A Typology of Imprecision

Solomon Marcus

Romanian Academy, Mathematics
Bucureşti, Romania
E-mail: solomon.marcus@imar.ro

1 Introduction

For about 300 years, scientific language takes its distance in respect to ordinary language, just because the latter is no longer able to face the requirements of rigor and precision of the former. In these conditions, it could look strange to make of imprecision a goal of scientific investigation. The explanation of this paradox is that there is no precise definition of precision, while the difference between mathematics and the other fields is that mathematics is dealing with exact approximations, while in absence of mathematics we cope with ... approximate exactness.

When we relate mathematics to precision, we don't have in view its object of investigation, but only the nature of its approach. We are looking for a precise approach to a world of imprecision. As a matter of fact, poetry too involves a combination of precision and imprecision, as Baudelaire observed: the psychological states determined by a work of art are imprecise, but the means used to obtain this work are precise.

2 Abstraction, Approximation, and Generality

The first type of imprecision is *abstraction*, the move from five apples, five boxes, five stones to "five", with no other specification. The imprecise nature of abstraction is related to the imprecision related to its various ways of instantiation. In absence of abstraction, mathematics is not possible.

The second type (historically speaking) of imprecision considered in mathematics was *approximation*, from its elementary form (approximation

of an irrational number by rational ones), to approximation of the sum of a series by a finite sum, then to approximation of a function by another one (typical example: approximation of a continuous function by polynomials), approximation of the solution of a differential equation, etc. The traditional approximation is that where we are looking for finite approximations of countable infinite structures, then for countable approximations of non-countable infinite structures, etc. The novelty brought by computer science and by the modern mathematics is the increasing importance of approximation of finite structures by means of infinite ones. Typical in this respect is grammatical inference. In formal language theory, the simplest example is the fact that the finite language of powers of x from 1 to n cannot be generated by essentially less than n Chomskian rules (of a bounded size, e.g., regular), while the infinite language of all powers of x going from 1 to infinite needs only two rules in order to generate it by a Chomskian grammar.

The next types of imprecision were *generality* and *genericity*; may b e they are concomitant with *approximation*, but their typical form is related to the birth of algebra, when we replace 5 by x. Let us observe also the importance of generality and genericity in natural languages.

3 Randomness, Vagueness (Fuzziness), and Ambiguity

The next type of imprecision coming into the attention of mathematics was (in the 15th and the 16th century) *randomness*. It gave rise to probability theory, developed by Pascal, Bernoulli, Laplace and others and reaching its modern form in the 20th century, with Kolmogorov.

Already the Greek antiquity, then the modern times brought into attention the *paradox* (*aporia*), which, interpreted as a transgression of one of the principles of Aristotelian logic, is a kind of imprecision, because it involves the simultaneity of two opposite states (in case of transgression of the principle of non-contradiction), or the lack of clarity, when the principle of identity is transgressed, between identity and alterity. Similarly for the transgression of principle of excluded middle, as it appears already in non-Euclidean geometry.

"Vagueness" was the title of a famous article by Black (1937) and we adopt the hypothesis according to which its mathematical model is Zadeh's fuzziness (1965). About their equivalence, we bring as argument the view proposed by Frege (1903): "The concept must have a sharp boundary. To the concept without a sharp boundary there would correspond an area that does not have a sharp boundary-line all around".

Surprisingly, *ambiguity*, one of the most frequent form of imprecision, st ill has no general mathematical representation. The word is so popular that it may be a label for any type of imprecision. A reason of this gap could be the fact that "ambiguity" itself is very ambiguous; it may mean "non-specificity"; "dissonance" or "confusion"; "loss of information", etc. (see, in this respect, Klir, 1987). Specific types of ambiguity were however investigated; see Empson (1930) for literary-artistic ambiguity and Marcus, ed. (1981, 1983) for contextual ambiguity, while ambiguity in formal grammars is presented in Rozenberg-Salomaa (eds.) (1997).

4 Learn to Combine Various Types of Imprecision

For a combination of some of the above considered types of imprecision, consider the statement "If it will not rain and if it will not be cold, I will wait for you tomorrow around 5 p.m., near the corner of the University". We have here randomness (rain), fuzziness (cold), approximation (around 5 p.m.), and ambiguity (if the building of the University has several corners). This is the typical situation with imprecision. Life situations bring together several types of imprecision and we have to learn to cope with them in combination, not only each of them on its own, as it happens in most cases.

5 Negligibility, Indiscernibility and Roughness

Negligibility concerns the global behavior of a set, of a function, of a class of sets or of functions or of another entity for which the local-global distinction makes sense. Negligibility refers usually to cardinality, measure or topology, but other types are also possible. Examples: A real monotonous function defined on the real interval $[a, b]$ is continuous in each point of $[a, b]$, except a countable set; it is differentiable in each point of $[a, b]$, except a set of Lebesgue measure zero. Any real function defined on $[a, b]$, which is in $[a, b]$ the limit of a sequence of continuous functions is continuous in each point of $[a, b]$, except a set of first Baire category (i.e., which is a countable union of rare sets). In the first example, the exceptional set is negligible in respect to cardinality, in the second example it is negligible in respect to Lebesgue measure, in the third example it is negligible in respect to Baire category (which is of a topological nature). In the theory of formal languages, "regular" may be considered negligible in respect to "non-regular context-free", while "context-free" may be considered negligible in respect to "non context-free, but context-sensitive", negligibility being here considered in respect to Chomsky hierarchy of languages. Finite sub-languages of a language L are negligible in respect to L if L is regular, but infinite. It was proved by Marcus–Păun that some theorems concerning convexity of sets, where negligibility is related to measure or to topology, have their counterpart in problems related to convexity in formal languages, where negligibility is considered just in respect to Chomsky hierarchy. Negligibility is considered also in the theory of recursive functions, mainly in respect to topology.

Indiscernibility was approached by Pawlak (1982) by means of his concept of a rough set, today a very well-known topic in Computer Science and in Artificial Intelligence. It is considered in respect to an information system, consisting of a set of objects and a set of criteria (such as size, color, material, weight etc.). Each criterion has various possible values; for instance, color may be green, red, yellow, blue, black, white etc. To each object we associate its values in respect to the considered criteria. Selecting a value for each criterion, we may consider the set A of objects with the respective values; it will be approximated from its interior by the set $int(A)$ of objects having with certainly the respective values and from its exterior by the set $ext(A)$ of objects which possibly have the respective values. We may always assert that $int(A)$ is contained in $ext(A)$; their ordered pair is a rough set. If $int(A) = ext(A)$, then we have a usual set. Taking into account that this approach is based on the idea of similarity (two objects having the same values in respect to all criteria are similar) and observing that similarity is a tolerance relation (i.e., reflexive and symmetric) rather than an equivalence relation, as it is considered in Pawlak's approach, Marcus (1993) considered tolerance rough sets (see synonymy in natural languages and the relation "smaller than" considered by Zeeman in its topology of the brain).

Rough sets showed their relevance in approaching other types of imprecision, such as fuzziness (Pawlak–Skowron 1993), evidence (Shafer 1976; Skowron 1993), inconsistency (Grzymala-Busse 1992) and vagueness (Pawlak 1992).

6 Plausibility, Possibility, Credibility, Uncertainty

The first three of them are favored topics in Logic and in Artificial Intelligence, as it can be seen, for instance, in Klir(1987), Klir and Folger (1988). Uncertainty is among the most fashionable topics; see, for instance, Kline (1980), Klir (1987), Klir and Folger (1988), Smithson (1989). Kline's slogan "Mathematics: the loss of certainty" was preceded by Heisenberg's uncertainty principle (1927) and has been followed by a similar slogan by Prigogine ("La fin des certitudes"). Both physics and mathematics are involved here. Gödel's incompleteness theorem seems to be a basic motivation of this pessimistic view. As a matter of fact, it is not the loss of certainty the right phenomenon we have to face here, but the fact that the previous feeling of certainty was determined by our ignorance; as soon as we became aware of how things happen, the mistake was no longer possible. The same thing is valid in respect to certainty in physics.

But we can go further and ask: is certainty the natural state of human psychic? The answer is rather negative. Human beings are characterized by a state of tension, of restlessness stimulating him to learn more and more and to be more and more creative, in order to diminish the gap between their limitations and life's and world's mystery.

7 Absence of Cohesion and Lack of Coherence

Within a social group G, the link between two parts A and B of G can be

evaluated by the product between the cardinal number of the common part of A and B and the cardinal number of the symmetric difference of A and B (Bunge 1971). In this product, the first factor refers to the common participation, while the second one to the heterogeneity of A and B. From another direction, related to language aspects and to linguistics, we observe that cohesion refers rather to the syntactic aspect, while coherence refers to the semantic one. We can speak of the cohesion of a text if its parts are organically related. The first idea coming in mind is the topological notion of connectedness. Connectedness of what? Lipsky (1974), Saloni–Trybulec (1974) and Brainerd (1977) proposed as a mathematical model of the cohesion of a statement s, the connectedness of the dependency and subordination graph G associated to s (see, in this respect, the last chapter in our *Algebraic Linguistics*, Academic Press, New York, 1967). If we proceed in this way, then the lack of cohesion is measured by the smallest positive integer n such that there exist, in G, n arcs with the property that if we delete them, then the obtained graph is connected. When $n = 0$, the statement s has the cohesion property. The proposition (having as its mathematical model a rooted tree) is the simplest example of linguistic cohesion.

We can extend this approach to obtain a model of lack of coherence, if we use an idea by Irena Bellert (1970), permitting to associate to the statement s not only a syntactic graph, as above, but also a semantic graph; it can be obtained by defining a relation of semantic dependency. A term a of s is semantically dependent of a term b of s if the semantic interpretation of a depends on the semantic interpretation of b. The reflexive and transitive closure of this relation leads to the semantic graph H of s. If H is connected, then s is coherent. If not, then a measure of the lack of coherence of s is given by the smallest positive integer n such that there exist n arcs of H with the property that if we delete them from H, then the remaining graph is connected.

For a global view on textual cohesion and textual coherence see S. Marcus, "Textual cohesion and textual coherence", *Revue Roumaine de Linguistique*, 25, 2 (1980), 101–112.

8 Measures of Graduality

The graduality $f(A)$ of a set A, conceived as something opposed to the idea of precise set A (as in Cantorian set theory), should fulfill some intuitive requirements: $f(A) = 0$ iff A is precise; if B is more gradual than A, then $f(A) < f(B)$ or $f(A) = f(B)$; $f(A)$ takes the maximal value iff A is maximally gradual (but let us observe that it may happen that maximality does not exist). A. de Luca and S. Termini (1972) define a measure of graduality taking as a model the entropy:

$$f(A) = -\sum_{x \in X} (m(A,x) log_2(m(A,x)) + (1 - m(A,x)) log_2(1 - m(A,x))),$$

where, if A is not more gradual than B, then $m(A,x) \leq m(B,x)$ when $m(B,x) \leq 1/2$, and $m(A,x) \leq m(B,x)$ when $m(B,x) \geq 1$, for any $x \in X$. Maximal graduality corresponds to the degree of belongingness equal to $1/2$ for any $x \in X$.

Another measure was proposed by A. Kaufmann (1975):

$$f(A) = \sum_{x \in X} (|m(A,x) - m(C,x)|),$$

where C is a precise set for which $m(C,x)$ is equal to zero if $m(A,x) \leq 1/2$, and $m(C,x) = 1$ if $m(A,x) > 1/2$.

A larger class of measures of graduality was proposed by S.G. Loo (1977). According to R.R. Yager (1979), the most natural way to express graduality of A is to require the absence of a sharp distinction between A and its complementary set $c(A)$. If A is gradual, then $c(A)$ is given by a mapping $c : [0,1] \to [0,1]$, associating to each value $m(A,x)$ a value $c(m(A,x))$ expressing the degree of belongingness of x to $c(A)$. The mapping c is required to be non-increasing, to take value 1 in origin and value zero in 1. This means that, if A is precise, then c becomes the classical complementary set. Sometimes it is also required to c to be continuous and involutive: $c(c(a)) = a$ for any a between 0 and 1. See for more G.J. Klir, "Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like", *Fuzzy Sets and Systems*, 24 (1987), 140–160.

It seems that there are distinctions which are not yet considered. It may happen that a graduality is not associated to a phenomenon of imprecision, in respect to the distinction between a property and its negation. The property of water to be liquid is not vague (fuzzy) in Zadeh's sense, because the move from liquid to non-liquid is exactly at zero or at 100 degrees. However, the same move is gradual, the water at 70 degrees is nearer to gaseous state than the water at 60 degrees. By contrast, an explosion of a plain during a flight, following the clash with another plain, is not at all gradual.

9 Confidence, Plausibility and Ignorance

The already mentioned paper by Klir (1987) reminds some other proposed measures of various types of imprecision. Starting from a basic probability m associating to each subset A of X a number $m(A)$ between 0 and 1, such that $m(O) = 0$ and $\sum_{A \subseteq X} m(A) = 1$, one can say that $m(A)$ defines the degree of confidence that a specific element of X belongs to A. The corresponding "measure of confidence" is given by a mapping f associating to each part A of X a number $f(A)$ between 0 and 1, such that $f(A) = \sum_{B \subseteq A} m(B)$. The number $f(A)$ gives the total degree of confidence that a considered element belongs to A or to an arbitrary subset of A.

The "measure of plausibility" is given by a mapping g associating to each part A of X a number $g(A)$ between 0 and 1, such that $g(A) = \sum_{B \cap A \neq \emptyset} m(B)$.. There is a link between the measure of confidence and the measure of plausibility: $g(A) = 1 - f(X - A)$.

The "total ignorance" is expressed by $m(X) = 1$ and $m(A) = 0$ for any A different from X; therefore, $f(X) = 1$ and $f(A) = 0$ for any A different from X and $g(O) = 0, g(A) = 1$ for any non-empty A. As a particular case of the measure of plausibility, we get the "measure of possibility". For bibliographic references and for more details, s ee Klir (1987).

10 Hartley, Shannon, Renyi, Higashi, Klir on Ambiguity

We have already pointed out the ambiguity of the phenomenon of ambiguity. We will call into attention a few cases when the mathematics of ambiguity was successful.

A first approach belongs to Hartley (1928) and refers to ambiguity as non-specificity. It is given by $A(N) = k.log_b N$, where N is the total number of variants involved in a system, while k is a strictly positive constant. For $k = 1$ and $b = 2$, the measure $A(N)$ of non-specificity is evaluated in bits. Alfred Renyi has shown that the mapping $A(N)$ measuring the ambiguity involved in the selection of an element in a set can be structurally characterized by three properties: $A(N \times M) = A(N) + A(M)$ (additivity) $(N, M = 1, 2, 3, ...)$; $A(N) \leq A(N+1)$ (monotonicity); $A(2) = 1$ (normalization).

Other measures of non-specificity were proposed by M. Higashi and G.J. Klir (1982, 1983). Given a normalized distribution $p = (p(x); xinX)$, $max(p(x); xinX) = 1$, of possibilities on X, the measure of non-specificity is given by the integral from 0 to 1 of the logarithm in base 2 of the cardinal number of the section $c(p, x)$ associated with x.

The measure of the classical ambiguity conceived as dissonance or confusion is just Shannon's entropy in respect to a probability distribution.

Ambiguity as loss of information has been investigated by J.L. Dolby (1977). For more details, see Klir (1987).

11 Contextual Ambiguity in Languages and in Medical Diagnosis

Given a finite non-empty alphabet A and a language L on A, the word x on A contextually dominates the word y on A in respect to L if for any two words u and v, such that $uxv \in L$, we have $uyv \in L$. In other words, x contextually dominates y in respect to L if any context accepting $x \in L$ accepts also $y \in L$. The interpretation of this fact is: when x contextually dominates y in respect to L, the contextual ambiguity of x in respect to L is not larger than the contextual ambiguity of y in respect to L. Things become very clear when x and y are of length one, i.e, elements of A, interpreted as the vocabulary of a natural language. Then, in French, for instance, we observe that 'beau' contextually dominates 'douce', but the converse is not true, because 'douce nuit' is well-formed in French, while 'beau nuit' is no longer well-formed. The relation of reciprocal domination is an equivalence relation, in respect to which strings on A are organized in equivalence classes, called 'distributional classes'. To any natural language we can associate the graph of its distributional classes. However, since a natural language is not a precise set (as a matter of fact, its status in respect to the typology of imprecision is not clear, because, for instance, it should be both finite and infinite; see Charles Hockett), we limit discussion to a precise subset of a natural language, i.e., to what is called one of its levels of grammaticality. We also limit the discussion to the distributional classes of the vocabulary of the language, in order to cope with a finite graph. In the graph G so obtained, we have a line from the vertex x to the vertex y if any element in the distributional class x contextually dominates the elements in the distributional class y. The number of different lines leading to the same vertex a is a measure of the contextual ambiguity of words in a, called their "index of ambiguity".

In the monograph *Contextual ambiguities in natural and in artificial languages* (ed. S.

Marcus), volume 1, 1981, volume 2, 1983, Communication and Cognition, Ghent, Belgium, contextual ambiguity is examined in English, French, Hungarian and Romanian and in Fortran IV, Assiris and Pascal. It was shown that in the considered programming languages contextual ambiguity is very poor, while in natural languages it is very rich.

In English, in respect to a level of grammaticality higher than that of noun groups and of verb groups, there are 14 types of contextual ambiguity, represented, in the increasing order of their index of ambiguity, by words like 'receive'(index=1), 'isolate'(index=2), 'often'(3), 'death'(4), 'rain'(5),'see'(6), 'cry'(7), 'iron'(8), 'hang'(9), 'round'(10), 'fly'(11), 'run'(12), 'wash'(13), 'call'(14).

The same mathematical model can be applied to medical diagnosis, if the alphabet A is formed by various possible symptoms of a disease. Words on A include in this case various possible clinical examinations (syndroms). We can define the degree of ambiguity of a diagnosis, having its lower limit in the case of pathognomonic symptoms, able to indicate with no ambiguity a specific disease. See, for more, Eugen Celan - Solomon Marcus, Le diagnostique comme langage, *Cahiers de Linguistique Theorique et Appliquee*, 10, 2 (1973), 163–173.

12 Eubulides, Zadeh and Sugeno on Graduality

The imprecision resulting from absence of a clear (sharp) border between a property and its negation, between a set and its complement has been observed already by the old Greeks; see Eubulides' paradox of the heap of grains and of baldness. However, Zadeh's (1965) fuzziness, where the characteristic function of a set A contained in the total set X is replaced by a mapping associating to each element in X a number in the compact interval $[0, 1]$ (value representing the degree of belongingness to A), does not capture successfully Eubulides' examples. Both fuzziness and roughness remain irrelevant in respect to these old paradoxes, as we have already shown (Marcus 1999).

An alternative to Zadeh's approach was proposed by M. Sugeno (1977), who considers, for each element $x \in X$, a mapping $g(x)$ associating to every part A of X a value $g(x, A) \in [0, 1]$, representing the degree of belongingness of x to A. We have $g(x, O) = 0, g(x, X) = 1$, while for A contained in B $g(x, A)$ is not larger than $g(x, B)$.

In all its variants, the definition of graduality uses non-graduality, i.e., sets which are no longer gradual. For instance, the set of points where Zadeh's mapping takes a definite value is no longer fuzzy, it is a set in the classical sense. If however we allow for it to be fuzzy, then we move the problem to the next level. The last meta-level will always be no longer fuzzy.

Another difficulty is related to the behavior of the values 0 and 1. For most properties, there is no way to assign these values. How could we assign the values 0 and 1 in the case of a property such as 'beautiful' or 'clever'? How can we assign the value 1 in the case of 'non-bald'?

13 Numerical Negligibility and the Crisis of Classical Probability

We have in view a history beginning with Leibniz, who refers to infinitely small quantities,different from zero, but smaller than any number of the form 1/n, where n is a strictly positive integer. Mathematicians in the XVIIIth and XIXth centuries were not able to make meaningful Leibniz's proposal, so they replaced it with another version, where the infinitely small is no longer a fixed quantity, but a function which, in some point, has zero as its limit. One can introduce also the order of an infinitely small. The function f is infinitely small of order n at the point a if the ratio between $f(x)$ and the power n of $x - a$ has the limit zero when x is approaching a. A fundamental problem in mathematical analysis is to make the error of approximation of a function to be infinitely small of order as high as we want. Higher is this order, better is the approximation. A typical example is the approximation of a continuous function by polynomial functions.

The above phenomenon concerns numerical negligibility of various orders. The numerical negligibility conceived by Leibniz received a coherent interpretation from Abraham Robinson, with his non-standard analysis, in the '60th of the XXth century. He conceived a universe larger than the universe of real numbers, under the form of a totally ordered field larger than the field of real numbers; in contrast with the latter, the former is no longer archimedean (i.e., it is no longer true that for any two positive elements a and b there exists a natural number n such that $na > b$). It can be observed that Leibniz's infinitely small is just the negation of Archimede's axiom. To give only one example of a situation requiring just a non-standard approach, let us consider the classical notion of probability, where the probability of a number in $[0, 1]$ to be in a subinterval of length $m < 1$ is just equal to m. Then, what is the probability of the number $x \in [0, 1]$ to be equal to $1/2$? Since $1/2$ is contained in subintervals of length m with m as small as we want, it follows that the only coherent solution is to give to the probability of x to be equal to $1/2$ the value zero. This happens if we remain in the framework of classical analysis and classical probability theory. But we cannot be satisfied with this 'solution', because it is in contradiction with the intuitive fact that x can be really equal to $1/2$. The shortcoming of probability theory, conceived as a measure, is just this gap between impossibility and zero probability; they should be equivalent, but, unfortunately, impossibility implies zero probability, while the converse is not true. The solution is to consider that the probability of x to be equal to $1/2$ is just an infinitely small in the Leibnizian sense, which is possible in the non-Archimedean framework conceived by Robinson.

14 Randomness and Its Intuitive Base

At the beginning of the XXth century, Emile Borel defined an infinite random sequence r on the binary alphabet $\{0, 1\}$ by the property that each of the binary digits 0 and 1 has the same probability of appearance in r. More precisely, denoting by $p(0, n)$ ($p(1, n)$) the probability to have zero (one) in the prefix of length n of r, the limit of $p(0, n)$ ($p(1, n)$) when n tends to infinite is equal to $1/2$ ($1/2$ respectively). Similarly one can define the randomness of an infinite sequence on an arbitrary finite alphabet. This type of randomness was called by Borel 'normality'. Real numbers in their decimal writing are infinite sequences on the alphabet $\{0, 1, 2, ..., 9\}$. The basic result obtained by Borel was that almost all real numbers are normal. 'Almost' means here 'except a set of Lebesgue measure zero'. On the other hand, it was shown later, by J.C. Oxtoby and S. Ulam (*Annals of Math.*, 42, 1941, 874–920) that the set of non-normal real numbers, which is negligible in respect to Lebesgue measure (according to Borel's theorem) is no longer negligible in respect to Baire category. In other words, normality is an exceptional property, in the sense that normal numbers form a set of first Baire category (i.e., it is a countable union of rare sets), while non-normal numbers no longer have this property. This fact shows how misleading can be, from an intuitive viewpoint, the mathematical terminology.

There is also another trap of theorems like those by Borel and Oxtoby-Ulam: they give an information of a global nature, having no counterpart from a local viewpoint. For instance, for most familiar numbers such as square root of 2 or number π we are ignorant whether they are or not normal. This feature is characteristic for most, if not all theorems involving global negligibility.

Defining randomness as normality proved to be unsatisfactory, as it can be seen on the sequence obtained by writing, in lexicographic order, all finite sequences of length 1, 2, 3, ... on a given alphabet. The resulting infinite sequence will be obviously normal, despite the fact that it was written according to a very clear rule. Such a situation is in conflict with the most primitive representation of randomness, as absence of rule. The considered example shows also why a stronger requirement such as to require to all sequences of the same length on the considered alphabet to have the same probability of appearance in the considered infinite sequence is also not acceptable as a mathematical model of randomness. Richard von Mises (1919) proposed then to impose Borel's property (the law of large numbers) also to some subsequences, according to some rules. But this restriction too proved to be insufficient. Some next proposals (A. Wald, 1937; A. Church, 1940) tried to impose to the selection rules proposed by von Mises a constructive character. The critical analysis of all these attempts, made by Michiel van Lambalgen (Von Mises's definition of random sequences reconsidered, *J. of Symbolic Logic*, 52, 1987, 3, 725–755) reaches the conclusion that in the framework of classical mathematics (Platonistic, as he calls it) there is no satisfactory possibility to formalize randomness. Trying however to recuperate von Mises' intuitions, Lambalgen finds that the most acceptable solution is to adopt the approach based on sequential tests, proposed by P. Martin-Loef (1966) and studied further by C.P. Schnorr (1971). If an infinite sequence is random according to Martin-Loef, then almost all its subsequences are also random.

Progress in this direction is looking for more and more weaker conditions, that however still refuse to be sufficient conditions of randomness. Pure(total) randomness seems to can be only approximated, asymptotically approached.

What is, intuitively speaking, randomness? equal preference, absence of rule, total imprevisibility, highest possible complexity ? Can they be simultaneously satisfied? It seems that the answer is negative.

15 Disorder as Entropy versus Order as Information

In a thermodynamic perspective, information was identified, in the second half of the XIXth century, with order and organization, as opposed to disorder, chaos and entropy. The notion of entropy, introduced by Clausius and reconsidered by Boltzmann and Helmholtz, leads to the evaluation of the thermodynamic order as the difference between the maximum possible entropy and the real entropy; just this order expresses the thermodynamic meaning of information. The second principle of thermodynamics indicates the trend of the physical world towards the increasing of entropy, i.e., of disorder. But, as Prigogine showed, within this ocean of increasing entropy the human being creates an island of decreasing entropy. According to George Birkhoff, the artistic beauty of an object is given by the ratio between its order and its complexity. According to Karl Popper, a statement says about the empirical reality just what it interdicts to it. Both Birkhoff and Popper wrote in the thirties of the XXth century, i.e., 20 years before Shannon's information theory, based on the same philosophy. Information means reduction of disorder.

16 Randomness for Finite and for Infinite Strings

Roughly speaking, according to Kolmogorov and Chaitin, a finite string x over a finite non-empty alphabet A is random if no computer progra m describing x is shorter than x. This definition is related to a way to look at the algorithmic complexity of x, in respect to which randomness expresses the highest possible complexity of x. Larger is the difference between the length of x and the length of the shortest possible computer program describing x, smaller is the algorithmic complexity of x. When this difference is zero, x is random.

Surprisingly, this way to look at complexity is similar to the way a specific type of poetry is considered by some literary critics, as a text in which nothing can be deleted, added or modified and no abstract is possible. So, poetry corresponds to highest complexity.

There is a huge discrepancy between the global and the local behavior of randomness of finite strings. On the one hand, in some sense (which is specified mathematically; see the above idea of global negligibility) most finite strings are random; on the other hand, no instantiation of a random finite string is possible. By analogy, we may have an idea of this phenomenon if we think to what means to consider an arbitrary triangle; as soon as you try to represent it on a piece of paper, it is no longer arbitrary. It is interesting to observe the basic difference between The Kolmogorov- Chaitin approach and Shannon's approach in his 'information theory'. The former is dealing with individual entities, while the latter considers the global aspect. Shannon's theory makes sense in respect to a probability distribution in a system having various possible states, while Kolmogorov and Chaitin refer to individual strings over A.

Starting from randomness in finite strings, we may obtain a natural approach to randomness of an infinite string s over A. We define the randomness of s by requiring the randomness of all prefixes of s.

Infinite random strings on A have nice properties, that may be considered far from our intuitive expectations: If s is random, then any possible finite string over A occurs infinitely many times in s. This means, for instance, if A is the alphabet of English, that s will include infinitely many times the whole work of Shakespeare and of any other English writer. But this fact shows that global randomness implies local non-randomness. On the other hand, if we start with a random finite string x and we consider the infinite string obtained by concatenation of x with itself infinitely many times, $xxx...x...$, then we obtain a periodic infinite string, i.e., a non-random string.

17 The Analytic Approach to Deterministic Chaos

Besides the traditional chaos, associated with probabilistic systems, there is also the deterministic chaos, associated with deterministic systems. Small differences of initial conditions of a dynamical system may lead to huge differences in the behavior of the system. The analytic approach to this phenomenon is possible in the following way:

Let I be a real compact interval and let f be a continuous mapping from I into I. A point $p \in I$ is considered periodic for f if there exists a strictly positive integer n such that the value of the iterated of order n of f in p is equal to $f(p)$. The smallest n with this property is considered the period of p. Given a strictly positive number a, f is considered a-chaotic if there exists a perfect subset S (i.e., a set which is identical to the set of its accumulation points) of I with the property that, for any two distinct points x and y of S and for any periodic point p of f, the the following properties take place: the difference between the iterates of order n of f in x and y has, in absolute value, its inferior limit (when n tends to infinite) equal to zero, while its superior limit remains larger than or equal to a; the difference between the iterates of order n of f in the points x and p has, in absolute value, its superior limit (when n tends to infinite) larger than or equal to a. As it was shown by K. Jankova and J. Smital (A characterization of chaos, *Bull. of Australian Math. Soc.*, 34 (1986), 283–292), this notion is equivalent to that considered previously by Li and J. Yorke in 1975, in order to approach a problem in biology.

As it was shown by J. Smital(Chaotic functions with zero topological entropy, *Trans. Amer. Math. Soc.*, 297 (1986), 269–282), any continuous function f which is chaotic for no a strictly positive has the following property: for any $x \in I$ and for any a strictly positive, there is a periodic point p, such that the difference between the iterates of order n of f in x and p has, in absolute value, its superior limit (when n tends to infinity) strictly inferior to a. In other words, for a continuous mapping f from I into I which is not chaotic, any trajectory can be approximated by cycles. Practically, this behavior cannot be distinguished from the asymptotic periodicity of the trajectories. Non-chaotic continuous mappings can serve as deterministic mathematical models of some real processes.

18 Smale's Horseshoe

Almost any system with a chaotic behavior includes as one of its components a certain dynamical system (or its continuous variant) discovered by Stephen Smale and known under the name 'Smale horseshoe'. Chronologically, this was one of the first dynamical systems where the sensible dependency on the initial conditions has been understood in a rigorous and complete way.

Let us consider a mapping f defined on the interval $[0, 1]$ and associating to each $x \in [0, 1]$ the fractional part of $2x$, i.e., the difference between $2x$ and the largest integer which is not larger than $2x$. The mapping f defines one of the simplest dynamical systems; looking at its functioning, we will be nearer to the understanding of the 'paradox of the deterministic randomness'. Let us represent x in base 2. We have $f(x) = 2x$ if $x < 1/2, f(x) = 0$ if $x = 1/2, f(x) = 2x - 1$ if $1/2 < x < 1$ and $f(1) = 0$. It follows, for instance, that $f(0,11) = 0.1$, $f(0,111) = 1.11 - 1 = 0.11$ and, by induction, if $x = 0.11...1$, where 1 appears $n + 1$ times after coma, then $f(x) = 0.11...1$, where 1 appears n times after coma. In other words, the mapping f moves with one digit at right the position of the coma and replaces with zero the first occurrence of 1.

This dynamical system is defined by the iterative application of the mapping f, iteration made possible by the fact that the values of f are situated in the interval of definition of f. The system has inputs and outputs. To an input x between 0 and 1 corresponds an output $f(f...f(x)...)$, where the number of left parentheses is equal to the number of right parentheses and both are equal to the number of applications of the mapping f. Let us consider the input $x = 0.11...1$, where 1 appears 30 times after coma. With each application of f, the number of occurrences of 1 after coma diminishes with one, so, after 30 iterations of f, the 'initial condition', under the form of the input x, will be completely neutralized. In other words, after a sufficient large iterations of f, the result is no longer dependent on the starting point (the initial state of the system) and the behavior gets a random aspect.

This is the type of mechanism explaining the behavior of most chaotic dynamical systems.

19 Chaos in the Evolution of a Population

Let us refer to the evolution of a population of a given species (human beings included). Its evolution in time is described by a mapping iteratively applied to some initial data, related to the state of the considered population at a given initial moment. The first application gives the evolution of the population after one year, n iterations gives the evolution after n years. But how should we choose the form of the mapping? Accepting that the population is growing with some percentage, the same every year, we are lead to a linear function, i.e., of the form $f(x) = ax$, where a is a constant defining the rate of growth. If, for instance, $x = 1000$ and the rate is equal to 1.1, then, after a year, the population will be 1.100; after one more year, 1210 etc. This was the scenario proposed by THomas Robert Malthus in respect to the population growth. There is no room in this scenario for various economic, social, psychological, moral parameters. Already in the first part of the XXth century, Vito Volterra has investigated, by means of the theory of differential equations, the evolution of some species. Researchers agree now that the mapping f should be selected in order to fulfill the following three conditions: rapid growth if the considered population is small; reduction of the growth until some values near to zero, if the population is of an intermediary size; diminution, if the population is very large. A function satisfying these requirements could be $f(x) = ax(1 - x)$. In this case, the population will be expressed by a number between 0 and 1.

To the function above one associates an equation with differences, called the 'logistic equation', considered successively with various modifications, in order to adapt it to various situations. Most authors agreed that, after some growth and some oscillation, a population has a trend of stability around an equilibrium value. THis idea is firmly expressed by J. Maynard Smith ("Mathematical ideas in biology", 1968), according to which populations get stable or oscillates with 'a regular enough periodicity' around an equilibrium point.

From another direction, related to meteorology (Lorenz) and further developed by Smale and Yorke, we reach the mapping already considered (Smale horseshoe). On the other hand, Robert May tried to investigate a population of fishes, by means of the already mentioned logistic equation. Its numerical investigation revealed surprising results. At the frontier between stability and oscillation. For $a = 2.7$, the population proved to be equal to 0.6292. By successive increasing the value of the parameter a, the population is increasing. As soon as the parameter a crosses the value 3, the imaginary population of fishes considered by May begins to oscillate, some times with a period of two years, other times with a period of four years. But beyond some point this periodicity becomes chaos. James Yorke has investigated this behavior in a paper with the significant title "Period three implies chaos" (published in *American Math. Monthly*).

*

Many other things remain to be said. Unfortunately, we have to stop here. For a common denominator of many of these types of imprecision, under the form of conjugate pairs, and for more precise bibliographic references, see our papers: Solomon Marcus: Imprecision, between variety and uniformity: the conjugate pairs. In J.J. Jadacki, W. Strawinski, eds. *In the World of Signs*, Pozna n Studies in the Philosophy of the Sciences and the Humanities 62, 1998, 59–72

, Rodopi, U.S.A. Solomon Marcus, *Controverses in science and in engineering* (in Romanian) Technical Publ. House, Bucharest, 1990.