CS 229r Information Theory in Computer Science Apr 30, 2019

Lecture 25

Instructor: Madhu Sudan *Scribe: Eric Bornstein*

1 Today

We will be talking about Quantum Information Theory:

- Basics of Quantum Mechanics: state, transformations, observations
- Basic Quantum Interactions tasks
- Some fundamental limits (without proof)
- Protocols for Quantum Communication

2 Short Summary of Quantum Mechanics

2.1 Quantum State

Definition 1 (Qubit)**.** An electron is represented by a qubit, which is unit vector in $\mathbb{C}^2$. The standard orthonormal basis vectors are $|0\rangle$ and $|1\rangle$ (spin down and spin up). The inner product is $\langle 0|1\rangle = 0$ and $\langle 0|0\rangle = \langle 1|1\rangle = 1$.

Definition 2 (n Qubit System)**.** A system of n qubits is a unit vector in $\mathbb{C}^{2^n}$. The standard orthonormal basis vectors are now $|x\rangle$ for $x \in \{0,1\}^n$. The inner product is $\langle x|y\rangle = 1_{x=y}$ for $x, y \in \{0,1\}^n$.

Each standard basis vector is a possible state for the electrons. So, a general n qubit system will have state $\psi = \sum_{x\in\{0,1\}^n} c_x|x\rangle$ where $\sum_x |c_x|^2 = 1$. Throughout this lecture, we will only be concerned with $c_x \in \mathbb{R}$, so we can replace $|c_x|^2$ with c_x^2.

These called are pure states. A mixed state is a distribution over qubits / pure states i.e $\psi = \sum_i p_i\psi_i$ where $p_i \geq 0$, $\sum_i p_i = 1$ and ψ_i are pure states.

2.2 Quantum Operations

We can perform any unitary operation T, meaning T is linear and preserves inner product. So,

$$\langle T(|x\rangle)|T(|y\rangle)\rangle = \langle x|y\rangle$$

meaning that $T(|x\rangle)$ is a unit vector for $x \in \{0,1\}^n$ and that $T(|x\rangle)$ and $T(|y\rangle)$ are orthogonal if $x \neq y \in \{0,1\}^n$ Also,

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

These are all invertible operations. However, some computations we would like to perform are not invertible operations. For example, given bits x and y, we would want to compute $x \wedge y$, the and of the two bits. When working with qubits, we instead use (x, y, z) and compute $(x, y, z \oplus (x \wedge y))$, which is an invertible map. So, our unitary operator on C^{2^3} is given by $|xyz\rangle \mapsto |xy(z \oplus (x \wedge y))\rangle$.

Other examples include

- The bit flip or X gate is given by $|0\rangle \mapsto |1\rangle$ and $|1\rangle \mapsto |0\rangle$.
- Reflection of $|1\rangle$ axis or Z gate is given by $|0\rangle \mapsto |0\rangle$ and $|1\rangle \mapsto -|1\rangle$
- The Hadamard or H gate is given by $|0\rangle \to \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ and $|1\rangle \to \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle$. Here, $\frac{1}{\sqrt{2}}$ is a normalizing constant.

2.3 Measurement

If we have a qubit $\alpha|0\rangle + \beta|1\rangle$. If we measure, we will get 0 with probability α^2 and 1 with probability β^2. These probabilities add to 1 because the qubit is a unit vector. Notice, that we are not revealing what α and β are. This is analogous to getting an $X \sim Bern(p)$. Although getting X gives us information of p, it will not reveal the value of p.

Similarly, when measuring an n-qubit system: $\psi = \sum_{x\in\{0,1\}^n} c_x|x\rangle$, we get each x with probability c_x^2. Measuring a mixed state $\psi = \sum_i p_i\psi_i$ is equivalent to first picking a ψ_i, where ψ_i has a p_i probability of being chosen, and then measure a qubit whose value is ψ_i. When we measure, we will only get n bits.

Measurement is distinct from all other operators as it is not unitary.

2.3.1 Partial Measurement

We can also observe some of the qubits in a quantum system. If $\psi = \sum_{x\in\{0,1\}^a, y\in\{0,1\}^b} c_x|xy\rangle$ we can measure just the x coordinates. We will see α with probability $\sum_{y\in\{0,1\}^b} c_{\alpha y}^2$. The resulting system has b undetermined qubits and is of the form $\frac{\sum_{y\in\{0,1\}^b} c_{\alpha y}|y\rangle}{\sum_{y\in\{0,1\}^b} c_{\alpha y}^2}$.

3 Quantum Information Theory

You might wonder why we don't just use classical bits to simulate the distribution of a qubit by keeping track of c_x for each $x \in \{0,1\}^n$. However, to simulate n qubit system using bits, we think we need 2^n time and space. Because a quantum system simulates itself far better than classical bits do, maybe quantum computers can also compute other functions in polynomial time that take classical bits exponential time.

Pure n-qubit systems have about 2^n complex degrees of freedom. If we were able to observe all the c_x, we could use a n qubit system to compress 2^n complex values. However, because we cannot just measure the c_x an n-qubit system doesn't seem to have 2^n "carrying capacity" of information. A mixed quantum state is a distribution over pure states, so it can be represented by 2^{2^n} bits. But, mixed states have at most 2^{2n} real number degrees of freedom. See the example below of two mixed states represented by different mixtures that are indistinguishable.

3.1 Density Matrix

Definition 3 (Density Matrix)**.** For each pure state $\psi = \sum_{x\in\{0,1\}^n} c_x|x\rangle$ where $c_x \in \mathbb{C}$, we associate a matrix M with entries $M_{x,y} = c_x\overline{c_y}$.

If $\vec{c}_x$ is the vector of coefficients, then $M = \vec{c}_x\vec{c}_x^{\dagger}$ the outer product of the vector of coefficients with itself where $\vec{c}_x^{\dagger}$ is the conjugate transpose of $\vec{c}_x$.

If $\psi = p_i\psi_i$ is a mixed state of pure qubits, then its density matrix is $M = \sum_i p_iM_i$ where M_i is the density matrix of ψ_i.

Fact 4. *If ψ and ϕ correspond to the same density matrix, there is nothing that we can do to distinguish ψ and ϕ.*

Exercise 5. *Show that*

$$\phi = \begin{cases} \cos\theta|0\rangle + \sin\theta|1\rangle & wp\ 1/2 \\ \cos\theta|0\rangle - \sin\theta|1\rangle & wp\ 1/2 \end{cases}$$

$$\psi = \begin{cases} |0\rangle & wp\ \cos^2\theta \\ |1\rangle & wp\ \sin^2\theta \end{cases}$$

are indistinguishable.

Proof. The first density matrix is

$$M = \frac{1}{2}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}^T + \frac{1}{2}\begin{pmatrix}\cos\theta\\ -\sin\theta\end{pmatrix}\begin{pmatrix}\cos\theta\\ -\sin\theta\end{pmatrix}^T = \begin{pmatrix}\cos^2\theta\\ \sin^2\theta\end{pmatrix} = \cos^2\theta\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}^T + \sin^2\theta\begin{pmatrix}0\\1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}^T$$

which is the density matrix of the second system. □

Question: If we have sequence of pure states $\psi_1, \psi_2, \ldots, \psi_N$ drawn from mixed state ψ, how many qubits are needed to convey $\psi_1, \ldots, \psi_N$ to someone who knows ψ?

Answer: The Quantum Entropy of ψ multiplied by N.

This is roughly what quantum entropy captures.

Definition 6 (Quantum Entropy)**.** The Quantum Entropy of a system ψ is the $H(X)$ where $X \sim Mult(1, \lambda_1, \ldots, \lambda_n)$ and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of M, the density matrix for ψ. This distribution is drawing 1 element from 1 to n where the probability of picking i is λ_i.

It is not obvious that this definition makes sense. However, it happens to be true that all the eigenvalue are non-negative and $\sum_i \lambda_i = 1$.

This is actually a generalization of Shannon Entropy.

Exercise 7. *Show that this definition agrees with the definition of Shannon Entropy on classical bit systems.*

Proof. Classical bits $b_1, \ldots, b_n$ can be represented as a pure state $|b_1, \ldots, b_n\rangle$. A distribution over classical bits is then $|x\rangle$ with probability p_x where $p_x \geq 0$ and $\sum_x p_x = 1$. The density matrix for the pure states are 0 except entry $M_{x,x} = 1$. Thus, the distribution density matrix is 0 except the diagonal has entries $M_{y,y} = p_y$ for $y \in \{0,1\}^n$. The eigenvalues are just the diagonal entries, so the Quantum Entropy is just the entropy of the p_x, which is the Shannon Entropy. □

4 Quantum Communication

For quantum communication, Alice and Bob each have their own qubit registers. The public channel is a set of qubit registers. The messages take the form of performing a unitary operation on the combined set of public registers and one's private registers. See Figure 4.

There are three resources that we can use. Each has a representation with a minus sign to show consumption:

- $-[c \to c]$ Sending a bit.
- $-[q \to q]$ Using a quantum channel. We quantify qubits based on how many of the qubits in the public channel need to be used.
- $-[qq]$ Using a shared entanglement. This is sort of the analog to private randomness. The basic entangle pair is a system of $\frac{1}{\sqrt{2}}|00\rangle_{AB} + \frac{1}{\sqrt{2}}|11\rangle_{AB}$. Here, Alice gets the first qubit and Bob gets the second qubit. Alice cannot choose what to send Bob, but if they both measure the qubit, they will have the same bit.

Similarly, there are three objectives

- $+[c \to c]$ Sending a bit.
- $+[q \to q]$ Sending a qubit.
- $+[qq]$ Generating a shared entanglement.

Proposition 8. $[q \to q] \geq [c \to c]$. *This means that we can send a classical bit by sending a qubit.*

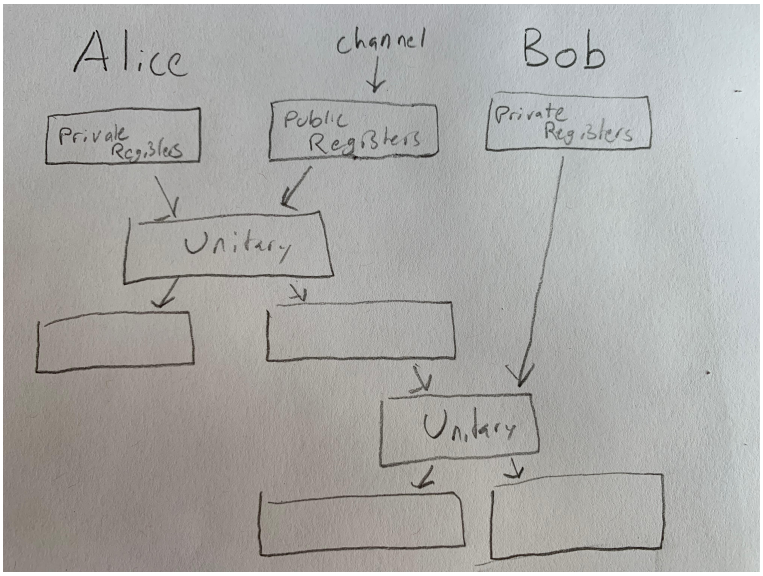

Figure 1: Quantum Communication Channel

Proof. If we want to send b, we send $|b\rangle$. This is always measured as b. □

Proposition 9. $[q \to q] \geq [qq]$

Proof. We entangle two qubits, then send one of the entangled qubits. □

Lemma 10 (Super Dense Coding)**.** $[qq] + [q \to q] \geq 2[c \to c]$

Proof. The shared entanglement is $\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})$. Alice will take her two bits and choose one of four unitary operators to apply to her entangled qubit.

1. Do nothing: $|0\rangle \mapsto |0\rangle, |1\rangle \mapsto |1\rangle$
2. Flip bits: $|0\rangle \mapsto |1\rangle, |1\rangle \mapsto |0\rangle$
3. Reflection of $|1\rangle$ axis: $|0\rangle \mapsto |0\rangle, |1\rangle \mapsto -|1\rangle$
4. Perform 2 and 3: $|0\rangle \mapsto |1\rangle, |1\rangle \mapsto -|0\rangle$

Alice then sends her qubit.

Bobs qubits are now proportional to either $|00\rangle + |11\rangle, |10\rangle + |01\rangle, |00\rangle - |11\rangle$ or $|10\rangle - |01\rangle$. These are orthogonal. So, he can apply a Unitary operation to distinguish which case he is in. □

Lemma 11 (Teleportation)**.** $[qq] + 2[c \to c] \geq [q \to q]$

Theorem 12. *This is all that is possible. Specifically, if* $\alpha[q \to q] + \beta[qq] + \gamma[c \to c] \geq 0$ *for* $\alpha, \beta, \gamma \in \mathbb{R}$ *Then* (α, β, γ) *is a conic combination of the four above* $(1, 0, -1), (1, -1, 0), (1, 1, -2), (-1, 1, 2)$ *and tuples* $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ *corresponding to wasting resources.*

The conic section of $x_1, \ldots, x_n$ is the set $\{\sum_i \alpha_i x_i : \alpha_i \geq 0\}$

Corollary 13. *Holevo's Theorem is an immediate consequence. It states that* $a[q \to q] \geq b[c \to c]$ *iff* $a \geq b$,

5 Conclusion

Harvard offers many courses related to Theoretical Computer Science. Madhu will post a list of course that students in this class may be interested in taking.

References

[1] Quantum Information Theory (2nd Edition), by Mark M. Wilde. Available as arxiv:1106.1445