CS 229r Information Theory in Computer Science Jan 29, 2019

Lecture 25

Instructor: Madhu Sudan *Scribe: Arsen Vasilyan*

1 Basics of quantum mechanics

1.1 The double slit experiment

Quantum mechanics can be thought of as probability theory with negative numbers. Sometimes particles behave in very strange ways, and one important example of this is the two-slit experiment. Let a photon travel towards a screen with two slits A and B and consider the probability it will be detected by a detector on the other side of the screen in the following three situations:

1. Both slits A and B are open.
2. Only slit A is open.
3. Only slit B is open.

Experiment shows that if the detector is placed in a certain location, then the probability of detecting the proton in case (1) is zero, but the probability of detecting it in cases (2) and (3) is non-zero. In other words, opening a slit makes us less likely to see an electron in a certain place.

If we think of the photon as a particle that sometimes goes through one slit and sometimes through the other, we cannot explain such behavior. Instead, quantum mechanics explains it through the concepts of complex amplitudes that have the following properties:

- For every location, there is a certain amplitude of observing the electron there. This is analogous to the probability density in the classical world.
- **Connection to probability**: the square of the absolute value of an amplitude equals to the probability of observing the electron.
- **Superposition**: the amplitudes arising when both of the slits are open equal to the sum of the amplitudes arising when only A is open and the amplitudes arising when only B is open.

These amplitudes can be negative, and therefore it is possible that they cancel each other out at some points in space.

1.2 A qubit

Having roughly sketched the physics behind quantum computation, we turn to the simplest quantum system: the qubit. A qubit is the quantum analogue of the classical bit.

Recall that our uncertainty whether a bit equals to zero or one can be represented as a real number p between zero and one. In the case of qubit, a quantum state of a qubit can be represented a a vector in[1] $\mathbb{C}^2$. Specifically, we write a state ψ of a qubit as a linear combination:

$$\psi = \alpha|0\rangle + \beta|1\rangle$$

Here $|0\rangle$ and $|1\rangle$ is a notation used in quantum mechanics for a basis in $\mathbb{C}^2$. The values α and β are the amplitudes, which satisfy

$$|\alpha|^2 + |\beta|^2 = 1$$

[1] $\mathbb{C}$ is the field of complex numbers. For the most of these notes, a reader can think of all the numbers as being real numbers.

For n quits, a state ψ is a vector in $\mathbb{C}^{2^n}$. It can be represented as:

$$\psi = \sum_{s \in \{0,1\}^n} \alpha_s |s\rangle$$

Analogously, here α_s is the amplitude of string s. Note that each of the 2^n strings has its own amplitude. Again, we have a normalization condition:

$$\sum_{s \in \{0,1\}^n} |\alpha_s|^2 = 1$$

1.3 Operations on qubit

If we have access to a quantum state and have an infinite amount of computational power, we can apply any *unitary* transformation to it. An operator T is unitary if:

- For every unit vector $|x\rangle$, the result of applying the operator $T|x\rangle$ is also a unit vector.
- For every pair of orthogonal vectors $|x\rangle$ and $|y\rangle$, it is the case that $T|x\rangle$ and $T|y\rangle$ are orthogonal.
- T is a linear operator, i.e. for every ϕ and ψ we have that:

  $$T(\phi + \psi) = T\phi + T\psi$$

  And for every complex number c, we have:

  $$T(c \cdot \psi) = c \cdot T(\psi)$$

Almost any Boolean operator can be expressed through unitary operations. The only caveat is that unitary operations have to be reversible, but some Boolean operations are not reversible, say the following set x to $x \wedge y$. To get around this issue, we use a "fresh" qubit z that is in state $|0\rangle$ and execute the following unitary operation:

- Set z to $(x \wedge y) \oplus z$
- Let the values of x and y remain the same.

This operator takes a basis state $|xyz\rangle$ to $|xy(x \wedge y \oplus z)\rangle$. This logical operation is reversible, and therefore the operator is unitary.

Some other interesting operators are:

- The flipping gate:
  - Takes $|0\rangle$ to $|1\rangle$.
  - Takes $|1\rangle$ to $|0\rangle$.
- The following operator:
  - Takes $|0\rangle$ to $|0\rangle$.
  - Takes $|1\rangle$ to $-|1\rangle$.
- The Hadamard gate:
  - Takes $|0\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$.
  - Takes $|1\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

  Here, the factor $\frac{1}{\sqrt{2}}$ makes sure the normalization conditions are satisfied.

1.4 Measurement

Given a qubit, we can apply unitary transformations to it and obtain a state $\alpha|0\rangle + \beta|1\rangle$. Then, we can *measure* it, which leads to the following:

- With probability $|\alpha|^2$ we receive "0", and with probability $|\beta|^2$ we receive "1".
- If we receive "0" the quantum state *collapses* to the state $|0\rangle$.
- If we receive "1" the quantum state collapses to the state $|1\rangle$.

For example, by measuring this way we cannot distinguish $\alpha|0\rangle + \beta|1\rangle$ from $\alpha|0\rangle - \beta|1\rangle$. To accomplish this, we should "rotate" the state first into an appropriate basis, and only then perform the measurement.

Analogously, when measuring the state:

$$\psi = \sum_x c_x|x\rangle$$

We receive "x" with probability $|c_x|^2$, which collapses the state to $|x\rangle$.

Finally, for systems consisting of more than one qubit, we can perform a richer class of "partial" measurements. Given a system in the state:

$$\sum c_{xy}|xy\rangle$$

We can measure x alone, and receive "α"" with probability $\sum_y |c_{\alpha y}|^2$, and as a result the state will collapse to:

$$\frac{\sum_y c_{\alpha y}|y\rangle}{\sum_z |c_{\alpha z}|^2}$$

2 Quantum information

2.1 How much information do qubits contain?

To simulate an n-qubit system, we seem to need 2^n time and space on a classical machine. This is the potential complication of quantum computation, but also the source of its power.

How much (classical) information can we transmit to someone by giving them n qubits? On one hand, we saw that n qubits have (roughly, ignoring normalization) 2^n degrees of freedom, which we can carefully control through unitary transformations. On the other hand, the quantum measurement cannot read off the whole state, but only see some outcomes with probabilities associated with the quantum state.

Before turning to this question, we address one more complication. We can be uncertain about which quantum state a system is in. This uncertainty is the usual classical uncertainty and not the quantum superposition. For example, we can believe that with probability $1/2$ the state of the system is $|0\rangle$ and with probability $1/2$ it is in the state $|1\rangle$. Such states are called *mixed states*, as opposed to the *pure states* we have been studying so far.

So, how many degrees of freedom does a mixed state have. On the first sight, one can think that it has roughly 2^{2^n} degrees of freedom: one degree of freedom for each pure state. However, it turns out that it really only has roughly 2^{2n} degrees of freedom, for the following reason: there are mixtures of pure states that cannot be distinguished from each other, even given access to infinitely many copies of these mixed states. For this reason, we say that those mixed states are identical.

One can decide if two mixtures of pure states give rise to the same mixed state by computing the *density matrices* of the two mixtures and checking that they are the same. The density matrix is defined as follows:

- View the pure state v as a column vector in $\mathbb{C}^n$. Then, the density matrix of v is vv^T, where v^T is the conjugate transpose of v.
- For a mixture of pure states, the density matrix is defined as the convex combination of the rank-1 density matrices of the pure states it is a mixture of, where weights are the same as in the mixture.

Density matrices have the following important properties:

1. It can be shown that for any set of unitary transformations and measurements, the results of the measurements depend only on the density matrix of the state.

2. The density matrix also has to be *Hermitian* (i.e. its conjugate transpose should equal to itself), positive definite and its trace should equal to 1. This is true, because each of the rank-1 components has these properties, and they are preserved when one takes a convex combination.

3. It can also be shown that every Hermitian matrix with trace 1 corresponds to the density matrix of some mixed state.

4. Classical probability distributions can also be expressed through density matrices as mixtures of different basis vectors. In other words, density matrices whose off-diagonal terms equal to zero correspond to classical probability distributions.

Finally, the following example demonstrates a pair of mixtures that are indistinguishable: ϕ is $\cos\theta|0\rangle + \sin\theta|1\rangle$ with probability $1/2$ and is $\cos\theta|0\rangle - \sin\theta|1\rangle$ with probability $1/2$. ψ is $|0\rangle$ with probability $\cos^2\theta$ and is $|1\rangle$ with probability $\sin^2\theta$. The reader can check that these two distributions give rise to the same density matrices, and therefore give rise to the same mixed state.

2.2 The von Neumann entropy

There is a quantity in quantum information which has roughly the following operational meaning. Suppose one is given a sequence of pure states $\psi_1, ..., \psi_N$ all drawn from a mixed state Ψ. Then, how many qubits are needed to convey $\psi_1, ..., \psi_N$ to someone who knows Ψ?

There are two key differences between this setup and the classical communication setup:

1. We have quantum states instead of bits.

2. We communicate quantum bits, instead of classical bits.

To define the von Neumann entropy, first observe that the density matrix M of Ψ has real positive eigenvectors, since M is Hermitian and positive definite. Additionally, the eigenvectors sum to one, since the trace of M is one. The von Neumann entropy is defined as the entropy of the distribution formed by this eigenvalues.

Note that the von Neumann entropy is strict generalization of the classical entropy, for the following reason. For a classical system we have the probabilities on the diagonal of the density matrix, and all the other values in the matrix equal to zero. Therefore, the probabilities precisely equal to the eigenvalues of the density matrix, and the two definitions coincide.

2.3 More quantum communication

Now, we make precise what quantum communication is. There are two players: Alice and Bob, who communicate through a channel. There are three sets of qubits: Alice's qubits A, Bob's qubits B and the qubits that belong to the channel Z. Alice and Bob take turns, and at each step one of them applies an arbitrary unitary operator the union of the channel qubits and his/her qubits. For instance, when it is Alice's turn, she applies an arbitrary unitary transformation to the set of qubits $A \cup Z$.

More generally, Alice and Bob can also communicate using classical channel together with a quantum channel. Overall there are three kinds of resources Alice and Bob can use:

1. Classical communication.

2. Quantum communication, which formally equals to $|Z|$, the number of qubits in the channel.

3. Shared entanglement, which equals to the number of entangled pairs of qubits Alice and Bob start out with. This is the quantum analogue of shared randomness.

These resources are in an exact correspondence with the three objectives:

1. Send a classical bit.
2. Send a qubit.
3. Share entanglement.

We use the following symbols for the resources and objectives above: $[C \to C]$ represents a classical bit, $[q \to q]$ represents a qubit and $[qq]$ represents a shared entangled pair of qubits.

In general, we would like to know how much of the resources can be used to achieve a given set of objectives. For instance, the two of the following are immediate:

$$[q \to q] \geq [C \to C] \tag{1}$$

$$[q \to q] \geq [qq] \tag{2}$$

The Equation 1 means that by sending a qubit, Alice can send a classical bit to Bob. The Equation 2 means that by sending a qubit, Alice can establish one shared entangled pair between her and Bob. Both of these relations are trivial. Using this notation, we present two non-trivial ones.

Super-dense coding:

$$[qq] + [q \to q] \geq 2[C \to C] \tag{3}$$

Quantum teleportation

$$[qq] + 2[C \to C] \geq [q \to q] \tag{4}$$

In other words, one entangled pair and one transmitted qubit suffice to communicate two classical bits, and one entangled pair and two classical bits suffice to send one qubit.

Finally, we also have the trivial "resource-wasting" relations:

$$[q \to q] \geq 0 \tag{5}$$

$$[C \to C] \geq 0 \tag{6}$$

$$[qq] \geq 0 \tag{7}$$

Clearly, given these relations we can combine them and obtain many others of the form:

$$\alpha[q \to q] + \beta[qq] + \gamma[C \to C] \geq 0$$

Formally, we say that the relations obtainable from inequalities (1-7) is in the *conic hull* of (1-7), which contains everything all inequalities that can be obtained by the convex combination of (1-7) and scaling by a constant.

We state the following highly non-trivial theorem:

Theorem 1. *The relations in the conic hull of (1-7) are the only true relations.*

2.4 Super-dense coding

Finally, we show how one can achieve super-dense coding. Suppose Alice and Bob start with an entangled pair of qubits:

$$\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$$

Now, Alice wishes to communicate a number x in $\{1, 2, 3, 4\}$, for each of the values of x, Alice applies a different unitary operator to her qubit:

- If $x = 1$, she applies the unitary transformation that sends $|0\rangle \to |0\rangle$ and $|1\rangle \to |1\rangle$.
- If $x = 2$, she applies the unitary transformation that sends $|0\rangle \to |1\rangle$ and $|1\rangle \to |0\rangle$.
- If $x = 3$, she applies the unitary transformation that sends $|0\rangle \to |0\rangle$ and $|1\rangle \to -|1\rangle$.
- If $x = 4$, she applies the unitary transformation that sends $|0\rangle \to |1\rangle$ and $|1\rangle \to -|0\rangle$.

Then Alice sends her qubit to Bob. Then, Bob has a pair of qubits:

- If $x = 1$, then $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$
- If $x = 2$, then $\frac{1}{\sqrt{2}}(|10\rangle + |01\rangle)$
- If $x = 3$, then $\frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$
- If $x = 4$, then $\frac{1}{\sqrt{2}}(|10\rangle - |01\rangle)$

These vectors are pairwise orthogonal, and therefore Bob can determine which one it is by applying the appropriate unitary transformation and then performing a measurement.

This completes the proof that Alice can transmit two classical bits to Bob by using a shared entangled pair and sending one qubit.