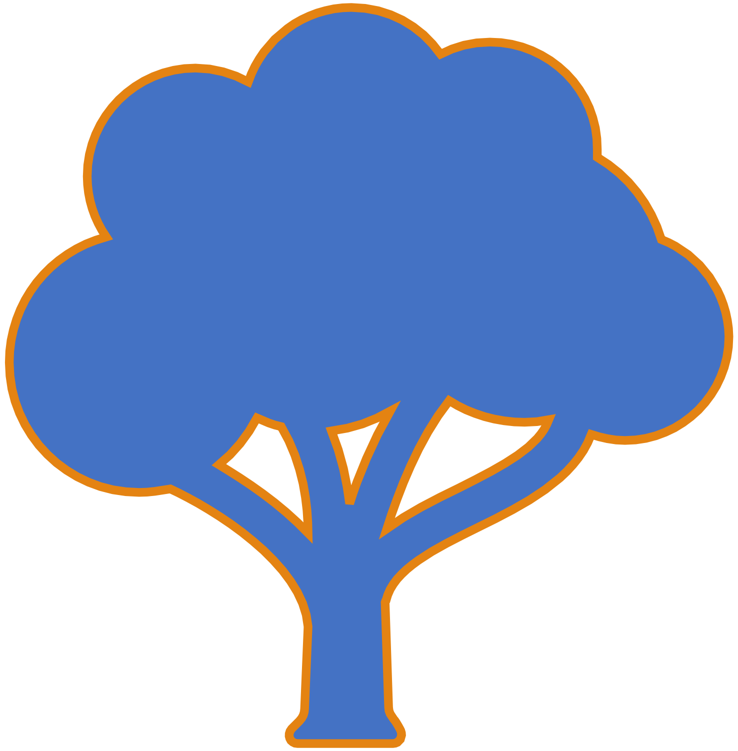


Decision Tree

BY HITESH

2019201039

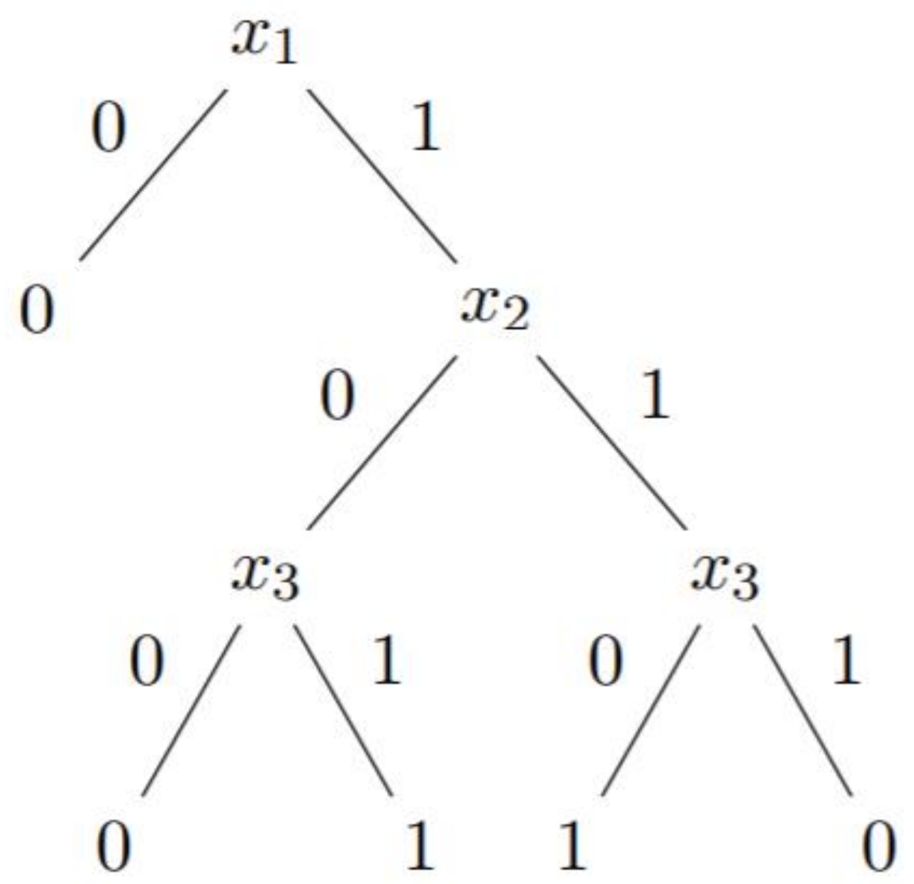
Decision Tree

- A computational problem is a function whose domain is one set of bit sequences and whose range is another set of bit sequence.
-

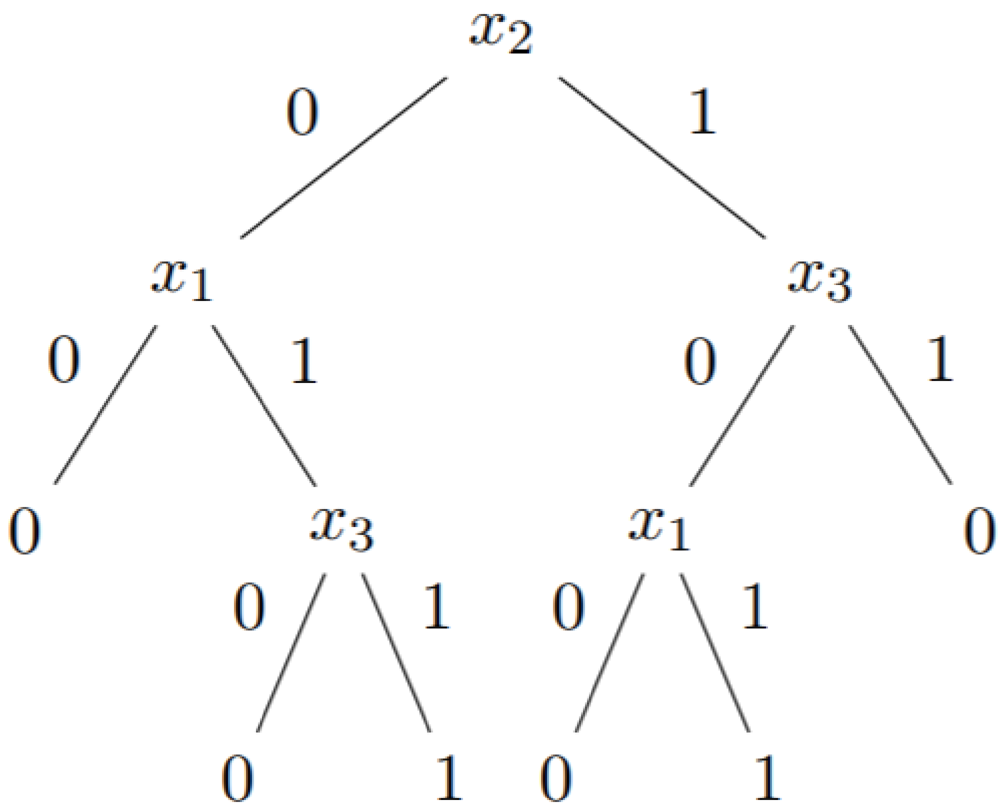
  We model such problems as functions $f : \{0, 1\}^n \to \{0, 1\}$ m for some a priori fixed lengths n and m.
- A decision problem is a computational problem with a yes/no answer, like "is graph G connected?". Such problems are represented by functions $f : \{0, 1\}^* \to \{0, 1\}$ and $f : \{0, 1\}^n \to \{0, 1\}$ in uniform and non-uniform complexity, respectively

Decision Tree

- A decision tree for inputs of length n is a rooted binary tree whose vertices and edges are labeled as follows. Each of its internal nodes is labeled by one of the variables x1, . . . , xn and its two outgoing edges are labeled by the values 0 and 1 respectively. Each leaf is labeled by a 0 or by a 1. Here is an example:



Decision Tree



x1 and (x2 xor x3).

Decision Tree

- The first decision tree is in some sense preferable than the second one as it is smaller. This is our first example of a complexity measure: The size of a decision tree is its number of leaves.

- It is not difficult to see that any function $f : \{0, 1\}\ n \rightarrow \{0, 1\}$ can be computed by a sufficiently large decision tree.

Definition 12.1 *(Decision tree complexity)* The cost of tree t on input x, denoted by $cost(t,x)$, is the number of bits of x examined by t.

The *decision tree complexity* of a function f, is defined as

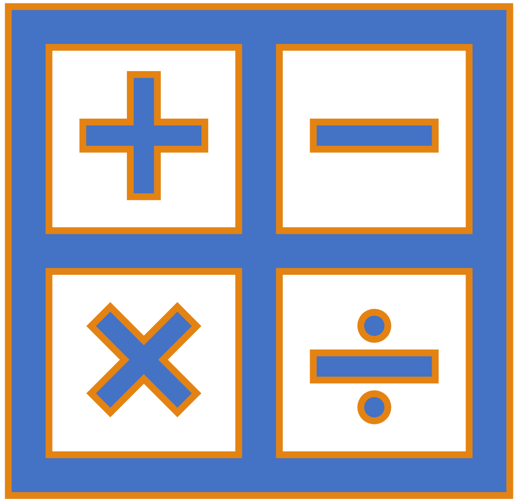
$$D(f) = \min_{t \in \mathcal{T}_f} \max_{x \in \{0,1\}^n} cost(t,x)$$

where $\mathcal{T}_f$ denotes the set of all decision trees that compute f.

Decision Tree Complexity

Question

- Suppose that we are given an m-vertex graph G as input, representedas a binary string of length n = mc2 binary string, with the eth coordinate equal to 1 if the edge e is in G, andequal to 0 otherwise. We wouldlike to know how many bits of the adjacency matrix a decision tree algorithm might have to inspect to determine whether G is connected?



Definition 12.3 *(Certificate complexity)* Let $f : \{0,1\}^n \to \{0,1\}$ and $x \in \{0,1\}^n$. A 0-*certificate* for x is a subset $S \subseteq \{0,1\}^n$, such that $f(x') = 0$ for every x' such that $x'|_S = x|_S$ (where $x|_S$ denotes the substring of x in the coordinates in S). Similarly, if $f(x) = 1$ then a 1-*certificate* for x is a subset $S \subseteq \{0,1\}^n$ such that $f(x') = 1$ for every x' satisfying $x|_S = x'|_S$.

The *certificate complexity* of f is defined as the minimum k such that every string x has a $f(x)$-certificate of size at most k. (Note that a string cannot have both a 0-certificate and a 1-certificate.)

Certificate Complexity

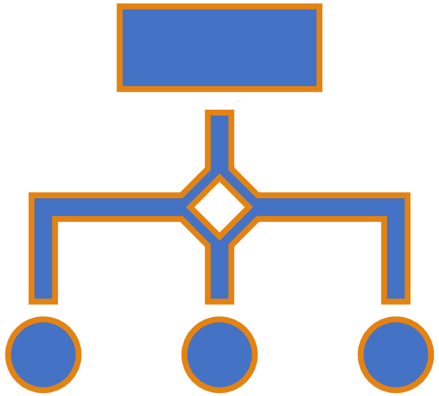
Certificate Complexity

Consider a function f : {0, 1}^n→ {0, 1}.

If f(x)=0 then a 0-certificate for x is a sequence of bits in x that proves f(x)=0.

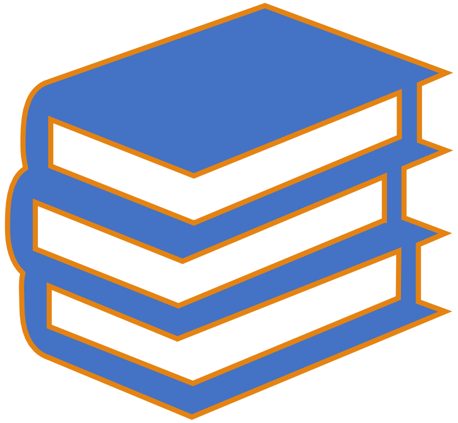
If f(x)=1, then a 1-certificate is a sequence of bits in x that proves f(x)=1.

Example



- Graph connectivity: Let f be the graph connectivity function. Recall that for an m-vertex graph, the decision tree complexity of f is $D(f) = m*(m-1)\backslash 2$.
- A 1-certificate for a graph G is a set of edges whose existence in G implies that it is connected. Thus every connected m-vertex graph G has a 1-certificate of size $m - 1$—any spanning tree for G.
- A 0-certificate for G is a set of edges whose nonexistence forces G to be disconnected—a cut. Since the number of edges in a cut is maximized when its two sides are equal, every m-vertex graph has a 0-certificate of size at most $(m/2)^2 = m^2/4$.

Resources



1. https://users.cs.duke.edu/~reif/courses/complectures/Arora/lec17.pdf
2. http://www.cse.cuhk.edu.hk/~andrejb/csci5170/notes/19L01.pdf
3. Book: COMPUTATIONAL COMPLEXITY by Sanjeev Arora

Thank You