CSE 599d - Quantum Computing
Grover's Algorithm

Dave Bacon
Department of Computer Science & Engineering, University of Washington

After Peter Shor demonstrated that quantum computers could efficiently factor, great interest arose in finding other problems for which quantum algorithms could outperform the best known classical algorithms. One of the early offshoots of this work was an algorithm invented by Lov Grover in 1996. Here we will describe Grover's algorithm and show that it is, in a query complexity manner, the optimal quantum algorithm.

I. GROVER'S ALGORITHM

Suppose that we have a function $f(x)$ from $\{0,1\}^n$ to $\{0,1\}$ which is zero on all inputs except for a single (marked) item x_0: $f(x)=\delta_{x,x_0}$. By querying this function you wish to find the marked item x_0. This is like finding a needle in a haystack. Certainly if you have no information about the particular x_0, then finding this marked item is very difficult. In the worst case it will take 2^n-1 queries to find x_0 for a deterministic algorithm. Well what if we turn this haystack into a quantum haystack? This is the question that Grover asked.

In the quantum version of Grover's problem we have access to a function which computes $f(x)$ in our standard reversible manner:

$$U_f=\sum_{x,y\in\{0,1\}^n}|x\rangle\langle x|\otimes|y\oplus f(x)\rangle\langle y|. \tag{1}$$

Using our standard trick of phase kickback, we can feed $|-\rangle=\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$ into the y register of this unitary. If we do this, then the effect on the first register is still unitary and is given by

$$V_f=\sum_{x\in\{0,1\}^n}(-1)^{f(x)}|x\rangle\langle x|=\sum_{x\in\{0,1\}^n}(-1)^{\delta_{x,x_0}}|x\rangle\langle x| \tag{2}$$

We will therefore assume that we have query access to V_f.

Consider the two vectors $|x_0\rangle$ and $|\psi\rangle=\frac{1}{\sqrt{2^n}}\sum_{x\in\{0,1\}^n}|x\rangle$. These vectors are not orthogonal, $\langle x_0|\psi\rangle=\frac{1}{\sqrt{2^n}}$, however they are linearly independent. Thus we can for a two dimensional basis for the subspace of vectors which are linear superpositions of these two basis elements. One such choice of basis consists of $|x_0\rangle$ and

$$|\psi'\rangle=\frac{1}{\sqrt{2^n-1}}\sum_{x\in\{0,1\}^n,x\neq x_0}|x\rangle \tag{3}$$

We note that

$$|\psi\rangle=\sqrt{\frac{2^n-1}{2^n}}|\psi'\rangle+\frac{1}{\sqrt{2^n}}|x_0\rangle. \tag{4}$$

Now notice that V_f preserve the subspace $|x_0\rangle,|\psi'\rangle$:

$$\begin{aligned}V_f|x_0\rangle&=-|x_0\rangle\\ V_f|\psi'\rangle&=|\psi'\rangle.\end{aligned} \tag{5}$$

Now consider the unitary $W=2|\psi\rangle\langle\psi|-I$. This is unitary? Yes because $(2|\psi\rangle\langle\psi|-I)(2|\psi\rangle\langle\psi|-I)^\dagger=4|\psi\rangle\langle\psi|-2|\psi\rangle\langle\psi|-2|\psi\rangle\langle\psi|+I=I$. W also preserves the subspace spanned by $|x_0\rangle$ and $|\psi\rangle$:

$$\begin{aligned}W|x_0\rangle&=(2|\psi\rangle\langle\psi|-I)|x_0\rangle=\frac{2}{\sqrt{2^n}}|\psi\rangle-|x_0\rangle=\frac{2\sqrt{(2^n-1)}}{2^n}|\psi'\rangle+\left(\frac{2}{2^n}-1\right)|x_0\rangle\\ W|\psi'\rangle&=(2|\psi\rangle\langle\psi|-I)|\psi'\rangle=2\sqrt{\frac{2^n-1}{2^n}}|\psi\rangle-|\psi'\rangle=\left(\frac{2(2^n-1)}{2^n}-1\right)|\psi'\rangle+\frac{2\sqrt{2^n-1}}{2^n}|x_0\rangle\\ &=-\left(\frac{2}{2^n}-1\right)|\psi'\rangle+\frac{2\sqrt{2^n-1}}{2^n}|x_0\rangle\end{aligned} \tag{6}$$

We can rewrite this as a rotation about an angle θ:

$$\begin{aligned} W|x_0\rangle &= -\cos\theta|x_0\rangle + \sin\theta|\psi'\rangle \\ W|\psi'\rangle &= \sin\theta|x_0\rangle + \cos\theta|\psi'\rangle \end{aligned} \tag{7}$$

where

$$\sin\theta = \frac{2\sqrt{2^n-1}}{2^n} \tag{8}$$

If we combine W and V_f together, we can form what is called Grover's iterate, $G = WV_f$. This acts as

$$\begin{aligned} G|x_0\rangle &= \cos\theta|x_0\rangle - \sin\theta|\psi'\rangle \\ G|\psi'\rangle &= \sin\theta|x_0\rangle + \cos\theta|\psi'\rangle \end{aligned} \tag{9}$$

Thus we see that the Grover's iterate takes $|\psi'$ and rotates it towards $|x_0\rangle$.

Now we can describe Grover's algorithm. We are trying to find $|x_0\rangle$. We cannot prepare $|\psi'\rangle$ or $|x_0\rangle$, but we can prepare the superposition of them $|\psi\rangle$ with no knowledge of $|x_0\rangle$. This has most of its amplitude on $|\psi'\rangle$. Now by applying the Grover iterate, we can rotate this initial state towards the state with more amplitude on $|x_0\rangle$. Notice that each Grover iterate requires a query of the function. So is it possible to use Grover's iterate to perform this rotation with enough fidelity and few enough queries to beat out the classical case? Well let's see!

The initial state is

$$|\psi\rangle = \sqrt{\frac{2^n-1}{2^n}}|\psi'\rangle + \frac{1}{\sqrt{2^n}}|x_0\rangle. \tag{10}$$

Apply G k times will achieve the rotation

$$\begin{aligned} G^k|x_0\rangle &= \cos k\theta|x_0\rangle - \sin k\theta|\psi'\rangle \\ G^k|\psi'\rangle &= \sin k\theta|x_0\rangle + \cos k\theta|\psi'\rangle \end{aligned} \tag{11}$$

For what k is this just the rotation which swaps $|x_0\rangle$ and $|\psi'\rangle$ (with a phase)? This is when $\sin\theta k = \frac{\pi}{2}$. Now lets be physicists and use the small angle approximation for sin: $\sin\theta = \frac{2\sqrt{2^n-1}}{2^n}$ implies $\theta \approx \frac{2\sqrt{2^n-1}}{2^n}$ for large n. Thus we need

$$k\frac{2\sqrt{2^n-1}}{2^n} \approx \frac{\pi}{2} \tag{12}$$

or

$$k \approx \frac{\pi}{4}\frac{2^n}{\sqrt{2^n-1}} \approx \frac{\pi}{4}\sqrt{2^n} \tag{13}$$

Thus we see that if we perform $\sqrt{2^n}$ rotations, we can approximately flip $|x_0\rangle$ and $|\phi'\rangle$. If we originally started with $|\psi\rangle$, then with very high probability we will obtain $|x_0\rangle$ after this rotation. We haven't dotted our probabilities or crossed our bounds here, but with a little work this can be done. What is important to notice is that there is a quadratic speedup in the number of queries needed to search compared to the classical situation.

What about the circuit costs for Grover's algorithm? Well there is the query cost, which we've already computed. And then there is W. How do we efficiently implement $W = 2|\psi\rangle\langle\psi| - I$? Well notice that this is $W = H^{\otimes n}(2|0\rangle\langle 0| - I)H^{\otimes n}$ where H is our friend the Hadamard gate. Thus we need to know how to implement $(2|0\rangle\langle 0| - I$ efficiently. This unitary maps all of states to themselves, with $|0\rangle\langle 0|$ acquiring no phase and all of the other computational basis states acquiring a -1 phase. It is useful to introduce a global phase which reverse this: $-2|0\rangle\langle 0| + I$. One way to implement this gate is as follows. First note that you can construct a multiple controlled operation using Tofolli gates, a single controlled gate, and some extra ancilla workspace qubits initialized to $|0\rangle$ (they will also end up being $|0\rangle$.) To do this, for the qubits you want to condition on, compute the AND of the first two qubits and put it into an ancilla workspace qubit using a Tofolli, then compute the AND of this qubit and the third qubit into a second workspace qubit using a Tofolli. Continuing in this way you see that in n qubits you can obtain the AND of all the control bits calculated in some ancilla register. Conditional on this qubit perform the desired controlled gate on the target qubit. Then run in reverse the Tofolli circuit you have perform, thus erasing the garbage in the aniclla qubits. Now to implemen $-2|0\rangle\langle 0| + I$, notice that if you perform a $n-1$ qubit controlled Z (Z is the Pauli phase operator), then this gate is $-2|1^n\rangle\langle 1^n| + I$. Simply applying $X^{\otimes n}$ before and after this gate turns this into the desired gate (up to the global phase.) Thus we see that we can implement W using $O(2n)$ elementary gates. Note that this operation W is sometimes called the "invert about the mean" operation. I leave it to you to determine why it has this name.

II. OPTIMALITY OF GROVER'S ALGORITHM

So we have seen that Grover's algorithm offers a speedup which is quadratic in finding the needle in the haystack. A good question to ask is whether we could have done any better. There is a distinct way in which we could not have done better and that is if we require ourselves to query f in the manner we have described then Grover's algorithm is optimal. Let's prove this.

Suppose we start our attempt to determine f by starting in the state $|\psi\rangle$. Now we apply a sequence of oracle calls followed by arbitrary (but independent of the oracles) unitaries. Suppose the oracle we use is the phase-kicked back oracle

$$V_f = I - 2|x_0\rangle\langle x_0| \tag{14}$$

Then the algorithm will produce the state,

$$|\phi^k_{x_0}\rangle = U_k V_f U_{k-1} V_f \cdots U_2 V_f U_1 V_f |\psi\rangle \tag{15}$$

Notice that V_f depends on x_0. If we hadn't called the oracle, we would be in the state

$$|\phi^k\rangle = U_k U_{k-1} \cdots U_2 U_1 |\psi\rangle. \tag{16}$$

This setup is the most general setup for querying the oracle. We want to show that when we do this it is impossible to make the $|\phi^k_{x_0}\rangle$ orthogonal enough such that a measurement can distinguish these different states (for different x_0 values) unless we make k big enough (i.e the number of queries is high enough.) We do this by showing that we need k large to make

$$r_k = \sum_{x_0 \in \{0,1\}^n} || |\phi^k_{x_0}\rangle - |\phi^k\rangle ||^2 \tag{17}$$

is small unless k is large enough. Then we will show that this implies that when this quantity is small there is no way to identify the x_0 with high probability. Why do we choose $|\phi^k\rangle$? Well this is the state that would exist if no oracle calls were made, so measuring our distance from this state is measuring how much effect the oracle is having.

So first we want to bound r_k. To do this we show that $r_k \leq 4k^2$. This is clearly true for $k = 0$. We proceed inductively. First we use the fact that U_{k+1} is unitary:

$$|| |\phi^{k+1}_{x_0}\rangle - |\phi^{k+1}\rangle ||^2 = || U_{k+1} V_f |\phi^k_{x_0}\rangle - U_{k+1}|\phi^k\rangle ||^2 = || V_f |\phi^k_{x_0}\rangle - |\phi^k\rangle ||^2 \tag{18}$$

The fact that this expression is equal is a consequence of the fact that the $|\phi^k_{x_0}\rangle$ and $|\phi^k\rangle$ states cannot be made more orthogonal by a unitary rotation. The only operator which can make them more orthogonal is V_f. Now we reexpress this as

$$|| |\phi^{k+1}_{x_0}\rangle - |\phi^{k+1}\rangle ||^2 = || V_f (|\phi^k_{x_0}\rangle - |\phi^k\rangle) + (V_f - I)|\phi^k\rangle ||^2 \tag{19}$$

Next use the inequality $|| |a\rangle + |b\rangle ||^2 \leq || |a\rangle ||^2 + || |b\rangle ||^2 + 2 || |a\rangle || \, || |b\rangle ||$ to obtain

$$|| |\phi^{k+1}_{x_0}\rangle - |\phi^{k+1}\rangle ||^2 \leq || V_f (|\phi^k_{x_0}\rangle - |\phi^k\rangle) ||^2 + || (V_f - I)|\phi^k\rangle ||^2 + 2 || V_f (|\phi^k_{x_0}\rangle - |\phi^k\rangle) || \, || (V_f - I)|\phi^k\rangle || \tag{20}$$

Since V_f is unitary,

$$|| |\phi^{k+1}_{x_0}\rangle - |\phi^{k+1}\rangle ||^2 \leq || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) ||^2 + || (V_f - I)|\phi^k\rangle ||^2 + 2 || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) || \, || (V_f - I)|\phi^k\rangle || \tag{21}$$

Now $V_f - I = -2|x_0\rangle\langle x_0|$, so this becomes

$$|| |\phi^{k+1}_{x_0}\rangle - |\phi^{k+1}\rangle ||^2 \leq || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) ||^2 + || -2|x_0\rangle\langle x_0||\phi^k\rangle ||^2 + 2 || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) || \, || -2|x_0\rangle\langle x_0|\phi^k\rangle || \tag{22}$$

or

$$|| |\phi^{k+1}_{x_0}\rangle - |\phi^{k+1}\rangle ||^2 \leq || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) ||^2 + 4 || \langle x_0|\phi^k\rangle |x_0\rangle ||^2 + 4 || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) || \, || \langle x_0|\phi^k\rangle |x_0\rangle || \tag{23}$$

Summing this over x_0 we obtain

$$r_{k+1} \leq r_k + \sum_{x_0 \in \{0,1\}^n} \left(4 || \langle x_0|\phi^k\rangle |x_0\rangle ||^2 + 4 || (|\phi^k_{x_0}\rangle - |\phi^k\rangle) || \, || \langle x_0|\phi^k\rangle |x_0\rangle || \right) \tag{24}$$

Now $\sum_{x_0\in\{0,1\}^n} 4||\langle x_0|\phi^k\rangle|x_0\rangle||^2 = 1$, this becomes

$$r_{k+1} \leq r_k + 4 + \sum_{x_0\in\{0,1\}^n} 4||(|\phi^k_{x_0}\rangle - |\phi^k\rangle)||\ ||\langle x_0|\phi^k\rangle|x_0\rangle|| \tag{25}$$

or

$$r_{k+1} \leq r_k + 4 + \sum_{x_0\in\{0,1\}^n} 4||(|\phi^k_{x_0}\rangle - |\phi^k\rangle)||\ |\langle x_0|\phi^k\rangle| \tag{26}$$

Now the Cauchy-Schwarz inequality for a real space is

$$\left(\sum_{i=1}^n x_i y_i\right)^2 \leq \sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j^2 \tag{27}$$

Applying this to our sum we obtain

$$r_{k+1} \leq r_k + 4 + 4\left(\sum_{x_0\in\{0,1\}^n} ||(|\phi^k_{x_0}\rangle - |\phi^k\rangle)||^2\right)^{\frac{1}{2}} \left(\sum_{x_0\in\{0,1\}^n} |\langle x_0|\phi^k\rangle|^2\right)^{\frac{1}{2}} \tag{28}$$

or

$$r_{k+1} \leq r_k + 4 + 4\sqrt{r_k} \tag{29}$$

Now from our inductive hypothesis $r_k \leq 4k^2$, so $r_{k+1} \leq 4k^2 + 4 + 42k = 4(k^2 + 2k + 1) = 4(k+1)^2$. Thus we have showed that $r_k \leq 4k^2$ for all k.

The next part of the proof is to show that if r_k is small then we cannot distinguish the different vectors $|\phi^k_{x_0}\rangle$ with high probability. Suppose that we wish to identify x_0 with probability greater than one half. Then, without loss of generality assume that the basis we use to make this distinction is the computational basis. In order for this to succeed we require that $|\langle x_0|\phi^k_{x_0}\rangle|^2 \geq \frac{1}{2}$. This in turn implies that $|||\phi^k_{x_0}\rangle - |x_0\rangle||^2 = ((\langle\phi^k_{x_0}|-\rangle x_0|)(|\phi^k_{x_0}\rangle - |x_0\rangle) = 2 - \langle\phi^k_{x_0}|x_0\rangle - \langle x_0|\phi^k_{x_0}\rangle$. We can always adjust the global phase of the computational basis states to make the last two terms sum to a real number and our bound on probability then yields

$$|||\phi^k_{x_0}\rangle - |x_0\rangle||^2 \leq 2 - \sqrt{2} \tag{30}$$

Now we want to show that if this is true that r_k must be large. So we do the obvious and stick in one of those $|x_0\rangle$ states into our expression for r_k:

$$r_k = \sum_{x_0\in\{0,1\}^n} |||\phi^k_{x_0}\rangle - |x_0\rangle + |x_0\rangle - |\phi^k\rangle||^2 \tag{31}$$

which is bounded by

$$r_k \geq \sum_{x_0\in\{0,1\}^n} \left[|||\phi^k_{x_0}\rangle - |x_0\rangle||^2 + |||\phi^k\rangle - |x_0\rangle||^2 - 2|||\phi^k_{x_0}\rangle - |x_0\rangle||\ |||\phi^k\rangle - |x_0\rangle||\right] \tag{32}$$

The Cauchy-Schwarz inequality implies

$$\sum_{x_0\in\{0,1\}^n} |||\phi^k_{x_0}\rangle - |x_0\rangle||\ |||\phi^k\rangle - |x_0\rangle|| \leq \left[\sum_{x_0\in\{0,1\}^n} |||\phi^k_{x_0}\rangle - |x_0\rangle||^2 \sum_{x_0\in\{0,1\}^n} |||\phi^k\rangle - |x_0\rangle||^2\right]^{\frac{1}{2}} \tag{33}$$

Thus

$$r_k \geq \left(-\left[\sum_{x_0\in\{0,1\}^n} |||\phi^k_{x_0}\rangle - |x_0\rangle||^2\right]^{\frac{1}{2}} + \left[\sum_{x_0\in\{0,1\}^n} |||\phi^k\rangle - |x_0\rangle||^2\right]^{\frac{1}{2}}\right)^2 \tag{34}$$

Now let's bound that last term

$$\sum_{x_0\in\{0,1\}^n} |||\phi^k\rangle - |x_0\rangle||^2 = \sum_{x_0\in\{0,1\}^n} \langle\phi^k|\phi^k\rangle + \langle x_0|x_0\rangle - \langle\phi^k|x_0\rangle - \langle x_0|\phi^k\rangle = 22^n - \sum_{x_0\in\{0,1\}^n} \langle\phi^k|x_0\rangle + \langle x_0|\phi^k\rangle \tag{35}$$

Which we can bound as

$$\sum_{x_0\in\{0,1\}^n} |||\phi^k\rangle - |x_0\rangle||^2 \geq 22^n - 2\sum_{x_0\in\{0,1\}^n} |\langle\phi^k|x\rangle| \tag{36}$$

Cauchy-Schwarz yields

$$\left[\sum_{x\in\{0,1\}^n} |\langle x|\phi^k\rangle|\right]^2 \leq \sum_{x\in\{0,1\}^n} 1 \sum_{x\in\{0,1\}^n} |\langle x|\phi^k\rangle|^2 = 2^n \tag{37}$$

Thus we find

$$\sum_{x_0\in\{0,1\}^n} |||\phi^k\rangle - |x_0\rangle||^2 \geq 2(2^n - \sqrt{2^n}) \tag{38}$$

Back to our original equation we find

$$r_k \geq \left[((2-\sqrt{2})2^n)^{\frac{1}{2}} + (2(2^n - \sqrt{2^n}))^{\frac{1}{2}}\right]^2 \geq c2^n \tag{39}$$

where we work for sufficiently large n and c is a constant less than $c = (\sqrt{2-\sqrt{2}} - \sqrt{2})^2 \approx 0.42$.

Now putting things together, we know that in order to succeed with probability greater than one half for all inputs, we require that $r_k \geq c2^n$. But we know that $r_k \leq 4k^2$. This implies that we need

$$k \geq \sqrt{\frac{c2^n}{4}} \tag{40}$$

queries in order for the measurements to so succeed. Thus we see that we obtain a bound which is like Grover's algorithm: $\Omega(\sqrt{2^n})$ queries are needed quantum mechanically.

III. GROVER'S ALGORITHM FOR MULTIPLE NEEDLES IN A HAYSTACK

So far I have worked with a function which maps $\{0,1\}^n \to \{0,1\}$ and has only one marked item. But more generally we can design our Grovers algorithm to work with a function $\{0,1,\ldots,N-1\} \to \{0,1\}$ which is 0 on $N-M$ on this domain and is 1 on the rest of the M points of this domain. The goal then will be to find a single one of the inputs x such that $f(x)=1$. For now we will assume that we know M. Let S be the set of x such that $f(x)=1$. Then we can define the states

$$|x_0\rangle = \frac{1}{\sqrt{M}}\sum_{x\in S}|x\rangle \tag{41}$$

and

$$|\phi' = \frac{1}{\sqrt{N-M}}\sum_{x\notin S}|x\rangle \tag{42}$$

and again produce a Grover's iterate which rotates on this space. The argument is nearly identical, so I will not repeat it. The important point, though, is that the state will rotate by an angle satisfying

$$\sin\theta = \frac{2\sqrt{M(N-M)}}{N} \tag{43}$$

on this basis (with identical expressions for the rotation matrix.) We can again start in the equal superposition state $|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle = \sqrt{\frac{N-M}{N}} |\phi'\rangle + \sqrt{\frac{M}{N}} |x_0\rangle$. To maximize that we will measure a state from $|x_0\rangle$ we need to rotate through the angle

$$k\theta = \arccos^{-1} \sqrt{\frac{M}{N}} \tag{44}$$

where k is the number of such iterations and we will assume $M \leq \frac{N}{2}$ for the moment. The number of iterations needed is

$$k = \frac{1}{\theta} \arccos^{-1} \sqrt{\frac{M}{N}} \tag{45}$$

Normally we could calculate this number and run for this number of iterations. To obtain a bound on the number of iterations, it is enough to note that the arccos is at most $\frac{\pi}{2}$, so $k \leq \frac{\pi}{2\theta}$. Now $\sin^2 \frac{\theta}{2} = \frac{1}{2}(1 - \cos\theta) = \frac{M}{N}$. So since $\frac{\theta}{2} \geq \sin \frac{\theta}{2}$ (θ is positive) this implies that

$$\theta \geq 2\sqrt{\frac{M}{N}} \tag{46}$$

Putting this together we find that the number of iterations is bounded by

$$k \leq \frac{\pi}{4} \sqrt{\frac{N}{M}} \tag{47}$$

Now what happens in the case where M is greater than $N/2$? Well one way to deal with this case (assuming we know M) is simply to double the size of where we are searching and then to use the $M \leq N/2$ algorithm. This can be done with a quantum circuit which uses just a single extra qubit: just use this extra qubit as a control for the unitary you are querying.