CSCI 1590
Intro to Computational Complexity
Circuits

John E. Savage

Brown University

March 4, 2009
Summary

1. Review

2. Circuit Families and Simple Linear Lower Bound

3. The Gate Elimination Method

4. Neciporuk’s Formula Size Lower Bound
Theorem

Let $\Omega$ be a complete basis of fan-in $r$. The depth of any $f : B^* \mapsto B$ with formula size $L(f) \geq 2$ satisfies

$$\log_r L(f) \leq D(f) \leq d(\Omega) \log_r L(f)$$

Let $\Omega_a$ and $\Omega_b$ be complete bases and $L_a(f)$ and $L_b(f)$ be formula size of $f$ over them. Then there is a constant $e$ such that

$$L_a(f) \leq [L_b(f)]^e$$

Let $0 < \epsilon < 1$. The fraction of Boolean functions $f : B^n \mapsto B$ for which

$$C(f) \geq \frac{2^n}{n} (1 - \epsilon) - 2n^2$$

is at least $1 - 2^{-\alpha}$ where $\alpha = \frac{\epsilon}{2} 2^n$ when $n \geq N_0$. 
Definition

A Time-$r(n)$ (Space-$r(n)$) **uniform circuit family** is a circuit family for which there is a DTM $M$ that for $n$ in unary writes the description of $C_n$ on its output tape in time (space) $r(n)$.

A **log-space uniform family of circuits** is a Space-$O(\log n)$ uniform circuit family.

Theorem

Let $f : B^n \leftrightarrow B$ depend on each of its variables. Over a basis of fan-in $r$, $C(f) \geq (n - 1)/(r - 1)$ and $D(f) \geq \log_r n$. 
The Gate Elimination Method

Definition

\( f : B^n \mapsto B \) is in \( Q_{2,3}^{(n)} \) if

- For every pair of variables \( x_j \) and \( x_k \), \( f \) has at least three distinct subfunctions as the two variables range over all four values.
- For each \( f \) in \( Q_{2,3}^{(n)} \) and each \( x_i \) there is a value \( c_i \) such that the subfunction obtained by assigning \( c_i \) to \( x_i \) is in \( Q_{2,3}^{(n-1)} \).
The Gate Elimination Method

Lemma

For $n \geq 3$, $Q_{2,3}^{(n)}$ contains $f_{\text{mod } 3,c}^{(n)}$ where

$$f_{\text{mod } 3,c}^{(n)}(x_1, x_2, \ldots, x_n) = ((y + c) \text{ mod } 3) \text{ mod } 2$$

for $c \in \{0, 1, 2\}$ where $y = x_1 + x_2 + \ldots + x_n$.

Proof

The functions $f_{\text{mod } 3,c}^{(1)}$ for $c \in \{0, 1, 2\}$ are $x_1$, $x_1$ and 0, respectively.

For $n = 2$, $f_{\text{mod } 3,c}^{(1)}$ has value 1 exactly when $y = 1, 0, 2$ for $c \in \{0, 1, 2\}$, respectively.

We now show that the property holds for $n \geq 3$. 
Proof (cont.)

In \( f_{\text{mod } 3,c}^{(n)}(x_1, x_2, \ldots, x_n) \) fix any two variables and let \( y^* \) be the sum of the remaining \( n - 2 \) variables and \( c^* \) be the sum of \( c \) and the values of the two fixed variables. Then,
\[
((y + c) \mod 3) \mod 2 = (((y^* \mod 3 + c^* \mod 3) \mod 3) \mod 2).
\]

Since the value of \( y^* \mod 3 \) is in \( \{0, 1, 2\} \) and \( c^* \mod 3 \) has values in \( \{0, 1, 2\} \), from above \( (((y^* \mod 3 + c^* \mod 3) \mod 3) \mod 2) \) is one of 3 different functions.

When \( c_i = 0 \), the resulting subfunction is in \( Q_{2,3}^{(n-1)} \).
The Gate Elimination Method

Theorem

Over the basis of all Boolean functions on 2 inputs, \( f \in Q_{2,3}^{(n)} \) for \( n \geq 3 \) has \( C(f) \geq 2n - 3 \).

Proof.

\( f \) depends on each of its variables because, if not, there is an \( x_i \) such that \( f \) doesn’t depend on it. If so, then picking any second variable, \( f \) has at most two subfunctions, a contradiction.

Some input has fan-out \( \geq \) two. If not, consider gate \( g \) with longest path to output. Both of its inputs are from variables. If each has fan-out one, \( f \) has at most two subfunctions on these inputs, a contradiction. Thus, for \( n = 3 \), there are at least 4 edges from inputs. The simple linear lower bound implies \( C(f) \geq 3 \).

Assume that \( C(f) \geq 2n - 3 \) for \( n \leq k \). For \( n = k + 1 \), fix an input with fan-out 2, thereby deleting two gates.
Lower Bounds to Formula Size

- **Goal:** To derive quadratic lower bounds to formula size
- **Approach:** Count number of times each variable must be used to compute a function using a formula.

### Test Function

Consider the **indirect storage access function** $f_{ISA}^{(k,l)} : B^n \mapsto B$ shown below where $n = k + l2^k + 2^l$, $a$ is a binary $k$-tuple, $x_j$ is a binary $l$-tuple, and $y$ is a binary $L$-tuple and where $K = 2^k$ and $L = 2^l$. Let $b = |x_{|a|}|$. Then,

$$f_{ISA}^{(k,l)}(a, x_{K-1}, \ldots, x_0, y) = y_b$$

- $a$ is an address that chooses one of $K$ words. The chosen word then chooses one bit in $y$. 
Let \( f_{\text{mux}}^{(k)} : B^n \mapsto B \), \( n = k + 2^k \), be the multiplexer function that uses the \( k \)-bit address \( a \) to select as output the \( |a| \)th of \( 2^k \) inputs \( y_0, y_1, \ldots, y_{2^k-1} \). Problem 9.24 of the book says this circuit can be realized with formula size \( 3 \cdot 2^k - 2 \) using \( 2(2^k - 1) \) copies of address variables.

The formula diagrammed below realizes \( f_{\text{ISA}}^{(k,l)} : B^{k+l2^k+2^l} \mapsto B \) with a formula of size \( O\left(\frac{n^2}{\log n}\right) \), \( n = k + l2^k + 2^l \).
Neciporuk’s lower bound method provides a formula size lower bound proportional to this upper bound.

**Definition**

Given $f : B^n \rightarrow B$, partition its $n$ variables $X$ into $p$ disjoint sets $X_1, X_2, \ldots, X_p$.

Let $r_j(f)$ be the number of different subfunctions of $f$ over $X_j$ when the variables in $X - X_j$ range over all values.

- We derive a lower bound on $L(f)$ in terms of $r_i(f)$ for $1 \leq i \leq p$.
- The $r_i(f)$’s depend on the partition used. Choose it wisely!
Theorem

Let $\Omega$ be a complete basis of fan-in $d$ and let $c_\Omega = 1/(d + 2)$. Then, for every $f : B^n \rightarrow B$ its formula size satisfies

$$L_\Omega(f) \geq c_\Omega \sum_{i=1}^{p} \log_2 r_i(f)$$
Proof

Let $T$ be smallest formula (tree) for $f$. Let $T_j$ be (bold) paths from variables in $X_j$ to root of $T$. **Controller (combiner)** vertices have one (two or more) input(s) from $X_j$. ($X_j = \{x_3^{(1)}, x_1^{(1)}, x_1^{(2)}, x_3^{(2)}\}$ in example.)

Let $n_j$ be the number of instances of vars in $X_j$ used in $T$. Then, $L(f) = n_1 + n_2 + \ldots + n_p$. Then, $T_j$ has as most $n_j - 1$ vertices with two or more inputs. It also has at most $2(n_j - 1)$ edges between vertices plus one associated with the output. (p. 394 Savage book & p. 7 Lect 12.)
Proof (cont.)

A **controller** computes one of the four different functions \((0, 1, x, \bar{x})\) in one variable \(x\), determined by constants in \(X - X_j\). In \(T_j\) two or more controllers per edge behave as one controller. Thus, \(T_j\) has at most \(2n_j - 1\) controllers.

A **combiner** has at least two inputs in \(X_j\) or at most \(d - 2\) inputs whose values are determined by constants in \(X - X_j\). Thus, a combiner can compute one of \(\leq 2^{d-2}\) functions. It has at most \(n_j - 1\) combiners.

The number of different functions associated with \(T_j\) is at most \(4^{2n_j - 1} 2^{(d-2)(n_j-1)} \leq 2^{(d+2)n_j}\). Since \(r_j(f)\) is number of different subfunctions of \(f\) over \(X_j\), \((d + 2)n_j \geq (\log_2 r_j(f))\) and the theorem follows.