Information Theory Meets Quantum Physics

The magic of wave dynamics

Apoorva Patel

Centre for High Energy Physics Indian Institute of Science, Bangalore

30 April 2016 (Commemorating the 100th birthday of Claude Shannon) IISc-IEEE Advanced Information Theory Workshop, Bangalore



Any physical process evolves an initial state to a final state under the influence of certain interactions.

Initial state — Interactions — Final state



Any physical process evolves an initial state to a final state under the influence of certain interactions.

Initial state — Interactions — Final state

A computer processes the given input to an output according to specified instructions.

Input — Processing — Output



Any physical process evolves an initial state to a final state under the influence of certain interactions.

Initial state — Interactions — Final state

A computer processes the given input to an output according to specified instructions.

Input — Processing — Output

It purposefully processes meaningful information. The meaning arises from mapping physical properties (hardware) to mathematical terminology (software).



Any physical process evolves an initial state to a final state under the influence of certain interactions.

Initial state — Interactions — Final state

A computer processes the given input to an output according to specified instructions.

Input — Processing — Output

It purposefully processes meaningful information. The meaning arises from mapping physical properties (hardware) to mathematical terminology (software).

The availability of different physical interactions makes it possible to design different types of computers.

Communication is the special case where processing is limited to coding and decoding.



Designing a Computer

Computation is Processing and Communication of Semantic Information expressed using a Language.

One finds examples of many types of information processing systems in the physical world.

How can a general information theory be systematically developed to cover all types of computational schemes?



Designing a Computer

Computation is Processing and Communication of Semantic Information expressed using a Language.

One finds examples of many types of information processing systems in the physical world.

How can a general information theory be systematically developed to cover all types of computational schemes?

How can one design the optimal computer for a given task?

Optimisation requires minimisation of physical resources (versatile composition) and control of errors (digitisation).

These criteria often conflict, and trade-offs are necessary.



Designing a Computer

Computation is Processing and Communication of Semantic Information expressed using a Language.

One finds examples of many types of information processing systems in the physical world.

How can a general information theory be systematically developed to cover all types of computational schemes?

How can one design the optimal computer for a given task?

Optimisation requires minimisation of physical resources (versatile composition) and control of errors (digitisation).

These criteria often conflict, and trade-offs are necessary.

Efficiency of information processing depends on both the available hardware and the possible software.

Software: What is the task? What is the algorithm? Hardware: How are the operations implemented?



Technical Terms

Data: They describe a particular realisation of the physical system, amongst its many possible states.

Information: It is the abstract mathematical property obtained by detaching all the physical characteristics from data.

Knowledge: It is obtained by adding a sense of purpose to the abstract information.



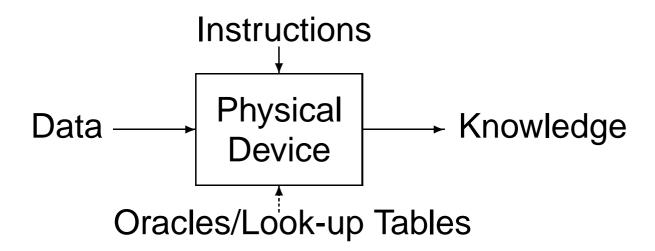
Technical Terms

Data: They describe a particular realisation of the physical system, amongst its many possible states.

Information: It is the abstract mathematical property obtained by detaching all the physical characteristics from data.

Knowledge: It is obtained by adding a sense of purpose to the abstract information.

Information = Data - Physical Realisation Knowledge = Information + Interpretation





Importance of Physics

- 1. Abstract information can be manipulated with precise mathematical rules, without going into nitty-gritty of its origin or meaning.
- 2. The manipulations can only be implemented using physical devices.
- 3. The interpretation of a language has to be established through physical properties.

Physical hardware properties govern the expression and the grammar of the language.



Importance of Physics

- 1. Abstract information can be manipulated with precise mathematical rules, without going into nitty-gritty of its origin or meaning.
- 2. The manipulations can only be implemented using physical devices.
- 3. The interpretation of a language has to be established through physical properties.

Physical hardware properties govern the expression and the grammar of the language.

Abstract information theory does not tell us what physical realisation would be appropriate for a particular message, nor does it tell us the best way of implementing a computational task.

These choices have to be made by analysing the type (and not the amount) of information, and checking how that can be mapped to the available physical resources.

An Example from Biology

A plant attracts an insect to its flower.

Purpose: Pollination for plant, food for insect.

Task: Convey existance, direction and distance.



An Example from Biology

A plant attracts an insect to its flower.

Purpose: Pollination for plant, food for insect.

Task: Convey existance, direction and distance.

Existence: Expressed by the three-dimensional structure of fragrant molecules, using lock-and-key mechanism.

Direction: Plant releases millions of fragrant molecules.

Insect detects gradients (parallax) with multiple receptors.

Distance: Infered from concentration of fragrant molecules.

Language: Arrived at by millions of years of coevolution, and stored in the genomes.

Resources: Massively redundant communication process optimised by going down all the way to atomic scale.



An Example from Biology

A plant attracts an insect to its flower.

Purpose: Pollination for plant, food for insect.

Task: Convey existance, direction and distance.

Existence: Expressed by the three-dimensional structure of fragrant molecules, using lock-and-key mechanism.

Direction: Plant releases millions of fragrant molecules.

Insect detects gradients (parallax) with multiple receptors.

Distance: Infered from concentration of fragrant molecules.

Language: Arrived at by millions of years of coevolution, and stored in the genomes.

Resources: Massively redundant communication process optimised by going down all the way to atomic scale.

Mobile phones convey existence (followed by two-way communication), but direction and distance are ignored.



Minimal Language

The language with the smallest set of building blocks (for a given task) is unique in the optimisation procedure.

- Largest tolerance against errors.
 (Discrete variables are spread as far apart as possible in the available range of physical hardware properties.)
- Smallest instruction set.
 (Number of possible transformations is limited.)
- High density of packing and quick operations.
 (These more than make up for the increased depth of computation.)
- Simplest language, without need of translation. (Simple physical responses of the hardware can be used.)



Minimal Language

The language with the smallest set of building blocks (for a given task) is unique in the optimisation procedure.

- Largest tolerance against errors.
 (Discrete variables are spread as far apart as possible in the available range of physical hardware properties.)
- Smallest instruction set.
 (Number of possible transformations is limited.)
- High density of packing and quick operations.
 (These more than make up for the increased depth of computation.)
- Simplest language, without need of translation. (Simple physical responses of the hardware can be used.)
- Boolean algebra provides the minimal classical language for encoding information as 1-dimensional sequences.



General Computational Framework

Generalise the concept of a language, from a "sequence of letters" to a "collection of building blocks".



General Computational Framework

Generalise the concept of a language, from a "sequence of letters" to a "collection of building blocks".

Collections: The building blocks can be arranged in the space-time in many different ways.

Building blocks: Physical properties (generically encoded using groups) express the meaning of the building blocks. Processing: Allowed changes in the properties of building blocks exhaust the possible manipulations of information.

Group representations fix the structure of the language, and group transformations provide the rules for processing information.

...contd.



For a given task, the smallest discrete group that can implement it is the ideal candidate for the optimal language.

In digitisation of continuous spaces (of physical properties), a discrete state is associated not with just a point on the manifold but with a finite neighbourhood of a point.



For a given task, the smallest discrete group that can implement it is the ideal candidate for the optimal language.

In digitisation of continuous spaces (of physical properties), a discrete state is associated not with just a point on the manifold but with a finite neighbourhood of a point.

For a d-dim space, the simplest building block is a simplex, i.e. a set of (d+1) points.

The dimension of a group is the number of its generators. In the dual description, the minimal building block set is the d-dim fundamental representation and the 1-dim identity representation.



For a given task, the smallest discrete group that can implement it is the ideal candidate for the optimal language.

In digitisation of continuous spaces (of physical properties), a discrete state is associated not with just a point on the manifold but with a finite neighbourhood of a point.

For a d-dim space, the simplest building block is a simplex, i.e. a set of (d+1) points.

The dimension of a group is the number of its generators. In the dual description, the minimal building block set is the d-dim fundamental representation and the 1-dim identity representation.

In the smallest discrete realisation, the entire group is replaced by a single simplex. A collection of simplices can then represent any quantity to the desired precision (e.g. the place value system for numbers).



Types of Collections

0 – dim: Multiple signals at the same point in space and time, i.e. superposition. Different states of internal degrees of freedom encode different signals. Only one signal can be extracted at a time.

Examples: Radio broadcasts, Mobile phones.



Types of Collections

0 – dim: Multiple signals at the same point in space and time, i.e. superposition. Different states of internal degrees of freedom encode different signals. Only one signal can be extracted at a time.

Examples: Radio broadcasts, Mobile phones.

1 - dim: Building blocks arranged as an ordered sequence. This tensor product structure is routinely used in conventional information theory.

Examples: Written languages (sequence in space), Spoken languages (sequence in time).



Types of Collections

0 – dim: Multiple signals at the same point in space and time, i.e. superposition. Different states of internal degrees of freedom encode different signals. Only one signal can be extracted at a time.

Examples: Radio broadcasts, Mobile phones.

1 - dim: Building blocks arranged as an ordered sequence. This tensor product structure is routinely used in conventional information theory.

Examples: Written languages (sequence in space), Spoken languages (sequence in time).

2 - dim: Combination of multiple ordered sequences. Information can reside in correlations amongst sequences without being present in any individual sequence.

Examples: Eyes and ears, Space-time codes, Parallax and gradient detection.



3 – *dim*: Building blocks encode structural information, for establishing lock and key mechanisms. Examples: Proteins and other biomolecules.



3 – *dim*: Building blocks encode structural information, for establishing lock and key mechanisms. Examples: Proteins and other biomolecules.

4 - dim: Complete description of any space-time event in our universe.

Example: Global positioning system (GPS).



3 – *dim*: Building blocks encode structural information, for establishing lock and key mechanisms.

Examples: Proteins and other biomolecules.

4 - dim: Complete description of any space-time event in our universe.

Example: Global positioning system (GPS).

Simultaneous use of multiple dimensionalities is also possible, with building blocks having multiple properties! Examples: Quantum computers (0 and 1 dim), Proteins (1 and 3 dim).



3 – dim: Building blocks encode structural information, for establishing lock and key mechanisms.

Examples: Proteins and other biomolecules.

4 - dim: Complete description of any space-time event in our universe.

Example: Global positioning system (GPS).

Simultaneous use of multiple dimensionalities is also possible, with building blocks having multiple properties! Examples: Quantum computers (0 and 1 dim), Proteins (1 and 3 dim).

Particle features: Countability, Order, Number density, Shape, Structure.

Wave features: Superposition, Differential analysis (parallax), Interference (multiple paths).



Types of Building Blocks

A combination of properties, in both internal and external space, can be used to describe distinguishable objects.



Types of Building Blocks

A combination of properties, in both internal and external space, can be used to describe distinguishable objects.

1-dim: Groups with a single generator include cyclic groups, integers and the real line. The minimal simplex is $Z_2=\{0,1\}$. The binary language can be extended to a d-dim situation, as the Cartesian product $(Z_2)^d$, and is therefore convenient as a general purpose language.



Types of Building Blocks

A combination of properties, in both internal and external space, can be used to describe distinguishable objects.

1-dim: Groups with a single generator include cyclic groups, integers and the real line. The minimal simplex is $Z_2 = \{0,1\}$. The binary language can be extended to a d-dim situation, as the Cartesian product $(Z_2)^d$, and is therefore convenient as a general purpose language.

2 – *dim*: The simplex is a triangle. Triangulation (or its dual hexagonal form) is useful for discrete description of arbitrary surfaces. Graphene structure may become useful in atomic scale lithography.



Types of Building Blocks (contd.)

3 - dim: The simplex is a tetrahedron. At molecular scale, sp^3 -hybridised orbitals provide its dual form. Tetrahedral geometry of carbon provides a convenient way to understand the 3-dim language of proteins.



Types of Building Blocks (contd.)

3-dim: The simplex is a tetrahedron. At molecular scale, sp^3 -hybridised orbitals provide its dual form. Tetrahedral geometry of carbon provides a convenient way to understand the 3-dim language of proteins.

SU(2): Description of quantum bits is based on this group with three generators. Arbitrary states of a qubit (including mixed states) can be fully described using a density matrix, which is a linear combination of the four operators $\{1, \sigma_x, \sigma_y, \sigma_z\}$. (Four real numbers \equiv Two complex numbers)



Types of Building Blocks (contd.)

3-dim: The simplex is a tetrahedron. At molecular scale, sp^3 -hybridised orbitals provide its dual form. Tetrahedral geometry of carbon provides a convenient way to understand the 3-dim language of proteins.

SU(2): Description of quantum bits is based on this group with three generators. Arbitrary states of a qubit (including mixed states) can be fully described using a density matrix, which is a linear combination of the four operators $\{1, \sigma_x, \sigma_y, \sigma_z\}$. (Four real numbers \equiv Two complex numbers)

Larger groups have been used in error correcting codes and cryptography, but not for processing information.



Types of Processing

0 - dim: Only mathematical operation allowed with superposition is addition.

It is commutative and produces interference effects.



Types of Processing

0 - dim: Only mathematical operation allowed with superposition is addition. It is commutative and produces interference effects.

1 - dim: Z_2 is a field—the smallest one. It allows two different commutative operations, addition and multiplication, which are the basis of all our arithmetic.



Types of Processing

0 - dim: Only mathematical operation allowed with superposition is addition. It is commutative and produces interference effects.

1 - dim: Z_2 is a field—the smallest one. It allows two different commutative operations, addition and multiplication, which are the basis of all our arithmetic.

d > 1: In higher dimensions, addition generalises to translation, and multiplication to scale transformation. But rotations appear as well (non-commutative for d > 2). Algebra generated by lattice transformations is much more powerful than common arithmetic.

More and more group operations become possible with increasing dimensionality.



Quantum Computation

Inevitable: Technological progress in controlling devices with shrinking sizes is taking us to the quantum domain.

Breakthrough: Quantum algorithms are better at carrying out computational tasks than their classical counterparts.



Quantum Computation

Inevitable: Technological progress in controlling devices with shrinking sizes is taking us to the quantum domain.

Breakthrough: Quantum algorithms are better at carrying out computational tasks than their classical counterparts.

Classical (particle dynamics):

Discrete Boolean logic is implemented using digital circuits.

Wave (classical wave dynamics):

Analog variables can superpose, interfere, disperse etc., and are convenient for differentiation and integration.

Waves have been widely used in communications, and in simple analog computers (e.g. RLC circuits), but they have been left out of digital computation.



Quantum Computation

Inevitable: Technological progress in controlling devices with shrinking sizes is taking us to the quantum domain.

Breakthrough: Quantum algorithms are better at carrying out computational tasks than their classical counterparts.

Classical (particle dynamics):

Discrete Boolean logic is implemented using digital circuits.

Wave (classical wave dynamics):

Analog variables can superpose, interfere, disperse etc., and are convenient for differentiation and integration.

Waves have been widely used in communications, and in simple analog computers (e.g. RLC circuits), but they have been left out of digital computation.

Quantum (particle+wave dynamics):

Unitary evolution is implemented using quantum states. Combination of particle and wave properties produces unusual correlations called entanglement.

Basics

The simplest quantum system is a qubit, with two basis vectors $|0\rangle$ and $|1\rangle$ (e.g. $|\uparrow\rangle$ and $|\downarrow\rangle$ for an electron spin). A generic qubit state is a 2-dim complex unit vector.

$$|q\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

A quantum register is an ordered string of n qubits. It is a complex unit vector in the 2^n -dim Hilbert space.

$$|x\rangle = \sum_{i_1, i_2 \dots i_n = 0}^{1} c_{i_1 i_2 \dots i_n} |x_{i_1}\rangle |x_{i_2}\rangle \dots |x_{i_n}\rangle, \sum_{i_1, i_2 \dots i_n = 0}^{1} |c_{i_1 i_2 \dots i_n}|^2 = 1.$$



Basics

The simplest quantum system is a qubit, with two basis vectors $|0\rangle$ and $|1\rangle$ (e.g. $|\uparrow\rangle$ and $|\downarrow\rangle$ for an electron spin). A generic qubit state is a 2-dim complex unit vector.

$$|q\rangle = \alpha|0\rangle + \beta|1\rangle, \quad |\alpha|^2 + |\beta|^2 = 1.$$

A quantum register is an ordered string of n qubits. It is a complex unit vector in the 2^n -dim Hilbert space.

$$|x\rangle = \sum_{i_1, i_2 \dots i_n = 0}^{1} c_{i_1 i_2 \dots i_n} |x_{i_1}\rangle |x_{i_2}\rangle \dots |x_{i_n}\rangle, \sum_{i_1, i_2 \dots i_n = 0}^{1} |c_{i_1 i_2 \dots i_n}|^2 = 1.$$

A generic instruction is a rotation of the quantum state vector in the Hilbert space. It is a unitary transformation that is deterministic and fully reversible.

A measurement is a projection. In the computational basis, it yields the state $|x_{i_1}\rangle|x_{i_2}\rangle\dots|x_{i_n}\rangle$ with probability $|c_{i_1i_2...i_n}|^2$. This operation is probabilistic and irreversible.



Communication Tasks

Singlet Bell State: $|\psi_-\rangle=(|0\rangle|0\rangle-|1\rangle|1\rangle)/\sqrt{2}$ Individual qubits behave randomly and carry no information. But jointly the two qubits are perfectly correlated.



Communication Tasks

Singlet Bell State: $|\psi_-\rangle=(|0\rangle|0\rangle-|1\rangle|1\rangle)/\sqrt{2}$ Individual qubits behave randomly and carry no information. But jointly the two qubits are perfectly correlated.

Dense coding: A Bell state exists spread over two locations. One of the four operators $\{I, X, Z, XZ\}$ is applied to the half Bell state at one end, and the qubit is sent to the other end. Joint Bell basis measurement of the two qubits determines which of the four operators was applied.

$$|\psi_{-}\rangle \longrightarrow \{|\psi_{-}\rangle, |\phi_{-}\rangle, |\psi_{+}\rangle, |\phi_{+}\rangle\}$$



Communication Tasks

Singlet Bell State: $|\psi_{-}\rangle=(|0\rangle|0\rangle-|1\rangle|1\rangle)/\sqrt{2}$ Individual qubits behave randomly and carry no information. But jointly the two qubits are perfectly correlated.

Dense coding: A Bell state exists spread over two locations. One of the four operators $\{I, X, Z, XZ\}$ is applied to the half Bell state at one end, and the qubit is sent to the other end. Joint Bell basis measurement of the two qubits determines which of the four operators was applied.

$$|\psi_{-}\rangle \longrightarrow \{|\psi_{-}\rangle, |\phi_{-}\rangle, |\psi_{+}\rangle, |\phi_{+}\rangle\}$$

Teleportation: A Bell state exists spread over two locations. The unknown state to be teleported from one end is jointly measured with the half Bell state in Bell basis. The two-bit measurement result sent to the other end recreates the unknown state from the other half of the Bell state.

$$|\alpha\rangle|\psi_{-}\rangle = (|\psi_{-}\rangle|\alpha\rangle + |\phi_{-}\rangle X|\alpha\rangle + |\psi_{+}\rangle Z|\alpha\rangle + |\phi_{+}\rangle ZX|\alpha\rangle)/2$$



Superadditivity: More information can be sent through an n-product channel than n times the amount that can be sent through a single use of a channel. Quantum correlations between the signals provide the extra channel capacity.

Zero capacity channels can be combined to obtain a nonzero communication rate! (Signal and noise can be separated when they have different correlation scales.)

Perfect entanglement between inputs to the channels can even eliminate the noise.



Superadditivity: More information can be sent through an n-product channel than n times the amount that can be sent through a single use of a channel. Quantum correlations between the signals provide the extra channel capacity.

Zero capacity channels can be combined to obtain a nonzero communication rate! (Signal and noise can be separated when they have different correlation scales.)

Perfect entanglement between inputs to the channels can even eliminate the noise.

Classical vs. Quantum:

Shannon entropy is generalised to von Neumann entropy.

$$H(\lbrace p_i \rbrace) = -\sum_i p_i \log p_i \longrightarrow S(\rho) = -\text{Tr}(\rho \log \rho).$$

But a bit and a qubit are incomparable units of information.



Superadditivity: More information can be sent through an n-product channel than n times the amount that can be sent through a single use of a channel. Quantum correlations between the signals provide the extra channel capacity.

Zero capacity channels can be combined to obtain a nonzero communication rate! (Signal and noise can be separated when they have different correlation scales.)

Perfect entanglement between inputs to the channels can even eliminate the noise.

Classical vs. Quantum:

Shannon entropy is generalised to von Neumann entropy.

$$H(\lbrace p_i \rbrace) = -\sum_i p_i \log p_i \longrightarrow S(\rho) = -\text{Tr}(\rho \log \rho).$$

But a bit and a qubit are incomparable units of information.

Shannon's theorems:

- Data compression is analogous to the classical case.
- Channel capacity is not analogous to the classical case.



Superadditivity: More information can be sent through an n-product channel than n times the amount that can be sent through a single use of a channel. Quantum correlations between the signals provide the extra channel capacity.

Zero capacity channels can be combined to obtain a nonzero communication rate! (Signal and noise can be separated when they have different correlation scales.)

Perfect entanglement between inputs to the channels can even eliminate the noise.

Classical vs. Quantum:

Shannon entropy is generalised to von Neumann entropy.

$$H(\lbrace p_i \rbrace) = -\sum_i p_i \log p_i \longrightarrow S(\rho) = -\text{Tr}(\rho \log \rho).$$

But a bit and a qubit are incomparable units of information.

Shannon's theorems:

- Data compression is analogous to the classical case.
- Channel capacity is not analogous to the classical case.



Power of One Qubit

The DQC1 model has one control qubit, n others in a fully mixed state, and a unitary operation between them.

$$\rho_i = \frac{1}{2}(I + \alpha X) \otimes \frac{1}{2^n} I^{\otimes n} \xrightarrow{Control - U_n} \rho_f = \frac{1}{2^{n+1}} \begin{pmatrix} I^{\otimes n} & \alpha U_n^{\dagger} \\ \alpha U_n & I^{\otimes n} \end{pmatrix}$$

Measurement of the control qubit gives:

$$\langle X \rangle = \alpha \operatorname{Re}(\operatorname{Tr} U_n)/2^n, \ \langle Y \rangle = \alpha \operatorname{Im}(\operatorname{Tr} U_n)/2^n$$



Power of One Qubit

The DQC1 model has one control qubit, n others in a fully mixed state, and a unitary operation between them.

$$\rho_i = \frac{1}{2}(I + \alpha X) \otimes \frac{1}{2^n} I^{\otimes n} \xrightarrow{Control - U_n} \rho_f = \frac{1}{2^{n+1}} \begin{pmatrix} I^{\otimes n} & \alpha U_n^{\dagger} \\ \alpha U_n & I^{\otimes n} \end{pmatrix}$$

Measurement of the control qubit gives:

$$\langle X \rangle = \alpha \operatorname{Re}(\operatorname{Tr} U_n)/2^n, \ \langle Y \rangle = \alpha \operatorname{Im}(\operatorname{Tr} U_n)/2^n$$

There exist $2^n \times 2^n$ unitary operations U_n that can be implemented using poly(n) one- and two-qubit operations, but whose trace cannot be computed with a known poly(n) classical algorithm. That makes the DQC1 model more powerful than its classical counterpart.

There is no entanglement between the control qubit and the mixed state in this example.



Database Search

Classical:

Binary tree search is the optimal classical algorithm. A sorted database of N items can be searched using $\log_2 N$ binary questions.

An unsorted database of N items can be searched using N/2 binary questions with memory, and using N binary questions without memory.

Quantum/Wave:

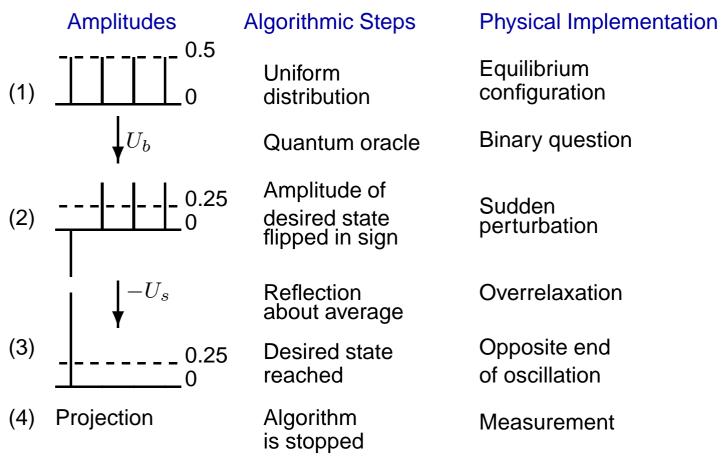
Wave mechanics works with amplitudes and not with probabilities. Superposition of amplitudes can yield constructive as well as destructive interference. Optimal search solutions differ from the classical ones.

Grover's algorithm: An unsorted database of N items can be optimally searched using $(\pi/4)\sqrt{N}$ binary questions.



Grover's Database Search

The steps of the algorithm for the simplest case of 4 items in the database. Let the first item be desired by the oracle.



(Dashed line denotes the average amplitude.)

One binary question unambiguously distinguishes 4 items.

References

All my papers are available at http://arXiv.org/

```
Minimal Languages:
```

quant-ph/0503068: The Future of Computation,

Proceedings of QICC 2005, Kharagpur, Allied Publishers (2006) 197-206

0705.3895[q-bio.GN]: Towards Understanding the Origin of Genetic Languages,

Chapter 10 in "Quantum Aspects of Life",

Eds. D. Abbott et al., Imperial College Press (2008) 187-219

Wave Algorithms:

0711.4733[quant-ph]: A Coupled Oscillator Model

for Grover's Quantum Database Search Algorithm,

Intel International Science and Engineering Fair, Albuquerque, USA (2007)

1108.1659[quant-ph]: Quantum Computation: Particle and Wave Aspects of Algorithms,

Resonance 16 (2011) 821-835

