IAS/Park City Mathematics Series
Volume 3, 1993

Quantum Computation

Ran Raz

LECTURE 1
Introduction

Let us start with one example for an algorithm.

(1) Input: natural numbers a, b
(2) If $a > b$, $\quad a := a - b$
(3) If $b > a$, $\quad b := b - a$
(4) If $a = b$, $\quad$ output a, STOP
otherwise, GOTO 2

This is the well known Euclid's algorithm for finding $GCD(a, b)$. It is a very simple algorithm that can be run on any computer. The remarkable thing about it is that it was discovered more than two thousand years ago, much before computers were invented, and much before people understood what an algorithm is. How did Euclid know what operations are legitimate for an algorithm to perform, before knowing what an algorithm is ? The simple answer is that algorithms were performed by people (sometimes using some additional memory devices). An algorithm is legitimate if it uses operations that can be performed efficiently by us.

Today, algorithms run on computers, and one remarkable thing is that the same computer can run any given algorithm. That is, we do not need a different computational machine for every algorithm. There is one machine that can run all of them. In that sense, the computer is a universal computational machine. Computers work fast, but, at least in principle, we (i.e., people) can also perform any algorithm that the computer runs, by performing the exact same operations.

[1]Faculty of Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel.
E-mail address: `ranraz@wisdom.weizmann.ac.il`.

July 21, 2000

Based on lecture notes taken by Yoad Lustig.

This rises the question of whether or not one can construct a machine that performs certain computations much more efficiently, using operations that we (or today's computers) are not able to (efficiently) simulate. Can nature perform certain computations much more efficiently than computers ? Is it possible to construct machines that perform certain complicated computations in some magical fast way ? It used to be assumed that the answer for these questions is negative, and that the computer is a universal machine also in the sense that it can simulate (efficiently) any other computational device. This assumption is sometimes referred to as *Church thesis*.

Quantum computation is an attempt to define a new computational model (and maybe even to construct new computational machines), based on the principles of quantum mechanics. Church thesis may be true for classical machines. However, in the previous century, a new physics was discovered; quantum physics. It has been observed (not so lately) that there is no simple way to simulate a quantum system by a (classical) computer. This brings us to the starting point of quantum computation: If there is no simple way to simulate a quantum system by a (classical) computer then maybe one can construct a quantum machine that performs certain computations much more efficiently than (classical) computers. Today, we have a strong indication that indeed, at least in theory, the computational power of quantum devices may be stronger than the one of classical computers.

The area of quantum computation is young and popular, and it has attracted a lot of attention in recent years. The area may be interesting for various reasons:

(1) From a practical point of view; maybe at some point it will be possible to construct *quantum computers* that will perform certain computations much faster than classical ones.
(2) From a theoretical point of view; we would like to understand the "computational power of nature".
(3) From the point of view of quantum physics; the research may help to get a better understanding of certain aspects of quantum mechanics, and in particular, a better understanding of some of the well known paradoxes of the theory.

1.1. Classical Deterministic Machines

Consider a computer with a total number of n memory bits. The number of states that these bits can be in is 2^n (we can also consider any other classical deterministic machine or system with 2^n states). We will name the 2^n states by the numbers $0, 1, ..., 2^n - 1$. The machine starts from an initial state (say, state 0), and the evolution in time is given by a transition function from the set of states to the set of states. For example,

$$\begin{aligned} 0 &\to 1 \\ 1 &\to 3 \\ 2 &\to 2 \\ 3 &\to 3 \end{aligned}$$

In each time step, the state of the machine changes deterministically, according to the transition function.

We will represent each state by a 2^n dimensional vector, called the *state vector*. For every $i \in \{0, ..., 2^n - 1\}$, state i is represented by a vector that has 1 in the i^{th}

coordinate and 0 in all other coordinates. That is,

$$\text{state } 0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{state } 1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \ldots\ldots\ldots$$

We will represent the transition function by a $2^n \times 2^n$ dimensional matrix, called the *transition matrix*. In column i of the matrix, we will have 1 in the j^{th} coordinate if the transition function sends state i to state j, and we will have 0 in all other coordinates. That is, we will have exactly one 1 in each column. For example, the above described transition function is represented by the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

When we apply the transition matrix on a given state vector, we obtain the state vector in the next time step.

1.2. Classical Probabilistic Machines

A well known and well studied idea is to use randomness in the computation. That is, the computation is probabilistic. The state of the machine at any given time is hence not fixed in advance. The state is a random variable with some probability to be in each of the basic states $\{0, ..., 2^n - 1\}$. We will represent the state of the machine by the state vector

$$\begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{2^n-1} \end{pmatrix}$$

where p_i is the probability for the machine to be in the basic state i. We will use the term *basic state* to refer to one of the basic states $\{0, ..., 2^n - 1\}$. We will use the term *state* to refer to the actual state of the machine as a random variable. We think of the state also as a linear combination of the 2^n basic states. The probabilities p_i satisfy:

(1) $\forall i \;:\; p_i \geq 0$.
(2) $\sum_{i=0}^{2^n-1} p_i = 1$.

That is, all the elements p_i are positive and the L_1-norm of the state vector is 1.

When we *measure* the system, that is, when we check in which basic state the machine actually is, we obtain one of the 2^n basic states. Each basic state i is obtained with probability p_i. If the basic state i was obtained then the machine continues from that state i, that is, we think of the state of the machine as changing from its original state (before we measured) to the basic state i that was obtained. In particular, if immediately after that we measure the system again we will get again the same basic state i.

For example, assume that a ball can be in two basic states, 0 and 1. If the system is probabilistic the ball can be in the state

$$\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

When we measure the system, we find the ball in each basic state with probability 1/2, and the state of the system changes from its current state to the basic state that was obtained.

The transition function of a probabilistic machine is a probabilistic function. For every two basic states i, j, we will have a certain probability, $p_{j|i}$, for transition to state j, given that the machine is in state i. As before, we represent the transition function by a $2^n \times 2^n$ dimensional matrix. In column i and row j of the matrix, we will have the transition probability $p_{j|i}$. When we apply the transition matrix on a given state vector, we obtain the state vector in the next time step. Note that the entries of the transition matrix are positive real numbers and the sum of the entries in each column of the matrix is 1. That is, the matrix is a *stochastic matrix*. Note that a stochastic matrix is a general matrix with positive entries that preserves the L_1-norm of all vectors.

1.3. Quantum Systems

In a quantum system with 2^n basic states, the general state of the system is a linear combination over $\mathbb{C}$ (i.e., over the complex numbers) of the 2^n basic states. The state is usually normalized to be of L_2-norm 1. We will represent the state by a unit vector in the vector space $\mathbb{C}^{2^n}$, (where by unit vector we mean to a vector with L_2-norm 1). That is, the state is

$$\begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2^n-1} \end{pmatrix} = c_0 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots\ldots\ldots + c_{2^n-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

where

$$\sum_{i=0}^{2^n-1} |c_i^2| = 1.$$

This is similar to the state vector in the classical probabilistic case, except that we require $c_i \in \mathbb{C}$, rather than $p_i \in \mathbb{R}^+$, and $\sum_{i=0}^{2^n-1} |c_i^2| = 1$, rather than $\sum_{i=0}^{2^n-1} p_i = 1$.

Thus, the state of the system is a linear combination (over $\mathbb{C}$) of the basic states. A linear combination, in this context, is also referred to as *superposition*. That is, we say that the state of the system is a superposition of the basic states. Consider for example the ball from the previous example. If the system is quantum the ball can be in the superposition

$$(1/\sqrt{2}) \begin{pmatrix} 1 \\ 0 \end{pmatrix} - (1/\sqrt{2}) \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

1.4. Dirac's Ket Notation

In the vector space $\mathbb{C}^d = \mathbb{C}^{2^n}$ (where $d = 2^n$), each vector

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_d \end{pmatrix}$$

has a dual vector

$$(a_0^*,\ a_1^*,\ ...,\ a_d^*)$$

in the dual vector space, where a_i^* is the complex conjugate of a_i. The inner product of two vectors a, b is defined by

$$\langle a, b\rangle = \sum_{i=0}^{d-1} a_i^* b_i = \langle b, a\rangle^*.$$

The norm of a vector a is defined by

$$\|a\| = \sqrt{\langle a, a\rangle} = \sqrt{\sum_{i=0}^{d-1} |a_i|^2}.$$

To denote vectors in that space, we will use Dirac's *ket* notation: The vectors of the standard base for C^{2^n} are denoted by

$$|0\rangle, |1\rangle, ..., |2^n - 1\rangle.$$

That is,

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad .... \quad |2^n - 1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

These 2^n vectors are the basic states. An arbitrary vector c is denoted by

$$|c\rangle = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2^n-1} \end{pmatrix} = \sum_{i=0}^{2^n-1} c_i |i\rangle$$

These vectors are called *ket vectors*. The dual vectors are called *bra vectors*, and the dual to a ket vector $|c\rangle$ is a bra vector denoted by $\langle c|$. The vectors are called bra and ket, because putting the two together we get bra-ket, which refers to the inner product between the two. The inner product of two vectors $|a\rangle$ and $|b\rangle$ is denoted by $\langle a|b\rangle$. Note that for every i, j,

$$\langle i|j\rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

1.5. Quantum Measurement

Thus, a quantum state is a linear combination (or, superposition)

$$|c\rangle = c_0|0\rangle + \dots\dots\dots + c_{2^n-1}|2^n - 1\rangle$$

For example, a ball may be in the state

$$\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$$

and Schroedinger's cat may be in the state

$$\frac{1}{\sqrt{2}}|dead\rangle - \frac{1}{\sqrt{2}}|alive\rangle$$

Schroedinger brought this example as a joke (he never saw his cat in such a state), but, at least in theory, this state is possible.

What happens when we measure the system ? That is, what happens when we check in which basic state the system is ? The answer is that if a system is in state $|c\rangle$, when we measure the system (according to the standard base of basic states) we will find the system in each state $|i\rangle$ with probability

$$|c_i|^2.$$

(Since $\sum_{i=0}^{2^n-1} |c_i|^2 = 1$, the probabilities sum to 1). Moreover, if we measured a system and found it in the basic state $|i\rangle$ the state of the system changes from the original state $|c\rangle$ to the state $|i\rangle$. From that point on, the system continues from state $|i\rangle$. The original state $|c\rangle$ is forgotten forever. In particular, if immediately after that we measure the system again we will get the same basic state $|i\rangle$ (with probability 1).

This is the famous *collapse* of the state of a quantum system (which is a very problematic concept in quantum mechanics). Note, however, that a similar collapse occurs also in the classical probabilistic case. When we measure the state of a classical probabilistic system, we find it in the basic state i with probability p_i, and from that point on, the state changes to the basic state i. The difference is that in the quantum case, the probability to get each basic state $|i\rangle$ is $|c_i|^2$, rather than p_i in the classical probabilistic case. But this difference can be considered to be just a different representation (or normalization) of the same thing. So, what is the real difference between the quantum case and the classical probabilistic case ?

1.6. Transition Matrix and Interference

The major difference between a quantum system and a classical probabilistic system is in the evolution in time, that is, in the dynamics of the system. In a classical probabilistic system, the transition matrix is a general matrix that preserves positiveness and L_1-norm, that is, a stochastic matrix. In the quantum case, the transition matrix is a general matrix that preserves L_2-norm over $\mathbb{C}$ (recall that in a quantum system, the state vector is a unit vector in C^{2^n}). Such a matrix is called a *unitary matrix*. Each one of the following can serve as an alternative definition for a unitary matrix.

A matrix U is unitary iff:

(1) U preserves the L_2-norm of every vector.
(2) U is a mapping from the unit sphere to the unit sphere.

(3) U maps every orthonormal base to an orthonormal base.
(4) The rows of U form an orthonormal base.
(5) The columns of U form an orthonormal base.
(6) $UU^\dagger = I$, (where I is the identity matrix, and $U^\dagger$ is the matrix U conjugated and transposed, i.e., $(U^\dagger)_{ij} = (U_{ji})^*$).

To get a better understanding of the difference between a quantum system and a classical probabilistic system, consider the following example for a unitary matrix, the, so called, Hadamard matrix of order 2.

$$H_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

If we apply the matrix H_2 on the vector $|0\rangle$ we obtain

$$H_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

If we apply the matrix H_2 on the vector $|1\rangle$ we obtain

$$H_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

That is, if we apply H_2 on the state $|0\rangle$ and then measure the system we obtain the state $|0\rangle$ with probability 1/2 and the state $|1\rangle$ with probability 1/2. If we apply H_2 on the state $|1\rangle$ and then measure the system we obtain the same results. However, if we apply H_2 on linear combinations of $|0\rangle$ and $|1\rangle$ we obtain different results. For example,

$$H_2 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

That is, if we apply H_2 on a certain linear combination of $|0\rangle$ and $|1\rangle$ and then measure the system we obtain the state $|0\rangle$ with probability 1. This phenomena, which is special for quantum systems, is called *interference*.

1.7. Measurement According to a Different Base

Quantum mechanics allows in certain cases to measure a system according to different orthonormal bases. Let

$$E = \{|e_0\rangle, |e_1\rangle, \ldots\ldots, |e_{2^n-1}\rangle\}$$

be an orthonormal base. Given a vector $|c\rangle$, We can represent $|c\rangle$ in the new base E as

$$|c\rangle = \sum_{i=0}^{2^n-1} c'_i |e_i\rangle$$

where c'_i is given by

$$c'_i = \langle e_i | c \rangle$$

The rule of the quantum measurement is simple: If we measure $|c\rangle$ according to the base E we obtain each vector $|e_i\rangle$ with probability $|c'_i|^2$, and from that point on, the system continues from the state $|e_i\rangle$, that is, the state $|c\rangle$ collapses to $|e_i\rangle$. Note that $|c'_i|$ is just the length of the projection of the vector $|c\rangle$ on the vector $|e_i\rangle$.

Another variant of quantum measurement is the *partial measurement*. This topic will be discussed in the next lecture.

1.8. The Polarizers Experiment

An interesting example for the power of measurements according to different bases is given by a famous experiment that involves light polarizers. Consider a photon (i.e., light particle) that moves on the z axes. The photon may be polarized to the direction of the x axes or to the direction of the y axes. We denote these two basic states by $|x\rangle$ and $|y\rangle$. The actual polarization of the photon may be any linear combination of these two basic states. Assume for simplicity that the linear combination is over $\mathbb{R}$, rather than over $\mathbb{C}$. Each such linear combination is interpreted as a photon polarized to a different angle in the x, y plan. More precisely, the linear combination

$$|\theta\rangle = \cos\theta \cdot |x\rangle + \sin\theta \cdot |y\rangle$$

is interpreted as a photon polarized to the direction that forms an angle θ with the x axes.

A polarizer measures a photon according to the basic states $|x\rangle$, $|y\rangle$, and the photon passes the polarizer only if it is found out to be polarized to the direction $|x\rangle$. That is, half of the photons pass the polarizer, and all the photons that pass are polarized to the direction of the x axes. When we turn the polarizer to an angle θ, we perform a measurement according to the base $|\theta\rangle, |\theta + \pi/2\rangle$. As before, half of the photons pass the polarizer, and all the photons that pass are in the state $|\theta\rangle$ (i.e., polarized to the direction $|\theta\rangle$). Different angles of the polarizer perform measurements according to different bases and produce photons polarized to different angles.

When we take two such polarizers and place them one behind the other, the amount of light that passes both polarizers depends on the angle between them. If the angle between the two polarizers is 0 then every photon that passes the first polarizer will also pass the second one. Hence, half of the photons will pass both polarizers. If the angle between the two polarizers is $\pi/2$ then no photon that passes the first polarizer will pass the second one. If the angle is $\pi/4$ then each photon that passes the first polarizer will pass the second one with probability $1/2$ (and hence $1/4$ of the photons will pass both polarizers).

Thus, if the angle between the two polarizers is $\pi/2$ no photon will pass both polarizers, and we will see no light. This happens because every photon that passes the first polarizer is in state $|x\rangle$ and hence it cannot pass the second polarizer (that passes only photons in state $|y\rangle$). The interesting thing happens when we place a third polarizer in between the first two polarizers. Assume that the third polarizer is in angle $\pi/4$ to the first two polarizers. A photon that passes the first polarizer will pass the third one with probability $1/2$ (as the angle between them is $\pi/4$). After passing the third polarizer, the photon is polarized to the direction $|\pi/4\rangle$, and hence the photon will pass the second polarizer with probability $1/2$ (since the angle between the third polarizer and the second one is also $\pi/4$). Thus, the total probability for a photon to pass all three polarizers is $1/8$.

To summarize: Before we place the third polarizer, a photon has no chance to pass the first two polarizers. After we place the third polarizer in between the first two, a photon has a chance of $1/8$ to pass all three polarizers.

1.9. Historical Background

Quantum physics was developed in the first three decades of the previous century (1900-1927, according to Sakurai). In 1982 and 1984 Benioff and Feynman suggested to construct computational machines, based on the principles of quantum mechanics. Feynman was interested in such a machine because he saw no way to efficiently simulate a quantum system by a classical computer, and hence such a machine may be crucial for physical simulations. The model of quantum computation was defined in 1985, and studied since then (and mainly since 1992). In 1994, Shor presented a quantum polynomial time algorithm for factoring integers. Since we believe that there is no classical polynomial time algorithm for factoring, Shor's result gives a strong indication that classical computers cannot simulate quantum systems, and that quantum computers, if ever built, will be more powerful than classical ones. In 1996, Grover presented a quantum database search algorithm that finds an element in a database in time (roughly) $O(\sqrt{N})$, where N is the size of the database. In particular, this shows how to solve any NP problem in time $O(2^{n/2})$, where n is the number of non-deterministic bits used. This is significantly better than $O(2^n)$, which is the best known classically. Another remarkable achievement is the development of quantum error correcting codes and quantum fault-tolerant computations. It was shown that quantum computations may be possible to perform even if some noise exists in the system and even if some interaction between the quantum system and its environment does exist. Quantum error correcting codes and quantum fault-tolerant computations may be crucial for constructing large quantum computational devices. A quantum circuit with 7 quantum bits has already been implemented. It is still not clear how to construct quantum circuits with a significantly larger number of bits, and at this point it seems very unlikely that a quantum computer will be constructed in the near future.