Reasoning Engines for Rigorous System Engineering

Block 3: Quantified Boolean Formulas and DepQBF

1. Basics of Quantified Boolean Formulas

Uwe Egly Florian Lonsing

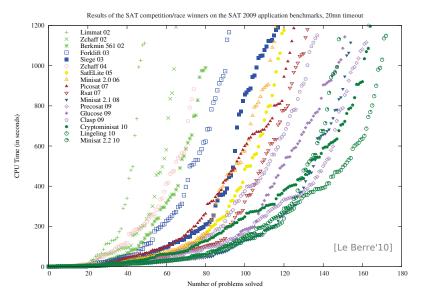
Knowledge-Based Systems Group Institute of Information Systems Vienna University of Technology



- I. Basics of Quantified Boolean Formulas
- II. Basic Deduction Concepts for Quantified Boolean Formulas
- III. Inside Search-Based QBF Solvers
- IV. DepQBF in Practice

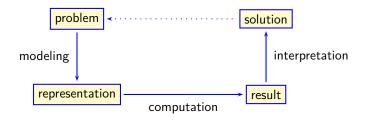
Results of the SAT 2009 application benchmarks

for leading solvers from 2002 to 2010



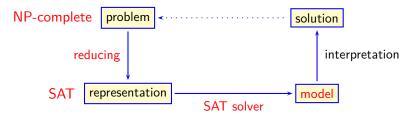
Success story of SAT: Why is it important?

Allows us to implement problem solving programs rapidly



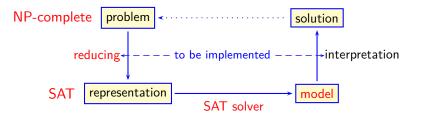
Success story of SAT: Why is it important?

Allows us to implement problem solving programs rapidly



Success story of SAT: Why is it important?

Allows us to implement problem solving programs rapidly

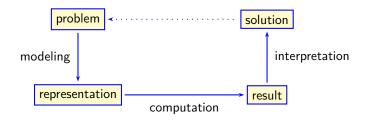


What if my problem is more difficult than SAT?

- We know how to implement solvers for NP-complete problems, e.g., planning, SAT for some equational logics, ...
- Prototypical implementation: reduce problem to a SAT problem and solve it with a "good" SAT solver
- Problem: What happens if the problem is to hard to be efficiently (polynomially) reduced to SAT?
- Solution: Use a more "expressive SAT problem" based on Quantified Boolean Formulas (QBFs)
- QBFs admit Boolean quantifiers in formulas and enable succinct problem representations for problems "harder than NP"

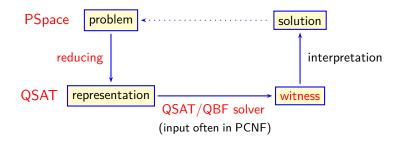
The lazy programmer's approach again

Allows us to implement problem solving programs rapidly



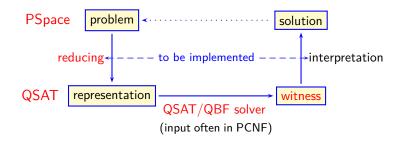
The lazy programmer's approach again

Allows us to implement problem solving programs rapidly



The lazy programmer's approach again

Allows us to implement problem solving programs rapidly



The plan

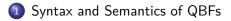
We discuss in the following

- the representation language ("QBFs") and some of its properties (like syntax and semantics),
- the concept of a witness and
- Ithe translation of representations to inputs of solvers.

Later in the course, we learn

- how we can reason using QBFs,
- how DepQBF works internally (using some of the reasoning methods), and
- how you can use it and even integrate it into your work-flow of problem solving.

Outline





- 3 Normal Form Translation for QBFs
- 4 Compact Representation with QBFs

Syntax of Quantified Boolean Formulas (QBFs) The language $\mathcal{L}_{\mathcal{P}}^{\text{pre}}$ of prenex QBFs

The simplest possibility to define QBFs (wrt Boolean variables \mathcal{P}) is:

- B1 Given a propositional formula φ over \mathcal{P} . Then $\varphi \in \mathcal{L}_{\mathcal{P}}^{pre}$.
- S1 If $\Phi \in \mathcal{L}_{\mathcal{P}}^{pre}$, then $Q_{\boldsymbol{p}} \Phi \in \mathcal{L}_{\mathcal{P}}^{pre}$, where $Q \in \{\forall, \exists\}$ and $\boldsymbol{p} \in \boldsymbol{\mathcal{P}}$.
- ▶ If there are quantifiers in a formula, then they occur at the beginning.

Example

Let φ be the propositional formula $(p \to q) \to r$ over propositional variables p, q, r. E. g., $\forall p \varphi \in \mathcal{L}_{\mathcal{P}}^{pre}$ has free variables q and r. An example for a closed formula (= without free variables) is $\forall p \exists q \forall r \varphi \in \mathcal{L}_{\mathcal{P}}^{pre}$.

Syntax of Quantified Boolean Formulas (QBFs) The language $\mathcal{L}_{\mathcal{P}}$ of arbitrary QBFs

Let \mathcal{P} be a set of propositional (Boolean) variables.

Inductive definition of the set $\mathcal{L}_{\mathcal{P}}$ of arbitrary QBFs (wrt \mathcal{P}) B1: For every propositional variable $p \in \mathcal{P}, p \in \mathcal{L}_{\mathcal{P}}$. B2: For every truth constant $t \in \{\bot, \top\}, t \in \mathcal{L}_{\mathcal{P}}$. S1: If $\Phi \in \mathcal{L}_{\mathcal{P}}$, then $\neg \Phi \in \mathcal{L}_{\mathcal{P}}$. S2: If $\Phi_1 \in \mathcal{L}_{\mathcal{P}}$ and $\Phi_2 \in \mathcal{L}_{\mathcal{P}}$, then $\Phi_1 \circ \Phi_2 \in \mathcal{L}_{\mathcal{P}}$ ($\circ \in \{\land, \lor, \rightarrow\}$). S3: If $\Phi \in \mathcal{L}_{\mathcal{P}}$, then $Qp \Phi \in \mathcal{L}_{\mathcal{P}}$ ($Q \in \{\forall, \exists\}$ and $p \in \mathcal{P}$).

Further connectives like \leftrightarrow or \oplus can be defined if necessary.

Some observations and examples

Observation 1

QBFs are allowed to be in non-prenex form, i.e., quantifiers are not only allowed in an initial prefix, but also deeply inside QBFs.

Example

$$\forall p \left(\left(\exists q \, (p \land q) \right) \to \exists r \, (r \lor p) \right)$$

Observation 2

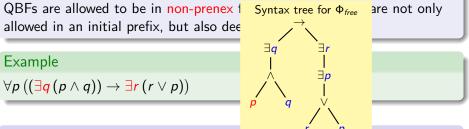
Free variables are allowed, i.e., there may be occurrences of propositional variables which have no quantification.

Example

$$\Phi_{free}: \quad (\exists q (p \land q)) \to \exists r \exists p (r \lor p)$$

Some observations and examples

Observation 1



Observation 2

Free variables are allowed, i.e., there may be occurrences of propositional variables which have no quantification.

Example

$$\Phi_{free}: \quad (\exists q (p \land q)) \to \exists r \exists p (r \lor p)$$

Normal forms

Prenex normal form (PNF), prefix, matrix, PCNF, closed

• Let
$$Q_i \in \{\forall, \exists\}$$
 and $p_i \in \mathcal{P}$. A QBF

$$\Phi = \mathsf{Q}_1 p_1 \dots \mathsf{Q}_n p_n \psi$$

is in prenex (normal) form (PNF) if ψ is purely propositional.

- $Q_1p_1Q_2p_2\cdots Q_np_n$ is the prefix of Φ ; ψ is the matrix of Φ .
- Φ is in PCNF if ψ is in CNF.
- Φ is closed if the variables in ψ are in $\{p_1, \ldots, p_n\}$.

Convention: Each quantifier binds another variable and bound variables do not occur free.

Examples for normal forms

closed, non-prenex

open, non-prenex

closed, PCNF

alternative notation 1

alternative notation 2

 $(\forall x \forall y \ (x \to y)) \land (\exists u \exists v \ (u \land v))$ $(\forall x \forall y \ (x \to y)) \land (\exists u \ (u \land v))$ $\forall x \forall y \exists z \ ((z \lor x \lor y) \land (\neg z \lor x \lor y))$ $\forall x \ y \ \exists z \ ((z \lor x \lor y) \land (\neg z \lor x \lor y))$ $\forall P \ \exists Q \ ((z \lor x \lor y) \land (\neg z \lor x \lor y))$ if $P = \{x, y\} \text{ and } Q = \{z\}$

Generating a prenex form (cf predicate logic): $\mathcal{L}_{\mathcal{P}} \mapsto \mathcal{L}_{\mathcal{P}}^{pre}$

x not free in Ψ , y not free in Φ

Apply the following rules until a PNF is obtained

- $R_1 \quad \mathsf{Q} x \, \Phi \circ \mathsf{Q} y \, \Psi \quad \Rightarrow \quad \mathsf{Q} x \mathsf{Q} y \, (\Phi \circ \Psi)$
- $R_2 \quad (\mathsf{Q} x \, \Phi) o \Psi \quad \Rightarrow \quad \mathsf{Q}^- x \, (\Phi o \Psi) \qquad x ext{ not free in } \Psi$
- $R_3 \quad \Phi \to \left(\mathsf{Q} y \, \Psi \right) \quad \Rightarrow \quad \mathsf{Q} y \left(\Phi \to \Psi \right) \qquad \quad y \text{ not free in } \Phi$
- $R_4 \quad \forall x \, \Phi \land \forall \, y \, \Psi \quad \Rightarrow \quad \forall x \, \big(\Phi \land \Psi[y/x] \big)$
- $R_5 \quad \exists x \, \Phi \lor \exists y \, \Psi \quad \Rightarrow \quad \exists x \, (\Phi \lor \Psi[y/x])$

Remarks

- $Q \in \{\forall, \exists\}$, (Q, Q^-) is (\forall, \exists) or (\exists, \forall) and $\circ \in \{\land, \lor\}$
- In general, the PNF of Φ is not unique (depends, e.g., on rule choice: R₁ vs R₄ if both are applicable)
- Φ and all of its prenex forms are logically equivalent. (Why?)

The semantics of QBFs

- Based on an interpretations / represented as a set of atoms.
- An atom p is true under I iff $p \in I$.

Inductive definition of the truth value, $\nu_I(\Phi)$, of a QBF Φ under *I*: (a) if $\Phi = \top$, then $\nu_I(\Phi) = 1$; (b) if $\Phi = p \in \mathcal{P}$, then $\nu_I(\Phi) = 1$ if $p \in I$, and $\nu_I(\Phi) = 0$ otherwise; (c) if $\Phi = \neg \Psi$, then $\nu_I(\Phi) = 1 - \nu_I(\Psi)$; (c) if $\Phi = (\Phi_1 \land \Phi_2)$, then $\nu_I(\Phi) = min(\{\nu_I(\Phi_1), \nu_I(\Phi_2)\})$; (c) if $\Phi = \forall p \Psi$, then $\nu_I(\Phi) = \nu_I(\Psi[p/\top] \land \Psi[p/\bot])$; (c) if $\Phi = \exists p \Psi$, then $\nu_I(\Phi) = \nu_I(\Psi[p/\top] \lor \Psi[p/\bot])$.

Truth conditions for \bot , \lor , \rightarrow , \leftrightarrow follow from the above "as usual".

The semantics of QBFs (cont'd)

Notations

- Φ is true under *I* iff $\nu_I(\Phi) = 1$; otherwise Φ is false under *I*.
- If $\nu_I(\Phi) = 1$, then *I* is a model of Φ (and Φ is satisfiable).
- If Φ is true under any interpretation, then Φ is valid.
- Two sets of QBFs (or ordinary Boolean formulas) are logically equivalent iff they possess the same models.

Observations

- A closed QBF is either valid or unsatisfiable, because it is either true under each interpretation *I* or false under each *I*.
- Hence, for closed QBFs, there is no need to refer to particular interpretations.

Let
$$\Phi$$
 be $\exists x ((\neg x \lor y) \land (x \lor \neg y))$ and $I = \{y\}$

$$\nu_{l}(\Phi) = \nu_{l}(\exists x ((\neg x \lor y) \land (x \lor \neg y)))$$

Let
$$\Phi$$
 be $\exists x ((\neg x \lor y) \land (x \lor \neg y))$ and $I = \{y\}$
 $\nu_I(\Phi) = \nu_I(\exists x ((\neg x \lor y) \land (x \lor \neg y)))$
 $= \nu_I((\neg \top \lor y) \land (\top \lor \neg y) \lor (\neg \bot \lor y) \land (\bot \lor \neg y))$

Let
$$\Phi$$
 be $\exists x ((\neg x \lor y) \land (x \lor \neg y))$ and $I = \{y\}$
 $\nu_{I}(\Phi) = \nu_{I}(\exists x ((\neg x \lor y) \land (x \lor \neg y)))$
 $= \nu_{I}((\neg \top \lor y) \land (\top \lor \neg y) \lor (\neg \bot \lor y) \land (\bot \lor \neg y))$
 $= \max\{\min\{\nu_{I}(\neg \top \lor y), \underbrace{\nu_{I}(\top \lor \neg y)}_{=1}\}\min\{\underbrace{\nu_{I}(\neg \bot \lor y)}_{=1}, \nu_{I}(\bot \lor \neg y)\}\}$

Let
$$\Phi$$
 be $\exists x ((\neg x \lor y) \land (x \lor \neg y))$ and $I = \{y\}$
 $\nu_{I}(\Phi) = \nu_{I}(\exists x ((\neg x \lor y) \land (x \lor \neg y)))$
 $= \nu_{I}((\neg \top \lor y) \land (\top \lor \neg y) \lor (\neg \bot \lor y) \land (\bot \lor \neg y))$
 $= \max\{\min\{\nu_{I}(\neg \top \lor y), \underbrace{\nu_{I}(\top \lor \neg y)}_{=1}\}\min\{\underbrace{\nu_{I}(\neg \bot \lor y)}_{=1}, \nu_{I}(\bot \lor \neg y)\}\}$
 $= \max\{\nu_{I}(\neg \top \lor y), \nu_{I}(\bot \lor \neg y)\}$

Let
$$\Phi$$
 be $\exists x ((\neg x \lor y) \land (x \lor \neg y))$ and $I = \{y\}$
 $\nu_{I}(\Phi) = \nu_{I}(\exists x ((\neg x \lor y) \land (x \lor \neg y)))$
 $= \nu_{I}((\neg \top \lor y) \land (\top \lor \neg y) \lor (\neg \bot \lor y) \land (\bot \lor \neg y))$
 $= \max\{\min\{\nu_{I}(\neg \top \lor y), \underbrace{\nu_{I}(\top \lor \neg y)}_{=1}\}\min\{\underbrace{\nu_{I}(\neg \bot \lor y)}_{=1}, \nu_{I}(\bot \lor \neg y)\}\}$
 $= \max\{\nu_{I}(\neg \top \lor y), \nu_{I}(\bot \lor \neg y)\}$
 $= \max\{\nu_{I}(y), \nu_{I}(\neg y)\} = 1$

- I contains (some) free variables of Φ .
- The evaluation result here is independent from *I*.
- A similar evaluation of $\forall x ((\neg x \lor y) \land (x \lor \neg y))$ results 0.

More examples of QBF evaluations

Let φ be $(p ightarrow q) \land (q ightarrow p)$

- $\exists p \exists q \varphi$ is true (since φ is sat and all its variables are bound)
- $\forall p \forall q \varphi$ is false (since φ is not valid and all its vars are bound)
- $\blacksquare \exists q \forall p \varphi \text{ is false}$
- $\forall p \exists q \varphi$ is true \blacktriangleright quantifier ordering matters!

Satisfiability and validity can be expressed in QBFs:

- $\exists V \psi(V)$ is true iff ψ is satisfiable.
- $\forall V \psi(V)$ is true iff ψ is valid.

Certificates for QBFs: an appetizer

A certificate provides evidence of satisfiability of a QBF

 One possibility to certify the truth of a closed QBF: Witness functions/formulas (WFs) for existential quantifiers which depend on (some) dominating universal quantifiers.

Example: $\forall x_1 \forall x_2 \exists y (x_1 \lor x_2 \lor \neg y) \land (\neg x_1 \lor y)$

- Solution Take $y = x_1$: $\forall x_1 \forall x_2 (x_1 \lor x_2 \lor \neg x_1) \land (\neg x_1 \lor x_1)$ becomes true.
- This can be checked with a validity checker for propositional logic.
 - WFs are sometimes the constructed solution to a problem.

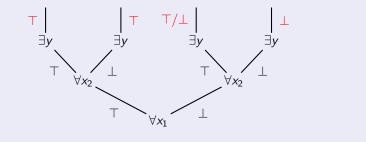
For a broader discussion, see V. Balabanov, J.-H. R. Jiang: Resolution proofs and Skolem functions in QBF evaluation and applications CAV 2011. [link]

Certificates for QBFs: an appetizer (cont'd)

A certificate provides evidence of satisfiability of a QBF

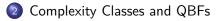
■ Others are e.g. tree-like strategies for the choice of truth values of ∃ quantifiers depending on dominating ∀ ones.

Example: $\forall x_1 \forall x_2 \exists y (x_1 \lor x_2 \lor \neg y) \land (\neg x_1 \lor y)$



Outline

Syntax and Semantics of QBFs



Normal Form Translation for QBFs



For which classes of problems do we need QBFs?

- NP-complete problems can be efficiently reduced to SAT.
- Q: Why is another SAT formalism based on QBFs needed?
- A: There are even "harder" problems than SAT.
 A Garey-Johnson like compendium of such problem can be found here [link]
- The SAT problem for QBFs provides a target formalism to which such computationally hard problems can be reduced.

Informal definition of important complexity classes

class	model of computation	expense wrt resource
Р	deterministic	polynomial time
NP	non-deterministic	polynomial time
PSPACE	deterministic	polynomial space
NPSPACE	non-deterministic	polynomial space
EXPTIME	deterministic	exponential time
NEXPTIME	non-deterministic	exponential time

Relations between some complexity classes

- $P \subseteq_{=?} NP \subseteq_{=?} PSPACE$
- PSPACE = NPSPACE
- PSPACE $\subseteq_{=?}$ EXPTIME

- $P \subset EXPTIME$
- $\blacksquare \text{ NP} \subset \text{NEXPTIME}$

The polynomial hierarchy (PH)

The PH consists of classes Σ_k^P , Π_k^P , and Δ_k^P , where

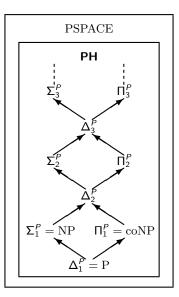
$$\Sigma_0^P = \Pi_0^P = \Delta_0^P = \mathrm{P};$$

and for $k \geq 1$:

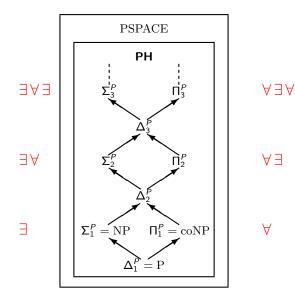
$$\begin{array}{rcl} \Delta^P_{k+1} &=& \mathrm{P}^{\Sigma^P_k};\\ \Sigma^P_{k+1} &=& \mathrm{N}\mathrm{P}^{\Sigma^P_k};\\ \Pi^P_{k+1} &=& \mathrm{co} - \Sigma^P_{k+1} \end{array}$$

 A^{B} : The set of decision problems solvable by a Turing machine in class A augmented by an oracle for some complete problem in class B.

The polynomial hierarchy (PH) (cont'd)



The polynomial hierarchy (PH) (cont'd)



Prenex QBFs and complexity classes (Wrathall 1976)

Eval. problems for prenex QBFs and their complexities

Given a propositional formula φ with its atoms partitioned into $i \ge 1$ pairwise distinct sets V_1, \ldots, V_i , deciding whether $\exists V_1 \forall V_2 \ldots Q_i V_i \varphi$ is true is \sum_i^P -complete, where $Q_i = \exists$ if i is odd and $Q_i = \forall$ if i is even, Dually, deciding whether $\forall V_1 \exists V_2 \ldots Q'_i V_i \varphi$ is true is \prod_i^P -complete, where $Q'_i = \forall$ if i is odd and $Q_i = \exists$ if i is even.

Examples of evaluation problems (EPs)

- The EP of $\exists V_1 \varphi(V_1)$ is Σ_1^P -complete (= NPC)
- The EP of $\forall V_1 \varphi(V_1)$ is Π_1^P -complete (= co-NPC)
- The EP of $\forall V_1 \exists V_2 \forall V_3 \varphi(V_1, V_2, V_3)$ is Π_3^P -complete

Important for reductions: If we know the complexity of our problem, we can choose the appropriate quantifier prefix for the target QBF.

How to handle non-prenex QBFs?

Extend the complexity landscape to arbitrary closed QBFs

- Take the maximal number of quantifier alternations along a path in the syntax tree of a QBF into account.
- Almost all QBFs can be translated into equivalent QBFs in PNF without increasing the number of quantifier alternations. (Which are the problematic QBFs?)
- The translation procedure is fast but non-deterministic, but
- ... can heavily influence the performance of QBF solvers.
- Details in E. et al. Comparing Different Prenexing Strategies for Quantified Boolean Formulas. Proc. SAT 2003, pp. 214-228.

Outline

Syntax and Semantics of QBFs







Generating prenex conjunctive normal forms (PCNFs)

A QBF in prenex conjunctive normal form (PCNF)

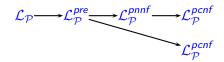
- starts with a quantifier prefix and
- consists of a conjunction of clauses (=disjunction of literals) (often represented as a set of clauses).
- Clauses are often represented as sets of literals.

Why are formulas in PCNF necessary?

- Most QBF solvers require the input being in PCNF!
- Translation procedure is required.
 - This procedure can be based on distributivity or Tseitin.

Languages for QBFs and their conversion

- $\mathcal{L}_\mathcal{P}$: arbitrary QBFs
- $\mathcal{L}_{\mathcal{P}}^{nnf}$: QBFs in negation normal form
- $\mathcal{L}_{\mathcal{P}}^{\textit{pre}}$: QBFs in prenex form with an unrestricted matrix
- $\mathcal{L}_{\mathcal{P}}^{ponf}$: QBFs in prenex form with a matrix in negation normal form $\mathcal{L}_{\mathcal{P}}^{pcnf}$: QBFs in prenex form with a matrix in conjunctive normal form $\mathcal{L}_{\mathcal{P}}^{pdnf}$: QBFs in prenex form with a matrix in disjunctive normal form



Traditional translation

Tseitin translation

Generating PCNFs (cont'd)

The Tseitin-based algorithm works in three steps:

- Generate a prenex form Ψ_p: Q_iX_i···Q_kX_kψ of the input QBF Ψ. Then the matrix ψ is purely propositional.
- 2 Use Tseitin's translation to transform ψ into CNF.
- Place the ∃ quantifiers for the newly introduced variables l₁,..., l_m abbreviating φ₁,..., φ_m "correctly", e.g.,
 - \blacksquare place all the new \exists at the end of the quantifier prefix, or
 - place ∃ℓ_i (1 ≤ i ≤ m) after all quantifiers of those variables which occur in φ_i.

Contrary to propositional logic, Ψ is logically equivalent to its PCNFs!

Outline

Syntax and Semantics of QBFs

2 Complexity Classes and QBFs

3 Normal Form Translation for QBFs



Compact Representation with QBFs

Tricky use of Boolean quantification

Trick 1: Introduce abbreviations for sub-formulas

• Given propositional formula φ :

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

Idea: Introduce a "definition" to abbreviate a ∨ ¬b ∨ c.
Obtain a QBF Φ:

 $\exists y (y \leftrightarrow a \lor \neg b \lor c) \land (y \lor d) \land (y \lor \neg e) \land (y \lor f)$

• $a \lor \neg b \lor c$ occurs only once!

• Φ is logically equivalent to φ (mainly because of $\exists y$).

Most examples from U. Bubeck, H. Kleine Büning: Encoding Nested Boolean Functions as Quantified Boolean Formulas. JSAT 8:101-116 (2012). [link]

Tricky use of Boolean quantification (cont'd)

Trick 2: "Unify" conjunctively connected instances

• Given propositional formula φ :

$$\varphi_1(\psi_1,\pi_1) \wedge \varphi_1(\psi_2,\pi_2) \wedge \varphi_1(\psi_3,\pi_3)$$

• We have three different instances of $\varphi_1(\psi, \pi)$.

Obtain a QBF Φ:

$$orall u orall v \ ig(\bigvee_{i=1}^3 ((u \leftrightarrow \psi_i) \land (v \leftrightarrow \pi_i)) ig)
ightarrow arphi_1(u,v)$$

• φ_1 occurs only once!

• Φ is logically equivalent to φ .

Tricky use of Boolean quantification (cont'd)

Trick 3: Non-copying iterative squaring

• Given formula $\Psi(x_0, x_n)$ with $n = 2^i$:

$$\exists x_1 \cdots \exists x_{n-1} (\varphi(x_0, x_2) \land \varphi(x_2, x_3) \land \cdots \land \varphi(x_{n-1}, x_n))$$

Idea: Take y in the middle and split the formula:

$$\Psi_{2^{i}}(x_{0}, x_{n}) \colon \exists y (\Psi_{2^{i-1}}(x_{0}, y) \land \Psi_{2^{i-1}}(y, x_{n}))$$

• Use Trick 2 and get
$$\Psi_{2^i}(x_0, x_n)$$
:

 $\exists y \forall u \forall v \left[\left(\left((u, v) \leftrightarrow (x_0, y) \right) \lor \left((u, v) \leftrightarrow (y, x_n) \right) \right) \rightarrow \Psi_{2^{i-1}}(u, v) \right]$

This can be used to model bounded model checking.

- All problems from the polynomial hierarchy can be handled by the "lazy programmer approach".
- Complexity results for the original problem provide appropriate quantifier alternations for the target QBF.
- QBFs have to be translated into the input format of QBF solvers.

Now we can start to reason with QBFs.