A VERY BRIEF HISTORY OF SECRECY

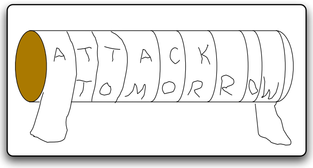
ARTUR EKERT

Human desire to communicate secretly is at least as old as writing itself and goes back over four thousand years, to the beginnings of our civilisation. Both Sumerian cuneiform and Egyptian hieroglyphs were sometimes altered to disguise their original meaning, however, it was the invention of the alphabet, a handful of symbols that represent sounds, that paved the way for the art, and eventually science, of secret communication. It is much easier to play with a set of thirty odd symbols rather than hundreds of pictograms.

1.1. **Transpositions and substitutions.** There are basically two operations on letters that can be used to disguise written text; these are transpositions and substitutions.

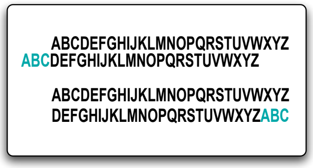
Scytale
Spartans
400BC

TRANSPOSITION



SCYTALE - SPARTA 400 BC

SUBSTITUTION


ABCDEFGHIJKLMNOPQRSTUVWXYZ
ABCDEFGHIJKLMNOPQRSTUVWXYZ
ABCDEFGHIJKLMNOPQRSTUVWXYZ
DEFGHIJKLMNOPQRSTUVWXYZABC

CAESAR CIPHER - ROME 50 BC

The first one was popular among the Spartans, the most warlike of the Greeks, who around 400 BC employed a device known as the scytale. It was used for communication between military commanders and consisted of a tapered baton around which was wrapped a spiral strip of parchment or leather containing the message. Words were then written lengthwise along the baton, one letter on each revolution of the strip. When unwrapped, the letters of the message appeared scrambled and the parchment was sent on its way. The receiver wrapped the parchment around another baton of the same shape and the original message reappeared.

substitution ciphers
Julius Caesar
100-44 BC

The first documented use of substitution of letters for secret communication appears in Julius Caesar's *Gallic Wars*. Caesar allegedly used a simple substitution method - each letter was replaced by the letter that followed it alphabetically by three places. The letter A was replaced by D, the letter B by E, and so on. For example, the English word COLD after the Caesar substitution appears as FROG. This method is still called the Caesar cipher, regardless the size of the shift used for the substitution.

Simple substitution ciphers are easy to break. For example, the Caesar cipher with 26 letters admits any shift between 0 and 25, so it has 26 possible substitutions. One can easily try them all, one by one. It is surprising that Caesar was able to keep any messages secret with this simple method, but evidently it worked well enough. The most general form of one-to-one substitution, not restricted to the shifts, can generate a staggering number,

Stirling's formula
$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$

$$26! \text{ or } 403,291,461,126,605,635,584,000,000 \approx 4 \times 10^{26} \tag{1}$$

possible substitutions. And yet, ciphers based on one-to-one substitutions, also known as *monoalphabetic* ciphers, can be easily broken by frequency analysis. The method was proposed by the 9th century polymath from Baghdad, al-Kindi, often called the philosopher of the Arabs.

frequency of letters
Al-Kindi
800-873

Al-Kindi noticed that if a letter in a message is replaced with a different letter or symbol then the new letter will take on all the characteristics of the original one. A simple substitution cipher cannot disguise certain features of the message, such as the relative frequencies of the different characters. Take the English language: the letter E is the most common letter, accounting for 12.7% of all letters, followed by T (9.0%), then A (8.2%) and so on. This means that if E is replaced by a symbol X, then X will account for roughly 13% of symbols in the concealed message, thus one can work out that X actually represents E. Then we look for the second most frequent character in the concealed message and identify it with the letter T, and so on. If the concealed message is sufficiently long then it is possible to reveal its content simply by analysing the frequency of the characters.

1.2. **Renaissance cryptography.** In the fifteenth and the sixteenth centuries, monoalphabetic ciphers were gradually replaced by more sophisticated methods. At the time Europe, Italy in particular, was a place of turmoils, intrigues and struggles for political and financial power, and the cloak-and-dagger atmosphere was ideal for cryptography to flourish.

multiple substitutions
concentric discs
Leone Battista Alberti
1404-1472

In the 1460s Leone Battista Alberti, better known as the Renaissance architect, invented a device based on two concentric discs that simplified the use of Caesar ciphers. The substitution - i.e. the relative shift of the two alphabets - is determined by the relative rotation of the two disks. It is believed that Alberti also considered changing the substitution within one message by turning the inner disc in his device this way he discovered the so-called *polyalphabetic ciphers*, which are based on superpositions of Caesar ciphers with different shifts. For example, the first letter in the message can be shifted by 7, the second letter by 14, the third by 19, the fourth again by 7, the fifth by 14, the sixth by 19, and so on repeating the shifts 7, 14, 19 throughout the whole message. The sequence of numbers - in this example 7, 14, 19 - is usually referred to as a *cryptographic key* and can be memorised as a keyword made out of the 7th, the 14th and the 19th letter of the alphabet, i.e. HOT. Using this particular key we transform the message SELL into its concealed version, which reads ZSES.

In technical terms the message to be concealed is often called the *plaintext* and the operation of disguising it is known as *encryption.* The encrypted plaintext is called the ciphertext or cryptogram. Our example illustrates the departure from a simple substitution; the repeated L in the plaintext SELL is enciphered differently in each case. Similarly, the repeated S in the ciphertext represent a different letter in the plaintext: the first S corresponds to the letter E and the second to the letter L. This makes the straightforward frequency analysis of characters in ciphertexts obsolete. Indeed, polyalphabetic ciphers invented by the main contributors to the field at the time, such as Johannes Trithemius, Blaise de Vigenere, and Giovanni Battista Della Porta, were considered unbreakable for at least another 200 years.

1.3. **(Not so) unbreakable.** The first description of a systematic method of breaking polyalphabetic ciphers was published in 1863 by the Prussian colonel Friedrich Wilhelm Kasiski, but, according to some sources, Charles Babbage had worked out the same method in private sometime in the 1850s. The basic idea of breaking polyalphabetic ciphers is based on the observation that if we use l different substitutions in a periodic fashion then every lth character in the cryptogram is enciphered with the same monoalphabetic cipher. In this case we have to find l, the length of the key, and apply frequency analysis to sub-cryptograms composed of every lth character of the cryptogram.

Friedrich Kasiski
1805-1881

Charles Babbage
1791-1871

But how do we find l? We look for repeated sequences in the ciphertext. If a sequence of letters in the plaintext is repeated at a distance which is a multiple of l, then the corresponding ciphertext sequence is also repeated. For example, for $l = 3$, with the $7, 14, 19$ (HOT) shifts we encipher

T	O	B	E	O	R	N	O	T	T	O	B	E
H	O	T	H	O	T	H	O	T	H	O	T	H
A	C	U	L	C	V	U	C	M	A	C	U	L

The repeated sequence ACUL is a giveaway. The repetition appears at a distance 9 thus we can infer that possible values of l are 9 or 3 or 1. We can then apply frequency analysis to the whole cryptogram, to every third character and to every ninth character; one of them will reveal the plaintext. This trial and error approach is getting more difficult for large values of l, i.e. for long keys, however, there are number of alternative techniques that can be used in such cases (see Exercises).

In the 1920s electromechanical technology transformed the original Alberti's disks into rotor machines in which an encrypting sequence with an extremely long period of substitutions could be generated, by rotating a sequence of rotors. Probably the most famous of them is the Enigma machine, patented by Arthur Scherbius in 1918. A notable achievement of cryptanalysis was the breaking of the Enigma in 1933. In the winter of 1932, Marian Rejewski, a twenty-seven year old cryptanalyst working in the Cipher Bureau of the Polish Intelligence Service in Warsaw, mathematically determined the wiring of the Enigma's first rotor. From then on, Poland was able to read thousands of German messages encrypted by the Enigma machine. In July 1939 Poles passed the Enigma secret to French and British cryptanalysts. After Hitler invaded Poland and France the effort of breaking Enigma ciphers continued at Bletchley Park in England. A large Victorian mansion in the centre of the park (now a museum) housed the Government Code and Cypher School, the precursor of the Government Communication Headquarters, and was the scene of many spectacular advances in modern cryptanalysis.

Enigma
Marian Rejewski
1905-1980

So, yet again ingenuity of code-makers was matched by the ingenuity of code-breakers. It is clear that if we are to construct a perfect cipher we must wipe out any patterns, such as frequency of individual letters, pairs of letters (digrams), triplets of letters (trigrams), and so on. Can this be done? Yes, it can!

1.4. **Truly unbreakable?** Despite its long history, cryptography only became part of mathematics and information theory in the late 1940s, mainly as a result of the work of Claude Shannon, of Bell Laboratories in the U.S. He showed that truly unbreakable ciphers do exist and, in fact, they had been known for over 30 years. They were devised in about 1918 by an American Telephone and Telegraph engineer Gilbert Vernam and Major Joseph Mauborgne of the US Army Signal Corps, and are called either one-time pads or Vernam ciphers.

One-time pads

Both the original design and the modern version of one-time pads are based on the binary alphabet. The message and the key are sequences of 0's and 1's of the same length. Each bit of the message M, or the plaintext, is then combined with the respective bit of the key K using addition modulo 2, defined as: $0 \oplus 0 = 0,\ 0 \oplus 1 = 1 \oplus 0 = 1,\ 1 \oplus 1 = 0$. This operation is also known as the 'exclusive or' (XOR). For example, given the key (1100101) the sender can encrypt his binary message, say (1011100), by combining each bit of the message with the respective bit of the key.

$$\begin{array}{cccccccc} & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ \oplus & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ \hline & 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{array}$$

The resulting cryptogram (0111001), can be then publicly transmitted to the receiver, who can recover the message by adding (modulo 2 again) the cryptogram and the key. This is, by the way, how we add two binary strings of the same length. Message M is added, bit by bit, to the key K to get cryptogram $C = M \oplus K$, which is then decrypted as

$$C \oplus K = (M \oplus K) \oplus K = M \oplus (K \oplus K) = M \tag{2}$$

because for any binary string K, $K \oplus K = 0$ and the binary addition is associative, that is, for any binary strings A, B, C of the same length we have $A \oplus (B \oplus C) = (A \oplus B) \oplus C$.

In one-time-pads the key is a random sequence of 0's and 1's, and therefore the cryptogram—the plaintext plus the key—is also random and completely scrambled unless one knows the key. Both the sender and the receiver must have exact copies of the key beforehand; the sender needs the key to encrypt the plaintext, the receiver needs the key to recover the plaintext from the cryptogram. An eavesdropper, who has intercepted the cryptogram and knows the general method of encryption but not the key, will not be able to infer anything useful about the original message. The secrecy of communication depends entirely on the secrecy of the key. If the key is secret, the same length as the message, truly random, and never reused, then one-time-pad is unbreakable. However, failure to comply with any of these requirements may result is serious security breaches. For example, National Security Agency, in a project VENONA, managed to decrypt some of KGB communications because of Soviet's reuse of keys in one-time pad encryptions.

1.5. **Key distribution problem.** One-time pads are not the most practical ciphers on the planet but, if used properly, they are unbreakable! Does it solve the problem of perfect security? Well, not quite. There is a snag. All one-time pads suffer from a serious practical drawback, known as the *key distribution problem.* Potential users have to agree secretly and in advance on the key - a long, random sequence of 0's and 1's. Once they have done this they can use the key for enciphering and deciphering, and the resulting cryptograms can be transmitted publicly, for example, broadcasted by radio, posted on Internet or printed in a newspaper, without compromising the security of messages. But the key itself must be established between the sender and the receiver by means of a very secure channel - for example, a very secure telephone line, a private meeting or hand-delivery by a trusted courier. Such a secure channel is usually available only at certain times and under certain circumstances. So users far apart, in order to guarantee perfect security of subsequent crypto-communication, have to carry around with them an enormous amount of secret and meaningless information (cryptographic keys), equal in volume to all the messages they might later wish to send. This is, to say the least, not very convenient!

Furthermore, even if a "secure" channel is available, this security can never be truly guaranteed. A fundamental problem remains because, in principle, any classical private channel can be monitored passively, without the sender or receiver knowing that the eavesdropping has taken place. This is because classical physics - the theory of ordinary-scale bodies and phenomena such as paper documents, magnetic tapes and radio signals - allows all physical properties of an object to be measured without disturbing those properties. Since all information, including cryptographic keys, is encoded in measurable physical properties of some object or signal, classical theory leaves open the possibility of passive eavesdropping, because in principle it allows the eavesdropper to measure physical properties without disturbing them.

1.6. **Public keys.** Cryptologists and mathematicians tried very hard to eliminate the key distribution problem. The 1970s, for example, brought a clever mathematical discovery in the shape of "public key" systems. The systems avoid the key distribution problem but unfortunately their security depends on unproven mathematical assumptions, such as, for example, the difficulty of factoring large numbers. For example, RSA — a very popular public key cryptosystem named after the three inventors, Ron Rivest, Adi Shamir, and Leonard Adleman — gets its security from the difficulty of factoring large numbers. However, if and when mathematicians or computer scientists come up with fast and clever procedures for factoring, the whole privacy and discretion of RSA could vanish overnight. Or if we build a quantum computer. In the 1980s physicists asked what would happen if you could build a computer out of individual atoms and molecules. The world of atoms and molecules is governed by quantum mechanics and some quantum phenomena can support qualitatively new types of computation. Moreover, there are problems which are intractable to all classical computers but which can be efficiently solved on a quantum computer, and factoring is one of them! Thus once a quantum factorisation engine (a special-purpose quantum computer for factoring large numbers) is built public key systems will become insecure. Admittedly, that day is probably decades away, but can anyone prove, or give any reliable assurance, that it is? Confidence in the slowness of technological progress is all that the security of the RSA system now rests on.

1.7. **Quantum key distribution.** But "what quantum taketh away quantum giveth back". Distribution of specially prepared photons has helped to provide a different solution to the key distribution problem. Unlike all classical cryptography it relies on the laws of physics rather than on ensuring that successful eavesdropping would require excessive computational effort. This solution is immune even to attacks from quantum computers. More than that, it allows to distribute a cryptographic key in such a way that its bit values "do not exist" until they are measured by the legitimate users. An eavesdropper cannot elicit any information from the photons while in transit from the source to the legitimate users, simply because there is no information encoded there! This is very puzzling, to say the least. And we shall see how it leads to the most paranoid form of security, namely, devices of unknown or dubious provenance, even those that are manufactured by our enemies, can be safely used for secure key distribution. This is a truly remarkable feat, also referred to as the "device independent" key distribution. Moreover the security of this method transcends the borders of quantum theory—even if one day quantum physics is refuted and superseded by a new theory, even then, as long as the new theory does not admit any instant communication, we can guarantee secure key distribution. The sheer fact that we can make sensible statements about security of devices operating according some yet to be discovered laws of physics is amazing.

Notes & Exercises

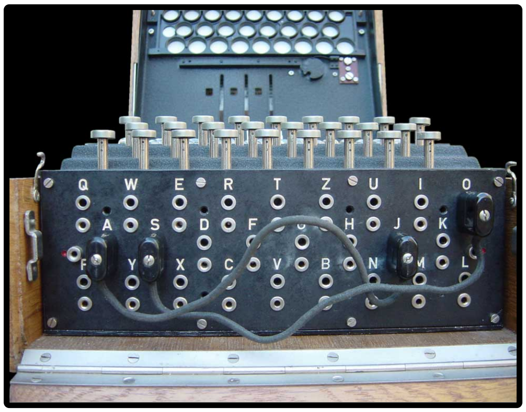
(1) **Factorials.** The exclamation mark in $n!$ is not intended to indicate surprise, but factorials do grow surprisingly fast. If you take a deck of cards and shuffle it you will get one of $52! \approx 8 \times 10^{67}$ different possible arrangements. Do you have any feeling how huge this number really is? The result of the shuffle is, most likely, an order of cards that never ever occurred before, and here by 'before' I mean in the whole history of card playing! For if all 6.7 billion people on Earth started shuffling cards, producing one shuffled deck every ten seconds or so, then it would take much more than the age of the universe (estimated at 13.7 billion years) to even have a chance of producing all possible orders.

(2) **Stirling's formula.** For large n a pretty good estimation of $n!$ was first proposed in the 18th century by a French mathematician, Abraham de Moivre. It was later improved by his Scottish colleague, James Stirling, and it is known today as Stirling's formula,

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \tag{3}$$

where e is the Euler number $e = 2.718281828...$, the basis of the natural logarithms. The number e is as ubiquitous as π, and appears in many common mathematical expressions, so we shall see it over and over again.

(3) **Enigma.** An Enigma machine, that was used for secret communication before and during World War II, implemented several permutations to the alphabet of 26 letters (A, B, C, ... X, Y, Z). One of the permutations was a plugboard permutation. The Enigma plugboard, shown below, was an array of 26 jacks (each jack consisting of a pair of holes), one for each letter, and six electrical cables, each of which can be plugged into two of the jacks so as to interchange those two letters. The effect of the plugboard was to swap six pairs of letters and let the remaining 14 letters be unchanged. The picture below shows only two cables, one swaps A and J, and another S and O.



Assume that only two cables are being used. How many different plugboard permutations are there? If your answer is $\frac{(26\times25)}{2} \times \frac{(24\times23)}{2}$ then think again; the correct answer is half of this number. What is the answer when n cables are being used? Clearly, given that we have 26 letters/ jacks we can use at most 13 cables.

(4) **Friedman's test.** Recall that in order to break a polyalphabetic cipher it is enough to find the length or the key. One of the simplest and yet effective methods to estimate the length of the key, for large keys, was proposed around 1925 by an American cryptologists William F. Friedman. The Friedman test is based on the collision probability P_c, also known as the index of coincidence, which is the probability of two letters randomly selected from a text being the same. If the letters or symbols x

from a particular alphabet appear in a text with respective frequencies $p(x)$ then the the collision probability is given by

$$P_c(\Sigma) = \sum_{x \in \Sigma} p(x)^2. \tag{4}$$

We can compute the collision probability P_c for the English language using relative frequencies of letters in a typical English text, i.e 12.7 % for E, 9.0% for T and so on. This gives $P_c = 0.065$. One-to-one substitutions do not change the relative frequencies of symbols thus for any monoalphabetic cipher $P_c = 0.065$. In contrast for a random sequence of letters $p(x) = \frac{1}{26}$ we obtain $P_c = 0.038$. How about the collision probability of a cryptogram resulting from encrypting an English text with some key of length l? Let us write the cryptogram in a table containing l columns and notice that each column forms a monoalphabetic cipher,

$$\begin{array}{llllll} C_1 & C_2 & C_3 & ... & C_{l-1} & C_l \\ C_{l+1} & C_{l+2} & C_{l+3} & ... & C_{2l-1} & C_{2l} \\ C_{2l+1} & C_{2l+2} & C_{2l+3} & ... & C_{3l-1} & C_{3l} \\ ... & ... & ... & ... & ... & ... \end{array} \tag{5}$$

If the cryptogram contains n symbols and we choose any two of them at random then out of $n(n-1)/2$ possible pairs of symbols $n\left(\frac{n}{l}-1\right)/2$ pairs are composed out of symbols from the same column and the remaining $n\left(n-\frac{n}{l}\right)/2$ pairs are composed out of symbols from different columns. In the first case the collision probability is 0.065 in the latter 0.038, hence the expected number of pairs containing identical symbols is

$$\frac{n(n-l)}{2l} 0.065 + \frac{n^2(l-1)}{2l} 0.038, \tag{6}$$

which gives

$$P_c = \frac{0.027n}{l(n-1)} + \frac{0.038n - 0.065}{n-1}, \tag{7}$$

and effectively the length of the key,

$$l = \frac{0.027n}{(n-1)P_c - 0.038n + 0.065}. \tag{8}$$

Given a cryptogram all we have to do is to estimate P_c by counting the frequency of each letter appearing in the n-character long cryptogram and then use the formula above to estimate the length of the key. You may then write the cryptogram in l columns and calculate P_c for each column - if you guessed l correctly then for each columns P_c will be approximately 0.065, otherwise you get approximately 0.038. The table below shows the correspondence between l and P_c for some selected values of l,

l	1	2	3	4	5	10	large
P_c	0.065	0.052	0.047	0.045	0.044	0.041	0.038

Simple but powerful. Friedman transformed the field of cryptology from a set of tricks and heuristic rules to something more scientific and the Friedman statistical test proved very useful in breaking more advanced ciphers which relied on mechanical encryption.

(5) Some time ago I was invited by one of the editors of the New Scientist to write a popular article about quantum cryptography. Soon after it was published the editors received a letter, from a Scottish schoolboy, who suggested the following solution to the key distribution problem. Let Alice and Bob pick up, secretly and independently from each other, binary strings A and B, respectively. Alice holds a secret key K, of the same length as A and B, which she wants to send to Bob. She adds $A \oplus K$ and sends the resulting string to Bob, who then adds his secret string B and sends

$B \oplus A \oplus K$ back to Alice. Alice then adds A again, which gives $B \oplus K$, because $A \oplus A = 0$. She sends $B \oplus K$ to Bob, who adds B and recovers K. The strings A, B are secret, their values never communicated to anyone, and the key K is indeed transferred from Alice to Bob. What is wrong with this key distribution scheme?

(6) **Groups.** The set $\{0,1\}^n$ together with the bit by bit addition $\oplus$ forms a group Z_2^n. Recall that a set G with an operation '$\cdot$', which assigns to any pair of elements $a, b \in G$ an element $a \cdot b \in G$, is called a group if the operation is associative, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, if there exist a unit element $e \in G$ such that for all $g \in G$, $g \cdot e = e \cdot g = g$, and if each element $g \in G$ has an inverse element, written g^{-1}, such that $g \cdot g^{-1} = g^{-1} \cdot g = e$. The group is said to be Abelian if, for all $a, b \in G$ we have $a \cdot b = b \cdot a$. Group G is finite if the number of elements in G is finite. The order of a finite group G, denoted as $|G|$, is the number of elements it contains. The group Z_2^n is Abelian and $|Z_2^n| = 2^n$.

(7) **Groups and fields.** We will hardly ever venture outside the binary alphabet thus let me elaborate on some of its features. We can define two elementary operations on binary digits: addition $\oplus$ and multiplication ×,

$$\begin{aligned} &0 \oplus 0 = 0, \quad 0 \oplus 1 = 1, \quad 0 \oplus 0 = 1, \quad 1 \oplus 1 = 0, \\ &0 \times 0 = 0, \quad 0 \times 1 = 0, \quad 1 \times 0 = 0, \quad 1 \times 1 = 1. \end{aligned}$$

The multiplication is also known as the logical AND ($\wedge$). The set $\{0,1\}$ together with the addition $\oplus$ forms a group Z_2 and the same set with both addition and multiplication is the simplest example of what is called a finite field. Various notation are used for this field, e.g. $\mathbb{F}_2$, $GF(2)$, $Z/2$ and Z_2. Recall that a field is a set that is a commutative group with respect to two compatible operations, addition and multiplication, with "compatible" meaning distributivity,

$$a \times (b + c) = (a \times b) + (a \times c).$$

The additive identity (0) has no multiplicative inverse (one cannot divide by 0).

(8) **Distance.** We can quantify how "close to" or "far away from" each other any two binary strings are by counting the number of binary places in which they differ. The resulting distance is known as the Hamming distance. For example, $x = 0110$ and $y = 1100$ differ in the first and the third place (counting from the left to the right), which means that they are separated by the Hamming distance 2. The Hamming weight of a binary string is the number of non-zero entries i.e. the number of entries that are 1. Check that the Hamming distance $d(\cdot,\cdot)$ is a proper distance, i.e. it is always non-negative, symmetric $d(x,y) = d(y,x)$, satisfies the triangle inequality $d(x,y) \leq d(x,z) + d(z,y)$ and $d(x,y) = 0$ implies $x = y$. Any set with a distance (a metric) is called a metric space.

(9) **Binary vectors.** Binary strings on length n can be also viewed as vectors in the n-dimensional vector space over the field $\mathbb{F}_2$. The vector space is spanned, for example, by the standard basis

$$e_1 \equiv (1,0,...,0),\ e_2 \equiv (0,1,...,0),\ \ldots, e_n \equiv (0,0,...,1)$$

and every binary string x can be expressed as $x = \sum_k c_k\, e_k$, where $k = 1...n$ and $c_k = 0,1$. We can define a scalar product of two binary strings x and y as

$$x \cdot y = (x_1 \times y_1) \oplus (x_2 \times y_2) \oplus \ldots (x_n \times y_n).$$

For example, if $x = 0110$ and $y = 1100$ then $x \cdot y = (0 \times 1) \oplus (1 \times 1) \oplus (1 \times 0) \oplus (0 \times 0) = 1$. Two strings x and y are orthogonal when $x \cdot y = 0$. Here you should be careful with your intuition - a string with an even number of ones can be orthogonal to itself. Linear transformations can be written using matrices over the field $\mathbb{F}_2$, for example,

$$\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0 & 1 & 1\\1 & 1 & 0\end{bmatrix}\begin{bmatrix}0\\1\\1\end{bmatrix}.$$

(10) **Shannon's Ultimate Machine.** In between lying down the foundations of modern information theory and cryptography Claude Shannon designed and constructed many ingenious but spectacularly useless devices, such as a petrol-powered pogo-stick, THROBAC (Thrifty ROman numerical BAckward looking Computer) a calculator that calculated in Roman numerals, a three-ball juggling machine with two hands that bounce-juggles three balls on a drumhead, and a two-seated unicycle, just to mention few of his remarkable products. He was in particular proud of his "Ultimate Machine" based on an idea of Mervin Minsky, which he built in the early fifties. The operation and spirit of the machine are best described by Arthur C. Clarke in his *Voice Across the Sea*,

> "Nothing could be simpler. It is merely a small wooden casket, the size and shape of a cigar box, with a single switch on one face. When you throw the switch, there is an angry, purposeful buzzing. The lid slowly rises, and from beneath it emerges a hand. The hand reaches down, turns the switch off and retreats into the box. With the finality of a closing coffin, the lid snaps shut, the buzzing ceases and peace reigns once more. The psychological effect, if you do not know what to expect, is devastating. There is something unspeakably sinister about a machine that does nothing – absolutely nothing – except switch itself off."

(11) **Modular Arithmetic.** Consider the set $Z_5 = \{0, 1, 2, 3, 4\}$. Define additions and multiplication modulo 5 on the set as the remainder left after dividing the sum or the product by 5. For example, $3 + 4 \bmod 5 = 2$ because $3 + 4 = 7$ divided by 5 leaves a remainder of 2; $4 \times 4 \bmod 5 = 1$ because $4 \times 4 = 16$ divided by 5 leaves a remainder of 1. Show that Z_5 together with such defined addition and multiplication forms a field. How about the set $Z_6 = \{0, 1, 2, 3, 4, 5\}$ together with additions and multiplication modulo 6 - is it a field?