Introduction

Boaz Barak

**Disclaimer:** The MS-Word version of this text has significant formatting issues compared to the PDF and HTML versions.

# Introduction

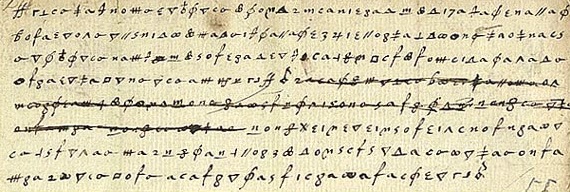
**Additional reading:** Sections 2.1 (Introduction) and 2.2 (Shannon ciphers and perfect security) in the Boneh Shoup book. Chapters 1 and 2 of Katz-Lindell book.[[1]](#footnote-21)

Ever since people started to communicate, there were some messages that they wanted kept secret. Thus cryptography has an old though arguably *undistinguished* history. For a long time cryptography shared similar features with Alchemy as a domain in which many otherwise smart people would be drawn into making fatal mistakes. Indeed, the history of cryptography is littered with the figurative corpses of cryptosystems believed secure and then broken, and sometimes with the actual corpses of those who have mistakenly placed their faith in these cryptosystems. The definitive text on the history of cryptography is David Kahn’s “The Codebreakers”, whose title already hints at the ultimate fate of most cryptosystems.[[2]](#footnote-23) (See also “The Code Book” by Simon Singh.)

We recount below just a few stories to get a feel for this field. But before we do so, we should introduce the **cast of characters**. The basic setting of “encryption” or “secret writing” is the following: one person, whom we will call **Alice**, wishes to send another person, whom we will call **Bob**, a **secret** message. Since Alice and Bob are not in the same room (perhaps because Alice is imprisoned in a castle by her cousin the queen of England), they cannot communicate directly and need to send their message in writing. Alas, there is a third person, whom we will call **Eve**, that can see their message. Therefore Alice needs to find a way to *encode* or *encrypt* the message so that only Bob (and not Eve) will be able to understand it.

## Some history

1587, Mary the queen of Scots, and the heir to the throne of England, wanted to arrange the assassination of her cousin, queen Elisabeth I of England, so that she could ascend to the throne and finally escape the house arrest under which she had been for the last 18 years. As part of this complicated plot, she sent a coded letter to Sir Anthony Babington.



Snippet from encrypted communication between queen Mary and Sir Babington

Mary used what’s known as a *substitution cipher* where each letter is transformed into a different obscure symbol (see maryscottletterfig). At a first look, such a letter might seem rather inscrutable- a meaningless sequence of strange symbols. However, after some thought, one might recognize that these symbols *repeat* several times and moreover that different symbols repeat with different frequencies. Now it doesn’t take a large leap of faith to assume that perhaps each symbol corresponds to a different letter and the more frequent symbols correspond to letters that occur in the alphabet with higher frequency. From this observation, there is a short gap to completely breaking the cipher, which was in fact done by queen Elisabeth’s spies who used the decoded letters to learn of all the co-conspirators and to convict queen Mary of treason, a crime for which she was executed. Trusting in superficial security measures (such as using “inscrutable” symbols) is a trap that users of cryptography have been falling into again and again over the years. (As in many things, this is the subject of a great XKCD cartoon, see XKCDnavajofig.)



XKCD’s take on the added security of using uncommon symbols

The [Vigenère cipher](https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher) is named after Blaise de Vigenère who described it in a book in 1586 (though it was invented earlier by Bellaso). The idea is to use a collection of substitution cyphers - if there are different ciphers then the first letter of the plaintext is encoded with the first cipher, the second with the second cipher, the with the cipher, and then the letter is again encoded with the first cipher. The key is usually a word or a phrase of letters, and the substitution cipher is obtained by shifting each letter positions in the alphabet. This “flattens” the frequencies and makes it much harder to do frequency analysis, which is why this cipher was considered “unbreakable” for 300+ years and got the nickname “le chiffre indéchiffrable” (“the unbreakable cipher”). Nevertheless, Charles Babbage cracked the Vigenère cipher in 1854 (though he did not publish it). In 1863 Friedrich Kasiski broke the cipher and published the result. The idea is that once you guess the length of the cipher, you can reduce the task to breaking a simple substitution cipher which can be done via frequency analysis (can you see why?). Confederate generals used Vigenère regularly during the civil war, and their messages were routinely cryptanalzed by Union officers.



Confederate Cipher Disk for implementing the Vigenère cipher



Confederate encryption of the message “Gen’l Pemberton: You can expect no help from this side of the river. Let Gen’l Johnston know, if possible, when you can attack the same point on the enemy’s lines. Inform me also and I will endeavor to make a diversion. I have sent some caps. I subjoin a despatch from General Johnston.”

The *Enigma* cipher was a mechanical cipher (looking like a typewriter, see enigmafig) where each letter typed would get mapped into a different letter depending on the (rather complicated) key and current state of the machine which had several rotors that rotated at different paces. An identically wired machine at the other end could be used to decrypt. Just as many ciphers in history, this has also been believed by the Germans to be “impossible to break” and even quite late in the war they refused to believe it was broken despite mounting evidence to that effect. (In fact, some German generals refused to believe it was broken even *after* the war.) Breaking Enigma was an heroic effort which was initiated by the Poles and then completed by the British at Bletchley Park, with Alan Turing (of the Turing machines) playing a key role. As part of this effort the Brits built arguably the world’s first large scale mechanical computation devices (though they looked more similar to washing machines than to iPhones). They were also helped along the way by some quirks and errors of the German operators. For example, the fact that their messages ended with “Heil Hitler” turned out to be quite useful.



In the *Enigma* mechanical cipher the secret key would be the settings of the rotors and internal wires. As the operator types up their message, the encrypted version appeared in the display area above, and the internal state of the cipher was updated (and so typing the same letter twice would generally result in two different letters output). Decrypting follows the same process: if the sender and receiver are using the same key then typing the ciphertext would result in the plaintext appearing in the display.

Here is one entertaining anecdote: the Enigma machine would never map a letter to itself. In March 1941, Mavis Batey, a cryptanalyst at Bletchley Park received a very long message that she tried to decrypt. She then noticed a curious property— the message did *not* contain the letter “L”.[[3]](#footnote-37) She realized that the probability that no “L”’s appeared in the message is too small for this to happen by chance. Hence she surmised that the original message must have been composed *only* of L’s. That is, it must have been the case that the operator, perhaps to test the machine, have simply sent out a message where he repeatedly pressed the letter “L”. This observation helped her decode the next message, which helped inform of a planned Italian attack and secure a resounding British victory in what became known as “the Battle of Cape Matapan”. Mavis also helped break another Enigma machine. Using the information she provided, the Brits were able to feed the Germans with the false information that the main allied invasion would take place in Pas de Calais rather than on Normandy.

In the words of General Eisenhower, the intelligence from Bletchley park was of “priceless value”. It made a huge difference for the Allied war effort, thereby shortening World War II and saving millions of lives. See also [this interview with Sir Harry Hinsley](http://www.cix.co.uk/~klockstone/hinsley.htm).

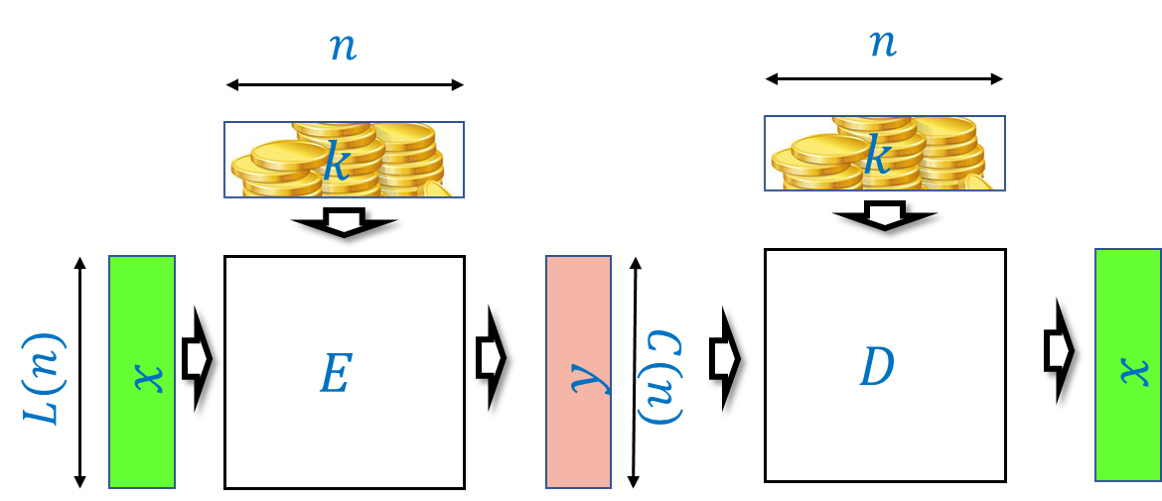
## Defining encryptions

Many of the troubles that cryptosystem designers faced over history (and still face!) can be attributed to not properly defining or understanding what are the goals they want to achieve in the first place. We now turn to actually defining what is an encryption scheme. Clearly we can encode every message as a string of bits, i.e., an element of for some . Similarly, we can encode the *key* as a string of bits as well, i.e., an element of for some . Thus, we can think of an encryption scheme as composed of two functions. The *encryption function* maps a secret key and a message (known also as *plaintext*) into a *ciphertext* for some . We write this as . The *decryption function* does the reverse operation, mapping the secret key and the cyphertext back into the plaintext message , which we write as . The basic equation is that if we use the same key for encryption and decryption, then we should get the same message back. That is, for every and ,

This motivates the following definition which attempts to capture what it means for an encryption scheme to be *valid* or “make sense”, regardless of whether or not it is *secure*:

Let and be two functions mapping natural numbers to natural numbers. A pair of polynomial-time computable functions mapping strings to strings is a *valid private key encryption scheme* (or *encryption scheme* for short) with plaintext length function and ciphertext length function if for every , and , and

We will often write the first input (i.e., the key) to the encryption and decryption as a subscript and so can write eqvalidenc also as .



A private-key encryption scheme is a pair of algorithms such that for every key and plaintext , is a ciphertext of length . The encryption scheme is *valid* if for every such , . That is, the decryption of an encryption of is , as long as both encryption and decryption use the same key.

The validity condition implies that for any fixed , the map is one to one (can you see why?) and hence the ciphertext length is always at least the plaintext length. Thus we typically focus on the plaintext length as the quantity to optimize in an encryption scheme. The *larger* is, the better the scheme, since it means we need a shorter secret key to protect messages of the same length.

*A note on notation:* We will always use to denote natural numbers.

The number will often denote the length of our secret key. The length of the key (or another closely related number) is often known as the *security parameter* in the literature. Katz-Lindell also uses to denote this parameter, while Boneh-Shoup and Rosulek use for it. (Some texts also use the greek lettter for the same parameter.) We chose to denote the security parameter by as to correspond with the standard algorithmic notation for input length (as in or time algorithms).

We often use to denote the length of the message, sometimes also known as “block length” since longer messages are simply chopped into “blocks” of length and also appropriately padded.

We will use to denote the secret key, to denote the secret plaintext message, and to denote the encrypted ciphertext. Note that and are not numbers but rather bit strings of lengths and respectively.

For simplicity, we denote the space of possible keys as and the space of possible messages as for . Boneh-Shoup uses a more general notation of for the space of all possible keys and for the space of all possible messages. This does not make much difference since we can represent every discrete object such as a key or message as a binary string. (One difference is that in principle the space of all possible messages could include messages of unbounded length, though in such a case what is done in both theory and practice is to break these up into finite-size blocks and encrypt one block at a time.)

## Defining security of encryption

encryptiondef says nothing about security and does not rule out trivial “encryption” schemes such as the scheme that simply outputs the plaintext as is. Defining security is tricky, and we’ll take it one step at a time, but lets start by pondering what is secret and what is not. A priori we are thinking of an attacker Eve that simply sees the ciphertext and does not know anything on how it was generated. So, it does not know the details of and , and certainly does not know the secret key . However, many of the troubles past cryptosystems went through was caused by them relying on “security through obscurity”— trusting that the fact their *methods* are not known to their enemy will protect them from being broken. This is a faulty assumption - if you reuse a method again and again (even with a different key each time) then eventually your adversaries will figure out what you are doing. And if Alice and Bob meet frequently in a secure location to decide on a new method, they might as well take the opportunity to exchange their secret messages..

These considerations led Auguste Kerckhoffs in 1883 to state the following principle:

*A cryptosystem should be secure even if everything about the system, except the key, is public knowledge.*[[4]](#footnote-45)

Why is it OK to assume the key is secret and not the algorithm? Because we can always choose a fresh key. But of course that won’t help us much if our key is “1234” or “passw0rd!”. In fact, if you use *any* deterministic algorithm to choose the key then eventually your adversary will figure this out. Therefore for security we must choose the key at *random* and can restate Kerckhoffs’s principle as follows:

*There is no secrecy without randomness*

This is such a crucial point that is worth repeating:

*There is no secrecy without randomness*

At the heart of every cryptographic scheme there is a secret key, and the secret key is always chosen at random. A corollary of that is that to understand cryptography, you need to know some probability theory. Fortunately, we don’t need much of probability- only probability over finite spaces, and basic notions such as expectation, variance, concentration and the union bound suffice for most of we need. In fact, understanding the following two statements will already get you much of what you need for cryptography:

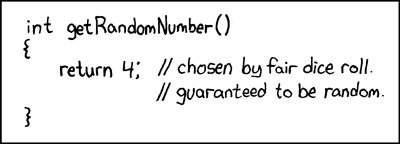
* For every fixed string , if you toss a coin times, the probability that the heads/tails pattern will be exactly is .
* A probability of is really really small.

### Generating randomness in actual cryptographic systems

How do we actually get random bits in actual systems? The main idea is to use a two stage approach. First we need to get some data that is *unpredictable* from the point of view of an attacker on our system. Some sources for this could be measuring latency on the network or hard drives (getting harder with solid state disk), user keyboard and mouse movement patterns (problematic when you need fresh randomness at boot time ), clock drift and more, there are some other sources including audio, video, and network. All of these can be problematic, especially for servers or virtual machines, and so hardware based random number generators based on phenomena such as thermal noise or nuclear decay are becoming more popular. Once we have some data that is unpredictable, we need to estimate the *entropy* in it. You can roughly imagine that has bits of entropy if the probability that an attacker can guess is at most . People then use a *hash function* (an object we’ll talk about more later) to map into a string of length which is then hopefully distributed (close to) uniformly at random. All of this process, and especially understanding the amount of information an attacker may have on the entropy sources, is a bit of a dark art and indeed a number of attacks on cryptographic systems were actually enabled by weak generation of randomness. Here are a few examples.

One of the first attacks was on the SSL implementation of Netscape (*the* browser at the time). Netscape use the following “unpredicatable” information— the time of day and a process ID both of which turned out to be quite predictable (who knew attackers have clocks too?). Netscape tried to protect its security through “security through obscurity” by not releasing the source code for their pseudorandom generator, but it was reverse engineered by [Ian Goldberg and David Wagner](https://www.cs.berkeley.edu/~daw/papers/ddj-netscape.html) (Ph.D students at the time) who demonstrated this attack.

In 2006 a programmer removed a line of code from the procedure to generate entropy in OpenSSL package distributed by Debian since it caused a warning in some automatic verification code. As a result for two years (until this was discovered) all the randomness generated by this procedure used only the process ID as an “unpredictable” source. This means that all communication done by users in that period is fairly easily breakable (and in particular, if some entities recorded that communication they could break it also retroactively). This caused a huge headache and a worldwide regeneration of keys, though it is believed that many of the weak keys are still used. See [XKCD’s take](http://www.xkcd.com/424/) on that incident.



XKCD Cartoon: Random number generator

In 2012 two separate teams of researchers scanned a large number of RSA keys on the web and found out that about 4 percent of them are easy to break. The main issue were devices such as routers, internet-connected printers and such. These devices sometimes run variants of Linux- a desktop operating system- but without a hard drive, mouse or keyboard, they don’t have access to many of the entropy sources that desktop have. Coupled with some good old fashioned ignorance of cryptography and software bugs, this led to many keys that are downright trivial to break, see [this blog post](https://freedom-to-tinker.com/blog/nadiah/new-research-theres-no-need-panic-over-factorable-keys-just-mind-your-ps-and-qs/) and [this web page](https://factorable.net/) for more details.

After the entropy is collected and then “purified” or “extracted” to a uniformly random string that is, say, a few hundred bits long, we often need to “expand” it into a longer string that is also uniform (or at least looks like that for all practical purposes). We will discuss how to go about that in the next lecture. This step has its weaknesses too and in particular the Snowden documents, combined with observations of Shumow and Ferguson, strongly suggest that the NSA has deliberately inserted a *trapdoor* in one of the pseudorandom generators published by the National Institute of Standards and Technologies (NIST). Fortunately, this generator wasn’t widely adapted but apparently the NSA did pay 10 million dollars to RSA security so the latter would make this generator their default option in their products.

## Defining the secrecy requirement.

Defining the secrecy requirement for an encryption is not simple. Over the course of history, many smart people got it wrong and convinced themselves that ciphers were impossible to break. The first person to truly ask the question in a rigorous way was Claude Shannon in 1945 (though a partial version of his manuscript was only declassified in 1949). Simply by asking this question, he made an enormous contribution to the science of cryptography and practical security. We now will try to examine how one might answer it.

Let me warn you ahead of time that we are going to insist on a *mathematically precise definition* of security. That means that the definition must capture security in all cases, and the existence of a single counterexample, no matter how “silly”, would make us rule out a candidate definition. This exercise of coming up with “silly” counterexamples might seem, well, silly. But in fact it is this method that has led Shannon to formulate his theory of secrecy, which (after much followup work) eventually revolutionized cryptography, and brought this science to a new age where Edgar Allan Poe’s maxim no longer holds, and we are able to design ciphers which human (or even nonhuman) ingenuity cannot break.

The most natural way to attack an encryption is for Eve to guess all possible keys. In many encryption schemes this number is enormous and this attack is completely infeasible. For example, the theoretical number of possibilities in the Enigma cipher was about which roughly means that even if we built a filled the milky way galaxy with computers operating at light speed, the sun would still die out before it finished examining all the possibilities.[[5]](#footnote-54) One can understand why the Germans thought it was impossible to break. (Note that despite the number of possibilities being so enormous, such a key can still be easily specified and shared between Alice and Bob by writing down digits on a piece of paper.) Ray Miller of the NSA had calculated that, in the way the Germans used the machine, the number of possibilities was “only” , but this is still extremely difficult to pull off even today, and many orders of magnitudes above the computational powers during the WW-II era. Thus clearly, it is sometimes possible to break an encryption without trying all possibilities. A corollary is that having a huge number of key combinations does not guarantee security, as an attacker might find a shortcut (as the allies did for Enigma) and recover the key without trying all options.

Since it is possible to recover the key with some tiny probability (e.g. by guessing it at random), perhaps one way to define security of an encryption scheme is that an attacker can never recover the key with probability significantly higher than that. Here is an attempt at such a definition:

# 

An encyption scheme is *-secure* if no matter what method Eve employs, the probability that she can recover the true key from the ciphertext is at most .

When you see a mathematical definition that attempts to model some real-life phenomenon such as security, you should pause and ask yourself:

1. Do I understand mathematically what is the definition stating?
2. Is it a reasonable way to capture the real life phenomenon we are discussing?

One way to answer question 2 is to try to think of both examples of objects that satisfy the definition and examples of objects that violate it, and see if this conforms to your intuition about whether these objects display the phenomenon we are trying to capture. Try to do this for securefirstattemptdef

You might wonder if securefirstattemptdef is not *too strong*. After all how are we going ever to prove that Eve cannot recover the secret key no matter what she does? Edgar Allan Poe would say that there can always be a method that we overlooked. However, in fact this definition is too *weak*! Consider the following encryption: the secret key is chosen at random in but our encryption scheme simply ignores it and lets and . This is a valid encryption, but of course completely insecure as we are simply outputting the plaintext in the clear. Yet, no matter what Eve does, if she only sees and not , there is no way she can guess the true value of with probability better than , since it was chosen completely at random and she gets no information about it. Formally, one can prove the following result:

# 

Let be the encryption scheme above. For every function and for every , the probability that is exactly .

# 

This follows because and hence which is some fixed value that is independent of . Hence the probability that is . QED

The math behind the above argument is very simple, yet I urge you to read and re-read the last two paragraphs until you are sure that you completely understand why this encryption is in fact secure according to the above definition. This is a “toy example” of the kind of reasoning that we will be employing constantly throughout this course, and you want to make sure that you follow it.

So, trivialsec is true, but one might question its meaning. Clearly this silly example was not what we meant when stating this definition. However, as mentioned above, we are not willing to ignore even silly examples and must amend the definition to rule them out. One obvious objection is that we don’t care about hiding the key- it is the *message* that we are trying to keep secret. This suggests the next attempt:

# 

An encryption scheme is *-secure* if for every message no matter what method Eve employs, the probability that she can recover from the ciphertext is at most .

Now this seems like it captures our intended meaning. But remember that we are being anal, and truly insist that the definition holds as stated, namely that for every plaintext message and every function , the probability over the choice of that is at most . But now we see that this is clearly impossible. After all, this is supposed to work for *every* message and *every* function , but clearly if is the all-zeroes message and is the function that ignores its input and simply outputs , then it will hold that with probability one.

So, if before the definition was too weak, the new definition is too strong and is impossible to achieve. The problem is that of course we could guess a fixed message with probability one, so perhaps we could try to consider a definition with a *random* message. That is:

# 

An encyption scheme is *-secure* if no matter what method Eve employs, if is chosen at random from , the probability that she can recover from the ciphertext is at most .

This weakened definition can in fact be achieved, but we have again weakened it too much. Consider an encryption that hides the last bits of the message, but completely reveals the first bits. The probability of guessing a random message is , and so such a scheme would be “ secure” per securethirdattemptdef but this is still a scheme that you would not want to use. The point is that in practice we don’t encrypt random messages— our messages might be in English, might have common headers, and might have even more structures based on the context. In fact, it may be that the message is either “Yes” or “No” (or perhaps either “Attack today” or “Attack tomorrow”) but we want to make sure Eve doesn’t learn which one it is. So, using an encryption scheme that reveals the first half of the message (or frankly even only the first bit) is unacceptable.

## Perfect Secrecy

So far all of our attempts at definitions oscillated between being too strong (and hence impossible) or too weak (and hence not guaranteeing actual security). The key insight of Shannon was that in a secure encryption scheme the ciphtertext should not reveal *any additional information* about the plaintext. So, if for example it was a priori possible for Eve to guess the plaintext with some probability (e.g., because there were only possibilities for it) then she should not be able to guess it with higher probability after seeing the ciphertext. This can be formalized as follows:

An encryption scheme is *perfectly secret* if there for every set of plaintexts, and for every strategy used by Eve, if we choose at random and a random key , then the probability that Eve guesses after seeing is at most .

In particular, if we encrypt either “Yes” or “No” with probability , then Eve won’t be able to guess which one it is with probability better than half. In fact, that turns out to be the heart of the matter:

An encryption scheme is perfectly secret if and only if for every two distinct plaintexts and every strategy used by Eve, if we choose at random and a random key , then the probability that Eve guesses after seeing is at most .

The “only if” direction is obvious— this condition is a special case of the perfect secrecy condition for a set of size .

The “if” direction is trickier. We need to show that if there is some set (of size possibly much larger than ) and some strategy for Eve to guess (based on the ciphertext) a plaintext chosen from with probability larger than , then there is also some set of size two and a strategy for Eve to guess a plaintext chosen from with probability larger than .

Let’s fix the message to be the all zeroes message and pick at random in . Under our assumption, it holds that for random key and message ,

On the other hand, for every choice of , is a fixed string independent on the choice of , and so if we pick at random in , then the probability that is at most , or in other words

Thus in particular, due to linearity of expectation, there *exists* some satisfying

(Can you see why? This is worthwhile stopping and reading again.) But this can be turned into an attacker such that for . the probability that is larger than . Indeed, we can define to output if and otherwise output a random message in . The probability that equals is higher when than when , and since outputs either or , this means that the probability that is larger than . (Can you see why?)

# 

The proof of twotomanythm is not trivial, and is worth reading again and making sure you understand it. An excellent exercise, which I urge you to pause and do now is to prove the following: is perfectly secret if for every plaintexts , the two random variables and (for randomly chosen keys and ) have precisely the same distribution.

Prove that a valid encryption scheme with plaintext length is perfectly secret if and only if for every and plaintexts , the following two distributions and over are identical:

* is obtained by sampling and outputting .
* is obtained by sampling and outputting .

We only sketch the proof. The condition in the exercise is equivalent to perfect secrecy with . For every , if and are identical then clearly for every , since these correspond applying on the same distribution . On the other hand, if and are not identical then there must exist some ciphertext such that (or vice versa). The adversary that on input will guess that is an encryption of if and otherwise will toss a coin will have some advantage over in distinguishing an encryption of from an encryption of .

We summarize the equivalent definitions of perfect secrecy in the following theorem, whose (omitted) proof follows from twotomanythm and perfectsecrecyequiv as well as similar proof ideas.

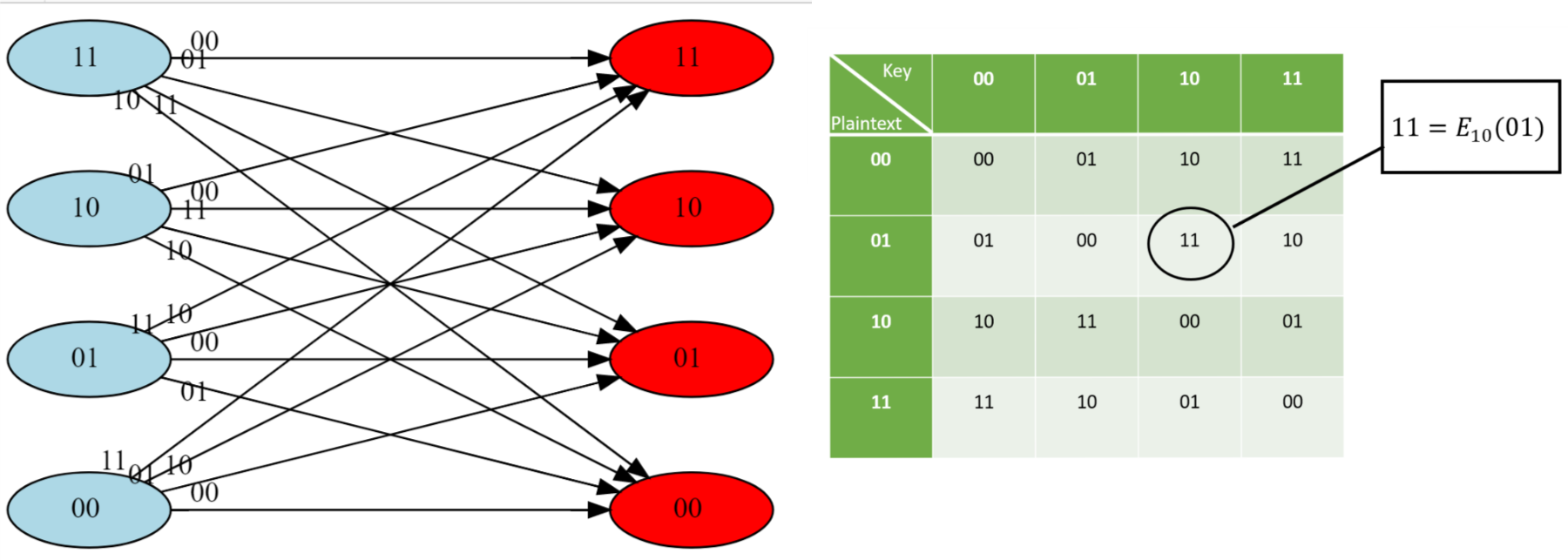
Let be a valid encryption scheme with message length . Then the following conditions are equivalent:

1. is perfectly secret as per perfectsecrecydef.
2. For every pair of messages , the distributions and are identical.
3. (Two-message security: Eve can’t guess which of one of two messages was encrypted with success better than half.) For every function and pair of messages ,
4. (Arbitrary prior security: Eve can’t guess which message was encrypted with success better than her prior information.) For every distribution over , and ,

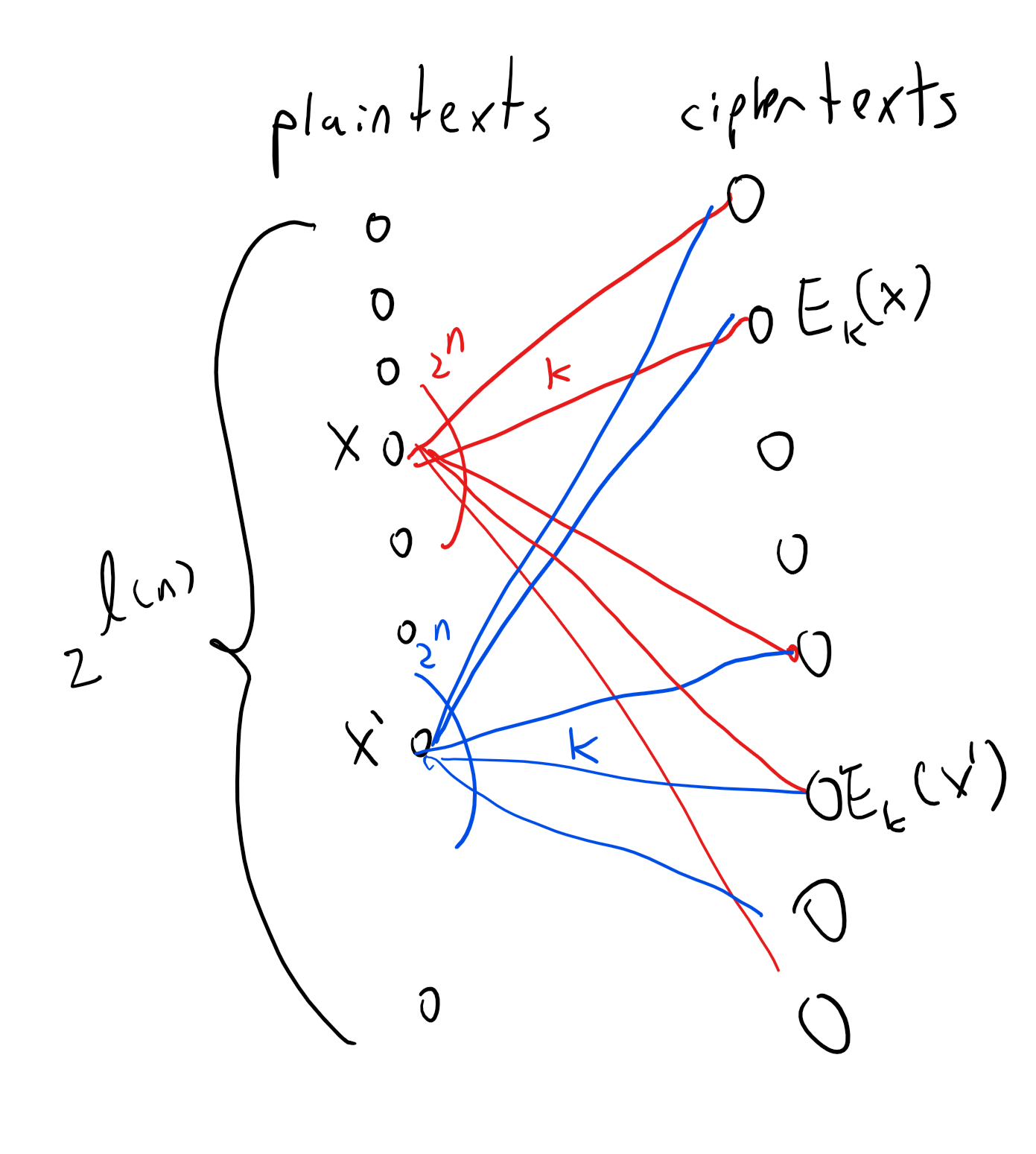
where we denote to be the largest probability of any element under .

### Achieving perfect secrecy

So, perfect secrecy is a natural condition, and does not seem to be too weak for applications, but can it actually be achieved? After all, the condition that two different plaintexts are mapped to the same distribution seems somewhat at odds with the condition that Bob would succeed in decrypting the ciphertexts and find out if the plaintext was in fact or . It turns out the answer is yes! For example, onetimepadtwofig details a perfectly secret encryption for two bits.



A perfectly secret encryption scheme for two-bit keys and messages. The blue vertices represent plaintexts and the red vertices represent ciphertexts, each edge mapping a plaintext to a ciphertext is labeled with the corresponding key . Since there are four possible keys, the degree of the graph is four and it is in fact a complete bipartite graph. The encryption scheme is valid in the sense that for every , the map is one-to-one, which in other words means that the set of edges labeled with is a *matching*.



For any key length , we can visualize an encryption scheme as a graph with a vertex for every one of the possible plaintexts and for every one of the ciphertexts in of the form for and . For every plaintext and key , we add an edge labeled between and . By the validity condition, if we pick any fixed key , the map must be one-to-one. The condition of perfect secrecy simply corresponds to requiring that every two plaintexts and have exactly the same set of neighbors (or multi-set, if there are parallel edges).

In fact, this can be generalized to any number of bits:[[6]](#footnote-71)

# 

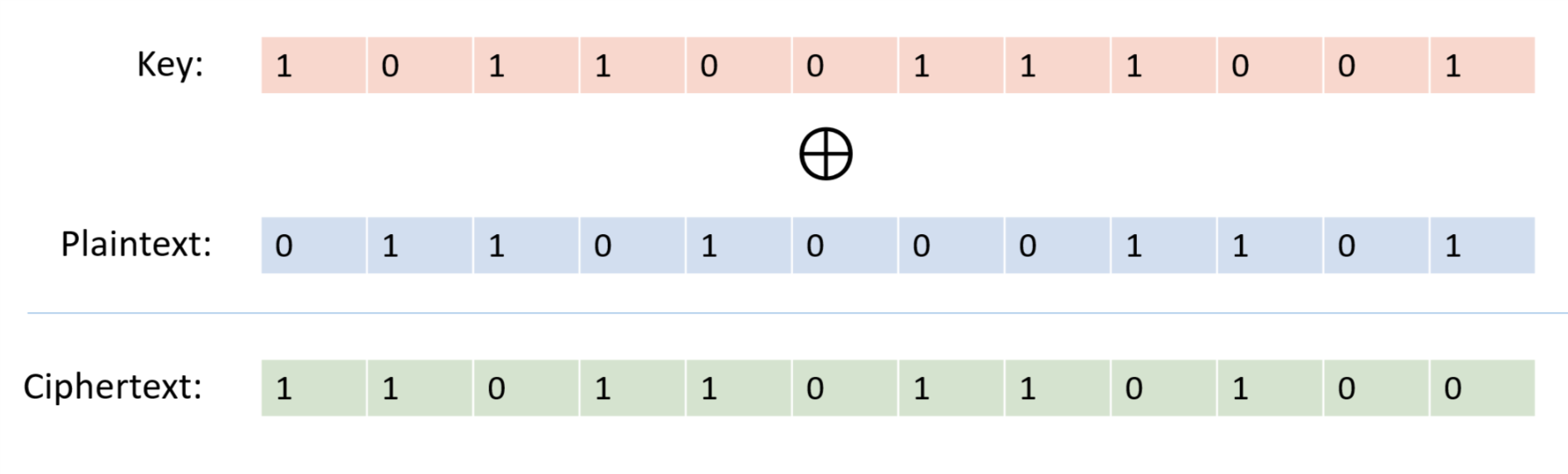
There is a perfectly secret valid encryption scheme with .

# 

Our scheme is the [one-time pad](https://en.wikipedia.org/wiki/One-time_pad) also known as the “Vernam Cipher”, see onetimepadfig. The encryption is exceedingly simple: to encrypt a message with a key we simply output where is the bitwise XOR operation that outputs the string corresponding to XORing each coordinate of and .

For two binary strings and of the same length , we define to be the string such that for every . The encryption scheme is defined as follows: and . By the associative law of addition (which works also modulo two), , using the fact that for every bit , and . Hence form a valid encryption.

To analyze the perfect secrecy property, we claim that for every , the distribution where is simply the uniform distribution over , and hence in particular the distributions and are identical for every . Indeed, for every particular , the value is output by if and only if which holds if and only if . Since is chosen uniformly at random in , the probability that happens to equal is exactly , which means that every string is output by with probability .



In the *one time pad* encryption scheme we encrypt a plaintext with a key by the ciphertext where denotes the bitwise XOR operation.

# 

The argument above is quite simple but is worth reading again. To understand why the one-time pad is perfectly secret, it is useful to envision it as a bipartite graph as we’ve done in onetimepadtwofig. (In fact the encryption scheme of onetimepadtwofig is precisely the one-time pad for .) For every , the one-time pad encryption scheme corresponds to a bipartite graph with vertices on the “left side” corresponding to the plaintexts in and vertices on the “right side” corresponding to the ciphertexts . For every and , we connect to the vertex with an edge that we label with . One can see that this is the complete bipartite graph, where every vertex on the left is connected to *all* vertices on the right. In particular this means that for every left vertex , the distribution on the ciphertexts obtained by taking a random and going to the neighbor of on the edge labeled is the uniform distribution over . This ensures the perfect secrecy condition.

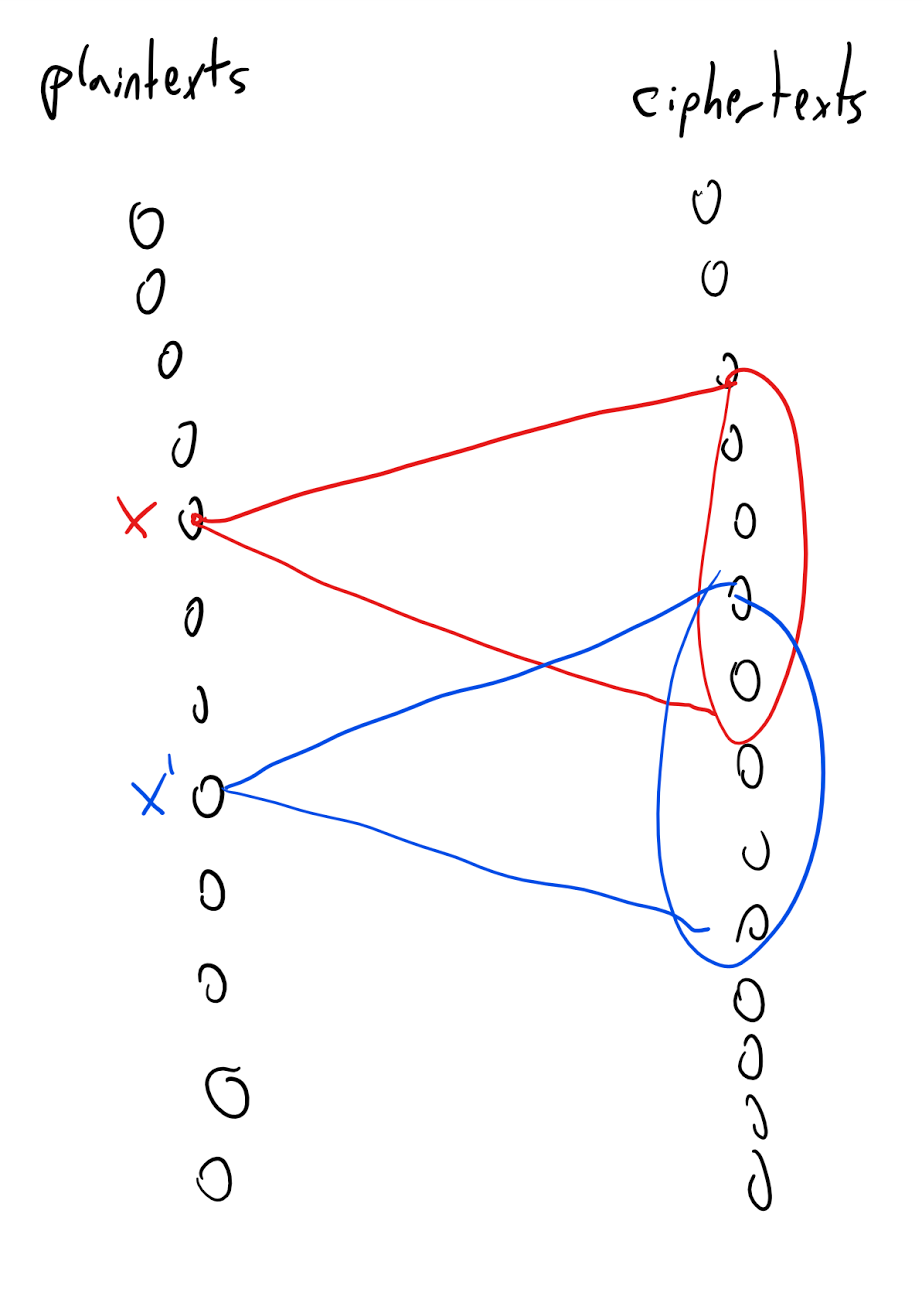
## Necessity of long keys

So, does onetimepad give the final word on cryptography, and means that we can all communicate with perfect secrecy and live happily ever after? No it doesn’t. While the one-time pad is efficient, and gives perfect secrecy, it has one glaring disadvantage: to communicate bits you need to store a key of length . In contrast, practically used cryptosystems such as AES-128 have a short key of bits (i.e., bytes) that can be used to protect terabytes or more of communication! Imagine that we all needed to use the one time pad. If that was the case, then if you had to communicate with people, you would have to maintain (securely!) huge files that are each as long as the length of the maximum total communication you expect with that person. Imagine that every time you opened an account with Amazon, Google, or any other service, they would need to send you in the mail (ideally with a secure courier) a DVD full of random numbers, and every time you suspected a virus, you’d need to ask all these services for a fresh DVD. This doesn’t sound so appealing.

This is not just a theoretical issue. The Soviets have used the one-time pad for their confidential communication since before the 1940’s. In fact, even before Shannon’s work, the U.S. intelligence already knew in 1941 that the one-time pad is in principle “unbreakable” (see page 32 in the [Venona document](http://nsarchive.gwu.edu/NSAEBB/NSAEBB278/01.PDF)). However, it turned out that the hassle of manufacturing so many keys for all the communication took its toll on the Soviets and they ended up reusing the same keys for more than one message. They did try to use them for completely different receivers in the (false) hope that this wouldn’t be detected. The [Venona Project](https://en.wikipedia.org/wiki/Venona_project) of the U.S. Army was founded in February 1943 by Gene Grabeel (see genegrabeelfig), a former home economics teacher from Madison Heights, Virgnia and Lt. Leonard Zubko. In October 1943, they had their breakthrough when it was discovered that the Russians were reusing their keys. In the 37 years of its existence, the project has resulted in a treasure chest of intelligence, exposing hundreds of KGB agents and Russian spies in the U.S. and other countries, including Julius Rosenberg, Harry Gold, Klaus Fuchs, Alger Hiss, Harry Dexter White and many others.



Gene Grabeel, who founded the U.S. Russian SigInt program on 1 Feb 1943. Photo taken in 1942, see Page 7 in the Venona historical study.



An encryption scheme where the number of keys is smaller than the number of plaintexts corresponds to a bipartite graph where the degree is smaller than the number of vertices on the left side. Together with the validity condition this implies that there will be two left vertices with non-identical neighborhoods, and hence the scheme does *not* satisfy perfect secrecy.

Unfortunately it turns out that that such long keys are *necessary* for perfect secrecy:

### 

For every perfectly secret encryption scheme the length function satisfies .

### 

The idea behind the proof is illustrated in longkeygraphfig. We define a graph between the plaintexts and ciphertexts, where we put an edge between plaintext and ciphertext if there is some key such that . The *degree* of this graph is at most the number of potential keys. The fact that the degree is smaller than the number of plaintexts (and hence of ciphertexts) implies that there would be two plaintexts and with different sets of neighbors, and hence the distribution of a ciphertext corresponding to (with a random key) will not be identical to the distribution of a ciphertext corresponding to .

Let be a valid encryption scheme with messages of length and key of length . We will show that is not perfectly secret by providing two plaintexts such that the distributions and are not identical, where is the distribution obtained by picking and outputting .

We choose . Let be the set of all ciphertexts that have nonzero probability of being output in . That is, . Since there are only keys, we know that .

We will show the following claim:

**Claim I:** There exists some and such that .

Claim I implies that the string has positive probability of being output by and zero probability of being output by and hence in particular and are not identical. To prove Claim I, just choose a fixed . By the validity condition, the map is a one to one map of to and hence in particular the *image* of this map which is the set has size at least (in fact exactly) . Since , this means that and so in particular there exists some string in . But by the definition of this means that there is some such that which concludes the proof of Claim I and hence of longkeysthm.

There is a sense in which both our secrecy and our impossibility results might not be fully convincing, and that is that we did not explicitly consider algorithms that use *randomness* . For example, maybe Eve can break a perfectly secret encryption if she is not modeled as a deterministic function but rather a *probabilistic* process. Similarly, maybe the encryption and decryption functions could be probabilistic processes as well. It turns out that none of those matter.

For the former, note that a probabilistic process can be thought of as a *distribution* over functions, in the sense that we have a collection of functions mapping to , and some probabilities (non-negative numbers summing to ), so we now think of Eve as selecting the function with probability . But if none of those functions can give an advantage better than , then neither can this collection (this is related to the *averaging principle* in probability).

A similar (though more involved) argument shows that the impossiblity result showing that the key must be at least as long as the message still holds even if the encryption and decryption algorithms are allowed to be probabilistic processes as well (working this out is a great exercise).

### Amplifying success probability

longkeysthm implies that for every encryption scheme with , there is a pair of messages and an attacker that can distinguish between an encryption of and an encryption of with success better than . But perhaps Eve’s success is only marginally better than half, say ? It turns out that’s not the case. If the message is even somewhat larger than the key, the success of Eve can be very close to :

Let be an encryption scheme with . Then there is a function and pair of messages such that

As in the proof of longkeysthm, let and let and be the set of size at most of all ciphertexts corresponding to . We claim that

We show this by arguing that this bound holds for every fixed , when we take the probability over , and so in particular it holds also for random . Indeed, for every fixed , the map is a one-to-one map, and so the distribution of for random is uniform over some set of size . For every , the probability over that is equal to

thus proving eqlongkeyprobproof.

Now, for every , define to be . By eqlongkeyprobproof, the expectation of over random is at most and so in particular by the averaging argument *there exists* some such that . Yet that means that the following adversary will be able to distinguish between an encryption of and an encryption of with probability at least :

* **Input:** A ciphertext
* **Operation:** If , output , otherwise output .

The probability that is equal to , while the probability that is equal to . Hence the overall probability of guessing correctly is

## Bibliographical notes

Much of this text is shared with [my Introduction to Theoretical Computer Science textbook](https://introtcs.org).

Shannon’s manuscript was written in 1945 but was classified, and a partial version was only published in 1949. Still it has revolutionized cryptography, and is the forerunner to much of what followed.

The Venona project’s history is described in [this document](http://nsarchive.gwu.edu/NSAEBB/NSAEBB278/01.PDF). Aside from Grabeel and Zubko, credit to the discovery that the Soviets were reusing keys is shared by Lt. Richard Hallock, Carrie Berry, Frank Lewis, and Lt. Karl Elmquist, and there are others that have made important contribution to this project. See pages 27 and 28 in the document.

In a [1955 letter to the NSA](https://www.nsa.gov/news-features/declassified-documents/nash-letters/assets/files/nash_letters1.pdf) that only recently came forward, John Nash proposed an “unbreakable” encryption scheme. He wrote *“I hope my handwriting, etc. do not give the impression I am just a crank or circle-squarer…. The significance of this conjecture [that certain encryption schemes are exponentially secure against key recovery attacks] .. is that it is quite feasible to design ciphers that are effectively unbreakable.”*. John Nash made seminal contributions in mathematics and game theory, and was awarded both the Abel Prize in mathematics and the Nobel Memorial Prize in Economic Sciences. However, he has struggled with mental illness throughout his life. His biography, [A Beautiful Mind](https://en.wikipedia.org/wiki/A_Beautiful_Mind_(book)) was made into a popular movie. It is natural to compare Nash’s 1955 letter to the NSA to the 1956 letter by [Kurt Gödel to John von Neumann](https://www.cs.cmu.edu/~aada/courses/15251s15/www/notes/godel-letter.pdf). From the theoretical computer science point of view, the crucial difference is that while Nash informally talks about exponential vs polynomial computation time, he does not mention the word “Turing Machine” or other models of computation, and it is not clear if he is aware or not that his conjecture can be made mathematically precise (assuming a formalization of “sufficiently complex types of enciphering”).

1. In the current state of these lecture notes, almost all references and credits are omitted unless the name has become standard in the literature, or I believe that the story of some discovery can serve a pedagogical point. See the Katz-Lindell book for historical notes and references. This lecture shares a lot of text with (though is not identical to) my lecture on cryptography in the [introduction to theoretical computer science](http://introtcs.org) lecture notes. [↑](#footnote-ref-21)
2. Traditionally, *cryptography* was the name for the activity of *making* codes, while *cryptoanalysis* is the name for the activity of *breaking* them, and *cryptology* is the name for the union of the two. These days *cryptography* is often used as the name for the broad science of constructing and analyzing the security of not just encryptions but many schemes and protocols for protecting the confidentiality and integrity of communication and computation. [↑](#footnote-ref-23)
3. Here is a nice exercise: compute (up to an order of magnitude) the probability that a 50-letter long message composed of random letters will end up not containing the letter “L”. [↑](#footnote-ref-37)
4. The actual quote is “Il faut qu’il n’exige pas le secret, et qu’il puisse sans inconvénient tomber entre les mains de l’ennemi” loosely translated as “The system must not require secrecy and can be stolen by the enemy without causing trouble”. According to Steve Bellovin the NSA version is “assume that the first copy of any device we make is shipped to the Kremlin”. [↑](#footnote-ref-45)
5. There are about atoms in the galaxy, so even if we assumed that each one of those atoms was a computer that can process say decryption attempts per second (as the speed of light is meters per second and the diameter of an atom is about meters), then it would still take seconds, which is about years to exhaust all possibilities, while the sun is estimated to burn out in about 5 billion years. [↑](#footnote-ref-54)
6. The one-time pad is typically credited to Gilbert Vernam of Bell and Joseph Mauborgne of the U.S. Army Signal Corps, but Steve Bellovin discovered an earlier inventor [Frank Miller](http://www.cs.columbia.edu/~CS4HS/talks/FrankMillerOneTimePad.pdf) who published a description of the one-time pad in 1882. However, it is unclear if Miller realized the fact that security of this system can be mathematically proven, and so theorem below should probably be still be credited to Vernam and Mauborgne. [↑](#footnote-ref-71)