

Outline of Information Theory

January 2, 2012

Contents

| | |
|---|---|
| MATH>Information Theory | 1 |
| MATH>Information Theory>Reversible | 2 |
| MATH>Information Theory>Coding | 2 |
| MATH>Information Theory>Error Prevention | 3 |
| MATH>Information Theory>Information | 3 |
| MATH>Information Theory>Information>Unit | 4 |
| MATH>Information Theory>Information Channel | 4 |
| MATH>Information Theory>Noise | 5 |

Note: To look up references, see the Consciousness Bibliography, listing 10,000 books and articles, with full journal and author names, available in text and PDF file formats at http://www.outline-of-knowledge.info/Consciousness_Bibliography/index.html.

MATH>Information Theory

information theory

Physical medium can store bits or qubits {information science} {information theory, mathematics}. Physical media can transform, to input, process, or output information.

memory

Because information uses bits, two-states devices, such as on-off switches, can store information. Devices like switch series can store bit series. More switches can hold more bits.

memory: address

Switches have addresses, so processing can access them and use relations between them.

information

Information amount is alternative number.

transfer

On or off signal series can transmit information-bit series. Signals have sequence numbers, so processing can access them and use relations between them.

transfer: capacity

Carrier-wave frequency has number of on-off positions per second, which is information carrying capacity.

cross-section

Information channels have cross-sections, which can hold waves or carriers, with total channel capacity. Surface area limits information capacity. Information transfer always flows through surface, even from three-dimensional space region.

algorithmic information theory

Computability theory (Turing) and information theory (Shannon) have relations {algorithmic information theory} [Chaitin, 1987] (Andrei N. Kolmogorov) (Ray J. Solomonoff) [1960]. Strings (patterns) have complexity (Kolmogorov complexity). String information is the smallest possible program that generates the string.

A universal computer running a random program has a probability {algorithmic probability} (Solomonoff) [1960] of outputting a string. Strings represent patterns and so algorithmic probability helps study induction.

efficiency in communication

In contexts, efficient communication {efficiency, communication} requires that linguistic-unit-contrast number is inversely proportional to frequency.

ensemble of information

At information-channel ends are possible-event sets {ensemble, information}. Events have probability. Information exchange changes probabilities, so people know more about events.

redundancy of information

Data repetition {redundancy} adds more information about context and state. For example, whole message can repeat to double information. Information channels can carry copies or repeats of same information. Parallel information channels can carry same information. Redundancy can overcome noise. Repeating message eliminates transient errors but not systematic errors.

sampling theorem

Theorems {sampling theorem} {Logan's zero-crossing theorem} describe how to extract information from data.

MATH>Information Theory>Reversible

reversible computing

Fredkin gate (Edward Fredkin) and Toffoli gate (Tommaso Toffoli) are reversible circuits that preserve information {reversible computing}.

controlled NOT gate

If first switch is at 1, second switch NOTs second input bit {controlled NOT gate}. If first switch is at 0, second switch transfers same second input bit. First input bit always transfers. Passing two-bit signals through controlled NOT gates twice restores two-bit signals, allowing reversible computing.

Landauer principle

Erasing information releases energy {Landauer's principle} {Landauer principle}.

MATH>Information Theory>Coding

coding length

In information theory, symbol coding-lengths {coding length} are inversely proportional to symbol probability, compared to other possible symbols. Number of bits needed equals negative of base-2 logarithm of probability. More-frequently-occurring symbols can use shorter-length strings, while less-frequently-occurring symbols can use longer strings {variable length code}. String length can increase to provide more redundancy {geometric code}.

compression in coding

Instead of binary code, codes can represent series, to make total length shorter {compression, information}. Instead of using 0 series, code can denote series length. For example, 000000000000000 can have code 1111, because number of 0's is 15.

symbol number

Compression requires that code has few symbols, allowing more repetition.

predictability

Series make predictability high. If predictability is high, number of possible states is less, and code can use fewer information bits.

arithmetic coding

Symbol probability can be relative symbol memory-amount needed.

amount

Maximum compression is about 100 times.

no compression

If system can have new elements, bits must be independent, allowing no compression.

Gray code

Information can change from analog into digital form {Gray code}. Analog changes can change digital code by one.

Huffman code

String length can be inversely proportional to symbol or state probability {Huffman code}.

instantaneous code

Coding {instantaneous code} can use no prefixes, so no reverse processing is necessary.

modulo-37 arithmetic

Modulo 37 can progressively digitize message, because 37 is prime {modulo-37 arithmetic}.

MATH>Information Theory>Error Prevention

error prevention

Information transfer can prevent errors {error prevention}. For fewest errors, strings and blocks are approximately same length.

check bit

Extra bits {check bit} in bytes can be for detecting message errors. Check bits are independent.

error correction

Coding methods {error correction} can correct errors automatically. Error-correcting code can perform same operation three times and use majority result.

Hamming code

Ordered parity-check series, using parity bits for overlapping strings, can find error positions {Hamming code}. Programs can check strings at all positions. For position errors, programs add 1 at position.

logical sum checking

Error checking {logical sum checking} can find logical sum of bits.

parity checking

Error checking {parity checking} can compare check bit to sum of other bits. Parity checks have order {syndrome, parity}, typically by position.

rectangular code

Error-correcting codes {rectangular code} can code message in arrays, with check bits for rows and columns. The same check bit can be for row and column {triangular code}. Three-dimensional arrays can have line check bits {cubic code}. Codes with higher dimension are better for error prevention.

Reed-Solomon code

Parity checking {Reed-Solomon code} can use Galois field theory.

source symbol

Symbols {source symbol} added to predicted symbols in messages can make error codes. Source symbol adds to predicted symbol at receiving end, to generate original string, which will be correct even if prediction is in error.

weighted check sum

Error checking {weighted check sum} can use bit frequencies.

MATH>Information Theory>Information

information and theory

Positions have a finite number of possible states {information}. Positions can be static, as in memories or registers, or moving, as in information channels. Mathematical rules {information theory, data} describe storing, retrieving, and transmitting data.

information extraction from data

States differ from other states, so information extraction at locations notes differences, rather than measuring amounts. Information is any difference, change, or possible-set selection.

Sampling theorems, such as Logan's zero-crossing theorem, describe how to extract information from data.

probability

States have probabilities of being at locations. If location has states at random, there is no information, even if states have known transitions. Non-random conditional probability is information.

system

Finite systems have finite numbers of elements, which have finite numbers of states. Systems are information spaces, and distributions are information sets. Highest probability has the most possible states. Some outputs are typically more probable than others.

dependence

Difference between sum of independent entropies and actual system entropy measures dependence. System subsets can depend on whole system {mutual information}.

data

Memories, registers, and information flows have state series {data}.

context

Preceding, following, adjacent, and related states define state environment {context, state} {data context}. Information meaning comes from context. Contexts have codes or contrasts. Syntax defines context relations.

code

Contexts have possible-symbol sets {code} {contrast, data}. Symbols have probabilities of being in contexts, which are information amounts.

MATH>Information Theory>Information>Unit

bit of information

The smallest information amount {bit, information} involves one position that can have two equally probable states, so states have probability 1/2. If one position has one possible state, state probability is 1, but this situation has no differences and no information. If one position has three equally probable states, states have probability 1/3, requiring 1.5 information bits. If one position has four equally probable states, states have probability 1/4, requiring two information bits. If two positions each have two equally probable states, pairs have probability 1/4, requiring two information bits.

quantum bit

Smallest quantum-information amount {quantum bit}| {qubit} involves 0 and 1 superposition.

model

Sphere points, with 0 and 1 at pole, can represent superposition. Sphere points have probabilities of obtaining 0 or 1 at decoherence.

information

Qubits have one quantum information bit {Holevo chi}, because output is 0 or 1 after decoherence. This information bit is quantum equivalent of information-theory information bit (Shannon).

entanglement

Quantum particles can be in systems with overall quantum states, so quantum-particle states interact by entanglement.

decoherence

Isolated systems can maintain quantum states, as in superconductivity and quantum Hall effect. Measurements, gravity, and other interactions with larger systems can collapse wavefunctions and cause wave decoherence.

uses

Quantum states can teleport, because entanglement can transfer to another quantum system. Quantum states can use entanglement for cryptography keys. Quantum-mechanical computers use entangled qubits. Quantum computers can find integer prime factors in same way as finding quantum-system energy levels. Quantum error correction can eliminate noise and random interactions of quantum system with environment, by correcting states without knowing what they are. However, unknown-state quantum bit cannot duplicate.

MATH>Information Theory>Information Channel

information channel

Two ensembles can link on paths {channel, information} {information channel} {communication channel} that carry information. Information channel transmits output sets, for information categories. Information-channel receiver knows output set, how to process outputs, how to correct errors, and how to mitigate noise. Communication-channel

input transforms into output, typically by selecting from available outputs. Physical information channels use frequency ranges, directions, times, and physical media.

bandwidth

Main frequency limits higher frequency or amplitude range {bandwidth}.

carrier frequency

Main frequency {carrier frequency} can carry information. Information can be in higher frequencies superimposed on main frequency {frequency modulation, data}. Information can be in main-frequency amplitude variations {amplitude modulation, data}.

channel capacity

Channels can carry numbers of bits each second {channel capacity}. Channel capacity depends on carrier frequency. Higher frequencies can carry more information. Channel capacity depends on carrying method. Older methods are amplitude modulation and frequency modulation.

MATH>Information Theory>Noise

noise

Communication-channel random disturbances {noise} can happen at encoding or decoding and interfere with selecting correct symbol from possible symbols. Noise decreases information. More noise requires more redundancy, to overcome information loss by adding information. Codes {error correcting code} can correct errors automatically, by adding information to overcome noise.

1 over f noise

Sounds can have loudness related to frequency reciprocal {1 over f noise} {1/f noise}, which is music, time-measurement, flow, and other rhythmic-event noise. 1/f noise is self-symmetric and fractal.

1 over f squared noise

Sounds can have loudness related to frequency-squared reciprocal {1 over f squared noise} {1/f² noise}, which is music noise. 1/f² noise is self-symmetric and fractal.

white noise

Sounds {white noise} can be purely random and not depend on loudness or frequency.