At the Boundaries of Man’s Power: Play

RENÉ THOM

To understand the world, to act on the world, such are undoubtedly the goals of Science. At first glance, one might think that these two goals are inextricably bound. For in order to act, isn’t it first necessary to have a good grasp of the situation? And inversely, isn’t action indispensable in order to achieve a good understanding of phenomena? Such would undoubtedly be the case if St. Thomas’ renowned “adaequatio rei et intellectus” always reigned in our universe. But the universe in its immensity and the human mind in its frailty are far from always offering us such a perfect fit. There are many examples of perfectly understood situations in which there is no possible course of action. For example, there is the man on the roof of a flooding house who watches the water rise to inundate him. Inversely, there are situations where one can act efficaciously without understanding why. As a proof of this, one could cite, almost without exaggeration, the entire history of medication. For example, the clinical properties of aspirin have been known and exploited for a long time, but a theoretical explanation on the molecular level has been only recently proposed.

All such circumstances in which there is flagrant inadequation between our possible courses of action on the one hand, and our capacity for analysis on the other, are sources of man’s “unhappy consciousness,” for man attains “happy consciousness,” his full and total realization of personality, only in reasoned action whose goal and efficacy are clearly apparent. The question of man’s power touches, then, on ethics. Epictetus understood this at the beginning of his Enchiridion where he invites us to make the distinction in our affairs between what depends on us (ta eph émin) and what does not depend on us (ta ouk eph émin), for we can be held responsible only for what we do with full understanding. Where we cannot act, there remains only to show courage against misfortune and to accept stoically destiny’s verdict.

This view of things is obviously too simplistic, for the world is teeming with situations in which we can clearly intervene, but without really knowing how the effect of our intervention is going to manifest itself. This is quite evident in social interaction. Who has not, in offering an excuse, uttered words which, in
fact, worsened the blunder? Even with respect to ourselves, it is often difficult to know beforehand the effect of a given decision on our ensuing behavior.

These ambiguous situations are not at all restricted to socio-psychological interaction. They are found just as frequently in the natural sciences, and it is because of them that the concept of “black box” has been created in Systems Theory.

Such a system is contained in a box with opaque walls, and we can know the system only by its explicit interaction with the outside world. This occurs at the inputs and outputs of the system. We shall consider henceforth that to specify an input is to give a point in the Euclidian space $V$ having $k$ dimensions (a system of $k$ real numbers, for example, the system of electrical intensities). In the same manner, the output will be a point in the Euclidian space $Y$ with $p$ dimensions. In principle, we can fix the value $u$ of the input (this is moreover the only theoretical means to act on the system). The system ($S$) then responds to an output $Y$. In general, a change in the value of $u$ leads to a change in the value of $Y$, but it is impossible in most cases to foresee with certainty the value of the variation at the output. Ordinarily there exists on the space $Y$ a gain function: $G: Y \rightarrow \mathbb{R}$ expressing the return which an observer anticipates from a value of $y$ of $Y$. It is a question then of varying the entry $u$ to $(u + \Delta u)$ in order to render maximum the gain $G (y + \Delta y)$.

Man’s zone of influence in nature is not then limited by the abrupt barrier in Epictetus’ formulation, but rather by a thick and fluctuating strip, a string of black boxes whose inputs can be modified, but whose outputs are not immediately foreseeable. This “no man’s land,” the boundary of human action, is the domain of the player.

Challenging the moralist’s fatalism, the player, confronting any situation, thinks that there is always something to be done. One can hardly accuse this player of being irrational, for, as we shall see further on, if man has acquired the power now at his disposition, it is because he has played successfully with his environment. There is little to be gained from fatalism, except, perhaps, a “good death.” Nevertheless, each strategy in an incompletely understood situation obviously implies risk. At this point, the ethical side of play appears, notably from the point of view of responsibility. If someone acts with good intentions but unleashes a real catastrophe, should he be cited for stupidity or for bad luck?

Such considerations reveal a primordial interest in techniques of risk evaluation. An explorative strategy, initially prudent, then more daring, can do a lot to uncover, to localize the “catastrophic” entries of the system, those for which a slight variation at the input leads to an abrupt and disproportionate variation at the output. Current nuclear controversy is implicated in this problematic, with the aggravating circumstance that it is not we who will suffer from our present decisions, but our grandchildren or our great-grandchildren. Finally, in certain particularly perverse cases, the player can find himself, without knowing it, in the situation of the demolition man who, while checking the ground, may inadvertently explode a mine. One can easily find analogues to this situation in certain experimental practices, notably in medicine.
In the final analysis, what justifies the player’s stance is the fact that the only conceivable way to expose a black box is to play with it. Every great technical and scientific success consists of a black box rendered explicit. Take, for instance, the “Galilean epistemological break.” What is at the origin of modern science, what is responsible for its prodigious success is not—as is too often foolishly repeated—the experimental method. For, to the extent that experimentation is a game, man played—and very effectively—long before Galileo, notably with stone, fire and metal. What is important with Galileo is a purely mathematical notion, the notion of function. Unknown in ancient mathematics, the concept of function was forged among Italian algebraists of the XVIth century. It emerged through the XVIIth century to take on its modern explicit form only with Leibniz.

This notion renders a perfect account of a particularly simple type of black box, one in which the input \( u \in Y \) determines completely (and uniquely) the output \( y \in Y \). In other words, these are systems whose past history does not affect the output. They are without memory, therefore without an “internal state.” Once this was understood, there remained only to search out those phenomena which admit of such a description (the movement of heavy bodies, for example). This description comes about more by “thought experiment” than by true experimentation. Although the black box is reduced here to a pure transition mechanism: input \( \rightarrow \) output, it is nonetheless capable of great complexity. For the general system, however, the output is not a unique function of the input. All the past history of the system affects the value of the output. To an input, there can correspond an infinity of possible outputs. Between these extreme cases, there are black boxes for which an input has a finite number of corresponding outputs. This type of black box is governed by the theory of elementary catastrophes.

It is hoped that this newly exposed type of black box will permit the interpretation of a whole series of phenomena which, until now, have resisted classical quantitative analysis.

The hermeneutic approach

It should, then, be clear that the essential task of the scientist is the explicitation [dévoilement] of black boxes. It is a task of interpretation, a hermeneutic task. Science conceived in this way will perhaps be in a better position to counter the cold judgement pronounced on it in 1929 by Heidegger: “Science does not think.” The scientist, like the philosopher, is interested in unveiling [dévoilement], in etymology \( \alpha-\lambda \nu\delta\epsilon \eta \alpha \). This is clearly something other than the simple statement of brute fact.

What are the useful techniques for this task of interpretation? The analytic-reductionist method is the one closest at hand. It consists in breaking the black box in order to see what is inside. This brutal method, quite military in spirit, seems to have had as protagonist Alexander cutting the Gordian knot. Indisputably it has enjoyed some success. But to be truly effective, the process requires that certain conditions be verified. It is first necessary that the
elements resulting from the destruction of the system be stable, reproducible and therefore identifiable. It is then necessary for the internal dynamic of these elements to be sufficiently transparent so that it can be formalized with rather close approximation. This condition is paradoxical, for the elements which resisted the release of energy associated with the system's destruction probably have a barrier more opaque, more impermeable, than the initial black box . . . moreover, beware of artifacts! Finally, the interactions between these elements must not lead to an overly complicated graph of interaction, and they must be individually modeled in a quantitative manner. If these conditions are not satisfied, then it is not certain that knowledge of the system's autonomy will yield much information about its physiology. The theoretical stagnation in neurophysiology due to the problem of brain function reminds us that it is not sufficient to grasp how a system is contructed in order to understand how it functions.

To this first brutal method, which has the sometimes major drawback of destroying its own object of study, a gentler method is preferable, one that is more respectful of the facts and beings examined.

This approach, which we shall call “hermeneutic,” consists, grosso modo, in reversing Auguste Comte's famous law of three stages: if the behavior of a system cannot be described by a simple explicitly formulated law, let us first try to describe this behavior qualitatively with respect to tendencies, properties of an abstract character, which control it. And if we are not able to explain the facts by these tendencies, then, as a last resort, let us imagine that a “spirit,” a psychic entity (the ghost in the machine) at least partially controls the system, and let us try to put ourselves in its place.

Confronting this method, one may accuse us of irrationalism, of apparently abandoning all which has given force to modern science. A possible response is that Galilean physicism did not erase Aristotle's “qualities,” but rather hid them in mathematical formalism. If heavy bodies fall, this can be either because they have a tendency to go toward their natural place—the center of the earth—or because they are subject to the potential function $V = + gz$ (where $z$ = height, $g$ = gravity constant). The two explanations are equally verbal, but the second has the advantage of being quantitatively precise. The particle physicist who classically determines the potential function of interaction (the inverse problem) from trajectories produced by collisions (diffusion) does nothing more than expose the “deep tendencies” underlying a whole set of phenomena. A simple linear potential function, for instance, of that of weight: $V = gz$, can be interpreted as the intention of a psychical entity (the tendency toward Aristotle's natural place). Where the tendencies are numerous, conflicting and interwoven, it is very natural to try to organize them in a unique “subjectivity” for which they would be the fleeting components.

Basically, it is necessary to resolve the problem of synthesis (synthesis being defined in general by those mathematical tools called analytical extension or the analytical application of group theory), and where mathematical synthesis fails, there remains only “subjective synthesis,” as Comte himself discovered.

Faced with a local enigmatic situation, universal reason—the logos—is not
sufficient. There must be recourse to cleverness, to that form of wily intelligence which the Greeks called métis\(^1\). Typically, at the outset, every great mathematical success is due to craftiness. This is a paradoxical situation because mathematics, the science of exemplary rationality, progresses more by cleverness than by general methods of broad application.

Now cunning plays a fundamental role in games. It is by cunning reflection that one determines the winning strategies in a game. The hermeneutic task before a particularly enigmatic black box can be compared to a game in which the interpreter and the “spirit in the box” are the players. The interpreter wins when he succeeds in bringing to light the strategy of the system’s internal demon. At that point the black box will be exposed [dévoilée].

This presents another justification of the “subjective” synthesis. As a matter of fact, the semantic study of language reveals that the most complex semantic concept is that of the human individual (localized by a proper noun). If, then, there exists for a given situation, the hope of rendering it intelligible, it will be by comparison to the behavior of a human psychical entity.\(^2\) If this comparison fails, there remains little hope of finding psychological means which would permit subjective simulation of the system’s internal mechanisms.

The following fact is ample proof that such quasi-psychological interpretation, even in pure science, is well founded as the basis of the hermeneutic task: two types of black box used by science go beyond the notion of function. They are the model of elementary catastrophes and statistical interpretation. These are two particular cases of “psychologizing” activity, but to prove the point, let us invoke the general model of Game Theory for two person games.\(^3\)

Each of two players, Pierre and Jean, has at his disposition a space which is his own: \(P\) for Pierre; \(J\) for Jean. To play, each player chooses a point in his space: \(p\) in \(P\); \(j\) in \(J\). Once these choices are made (independently of each other), a third party (the “bank”) determines the respective gains of the two players \(G_p; G_j\) as the functions \(G_j(p;j)\); \(G_p(p;j)\) of the chosen points. The goal of each player is to maximize his gains.

If, however, the presence of Pierre is unknown to Jean, Jean is confronted with a black box whose input is the point \(j\) of \(J\) and whose output is the payoff \([\text{gain}]\ G_j(p;j)\). Turning the problem around, under what circumstances can a black box with a real output be compared to a two-person game?

Two rather classical cases in science sustain this comparison. First, when the black box is of the type defined in the theory of elementary catastrophes, the system has an internal space \(S\) and the point \(s\) which represents a state tending toward a maximum of the potential function \(V(s;u)\). \(S\) then becomes comparable to an internal player (Pierre), whose gain function is precisely the potential function \(V(s;u)\).

Here, the interpretation of the black box is made by granting its internal “spirit” a permanent intentionality defined by the potential function \(V(s;u)\), but this function itself can have a relatively complex psychological structure.

Let us consider now the case of statistical interpretation. To an input \(u\), there corresponds a cloud of points in the output \(Y\). This cloud is centered in a point \(y_o = \varphi(u)\), and for each possible output corresponding to \(u\), we write:
\[ Y = y_0 + \delta \] where \( y_0 \) is the “signal” and “\( \delta \)” the “noise” (\( \delta \) is generally assumed to be small).

The inclusion of noise distribution (\( \delta \)) in a classical measure like that of Gauss is equivalent to saying that in a box, aside from the relation \( u \to \varphi (u) \), there is a demon having a Euclidean space \( S \) in which the point \( s \) chosen by the demon moves in a stochastic manner, in this case according to an ergodic, mixing dynamic which satisfies statistics’ “central limit law” (here, the “law of large numbers”). The payoff \([gain]\) for the demon is then defined by a function \( \Delta = S \to \mathbb{R} \), assumed to be linear. Such a dynamic can be generated in a very deterministic manner, for instance by the so-called Anosov systems. . . . In this case, in attempting interpretation, the player has conferred on his assumed opponent only a very rudimentary psychical entity, that of the drunken sailor whose erratic staggering generates Brownian movement.

Between this case and the permanent intentionality defined by a potential function, there is undoubtedly a whole class of intermediate dynamics to be discovered. Situated between the stubborn determinism of a potential function and the gratuitous spontaneity of “arbitrary” choice, these dynamics will be more apt to simulate the real behavior of the human psychical entity. Perhaps Qualitative Dynamics will be able in the near future to aid in the exploration of this obscure domain.

**The suggestiveness of conflict**

The necessity of “subjective synthesis” presupposes a fundamental psychological phenomenon: the suggestiveness of conflict.

Every manifestly indeterminate natural, or socio-cultural, situation has a considerable attraction for the mind.

This attention given to indeterminacy may become upsetting if the deterministic outcome of the indeterminate process is a potential menace to our safety. On the other hand, if the process has no perceptible effect on us, it still remains a source of considerable fascination. There is undoubtedly a general dynamic explanation for this attraction \([prégnance]\) of chance: every indeterminate situation is analogically simulated by a body in an unstable position, for example, a cone lying vertically on its apex. Indeterminacy resolves itself by passing to a stable state—the fall of the cone on one of its generating lines. In this “catastrophe,” the passage from a high energy level (metastable equilibrium) to a lower stable level frees energy which can, when disseminated in the surrounding environment, provoke secondary catastrophes, by virtue of the “contagiousness of catastrophes” principle. These secondary catastrophes may be perceived as dangerous to our organisms or to our interests.

It is important then to have as exact an understanding as possible of all the various outcomes of an indeterminate situation in order to prevent the occurrence of dangerous catastrophes, or, in any event, to foresee the spatial and causal propagation of possible outcomes. In this examination of possible outcomes, there naturally appears the concept of “tendency.” The choice of the
most probable outcomes corresponds to an innate "tendency" of the system, and it is natural to interpret the initial indeterminate situation as a result of conflict between diverse "tendencies" vying for resolution. The paradigm of all indeterminate situations is "conflict." Man has a very deep phylogenetic and cultural experience of conflict. However, this experience does not nearly exhaust the totality of conflictual situations which can occur in a cultural or natural environment. For this reason, observation of a conflictual situation is always rich in lessons, whether by satisfaction of a realized anticipation or by surprise at an enigma to be resolved. It should also be noted that the interpretation of an uncertain situation as being due to a conflict of antagonistic tendencies moving toward equilibrium often furnishes extremely valuable global intuitions about the process.

As soon as it is recognized that there is a conflict in a process between abstract tendencies, or more generally between anthropomorphic "agents," the "players," one can identify with one of the players and try to imagine winning strategies for these players. However, it is difficult to consider simultaneously the strategies of each player. By taking the side of one or another player, the problem is considerably simplified. It is an empirical fact that it is difficult to take into account simultaneously the interests of all the players in a game. Hence, in chess commentaries, the commentators are generally very hard on the losers. All identify with the winner.

The necessity of taking a side in such a conflict is due partially to our inability to view things objectively. By identifying with a player, one benefits from all the potential affect involved in this identification. This leads us naturally to consider the whole problem raised by play as a cultural activity of man.

**Anthropology of play**

If it is true, as I have maintained elsewhere, that human consciousness exists only in projections [en projet], then to maintain itself in existence in the absence of a natural object, consciousness must forge an imaginary object [fictif] to act upon in order to achieve the state of "happy consciousness" which accompanies every desired and well-planned action. This is already evident among animals. The cat who plays with a ball of string as if it were a mouse projects the image of the prey on an obviously inedible object. (We would be tempted to mention here the "play of nature," spectacular realizations of exceptional dynamic situations where the problematic of indeterminacy returns.) By identifying with one of the agents [actants] in an external conflict, man succeeds in projecting himself onto that external person, and there he lives the conflict by surrogate. More exactly, a first examination of the situation determines whether we are in potential danger. If so, we will identify with the apotropaic tendency which favors a benign outcome, one which spares the organism: for example, to avoid a projectile heading toward us. In such a case, the self is automatically engaged in an attempt at self-defense. In such a
struggle, consciousness will be happy or unhappy according to how it views the chances of success or failure. But let us assume that the conflictual situation reveals no danger for us. We will not for that reason cease to be interested in it, and, most of the time, we will identify with one of the agents in the conflict. This imaginary participation will be for us a source of undeniable pleasure: “Suave, mari magno . . .”

One more step and we end up in those socially organized and pure conflictual situations called games of chance. There, the possible outcomes (like the numbers of a lottery) are equiprobable and perfectly interchangeable. One can identify with an “agent” as abstract as a number only by the procedure of a “wager” which consists in placing some money on an area marked by a corresponding outcome. This is the case of a player who bets on a roulette number. There is hardly any doubt that at the basis of the psychological mechanism of betting there is the seemingly magic desire of the player to affect the outcome of the process. (In a football game, the supporters of a team can affect their team’s morale with loud cheering.)

The relationship between art and play should be mentioned here. The spectacle of the funambulist fascinates us because we accept as our own the acrobat’s struggle against gravity. Might not a work of art simply be a refusal to make any choices? This is at least true of the technological work of art: viaduct or tunnel, the work of art brings about the simultaneous passage of two agents [actants] in conflict, the railroad track on one hand, river or mountain on the other. In each of these cases we recognize the phenomenon of “stabilization of thresholds” developed in Structural Stability and Morphogenesis. The artist resists the fatal outcome, the fall into low level attractors: “un coup de dés jamais n’abolira le hasard.”

Theater suggests itself here as fictional conflict offered as spectacle. As long as there is a reversible situation in the plot, the comic is in view. On the other hand, as soon as irreversible outcomes appear, the comic turns tragic. It is the ruin of the player. If we sense the irreversible, it will be by way of Aristotle’s two tragic emotions: terror, if we are subjugated, or pity, if we retain consciousness of our safety.

The beginning of the irreversible fall is vertigo. Now vertigo is really play only if it is practiced in a periodic and reversible fashion. Otherwise it is the lure of death.

But let us return to play with a hermeneutic function: the game of exposing black boxes. The problem is to find the strategies of varying the input which are most apt to reveal significant behavior in the output. In the absence of pre-existing theory, we are reduced to tinkering without much internal necessity. That is the great vice of modern experimental sciences like biology. The scientist tinkers with a system without any preconceptions: “if I do this, the system reacts by doing that.” Following this method, one could experiment quasi-indefinitely with a system without ever showing a significant behavior of outputs. It follows that one would never succeed in building a valid theory. It might be said in defense of the experimentors that they are pursuing the game of discovery with nature. That would be fine if experimentation had not become relatively
expensive. The scientist should be held responsible for the social cost of his experiments. An experiment which does not lead to the reinforcement of a known theory (or to its rejection) should be judged worthless. Disdain for theory encountered among experimentors has its source in the analytic-reductionist attitude. If we are to discover the right strategy, it is necessary to identify with one of the system’s permanent factors, to take its place in an almost amorous identification. Now how can one love what one has just irreversibly broken?

All modern science is thus founded on the postulate of the mindlessness of things. If this postulate appears rather well founded in physics (where theoretical difficulties most often arise from the infinite number of entities to be considered), it is not so in biology (nor a fortiori in the social sciences). The phenomenon of certain living species adapting to our chemical or biological extermination treatments should make us more humble. Rather than automatically attributing these adaptations to neo-Darwinian chance in the occurrence of quickly realized favorable mutations, one would do better to wonder whether the imitative structure [structures simulatrices] of human intelligence is at work. It is very upsetting to think that in nature there may be occurrences whose behavior could imitate and exceed our own intelligence and thus create an obstacle to our best laid plans. For then our capacity for progress in the exposing [dévoulement] of nature would vanish, and a very sad world, a world without play, truly humanity’s tomb, would come into being. With such a prospect, it is useless to evoke the existence of “extraterrestrials” who might dominate us. It is sufficient to imagine the existence of quasi-Platonic beings of an abstract nature who might play this role. All well informed science should accept this possibility and hold itself ready to meet the challenge.

Institut des Hautes Etudes Scientifiques

NOTES

2. Concerning this, we should remember Oedipus’ answer to the riddle of the sphinx: man. Isn’t this the paradigmatic answer to all riddles?