1. Complexity Measure

Definition 1. A formal complexity measure is a function \( k \) from \( B_n \) to \( \mathbb{N} \) where \( B_n \) is a set of all boolean functions over \( \Sigma^n \) such that

1. \( k(x_i) = 1 \) for all \( 1 \leq i \leq n \),
2. \( k(f) = k(\neg f) \) for \( f \in B_n \), and
3. \( k(f \lor g) \leq k(f) + k(g) \) for \( f, g \in B_n \).

By the definition and using the rules of deMorgan \( k(f \land g) = k(\overline{f \lor g}) \leq k(f) + k(g) \) also holds. Since Fsize satisfies above conditions, it is also a formal complexity measure. In addition, Fsize is the largest formal complexity measure as in the following theorem.

Theorem 1. Let \( k \) be a formal complexity measure, then for every function \( f \in B_n \)

\[
Fsize(f) \geq k(f).
\]

Proof. By induction on \( \ell = Fsize(f) \), let’s start with the case \( \ell = 1 \), if \( \ell = 1, f(x_1, \ldots, x_n) = x_i \), \( Fsize(f) = 1 = k(f) \) by definition. Let \( \ell = Fsize(f) > 1 \) and let \( F \) be an optimal formula for \( f \). Consider the formula tree of \( F \). Without loss of generality, the last gate of \( F \) is an \( \lor \) gate, otherwise the rule of deMorgan can be considered. Let \( G \) and \( H \) be the subformulas that feed into this gate, and let \( g \) and \( h \) be the functions computed by them. Thus \( f = g \lor h \). Since \( F \) is optimal, \( G \) and \( H \) are optimal formulas for \( g \) and \( h \). Thus FSize\((g) = Size(G) \) and FSize\((h) = Size(H) \). Thus

\[
FSize(f) = Size(F) = Size(G) + Size(H) = FSize(g) + FSize(h).
\]

By the induction hypothesis,

\[
Fsize(f) = Fsize(g) + Fsize(h) \geq k(g) + k(h) \geq k(g \lor h) = k(f).
\]

Above theorem provides a tool to lower bound formula size. Define a complexity whose value is “easy” to bound, this yields a lower bound on formula size. We now define a complexity measure, this measure is also known as Krapchenko’s measure.

Definition 2. Given a function \( f \), let \( A \subseteq f^{-1}(0) \), \( B \subseteq f^{-1}(1) \), a set of neighbors \( \langle a, b \rangle \in A \times B \) is defined as

\[
N(A, B) = \{ \langle a, b \rangle | a \in A, b \in B, a \text{ and } b \text{ differ in exactly one bit} \}
\]

In other words, \( N(A, B) \) is a set of \( \langle a, b \rangle \) pairs where \( f(a) = 0 \) and \( f(b) = 1 \) and there is only one bit of difference between \( a \) and \( b \). Let

\[
N(A, B) = |N(A, B)|^2
\]

The complexity measure is defined as

\[
k_{AB} = \frac{|N(A, B)|^2}{|A||B|}.
\]

The complexity measure is defined as

\[
k(f) = \max \left\{ k_{AB} | A \subseteq f^{-1}(0), B \subseteq f^{-1}(1) \right\},
\]

where the maximum is taken over all possible sets \( A \subseteq f^{-1}(0) \) and \( B \subseteq f^{-1}(1) \).
Theorem 2. $k$ is a formal measure.

Proof. To prove we have to show that $k$ satisfies all three conditions of formal complexity measure. Hence, the theorem will be proved on three parts. Let $f(x_1, \ldots, x_n) = x_i$, $f^{-1}(0) = \{0,1\}^{n-1}0\{0,1\}^{n-i}$ and $f^{-1}(1) = \{0,1\}^{n-1}1\{0,1\}^{n-i}$.

First, we claim that $k(f) = k(x_i) = 1$. Let $A \subseteq f^{-1}(0)$, $B \subseteq f^{-1}(1)$. For $a \in A$, there is only one candidate $b$ is only obtained by flipping $i$th bit of $a$. Thus, $N(A, B) \leq |A|$ and $N(A, B) \leq |B|$ and following holds

$$k_{AB} = \frac{|N(A, B)|^2}{|A||B|} \leq 1.$$  

If $A = f^{-1}(0)$ and $B = f^{-1}(1)$, then for every $a \in A$, there exists a corresponding neighbor $b \in B$. Thus, $N(A, B) = |B| = |A|$ and $k_{A, B} = 1$. Therefore, $\max k_{AB} = k(f) = 1$.

Second, $k(f) = k(\bar{f})$. By the definition, above measure is symmetric. In addition $f^{-1}(0) = \neg f^{-1}(1)$ and vice versa. For this reason, we can conclude that $k(f) = k(\bar{f})$.

Lastly, $k(f \lor g) \leq k(f) + k(g)$. Let’s denote $h = f \lor g$ and let $S_0$ and $S_1$ be two sets such that $S_0 \subseteq h^{-1}(0)$, $S_1 \subseteq h^{-1}(1)$. Then

$$k_{s_0s_1} = \frac{|N(S_0, S_1)|^2}{|S_0||S_1|}$$

and $k(h) = \max \{k_{s_0s_1} | S_0 \subseteq h^{-1}(0), S_1 \subseteq h^{-1}(1)\}$. Since $h = f \lor g$, $f^{-1}(0) \cap g^{-1}(0) = h^{-1}(0)$ and $S_0 \subseteq f^{-1}(0)$ and $S_0 \subseteq g^{-1}(0)$. Let’s partition $S_1$ into two disjoint sets $C$ and $D$ such that $C \cup D = S_1$, $C \subseteq f^{-1}(1), D \subseteq g^{-1}(1)$\(^1\). Then $N(S_0, S_1)$ is the union of $N(S_0, C)$ and $N(S_0, D)$ where $N(S_0, C)$ and $N(S_0, D)$ are disjoint. By the definition, following holds

$$k(f) \geq k_{s_0C} = \frac{|N(S_0, C)|^2}{|S_0||C|}$$

and

$$k(g) \geq k_{s_0D} = \frac{|N(S_0, D)|^2}{|S_0||D|}.$$ 

It is sufficient to prove that

$$\frac{|N(S_0, S_1)|^2}{s_0s_1} \leq \frac{p^2}{s_0C} + \frac{q^2}{s_0D}$$

\(^1\)Every element in a set $f^{-1}(1) \cap g^{-1}(1)$ should be placed on $C$ or $D$ not on both.
where $|C| = c$, $|D| = d$, $|N(S_0, C)| = p$, and $|N(S_0, D)| = q$. Further, $S_1 = C \cup D$ and $C \cap D = \emptyset$ thus $s_1 = c + d$ and we have $N(S_0, S_1) = N(S_0, C) + N(S_0, D)$. Thus,

\[
\begin{align*}
\frac{|N(S_0, S_1)|^2}{s_0s_1} &\leq \frac{p^2}{s_0c} + \frac{q^2}{s_0d} \\
\Rightarrow \frac{(p + q)^2}{s_0(c + d)} &\leq \frac{p^2}{s_0c} + \frac{q^2}{s_0d} \\
\Rightarrow \frac{(p + q)^2}{(c + d)} &\leq \frac{dp^2 + cq^2}{cd}
\end{align*}
\]

Consequently,

\[
\frac{cd(p + q)^2}{s_0s_1} \leq (c + d) \left( \frac{dp^2 + cq^2}{cd} \right) \Rightarrow 2cdpq \leq d^2p^2 + c^2q^2 \Rightarrow 0 \leq (dp - cq)^2
\]

Consequently, $k(f \lor g) = \frac{(p+q)^2}{s_0s_1} \leq \frac{p^2}{s_0c} + \frac{q^2}{s_0d} \leq k(f) + k(g)$. \qed

We can use Krapchenko’s measure to bound the formula complexity of the parity function.

**Theorem 3.** For the parity function $f(x_1, \ldots, x_n) = \bigoplus_{i=1}^n x_i$, $\theta(n^2) = Fsize(f)$.

**Proof.** First, we will prove the lower bound as $n^2 \leq Fsize(f)$. Let $A = f^{-1}(0)$ and $B = f^{-1}(1)$. Because flipping any one bit causes the function to return different value, there are $n$ neighbors for every $a \in A$ and vice versa. As a result, $N(A, B) = n|A| = n|B|$. Therefore,

\[
k_{AB} = \frac{|N(A, B)|^2}{|A||B|} = \frac{n|A||n|B|}{|A||B|} = n^2
\]

As a consequence, for parity function $f : \Sigma^n \rightarrow \Sigma$, every formula requires at least $n^2$ size, $n^2 \leq Fsize(f)$.

Moreover, the upper bound for $Fsize$ can be derived for $n = 2^k$ for $k \in \mathbb{N}$. The parity function for $\Sigma^n$ can be decomposed into four functions for $\Sigma^{n/2}$ as

\[
f_n(x) = (f_{n/2}(x_\ell) \land \neg f_{n/2}(x_r)) \lor (\neg f_{n/2}(x_\ell) \land f_{n/2}(x_r)),
\]

where $x \in \Sigma^n$ and $x_\ell, x_r \in \Sigma^{n/2}$ where $x = (x_\ell, x_r)$. With $Fsize(f_n) \leq 4Fsize(f_{n/2})$ and $Fsize(f_1) = 1$, we can conclude that $Fsize(f_n) \leq 4^{\log_2 n} = 2^{2\log_2 n} = 2^{\log_2 n^2} = n^2$. \qed

We now observe the Krapchenko’s measure can not yield bounds that are bigger than $n^2$.

**Lemma 1.** For all $f : \Sigma^n \rightarrow \Sigma$, $k(f) \leq n^2$.

**Proof.** For any $a \in A$ or $b \in B$, there are at most $n$ neighbors. Thus, $N(A, B) \leq \min \{n |A|, n |B| \}$ and

\[
k_{AB} = \frac{|N(A, B)|^2}{|A||B|} \leq \frac{n |A||n|B|}{|A||B|} = n^2.
\]

\qed