(46) Now $\alpha^{(t+1)} = 2\mu^{(t)} + \alpha^{(t)} \geq \frac{1}{\sqrt{N}} + \alpha^{(t)}$

Of Classical computation

Computation is a physical process finite in time and fixed set of states. (which are distinguishable)

~~Inter~~ Information is an interpretation of a particular system's state.

Shannon's information theory: $I = -\sum_{i=1}^{n} P_i \log_b (P_i)$

amount of information stored in n states, P_i = probability of state i. b is some base. This formula defines some kind of entropy of the system. For computational purposes, $P_i = \frac{1}{n}$,

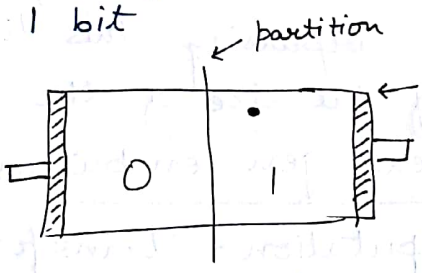
$$I = -\sum_{i=1}^{n} \frac{1}{n} \cdot \log_b\left(\frac{1}{n}\right) \Rightarrow I = \sum_{i=1}^{n} \frac{1}{n} \log_b (n)$$

$$\Rightarrow I = n \cdot \frac{1}{n} \log_b (n) = \log_b (n) \quad \text{where } b \geq 2$$

Bit — smallest amount of information that can be stored in a physical system with 1 state.

bit $= \log_2 (2) = 1$ bit

Szilard's engine -



Leo Szilard 1929

There is a particle of gas. If it is to the left of the partition, the system has a value 0 and elsewise for right. This is not a computational process as the states are not distinguishable.

Assign value 0: Remove partition, move right piston until it reaches the mid, and replace partition. Work done $= k_B T \ln \frac{V_i}{V_f} = k_B T \ln 2$

or such energy is dissipated to assign a value to the system.

(47) NOT gate - maps state 0 to state 1 or the particle is confined to the right side of the engine. Information in some sense equals energy in this case.

Why computers are not source engines as above? Certain characteristics of computational system -

① Information capacity (Shannon formula)
② Speed of switching between states
③ Universality (variety of tasks solved).

Analog system drawbacks - ① Error correction ② Universality. Digitizing this removes these errors.

Computer generations - ① Vacuum tubes ② Transis ③ Integrated circuits ④ Microprocessors.

Smaller elements have less inertia (inductance), so switching process is faster, consume less energy, and allow packing more elements into something smaller. The smallest element of implementing this is the p-n junction (200 times larger than the hydrogen atom). For surpassing this limit, we need something of the size of the atom or even smalle Thus the next gen computers have to be quantu

---

Goal of computation - transforming information to - find value of some function, and since functions have exact definitions, the goal becomes clearer. Thus, algorithms may be defined. Algorithms are deterministic Turing machines (DTMs).

- Computability - a function is computable wher there exists a DTM for an algorithm. And ther are an infinite countable number of DTMs out there.

How many functions of $f: \mathbb{N} \to \{0,1\}$ are there?

(48) Defining the function f : $0, f(1)*f(2)*f(3)...$

Hence there are continuum number of functions. Thus if there are ~~con~~ countable algorithms and continuum functions ⇒ some functions are uncomputable. ex Halting problem.

Church-Turing thesis – A DTM can analyse the behaviour of another DTM. So there can be written a computer program such that it can't be analysed by classical computation.

Complexity classes – P = polynomial time; NP = Non-deterministic polynomial. (it is hard to solve the problem but if solved, it is easy to check it). Notions about NP-Hard and NP-Complete (Intersection of NP-Hard and NP).

Quantum computing – The next gen is quantum:

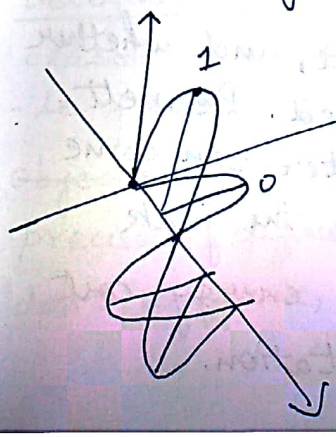
① Base element is going to be quantum

② A classical system with several quantum units is hard to model.

Quantum effects – probabilitic notions, superposition, interference, wave function.

Wave function is intact for a closed system. Measurement makes us open to the system, and the system collapses again. Observation alters the wave function of the observer, and not that of the system.

Multiverse Interpretation of quantum mechanics – answers why observers don't see each other.



The electromagnetic wave propagation with oscillations orthogonal to the plane of propagation.

Hilbert space is a field with some dot or inner product.

(49) $|0\rangle\{|s_1\rangle + |s_2\rangle\} = |0\rangle|s_1\rangle + |0\rangle|s_2\rangle$ interpreted as two copies of observer observing different states of the original system. Multiverse point of view, involving entanglement.

Of Preskill's notes

Caltech

Quantum computation and quantum information from the perspective of a theoretical physicist

Physics of information and computation:

① Landauer's principle – same as Szilard's engine erasure procedure. Erasure of information is a dissipative process. In Szilard's engine, some work needs to be done to erase the already existing information encoded in the physical state of the system. $W = kT \ln 2$

② Reversible computation – Consider an irreversible logic operation $(a, b) \to \neg(a \wedge b)$ (NAND gate), since about one bit is erased by the bit, according to Landauer's Principle, at least $W = kT \ln 2$ is needed. Assuming finite supply of batteries, this operation has a finite time. Erasing information requires power

Charles Bennett in 1973

Computation using reversible steps, in principle, requires no dissipation and no power dissipation.

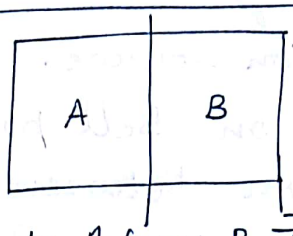
Toffoli gate: $(a, b, c) \to (a, b, c \oplus a \wedge b)$

if $c = 1$, NAND operation

A question of extra junk comes up here, and whether the energy cost has just been delayed. Bennett showed that the reversible computer can come back to initial configuration, removing the junk without any extra cost. There is no energy cost associated with reversible computation.

50) — The consideration then is that modern computers dissipate more than $kT \ln 2$ energy per gate. As size of components shrink, it is important to beat this limit else components melt. Then reversible computation shall be the only option.

Maxwell's demon —



Maxwell's box with a gas, and a demon at the partition.

The demon allows fast molecules from A to B and cool ones to A from B. This implies A gets hotter and B gets cooler, or transfer of heat without expenditure of energy, violating second law of thermodynamics.

Resolution — Demon must collect and store information about the molecules. If the demon has a finite memory capacity, the gas can't be made to cool indefinitely; at some time information must be erased, and thus we pay the energy cost to cool the gas. For thermodynamic accounting before the erasure, there must be some entropy associated with the recorded information.

Leo Szilard in 1929, in his analysis of Maxwell's demon, inverted the concept of bit of information and associated entropy $\Delta S = k \ln 2$ (isothermal) with acquisition of 1 bit.

MORAL — information is physical

Quantum information — Classical ideas about information needed revision once QFT came out. There is no place for randomness in deterministic classical dynamics. Only a chaotic classical system can exhibit behaviour indistinguishable from true randomness. Moreover, noncommuting ~~objects~~ observables can't simultaneously have precisely defined values (Hisenberg). So if A and B don't commute, measuring one disturbs the other.

(51) Quantum information can't be copied with perfect fidelity (No cloning ~~theorem~~ principle by Wootters and Zurek and Dieks in 1982). Copying quantum information would disturb the original information source.

John Bell's work on Bell pairs is the most important distinction between quantum and classical information.

Efficient quantum algorithms –

The factoring problem is intractable (hard to solve but easy to verify). $n = pq$

Given p and q, it requires $\log_2 p \cdot \log_2 q$ operations to verify n. But given n, the problem is Superpolynomial in $\log(n)$. The state of the art (number field sieve) requires:

$$\text{time} \simeq \exp\left[c(\ln n)^{1/3}(\ln \ln n)^{2/3}\right]$$

$$c = (64/9)^{1/3} \approx 1.9$$

Shor's algorithm – $O\left[(\ln n)^3\right]$

Quantum complexity –

Idea that a quantum system could perform computation came from Paul Benioff and Ric. Feynman in 1982.

Benioff: As circuitry becomes smaller, quantum effects jump in

Feynman's: mathematical description of quantum information and computation:–

considering a 2D Hilbert space (a 2D complex vector space with inner product), unit of quantum information is the qubit, with orthonormal basis $|0\rangle$ and $|1\rangle$.

Ensemble of N qubits = vector in space of dimension 2^N. (1 qubit needs 2D space to be represent)

Orthonormal basis of this space = states in which each qubit is a $|0\rangle$ or a $|1\rangle$, labelled by binary strings $|01100\ldots1001\rangle$ or numbers like $|0\rangle, |1\rangle, |2\rangle, |3\rangle, \ldots, |2^N-1\rangle$. (No. They represent possible outputs refer Page 54).

General normalized vector in this space. $= \sum_{x=0}^{2^N-1} a_x |x\rangle$

a_x are complex numbers such that $\sum_{x=0}^{2^N-1} |a_x|^2 = 1$ (probabilities sum = 1).

Measurement by projection ⇒ probability of getting $|x\rangle = |a_x|^2$

Computation – Assemble N qubits, perform a unitary transformation, and measure all the qubits onto the $\{|0\rangle, |1\rangle\}$ basis. Measurement outcome is classical, with some probabilistic measure.

Feynman had this idea that since a classical computer will not be able to deal with huge matrices, hence a quantum computer could do things impossible for a classical one.

– We could settle for a classical probabilistic algorithm in which the outcome is not uniquely determined by the input, but the outcomes have a probability distribution attached to outputs similar to that of a quantum computer. But Bell's theorem suggests such an algorithm is not possible.

If a quantum system of 3N qubits is prepared, and they are kept in 3 separate systems in 3N qubits each, it is impossible to determine the state of a system by measurements in one subsystem only. All the information that distinguishes one state from other is in the nonlocal correlations between

(53) measurements in subsystems 1, 2, and 3.

information quantification using entropy:

$$S \cong N - 2^{-(N+1)}$$

as N gets large, entropy increases, or we can access lesser information by looking at each subsystem separately.

- Measurements give very little information until the correlations of measurements of other subsystems are studied (collective measurement). Nonlocal correlations = entangled.
- This is why classical simulation of quantum systems requires vast resources. When different parts are entangled, they can't be separated and studied.

Note: identical copies of a quantum state can be prepared by a friend, who doesn't tell us what she/he did.

Quantum parallelism: Deutsch

Consider a black box that takes input x and outputs f(x). There are four possible combinations: $f(0) = 0$, $f(0) = 1$, $f(1) = 0$, $f(1) = 1$

One computation takes 24 hours. It takes 48 hours to determine whether the function is constant $[f(0) = f(1)]$ or balanced $[f(0) \neq f(1)]$

Quantum equivalent: The operation must be invertible and Unitary. Considering a unitary transformation such that:

$$U_f : |x\rangle |y\rangle \longrightarrow |x\rangle |y \oplus f(x)\rangle$$

Deutsch problem - Can we get the answer in 24 hours?

(54) Superposition or quantum parallelism –

$|y\rangle$ is prepared in state $\frac{1}{\sqrt{2}}\{|0\rangle - |1\rangle\}$ or Hadamard to $|1\rangle$.

$$U_f : |x\rangle \cdot \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \longrightarrow |x\rangle \frac{1}{\sqrt{2}}\left(|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle\right)$$

$$\left.\begin{aligned} |0 \oplus f(x)\rangle &= (-1)^{f(x)}|0\rangle \\ |1 \oplus f(x)\rangle &= (-1)^{f(x)}|1\rangle \end{aligned}\right\} \text{ sign change depending on value of } f(x).$$

$$U_f : |x\rangle \cdot \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \rightarrow |x\rangle \cdot \frac{1}{\sqrt{2}} \cdot (-1)^{f(x)}(|0\rangle - |1\rangle)$$

Preparing the first qubit as $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

$$U_f : \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \cdot \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \longrightarrow$$

$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \cdot \frac{1}{\sqrt{2}} \cdot (-1)^{f(x)}(|0\rangle - |1\rangle)$$

($x = 0$ and $x = 1$) → since value of x is determined in first qubit

$$\Rightarrow \frac{1}{\sqrt{2}}\left((-1)^{f(0)} \cdot |0\rangle + (-1)^{f(1)} \cdot |1\rangle\right) \cdot \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$$

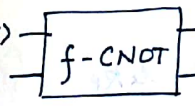
Measurement projecting the first qubit onto basis:

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$$

$-$ ≡ Balanced or $f(0) \neq f(1)$ or $f(0) = \overline{f(1)}$

$+$ ≡ constant or $f(0) = f(1)$

Notes:



Detailed working:

$$\frac{1}{\sqrt{2}}(|0\rangle_1^* + |1\rangle_1^*) \cdot \frac{1}{\sqrt{2}}(|0_2\rangle - |1_2\rangle)$$

$$\frac{1}{2}(|0_1 0_2\rangle - |0_1 1_2\rangle + |1_1 0_2\rangle - |1_1 1_2\rangle)$$

$$= \frac{1}{2}(|0_1 f(0_1)\rangle - |0_1 \overline{f(0_1)}\rangle + |1_1 f(1_1)\rangle - |1_1 \overline{f(1_1)}\rangle)$$

where in a given ket, $|q_1 q_2\rangle$ such that $q_1 \in$ first qubit and $q_2 \in$ second qubit. This expression is just the entanglement of 4 possibilities of 2 qubit system

$|q_1 q_2\rangle \rightarrow |q_1 \cdot (q_2 \oplus f(q_1))\rangle$

such that $|0_1 0_2\rangle = |0_1 (0_2 \oplus f(0_1))\rangle$ and $0 \oplus f(0) = f(0)$ while $1 \oplus f(0) = \overline{f(0)}$

(55) Final function in superposition =

$$\frac{1}{2}\left(|0,f(0)\rangle - |0,\overline{f(0)}\rangle + |1,f(1)\rangle - |1,\overline{f(1)}\rangle\right)$$

Case I: $f(x)$ is constant: $f(0) = f(1)$

$$\frac{1}{2}\left(|0\rangle + |1\rangle\right)\left(-|\overline{f(0)}\rangle + |f(0)\rangle\right)$$

= A Hadamard on first qubit $\Rightarrow |0\rangle$

Case 2: similarly: if $f(0) = \overline{f(1)}$

$$\frac{1}{2}\left(|0\rangle - |1\rangle\right)\left(|f(0)\rangle - |\overline{f(0)}\rangle\right)$$

Hadamard on first qubit $\Rightarrow |1\rangle$

All this is possible because the quantum computer can act on a superposition of $|0\rangle$ and $|1\rangle$. or extracting global information

On global information –

In general, a function of N bits, having 2^N diff possibilities needs 2^N calculations, infeasible for large N. For quantum computer,

$$U_f : |x\rangle|0\rangle \longrightarrow |x\rangle|f(x)\rangle$$

Input: $\left[\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\right]^N = \frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle$

where $|x\rangle$ denotes the possible states like

$$\frac{1}{2^{2/2}}(|0_1\rangle + |1_1\rangle)(|0_2\rangle + |1_2\rangle) = \frac{1}{2}\left(|0_1 0_2\rangle + |0_1 1_2\rangle + |1_1 0_2\rangle + |1_1 1_2\rangle\right)$$

in a 2-qubit system.

By computing $f(x)$ only once, calculating,

$$= \frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle|f(x)\rangle$$

Global properties of $f(x)$ are encoded in this system. Massive quantum parallelism.

Quantum information can be encoded in nonlocal correlations (or entanglements) among different parts of a physical system. ex in $\frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle |f(x)\rangle$ information is stored as correlation between input and output.

- if input register is measured, $|x_0\rangle$ maybe an outcome such that $|f(x_0)\rangle$ is created in the output register. Now all the correlations are lost, i.e. equation $\frac{1}{2^{N/2}} \sum_{x=0}^{2^N-1} |x\rangle |f(x)\rangle$ is lost and there is no opportunity to measure $f(y_0)$ for any $y_0 \neq x_0$. In this case, there is no advantage over a classical one.

Complexity - According to the classical complexity theory, if someone provides unlimited amount of memory and waits unlimited amount of time, anything that can be called computation can be run on a classical computer.

A general notion of complexity comes from making the complexity theory machine-independent, or defining the exponential and polynomial times such that, if $T(A)$ is the function denoting max. time taken as a function of input length, then:

$T(A) \leq$ polynomial time for solvable problems

poly $\leq T(A) \leq$ expo. for harder problems.

↑ in classical computation.

- Thus, one universal (classical) computer can simulate another with at worst "polynomial" overhead".

If it is true that no simulation of a quantum computer with polynomial overhead is possible, the result could shake the foundations of computer science. It is not yet proved that classical Turing machines is not an appropriate model of computations in the physical world.

(57) Burning question – If the quantum classification of complexity different from the classical classification? Yes – suspected, not proved

Quantum error correction – there is information hidden in the nonlocal correlations between subsystems.

The quantum system is NOT isolated, it is in contact with the environment. Interactions between the system and environment sets up correlations between them, re-encoding information. No longer is the information accessible by only observing the system.

Even to Schrödinger, the proponents seemed blemish:

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}\left(|\text{dead}\rangle + |\text{alive}\rangle\right)$$

– Such a state is possible, but extremely unstable. If someone prepares the state $|\text{cat}\rangle$, the information stored in correlations is immediately lost as new correlations are established with the environment, so information in $|\text{cat}\rangle$ becomes completely inaccessible.

The environment continually measures the state of cat, projecting either to $|\text{alive}\rangle$ or to $|\text{dead}\rangle$. This is called decoherence (loss of information encoded in correlations within an isolated system when it begins to interact with the environment.

There has to be a way to prevent this degradation of quantum information, handled by quantum error correction.

(58) Problem 2: unitary transformations are a continuum (consider a Bloch sphere and a vector rotating on it).

The application of such a transformation might not be flawless, causing an error: $U = U_0(1 + O(\varepsilon))$. After $1/\varepsilon$ applications, the error is serious.

– Modern digital circuits have similar problems but achieve accuracy by energy dissipation. After each operation, the bit is cooled, or heat from the logic gate operation is allowed to dissipate to the environment, and performance is not compromised.

– Dissipation in quantum computers is not possible, as that establishes a connection with the environment. We can throw away information about the errors, but that discard of information will still be a dissipative process.