CS 410/510 Introduction to Quantum Computing

Lecture 10

Portland State U, Spring 2017 *May 04, 2017*
Lecturer: Fang Song *Scribe: Asher Toback*

Version: October 9, 2017

In this lecture we'll be discussing why Grover's search is the optimal searching algorithm in the query model of quantum computing. It's assumed the reader already has a working knowledge of how Grover's Searching Algorithm works; namely, that it can be done within $O(\sqrt{N})$ querries, where $N = 2^n$. We'll show that any searching algorithm must take a minimum of this many querries, and thus Grover's is optimal. We'll do this by arguing what a general searching algorithm must look like, prove two lemmas, and then use these lemmas to prove the optimality of Grover. Let's begin.

1 General Search Algorithm Setup

We'll be concerning ourselves with the strictest search algorithm, where there is only one marked element that we're searching for. Namely, we're given a boolean function f

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

where $\exists! r$ such that $f(r) = 1$. We also assume we have access to an oracle O_f such that

$$|x\rangle \xrightarrow{O_f} (-1)^{f(x)} |x\rangle .$$

Because r is the only element such that $f(r) = 1$, we see that if we had a superposition $\sum \alpha_x |x\rangle$ and applied O_f you get

$$\sum \alpha_x |x\rangle \xrightarrow{O_f} \begin{cases} -\alpha_r |r\rangle & \text{if } x = r \\ \alpha_x |x\rangle & \text{if } x \neq r \end{cases}$$

Now, we're assuming α_r is small, otherwise we wouldn't need a searching algorithm. Thus, when we apply O_f to our superposition, the superposition won't be affected very much. So, intuitively, we must spend many querries to find r. We'll show that it requires $\sqrt{N}$.

2 Theorem Setup

We claim that any k-query quantum search algorithm has the circuit form.

$$|0^m\rangle \xrightarrow{U_0} \xrightarrow{O_f} \xrightarrow{U_1} \xrightarrow{O_f} \dots \xrightarrow{U_k} \xrightarrow{O_f} |\psi_r^{(k)}\rangle$$

Namely, you apply an arbitrary unitary followed by a call to the oracle to the input state, and you do this k times. It should be noted that each unitary could be different. We then compare this to the circuit:

$$|0^m\rangle \xrightarrow{U_0}\xrightarrow{\mathbb{I}}\xrightarrow{U_1}\xrightarrow{\mathbb{I}} \ldots \xrightarrow{U_k}\xrightarrow{\mathbb{I}} |\phi^{(k)}\rangle$$

Where $\mathbb{I}$ is the identity matrix and the unitaries applied are identical to the ones applied in the previous circuit. Now, we will try to show that $\exists r$ such that $\| |\psi_r^{(k)}\rangle - |\phi^{(k)}\rangle \| \leq 2k/2\sqrt{N}$, where $\|.\|$ is the euclidean distance metric. Intuitively this is saying that these two processes are practically indistinguishable. However, since we aren't querrying the oracle in the second circuit, this means we aren't learning anything about r, the marked element. Thus, we're essentially saying we learn nothing about r from $|\psi_r^{(k)}\rangle$ unless we query enough times so that $2k/2\sqrt{N}$ is big. But this would mean we'd need to query close to $\sqrt{N}$ times, which is just what we want to show. In order to prove this we must prove two lemmas first. Before we do, let's introduce two simplifying notations. First:

$$D_r^{(i)} := \| |\psi_r^{(i)}\rangle - |\phi^{(i)}\rangle \|$$

This is simply comparing the states after the i^{th} step of the two circuits. With this notation we then want to show that $\exists r$ such that $D_r^{(k)} \leq 2k/2\sqrt{N}$. And lastly:

$$E_r^{(j-1)} := \|O_f |\phi^{(j-1)}\rangle - |\phi^{(j-1)}\rangle \|$$

This equation can be thought of as the *error* the second circuit made by not querrying the oracle O_f. With these let's proceed.

3 Lemma 1

We'll first prove that

$$D_r^{(j)} \leq D_r^{(j-1)} + E_r^{(j-1)}$$

for $j = \{0, 1, ..., k-1\}$. Let's first rewrite $D_r^{(j)}$. Clearly the j^{th} step of each circuit can be thought of as applying U^j and either O_f or $\mathbb{I}$, depending on which circuit you're in, to the $(j-1)^{th}$ step of the circuit. Namely:

$$\begin{aligned} D_r^{(j)} &= \| |\psi_r^{(j)}\rangle - |\phi^{(j)}\rangle \| \\ &= \|U_j O_f |\psi_r^{(j-1)}\rangle - U_j \mathbb{I} |\phi^{(j-1)}\rangle \| \\ &= \|U_j O_f |\psi_r^{(j-1)}\rangle - U_j O_f |\phi^{(j-1)}\rangle + U_j O_f |\phi^{(j-1)}\rangle - U_j \mathbb{I} |\phi^{(j-1)}\rangle \| \end{aligned}$$

Then, by the triangle inequality:

$$\leq \|U_jO_f(|\psi_r^{(j-1)}\rangle - |\phi^{(j-1)}\rangle)\| + \|U_j(O_f\,|\phi^{(j-1)}\rangle - \mathbb{I}\,|\phi^{(j-1)}\rangle)\|$$

However, we know that unitaries don't effect the length of these vectors. Thus:

$$= \||\psi_r^{(j-1)}\rangle - |\phi^{(j-1)}\rangle\| + \|O_f\,|\phi^{(j-1)}\rangle - \mathbb{I}\,|\phi^{(j-1)}\rangle\|$$
$$= D_r^{(j-1)} + E_r^{(j-1)}$$

Which is just what we wanted to show.

4 Lemma 2

We'll now prove that

$$E_r^{(j)} \leq 2\left|\alpha_r^{(j)}\right|$$

for $j = \{0, 1, ..., k-1\}$. Let's first rewrite $E_r^{(j)}$ using it's definition.

$$E_r^{(j)} = \|O_f\,|\phi^{(j)}\rangle - |\phi^{(j)}\rangle\|$$

But, by analysis given in section 2 of this paper, we know that O_f only truly affects the marked element, so we know that we can rewrite this as:

$$= \|(\sum_{x\neq r}\alpha_x^{(j)}\,|x\rangle - \alpha_r^{(j)}\,|r\rangle) - (\sum_{x\neq r}\alpha_x^{(j)}\,|x\rangle + \alpha_r^{(j)}\,|r\rangle)\|$$
$$= 2\|\alpha_r^{(j)}\,|r\rangle\|$$

But $\||r\rangle\| = 1$. Thus,

$$= 2\left|\alpha_r^{(j)}\right|$$

Just as we wanted.

5 Theorem Proof

Let's now use these two lemmas to prove our claim. To reiterate, we want to prove that $D_r^{(k)} \leq 2k/2\sqrt{N}$. Let's first use the definition of $D_r^{(k)}$:

$$\||\psi_r^{(j)}\rangle - |\phi^{(j)}\rangle\| = D_r^{(j)}$$

Then, by consecutive applications of Lemma 1 we get:

$$\begin{gathered}\leq D_r^{(j-1)} + E_r^{(j-1)} \leq D_r^{(j-2)} + E_r^{(j-2)} + E_r^{(j-1)} \\ \leq ... \leq \\ \leq D_r^{(0)} + \sum_{j=0}^{k-1} E_r^{(j)}\end{gathered}$$

But we know that $D_r^{(0)} = 0$ because both circuits have the same initial state. Thus:

$$= \sum_{j=0}^{k-1} E_r^{(j)}$$

This should intuitively make sense. Essentially, the difference between the ending states of the two circuits is just the error incurred at each step by not applying O_f. Now that we've proved this, let's sum over all the elements in our space, of which we're trying to find r, and then apply Lemma 2.

$$\sum_{r \in \{0,1\}^n} D_r^{(k)} \leq \sum_r \sum_{j=0}^{k-1} E_r^{(j)}$$

We know this above statement is true because we literally just proved it. Then, re-ordering the sums and using Lemma 2:

$$= \sum_{j=0}^{k-1} \sum_r E_r^{(j)} \leq \sum_{j=0}^{k-1} \sum_r 2 \left| \alpha_r^{(j)} \right|$$

We now must find a clever way to rewrite $\sum_{r \in \{0,1\}^n} \left| \alpha_r^{(j)} \right|$. Luckily, we know that $\sum_{r \in \{0,1\}^n} \left| \alpha_r^{(j)} \right|^2 = 1$ because our vectors must be normalized. So, we essentially need to come up with a mathematical fact that says if we have

$$\begin{gathered}a_1^2 + a_2^2 + ... + a_N^2 = 1 \;\; \text{then} \\ a_1 + a_2 + ... + a_N \leq \;?\; = \sqrt{N}\end{gathered}$$

Luckily, there exists such a fact. It takes advantage of the Cauchy-Schwarz inequality, which is:

$$|\langle u| \, |v\rangle| \leq \|u\| * \|v\|$$

Where $\langle . | \, | . \rangle$ represents the inner product of our metric. In our case, we can use this in the following manner:

$$\left| \langle 1| \, |\alpha_x^{(j)}\rangle \right| \leq \sqrt{\sum_{x \in \{0,1\}^n} 1^2} = \sqrt{N}$$

We're now essentially finished. We can thus do the following rewrite:

$$D_r^{(k)} \leq \sum_{j=0}^{k-1} \sum_{r \in \{0,1\}^n} 2 \left| \alpha_r^{(j)} \right| \leq \sum_{j=0}^{k-1} 2\sqrt{N} = 2k\sqrt{N}$$

However, there's only 1 element r out of N that is marked. So, on average, the difference between the ending states of the two circuits is:

$$\| |\psi_r^{(k)}\rangle - |\phi^{(k)}\rangle \| = D_r^{(k)} \leq 2k\sqrt{N}/N = 2k/\sqrt{N}$$

This concludes our proof as this is just what we wanted to show.