Memory and Mental Activity

Renzo Beltrame[‡]

Abstract. In this brief note we discuss some methodological consequences of studying memory phenomena in systems where the level of mental activity is sufficiently sophisticated to have conscious memories. As is characteristic of scientific praxis, the facts and procedures in the study must be repeatable without any restrictions on principles or methods. We discuss some aspects of a theory of the physical systems with memory, and we indicate a way to characterise the act of consciously remembering in intelligent systems. We then show the importance of behaviour theories that use mental facts as intermediate explanatory elements to build a model with the necessary complexity. We finally show how behaviour theories that use only physical processes as intermediate explanatory elements are particularly suitable for giving a unifying point of view.

In a previous paper[1] I discussed some methodological consequences on the study of mental activities, which follow from a prerequisite of scientific praxis; i.e., the requirement that both the investigated facts and the procedures employed in studying them be repeatable without any restrictions on principles or methods.

As an example, we will discuss some methodological problems that we find in studying memory phenomena in systems where the level of mental activity is sufficiently sophisticated to allow conscious acts of remembering. Like the previous one, this paper too is only an extended note on some aspects of the problem and so we will not refer to the long and well-documented history of these problems.

The most relevant consequence of the full repeatability requirement is that we cannot involve psychical or mental facts in an experiment either as dependent variables, or as independent variables, or as one of the parameters that characterises the experiment. Because psychical and mental facts are private in character, identifying them requires the testimony or the description of the person who is the subject of the experiment; and facts that have someone's account, description or testimony as a constitutive element, do not satisfy the full repeatability requirement. Accounts, descriptions, and testimonies can only serve as indications to get back directly to the fact we want to assume as an experimental fact.

When we apply these conclusions to the study of man, we conclude that only those things that concern the body, or the physical transformations which man performs on other objects can be dependent variable, or independent variable, or one of the parameters that characterise an experiment.

In developing a theory we are conversely free to choose any kind of fact, whether physical or not, as intermediate elements to explain the experimental results. The constraints only concern the things we want to explain or predict in the theory. These and the things that cause them must be physical facts, because they must be subjected to ex-

perimental control. Consequently, the intermediate elements in the explanation must be physical facts only if we want to check them through the experiments[2].

If we decide to satisfy the requirement of full repeatability, we can then introduce psychical or mental facts only as intermediate, explanatory elements in a theory of human behaviour; and the behaviour must be a physical behaviour, because what characterises it must occur as a dependent variable in an experiment.

The elements of mental type that occur in the theory will be described in way that shall be suitable for these kinds of things; in our case as activities, and by giving the related constitutive operations[3]. We have a choice here and so this particular choice is neither unique, nor necessitated; though it is equally plain that some choice has to be made about the kinds of things we will introduce in the theory as intermediate, explanatory elements, and about the way to describe them.

Consequently, we have no hierarchy in the knowledge from which a privileged point of view follows; instead we have a circularity in the knowledge, which requires an explicit declaration of the point of view adopted in each particular case.

The constitutive operations too, by which the mental facts are described, cannot be directly observable: they too can be neither dependent variables, nor independent variables, nor characterising parameters in a scientific experiment. But this fact is not a limitation from a methodological point of view. The validity of the explanatory, intermediate element that enters a theory, depends on how good the predictions are that the theory allows; and the predictions must concern, as we saw above, a physical behaviour, if we have decided to satisfy the requirement of full repeatability.

We can put a correspondence between the occurrence of non physical explanatory elements and the occurrence of certain physical facts. This correspondence is clearly part of the theory.

The linguistic behaviour offers a significant example of this situation. In this case, even the interpretation as word of the sounds and/or of the writing is really part of the theory that explains and foresees the behaviour. Thus the semantic convention that fixes a correspondence between some sounds or writings (considered as designating things, and their corresponding designated things), is also part of the same theory.

Theories of this type link mental facts, which are described, for example, by giving their constitutive operations, with pieces of linguistic behaviour. We may thus use the physical part of the linguistic behaviour to single out mental facts; but this procedure requires much caution, because we must usually look upon a very large context to obtain a reliable individuation[4].

The most favourable situation occurs when we have two theories that explain the same experimental facts, and one of them uses only physical processes to explain the observed behaviour, whilst the other uses mental facts too. In this case it becomes possible to pose a one to one correspondence between certain physical processes, or certain groups of physical processes, and certain mental activities, or the constitutive operations by which these activities are described.

This correspondence offers other advantages besides the immediately evident ones that follow by having a correspondence between mental facts, which are not directly observable, and physical processes, which are. The groupings that it introduces in the physical processes have a precise functional meaning in the theory that suggested them, and this meaning becomes a leading factor in allocating functional equivalencies among processes that occur in different biological architecture. Therefore, more general theories, which include different species, become easier to develop.

Returning now to our specific theme, we usually speak of memory phenomena as concerning physical systems, when in the theory that describes the behaviour of the system, the values of the variables associated with the system depend, at a certain instant of time, both on the values that these and other variables assume at the same instant, and on the values that they assumed in the past[5].

The most immediate way to describe memory phenomenon in a physical system is to imagine the processes that occur in the system to induce some modifications in the material of which the system is made. The modifications chosen usually satisfy a *locality* principle[6]; i.e., the changes in each part of the material only depend on what happened in the past to that part.

Modifications of this type are usually thought of as being permanent too; i.e., we assume that their effects on the behaviour of the system manifest themselves without any change after any interval of time where no further modification occurred in the material. The technique offers several examples of objects (e.g., magnetic disks commonly employed in computers) where this way of considering a physical system with memory is particularly evident.

If we imagine the parts of a physical system as interacting among themselves — i.e., if we refer to a scheme where the change in a physical quantity at a certain point in the system is considered as the cause of the changes of the same or of another physical quantity at a different point — we have to decide if the delay in which the effect follows the cause is significant or not[7].

When this delay is significant, the values of a physical quantity at a certain point and time depend on the values that the same or other physical quantities assume at different points and at past instants of time. If the interaction among the parts of a system is active for a long time, each of these values depends on analogous values assumed in other points, and at certain, still anterior, instants of time; and so on. To describe the evolution of the system we must then know the system’s history over very long intervals of time; even its whole life. On the other hand, the system's history ceases to affect its evolution, when the interaction among its parts stops for a sufficiently long interval of time.

The description of the processes that occur in a system of this type therefore requires the knowledge of the system's history until a situation where the system did not vary for a sufficiently long interval of time to ensure that it should stay in that state as long as external causes did not force it to change[8].

We have here a different way to describe memory phenomena in physical systems; and it is interesting to outline that in this case the system shows phenomena of memory without us having to assume changes in the material from which it is made.

The two ways of describing memory phenomena — one employing permanent changes in the material the system is made, the other using the delay in the interaction among the various parts of the system — provide different and complementary facilities, though both allow us to describe memory phenomena that span over the whole life of the system. We have highlighted the second facility, because the first is more widely employed, especially in biological systems[9].

We will avoid dealing with the very interesting, and often complex, problems that the mathematical treatment of physical systems with memory involves, and we will return to our original problem of discussing some facts about memory in living beings that show a level of mental activity so sophisticated to ensure the presence of conscious memories.

Clearly we can apply all the methods devised to study physical systems with memory to the living beings, because they too can be considered as physical systems. In doing this we develop a theory of their behaviour where all the intermediate, explanatory elements are physical processes.

Nevertheless, we can tackle the problems of memory in human beings from the other point a view. We can decide to develop a theory that will explain the physical behaviour observed, by using mental facts as intermediate explanatory elements. We shall subsequently choose to describe these mental facts by regarding them as activities and by giving their constitutive operations.

In the framework of a theory of this type, it was proposed that, when we speak of a thing as being remembered, a certain mental fact will be considered as a repetition of another mental fact, and the latter will be considered as having occurred in the past[10]. It was thus proposed that a mental fact will become a conscious memory as a result of a mental categorisation, which follows the scheme described above.

As a consequence of this hypothesis, we expect a lack of conscious memory of those facts a subject cannot actually produce as mental facts for disease reasons, although they occurred many times in his past. The subject, in fact, cannot attain the mental fact, which should be considered as a repetition of one that occurred in the past. And achromatopsias are known, which follow from brain lesions, where an adult man loss the ability not only to perceive, but also to remember colours; even if he had perceived and remembered colours several times before incurring the disease.

Thinking of one thing as being a repetition of another implies that the two things are equal; and, depending on the previous choices, the equality will concern the constitutive operations of the mental facts. Furthermore, what is thought to have occurred in the past is used as paradigm in the comparison that is part of the constitutive operations of the equality.

The categorisation activity that we perform by thinking of something as a memory, conforms to this scheme. We become aware of this fact, when we find a disagreement with the paradigm. We may find by different means this disagreement, e.g., by means of factual or document checking, testimonies, etc., after having considered something as a memory. In these cases we usually decide to explain the failure of those equalities we expect because of the applied mental categories, by inserting suitable causes.

Expectations usually arise from mental categorisation[11]; and, when no check occurs, the subsequent behaviour continues as if the expected consequences hold[12]. This quite general fact assumes particular relevance in our case.

What is considered as a memory (and is thus considered as a repetition of something that occurred in the past), is considered as a repetition of something that occurred in the past, concerning the subsequent behaviour too. The stimulus is thus weakened to check whether the equalities really exist, which we would expect because of the applied mental categories; and this effect will become progressively strong when such situation repeats. Motivations, of which the person might not be completely aware, can strengthen the tendency to do no check.

Furthermore, a subsequent memory can base itself on a previous one, rather than on the original situation; i.e., the constitutive operations that occurred in the previous memory will promote the categorisation of the actual constitutive operations as a memory, instead of the original ones.

The main consequence is that we can have facts with the character of good memories, which may either result as not having occurred, or, when a check is performed, re-

veal significant differences from those ones a person considers as memories. Furthermore the person will consider these facts as really pertaining to his past life, with all the relevant consequences on his behaviour.

We explain certain behaviours in human beings as being the result of having something as a conscious memory. Therefore, in a theory where we explain the observed behaviour only by means of physical processes that occur in the biological substrate of the individual, some of these processes should be considered as the counterpart of the categorial facts we discussed above. It is therefore necessary that those processes be explicitly present in the theory, which corresponds to considering something as a memory.

This very general consequence holds for all the mental activities. It only becomes particularly evident when mental categories are involved, because the necessity now clearly arises of having different series of physical processes: one as the counterpart of the constitutive operations of a particular mental thing, and the other as the counterpart of the constitutive operations by which that thing is considered as cause, or as an effect, or as a memory, etc.

Nevertheless we have experimental evidence that things are considered as a memory, which are plainly different for an external observer. An adequate theory of the behaviour must therefore allow that things to which different physical processes correspond, might be categorised as memories; and it is not necessary to assume that equal physical processes — i.e., the repetition of the physical process that is the counterpart of a mental fact — imply the occurrence of the categorisation as a memory.

A theory thus seems preferable where the activity of categorising something as a memory occurs following rules that do not strictly depend on the equality of physical processes that are the counterpart of the mental fact we are considering as a memory[13].

On the other hand, in studying systems that we have associated with a mental activity, we find that the physical counterpart of the conscious activity only covers a part of the physical processes that occur in the system. It is sufficient to observe, for this purpose, that the conscious activity is, by definition, filtered by the selective function of the attention.

Thus a theory would exclude a significant part of the physical processes that occur in such living beings, if it only considered the conscious activity. Thus, as general consequence, both the behaviour and the occurrence of the mental facts, which we introduce as intermediate explanatory elements in a theory, cannot be entirely explained by the mental facts that previously occurred to the subject[14].

We do not find any methodological obstacle to imagining a theory of the behaviour where all the intermediate explanatory elements are physical processes that occur in the biological material of the particular individual, we will have memory phenomena and chaining rules among the processes that can only depend on the material and architecture of the biological system, and on the present and past physical processes that occurred in the system.

The practical difficulties are quite a different thing. It is not easy, when the system is really complex, yet to find a good definition of the state of the system; i.e., a reasonably limited set of physical quantities whose value characterises the system at an instant of time. Nevertheless this approach is particularly fruitful, because it unifies the treatment of all the facts, which are relevant for a theory of the behaviour[15].

We will then conclude this short note by outlining that, in the study of systems to which we ascribe a mental activity, a theory of their behaviour assumes great importance where the intermediate explanatory elements might be mental facts, since this approach

favours a model with a satisfactory degree of complexity. To ensure a unifying point of view and a good completeness, we need, instead, a theory where all the intermediate explanatory elements are physical facts or processes that occur in the individual or species under study.

Notes

‡ NATIONAL RESEARCH COUNCIL OF ITALY – CNUCE INSTITUTE – Via Santa Maria 36, I-56126 Pisa (Italy) - *phone* +39 (50) 59 3211 - *email*:: r.beltrame@cnuce.cnr.it

A first version of this paper appeared as CNUCE Report C92-15

[1] R. Beltrame, “On Brain and Mind”, *Methodologia,* **10** (1992), pp. 7-13.

[2] A theory must usually give a certain number of provisions that experiments shall verify, to be considered a good theory. However repeatability does not imply this requirement.

[3] An interesting characterisation of physical, psychical, and mental, things, and of other distinctions related to the study of the mental activity, were proposed, in the framework adopted for this work, in S.Ceccato, *Un tecnico tra i filosofi*, Vol. II, Marsilio, Padova, 1966, pp.1-65 (in Italian).

[4] The correspondence is not in general so simple that we may fix it as a one to one correspondence at the level of single, isolated words of a language. In fact, the correspondence that one would put between single words of a language and mental facts is not, in general, *context free*, but strongly dependent on the context.

[5] This dependence can only involve values assumed at certain instants of time in the past, or the way of varying in certain intervals of time.

[6] A general theory of constitutive relations in continuum mechanics can be found in C. Truesdell and W. Noll, “The Non-Linear Field Theories of Mechanics”, *Encyclopaedia of Physics*, Vol. III/3, Springer-Verlag, Berlin 1965.

[7] When the effect in the interaction follows the cause with a certain delay, it is usual to speak of delayed action, or of delayed interaction, both if the cause and the effect occur at the same point, and if they occur at different points. When the delay is considered significant, it is often satisfactory to express this delay as a linear function of the distance between the two points where the changes of the physical quantities occur; and, in this case, the term ‘propagation speed’ often designates the constant rate in the linear function. Nevertheless, the reasons for introducing this concept in a theory, with the related problems about the thing that travels from one point to another, really concern the decision to write balance equations of certain physical quantities, which must hold at every instant of time both for the system, and for its parts. When the system occupies a region of space such that we can neglect the delay of the interaction, we can simplify the study of the system, because the knowledge of the external actions can be substituted by the knowledge of the values that the significant physical quantities assume on a closed surface that

envelops the system. This way of studying a physical system is usual in laboratory experiments; nevertheless it leaves the problem open of knowing when those physical quantities will assume given values on the closed surface that envelops the system; and this is the main problem when we study a system *in vivo*.

[8] We do not consider the so-called isolated systems. An isolated system will lie in the same state for ever, when it reaches a stationary state; i.e. a state where no change occurs in every part. Thus it is of no interest in this discussion.

[9] We recall that the occurrence of memory phenomena of this type is very frequent in natural systems. Systems without memory are nevertheless of high theoretical interest because of their simple mathematical treatment, and because the actual production of the artefact concerns systems with a behaviour strictly stereotyped, repetitive, and history independent.

[10] This characterisation was proposed in S. Ceccato, *La fabbrica del bello*, Rizzoli, Milano, 1987,
pp. 234-36 (in Italian).

[11] We are discussing here a fact that occurs generally in mental activities. For example, the particular perception that follows from a given stimulus pattern induces some expectations; and we become aware of this effect only when the consequences are negative, such as when we misjudge distances when driving a car. The expectations that a certain mental activity induces, usually have different effects on different behaviours, and this fact complicates the theory.

[12] We avoid talking about consequences that are assumed to be true or verified, because a check is implied, which was excluded by hypothesis.

[13] The probability of observing the repetition of a given sequence of physical processes will be very low in a system where remarkable memory phenomena result from a delay in the interaction among its parts too. Thus we will observe a low probability in the occurrence of the categorisation as a memory, although we will suppose that the repetition of sequences of physical processes promotes such a categorisation.

[14] It is often difficult to find a good definition of the state of a complex system; i.e., to find a set of physical quantities that characterises the system in a given configuration.

[15] The problem of specifying the moment and the causes of the occurrence of a mental fact is present in the studies of the Scuola Operativa Italiana, though it was often overcome by the stress given to the analysis of the mental constructs that were considered particularly significant. The conditions of the occurrence of a mental fact were called *dependencies*.