NAMUR
1958

2

Cybernetica

CYBERNETICA

est la revue de l'Association Internationale de Cybernétique. Elle paraît 4 fois par an.

is the review of the International Association for Cybernetics. It is issued four times a year.

Prix et conditions de vente — Price and conditions of sale.

Abonnement annuel — *Yearly subscription :*

membres de l'Association	150,– F. B.
members of the Association	150,– F. B.
non-membres :	300,– F. B.
non-members :	300,– F. B.

Par numéro — *Each number :*

membres de l'Association	50,– F. B.
members of the Association	50,– F. B.
non-membres :	100,– F. B.
non-members :	100,– F. B.

Toute correspondance concernant la revue est à adresser à l'Association Internationale de Cybernétique, 13, rue Basse Marcelle, Namur (Belgique).

All correspondance concerning the review is to be sent to the International Association for Cybernetics, 13, rue Basse Marcelle, Namur (Belgium).

Secrétaire de Rédaction : M. Roger DETRY.

CYBERNETICA

VOLUME I
N° 2 - 1958

Revue de l'Association Internationale de Cybernétique
Review of the International Association for Cybernetics

NAMUR

SOMMAIRE
CONTENTS

Requisite variety and its implications for the control of complex systems

by W. Ross Ashby,

Director of Research, Barnwood House (Gloucester)

Recent work on the fundamental processes of regulation in biology (Ashby, 1956) has shown the importance of a certain quantitative relation called the law of requisite variety*. After this relation had been found, we appreciated that it was related to a theorem in a world far removed from the biological — that of Shannon on the quantity of noise or error that could be removed through a correction-channel (Shannon and Weaver, 1949; theorem 10). In this paper I propose to show the relationship between the two theorems, and to indicate something of their implications for regulation, in the cybernetic sense, when the system to be regulated is extremely complex.

Since the law of requisite variety uses concepts more primitive than those used by entropy, I will start by giving an account of that law.

I

Variety.

Given a set of elements, its *variety* is the number of elements that can be distinguished. Thus the set

$$\{g\,b\,c\,g\,g\,c\}$$

has a variety of 3 letters. (If two observers differ in the distinctions they can make, then they will differ in their estimates of the variety. Thus if the set is

$$\{b\,c\,a\,a\,C\,a\,B\,a\}$$

* Traduction française : loi de la variété indispensable.

its variety in shapes is 5, but its variety in letters is 3. We shall not, however, have to treat this complication).

For many purposes the variety may more conveniently be measured by the logarithm of this number. If the logarithm is taken to base 2, the unit is the *bit*. The context will make clear whether the number or its logarithm is being used as measure.

Regulation and the pay-off matrix.

Regulation achieves a « goal » against a set of disturbances. The disturbances may be actively hostile, as are those coming from an enemy, or merely irregular, as are those coming from the weather. The relations may be shown in the most general way by the formalism that is already well known in the theory of games (Neumann and Morgenstern, 1947).

A set D of disturbances d_i can be met by a set R of responses r_j. The outcomes provide a table or matrix

		R			
		r_1	r_2	r_3	...
	d_1	z_{11}	z_{12}	z_{13}	...
	d_2	z_{21}	z_{22}	z_{23}	...
D	d_3	z_{31}	z_{32}	z_{33}	...
	d_4	z_{41}	z_{42}	z_{43}	...
	...	...	...	...	...

in which each cell shows an element z_{ij} from the set Z of possible outcomes.

It is not implied that the elements must be numbers(though the possibility is not excluded). The form is thus general enough to include the case in which the events d_i and r_j are themselves vectors, and have a complex internal structure. Thus the disturbances D might be all the attacks that can be made by a hostile army, and the responses R all the counter-measures that might be taken. What is required at this stage is that the sets are sufficiently well defined so that the facts determine a single-valued mapping of the product set $D \times R$ into the set Z of possible outcomes. (I use here the concepts as defined by Bourbaki, 1951).

The « outcomes » so far are simple events, without any implication desirability. In any real regulation, for the benefit of some defined person or organism or organisation, the facts usually determine a further mapping of the set Z of outcomes into a set E of values. E may be as simple as the 2-element set {good, bad}, and is commonly an ordered set, representing the preferences of the organism. Some subset of E is then defined as the « goal ». The set of values, with perhaps a scale of preference, is often obvious in human affairs ; but in the biological world, and in the logic of the subject, it must have explicit mention. Thus if the outcome is « gets into deep water », the valuation is uncertain until we know whether the organism is a cat or a fish.

In the living organisms, the scale of values is usually related to their « essential variables » — those fundamental variables that must be kept within certain « physiological » limits if the organism is to survive. Other organisations also often have their essential variables : in an economic system, a firm's profits is of this nature, for only if this variable keeps positive can the firm survive.

Given the goal — the « good » or « acceptable » elements in E — the inverse mapping of this subset will define, over Z, the subset of « acceptable outcomes ». Their occurrence in the body of the table or matrix will thus mark a subset of the product set $D \times R$. Thus is defined a binary relation S between D and R in which « the elements d_i and r_j have the relation S » is equivalent to « r_i, as response to d_i, gives an acceptable outcome ».

Control.

In this formulation we have considered the case in which the regulator acts so as to limit the outcome to a particular subset, or to keep some variables within certain limits, or even to hold some variables constant. This reduction to constancy must be understood to include all those cases, much more numerous, that can be reduced to this form. Thus if a gun is to follow a moving target, the regulation implied by accuracy of aim may be represented by a keeping at zero of the difference between the gun's aim and the target's position. The same remark is clearly applicable to all cases where an unchanging (constant) relation is to be maintained between one variable that is independent and another variable that is controlled by the regulator.

Thus, as a special instance, if a variable y (which may be a vector) is to be controlled by a variable a, and if disturbance D has access to the system so that y is a function of both the control a and the value of disturbance D, then a suitable regulator that has access to the disturbance may be able to counter its effects, remove its effect from y, and thus leave y wholly under the control of a. In this case, successful regulation by R is the necessary and sufficient condition for successful control by a.

Requisite variety.

Consider now the case which, given the table of outcomes (the pay-off matrix), the regulator R has the best opportunities for success. (The other cases occur as degenerate forms of this case, and need not be considered now in detail).

Given the table, R's opportunity is best if R can respond knowing D's value. Thus, suppose that D must first declare his (or its) selection d_i; a particular row in the table is thereby selected. When this has been done, and knowing D's selection, R selects a value r_j, and thus selects a particular column. The outcome is the value of Z at the intersection. Such a table might be :

		R		
		r_1	r_2	r_3
	d_1	c	a	d
	d_2	b	d	a
D	d_3	c	d	c
	d_4	a	a	b
	d_5	d	b	b

If outcomes a, b count as Good, and c, d as Bad, then if D selects d_1, R must select r_2; for only thus can R score Good. If D selects d_2, R may chose r_1 or r_3. If D selects d_3, then R cannot avoid a Bad outcome ; and so on.

Nature, and other sources of such tables, provides them in many forms, ranging from the extreme at which every one of R's responses results in Good (these are distinctly rare !), to those hopeless situations in which every one of R's responses leads to Bad. Let us set aside these less interesting cases, and consider the case, of central importance, in which each column has all its elements different. (Nothing is assumed here about the relation between the contents of one column and those of another). What this implies is that if the set D had a certain variety, the outcomes in any one column will have the same variety. In this case, if R is inactive in responding to D (i. e. if R adheres to one value r_j for all values of D), then the variety in the outcomes will be as large as that in D. Thus in this case, and if R stays constant, D can be said to be exerting full control over the outcomes.

R, however, aims at confining the actual outcomes to some subset of the possible outcomes Z. It is necessary, therefore, that R acts so as to lessen the variety in the outcomes. If R does so act, then there is a quantitative relation between the variety in D, the variety in R, and the smallest variety that can be achieved in the set of actual outcomes ; namely, *the latter cannot be less than the quotient of the number of rows divided by the number of columns* (Ashby, 1956 ; S.11/5).

If the varieties are measured logarithmically, this means that if the varieties of D, R, and actual outcomes are respectively V_d, V_r, and V_o then the minimal value of V_o is $V_d - V_r$. If now V_d is given, V_o's minimum can be lessened *only by a corresponding increase in* V_r. This is the law of requisite variety. What it means is that restriction of the outcomes to the subset that is valued as Good demands a certain variety in R.

We can see the relation from another point of view. R, by depending on D for its value, can be regarded as a channel of communication between D and the outcomes (though R, by acting as a regulator, is using its variety subtractively from that of D). The law of requisite variety says that *R's capacity as a regulator cannot exceed its capacity as a channel for variety.*

The functional dependencies can be represented as in Fig. 1. (This diagram is necessary for comparison with Figs. 2 and 3).

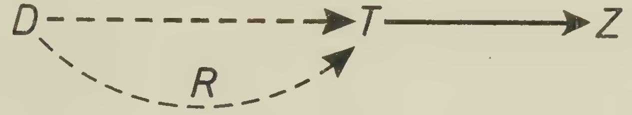

FIG. 1

The value at D threatens to transmit, via the table T to the outcomes Z, the full variety that occurs at D. For regulation, another channel goes through R, which takes a value so paired to that of D that T gives values at Z with reduced variety.

Nature of the limitation.

The statement that some limit cannot be exceeded may seem rash, for Nature is full of surprises. What, then, would we say if a case were demonstrated in which objective measurements shows that the limit was being exceeded ? Here we would be facing the case in which appropriate effects were occuring without the occurence of the corresponding causes. We would face the case of the examination candidate who gives the appropriate answers before he has been given the corresponding questions ! When such things have happened in the past we have always looked for, and found, a channel of communication which has accounted for the phenomenon, and which has shown that the normal laws of cause and effect do apply. We may leave the future to deal similarly with such cases if they arise. Meanwhile, few doubt that we may proceed on the assumption that genuine overstepping of the limitation does not occur.

Examples in biology.

In the biological world, examples that approximate to this form are innumerable, though few correspond with mathematical precision. This inexactness of correspondence does not matter in our present context, for we shall not be concerned with questions involving high accuracy, but only with the existence of this particular limitation.

An approximate example occurs when a organism is subject to attacks by bacteria (of species d_i) so that, if the organism is to survive, it must produce the appropriate anti-toxin r_j. If the bacterial species are all different, and if each species demands a different anti-toxin, then clearly the organism, for survival, must have at least as many anti-toxins in its repertoire of responses as there are bacterial species.

Again, if a fencer faces an opponent who has various modes of attack available, the fencer must be provided with at least an equal number of modes of defence if the outcome is to have the single value : attack parried.

Analysis of Sommerhoff.

Sommerhoff (1950) has conducted an analysis in these matter that bears closely on the present topic. He did not develop the quantitative relation between the varieties, but he described the basic phenomenon of regulation in biological systems with a penetrating insight and with a wealth of examples.

He recognises that the concept of « regulation » demands variety in the disturbances D. His « coenetic variable » is whatever is responsible for the values of D. He also considers the environmental conditions that the organism must take into account (but as, in his words, these are « epistemically dependent » on the values of the coenetic variable, our symbol D can represent both, since his two do not vary independently.) His work shows, irrefutably in my opinion, how the concepts of co-ordination, integration, and regulation are properly represented in abstract form by a relation between the coenetic variable and the response, such that the outcome of the two is the achievement of some « focal condition » (referred to as « goal » here). From our point of view, what is important is the recognition that without the regulatory response the values at the focal condition would be more widely scattered.

Sommerhoff's diagram (Fig. 2) is clearly similar. (I have modified it slightly, so as to make it uniform with Figs. 1 and 3).

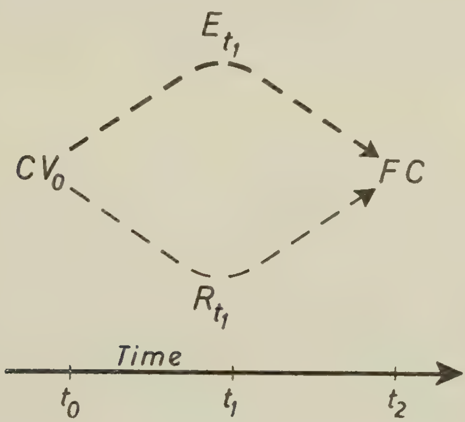

FIG. 2

His analysis is valuable as he takes a great many biological examples and shows how, in each case, his abstract formulation exactly catches what is essential while omitting the irrelevant and merely special. Unfortunately, in stating the thesis, he did what I did in 1952 — used the mathematical language of analysis and continuous functions. This language now seems unnecessarily clumsy and artificial ; for it has been found (Ashby, 1956) that the concepts of set theory, especially as expounded by Bourbaki (1951), are incomparably clearer and simpler, while losing nothing in rigour. By the change to set theory, nothing in fact is lost, for nothing prevents the elements in a set from being numbers, or the functions from being continuous, and the gain in generality is tremendous. The gain is specially marked with biological material, in which non-numerical states and discontinuous functions are ubiquitous.

Let me summarise what has been said about « regulation ». The concept of regulation is applicable when there is a set D of disturbances, to which the organism has a set R of responses, of which on any occasion it produces some one, r_j say. The physico-chemical or other nature of the whole system then determines the outcome. This will have some value for the organism, either Good or Bad say. If the organism is well adapted, or has the know-how, its response r_j, as a variable, will be such a function of the disturbance d_i that the outcome will always lie in the subset marked as Good. The law of requisite variety then says that such regulation cannot be achieved unless the regulator R, as a channel of communication, has more than a certain capacity. Thus, if D threatens to introduce a variety of 10 bits into the outcomes, and if survival demands that the outcomes be restricted to 2 bits, then at each action R must provide variety of at least 8 bits.

Ergodicity

Before these ideas can be related to those of the communication theory of Shannon, we must notice that the concepts used so far have not assumed ergodicity, and have not even used the concept of probability.

The fact that communication theory, during the past decade, has tended to specialise in the ergodic case is not surprising when we consider that its application has been chiefly to telephonic and other communications in which the processes go on incessantly and are usually stationary statistically. This fact should not, however, blind us to the fact that many important communications are non-ergodic, their occurence being specially frequent in the biological world. Thus we frequently study a complex biological system by isolating it, giving it a stimulus, and then observing the complex trajectory that results. Thus the entomologist takes an ant-colony, places a piece of meat nearby, and then observes what happens over the next twenty-four hours, without disturbing it further. Or the social psychologist observes how a gang of juvenile criminals forms, becomes active, and then breaks up. In such cases, even a single trajectory can provide abundant information by the comparison of part with part, but the only ergodic portion of the trajectory is that which occurs ultimately, when the whole has arrived at some equilibrium, in which nothing further of interest is happening. Thus the ergodic part is degenerate. It is to be hoped that the extension of the basic concepts of Shannon and Wiener to the non-ergodic case will be as fruitful in biology as the ergodic case has been in commercial communication. It seems likely that the more primitive concept of « variety » will have to be used, instead of probability ; for in the biological cases, systems are seldom isolated long enough, or completely enough, for the relative frequencies to have a stationary limit.

Among the ergodic cases there is one, however, that is obviously related to the law of requisite variety. It is as follows:

Let D, R, and E be three variables, such that we may properly observe or calculate certain entropies over them. Our first assumption is that if R is constant, all the entropy at D will be transmitted to, and appear at, E. This is equivalent to

$$H_r(E) = H_r(D) \tag{1}$$

By writing $H\ (D,R)$ in two forms we have

$$H(D) + H_d(R) = H(R) + H_r(D)$$

Use of (1) gives

$$\begin{aligned} H(D) + H_d(R) &= H(R) + H_r(E) \\ &= H(R,E) \\ &\leqslant H(R) + H(E) \end{aligned}$$

i.e. $$H(E) \geqslant H(D) + H_d(R) - H(R) \tag{2}$$

The entropy of E thus has a certain minimum — the expression on the right of (2). If $H(E)$ is the entropy of the actual outcomes, then, for regulation, it may have to be reduced to a certain value. Equation (2) shows what can reduce it ; it can be reduced :

(i) by making $H_d(R) = 0$, i.e. by making R a determinate function of D,

(ii) by making $H(R)$ larger.

If $H_d(R) = 0$, and $H(R)$ the only variable on the right of (2), then a decrease in $H(E)$ demands at least an equal increase in $H(R)$. This conclusion is clearly similar to that of the law of requisite variety.

A simple generalisation has been given (Ashby, 1956) in which, when R remains constant, only a certain fraction of D's variety or entropy shows in the outcomes or in $H(E)$. The result is still that each decrease in $H(E)$ demands at least an equal increase in $H(R)$.

With this purely algebraic result we can now see exactly how these ideas join on to Shannon's. His theorem 10 uses a diagram which can be modified to Figure 3 (to match the two preceding Figures).

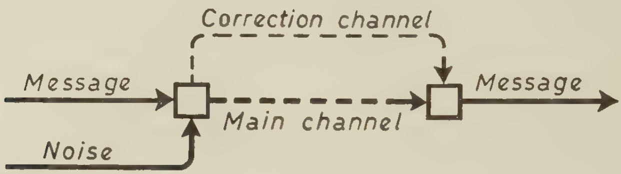

FIG. 3

Our 'disturbance D', which threatens to get through to the outcome, clearly corresponds to the noise ; and his theorem says that the amount of noise that can be prevented from appearing in the outcomes is limited to the entropy that can be transmitted through the correction channel.

The message of zero entropy.

What of the « message » ? In regulation, the « message » to be transmitted is a constant, i.e. has zero entropy. Since this matter is fundamental, let us consider some examples. The ordinary thermostat is set at, say, 70° F. « Noise », in the form of various disturbances, providing heat or cold, threatens to drive the output from this value. If the thermostat is completely efficient, this variation will be completely removed, and an observer who watches the temperature will see continuously only the value that the initial controller has set. The « message » is here the constant value 70.

Similarly, the homeostatic mechanism that keeps our bodies, in health, about 98° F is set at birth to maintain this value. The control comes from the gene-pattern and has zero entropy, for the selected value is unchanging.

The same argument applies similarly to all the regulations that occur in other systems, such as the sociological and economic. Thus an attempt to stabilise the selling price of wheat is an attempt to transmit, to the farmers, a « message » of zero entropy ; for this is what the farmer would receive if he were to ask daily « what is the price of wheat today » ? The stabilisation, so far as it is successful, frees the message from the effects of those factors that might drive the price from the selected value.

Thus, all acts of regulation can be related to the concepts of communication theory by our noticing that the « goal » is a message of zero entropy, and that the « disturbances » correspond to noise.

The error-controlled regulator.

A case in which this limitation acts with peculiar force is the very common one in which the regulator is « error-controlled ». In this case the regulator's channel for information about the disturbances has to pass through a variable (the « error ») which is kept as constant as possible (at zero) by the regulator R itself. Because of this route for the information, the more successful the regulator, the less will be the range of the error, and therefore the less will be the capacity of the channel from D to R. To go to the extreme : if the regulator is totally successful, the error will be zero unvaryingly, and the regulator will thus be cut off totally from the information (about D's value) that alone can make it successful — which is absurd. The error-controlled regulator is thus fundamentally incapable of being 100 percent. efficient.

Living organisms encountered this fact long ago, and natural selection and evolution have since forced the development of channels of information, through eyes and ears for instance, that supply them with information about D before the chain of cause and effect goes so far as to cause actual error. At the present time, control by error is widely used in industry, in servomechanisms and elsewhere, as a means to regulation. Some of these regulations by error-control are quite difficult to achieve. Immersed in the intricacies of Nyquist's theorem, transfer functions, and other technical details, the design engineer may sometimes forget that there is another way to regulation. May I suggest that he would do well to bear in mind what has been found so advantageous in the biological world, and to consider whether a regulation which is excessively difficult to design when it is controlled by error may not be easier to design if it is controlled not by the error but by what gives rise to the error.

This is a first application to cybernetics of the law of requisite variety and Shannon's theorem 10.

It is not my purpose in this paper, however, to explore in detail how the limitation affects simple regulators. Rather I want to consider its effect in matters that have so far, I think, received insufficient attention. I want to indicate, at least in outline, how this limitation also implies a fundamental limitation on the human intellect, especially as that intellect is used in scientific work. And I want to indicate, in the briefest way, how we scientists will sometimes have to readjust our ways because of it.

II

The limitations of the scientist.

In saying that the human intellect is limited, I am not referring to those of its activities for which there is no generally agreed valuation — I am not referring for instance, to the production of pictures that please some and displease others — for without an agreed valuation the concept of regulation does not exist. I refer rather to those activities in which the valuation is generally agreed on, and in which the person shows his capacity by whether he succeeds or fails in getting an acceptable outcome. Such is the surgeon, whose patient lives or dies ; such is the mathematician, given a problem, which he does or does not solve ; such is the manager whose business prospers or fails ; such is the economist who can or cannot control an inflationary spiral.

Not only are these practical activities covered by the theorem and and so subject to limitation, but also subject to it are those activities by which Man shows his « intelligence ». « Intelligence » today is defined by the method used for its measurement ; if the tests used are examined they will be found to be all of the type : from a set of possibilities, indicate one of the appropriate few. Thus all measure intelligence by the *power of appropriate selection* (of the right answers from the wrong). The tests thus use the same operation as is used in the theorem on requisite variety, and must therefore be subject to the same limitation. (D, of course, is here the set of possible questions, and R is the set of all possible answers). Thus what we understand as a man's « intelligence » is subject to the fundamental limitation : it cannot exceed his capacity as a transducer. (To be exact, « capacity » must here be defined on a per-second or a per-question basis, according to the type of test.)

The team as regulator.

It should be noticed that the limitation on « the capacity of Man » is grossly ambiguous, according to whether we refer to a single person, to a team, or to the whole of organised society. Obviously, that one man has a limited capacity does not impose a limitation on a team of n men, if n may be increased without limit. Thus the limitation that holds over a team of n men may be much higher, possibly n times as high, as that holding over the individual man.

To make use of the higher limitation, however, the team must be efficiently organised ; and until recently our understanding of organisation has been pitifully small. Consider, for instance, the repeated attempts that used to be made (especially in the last century) in which some large Chess Club played the World Champion. Usually the Club had no better way of using its combined intellectual resources than either to take a simple majority vote on what move to make next (which gave a game both planless and mediocre), or to follow the recommendation of the Club's best player (which left all members but one practically useless). Both these methods are grossly inefficient. Today we know a good deal more about organisation, and the higher degrees of efficiency should soon become readily accessible. But I do not want to consider this question now. I want to emphasise the limitation. Let us therefore consider the would-be regulator, of some capacity that cannot be increased, facing a system of great complexity. Such is the psychologist, facing a mentally sick person who is a complexly interacting mass of hopes, fears, memories, loves, hates, endocrines, and so on. Such is the sociologist, facing a society of mixed races, religions, trades, traditions, and so on. I want to ask : given his limitation, and the complexity of the system to be regulated, what scientific strategies should he use ?

In such a case, the scientist should beware of accepting the classical methods without scrutiny. The classical methods have come to us chiefly from physics and chemistry, and these branches of science, far from being all-embracing, are actually much specialised and by no means typical. They have two peculiarities. The first is that their systems are composed of parts that show an extreme degree of homogeneity : contrast the similarity between atoms of carbon with the dissimilarity between persons. The second is that the systems studied by the physicist and chemist have nothing like the richness of internal interaction that have the systems studied by the sociologist and psychologist.

Or take the case of the scientist who would study the brain. Here again is a system of high complexity, with much heterogeneity in the parts, and great richness of connexion and internal interaction. Here too the quantities of information involved may well go beyond the capacity of the scientist as a transducer.

Both of these qualities of the complex system — heterogeneity in the parts, and richness of interaction between them — have the same implication : the quantities of information that flow, either from system to observer or from part to part, are much larger than those that flow when the scientist is physicist or chemist. And it is because the quantities are large that the limitation is likely to become dominant in the selection of the appropriate scientific strategy.

As I have said, we must beware of taking our strategies slavishly from physics and chemistry. They gained their triumphs chiefly against systems whose parts are homogeneous and interacting only slightly. Because their systems were so specialised, they have developed specialised strategies. We who face the complex system must beware of accepting their strategies as universally valid. It is instructive to notice that their strategies have already broken down in one case, which is worth a moment's attention. Until about 1925, the rule « vary only one factor at a time » was regarded as the very touchstone of the scientific method. Then R. A. Fisher, experimenting with the yields of crops from agricultural soils, realised that the system he faced was so dynamic, so alive, that any alteration of one variable would lead to changes in an uncountable number of other variables long before the crop was harvested and the experiment finished. So he proposed formally to vary whole sets of variables simultaneously — not without peril to his scientific reputation. At first his method was ridiculed, but he insisted that his method was the truly scientific and appropriate one. Today we realise that the rule « vary only one factor at a time » is appropriate only to certain special types of system, not valid universally. Thus we have already taken one step in breaking away from the classical methods.

Another strategy that deserves scrutiny is that of collecting facts « in case they should come in useful some time » — the collecting of truth « for truth's sake ». This method may be efficient in the systems of physics and chemistry, in which the truth is often invariant with time ; but it may be quite inappropriate in the systems of sociology and economics, whose surrounding conditions are usually undergoing secular changes, so that the parameters to the system are undergoing changes — which is equivalent to saying that the systems are undergoing secular changes. Thus, it may be worth while finding the density of pure hafnium, for if the value is wanted years later it will not be changed. But of what use today, to a sociologist studying juvenile delinquency, would a survey be that was conducted, however carefully, a century ago ? It *might* be relevant and helpful ; but we could know whether it was relevant or not only *after* a comparison of it with the facts of today ; and when we know these, there would be no need for the old knowledge. Thus the rule « collect truth for truth's sake » may be justified when the truth is unchanging ; but when the system is not completely isolated from its surroundings, and is undergoing secular changes, the collection of truth is futile, for it will not keep.

There is little doubt, then, that when the system is complex, the scientist should beware of taking, without question, the time-honored strategies that have come to him from physics and chemistry, for the systems commonly treated there are specialised, not typical of those that face him when they are complex.

Another common aim that will have to be given up is that of attempting to « understand » the complex system ; for if « understanding » a system means having available a model that is isomorphic with it, perhaps in one's head, then when the complexity of the system exceeds the finite capacity of the scientist, the scientist can no longer understand the system — not in the sense in which he understands, say, the plumbing of his house, or some of the simple models that used to be described in elementary economics.

Operational research.

It will now be obvious that the strategies appropriate to the complex system are those already getting well known under the title of « operational research ». Scientists, guided doubtless by an intuitive sense of what is reasonable, are already breaking away from the classical methods, and are developing methods specially suitable for the complex system. Let me review briefly the chief characteristics of « operational » research.

Its first characteristic is that its ultimate aim is not understanding but the purely practical one of control. If a system is too complex to be understood, it may nevertheless still be controllable. For to achieve this, all that the controller wants to find is some action that gives an acceptable result ; he is concerned only with what happens, not with why it happens. Often, no matter how complex the system, what the controller wants is comparatively simple : has the patient recovered ? — have the profits gone up or down ? — has the number of strikes gone up or down ?

A second characteristic of operational research is that it does not collect more information than is necessary for the job. It does not attempt to trace the whole chain of causes and effects in all its richness, but attempts only to relate controllable causes with ultimate effects.

A third characteristic is that it does not assume the system to be absolutely unchanging. The research solves the problems of today, and does not assume that its solutions are valid for all time. It accepts frankly that its solutions are valid merely until such time as they become obsolete.

The philosopher of science is apt to look somewhat askance at such methods, but the practical scientist knows that they often achieve success when the classical methods bog down in complexities. How to make edible bread, for instance, was not found by the methods of classical science — had we waited for that we still would not have an edible loaf — but by methods analogous to those of operational research : if a variation works, exploit it further ; ask not *why* it works, only *if* it works. We must be careful, in fact, not to exaggerate the part played by classical science in present-day civilisation and technology. Consider, for instance, how much empirical and purely practical knowledge plays a part in our knowledge of metallurgy, of lubricants, of house-building, of pottery, and so on.

What I suggest is that measurement of the quantity of information, even if it can be done only approximately, will tell the investigator where a complex system falls in relation to his limitation. If it is well below the limit, the classic methods may be appropriate ; but should it be above the limit, then if his work is to be realistic and successful, he must alter his strategy to one more like that of operational research.

My emphasis on the investigator's limitation may seem merely depressing. That is not at all my intention. The law of requisite variety, and Shannon's theorem 10, in setting a limit to what can be done, may mark this era as the law of conservation of energy marked its era a century ago. When the law of conservation of energy was first pronounced, it seemed at first to be merely negative, merely an obstruction ; it seemed to say only that certain things, such as getting perpetual motion, could not be done. Nevertheless, the recognition of that limitation was of the greatest value to engineers and physicists, and it has not yet exhausted its usefulness. I suggest that recognition of the limitation implied by the law of requisite variety may, in time, also prove useful, by ensuring that our scientific strategies for the complex system shall be, not slavish and inappropriate copies of the strategies used in physics and chemistry, but new strategies, genuinely adapted to the special peculiarities of the complex system.

References.

Ashby, W. Ross, *Design for a brain.* 2nd. imp. Chapman & Hall, London, 1954.

Ashby, W. Ross, *An introduction to cybernetics.* Chapman & Hall, London, 1956.

Bourbaki, N., *Théorie des ensembles. Fascicule de résultats.* A.S.E.I. N° 1141. Hermann et Cie, Paris, 1951.

Neumann, J. (von) and Morgenstern, O., *Theory of games and economic behaviour.* Princeton, 1947.

Shannon, C. E. and Weaver, W., *The mathematical theory of communicacation.* University of Illinois Press, Urbana, 1949.

Sommerhoff, G., *Analytical biology.* Oxford, University Press, London, 1950.

Les manutentions automatisées : leur aspect social et technique.

par Ch. Gilbert ROBBA,

Ingénieur-Conseil en organisation industrielle.
Expert agréé R. K. W. (Deutschland).
Betriebsingenieur bei LIPOWA *Polstermöbelfabrik.*
Edingen-Mannheim (Deutschland).

Qu'il s'agisse des méthodes allemandes (Refa), du Bureau des Temps Élémentaires français, des systèmes Taylor ou Bedeau, le principe de l'étude et de la rationalisation du travail reste le même.

Il faut connaître le prix de la matière première, les temps requis pour la fabrication d'un objet déterminé, la place et les machines nécessaires, la manière d'échelonner le travail, de l'organiser et de le répartir suivant les possibilités et les disponibilités, au mieux des aptitudes de chacun, cadres et personnel, enfin obtenir une coordination du début à la fin entre tous les moyens mis en œuvre.

Ces attributions se placent dans le cadre des activités du personnel technique préposé à l'organisation du travail et à son ordonnancement.

Il est nécessaire que ce personnel puisse prévoir toutes les éventualités pouvant surgir, évaluer le coût de la fabrication, les possibilités d'améliorer les moyens de production et la mise en œuvre de tout ce qui peut contribuer à l'amélioration de la qualité, des prix, des conditions de travail des ouvriers ; c'est là l'aspect industriel et technique de l'organisation.

Parallèlement à cela, il y a l'aspect social et humain : sécurité du personnel, réactions psychologiques et physiques occasionnées par le travail, protection contre les accidents, l'incendie, les explosions, les maladies professionnelles, contrôle de la salubrité et de l'hygiène des locaux et en général tout ce qui peut, à brève ou longue échéance, nuire à la santé physique et morale du personnel. Cela exige de nombreux spécialistes !

Ce travail social aide à maintenir un bon climat, favorable à la bonne marche de l'entreprise et à la productivité.

L'automatisation et la rationalisation ne peuvent négliger ces problèmes, car si elles ont pour but de réduire les prix de revient et d'augmenter le capital de l'entreprise, elles ont aussi un autre but qui est de libérer l'homme des tâches pénibles et monotones et de produire mieux avec un minimum d'efforts, partant de vivre et de travailler plus confortablement. Une amélioration des conditions physiques et morales de l'ouvrier !

Bien entendu, il n'est pas question ici d'entrer dans des considérations politiques ou philosophiques ; le problème est étudié ici très objectivement et de façon strictement technique par rapport à son influence sur la production, rien de plus.

Certes le chef d'entreprise a le droit et même le devoir de viser au développement de l'entreprise ; c'est le but normal et incontesté de toute entreprise industrielle ou commerciale. Le patron cherche avec raison à amortir les investissements, mais il a aussi la responsabilité morale de son personnel ; c'est une question de bon sens et de bons rapports sociaux.

Une rationalisation n'est pleinement rentable qu'avec la participation de tous les éléments de l'entreprise et elle n'est applicable dans de bonnes conditions que si tout le monde s'accorde à s'y intéresser ; pour cela il faut que chacun y trouve son compte.

Avant d'aller plus avant, il est bon d'informer le personnel des buts poursuivis et des avantages à en espérer ; c'est l'aspect d'équipe de la rationalisation. Après plus de 25 ans d'expérience dans ce domaine, j'ai pu constater souvent que les entreprises saines et en plein développement étaient celles qui avaient tenu compte de ces principes.

Contrairement à certaines croyances, colportées par des gens mal informés ou de mauvaise foi, la mécanisation et la rationalisation ne visent pas à diminuer l'ouvrier ; leur but est de permettre de réduire le prix de vente et de mettre à la portée du plus grand nombre de consommateurs des choses qui étaient autrefois inaccessibles à l'ouvrier : télévision, motos, autos, machines à laver, réfrigérateurs, etc...

La machine libère l'homme et lui laisse plus de temps pour penser ; c'est un fait qu'en général et de plus en plus, l'ouvrier cherche à mieux comprendre ce qu'il fait ; il est incontestable que depuis 20 ans, le niveau des connaissances techniques de l'ouvrier a considérablement augmenté.

La mécanisation l'a mis en contact avec des machines compliquées auxquelles il s'intéresse.

L'ouvrier moderne apporte de plus en plus à l'entreprise, non seulement sa présence physique, mais encore son bon sens, ses tours de main, ses idées personnelles qui sont le plus souvent dignes d'être prises en considération et sont utiles pour régler les détails de la fabrication.

Le développement de la technique lui donne de nombreuses occasions d'affirmer ses connaissances et son désir de se familiariser avec les techniques nouvelles, en un mot, il n'est pas systématiquement contre ce qui est nouveau et en particulier contre l'automation et la rationalisation ; il accepte d'y participer à condition qu'elles lui apportent des avantages et que le but poursuivi soit bien défini.

Dans les usines de meubles d'Europe et particulièrement en Allemagne, j'ai pu remarquer que la rationalisation a presque toujours apporté des améliorations de salaires, de sécurité, de logement, de climat social.

La mécanisation et la rationalisation ne sont pas mauvaises en soi ; elles sont, du point de vue technique, indispensables à la rentabilité d'une entreprise et restent sur le plan humain ; comme toutes les actions de l'homme, il y a diverses façons de la comprendre et de l'appliquer.

Depuis quelque temps on parle beaucoup d'automation. Je ne veux pas épiloguer ici sur la logique étymologique de ce barbarisme, je veux seulement considérer le fait en soi. Disons en bref qu'il s'agit là d'une appellation d'origine étrangère employée pour désigner l'ensemble des techniques mises en œuvre pour obtenir une coordination harmonieuse de la production.

On y retrouve les branches professionnelles de la mécanique, de l'électronique et de l'électricité, de la chimie, de la force hydraulique et pneumatique, de l'organisation en général, qui apportent dans la planification et l'organisation du travail les moyens les plus modernes pour rendre le travail plus rentable, plus facile et plus précis.

Si nous voulons comprendre les raisons de son apparition, disons qu'elles sont d'ordre économique par la nécessité d'augmenter la production commerciale, par le rôle qu'elle joue dans la concurrence industrielle, et d'ordre social ; parce qu'elle améliore et facilite les conditions de travail de l'ouvrier.

Il reste un secteur particulièrement intéressant à observer qui se rattache directement à l'automation ; c'est le problème de l'automatisation et de la mécanisation des transports et des manutentions.

Ce secteur nous intéresse particulièrement car c'est là que l'on retrouve les travaux les plus pénibles pour l'ouvrier et les plus coûteux pour l'entreprise ; de ce fait, les uns et les autres ont le même intérêt à voir se développer les efforts faits par les techniciens de la manutention.

Il me semble utile, avant d'aborder ce sujet, de parler d'une autre science se rattachant directement à l'étude du travail, la cybernétique ; c'est une des formes de connaissance et d'analyse rendue indispensable par les problèmes modernes que pose l'organisation rationnelle du travail. C'est l'étude et la science des actes dits « gouvernés » ; c'est une chose très ancienne que Platon lui-même étudiait et qui est aujourd'hui remise en actualité par la technique moderne. Comme l'automation est une nouvelle conception du travail, elle doit faire appel à toutes les connaissances permettant d'analyser le problème et de faire, ensuite, une synthèse des choses, des idées et des moyens les plus favorables à la réalisation des buts poursuivis.

Disons donc que la cybernétique et tout ce qui s'y rattache est un aspect de la rationalisation, l'étude des actes et des mouvements des temps opérationnels, leurs répercussions sur l'organisme humain, leur influence morale et physique sur l'ensemble de l'organisme, en un mot un sondage dans l'individu actif et pensant pour connaître les relations entre les besoins du travail et les possibilités humaines. Le diagramme des dépenses énergétiques, la fatigue du cerveau, du cœur, des nerfs, les limites d'effort, les phénomènes de l'accoutumance, la mécanisation des gestes, les réflexes, doivent être connus et étudiés pour fixer les limites de la résistance du personnel.

Ceci fait aussi partie de la rationalisation du travail. C'est l'étude inséparable du travail, de ses conséquences, des moyens disponibles.

Avant de clore ce chapitre, je veux encore parler de l'accoutumance de l'exécutant à certains gestes et mouvements qui pose des problèmes très importants pour l'organisation du travail :

a) Un travail, au début, nécessite une période plus ou moins longue de mise au courant, période d'information, de coordination, et d'étude ; pendant ce laps de temps la tension de l'exécutant est plus élevée jusqu'au moment où il sera familiarisé avec ce nouveau travail.

b) Vient ensuite une période de rodage au cours de laquelle l'automatisation des mouvements va se développer plus ou moins vite suivant les sujets, parallèlement à cette prise de confiance ; la tension générale va baisser, la fatigue deviendra moins grande.

c) Arrive la période des mouvements synchronisés et automatisés par réflexes conditionnés (loi de Pawlow). Ces réflexes sont si bien ordonnés, dans certains cas, qu'ils ne nécessitent plus qu'une très faible concentration, donc une intervention minimum du cerveau ;

on assiste alors à une sorte de synchronisation des nerfs moteurs et de la vue, en circuit raccourci. Je me souviens, par exemple, avoir vu des ouvrières monter du petit appareillage électrique très compliqué avec une adresse et une facilité apparente déconcertantes ; la sureté du geste dans le choix des pièces placées dans plus de dix boites devant elles, la précision dans le mouvement, la régularité dans le geste, donnaient l'impression d'une automatisation des réflexes si poussée, que l'intervention du cerveau était presque nulle. Il m'a suffi de déplacer une de ces boîtes pour voir se désorganiser en deux secondes ce réseau de réflexes et la femme passer de l'état d'automatisation où elle se trouvait, à une action réfléchie, nécessitant de toute évidence un effort important du cerveau pour rétablir le synchronisme des mouvements.

De telles dépenses énergétiques doivent être connues et mesurées, car elles ont une influence très grande sur l'état physique du personnel.

Du point de vue travail, arrivé à ce stade, il faut se méfier, faire des contrôles fréquents ; l'ouvrier prend trop confiance en ses réflexes et on constate souvent une baisse dans la qualité du travail allant quelquefois jusqu'à l'oubli d'une pièce ou d'une opération importante.

C'est qu'il y a rupture des éléments de contrôle et en particulier l'évaluation du degré de fatigue physique ; le synchronisme du cerveau et des nerfs moteurs est en déséquilibre à tel point que l'ouvrier arrive à perdre la notion du temps et de sa propre limite de dépense physique.

Cela fait partie des attributions des gens chargés de la préparation de l'étude et de la répartition du travail que d'étudier non seulement le temps nécessaire à l'exécution d'un travail déterminé, mais encore le degré d'énergie et de tension qu'il nécessite.

Par la même occasion, disons en passant, que dans les usines où un travail en chaîne est organisé, il est indispensable de respecter une pose de quelques minutes par heure pour ne pas voir la qualité du travail baisser et ne pas créer une fatigue excessive.

* * *

Après cette vue générale du problème de l'organisation du travail, parlons des manutentions automatisées.

Pour la bonne compréhension des exemples, j'ai pris une branche de l'industrie que je connais bien : *les fabriques de meubles*.

Les ennemis numéro 1 dans cette branche sont le manque de place, la proportion énorme des manutentions ! Pour ces deux raisons on peut estimer à 40 % le nombre d'heures improductives.

Il est nécessaire de mieux connaître les ressources offertes par l'ensemble de la technique et des services spécialisés dans la rationalisation du travail pour pouvoir réduire au minimum les temps « morts ».

Les spécialistes de la manutention ont des solutions pour presque toutes les opérations de transport et de manutention, avant et après les fabrications, c'est-à-dire les manutentions et transports *visibles*.

Il existe une multitude de manutentions sournoisement cachées dans le déroulement même des opérations de fabrication ; elles sont beaucoup moins connues et c'est pourtant celles qui coûtent le plus cher et nécessitent le plus de dépense d'énergie.

Ceci est également vrai pour l'industrie du meuble.

Il faut construire des appareils pour effectuer les manutentions cachées, celles qui sont incorporées dans les processus même de fabrication ; elles sont innombrables !

Pour construire ces machines, il faut connaître à fond toutes les phases de la fabrication et c'est pourquoi, depuis quelques années, on a pu voir des ingénieurs se spécialiser dans la rationalisation des fabriques de meubles et en particulier dans l'étude et la construction d'appareils spéciaux de transport et de manutention automatique spécialement adaptés à cette fabrication.

On attend de ceux-ci qu'ils soient en même temps techniciens du bois, constructeurs de machines, spécialistes de la ventilation et du séchage, du laquage, de l'emploi des tissus, et en plus conseils en organisation du travail !

En effet, la variété des produits utilisés, le nombre des fabrications et leur variété, la quantité de matériel employée, sont tels, qu'à chaque minute se posent des problèmes nouveaux.

Les appareils construits devront être *pratiques*, *rentables*, *utilisables pour* l'ensemble des fabrications, rapidement *amortissables*.

Prenons par exemple une fabrication précise, celle des canapés-lits et des fauteuils (meubles rembourés et tapissés). Quelles sont par ordre d'urgence et de rentabilité les opérations à mécaniser ?

a) Transports et manutentions.

b) Avance du travail.

c) Préparation du plan de travail.

d) Approvisionnement des postes de travail.

Je parle ici, bien entendu, plus particulièrement des manutentions cachées.

Il faut, avant tout, éviter les déplacements du personnel ; il ne devra plus courir vers son travail, c'est le travail qui viendra vers lui ;

c'est la ligne de conduite qui a été suivie dans l'élaboration et la construction des appareils dont il est fait description par la suite. Ces appareils ont été spécialement étudiés pour ces travaux et ces fabrications spéciales.

Le principe de base est le suivant.

L'ouvrier a un poste de travail fixe ; le travail est divisé en périodes opératoires de temps égal (compte-tenu des dépenses énergétiques nécessaires pour chaque opération déterminée). Les appareils de transport et de manutention sont automatiques et commandés par un système d'armoire électronique réglable. Ces armoires, dites de progression, sont exactement pré-réglées d'après les temps établis par les services de préparation et de répartition des tâches et des temps de travail ; après ce pré-réglage il n'y aura donc plus à intervenir au cours des fabrications.

Les transporteurs sont mis, au temps prévu, en fonctionnement à la fin de chaque période opérationnelle, amenant à l'ouvrier le travail suivant et poussant vers le poste suivant le travail qu'il vient de terminer.

Autrefois on utilisait des bandes à avances continues, mais pour certains travaux la mobilité est nuisible et cette avance par bonds après temps d'arrêt réglable est de beaucoup plus préférable.

Chaque poste de travail est approvisionné pour tout ce qui lui est nécessaire, par le même principe : bandes transporteuses, annexes à avance par bonds successifs et synchronisées avec les temps d'avance de la bande générale de travail. Tous les transports et les déplacements de personnel sont ainsi supprimés à 100 %.

On a constaté dans ces conditions une amélioration de la prodution, en un mois, de 50 à 70% et une diminution de la fatigue physique de 20 à 30 %.

A titre d'exemple, il fallait autrefois de 250 à 300 minutes pour terminer un canapé-lit ; avec cette nouvelle méthode 115 à 150 minutes (suivant les modèles) suffisent.

Pour un fauteuil normal on admettait autrefois 120 à 200 minutes (suivant les modèles) ; aujourd'hui 45 à 60 minutes suffisent.

On a réalisé aussi une économie des surfaces de travail de 30 à 60 % ainsi qu'une amélioration considérable de la qualité par la spécialisation des postes de travail.

* * *

Disons quelques mots, pour terminer, de la préparation et de la répartition du travail.

Le temps de travail général, pour une pièce déterminée, étant calculé, on le subdivise en postes de travail ; le tout est enregistré par une méthode spéciale, sur un magnétophone à bande ; l'avance de cette bande enregistrée étant constante, il est facile de faire correspondre cet enregistrement au temps de travail prévu (enregistrement consistant en indications de détail concernant chaque poste de travail) ; au cours du déroulement des opérations, ce magnétophone donnera par haut-parleur les indications nécessaires pour chaque poste de travail.

J'ai même combiné, dans certains cas, les indications et des disques enregistrés sur la bande avec musique.

Par un procédé récemment breveté, le magnétophone est synchronisé avec un système de relais électroniques et l'armoire de commande des bandes.

On obtient ainsi, par enregistrement, le plan de travail parlé et chronométré, la commande automatique et synchronisée des bandes de travail et d'approvisionnement, le calcul des temps, le contrôle de la production, enfin sur simple bobine conservée aux bureaux de préparation, une programmation permanente de tous les modèles fabriqués dans la firme, programme établi une fois pour toute.

Un seul opérateur peut donc diriger tout le travail de l'atelier, il lui suffit de placer une bobine sur le magnétophone ; il peut ainsi coordonner tous les services.

Ce système qui semble très compliqué est d'une extrême simplicité.

Le même système est employé pour la préparation des fournitures ; la liste de celles-ci est enregistrée sur une bande et par haut-parleur réclame aux services compétents, avec indication des postes de travail sur lesquels le matériel doit être dirigé ; l'avance des bandes transporteuses est signalée au personnel 10 ou 15 secondes avant la mise en route par les haut-parleurs et par une lampe d'alarme commandée par l'armoire électronique. En réalisant ces appareils, les données énoncées au début de cet article ont été respectées.

Libérer l'ouvrier des tâches ingrates, améliorer la production en quantité et en qualité, assurer un climat social favorable, augmenter les gains, produire plus avec moins d'efforts,... n'est-ce pas là un aspect rassurant et satisfaisant de la mécanisation et de l'automatisation des manutentions ?

CYBERNETICA (Namur)
Vol I — N° 2 — 1958

Cybernétique et physiologie de la conscience.

par Paul CHAUCHARD,

Docteur en Médecine,
Directeur-adjoint, École des Hautes Études (Sorbonne).

L'homme est arrivé à construire des machines douées d'un comportement autonome et suseptibles d'un fonctionnement intrinsèque complexe ressemblant à notre pensée ; on est tenté d'opposer cette pensée mécanique matérielle à la vraie pensée spirituelle. En fait notre pensée est tout aussi dépendante de la matière, la physiologie du cerveau, d'où elle résulte et dont elle est indissociable, tandis que la pensée de la machine ne lui appartient pas en propre, elle est le fruit de l'esprit du constructeur. Il est très remarquable que devant un but commun, l'accomplissement de certaines opérations mentales ou d'un mode de comportement, l'ingénieur, sans le savoir, ait dû recourir à des structures matérielles organisées analogues à celles qui ont été autoconstruites au cours de l'évolution millénaire par la matière vivante douée d'invention inconsciente et d'autoorganisation.

Longtemps, une manière de voir *mécaniste* qui cherchait à localiser les fonctions mentales dans un neurone donné ou un produit cellulaire, nous a empêchés de concevoir comment le cerveau pensait, ce qui conduisait certains à se réfugier dans la fausse solution *idéaliste*, d'une machine au service de l'esprit, ce qui soulève l'insoluble problème des rapports de l'esprit et de la matière. Il est assez curieux que ce dualisme, séparant dans l'unité indécomposable du vivant le cerveau et la pensée, ne différait pas tellement dans son principe du premier dualisme, cette fois mécaniste, distinguant la pensée comme une production distincte du cerveau.

L'objectivité de la neurophysiologie du 20e siècle qui a voulu rester sur son strict terrain en explorant dans le détail la physiologie du cerveau avec les techniques convergentes de Pavlov, de Lapicque, puis l'étude oscillographique des pulsations nerveuses née de la réflexologie de Sherrington, nous a permis de bien comprendre les mécanismes simples de la propagation et de l'aiguillage des pulsations nerveuses dans les structures perpétuellement fluctuantes qui associent entre eux les milliards de neurones de l'écorce et des centres de la base. Rien ne diffère fondamentalement entre l'écorce même humaine et un centre inférieur comme la moelle, sinon la complexité; le plus grand nombre d'éléments et cette complexification, non d'éléments indépendants, mais de neurones interconnectés de façon variable, suffit du point de vue scientifique à expliquer l'*émergence* de propriétés nouvelles ; en particulier aux réflexes innés donnés avec la structure anatomique se surajoutent toutes les possibilités d'associations accidentelles des réflexes conditionnés basées sur la mise en jeu à un degré jamais atteint des processus de mémoire, cette sorte de sensibilisation qui rouvre devant un signal déjà rencontré, les structurations cérébrales correspondantes, conception physiologique de la mémoire toujours matérielle mais combien éloignée des grossiers « magasins d'images » jadis postulés.

Pour passer de cette microphysiologie cérébrale à la pensée qui nous semble un processus spirituel d'un tout autre ordre, la prise en considération des machines à penser s'est montrée très éclairante.

Cerveau et machine transmettent et aiguillent dans des structures modifiables automatiquement par autorégulation, des messages d'information transformés en un code d'information électrique. Plus il y aura d'éléments interconnectés, de possibilités de schémas structuraux, de cablages différents et plus les possibilités fonctionnelles de la pensée seront importantes. On comprend ainsi très aisément la différence cérébrale entre singe et homme ; il manque au singe un nombre suffisant de neurones pour pouvoir disposer d'une pensée humaine et ceci pour l'analyse des messages sensoriels, la commande motrice volontaire notamment celle du langage, comme pour les associations intra-cérébrales, base de la pensée pure.

En dotant l'homme par la complexification des centres cérébraux d'une possibilité infiniment supérieure d'expression vocale articulaire, l'évolution a permis sur les mêmes mécanismes cérébraux élémentaires, l'établissement d'un nouveau code infiniment plus commode pour la pensée, le remplacement des associations d'images de la pensée animale, par les associations verbales du second système de signalisation seules susceptibles d'une pensée abstraite infiniment capable de progrès.

De cette analogie de détail entre fonctionnement cérébral et fonctionnement de la machine qui nous fait comprendre la base matérielle de la pensée, il faut se garder de conclure à une identité et de s'imaginer qu'en augmentant le nombre des pièces de la machine, ce qui en accroitra incommensurablement les pouvoirs, il en résultera l'apparition d'une vraie intelligence, d'une vraie conscience réfléchie. Ce n'est pas en effet n'importe quelle complexification qui peut faire émerger les phénomènes de conscience, la vraie subjectivité. Celle-ci est le propre de la matière vivante, non en vertu de quelque mystérieuse propriété métaphysique surajoutée mais en raison de sa constitution elle-même. La cellule est une machine à comportement douée d'une certaine *individualité*, d'une *présence* active au monde où elle assure sa conservation. Le système nerveux de l'être supérieur pluricellulaire ne fait que lui assurer un comportement plus adapté, une plus grande présence au monde et à lui-même. Psychologie du comportement animal et neurophysiologie cérébrale ont passé par trois stades successifs, d'abord une période préscientifique où l'on utilisait anthropomorphiquement le recours explicatif du vocabulaire psychologique ; on croyait expliquer une réaction en parlant de conscience et de volonté, en assimilant comportement animal et humain, puis la phase d'analyse scientifique a éliminé rigoureusement tout appel au psychologique, toute comparaison de l'animal et de l'homme. Grâce à cette analyse, nous parvenons aujourd'hui à un nouveau stade (1), l'explication physiologique de la prise de conscience et de ses divers degrés, de sa progression dans la série animale de l'unicellulaire à l'homme, en raison des progrès des structures nerveuses intégratrices. Entre la conscience réfléchie humaine liée aux intégrations cérébrales, la conscience moindre d'un mammifère au cerveau moins intégré et non susceptible de verbaliser l'image de son moi pour en faire un « je », support commode de subjectivité, mais détenteur cependant de cette image synthétisée de toutes les impressions du corps et du monde extérieur actuelles et passées, l'obscure conscience inférieure d'un invertébré peu capable de maîtriser ses instincts ou ses tropismes et enfin l'individualité cellulaire, cette humble *bioconscience* qui ne dépend que de l'auto-harmonisation de la matière vivante cellulaire et est si peu intériorité, il y a certes un monde, mais le supérieur apparaît comme le développement progressif par complexification des mêmes structures vivantes, de ce qui existe au stade inférieur, l'émergence étant considérablement réduite si on fait abstraction du language produit social que ne possédait pas l'homme primitif. Parler de bioconscience cellulaire n'est donc en rien parler métaphysique mais exprimer cette réalité scientifique que le biologiste n'a pas affaire dans le fonctionnement cellulaire à des réactions indépendantes, mais à un comportement intégré et que cette intégration est de nature analogue à celle qui grâce au cerveau donnera la pleine conscience humaine.

Aujourd'hui on ne peut plus considérer notre conscience comme un épiphénomène, mais comme le phénomène cérébral essentiel ; on ne peut plus séparer l'homme conscient et l'animal machine, mais on suit dans toute la série animale objectivement et sans anthropomorphisme les progrès de la maitrise et de la prise de conscience, donc de l'intériorité, de la subjectivité.

La biologie qui doit à sa science de ne jamais négliger l'ensemble pour le détail, qui voit donc l'esprit dans le degré d'intégration individuelle, est qualifiée pour retrouver dans la matière inanimée cette « prévie » au sens de Teilhard de Chardin, un degré inférieur d'intégration, dans l'organisation qui est un aspect inférieur du « dedans des choses », de l'intériorité, très inférieur à la bioconscience cellulaire, ce qui s'accorde à la fois avec la philosophie aristotélo-thomiste de l'information homologuant la forme et l'âme, et la conception cybernétique moderne de l'information, assimilée à une négentropie. Mais cette matière inanimée, si organisée soit-elle par l'effort humain, dans des machines complexes, ne pourra faire sortir de vraie conscience que le peu qu'elle en possède dans ses rouages inanimés ; c'est donc une différence de nature qui séparera toujours la machine de l'être vivant et si dans le détail analytique, pensée de la machine et fonctionnement cérébral sont analogues, sur le plan de l'intégration source de conscience il y a, par contre, bien plus de parenté entre une amibe et l'homme qu'entre celui-ci et une machine à calculer.

Le spirituel reste pour la machine un épiphénomène dont elle est le support grâce à l'organisation donnée par le constructeur, tandis que le spirituel est intrinsèque au vivant. On a souvent signalé les différences intimes de fonctionnement entre les cellules nerveuses, les pulsations nerveuses et les tubes électroniques, les signaux des machines. Ce qui est fondamental c'est que la machine est faite de pièces inertes, support *passif* de signaux qui y sont conduits sans participation active, sans présence du rouage à son action ; au contraire dans le cerveau, les rouages vivants qui même au repos doivent être chimiquement actifs, dépenser de l'énergie pour le maintien de leur structure complexe, sont le support d'une réaction à l'excitation, l'impulsion n'est pas conduite passivement, mais est une autoexcitation, une réponse cellulaire. D'autre part la cellule est soumise en même temps à de multiples influences physico-chimiques dont elle fait la synthèse pour donner sa réponse.

Il ne faut pas dans le cerveau considérer seulement le message en cours d'aiguillage, mais en même temps, par d'autres messages actuels évoquant des souvenirs, toute la machine, tout le corps est présent à lui-même sous forme de l'image du corps et c'est la prise en charge éventuelle du message actuel par cette structuration cérébrale globale de l'image du corps dans l'*attention* qui est source de la prise de conscience que ne possède pas la machine. La mémoire des machines est simple circulation en rond d'impulsions ; cette sorte de mémoire élémentaire existe dans le cerveau, mais la vraie mémoire psycho-biologique est d'un tout autre ordre et réside dans une modification de la matière vivante qui la rend apte à reprendre électivement certaines structures fonctionnelles quand les circonstances se reproduisent : il n'y a qu'une possibilité. Ici encore la mémoire de l'unicellulaire si faible par rapport à la mémoire humaine est cependant de nature analogue tandis qu'entre la mémoire des machines et la mémoire humaine, il n'y a qu'une ressemblance superficielle de structures mécaniques.

La machine qui n'est pas limitée dans ses possibilités d'associations de rouages, base de son pouvoir analytique, restera toujours très insuffisante sur le plan de l'intégration n'atteignant jamais à la vraie autointégration organique. On retrouve dans les machines comme le cerveau ou plus généralement l'organisme, des structures localisables dont le fonctionnement global donne naissance à la pensée elle-même non localisée, mais ce n'est que dans l'organisme que l'ensemble possède une intégration qui accède à l'individualité et à la conscience, fonctions organiques quoique non localisables dans une structure matérielle.

Cette différence entre la machine et l'organisme éclate sur le plan *embryologique* : les rouages de la machine sont assemblés par le constructeur ou même une autre machine agissant de l'extérieur ; les cellules se différencient par influences mutuelles autoagissant dans l'organisme : toute l'architecture nerveuse repose sur les interattractions embryonnaires : tout se fait tout seul en vertu des propriétés de sensibilité de la matière vivante. Il y a donc ainsi création d'une société, d'une association harmonieuse de cellules, une intégration des présences individuelles cellulaires, source de la présence individuelle globale. A la passivité des rouages inertes s'oppose l'activité des cellules : une machine en construction ne fonctionne pas, l'organisme en croissance assure sa vie.

Nous pouvons donc *conclure* à la grande fécondité des comparaisons cybernétiques qui nous permettent de saisir, non seulement des modèles artificiels des structures cérébrales, mais plus encore la différence de nature des structures inertes et des structures vivantes. C'est en méditant sur les insuffisances des modèles cybernétiques que Lapicque (2) après avoir étudié toute sa vie les rouages de la machine nerveuse en insistant sur ses fonctions d'intégration, en était venu à insister sur les particularités psychogènes de la matière vivante, voyant dans la conscience proprement dite le résultat de l'intégration des éléments cellulaires de conscience qui ne pourront toujours que manquer à la machine.

Références.

(1) Chauchard, P., *Les mécanismes cérébraux de la prise de conscience.* Masson, Paris, 1956.

(2) Lapicque, L., *Comptes-rendus Académie des Sciences,* Paris, 1952.

Systèmes de Post, algorithmes de Markov.

par Jean PORTE,

Attaché de recherches au Centre National de la Recherche Scientifique (Paris).

1. INTRODUCTION.

Toute étude générale des procédés de calcul mécanique exige l'examen de la question suivante : *Quels sont les problèmes « mécanisables »* ? et, inversement, étant donnés certains procédés de calcul, la question se pose de savoir si ces procédés permettent de traiter tous les problèmes « mécanisables »...... Si par « calcul » on entend le calcul proprement dit d'une fonction arithmétique, l'examen des questions précédentes conduit à définir d'une part les *fonctions récursives* (les fonctions arithmétiques dont le calcul est « mécanisable ») et d'autre part les *machines de Turing* (des procédés de calcul en apparence particuliers, mais qui permettent de calculer toutes les fonctions récursives). Si maintenant par « calcul » on entend tout procédé permettant de passer d'une formule à une autre, des problèmes analogues se posent qui conduisent, en particulier, à la notion d'*algorithme*, introduite dans la science sous une forme précise par A. A. Markov et ses disciples.

L'attention du public scientifique a été attirée récemment sur la notion des « algorithmes de Markov » surtout depuis que Novikov a obtenu un « prix Lénine » pour son mémoire sur l'indécidabilité dans la théorie des groupes (voir Novikov 1955 et Markov 1956) [1], mémoire dans lequel la théorie de Markov joue un rôle essentiel.

Ce fait n'est d'ailleurs pas isolé. Beaucoup d'autres résultats

1. Les noms d'auteurs suivis d'une date renvoient à la liste des « références », qui se trouve à la fin de cet article.

importants dépendant de la Théorie de Markov peuvent être trouvés dans la littérature mathématique soviétique de ces dix dernières années. Malheureusement cette littérature est écrite en langue russe... Personnellement, ne lisant pas le russe, il ne m'a guère été possible de prendre connaissance de la Théorie de Markov et de ses applications qu'à partir des comptes-rendus publiés dans le *Mathematical Reviews* et dans le *Journal of Symbolic Logic*, et des prolongements que cette théorie a suscités chez des auteurs de langue française ou anglaise. Je n'ai donc pas l'intention de donner ici un exposé complet de la théorie de Markov.

Mais cette théorie a ses sources historiques dans les théories développées par Post et par Turing, et elle est en rapport étroit avec la théorie des fonctions récursives, due à Gödel et développée par Kleene et d'autres auteurs. Mon intention est d'exposer ces interconnections d'une façon aussi précise que possible.

On peut dire que ces quatre théories (*fonctions récursives, systèmes de Post, machines de Turing, algorithmes de Markov)* constituent quatre réponses au problème suivant : créer des concepts mathématiques précis adéquats aux notions vagues (mais intuitives) de « fonction d'entiers effectivement calculable » et de « raisonnement effectif » (ou de « raisonnement qui peut être fait de façon mécanique »). Ces quatre réponses sont en fait équivalentes, mais à condition de préciser en quel sens le mot « équivalent » est utilisé. Le meilleur moyen d'exposer avec précision des interconnexions me semble être de plonger toutes ces théories dans une *théorie générale des systèmes logistiques*, et je commencerai par exposer les éléments de cette dernière théorie.

Dans ce qui suit, le § 2 sera consacré à un résumé des notions fondamentales de la théorie des systèmes logistiques ; les §§ 3 et 4, au rappel des définitions et des résultats principaux de la théorie des fonctions récursives, le § 5 à un rappel de la théorie de Turing [1], les §§ suivants aux théories de Post et de Markov.

Les principaux résultats exposés dans cet article sont déjà connus mais j'en donnerai des démonstrations nouvelles et des généralisations.

1. Je ne suppose pas, en principe, que le lecteur connaisse l'article Porte 1956 — mais le § 5 ci-dessous sera plus clair pour ceux qui auront déjà lu les premiers § de cet article.

2. Les systèmes logistiques.

L'idée de systèmes logistiques provient d'un examen du raisonnement mathématique lui-même. Mais cet examen conduit à une abstraction et à des généralisations, de sorte que la théorie des systèmes logistiques, ainsi conçue, n'est plus une étude des raisonnements mathématiques, mais bien une branche particulière des mathématiques.

Si l'on s'astreint à développer un raisonnement de façon tout à fait rigoureuse, c'est-à-dire en énonçant tout ce que l'on se donne au départ et tout ce que l'on se permet, ce raisonnement prend la forme d'une suite de phrases qui possèdent les propriétés suivantes :

(a) chaque phrase est composée d'une suite finie de signes d'un certain alphabet, au moyen de règles données explicitement ;

(b) certaines de ces phrases, explicitées à l'avance, sont admises comme point de départ (ce sont les axiomes) ;

(c) toute phrase du raisonnement est obtenue à partir de certaines des précédentes au moyen de certaines règles de déduction précises énoncées à l'avance.

Il faut encore ajouter, dans les raisonnements mathématiques réels que :

(d) l'alphabet est fini (il y a un nombre fini de signes dans une casse d'imprimeur) ;

(e) on sait décider « effectivement » (ou « mécaniquement ») si une suite de signes est ou non une phrase (sous-entendu : pourvue de sens) ;

(f) il y a un nombre fini d'axiomes et un nombre fini de règles de déduction ;

(g) on sait toujours décider « effectivement » (ou « mécaniquement ») si une phrase donnée peut être déduite de telle, telle et telle phrase donnée par une seule application d'une règle de déduction.

Maintenant, il est clair que ces propriétés des raisonnements sont indépendantes de la forme matérielle des signes de l'alphabet : il est indifférent d'utiliser, par exemple, des trous dans une carte à la place des lettres des alphabets latin ou grec et des signes mathématiques usuels. Ceci nous conduit à la conception abstraite des des systèmes logistiques. *L'« alphabet »* ne sera plus un ensemble de signes concrets, mais un ensemble quelconque, **E**. Les « phrases », qui prendront maintenant le nom de « *formules* » seront certaines suites finies d'éléments de **E** — la notion de *suite finie,* notion fonmentale dans la théorie des systèmes logistiques, est une notion mathématique précise, dont la définition peut être trouvée, par exemple, dans Bourbaki 1942, § 1, n° 2 — [1]. Les « *axiomes* » constituent une partie de l'ensemble des formules. Les « *règles de déduction* » sont des relations n-aires ($n \geqslant 2$) sur l'ensemble des formules. Et ceci traduit abstraitement les conditions (a), (b) et (c).

Il est plus difficile de traduire abstraitement les conditions (e) et (g), où il est question d'« effectivité ». Dans une théorie abstraite générale, il est préférable de ne pas les conserver, quitte à particulariser ensuite la définition générale des systèmes logistiques (voir ci-dessus § 4). Pour des raisons qu'il n'est pas possible d'exposer ici, il est également préférable d'ignorer les conditions (d) et (f) dans la théorie abstraite générale [2].

En définitive nous obtenons les définitions suivantes, dont la première indique seulement des notations utiles :

Définition 1 — Si $a_1,\ldots a_n$ sont des objets quelconques, $<a_1, a_2\ldots a_n>$ est la *suite finie* des objets $a_1\ldots a_n$ pris dans cet ordre. Si **X** est un ensemble quelconque, les suites finies à éléments sans **X** seront dites les *mots* en **X** ; $\check{\mathbf{X}}$ est l'ensemble des mots en **X**. Le nombre d'éléments d'un mot est dit la *longueur* de ce mot. ϕ est le *mot vide* (qui ne contient aucun élément).

Définition 2 — On appelle *système logistique* toute structure [3] :

$$\mathbf{S} = <\mathbf{E}, \mathbf{F}, \mathbf{A}, \mathbf{R}_1,\ldots \mathbf{R}_k\ldots \mathbf{N}>$$

où :

— **E**, l'ensemble de base de la structure, est un ensemble fini ou

1. Une « suite finie » est la même chose qu'un « arrangement avec répétition ». On peut donc dire que la théorie des systèmes logistiques est une partie — très spécialisée — de « l'analyse combinatoire ».

2. On peut choisir entre plusieurs définitions, non équivalentes, pour définir les « systèmes logistiques ». C'est dans la suite de la théorie qu'il apparaît qu'on a choisi une définition à la fois assez générale (pour englober les différentes études particulières) et pas trop générale (de façon à obtenir des propriétés assez « fortes », utiles dans les applications). Je ne garantis pas que la définition adoptée ici se trouve exactement dans ce « juste milieu »...

3. En un sens très voisin de celui de Bourbaki 1939. Mais je considère les structures, non pas comme des éléments de « l'échelle des ensembles » (Bourbaki 1939, § 8), mais comme des suites de tels éléments — ce qui fournit une généralisation utile, en considérant parfois des suites infinies d'éléments de l'échelle des ensembles.

dénombrable quelconque, dit « alphabet ». Les éléments de **E** seront dits des « *symboles* »[1] ;

— **F** est une partie quelconque de $\check{\mathbf{E}}$, dite ensemble des *formules*[2] ;

— **A** est une partie quelconque de **F**, dite ensemble des *axiomes*[3] ;

— $\mathbf{R}_k, \ldots \mathbf{R}_n \ldots$ (en nombre fini ou dénombrable) sont des relations n_1-aires,... n_k-aires... ($n_1, \ldots n_n \geqslant 2$) dites *règles de déduction.* Une règle $(n + 1)$-aire est dite une *règle à n antécédents* ; le dernier argument est dit le *conséquent* de la règle.

— **N** est une numérotation de l'alphabet, c'est-à-dire une application bi-univoque de **E** sur un segment initial [1, 2....] de l'ensemble N*, des nombres entiers positifs (ce segment est identique à N* lorsque **E** est dénombrable).

La numérotation **N** sera utilisée au § 4 ci-dessous. Dès maintenant elle est nécessaire pour nous permettre d'établir une distinction entre les divers symboles[4].

Il est clair que les seuls systèmes logistiques qui pourront être considérés comme des abstractions adéquates des méthodes de raisonnement mathématique, seront ceux qui auront des propriétés traduisant abstraitement les propriétés (d), (e), (f), et (g) ci-dessus. Le rapport entre les systèmes logistiques et les raisonnements est ainsi analogue à celui qui existe entre les structures topologiques et l'espace physique : certaines structures topologiques sont des abstractions adéquates de l'espace physique ; la notion générale de structure topologique est une généralisation utile de ces abstractions[5].

Introduisons quelques définitions destinées à traduire abstraitement les notions concrètes de « déduction », « déductibilité », « théorème ».

---

1. J'emploierai toujours le mot « symbole » dans ce sens abstrait — par opposition au mot « signe », qui désignera les signes concrets (typographiques, par exemple), dans les rares cas où j'aurai à parler de ces derniers.

2. Beaucoup d'auteurs appellent « formules bien formées » ce que j'appelle « formules », et « formules » ce que j'appelle « mots ».

3. Le mot « axiome » aura donc un sens purement abstrait, désignant certaines suites finies de symboles d'un système logistique ; par opposition je pourrai employer le mot « postulat » pour désigner certaines propositions mathématiques concrètes.

4. Ceci remplace la distinction concrète entre les signes concrets, réalisée par le fait que chaque signe a une forme différente de tous les autres. Cette possibilité de « distinguer » se traduit abstraitement par le fait que la structure **S** n'a pas d'autre automorphisme que l'automorphisme identique. Cette dernière propriété est assurée par la numérotation **N**.

5. Voir le § 9 ci-dessous.

Définition 3 — Une formule y est *immédiatement déductible* dans **S** des formules $<x_1, \ldots x_l>$ s'il existe une règle $\mathbf{R}_k$ a l antécédents telle qu'on ait :

$$\mathbf{R}_k \; (x_1, \ldots x_l, \; y)$$

Définition 4 — Une *déduction dans* **S** *(de la formule y a partir des formules* $<x_1, \ldots x_n>$*)* est une suite, finie ou infinie, de formules

$$\Delta = <z_1, \ldots> \qquad \text{ou} \qquad \Delta = <z_1, \ldots z_m>$$

telle que :

1° : Si Δ est finie, z_m est identique a y

2° : pour tout z_p l'une des trois conditions suivantes et satisfaite :

(α) $z_p \in \mathbf{A}$

(β) $z_p \in \{x_1 \ldots x_n\}$

(γ) z_p est immédiatement déductible de certaines des formules qui la précédent dans Δ.

Définition 5 — On dira que y est *déductible dans* **S** *à partir de* $<x_1, \ldots x_n>$ s'il existe une déduction dans **S** de y à partir de $<x_1, \ldots x_n>$. On écrira

$$x_1, \ldots x_n \quad \Vdash_{\mathbf{S}} y$$

si y est déductible dans **S** de $<x_1 \ldots x_n>$ (On écrira simplement « $\Vdash$ » au lieu de « $\Vdash_{\mathbf{S}}$ » si aucune confusion n'est à craindre).

Il est clair que $\Vdash_{\mathbf{S}}$ est une relation binaire entre éléments de $\breve{\breve{\mathbf{F}}}$ et éléments de **F** ; elle est dite relation de déductibilité.

Définition 6 — On dira que y est une *thèse de* **S** et on écrira

$$\vdash_{\mathbf{S}} y$$

($\vdash y$ si aucune confusion n'est à craindre) si

$$\phi \Vdash \;\; y$$ [1]

1. Ainsi les thèses sont les formules déductibles de rien du tout ... Si on se rapporte à la Déf. 4, cela signifie que ce sont les formules qu'on obtient à partir des axiomes en appliquant les règles. C'est bien la notion abstraite correspondant à la notion concrète de « théorème ». Certains auteurs (en particulier Church 1956) appellent « théorèmes » ce que j'appelle « thèses » et « métathéorèmes » les propositions mathématiques concrète que j'appelle « théorèmes ».

L'ensemble des thèses, **T**, est une chose essentielle pour un système logistique donné. Cependant, avec les définitions adoptées, plusieurs systèmes logistiques différents peuvent avoir le même ensemble de thèses. Cela mène à la définition suivante :

Définition 7 — Deux systèmes logistiques seront dits *thétiquement équivalents* (ou *T-équivalents*) s'ils ont même alphabet, mêmes formules, mêmes thèses [1].

On peut affaiblir la notion de T-équivalence en introduisant la notion d'« élargissement », qui jouera un rôle important ci-dessous.

Définition 8 — Soient deux systèmes logistiques :

$\mathbf{S} = \langle \mathbf{E}, \mathbf{F}, \mathbf{R}_1, \ldots \mathbf{N} \rangle$ dont l'ensemble des thèses est $\mathbf{R}$
$\mathbf{S}' = \langle\langle \mathbf{E}', \mathbf{F}', \mathbf{R}_1', \ldots \mathbf{N}' \rangle$ » $\mathbf{R}'$

On dira que **S'** est un *élargissement* de **S** si

1° : $\mathbf{E} \subseteq \mathbf{E}'$
2° : $\mathbf{N}'a = \mathbf{N}a$ pour tout $a \in \mathbf{E}$ [2]
3° : $\mathbf{F} \subseteq \mathbf{F}'$
4° : $\mathbf{T} \subseteq \mathbf{T}'$

Définition 9 — **S'** sera dit un *élargissement conservatif* de **S** si **S'** est un élargissement de **S** et si, de plus

$$\mathbf{T}' \cap \mathbf{F} = \mathbf{T}$$

(c'est-à-dire si toute thèse de **S'** qui est une formule de **S** est aussi une thèse de **S**) [3].

Il résulte des définitions que tout système logistique est uu élargissement conservatif de lui-même, et que, si deux systèmes sont T-équivalents, chacun est un élargissement conservatif de l'autre.

---

1. Une autre suite d'« équivalence » entre systèmes logistiques est l'équivalence déductive : deux systèmes logistiques sont dits D-équivalents s'ils ont même alphabet, même numérotation de l'alphabet, mêmes formules et *même relation de déductibilité*. Deux systèmes logistiques D-équivalents sont aussi T-équivalents, mais la réciproque est fausse.

2. Cette condition est assez restrictive : avec les définitions adoptées ici, elle implique que **E'** ne peut être différent de **E** que si **E** est fini. Cela suffit pour les applications envisagées dans le présent travail. — Dans d'autres travaux il pourra être utile de généraliser la Déf. 2, en admettant que la numérotation peut utiliser des nombres ordinaux transfinis et pas seulement ces nombres finis.

3. D'après Rosenbloom 1950, cf. IV. Rosenbloom appelle « extension », ce que j'appelle « élargissement ».

Les élargissements conservatifs jouent un rôle essentiel dans la théorie de Post.

Pour saisir la connexion entre les systèmes logistiques et les algorithmes de Markov, il est nécessaire d'introduire une nouvelle action :

Définition 10 — On appellera *système dérivationnel* toute structure :

$$\mathbf{S} = <\mathbf{E}, \mathbf{F}, \mathbf{R}_1 \ldots \mathbf{N}, \mathbf{D}>$$

telle que :

1° : $S^o = <\mathbf{E}, \mathbf{F}, \mathbf{R}_1,\ldots \mathbf{N}>$ est un système logistique qui est dit le *système logistique sous-jacent* de **S**.

2° : **D** est une partie de $\check{\mathbf{F}}^*$ (ensemble des suites, finies ou infinies, à élément dans **F**) qui possède les propriétés suivantes :

(α) Tout élément de **D** est une déduction dans $\mathbf{S}^o$

(β) Définissons $\vdash\!\Vdash_{\mathbf{S}}$ à partir de **D**, comme la déductibilité a été définie plus haut à partir des déductions d'un système logistique (Déf. 4), alors, pour toute formule x :

$$\phi \Vdash_{\mathbf{S}^o} x \qquad \text{implique} \qquad \phi \vdash\!\Vdash_{\mathbf{S}} x$$

L'ensemble **D** est dit *ensemble des dérivations* ; la relation $\vdash\!\Vdash_{\mathbf{S}}$ (écrite « $\vdash\!\Vdash$ » si aucune confusion n'est à craindre) est dite *relation de dérivabilité*.

On peut dire qu'on définit un système dérivationnel en « ne gardant que » une partie des déductions d'un système logistique, sous la seule condition que ces dérivations soient suffisantes pour dériver toutes les thèses. Beaucoup d'auteurs appellent « déductibilité » dans un certain système logisitique ce qui est appelé ici « dérivabilité » dans un certain système dérivationnel dont le système logistique constitue une structure sous-jacente (voir par exemple Church 1956, § 13).

Une façon triviale de définir un système dérivationnel **S** à partir d'un système logistique $\mathbf{S}^o$, consiste à prendre **D** identique à l'ensemble des déductions de $\mathbf{S}^o$.

Rien dans la définition 2 n'exclut que **A** soit vide. Si **A** est vide, c'est-à-dire s'il n'y a pas d'axiomes, il est évident qu'il n'y a pas non plus de thèses et par suite tous les systèmes logistiques sans axiomes sont T-équivalents de façon triviale. Mais on peut encore définir la déductibilité dans un tel système logistique, ou utiliser ce système pour définir des systèmes dérivationnels. Aussi, loin d'être sans intérêt, les systèmes logistiques sans axiomes jouent un rôle fondamental dans beaucoup d'études [1], et en particulier dans la définition des algorithmes.

Définition 11 — Un système logistique ou dérivationnel sera dit *simple* si :

1° Il a au plus un axiome.

2° Chaque règle de déduction est une règle à un antécédent.

Un système logistique ou dérivationnel simple sera dit *O-simple* s'il n'a pas d'axiome, 1-*simple* s'il a un axiome.

Définition 12 — Un système dérivationnel sera dit *unitaire* s'il est 0-*simple* et si dans toute dérivation

$$\Delta = <z_1, \dots z_n>$$

il y a une formule et une seule qui n'est pas immédiatement déductible de la formule immédiatement précédente (cette formule est évidemment z_1).

Définition 13 — Un système dérivationnel sera dit *déterministe par rapport à un sous-ensemble* **G** *de* **F** si, pour toute formule $x \in$ **G**, il n'existe pas deux dérivations à partir de $<x>$

$$\Delta = <z_1, \dots z_i \dots , z_k>$$
$$\Delta' = <z'_1, \dots z'_{i1} \dots , z'_k>$$

telle que $z_i \neq z'_i$. Un système dérivationnel sera dit *absolument déterministe* s'il est déterministe par rapport à **F**.

Définition 14 — On appellera *algorithme* tout système dérivationnel unitaire, absolument déterministe, et tel qu'aucune dérivation n'est le début d'une autre [2].

1. Ainsi les systèmes de λ-conversion de Church, (voir Church 1941) sont des systèmes logistiques sans axiome, de même que les systèmes utilisés par Gödel et Kleene pour définir les fonctions récursives (voir Kleene 1952, ch. XI).

2. Cette définition est différente de celle de Markov 1954. Markov emploie le mot « algorithme » (sans qualificatif) dans un sens intuitif et vague, pour désigner des méthodes permettant d'obtenir une formule au moyen d'une autre par des moyens effectifs. L'« effectivité » ne joue aucun rôle dans la Déf. 14. Les algorithmes définis par la Déf. 14 constituent une généralisation des « algorithmes normaux » de Markov 1954, qu'on retrouvera ci-dessous (§ 7, Déf. 41) sous le nom d'« algorithmes bilatéraux ».

Il est équivalent de dire qu'un algorithme est un système dérivationnel.

$$\mathbf{S} = <\mathbf{E}, \mathbf{F}, \mathbf{A}, \mathbf{R}_1, \dots \mathbf{N}, \mathbf{D}>$$

où

1° **A** est vide.

2° $\mathbf{R}_1$ ... sont des règles à un antécédent.

3° Toute dérivation $\Delta = <z_1, \dots z_n> \in \mathbf{D}$ est obtenue en partant d'une formule quelconque z_1, d'où on tire z_2 en appliquant l'une des règles, etc....

4° Deux dérivations commençant par la même formule sont identiques.

Pour définir un algorithme à partir d'un système logistique o-simple, il suffit de préciser :

(a) quand plusieurs règles peuvent être appliquées à la même formule, laquelle de ces règles doit être appliquée ;

(b) quand une règle peut être appliquée de plusieurs façons à une formule donnée, de quelle façon la règle doit être appliquée.

C'est effectivement ainsi que certaines classes d'algorithmes seront définies ci-dessous. Mais, il y a des cas où les précisions *(a)* et *(b)* sont inutiles, une seule règle au plus étant applicable, et d'une seule façon, à chaque formule.

Définition 15 — Dans un système logistique simple, **S**, on appellera *dérivations simples* les déductions, $\Delta = <z_1, \dots>$ telles que

1° pour tout terme z_p autre que le premier, z_p est immédiatement déductible de z_{p-1} ;

2° $\Delta = <z_1, \dots z_n>$ n'est finie que si aucune formule n'est immédiatement déductible de z_n.

Définition 16 — Un système logistique o-simple sera dit *déterministe par rapport à* $\mathbf{G} \subseteq \mathbf{F}$ si le système dérivationnel obtenu en ajoutant l'ensemble $\mathbf{D}_s$ des dérivations simples à **S**, est un algorithme. Un système logistique 1-simple sera dit *déterministe* si le système dérivationnel obtenu en ajoutant à **S** l'ensemble $\mathbf{D}_{sa}$, des dérivations simples dont la première formule est l'axiome, est un algorithme (alors $\mathbf{D}_{sa}$ n'a qu'un élément).

Définition 17 — On appellera *fonction attachée à l'algorithme* **S** l'application, $\mathfrak{F}_{\mathbf{S}}$, d'une partie de **F** dans **F**, définie par

$$\mathfrak{F}_{\mathbf{S}}(z_1) = z_n \quad \text{si et seulement si} \quad < z_1, \dots z_n >$$

est une dérivation de **S**.

3. Les fonctions récursives.

Les fonctions récursives ont été créées dans le but d'avoir une abstraction adéquate à la notion de fonction d'entiers effectivement calculable. On définit des fonctions récursives de une ou plusieurs variables. Je me contenterai de rappeler les définitions et les résultats principaux relatifs aux fonctions récursives d'une variable.

Dans ce qui suit, il sera sous-entendu que *toutes les « fonctions » envisagées sont des applications de l'ensemble* N *des nombres naturels* (entiers positifs ou nuls), *ou d'une partie de* N, *dans* N ; les lettres x, y ... désigneront des nombres naturels quelconques.

Définition 18 — On appellera *opérateur* de μ-conversion une application de l'ensemble des fonctions dans lui-même définie de la façon suivante : soit f une fonction, $\overline{\mu}f$ est une nouvelle fonction définie par :

$$(\overline{\mu}f)x = y \quad \text{si et seulement si}$$

1° $fy = x$

2° pour tout nombre $y_0 < y$, fy_0 existe et $fy_0 \neq x$.

Si, pour un certain x, il n'existe pas de y satisfaisant aux conditions 1° et 2°, alors $(\overline{\mu}f)x$ n'existe pas.

Il s'ensuit que $\overline{\mu}f$ ne peut être définie partout sur N que si f est définie partout et prend toutes les valeurs possibles.

Définition 19 — On appellera *opérateur de μ-cconversion* la restriction de $\overline{\mu}$ aux arguments f qui sont des fonctions définies partout et prenant toutes les valeurs possibles. Autrement dit :

$(\mu f)x = (\overline{\mu}f)x$ pour tout x, si f est une fonction définie partout qui prend toutes les valeurs possibles — et, dans ce cas, il est clair que μf est une application de N entier dans N — si, au contraire, f ne prend pas toutes les valeurs possibles, ou n'est pas définie partout, alors « $\overline{\mu}f$ » n'a pas de sens.

Définition 20 — Considérons les fonctions s (« suivant ») et e (« excès sur un carré ») définies par

$$\left.\begin{aligned} sx &= x + 1 \\ ex &= x - (\text{valeur entière par défaut de } \sqrt{x})^2 \end{aligned}\right\} \text{pour tout } x \in \mathrm{N}$$

Considérons également les opérateurs d'addition et de composition définis par

$$\begin{aligned} (f + g)x &= gx && \text{pour tout } x \in \mathrm{N} \\ (fg)x &= f(gx) && \text{pour tout } x \in \mathrm{N} \end{aligned}$$

Alors on appellera *fonctions récursives de une variable* les fonctions qui peuvent être construites à partir de s et de e par un nombre fini quelconque d'applications des opérateurs d'addition composition et μ-conversion.

Définition 21 — *Les fonctions semi-récursives de une variable* sont les fonctions obtenues à partir de s et e par un nombre fini d'applications des opérateurs d'addition, composition et $\bar{\mu}$-conversion [1].

Ces définitions, d'apparence artificielle, ont l'avantage de la brièveté [2]. Il est clair que toute fonction récursive est « effectivement calculable » pour toutes les valeurs de la variable. La réciproque est connue sous le nom de « *Thèse de Church* » :

— *Toute fonction effectivement calculable pour toute valeur de la variable est une fonction récursive.*

Cette proposition est appelée une « thèse » et non un « théorème » [3], car elle ne peut pas être démontrée, puisqu'elle porte sur une notion (« fonction effectivement calculable ») qui est intuitive et vague. On ne peut que la vérifier expérimentalement. Le principe de cette vérification est le suivant : on cherche une définition aussi générale que possible d'une abstraction qui paraisse adéquate à la notion intuitive de fonction effectivement calculable ; on démontre

1. Dans l'article Porte 1956 j'appelais « semi-fonctions récursives » ce que j'appelle maintenant « fonctions semi-récursives » (« partial recursive functions » dans Kleene 1952, ch. XII). Ce changement de vocabulaire a plusieurs avantages — entre autre celui de faire disparaître une apparence trompeuse de trivialité dans l'énoncé du théorème 4.

2. Elles sont dues à Julia Robinson (v. Robinson 1950).

3. Ici le mot « thèse » est utilisé dans un sens entièrement différent de celui de la Déf. 6 du § 2.

alors que cette notion abstraite se confond avec celle des fonctions récursives [1].

Il est possible et commode, d'utiliser la Thèse de Church comme intermédiaire dans une pseudo-démonstration : on définit une fonction, on constate qu'elle est effectivement calculable, on en conclut (Thèse de Church) qu'elle est récursive, d'où, etc ... L'expérience montre que ces pseudo-démonstrations peuvent toujours être transformées en démonstrations rigoureuses — au prix de certaines longueurs.

Appliquée aux ensembles de nombres, la notion de fonction récursive fournit deux concepts importants : ceux d'ensemble récursif et d'ensemble récursivement énumérable.

Dans la suite du présent paragraphe, il sera sous-entendu que les *« ensembles » envisagés sont des parties de N.*

Définition 22 — On appelle *fonction caractéristique* de l'ensemble E, la fonction φE telle que

$$\begin{array}{lll} \varphi E x = 0 & \text{si} & x \in E \\ \varphi E x = 1 & \text{si} & x \in E \end{array}$$

Définition 23 — Un *ensemble* est dit *récursif* si sa fonction caractéristique est récursive.

Définition 24 — Un ensemble E est dit *récursivement énumérable* si c'est l'ensemble des valeurs prises par quelque fonction récursive.

Intuitivement — et en utilisant la Thèse de Church — dire qu'un ensemble est récursivement énumérable, cela signifie qu'on peut en trouver tous les éléments successivement par un procédé mécanique (il suffit de calculer successivement $f0$, $f1$, etc ... f étant une fonction récursive) ; dire qu'un ensemble est récursif, c'est dire qu'il existe un procédé mécanique pour décider si un nombre quelconque donné à l'avance fait partie de l'ensemble ou non (il suffit de calculer la valeur de la fonction caractéristique).

Voici les principaux théorèmes classiques de la théorie des fonctions récursives. On pourra en trouver les démonstrations — ainsi que celles des résultats cités au § 4 — dans les ouvrages mentionnés au § 9 ci-dessous.

Théorème 1 — Tout ensemble récursif est récursivement énumérable.

1. Pour les arguments en faveur de la Thèse de Church, voir Kleene 1952, ch. XII. Pour un avis opposé voir Kalmár 1957 — personnellement je ne peux pas considérer les arguments de Kalmár comme convaincants.

Théorème 2 — Il existe des ensemble récursivement énumérables qui ne sont pas récursifs (Théorème fondamental).

Théorème 3 — Toute fonction semi-récursive est récursive (trivial).

Théorème 4 — Si une fonction semi-récursive est une application de N (entier) dans N, c'est une fonction récursive.

Théorème 5 — L'ensemble des nombres sur lesquels une fonction semi-récursive est définie est un ensemble récursivement énumérable.

Théorème 6 — Si une fonction semi-récursive est définie par un ensemble récursif, elle peut être prolongée en une fonction récursive.

Théorème 7 — Il existe des fonctions semi-récursives qui ne peuvent pas être prolongées en des fonctions récursives.

Théorème 8 — L'ensemble des valeurs prises par une fonction semi-récursive est un ensemble récursivement énumérable.

Théorème 9 — L'ensemble des valeurs prises par une fonction récursive toujours croissante est un ensemble récursif.

Théorème 10 — Si un ensemble, E, récursivement énumérable, a pour complémentaire un ensemble récursivement énumérable, alors E est un ensemble récursif.

Théorème 11 — Tout ensemble fini est récursif. Le complémentaire d'un ensemble récursif est récursif. La réunion et l'intersection de deux ensembles récursifs sont récursifs.

On définit aussi des fonctions récursives et semi-récursives de plusieurs variables (voir par exemple Kleene 1952 ch. XI et XII). Au lieu de les définir directement, on peut se servir de *fonctions d'énumération des couples*. Une fonction d'énumération des couples est une application bi-univoque de $N \times N$ sur N. On vérifiera facilement que la fonction j_2 définie par

$$j_2 < x, y > = \frac{1}{2}\left((x + y)^2 + 3x + y\right)$$

est une telle fonction. A l'aide de cette fonction, on peut définir des fonctions d'énumérations des triplets, etc ... :

$$j_3 <x_1, x_2, x_3> = j_2 <x_1, j_2 <x_2, x_3>>$$
$$j_n <x_1, x_2, x_3, x_4> = j_2 <x_1. j_3 <x_2, x_3, x_4>>$$
etc ...

On peut maintenant poser :

Définition 25 — Une fonction f de deux variables est récursive s'il existe une fonction récursive d'une variable g, telle que

$$f(x,y) = g(j_2 <x,y>) \quad \text{pour tout } x, y \in \mathrm{N}$$

et de même pour 3, 4, ... n variables — analogue pour les fonctions semi-récursives.

Cette définition est équivalente à la définition directe de Kleene 1952 ch. XI. Elle parait dépendre du choix particulier de la fonction d'énumération des couples, mais en fait on aurait pu utiliser n'importe quelle fonction d'énumération « effectivement calculable » à la place de j_2 : on aurait obtenu une définition équivalente à la définition 22.

On peut maintenant définir les relations récursives (binaires, ... n-aires), en identifiant la relation à l'ensemble des couples (ou de triplets, etc ...) pour lesquels la relation est satisfaite.

Définition 26 — Une relation binaire R, définie sur N $\times$ N est dite récursive si sa fonction caractéristique est récursive — et de même pour les relations ternaires etc ...

Comme les fonctions à n arguments peuvent être identifiées à des relations $(n + 1)$-aires particulières, il faut s'assurer que les déf. 17 et 23 s'accordent avec cette identification. En fait, on a bien :

Théorème 12 — Toute relation récursive qui est fonctionnelle est une fonction récursive.

4. Systèmes logistiques et fonctions récursives.

Les notions de fonction récursive, ensemble récursif, etc ... définies seulement pour les membres naturels, s'étendent aux mots des systèmes logistiques, au moyen des *numérotations*.

Il est facile de numéroter tous les *mots* en **E** si **E** est un alphabet fini : il suffit de ranger les mots par longueur croissante, et (pour les mots de même longueur) dans l'ordre lexicographique, celui-ci étant défini par *l'ordre alphabétique* des symboles qui n'est autre que l'ordre de leurs numéros suivant la numérotation **N**. Exemple : soit un alphabet de deux symboles :

$$\mathbf{E} = \{a,b\}$$

avec $\quad \mathbf{N}a = 1 \quad$ et $\quad \mathbf{N}b = 2$

On aura alors, en appelant $\gamma(x)$ le numéro d'un mot [1] :

$$\left|\begin{array}{rl}
\gamma\ (\phi) & = 0 \\
\gamma < a > & = 1 \\
\gamma < b > & = 2 \\
\gamma <a,a> & = 2 \\
\gamma <a,b> & = 4 \\
\gamma <a,b> & = 5 \\
\gamma <b,b> & = 6 \\
\gamma <a,a,a> & = 7 \\
\text{etc} \ldots &
\end{array}\right.$$

On réalise ainsi une application bi-univoque de $\hat{\breve{\mathbf{E}}}$ sur N. Si l'alphabet était dénombrable, il serait encore possible de réaliser une telle numérotation, en employant des méthodes légèrement plus compliquées (parallèles à la manière classique de numéroter l'ensemble de toutes les suites finies d'entiers).

Plus généralement :

Définition 27 — On appellera *numérotation récursive* des mots d'un système logistique, toute fonction γ telle que :

1° γ est une application bi-univoque de $\hat{\breve{\mathbf{E}}}$ sur un ensemble récursif de nombres naturels.

2° Soit XY le mot obtenu en concaténant [2] les mots X et Y, l'application

$$\gamma(\mathrm{X}), \gamma(\mathrm{Y}) \to \gamma(\mathrm{XY})$$

est une fonction semi-récursive de deux variables.

3° L'application $\mathbf{N}s \to \gamma <s>$ (où s est un élément quelconque de $\mathbf{E}$) est une fonction semi-récursive qui peut être prolongée en une fonction récursive. Cette application pourra être quelconque si $\mathbf{E}$ est fini.

1. Dans ce cas on a $\gamma < a > = \mathbf{N}a$. Mais cela n'est pas le cas général. De toute façon, il ne faut pas confondre le symbole a avec le mot $< a >$ composé du seul symbole a — Remarquer qu'un mot tel que $< a, b, a >$ est généralement écrit *aba* par abus de langage.

2. Intuitivement l'opération de *concaténation* consiste à former un mot en écrivant joignant deux mots bout-à-bout ; on peut en donner une définition mathématique précise.

On appelle souvent les numérotations récursives des *numérotations de Gödel.*

Étant donné une numérotation particulière γ, on définira des ensembles récursifs de mots, etc, ... en posant :

Définition 28 — Un *ensemble de mots*, est dit *récursif par rapport à la numérotation* γ, si l'ensemble des numéros de ces mots est un ensemble récursif. On définira de même les ensembles de mots récursivement énumérables, les fonctions de mots récursives et semi-récursives, les relations récursives entre mots (par rapport à la numérotation γ).

La Déf. 28 est relative à une numérotation particulière. Mais cette relativité n'est qu'apparente, car on a

Théorème 13 — Si un ensemble de mots est récursif par rapport à une numérotation récursive particulière, il est récursif par rapport à toute numérotation récursive. Et la même propriété d'invariance s'applique aux autres notions définies par Déf. 28

On peut donc parler sans ambiguité d'un *ensemble récursif de mots* d'un système logistique, ou d'un ensemble récursivement énumérable de mots etc...

Définition 29 — Un système logistique

$$\mathbf{S} = <\mathbf{E}, \mathbf{F}, \mathbf{A}, \mathbf{R}_1, \ldots \mathbf{N}>$$

est dit *récursif* si les ensembles **F** et **A** et toutes les relations $\mathbf{R}_1$, etc... sont récursives.

On obtient alors les théorèmes fondamentaux :

Théorème 14 — Dans un système logistique récursif, l'ensemble des thèses est récursivement énumérable.

Théorème 15 — Il existe des systèmes logistiques récursifs dont l'ensemble des thèses n'est pas récursif.

Définition 30 — Un système logistique récursif dont l'ensemble des thèses est récursif est dit *décidable.*

La notion de système logistique récursif fournit la traduction abstraite des conditions (e) et (g) du § 2.

En utilisant la thèse de Church on peut dire que :

— dans un système logistique récursif, il existe un procédé mécanique permettant d'obtenir toutes les thèses les unes après les autres ;

— dans un système logistique décidable, il existe un procédé mécanique permettant de décider si une formule donnée est ou non une thèse.

5. Les machines de Turing.

Cette notion, inventée simultanément par Turing et par Post en 1936, est une tentative pour définir une notion abstraite adéquate à la notion intuitive de « calcul effectif » [1].

Je renvoie à mon article Porte 1956 pour la description des machines de Turing, considérées d'un point de vue intuitif. Je rappellerai seulement la notion tout a fait abstraite suivante [2] : on pourrait considérer les machines de Turing comme des cas particuliers des systèmes logistiques, si on avait généralisé les systèmes logistiques en considérant, à la place de $\check{\mathbf{E}}$, l'ensemble $\mathbf{E}^*$ des « mots doublement infinis » [3]. Alors :

Définition 31 — Une machine de Turing est une structure

$$M = <\mathbf{E},\ \mathbf{F},\ \mathbf{A}, \mathbf{R}_1, \dots \mathbf{N}>$$

où :

1° **E** est un ensemble fini comprenant les symboles suivants (dans l'ordre des numéros) :

1. Voir Post 1936, Turing 1936, Turing 1937. — Ces « machines » ont été définies sur le papier à une époque où il n'existait pas de machines à calculer universelles. Si on compare cette notion abstraite à ce qui existe réellement aujourd'hui, on peut dire qu'une machine de Turing est réalisée par une machine à calculer universelle *munie d'un programme*. De plus, dans la machine de Turing la mémoire devra être considérée comme potentiellement infinie (non bornée) ; d'autre part, les opérations élémentaires seraient beaucoup plus simples que celles qui sont réalisables par les machines réelles.

2. Pour des raisons de commodité, j'ai renversé les notations que j'avais utilisées dans Porte 1956 § 2 ; les symboles q_i (que j'appelle « explorateurs » et non « état de la machine » — d'un point de vue intuitif, ce seraient plutôt des « ordres » que des « états ») sont écrits à gauche du symbole exploré et non plus à droite. La définition des « machines simples » (Porte 1956, § 3) est modifiée ici de façon analogue.

3. Plus précisément, $\mathbf{E}^*$ est l'ensemble des familles (voir Bourbaki 1939, § 2, n° 14) $(x_i)_{i \in \mathbf{Z}}$ à éléments dans **E** — où Z est l'ensemble des entiers rationnels (positifs ou négatifs).

$s_1, \dots s_t$ — dits « symboles primitifs »
σ — dit « blank »[1], écrit aussi « s_0 » dans ce qui suit
$q_0, q_1, \dots q_r$ — dits « explorateurs » (q_0 est dit « l'explorateur final », q_1 est dit « l'explorateur initial »).

2° **A** est vide.

3° **F** (ensemble des formules) est l'ensemble des éléments de $\mathbf{E}^*$ qui satisfait aux conditions que, dans une formule :

— il n'y a qu'un nombre fini de symboles différents de σ ;
— il y a un explorateur et un seul.

4° Les règles, $\mathbf{R}_1$, ..., sont en nombre fini, et toutes les règles sont à un antécédent.

5° Chaque règle a l'une des trois formes suivantes : dans toute formule

— remplacer le segment $q_i s_k$ par $q_j s_l$
— » » $q_i s_k$ par $s_l q_j$
— » » $s_m q_i s_k$ par $q_j s_m s_l$

(où $i, j = 1, 2, \dots$ ou r et $k, l, m = 0, 1, \dots$ ou t — on peut avoir $i = j$ de même que $k = l$).

6° Pour i et k donnés, il y a au plus une seule règle s'appliquant aux formules qui contiennent le segment s_m, $q_i s_k$.

7° Aucune règle ne s'applique aux formules qui contiennent q_0.

Les conditions, 4°, 5°, 6°, 7° permettent de décrire toutes les règles d'une machine de Turing par une table à double entrée portant en colonnes les explorateurs, en lignes les *symboles explorés* (ceux qui suivent l'explorateur dans la formule), dans les cases le résultat de l'application de l'explorateur q_i au symbole s_k (certaines cases peuvent être vides — toutes les cases de la ligne q_0 sont vides) — voir Porte 1956, § 4.

Définition 32 — Une machine de Turing sera dite *semi-simple* si elle n'a qu'un seul symbole primitif, s_1, — qui sera alors écrit | et appelé « stroke »[1] [2].

1. « Blank », mot anglais signifiant à peu près « vide » ; « stroke », mot anglais désignant en général une barre verticale ou inclinée.
2. Les machines semi-simples sont simples au sens de Porte 1956, § 3, sous réserve de renversement des notations indiqué ci-dessus, note 2 de la p. 131.

Définition 33 — On appellera *segment actif* d'une formule plus petit segment de cette formule qui contient à la fois le symbole exploré et tous les symboles différents de σ [1].

Définition 34 — Dans une machine semi-simple, un segment de formule formé de $n + 1$ strokes ($n \geqslant 0$) est appelé le *numéral* correspondant au nombre entier n ; il sera représenté par « $\bar{n}$ ».

Définition 35 — La machine semi-simple **M** sera dite calculer la fonction f (de une variable) si les deux conditions suivantes sont vérifiées [1]

1° $$[q_1\bar{n}] \quad \Vdash \quad [q_0\bar{f}_n]$$

— où « [X] » représente l'une quelconque des formules dont le segment actif est X.

2° Si f n'est pas définie pour la valeur n de la variable, et si

$$[q_1\bar{n}] \quad \Vdash \quad Z$$

alors, la formule Z ne contient pas q_0.

Théorème 16 — Pour toute fonction semi-récursive de une variable, il existe au moins une machine semi-simple qui la calcule (voir Porte 1956, § 8).

Ce théorème sera largement utilisé dans ce qui suit.

6. Les systemes de Post.

Ce sont des systèmes logistiques particuliers :

Définition 36 — Un système de Post est un système logistique satisfaisant aux conditions suivantes :

1° L'alphabet est fini.

2° Tout mot est une formule.

3° Le nombre des axiomes est fini (il peut être nul).

4° Il y a un nombre fini de règles.

5° Toute règle est de la forme [2]

---

1. Définitions légèrement différentes de celles de Porte 1956.

2. Une règle est à proprement parler une relation, mais il est souvent commode de faire l'abus de langage qui consiste à la représenter en écrivant successivement la forme générale de l'antécédent (ou des antécédents séparés par des virgules) le

$$\mathbf{R}: \quad A_{10}X_{11}A_{11}X_{12}\ldots X_{1n_1}A_{1n_1},\ A_{20}X_{21}\ldots X_{2n_2}A_{2n_2},\ldots$$
$$\ldots A_{k0}X_{k1}\ldots X_{knk}A_{knk} \quad \Vdash \quad B_0Y_1B_1\ldots Y_kB_k$$

— où les A_{ij} et les B_i sont des mots fixes et les X_{ij}, X_j sont des mots variables quelconques.

On peut considérer les systèmes de Post comme une tentative de définir les « systèmes logistiques effectifs » indépendamment de la théorie des fonctions récursives. En effet, les conditions 1° à 5° de la Déf. 31 sont, intuitivement, au moins aussi restrictives que les conditions (d) à (g) des § 2 ci-dessus. Intuitivement, il apparait probable qu'elles ne soient pas *pas trop* restrictives, l'idée générale étant que, dans un raisonnement réel où rien n'est sous-entendu, on décide si une règle est ou non applicable à une certaine formule en examinant le contexte dans lequel cette formule est écrite : il semble donc probable qu'il suffise de « faire passer » ce contexte dans le système logistique pour avoir un système de Post. De façon plus précise, il apparait intuitivement probable que tout système logistique « effectif » ayant un alphabet fini possède un élargissement conservatif qui est un système de Post. En fait cela résultera du théorème 22 ci-dessous à condition d'utiliser la Thèse de Church. On peut aussi se fonder sur les remarques intuitives qui viennent d'être exposée pour considérer le théorème 22 comme une justification supplémentaire de la Thèse de Church.

On aura ci-dessous à considérer des classes spéciales de systèmes de Post.

Définition 37 — Un système de Post sera dit obéir à la *condition* K_1 si, dans chacune des règles, les mots Y_1 ... Y_p (utilisés pour décrire la forme du conséquent) sont certains des mots X_{01} ... X_{kn_k} [1].

Il est dit obéir à la condition K_2 si, dans chacune des règles l'un au moins des mots B_0, ... B_n n'est pas vide [2].

---

signe de déductibilité et la forme générale du conséquent, le nom de la règle étant écrit en marge. Par exemple, une écriture telle que

$$\mathbf{R}: \quad X \mid Y \Vdash Y\sigma\sigma$$

signifierait : $\mathbf{R}(U, V)$ si et seulement si il existe des mots X et Y tels que

$$U = X \mid Y \text{ et } V = Y\sigma\sigma$$

Cette convention sera adoptée dans la suite du travail présent.

1. Ainsi la condition K_1 admet des règles telles que la règle R de la note précédente mais exclut des règles telles que

$$\mathbf{R}': \quad Y\sigma\sigma \Vdash X \mid Y$$

2. Ainsi la condition K_2 admet des règles telles que $\mathbf{R}$ — note [1] — ou $\mathbf{R}'$ — note [2] — mais exclut des règles telles que

$$\mathbf{R}'': \quad X \mid Y \Vdash XY$$

Définition 38 — Un système de Post sera dit *unilatéral* s'il est simple [1] et si toutes les règles sont de la forme

R : $AX \Vdash XB$ ([2])

Il sera dit *bilatéral* s'il est simple et si toutes les règles sont de la forme

R : $XAY \Vdash XBY$ ([3])

Il sera dit *de troisième espèce* [4] s'il est simple et si toutes les règles sont de la forme

R : $AX \Vdash BXC$

Il est clair que tous les systèmes unilatéraux sont aussi de troisième espèce.

Théorème 17 — Tout système de Post est récursif (voir Kleene 1952, ch. XII)

Maintenant, il existe une connexion étroite entre les machines de Turing et les systèmes de Post bilatéraux. En effet, dans une machine de Turing, chaque règle ordonne de remplacer un certain segment d'une formule par un autre segment. La seule différence avec les systèmes bilatéraux est que, dans les machines de Turing, les formules sont infinies. Comme le segment actif d'une telle formule est fini (voir Déf. 31 et 33) on pourrait faire correspondre à chaque formule, X, d'une machine une formule, X^1, d'un système logistique, constituée par le segment actif de X ; mais, pour transcrire les règles, il faudra tenir compte du fait que, dans X, le segment actif est précédé et suivi d'autant de blanks qu'on en a besoin. On est ainsi conduit à considérer un système logistique bilatéral dans lequel les segments actifs des formules de la machine ont été traduits en des formules « scellées » par des symboles supplémentaires situés au début et à la fin. C'est la solution adoptée dans Kleene 1952, § 71. La solution ci-dessus est un peu différente de celle de Kleene, mais plus adaptée aux usages qui vont en être faits.

Soit une machine **M**, d'alphabet

$$\mathbf{E} = \{s_1, \ldots s_t, \sigma, q_0, q_1 \cdots q_r\}$$

1. C'est-à-dire s'il a au plus un axiome ; voir Déf. 11.
2. Post 1943 et Rosenbloom 1950 appellent « canoniques » les systèmes de Post satisfaisant aux conditions K_1 et K_2, « normaux » les systèmes de Post unilatéraux satisfaisant à K_2 (tout système unilatéral satisfait évidemment à K_1).
3. Les systèmes bilatéraux sont appelés « semi-Thue systems » dans Kleene. 1952, § 71.
4. A défaut d'un terme meilleur ...

On lui fera correspondre un système logistique $\mathbf{M}_p$ d'alphabet

$$\mathbf{E}_p = \mathbf{E} \cup \{\alpha, \omega\}$$

où α et ω sont deux nouveaux symboles dits respectivement *sceau initial* et *sceau final.*

A chaque règle de **M**, telle que, par exemple [1]

$$\begin{array}{c|c} & s_1 \\ \hline q_1 & \sigma \leftarrow q_2 \end{array}$$

on fera correspondre, dans M_n, $(t + 2)^2$ règles bilatérales de la forme

$$\mathbf{R}_k : \mathrm{YA}_k\mathrm{Y} \mathrel{|\!\!\!-\!\!\!|} \mathrm{XB}_k\mathrm{Y}$$

où les A_k sont tous les mots des formes

$$\left\{\begin{array}{l} s_i\, q_1\, s_1\, s_j \\ \alpha\, q_1\, s_1\, s_j \\ s_i\, q_1\, s_1\, \omega \\ \alpha\, q_1\, s_1\, \omega \end{array}\right. \qquad (i, j = 0, \ldots t\,;\ \text{avec } s_0 = \sigma)$$

et où les B_k traduisent l'opération de la machine, en respectant les principes suivant lesquels, dans B_k, on a

1° toujours au moins un s_i ($i = 0, 1, \ldots t$) après l'explorateur (c'est le symbole exploré) ;

2° jamais de σ immédiatement à droite de α ;

3° jamais de σ immédiatement à gauche de ω et après le symbole exploré.

On aura ainsi, par exemple, des règles telles que :

$$\begin{array}{lll} \mathbf{R}_1 : & \mathrm{X}\, s_2 q_1\, s_1\, s_3\, \mathrm{Y} \mathrel{|\!\!\!-\!\!\!|} & \mathrm{X}\, q_2\, s_2\, \sigma\, s_3\, \mathrm{Y} \\ \mathbf{R}_2 : & \mathrm{X}\, \sigma\, q_1\, s_1\, \sigma\, \mathrm{Y} \mathrel{|\!\!\!-\!\!\!|} & \mathrm{X}\, q_2\, \sigma\, \sigma\, \sigma\, \mathrm{Y} \\ \mathbf{R}_3 : & \mathrm{X}\, \alpha\, q_1\, s_1\, s_3\, \mathrm{Y} \mathrel{|\!\!\!-\!\!\!|} & \mathrm{X}\, \alpha\, q_2\, \sigma\, \sigma\, s_3\, \mathrm{Y} \end{array}$$

(introduction d'un blank à gauche) et :

$$\mathbf{R}_4 : \mathrm{X}\, s_2\, q_1\, s_1\, \omega\, \mathrm{Y} \mathrel{|\!\!\!-\!\!\!|} \mathrm{X}\, q_2\, s_2\, \omega\, \mathrm{Y}$$

(suppression d'un blank à droite).

1. Voir Porte 1956, § 4, pour cette notation.

La définition précise de $\mathbf{M}_p$ n'est pas difficile à énoncer en énumérant tous les cas possibles ; mais elle est longue. Je poserai simplement :

Définition 39 — On appelle *système de Post bilatéral associé à la machine de Turing* **M**, le système $\mathbf{M}_p$ défini de la façon esquissée ci-dessus.

On a alors, par construction :

Théorème 18 — Quelle que soit **M**, M_p est un système o-simple déterministe par rapport aux mots de la forme $\alpha X \omega$ où X est le segment actif d'une formule de **M**.

Théorème 19 — On a dans **M**

$$[X] \;\Vdash\; [Y]$$

[X] et [Y] désignant des formules de segments actifs X et Y, si et seulement si

$$\alpha X \omega \;\Vdash\; \alpha Y \omega \qquad \text{dans } \mathbf{M}_p$$

Les systèmes $\mathbf{M}_p$ rendent donc exactement les mêmes services que les machines **M**. On peut les composer de la même façon (voir Porte 1956, § 5), etc ... On a en particulier, d'après le théorème 16 :

Théorème 20 — Pour toute fonction semi-récursive f, il existe un système de Post bilatéral sans axiome d'alphabet

$$\mathbf{E} = \{ |, \sigma, q_0, q_1, \dots q_r, \alpha, \omega \}$$

tel que

1° :
$$\alpha\, q_1\, \bar{n}\, \omega \;\Vdash\; \alpha\, q_0\, \overline{fn}\, \omega$$

pour tout nombre naturel n tel que fn existe,

2° : si n est un nombre naturel tel que fn n'existe pas et si

$$\alpha\, q_1\, \bar{n}\, \omega \;\Vdash\; X$$

alors X ne contient pas q_0 — (Il suffit de considérer le système $\mathbf{M}_p$ associé à la machine **M** dont l'existence est affirmée par le théorème 16).

Théorème 21 — Tout système logistique récursif **U**. dont l'alphabet ne contient qu'un seul symbole et où ϕ n'est pas une thèse possède un élargissement conservatif qui est un système de Post bilatéal.

Soit $|$ le symbole unique de **U** ; les thèses de **U** sont des numérales qui constituent un ensemble récursivement énumérable, par rapport à toute numérotation récursive. Mais la numérotation définie par

$$\gamma(\bar{n}) = n$$

est récursive. Il suffit de trouver un système bilatéral tel que ses thèses constituent l'ensemble :

$$\mathbf{T} = \{ \overline{f0}, \overline{f1} \quad , \dots \overline{fn}, \quad \dots \}$$

où f est une certaine fonction récursive.

Soit **M** une machine de Turing qui calcule f ; soit $\mathbf{M}_p$ le système bilatéral associé, d'alphabet

$$\mathbf{E}_p = \{ \,|, \sigma, q_0, q_1, \dots q_r, \alpha, \omega \}$$

Une solution est fournie par le système $\mathbf{M}'_p$ défini de la façon suivante :

— alphabet :

$$\mathbf{E}'_p = \{ \,|, \sigma, q_0, q_1, \dots q_r, \alpha, \omega \} = \mathbf{E}_p$$

— axiome

$$\mathbf{A} : \vdash \alpha\, q_1 \,|\, \omega$$

— règles : celles de $\mathbf{M}_p$, plus

$$\begin{array}{lllll} \mathbf{R}_1 : & X \alpha q_1 Y & \Vdash & X \alpha q_1 | Y \\ \mathbf{R}_2 : & X \alpha q_0 | Y & \Vdash & X | \alpha q_0 Y \\ \mathbf{R}_3 : & X \alpha q_0 \omega Y & \Vdash & XY \end{array}$$

En effet, par **A** et $\mathbf{R}_1$, on obtient

$$\vdash \alpha\, q_1\, \bar{n}\, \omega \tag{1}$$

pour tout numéral $\bar{n}$. Ensuite, par les règles de $\mathbf{M}_p$, on obtient

$$\vdash \alpha\, q_0\, \overline{fn}\omega \tag{2}$$

Puis, par $\mathbf{R}_2$:

$$\vdash \overline{fn}\, \alpha\, q_0\, \omega \tag{3}$$

et, par $\mathbf{R}_3$:

$$\vdash \overline{fn} \tag{4}$$

Inversement, on ne peut avoir $\vdash \overline{m}$ que par $\mathbf{R}_3$, donc, à partir de $\vdash \overline{m}\ \alpha\ q_0\ \omega$, ce qui ne peut provenir de $\vdash \alpha\ q_0\ \overline{m}\ \omega$ par $\mathbf{R}_2$ et, d'après le Théorème 20, on ne peut avoir $\alpha\ q_0\ \overline{m}\ \omega$ comme thèse que s'il existe n tel que $fn = m$.

Théorème 22 — Tout système logistique récursif, **S**, à alphabet fini possède un élargissement conservatif qui est un système de Post bilatéral.

La démonstration est élémentaire, mais longue, et je n'en donnerai que le principe. Considérons une certaine numérotation récursive des formules de **S**, soit γ. Les numéros des thèses constituent un ensemble récursivement énumérable $\{ f0, f1, \ldots fn, \ldots \}$ (théorème 14). Or on peut construire une machine de Turing, et par suite un système bilatéral associé, $\mathbf{S}_1$, tel que

$$\alpha\ q_1\ \overline{\gamma\ (\mathrm{X})}\ \omega \ \Vdash_{\mathbf{S}_1}\ \alpha\ q_0\ \mathrm{X}\ \omega$$

pour toute formule X de **S** (même si $\mathbf{X} = \phi$). Soir $\mathbf{S}_2$ le système bilatéral associé à la machine de Turing qui calcule f. Soit $\mathbf{S}_3$ le système obtenu en composant $\mathbf{S}_2$ et $\mathbf{S}_1$. On a dans $\mathbf{S}_3$

$\alpha q_1 \bar{n} \omega \Vdash_{\mathbf{S}_3} \alpha q_0 \mathrm{X} \omega$ (où q_1 est l'explorateur initial de $\mathbf{S}_2$ et donc de $\mathbf{S}_3$)

si et seulement si $\gamma(\mathrm{X}) = fn$. Construisons $\mathbf{S}_4$ en ajoutant à $\mathbf{S}_3$ l'axiome **A** et les règles $\mathbf{R}_1$ et $\mathbf{R}_3$ du théorème 21 (avec q'_1 à la place de q_1), ainsi que les règles, analogues à $\mathbf{R}_2$:

$$\mathbf{R}_{2\cdot 1} \quad \mathrm{X}\ \alpha\ q_0\ s_1\ \mathrm{Y} \quad \Vdash \quad \mathrm{X}\ s_1\ \alpha\ q_0\ \mathrm{Y}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\mathbf{R}_{2\cdot k} \quad \mathrm{X}\ \alpha\ q_0\ s_k\ \mathrm{Y} \quad \Vdash \quad \mathrm{X}\ s_k\ \alpha\ q_0\ \mathrm{Y}$$

Le système $\mathbf{S}_4$ donne la solution du problème.

Le théorème 23 indique que les systèmes de Post possèdent bien le caractère d'« universalité » qui avait été annoncé au début du § présent. On peut ajouter que la construction du système $\mathbf{S}_4$ est « effective », en ce sens que la démonstration fournit un moyen mécanique d'écrire les règles de $\mathbf{S}_4$ connaissant la définition du système **S**. Cette « effectivité » est d'ailleurs assez théorique, car la méthode conduira facilement à un système possèdant des millions de symboles et des millions de règles, même à partir d'un système **S** très simple (un système pour le calcul des propositions par exemple). En pratique, et si on veut seulement trouver un système de Post (non bilatéral) qui soit un élargissement conservatif de **S**, on arrive en général facilement à des systèmes assez simples en procèdant par tâtonnement. On trouvera un grand nombre de tels élargissements dans Rosenbloom 1950.

En combinant les théorèmes 17 et 23 on obtient :

Théorème 24 — Tout système de Post possède un élargissement conservatif bilatéral.

Le théorème suivant est une généralisation du théorème fondamental de Post 1943.

Théorème 25 — Tout système de Post possède un élargissement conservatif unilatéral.

D'après le théorème 24, il suffit de construire un élargissement conservatif unilatéral de chaque système de Post bilatéral. Soit **S** un système bilatéral d'alphabet

$$\mathbf{E} = \{a_1, \dots a_k\}$$

et de règles

$$\left|\begin{array}{ll} \mathbf{R}_1 : & X\,A_1\,Y \;\;\vdash\!\vdash\;\; X\,B_1\,Y \\ \dots\dots\dots\dots & \\ \mathbf{R}_l : & X\,A_l\,Y \;\;\vdash\!\vdash\;\; X\,B_l\,Y \end{array}\right.$$

(où les A_i et les B_i sont certains mots fixes).

Le système suivant, **S′**, donnera la solution

— *alphabet* :

$$\mathbf{E}| = \{ a_1, \dots a_k, \overline{a_1}, \dots \overline{a_k} \}$$

Et convenons que si Z est un mot de **S**, Z sera le mot de **S′** obtenu à partir de Z en remplaçant a_1 par $\bar{a}_1$,a_2 par $\bar{a}_2$, ... a_k par $\bar{a}_k$.

— *axiome* : celui de **S**.

— *règles* :

$$\left|\begin{array}{l} \mathbf{R}'_{a_1} : a_1\,X \;\;\vdash\!\vdash'\;\; X\bar{a}_1 \\ \dots\dots\dots\dots\dots\dots \\ \mathbf{R}'_{a_k} : a_k\,X \;\;\vdash\!\vdash'\;\; X\bar{a}_k \\ \mathbf{R}'_{\bar{a}_1} : \bar{a}_1\,X \;\;\vdash\!\vdash'\;\; X\,a_1 \\ \dots\dots\dots\dots\dots\dots \\ \mathbf{R}'_{\bar{a}_k} : \bar{a}_k X \;\;\vdash\!\vdash'\;\; X\,a_k \end{array}\right\}\text{« règles de transposition »}$$

$$\left|\begin{array}{l} \mathbf{R}'_1 : A_1 X \quad \Vdash' \quad X\overline{B_1} \\ \dots\dots\dots\dots\dots\dots \\ \mathbf{R}'_l : A_l X \quad \Vdash' \quad X\overline{B_l} \end{array}\right| \text{« règles de calcul »}$$

En effet on a alors dans $\mathbf{S}'$:

$$\begin{array}{llll} X A_i Y & \Vdash' \; A_i Y \overline{X} & \text{par} & \mathbf{R}'_{a_1}, \dots \mathbf{R}'_{\bar{a}_k} \\ A_i Y \overline{X} & \Vdash' \; Y \bar{\overline{X}} \overline{B_i} & \text{par} & \mathbf{R}'_i \\ Y \bar{\overline{X}} \bar{\overline{B}}_i & \Vdash' \; \bar{\overline{X}} \bar{\overline{B_i}} \overline{Y} & \text{par} & \mathbf{R}'_{a_1}, \dots \mathbf{R}'_{a_k} \\ \bar{\overline{X}} \bar{\overline{B}}_i \overline{Y} & \Vdash' \; X B_i Y & \text{par} & \mathbf{R}'_{\bar{a}_1}, \dots \mathbf{R}'_{a_k} \end{array}$$

donc

$$X A_i Y \quad \Vdash' \quad X B_i Y$$

Et il est immédiat que les règles de $\mathbf{S}'$ ne permettent d'obtenir en conclusion un mot de $\mathbf{S}$ qui n'ait pas été pris comme prémisse que si l'une des règles $\mathbf{R}'_i$ a été utilisée. Par suite, on voit facilement qu'on n'a

$$X \quad \Vdash' \quad Y$$

(X et Y étant des mots de S) que si l'on a aussi

$$X \quad \Vdash \quad Y$$

Le théorème 25 avait été démontré dans Post 1943, par une méthode différente, mais avec la restriction que le système de Post en question obéisse aux conditions K_1 et K_2 (déf. 37). Il est facile d'étendre la méthode de Post au cas où K_1 n'est pas satisfaite, mais c'est plus difficile dans le cas où K_2 n'est pas satisfaite, et où, par suite, le mot vide risque d'être une thèse.

7. Les algorithmes de Markov.

On peut en donner la définition générale suivante :

Définition 40 — On appellera *algorithme de Markov* un système dérivationnel (voir Déf. 10).

$$\mathbf{S} = \langle \mathbf{E}, \mathbf{A}, \mathbf{R}_1, \dots \mathbf{R}_k, \mathbf{N}, \mathbf{D} \rangle$$

où :

1° Le système logistique

$$S^0 = < \mathbf{E}, \mathbf{A}_1\ \mathbf{R}_1 \ldots \mathbf{R}_k, \mathbf{N} >$$

est un système de Post o-simple.

2° Lès dérivations sont définies de la façon suivante :

a) Un choix préalable général est fait entre les diverses applications possibles de la même règle à la même formule.

b) Certaines des règles sont distinguées et dites « conclusives ».

c) Les dérivations sont alors construites par application des règles dans un certain ordre : on essaie d'appliquer $\mathbf{R}_1$, si c'est impossible on passe à $\mathbf{R}_2$, etc ... Aussitôt qu'une règle s'applique on repart de la formule résultant de son application et on recommence au début : essayer d'appliquer $\mathbf{R}_1$ etc...

d) Quand une règle conclusive a été appliquée, la dérivation est terminée (Une dérivation peut aussi se terminer quand aucune règle n'est applicable).

On vérifie que la Déf. 40 définit bien des algorithmes au sens de la Déf. 14 [1].

Définition 41 — Un algorithme de Markov est dit *bilatéral* si le système de Post sous-jacent est bilatéral, le « choix général » dans l'application des règles (point 2° *a* de la Déf. 40) étant fait de la façon suivante :

Si une règle

$$\mathbf{R} : \mathrm{X\,A\,Y} \quad \mathrm{H} \Vdash \mathrm{X\,B\,Y}$$

peut s'appliquer de plusieurs façons à un mot

$$\mathrm{Z} = \mathrm{X\,A\,Y} = \mathrm{X'\,A\,Y'} \qquad (\mathrm{X} \neq \mathrm{X'}, \mathrm{Y} \neq \mathrm{Y'})$$

on choisit l'application dans laquelle X est la sous-formule la plus courte [2].

Définition 42 — Un algorithme de Markov est dit *unilatéral* si le système de Post sous-jacent est unilatéral. Dans ce cas il n'y a pas de choix à faire, chaque règle s'appliquant au plus d'une seule façon.

---

1. En fait, Markov et ses disciples n'ont considéré que les algorithmes bilarétaux (voir ci-dessus Déf. 41).

2. Autrement dit la règle **R** ordonne de remplacer par B la première (en partant de la gauche) occurence de A qui se trouve dans Z.

Définition 43 — Un algorithme de Markov est dit *de troisième espèce* si le système de Post sous-jacent est de troisième espèce. — Même remarque que dans la Déf. 42.

Définition 44 — Un algorithme S est dit calculer la fonction d'entiers f (application d'une partie de N dans N) s'il existe un symbole, $|$, dans l'alphabet de S tel que, les numérals étant définis comme dans la Déf. 34 et $\mathfrak{R}_{\mathrm{S}}$ comme dans la Déf. 17 :

1° $\mathfrak{R}_{\mathrm{S}}\ (\bar{n}) = \overline{fn}$ chaque fois que fn existe.

2° $\mathfrak{R}_{\mathrm{S}}\ (\bar{n})$ n'existe pas quand fn n'existe pas.

Théorème 26 — Pour toute fonction semi-récursive f, il existe un algorithme bilatéral qui calcule f (Detlovs 1953) [1].

Je donne simplement la solution sans démonstration. Partant du système $\mathbf{M}_p$ du théorème 19 (voir déf. 39), on obtient un algorithme de la façon suivante :

— *l'alphabet* est le même que celui de $\mathbf{M}_p$;

— les *règles* sont celles de $\mathbf{M}_p$ (dites ici « règles de calcul »), plus deux « règles d'introduction », $\mathbf{R}^i_1$ et $\mathbf{R}^i_2$, et deux « règles de fin de calcul », $\mathbf{R}^f_1$ et $\mathbf{R}^f_2$ — à savoir :

$$\left|\begin{array}{l} \mathbf{R}^i_1 : \mathrm{X}\,\omega\,|\,\mathrm{Y} \quad \vdash\!\vdash \quad \mathrm{X}\,|\,\omega\,\mathrm{Y} \\ \mathbf{R}^i_2 : \mathrm{X}\,|\,\mathrm{Y} \quad \vdash\!\vdash \quad \mathrm{X}\,\alpha\,q_1\,\omega\,|\,\mathrm{Y} \end{array}\right.$$

$$\left|\begin{array}{l} \mathbf{R}^f_1 : \mathrm{X}\,\alpha\,q_0\,|\,\mathrm{Y} \quad \vdash\!\vdash \quad \mathrm{X}\,|\,\alpha\,q_0\,\mathrm{Y} \\ \mathbf{R}^f_2 : \mathrm{X}\,\alpha\,q_0\,\omega\,\mathrm{Y} \quad \vdash\!\vdash \quad \mathrm{X}\,\mathrm{Y} \end{array}\right.$$

— *ordre des règles* : $\mathbf{R}^i_1$ précède toutes les autres règles, $\mathbf{R}^i_2$ suit toutes les autres règles (à part cela l'ordre des règles est indifférent) ;

— la règle $\mathbf{R}^f_2$ est *conclusive*.

On vérifiera que l'algorithme opère de la façon suivante :

partant de	$\bar{n}$		
on obtient	$\alpha\,q_1\,\bar{n}\,\omega$	par	$\mathbf{R}^i_1$, $\mathbf{R}^i_2$
d'où	$\alpha\,q_0\,\overline{fn}\,\omega$	par les règles de calcul	
d'où finalement	$\overline{fn}$	par	$\mathbf{R}^f_1$, $\mathbf{R}^f_2$

1. En fait, le théorème de Detlovs est un peu différent du Théorème 26 et plus

Théorème 27 — Pour toute fonction semi-récursive f, il existe un algorithme unilatéral qui calcule f.

La solution (sans démonstration) est la suivante.

On part du système unilatéral $\mathbf{M}_{pu}$, défini à partir de $\mathbf{M}_p$ (voir déf. 39) en employant la méthode indiquée dans la démonstration du théorème 25. Soit

$$E = \{|, \sigma, q_0, q_1, \ldots q_r, a, \omega, \dagger, \bar{\sigma}, \bar{q}_0, \bar{q}_1, \ldots \bar{q}_r\, \bar{a}, \bar{\omega}, \}$$

l'alphabet de ce système. On construit l'algorithme de la façon suivante :

— *Alphabet* : ajouter à **E** quatre nouveaux symboles, δ, ψ, φ et ν.»

— *Règles* : modifier les règles du système $\mathbf{M}_{pu}$ (« règles de calcul » et « règles de transposition ») en y remplaçant partout | par ν. Ajouter quatre « règles d'introduction », $\mathbf{R}^i_1$, $\mathbf{R}^i_2$, $\mathbf{R}^i_3$ et $\mathbf{R}^i_4$, et cinq « règles de fin de calcul », $\mathbf{R}^f_1$, $\mathbf{R}^f_2$, $\mathbf{R}^f_3$, $\mathbf{R}^f_4$ et $\mathbf{R}^f_5$ — à savoir :

$$\left\{\begin{array}{lrcl}
\mathbf{R}^i_1 : & |\delta X & \Vdash & X\dagger \\
\mathbf{R}^i_2 : & |\bar{\omega} a q_1 X & \Vdash & X|\delta
\end{array}\right.$$

$$\left|\begin{array}{lrcl}
\mathbf{R}^i_3 : & \bar{\omega} a q_1 |X & \Vdash & X\dagger\bar{\omega} a q_1 \\
\mathbf{R}^i_4 : & |X & \Vdash & X\bar{\omega} a q_1|
\end{array}\right.$$

$$\left\{\begin{array}{lrcl}
\mathbf{R}^f_1 : & \bar{\omega} a q_0 \nu X & \Vdash & X\psi\nu\omega\bar{a}\bar{q}_0 \\
\mathbf{R}^f_2 : & \psi\nu X & \Vdash & X\psi\nu \\
\mathbf{R}^f_3 : & \bar{\omega} a q_0 \psi X & \Vdash & X\psi\varphi \\
\mathbf{R}^f_4 : & \nu\psi X & \Vdash & X| \\
\mathbf{R}^f_5 : & \varphi X & \Vdash & X
\end{array}\right.$$

— *Ordre des règles* : $\mathbf{R}^i_4$ et toutes les règles de transposition suivent toutes les autres règles (à part cela, l'ordre des règles est indifférent).

— La règle $\mathbf{R}^f_5$ est *conclusive*.

Ces algorithmes sont plus compliqués que les algorithmes bilatéraux définis ci-dessus (théorème 26) ; mais ils sont plus proches des systèmes de Post correspondants puisqu'une règle unilatérale

---

général. D'une part il se rapporte à la représentation d'un nombre entier n par n strokes, au lieu de $n + 1$; d'autre part, il s'applique aussi aux fonctions récursives de plusieurs variables.

ne peut s'appliquer que d'une seule façon. Le théorème suivant établit des rapports encore plus étroits entre une classe de systèmes de Post et une classe d'algorithmes.

Théorème 28 — Pour toute fonction semi-récursive f, il existe un algorithme de 3° espèce qui calcule f et dans lequel l'ordre des règles est indifférent.

La solution (sans démonstration encore) est la suivante. Partons du système de Post $\mathbf{M}_p$ (voir théorème 19 et déf. 39) d'alphabet

$$\mathbf{E} = \{|, \sigma, q_0, q_1 \dots q_r, \alpha, \omega\}$$

et de règles $\mathbf{R}_k : \; X A_k Y \;\; \Vdash \;\; X B_k Y$

On obtient un algorithme de la façon suivante :

— *Alphabet* : Ajouter à **E** un nouveau symbole : ν

— *Règles* : On aura encore des règles « d'introduction » « de fin de calcul », « de calcul » et « de transposition » :

Règles d'introduction :

$$\mathbf{R}^i_1 : \; | X \;\; \Vdash \;\; \omega\, \alpha\, q_1\, X\, \nu$$

$$\mathbf{R}^i_2 : \; \omega\, \alpha\, q_1 \, | X \;\; \Vdash \;\; \omega\, \alpha\, q_1\, X\, \nu$$

Règles de fin de calcul :

$$\mathbf{R}^f_1 : \; \alpha\, q_0\, \nu\, X \;\; \Vdash \;\; \alpha\, q_0\, X \, |$$

$$\mathbf{R}^f_2 : \; \alpha\, q_0\, \omega\, X \;\; \Vdash \;\; X$$

Règles de transposition : toutes les règles de la forme

$$\mathbf{R}^t_{abcd} : \; a\, b\, c\, d\, X \;\; \Vdash \;\; b\, c\, d\, X\, a$$

a, c et d sont des symboles quelconques autres que $|$, et où b est un symbole autre que $|$ ou un explorateur.

Règles de calcul : à chaque règle

$$\mathbf{R}_k : \; X A_k Y \;\; \Vdash \;\; X B_k Y$$

de $\mathbf{M}_p$, correspond dans l'algorithme la règle

$$\mathrm{R}^c_k : \; \overline{A}_k X \;\; \Vdash \;\; X \overline{B}_k$$

où $\overline{A}_k$ et $\overline{B}_k$ sont les mots obtenus à partir de A_k et B_k en remplaçant partout $|$ par ν.

— *L'ordre des règles* est *indifférent*, car il est facile de vérifier que, par deux règles quelconques de l'algorithme :

$$\mathbf{R}' : A'X \;\; \vdash\!\vdash \;\; XB'$$

$$\mathbf{R}'' : A''X \;\; \vdash\!\vdash \;\; XB''$$

A′ n'est jamais le début de A″ [1].

— La règle $\mathbf{R}_2^f$ est *conclusive.*

8. Remarques finales.

J'espère avoir montré les rapports étroits qui existent entre les algorithmes de Markov et les systèmes de Post, le fait que ces rapports ne se comprennent bien que replacés dans une théorie générale des systèmes logistiques, et le fait que l'étude de ces rapports mène à une généralisation des « algorithmes normaux » définis par Markov (les algorithmes bilatéraux) qui peut être de quelque intérêt.

Je ne me dissimule pas que les considérations qui précèdent sont quelque peu arbitraires et manquent de généralité. Ainsi des algorithmes satisfaisant aux conditions des théorèmes 26, 27 et 28 auraient pu être construits autrement — peut-être plus simplement — peut-être par des méthodes plus uniformes. En fait, si nous disons qu'une classe **K** de systèmes de Post 1-simples est « universelle » lorsque tout système récursif d'alphabet fini a un élargissement conservatif dans **K**, si nous disons qu'une classe d'algorithmes est « universelle » lorsqu'elle contient des algorithmes qui calculent chacune des fonctions semi-récursives, et si nous convenons d'appeler $\mathbf{K}^0$ la classe de tous les algorithmes définis à partir des systèmes **K** en fixant :

— le choix à faire dans chaque application des règles de façon uniforme (voir § 7, déf. 40, 2° *a* — la méthode utilisée pour les algorithmes bilatéraux pourrait facilement être considérée comme un cas particulier d'une définition générale) ;

— l'ordre des règles, de façon arbitraire ;

— le choix des règles conclusives, de façon arbitraire.

Alors les systèmes bilatéraux et les systèmes unilatéraux constituent deux classes universelles de système de Post $\mathbf{K}_1$ et $\mathbf{K}_2$

---

1. Comparer Schützenberger 1956.

(théorèmes 22 et 25), telles que K_1^0 et K_2^0 sont des classes universelles d'algorithmes (théorèmes 26 et 27). D'où les problèmes :

Problème I : Ce fait est-il général ?

Problème II : Si oui au probl. I, peut-on trouver un procédé uniforme, valable pour toutes les classes universelles **K**, pour définir ceux des algorithmes K^0 qui calculent une fonction semi-récursive donnée, f, connaissant le système de Post appartenant à **K** qui a comme thèse les numéraux $\overline{f0}$, $\overline{f1}$ etc ... ?

Ce ne sont là que quelques-uns des problèmes non résolus soulevés par les définitions et les résultats exposés ci-dessus.

Note bibliographique.

Voici les principaux ouvrages ou mémoires dans lesquels on peut trouver facilement des renseignements supplémentaires relatifs aux questions traitées dans l'article présent.

Systèmes logistiques. La théorie générale esquissée au § 2 ci-dessus n'est exposée nulle part. Elle est contenue en partie, au moins implicitement, dans tous les bons manuels de logique. Voir Church 1956 et Kleene 1952, aussi Rosenbloom 1950 et Tarski 1956.

Fonctions récursives. Exposés complets dans Kleene 1952, et Péter 1951. Un manuel élémentaire est Grzegorczyk 1957, malheureusement écrit en polonais. La définition utilisée au § 3 ci-dessus est empruntée à Robinson 1950.

Machines de Turing. Exposé clair et assez complet dans Kleene 1952 ch. XIII On pourra ainsi consulter Porte 1956 ainsi que les mémoires originaux (Post 1936, Turing 1936, Turing 1937).

Systèmes de Post. Le mémoire fondamental est Post 1943 ; on trouvera beaucoup de renseignements complémentaires dans Robinson 1950. Voir aussi Lorenzen 1955.

Algorithmes de Markov. La littérature est surtout en langue russe. Le travail fondamental est Markov 1954. On pourra consulter Curry 1956, Rose 1957 et Riguet 1956.

Note ajoutée sur épreuve.

Un ouvrage récent « *Computability and unsolvability* » by Martin Davis (Mc-Graw Hill, New York, 1958) traite d'un grand nombre de questions étudiées dans l'article présent et dans mon article PORTE 1956. Le lecteur aura intérêt à comparer la rédaction de R. DAVIS et la mienne ; elles sont complètement indépendantes l'une de l'autre.

REFERENCES.

BOURBAKI, N., *Théorie des ensembles, Fascicule de résultats*. Hermann et Cie, Paris, 1939.

BOURBAKI, N., *Algèbre. Ch. I : Structures algébriques*. Hermann et Cie, Paris, 1942.

CHURCH, Alonzo, *The calculi of λ-conversion*. Univ. Press, Princeton, 1941.

CHURCH, Alonzo, *Introduction to mathematical logic*. Vol. I. Univ. Press, Princeton, 1956.

CURRY, H. B., *Review of Markov 1954*. *Math. Rev.*, vol. 17 (1956), pp. 1038-1039.

DETLOVS, V. K., *Normalnyie algorifmy i rekursivnyie funktsii*. Dokl. Akad. Nauk SSSR, vol. 90 (1953), pp. 723-725.

GRZEGORCZYK, Andrzej, *Zagadnienia rozstrzygalności*. Panstwowe wydawnictwo naukowe, Varsovie, 1957.

KALMÁR, László, *An argument against the plausibility of Church's thesis*. Communication au colloque « *Constructivity in Mathematics* » (Amsterdam, août 1957). Sous-presse.

KLEENE, S. C., *Introduction to metamathematics*. Van Nostrand, New York, 1952.

LORENZEN, Paul, *Einführung in die operative Logik und Mathematik*. Springer, Berlin, 1955.

MARKOV, A. A., *Teoriia Algorifmov*. Trudy Mat. Inst. Steklov, n° 42. Izdat. Akad. Nauk SSSR. Moscou, 1954.

MARKOV, A. A., *Review of Novikov 1955*. *Math. Rev.* vol. 17 (1956), pp. 706-708.

NOVIKOV, P. S., *Ob Algorifmitcheskoi nerazechimosti problemy tojdestva slov v teorii grupp*. Trudy Mat. Inst. im. Steklov. n° 44. Izdat Akad. Nauk SSSR. Moscou, 1955.

PÉTER, Rósza, *Rekursive Functionen*. Akadémiai Kiadó, Budapest, 1951.

PORTE, J., *Une simplification de la Théorie de Turing*. Comptes-rendus du 1er Congrès International de Cybernétique (Namur, 1956), sous presse.

POST, Emil L., *Finite combinatory process ; formulation I*. Jour. Symb. Logic, vol. 1 (1936), pp. 103-105.

POST, Emil L., *Formal reductions of the general combinatorial problem*. Amer. Jour. of Math., vol. 65 (1943), pp. 197-215.

RIGUET, Jacques, *Algorithmes de Markov et Théorie des machines*. Comptes Rendus Ac. Sciences, t. 242 (1956), pp. 435-437.

ROBINSON, Julia, *General recursive functions*. Proc. Amer. Math. Soc., vol. 1 (1950), pp. 703-718.

ROSE, Gene F., *Review of Markov 1954*. *Journal of Symbolic Logic*, vol. 22 (1957), pp. 77-79.

ROSENBLOOM, P. C., *Elements of Mathematical Logic.* Dover, New-York, 1950.

SCHÜTZENBERGER, M. P., *Une théorie algébrique du codage.* Comptes-rendus Académie des Sciences, t. 242 (1956), pp. 862-864.

TARSKI, Alfred, *Logic, Semantics, Metamathematics.* Clarendon, Oxford, 1956.

TURING, A. M., *On computable numbers, with an application to the Entscheidungsproblem.* Proc. London Math. Soc., ser. 2, vol. 42 (1936-1937), pp. 230-265.

TURING, A. M., *On computable numbers. A correction.* Proc. London Math. Soc., ser. 2, vol. 43 (1947), pp. 544-546.

ÉDITIONS J. DUCULOT, S. A., GEMBLOUX *(Imprimé en Belgique)*.