Soonho Kong

Programming Research Laboratory Seoul National University

March 21, 2008

¹Work done with Dr. Ting Zhang in Microsoft Research Asia < ≥ > < ≥ >

- **QBF** Problem
- Motivation

- 3 Cost Minimizing Search
- Performance
- **Future Work**
- References

Quantified Boolean Formulae

- Quantifiers: ∀,∃
- All variables are Boolean variables.
- Logical Operator: ∧, ∨, ¬



Example

$$\psi_1 = \forall x \exists y (x \vee \neg y)$$

True



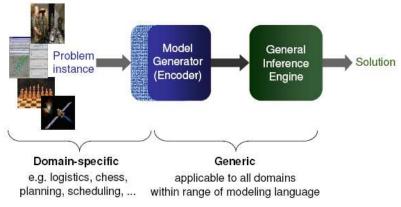
Example

$$\psi_2 = \forall x \exists y (x \land \neg y)$$
False

- The QBF problem is to determine its truth value.
- In complexity theory, it is PSPACE problem.
- At least as hard as NP.



Why is QBF solving important?



[Credit: A. Sabharwal, B. Selman, Beyond Traditional SAT Reasoning, AAAI 2007 Tutorial]



QBF compactly encode many problems in CSE.

- Automated Plannig
- Scheduling
- Bounded Model Checking(BMC)
- Electronic Design Automation

Improvement in solving engine affects many application areas.



- QBF provides a compacter encoding than SAT.
- Transforming QBF into equivalent SAT increases space exponentially.
 - $\blacksquare \mathsf{QBF} : \forall x \exists y \exists z (x \vee \neg y \vee z)$
 - SAT : $(\exists y \exists z (\top \lor \neg y \lor z)) \land (\exists y \exists z (\bot \lor \neg y \lor z))$

Existing Approach

- Search-based
- Q-Resolution
- Symbolic Skolemization



Search-based QBF solving is dominant.

- Search-based technique is dominant in SAT solving.
- People tried to apply and extend those techniques to QBF solving problem.



Search-based QBF solvers require normalized input.

- Prenex Form
- Conjunctive Normal Form(CNF)



Future Work

Normalized Input - Prenex Form

Prenex Form

- \blacksquare (X): $\forall x(\exists y(x \lor y) \land (\exists z(x \lor z)))$
- \blacksquare (O) : $\forall x \exists y \exists z ((x \lor y) \land (x \lor z))$



Normalized Input - CNF

Conjunctive Normal Form(CNF)

- Literal : Variable or Negation of Variable x, ¬x
- Clause : Disjucntion of Literals $(x \lor \neg y \lor z)$
- CNF : Conjunction of Clauses $(x \lor \neg y \lor z) \land (\neg x \lor y \lor \neg z)$

Future Work

Why should be "prenex and CNF"?



Why "prenex and CNF"?

Binary Constraint Propagation(BCP)

BCP propagate one assignment and deduce another another assignment without searching the space.



BCP is a "Founding block" for

- non-chronological backtracking
- conflict-driven learning



Binary Constraint Propagation(BCP)

$$\forall x \exists y \exists z ((x \vee \neg y \vee z) \wedge (\neg x \vee \neg y \vee \neg z))$$

Binary Constraint Propagation(BCP) Example

$$\forall x \exists y \exists z ((x \vee \neg y \vee z) \wedge (\neg x \vee \neg y \vee \neg z))$$

$$\mathsf{Try} \ x = \top.$$



Future Work

Outline

Binary Constraint Propagation(BCP) Example

$$\exists y \exists z ((\top \vee \neg y \vee z) \wedge (\bot \vee \neg y \vee \neg z))$$



Future Work

$$\exists y \exists z (\neg y \lor \neg z)$$



Binary Constraint Propagation(BCP) Example

$$\exists y \exists z (\neg y \lor \neg z)$$

Try
$$y = \top$$



Future Work

Binary Constraint Propagation(BCP)

$$\exists z(\bot \lor \neg z)$$



Binary Constraint Propagation(BCP) Example

Binary Constraint Propagation(BCP)

$$\exists z(\neg z)$$

Now, z must be \perp and it makes the formula satisfied.



Binary Constraint Propagation(BCP) Example

Binary Constraint Propagation(BCP)

$$\forall x \exists y \exists z ((x \vee \neg y \vee z) \wedge (\neg x \vee \neg y \vee \neg z))$$

Go back to try $x = \bot$ because of $\forall x$.



Binary Constraint Propagation(BCP)

$$\forall x \exists y \exists z ((\bot \lor \neg y \lor z) \land (\top \lor \neg y \lor \neg z))$$

Binary Constraint Propagation(BCP) Example

$$\exists y \exists z (\neg y \lor z)$$

Try
$$y = \top$$
.



Binary Constraint Propagation(BCP)

$$\exists z(\bot \lor z)$$



Binary Constraint Propagation(BCP) Example

Binary Constraint Propagation(BCP)

$$\exists z(\lor z)$$

Now we know that z should be \top and it makes the formula satisfied.



Binary Constraint Propagation(BCP) Example

Binary Constraint Propagation(BCP)

$$\forall x \exists y \exists z ((\bot \lor \neg y \lor z) \land (\top \lor \neg y \lor \neg z))$$

Now we know that the formula is true.



Problem

Outline

It is necessary to have prenex and CNF to conduct BCP.



Problem

Outline

However,

QBF encodings from the real world applications never come in prenex and CNF.



References

Problem

Outline

Prenex converting and CNF converting are not free.



References

Problem - CNF

Converting into equivalent CNF takes exponential time.



Tseytin introduced linear-time CNF converting method.

However,

it introduces new existentially quantified variables.



Our experiment shows total number of variables increase by four times.



Converting into prenex form enlarges the search space by

- widening quantifier scope
- introducing dependency to the variables which were independent



Problem - prenex form

$$\forall x \psi_1(x) \land \exists y \psi_2(y) \land \forall z \psi_3(z) \land \exists w \psi_4(w)$$

is converted into

$$\forall x \exists y \forall z \exists w (\psi_1(x) \land \psi_2(y) \land \psi_3(z) \land \psi_4(w))$$



Outline

Is there another way to solve QBF without converting input into prenex and CNF?



Outline

Is there another direction in QBF solving which is not based on boolean constraint propagation?



Outline

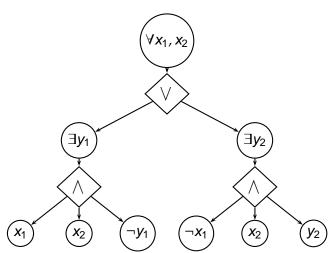
Yes, there is! Naive algorithm solves QBF without BCP.



Motivation - Naive Algorithm

Example

Outline





Motivation - Naive Algorithm

But its performance is poor. 2 - 4 orders of magnitude slower than the state-of-the-art solvers.



Outline

Is there an "efficient" search-based method to solve QBF without conducting boolean constraint propagation?



Cost Minimizing Search Based on the same naive algorithm shown before.



Preprocessing

Outline

Push Quantifiers as Further as Possible to Reduce Search Space

$$\forall x(\varphi \land \psi) \equiv (\forall x\varphi) \land \psi, \qquad \forall x(\varphi \lor \psi) \equiv (\forall x\varphi) \lor \psi,$$
$$\exists x(\varphi \land \psi) \equiv (\exists x\varphi) \land \psi, \qquad \exists x(\varphi \lor \psi) \equiv (\exists x\varphi) \lor \psi.$$

 $\varphi, \psi \in \mathsf{QBF}$ and x a variable not occurring free in ψ . For example

$$\forall x \exists y \exists z ((x \lor y) \land (x \lor z)) \rightarrow \forall x (\exists y (x \lor y) \land (\exists z (x \lor z)))$$



ROPAS

Observation

Outline

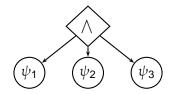
We can reduce the time by exploiting "short-circuit" property of \land and \lor operator.



Observation

Outline

Once we found a desired result from the subnode. We do not have to continue the evaluation.



If ψ_1 is false, then whole formula turns out to be false.

Observation

Outline

Since operators \land and \lor are *commutative*, we can *sort* the operands and exploit short-circuit property.



Considerations

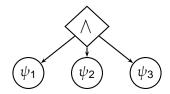
Outline

Consider two things:

- Truth Probability
- 2. Expected Search Space

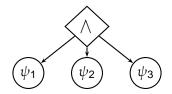


References



$$P(\psi_1) = 0.8, P(\psi_2) = 0.6, P(\psi_3) = 0.3$$

It is better to sort ψ_i to ψ_3, ψ_2, ψ_1 and try ψ_3 first.



$$S(\psi_1) = 300, S(\psi_2) = 60, S(\psi_3) = 90$$

It is better to sort ψ_i to ψ_2, ψ_3, ψ_1 and try ψ_2 first.

Cost

Outline

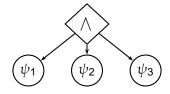
In general, truth probability and expected search space vary at the same time. We need a *single measure* to sort the subnode ψ_i s. We introduce the concept of *cost*.

Cost = Probability of Being Useless × Expected Search Space



Cost Minimizing Search

Conduct depth-first cost minimizing search



$$P(\psi_1) = 0.8$$
 $P(\psi_2) = 0.6$ $P(\psi_3) = 0.3$ $S(\psi_1) = 300$ $S(\psi_2) = 60$ $S(\psi_3) = 90$ $C(\psi_1) = 24$ $C(\psi_2) = 36$ $C(\psi_3) = 27$ It is better to sort ψ_i to ψ_1, ψ_3, ψ_2 and try ψ_1 first.



Probability Calculation

1
$$\varphi \equiv x \mid \neg x$$

$$P(x) = P(\neg x) = \frac{1}{2}$$

$$\varphi \equiv \bigwedge_{i=1}^n \varphi_i$$

$$P\left(\bigwedge_{i=1}^n \varphi_i\right) = \prod_{i=1}^n P(\varphi_i)$$

$$\mathbf{3} \ \varphi \equiv \bigvee_{i=1}^{n} \varphi_{i}$$

$$\mathsf{P}\bigg(\bigvee_{i=1}^n \varphi_i\bigg) = 1 - \prod_{i=1}^n \left(1 - \mathsf{P}(\varphi_i)\right)$$



Probability Calculation

$$\mathbf{4} \ \varphi \equiv \forall \bar{\mathbf{x}} \psi$$

$$\mathsf{P}(\forall \bar{\mathsf{x}}\psi) = \prod_{i=1}^{2^{|\bar{\mathsf{x}}|}} \mathsf{P}(\psi) = (\mathsf{P}(\psi))^{(2^{|\bar{\mathsf{x}}|})}$$

$$\varphi \equiv \exists \bar{\mathbf{x}} \psi$$

$$\mathsf{P}(\neg \forall \bar{\mathsf{x}} \neg \psi) = \mathsf{1} - \mathsf{P}(\forall \bar{\mathsf{x}} \neg \psi) = \mathsf{1} - (\mathsf{1} - \mathsf{P}(\psi))^{(2^{|\bar{\mathsf{x}}|})}$$



Expected Search Space Calculation

1
$$\varphi \equiv x \mid \neg x$$

$$S(x) = S(\neg x) = 1$$



$$\mathbf{2} \ \varphi \equiv \bigwedge_{i=1}^{n} \varphi_{i}$$

$$S\left(\bigwedge_{i=1}^{n} \varphi_{i}\right) = \sum_{i=1}^{n} \widehat{P}_{i}(\varphi)\widehat{S}_{i}(\varphi) + \left(\prod_{j=1}^{n} P(\varphi_{j})\right)\left(\sum_{j=1}^{n} S(\varphi_{j})\right)$$

$$\widehat{P}_{i}(\varphi) = \left(\prod_{j=1}^{i-1} P(\varphi_{j})\right)(1 - P(\varphi_{i})),$$

$$\widehat{S}_{i}(\varphi) = \sum_{i=1}^{i} S(\varphi_{j})$$

where $P_i(\varphi)$ is the probability that $\llbracket \varphi_i \rrbracket = \top$ for any $1 \le i \le i - 1$ and $\llbracket \varphi_i \rrbracket = \bot$.



Expected Search Space Calculation

2 $\varphi = \bigwedge_{i=1}^{n} \varphi_i$ Introducing an imaginary element φ_{n+1} such that $P(\varphi_{n+1}) = 0$ and $S(\varphi_{n+1}) = 0$, we simplify to

$$S\left(\bigwedge_{i=1}^{n} \varphi_{i}\right) = \sum_{i=1}^{n+1} \widehat{P}_{i}(\varphi)\widehat{S}_{i}(\varphi)$$



Expected Search Space Calculation

3 $\varphi \equiv \bigvee_{i=1}^{n} \varphi_{i}$

$$S\left(\bigvee_{i=1}^{n} \varphi_{i}\right) = \sum_{i=1}^{n+1} \widetilde{P}_{i}(\varphi) \widehat{S}_{i}(\varphi)$$
$$\widetilde{P}_{i}(\varphi) = \left(\prod_{j=1}^{i-1} \left(1 - P(\varphi_{j})\right)\right) P(\varphi_{i})$$

where $P(\varphi_{n+1})$ is an imaginary element set to 1.

4 $\varphi \equiv \forall \bar{x} \psi$ In the general case where \bar{x} has $|\bar{x}|$ elements, we have

$$\mathsf{S}(\forall ar{x}\psi) = (\mathsf{1} + \mathsf{P}(\psi))^{|ar{x}|}\,\mathsf{S}(\psi)$$

 $\varphi \equiv \exists \bar{x} \psi$

$$S(\exists \bar{x}\psi) = (2 - P(\psi))^{|\bar{x}|} S(\psi) \implies \text{ for all } y \in \mathbb{R}$$

Soonho Kona

Cost Calculation

- 1 $\varphi \equiv x \mid \neg x$ Leaf nodes have no child and hence no cost assigning is needed.
- $\varphi \equiv \bigwedge_{i=1}^n \varphi_i$

$$C(\varphi_i) = P(\varphi_i)S(\varphi_i)$$

3 $\varphi \equiv \bigvee_{i=1}^n \varphi_i$

$$C(\varphi_i) = (1 - P(\varphi_i))S(\varphi_i)$$

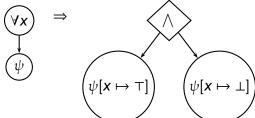
 $\varphi \equiv \forall \bar{x} \psi$

$$C(\psi_i) = P(\psi_i)S(\psi_i)$$

 $\varphi \equiv \exists \bar{\mathbf{x}} \psi$

$$C(\psi_i) = (1 - P(\psi_i))S(\psi_i)$$

During the evaluation quantifier will be replaced with logical operator \land (or \lor) and instantiated child node.



Quantifier Node

Therefore, we can apply CMS for the quantifier.

Instead of sorting the subnodes, we control the variable instantiation order.



Variable Instantiation Order

If a quantifier node has *n* variables, then there could be 2^n variable instantiations.

$$\forall \{x_1, x_2, \dots, x_n\} \psi$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \top, \dots, x_n \mapsto \top] = ?$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \top, \dots, x_n \mapsto \bot] = ?$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \bot, \dots, x_n \mapsto \top] = ?$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \bot, \dots, x_n \mapsto \bot] = ?$$

4 D > 4 A > 4 E > 4 E > B 9 Q P

Variable Instantiation Order

Naive approach would take 2ⁿ cost calculations to find minimal cost variable instantiation.

That is prohibitively expensive.



Variable Instantiation Order

We approximate the cost of an instantiation

$$C(\psi[x_1 \mapsto b_1][x_2 \mapsto b_2] \cdots [x_n \mapsto b_n]) = \frac{\sum_{i=1}^n C(\psi[x_i \mapsto b_i])}{n}$$

Using this approximation we only need to compute 2*n* single-variable instantiations.

Variable Instantiation Order

However, it is not efficient enough as we still have $2^{|\bar{x}|}$ combinations to compute and sort.

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \top, \dots, x_n \mapsto \top]) = ?$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \top, \dots, x_n \mapsto \bot]) = ?$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \bot, \dots, x_n \mapsto \top]) = ?$$

$$C(\psi[x_1 \mapsto \top, x_2 \mapsto \bot, \dots, x_n \mapsto \bot]) = ?$$

Here we need another heuristic.



Variable Instantiation Order - Preference Function

First we fix an instantiation order for every single variable by the notion of preference function

$$\mathsf{Pref}(x_i) = \left\{ \begin{array}{ll} \top & \text{if } \mathsf{C}(\psi[x_i \mapsto \top]) \leq \mathsf{C}(\psi[x_i \mapsto \bot]) \\ \bot & \text{otherwise} \end{array} \right.$$

For a variable x, Pref(x) returns the Boolean assignment which has the lower cost.

For example,
If
$$C(\psi[x_i \mapsto \top]) = 10$$
 and $C(\psi[x_i \mapsto \bot]) = 3$,
then $Pref(x_i) = \bot$.



ROPAS

Variable Instantiation Order - New Algorithm

We use Pref to modify the evaluation function []. as follows.

Outline

Variable Instantiation Order

In this algorithm, variable order induces instantiation order.

Example

$$\varphi = \forall x_1, x_2, x_3 \psi$$

where $\operatorname{Pref}(x_1) = \top$, $\operatorname{Pref}(x_2) = \bot$, and $\operatorname{Pref}(x_3) = \bot$.

Instantiation Order		
$\{[x_1 \mapsto \top],$	$[x_2 \mapsto \bot],$	$[x_3 \mapsto \bot]$
$\{[x_1 \mapsto \top],$	$[x_2 \mapsto \bot]$,	$[x_3 \mapsto \top]$
$\{[x_1 \mapsto \top],$	$[x_2 \mapsto \top]$,	$[x_3 \mapsto \bot]$
	•••	
$\{[x_1 \mapsto \bot],$	$[x_2 \mapsto \top]$,	$[x_3 \mapsto \top]$



Variable Instantiation Order

We introduce the significance of variable x_i as

$$\mathsf{Diff}(x_i) = \mathsf{C}(\psi[x_i \mapsto \neg \mathsf{Pref}(x_i)]) - \mathsf{C}(\psi[x_i \mapsto \mathsf{Pref}(x_i)]).$$

For example,
If
$$C(\psi[x_i \mapsto \top]) = 10$$
 and $C(\psi[x_i \mapsto \bot]) = 3$,
then $Pref(x_i) = \bot$ and $Diff(x_i) = 7$.

By Sorting \bar{x} by Diff in descending order, we can have better instantiation order.



ROPAS

$$\begin{array}{c|cccc} & x_1 & x_2 & x_3 \\ \hline C(\psi[x_i \mapsto \top]) & 3 & 4 & 5 \\ C(\psi[x_i \mapsto \bot]) & 2 & 7 & 3 \\ \end{array}$$

Variable Instantiation Order - Example

	<i>X</i> ₁	X 2	X 3
$C(\psi[x_i \mapsto \top])$	3	4	5
$C(\psi[x_i \mapsto \bot])$	2	7	<u>3</u>
$\overline{\text{Diff}(x_i)}$	1	3	2

Pref is marked and Diff is calculated.



$$\begin{array}{c|cccc} & x_2 & x_3 & x_1 \\ \hline C(\psi[x_i \mapsto \top]) & 4 & 5 & 3 \\ C(\psi[x_i \mapsto \bot]) & 7 & 3 & 2 \\ \hline Diff(x_i) & 3 & 2 & 1 \\ \hline \bar{x} \text{ is sorted by Diff.} \end{array}$$

Ins	Cost		
$\{[x_2\mapsto \top],$	$[x_3 \mapsto \bot],$	$[x_1 \mapsto \bot]$	3
$\{[x_2 \mapsto \top],$	$[x_3 \mapsto \bot]$,	$[x_1 \mapsto \top]$	3.3
$\{[x_2 \mapsto \top],$	$[x_3 \mapsto \top]$,	$[x_1 \mapsto \bot]$	3.6
	• • •		• • •
$\{[x_2\mapsto \bot],$	$[x_3 \mapsto \top]$,	$[x_1 \mapsto \top]$	5

Experiment - Configuration

Almost the same configuration with QBFEVAL'08

■ CPU : Intel Pentium D 3.0 GHz

RAM: 4GB

■ OS : Ubuntu/GNU Linux 7.10

■ Timeout : 600 sec



- QBF 1.0 benchmark from QBFLIB.
- 92 instances of non-prenex and non-CNF QBF
- generated from the computation of state space diameters of four models:
 - chain of inverters (Ring, 20 instances)
 - semaphore-based mutual exclusion protocol (Semaphore, 16 instances)
 - chain of inverters (Counter, 45 instances)
 - distributed mutual exclusion protocol (DME, 11 instances)



State-of-the-art Solvers

Three state-of-the-art solvers are chosen:

- QuBE6.0
- sKizzo v0.8.2-beta
- Quantor 3.0 with Picosat-632.

They only accept prenex and CNF instances. We used the converter provided by QBFLIB and did not include conversion time in running time.



Outline

Model	Quantor	sKizzo	QuBE	CMS
Ring	10/20	11/20	14/20	20/20
Semaphore	16/16	14/16	13/16	16/16
Counter	28/45	33/45	29/45	27/45
DME	0/11	0/11	6/11	0/11
Total	54/92	58/92	62/92	63/92

Table: Number of solved instances



Outline

Model	Quantor	sKizzo	QuBE	CMS
Ring	77,402	91,780	250,351	973,093
Semaphore	336,596	248,583	219,908	336,596
Counter	341,074	707,679	409,578	308,826
DME	0	0	633,770	0
Total	755,072	1,048,042	1,513,607	1,618,515

Table: Total size of solved instances (in bytes)



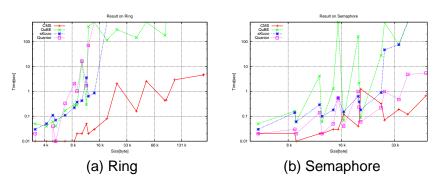


Figure: Performance comparison between *CMS* and three state-of-the-art solvers on QBF 1.0 test suite



Result on Counter

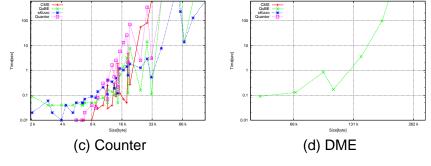


Figure: Performance comparison between *CMS* and three state-of-the-art solvers on QBF 1.0 test suite



Result on DME

- Using better search algorithm such as A* instead of depth-first search.
- Using BDD instead of tree representation.
- Extending to support more logical operator such as XOR and EQUIV which frequently occur in electronic design.



References

Outline

Giunchiglia, E., Narizzano, M., Tacchella, A.: Backjumping for quantified boolean logic satisfiability. In: IJCAI. (2001) 275–281

Zhang, L., Malik, S.: Conflict driven learning in a quantified boolean satisfiability solver.

In: Proceedings of the 2002 IEEE/ACM international conference on Computer-aided design (ICCAD'02), New York, NY, USA, ACM (2002) 442–449

Benedetti, M.: Evaluating QBFs via symbolic Skolemization.

In: 11th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning



References



Resolve and expand.

In: Proceedings of 7th International Conference on Theory and Applications of Satisfiability Testing (SAT'04), Springer-Verlag (2004)

Benedetti, M.: sKizzo: A QBF decision procedure based on propositional skolemization and symbolic reasoning. Technical report. ITC-Irst (2004)

Tseytin, G.S.:

On the complexity of derivation in propositional calculus. In: Studies in Constructive Mathematics and Mathematical Logic II. Volume 8 of Zapiski Nauchnykh Seminarov LOMI. Nauka, Leningrad (1968) 235–259 In Russian.



Future Work

- Giunchiglia, E., Narizzano, M., Tacchella, A.: QuBE: A system for deciding quantified boolean formulas satisfiability.
 - In: Proceedings of 1st International Joint Conference on Automated Reasoning (IJCAR'01), (2001) 364-369
- Giunchiglia, E., Narizzano, M., Tacchella, A.: Quantified Boolean Formulas Satisfiability Library (QBFLIB) (2001) www.qbflib.org.
- Peschiera, C., Pulina, L., Tacchella, A.: QBF solver evaluation portal http://www.qbflib.org/qbfeval.



Thank you

Question?

