(Freely extracted and translated from Chapter 7 of the book: F. Luccio and L. Pagli, Algoritmi, divinità e gente comune, Edizioni ETS, Pisa 1999).

RANDOMNESS

How is observed that, if the correct result is very difficult to obtain, a probably correct result can be accepted.

PASSER-BY. But what kind of life would you like then?
PEDDLER. Any kind, just as God would send it to me, with no other conditions.
PASSER-BY. Any life at random, without knowing anything about it in advance, just as we don't know anything about the new year?
PEDDLER. Precisely.
PASSER-BY. That's I would like too if I were to live all over again; and that's what everyone would like. But this means that up until the end of this year, Fortune has treated everyone badly. And it's clear that everyone thinks that he was allotted more, and greater, evil than good; if to live the same life all over again, with all its good and all its evil, no one would want to be born anew. The life that's beautiful is not the life we know, but the life we don't know; not the past life, but the future. With the new year, Fortune will start treating you and me and all the others well, and the happy life will begin. Isn't it true?
PEDDLER. Let's hope so.

This famous dialogue was imagined by Giacomo Leopardi[1]. In the last sentences he was expressing hope as mere irony, as chance will remain an enemy. But not everybody is as pessimistic as Leopardi, nor is convinced that chance can be characterized in any way, if it exists at all. Let us then look after this fellow, who will be the main character of the following discussion.

The ancients did not teach us much on this matter limiting their speculation, without particular acuteness, on what happens when playing dice[2]. In fact, the

[1] G. Leopardi, *Operette Morali: Essays and Dialogues,* translated by G. Cecchetti, University of California Press, Berkeley, 1982.

[2] Even giants of mathematics like Pascal and Fermat stopped there, but in fact the whole history of chance is quite unclear. Democritus is often quoted as being the first assertor of a fundamental role that chance plays in natural phenomena, although probably no ancient thinker was more mechanistic than he. The confusion is due to Dante, who met the philosopher in Limbo, and designated him as the one *che il mondo a caso pone* ("who puts the world on chance": *La Divina Commedia, Inferno,* Canto IV, v.136). Maybe Dante was disappointed by the absence of a spiritual end in atomism; nevertheless the confusion continues nowadays, to such an extent that J. Monod dedicated to Democritus a scheme of evolution based on chance, in a famous book of biology (*Le hasard et la nécessité*, Ed. du Seuil, Paris 1970). In any case we can affirm that even the theories built on the principle of cause have accepted chance as the warrantor of freedom against the arrogance of rules. A collection of conversations edited by É. Noël (*Le hasard aujourd'hui*, Ed. du Seuil, Paris 1991) allows to understand the role of chance in a variety of disciplines.

scientific theory of chance is rather recent, and it is this we refer to, although the example of dice is still useful. Let us take the viewpoint of the caster to exclude that the game is consciously altered (otherwise the caster would cheat himself). If a sequence of results is random, the caster cannot foresee any one of them beforehand. Using a die with numbered faces the results in a sequence are numbers between 1 and 6. An immediate consequence of randomness is that, if the sequence is very long, all the values from 1 to 6 must appear in it approximately the same number of times. Indeed, if this were not the case at a certain point it would be possible to forecast the future by observing the past ones. For example, if the value 3 had appeared more frequently than the other ones the caster could bet on it, because the die is probably unbalanced, or he himself influences the casting unconsciously. Very complicated consequences derive from very simple arguments when randomness is involved, at a point that a rigorous foundation of probability theory was laid down only in the twentieth century by Andrej Kolmogorov, who also gave the new definition of randomness presented below in an intuitive fashion. [3]

We know from mathematical logic that the possibly infinite set of valid sentences built with the alphabet and the rules of a sufficiently powerful language includes the sequences of characters representing all possible algorithms[4]. Once the language has been fixed, a sequence is *random* according to Kolmogorov if *cannot be generated by an algorithm represented in turn with a sequence shorter than the original one.* Clearly this statement must be set in a strictly mathematical context, but its main consequence can be simply stated: *the most economical way of defining a random sequence is to write it down*.

According to the above definition a caster of dice might predict a sequence of results using an *ad hoc* algorithm; but if the shortest sequence expressing a prediction algorithm is longer than the sequence of results itself, the latter must be considered random. In such an intuitive context consider the world of well-formed sentences in English. The sentence:

It is an ancient Mariner,
and he stoppeth one of three

is to be regarded as a random sequence, because its shortest univocal definition is:

First two verses of "The Rime of the Ancient Mariner" by Samuel T. Coleridge

and this sequence is longer than the former one. With the same reasoning, the whole poem is not random. For example let us establish a communication link with a friend

[3] To investigate this topic thoroughly, without leaving an elementary and pleasant level, we suggest the essay of I. Ekeland: *Au hasard*, Ed. du Seuil, Paris 1991.
[4] An *algorithm* is the specification of the steps to be performed on a predefined "abstract machine", for solving an arbitrary problem. Informally a computer program for locating an address in the Web, or the specification of the arithmetic steps for computing the square root of an integer, or the instructions for using a library catalogue, or for cooking a Chinese dish, are examples of algorithms in their specific fields.

living on the Moon (high transmission costs) where he has access to a library identical to the one we have (same reference system). For communicating the two verses above we better transmit the verses as they are, while for communicating the whole poem we will merely transmit author and title.

An encouraging fact is that Kolmogorov random sequences have identical properties as the ones non-predictable in the previously existing probabilistic paradigms. In particular a long sequence generated by casting a die, which does not admit a prediction algorithm shorter than the sequence itself, must contain the values from 1 to 6 with approximately equal frequencies.

This algorithmic approach leads to a new understanding of randomness in terms of *decidability*, a fascinating and subtle branch of mathematical logic. Invoking a sophisticated version of Epimenides' paradox it can be proved that it is undecidable to establish if an arbitrary sequence is random. I.e., no algorithm may exist to make such a decision[5]. This gives rise to an awkward situation. A caster of dice may easily realize that some particular sequences are not random, but in general cannot decide if an arbitrary sequence is random. More precisely, if he is unable to find a simple generating rule, he can decide that the sequence is random *to the best of his knowledge,* but is not authorized to apply the theorems of probability calculus that would lead to invalid results if the sequence is, in fact, not random.

In the strange world of randomness we observe another delicate fact. A simple arithmetic argument indicates that the number of sequences of a predetermined length is greater than the number of all the sequences of smaller length, independently of the alphabet used. For example with the 26 characters of the English alphabet we can build 26 sequences of length one (i. e., of one character); 26x26 = 676 sequences of length two, obtained by inserting each one of the 26 characters in front of each of the sequences of length one; 26x26x26 = 17,576 sequences of length three, and so on. This is an instance of what is called *exponential growth*. In our example the sequences of length three are much more numerous than the ones of length one plus the ones of length two (17,576 is much greater than 26 + 676). This difference grows enormously with the length of the sequences and gives an incontestable proof of the *existence* of random sequences of any length. In fact, if all the sequences of length *n* were non-random, each one of them would admit a generating algorithm represented by a sequence shorter that *n*. That is impossible because the number of such sequences is not sufficient. However this proof is not constructive, in that it does not teach us how to *generate* random sequences. We then remain in doubt whether such sequences, although existing, can actually be found.

[5] "The Cretans are liars". This statement, attributed to Epimenedes by St. Paul, is a provocation: not against the Cretans who are known for their sense of humor, but against the logical bases of reasoning. In fact, Epimenides was born in the island of Crete: if the statement were true (Cretans indeed do lie), the author would have lied, thereby implying that Cretans do not lie; on the other hand, if the statement were false, the author would have said the truth, hence the statement would be true. The same logical structure has been crucially used in the twentieth century to prove mathematical undecidability of several problems, for which an assumption of decidability would give rise to a statement that is contemporarily true and false.

In essence, random sequences should be generated by chance, however the question of the *existence of chance* remains open. People have different opinions on this intriguing problem. Carl Gustav Jung and other famous psychologists had no doubt on the existence of chance as a threat to rational behaviour. Physicists assert its existence to support some of their theories, until new theories will possibly arise. Biologists are still discussing. We will escape the perils of this discussion, by personifying chance as a deity of our private Olympus: as the ancient Greeks would have done; as several African cultures are still doing. For us, chance exists and that's that! So, for underlining its supernatural essence, and distinguishing it from the various meanings of the same word, we will henceforth write it with a capital initial.

To proceed we must speak a bit of algorithms. The classical theory of algorithms is based upon several mathematical hypotheses, whose mere attempt of explication would bore the reader to death. The simplest and most important hypothesis, however, is very simple: *if a correct algorithm generates a result, this result is certainly correct.* But is this always true, or indispensable? Is it even possible?

This question is particularly relevant for the class of problems that cannot be solved in a reasonable amount of time even using an extremely powerful computer on a modest amount of input data. A class with the disheartening property of including many important problems. If a goal is practically unreachable we have to pursue less ambitious objectives: indeed we can content ourselves with procedures which terminate rapidly, but whose solution is correct *only* in the great majority of cases. Let us discuss how this can be attained, and what is the real significance of this limitation.

The key elements with which to proceed are Chance and some very elementary concepts of probability theory, which we will illustrate without pretending to be rigorous. The first and obligatory step is to enter into the world of gambling, as the mathematicians of the seventeenth century did while giving birth to the calculus of probability. This will give us occasion of reviewing some mathematical properties of games, without scorning the idea of realizing a good profit. For this purpose, nothing compares to roulette.

With unlimited capital there is a mathematical certainty of realizing a win equal to the initial stake. The gambler has simply to always bet on red, according to the strategy of *double or nothing* of doubtful reputation. In fact, if the roulette is ruled by Chance (i. e., is not rigged), the probability that the ball stops on the red is equal to the probability that it stops on the black. In mathematics, the probability p of an event is expressed by a number between 0 and 1 or, if one prefers, by a fraction. Thus $p=0$ indicates that the event will never occur; $p=1$ indicates that the event will certainly occur; p=1/2 will indicate that the event will occur one every two times, as in the case of red and black in the roulette game. [6] If the gambler plays double or nothing, in the case of a red outcome he cashes a prize equal to the stake, and stops; in the case of a

[6] More precisely the probabilities of having a red or a black outcome are both equal to 18/37, because the roulette generally includes 37 numbers, of which 18 are red, 18 are black, and there is a "zero" that is neither red nor black. We have assumed that the zero is not present to simplify the computation, without altering the substance of the results: the probabilities of red or black are then 18/36=1/2.

black outcome he continues to bet on the red, doubling the stake at each round until a red outcome occurs. For this purpose the gambler must have an amount of money available that grows exponentially with the number of consecutive unfavourable rounds. For example, if the initial stake is S and the red appears only in the fifth round, the total amount of the stakes is $S + 2S + 4S + 8S + 16S$, equal to $31S$. The win is $16S$ equal to the last stake, and since also this stake is recuperated, the total amount cashed is $32S$, with a gain of S on the total investment. Not much if compared to the amount of the financial commitment, but satisfactory in the light of the simple reasoning that follows.

Each round has two possible outcomes (red and black). On the average one loses once every two times. Two rounds have four possible outputs: red-red, red-black, black-red, black-black, and one loses only in the last case, that is once every four times. Three rounds have eight possible outcomes, and one loses only with black-black-black, that is once every eight times, and so on. All this is valid if the rounds are mutually independent, that is, if each ball launch is not influenced by the previous ones. In mathematical terms we have to compute the probability of occurrence of a sequence of independent events that equals the product of the probabilities of the single events. That is, we multiply the probabilities of a black outcome in, say, three rounds, obtaining a total probability of loosing equal to $1/2 \times 1/2 \times 1/2 = 1/8$. As we can see the result decreases rapidly with an increasing number of events because the product of different factors decreases if each factor is smaller than one, as is the case with probabilities.

Proceeding this way we can immediately compute that, when playing double or nothing, the probability of failure for twenty consecutive rounds (i. e., the ball never falls into a red cell in twenty launches) is less than one over one million; in thirty rounds is less than one over one billion; in forty rounds is less than one over one thousand billions. This last value is extremely small, and in any case is smaller than the probability that the game comes to an end for a different reason, such as the casino is at the epicenter of an earthquake of magnitude 9 on the Richter scale or the croupier is stricken by an intolerable attack of nettle-rash. If we play the game for forty rounds, the probability of loss can therefore be neglected. Unfortunately, the casinos impose an upper bound to the amount that may be staked!

When playing double or nothing, Chance does not intervene in the algorithm itself but only in the situations that the algorithm has to face. The actions to be taken are established beforehand and take into account all the possible outcomes of the launches. That is, the algorithm is deterministic and is designed to face Chance. Much less common in our way of thinking, and much more interesting in an algorithmic sense, is the use of randomisation as an ingredient of computation. If an algorithm relies on Chance for some of its actions it thus becomes *randomised*, even if this adjective expresses the mechanism only partially. To understand this new situation, let us recall a concrete example taken from an ancient chronicle of Central Asia (oral tradition).

Massud, a young shepherd of the steppe, is consumed by the desire to know if Pardis, his betrothed, is as beautiful as all her relatives swear. By modesty and tradition Pardis always appears in public with a veiled face, and only women can see her in privacy. Massud wishes to ask some of them tactfully, but does not know who are the right women to ask. He fears the opinion of relatives to be exceedingly positive, the judgment of Fatima (the belly dancer who used to be his secret girl friend) to be exaggeratedly negative, the opinions of persons with similar taste and habits to be meaninglessly alike. In essence, Massud wishes to collect reliable opinions from independent witnesses chosen at random.

To proceed with a scientific method, Massud establishes some rules based on his knowledge of human psychology. A woman's opinion that Pardis is unpleasant to the sight must unfortunately be taken for sure. An opinion that Pardis is beautiful, instead, can only be trusted with a certain probability. In the first case Massud will terminate his inquiry, and try to escape his commitment for marriage without major damages. In case of a positive opinion, instead, he will continue to make inquiries with other women until he is convinced that his betrothed is really beautiful. But how many women will have to be interrogated in order to convince him? The roulette example is clear: if the probability that Pardis is really beautiful when a woman asserts it is equal to 1/2, then after forty favourable and independent judgments the probability that Pardis is ugly is less than one over one thousand billion. Massud can wisely accept such a result, thinking that there is a higher probability to loose his own sight at the moment of lifting Pardis' veil after marriage due to the concurrent attack of two bumblebees of the steppe, one for each eye. Note how powerful the method is. Be the suspicious fiancé unsatisfied with the value of one over one thousand billion, he could interrogate with little effort a greater number of women, until the probability of ugliness drops below any desired value, however small it might be.

The nature of a well-designed randomised algorithm is thus revealed: using independent random choices we can reach correct solution with certainty in short time (Pardis is unpleasant). The solution may be wrong (Pardis is beautiful) but this happens with an arbitrarily small probability. Should an error occur, however, the algorithm gives no indication of it, and we would be left with no other alternative but suicide to avoid living accompanied by such a bad luck.

Although Masud decided to ascertain the beauty of his bride using probability calculus, not all grooms trust mathematical reasoning. Instead the above concepts may be applied unchanged to the solution of an equally important but less dramatic problem arising for example in the construction of secret codes: the determination of whether a very large arbitrary number N (say of one thousand digits) is prime.[7] Since

[7] We recall that a *prime number* (or simply a *prime*) is a positive integer that has no integer divisors but itself and unity. Otherwise the number is *composite*. For instance 5 is prime, while 6 is composite because can be divided by 2 and 3. The smallest primes are: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, … Euclid showed that there are infinite primes. Gauss and others showed that primes are encountered frequently among the integers. Primes are the unquestioned main characters of Number Theory: their fascination depends to a large extent upon the extreme difficulty of capturing their essence, even if

all known deterministic procedures to solve the problem are exceedingly slow we can invoke a fast randomised algorithm whose answer could be wrong with an arbitrarily small probability. The first author of such an algorithm was Michael Rabin, a famous mathematician who apparently never met Masud.

Not to intimidate the reader we merely present the general structure of Rabin's algorithm whose mathematical details are in fact not simple. The basic idea is the one of executing a random move, in fact, choosing a random number K, and checking whether N is prime by making a test T whose result depends on K and N (K is the choice of a woman, N is Pardis, T is the woman's opinion of Pardis). As T could give a wrong answer on the nature of N, due to the intrinsic randomness of K, the test should be designed so that the error probability is kept under control. For our purpose T must always give a right answer if N is composite (Pardis is unpleasant), and it may give a wrong answer with probability at most 1/2 (i. e., for no more than one half of the possible choices of K) when it asserts that N is prime (Pardis is beautiful). The merit of Rabin was to discover a sophisticated mathematical test with the above properties that can be performed very rapidly.

The structure of the algorithm should now be clear. An integer K is chosen at random, and the test T is applied to K and N. If T answers that N is composite the algorithm halts with certainty; if T answers that N is prime, such an answer can be wrong with probability at most 1/2. We repeat the experiment many times, with random choices of K. If, for one of them, the test answers that N is composite, we take this answer for sure and stop. If we do not obtain such an answer for (say) forty consecutive times, we can content ourselves by saying that N is prime with an error probability less than one over one thousand billions. Even the most pessimistic user can perform a sufficient number of trials to guarantee an extremely small error probability: smaller, for instance, than the probability of the algorithm failing for any other reason such as a hardware crash.

Algorithm designers have discovered the power of Chance only recently and have thought us how to take advantage of it. Many other professionals, however, could do the same, thereby exercising a scientific control over their decisions. In particular judges often have Chance on their side in courts without even being aware of it, with the exception of judge Bridlegoose who, according to Rabelais, gave his sentences by casting dice. When asked by the President of the High Court for the reasons of his unusual behavior, Bridlegoose answered that he was not doing anything different from what all the other judges were doing:

But when you have done all these fine things, quoth Trinquamelle, how do you my friend, award your decrees, and pronounce judgement?

---

some of their properties have been discovered. In particular nobody knows a formula to express all of them.

Even as your other worships, answered Bridlegoose; for I give out sentence in his favour unto whom hath befallen the best chance by dice, judiciary, tribunian, pretorail what comes first. [8]

Even without being fans of Bridlegoose, we believe that a probabilistic evaluation of testimony would be desirable in all trials. This is not intended for the benefit of judges, as we are not so arrogant to teach them what to do, but for ordinary people who generally know about judicial proceedings from TV movies. Was Alfred Dreyfus really guilty? We believe he was not, as all the various testimonies given against the poor officer were apparently inconsistent. Nevertheless, maybe those testimonies were right. Or other more convincing evidence of his guilt existed but was false. Or maybe it was right, but the judge (secretly a leftist) concealed it and pronounced an apparently unjust sentence, to raise a wave of indignation against the army. This chain of hypotheses could go on without an end. Let us then ask ourselves at which point is it proper to stop a trial and give a sentence with the aid of randomised algorithms.

Assume that, in a liberal body of laws, a testimony in favor of the defendant is sufficient to find him not guilty, and a testimony against him has to be taken with caution. The stricter the proof of guilt, the lower the probability that the witnesses' assertions accusing the defendant are deceptive. Note that the value of such a probability must be assigned by the judge. If n independent testimonies of guilt have been produced, and $p_1, p_2, \ldots, p_n$ are the probabilities that they are false, the probability that the defendant is unjustly condemned is given by the product $p_1 \times p_2 \times \ldots \times p_n$, in the same way that Rabin's algorithm can wrongly declare a composite number to be prime. The responsibility of the judge is now apparent. No sentence is absolutely certain, however, enough testimonies must be accumulated until the probability of error becomes smaller than the probability that other accidents render the sentence vain, for example, a power blackout that interrupts the execution by electrocution. If the testimonies are independent (otherwise we cannot multiply the corresponding probabilities), a small number of them is sufficient to attain a very small probability of error.

Condemning an innocent is very sad. But judges can do better than Bridlegoose who, after all, gave a correct sentence in every other trial.

[8] F. Rabelais, *Gargantua and Pantagruel,* III, 39, transl. By Sir T. Urquhart and P. Motteux, William Benton Publ., Chicago 1952.