Quantifier Rewriting and Equivalence Models for Quantified Horn Formulas

Uwe Bubeck[1], Hans Kleine Büning[2], and Xishun Zhao[3]

[1] International Graduate School Dynamic Intelligent Systems, Universität Paderborn, 33098 Paderborn (Germany), bubeck@upb.de
[2] Department of Computer Science, Universität Paderborn, 33098 Paderborn (Germany), kbcsl@upb.de
[3] Institute of Logic and Cognition, Zhongshan University, 510275, Guangzhou (P.R. China), hsdp08@zsu.edu.cn

Abstract. In this paper, quantified Horn formulas with free variables ($QHORN^*$) are investigated. The main result is that any quantified Horn formula Φ of length $|\Phi|$ with free variables, $|\forall|$ universal quantifiers and an arbitrary number of existential quantifiers can be transformed into an equivalent formula of length $O(|\forall| \cdot |\Phi|)$ which contains only existential quantifiers. Moreover, it is shown that quantified Horn formulas with free variables have equivalence models where every existential quantifier is associated with a monotone Boolean function.
The results allow a simple representation of quantified Horn formulas as purely existentially quantified Horn formulas ($\exists HORN^*$). An application described in the paper is to solve $QHORN^*$-SAT in $O(|\forall| \cdot |\Phi|)$ by using this transformation in combination with a linear-time satisfiability checker for propositional Horn formulas.

1 Introduction

Quantified Boolean Formulas (QBF) offer a concise way to represent formulas which arise in areas such as planning, scheduling or verification. The ability to provide compact representations for many Boolean functions does however come at a price: determining the satisfiability of formulas in QBF is PSPACE-complete, which is assumed to be significantly harder than the NP-completeness of the propositional SAT problem. However, continued research and the lifting of propositional SAT techniques to $QBFs$ have recently produced interesting improvements (see, e.g., [2, 3, 9, 10]) and have lead to the emergence of more powerful QBF-SAT solvers [8].

Furthermore, the satisfiability problem is known to be tractable for some restricted subclasses like $QHORN$ [5] or $Q2-CNF$ [1]. Those classes are defined by imposing restrictions on the syntactic structure of the formula matrices. The interesting question which we are investigating in this paper is how such a syntactic restriction is affecting the structure of the set of satisfying truth value assignments to the existentially quantified variables.

A suitable concept for describing the satisfiable truth value assignments to the existential variables is the notion of models for formulas in QBF, which has been introduced in [6]: for a quantified formula Φ with existential variables $\mathbf{y} = y_1, ..., y_m$, let $M = (f_{y_1}, ..., f_{y_m})$ be a mapping which maps each existential variable y_i to a propositional formula f_{y_i} over universal variables whose quantifiers precede the quantifier of y_i. Then M is a satisfiability model for Φ if the resulting formula $\Phi[\mathbf{y}/M] := \Phi[y_1/f_{y_1}, ..., y_m/f_{y_m}]$, where simultaneously each existential variable y_i is replaced by its corresponding formula f_{y_i} and the existential quantifiers are dropped from the prefix, is true.

The concept of models for closed formulas can easily be extended to quantified Boolean formulas with free variables (QBF^*) [7]. In that case, the propositional formulas f_{y_i} may also contain free variables. For QBF^* formulas, an interesting additional constraint is to require that Φ and $\Phi[\mathbf{y}/M]$ must be equivalent. Formally, *equivalence models* are defined as follows: let $\Phi(\mathbf{z}) = Q\,\phi(\mathbf{x}, \mathbf{y}, \mathbf{z})$ be a quantified Boolean formula with prefix Q and matrix ϕ, universal variables $\mathbf{x} = x_1, ..., x_n$, existential variables $\mathbf{y} = y_1, ..., y_m$ and free variables $\mathbf{z} = z_1, ..., z_r$. For propositional formulas f_{y_i} over $\mathbf{z}$ and over universal variables whose quantifiers precede $\exists y_i$, we say $M = (f_{y_1}, ..., f_{y_m})$ is an equivalence model for $\Phi(\mathbf{z})$ if and only if $\Phi(\mathbf{z}) \approx \forall x_1 ... \forall x_n\, \phi(x_1, ..., x_n, \mathbf{y}, \mathbf{z})[\mathbf{y}/M]$.

In this paper, we focus on the class of quantified Horn formulas with free variables ($QHORN^*$). While it has been previously known that closed quantified Horn formulas have equivalence models consisting of monotone monomials and the constant functions 0 or 1 (see [6]), the structure of $QHORN^*$ equivalence models has been an open problem. In Section 2, we show that $QHORN^*$ formulas have equivalence models where the propositional formulas f_{y_i} contain only conjunctions and disjunctions of positive literals, i.e. the f_{y_i} are monotone.

In the second part of this paper, we turn to the universal quantifiers. Section 3 demonstrates that we can eliminate all universal quantifiers and the corresponding universal variables in $QHORN^*$ formulas without significantly increasing the length of the formula. To be more precise, we show that a quantified Horn formula Φ of length $|\Phi|$ with free variables, $|\forall|$ universal quantifiers and an arbitrary number of existential quantifiers can be transformed into an equivalent formula of length $O(|\forall| \cdot |\Phi|)$ which contains only existential quantifiers.
We finally explain how this transformation can be used to solve the satisfiability problem for $QHORN^*$ in time $O(|\forall| \cdot |\Phi|)$ with a very simple algorithm.

We need some additional notation: for $\Phi \in QBF^*$, $\Phi(\mathbf{z}) = Q\,\phi(\mathbf{x}, \mathbf{y}, \mathbf{z})$ with $Q = \forall x_{1,1} ... \forall x_{1,n_1} \exists y_{1,1} ... \exists y_{1,m_1} ... \forall x_{r,1} ... \forall x_{r,n_r} \exists y_{r,1} ... \exists y_{r,m_r}$, $n_i \geq 1$ and $m_i \geq 1$ for $i = 1, .., r$, we combine successive quantifiers of the same kind and simply write $Q = \forall X_1 \exists Y_1 ... \forall X_r \exists Y_r$ with $X_i = (x_{i,1}, ..., x_{i,n_i})$ and $Y_i = (y_{i,1}, ..., y_{i,m_i})$. Another notation that we use is $ab := (a_1, ..., a_m, b_1, ..., b_n)$ to denote the concatenation of two tuples $a = (a_1, ..., a_m)$ and $b = (b_1, ..., b_n)$.

2 Equivalence Models for $QHORN^*$ Formulas

Definition 1. *Let $M = (f_{y_1}, ..., f_{y_m})$ be an equivalence model for a quantified Boolean formula $\Phi \in QBF^*$. Then M is a* ***monotone equivalence model*** *if and only if the functions f_{y_i}, $1 \leq i \leq m$, do not contain negative occurrences of atoms, i.e. the f_{y_i} can be written using only positive literals and the reduced operator set $\{\wedge, \vee\}$ as well as $f_{y_i} = 0$ or $f_{y_i} = 1$.*

Theorem 1. *Any formula $\Phi \in QHORN^*$ has a monotone equivalence model $M = (f_{y_1}, ..., f_{y_m})$ which satisfies the following properties: $\Phi[\mathbf{y}/M] \in QHORN^*$, and any clause in Φ with a positive existential variable contributes only tautological clauses to $\Phi[\mathbf{y}/M]$.*

Proof: Due to space limitations, we omit the proof. It is by induction on the number of quantifiers and can be outlined as follows: in the induction base, when there is only one single existential variable, a monotone model is chosen such that tautological clauses are created when the model is substituted for positive instances of the existential variable. On the other hand, the negative instances produce the set of possible Q-resolvents, which is known to be equivalent to the original formula. In the induction step, we compose a monotone equivalence model for $k-1$ existential variables and a monotone equivalence model for the k-th existential variable to obtain a model for all k existential variables.

Unfortunately, these models may be of exponential length, as a short argument reveals. It is known (see, e.g., [5]) that converting a quantified Horn formula into an equivalent propositional Horn formula may result in an exponential increase in length required. But if we had an equivalence model of polynomial length, we could use it to eliminate the existential variables, which in turn would allow us to drop the universal variables, and we would end up with a purely propositional formula of polynomial length. That contradiction leads to the conclusion that the equivalence model may have exponential length. Nevertheless, knowing about the monotony of those models is useful, as it allows for concise and elegant proofs.

3 Eliminating Universal Quantifiers

As converting a quantified Horn formula into an equivalent propositional Horn formula may result in an exponentially longer formula, it is not a practical way to go. As discussed above, eliminating only the existential quantifiers is not practical either. But as it turns out, eliminating just the universal quantifiers does not lead to this exponential increase in length. We will now prove this surprising result and present the corresponding algorithm.

Definition 2. *A formula $\Phi \in QHORN^*$ is an* ***existentially quantified Horn formula*** *with free variables if it is of the form $\Phi(\mathbf{z}) = \exists y_1 ... \exists y_m\, \phi(\mathbf{z})$ $(m \geq 0)$, i.e. if it does not contain universally quantified variables. The class of all such formulas we denote by $\exists HORN^*$.*

The goal of the following investigation is to transform an arbitrary formula in $QHORN^*$ into an equivalent formula in $\exists HORN^*$ with a polynomial increase in length. The method that we present is a specialization of the known exponential method of expanding universal quantifiers in general QBF^* formulas. We first present the general technique and then show how it can be refined for $QHORN^*$.

3.1 Eliminating Universal Quantifiers in QBF^* Formulas

The general method for expanding universal quantifiers is rather straightforward: two copies of the original matrix are generated, one for the universally quantified variable being true, and one for that variable being false. As explained in [2], existential variables which depend on that universal variable need to be duplicated as well. For example, the formula $\exists y_1 \forall x \exists y_2 \, \phi(x, y_1, y_2)$ is expanded to $\exists y_1 \exists y_2 \exists y_2' \, \phi(0, y_1, y_2) \wedge \phi(1, y_1, y_2')$. For multiple universal quantifiers, we successively expand each universal quantifier, starting with the innermost.

Based on this informal description, we now provide a formal representation.
Let $\Phi \in QBF^*$ with $\Phi(\mathbf{z}) = \forall X_1 \exists Y_1 ... \forall X_r \exists Y_r \, \phi(X_1, ..., X_r, Y_1, ..., Y_r, \mathbf{z})$, where $X_i = (x_{i,1}, ..., x_{i,n_i})$ and $Y_i = (y_{i,1}, ..., y_{i,m_i})$ ($n_i \geq 1$ and $m_i \geq 1$, $i = 1, ..., r$, $r \geq 1$). Without loss of generality, we assume that the outermost quantifiers are universal and the innermost quantifiers are existential.

The expanded formula is then given as $\Phi_{\exists \exp}(\mathbf{z}) :=$

$$\bigwedge_{A_1 \in \{0,1\}^{n_1}} \exists Y_1^{A_1} \left(\bigwedge_{A_2 \in \{0,1\}^{n_2}} \exists Y_2^{A_2} .... \left(\bigwedge_{A_r \in \{0,1\}^{n_r}} \exists Y_r^{A_r} \, \phi(A_1 ... A_r, Y_1 ... Y_r, \mathbf{z}) \right) ... \right)$$

The tuples A_i represent the possible truth value assignments to the universal variables $x_{i,1}, ..., x_{i,n_i}$. The expression $\bigwedge_{A_i \in \{0,1\}^{n_i}}$ should be understood as a conjunction of 2^{n_i} clauses, one for each truth value assignment. Finally, $\exists Y_i^{A_i}$ is an abbreviation for $\exists y_{i,1}^{A_i} ... \exists y_{i,m_i}^{A_i}$, the copies of the i-th sequence of existential quantifiers. The index A_i is used to have a unique name for each of those copies. For example, four copies of $y_{i,j}$ would be named $y_{i,j}^{(0,0)}$, $y_{i,j}^{(0,1)}$, $y_{i,j}^{(1,0)}$ and $y_{i,j}^{(1,1)}$.

If there are n universal quantifiers in a formula Φ, its expansion $\Phi_{\exists \exp}$ contains 2^n copies of the formula's original matrix. This exponential increase in length makes the method unusable in general, but we can significantly simplify it:

3.2 Special Case: $QHORN^*$ Formulas

Definition 3. *Let* $\Phi(\mathbf{z}) = \forall X_1 \exists Y_1 ... \forall X_r \exists Y_r \, \phi(X_1, ..., X_r, Y_1, ..., Y_r, \mathbf{z})$ *with* $X_i = (x_{i,1}, ..., x_{i,n_i})$ *and* $Y_i = (y_{i,1}, ..., y_{i,m_i})$ *(*$n_i \geq 1$ *and* $m_i \geq 1$, $i = 1, ..., r$, $r \geq 1$*) be a quantified Horn formula whose outermost quantifiers are universal and whose innermost quantifiers are existential.*

Then we define the formula $\Phi_{\exists\text{poly}}(\mathbf{z})$ as

$$\Phi_{\exists\text{poly}}(\mathbf{z}) := \bigwedge_{A_1 \in Assign_1} \exists Y_1^{A_1}$$

$$\left(\bigwedge_{A_2 \in Assign_2(A_1)} \exists Y_2^{A_2} \ldots. \left(\bigwedge_{A_r \in Assign_r(A_1 \ldots A_{r-1})} \exists Y_r^{A_r} \, \phi(A_1 \ldots A_r, Y_1 \ldots Y_r, \mathbf{z}) \right) \ldots \right)$$

with the restricted set of possible assignments

$$Assign_1 = MaxOneZero(n_1)$$

$$Assign_i(A_1, \ldots, A_{i-1}) = \begin{cases} MaxOneZero(n_i) & \text{, if } A_1 \ldots A_{i-1} = \{1\}^{n_1 + \ldots + n_{i-1}} \\ \{1\}^{n_i} & \text{, else} \end{cases}$$

where

$$MaxOneZero(n) = \{(a_1, \ldots, a_n) \,|\, \exists i : \, a_i = 0 \, \text{and} \, a_j = 1 \, \text{for} \, j \neq i\} \cup \{(1, \ldots, 1)\}$$

is the set of n-ary tuples of binary values with at most one component being zero.

The only difference between the formula $\Phi_{\exists\text{poly}}$ and the expansion $\Phi_{\exists\text{exp}}$ for general QBF^* formulas is that for quantified Horn formulas, not all possible truth value assignments to the universally quantified variables have to be considered. For Horn formulas, we discard assignments where more than one universally quantified variable is false.

Lemma 1. *$\Phi_{\exists\text{poly}}$ is equivalent to Φ.*

Proof: We need to show that for a quantified Horn formula Φ with free variables $\mathbf{z}$, $\Phi_{\exists\text{exp}}(\mathbf{z}) \approx \Phi_{\exists\text{poly}}(\mathbf{z})$ holds. Due to space considerations, we also have to omit this proof. The main proof idea is as follows: if the matrix of $\Phi_{\exists\text{poly}}$ can be satisfied when exactly one universal variable is false, we can compensate for an additional universal variable being false by modifying the truth value assignment to the existential variables.

In the definition of $\Phi_{\exists\text{poly}}$, we can observe that there is one instantiation of the matrix of the original formula for each possible assignment to the universal variables where either all of those variables are true, or exactly one of them is false. There are $n + 1$ such assignments. Furthermore, the previous lemma has shown that $\Phi_{\exists\text{poly}}$ is equivalent to $\Phi_{\exists\text{exp}}$, which in turn is equivalent to Φ, so we immediately have the following theorem:

Theorem 2. *For any quantified Horn formula $\Phi \in QHORN^*$ with free variables, there exists an equivalent formula $\Phi' \in \exists HORN^*$ without universal quantifiers. The length of Φ' is bounded by $|\forall| \cdot |\Phi|$, where $|\forall|$ is the number of universal quantifiers in Φ, and $|\Phi|$ is the length of Φ.*

3.3 The Transformation Algorithm

Listing 1 presents an algorithm to transform Φ into $\Phi_{\exists\text{poly}}$ as described above. In the main loop of the algorithm, one universal variable $x_{i,j}$ is given the value false, while all the others are true. For any such assignment $A_{\mathbf{x}}$, all universal variables which are dominated by $x_{i,j}$ (i.e. their corresponding quantifier follows $\forall x_{i,j}$) have to be replaced by independent new variables $\mathbf{y}'$. Then, the matrix of the original formula has to be duplicated, with $A_{\mathbf{x}}$ being substituted for $\mathbf{x}$ and $\mathbf{y}'$ being substituted for $\mathbf{y}$. After executing the main loop, one additional copy is needed for the case where all universal variables are true.

Listing 1: The Transformation Algorithm

```
// Input: Φ ∈ QHORN*, Φ(z) = ∀X₁∃Y₁...∀X_r∃Y_r φ(X₁,...,X_r,Y₁,...,Y_r,z),
//        where X_i = (x_{i,1},...,x_{i,n_i}) and Y_i = (y_{i,1},...,y_{i,m_i})
// Output: Φ_∃poly ∈ ∃HORN* with Φ_∃poly ≈ Φ

φ_∃poly = ∅;
for (i = 1 to r) do
  for (j = 1 to n_i) do A_{x_{i,j}} = 1;
for (i = 1 to r) do {
  for (j = 1 to n_i) do {
    A_{x_{i,j}} = 0;
    for (k = i to r) do
      for (l = 1 to m_k) do y'_{k,l} = new ∃-var;
    φ_∃poly = φ_∃poly ∪ φ[x/A_x, y/y']; // (*)
    A_{x_{i,j}} = 1;
  }
  for (l = 1 to m_i) do y'_{i,l} = new ∃-var;
}
φ_∃poly = φ_∃poly ∪ φ[x/A_x, y/y']; // (*)
```

The lines marked with (*) need time $O(|\Phi|)$. They are executed $n_1 + ... + n_r + 1 = |\forall| + 1$ times, so the algorithm in total requires time $O(|\forall| \cdot |\Phi|)$.

3.4 Application: Satisfiability Testing

Let $\Phi(\mathbf{z}) \in \exists HORN^*$ be an existentially quantified Horn formula of the form $\Phi(\mathbf{z}) = \exists y_1 ... \exists y_m\, \phi(y_1, ..., y_m, \mathbf{z})$. Then $\Phi(\mathbf{z})$ is satisfiable if and only if its matrix $\phi(y_1, ..., y_m, \mathbf{z})$ is satisfiable. The latter is a purely propositional formula, therefore a SAT solver for propositional Horn formulas can be used to determine the satisfiability of an arbitrary formula in $\exists HORN^*$.

That observation leads to the following algorithm for determining the satisfiability of a formula $\Psi \in QHORN^*$:

1. Transform Ψ into $\Psi_{\exists\text{poly}} \in \exists HORN^*$ with $|\Psi_{\exists\text{poly}}| = O(|\forall| \cdot |\Psi|)$. This requires time $O(|\forall| \cdot |\Psi|)$ as discussed in Section 3.3.

2. Determine the satisfiability of $\psi_{\exists\text{poly}}$, which is the purely propositional matrix of $\Psi_{\exists\text{poly}}$. It is well known (see [4]) that SAT for propositional Horn formulas can be solved in linear time, in this case $O(|\psi_{\exists\text{poly}}|) = O(|\forall| \cdot |\Psi|)$.

In total, the algorithm requires time $O(|\forall| \cdot |\Psi|)$. The best existing algorithm presented in [5] has the same complexity, but is significantly more complicated.

4 Conclusion and Outlook

In this paper, we have given two new properties of $QHORN^*$ formulas:
(1) they can easily be transformed into equivalent purely existentially quantified formulas of length $O(|\forall| \cdot |\Phi|)$, (2) they have monotone equivalence models. Both properties characterize $QHORN^*$ as a rather simple subclass of QBF^* formulas.

Based on these results, a new algorithm for determining the satisfiability of $QHORN^*$ formulas in time $O(|\forall| \cdot |\Phi|)$ has been presented. Further investigation is needed to determine whether the simplicity of this new algorithm also translates to better real-world performance than the previously known algorithm.

References

[1] B. Aspvall, M. Plass, and R. Tarjan. *A Linear-Time Algorithm for Testing the Truth of Certain Quantified Boolean Formulas.* Information Processing Letters, 8(3):121–123, 1979.

[2] A. Biere. *Resolve and Expand.* Proc. 7th Intl. Conf. on Theory and Applications of Satisfiability Testing (SAT'04). Springer LNCS, 2004.

[3] M. Cadoli, A. Giovanardi, and M. Schaerf. *An algorithm to evaluate quantified boolean formulae.* Proc. 16th Natl. Conf. on AI (AAAI-98), 1998.

[4] W. Dowling and J. Gallier. *Linear-Time Algorithms for Testing the Satisfiability of Propositional Horn Formulae.* J. of Logic Programming, 1(3):267–284, 1984.

[5] H. Kleine Büning and T. Lettmann. *Propositional Logic: Deduction and Algorithms.* Cambridge University Press, Cambridge, UK, 1999.

[6] H. Kleine Büning, K. Subramani, and X. Zhao. *On Boolean Models for Quantified Boolean Horn Formulas.* Proc. 6th Intl. Conf. on Theory and Applications of Satisfiability Testing (SAT'03). Springer LNCS 2919 93–104, 2004.

[7] H. Kleine Büning and X. Zhao. *Equivalence Models for Quantified Boolean Horn Formulas.* Proc. 7th Intl. Conf. on Theory and Applications of Satisfiability Testing (SAT'04). Springer LNCS, 2004.

[8] D. Le Berre, M. Narizzano, L. Simon, and A. Tacchella. *The second QBF solvers comparative evaluation.* Proc. 7th Intl. Conf. on Theory and Applications of Satisfiability Testing (SAT'04). Springer LNCS, 2004.

[9] R. Letz. *Advances in Decision Procedures for Quantified Boolean Formulas.* Proc. IJCAR 2001 Workshop on Theory and Applications of QBF, 2001.

[10] J. Rintanen. *Partial implicit unfolding in the Davis-Putnam procedure for quantified boolean formulae.* Intl. Conf. on Logic for Programming, Artificial Intelligence and Reasoning (LPAR'01), 2001.