Proof Complexity of Quantified Boolean Formulas

Olaf Beyersdorff

Friedrich Schiller University Jena, Germany

Oaxaca, August 2018

Quantified Boolean Formulas (QBF)

What's different in QBF from propositional proof complexity?

Quantified Boolean Formulas (QBF)

What's different in QBF from propositional proof complexity?

- Quantification!

Quantified Boolean Formulas (QBF)

What's different in QBF from propositional proof complexity?

- Quantification!
- Boolean quantifiers ranging over 0/1

Quantified Boolean Formulas (QBF)

What's different in QBF from propositional proof complexity?

- Quantification!
- Boolean quantifiers ranging over $0/1$

Why QBF proof complexity?

- driven by QBF solving
- shows different effects from propositional proof complexity
- connects to circuit complexity, bounded arithmetic, ...

QBF proof complexity vs solving

Impact for proof complexity

different resolution systems defined that capture ideas in solving:

- CDCL
- expansion of universal variables
- dependency schemes

Impact for solving

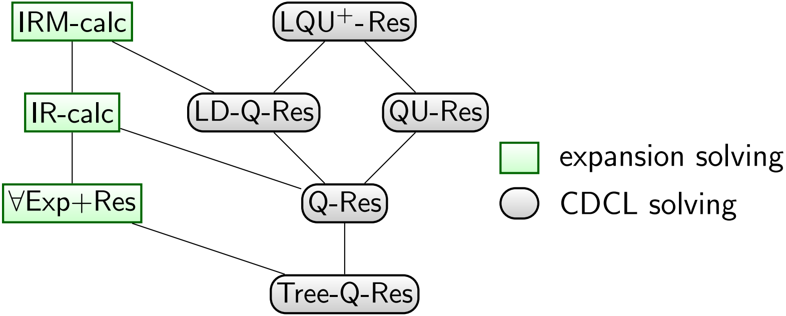
- proves soundness of new algorithmic approaches
- upper/lower bounds suggest new directions in solving

Interesting test case for algorithmic progress

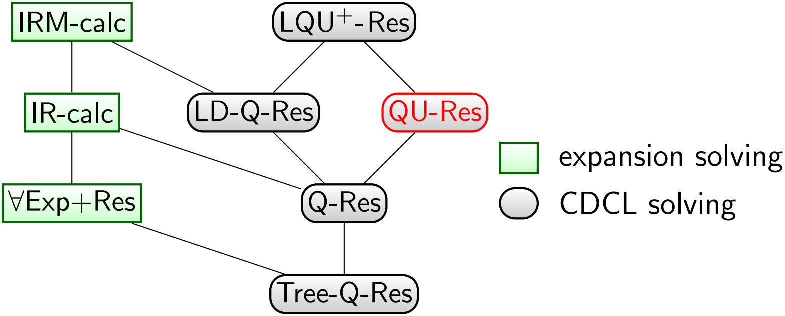
SAT revolution

SAT	NP	main breakthrough late 90s
QBF	PSPACE	reaching industrial applicability now
DQBF	NEXPTIME	very early stage

QBF resolution systems



QBF resolution systems



A core system: QU-Resolution

$=$ Resolution $+$ $\forall$-reduction [Kleine Büning et al. 95, V. Gelder 12]

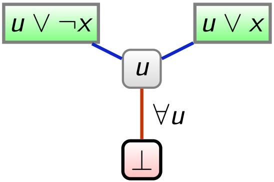
Rules

- Resolution: $$\frac{x \vee C \qquad \neg x \vee D}{C \vee D}$$ ($C \vee D$ is not tautological.)
- $\forall$-Reduction: $$\frac{C \vee u}{C}$$ (u universally quantified)

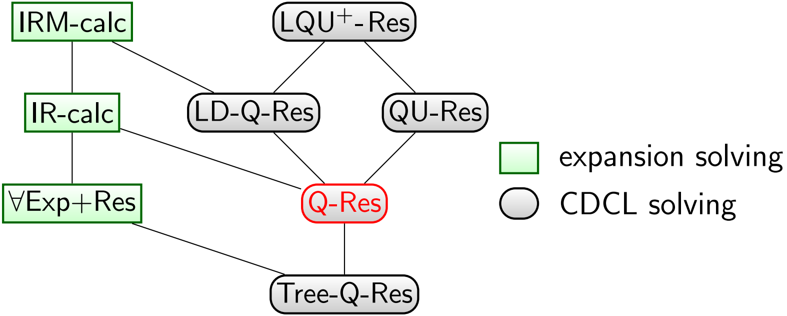
  C does not contain variables right of u in the quantifier prefix.

Example

$\forall \mathbf{u} \exists \mathbf{x}$



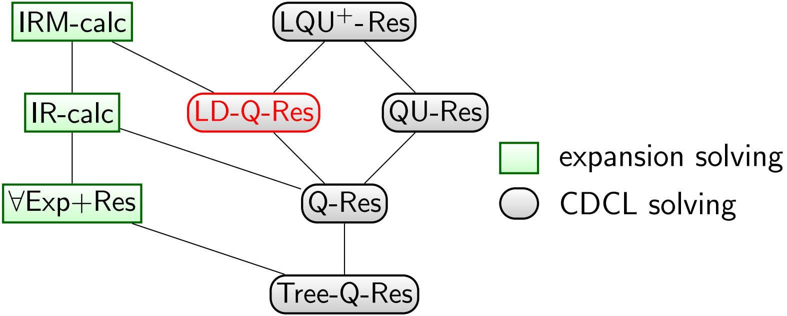
Further systems at a glance



Q-resolution (Q-Res)

- pivots in resolution rule must be existential
- otherwise same rules as QU-Res
- first QBF resolution system [Kleine Büning et al. 95]

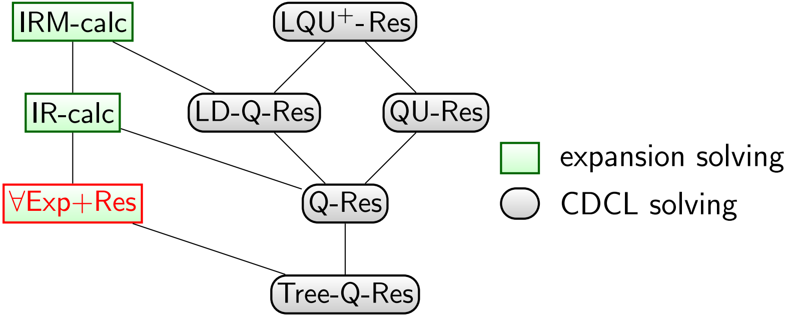
Further systems at a glance



Long-distance resolution (LD-Q-Res)

- allows certain resolution steps forbidden in Q-Res
- merges universal literals u and $\neg u$ in a clause to u^*
- introduced by [Zhang & Malik 02] [Balabanov & Jiang 12]

Expansion based calculi



∀Exp+Res

- expands universal variables (for one or both values 0/1)
- introduced by [Janota & Marques-Silva 13]

$\forall$Exp+Res

Annotated literals

couple together existential and universal literals: l^{α}, where

- l is an existential literal.
- α is a partial assignment to universal literals.

Rules of $\forall$Exp+Res

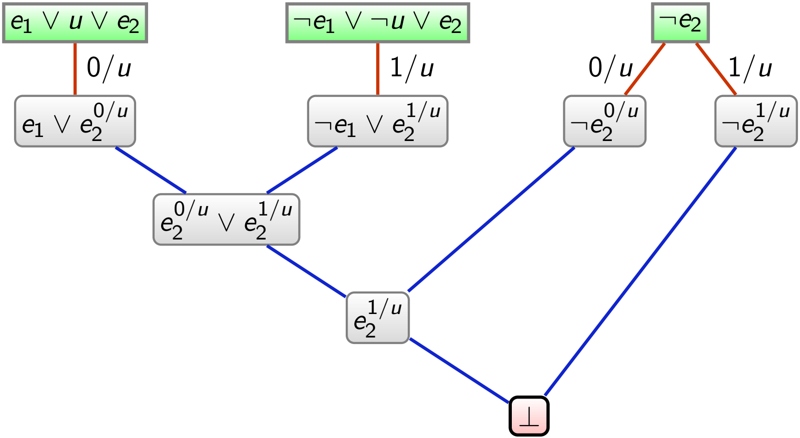
$$\frac{C \text{ in matrix}}{\left\{ l^{[\tau]} \mid l \in C, l \text{ is existential} \right\}} \text{ (Axiom)}$$

- τ is a complete assignment to universal variables that falsifies all universal literal in C.
- $[\tau]$ restricts τ to variables left of l in the prefix.

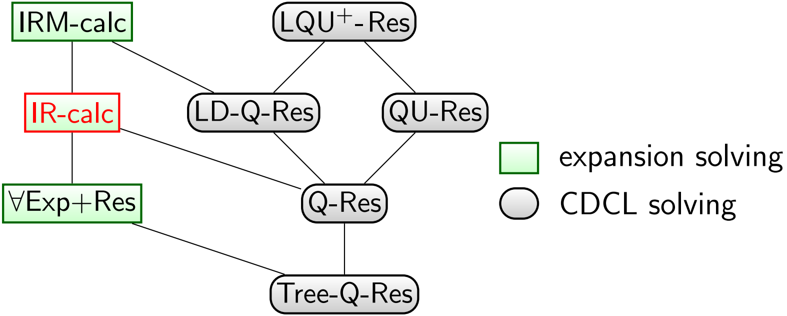
$$\frac{x^{\tau} \vee C_1 \qquad \neg x^{\tau} \vee C_2}{C_1 \cup C_2} \text{ (Resolution)}$$

Example proof in ∀Exp+Res

$\exists \mathbf{e_1} \forall \mathbf{u} \exists \mathbf{e_2}$



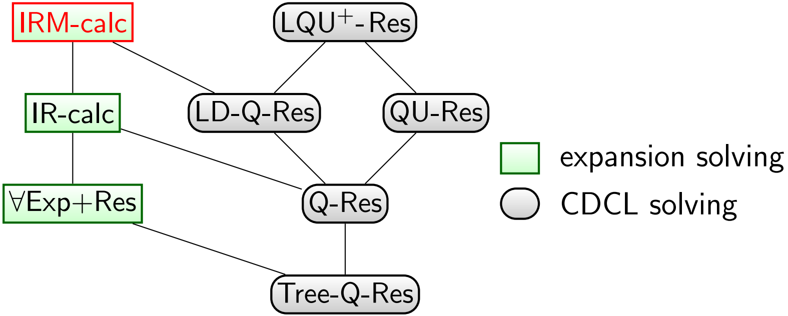
Further expansion-based systems at a glance



IR-calc

- Instantiation + Resolution
- ‘delayed’ expansion
- introduced by [B., Chew, Janota 14]

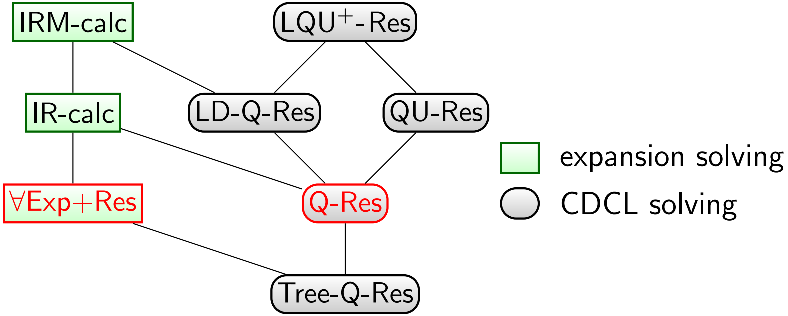
Further expansion-based systems at a glance



IRM-calc

- Instantiation + Resolution + Merging
- allows merged universal literals u^*
- introduced by [B., Chew, Janota 14]

CDCL vs expansion systems



- Q-Res and ∀Exp+Res are incomparable.
- But tree ∀Exp+Res simulates tree Q-Res (and is stronger). [Janota & Marques-Silva 15]
- ∀Exp+Res even simulates Q-Res on QBFs of bounded quantifier complexity. [Clymo 18]

CDCL vs expansion systems

Q-Res and $\forall$Exp+Res are incomparable

1. construct formulas that are easy in Q-Res, but require exponentially many expansions of universal variables [Janota & Marques-Silva 15]
2. construct Parity formulas hard in Q-Res, but easy in $\forall$Exp+Res

- uses the concept of strategy extraction

Strategy extraction

Game semantics of QBF

- $\exists$ and $\forall$ assign variables in order of the prefix.
- $\forall$ wins if a clause falsifies, otherwise $\exists$ wins.
- $\forall$ has a winning strategy iff the QBF is false.

Strategy extraction

- in QBF solving: return true/false + a strategy for $\exists/\forall$, witnessing the answer.
- for QBF calculi: given a refutation of a false QBF, compute a winning strategy for $\forall$

Strategy extraction for QBF calculi

$\forall$ winning strategies can be efficiently extracted

- in polynomial time for all QBF resolution systems
- in AC^0 for QU-Res and Q-Res

Lower bound idea

- Construct false QBFs without easily computable winning strategies
- These formulas must have large proofs.

Hard formulas for QU-Res

- Let ϕ_n be a propositional formula computing $x_1 \oplus \ldots \oplus x_n$.
- Consider the QBF $\exists x_1, \ldots, x_n \forall z. (z \vee \phi_n) \wedge (\neg z \vee \neg \phi_n)$.
- The matrix of this QBF states that z is equivalent to the opposite value of $x_1 \oplus \ldots \oplus x_n$.
- The unique strategy for the universal player is therefore to play z equal to $x_1 \oplus \ldots \oplus x_n$.

Defining ϕ_n

- Let $\mathsf{xor}(o_1, o_2, o)$ be the set of clauses $\{\neg o_1 \vee \neg o_2 \vee \neg o,\ o_1 \vee o_2 \vee \neg o,\ \neg o_1 \vee o_2 \vee o,\ o_1 \vee \neg o_2 \vee o\}$.
- Define

$$\text{QPARITY}_n = \exists x_1, \ldots, x_n \forall z \exists t_2, \ldots, t_n.\ \mathsf{xor}(x_1, x_2, t_2) \cup \bigcup_{i=3}^{n} \mathsf{xor}(t_{i-1}, x_i, t_i) \cup \{z \vee t_n, \neg z \vee \neg t_n\}$$

The exponential lower bound

$$\text{QPARITY}_n = \exists x_1, \ldots, x_n \forall z \exists t_2, \ldots, t_n.\ \text{xor}(x_1, x_2, t_2) \cup \bigcup_{i=3}^{n} \text{xor}(t_{i-1}, x_i, t_i) \cup \{z \vee t_n, \neg z \vee \neg t_n\}$$

Theorem (B., Chew & Janota 15)

QPARITY_n *require exponential-size QU-Res refutations.*

Proof idea

- By [Balabanov & Jiang 12] we extract strategies from any Q-Res proof as bounded-depth circuits in polynomial time.
- But $\text{PARITY}(x_1, \ldots x_n)$ requires exponential-size bounded-depth circuits [Håstad 87].
- Therefore QU-Res proofs must be of exponential size. □

Beyond QBF Resolution

So far we looked at QBF Resolution systems

- What about Cutting Planes, Polynomial Calculus, Frege etc.?
- Can we find stronger calculi that still have strategy extraction?

From propositional proof systems to QBF

A general ∀red rule

- Fix a prenex QBF Φ.
- Let $F(\bar{x}, u)$ be a propositional line in a refutation of Φ, where u is universal with innermost quant. level in F

$$\frac{F(\bar{x}, u)}{F(\bar{x}, 0)} \qquad \frac{F(\bar{x}, u)}{F(\bar{x}, 1)}$$

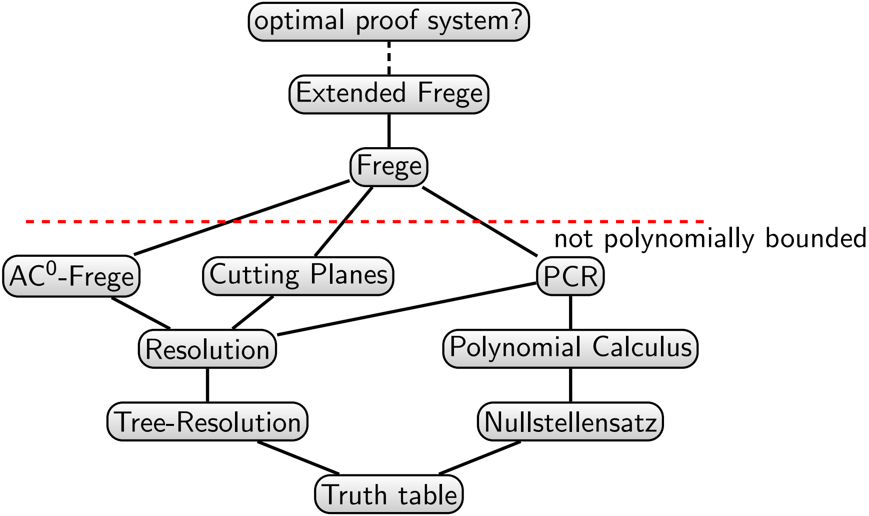
New QBF proof systems

For any 'natural' line-based propositional proof system P define the QBF proof system $Q\text{-}P$ by adding ∀red to the rules of P.

Proposition (B., Bonacina & Chew 16)

$Q\text{-}P$ is sound and complete for QBF.

Important propositional proof systems



Do we get strategy extraction?

QBF systems with efficient strategy extraction

- all QBF Resolution systems
- Q-CP, Q-PC
- Q-Frege, Q-EF

General proof checking format

- QRAT [Heule, Seidl, Biere 14]

Stronger systems without strategy extraction

- sequent systems G_0, G_1, ...
- formulas not necessarily prenex

Which lower bound techniques apply?

Techniques for propositional proof systems

- size-width relation [Ben-Sasson & Wigderson 01]
- feasible interpolation [Krajíček 97]
- game-theoretic techniques [Pudlák, Buss, Impagliazzo,. . .]
- proof complexity generators [Krajíček, Alekhnovich et al.]

In QBF proof systems

- size-width relations fail for QBF resolution systems [B., Chew, Mahajan, Shukla 16]
- feasible interpolation holds for QBF resolution systems [B., Chew, Mahajan, Shukla 17]
- game-theoretic techniques work for weak tree-like systems [B., Chew, Sreenivasaiah 17] [Chen 16]

Genuine QBF lower bounds

Propositional hardness transfers to QBF

- If $\phi_n(\vec{x})$ is hard for P, then $\exists\vec{x}\,\phi_n(\vec{x})$ is hard for $Q\text{-}P$.
- propositional hardness: not the phenomenon we want to study.

Genuine QBF hardness

- in $Q\text{-}P$: just count the number of $\forall$red steps
- can be modelled precisely by allowing NP oracles in QBF proofs [Chen 16; B., Hinde & Pich 17]

QBF systems with only genuine lower bounds

A relaxation of a quantifier prefix

- can turn $\forall$ into $\exists$
- move $\forall$ to the left

The QBF system $Q\text{-}P^{\Sigma_k^p}$ has the rules:

- of the propositional system P
- $\forall$-reduction
- $\dfrac{C_1 \ \ldots \ C_l}{D}$ for any l,
  where the quantifier prefix Π is relaxed to a Σ_k^b-prefix Π' such that $\Pi'.\bigwedge_{i=1}^{l} C_i \models \Pi'.D \wedge \bigwedge_{i=1}^{l} C_i$

Genuine hardness results

Theorem [B., Hinde, Pich 17]

- For every odd k there exist QBFs that are easy in $Q\text{-}Res^{\Sigma_k^p}$, but require exponential-size proofs in $Q\text{-}Res^{\Sigma_{k-1}^p}$.
- There exist QBFs that require exponential-size proofs in $Q\text{-}Res^{\Sigma_k^p}$ for all k.

Theorem [B., Blinkhorn, Hinde 18]

Random QBFs (in a suitable random model) require exponential-size proofs in $Q\text{-}Res^{\mathsf{NP}}$, $Q\text{-}CP^{\mathsf{NP}}$ and $Q\text{-}PC^{\mathsf{NP}}$.

Theorem [B., Bonacina, Chew 16]

There exist QBFs that require exponential-size proofs in $Q\text{-}\mathsf{AC}^0[p]\text{-}Frege^{\mathsf{NP}}$.

Characterisations

Theorem [B. & Pich 16]

- super-polynomial lower bounds for *Q-Frege*$^{\mathsf{NP}}$ iff PSPACE $\not\subseteq$ NC^1
- super-polynomial lower bounds for *Q-EF*$^{\mathsf{NP}}$ iff PSPACE $\not\subseteq$ P/poly

A new lower bound technique

Semantic lower bound technique for QBF

- applies to all QBF systems of the form Q-P
- measures the complexity of strategies

Response map

A response map R for a proof system Q-P is a function

$$R : (L, \alpha) \mapsto \beta \qquad \text{where}$$

- L is a line in Q-P
- α is a total assignment to the existential variables of L
- β is a total assignment to the universal variables in L

such that if $L|_\alpha$ is not a tautology, then $L|_{\alpha \cup \beta}$ is false.

Example: Resolution

- lines are clauses, e.g. $L = \underbrace{x_1 \vee \neg x_2}_{\textit{existential}} \vee \underbrace{u_1 \vee u_2}_{\textit{universal}}$
- map (L, α) to $(u_1/0, u_2/0)$.
- Response is independent of α.

Strategy extraction algorithm

Round-based strategy extraction

- Fix a response map R for $Q\text{-}P$.
- Let π a $Q\text{-}P$ refutation for $\Phi = \exists E_1 \forall U_1 \cdots \exists E_n \forall U_n\ \phi$.
- $\exists$ player chooses an assignment α_1 for E_1.
- $\forall$ player searches for the first line L in π which only contains variables from $E_1 \cup U_1$ and is not a tautology under α_1.
- $\forall$ responds by $R(L, \alpha_1)$.
- iteratively continue with E_2, U_2 ...

The cost of strategies

Definition

- Fix a winning strategy S for a QBF Φ and consider the size of its range (in each universal block).
- The cost of Φ is the minimum of this range size over all winning strategies.

Intuition

Strategies that require many responses of the universal player (in one block) are costly.

Example

Equality formulas

$$\exists x_1 \cdots x_n \forall u_1 \cdots u_n \exists t_1 \cdots t_n$$
$$\left(\bigwedge_{i=1}^{n} (x_i \vee u_i \vee \neg t_i) \wedge (\neg x_i \vee \neg u_i \vee \neg t_i) \right) \wedge \left(\bigvee_{i=1}^{n} t_i \right).$$

- The only winning strategy for these formulas is $u_i = x_i$ for $i = 1, \ldots, n$.
- The cost (=size of the range of the winning strategy) is 2^n.

Capacity

Capacity of lines and proofs

- Let L be a line in $Q\text{-}P$.
- The capacity of a line L is the size of the minimal range of $R(L, .)$ over all response maps R for $Q\text{-}P$.
- The capacity of a $Q\text{-}P$ proof is the maximum of the capacity of its lines.

Example

- Clauses have capacity 1 (require only one response).
- Resolution proofs have always capacity 1.

The central connection

The Size-Cost-Capacity Theorem [B., Blinkhorn, Hinde 18]

For each $Q\text{-}P^{\mathsf{NP}}$ proof π of a QBF ϕ we have

$$|\pi| \geq \frac{cost(\phi)}{capacity(\pi)}.$$

Example: Equality formulas in resolution

$\exists x_1 \cdots x_n \forall u_1 \cdots u_n \exists t_1 \cdots t_n$
$[\bigwedge_{i=1}^{n}(x_i \vee u_i \vee \neg t_i) \wedge (\neg x_i \vee \neg u_i \vee \neg t_i)] \wedge \bigvee_{i=1}^{n} t_i$

- cost $= 2^n$
- capacity $= 1$
- $\Rightarrow$ proofs in *Q-Res* are of size 2^n.

The central connection

The Size-Cost-Capacity Theorem [B., Blinkhorn, Hinde 18]

For each $Q\text{-}P^{\mathsf{NP}}$ proof π of a QBF ϕ we have

$$|\pi| \geq \frac{\mathit{cost}(\phi)}{\mathit{capacity}(\pi)}.$$

Intuition on the proof

- cost counts the number of necessary responses of universal winning strategies
- these can be extracted from the proof (by the round-based strategy extraction algorithm)
- capacity gives an upper bound on how many responses can be extracted per line

The central connection

The Size-Cost-Capacity Theorem [B., Blinkhorn, Hinde 18]

For each $Q\text{-}P^{\mathsf{NP}}$ proof π of a QBF ϕ we have

$$|\pi| \geq \frac{cost(\phi)}{capacity(\pi)}.$$

Remarks

- lower bound technique with semantic flavour
- works for all base systems P (under very mild assumptions)
- always produces 'genuine' QBF lower bounds on the number of $\forall$-reduction steps

In other QBF systems

Cutting planes

- capacity of lines is still 1
- the best response for a line

$$\underbrace{a_1 x_1 + \ldots a_m x_m}_{\textit{existential}} + \underbrace{b_1 u_1 + \ldots b_n u_n}_{\textit{universal}} \geq C$$

is to play $u_i = 0$ if $b_i > 0$ and 1 otherwise

Corollaries

- For each $Q\text{-}CP$ proof π of a QBF ϕ we have $|\pi| \geq cost(\phi)$.
- Equality formulas require $Q\text{-}CP$ proofs of size 2^n.

Polynomial Calculus (with Resolution)

Capacity is non-constant

- consider $x(1-u)+(1-x)u=0$
- winning strategy is $u = 1-x$.
- requires 2 responses, hence capacity of the line is 2.

Lemma

If π is a *Q-PC* proof where each line contains at most M monomials, then $\mathit{capacity}(\pi) \leq M$.

Corollary

For each *Q-PC* proof π of a QBF ϕ we have $|\pi| \geq \sqrt{\mathit{cost}(\phi)}$.

Frege

Capacity can be exponential

- Consider $\bigvee_{i=1}^{n} [(x_i \vee u_i) \wedge (\neg x_i \vee \neg u_i)]$.
- The unique winning response is to play $u_i = x_i$ for all $i \in [n]$.
- Capacity of this line is 2^n.

Proposition

Equality formulas are easy in *Q-Frege* .

Application: Hard random formulas in QBF

Random QBFs

- Pick clauses $C_i^1, \ldots, C_i^{cn}$ uniformly at random
- for each C_i^j choose 1 literal from the set $X_i = \{x_i^1, \ldots, x_i^m\}$ and 2 literals from $Y_i = \{y_i^1, \ldots, y_i^n\}$.
- Define $Q(n, m, c)$ as

$$\exists Y_1 \ldots Y_n \forall X_1 \ldots X_n \exists t_1 \ldots t_n. \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{cn} \left(\neg t_i \vee C_i^j \right) \wedge \bigvee_{i=1}^{n} t_i$$

Remarks

- All clauses contain existential and universal literals.
- Rightmost quantifier block is existential.

Hardness of the random QBFs

$$Q(n,m,c) = \exists Y_1 \dots Y_n \forall X_1 \dots X_n \exists t_1 \dots t_n . \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{cn} \left(\neg t_i \vee C_i^j\right) \wedge \bigvee_{i=1}^{n} t_i$$

Theorem [B., Blinkhorn, Hinde 18]

Let $1 < c < 2$ and $m \leq (1-\epsilon)\log_2(n)$ for some $\epsilon > 0$.
With high probability, $Q(n,m,c)$ is false and requires size $2^{\Omega(n^\epsilon)}$ in QU-Resolution, $Q\text{-}CP$, and $Q\text{-}PCR$.

Proof idea

$Q(n,m,c)$ is false iff all QBFs $\Psi_i = \exists Y_i \forall X_i \bigwedge_{j=1}^{cn} C_i^j$ are false.

1. Show that Ψ_i is false whp.
2. Show that Ψ_i requires non-constant winning strategies whp.

Dependency schemes

In QBF

- quantifier prefix is linear.
- in a prefix $\forall u \ldots \exists x$ the variable x depends on u
- relevant e.g. for universal reduction: in $\{u, x\}$ we cannot reduce u

Idea of dependency schemes

- use preprocessing to identify spurious dependency schemes
- computes a binary relation indicating variable (in)dependence
- adapt rules to the new dependency relation
- e.g. allow u to be reduced in clause C if no literal in C depends on u

Dependency schemes in solving and proof complexity

Identify dependency schemes

- standard dependency scheme, used in DepQBF
  [Samer & Szeider 09], [Lonsing & Egly 17]
- reflexive resolution path dependency scheme
  [Slivovsky & Szeider 14]

Soundness

- provide calculi modelling dependency-aware solving
  [Peitl, Slivovsky & Szeider 16], [Slivovsky & Szeider 16]
- establish general conditions for soundness of the resulting systems [B. & Blinkhorn 16]

Dependency schemes in solving and proof complexity

Advantages of dependency schemes [Blinkhorn & B. 17]

- first separation of Q-Resolution and Q-Resolution with dependency schemes
- show further advantage of dependency schemes when used as inprocessing scheme
- applies to further systems, including expansion

Conclusion

- QBF vs propositional proof complexity: different picture
- New semantic QBF technique, based on strategy extraction
- Yields genuine QBF lower bounds

Challenges in QBF proof complexity

- Characterise reasons for hardness in QBF Resolution
- Find more hard QBF families
- Understand randomness in QBF
- Quantify gain from dependency solving
- Model precisely QBF solving and guide developments