CS 880: Quantum Information Processing 9/20/10

Lecture 7: Query Lower Bounds

Instructor: Dieter van Melkebeek Scribe: Cong Han Lim

Last class, we covered Grover's quantum search algorithm, which gives us a quadratic speedup over classical and probabilistic algorithms. Today, we review some applications of Grover's algorithm and prove the optimality of the algorithm (up to a constant factor).

1 Applications of Grover's Algorithm

1.1 Amplitude Amplification

Suppose we are given a function f and a quantum algorithm A that produces a superposition over the base states, which can be partitioned into a good set ($f(x) = 1$) and a bad set of states ($f(x) = 0$):

$$\text{Output of } A = \sum_{x:f(x)=1} \alpha_x |x\rangle |\text{Garbage}(x)\rangle + \sum_{x:f(x)=0} \alpha_x |x\rangle |\text{Garbage}(x)\rangle$$

(Grover's quantum search algorithm produces precisely such an output, with the additional restriction that our α_x are uniform within each of the partitions). Our goal is to boost the probability of observing a good state, which is

$$\Pr[\text{observing a good state}] = \sum_{x \in \text{GOOD}} |\alpha_x|^2 .$$

This can be done by applying the technique of Amplitude Amplification, which we will briefly outline since its a generalization of Grover's algorithm. The notation here is the same as the one used in our previous lecture.

Consider the state after we apply A. Just as in Grover's algorithm we can think of the state as a point on the unit circle on $\mathbb{R}^2$ with axes denoted by B and C (bad and correct), where the angle between the point and the axes give the probabilities of observing a bad or good state.

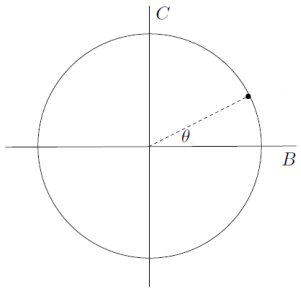


Figure 1: State after one run of A

As in Grover's algorithm, in each step we will be applying

1. Flips about the B axis
2. Rotation about the axis defined by the point.

To implement the flips, we simply need to apply U_f (phase kickback). For the rotations, we will

1. apply A^{-1} to 'shift' the axis defined by the point to the B axis,
2. perform a reflection about the B axis, denoted by $U_{|0^n\rangle}$, and finally
3. apply A.

Hence, we can describe each iteration of amplitude amplification as:

$$AU_{|0^n\rangle}A^{-1}U_f$$

(Note that if we consider Grover's algorithm in this framework, the A here is simply the Hadamard gate $H^{\otimes n}$.)

Repeating the same analysis as last lecture, we know the number of iterations required is $O(\frac{1}{\sqrt{p}})$, where p is the probability of observing a good state after one run of A. This is again a quadratic speedup over the classical case which requires $\Omega(\frac{1}{p})$ trials.

1.2 Finding a Witness for an NP Problem

Consider the Satisfiability problem. In the classical setting, the brute force approach would take $O(2^n)$ trial to obtain a valid assignment. However, using Grover's algorithm, we can search over the space of all assignments and obtain a valid assigment in $O(2^{\sqrt{n}})$ trials. While this does not necessarily mean that for any classical algorithm for SAT we can always find a quantum algorithm that gives a quadratic speedup, this has been true for known algorithms. Current known deterministic methods for solving SAT gives us a search space that we can recast in a quantum setting to allow Grover's algorithm to work efficiently.

1.3 Unstructured Database Search

One can view Grover's algorithm as a way to search over an unstructured database where the keys are precisely the boolean strings x of length n, corresponding to the 2^n base states $|x\rangle$. While this is often presented as an application of the algorithm (Grover's original paper does this), this is impractical in reality. Firstly, it is unlikely that a set of real-world data has no intrinsic structure. Secondly, both the number of quantum gates required and the time needed to transform the real-world data into the appropriate form for the algorithm will be at least linear in the number of inputs.

1.4 Deciding $OR(f)$

We have been considering the search problem of finding an input x such that a given function $f : \{0,1\}^n \to \{0,1\}$ evaluates to $f(x) = 1$. We can consider a related decision problem $OR(f)$: Given the function $f : \{0,1\}^n \to \{0,1\}$, does there exist an $x \in \{0,1\}^n$ such that $f(x) = 1$? (This is equivalent to computing the boolean OR over all possible $f(x)$).

It is clear that the search problem is at least as difficult as deciding $OR(f)$, and we will make use of this fact to obtain a lower bound for quantum search.

2 Tight Lower Bound for Quantum Search

From the previous lecture we know that Grover's algorithm runs in $O(\sqrt{N})$, where $N = 2^n$ denotes the number of binary strings of length n. We will show that this is optimal by proving the following theorem:

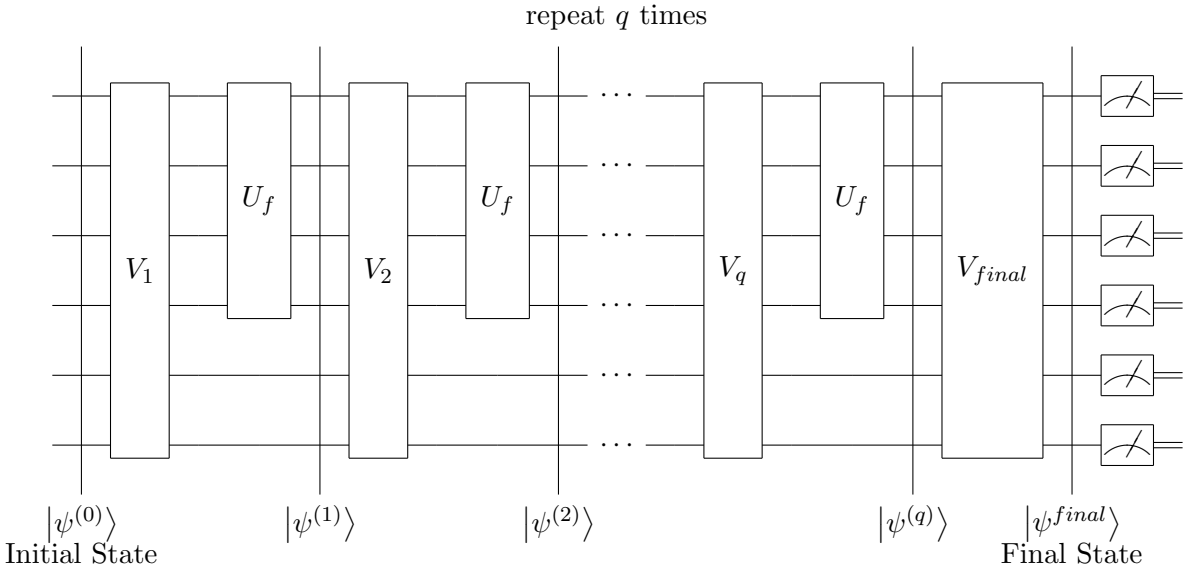
Theorem 1. *Any quantum black-box algorithm that decides $OR(f)$ with constant error $\epsilon < \frac{1}{2}$ needs to make $\Omega(\sqrt{N})$ queries to f.*

2.1 Structure of Quantum Circuit

To prove Theorem 1, we first need to consider the structure of any quantum circuit that makes q many queries. Each quantum circuit consists of two types of operators:

1. Unitary operators V_i, independent of f,
2. Oracle U_f, which provides black-box access to f.

We will also assume that we postpone observation of the system till the end. Hence, our quantum circuits have the following form:



Hence, the state after the ith U_f gate $\left|\psi^{(i)}\right\rangle$ is given by

$$\left|\psi^{(i)}\right\rangle = (U_f \otimes I) \cdot V_i \cdot \ldots \cdot (U_f \otimes I) \cdot V_1 \cdot \left|\psi^{(0)}\right\rangle,$$

where $\left|\psi^{(0)}\right\rangle = |0 \ldots 0\rangle$ without loss of generality.

2.2 Proof Idea for Theorem 1

Given any two functions f, g such that $OR(f) \neq OR(g)$, the corresponding final states have to be almost orthogonal for us to observe the correct answer with high probability. Therefore, we can

prove the theorem by picking an 'adversarial' set of functions f such that any quantum circuit that correctly decides $OR(f)$ for this set will require $\Omega(\sqrt{N})$ queries.

We let $\left|\psi^{(i)}\right\rangle$ denote the states for the function f such that $OR(f) = 0$ and $\left|\psi_{\tilde{x}}^{(i)}\right\rangle$ for function $f_{\tilde{x}}$ such that $f_{\tilde{x}}(x) = 1 \Leftrightarrow x = \tilde{x}$. Note that $\left|\psi_{\tilde{x}}^{(0)}\right\rangle = \left|\psi^{(0)}\right\rangle$, but $\left|\psi_{\tilde{x}}^{final}\right\rangle$ and $\left|\psi^{final}\right\rangle$ have to be nearly orthogonal. Only the oracle gate U_f can affect the angles, and we will show that any query can increase the angles between $\left|\psi_{\tilde{x}}^{(i)}\right\rangle$ and $\left|\psi^{(i)}\right\rangle$ by a small factor on average over all $\tilde{x}$. This means many queries are required, giving us a lower bound.

2.3 Proof of Theorem 1

We will begin by making formal the 'almost orthogonal' condition. Consider the distance between these two probability distributions:

$$\begin{aligned} D\left(\Pr_{\psi^{final}}, \Pr_{\psi_{\tilde{x}}^{final}}\right) \geq & \left|\Pr[\text{algorithm outputs 0 on } f] - \Pr[\text{algorithm outputs 0 on } f_{\tilde{x}}]\right| + \\ & \left|\Pr[\text{algorithm outputs 1 on } f] - \Pr[\text{algorithm outputs 1 on } f_{\tilde{x}}]\right| \end{aligned}$$

Since we want the error rate of the algorithm to be a constant factor ϵ, we obtain

$$D\big(\Pr_{\psi^{final}}, \Pr_{\psi_{\tilde{x}}^{final}}\big). \geq 2(1-2\epsilon)$$

which implies we need

$$\left\| \left|\psi^{final}\right\rangle - \left|\psi_{\tilde{x}}^{final}\right\rangle \right\| \geq 2(1-\delta) \tag{1}$$

for some $\delta > 0$ that is dependent only on ϵ.

Exercise 1. *Verify Equation 1 and determine $\delta(\epsilon)$ (there is a simple expression for $\delta(\epsilon)$).*

Finally, we can begin the proof of Theorem 1.

Proof. We will now put an upper bound on how much the norm $\left\| \left|\psi^{(i)}\right\rangle - \left|\psi_{\tilde{x}}^{(i)}\right\rangle \right\|$ in can change in any one step of the quantum algorithm (where each step consists of a unitary gate V_i and the U_f gate directly after it), thereby showing any quantum algorithm requires $q = \Omega(\sqrt{N})$ iterations to satisfy Equation 1. Since

$$\begin{aligned} \left|\psi^{(i)}\right\rangle &= (I \otimes I)V_i \left|\psi^{(i-1)}\right\rangle \\ \left|\psi_{\tilde{x}}^{(i)}\right\rangle &= (U_{f_{\tilde{x}}} \otimes I)V_i \left|\psi_{\tilde{x}}^{(i-1)}\right\rangle, \end{aligned}$$

we have

$$\left\| \left|\psi^{(i)}\right\rangle - \left|\psi_{\tilde{x}}^{(i)}\right\rangle \right\| = \left\| \left((I \otimes I)V_i \left|\psi^{(i-1)}\right\rangle - (U_{f_{\tilde{x}}} \otimes I)V_i \left|\psi_{\tilde{x}}^{(i-1)}\right\rangle \right) \right\|$$

This equation can be simplified by applying the triangle inequality and removing the unitary terms (which does not affect the norm):

$$\begin{aligned} \left\| \left|\psi^{(i)}\right\rangle - \left|\psi_{\tilde{x}}^{(i)}\right\rangle \right\| &\leq \left\| (I \otimes I)V_i \left|\psi^{(i-1)}\right\rangle - (U_{f_{\tilde{x}}} \otimes I)V_i \left|\psi^{(i-1)}\right\rangle \right\| + \left\| (U_{f_{\tilde{x}}} \otimes I)V_i \left(\left|\psi^{(i-1)}\right\rangle - \left|\psi_{\tilde{x}}^{(i-1)}\right\rangle \right) \right\| \\ &= \underbrace{\left\| (I \otimes I)V_i \left|\psi^{(i-1)}\right\rangle - (U_{f_{\tilde{x}}} \otimes I)V_i \left|\psi^{(i-1)}\right\rangle \right\|}_{B} + \left\| \left|\psi^{(i-1)}\right\rangle - \left|\psi_{\tilde{x}}^{(i-1)}\right\rangle \right\|. \end{aligned} \tag{2}$$

We now analyze the term B in Equation (2), which gives us an upper bound on the change in norm. Let

$$V_i \left|\psi^{(i-1)}\right\rangle = \sum_z \alpha_z |z\rangle$$

where $|z\rangle = |xbu\rangle$, such that $|xb\rangle$ represents the input into U_f and $|b\rangle$ represents the ancilla qubit that records the output of U_f. Note that

$$U_{f_{\tilde{x}}} \otimes I : |xbu\rangle \mapsto \begin{cases} \left|x\bar{b}u\right\rangle & \text{if } x = \tilde{x} \\ |xbu\rangle & \text{if } x \neq \tilde{x} \ . \end{cases}$$

We can now proceed to bound B:

$$\begin{aligned} B &= \left\| \left[(I \otimes I) - (U_{f_{\tilde{x}}} \otimes I) \right] \sum \alpha_z |z\rangle \right\| \\ &= \left\| \left[(I \otimes I) - (U_{f_{\tilde{x}}} \otimes I) \right] \sum_{z=\tilde{x}bu} \alpha_z |z\rangle \right\| \\ &= \sqrt{\sum_{z=\tilde{x}bu} \left(\alpha_{\tilde{x}bu} - \alpha_{\tilde{x}\bar{b}u}\right)^2} \\ &\leq \sqrt{\sum_{z=\tilde{x}bu} 2\left(|\alpha_{\tilde{x}bu}|^2 + \left|\alpha_{\tilde{x}\bar{b}u}\right|^2\right)} \\ &= \sqrt{4 \sum_{z=\tilde{x}bu} |\alpha_{\tilde{x}bu}|^2} \\ &= 2\sqrt{\Pr[i^{\text{th}} \text{ query for } f \equiv 0 \text{ is } \tilde{x}]}. \end{aligned} \tag{3}$$

We return to the term $\left\| |\psi^{final}\rangle - \left|\psi^{final}_{\tilde{x}}\right\rangle \right\|$ which, by removing unitary transformations, is simply $\left\| |\psi^{(q)}\rangle - \left|\psi^{(q)}_{\tilde{x}}\right\rangle \right\|$. By combining Equations (2) and (3), this gives us

$$\begin{aligned} \left\| \left|\psi^{final}\right\rangle - \left|\psi^{final}_{\tilde{x}}\right\rangle \right\| &= \left\| \left|\psi^{(q)}\right\rangle - \left|\psi^{(q)}_{\tilde{x}}\right\rangle \right\| \\ &\leq 2\sum_{i=1}^{q} \sqrt{\Pr[i^{\text{th}} \text{ query for } f \equiv 0 \text{ is } \tilde{x}]} + \underbrace{\left\| \left|\psi^{(0)}\right\rangle - \left|\psi^{(0)}_{\tilde{x}}\right\rangle \right\|}_{=0}. \end{aligned} \tag{4}$$

which is true for every possible $\tilde{x}$. While the probability term in Equation (4) might be large for particular $\tilde{x}$, on average they have to be small since they add up to 1. So, we sum over all $\tilde{x}$ and apply Cauchy-Schwarz inequality to obtain:

$$\begin{aligned} \sum_{\tilde{x}} \left\| \left|\psi^{final}\right\rangle - \left|\psi^{final}_{\tilde{x}}\right\rangle \right\| &\leq 2\sum_{i=1}^{q} \sum_{\tilde{x}} \sqrt{\Pr[i^{\text{th}} \text{ query for } f \equiv 0 \text{ is } \tilde{x}]} \\ &\leq 2\sum_{i=1}^{q} \left(1 \cdot \sqrt{N}\right) \\ &\leq 2q\sqrt{N} \end{aligned} \tag{5}$$

Finally, we combine the lower and upper bounds (Equations (1) and (5) respectively) to get

$$N \cdot 2(1-\delta) \leq 2q\sqrt{N} \qquad \Rightarrow \qquad q \geq (1-\delta)\sqrt{N},$$

so $q = \Omega(\sqrt{N})$, as desired. □

2.4 Conclusion

In the proof above, we used an adversial argument - we chose a relatively small set of functions that is easy to analyze, where distinguishing between those with different outputs requires many queries. This is a simplified form of the quantum adversial argument, which is a generalization of the method used in the classical setting.

In the next lecture, we will outline two other methods to obtain lower bounds in the quantum black-box model - the generalized quantum adversial method and the polynomial method.