Asking Pythia

Alapan Chaudhuri and Pratishtha Abrol

November 13, 2021

Contents

Introduction

Sophocles is wise, Euripides is wiser, but of all men Socrates is wisest.

- Pythia

Introduction

Query complexity is the study of complexity of one of the fundamental paradigms of computation, namely — queries.

So, what is a query? It is a mapping from "structures of one signature to structures of another vocabulary".

Introduction

Introduction

So, why is this of any importance?

- Simplicity and its importance in studying complexity of computable functions
- Ability to analyse (all major) quantum algorithms

Basics

In query complexity we deal with the notion of a blackbox. Imagine that we are provided with a blackbox x computing some unknown function, and all we are allowed to do is to make queries to the blackbox.

Our goal would be to calculate some property $f(x)$ of this blackbox with minimal queries.

Queries entail feeding in an input i to x, and observe the output $x(i)$.

Basics

Considering x as a function, we can agree that x is completely determined by the outputs to all the different types of inputs that it accepts.

Say that the domain of x is $[n] = \{1, 2, 3, \ldots, n\}$. Then, we can use a string s to represent x where:

- $\text{len}(s) = n$
- $s_i = x(i)$

Basics

Typically, the blackbox x will output Boolean outputs, which makes the string x a Boolean string in $\{0,1\}^n$.

However, it is also possible to consider blackboxes with non-Boolean outputs, in which case the string x will be in Σ^n for some alphabet Σ.

Basics

This beautifully transforms the problem into computing some function f of the input x by computing as few bits of x as possible.

Thus,

$$f : \{0, 1\}^n \to \{0, 1\}$$

where f is a known function.

For example, f might be the **OR** function, in which case our goal would be to determine whether the input string x is all zeros (equivalently, whether the input blackbox x ever returns a 1).

Basics

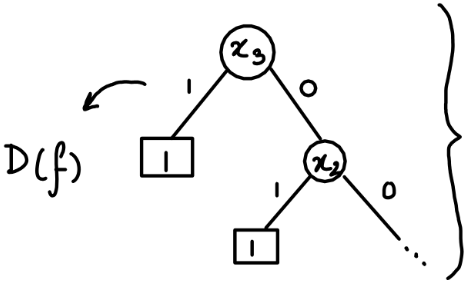
Can we model this problem better?

Umm yes, we shall use **Decision Trees**!

Basics

$$f : \{0,1\}^n \rightarrow \{0,1\}$$

$$x = \begin{array}{|c|c|c|c|c|c|}\hline 0 & 1 & 1 & 0 & 0 & 0 \\ \hline \multicolumn{1}{c}{x_1} & \multicolumn{1}{c}{x_2} & \multicolumn{1}{c}{x_3} & \multicolumn{1}{c}{x_4} & \multicolumn{1}{c}{x_5} & \multicolumn{1}{c}{x_6} \end{array}$$



height of D(f) = cost / complexity

Basics

If D is a decision tree, we use $D(x)$ to denote the output the decision tree when run on x. So, basically it is the leaf node we reach across the tree.

Furthermore, we say D computes f iff $D(x) = f(x)\ \forall x \in \{0,1\}^n$.

Determinism

Thus, as described above a deterministic query algorithm will be defined as a decision tree.

Definition

A (deterministic) decision tree on inputs of size n is either a leaf in $0, 1^n$ or a tuple (i, D_0, D_1) where $i \in [n]$ and where D_0 and D_1 are decision trees (on inputs of size n) which do not use i in any internal tuple.

Determinism

On that note, we define the deterministic query complexity as follows.

Definition

The decision tree complexity of $f : \{0,1\}^n \rightarrow \{0,1\}$ is namely,

$$\mathit{min}\ \text{height}(D)\ \ \forall D \in D_f$$

where D_f is the set of all decision trees that compute f.

Randomness

A randomized algorithm can make random queries to the input x, and must still compute $f(x)$.

It will basically be a probability distribution over deterministic decision trees and which deterministic algorithm (tree) we are going to run will depends on the random seed of the randomized algorithm.

$$R = 0.1 \times (D_1) + 0.9 \times (D_2)$$

Randomness

In this case, we have two notions for applicable complexity measures.

Definition

$\text{height}(R) = \max \text{height}(D) \ \forall D \in \text{support}(R)$

Randomness

The second one is as follows.

Here, we are still considering the worst-case input, but we average out over the internal randomness of R when measuring the query cost on that input. In the former case, we take the worst case over both inputs and randomness.

Definition

$\text{cost}(R) = E[height(R, x)]$

Randomness

If R is a randomized decision tree, we let $R(x)$ denote the random variable we get by picking a decision tree according to R (the distribution) and running it on x.

We say R computes f to error ϵ if $Pr[R(x) \neq f(x)] \leq \epsilon$ for all x.

Randomness

If we relax our cost measure but not the correctness measure, we get a called zero-error randomized query complexity, denoted $R_0(f)$.

Zero-error randomized algorithms (also called Las Vegas algorithms) always output the right answer, and for all inputs they terminate quickly on average, but they may use a very large number of queries if you get unlucky.

Randomness

Alternatively, we can relax the correctness measure but not the cost measure.

Thus, we have:

- $R_0(f) = \min \text{cost}(R)$ where R computes f to error 0
- $R(f) = \min \text{height}(R)$ where R computes f to error $1/3$

Randomness

Thus, given $\epsilon, \epsilon' < 1/2$ we have:

Theorem

$R_\epsilon(f) = \Theta(R_{\epsilon'}(f))$

Further, we have $R_0(f) \leq D(f)$.

Intro to Quantum Systems

Quantum Theory can be described as an extension of Classical Probability.

Just as Classical Probability Theory, Quantum Mechanics can be fundamentally formulated by pure mathematical constructs alone without any particular appeal to experiment.

It is what we get upon conserving the L_2 norm rather than the L_1 norm (as in Classical Probability) along with addition of the continuity axiom and the idea of measurement.

Intro to Quantum Systems

Thus, the axiomatic formulation of Quantum Mechanics can be stated as follows.

- The state of a system encodes probability of outcomes in a vector, say $[\alpha_1, \alpha_2, \ldots \alpha_n]$, such that $\sum_{\forall i} |\alpha_i|^2 = 1$, or that the L_2 norm is preserved.
- There exists a continuous reversible transformation on a system between any two pure states of that system. This is called the axiom of continuity.
- Measurement in a standard basis results in a collapse of the state to whatever outcome is obtained. The outcome is governed by the probability distribution.