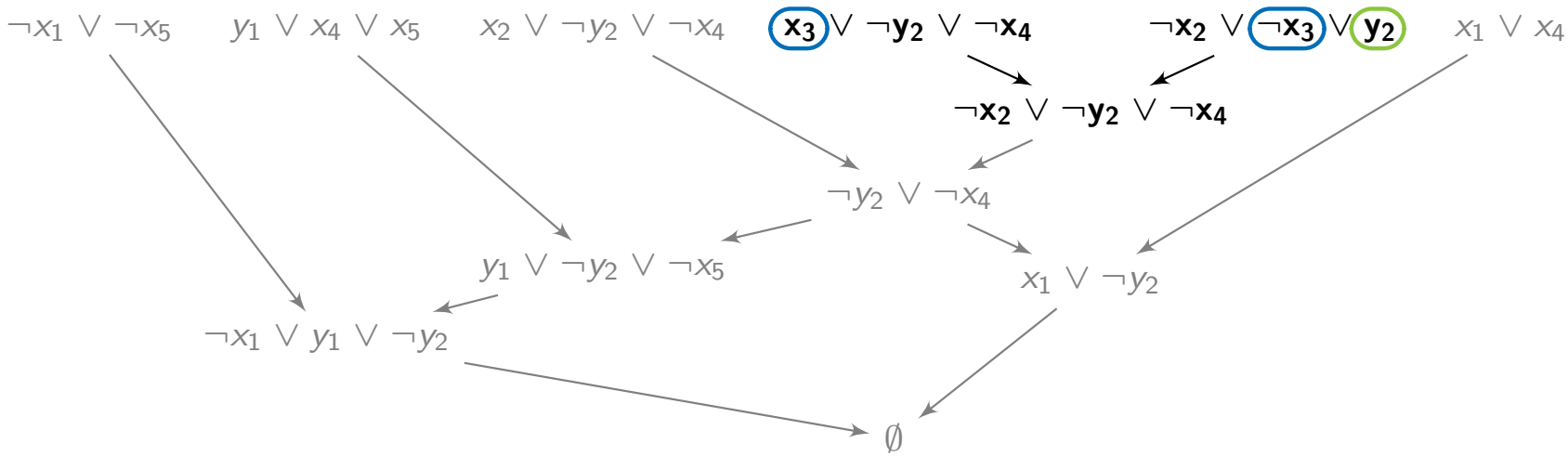
Q-Resolution Proof DAG



Input Formula

$$\exists x_1 \forall y_1 \exists x_2 x_3 \forall y_2 \exists x_4 x_5 . (\neg x_1 \vee \neg x_5) \wedge (y_1 \vee x_4 \vee x_5) \wedge (x_2 \vee \neg y_2 \vee \neg x_4) \wedge (x_3 \vee \neg y_2 \vee \neg x_4) \wedge (\neg x_2 \vee \neg x_3 \vee y_2) \wedge (x_1 \vee x_4)$$

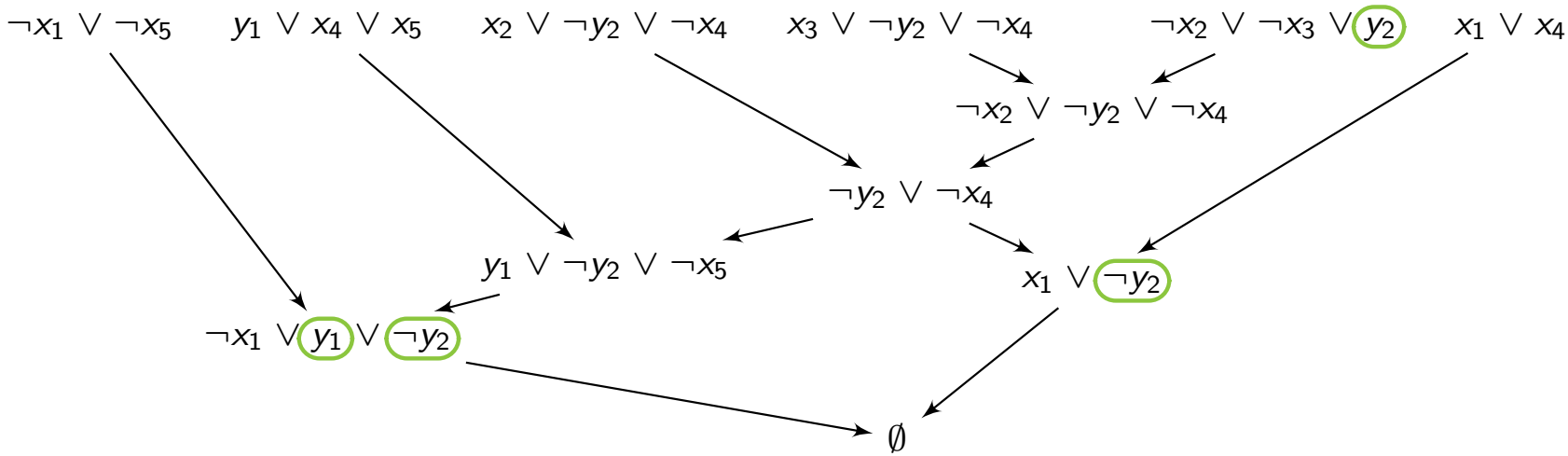
Q-Resolution Proof DAG



Input Formula

$$\exists x_1 \forall y_1 \exists x_2 x_3 \forall y_2 \exists x_4 x_5 . (\neg x_1 \vee \neg x_5) \wedge (y_1 \vee x_4 \vee x_5) \wedge (x_2 \vee \neg y_2 \vee \neg x_4) \wedge (x_3 \vee \neg y_2 \vee \neg x_4) \wedge (\neg x_2 \vee \neg x_3 \vee y_2) \wedge (x_1 \vee x_4)$$

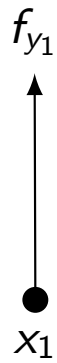
Q-Resolution Proof DAG

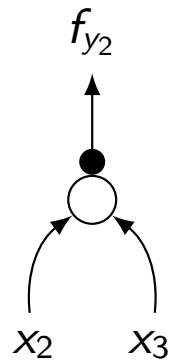


Extracted Herbrand Functions

$$\left.\begin{aligned} f_{y_1}(x_1) &= \neg x_1 \\ f_{y_2}(x_1, x_3) &= \neg x_2 \vee \neg x_3 \end{aligned}\right\} \text{Certificate}$$

Extracted Certificate: AIG Representation



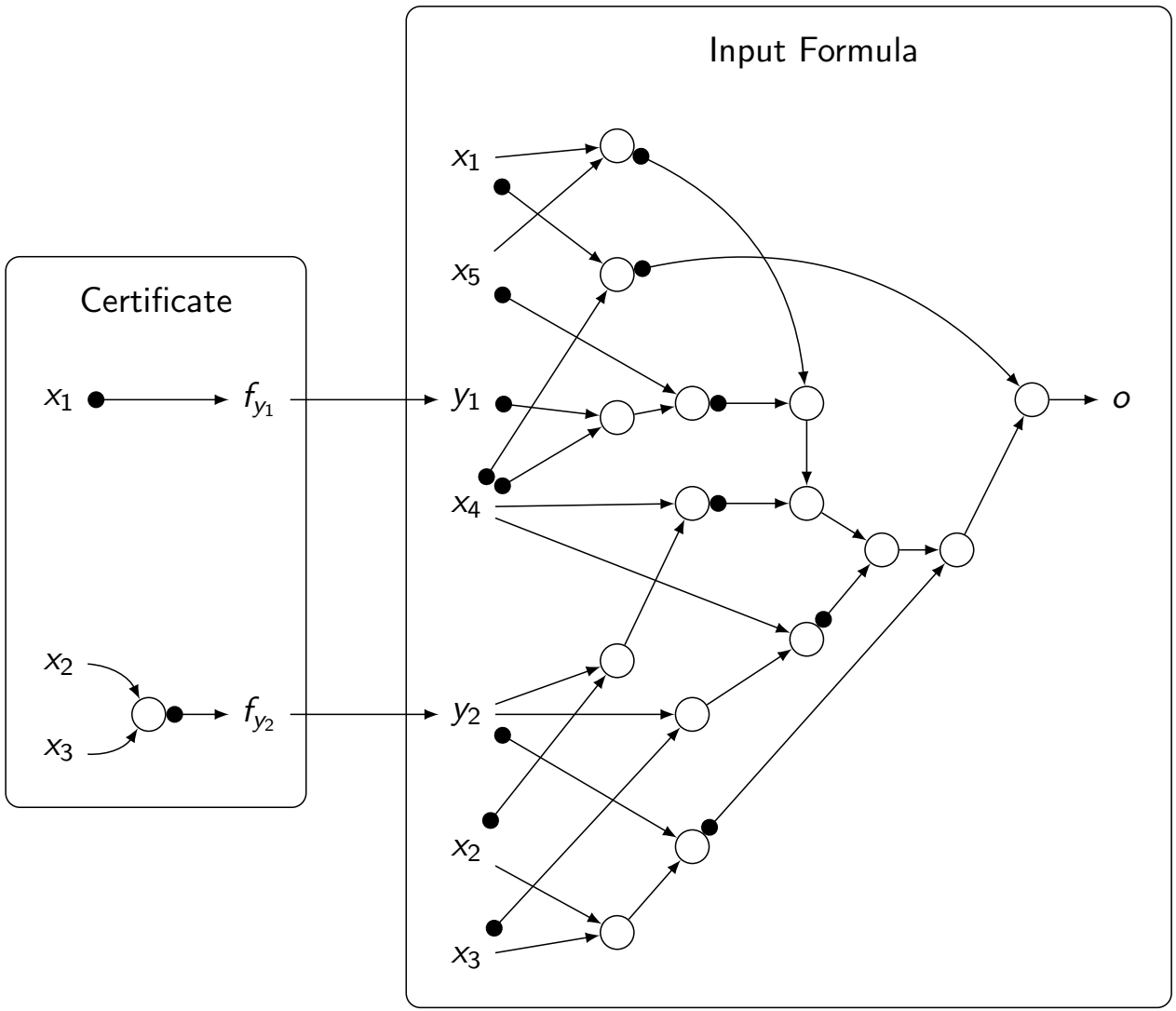


$$f_{y_1}(x_1) = \neg x_1$$

$$\begin{aligned} f_{y_2}(x_2, x_3) &= \neg x_2 \vee \neg x_3 \\ &= \neg(x_2 \wedge x_3) \end{aligned}$$

Certification by Example

Herbrandization



Experimental Results

Benchmarks: QBFEVAL'10 set (568 formulas)
Limits: 1800 seconds and 7 GB limits

1. **Proof Extraction, Checking**
   - out of 362 solved instances, 348 proofs extracted and checked by QRPcheck
   - 14 instances lost due to memory out

Experimental Results

Benchmarks: QBFEVAL'10 set (568 formulas)
Limits: 1800 seconds and 7 GB limits

1. **Proof Extraction, Checking**
   - out of 362 solved instances, 348 proofs extracted and checked by QRPcheck
   - 14 instances lost due to memory out
2. **Certificate Extraction**
   - out of 348 proofs, 337 certificates extracted
   - 11 instances lost due to memory out
   - AND-Gates: max. 147 Mill., avg. 8 Mill., med. 369

Experimental Results

Benchmarks: QBFEVAL'10 set (568 formulas)
Limits: 1800 seconds and 7 GB limits

1. **Proof Extraction, Checking**
   - out of 362 solved instances, 348 proofs extracted and checked by QRPcheck
   - 14 instances lost due to memory out
2. **Certificate Extraction**
   - out of 348 proofs, 337 certificates extracted
   - 11 instances lost due to memory out
   - AND-Gates: max. 147 Mill., avg. 8 Mill., med. 369
3. **Skolemization/Herbrandization**
   - out of 337 certificates, 337 formulas skolemized/herbrandized
   - Clauses: max. 441 Mill., avg. 25 Mill., med. 71000

Experimental Results

Benchmarks: QBFEVAL'10 set (568 formulas)
Limits: 1800 seconds and 7 GB limits

1. **Proof Extraction, Checking**
   - out of 362 solved instances, 348 proofs extracted and checked by QRPcheck
   - 14 instances lost due to memory out
2. **Certificate Extraction**
   - out of 348 proofs, 337 certificates extracted
   - 11 instances lost due to memory out
   - AND-Gates: max. 147 Mill., avg. 8 Mill., med. 369
3. **Skolemization/Herbrandization**
   - out of 337 certificates, 337 formulas skolemized/herbrandized
   - Clauses: max. 441 Mill., avg. 25 Mill., med. 71000
4. **Certificate Validation**
   - out of 337 skolemized/herbrandized formulas, 275 checked successfully
   - 45 (17) certificates not validated due to memory (time) out
     $\rightarrow$ out of these 62, 57 instances were satisfiable
   - $>$ 70% of the total runtime

Conclusion

Summary

- complete and solver-independent framework
- certification and validation of true and false QBF
- certificates for over 90% of solved instances extracted
  $\rightarrow$ 100% if memory limit is lifted
- over 80% of all extracted certificates validated
- certificate validation is still challenging

Future Work

- optimize certificate validation process
  $\rightarrow$ employ incremental SAT-checking
- support for advanced dependency schemes
  (key feature of DepQBF)

References


Valeriy Balabanov and Jie-Hong R. Jiang.
Resolution Proofs and Skolem Functions in QBF Evaluation and Applications.
In *Proc. of the 23rd International Conference on Computer Aided Verification (CAV 2011)*, volume 6806 of *Lecture Notes in Computer Science*, pages 149–164. Springer, 2011.

Appendix

QBFEVAL'10

Runtime Overview

	Instances				Total Time [s]			
	sv	ch	ex	va	DepQBF	QRPcheck	QRPcert	PicoSAT
sat	157	153	143	86	701.8	80.1	30.9	3247.0
unsat	205	195	194	189	4241.9	1011.5	86.8	1090.0
total	**362**	**348**	**337**	**275**	**4943.7**	**1091.7**	**117.6**	**4337.0**

Comparison of Proof, Certificate, Prop. Formula Sizes

	Proof				Certificate		Prop. Formula			
	vertices		literals		AND-Gates		variables		clauses	
	avg	med	avg	med	avg	med	avg	med	avg	med
sat	308k	1k	117M	626k	20M	24k	20M	62k	59M	183k
unsat	135k	2k	14M	146k	170k	193	336k	23k	846k	55k
total	**211k**	**2k**	**60M**	**175k**	**8M**	**369**	**8M**	**28k**	**25M**	**71k**

Certificate Statistics

	In	Out	AND-Gates			AND-Gates (shared) [%]		
	avg.	avg.	max.	avg.	med.	max.	avg.	med.
sat	125	3k	147M	20M	24k	98.1	65.2	66.8
unsat	20k	95	10M	170k	193	49.5	23.0	23.7
total	**12k**	**1k**	**147M**	**8M**	**369**	**98.1**	**40.9**	**46.6**