

08. Quantum Information Theory: Basics

- 1. Quantum Information
- 2. Qubits
- 3. Qubit Transformations
- 4. The No-Cloning Theorem
- 5. Quantum Computation

1. Quantum Information

Analogy with Shannon Information

- Recall: Shannon info is associated with classical communication protocols.
- Suggests: Replace a classical communication device with a quantum communication device...

<u>Shannon information</u> : What is produced by a <i>classical</i> source that is required to be reproducible at a receiver if the transmission is be counted a success.	<u>Quantum information</u> : What is produced by a <i>quantum</i> source that is required to be reproducible at a receiver if the transmission is be counted a success.
<u>Shannon information content</u> of the value x_i of a random variable $X = \{x_i, p_i\}$ is given by $-\log_2 p_i$.	<u>Quantum information content</u> of a vector state $ \psi_i\rangle$ of an ensemble $\{ \psi_i\rangle, p_i\}$ is given by $-\ln p_i$.
<u>Shannon entropy</u> $S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$ of a random variable X is the average value of the Shannon info content $-\log_2 p_i$ of x_i .	<u>von Neumann entropy</u> $S_{\text{vN}}(\rho) = -\sum_i p_i \ln p_i$ of a density operator $\rho = \sum_i p_i \psi_i\rangle \langle \psi_i $ is the average value of the quantum info content $-\ln p_i$ of $ \psi_i\rangle$.
<u>A bit (unit of measure sense)</u> is the average Shannon info content in a binary random variable.	<u>A "qubit" (unit of measure sense)</u> is the average quantum info content in a 2-dim density operator ρ .

Uncertainty interpretation

- *Shannon Entropy:*

$$S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$$

- $X = \{x_1, \dots, x_n\}$, where x_i is a state produced by a classical information source, and p_i is a probability distribution over such states.



Expected value of Shannon information gained upon measurement of X with outcome x_i .

- *von Neumann Entropy:*

$$S_{\text{vN}}(\rho) = -\sum_i p_i \ln p_i$$

- $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle$ is a vector state produced by a quantum information source, and p_i is a probability distribution over such states.



Expected value of quantum information gained upon measurement of ρ with outcome $|\psi_i\rangle$.

*But: $S_{\text{vN}}(\rho)$ measures mixedness of ρ .
And: Mixedness does not necessarily entail uncertainty...*

"Information" compression interpretation

- *Shannon Entropy:*

$$S_{\text{Shan}}(X) = -\sum_i p_i \log_2 p_i$$

- $X = \{x_1, \dots, x_n\}$, where x_i is a state produced by a classical information source, and p_i is a probability distribution over such states.



Specifies the minimal number of bits required to encode the output of a Shannon information source.

- *Bit as unit of Shannon information.*

- *von Neumann Entropy:*

$$S_{\text{vN}}(\rho) = -\sum_i p_i \ln p_i$$

- $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle$ is a vector state produced by a quantum information source, and p_i is a probability distribution over such states.



Specifies the minimal number of qubits required to encode the output of a quantum information source.

- *Qubit as unit of quantum information.*

Is quantum information different from Shannon information?

Claim: No *fundamental* difference between quantum and Shannon information: just a difference in types of *sources*.

Information (in communication theory) = What is produced by a source that is required to be reproducible at the receiver if the transmission is to be counted a success.

I. Classical source

- Abstractly: Produces letters from a set $\{x_1, x_2, \dots, x_n\}$ with probabilities $p_i = p(x_i)$.
- Messages = sequences of letters. Ex: $x_7x_3x_4\dots$
- Concretely: Produces *physical systems* (e.g., on-off switches) in *classical states* $\{x_1, x_2, \dots, x_n\}$.
- Output = sequence of physical systems in classical states. Ex: $x_7x_3x_4\dots$

II(a). Quantum Non-Entanglement Source

- Produces *physical systems* (e.g., electrons) in *non-entangled quantum states* $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$.
- Output = sequence of physical systems in quantum pure states. Ex: $|\psi_7\rangle|\psi_3\rangle|\psi_4\rangle\dots$

II(b). Quantum Entanglement Source

- Produces *physical systems* (i.e., electrons) in *entangled quantum states* which include other systems inaccessible to the source.
- Output = sequence of physical systems in quantum entangled states.

Example of II(b):

$B = \{B_1, B_2, \dots\} = \{\text{electrons produced by source}\}$

$A = \{A_1, A_2, \dots\} = \{\text{electrons entangled with source electrons}\}$

$C = \{C_1, C_2, \dots\} = \{\text{"target" electrons at receiver}\}$

- Suppose: Electron B_i is produced at source in entangled state $|\psi\rangle_{A_i B_i}$ with electron A_i .
- Goal: To reproduce this entangled state at receiver, but between A_i and C_i : $|\psi\rangle_{A_i C_i}$
- In general: If source produces systems in the sequence of states

$$|\psi\rangle_{A_i B_i} |\psi'\rangle_{A_j B_j} |\psi''\rangle_{A_k B_k} \dots,$$

then successful transmission occurs if receiver produces systems in the sequence of states

$$|\psi\rangle_{A_i C_i} |\psi'\rangle_{A_j C_j} |\psi''\rangle_{A_k C_k} \dots .$$

2. Qubits

Four senses of "bit"!

- (1) *Unit of measure sense.* A unit of Shannon information: One bit = average amount of *Shannon info content* in a *binary random variable* (i.e., a random variable with just two possible values).
- (2) *Value sense.* A value of a binary random variable: 0 or 1.
- (3) *Physical system sense.* A 2-state classical physical system.
- (4) *Physical state sense.* A state of a 2-state classical system.

Four senses of "qubit"!

- (1) *Unit of measure sense.* A unit of quantum information: One qubit = average amount of *quantum info content* in a 2-dim density operator state (i.e., a density operator corresponding to an ensemble with just two possible vector states).
- (2) *Value sense.* An eigenvalue of a 2-dim density operator.
- (3) *Physical system sense.* A 2-state quantum physical system.
- (4) *Physical state sense.* A state of a 2-state quantum system.

General form of a qubit

$$|Q\rangle = a|0\rangle + b|1\rangle, \text{ where } |a|^2 + |b|^2 = 1$$

According to the Eigenvalue-eigenvector Rule

- $|Q\rangle$ has no determinate value (of Hardness, say).
- It's value only becomes determinate (0 or 1; *hard* or *soft*) when we measure it.
- All we can say about $|Q\rangle$ is:
 $\Pr(\text{value of } |Q\rangle \text{ is } 0) = |a|^2$
 $\Pr(\text{value of } |Q\rangle \text{ is } 1) = |b|^2$

Common Claim: A qubit $|Q\rangle = a|0\rangle + b|1\rangle$ encodes an arbitrarily large amount of quantum information, but at most only one bit's worth of (Shannon) information in a qubit is *accessible*.

Why?

- a and b encode an arbitrarily large amount of "information".
- But the outcome of a measurement performed on $|Q\rangle$ is its collapse to either $|0\rangle$ or $|1\rangle$, which each encode just one classical bit.

Shannon?

3. Qubit Transformations

Single-Qubit Transformations

- Let $|0\rangle$ and $|1\rangle$ be given the matrix representations: $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- Define the following operators that act on $|0\rangle$ and $|1\rangle$:

$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$Y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
<i>Identity</i>	<i>Negation</i>	<i>Negation/Phase-change</i>	<i>Phase-change</i>

$I 0\rangle = 0\rangle$	$X 0\rangle = 1\rangle$	$Y 0\rangle = - 1\rangle$	$Z 0\rangle = 0\rangle$
$I 1\rangle = 1\rangle$	$X 1\rangle = 0\rangle$	$Y 1\rangle = 0\rangle$	$Z 1\rangle = - 1\rangle$

$$\mathfrak{H} = \begin{pmatrix} \sqrt{1/2} & \sqrt{1/2} \\ \sqrt{1/2} & -\sqrt{1/2} \end{pmatrix}$$

Hadamard operator

$$\mathfrak{H}|0\rangle = \sqrt{1/2} (|0\rangle + |1\rangle)$$

$$\mathfrak{H}|1\rangle = \sqrt{1/2} (|0\rangle - |1\rangle)$$

*Takes a basis qubit and
outputs a superposition*

Multi-Qubit Transformations

- Let $\{|0\rangle_1, |1\rangle_1\}, \{|0\rangle_2, |1\rangle_2\}$ be bases for the single qubit state spaces $\mathcal{H}^{(1)}, \mathcal{H}^{(2)}$.

- Then: A basis for the 2-qubit state space $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ is given by

$$\{|0\rangle_1|0\rangle_2, |0\rangle_1|1\rangle_2, |1\rangle_1|0\rangle_2, |1\rangle_1|1\rangle_2\}$$

- Aside: Another basis for $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$ is given by

$$\{|\Psi^+\rangle, |\Psi^-\rangle, |\Phi^+\rangle, |\Phi^-\rangle\},$$

where:

$$|\Psi^+\rangle = \sqrt{1/2} (|0\rangle_1|0\rangle_2 + |1\rangle_1|1\rangle_2)$$

$$|\Psi^-\rangle = \sqrt{1/2} (|0\rangle_1|0\rangle_2 - |1\rangle_1|1\rangle_2)$$

$$|\Phi^+\rangle = \sqrt{1/2} (|1\rangle_1|0\rangle_2 + |0\rangle_1|1\rangle_2)$$

$$|\Phi^-\rangle = \sqrt{1/2} (-|1\rangle_1|0\rangle_2 + |0\rangle_1|1\rangle_2)$$

*The "Bell basis" for $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$.
Each basis vector is an entangled state!*

- Let $|0\rangle_1|0\rangle_2$, $|0\rangle_1|1\rangle_2$, $|1\rangle_1|0\rangle_2$, $|1\rangle_1|1\rangle_2$ be given the matrix representations:

$$|0\rangle_1|0\rangle_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad |0\rangle_1|1\rangle_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad |1\rangle_1|0\rangle_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad |1\rangle_1|1\rangle_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

- Define the 2-qubit "Controlled-NOT" operator by:

$$C_{NOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \begin{aligned} C_{NOT}|0\rangle_1|0\rangle_2 &= |0\rangle_1|0\rangle_2 \\ C_{NOT}|0\rangle_1|1\rangle_2 &= |0\rangle_1|1\rangle_2 \\ C_{NOT}|1\rangle_1|0\rangle_2 &= |1\rangle_1|1\rangle_2 \\ C_{NOT}|1\rangle_1|1\rangle_2 &= |1\rangle_1|0\rangle_2 \end{aligned}$$

Acts on two qubits:

- *Changes the second if the first is $|1\rangle$.*
- *Leaves the second unchanged otherwise.*

4. The No-Cloning Theorem

Claim: Unknown qubits cannot be "cloned".

- In particular, there is no (unitary, linear) operator U such that

$$U|v\rangle_1|0\rangle_2 = |v\rangle_1|v\rangle_2, \text{ where } |v\rangle_1 \text{ is an unknown qubit.}$$

Proof: Suppose there is such a U .

- Then: $U|a\rangle_1|0\rangle_2 = |a\rangle_1|a\rangle_2$ and $U|b\rangle_1|0\rangle_2 = |b\rangle_1|b\rangle_2$, for unknown qubits $|a\rangle_1, |b\rangle_1$.

- Let: $|c\rangle_1 = \alpha|a\rangle_1 + \beta|b\rangle_1$, where $|\alpha|^2 + |\beta|^2 = 1$

- Then: $U|c\rangle_1|0\rangle_2 = U(\alpha|a\rangle_1|0\rangle_2 + \beta|b\rangle_1|0\rangle_2)$

$$= (\alpha U|a\rangle_1|0\rangle_2 + \beta U|b\rangle_1|0\rangle_2), \quad \text{since } U \text{ is linear}$$

$$= \alpha|a\rangle_1|a\rangle_2 + \beta|b\rangle_1|b\rangle_2$$

- But: By definition, U acts on $|c\rangle_1$ according to:

$$U|c\rangle_1|0\rangle_2 = |c\rangle_1|c\rangle_2$$

$$= (\alpha|a\rangle_1 + \beta|b\rangle_1)(\alpha|a\rangle_2 + \beta|b\rangle_2)$$

$$= \alpha^2|a\rangle_1|a\rangle_2 + \alpha\beta|a\rangle_1|b\rangle_2 + \beta\alpha|b\rangle_1|a\rangle_2 + \beta^2|b\rangle_1|b\rangle_2.$$

- So: There can be no such U .

- Note: Known qubits (like $|1\rangle_1$) can be cloned (ex: $C_{NOT}|1\rangle_1|0\rangle_2 = |1\rangle_1|1\rangle_2$).

5. Quantum Computation

- Can classical computations be done using qubits instead of classical bits?
 - *Can transformations on qubits be defined that reproduce the transformations on bits that are needed to implement a classical computer.*
- To implement a classical computer, it suffices to have an *AND* transformation and a *NOT* transformation on classical bits defined by the following:

$$0 \text{ AND } 0 = 0$$

$$\text{NOT } 0 = 1$$

$$0 \text{ AND } 1 = 0$$

$$\text{NOT } 1 = 0$$

$$1 \text{ AND } 0 = 0$$

$$1 \text{ AND } 1 = 1$$

- *AND takes two input bits and produces one output bit.*

- *NOT takes one input bit and produces one output bit.*

- Initial problem: Transformations on qubits are *reversible*: the number of input qubits *always* must equal the number of output qubits.

Why? Qubit transformations are operators on vector spaces. And an operator defined on an n -dim vector space (e.g., n -qubit space) that acts on n -dim vectors (e.g., n qubits) can only spit out n -dim vectors.

Solution: The Controlled-controlled-*NOT*, CC_{NOT} , operator.

- Changes the third target qubit if the first two qubits are $|1\rangle|1\rangle$, and leaves it unchanged otherwise.

$$\begin{aligned}
 CC_{NOT}|0\rangle|0\rangle|0\rangle &= |0\rangle|0\rangle|0\rangle & CC_{NOT}|0\rangle|1\rangle|1\rangle &= |0\rangle|1\rangle|1\rangle & CC_{NOT}|1\rangle|1\rangle|0\rangle &= |1\rangle|1\rangle|1\rangle \\
 CC_{NOT}|0\rangle|0\rangle|1\rangle &= |0\rangle|0\rangle|1\rangle & CC_{NOT}|1\rangle|0\rangle|0\rangle &= |1\rangle|0\rangle|0\rangle & CC_{NOT}|1\rangle|1\rangle|1\rangle &= |1\rangle|1\rangle|0\rangle \\
 CC_{NOT}|0\rangle|1\rangle|0\rangle &= |0\rangle|1\rangle|0\rangle & CC_{NOT}|1\rangle|0\rangle|1\rangle &= |1\rangle|0\rangle|1\rangle & &
 \end{aligned}$$

$$CC_{NOT} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad |0\rangle|0\rangle|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |0\rangle|0\rangle|1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad |1\rangle|1\rangle|1\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

- Claim: CC_{NOT} implements *AND* and *NOT* on qubits.

- To implement *AND*, act with CC_{NOT} on a 3-qubit state in which the last qubit is $|0\rangle$:

$$CC_{NOT}|x\rangle|y\rangle|0\rangle = |x\rangle|y\rangle|x \text{ AND } y\rangle$$

- To implement *NOT*, act with CC_{NOT} on a 3-qubit state in which the first two qubits are $|1\rangle|1\rangle$:

$$CC_{NOT}|1\rangle|1\rangle|x\rangle = |1\rangle|1\rangle|\text{NOT } x\rangle$$

So: Any classical computation can be done using qubits instead of bits.

- In particular: Any classical function that takes n input bits and produces k output bits can be implemented using arrays of primitive CC_{NOT} "gates".

Is quantum computation any different from classical computation?

Claim: No *fundamental* difference between quantum and classical computation: just a difference in *hardware*.

- A quantum computer can compute anything that a classical computer can.
- A classical computer can compute anything that a quantum computer can.
 - *Any computation implemented using qubits can be implemented using bits and a probabilistic algorithm.*
 - Intuitively: *There are probabilistic classical 2-state systems that can simulate the output of quantum 2-state systems, (although perhaps not as efficiently).*

But: *Quantum computation is potentially much more efficient than classical computation...*

Ex. A Qubit-Based Function Calculator

- Let $|x\rangle_{(n)}$ represent n input qubits that encode the number x .
 - Example: $|1\rangle|1\rangle|0\rangle$ represents 6.
- Let $|0\rangle_{(k)}$ represent k $|0\rangle$ qubits (the output register).
- Let $|f(x)\rangle_{(k)}$ represent k output qubits that encode the number $f(x)$.
- Define an operator U_f that acts on $(n+k)$ qubits in the following way:

$$U_f|x\rangle_{(n)}|0\rangle_{(k)} = |x\rangle_{(n)}|f(x)\rangle_{(k)}$$

- Now: Feed U_f a *superposition* of all possible numbers x it can take as input.
- Result: A superposition of all possible values of the function in a *single* computation!

Two Steps:

1. Prepare a superposition of all possible numbers x that can be encoded in n bits:

(i) Start with an n -qubit state $|0\rangle_1|0\rangle_2\cdots|0\rangle_n$

(ii) Now apply a Hadamard transformation to each qubit:

$$(\mathfrak{H}_1 \otimes \mathfrak{H}_2 \otimes \cdots \otimes \mathfrak{H}_n) |0\rangle_1 |0\rangle_2 \cdots |0\rangle_n$$

$$= (\sqrt{1/2})^n \{(|0\rangle_1 + |1\rangle_1)(|0\rangle_2 + |1\rangle_2) \cdots (|0\rangle_n + |1\rangle_n)\}$$

$$= (\sqrt{1/2})^n \{ \underbrace{|0\rangle_1 |0\rangle_2 \cdots |0\rangle_n}_{\text{The first term encodes the binary number for 0, or } |0\rangle_{(n)}} + |0\rangle_1 |0\rangle_2 \cdots |1\rangle_n + \cdots + \underbrace{|1\rangle_1 |1\rangle_2 \cdots |1\rangle_n}_{\text{The last term encodes the binary number for } 2^n-1, \text{ or } |2^n-1\rangle_{(n)}} \}$$

The first term encodes the binary number for 0, or $|0\rangle_{(n)}$

Each term in between is the binary number for each number between 0 and 2^n-1 .

The last term encodes the binary number for 2^n-1 , or $|2^n-1\rangle_{(n)}$

$$= (\sqrt{1/2})^n \sum_{x=0}^{2^n-1} |x\rangle_{(n)}$$

So the entire sum is a superposition that encodes all numbers x such that $0 \leq x < 2^n$.

Two Steps:

2. Now attach a k -qubit output register $|0\rangle_{(k)}$ and apply U_f :

$$\begin{aligned} U_f(\sqrt{1/2})^n \sum_{x=0}^{2^n-1} |x\rangle_{(n)} |0\rangle_{(k)} &= (\sqrt{1/2})^n \sum_{x=0}^{2^n-1} U_f |x\rangle_{(n)} |0\rangle_{(k)} \\ &= (\sqrt{1/2})^n \sum_{x=0}^{2^n-1} |x\rangle_{(n)} |f(x)\rangle_{(k)} \end{aligned}$$

A superposition of all possible values $f(x)$, for $0 \leq x < 2^n$, of the function f . And we've effectively calculated them all with just a single application of U_f .

The Catch: None of these values of f is accessible until we make a measurement!

The Task for Quantum Algorithm construction

Given a problem, first construct an appropriate superposition of solutions; and then manipulate the superposition so that the relevant terms acquire high probability.