CS 264: Quantum Computation **September 24, 2009**

Lecture 1

Instructor: Sandy Irani *Scribe: Sandy Irani*

1 Overview

This class will mostly be about what we could and could not do if we could build a large-scale quantum computer. However, before we even begin that discussion we need to define what a quantum computer is. In computer science, its useful to work with a well-defined but somewhat abstract model of a machine (like the Turing Machine) so that when we talk about the resources required to solve a particular problem, these quantities can be made precise. We also like to abstract away details about the actual implementation so that when we talk about the complexity of solving a particular problem, it doesnt matter which particular machine we are using. We will do the same for quantum computation. We start with a very brief introduction to the basic postulates of quantum mechanics.

Once the model of computation is in place, we will be well positioned to define complexity classes based on the model of computation. These are sets of problems that can be solved using a particular computational model and a particular set of resources. Typically, we think of the resources as scaling with the size of the problem we are trying to solve. So, for example, the class P is the set of problems that can be solved on a classical computer in time that is bounded by polynomial in the number of bits required to specify the input to the problem. We will define the analogous class BQP for quantum computers and other complexity classes based on our quantum model.

The next part of the class will be dedicated to discussing quantum algorithms. Problems that we could solve efficiently if we could physically implement our model of computation on a large scale. The current set of techniques for algorithm design is somewhat limited, so there are only a few broad categories of problems that we know of for which quantum algorithms are more efficient than the best known classical algorithms. Next we will turn to limitations on quantum computation problems that we believe we will not be able to solve, even if large scale quantum computers could be built. Since proving out-right lower bounds is often very difficult, these will typically be reductions proving that solving a given problem is as hard as solving another set of problems.

2 The Postulates of Quantum Mechanics

Quantum mechanics is a mathematical theory framework for describing nature. It is not a description of nature on its own. The postulates of quantum mechanics tell me nothing about any particular physical system but just provide a means of describing the states of a physical system, how those states change over time and what information can be extracted

from the system. Typically, quantum mechanics is taught from a historical perspective which starts from the state of science towards the end of the 19th century when experimental phenomena was observed which could not be explained by the current classical theories. A typical quantum mechanics course will then go on to explain the progression of theoretical developments that culminated in the postulates of quantum mechanics. The core ideas can be difficult to grasp in this treatment since they are introduced in the context of more complicated physical systems like the hydrogen atom. That may be appropriate for a physics or chemistry course in which one of the goals is to understand the structure of atoms, but for our purposes, we can understand that basic tenets of quantum theory in a much more abstract (and simpler) framework. We will then go back and touch very briefly on a bit of the history and what might be required for a physical implementation of the abstract model.

2.1 Superposition

At its core, quantum mechanics can be understood as a generalization of probability theory. In a classical computer we store bits of information. A bit is a piece of information that is either 0 or 1. How that bit is stored (whether it be in silicon, a vacuum tube or a light switch) is abstracted away when we discuss classical algorithms. All we need to know is that a bit can be in a 0 state or a 1 state.

If I have a probability distribution over the possible states, this distribution can be described by a pair (p_0, p_1) in which the probability that the bit is 0 is p_0 and the probability the bit is 1 is p_1. We, of course, require that p_0 and p_1 both be non-negative and $p_0 + p_1 = 1$. The latter statement says that the L_1 norm of the vector is 1.

We could similarly have n bits which could be in one of 2^n different states. A probability distribution would be defined by a vector of length 2^n in which for each x (a string of length n) p_x is the probability that the set of bits is in the state corresponding to string x. Again, we require that the probabilities are non-negative and sum to 1. This latter fact again means that the vector describing the distribution has L_1 norm equal to 1.

In the quantum world, we describe the states of a single quantum bit (qubit) by a pair of numbers which can be complex. More generally, if I have a system that has k distinguishable states: $|1\rangle, \ldots, |k\rangle$, then the state system is described by a vector of length k, but now we allow the numbers to be complex. Thus, if the system conststs of n qubits, a state of the system is described by a vector in $\mathcal{C}^{2^n}$, $(\alpha_{X_0}, \ldots, \alpha_{X_{2^n}})$, where the set $\{x_i\}$ are all binary strings of length n. We say that α_{x_i} is the *amplitude* of x_i.

There is an important distinction here between our quantum model of information and the probabilistic one. Assuming we have access to a source of perfectly random bits, when we perform a probabilistic computation then for any particular execution of the algorithm, if we stop the computation at a particular point in its execution, the state of the system is actually in exactly one particular state (i.e. the bits have a particular setting). Before we examine the contents of memory, our knowledge of the state of the system is described by a probability distribution, but the state itself is one of the possible settings for the bits in memory. The probability distribution tells us the frequencies with which we would observe particular values stored in the qubits if we were to perform the computation and read the contents of memory at the same point in the computation many times over.

In quantum mechanics, the state of the system partially exists simulataneously in different states. In our n qubit example, the actual state would be a superposition of the 2^n different possible n-bit strings. What does it mean for the value for the amplitude of x to be $-1/4$ or even $-i/4$? How do we get an actual probability? To explain this, we need to introduce the idea of measurement.

2.2 Measurement

In our quantum system of n qubits, the state contains a great deal of information. 2^n complex numbers are required to represent a state of the system. However, all this information is not necessarily accessible. We can only learn about the state of a system through measurement. There are many different kinds of measurement one might do on a state of n qubits, but lets just say for now that the measurement we perform is to determine the value of the n qubits. Then the probability that the outcome of my measurement is x_i is $|\alpha_{x_i}|^2$. Note that this immediately places a restriction on the states. There are 2^n possible outcomes of this measurement, each with a well-defined probability. Thus, we require that these probabilities sum to 1. Thus, the state of the system must have the property that

$$\sum_{x \in \{0,1\}^n} |\alpha_x|^2 = 1.$$

In other words, when we view the amplitudes over the n-bit strings as a vector, we require that the L_2 norm of the vector be 1.

Another surprising feature of quantum mechanics is what happens to the state of the system as a result of the measurement. If the outcome of this measurement is x_i, then afterwards the state of the system is exactly x_i. In the vector notation we used above, the state would collapse to the vector which is all 0s, except with a 1 in the i^{th} location. In other words, the resulting state must be consistent with the outcome of the measurement.

We dont necessarily have to measure the entire string. We could have measured just the first qubit. In this case, the outcome of the measurement is either 0 or 1. The probability of getting a 0 is the square of the amplitude of those states which have a 0 in the first bit:

$$p_0 = \sum_{x \in 0\{0,1\}^{n-1}} |\alpha_x|^2.$$

p_1 is defined similarly and $p_0 + p_1 = 1$. After the measurement, the state must be consistent with that outcome, so if the outcome was 0, then the new amplitude α'_x is zero if the first bit of x is 1 and is $1/\sqrt{p_0}\alpha_x$ is the first bit of x is 0.

2.3 Evolution

The third postulate of quantum mechanics is that the evolution of a quantum system over time is unitary. Any operator that describes how a quantum system changes over time must be a unitary operator. In our matrix notation, the operator is described by a unitary matrix. There are a variety of ways to define a unitary operator, but at this point, probably the most

natural is that it preserves the length of any vector. Thus, if we have a state whose L_2 norm is 1 and the system changes over time, the L_2 norm of the state will remain 1. In the n-qubit example, if we measure the value of each qubit, the sum of the probabilities over all the possible outcomes will sum to 1. The third postulate states that this will continue to be the case, even as the state evolves over time.

Although these postulates may seem mysterious and arbitrary, there are a couple of features that make sense mathematically. First of all, why should amplitudes necessarily be complex? Why not just allow for negative amplitudes? Consider the two states

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \qquad\qquad |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

We could have a unitary transformation that changes the $|+\rangle$ state to the $|-\rangle$ state over some period of time. Since time is continuous, it would be natural for this transformation to take place in a continuous manner over time. In order to express a continuous range between these two possibilities ($|+\rangle$ and $|-\rangle$), we need to express it as $1/\sqrt{2}(|0\rangle + e^{i\theta}|1\rangle)$.

Another question to ask is why should it be the L_2 norm that is one? Why shouldn't the probability of an outcome of a measurement be the amplitude to the fourth power? It turns out that the L_1 and L_2 norms are the only two that allow for non-trivial norm-preserving linear transformations.

2.4 Computing

So how do we use a quantum mechanical system to compute? In order to compute we need a means of storing information, manipulating that information and then extracting an answer. These are the three components described above. We can think of information as being stored in a set of n qubits in which case, the state of our computation is represented by a vecor in C^{2^n}. We manipulate the information stored by applying some unitary transformation to the information stored (this is the computation) and then we perform a measurement to read the outcome of the computation.

If the computation is deterministic and classical, then a state vector must be all 0's except for one 1. A computation over time is described by a permutation matrix which is all 0's except for one 1 in every column and one 1 in every row. If the computation is classical but probabilistic, a state is a vector whose L_1 norm and the transformation over times can be described by a stochastic matrix (non-negative, real, columns sum to 1) which preserves the L_1 norm. A quantum state is a superposition represented with complex entries in the vector whose L_2 norm is 1. An operation is a unitary matrix (L_2 norm preserving).

3 More on Notation and Linear Algebra

A finite dimensional Hilbert Space is a complex vector space along with an inner product. If a quantum system has k distinguishable states, $|1\rangle, \ldots, |k\rangle$, then the set of possible states are the unit vectors in a Hilbert space of dimension k. The states $|1\rangle, \ldots, |k\rangle$ form an

orthonormal basis of the Hilbert space, meaning that any state can be expressed as a linear combination of the $|j\rangle$'s with complex coefficients:

$$|\psi\rangle = \alpha_1|1\rangle + \cdots + \alpha_k|k\rangle,$$

and the inner product of two basis states $(|i\rangle, |j\rangle) = \delta_{ij}$. Once we have specified a particular basis, we can express the state $|\psi\rangle$ as a column vector in this basis:

$$|\psi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix}.$$

The inner product of two states, $|\psi\rangle = \sum_{j=1}^{k} \alpha_j|j\rangle$ and $|\psi'\rangle = \sum_{j=1}^{k} \beta_j|j\rangle$ is defined in the usual way:

$$(|\psi\rangle, |\psi'\rangle) = \sum_{j=1}^{k} \alpha_j^* \beta_j.$$

In Dirac notation, the vector dual of $|\psi\rangle$ is denoted $\langle\psi|$. If $|\psi\rangle$ and $\langle\psi|$ are expressed in the same basis then

$$\langle\psi| = (\alpha_1^*, \ldots, \alpha_k^*),$$

and

$$\langle\psi|\psi'\rangle = (\alpha_1^*, \ldots, \alpha_k^*) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = (|\psi\rangle, |\psi'\rangle).$$

Thus, our normalization criterion for a state can be expressed as $\langle\psi|\psi\rangle = 1$.

Once we fix a particular basis, we can also express a linear operator in matrix form. If A transforms vectors in k-dimensional Hilbert space V into vectors from V, it is a $k \times k$ matrix. The *adjoint* of A is the unique linear operator that satisfies

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle),$$

for all $|v\rangle$ and $|w\rangle$ in V. In other words, if $|\psi'\rangle = A|\psi\rangle$, then $\langle\psi'| = \langle\psi|A^\dagger$. If A and $A^\dagger$ are expressed in the same basis then $A^\dagger$ is obtained by taking the conjugate-transpose of A (transposing the matrix A and taking the complex-conjugate of each entry).

The third postulate states that the evolution of a quantum system over time, can be expressed as a unitary operator U such that the transformation applied to state $|\psi\rangle$ is $U|\psi\rangle$. The following set of conditions are equivalent:

- U is norm preserving.
- U preserves the inner product: the inner product of $|\psi\rangle$ and $|\psi'\rangle$ is the same as the inner product of $U|\psi\rangle$ and $U|\psi'\rangle$.

- $U^\dagger U = U^\dagger U = I$.
- When U is expressed as a matrix, the rows form an ortho-normal basis.
- When U is expressed as a matrix, the columns form an ortho-normal basis.
- The eigenvalues of U are roots of unity, $e^{i\theta}$.

The Dirac notation can also be used to express operators using the outer-bracket notation: $|\psi'\rangle\langle\psi|$. This operator applied to state $|\phi\rangle$ yields $\langle\psi|\phi\rangle|\psi'\rangle$. Linear operators can be built up by taking linear combinations of operators expressed in the outer-product notation.

While we will typically be discussing qubits in the $|0\rangle$, $|1\rangle$ basis (also called the *standard* or *computational* basis), there are others that will come up. For example the $+/-$ basis is

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

Any state $\alpha|0\rangle + \beta|1\rangle$ can be expressed as $\alpha'|+\rangle + \beta'|-\rangle$, where

$$\begin{aligned} \alpha' &= \alpha\langle+|0\rangle + \beta\langle+|1\rangle \\ \beta' &= \alpha\langle-|0\rangle + \beta\langle-|1\rangle \end{aligned}$$