Notes 8.370/18.435 Fall 2022

Lecture 25 Prof. Peter Shor

In this lecture, we will give a lower bound that shows that Grover's algorithm is nearly optimal. In particular, we assume that you have a quantum state space with basis $\{|1\rangle, |2\rangle, \ldots |N\rangle\}$ and an oracle function O_x such that

$$O_x |y\rangle = \begin{cases} -|y\rangle & \text{if } y = x \\ |y\rangle & \text{if } y \neq x\,. \end{cases}$$

We will show that any algorithm which always returns the output $|x\rangle$ must call the oracle on the order of $\sqrt{N}$ times. We first show that this is true if the algorithm is required to give the answer $|x\rangle$ with probability 1, and then show that the proof still works if we just require the answer with probability at least $1 - \epsilon$.

One amazing thing about this proof is that I believe it was developed independently from Grover's algorithm, starting with the motivation of whether you can show that quantum computers can't solve NP-complete problems (although it was published after Grover, and the original paper mentions Grover.

We model our algorithm as follows. First, by the principle of deferred measurement (see homework), we can assume that we postpone all the measurements in our algorithm until the end. The algorithm will then consist of unitary operations and calls to the oracle. We will define U_j so that the algorithm alternates calling the oracle O_x and implementing the unitary U_k.Suppose the algorithm requires t calls to the oracle in the worst case. Then, we define the algorithm to be:

$$|\psi_t^x\rangle = U_t O_x U_{t-1} O_x U_{t-2} \ldots U_1 O_x |\psi_0\rangle$$

If not all runs of the algorithm use the same number of calls to the oracle, we can always add more dummy calls to the oracle that in essence do nothing.

Now, suppose we are able to identify x with probability 1 after this algorithm. Then it must be the case that all the $|\psi_t^x\rangle$ are orthonormal. We will bound how quickly they are able to become distinct.

Define the state of the quantum computer after k oracle calls to be

$$|\psi_k^x\rangle = U_k O_x U_{k-1} O_x U_{k-2} \ldots U_1 O_x |\psi_0\rangle$$

We will define a measure of how quickly the states $|\psi_k^x\rangle$ are diverging from each other. In order to do this, let

$$|\psi_t\rangle = U_t U_{t-1} U_{t-2} \ldots U_1 |\psi_0\rangle\,,$$

so $|\psi_t\rangle$ is what we get when we run the algorithm with no oracle calls (or equivalently, when there are no solution states). Now, define

$$D_k = \sum_x \big|\, |\psi_k^x\rangle - |\psi_k\rangle \,\big|^2.$$

We will use D_k as a measure of how far the states $|\psi_k^x\rangle$ are from each other. In other to get a bound on the number of oracle calls we need, we will need to prove two things:

1. an upper bound on D_k,
2. a proof that if D_t is too small, then we cannot distinguish the $|\psi_x\rangle$ from each other well.

We first prove the upper bound for D_k.

We have

$$\begin{aligned} D_{k+1} &= \sum_x \big|U_{k+1}O_x\,|\psi_k^x\rangle - U_{k+1}\,|\psi_k\rangle)\big|^2 \\ &= \sum_x \big|O_x(|\psi_k^x\rangle - |\psi_k\rangle) + (O_x - I))\,|\psi_k\rangle\,\big|^2 \\ &\le \sum_x \big|O_x(|\psi_k^x\rangle - |\psi_k\rangle)\big|^2 + \big|(O_x - I))\,|\psi_k\rangle\,\big|^2 + 2\big|(O_x - I))\,|\psi_k\rangle\,\big| \cdot \big|O_x(|\psi_k^x\rangle - |\psi_k\rangle)\big|, \end{aligned}$$

where we have added and subtracted the term $O_x\,|\psi_k\rangle$, and used the inequality $|a+b|^2 \le |a|^2 + |b|^2 + 2|b||a|$.

We will deal with the three terms in this inequality in order. For the first term, because unitary transformations preserve length,

$$\sum_x \big|O_x(|\psi_k^x\rangle - |\psi_k\rangle)\big|^2 = \sum_x \big|\,|\psi_k^x\rangle - |\psi_k\rangle\,\big|^2 = D_k$$

For the second term, we have $(O_x - I))\,|\psi_k\rangle = -2\,\langle x|\psi_k\rangle\,|x\rangle$ because $O_x\,|y\rangle = |y\rangle$ if $y \ne x$ and $O_x\,|x\rangle = -2\,|x\rangle$. Thus,

$$\sum_x \big|(O_x - I))\,|\psi_k\rangle\,\big|^2 = \sum_x |2\,\langle x|\psi_k\rangle\,|^2 = 4;$$

because $\{|x\rangle\}$ is a basis Finally, we have

$$2\sum_x \big|(O_x - I))\,|\psi_k\rangle\,\big| \cdot \big|O_x(|\psi_k^x\rangle - |\psi_k\rangle)\big| = 2\sum_x \big|\,\langle x|\psi_k\rangle\,\big|2 \cdot \big|\,|\psi_k^x\rangle - |\psi_k\rangle)\big|.$$

Now, we use the Cauchy-Schwarz inequality, $v \cdot w \le |v||w|$, on this term. Thus,

$$\begin{aligned} 2\sum_x \big|2\,\langle x|\psi_k\rangle\,\big| \cdot \big|\,|\psi_k^x\rangle - |\psi_k\rangle)\big| &\le 4\sqrt{\sum_x \big|\,\langle x|\psi_k\rangle\,\big|^2}\sqrt{\sum_x \big|\,|\psi_k^x\rangle - |\psi_k\rangle)\big|^2} \\ &= 4\sqrt{D_k}\,. \end{aligned}$$

We thus have, substituting into the equation for D_{k+1},

$$D_{k+1} \le D_k + 4\sqrt{D_k} + 4$$

We show by induction that this gives $D_k \le 4k^2$. It's clear that $D_0 = 0$. Now, assume that this equation holds true for D_k. We have

$$D_{k+1} \le D_k + 4\sqrt{D_k} + 4 \le 4k^2 + 8k + 4 = 4(k+1)^2,$$

and we are done.

The next thing we need to do is show that D_t has to be large in order to identify x. We will assume that we find x with certainty; you can modify the proof so that this is not necessary. If we find x with certainty, then the $|\psi_t^x\rangle$ are othonormal for all N values of x. We will then show that $D_t \geq 2N - 2\sqrt{N}$.

Let's look at the equivalent problem where we have $|e_1\rangle$, $|e_2\rangle$, $|e_3\rangle$, ..., $|e_N\rangle$, and we want to find the vector $|v\rangle$ that minimizes $\sum_{i=1}^{N} |\,|e_i\rangle - |v\rangle\,|^2$. It turns out that $|v\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |e_i\rangle$, and one can calculate that for this value of $|v\rangle$, $\sum_{i=1}^{N} |\,|e_i\rangle - |v\rangle\,|^2 = 2N - 2\sqrt{N}$. Since $D_t \leq 2t^2$, this shows that $2t^2 \geq 2N - 2\sqrt{N}$, which implies that $t > \sqrt{N} - 1$.

We now show that this is indeed the minimum. We have that

$$|\,|e_i\rangle - |v\rangle\,|^2 = \langle e_i|e_i\rangle + \langle v|v\rangle + \langle e_i|v\rangle + \langle v|e_i\rangle \geq 2 - 2|\,\langle e_i|v\rangle\,|$$

so

$$\sum_{i=1}^{N} |\,|e_i\rangle - |v\rangle\,|^2 \geq 2N - 2\sum_i |\,\langle e_i|v\rangle\,|.$$

By Cauchy-Schwarz,

$$\sum_{i=1}^{N} |\,\langle e_i|v\rangle\,| \cdot 1 \leq \sqrt{\sum_{i=1}^{N} |\,\langle e_i|v\rangle\,|^2}\sqrt{\sum_{i=1}^{N} 1} = \sqrt{N}$$

Thus, we have

$$\sum_{i=1}^{N} |\,|e_i\rangle - |v\rangle\,|^2 \geq 2N - 2\sqrt{N},$$

and we are done.

In fact, you can show that even if you only require the algorithm to succeed half the time, it still must take order $\sqrt{N}$ steps; you can see the proof of this in the Nielsen and Chuang.