Lecture 2: Overview of quantum information (continued)

January 12, 2006

In the previous lecture we started discussing the basics of quantum information, beginning with the example of single qubit systems. In this lecture we will continue this discussion. In particular, we will discuss multiple qubit systems and a more convenient notation for describing superpositions.

Multiple qubits

In order to talk about what happens when we have multiple qubits, it will be helpful to briefly return to the probabilistic model from before. Suppose that X and Y are devices that implement bits. Then there are 4 possible states of the pair (X, Y), namely 00, 01, 10, and 11. Thus, our set of states Σ corresponding to this pair is now $\{00, 01, 10, 11\}$. In the probabilistic model we represent our knowledge of the state of the pair (X, Y) with a 4 dimensional probability vector. For example we could have the following probability vector:

$$\begin{pmatrix} \frac{1}{8} \\ \frac{1}{2} \\ 0 \\ \frac{3}{8} \end{pmatrix} \begin{array}{ll} \leftarrow & \text{probability associated with state } 00 \\ \leftarrow & \text{probability associated with state } 01 \\ \leftarrow & \text{probability associated with state } 10 \\ \leftarrow & \text{probability associated with state } 11 \end{array}$$

(The vector indices are labeled by the states in the order given by binary notation.) Operations again correspond to stochastic matrices, but this time the matrices are 4×4 matrices. For example, the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

is stochastic. It happens to correspond to the operation where you do nothing if the first bit is 0, but if the first bit is 1 then replace the second bit with a random bit.

The quantum variant works in an analogous way. If we have two qubits (X, Y), then a superposition of these two qubits is a 4 dimensional vector with Euclidean length equal to 1. For example:

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{i}{2} \\ -\frac{1}{2} \end{pmatrix}$$

Measurements work the same way as before, except that the outcome will be two bits. For example, measuring the previous superposition gives results as follows:

$$\begin{aligned}
&00 \quad \text{with probability } \left|\tfrac{1}{\sqrt{2}}\right|^2 = \tfrac{1}{2} \\
&01 \quad \text{with probability } |0|^2 = 0 \\
&10 \quad \text{with probability } \left|\tfrac{i}{2}\right|^2 = \tfrac{1}{4} \\
&11 \quad \text{with probability } \left|-\tfrac{1}{2}\right|^2 = \tfrac{1}{4}
\end{aligned}$$

We will see next lecture how it works when you just measure one qubit. Unitary operations also work the same way as before, but this time are 4×4 matrices. For example, here is a 2 qubit unitary operation called the controlled-NOT:

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

The same pattern is used for 3 qubits, 4 qubits, etc. The dimension of the vectors and matrices grows exponentially: 8 dimensional vectors for 3 qubits, 16 dimensional vectors for 4 qubits, etc.

By the way, there is no reason why you cannot consider the model for any other choice of Σ, instead of Σ corresponding to all possible strings of a given length. Typically, however, we will focus on the case where $\Sigma = \{0,1\}^n$ for some positive integer n.

Tensor products

Returning again briefly to the probabilistic model, let us suppose that as before X and Y are devices implementing bits, and the two devices are completely uncorrelated with one another—let us say that the probability vector corresponding to X is

$$v = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

and the probability vector corresponding to Y is

$$w = \begin{pmatrix} \frac{1}{4} \\ \frac{3}{4} \end{pmatrix}.$$

Then the 4 dimensional probability vector corresponding to the pair (X, Y) is easily determined by multiplying the corresponding probabilities. In particular, the resulting vector is

$$v \otimes w = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{2} \\ \frac{1}{12} \\ \frac{1}{4} \end{pmatrix}.$$

The operation $\otimes$ is called the *Kronecker product* or the *tensor product*. (It is most common in quantum computing to use the term *tensor product* to refer to this operation.) In general, for any two matrices

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,l} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k,1} & b_{k,2} & \cdots & b_{k,l} \end{pmatrix}$$

we define $A \otimes B$ to be the $nk \times ml$ matrix

$$A \otimes B = \begin{pmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,m}B \\ a_{2,1}B & a_{2,2}B & \cdots & a_{2,m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}B & a_{n,2}B & \cdots & a_{n,m}B \end{pmatrix}.$$

The definition works for vectors by thinking of them as matrices with only one column.

The tensor product satisfies many nice properties. For example, it is an associative operation; $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ for any choice of matrices A, B and C. Thus, it makes sense to talk about products such as $A \otimes B \otimes C \otimes \cdots \otimes Z$ without including parentheses, because it doesn't matter in which order the products are evaluated. Next, we have

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

for any choice of matrices A, B, C and D (assuming the sizes of the matrices are such that the products AC and BD make sense). The distributive law holds for tensor products;

$$A \otimes (B + C) = A \otimes B + A \otimes C \quad \text{and} \quad (A + B) \otimes C = A \otimes C + B \otimes C.$$

Also, for matrices A and B and any scalar α, we have

$$(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B).$$

In other words, scalars "float freely" through the tensor product. A word of warning, however, is that the tensor product is not commutative; in general it may be the case that $A \otimes B \neq B \otimes A$.

Not every probability vector v representing a distribution of (X, Y) can be written as a tensor product. For example,

$$v = \begin{pmatrix} 1/2 \\ 0 \\ 0 \\ 1/2 \end{pmatrix}$$

cannot be written as a tensor product. In this distribution we have

$$\Pr[\text{state of } (\mathsf{X}, \mathsf{Y}) \text{ is } 00] = \Pr[\text{state of } (\mathsf{X}, \mathsf{Y}) \text{ is } 11] = \frac{1}{2}.$$

We say that X and Y are *correlated* in this case. The only way a probability vector can be written as a tensor product is when the associated systems are *uncorrelated* (or independent).

As you might have guessed, we do exactly the same thing in the quantum case as in the classical, probabilistic case. If X and Y are qubits having associated superpositions

$$v = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{and} \quad w = \begin{pmatrix} \gamma \\ \delta \end{pmatrix},$$

then the superposition of the pair (X, Y) is

$$v \otimes w = \begin{pmatrix} \alpha\gamma \\ \alpha\delta \\ \beta\gamma \\ \beta\delta \end{pmatrix}.$$

The superposition

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

is an example of a superposition that cannot be written as a tensor product. In the quantum case, this type of correlation between X and Y is special and we call it *entanglement*. We will talk about entanglement a lot during the course.

We use tensor products for independent (or uncorrelated) operations as well. For example, suppose that we have two single-qubit unitary transformations such as

$$U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \text{NOT} \quad \text{and} \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = H,$$

and we perform transformation U on X and V on Y. Then the effect of these two independent operations on any superposition of the pair (X, Y) is determined by the matrix

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

This matrix describes the effect of the two independent unitary operations even on entangled states. For example, applying U to X and V to Y when the pair (X, Y) is in the entangled superposition

from above result in the superposition

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

Dirac Notation

Because the dimension of vectors representing superpositions and matrices representing unitary transformations grows exponentially in the number of qubits, it quickly becomes difficult to write these things down with the notation we have been using. One way to avoid this problem is to use the *Dirac notation*, named after its inventor Paul Dirac (who made many important contributions to quantum mechanics and mathematical physics). The notation is simple but very convenient.

In the Dirac notation, column vectors are represented by "kets", such as

$$|0\rangle \stackrel{\text{def}}{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle \stackrel{\text{def}}{=} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We use symbols such as $|\phi\rangle$ and $|\psi\rangle$ to represent arbitrary vectors (even when the symbols ϕ and ψ have been assigned no meaning by themselves). Any vector indexed by the set $\{0, 1\}$ can be represented by a linear combination of $|0\rangle$ and $|1\rangle$, because $\{|0\rangle, |1\rangle\}$ is a basis for this space of vectors. For instance, we would write

$$\frac{1}{\sqrt{2}}|0\rangle + \frac{i}{\sqrt{2}}|1\rangle$$

to represent the vector

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}.$$

Juxtaposition of kets implicitly refers to the tensor product:

$$|\psi\rangle|\phi\rangle \stackrel{\text{def}}{=} |\psi\rangle \otimes |\phi\rangle.$$

For spaces indexed by $\{00, 01, 10, 11\}$ we define

$$|00\rangle \stackrel{\text{def}}{=} |0\rangle|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle \stackrel{\text{def}}{=} |0\rangle|1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{etc.}$$

The pattern continues in this way for any number of bits. For example, $|1010\rangle$ is a 16 dimensional vector with a 1 in the position indexed by 1010 in binary (which is the eleventh entry because we start with 0000). The vector

$$\frac{1}{\sqrt{2}}|000000\rangle + \frac{1}{\sqrt{2}}|111111\rangle$$

would be written

$$\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \vdots \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \left.\begin{matrix} \\ \\ \\ \\ \\ \end{matrix}\right\} \text{62 zeroes}$$

in the usual vector notation. An arbitrary vector with entries indexed by $\{0,1\}^n$, which perhaps refers to a superposition of n qubits, can again be written as a linear combination of the elements in the basis

$$\{|x\rangle \,:\, x \in \{0,1\}^n\},$$

for instance as

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

Example 1. Let us suppose that we have two qubits X and Y in the superposition

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

Using the Dirac notation we write this superposition as

$$\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle\,.$$

Suppose now that we perform a Hadamard transform to the first qubit and do nothing to the second qubit. We can determine the effect of these operations on the above superposition of (X,Y) by computing

$$H \otimes I = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

and multiplying:

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

which is written

$$\frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle - \frac{1}{2}|11\rangle$$

in the Dirac notation.

However, we can perform this computation directly and more easily by not converting back and forth between notations, and instead just sticking with the Dirac notation. Let us start by noting that

$$H|0\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \quad \text{and} \quad H|1\rangle = \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle .$$

The starting superposition is

$$\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$$

which is equivalent to

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

The superposition after performing a Hadamard transform on the first qubit and doing nothing (performing the identity operation) to the second qubit is

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

Substituting for $H|0\rangle$ and $H|1\rangle$ and using the distributive law, we get

$$\begin{aligned} &\frac{1}{\sqrt{2}}(H|0\rangle)|0\rangle + \frac{1}{\sqrt{2}}(H|1\rangle)|1\rangle \\ &\qquad = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle\right)|0\rangle + \left(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\right)|1\rangle \\ &\qquad = \frac{1}{2}|0\rangle|0\rangle + \frac{1}{2}|1\rangle|0\rangle + \frac{1}{2}|0\rangle|1\rangle - \frac{1}{2}|1\rangle|1\rangle \\ &\qquad = \frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle - \frac{1}{2}|11\rangle . \end{aligned}$$

The two computations of course agree. The second method is much easier (once you know the basics of how it works), particularly for larger numbers of qubits.

For every ket $|\psi\rangle$ there is a corresponding object $\langle\psi|$, called a "bra". You may think that this is a strange name for a mathematical object, but the names "bra" and "ket" are derived from the the

fact that when you put a bra and a ket together, you get a "bracket". For this to make sense you need to know what a bra is—for any vector $|\psi\rangle$ we define

$$\langle\psi| = (|\psi\rangle)^\dagger,$$

which is the conjugate transpose of $|\psi\rangle$. In other words, $\langle\psi|$ is the row vector you get by transposing $|\psi\rangle$ and taking the conjugate of each of its entries. For instance:

$$|\psi\rangle = \begin{pmatrix} \frac{1+i}{2} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad \Rightarrow \quad \langle\psi| = \begin{pmatrix} \frac{1-i}{2} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Now, when you juxtapose a bra and a ket, the implicit operation is matrix multiplication (thinking of the vectors as matrices with only one row or one column). A row vector times a column vector results in a scalar, and this scalar will be the *inner product* (or *bracket*) of the vectors involved. For instance, if

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{and} \quad |\phi\rangle = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$$

then

$$\langle\psi|\phi\rangle \stackrel{\text{def}}{=} \langle\psi|\,|\phi\rangle = \begin{pmatrix} \overline{\alpha} & \overline{\beta} \end{pmatrix} \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \overline{\alpha}\gamma + \overline{\beta}\delta.$$

When you have an expression such as

$$|\psi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x\,|x\rangle$$

it is easy to express $\langle\psi|$ using similar notation; it is

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

When you juxtapose a ket and a bra in the opposite order, such as

$$|\psi\rangle\,\langle\phi|\,,$$

you do not get a scalar—a column vector times a row vector gives you a matrix. It is easy to determine the action of this matrix on another vector. For instance,

$$|\psi\rangle\,\langle\phi|\,|\gamma\rangle = |\psi\rangle\,\langle\phi|\gamma\rangle = \langle\phi|\gamma\rangle\,|\psi\rangle\,.$$

Later on when we wish to speak at a higher level of abstraction about computational problems, algorithms, etc., we may refer to $|x\rangle$ where x is some arbitrary mathematical object (such as a matrix, a graph, or a list of numbers). In this case the interpretation is that we are implicitly referring to the encoding of x with respect to some agreed upon encoding scheme.