A Theorist's Toolkit **(CMU 18-859T, Fall 2013)**

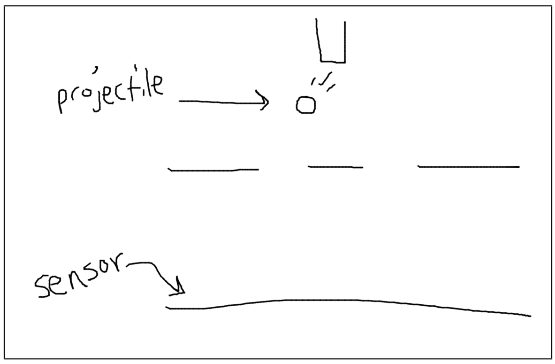
Lecture 22: Quantum Computing

November 20, 2013

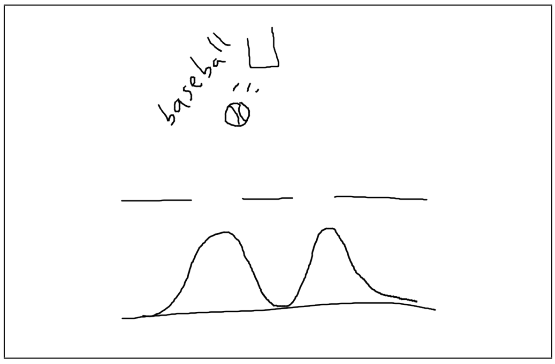
Lecturer: John Wright *Scribe: Elara Willett*

1 MOTIVATION

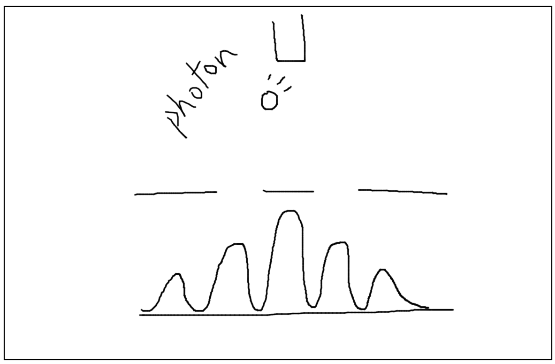
Quantum particles, such as photons, can occupy two states simultaneously. This principle, called *quantum superposition*, can be observed in the famous double-slip experiment. Suppose we have a wall with two slits. On one side, we project particles from a source between the two slits toward the wall. On the other side, there is a sensor that measures where the particles intersect a plane parallel to the wall.



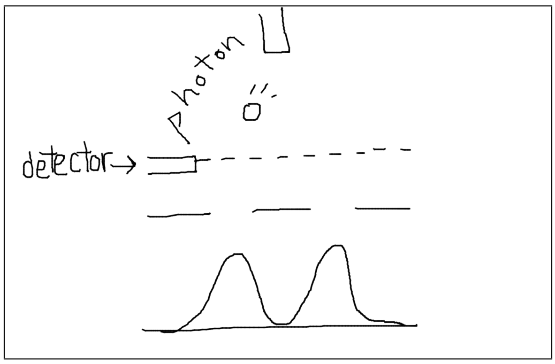
If we do this experiment with a baseball instead of a particle, the baseball intersects the plane at points directly across from the slits. So the baseball's measured location distribution looks like:



On the other hand, if we do this experiment with a photon, we find that the photon is most likely to be measured directly between the slits! The photon's measured location distribution looks like:



Moreover, if we add a detector into our experiment that measures the photon before it passes through the wall, then the quantum superposition goes away. That is, the distribution becomes:



This experiment shows that quantum particles can exist in two states simultaneously. The states in this experiment are Left and Right. In classical physics, a projectile can only travel through the left slit or the right slit, not both. That is, the classical projectile exists in only one of the two states, where quantum particles can exist in both states. Furthermore, when we check the state of a quantum particle, its superposition collapses, and it exists in only one state.

In the 1980's physicists suggested that to carry out this type of quantum experiment, we might build a quantum computer.

2 QUANTUM MECHANICS

The basic unit in quantum computers is called a qubit. There are two basic qubits corresponding to two states. The basic quits are the zero qubit, written $|0\rangle$ and the one quit, written $|1\rangle$. We also give names to two other important qubits,

$$|+\rangle := \tfrac{|0\rangle+|1\rangle}{\sqrt{2}} \qquad \text{and} \qquad |-\rangle := \tfrac{|0\rangle-|1\rangle}{\sqrt{2}}$$

A generic qubit is written as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha, \beta \in \mathbb{R}$ and $\alpha^2 + \beta^2 = 1$. The coefficients, α and β, are called *amplitudes*, and we can think of them as funny probabilities over the states.

(In general, α and β are allowed to be complex and then they must satisfy $|\alpha|^2+|\beta|^2 = 1$. However for this lecture, we will assume α and β are real.)

2.1 Physical Representations

The qubits can be physically represented by quantum particles that can occupy two states (or a superposition of the two states). For example, we could use

- a photon, where $|1\rangle$ and $|0\rangle$ are the polarizations of the photon, or
- an electron orbiting a nucleus, where $|1\rangle$ and $|0\rangle$ are the excited state and the ground state, respectively.

2.2 Mathematical Representations

We can represent qubits as unit vectors in $\mathbb{R}^2$. Namely, $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$. A generic qubit is

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix}\alpha\\\beta\end{pmatrix}$$

where $\alpha^2 + \beta^2 = 1$. Also,

$$\langle\psi| = \begin{pmatrix}\alpha\\\beta\end{pmatrix}^T = \begin{pmatrix}\alpha & \beta\end{pmatrix}$$

In the standard *bra-ket notation*, column vectors are called 'ket' and row vectors are called 'bra'. That is, $|x\rangle$ is read 'ket-x' and $\langle x|$ is read 'bra-x'.

Let $|x_1\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ and $|x_2\rangle = \beta_0|0\rangle + \beta_1|1\rangle$. Then we define

$$\langle x_1|x_2\rangle := \langle x_1||x_2\rangle = \begin{pmatrix}\alpha_0 & \alpha_1\end{pmatrix}\begin{pmatrix}\beta_0\\\beta_1\end{pmatrix} = \text{dot product}$$

Notice that $\alpha^2 + \beta^2 = 1$ is an equivalent condition to $\langle\psi|\psi\rangle = 1$, so for all ψ, $\langle\psi|\psi\rangle = 1$.

Furthermore, $\langle 0|1\rangle = 0$ and $\langle +|-\rangle = 0$, so both $|0\rangle, |1\rangle$ and $|+\rangle, |-\rangle$ form orthonormal bases for $\mathbb{R}^2$.

2.3 Measurement

Measurement of a qubit $|\psi\rangle$ should give:

- $|0\rangle$ if $|\psi\rangle = |0\rangle$,
- $|1\rangle$ if $|\psi\rangle = |1\rangle$, and
- $\begin{cases} |0\rangle & \text{w.p. } \alpha^2 \\ |1\rangle & \text{w.p. } \beta^2 \end{cases}$ if $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$.

A qubit $|\psi\rangle$ will collapse to the observed state once it is measured. That is, once you observe a state, the qubit becomes the state you observed.

We are particularly interested in $|+\rangle$ and $|-\rangle$ because for these qubits we observe $|0\rangle$ with probability $\frac{1}{2}$ and $|1\rangle$ with probability $\frac{1}{2}$. Thus, measuring $|+\rangle$ or $|-\rangle$ is like flipping a fair coin, and if we could construct these qubits, we could generate coin flips.

2.4 Evolution

In quantum mechanics, particles undergo transformations of state. A transformation operator U is a mapping from $|\psi\rangle$ to $U|\psi\rangle$.

The transformation operator U is realizable by quantum particles if and only if the following conditions hold:

1. U is linear. That is, $U(|x_1\rangle + |x_2\rangle) = U|x_1\rangle + U|x_2\rangle$. Equivalently, U can be written as a matrix.
2. U preserves state. That is, U maps valid qubits to valid qubits. Precisely, if $|\psi'\rangle = U|\psi\rangle$ and $\langle\psi|\psi\rangle = 1$ then $\langle\psi'|\psi'\rangle = 1$. Equivalently,

$$(U|\psi\rangle)^T U|\psi\rangle = 1 \Longrightarrow \langle\psi|U^T U|\psi\rangle = 1 \Longrightarrow U^T U = I \Longrightarrow U \text{ is a unitary matrix.}$$

Conditions 1. and 2. are necessary and sufficient conditions for inclusion in the set of transformations in quantum mechanics. Two examples of operators are:

Example 2.1. $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *maps* $|1\rangle \to |0\rangle, |0\rangle \to |1\rangle$.
Alternatively, we write $X = |1\rangle\langle 0| + |0\rangle\langle 1|$ *Then* $X|0\rangle = |1\rangle\langle 0||0\rangle + |0\rangle\langle 1||0\rangle = |1\rangle$ *and similarly* $X|1\rangle = |0\rangle$ *as desired.*

Example 2.2. $H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ *maps* $|1\rangle \to |-\rangle, |0\rangle \to |+\rangle$.
This is called the ***Hadamard operator****. Alternatively, we write* $H = |+\rangle\langle 0| + |-\rangle\langle 1|$.

2.5 Multiple Qubit Systems

In a two qubit system, the basic elements are $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. We will sometimes denote $|xy\rangle$ by $|x\rangle \otimes |y\rangle$. These elements can be represented in $\mathbb{R}^4$ by

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

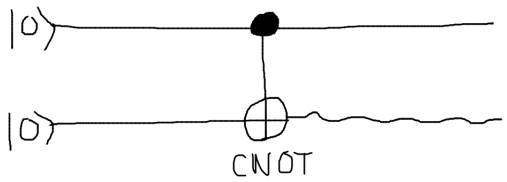
These vectors form an orthonormal basis in $\mathbb{R}^4$. A generic element in a two qubit system is a unit vector in $\mathbb{R}^4$.

In general, an n qubit system is represented by unit vectors in $\mathbb{R}^{2^n}$. The vector representations of states $|x\rangle$ where $x \in \{0,1\}^n$ form a basis in $\mathbb{R}^{2^n}$. A generic qubit is written as $\sum_{x\in\{0,1\}^n} \alpha_x |x\rangle$, where $\sum_{x\in\{0,1\}^n} \alpha_x^2 = 1$.

Here is an example of evolution on a two-bit system:

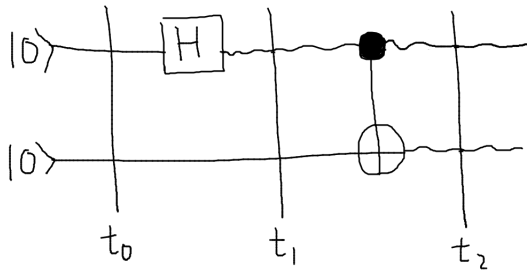
Example 2.3. $CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ *maps* $\begin{matrix} |00\rangle \to |00\rangle & |10\rangle \to |11\rangle \\ |01\rangle \to |01\rangle & |11\rangle \to |10\rangle \end{matrix}$

That is, if the top bit is $|1\rangle$ *then the bottom bit is flipped.*



We can combine multiple operators to get a new operator. When we combine the Hadamard operator and the $CNOT$ operator as depicted below the result is an EPR pair.

Definition 2.4. An EPR (Einstein-Podolsky-Rosen) pair is a pair of two quantum particles such that the first particle is observed as a $|0\rangle$ with probability $\frac{1}{2}$, and given the observation of the first particle, the state of the second particle is known.



Example 2.5. *At* t_0*, we have* $|0\rangle \otimes |0\rangle$*. Then the top qubit is affected by a Hadamard operator, which maps* $|0\rangle \to |+\rangle$*, so at* t_1*, the state of the system is*

$$|+\rangle \otimes |0\rangle = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \otimes |0\rangle = \frac{|00\rangle + |10\rangle}{\sqrt{2}}$$

Then the qubits are affected by a CNOT operator, which maps $|00\rangle \to |00\rangle$ *and* $|10\rangle \to |11\rangle$*, so at* t_2*, we have*

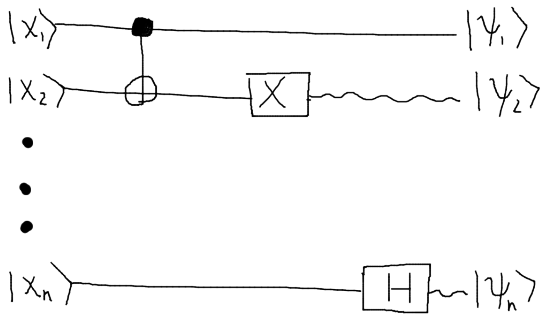
$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

Suppose now we measure the top particle. If we observe $|0\rangle$ then we know the bottom particle is also $|0\rangle$. Similarly, if we observe the top particle in state $|1\rangle$ then we know the bottom particle is also $|1\rangle$. Since these events occur with equal probability, the system is an EPR pair.

3 Quantum Computing

3.1 Quantum Circuits

Given an input x and a function f, we would like to design a circuit that computes $f(x)$. A model for quantum circuits is depicted below. The gates are unitary operators.



- Given x, input $|x_1\rangle, |x_2\rangle, ..., |x_n\rangle$ followed by $poly(n)$ free $|0\rangle$'s.
- Output observation $b = b_1, b_2, ..., b_n$ of $|\psi_{final}\rangle$.

In the classical circuit model, we restrict ourselves to sets of basic gates, such as $\{\neg, \vee_2, \wedge_2\}$. Similarly, quantum circuits are restricted to a "simple" universal family of gates, that is, a very small set of gates from which we can get all unitary operators. An example of such a family is the set of all 1-qubit gates and CNOT.

Let C be a quantum circuit. We denote the number of gates in C by $|C|$. Our goal is to make a circuit with as few gates as possible. We judge the hardness of a function by the number of gates required by any circuit computing the function. That motivates the complexity class BQP.

Definition 3.1. BQP is the set of all languages L such that for all n, there exists a circuit C such that $|C| \leq poly(n)$ and

- If $x \in L$, $C(x) = 1$ with probability at least $\frac{2}{3}$.
- If $x \notin L$, $C(x) = 0$ with probability at least $\frac{2}{3}$.

Recall BPP (bounded-error probabilistic polynomial time) is the class of decision problems computable in polynomial time with probability of error at most $\frac{1}{3}$ on a probabilistic Turing Machine. Using $|0\rangle$ to represent the zero bit, $|1\rangle$ to represent the one bit, and $|+\rangle$ for random bits, quantum circuits can easily model probabilistic Turing machines, which tells us $BPP \subseteq BQP$. Also, it can be shown that $BQP \subseteq PSPACE$. This reduction relies on having $poly(n)$ free $|0\rangle$ qubits as input. And so we have

$$P \subseteq BPP \subseteq BQP \subseteq PSPACE$$

It is also true that $P \subseteq NP \subseteq PSPACE$. However, BPP is *not* believed to be related to NP. That is, it is thought that neither class is a subset of the other.

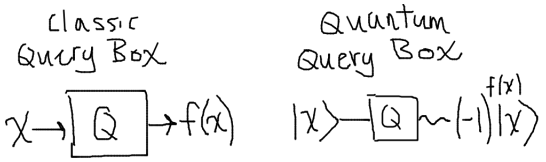
Some problems are much easier in quantum circuits then classical circuits. For example, Shor's algorithm factors an integer n in $poly(\log n)$-time and solves discrete logarithms in $poly$-time. The basis of his algorithm is the fact that in quantum computers, computing Fourier transforms of boolean functions can be done efficiently. In the classical model, factoring is believed to be hard, and important cryptographic protocols rely on its difficulty.

3.2 Search

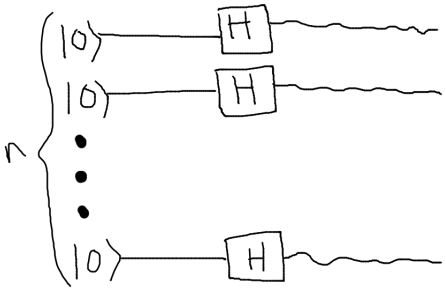
Another problem that is more efficient on quantum machines is searching for a particular item in an unordered list. Precisely, we define the problem as follows. Let $f : \{0,1\}^n \to \{0,1\}$ be a hidden function with the property that there exists a unique z such that $f(z) = 1$. We are given query access to the function, and our goal is to find z. Let $N = 2^n$.

Classically, we need $\Omega(N)$ queries as we pretty much have to try everything. In 1996, Grover showed that using a quantum computer, we only need $O(\sqrt{N})$ queries. Moreover, it was later shown that that we need $\Theta(\sqrt{N})$ queries.

The quantum version of a query box is an operator Q that on input $|x\rangle$ outputs $(-1)^{f(x)}|x\rangle$. So if $x = z$ the box outputs $-|x\rangle$, otherwise the box outputs $|x\rangle$.



For Grover's Algorithm, we need to compute an initial state $|x_0\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \{0,1\}^n} |x\rangle$. To do so, we send $n = \log N$ $|0\rangle$ qubits through Hadamard gates. Then



$$|0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle \to \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \otimes \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) \otimes \cdots \otimes \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) = \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} |x\rangle$$

where the equality follows from simply multiplying out the terms, which gives us every combination of $|1\rangle$ and $|0\rangle$ in length n qubits.

Grover's Algorithm also relies on an operator called the 'Grover Flip'. Given a qubit $|\psi\rangle = \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle$, calculate the mean of the amplitudes $\mu = \frac{1}{N} \sum_{x \in \{0,1\}^n} \alpha_x$. Then flip the altitudes over the mean. Precisely, $\alpha_x \to 2\mu - \alpha_x$. So

$$|\psi\rangle \to \sum_{x \in \{0,1\}^n} (2\mu - \alpha_x)|x\rangle$$

To check this is a unitary operator,

$$\begin{aligned}
\sum_{x \in \{0,1\}^n} (2\mu - \alpha_x)^2 &= \sum_{x \in \{0,1\}^n} (4\mu^2 - 4\mu\alpha_x + \alpha_x^2) \\
&= \sum_{x \in \{0,1\}^n} 4\mu^2 - \sum_{x \in \{0,1\}^n} 4\mu\alpha_x + \sum_{x \in \{0,1\}^n} \alpha_x^2 \\
&= N4\mu^2 - 4\mu \sum_{x \in \{0,1\}^n} \alpha_x + 1 \\
&= N4\mu^2 - 4\mu(N\mu) + 1 \\
&= 1
\end{aligned}$$

Now, we present Grover's Algorithm,

Gover's Algorithm:

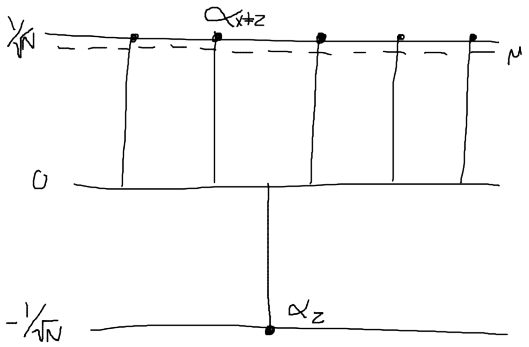
1. Initialize the start state $|x_0\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \{0,1\}^n} |x\rangle$.
2. For $i = 0, ..., \sqrt{N}$
   - $|x_i\rangle \to Q|x_i\rangle = |y_i\rangle$. (Quantum Query Box)
   - $|y_i\rangle \to GF|y_i\rangle = |x_{i+1}\rangle$. (Grover Flip)
3. Output observation of $|x_{\sqrt{N}}\rangle$.

Claim 3.2. *Grover's algorithm outputs z with constant probability.*

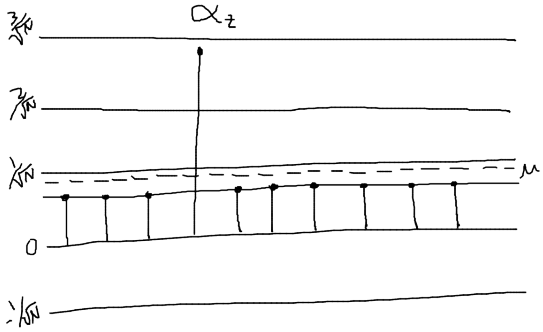
Proof. (Sketch) Consider $|y_1\rangle = Q|x_0\rangle$. Since Q is linear,

$$|y_1\rangle = \frac{1}{\sqrt{N}} \sum_{x \in \{0,1\}^n} Q|x\rangle = \frac{1}{\sqrt{N}}(-|z\rangle + \sum_{x \neq z} |x\rangle)$$

Then the mean of the amplitudes of $|y_1\rangle$ is very close to $\frac{1}{\sqrt{N}}$, and the only state far from the mean is $|z\rangle$. The amplitude α_z is about distance $\frac{2}{\sqrt{N}}$ from the mean.



Then when we preform the Grover Flip, all amplitudes $\alpha_{x\neq z}$ remain close to $\frac{1}{\sqrt{n}}$ except α_z, which is mapped to about $\frac{3}{\sqrt{N}}$.



If we were to observe the state of the system after this first iteration, $|x_1\rangle$ has probability about $(\frac{3}{\sqrt{N}})^2 = \frac{3^2}{N}$ of being observed as z. In the next iteration, the $\alpha_z \approx -\frac{3}{\sqrt{N}}$ in $|y_1\rangle$, and so $\alpha_z \approx \frac{5}{\sqrt{N}}$ in $|x_2\rangle$. Continuing this reasoning, each iteration, the amplitude α_z increases by about $\frac{2}{\sqrt{N}}$. Thus, in $\sqrt{N}$ iterations we have a constant amplitude α_z, and so a constant probability of observing z. □

As usual, we could repeat this procedure a constant number of times to observe z with high probability.