THE CODE BOOK

The Science of Secrecy from Ancient Egypt to Quantum Cryptography

(Simon Singh)

Freshman Seminar, Winter 2006

February 28, 2006

Contents

1 January 26, 2006

Introduction

- "Detective stories or crossword puzzles cater for the majority; the solution of secret codes may be the pursuit of a few".
- It was the threat of enemy interception that motivated the development of codes and ciphers, the history of which is the story of the centuries-old battle between codemakers and codebreakers.
- Two main objectives: to chart the evolution of codes (including the impact on history and science), and to show that today it is more relevant than ever (privacy versus a police state and security of internet commerce). A code is constantly under attack from codebreakers. There is an analogy: codemaker vs. codebreaker; antibiotic vs. bacteria.
- First World War = the chemist's war (mustard gas, chlorine); Second World War = the physicist's war (atomic bomb); Third World War = the mathematician's war (information).
- There is another purpose for the science of cryptography besides disguising messages: to uncover the meaning of unintentionally indecipherable archeological texts.
- Some terminology: code (a word or phrase is replaced with a word, number, or symbol, e. g. codeword), cipher (each letter in a phrase is replaced by another letter, or number, or symbol), plaintext (the message), ciphertext (the encrypted message).
- The science of secrecy is largely a secret science (National Security Agency). Research is classified until it is no longer deemed helpful to adversaries.
- Is there a "quantum computer"?

1.1 Chapter 1—The Cipher of Mary Queen of Scots

- On trial for treason. Not for the first time, a life hung on the strength of a cipher. Mary was accused of plotting the assassination of Queen Elizabeth. Elizabeth needed proof of her complicity, otherwise she had reasons not to execute another queen, who was in fact a cousin. The proof rested on communications between the conspirators and Mary, which were encrypted. Unfortunately for Mary, the messages were intercepted and the code was broken.

1.1.1 The Evolution of Secret Writing

- The Greeks and Persians—fifth century B. C. is one story. Another is about a message written on a shaved head. This was clearly a period of history that tolerated a certain lack of urgency.
- Strategy was: Hiding the message (*steganography*). Fundamentally weak. Other examples: a swallowed silk ball, a hard-boiled egg, writing with invisible ink. This was used for 2000 years despite the obvious lack of security.
- Cryptography: Hiding the meaning (*encryption*). An example of a combination is the microdot. Cryptography was developed in parallel with steganography. It had the obvious advantage that without knowing the scrambling protocol, the enemy could not easily determine the message.
- Two branches of cryptography:

  transposition ("Rail fence" transposition—put letters of the plaintext alternately on two lines, then follow one line by the other in the ciphertext; Spartan scytale 500 B. C.—write a message on a belt and wrap it around a wooden staff of predetermined size).

substitution (mlecchita-vikalpā—the art of secret writing, one of 64 arts recommended for women in *Kāma-Sūtra*, this one to help them conceal the details of their liaisons; example: pair letters of the alphabet at random, then substitute each letter of the plaintext with its partner. In transposition, each letter retains its identity but changes its position, whereas in substitution each letter changes its identity but retains its position. The *Caesar (shift) cipher* is based on a cipher alphabet that is shifted a certain number of places (in Caesar's case three) relative to the plain alphabet. There are 25 distinct shift ciphers. If you allow the cipher alphabet to be any rearrangement of the plain alphabet then you have over 400,000,000,000,000,000,000,000,000 such distinct ciphers. Although this is impossible to break by using brute-force, it is not feasable because the key is not "simple."

- More terminology: plain alphabet, cipher alphabet, algorithm (=the general encrypting method), key (=the exact details of a particular encryption)
- Kerckhoffs' Principle: The security of a crypto-system must not depend on keeping secret the crypto-algorithm, but only on keeping secret the key.
- Large variety of keys: keyword, keyphrase. If a key can be committed to memory, it is less likely to fall into enemy hands. To use a keyphrase, begin by removing any spaces and repeated letters, then follow by the remaining letters of the alphabet, in their correct order.
- The substitution cipher dominated the art of secret writing throughout the first millennium A. D. The breakthrough which allowed these codes to be broken occurred in the East, and required a brilliant combination of linguistics, statistics, and religious devotion.

1.1.2 The Arab Cryptanalysts

- The year 750 heralded the golden age of Islamic civilization. The arts and sciences flourished in equal measure. The legacy of Islamic scientists is evident from the number of Arabic words that pepper the lexicon of modern science, such as *algebra, alkaline, zenith*. The social order relied on an effective system of administration, which in turn relied on secure communication achieved through the use of encryption.
- This *monoalphabetic substitution cipher* used symbols as well as letters.
- The Arab scholars invented *cryptanalysis*, the science of unscrambling a message without knowledge of the key. They cracked the monoalphabetic substitution cipher after several centuries of its successful use. This would not have been possible in a society until it had reached a sufficiently sophisticated level of scholarship in *mathematics, statistics, and linguistics*.
- The innocuous observation that some letters are more common than others in written documents would lead to the first great breakthrough in cryptanalysis. The method, called *frequency analysis* is described in a treatise by *Abu Yusuf Ya'qub ibn Is-haq ibn as-Sabbah ibn 'omran ibn Ismail al-Kindi* (let's call him al-Kindi for short) in the ninth century.
- Table 1 on page 19 give the relative frequency of each letter of the English alphabet and is based on newspaper articles and novels. *e* is the most common letter, followed by *t* and then *a*, and so on.
- In general, short texts are likely to deviate significantly from the standard frequencies, and if there are less than a hundred letters, then decipherment will be very difficult. A counterexample: *La Disparition*, and its translation into English (Appendix A, a 200-page novel that did not use any words containing the letter *e*!
- Besides logical thinking, frequency analysis requires guile, intuition, flexibility, and guesswork.

1.1.3 Cryptanalyzing a Ciphertext

- In this example (pp. 20–25), an unquestioning application of frequency analysis would lead to gibberish.
- Start with the most frequently occurring letters and determine if they might be vowels or consonants. Other cryptanalytic tips appear in Appendix B.
- Often, a cryptographer will remove all the spaces to make it harder for an enemy interceptor to unscramble the message. There is a technique to apply here: see p. 23.

1.1.4 Renaissance in the West

- Between 800 and 1200, while Arab scholars enjoyed a vigorous period of intellectual achievement, Europe was firmly stuck in the Dark Ages. The only European institutions to encourage the study of secret writing were the monasteries, where monks would study the Bible in search of hidden meanings. (*The Bible Code* 1997 Michael Drosnin—Appendix C).
- The Old Testament contained deliberate and obvious examples of cryptography (for example, *atbash*, a Hebrew substitution cipher—p. 26). Atbash and other similar biblical ciphers were probably intended only to add mystery, rather than to conceal meaning, but they were enough to spark an interest in serious cryptography.
- 13th Century—Cryptography was introduced into Western Civilization by European monks: *Epistle on the Secret Works of Art and the Nullity of Magic* Roger Bacon (English Franciscan monk)
- 14th Century—Used to keep scientific discoveries secret: *Treatise on the Astrolabe* Geoffrey Chaucer.
- 15th Century—The revival in the arts, sciences and scholarship during the Renaissance nurtured the capacity for cryptography, while an explosion in political machinations offered ample motivation for secret communication. Each state had a cipher office, and each ambassador had a cipher secretary.
- Simultaneously, the science of cryptanalysis was beginning to emerge in the West. Some of the characters, with anecdote.
  - Giovanni Soro—Venetian (and Vatican) cipher secretary. Soro might have been reluctant to point out the weaknesses of the Papal cipher, because this would only have encouraged the Vatican to switch to a more secure cipher, one that Soro might not have been able to break.
  - Philibert Babou—cryptanalyst to the King of France. Babou had a reputation for being incredibly persistent, working day and night and persevering for weeks on end in order to crack an intercepted message. The gave the king ample opportunity to carry on a long-term affair with his wife.
  - François Viète—another cryptanalyst to the King of France. He took particular pleasure in cracking Spanish ciphers. The astonished Spanish cryptographers went so far as to complain to the Vatican that Viéte was an "archfiend in league with the devil." The Pope, was had been reading Spanish ciphers for years, rejected the Spanish petition, and the Spanish cryptographers became the laughingstock of Europe.
- This was a period of transition, with cryptographers still relying on the monoalphabetic substitution cipher, while cryptanalysts were beginning to use frequency analysis to break it.
- Simple improvements to the security of the monoalphabetic cipher: *nulls* (symbols or letters that were not substitutes for actual letters, merely blanks that represented nothing); deliberately misspelled words, *codewords* (each word is represented by another word or symbol).
- Terminology revisited: *code* (substitution at the level of words), *cipher* (substitution at the level of letters), *encipher* (scramble a message using a cipher), *encode* (scramble a message using a code), *decipher* (unscrambling an enciphered message), *decode* (unscrambling an encoded message), *encrypt*

(refers to scrambling with respect to both codes and ciphers), *decrypt* (refers to unscrambling with respect to both codes and ciphers). Though technically not accurate, the word "codebreaking" might be used to describe a process that is really "cipher breaking."

- Two major defects of codes: compiling a codebook is a major task and carrying it around is a major inconvenience (essentially a dictionary); the consequences of having the codebook captured by the enemy are devastating.
- As a compromise, in the 16th century, cryptographers sometimes relied on a *nomenclator* (a system of encryption that relies on a cipher alphabet, and a limited list of codewords). In effect, a nomenclator is not much more secure than a straightforward cipher because the bulk of the message can be deciphered using frequency analysis, and the remaining encoded words can be guessed from the context.
- The best cryptanalysts were capable of coping with the nomenclator, dealing with badly spelled messages, and the presence of nulls, and thus were able to break the majority of encrypted messages. Their skills influenced the decisions of their masters and mistresses, thereby affecting Europe's history at critical moments.

1.1.5 The Babington Plot

- Mary Queen of Scots was one of the most significant figures of the sixteenth century—Queen of Scotland, Queen of France, pretender to the Engllish throne—yet her fate would be decided by a slip of paper, the message it bore, and whether or not that message could be deciphered.
- The story of the Plot is detailed on pp. 32–44. Mary was jailed by her cousin Queen Elizabeth. Unknown to Mary, a plan to rescue her, under the leadership of Anthony Babington, was hatched in the taverns of London. The plan also included the assassination of the Queen. The plan could not proceed without Mary's blessing, thus the need to communicate with her. Babington's letter used a nomenclator consisting of 23 symbols as substitutes for the alphabet, 35 symbols representing words or phrases, four nulls, and a symbol signifying that the next symbol represented a double letter. One of Mary's replies, in which she agreed with the plot, effectively became her death sentence when it was deciphered by the Queen's cipher secretary.

1.2 Chapter 2—Le Chiffre Indéchiffrable

- The tragic execution of Mary Queen of Scots was a dramatic illustration of the weaknesses of monoalphabetic substitution, and in the battle between cryptographers and cryptanalysts it was the latter that had gained the upper hand. The onus was on the former to concoct a new, stronger cipher.
- Sometime in the 1460s, the Florentine polymath Leon Battista Alberti proposed using two or more cipher alphabets (switching between them during encipherment). The method was developed further by others, most notably, Blaise de Vigenère, a French diplomat.
- The Vigenère cipher uses 26 alphabets, each shifted by one letter. The algorithm to determine which alphabet (row in a *Vigenère square*) is used for which letters in the message is given by a *keyword*. If the keyword has five letters, the message is encrypted by cycling through five rows of the Vigen‘ere square.
- As well as being invulnerable to frequency analysis, the Vigen‘ere cipher has an enormous number of keys. The method was published in a treatise in 1586, the same year that the cipher of Mary Queen of Scots was broken. If only Mary's secretary had known of the treatise, then Mary's messages would have resisted decipherment, and Mary's life might have been spared.
- Despite its great potential, the Vigenère cipher remained largely neglected for two centuries, possibly because of the additional effort required because of its polyalphabetic nature.

1.2.1 From Shunning Vigenère to the Man in the Iron Mask

- Monoalphabetic ciphers were quick, easy to use, and secure against people unschooled in cryptanalysis. For military purposes something better was needed, but there was reluctance to adopt the polyalphabetic cipher because of its complexity. Consequently, an intermediate cipher was sought.
- One candidate was the *homophonic substitution cipher*. Each letter is replaced by a variety of substitutes, the number of potential substitutes being proportional to the frequency of the letter. The effect is to balance out the frequencies. If no symbol appears more frequently than any other, this is a perfect security system, right? Wrong! Each letter in the English language has its own personality, defined according to its relationship with all the other letters, and these traits can still be discerned even if the encryption is by homophonic substitution. Although the homophonic cipher is breakable, it is much more secure than a straightforward monoalphabetic cipher.
- Counter-intuitively, the homophonic substitution cipher is closer to a monoalphabetic cipher, albeit an enhanced one. Indeed, for example, the letter *a* can be represented by eight numbers, but each of these numbers represents only one plaintext letter. In a polyalphabetic cipher, a plaintext letter will also be represented by different symbols, and even more confusingly, these symbols will represent different letters during the course of an encipherment.
- By tweaking the basic monoalphabetic cipher in various ways, such as adding homophones, it became possible to encrypt messages securely, without having to resort to the complexities of the polyalphabetic cipher. An example of this was the great cipher of Louis XIV, which held for two centuries in regards to a message about the Man in the Iron Mask.

2 February 2, 2006

- The Great Cipher was so secure that it defied the efforts of all enemy cryptanalysts attempting to steal French secrets. Historians knew that the papers encrypted by the Great Cipher would offer a unique insight into the intrigues of sevententh-century France. Unfortunately, its details vanished with its inventors (Antoine and Bonaventure Rossignol, late 1600s).
- When Étienne Bazeries broke this code by a three-year effort in the 1890s, he became the first person for 200 years to witness the secrets of Louis XIV; in particular, the identity of the Man in the Iron Mask. How was this broken? The encrypted letters contained thousands of numbers, but only 587 different ones. After months of painstaking effort, Bazeries concluded that the Great Cipher was not a homophonic cipher. There are 676 possible *digraphs*, or pairs of letters. The proximity of 676 to 587 suggested an approach (frequency analysis at the level of pairs of letters) which however also failed after months of work. However, upon reflecting on digraphs, Bazeries considered the possibility that the numbers (symbols) represented not pairs of letters, but syllables. This proved to be a crucial breakthrough. Once you know a part of the plaintext, you can extrapolate, as in solving crossword puzzles, and fill in other parts of the puzzle.
- The Mask has been the subject of much speculation ever since he was first imprisoned. The most popular conspiracy theory relating to the Mask suggests that he was the twin of Louis XIV, condemned to imprisonment in order to avoid any controversy over who was the rightful heir to the throne. This theory persisted despite the evidence revealed in one of Bazeries's decipherments.

2.0.2 The Black Chambers

- By the 1700s, cryptanalysis was becoming industrialized and each European power had its own Black Chamber, a nerve center for deciphering messages and gathering intelligence. The most celebrated, disciplined, and efficient was in Vienna. Letters were copied first in order not to interrupt the smooth running of the postal service

- The Black Chambers were effectively making all forms of monoalphabetic cipher insecure, forcing cryptographers to adopt the more complex but more secure Vigenère cipher. In addition to the more effective cryptanalysis, the development of the telegraph also encouraged the move toward securer forms of encryption.
- As revolutionary as Morse code was for sending messages, guarding them was an issue. The code itself is not a form of cryptography, because there is no concealment of the message. The obvious solution was to encipher the message before giving it to the telegraph operator.
- The polyalphabetic Vigenère cipher was clearly the best way to ensure secrecy for important business communicaitons. It was considered unbreakable, and became known as *le chiffre indéchiffrable*. Cryptographers had, for the time being at least, a clear lead over the cryptanalysts. (c. 1850)

2.0.3 Mr. Babbage Versus the Vigenère Cipher

- Charles Babbage was an eccentric British genius and one of the most intriguing figures of the 19th century. Among his many contributions were efficiency in setting postal rates and providing a template for modern computers.
- Babbage was fascinated with the idea of building a machine of faultlessly calculating mathematical tables used for astronomical, engineering and navigational calculations, to a high degree of accuracy. He designed such a machine in 1823 but although he was a brilliant innovatgor, he was not a great implementer, which resulted in loss of government funding. The scientific tragedy here was that Babbage's machine would have been a stepping stone to the "Analytical Engine," which would have been programmable, hence a template for modern computers.
- Babbage broke the Vigenère cipher, the greatest breakthrough in cryptanalysis since frequency analysis in the ninth century. An example of his method is given on pp. 69–77.
- It is possible that the reason Babbage did not publicize his method (1854), which came to be known as the Kasiski Test, after the person who independently discovered the method a decade later, was not necessarily that he was prone to leave projects undone and not publishing his discoveries, but rather because of the tradition of hushing up codebreaking achievements in the interests of national security, a practice that continues to this day.

2.0.4 From Agony Columns to Buried Treasure

- Thanks to the breakthroughs by Babbage and Kasiski, cryptanalysts regained control in the communications war. During the last half of the 19th century, professional cryptography was in disarray. Ironically, this same period witnessed an enormous growth of interest in ciphers among the general public. The ciphers used by the general public would not have withstood attack by a professional cryptanalyst, but they were sufficient to guard against the casual snooper.
- Examples of the interest in ciphers among the general public: "agony columns" (lovers sending encrypted forbidden messages to each other via the personal columns of newspapers); pinprick encryption (conveying a message by pricking tine holes under particular letters in an otherwise innocuous page of text—newspapers were free of postage in 1800s Britain); cryptographic fiction where ciphers played key roles: (Jules Verne—*Journey to the Center of the Earth*; Sir Arthur Conan Doyle—*The Adventure of the Dancing Men* (Sherlock Holmes); Edgar Allan Poe—*The Gold Bug*). The Gold Bug was "realized" in the 19th century by the case of the *Beale papers* (pp. 82–98), which involved Wild West escapades, a cowboy who amassed a vast fortune, a buried treasure worth $20 million and a mysterious set of encrypted papers describing its whereabouts.
- There were three Beale ciphers (pp. 87-88-89), each containing an array of numbers. The first one described the location of the buried treasure, the second outlined the contents of the treasure, and the third listed the relatives who should receive a share of the treasure.

- One of the possible ciphers used for the Beale papers is a so-called *book cipher*, in which a book, or any other piece of text, is itself the key. The words of the text are numbered and then any number is associated with the first letter of the corresponding word. Some letters would have more than one number associated with them and some letters might not have any number associated with them (if the text is short, for example). To crack this cipher, one would have to test all known books (a formidable task nowadays). The key for the second Beale cipher was determined to be the *Declaration of Independence*, revealing the worth of the treasure to be $20 million at today's bullion prices.

- Despite the efforts by the "Beale Cypher and Treasure Association," amateur treasure hunters, and professional cryptanalysts, the first and third Beale ciphers have remained a mystery for over a century, and the gold, silver, and jewels have yet to be found.

- The first cipher contains numbers as high as 2906 whereas the Declaration of Independence contains only 1322 words. How did Beale come up with something that is so formidable during a period in which the codebreakers were ahead? Beale may have created a special keytext (an original essay on the subject of buffalo hunting, for example). This is much more secure than a text based on a published book, but highly impractical. Other possible explanations include sabotage and the possibility that the treasure was found long ago and spirited away without being spotted by local residents. Carrying this to an extreme, it could be that NSA ("Never Say Anything") has the treasure. One cannot exclude the possibility that the Beale papers are a hoax. There is evidence to support both the fabrication of the whole story as well as the probity of the ciphers. (Dictionary.com: probity PRO-bih-tee noun: Tried virtue or integrity; approved moral excellence; honesty; uprightness.)

- Advice from the author of the 1885 pamphlet concerning the whole affair: *... devote only such time as can be spared from your legitimate business to the task, and if you can spare no time, then let the matter alone.... Never, as I have done, sacrifice your own and your family's interest to what may prove an illusion; but as I have already said, when your day's work is done, and you are comfortably seated by your good fire, a short time devoted to the subject can injure no one, and may bring its reward.*

2.1 Chapter 3—The Mechanization of Secrecy

- At the end of the 19th century, cryptography was in disarray. A new cipher was sought what could exploit the immediacy of the telegraph. Enter Marconi, the inventor of the (wireless) radio. Furthermore, Marconi also shattered the myth that radio transmission was limited by the horizon.

- The all-pervasive property of radio (emanate in all directions and reach receivers wherever they may be) was also its greatest weakness. If the enemy were going to be able to intercept every radio message, then cryptographers had to find a way of preventing them from deciphering these messages. Together, the advent of radio and the outbreak of the First World War intensified the need for effective encryption. However, the period 1914–1918 was a catalogue of cryptographic failures.

- One of the most famous (infamous?) wartime ciphers was the German ADFGVX cipher of 1918, a mixture of a substitution and transposition (Appendix F). This was broken by a Frenchman Georges Painvin (who lost 15 kilograms of weight in the process), leading to the defeat of the German army in a battle in which the element of surprise was lost.

- Despite the success of the cryptanalysts during the First World War, there was the problem of dealing with the sheer volume of radio traffic. Of all the wartime cryptanalysts, the French were the most effective, having learned of its value from the humiliating French defeat in the Franco-Prussian war in 1870. Auguste Kerckhoffs, of Dutch origin, spent most of his life in France and during this time wrote his treatise *La Cryptographie militaire*, which provided the French with an exceptional guide to the principles of cryptanalysis.

- In sharp contrast to the French (and British and Americans who also made important contributions to Allied cryptanalysis), the Germans were rather late in developing deciphering skills. This supremacy

of the Allied codebreakers and their influence on the First World War are best illustrated by the decipherment of a German telegram that was intercepted by the British on January 17, 1917. This decipherment showed how cryptanalysis can affect the course of war at the very highest level, and demonstrates the potentially devastating repercussions of employing inadequate encryption. This decipherment forced America to rethink its policy of neutrality, thereby shifting the balance of the war. The details of this story are given on pp. 107–115.

2.1.1 The Holy Grail of Cryptography

- Towards the end of the First World War, when cryptographers were in a state of utter despair, scientists in America made an astounding breakthrough. The fundamental weakness of the Vigenère cipher is its cyclical nature. The decipherment can be done by considering as many monoalphabetic ciphers as there are letters in the keyword. Thus, the longer the keyword (or keyphrase), the more difficult will be the decipherment, and the cryptanalytic technique developed by Babbage and Kasiski will not work. However, a key that is as long as the message is not sufficient to guarantee security. (An example is given on pp. 116-119.)
- The reason the decipherment in this example was possible is that the key consisted of meaningful words. In 1918, cryptographers began experimenting with keys that were devoid of structure, that is, random. The result was a theoretically unbreakable cipher, the *onetime pad cipher*.
- Two identical thick pads are compiled consisting of hundreds of sheets of paper, each sheet bearing a unique key in the form of lines of randomly sequenced letters. Each key is used only once, then destroyed by both sender and intended receiver.
- The first hurdle for the cryptanalyst is that there is no repetition in a random key because the cyclical aspect is lost. Hence the method of Babbage and Kasiski cannot break the onetime pad cipher. Also, the number of possible keys is so large that it is humanly or mechanically infeasible for the cryptanalyst to consider an exhaustive search of all possible keys. Rubbing salt in the wound, even if somehow all keys could be tested, the result would be that one and the same ciphertext could generate more than one message which was meaningful. (See the example on p. 122.)
- It can be proved mathematically that it is impossible for a cryptanalyst to crack a message encrypted with a ontime pad cipher. Hence, the onetime pad cipher was considered the "Holy Grail of cryptography."
- From Wikipedia: In Christian mythology, the Holy Grail was the dish, plate, cup or vessel used by Jesus at the Last Supper, said to possess miraculous powers.
- From Wikipedia: *Holy Grail as a Casual metaphor*: The legend of the Holy Grail is the basis of the use of the term holy grail in modern-day culture. This or that "holy grail" is seen as the distant, all-but-unobtainable ultimate goal for a person, organization, or field to achieve. For instance, cold fusion or anti-gravity devices are sometimes characterized as the "holy grail" of applied physics.
- From Wikipedia: Examples of "Holy Grails."

  Art Something new

  Astronomy Detection of a pulsar in orbit around a black hole

  Biology Biological immortality in humans; Assembling a cell from scratch using simple precursor materials (either manually or via random seeding); A complete understanding of the genetic code; Accurately predicting the three-dimensional structure of a protein from the sequence of its amino acids; A leafy plant that can absorb glucose (or a similar energy source) through its roots and that can use that nutrient for growth.

  Medicine A cure or treatment for all forms of Cancer; A cure or vaccine for AIDS; A cure or vaccine for the common cold.

Computer science Artificial consciousness; A polynomial time algorithm for NP-complete problems, or a proof of its nonexistence.

Cosmology The precise Age of the Universe

Economics Determining the cause of the Great Depression; Solving the equity premium puzzle; Creating an equivalent of Black-Scholes for futures contract pricing.

Historical linguistics Reconstruction of the Proto-World language; Deciphering Linear A

Mathematics A Unified theory of mathematics; **A PROOF OF THE RIEMANN HYPOTHESIS**

Physics A theory of everything; A unified field theory

Applied physics Anti-gravity devices; Efficient cold or controlled fusion; Faster-than-light travel or communication; Time travel; Room temperature superconductors.

High-energy physics The observation of a Higgs particle; The observation of a Magnetic Monopole.

Optics A self-focusing lens that employs a principle functionally similar to the optical Kerr effect but that works with the input of low intensity light.

Unsorted Detection of Extraterrestrial life .

- Although it is perfect in theory, the onetime pad cipher is flawed in practice, suffering from two defects. Supplying a large number of random keys is an immense task, primarily because it must be guaranteed to be random. Truly random keys can be created by harnessing natural physical processes, such as radioactivity, but this is impractical for day-to-day cryptography. The second problem is the logistical difficulty of distributing the onetime pads. You cannot cut down on the manufacture and distribution of keys by reusing onetime pads, as explained in Appendix G.

3 February 9, 2006

- The next breakthrough in the development of strengthened ciphers involved exploiting the very latest technology to produce machines to scramble messages.

3.0.2 The Development of Cipher Machines—from Cipher Disks to the Enigma

- Even though the cipher disk is a very basic device, it does ease encipherment, and it endured for five centuries (15th to 19th). It was used in the American Civil War and was popularized in the early 20th century radio drama Captain Midnight[1].

- The cipher disk's inventor, Leon Alberti, suggested changing the setting of the disk during the message (by using a keyword à la Vigenère), which in effect generates a polyalphabetic cipher instead of a monoalphabetic one.

- In 1918, Arthur Scherbius set out to replace the inadequate systems of (German) cryptography used in the First World War and developed an electrical version of the cipher disk which he called *Enigma*. Enigma consisted of three components: a keyboard for inputting plaintext letters, a scrambling unit that encrypts each plaintext letter, and a display board indicating the corresponding ciphertext letter.

- Instead of implementing a monoalphabetic cipher, this basic construction was enhanced in several ways: *more than one scrambler unit, rotation of the scramblers, a reflector, rearrangement of the scramblers, a plugboard.*

- Each time a letter is encrypted, the first scrambler rotates by one space. The second scrambler moves only after the first one has made a complete revolution. (Think of how an odometer works.) With a full alphabet of 26 letters, two scramblers effectively use 676 cipher alphabets (A third scrambler

[1] One of my childhood heros, along with Captain Marvel and Flash Gordon

increases this to 17,576). All of this is done with great efficiency and accuracy, thanks to the automatic movement of scramblers and the speed of electricity.

- The *initial settings*, dictated by a codebook, provide the keys for encrypted messages. In order to decipher the message, the receiver needs to have another (identically wired) Enigma machine and a copy of the codebook that contains the initial scrambler settings for that day. The reflector enables easy decipherment: the sender types in the plaintext to generate the ciphertext, and the receiver types in the ciphertext to generate the plaintext.

- It is quite possible that the enemy may capture an Enigma machine, but without knowing the initial settings, they cannot easily decrypt an intercepted message. On the other hand, one person working day and night could check 17,576 possible initial settings about 12 days. Hence the need for the enhancements listed above.

- The plugboard, which has by far the greatest effect on increasing the number of keys compared to having several scramblers and rearranging them, swaps six pairs of the 26 letters before they enter the first scrambler. Needless to say, this puts the number of possible keys out of the reach of a trial and error approach. Despite the relatively small contributions to the number of keys from the scramblers, they are still necessary since the plugboard on its own could be deciphered by frequency analysis.

- Although Enigma appeared to be unbreakable, its initial marketing attempts for business and military fell on deaf ears. Reasons for this were its high cost and the lack of enthusiasum on the part of the German military, which was oblivious to the damage caused by their insecure ciphers during the First World War. Three other inventors in three other countries (Netherlands, Sweden, USA), who also had the idea of a cipher machine based on rotating scramblers, saw their efforts come to nil.

- In the mid-1920s the mood in America was changing from paranoia to openness. By the time Herbert Hoover became president, the American Black Chamber (led by the flamboyant Herbert Yardley) was disbanded as part of Hoover's attempt to usher in a new era of trust in international affairs. At the same time, the German military were shocked into appreciating the value of Enigma, thanks to the publication of two British documents in 1923, Winston Churchill's (*The World Crisis*), and the official history of the First World War, by the British Royal Navy, both praising the proud achievements of British Intelligence.

- At the dawn of the Second World War, Scherbius's invention provided the German military with the most secure system of cryptography in the world. At times, it seemed that the Enigma machine would play a vital role in ensuring Nazi victory, but instead it was ultimately part of Hitler's downfall.

3.1 Chapter 4—Cracking the Enigma

- In the wake of their successes in the First World War, the Allies no longer feared anyone and seemed to lose their cryptanalytic zeal. In sharp contrast, Poland, sandwiched between two enemies (Russia and Germany), was desperate for intelligence information, so they formed a new cipher bureau (called *Biuro Szyfrów*) just after the First World War. *If necessity is the mother of invention, then perhaps adversity is the mother of cryptanalysis.* The Poles were successful until 1926 when they encountered Enigma.

- The first step toward breaking the Enigma cipher was due to a disaffected German, Hans-Thilo Schmidt, who sold documents containing Enigma information (for profit and to spite the country that he felt betrayed him and the brother he was jealous of) to a French secret agent, who then turned it over to the Poles under a prior military agreement they had with the French.

- The strength of the Enigma cipher depends not on keeping the machine secret, but on keeping the initial setting secret (recall that this is Kerckhoffs' principle). Hence the French cryptographers at the Bureau du Chiffre, who in the wake of the First World War suffered from overconfidence and lack of motivation, did not even bother trying to build a replica of the military Enigma machine. They

were convinced that achieving the next stage, finding the key required to decipher a particular Enigma message, was impossible.

- As well as revealing the internal wirings of the scramblers, Schmidt's documents also explained in detail the layout of the codebooks used by the Germans. On the first day of each month, the codebook would specify the *day key* consisting of (1) the plugboard settings, (2) the scrambler arrangement, and (3) the scrambler orientation.

- The process is reasonably secure, but it is weakened by the repeated use of a single day key to encrypt the hundreds of messages that might be sent each day. Therefore, as an extra precaution, the Germans took the step of using the day key settings to transmit a new *message key* for each message. The message keys would have the same plugboard settings and scrambler arrangement as the day key, but different scrambler orientations. The new scrambler orientation for each message was transmitted at the beginning of the message and enciphered according to the day key for that day. The message key is entered twice, just to provide a double-check for the receiver.

- To belabor the point, instead of using the single main cipher key to encrypt every message on a particular day, it is used to merely encrypt a new cipher key for each message. and then the actual message is encrypted according to the new cipher key.

- At first sight the system seemed to be impregnable, but the Polish cryptanalysts were undaunted. The Bureau Szyfrów organized a course on cryptography and invited twenty MATHEMATICIANS (each sworn to an oath of secrecy) to participate, the most gifted of which was Marian Rejewski, age 23.

- Rejewski's strategy for attacking Enigma focused on the fact that repetition is the enemy of security: *repetition leads to patterns, and cryptanalysts thrive on patterns.* The most obvious repetition in the Enigma encryption was in the message key, which was demanded in order to avoid mistakes caused by radio interference or operator error. But they did not see that this woud jeopardize the security of the machine.

- As each new message is intercepted, it is possible to identify relationships between the 1st and 4th letters of the repeated message key. All these relationships are reflections of the initial setting of the Enigma machine. If enough messages are accessed in a single day, then a complete alphabet of relationships between the 1st and 4th letters of each message could be constructed (see p. 151). The next question was whether there existed any way of determining the day key by looking at this table of relationships.

- Eventually, Rejewski began to look for one particular type of pattern in the relationship, which featured chains (I would call them cycles since they end with the same letter they started with) of letters (see p. 152). Rejewski then observed that the number of links in the chains is purely a consequence of the scrambler settings. The total number of possible scrambler settings is 105,456 (=6 X 17,576): which of these scrambler settings was associated with the numbers of links within a set of chains? The task had become one hundred billion times easier, certainly within the realm of human endeavor.

- It took Rejewski and his team an entire year to catalogue the chain lengths that were generated by each of the 105,456 possible scrambler settings. However, once this was completed, it was finally possible to begin unraveling the Enigma cipher. Once the first six letters of enough intercepted messages were analyzed to determine the table of relationships, Rejewski could then go to his catalogue and identify the scrambler part of the day key.

- There remained the task of determining the hundred billion possibilities for the plugboard settings, which surprisingly turned out to be straightforward as follows: Set the Enigma machine according to the newly established scrambler part of the day key and remove all cables from the plugboard; then type a piece of intercepted ciphertext into the Enigma, and repeat until some vaguely recognizable phrases occur, which will imply the swapping of certain letters and the non-swapping of certain other letters. This is repeated until all the plugboard settings, and therefore the day key, is determined.

- Following Rejewski's breakthrough, German communications became transparent. Rejewski's attack on Enigma is one of the truly great accomplishments of cryptanalysis. The Polish success in breaking the Enigma cipher can be attributed to three factors: fear, MATHEMATICS, and espionage.

- Even when the Germans made a minor alteration to the way they transmitted messages, Rejewski fought back. Rather than rewriting his catalogue, he devised a mechanized version of his cataloguing system, which could automatically search for the correct scrambler settings. The units, consisting of six Enigma machines working in parallel, and called *bombes*, were capable of finding the day key in two hours. The bombes effectively mechanized the process of decipherment, and were therefore a natural response to Enigma, which was a mechanization of encipherment.

- Now we encounter some intrigue. It turns out that for most of the 1930s, the chief of the Biuro, Major Gwido Langer, already had the Enigma day keys (which he obtained from Hans-Thilo Schmidt), but he withheld them from Rejewski believing he was preparing him for the inevitable time when the keys would no longer be available. Langer thought that Rejewski should practice self-sufficiency in peacetime, as preparation for what lay ahead.

- In December 1938, Rejewski's skills reached their limits, when German cryptographers enhanced Enigma's security by adding two scramblers and four more plugboard cables. To add insult to injury, by this time, the keys were no longer being provided to Langer by Schmidt. For seven years (1931–1938), Schmidt had supplied keys which were superfluous because of Polish innovation. Now, just when the Poles needed the keys, they were no longer available.

- Langer was determined that if Poland was invaded, then its cryptanalytic breakthroughs, which had so far been kept secret from the Allies, should not be lost. He passed them on to the French and British, who with their extra resources, might fully exploit the concept of the bombe. The French were particularly astonished, because the Polish work had been based on the results of French espionage. The French had handed the information from Schmidt to the Poles because they believed it to be of no value, but the Poles had proved them wrong.

4 February 16, 2006

4.0.1 The Geese that Never Cackled

- The Polish breakthroughs, despite the recent setback, boosted the morale of Allied cryptographers and demonstrated the value of employing MATHEMATICIANS. In anticipation of a deluge of encrypted intercepts as soon as the war started, the British Government Code and Cypher School (GC&CS) was relocated to larger facilities at Bletchley Park. Within five years, the staff grew from 200 to 7000.

- During the autumn of 1939, the scientists and mathematicians at Bletchley learned the intricacies of the Enigma cipher and rapidly mastered the Polish techniques. Although it took only a few hours to discover the Enigma settings for the day, this information became obsolete less than 24 hours later. Bletchley decipherments were of the utmost importance in, for example, the German invasion of Denmark and Norway in April 1940, the Battle of Britain, and the state of the German Air Force.

- The Bletchley cryptanalysts developed their own techniques for breaking Enigma: *cillies* (obvious message keys, for example three consecutive letters on the keyboard); repeated use of the same message keys; excluding certain scrambler arrangements to avoid a scrambler remaining in the same position meant that the codebook compilers reduced by half the number of possible scrambler arrangements; no swaps were being made between neighboring letters, again reducing the number of possible keys.

- Because the Enigma machine continued to evolve during the course of the war, the Bletchley cryptanalysts were continually forced to innovate, and to devise wholly new strategies. There were many great cryptanalysts and many significant breakthroughs, but it was Alan Turing who identified Enigma's greatest weakness and ruthlessly exploited it.

- Turing's way of coming to terms with the death of his friend and colleague was to focus on his scientific studies. Turing arrived at King's College, in Cambridge, during a period of intense debate about the nature of mathematics and logic, centering around the issue of *undecidability*, a notion developed by logician Kurt Gödel, who demonstrated that there are questions which were beyond the reach of logical proof. Turing made a significant contribution to this area in 1937 in an unsuccessful attempt to subdue Gödel's monster, but in the process, his *universal Turing machine* (a reincarnation of Babbage's Difference Engine No. 2) provided the blueprint for the modern programmable computer.

- Turing enjoyed a very satisfying academic and social life at King's College. However, in 1939, his academic career was brought to an abrupt halt when the GS&CS invited him to become a cryptanalyst at Bletchley. His first assignment was to find an alternative way to attack Enigma, one that did not rely on a repeated message key.

- When a piece of plaintext can be associated with a piece of ciphertext, this combination is known as a *crib*. From the vast library of decrypted messages which was accumulating at Bletchley, Turing was convinced that he could exploit the cribs to attack Enigma. The direct approach, of simply entering the crib and seeing if the correct ciphertext appeared, had the obvious disadvantage of being a trial and error and approach and there were too many possible settings to check. Following Rejewski's strategy, Turing proceeded by disentangling the settings, reducing the number of (preliminary) possibilities to a mere 1,054,560!

- Turing's method of breaking Enigma is described on pp. 171–174. The combination of crib, loops, and electrically connected machines resulted in a remarkable piece of cryptanalysis, and only Turing, with his unique background in mathematical machines, could ever have come up with it. His musings on the imaginary Turing machines were intended to answer esoteric questions about mathematical undecidability, but this purely academic research had put him in the right frame of mind for designing a practical machine capable of solving very real problems.

- Everything at the Government Code and Cypher School was top secret, so nobody outside of Bletchley Park was aware of Turing's remarkable achievement. If everything was going well, one of Turing's bombes might find an Enigma key within an hour. This is not to say that decipherment had become a formality. There were many hurdles to overcome before the bombes could even begin to look for a key. For example, to operator a bombe you first needed a crib.

- Turing's lifestyle while in academe was known and tolerated. Whether the military would have tolerated his homosexuality remains unknown. It has been commented: "Fortunately the authorities did not know that Turing was a homosexual. Otherwise we might have lost the war."

- The intelligence gathered at Bletchley impressed Winston Churchill, who in turn provided unlimited resources to the codebreakers, who had written directly to the prime minister. Henceforth, there were to be no more barriers to recruitment or materials, including a challenge to readers of the *Daily Telegraph* in the form of a crossword puzzle.

4.0.2 Kidnapping Codebooks

- Enigma traffic was not one giant communications system, but rather several distinct networks. For example, the German Army in North Aftrica, the German Navy, and the German Air Force used different codebooks, and the Navy used a more sophisticated version of the Enigma machine. Extra scramblers, a variable reflector, nonstereotypical messages and a new system for exchanging message keys all contributed to making German Naval communication impenetrable.

- The Polish experience and the case of Hans-Thilo Schmidt had taught Bletchley Park that if intellectual endeavor fails to break a cipher, then it is necessary to rely on espionage, infiltration and theft. At times, this resulted in helping to provide a crib, at others, whole codebooks were captured, enabling the Naval Enigma to be deciphered for an entire month.

- With the Naval Enigma transparent, the Battle of the Atlantic began to swing in favor of the Allies. In order not to lose the benefit of material that had been secretly captured, further precautions had to be taken so as not to arouse suspicion that the German security had been compromised, lest the Enigma machines would be updated and Bletchley would be back to square one.
- This was accomplished in one instance by the British purposely not attacking all possible targets, but when those unaffected targets were later destroyed by units unaware of the policy, the enemy shrugged it off as either the result of natural misfortune, or caused by infiltration of a British spy, when in fact the lion's share was due to the breaking of the code rather than espionage. To the German Navy, the breaking of Enigma was considered impossible and inconceivable.

4.0.3 The Anonymous Cryptanalysts

- Bletchley Park also succeeded in deciphering Italian and Japanese messages, an intelligence achievement with the codename *Ultra*. In 1944, Ultra played a major role in the Allied invasion of Europe. Throughout the war, the Bletchley codebreakers knew that their decipherments were vital, but the cryptanalysts were never given any operational details or told how their decipherments were being used.
- It has been argued, albeit controversially, that Bletchley Park's achievements were the decisive factor in the Allied victory. Historian David Kahn summarized the impact of breaking Enigma as follows: "It saved lives. Not only Allied and Russian lives but, by shortening the war, German, Italian,and Japanese lives as well. . . . That is the debt that the world owes to the codebreakers; that is the crowning human value of their triumphs."
- After the war, Bletchley's accomplishments remained a closely guarded secret. In those years, Britain routinely deciphered secret communications of their former colonies, to whom they had distributed thousands of Enigma machines.
- This secrecy was particularly unfair to the cryptanalysts, most of whom returned to civilian life, sworn to secrecy, unable to reveal their role in the Allied war effort. While those who had fought conventional battles could talk of their heroic achievements, those who had fought intellectual battles of no less significance had to endure the embarrassment of having to evade questions about their wartime activities. (What did you do in the war, daddy?)
- A book published in 1974 (*The Ultra Secret*, F. W. Winterbotham) was the signal that Bletchley personnel were at last free to discuss their wartime activities. Those who had contributed so much to the war effort could now receive the recognition they deserved. In addition, some cryptanalysts came to know for the first time that their ideas had such important consequences, most notably Rejewski himself was in this category. Others did not live long enough to enjoy this glory. Most notable in this category was the tragic case of Alan Turing, who committed suicide in 1954 at the age of 42 after his life style became public.

4.1 Chapter 5—The Language Barrier

- While the British were breaking Enigma and altering the course of the war in Europe, American codebreakers cracked the Japanese cipher known as Purple, leading to success in a number of operations. Despite these achievements, it is possible that if the relevant cipher machines had been used properly (without repeated message keys, without cillies, without restrictions on plugboard settings and scrambler arrangements, and without stereotypical messages which resulted in cribs) they might never have been broken.
- The British and Americans both properly used cipher machines which were more complex than Enigma and therefore they remained unbroken throughout the war. Still, although electromechanical encryption offered relatively high levels of security, it was painfully slow and not well suited to the hostile and intense environments, such as the islands in the Pacific.

- After realizing that the use of "King's English" (including profanities) did not work, it was suggested that if each battalion in the Pacific employed a pair of Native Americans as radio operators, then secure communications could be guaranteed. Although the enemy may have ways of acquiring a good command of English, including the profanities, such was not the case for Native American languages, and mere translation could act as a virtually unbreakable code.
- What was needed was a large number of Native Americans fluent in English and literate. Unfortunately, due to government neglect, the literacy rate was very low on most of the reservations, so attention was focused on the four largest tribes: Navajo, Sioux, Chippewa, Pima-Ppago. For at least two reasons, Navajo was the favored tribe (biggest and no prior contact with German students). The use of these Navajo code talkers, as they were to become known, obviated the need for codebooks which might fall into enemy hands.
- The initial landing parties for the invasion of Guadalcanal in August 1942 included the first group of code talkers to see action. The code talkers soon proved their worth on the battlefield, and became heros in the process. As the war in the Pacific intensified, the Navajo code talkers played an increasingly vital role. A Major General commented that "without the Navajos, the marines would never have taken Iwo Jima."
- The special role of the 420 Navajo code talkers in securing communications was classified information. Just like Turing and the cryptanalysts at Bletchley, the Navajo were ignored for decades, until 1968 when the Navajo code was declassified. In 1982, the U. S. Government declared August 14 to be "National Navajo code Talkers Day." The greatest tribute to the work of the Navajos is that their code is one of the very few throughout history that has never been broken.

4.1.1 Deciphering Lost Languages and Ancient Scripts

- The task that confronted Japanese crytpanalysts is similar to that which is faced by archaeologists attempting to decipher a long-forgotten language, perhaps written in an extinct script. Deciphering ancient texts seems an almost hopeless pursuit. The decipherment of ancient scripts is not part of the ongoing evolutionary battle between codemakers and codebreakers (since there are no codemakers).
- The most famous, and the most romantic, of all decipherments was the cracking of Egyptian hieroglyphics. Thanks to a classic piece of codebreaking, the hieroglyphs were eventually deciphered, and ever since archaeologists have been able to read firsthand accounts of the history, culture and beliefs of the ancient Egyptians.
- From Dictionary.com:

  > **hieroglyphic** adj.
  > 1. (a) Of, relating to, or being a system of writing, such as that of ancient Egypt, in which pictorial symbols are used to represent meaning or sounds or a combination of meaning and sound. (b) Written with such symbols.
  > 2. Difficult to read or decipher.

- For over 3000 years the ancient Egyptians used these scripts (hieroglyphics, hieratic and demotic)in every aspect of their lives. Then toward the end of the fourth century A. D., within a generation, the Egyptian scripts vanished. The spread of the Christian Church was responsible for the extinction of the Egyptian scripts, outlawing their use in order to eradicate any link with Egypt's pagan past, and replacing them with a script called Coptic. The final linguistic link to Egypt's ancient kingdoms had been broken, and the knowledge needed to read the tales of the pharaohs was lost.
- Hieroglyphics are actually phonetic, which is to say that the characters largely represent distinct sounds, just like the letters in the English alphabet. However, the idea of phonetic spelling was thought to be too advanced for such an ancient civilization, and 17th century scholars were convinced that hieroglyphics were *semagrams*, that is, characters representing whole ideas and nothing more than primitive picture writing.

- For a century and a half, from 1650 to 1800, generations of Egyptologists were influenced by the (flawed) ideas of a German Jesuit priest Athanasius Kircher, who was widely acknowledged to be the most respected scholar of his age. In 1799, French scholars under Napoleon's initiative encountered the single most famous slab of stone in the history of archaeology (the Rosetta Stone).

- From Wikipedia.

  > [Definition] The Rosetta Stone is a dark grey-pinkish granite stone (often incorrectly identified as basalt) with writing on it in two languages, Egyptian and Greek, using three scripts, Hieroglyphic, Demotic Egyptian and Greek. Because Greek was well known, the stone was the key to deciphering the hieroglyphs.
  >
  > [Use as metaphor] Rosetta Stone is also used as a metaphor to refer to anything that is a critical key to a process of decryption, translation, or a difficult problem, e.g., "the Rosetta stone of immunology", "thalamo-cortical rhythms, the Rosetta stone of a subset of neurological disorders", "Arabidopsis, the Rosetta stone of flowering time (fossils)".
  >
  > Today, there is a popular foreign language instructional software package called The Rosetta Stone. Also, the ESA space probe Rosetta is named after it, because it is hoped the mission will help unlock the secrets of how our solar system looked before the planets formed.

- The Rosetta Stone appeared to be the equivalent of a cryptanalytic crib and was potentially a means of unraveling the meaning of the ancient Eqyptian symbols. This effort has hindered by three facts: the stone was damaged; the Eqyptian part used an ancient Egyptian language which had not been spoken for 8 centuries, making it difficult to establish the sound of the words; the legacy of Kircher encouraged archaeologists to think in terms of semagrams, rather than phonograms. The stone (all 1500 pounds) has resided in the British Museum since the defeat of the French in 1802.

- The two key players in the eventual successful decipherment of the hieroglyphics of the Rosetta Stone were Thomas Young (b. 1773) and Jean-Francois champollion (b. 1791).

- Young's breakthrough came when he focused on a set of 7 hieroglyphs surrounded by a loop, called a *cartouche*. Using a hunch which was correct, this was deciphered as the name of the Pharaoh Ptolemy. He repeated his strategy in the successful decipherment of another cartouche, this time as the name of a Ptolemaic queen, Berenika. Despite this success, the influence of Kircher was taking its toll and he lost interest in hieroglyphics.

- Champollion was prepared to take Young's seminal ideas to their natural conclusion. At the age of 17, encouraged by the great French mathematician Jean-Baptiste Fourier, he presented ideas that were so innovative that he was immediately elected to the Academy in Grenoble. Sixteen years later, in 1822, he applied Young's approach to other cartouches, resulting in one whose decipherment was one of the greatest names of ancient times, Alexander. Despite this progress, the credit was still primarily due to Young. Nevertheless, the next cartouche to be deciphered turned out to be Rameses, one of the greatest pharaohs, breaking the theory that spelling was used only for foreign names.

- Champollion was also credited with pointing out that Coptic might be the language of hieroglyphics, and that scribes sometimes exploited the rebus principle. In the Rameses example, the first syllable is represented by a rebus image, a picture of the sun, while the remainder of the word is spelled more conventionally.

- From Wikipedia:

  > [Definition] A rebus (Latin: "by things") is a kind of word puzzle which uses pictures to represent words or parts of words, for example: H + picture of ear = Hear, or Here.

- In 1824, at the age of 34, Champollion published all his achievements in a book entitled *Précis du système hiéroglyphique*. For the first time in 14 centuries it was possible to read the history of the pharaohs, as written by their scribes. Who should receive the most credit for breaking the hieroglyphics code? On his first expedition to Egypt, four years before his untimely death at the age of 41, Champollion could read and interpret easily character by character the hieroglyphs which he had seen previously only in drawings and lithographs, and which 30 years earlier could only be guessed wildly at by Napoleon's expedition.

5 February 23, 2006

5.0.2 The Mystery of Linear B

- Having conquered hieroglyphics, archaeologists went on to decipher other ancient scripts. Linear B, a Cretan script from the Bronze Age, was deciphered without the aid of cribs, and is generally regarded as the greatest of all archaeological decipherments. From Wikipedia:

  > The Minoans were a pre-Hellenic Bronze Age civilization in Crete in the Aegean Sea, prior to Helladic or Mycenaean culture (i.e., well before what we know as Classical Greece). Their civilization flourished from approximately 2600 to 1450 BC.

- In 1900, Sir Arthur Evans, an eminent British archaeologist uncovered sets of clay tablets on Crete. The most recent set, called Linear B, contained the most tablets and was thus the target for decipherment. Writing was from left to right and there were 90 distinct characters, indicating syllabic writing, but it was not clear what language Linear B was written in. A nasty debate arose about whether the Minoans spoke Greek or their own non-Greek language. Evans died in 1940 without deciphering Linear B.

5.0.3 Bridging Syllables

- In the mid 1940s, Alice Kober concluded that Linear B represented a highly inflective language, meaning that the word endings are changed to reflect gender, tense, case. This led to her discovery that the symbols represented syllables. She also died without deciphering Linear B.

5.0.4 A Frivolous Digression

- Building on Kober's pioneering work, Michael Ventris, an architect working in his spare time, in 1952 published significant results in the decipherment of Linear B, providing strong evidence that the language of Linear B was Greek. Together with John Chadwick, a specialist in Greek philology (=linguistics), they convinced the rest of the world that Linear B is indeed Greek. This discovery had a profound impact on archaeologists' view of Minoan history. Since it has not yet been deciphered, it is now likely that Linear A (which preceded Linear B) was a distinctly different language from Linear B.

- The bulk of Linear B tablets are inventories, describing everyday transactions and implying a bureaucracy. From Dictionary.com: Domesday Book

  > The written record of a census and survey of English landowners and their property made by order of William the Conqueror in 1085-1086.

- These palace records might seem mundane, but they are inherently romantic because they are so intimately associated with the *Odyssey* and *Iliad*. While scribes in Knossos and Pylos recorded their daily transactions, the Trojan War was being fought. The language of Linear B is the language of Odysseus (a character in Odyssey).

5.1 Chapter 6—Alice and Bob Go Public

- In addition to Turing's bombes, which were used to crack the Enigma cipher, in 1943 the British also invented another codebreaking device, Colossus, to combat an even stronger form of encryption, namely the German Lorenz cipher. Consisting of electronic valves instead of electromechanical relay switches, the Colossus was considerably faster than the bombes, and it was programmable, making it a precursor to the modern digital computer. Because everything at Bletchley park was destroyed after the war, it was ENIAC (1945, University of Pennsylvania), not Colossus, that was considered the mother of all computers.

- The computer played a crucial role in the postwar battle between codemakers and codebreakers. Three significant differences between computer and mechanical encryption were: virtual vs. physical, sheer speed, numbers rather than letters. Before encryption, any message is converted into binary digits (bits). Even though we are dealing with computers and numbers, and not machines and letters, every encipherment, no matter how complex, can be broken down into combinations of the simple operations of substitution and transposition.

- A series of scientific, technological and engineering breakthroughs (transistor 1947, computers made to order by Ferranti 1951, IBM computer 1953, Fortran 1957, integrated circuit 1959) made computers, and computer encryption, far more widely available.

- In the 1960s, one of the primary concerns of businesses using encryption was standardization. In 1973, the National Bureau of Standards of the US requested proposals for a standard encryption system. Despite harassment by NSA, a version of an IBM product called Lucifer was adopted in 1976, and this system, called Data encryption Standard (DES), remains America's official standard a quarter of a century later.

- Despite standardization and the strength of DES, businesses still had to deal with one more major issue, a problem known as *key distribution*. As business networks grew in size, as more messages were sent, and as more keys had to be delivered, the distribution process by couriers became a horrendous logistical nightmare, and the overhead cost became prohibitive.

- The problem of key distribution has plagued cryptographers throughout history. No matter how secure a cipher is in theory, in practice it can be undermined by the problem of key distribution. Although it may seem to be a mundane issue, key distribution became the overriding problem for postwar cryptographers. Although computers transformed the implementation of ciphers, the greatest revolution in 20th century cryptography has been the development of techniques to overcome the problem of key distribution.

5.1.1 God Rewards Fools

- Whitfield Diffie, who had matured into one of the truly independent security experts in the 1970s, realized that whoever could find a solution to the key distribution problem would go down in history as one of the all-time great cryptographers. If governments and large corporations were having trouble coping with key distribution, then the public would find it impossible, and would effectively be deprived of the right to privacy. This further inspired Diffie with the idea of finding a solution, an effort he shared with Martin Hellman (Stanford).

- The whole problem of key distribution is a classic catch-22 situation: before two people can exchange a secret (the message), they must already share a secret (the key). For 2000 years it was considered an indisputable truth of cryptography that the distribution of keys is unavoidable, that key exchange was an inevitable part of encipherment.

- But is it really? Consider the following. Alice uses her own key to encrypt a message to Bob, who encrypts it again with his own key and returns it. When Alice receives the doubly encrypted message, she removes her own encryption and returns it to Bob, who can then remove his own encryption and

read the message. What's wrong with this picture? Encryption systems are far more sensitive than padlocks when it comes to order.

- Modular arithmetic is rich in one-way functions. In half an hour of frantic scribling, Hellman proved that Alice and Bob could agree on a key without meeting, thereby disposing of an axiom that had lasted for centuries. ("The muse whispered to me, but we all laid the foundations together") The Diffie-Hellman-Merkle key exchange scheme is one of the most counterintuitive discoveries in the history of science, and it forced the cryptographic establishment to rewrite the rules of encryption. It enables Bob and Alice to establish a secret via public discussion.

5.1.2 The Birth of Public Key Cryptography

- The Diffie-Hellman-Merkle key exchange scheme was a giant leap forward, but it was still not perfect. It needed some tweaking to preserve the spontaneity of e-mail by overcoming global time differences.
- After going through long periods of barren contemplation, in 1975 Diffie concocted a new type of cipher (in theory only but revolutionary nonetheless), one that incorporated an *asymmetric key*, rather than a *symmetric* one. The key distribution problem is avoided by the use of a *private key*, known only to an individual, and a *public key*, accessible to anyone. Anyone can send an encrypted message to Alice, but only Alice, who possesses her private key, can decrypt the message.
- To turn asymmetric ciphers from a great idea into a practical invention, somebody had to discover an appropriate mathematical function. By the end of 1976, no one could find such a function, and it seemed that Diffie's idea worked in theory but not in practice. Nevertheless, the team of Diffie, Hellman and Merkle had revolutionized the world of cryptography by creating a workable but imperfect key exchange system and proposing the concept of an asymmetric cipher—a perfect but as yet unworkable system.

5.1.3 Prime Suspects

- At the MIT Laboratory for Computer Science, Ron Rivest was excited about a paper by Diffie and Hellman describing the concept of asymmetric ciphers, so he drafted two colleagues, Adi Shamir, and Leonard Adleman, to join him in trying to find a one-way function that fitted the requirements of an asymmetric cipher. A year later, in 1977, Rivest wrote the outline of a scientific paper, a breakthrough, which had grown out of a yearlong collaboration with Shamir and Adleman. The system, a form of *public key cryptography*, was dubbed RSA, and went on to become the most influential cipher in modern cryptography.
- In the RSA system, Alice's public key is the product N of two (very large) prime numbers p and q. Anyone can encrypt a message to Alice by combining Alice's public key with a one-way function (also public knowledge), obtaining a one-way function (called Alice's one-way function), and then inserting the message into it, noting the result, and sending it to Alice. Alice (and only Alice) then uses the information p and q to reverse the one-way function and decipher the message. The most beautiful and elegant aspect of the RSA asymmetric cipher is that if N is large enough, it is virtually impossible to deduce p and q from N.
- If all the computers in the world worked together, a number as big as 10^{130} could be factored in about 15 seconds. Just to be on the safe side, for important banking transactions, N tends to be at least 10^{308}. The 15 seconds then becomes 1000 years! It follows that the only possible threat to RSA is that at some time in the future, somebody will find a quick way to factor very large numbers. This is very unlikely, and RSA seems secure for the foreseeable future.
- The great advantage of RSA public key cryptography is that it does away with all the problems associated with traditional ciphers and key exchange. It is now routine to ecrypt a message with a sufficiently large value of N so that all the computers on the planet would need longer than the age of the universe to break the cipher.

5.1.4 The Alternative History of Public Key Cryptography

- According to the British Government, public key cryptography was originally invented at the Government Communications Headquarters (GCHQ), the top-secret establishment formed from the remnants of Bletchley Park after the Second World War. This is a story of remarkable ingenuity, anonymous heroes and a government cover-up that endured for decades.

- In 1969, the military asked James Ellis, one of Britain's foremost government cryptographers, to look into ways of coping with the key distribution problem. Because he was involved in issues of national security, Ellis was sworn to secrecy throughout his career.

- Ellis suggested that the receiver, Alice, deliberately create noise, which she could measure before adding it to the communication channel that connects her with Bob. The key was the noise, and only Alice needed to know the details of the noise. This idea was similar to those of Diffie, Hellman and Merkle, except that he was several years ahead of them. By the end of 1969, he had proved to himself that public key cryptography was possible, and he had developed the concept of separate public keys and private keys.

- Ellis also knew that he needed to find a special one-way function, so MATHEMATICIANS were brought in. Three years later, in 1973, and four years before RSA was announced, Clifford Cocks, a recent PhD from Cambridge University, was going through the exact same thought processes as Rivest, Shamir, and Adleman.

- Cocks did not fully appreciate the significance of his discovery. He had no idea that he had made one of the most important cryptographic breakthroughs of the century. When he finally did realize what he had done, it struck him that his discovery might have disappointed G. H. Hardy, one of the great English mathematicians of the early part of the 20th century, who had proudly stated (*A Mathematician's Apology* 1940) "Real mathematics has no effects on war. No one has yet discovered any war-like purpose to be served by the theory of numbers."

- Although Cocks's idea was one of GCHQ's most potent secrets, it suffered from the problem of being ahead of its time. Cocks and Ellis had proved that the apparently impossible was possible, but nobody could find a way of making the possible practical. By 1975, James Ellis, Clifford Cocks, and a third collaborator Malcolm Williamson, had discovered all the fundamental aspects of public key cryptography, yet they all had to remain silent.

- GCHQ failed to patent public key cryptography because of the fear of having to reveal the details of their work, and also failed to block the Diffie-Hellman application for a patent in 1976 because of the inability to see the coming digital revolution and potential of public key cryptography. Meanwhile, in 1996, RSA Data Security, Inc., was sold for $200 million.

- Although GCHQ were the first to discover public key cryptography, it was the academics who were the first to realize the its potential. Moreover, the discovery by the academics was wholly independent of GCHQ's discovery, and on an intellectual par with it. Indeed, the academic environment is completely isolated from the top-secret domain of classified research, and academics do not have access to the tools and secret knowledge that may be hidden in the classified world. On the other hand, government researchers always have access to the academic literature. Information flows freely in one direction, but it is forbidden to send information in the opposite direction—a sort of one-way function.

- GCHQ went public only in 1997. Ellis, who died one month before this coming out, thus joined the list of British cipher experts, namely Charles Babbage and Alan Turing, whose contributions would never be recognized during their lifetimes.

6 March 2, 2006

6.1 Chapter 7—Pretty Good Privacy

- The success of the Information Age depends on the ability to protect information as it flows around the world, and this relies on the power of cryptography, and not just in government and the military. In the 21st century, the fundamental dilemma for cryptography is to find a way of allowing the public and business to use encryption in order to exploit the benefits of the Information age without allowing criminals to abuse encryption and evade arrest.

- Phil Zimmermann attempted to encourage the widespread use of strong encryption, and as a result, America's security experts panicked, the effectiveness of the billion-dollar national Security Agency was threatened, and he became the subject of an FBI inquiry and a grand jury investigation. According to Zimmermann, cryptographers have a duty to encourage the use of encryption and thereby protect the privacy of the individual.

- In theory, RSA offered an antidote to the Big Brother scenario, but in practice RSA encryption required a substantial amount of computing power in comparison with symmetric forms of encryption, such as DES. Consequently, in the 1980s it was only government, the military, and large businesses that owned computers powerful enough to run RSA. In contrast, Zimmermann believed that everybody deserved the right to privacy that was offered by RSA encryption, and he directed his political zeal toward developing an RSA encryption product for the masses.

- If Alice wants to send an encrypted message to Bob using Zimmermann's software product PGP, she begins by encrypting it with a symmetric cipher IDEA, similar to DES, and then she uses Bob's RSA public key to encrypt the IDEA key. Bob uses his RSA private key to decrypt the IDEA key, and then uses the IDEA key to decrypt the message. PGP creates private keys from random movement of a mouse, and has the additional feature of certifying digital signatures, thus guaranteeing both privacy and authorship.

- Credit for the certification of signatures, as well as the idea of using a combination of symmetric and asymmetric ciphers to speed up encryption belongs to Diffie and Hellman—-but Zimmermann was the first to put everything together in one easy-to-use encryption product, which was efficient enough to run on a moderately sized personal computer.

- Concerns about patent infringement and government restrictions led Zimmermann to post PGP on the internet, rather than sell it. These concerns were realized when PGP was labeled by RSA Data Security, Inc. as "banditware," and Zimmermann was accused of being an arms dealer.

6.1.1 Encryption for the Masses... Or Not?

- The debate about the positive and negative effects of encryption in the Information Age continued throughout the 1990s and is currently as contentious as ever. Restricting the use of cryptography would allow the police to spy on criminals, but it would also allow the police and everybody else to spy on the average citizen.

- The Supreme Court was sympathetic to the argument of law enforcers and ruled in 1967 that wiretaps could be employed as long as a court authorization was obtained first. (FISA anybody?) The Four Horsemen of the Infocalypse—drug dealers, organized crime, terrorists and pedophiles. In addition to domestic policing, there were issues of national security. The Echelon system of scanning e-mails,faxes, telexes and telephone calls would effectively be useless if all messages were strongly encrypted.

- The pro-encryption case is based on the belief that privacy is a fundamental human right. Well-known cases of unjustified wiretapping were perpetrated by Presidents Johnson, Nixon, and Kennedy. Civil libertarians argue that wiretapping is not a crucial element in most cases. (This may be true, but it is

a weak argument). Ironically, the greatest allies of the civil libertarians are big corporations. In just a few years from now, Internet commerce could dominate the marketplace. If Internet commerce is to thrive, consumers around the world must have proper security.

- In general, popular opinion appears to be swinging behind the pro-encryption alliance, who have been helped by a sympathetic media and a couple of movies, *Mercury Rising, Enemy of the State*. Some dubious attempts at compromises are *key escrow, trusted third parties*.

- Despite the debate, one aspect of future encryption policy seems certain, namely the necessity for *certification authorities*. Certification authorities can verify that a public key does indeed correspond to a particular person and can guarantee the validity of digital signatures. The deciding factor in the debate will be whom the public fears most—criminals or the government.

6.1.2 The Rehabilitation of Zimmermann

- In 1996, after three years of investigation, the U. S. Attorney General's Office dropped its case against Zimmermann. Even today, the legal issues surrounding the Internet are subject to debate and interpretation. On the other hand, Zimmermann eventually reached a settlement with RSA resolving the patent issue. PGP is now sold to businesses but it is still freely available to individuals who do not intend to use it for any commercial purpose. A laptop and its PGP software provide a level of security that is beyond the combined efforts of all the world's codebreaking establishments.

6.2 Chapter 8—A Quantum Leap into the Future

- The present state of cryptography is dominated by public key cryptography and the political debate that surrounds its use, and it is clear that the cryptographers are winning the information war. Moreover, not only do we have to guess which discoveries lie in the future, but we also have to guess which discoveries lie in the present. There may already be remarkable breakthroughs hidden behind the veil of government secrecy.

6.2.1 The Future of Cryptanalysis

- Despite the edge currently held by cryptographers (or perhaps because of it), cryptanalysts are still playing a valuable role in intelligence gathering—the NSA is the world's largest employer of MATHEMATICIANS. For one thing, the number of Internet users is far greater than the number that take adequate precautions in terms of privacy. For another, even if users employ the RSA cipher properly, there is still plenty that codebreakers can do to glean information from intercepted messages.

- Example: the *tempest attack*—detection of the electromagnetic signals emitted by the electronics in a computer's display unit. In other words, intercept the message as it is typed into the computer, before it is encrypted. Other examples: the use of viruses (which send private keys to the eavesdropper) and Trojan horses (such as a bogus copy of PGP); *backdoors*—encryption software which allows its designers to decrypt the messages. Although these are useful to cryptanalysts, the main goal is to find a way to decipher RSA, and this will require a major theoretical or technological breakthrough.

- Actually, all you have to do is figure out a shortcut to factoring a number N which is the product of two prime numbers. As we have seen, for very very super astronomically large N this is essentially impossible. What is being looked at now is the notion of a *quantum computer*, the description of which involves some notions from quantum mechanics, an explanation of how objects behave at the microscopic level.

- Some key words from quantum theory, which are discussed in this chapter: wave-particle duality, photons, superposition of states, many-worlds interpretation. Quantum mechanics has shown itself to be the most successful and practical scientific theory ever conceived. Of all its consequences, the most technologically important is potentially the quantum computer.

- As well as destroying the security of all modern ciphers, the quantum computer would herald a new era of computing power. It was David Deutsch who in 1984 advanced the idea that computers ought to obey the laws of quantum physics instead of classical physics.
- Ordinary computers operate at a relatively macroscopic level, where quantum laws and classical laws are almost indistinguishable. However, at the microscopic level (of very fast computers) the quantum laws prevail. Ordinary computers have to address several questions sequentially. Quantum computers can address them simultaneously by exploiting the laws of quantum physics.
- A quantum computer defies common sense. It can be thought of as either a single entity that performs the same calculation simultaneously on, say 128 numbers simultaneously; or as 128 entities, each in a separate universe, each performing just one calculation. Quantum computing is *Twilight Zone* technology. Because a quantum computer deals with 1's and 0's that are in a quantum superposition, they are called *qubits* instead of bits.
- No one has yet envisaged how to create a solid, practical quantum computer. Another problem at the beginning was how to program a quantum computer. However, in 1994, Peter Shor of AT&T Bell Laboratories in New Jersey succeeded in defining a useful program for a quantum computer. The latter however, doesn't yet exist! Another scientist from Bell Labs, Lov Grover, in 1996 discovered a program which searches a list at an incredibly high speed, fast enough to crack DES. Of course, a quantum computer still didn't exist at that time (or even today). A quantum computer, when it is ever built, will jeopardize the stability of the world!

6.2.2 Quantum Cryptography

- Quantum theory is at the heart of a new unbreakable cipher called quantum cryptography, which had its origins in the concept of quantum money (a brilliant but wholly impractical idea), which in turn depends on the concept of polarization of a photon to create dollar bills that can never be forged. Incorporating photons in its design made the dollar bill impossible to be measured accurately (an aspect of the uncertainty principle in quantum mechanics), and hence created a barrier to counterfeiting,
- Stephen Wiesner's original idea was passed on to Charles Bennett and Gilles Brassard and it evolved into an idea for an uncrackable cipher. If an encrypted message was represented and transmitted by a series of polarized photons, then no one would be able to measure it accurately (hence read it), and consequently it couldn't be decihered. An outline of the physics involved in quantum cryptography is given on pp. 341–344 and a summary is given on pp. 346–347.
- Responding to critics, Bennett constructed in 1988 an apparatus that showed that two computers, 30 cm. apart, could communicate in absolute secrecy. Since then, the challenge has been to build a quantum cryptographic system that operates over useful distances. (However, photons do not travel well!) In 1995 this was stretched to 23 km. using photons traveling over optic fiber. Through air, the best distance is 1 km.
- If quantum computers were to become a reality, then RSA and all other modern ciphers would be useless, and quantum cryptography would become a necessity in order to overcome a privacy gap. Quantum cryptography is the more advanced technology. It is currently possible to build a quantum cryptography link between the White House and the Pentagon. Perhaps there already is one.
- Quantum cryptography would mark the end of the battle between codemakers and codebreakers, and the codemakers emerge victorious. The claim that quantum cryptography is secure is qualitatively different from all previous such claims. Reason: if a message protected by quantum cryptography were ever to be deciphered, it would mean that quantum theory is flawed, which would have devastating implications for physicists, forcing them to reconsider their understanding of how the universe operates at the most fundamental level. If quantum cryptography systems can be engineered to operate over long distances, the evolution of ciphers will stop. The technology will be available to guarantee secure communications for government, the military, businesses and the public (including criminals).