1. Prove that the 2-bit NAND is a universal basis.

2. Prove that with MAJ, ∨, ∧, ¬ gates (arbitrary fan-in), there is a polynomial size, $O(1)$ depth circuit that computes PARITY.

3. Let $\text{SIZE}_{dt}(f)$ denote the smallest number of leaves in a decision tree computing $f$. Prove that for every $f : \{0, 1\}^n \rightarrow \{0, 1\}$ if $\rho$ is a random restriction with parameter $p$ (that is $\Pr[*] = p$), we have

$$E|_{\rho}[\text{SIZE}_{dt}(f_\rho)] \leq \text{SIZE}_{dt}(f)^{\log_2(1+p)}.$$

Prove that this bound cannot be improved by giving an example where the bound is sharp.

4. Let $C \subseteq \{0, 1\}^n$ be a set of size $m$, and $f : C \rightarrow \{0, 1\}$ be known to us. In the terminology of machine learning, $C$ is called a concept class, and $f(x)$ is called the label of the concept $x \in C$. The goal is to determine the label of an unknown $x$ by making the smallest possible number of queries about $x$. A query model specifies what queries are allowed. Some commonly used query models are coordinate queries $x_i$, parity queries $\oplus_{i \in S} x_i$, and threshold queries $\text{sgn}(t - \sum_{i=1}^n w_i x_i)$. Consider a query model $Q$ (e.g., the parity query model).

The deterministic query complexity of $C$, denoted by $dt^Q(C)$, is the smallest $k$ such that the label of every point $x \in C$ can be determined by making at most $k$ queries about $x$. Note that this corresponds to a decision tree of depth at most $k$, where the internal nodes correspond to queries, and the leaves correspond to the predicted labels.

The randomized query complexity of $C$, denote by $r dt^Q(C)$, is the smallest $k$ such that the label of every point $x \in C$ can be determined with error probability of at most $\frac{1}{3}$ by making at most $k$ random queries. Prove that

$$dt^Q(C) = O(r dt^Q(C) \log m).$$
5. Let \( k > 0 \) be a fixed constant, and let the threshold function \( t_k : \{0,1\}^n \rightarrow \{0,1\} \) be defined as \( t_k(x) = 1 \) if and only if \( \sum x_i \geq k \).

(a) Prove that the deterministic parity decision tree complexity\(^1\) of \( t_k \) is \( \Omega(n) \).

(b) Prove that there is a (one-sided error) randomized parity decision tree \( T_R \) of constant depth that computes \( t_k(x) \). That is

\[
\Pr_{R}[T_R(x) = t_k(x)] = 1 \quad \text{for all } x \text{ with } t_k(x) = 0.
\]

\[
\Pr_{R}[T_R(x) = t_k(x)] \geq \frac{1}{2} \quad \text{for all } x \text{ with } t_k(x) = 1.
\]

What does this tell us about the bound in the previous question?

(Hint: First solve the problem for \( k = 1 \). For general \( k \), randomly partition the variables into \( k + 1 \) disjoint sets, and focus on each of these sets.)

6. Follow Razborov-Smolensky proof strategy to show that \textsc{Parity} and \textsc{Maj} do not have polynomial size \( \text{AC}^0[\text{mod } 3] \) circuits. These are circuits similar to \( \text{AC}^0 \) but additionally “mod 3” gates with arbitrary fan-in are allowed. The “mod 3” gates output 0 if and only if the number of 1’s on the input wires is divisible by 3.

Hint: Try to work in the field \( \mathbb{F}_p \) (\( p \) is a prime) for an appropriate value of \( p \). Recall that \( \mathbb{F}_p = \{0,1,\ldots,p-1\} \) along with operations addition and multiplication mod \( p \) forms a field. Recall that for every \( a \in \mathbb{F}_p \) we have \( a^p = a \).

7. A function \( f : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\} \) naturally corresponds to a bipartite graph on \( 2 \times 2^n \) vertices. More formally there are two parts each with \( 2^n \) elements labeled with \( \{0,1\}^n \), and there is an edge between \( x \) from part I and \( y \) from part II, if and only if \( f(x,y) = 1 \). This graph is called \( k \)-Ramsey if there do not exist \( A,B \subseteq \{0,1\}^n \) with \( |A| = |B| = k \), and \( b \in \{0,1\} \), with \( f(x,y) = b \) for all \( x \in A, y \in B \).

(a) Prove that with probability at least \( \frac{1}{2} \) a random function defines a \( O(n) \)-Ramsey graph.

More precisely, show that there is a universal constant \( c > 0 \) (does not depend on \( n \)) such that a random function is \( cn \)-Ramsey with probability at least \( 1/2 \).

(b) Show that if \( f \) has an \( \text{AC} \) circuit of size \( s \), and depth \( d \), then it is not \( 2^{\Omega(n/\log^d(s))} \)-Ramsey.

---

\(^1\)i.e., smallest depth of a parity decision for \( t_k \).