Representations of Monotone Boolean Functions by Linear Programs

Mateus de Oliveira Oliveira[1], Pavel Pudlák[2]

[1]University of Bergen
[2]Czech Academy of Sciences

Work financed by the European Research Council, project FEALORA.
[1] Acknowledges support from the Bergen Research Foundation

July 10, 2017

MLP Gates

Weak MLP Gates

$A \in \mathbb{R}^{m \times k}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^k$, B, C nonnegative matrices in $\mathbb{R}^{m \times n}$.

Weak MLP Gates

$A \in \mathbb{R}^{m\times k}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^k$, B, C nonnegative matrices in $\mathbb{R}^{m\times n}$.

- MAX-RIGHT: $\ell(y) = \max\{c^T \cdot x \mid Ax \le b + By,\ x \ge 0\}$

Weak MLP Gates

$A \in \mathbb{R}^{m \times k}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^k$, B, C nonnegative matrices in $\mathbb{R}^{m \times n}$.

- MAX-RIGHT: $\ell(y) = \max\{c^T \cdot x \mid Ax \leq b + By,\ x \geq 0\}$
- MIN-RIGHT: $\ell(y) = \min\{c^T \cdot x \mid Ax \geq b + By,\ x \geq 0\}$

Weak MLP Gates

$A \in \mathbb{R}^{m \times k}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^k$, B, C nonnegative matrices in $\mathbb{R}^{m \times n}$.

- MAX-RIGHT: $\ell(y) = \max\{c^T \cdot x \mid Ax \leq b + By,\ x \geq 0\}$
- MIN-RIGHT: $\ell(y) = \min\{c^T \cdot x \mid Ax \geq b + By,\ x \geq 0\}$
- MAX-LEFT: $\ell(y) = \max\{(c + Cy)^T \cdot x \mid Ax \leq b,\ x \geq 0\}$

Weak MLP Gates

$A \in \mathbb{R}^{m \times k}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^k$, B, C nonnegative matrices in $\mathbb{R}^{m \times n}$.

- MAX-RIGHT: $\ell(y) = \max\{c^T \cdot x \mid Ax \leq b + By,\ x \geq 0\}$
- MIN-RIGHT: $\ell(y) = \min\{c^T \cdot x \mid Ax \geq b + By,\ x \geq 0\}$
- MAX-LEFT: $\ell(y) = \max\{(c + Cy)^T \cdot x \mid Ax \leq b,\ x \geq 0\}$
- MIN-LEFT: $\ell(y) = \min\{(c + Cy)^T \cdot x \mid Ax \geq b,\ x \geq 0\}$

Strong MLP Gates

MAX: $\ell(y) = \max\{(c + Cy)^T \cdot x \mid Ax \leq b + By,\ x \geq 0\}$

MIN: $\ell(y) = \min\{(c + Cy)^T \cdot x \mid Ax \geq b + By,\ x \geq 0\}$

Definition (MLP-Circuit Representation)

We say that an MLP circuit C represents a partial Boolean function $F : \{0,1\}^n \to \{0,1,*\}$ if the following conditions are satisfied for each $a \in \{0,1\}^n$.

1. $C(a) > 0$ if $F(a) = 1$.
2. $C(a) \leq 0$ if $F(a) = 0$.

Weak MLP gates vs Monotone Boolean Circuits

Theorem

Let $F : \{0,1\}^n \to \{0,1,*\}$ be a partial Boolean function, and let C be a Boolean circuit of size s representing F. Then for any weak type τ, F can be sharply represented by an MLP gate of type τ and size $O(s)$.

Weak MLP gates vs Monotone Boolean Circuits

1. Let $BPM_n : \{0,1\}^{n^2} \to \{0,1\}$ be the Boolean function that evaluates to 1 on an input $p \in \{0,1\}^{n^2}$ if and only if p represents a bipartite graph with a perfect matching.
2. The Boolean function $BPM_n : \{0,1\}^{n^2} \to \{0,1\}$ can be represented by a MAX-RIGHT MLP gate of size $n^{O(1)}$.
3. Monotone Boolean Circuits computing BPM_n must have size $n^{\Omega(\log n)}$ (Razborov 1985).
4. Corollary: MAX-RIGHT MLP gates cannot be polynomially simulated by monotone Boolean circuits.
5. The gap between the complexity of MAX-RIGHT MLP gates and the complexity of Boolean formulas computing the BPM_n function is even exponential, since Raz and Wigderson have shown a linear lower-bound on the depth of monotone Boolean circuits computing BPM_n (Raz-Wigderson 1992).

Monotone Span Programs

1. Monotone span programs (MSP) were introduced by Karchmer and Wigderson (Karchmer-Wigderson 1993).
2. Such a program, which is defined over an arbitrary field $\mathbb{F}$, is specified by a vector $c \in \mathbb{F}^k$ and a labeled matrix $A^\rho = (A, \rho)$ where
   1. A is a matrix in $\mathbb{F}^{m \times k}$,
   2. $\rho : \{1, ..., m\} \rightarrow \{p_1, ..., p_n, *\}$ labels rows in A with variables in p_i or with the symbol $*$ (meaning that the row is unlabeled).
3. For an assignment $p := w$, let $A^\rho_{\langle w \rangle}$ be the matrix obtained from A by deleting all rows labeled with variables which are set to 0.

Monotone Span Programs

A span program (A^ρ, c) represents a partial Boolean function $F : \{0,1\}^n \rightarrow \{0,1,*\}$ if the following conditions are satisfied for each $w \in \{0,1\}^n$.

$$F(w) = \begin{cases} 1 \Rightarrow & \exists y,\ y^T A^\rho_{\langle w \rangle} = c^T \\ 0 \Rightarrow & \neg\exists y,\ y^T A^\rho_{\langle w \rangle} = c^T \end{cases} \tag{1}$$

Theorem

Let $F : \{0,1\}^n \to \{0,1\}$ be a Boolean function. If F can be represented by an MSP of size s over the reals, then F can be represented by a MIN-RIGHT *MLP gate of size $O(s)$.*

1. It has been recently shown that there is a family of functions $\mathrm{GEN}_n : \{0,1\}^n \to \{0,1\}$ which can be computed by polynomial-size monotone Boolean circuits but which require monotone span programs over the reals of size $\exp(n^{\Omega(1)})$ (Cook et al. 2016).
2. On the other hand, monotone Boolean circuits can be polynomially simulated by weak MLP gates of any type
3. In particular, weak MLP gates of size polynomial in n can represent the function $\mathrm{GEN}_n : \{0,1\}^n \to \{0,1\}$. Therefore, we have the following corollary.
4. Corollary: Weak MLP gates cannot be polynomially simulated by monotone span programs over the reals.

Lovás-Schrijver Proof System

Lovás-Schrijver Proof System

A method to construct certificates of unsatisfiability (proofs) for sets of linear inequalities / CNF formulas.

1. Translate clauses into inequalities in the obvious way.
2. $x_i \;\rightarrow\; x_i$
3. $\overline{x}_i \;\rightarrow\; (1 - x_i)$
4. $(x_1 \vee \overline{x}_2 \vee x_3) \;\rightarrow\; x_1 + (1 - x_2) + x_3 \geq 1.$

Lovás-Schrijver Proof System

- Axioms:
  1. $0 \geq 0$, $1 \geq 0$, $1 \geq 1$
  2. $0 \leq p_j \leq 1$
  3. $p_i^2 - p_i = 0$ (integrality).
- Rules:
  1. *positive linear combinations of linear and quadratic inequalities*
  2. *multiplication:* given a linear inequality $\sum_i c_i p_i - d \geq 0$, and a variable p_j, derive

$$p_j(\sum_i c_i p_i - d) \geq 0 \quad \text{and} \quad (1 - p_j)(\sum_i c_i p_i - d) \geq 0.$$

1. A proof Π of an inequality $\sum_i c_i p_i - d \geq 0$ from Φ is a sequence of inequalities such that every inequality in the sequence is either an element of Φ or is derived from previous ones using some LS rule.
2. We say that Π is a refutation of the set of inequalities Φ, if the last inequality is $-d \geq 0$ for some $d > 0$.
3. The LS proof system is implicationally complete. This means that if an inequality $\sum_i c_i p_i - d \geq 0$ is semantically implied by an initial set of inequalities Φ, then $\sum_i c_i p_i - d \geq 0$ can be derived from Φ by the application of a sequence of LS-rules (Lovasz-Schrijver 1991).

Monotone Feasible Interpolation Theorem For LS

1. Let $\Phi(p,q) \cup \Gamma(p,r)$ be an unsatisfiable set of inequalities such that the variables $p = (p_1, \ldots, p_n)$ occur in Φ only with negative coefficients.
2. Let Π be an LS refutation of $\Phi(p,q) \cup \Gamma(p,r)$.
3. Then one can construct an MLP circuit C containing only MAX MLP gates which represents a Boolean function $F : \{0,1\}^n \to \{0,1\}$ such that for each $a \in \{0,1\}^n$,
    1. if $F(a) = 1$, then $\Phi(a,q)$ is unsatisfiable,
    2. if $F(a) = 0$, then $\Gamma(a,r)$ is unsatisfiable,
4. Additionally, the size of the circuit C is polynomial in the size of Π.

Monotone Feasible Interpolation: Who Cares?

1. Resolution: monotone boolean circuits (Krajicek 1997) .
2. Cutting Planes: monotone *real* circuits (Pudlak 1997). Monotone real circuits are circuits with Boolean inputs and outputs, but whose gates are allowed to be arbitrary 2-input functions over the reals.
3. Razborov's lower bound on the clique function has been generalized to monotone real circuits (Pudlak 1997, Cook-Haken 1999).
4. Nullstellensatz: Monotone Span Programs (Pudlak Sgall 1998).

Framework for proving lower bounds for proof systems

- Pick a monotone model of computation $\mathcal{M}$.
- Show that refutations of $\Phi(p,q) \cup \Gamma(p,r)$ can be efficiently translated into monotone $\mathcal{M}$-circuits which identify which of $\Phi(p,q)$ or $\Gamma(p,r)$ is unsatisfiable.
- Exhibit a family of formulas $\hat{\Phi}(p,q) \cup \hat{\Gamma}(p,q)$ requiring large $\mathcal{M}$-circuits to decide whether $\hat{\Phi}$ or $\hat{\Gamma}$ is unsatisfiable.
- Then refutations of the corresponding formula must be large.

1. Our interpolation theorem for LS proof systems is stated in terms of strong MLP gates.
2. Strong MLP gates can compute quadratic functions!
3. Lower bounds seem to be out of reach.
4. Better chance: Weak MLP gates.
5. Size of MLP gates computing monotone functions has some relations with the field of extended formulations.

Theorem (From Circuits to Gates)

Let C be an MLP circuit of size s where all gates in C are weak MLP gates of type τ. Then there is an MLP gate ℓ_C of type τ and size $O(s)$ such that for each $a \in \mathbb{R}^n$ for which $C(a)$ is defined, $\ell_C(a) = C(a)$.

Monotone Feasible Interpolation for Mixed LS

1. Let $\Phi(p, q) \cup \Gamma(p, r)$ be a set of inequalities where p, q range over 0s and 1s, r range over reals, and the common variables $p = (p_1, ..., p_n)$ occur in Φ only with negative coefficients.
2. Let Π be an LS-refutation of $\Phi(p, q) \cup \Gamma(p, r)$.
3. Then there exists a MAX-LEFT MLP gate ℓ that represents a Boolean function $F : \{0, 1\}^n \to \{0, 1\}$ such that for every $a \in \{0, 1\}^n$,
    1. if $F(a) = 1$, then $\Phi(a, q)$ is unsatisfiable, and
    2. if $F(a) = 0$, then $\Gamma(a, r)$ is unsatisfiable,
4. Additionally, the size of the MLP gate ℓ is polynomial in the size of Π.

Relation with Other Proof Systems

LS vs Other Proof Systems

1. Resolution
   1. $(A \vee x) \wedge (B \vee \overline{x}) \;\rightarrow\; A \vee B$
2. Cutting Planes
   1. Positive linear combinations of inequalities.
   2. Rounding rule: If c_i are integers, then from $\sum c_i p_i \geq d$ derive $\sum c_i p_i \geq \lceil d \rceil$.

LS vs Resolution

1. The LS proof system is strictly stronger than Resolution.
    1. Resolution proofs can be simulated by LS proofs with a linear blow up in size.
    2. Pigeonhole principle requires resolution proofs of exponential size (Haken 1985).
    3. Pigeonhole principle has LS proofs of polynomial size.

LS vs Cutting Planes

1. Problems stated in the 1990's.
2. Determine whether LS proofs can be superpolynomially more concise than Cutting Planes Proofs. (Solved in this work.)
3. Determine whether cutting-planes proofs can be superpolynomially more concise than LS proofs. (Still Open)

CP does not polynomially simulate LS

1. Cutting plane proofs can be interpolated in terms of monotone real circuits (Pudlák 1997)
2. Monotone real circuit separating unbalanced graphs on n vertices from perfect matchings must have size $n^{\Omega(\log n)}$ (Fu 1998, by a generalization of Razborov's lower bound for monotone Boolean circuits).
3. Therefore Unbalanced Graphs vs Perfect Matching Inequalities require superpolynomial cutting plane proofs. (Fu 1998)
4. Unbalanced Graphs vs Perfect Matching Inequalities have short Mixed LS proofs. (This work)
5. Bonus: By our monotone interpolation theorem for mixed LS, a single weak MLP gate can separate unbalanced graphs from perfect matchings.
6. Therefore weak MLP gates can be superpolynomially stronger than monotone real circuits.

Open Problems

1. Prove superpolynomial lower bounds the size of weak MLP gates.
2. What if we make reasonable restrictions on the allowed gates? Examples: Bound on coefficients, or on the number of internal variable of the MLP gate.
3. Strengthen connections with extended formulations.
4. Show that monotone real circuits can be superpolynomially more concise than weak MLP gates. This would show that the two models are incomparable.
5. Monotone semidefinite programming gates? Which proofs systems can be interpolated by this model?

Thank you!