Lecture 1: Overview of quantum information

January 10, 2006

References

Most of the material in these lecture notes is discussed in greater detail in the following two books, which I recommend you study if you are interested in quantum computation.

1. M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
2. A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

Introduction to Quantum Information

For the remainder of this lecture we will take a first look at quantum information, a concept upon which quantum computation is based.

A probabilistic model

It is helpful to start classically, with a model that will probably seem completely simple to everyone. Imagine that we have some physical device, called X, that has some finite, non-empty set Σ of possible *states*[1]. For example, we might have $\Sigma = \{0, 1\}$, in which case we would think of X as representing a bit. For the following discussion let us restrict ourselves to this example (but keep in mind that everything can easily be generalized to sets other than $\{0, 1\}$).

Suppose that we do not necessarily have complete information about the state of X, but instead represent our knowledge of its state by assigning probabilities to the different states. For example, we might have

$$\Pr[\text{state of } \mathsf{X} \text{ is } 0] = 1/4,$$

$$\Pr[\text{state of } \mathsf{X} \text{ is } 1] = 3/4.$$

Mathematically we can represent this type of knowledge about the state of X with a *probability vector*, which is a column vector whose entries are all nonnegative real numbers that sum to 1. In the case at hand, the associated probability vector is

$$v = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix}.$$

[1] Shortly we will change our terminology and use the term *classical states* to refer to elements of Σ, because the term *state* will be used in a different context. Nevertheless, for the time being we will stick with the term *state* when referring to elements of the set Σ.

The understanding is that the entries of v are indexed by Σ, and when we write such a vector in the above form we are using the most natural way of ordering the elements of Σ:

$$v = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix} \begin{array}{l} \leftarrow \text{entry indexed by } 0 \\ \leftarrow \text{entry indexed by } 1 \end{array}$$

We may write $v[0]$ and $v[1]$ to refer to the entries of v when necessary.

What happens when you look at X? Of course you will not see a probability vector v. Instead you will see some element of Σ. If our representation of the state of X by a probability vector v is in some way meaningful, you may as well imagine that the state you saw was determined randomly according to the probabilities associated with the various states. Notice that by looking at the state of X you effectively change the description of your knowledge of its state. Continuing with the example above, if you look and see that the state is 0, the description of your knowledge changes from v to a new probability vector w:

$$v = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix} \longrightarrow w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

You know that the state is 0, and the vector w represents this knowledge. If you saw that the state was 1 instead of 0, the vector would become

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

instead.

What sorts of operations can you imagine performing on X? There are not very many deterministic operations: you could initialize X to either 0 or 1, you could perform a NOT operation to X, or you could do nothing to X (which can still be considered an operation even though it has no effect). You could also perform an operation involving randomness—for instance perform a NOT operation with probability 1/100, and otherwise do nothing. I claim that any *physically meaningful* operation can be represented by a matrix, with the effect of the operation being determined by matrix-vector multiplication. For instance, these four matrices

$$\mathrm{INIT}_0 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathrm{INIT}_1 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathrm{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

represent the deterministic operators mentioned above. For example, if our knowledge of the state of X is represented by

$$v = \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix}$$

and we perform a NOT operation on X, the new probability vector that results is

$$w = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix} = \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}.$$

The probabilistic operation mentioned above is represented by the matrix

$$\begin{pmatrix} \frac{99}{100} & \frac{1}{100} \\ \frac{1}{100} & \frac{99}{100} \end{pmatrix}.$$

All of these matrices have the property that (i) all entries are nonnegative real numbers, and (ii) the entries in each column sum to 1. In other words, every column is a probability vector. Such matrices have a name: they are called *stochastic matrices*. In the simple model we are discussing, physically meaningful operations are described by stochastic matrices. It works the other way as well; any stochastic matrix describes some physically meaningful operation.

As mentioned before, this entire picture is easily generalized to the case where Σ is not necessarily $\{0,1\}$. In general the dimension of the vectors and matrices will be equal to the size of Σ.

Quantum bits (qubits)

The framework of quantum information works in a similar way to the simple probabilistic model we just saw, but with some key differences. Let us again imagine that we have a physical device called X. As before we imagine that there is some set Σ of possible states of X, and we will again consider for now just the simple case $\Sigma = \{0,1\}$. At this point, to avoid confusion let us now refer to elements of Σ as *classical states* rather than just *states*. Intuitively you can think of a classical state that you as a human can look at, touch, and recognize without ambiguity. The device X will represent the quantum analogue of a bit, which we call a *qubit*.

We will still represent our knowledge[2] of X with column vectors indexed by Σ, but this time they will not be probability vectors. Instead of representing probability distributions, the vectors represent what we call a *superposition* or just a *state* (by which we mean a *quantum state*). For example, here are a few vectors representing superpositions:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{3}{5} \\ \frac{4i}{5} \end{pmatrix}.$$

Notice that the entries in these vectors are not probabilities: they are not necessarily nonnegative (in fact they are not even necessarily real numbers), and they do not necessarily sum to 1. We call these numbers *amplitudes* instead of probabilities. The condition that replaces the probabilities summing to 1 in a probability vector is this: vectors representing superpositions have Euclidean length equal to 1. In the simple case at hand where $\Sigma = \{0,1\}$, this means that any vector representing a superposition has the form

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

[2] We discussed briefly in the lecture whether or not the column vectors represent knowledge in the same sense as the probabilistic model or something more "actual". My choice of the word "knowledge" is really only intended to stress the similarity with the probabilistic model; and although the question makes for an interesting philosophical discussion, I don't intend that this course will go in that direction. As soon as possible we will be treating everything mathematically and just thinking of these things as vectors.

for $\alpha, \beta \in \mathbb{C}$ satisfying $|\alpha|^2 + |\beta|^2 = 1$.

Similar to the probabilistic case, if you look at the qubit X you will not see a superposition. Instead, you will see either 0 or 1 just like before. The probability associated with the two possible outcomes is given by the absolute value squared of the associated amplitude—so if the superposition of X is represented by the vector

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

and you look at X, you will see 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$. This is why we have the condition $|\alpha|^2 + |\beta|^2 = 1$, because the probabilities have to sum to 1 for the model to make sense. The same rules apply as for the probabilistic case for determining the superposition of X after you look at it: the superposition becomes

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

depending on whether you see 0 or 1, respectively.

So far the model does not seem qualitatively different from the probabilistic model, but that changes a lot when the possible operations that can be performed are considered. Again the possible operations are represented by matrices; but now instead of being stochastic matrices, the matrices that represent valid physical operations correspond to *unitary* matrices. A matrix is unitary if and only if it preserves the Euclidean norm. Fortunately there is a very simple condition to check this: a matrix U is unitary if and only if

$$U^\dagger U = I,$$

where $U^\dagger$ is the conjugate transpose of U (meaning that you take the transpose of U and then take the complex conjugate of each of the entries). For example, these are unitary matrices:

$$H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{NOT} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

(for any real number θ in the case of R_θ). For example, if X is in a superposition described by

$$v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and the operation corresponding to the matrix H (called the Hadamard transform) is performed, the superposition becomes

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

If you measured X at this point you would see outcome 0 or 1, each with probability 1/2. If you didn't measure and instead applied the Hadamard transform again, the superposition would become

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

To recapitulate, these are the two things you can do to a qubit:

1. **Perform a measurement.** If the superposition of the qubit is

   $$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

   and a measurement is performed, the outcome is 0 or 1, with probabilities $|\alpha|^2$ and $|\beta|^2$, respectively. The superposition of the qubit becomes

   $$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

   depending on whether the measurement outcome was 0 or 1.

2. **Perform a unitary operation.** For any unitary matrix U, the operation described by U transforms any superposition v into the superposition Uv.

Later on in the course we will see that there are somewhat more general operations and measurements that can be performed, but this simple model will turn out to be sufficient for discussing quite a lot about quantum computing.

Example 1. Suppose your friend has a qubit that he knows is in one of the two superpositions

$$v_0 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{or} \quad v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

but he isn't sure which. How can you help him determine which one it is?

Measuring right away will not help—you would see a random bit in either case. Instead, you should perform the Hadamard transform and then measure. Performing the Hadamard transform changes the superpositions as follows:

$$Hv_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad Hv_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Now if you measure, you will see 0 (with certainty, meaning probability 1) if the original superposition was v_0 and you will see 1 (with certainty) if the original superposition was v_1.