Lecture Jan 16th + Jan 18th

Minghao Liu, 51102135

February 2017

1 OVERVIEW

In this lecture, we discussed: linear decision trees, connected sets and the complexity of the ELEMENT UNIQUENESS problem.

2 LECTURE DETAILS

2.1 LINEAR DECISION TREE

Definition: A linear decision tree is a kind of decision tree whose nodes can be viewed as an $(n + 1)$-tuple$(c_0, c_1, ..., c_n)$, where the c_i's form the coefficients of the linear combination $c_0 + c_1x_1 + c_2x_2 + ... + c_nx_n$. Note that $x_1, ..., x_n$ are inputs to the tree, and we make decisions at each node based on the value of the linear combination.

Some properties of linear decision trees:

1. Decisions at each internal node is based on the sign of the linear combination.

2. Inputs that follow a child edge either form a halfspace (for $<$ and $>$ sign) or a halfplane ($=$ sign), both of which are convex. Note that halfplanes have one less dimension than halfspaces in this case.

3. Since the intersection of convex sets are convex, by induction, it follows that the inputs that reach any internal node form a convex set.

2.2 CONNECTED SETS

Let F be a function, $F_t = \{x | F(x) = t\}$ be the inputs that get mapped to t. A set S is connected if for any pair of points $p, q \in S$, there exists a path through p and q that lies completely in S.

Lemma: Let $\#F_t$ be the number of connected components in F_t. Any linear decision tree that computes F has height at least $\left\lceil log_3 \left(\sum_{t \in Range(F)} \#F_t \right) \right\rceil$

Proof: Each leaf node of the linear decision tree represents a connected component having the same output. So the total number of leaves is at least $\sum_{t \in Range(F)} \#F_t$. And since each internal node has up to 3 branches, the height is at least $\lceil log_3(\sum_{t \in Range(F)} \#F_t) \rceil$. □

2.3 Complexity of Element Uniqueness

Element Uniqueness problem: Given a list of integers, determine if all elements are distinct.

Theorem: Any linear decision tree that solves Element Uniqueness has height $\Omega(nlogn)$.

Proof Idea/Intuition: For a list of n unique integers there are $n!$ permutations. If we can show that no two permutations lie in the same connected components, then there are $n!$ connected components in total; by the previous lemma, we've proved the lower bound on the height.

Proof: Let S be a list of n unique integers and T be any permutation of S. Since S and T are different permutations consisting of the same integers, there must be indices i and j such that $S_i < S_j$ and $T_i > T_j$. Any continuous path from S to T must contain a point V such that $V_i = V_j$ by the Intermediate Value Theorem. Since V is a *NO*-input to Element Uniqueness, S and T do not lie in the same connected component. Since they are arbitrary, there are at least $n!$ different connected components to YES inputs of Element Uniqueness. By the previous lemma, the height of the linear decision tree is at least $\lceil log_3(\sum_{t \in Range(F)} \#F_t) \rceil = \lceil log_3(n!) \rceil \in \Omega(nlogn)$. □