

Quantified Boolean Formulas: Complexity and Expressiveness

Uwe Bubeck

University of Paderborn

20.11.2012

Outline



- Introduction
- Free Variables and Equivalence
- Complexity
- Expressiveness
- Conclusion





Introduction



Quantified Boolean Formulas 1/2



QBF extends propositional logic by allowing universal and existential quantifiers over propositional variables.

Inductive definition:

- 1. Every propositional formula is a QBF.
- 2. If Φ is a QBF then $\forall x \Phi$ and $\exists y \Phi$ are also QBFs.
- 3. If Φ_1 and Φ_2 are QBFs then $\neg \Phi_1$, $\Phi_1 \land \Phi_2$ and $\Phi_1 \lor \Phi_2$ are also QBFs.

Quantified Boolean Formulas 2/2



In a closed QBF, every variable is quantified.

Semantics definition for closed QBF:

 $\exists y \Phi(y)$ is true if and only if $\Phi[y/0]$ is true or $\Phi[y/1]$ is true.

 $\forall x \ \Phi(x)$ is true if and only if $\Phi[x/0]$ is true and $\Phi[x/1]$ is true.





- Closed QBFs Φ and Ψ are logically equivalent ($\Phi \approx \Psi$)
- \Leftrightarrow they are satisfiability equivalent
- \Leftrightarrow they both evaluate to the same truth value.
- Every QBF can be transformed in linear time into a logically equivalent prenex formula $Q_1v_1 \dots Q_kv_k \phi$ by:
- 1. renaming quantified variables,
- 2. transformation into negation normal form (NNF),
- 3. moving quantifiers to the front by maxiscoping: $(Qv \Phi) \circ \Psi \approx Qv (\Phi \circ \Psi)$ for $\circ \in \{\Lambda, V\}$ and $Q \in \{\forall, \exists\}$.



A closed prenex QBF $Q_1v_1 \dots Q_kv_k \phi$ is true if and only if there exists a tree such that: [Samulowitz et. al., 2006]

- 1. Each inner node is labeled with a variable v_i , $1 \le i \le k$, its outgoing edges are labeled with $v_i = 0$ or $v_i = 1$.
- 2. Inner nodes labeled with v_i have two children if and only if $Q_i = \forall$, and one child otherwise.
- 3. For each path from root to leaf labeled with $(v_{i_1}, ..., v_{i_j})$, we have $1 \le i_1 < \cdots < i_j \le k$ (i.e. order as in the prefix).
- 4. On each path from root to leaf, the edge labels are a satisfying assignment to the propositional matrix ϕ .



- A closed prenex QBF $Q_1v_1 \dots Q_kv_k \phi$ is true if and only if there exists a tree such that: [Samulowitz et. al., 2006]
- Each inner node is labeled with a variable v_i, 1 ≤ i ≤ k, its outgoing edges are labeled with v_i = 0 or v_i = 1.
 Inner nodes labeled with v_i have two children if and only if Q_i = ∀, and one child otherwise.
- 3. For each path from root to leaf labeled with $(v_{i_1}, ..., v_{i_j})$, we have $1 \le i_1 < \cdots < i_j \le k$ (i.e. order as in the prefix).
- 4. On each path from root to leaf, the edge labels are a satisfying assignment to the propositional matrix ϕ .













Tree Models: Generalization



The requirement that variables must appear in the same order as in the prefix (rule 3) can be further relaxed:

by grouping similar quantifiers into quantifier blocks.

$$\underbrace{\underbrace{\forall v_1 \dots \forall v_{n_1}}_{S_1} \underbrace{\exists v_{n_1+1} \dots \exists v_{n_2}}_{S_2} \underbrace{\forall v_{n_2+1} \dots \forall v_{n_3}}_{S_3} \dots \phi}_{S_3}$$

• by considering variable dependency schemes.

Dependency Schemes 1/2



Consider the following formula:

 $\forall x \exists y_1 \forall u \exists y_2 \exists y_3 [(x \lor u \lor y_3) \land (y_1 \lor y_3) \land (\neg y_1 \lor \neg y_3) \land (\neg u \lor \neg y_3) \land (\neg y_1 \lor \neg y_2)]$

Which variables must be above y_2 in the tree without altering the truth value of the formula?

In general:

dependency decision problem is PSPACE-complete. [Samer / Szeider, 2007]

Dependency Schemes 2/2



$$\forall x \exists y_1 \forall u \exists y_2 \exists y_3 [(x \lor u \lor y_3) \land (y_1 \lor y_3) \land (\neg y_1 \lor \neg y_3) \land (\neg u \lor \neg y_3) \land (\neg y_1 \lor \neg y_2)]$$

One simple heuristic: recovering non-prenex structure: $\forall x \exists y_1 [[\forall u \exists y_3 (x \lor u \lor y_3) \land (y_1 \lor y_3) \land (\neg y_1 \lor \neg y_3) \land (\neg u \lor \neg y_3)]$ $\land [\exists y_2 (\neg y_1 \lor \neg y_2)]]$

- Identified informally by [Biere, 2004],
- Formalized in [Samer / Szeider, 2007], [Bubeck / Kleine Büning, 2007],
- Currently tightest scheme by [van Gelder, 2011].

Still no practicable algorithm for computing tight schemes!



Free Variables and Equivalence



Semantics of Free Variables 1/2



A focus of this talk is to allow free variables. Notation:

- $\Phi(z_1, ..., z_n)$ for formula Φ with free variables $z_1, ..., z_n$.
- QBF* is the class of quantified Boolean formulas with free variables.
- The valuation of a QBF* depends on the values of the free variables:

 $\Phi(z_1, ..., z_n) \in QBF^*$ is satisfied by a truth assignment $t: \{z_1, ..., z_n\} \rightarrow \{0, 1\}$ if and only if $\Phi(t(z_1), ..., t(z_n)) \coloneqq \Phi[z_1/t(z_1), ..., z_n/t(z_n)]$ is true.



Example:

 $\Phi(a, b, c, d) := \exists y (a \lor y) \land (b \lor y) \land (\neg y \lor c \lor d)$

Then

 $\Phi(0,1,0,1) = \exists y (0 \lor y) \land (1 \lor y) \land (\neg y \lor 0 \lor 1) \text{ is true,} \\ \Phi(0,1,0,0) = \exists y (0 \lor y) \land (1 \lor y) \land (\neg y \lor 0 \lor 0) \text{ is false,} \\ \text{etc.}$

 $\rightarrow \exists y (a \lor y) \land (b \lor y) \land (\neg y \lor c \lor d) \text{ is true if and only if}$ $(a \lor c \lor d) \land (b \lor c \lor d) \text{ is true.}$

\rightarrow Every QBF* is equivalent to a propositional formula.

Logical Equivalence

Propositional or quantified Boolean formulas α and β with (free) variables $z_1, ..., z_n$ are logically equivalent, written as $\alpha(z_1, ..., z_n) \approx \beta(z_1, ..., z_n)$, if and only if $\alpha(t(z_1), ..., t(z_n)) = \beta(t(z_1), ..., t(z_n))$ for each truth assignment *t* to $z_1, ..., z_n$.

→ Quantified variables are not directly considered for logical equivalence. They can be seen as local within the respective formula.

Satisfiability Equivalence



 $(a \lor \underline{c} \lor \underline{d}) \land (b \lor \underline{c} \lor \underline{d}) \approx \exists y (a \lor y) \land (b \lor y) \land (\neg y \lor \underline{c} \lor \underline{d})$

Without quantifier:

 $(a \lor c \lor d) \land (b \lor c \lor d) \not\approx (a \lor y) \land (b \lor y) \land (\neg y \lor c \lor d)$

Problem:

a = b = 1, c = d = 0, y = 1 satisfies left, but not right side.

Relaxation: satisfiability equivalence $(a \lor c \lor d) \land (b \lor c \lor d) \approx_{SAT} (a \lor y) \land (b \lor y) \land (\neg y \lor c \lor d)$ Existence of a satisfying assignment for one side implies there is <u>some</u> satisfying assignment for the other side.

Equivalence and Rewriting



For propositional formulas α , β , γ , it holds that $\beta \approx \gamma \Rightarrow \alpha \approx \alpha [\beta/\gamma]$.

→ Parts of a formula can be replaced with logically equivalent formulas.

But:
$$\beta \approx_{SAT} \gamma \Rightarrow \alpha \approx_{SAT} \alpha [\beta/\gamma].$$

 \rightarrow Satisfiability equivalence is sometimes too weak.





Complexity





- Well known: the decision problem for QBF and the satisfiability problem for QBF* are PSPACE-complete. [Meyer / Stockmeyer, 1973]
- → QBF* consequence and equivalence problems are also PSPACE-complete.
- \rightarrow QDNF* remains PSPACE-complete.

Some verification problems are also PSPACE-complete:

- Propositional LTL satisfiability
 [Sistla / Clarke, 1985]
- Symbolic reachability in sequential circuits [Savitch, 1970]

QBF* Complexity 2/2



Typical approaches to reduce the complexity:

- Restriction of the matrix to special classes of propositional formulas
- Bounding of quantifier alternations in the prefix:

k quantifier blocks, outermost existential:

 $\underbrace{\forall \dots \forall \exists \dots \exists \forall \dots \forall \dots}_{\bullet} \phi$

k quantifier blocks, outermost universal:

prefix type Π_k

prefix type
$$\Sigma_k$$

Polynomial-time Hierarchy



Well-known relationship with polynomial-time hierarchy:

For fixed $k \ge 1$, the satisfiability problem for QBF* with prefix type Σ_k is Σ_k^P -complete, and Π_k^P -complete for prefix type Π_k .

[Stockmeyer, 1976]

$$\Delta_0^P \coloneqq \Sigma_0^P \coloneqq \Pi_0^P \coloneqq P$$

$$\Sigma_{k+1}^P \coloneqq NP^{\Sigma_k^P}, \qquad \Pi_{k+1}^P \coloneqq co - \Sigma_{k+1}^P, \qquad \Delta_{k+1}^P \coloneqq P^{\Sigma_k^P}$$

Schaefer's Dichotomy Theorem



Consider a quantified constraint expression

 $Q_1y_1 \dots Q_ny_n f_1(x_{1,1}, \dots, x_{1,m_1}) \land \dots \land f_k(x_{k,1}, \dots, x_{k,m_k})$ with $Q_1, \dots, Q_n \in \{\forall, \exists\}$, Boolean functions $f_1, \dots, f_k \in C$ and arguments $x_{i,j} \in \{y_1, \dots, y_n\} \cup \{0,1\}$.

Dichotomy Theorem:

Let *C* be a finite set of constraints. If *C* is *Horn*, *anti-Horn*, *bijunctive / Krom* (equiv. to 2-CNF) or *affine* (equiv. to XOR-CNF) then $QSAT_c(C)$ is in P. Otherwise, $QSAT_c(C)$ is PSPACE-complete.

[Schaefer, 1978], [Dalmau, 1997] [Creignou / Khanna et al., 2001]

QHORN* and Q2-CNF*

- QHORN* is the class of QCNF* formulas with at most one positive literal per clause.

QHORN* satisfiability is decidable in quadratic time $O(|\forall| \cdot |\Phi|)$, where $|\forall|$ is the number of universal variables and $|\Phi|$ the formula length.

Algorithm: by unit propagation[Flögel et al. 1995]or by universal expansion[Bubeck / Kleine Büning, 2008]

Q2-CNF* satisfiability is decidable in linear time.
 Algorithm: by strongly connected components

[Aspvall et al., 1979]





Expressiveness



Other Representations



We can also define logical equivalence with other representations of Boolean functions:



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

 $\approx \exists y \, (a \lor y) \land (b \lor y) \land (\neg y \lor c \lor d)$

For each assignment to a, b, c, d, all three representations evaluate to the same truth value.

Main Theme:

How compact are encodings in QBF* (or subclasses) versus other logically equivalent representations?

Encoding Techniques:

1. Abbreviate exact repetitions by existential variables: $(\underline{A \lor \neg B \lor C \lor D}) \land (\underline{A \lor \neg B \lor C} \lor \neg E) \land (\underline{A \lor \neg B \lor C} \lor F)$ $\approx \exists y (y \leftrightarrow (\underline{A \lor \neg B \lor C})) \land (y \lor D) \land (y \lor \neg E) \lor (y \lor F)$ Simple implication if term occurs only in one polarity: $\exists y (y \rightarrow (A \lor \neg B \lor C)) \land (y \lor D) \land (y \lor \neg E) \lor (y \lor F)$

(quantified versions of [Tseitin, 1970] and [Plaisted, 1986])

Quantified Encodings 2/2



2. Compress conjunctions of renamings / instantiations by universal variables:

$$\phi(A_1, B_1) \land \phi(A_2, B_2) \land \phi(A_3, B_3) \approx \forall u \forall v \left(\bigvee_{i=1\dots 3} ((u \leftrightarrow A_i) \land (v \leftrightarrow B_i)) \right) \to \phi(u, v)$$

[Dershowitz et al., 2005], [Meyer/Stockmeyer, 1973]

3. Iterative Squaring (Extension of 2.): $\Phi(x_{0}, x_{n}) = \exists x_{1} \dots \exists x_{n-1} \phi(x_{0}, x_{1}) \land \phi(x_{1}, x_{2}) \land \dots \land \phi(x_{n-1}, x_{n})$ Encoding: $\Phi_{n}(x_{0}, x_{n}) \coloneqq \exists y (\Phi_{n/2}(x_{0}, y) \land \Phi_{n/2}(y, x_{n}))$ $\approx \exists y \forall u \forall v (((u \leftrightarrow x_{0}) \land (v \leftrightarrow y)) \lor ((u \leftrightarrow y) \land (v \leftrightarrow x_{n}))) \rightarrow \Phi_{n/2}(u, v)$ [Meyer/Stockmeyer, 1973]



Instantiations might also be nested:

 $((A_1 \wedge A_2) \vee (\neg A_1 \wedge \neg A_2)) \rightarrow ((B_1 \wedge B_2) \vee (\neg B_1 \wedge \neg B_2))$ is of the form $\psi(\phi(A_1, A_2), \phi(B_1, B_2)).$

Other example: Parity of *n* Boolean variables $f_0(p_1, p_2) \coloneqq (\neg p_1 \land p_2) \lor (p_1 \land \neg p_2)$ $f_1(p_1, p_2, p_3, p_4) \coloneqq f_0(f_0(p_1, p_2), f_0(p_3, p_4))$ $f_2(p_1, \dots, p_{16}) \coloneqq f_1(f_1(p_1, \dots, p_4), \dots, f_1(p_{13}, \dots, p_{16}))$ $\log_2 \log_2 n + 1$ definitions of size O(n).



General Definition: A Nested Boolean Function (NBF) is a sequence of functions $D(f_k) = (f_0, ..., f_k)$ with

- initial functions $f_0, ..., f_t$ defined by $f_i(x^i) \coloneqq \alpha_i(x^i)$ for a propositional formula α_i over $x^i \coloneqq (x^{i,1}, ..., x^{i,n_i})$
- compound functions $f_{t+1}, ..., f_k$ of the form $f_i(\mathbf{x}^i) \coloneqq f_{j_0}(f_{j_1}(\mathbf{x}_1^i), ..., f_{j_r}(\mathbf{x}_r^i))$ with previously defined functions $f_{j_0}, ..., f_{j_r}$ and matching tuples $\mathbf{x}_1^i, ..., \mathbf{x}_r^i$ over variables from \mathbf{x}^i or Boolean constants.

[Bubeck / Kleine Büning, 2012], [Cook / Soltys, 1999]

Nested Instantiations 3/3



- By clever combination of previously presented encoding techniques, every NBF can be transformed in linear time into a logically equivalent prenex QBF*.
 [Bubeck / Kleine Büning, 2012]
- The inverse direction is very simple by simulating quantifier expansion, but length increases slightly to O(|v| · |Φ|) for a QBF* Φ with |v| quantified variables.
- **Application:** Configuration Problems (→ Talk by Hans)
- Future Work: Solvers, interesting NBF subclasses

Circuits and Existential Quantifiers

There is a close connection between fan-out in Boolean circuits and existential quantification:

1. Transformation from circuit to formula



$$\exists y_1 \dots \exists y_8$$

 $(a \lor y_1) \land (c \lor y_2) \land$
 $(d \lor y_3) \land (b \lor y_4) \land$
 $(y_2 \land y_3 \rightarrow y_5) \land$
 $(y_1 \land y_5 \rightarrow y_6) \land$
 $(y_5 \land y_4 \rightarrow y_7) \land$
 $(y_6 \lor y_7 \rightarrow y_8) \land$

[Bauer / Brand et al., 1973] [Anderaa / Börger, 1981]

Circuits and Existential Quantifiers



There is a close connection between fan-out in Boolean circuits and existential quantification:

1. Transformation from circuit to formula



The Class ∃HORN^b

- The previous linear transformation produces

 existentially quantified formulas in CNF with at most
 one positive bound variable per clause (i.e. the bound
 variables respect the Horn property).

 We call such formulas ∃HORN^b.
- **General:** for QK \subseteq QCNF, QK^b is the class of formulas $\Phi(z_1, ..., z_n) = Q_1 v_1 ... Q_m v_m \wedge_i (\phi_i^b(v_1, ..., v_m) \vee \phi_i^f(z_1, ..., z_n))$ where $Q_1 v_1 ... Q_m v_m \wedge_i \phi_i^b(v_1, ..., v_m)$ is a formula in QK, and $\phi_i^f(z_1, ..., z_n)$ an arbitrary clause over free variables.

Circuits and Existential Quantifiers



2. Transformation from formula to circuit

Any ∃HORN^b formula can be transformed in polynomial time into a logically equivalent Boolean circuit.

[Anderaa / Börger, 1981] [Kleine Büning / Zhao / Bubeck, 2009]

 $\rightarrow \exists HORN^{b}$ and circuits have similar expressiveness.

HORN^b vs **Propositional Formulas**



- There are \exists HORN^b (even \exists HORN*) formulas for which there is no logically equivalent propositional CNF of polynomial length.
- [Kleine Büning / Lettmann, 1999]
- Is a polynomial-length transformation from $\exists HORN^{b}$ to arbitrary propositional formulas possible?
- Equivalent: Are circuits with unrestricted fan-out more expressive than with fan-out 1?



Can the expressive power of \exists HORN^b be enhanced by universal quantification?

Not significantly. For every $\Phi \in QHORN^{b}$ with $|\forall|$ universal quantifiers, there exists a logically equivalent $\exists HORN^{b}$ formula of quadratic length $O(|\forall| \cdot |\Phi|)$ which can be computed also in time $O(|\forall| \cdot |\Phi|)$. [Bubeck, 2010]

That means QHORN^b satisfiability is NP-complete.

Minimal Falsity and Quantification



A closed QCNF formula is minimal false (MF) iff it is false and removing an arbitrary clause makes it true.

Example:

 $\forall x \exists y_0 \exists y_1 (y_0 \lor z_0) \land (\neg y_0 \lor z_1) \land (\neg y_0 \lor z_2) \land (y_1 \lor z_3) \land (x \lor z_4)$

Bound parts: $\forall x \exists y_0 \exists y_1 (y_0) \land (\neg y_0) \land (\neg y_0) \land (y_1) \land (x)$

MF subformulas:

- $1. \quad \exists y_0 (y_0) \land (\neg y_0)$
- $2. \quad \exists y_0 (y_0) \land (\neg y_o)$
- 3. $\forall x (x)$

Minimal Falsity and Quantification



 $\forall x \exists y_0 \exists y_1 (y_0 \lor z_0) \land (\neg y_0 \lor z_1) \land (\neg y_0 \lor z_2) \land (y_1 \lor z_3) \land (x \lor z_4)$

Bound parts: $\forall x \exists y_0 \exists y_1 (y_0) \land (\neg y_0) \land (\neg y_0) \land (y_1) \land (x)$

MF subformulas:

Corresponding bound parts:

- 1. $\exists y_0 (y_0) \land (\neg y_0)$ 1. z_0, z_1
- 2. $\exists y_0 (y_0) \land (\neg y_0)$ 2. z_0, z_2
- 3. $\forall x (x)$ 3. z_4

For each MF subformula, one of the corresponding free parts must be satisfied.

So the formula is equivalent to $(z_0 \lor z_1) \land (z_0 \lor z_2) \land z_4$.

Minimal Falsity and Quantification



MF subformulas of the bound parts determine the role of the free parts in QCNF* formulas:

$$Q \bigwedge_{1 \le i \le q} (\phi_i^b \lor \phi_i^f) \approx \bigwedge_{\substack{\left(Q \ \phi_{i_1}^b \land \dots \land \phi_{i_r}^b \right) \in MF}} \left(\phi_{i_1}^f \lor \dots \lor \phi_{i_r}^f \right)$$

all MF subformulas

of the bound parts

[Bubeck / Kleine Büning, 2010]

How do restrictions on the structure of the MF skeleton influence the expressiveness?





Conclusion



Conclusion

- Free variables are nothing to be afraid of.
- Compared to propositional logic, there is a very rich set of QBF* subclasses.
- How do these subclasses differ in expressiveness?
 Is there a hierarchy PROP <_{poly-len} ∃HORN^b <_{poly-len} ∃CNF* <_{poly-len} ∀∃CNF* ... ?
- How are these classes related to other representations, e.g. circuits?

References 1/6

- S. Anderaa and E. Börger. The Equivalence of Horn and Network Complexity for Boolean Functions. 1981
- B. Aspvall, M. Plass and R. Tarjan.
 A Linear-Time Algorithm for Testing the Truth of Certain Quantified Boolean Formulas. 1979
- M. Bauer, D. Brand, M. Fischer, A. Meyer and M. Paterson.
 A Note on Disjunctive Form Tautologies. 1973
- A. Biere.

Resolve and Expand. 2004

• U. Bubeck.

Model-Based Transformations for Quantified Boolean Formulas. 2010



References 2/6

- U. Bubeck and H. Kleine Büning.
 Bounded Universal Expansion for Preprocessing QBF. 2007
- U. Bubeck and H. Kleine Büning.
 Models and Quantifier Elimination for Quantified Horn Formulas. 2008
- U. Bubeck and H. Kleine Büning.
 The Power of Auxiliary Variables for Propositional and Quantified Boolean Formulas. 2010
- U. Bubeck and H. Kleine Büning.
 Encoding Nested Boolean Functions as Quantified Boolean Formulas.
 2012
- S. Cook and M. Soltys.

Boolean Programs and Quantified Propositional Proof Systems. 1999



References 3/6

- N. Creignou, S. Khanna and M. Sudan.
 Complexity Classifications of Boolean Constraint Satisfaction Problems. 2001
- V. Dalmau

Some Dichotomy Theorems on Constant-free Boolean Formulas. 1997

- N. Dershowitz, Z. Hanna and J. Katz. Bounded Model Checking with QBF. 2005
- U. Egly, M. Seidl, H. Tompits, S. Woltran and M. Zolda.
 Comparing Different Prenexing Strategies for Quantified Boolean Formulas. 2004
- A. Flögel, M. Karpinski and H. Kleine Büning. Resolution for Quantified Boolean Formulas. 1995



References 4/6

- H. Kleine Büning and U. Bubeck.
 Theory of Quantified Boolean Formulas. 2009
- H. Kleine Büning and T. Lettmann.
 Propositional Logic: Deduction and Algorithms. 1999
- H. Kleine Büning, X. Zhao and U. Bubeck. Resolution and Expressiveness of Subclasses of Quantified Boolean Formulas and Circuits. 2009
- A. Meyer and L. Stockmeyer. *Word Problems Requiring Exponential Time*. 1973
- D. Plaisted and S. Greenbaum.

A Structure-preserving Clause Form Translation. 1986



References 5/6

• M. Samer and S. Szeider.

Backdoor Sets of Quantified Boolean Formulas. 2007

- H. Samulowitz, J. Davies, F. Bacchus.
 Preprocessing QBF. 2006
- W. Savitch.

Relationships Between Nondeterministic and Deterministic Tape Complexities. 1970

• T. Schaefer.

The Complexity of Satisfiability Problems. 1978

• A. Sistla and E. Clarke.

The Complexity of Propositional Linear Time Logics. 1985





• L. Stockmeyer.

The Polynomial-Time Hierarchy. 1976

• G. Tseitin.

On the Complexity of Derivation in Propositional Calculus. 1970

• A. van Gelder.

Variable Independence and Resolution Paths. 2011