THE CYBERNETIC ONTOLOGY OF ACTION

Valentin Turchin

The City College, the City University of New York

1 Introduction. Metaphysics now

Metaphysics is often viewed as almost opposite to physics and utterly useless for it, if not for any reasonable purpose. This attitude is a hangover from outdated forms of empiricism and positivism, namely the naive reflection-correspondence theory of language and truth, which sees language as an image, a replica of the world. It is easy to conclude from this theory that any expression of our language which cannot be immediately interpreted in terms of observable facts, is meaningless and misleading. This viewpoint in its extreme form, according to which all unobservables must be banned from science, was developed by the early nineteenth-century positivism (August Comte). From this perspective, metaphysics is definitely meaningless.

The history of the Western philosophy is, to a considerable extent, the history of struggle against the reflection-correspondence theory. Now we understand language as a hierarchical model of reality, i.e. a device which produces predictions, and not as a true picture of the world. This device, especially in its higher levels of structure, need not 'look like' the things it is about; it only should produce correct predictions. Therefore, the claim made by metaphysics is now read differently. To say that the real nature of the world is such and such means to propose the construction of a model of the world along such and such lines. Metaphysics creates a mental structure to serve as a basis for further refinements. Metaphysics is the beginning of physics; it provides fetuses for future theories. Even though a mature physical theory fastidiously distinguishes itself from metaphysics by formalizing its basic notions and introducing verifiable criteria, metaphysics, in a very important sense, *is physics*.

For a cybernetician metaphysics should be more than of a detached interest. On our agenda is the creation of universal models of the world, which would alow us, in particular, to interpret human thought expressed in natural language. How should we start this enterprise? What concepts must be taken as the basis? It is the business of metaphysics to give answers to these questions.

In this paper we present an argument for seeing *action* as the ultimate reality of the world, and therefore, for taking action as the elementary building element in the construction of world models. This is a truly cybernetic approach. Physics is concerned with the material of the world, the matter-energy aspect of it. Cybernetics abstracts from the material and concentrates on control, communication, information. All of these are actions.

Intuitively, we see the world as a collection of objects occupying some space and changing in time. Objects are seen as primary, change as something secondary, which could or could not take place. The metaphysics we develop here reverses this relationship. We modify Schopenhauer's formula as *The world is action and representation*, with action taking ontological precedence over representation, such as our perception of objects. Thus we take the concept of action *in abstracto* and on this basis try to interpret the fundamental concepts of our knowledge: what are objects, what is objective description of the world, what are space, time etc.

2 From Kant to Schopenhauer

Kant synthesized empiricism and rationalism by seeing knowledge as organization of sensations by our mind. Space, time, and other categories are not given us in sensation. They are our forms of perception, the way we organize sensations. This is how *synthetic judgments a priory* become possible. They reveal the working methods of our mind which are inborn and do not depend on sensations.

In the light of cybernetics, Kant's ideas are surprisingly modern. Now we say that sensations are at the input of our cognitive apparatus, the nervous system. This input is then processed by a huge hierarchical system. As the signals move up in the hierarchy, sensations become perceptions (there are no sharp boundaries, of course).

Mach and Einstein would be, probably, impossible without Kant. They used the Kantian principle of separating elementary facts of sensations and organizing these facts into a conceptual scheme. Einstein's analysis went further from the intuitive space-time picture given by classical mechanics down to the level of separate measurements, and resulted in reorganization of measurements into a different space, the four-dimensional space-time of the relativity theory. This space-time is now as counterintuitive as it was in 1905, even though we have accustomed to it. Hence what we call the paradoxes of the relativity theory. But they do not bother us. We use a bit less of neuronal models of the world, and a bit more of symbolic models, that is all.

In quantum mechanics, the physicists went even further. They rejected the idea of a material body located in the space and time as the ultimate reality. The space-time continuum is left as a mathematical construct, and this construct serves the purposes of relating micro-phenomena with macro-phenomena, where it has the familiar classical interpretation. But elementary particles lost their tangible character. In the relativity theory, observations (measurements) at least belong to the same universe as the basic conceptual scheme: the space-time continuum. In quantum mechanics, on the contrary, there is a gap between micro-world and macro-world, between what we believe *to really exist*, i.e. quantum particles and fields, and what we take as the basic observable phenomena, which are all expressed in macroscopical concepts: space, time and causality.

Here we face the most intriguing part of metaphysics: the concept of 'real existence'. Cybernetic epistemology, according to which all meaningful statements are hierarchical models of reality (see [1]), has a double effect on the concept of existence. On the one hand, theoretical concepts, such as mechanical forces, electromagnetic and other fields, and wave functions, acquire the same existential status as the material things we see around us. On the other hand, quite simple and trustworthy concepts like a heavy mass moving along a trajectory, and even the material things themselves, the egg we eat at breakfast, become as unstable and hazy as theoretical concepts.

One can argue that since every theory, in the last analysis explains and organizes observable facts, which all are, and will always be, macroscopic facts, there is simply no need in the concept of real, or ultimate, existence. This is formally true. But we still feel a need for our theory to give an answer to the question of ultimate existence. What is the ultimate reality of physics? This question is not meaningless. Its meaning is in the quest for a theory which would start with concepts believed to correspond to that ultimate reality, and then step by step construct observables from these "really existing" things. Somehow it seems that such a theory has better chances for success. If we have a theory of that kind, and it constructs the world from some things — call them 'ex-why-zeds' — and the theory is born out by experiment, then we can say that the ex-why-zeds *do really exist*, and that the world really consists of ex-why-zeds. Ontologically, this will be as certain as when we say that the apple is in a bowl on the basis of seeing it and touching it.

Suppose we are determined to construct a theory which is built as required above. How should we go about the construction of such a theory? We must go further down in the hierarchy of neuronal concepts. Space and time must not be put in the basis of the theory. They must be constructed and explained in terms of really existing things. An attempt must be made to identify the most essential, pervasive, primordial elements of experience. Kant's metaphysics had served as the philosophical basis for the modern theories of physics. We see now that a further movement down is required. Thus let us turn to the development of metaphysics after Kant.

Kant introduced the concept of the thing-in-itself for that which will be left of a thing if we take away everything that we can learn about it through our sensations. Thus the thing-in- itself has only one property: to exist independently of the cognizant subject. This concept is essentially negative; Kant did not relate it to any kind or any part of human experience. This was done by Schopenhauer. To the question 'what is the thing-in- itself?' he gave a clear and precise answer: it is *will*. The more one thinks about this answer, the more it looks as a revelation. My will is something I know from within. It is part of my experience. Yet it is absolutely inaccessible to anybody except myself. Any external observer will know about myself whatever he can know through his sense organs. Even if he can read my thoughts and intentions – literally, by deciphering brain signals – he will not perceive my will. He can conclude about the existence of my will by analogy with his own. He can bend and crush my will through my body, he can kill it by killing me, but he cannot in any way perceive my will. And still my will exists. It is a thing-in-itself.

What is then the rest of the world as we know it? Schopenhauer answers: a 'Vorstellung', representation. Schopenhauer's formula for all that exists is:

$$\textit{the world} = \textit{will} + \textit{representation}$$

3 Action ontology

Am Anfang war die Tat.[1]
Goethe.

Will is manifested in *action*. Will and action are inseparable. Our understanding of will is this: will is not a description of options the subject has, nor is it a list of preferences of the subject – these are all representations; will is the quality that allows to choose between the (possible) options and act. Action and will are two faces of essentially the same phenomenon, and in the philosophy we are developing, action is its perceivable part. A human subject that performs an action usually sees it from within as an action of his will. But not every action, of course, is an action of a human subject; we regard them as manifestations of some other, non-human, wills. We rewrite Schopenhauer's formula as follows:

$$\textit{the world} = \textit{action} + \textit{representation}$$

If we are looking for the ultimate undoubted reality, we must turn to action, and not to the space-time picture of the world. For a picture is only a picture, a representation which changes from one subject to another, from one theory to another; while action is an irrefutable reality.

Our knowledge of anything in the world starts with sensations. Sensations are not things. They do not have reality as things. Their reality is that of an event, an action. Sensation is an *interaction* between the subject and the object, a physical phenomenon.

Consider the concept of action in the context of physics. According to our present understanding of the world, all the variety of events we observe result from elementary acts of interaction between elementary particles. These acts constitute unquestionable reality, while both our theory, and our intuitive picture of the world, are only representations of reality.

Furthermore, it is the physical quantity of *action* that is quantized by Plank's constant h. This can be seen as an indication that action should have a higher existential status than space, time, or matter. Of course, it is not immediately clear whether the concept of action as we understand it intuitively and the physical quantity that has the dimension of energy by time and called 'action' are one and the same, or related at all. That the physicists use the word 'action' to denote this quantity could be a misleading coincidence. Yet the intuitive notion of an action as proportional to the intensity (intuitive understanding of energy) and the time does not seem unreasonable. Furthermore, it is operators, i.e., *actions* in the space of states, that represent observable (real!) physical quantities in quantum mechanics, and not the space-time states themselves!

[1] In the beginning there was the deed

Even if we reject these parallels and intuition as unsafe, it still remains true that neither space, nor time, nor matter are characterized by a single constant indestructible quantum, but a combination of these. Is it not natural to take this combination as a basis for the picture of the world — if not for a unifying physical theory?

The purpose of metaphysics is to find in our experience the most fundamental elements or aspects of the world. We take actions as such, which means that in the model of the world we are constructing the lowest level of representations consists of representations of actions.

Action and event are prominent features of reality, so philosophers explored them since very long. The idea that events may be, in some sense, more primary than space and time has been appearing, now and then. Russell [2], in his treatment of time and space, takes "as raw material" events,

> *which are to be imagined as each occupying a finite continuous portion of space-time. It is assumed that two events can overlap, and that no event recurs.*

The motive is to explain and justify the continuing use by the physicists of instants in time and points in space while rejecting at the same time Newton's conception of absolute time and space, which conferred to instants and points a great deal of ontological primacy. Using events, Russell defines instants as follows:

> *An 'instant' as I propose to define them, is a class of events having the following two properties: (1) all the events in the class overlap; (2) no event outside the class overlaps with every member of the class.*

Russell's event is, essentially, a set-theoretic concept. It is a set of instants of time. It remains so even when instant (element) is defined through set (event) as the intersection of a certain family (class) of sets. This class is nothing but the set of all events which include a given instant. Typically for a set-theoretic approach, all these concepts are static, do not really involve action.

Action proper is treated by several contemporary philosophers, including Aune, Davidson, Quine, Sellars and others. To quote Bruce Aune [3],

> *Perhaps the most controversial aspect of so called action theory is its subject matter. This subject matter is generally said to be (or to concern) actions, but different philosophers conceive of actions in radically different ways. For some philosophers, actions are abstract entities – states of affairs, propositions, sets, or even ordered pairs of some kind. For others, actions are distinctively concrete entities located in space and time. Another group of philosophers, among whom I include myself, have even denied that actions are required for a reasonable action theory, insisting that agents or actors will suffice as the theory's sole objects.*

For us, actions are concrete, in the sense that they constitute the primary reality and not man-made representation, such as propositions, sets etc., which are referred in the quote as 'abstract entities'. At the same time we cannot say that actions are located in space and time; it is space and time that are constructed from actions. The high ontological status of action which we maintain is usually rejected on the grounds that action is not 'fundamental' enough. Thus Aune writes:

> *According to an agent theory, although agents clearly exist and may truly be described as acting in this or that way, it is philosophically misleading to say that their actions or deeds also exist.*

Aune justifies his view-point by using the concept of fundamental realities, which is traced back to Aristotle. A pile of stones is less fundamental than the stones which make it up. If a pile exists at all, then it does so in some weaker form than the stones. We read further in Aune's paper:

> *Philosophers holding what are known as substance ontologies contend that the fundamental objects of our world are continuants – and nothing but continuants. A continuant is a thing like a man,a marble, or a tree: something that, as Aristotle said, persists in time and can undergo change. Philosophers holding an agent theory often accept such an ontology. If fundamental objects are continuants, then changes, events, and therefore actions are not fundamental objects; their existence is derivative at best.*

Our view is exactly the opposite: actions constitute the reality which we perceive through our sense organs; agents are representations, elements of a language, which we use to construct models explaining the observed actions. This view is *action ontology.*

The reason why we espouse this ontology is inseparable from our cybernetic epistemology. We do not take the concept of existence as given intuitively, nor do we think about it as a reflection of 'real' existence. For us, fundamental things, i.e. things that have the greatest degree of existence, are those we start with when we construct our models of the world: the cornerstones of construction. To pick up those cornerstones we look for the most certain, unquestionable facts of life, and we find that they are actions (the word *fact* itself comes from the Latin *facere*, which means to do or to make). Moreover, our knowledge itself, i.e. a collection of the world's models, is nothing but action; the only meaning of a model is in its operation. Representations are arbitrary: a matter of convenience; the changes in representations must correspond to the changes in reality if we want to have a true model. The translational relativity (symmetry) of space is the best illustration of this. When we think of an isolated point, we do not yet think about space. Space is created by certain actions: moves, or shifts. These shifts are a measurable reality. The points in space have meaning only with respect to some reference system. The same shift on a line can be seen as a shift from point 3 to point 7 or from point 12 to 16. The choice of a reference – i.e. representation – system is arbitrary.

3.1 Will

In our thought and language we distinguish two different classes of elements about which we say that they exist: those expressing what we know, or think we know, and those expressing what we are striving for and intend to do. We unite the elements of the first class to referred as *knowledge*, and the elements of the second class as *will*. They are not isolated from each other. Our goals and even our wishes depend on what we know about our environment. Yet they are not determined by it in a unique way. We clearly distinguish between the range of options we have and the actual act of choosing between them. As an American philosopher noticed, no matter how carefully you examine the schedule of trains, you will not find there an indication as to where you want to go.

Another way to describe the relation between will and knowledge is as a dichotomy between 'I' and 'not-I', or between subject and object. The border between them is defined by the phrase 'I can'. Indeed, the content of my knowledge is independent of my will in the sense that I cannot change it by simply having some intentions or preferences. On the contrary, I can change my intentions without any externally observable actions. I call it my will. It is the essence of my 'I'.

It is only *my* will, i.e. the will of the subject of knowledge, that exists as will. *Its* will, and *their* wills, if they exist (of course, they do), exist only as my representations.

If Kant's view of knowledge has a clear cybernetic interpretation, then even more so has Schopenhauer's view of the world. His formula is borne out by the practice of cyberneticians during the last decades. We try to understand ourselves by building cybernetic creatures and computer programs which model intelligent behavior. Our artificial models of intellect consist of two parts: a device that collects, stores and processes information; and a decision maker – another device that keeps certain goals and makes choices in order to reach these goals, using the information from the first device. Thinking about ourselves in those terms we speak about our knowledge and our will. It is there, and there is nothing beyond it.

3.2 Freedom

The concept of will assumes the existence of *freedom* to exercise the will. Thus recognizing will as a cornerstone of being, we do the same for freedom. For the mechanistic worldview of the nineteenth century freedom was a misconcept, a nuisance which escaped satisfactory definition within the scientific context. For us freedom is the very essence of the things, and, first of all, of the human person.

However, in many minds, science is still associated with the deterministic picture of the world, as it was in the nineteenth century. This picture, was as follows.

Very small particles of matter move about in virtually empty three-dimensional space. These particles act on one another with forces which are uniquely determined by their positioning and, possibly, velocities.The forces of interaction, in their turn, uniquely determine, in accordance with Newton's laws, the subsequent movement of particles. Thus each subsequent state of the world is determined, in a unique way, by its preceding state. Determinism was an intrinsic feature of the scientific worldview of that time. In such a world there was no room for freedom: it was illusory. Humans, themselves merely aggregates of particles, had as much freedom as wound-up watch mechanisms.

In the twentieth century the scientific worldview has undergone a radical change. It has turned out that subatomic physics cannot be understood within the framework of the naive realism of the nineteenth century scientists. The theory of relativity and, especially, quantum mechanics require that our worldview be based on critical philosophy, according to which all our theories and mental pictures of the world are only devices to organize and foresee our experience, and not the images of the world as it "really" is. Thus along with the twentieth-century's specific discoveries in the physics of the microworld, we must regard the inevitability of critical philosophy as a scientific discovery – one of the greatest of the twentieth century.

We now know that the notion that the world is "really" space in which small particles move along definite trajectories, is illusory: it is contradicted by experimental facts. We also know that determinism, i.e. the notion that in the last analysis all the events in the world must have specific causes, is illusory too. On the contrary, freedom, which was banned from the science of the nineteenth century as an illusion, became a part, if not the essence, of reality.

There is genuine freedom in the world. When we observe it from the outside, it takes the form of quantum-mechanical unpredictability; when we observe it from within, we call it our free will. We know that the reason why our behavior is unpredictable from the outside is that we have ultimate freedom of choice. This freedom is the very essence of our personalities, the treasure of our lives. It is given us as the first element of the world we come into.

Logically, the concept of free will is primary, impossible to derive or to explain from anything else. The concept of necessity, including the concept of a natural law, is a derivative: we call necessary, or predetermined, those things which cannot be changed at will.

3.3 Agent

When we speak of an action, we speak also of an *agent* that performs the action. An agent is the carrier of will, the entity that chooses between possible actions. We do not see agents, we see only what they are doing. But we use the concept of agent to create models of the world.

When we speak of actions of human beings we know very well what the agent is: just the person whose action it is. We reconstruct this notion, of course, starting from our own 'I'. When we speak of such animals as dogs, we again have no doubt in the validity of the concept *agent*. This reasoning can be continued down to frogs, worms, amebas, trees, and inanimate objects, without any convincing arguments for stopping. When we say: "the bomb exploded and the ship sank", are there any reasons to object against understanding this in the same way as if we were speaking abut people and dogs? After all, if the bomb was not very big, the ship might or might not sink, depending on the *ship* itself, the ship as a whole. Notice that even given a definite ship and a definite time, the result might not be uniquely predetermined.

And what about an *act* (sic!) of radioactive decay? It is definitely an action, but whose action is it? The physicist could say that the agents here are electrodynamic and chromodynamic fields. This make sense because of the theory the physicist has. If we do not have such a theory, we simply say that there is a special agent for each possible act of radioactive decay. At each moment in time this agent makes a choice: to decay or not to decay. This immediately explains the exponential law of radioactivity.

Introduction of agents is, speaking informally, our first theory of the world. The primary instance of an agent for a human being is itself. So, it is not surprising that in primitive societies the concept of agent is understood anthropomorphically: as something which is very similar, if not identical, to ourselves. Hence the animism of primitive thinking: understanding of all actions as initiated by various kinds of spirits or other imaginary creatures.

The development of modern science banned spirits from the picture of the world. But agents, cleared from anthropomorphism, still remain, even though the physicists do not call them so. What is Newtonian force if not an agent that changes, every moment, the momentum of a body? Physics concentrates on the description of the world in space and time; it leaves – at least at present – the concept of agent implicit. We need it explicitly because of our metaphysics based on the concept of action, not to mention the simple fact that cybernetics describes, among other things, the behavior of human agents. (This last field of application of cybernetics is, of course, one of the reasons for our metaphysics).

3.4 Emergence

Agents come into, and go out of, existence. For centuries philosophers grappled with a problem: how to distinguish simple ("quantitative") changes from the cases where something really "new" emerges. What does it mean to be "new", to emerge? In our theory this intuitive notion is formalized as the coming of a new agent into existence. An action can lead to an emergence of new agents.

Take radioactive decay. A neutron suddenly *chooses* to break down into a proton, electron and neutrino. Whatever agents could have been involved into the events around the neutron do not exist anymore. New agents emerge, such as the interaction between the newborn proton and electron.

In the case of complex actions, such as the birth of a baby, we can argue about the exact time of the event, because we have more than one reference system in which to describe actions. As a member of society, the baby emerges at birth. As an object of embryology it emerges at the moment of egg fertilization.

3.5 Event

When we ignore the agent, we speak of actions as events. Event is an action abstracted from the agent.

4 State

The idea of a *state* of the world, or some part of the world, is familiar to everybody who took elementary courses in science. It is usually considered so basic that there is no need and, probably, possibility of definition. But we need a definition, because the only indefinable element in our metaphysics is action.

Our definition is: a state of a part of the world is the set of actions that are possible in this state, with their probabilities, if this concept is applicable. Two states in which all the same actions are possible and equally probable are the same state, because there is no way to distinguish between them.

For example, if the state of affairs is such that there is an apple on the table in front of me, I can reach it and pick it up. If there is no apple this is impossible. If the moon is on the night sky, I can execute the action of observing it. For this purpose I rotate my head in a certain way and keep my eyes open. An atom is in an excited state when it can emit a photon.

There are states of a type different from what we have considered above, the states for which our definition is not suitable. If I feel pain, or am frustrated, or elated, angry, or complacent, this has no effect on the actions I can take. It affects only the choices I am going to make selecting from the same set of possible actions. Indeed, if my hand is over a gas heater and hurts (say, gently, for plausibility), I still have the choice between keeping the hand where it is, or withdrawing. But, obviously, the more it hurts, the more likely I am to withdraw it.

Thus we come to distinguish between:

- a physical state, which is a set of possible actions for the subject of knowledge and other agents; and
- a mental state, which influences the choices to be made by the subject, but does not alter the set of possible actions.

This dichotomy clearly reflects the fundamental feature of the world as consisting of the will and representation. A physical state is that of representation. A mental state is a state of the subject's will. When we use the concept of an agent in out model of the world we may endow agents with some 'mental' states. However, such a state remains physical for all other agents: part of the representation of the world. It is only the subject's mental states which belong to the category of will. Other agent's will is for me only a representation.

When speaking of "states" without any of the two adjectives, we shall mean physical states.

We have already had a number of reasons for considering action the most fundamental observable reality. Seeing action in the context of state provides more of it. The concept of state is a strong contender for the most basic role.

Indeed, the standard beginning of a theory is to introduce states of the objects of the theory, whatever they are, and then define actions, which are understood as certain changes of the state. But when we define an action as a change of the state, we introduce something new, which is not present in the idea of a state; change is an event i.e. an action abstracted from the agent that performs it. Thus we cannot avoid introducing action as an indefinable element. At the same time the state of the world can be defined through action, as we have just demonstrated.

5 The algebra of actions

Algebra is the part of mathematics that deals with operations. Operations are actions. Since action is the basic reality of the world, algebra is the beginning of all beginnings.

A set of actions is referred to as a *domain*. It is a formalization of the idea of 'a part of the world'. The models we are given by nature and construct artificially are never universal. They are always applicable only to some part of reality. This part is the domain of the model. Since states of that part of reality are defined by sets of actions, the domain of a model also defines the set of all states which can, in principle, exist: it the powerset (the set of all subsets) of the domain. The actual set of possible states may be a subset of this powerset.

When we apply a model (in particular, a theory), we assume that only those actions take place that are within the domain. Make an action which is not included in the domain, and the whole theory may become out of place. The states of the world are defined as subsets of the domain of the model. Other actions are ignored; they may be either irrelevant, when they have no impact on the legitimacy of the model, or prohibited, when they make the model inapplicable.

We call a *null action* the absence of any action. An action which is not null is a *non-null* action.

There are two ways to unite actions into a composite action.

If a_1 and a_2 are actions then their *sequential composition*, denoted as $(a_1; a_2)$, or just $a_1 a_2$, is the action which is performed by first performing a_1 and then, immediately, a_2.

We say that action a_1 is *sequential part of* action a, if there exists an action a_2 such that $a_1 a_2 = a$.

If a_1 and a_2 are actions then their *parallel composition*, denoted as or $(a_1 \parallel a_2)$, is the action which is performed by performing a_1 and a_2 in parallel.

We say that action a_1 is *parallel part of* action a, if there exists an action a_2 such that $a_1 \parallel a_2 = a$.

A composite action is referred to as a *process*.

Given an action a, we denote as $\bar{a}$ the action such that $a\bar{a}$ and $\bar{a}a$ is an action which returns the world to the same state as before this composite action. For a given a the inverse action $\bar{a}$ may or may not exist. If it does exist, a is *reversible*, otherwise *irreversible*.

A *domain* is *continuous* if for every action a there exist two actions a_1 and a_2 such that neither is null and the composite action $a_1 a_2$ is the same as a.

Note, that the requirement for $a_1 a_2$ and a is *to be the same*, and not just lead to the same state.

An action a is *elementary* if $a \neq a_1 a_2$ for any two non-null actions a_1 and a_2.

A domain is *discrete* if every action from it can be represented by a finite sequence of elementary actions.

A domain may have discrete and continuous subdomains.

6 Cognitive action

Among the possible actions which constitute a state of the world there are some which are performed by the subject of knowledge with the explicit purpose of increasing the subject's knowledge. We shall call them *cognitive actions*. Actually, there is no sharp boundary between purely cognitive actions and other actions which serve different purposes. Each action can be considered cognitive to the extent it takes part in the formation of a model of the world.

When I rotate my head to see the moon, this is, obviously, a good cognitive action. To take an apple is to a great extent a cognitive action for a baby who is just now forming the concept of an external object. It is also cognitive for a blind man who has no other way to know that an apple is there. But if you take an apple after you have seen it, this hardly adds much to your mental model of the world. The act is alimentary rather than cognitive.

6.1 Modeling scheme

The basic principle of the cybernetic epistemology is that knowledge is a *model* of (a part of) the world. Below we define the concept of a model as it is most commonly used (see Fig. 1).

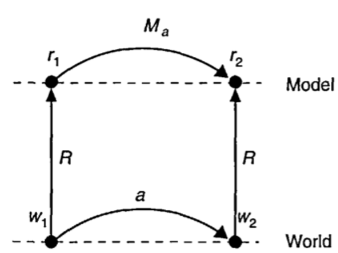

Figure 1: The modeling scheme

A model is a system which includes: (1) a certain subsystem, whose states will be referred to as *representations*; (2) a *representation procedure* R, such that if performed when the state of the world is w it causes a certain representation $r = R(w)$ in the system; (3) a family R_a of *modeling procedures* which depend on a possible action a in the world and convert one representation into another.

For the system to be a (correct) model, the modeling procedures must have the following property. Let the world be in a state w_1, and suppose that system S makes an action a, as a result of which the state of the world becomes w_2. Then the modeling procedure M_a applied to the representation $r_1 = R(w_1)$ produces the representation r_2 of the resulting state w_2:

$$M_a(R(w_1)) = r_2 = R(w_2)$$

.

Thus by applying M_a to r_1 the system can predict, to some extent, the development of events in the world resulting from its action a. This definition describes equally well the case of a dog catching in flight a piece of sugar, and an astronomer who computes the position of a planet in the sky. In the first case, the model is built in the material of nervous cells of the dog's brain, in the second case from the signs that the astronomer writes on paper when he makes computations.

Mathematically, a model is defined by the representation function $R(w)$ and the family of modeling functions $M_a(r)$, for possible actions a of the system. When we fix an action a, we have what is known as a *homomorphism.* Thus a model is a family of homomorphisms.

6.2 Observation

Observation is an abstraction of knowledge from the impact of our cognitive actions on the world. The state of the world w_i in Fig.1 and the action of using the procedure R are not really separable: it is an action of cognition, interaction of the subject and the object. When we speak of observation, we assume that there exist cognitive actions which only serve us to acquire knowledge, but have no influence on the phenomena we observe.

For instance, when we watch a party of billiards, the positions of the balls are registered by means of the light thrown on the balls and reflected into our eyes. We rightly believe that the effect of the lighting on the movements of the balls is negligible, so we speak about the play in a complete abstraction from the way we know about it.

This separation is not always possible. Quantum mechanics deals with actions so elementary that the means we use to know of them cannot be abstracted from. Our usual "classic" notions of space and time include the abstraction of observation. Indeed, the mental construction of a reference frame uses actions of shifting and waiting (doing nothing) which are assumed to have no effect on the studied processes; for quantum-mechanical processes this assumption is not valid, and the classic space-time frame of reference looses its legitimacy, becomes meaningless. One should not use in mental constructions and experiments the things known not to exist.

6.3 Objective description of the world

By objective description of the world we mean, first, a description in terms of some *objects*, and second, a description which is, as much as possible, "objective" in the usual sense, i.e. impersonal, not depending on the cognitive actions, or other features, of the subject of knowledge. As we shall see, the use of the same word in these two meanings is not accidental: a description can be "objective" because it is a description in terms of objects.

6.3.1 Object

Suppose I am aware of a tea-pot on the table in front of me. This is a result of my having the mechanism of abstraction in the brain. I recognize the image on my retina as belonging to a certain set of images, the abstraction 'tea-pot'.

But there is more to it. I perceive the tea-pot as an *object*. The object 'tea-pot' is certainly not a definite image on the retina of my eyes; not even a definite part of it. For when I turn my head, or walk around the table, this image changes all the time, but I still perceive the tea-pot as the same object. The tea-pot as an object must, rather, be associated with the transformation of the image on my retina which results from the changing position of my eyes. This is, of course, a purely visual concept. We can add to it a transformation which produces my tactile sensations given the position and movements of my fingers.

The general definition of an object suggested by this example consists of three parts.

(1) First we define a set R_{ob} of representations which are said to represent the same object; in our example this set consists of all images of the tea-pot when I look at it from different view-points, and possibly, my sensations of touching and holding it.

(2) Then from the set of all possible actions we separate a subset A_{cogn} of actions which will be referred to as *cognitive*; in our case A_{cogn} includes such actions as looking at the tea-pot, turning my head, going around the table, touching the tea-pot etc. – all those actions which are associated with the registration of the fact that a tea-pot is there.

(3) Finally, we define a family of functions $f_a(r)$, where for every cognitive action $a \in A_{cogn}$, the function

$$f_a : R_{ob} \to R_{ob}$$

transforms a representation $r \in R_{ob}$ into $f_a(r) = r'$ which is expected as a result of action a.

The most important part here is the third; the first two can be subsumed by it. We define an object b as a family of functions f_a:

$$b = \{f_a : a \in A_{cogn}\}$$

The set A_{cogn} is the domain of the index a; the set R_{ob} is the domain and co-domain of the functions of the family.

When I perceive an object b, I have a representation r which belongs to the set R_{ob}; I then execute some cognitive actions, and for each such action a I run my mental model, i.e. perform the transformation f_a on r. If this anticipated representation $f_a(r)$ matches the actual representation r' after the action a:

$$f_a(r) = r'$$

then my perception of the object b is confirmed; otherwise I may not be sure about what is going on. Observing a tea-pot I check my actual experience against what I anticipate as the result of the movements of my head and eyeballs. If the two match, I perceive the tea-pot as an object. If I travel in a desert and see on the horizon castles and minarets which disappear or turn topsy-turvy as I get closer, I say that this is a mirage, an illusion, and not a real object.

The concept of an object is naturally (one is tempted to say, inevitably) arises in the process of evolution. It is simply the first stage in the construction of the world's models. Indeed, since the sense organs of cybernetic animals are constantly moving in the environment, these actions are the first to be modeled. In the huge flow of sensations a line must be drawn between what is the result of the animal's own movements, and other changes which do not depend on the movements, are *objective*. Looking for objectivity is nothing else but *factoring out certain cognitive actions*. Function f_a factors out the action a by predicting what should be observed when the only change in the world is the subject's taking of the action a. If the prediction comes true, we interpret this as the same kind of stability as when nothing changes at all. The concept of object fixates a certain invariance, or stability, in the perception of a cybernetic system that actively explores its environment.

The metasystem transition from representations to their transformations is a step towards objectivity of knowledge. Actions and, in particular, sensations are intimately tied to the agent, the subject of knowledge. An object is a transformation and prediction of actions. The very fact that prediction is possible indicates that the transformation depends less on the subject of knowledge, the 'I', and more on the 'not-I'. This does not ensure a complete objectivity; alas, there is no such thing. But a jump from a representations to a transformation of representations verified by the practice of correct predictions, is the only way we know to increase the informally understood objectivity.

When we perceive a tea-pot as an object, we have a lot of cognitive actions to factor out: we can walk around it, grasp it, rotate it in from of our eyes etc. But often we observe things from afar and that is about all we can do, as, fro instance, when we observe a star and still call it an object. Well, from the viewpoint of our theory, we always associate with an object some kind of stability, and stability exist only with respect to action. In the case of a star, this is the stability with respect to varying conditions of observation. We can observe 'the same' star at different times and factor out the differences in time by taking into account the rotation of the sky around the Earth's axis. The same is true with respect to the movement of the observer around the Earth's surface. The more we know of astronomy and physics, the greater number of properties of the object will we discover, such as the constancy of the star's spectrum etc.

We also must include into the concept of cognitive actions the more sophisticated and esoteric actions which were not among those actions for which evolution created human brain, but emerge as a result of the development of science. We get involved in this kind of actions when we construct huge accelerators of elementary particles and set up experiments to explore how the particles interact. As an apple and other physical bodies are invariants in the processing of input information by the brain, so an electron and other elementary particles are invariants of the scientific symbolic models of the world. We can measure the charge of the electron in many different ways – which all are various cognitive actions – but after making all the computations required by the theory, we still come to the same number (within the error). The same with mass, spin, etc. So an electron is, for us, an object, as real as an apple. One could qualify this statement by noticing that the existence of electrons depends on the legitimacy of our physical theory, which is not absolute. True enough. But who are we to claim that the legitimacy of our brain as a collection of models is absolute?

6.3.2 Hierarchical modeling

We defined an object as a family of transformations $f_a(r)$ on the set of representations which predicts the representation resulting from a given cognitive action a. This is the same definition as the general definition of a model, where f_a is the modeling (prediction) function M_a. The specificity of an object is, first, in the domain of the index a, which is a set of cognitive actions A_{cogn}; and second, in the way the functions f_a of the family are further used.

The domain A_{cogn} includes only actions we deem external to the intuitively understood *essence* of the object. We call them cognitive because they allow us to separate the object from other phenomena, to see it from different sides – often literally – an at the same time not to change the object itself beyond recognition.

As for the use of f_a, it serves not to provide a needed prediction, but only to confirm, by checking the prediction against reality, that we do deal with a given object. Then on the basis of this information we, probably, will make a prediction which is needed as such. We see here a hierarchy of two models: a model that, having primary sensory data as input, recognizes an object, and a model which uses this object as input. In order to use a model as an object, it must be *objectified*, represented by a material object. This act of representation is of the same nature, and plays the same role as representation in models. In this way a hierarchy of models is constructed, where each next level is a representation of the transformation of the representations of the preceding level.

A hierarchical model of this kind is shown in Fig.2.
It consists of the ground level to be modeled (the world), and three levels of representations. This scheme is constructed by combining several simple modeling schemes shown in Fig.1. On the ground level we see, as in Fig.1, the states w_1

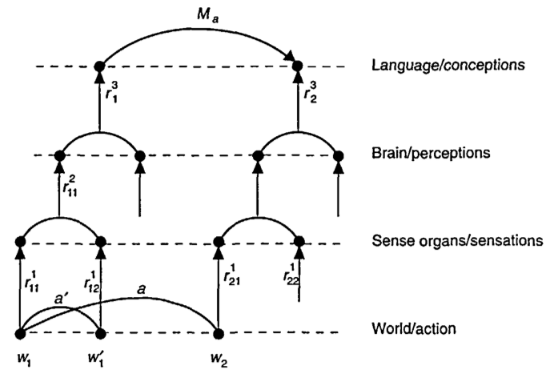

Figure 2: Hierarchical modeling scheme

and w_2; and the ultimate goal of the hierarchy is still to predict some features of w_2 on the basis of the initial state w_1 and the action a. In Fig.1 these features were expressed in terms of representation of the first level r_2; in the three-level scheme they are expressed in terms of representation of the third level r_2^3.

Now let us see how the goal of the model is achieved. The representation of w_1 is $r_1^1 1$; this might be, e.g., the image of a tea-pot on the retina of my eye. But I perceive this image as a part of *the object* tea-pot. The entire primary modeling scheme is reproduced here: it consists of the states of the world w_1 and w_1', the cognitive action a', and the representations r_{11}^1 and r_{12}^1. Here w_1' is an intermediate state of the world resulting from the cognitive action a' (e.g. that of my coming closer to the tea-pot). The transformation of r_{11}^1 into r_{12}^1 is part of my perception of the tea-pot when I make allowance for a'. The whole scheme of the perception of a tea-pot may be executed any number of times, depending on the time scale of the processes involved; in particular, it is possible that there is no time to check the model property for r_{12}^1, so that the tea-pot will be perceived on the basis of r_{11}^1, which is just a glimpse of it. In any case the transformation of r_{11}^1 into r_{12}^1 is objectified as a representation r_{11}^2 on the second level of representations.

In a similar manner the other primary models work, which are shown partially in Fig.2. A transformation of r_{11}^2 becomes represented by r_1^3, and the edifice of transformations and representations on the predicted side of the model is constructed analogously. The height of the hierarchy in Fig.2 is three, so on the third level we see the desired prediction M_a: if r_1^3 and the action a is taken, then r_2^3.

It is very important that representations of the i-th level are abstractions from *transformations* of representations of the $i-1$-st level, not from *representations themselves*. Indeed, if the representations of higher levels were abstractions of representations of lower levels, then all lower (intermediary) levels of representation would have been unnecessary. Let the representation function of the i-th level be $R_i(r_{i-1})$. Then we could combine the functions of each two neighboring levels into one:

$$R_i(r_{i-2}) = R_i(R_{i-1}(r_{i-2}))$$

Thus a hierarchy of representation levels in this case could have been only a question of expediency; it would not really give any additional power. Repeated abstraction leads to loss of contents, and ultimately to the concept 'something', about which we can say nothing. With the hierarchical modeling as we describe it, the creation of each new level is a *metasystem transition* (see [4]); it provides new possibilities and can be repeated indefinitely.

6.3.3 Sensations, perceptions, conceptions

We hypothesize that the hierarchical modeling scheme, as described above, is actually implemented in human brain and produces our many perceptions of objects in the world, where objects may be understood in the widest sense as islands of stability in the see of sensations. We do not know how many levels of this hierarchy is there, and we speak of a hierarchy in a very loose sense. Some models may use models of various lower levels and be used by more than one model of higher levels.

We intend the three-level modeling scheme in Fig.2 as a rough picture of human knowledge. The first representation level is constituted by sensations. All levels of the brain hierarchy we compress in one level: that of perceptions. The third level is that of human language and the models we construct in language on the basis of perceptions. Here again we merge many actual levels into one.

Sensations are produced by our sense organs. Perceptions are formed and used within the brain. Conceptions are created by ourselves while we create new, linguistic, models of the world. The triad *sensation, perception, conception* seems close in meaning to Kant's usage of these terms. We leave it to the reader, though, to judge on it.

7 Space and time

Among the most elementary actions known to us are small displacements "in space". We have put it in the quotes, because people have accustomed to imagine that some entity, called "space" exists as a primary reality, which creates the possibility of moving from one point of this space to another. Our analysis turns this notion topsy-turvy. Only actions constitute observable reality; space is nothing but a product of our imagination which we construct from small displacements, or shifts, of even smaller objects called points. If x is such a shift, then xx – the action x repeated twice – is a double shift, which we would call in our conventional wisdom a shift at the double distance in the same direction. On the other hand, we may want to represent a shift x as the result of another shift x' repeated twice: $x = x'x'$. It so happens that we can make three different kinds of shifts, call them x, y, z, none of which can be reduced to a combination of the other two. At the same time any shift w can be reduced to a properly chosen combination of shifts x, y, z. So we say that our space has three dimensions.

When we do nothing for a while we still feel that something has happened: we say that some "time" has passed. In terms of actions, doing nothing is a special type of action. If we denote it by t, then tt is an action of waiting for two times longer than with t.

We often say that all real processes take place in space and time. The meaning of such statements is that in addition to what really goes on, we imagine some reference actions of consecutive shifts ("in space") and waits ("in time") and establish relationships between these actions and actual objects and processes. Thus, in accordance with Kant's view, space and time are not observable realities, but our ways to organize experience.

When we think about space, we treat those shifts which create the space as instantaneous. In fact, however, all actions have a time component. Space, as we understand it intuitively, is a complete abstraction from time. Respectively, our intuitive time is abstracted form space and from everything that happens in space. This intuition, when formalized into a theory, gave rise to classical mechanics. It treats time as sort of a flow which goes on uniformly in every point of space.

If the ultimate reality is that of actions, then there must be as many times as there are agents. When an agent acts, the time that elapses is just one of the characteristics of the action. Thus whatever is happening, there must be its own time measure for it. This truth was dramatically demonstrated by Einstein's relativity theory.

When we measure time, we take some repetitious process, like the swinging of a pendulum, for a model of other processes. We may say, for instance, that John needs 80 'pendulums' of time to smoke up a cigarette. In terms of the modeling scheme (Fig.1), the state when John is lighting his cigarette is w_1; the state when he extinguishes it is w_2; the representation function R is registration of the current value of the counter; r_1 and r_2 are the states of the counter at the beginning and end of smoking.

8 Objectification

We often want to think and speak about an action or process as something definite, constant – in other words, as an object. Then we *objectify* it, i.e. replace the process, in reality or in our imagination, by an object. Objectification is a kind of metasystem transition: a process becomes an object to be manipulated by a metasystem. Strictly speaking, we should say 'becomes represented by an object', but 'becomes an object' is also admissible, because as a rule the representation would be such as to allow to reproduce the process, maybe even with variations. Hence a partial equivalence between a process and its objectification.

One common case of objectification is replacement of a process by its definition. For instance, we define algorithms as computational processes which we expect to be executed in a certain fixed manner. The definition of an algorithmic process is an object, usually a text in some formal language. The semantics, i.e. the meaning, of the language is provided by a machine which executes the process in accordance with its definition. The famous Turing machine is an example.

The mapping of an action onto its representation, which is made in the use of a model, is also a form of objectification: representations are objects.

8.1 Historical record

Consider a model M in which certain states of the world w_1 and w_2 are represented, respectively, by r_1 and r_2. Recall that states of (a part of) the world are, in fact, *cognitive actions* used by M to create representations of the world. Since w_1 and w_2 are actions, we can ask ourselves: how will M represent the composite action w_1w_2? Typically, each model would have a certain set, a tool-bag of "atomic" actions to register the states of the world. Therefore, to represent the sequential composition of those, a method must either exist in M, or to be built as an addition to M. Depending on the physical nature of the representations r_i, various methods may be used; for example, if r_i are symbols we can simply concatenate them: r_1r_2; then a sequence of state-registering actions $w_1, w_2, \ldots, w_n$ will be represented as $r_1, r_2, \ldots, r_n$. No matter what the actual method of combining representations is, we can think of the result as a sequence of representations. We shall refer to it as a *historical record.*

8.2 Memory

The subsystem of the subject of knowledge which keeps historical records is its *memory.*

Historical records add nothing to our formal concepts of model and knowledge. States are always actions, whether they are split into sequences or not. It is a technical detail, after all. I can consider every state of the world in my knowledge as a sequence of actions $w_1, w_2, \ldots, w_n$ which covers all my life from the moment I was born. Then my predictions will always be functions of the representation $r_1, r_2, \ldots, r_n$. This, of course, does not exclude predictions which take into account only the latest member r_n of the sequence. If the system has memory, then memory becomes a component of each current representation; its role in predicting, though, may vary in wide limits.

8.3 Present, Past, Future

The last atomic representation at the moment of making a prediction is referred to as *the present.* The whole sequence of memorized representation, the total historic record, is *the past.* The representation resulting from prediction is *the future.*

It is clear from this definition that the duration of time which is included in the present depends on the cognitive actions which determine the state of the world. In the context of writing a diary, the current day is the present. In the work of a computer, one millionth of a second may mean a long past state of affairs.

8.4 Real time vs. model time

Henri Bergson was first to notice and emphasize the difference between *real time*, in which we live and act, and the objectified time of history and physics. Imagine a pendulum which at each swing puts a mark on a moving tape. We have a historical record of 'how the time is moving'. This historic record is an object at every moment we look at it. We use it as a model of reality. We shall refer to the marks on the tape as representing a *model time*. It is very much different from the real time.

Real time is such that two moments of it never coexist. In model time the moments coexist as different objects in some space. Thus Bergson calls model time a projection of real time on space. Bergson's real time is irreversible. Model time is reversible: we read historical records equally well from left to right and right to left. The seemingly inconceivable feature of Feynman's diagrams, the movement of a particle in the direction opposite to time, is explained simply by the fact that the time of physical theories is model time, i.e. a spatial phenomenon.

The solution of the fundamental metaphysical problem is the key to the understanding of time. As long as you see the world as abiding, at each moment of time, in a certain state, your time will be reversible, because there always remains the possibility of returning to one of the passed states. You may invent some tricks to prevent such things from happening in your theory, but a theory which requires patches at the very base is seriously deficient.

If, on the contrary, the only reality of the world is action, then time is irreversible from the beginning, because a new action does not revoke a previous action, as a new state does, but builds on it. If in the sequence of two alternating states $w_1 w_2 w_1 w_2 w_1 w_2 \ldots$ etc. w_1 and w_2 are understood, literally, as states of the world, then each time a state returns, the world is the same, as if the time were running on and back. If the states are cognitive actions, the reality is continuously changing:

$$w_1,\ w_1 w_2,\ w_1 w_2 w_1,\ w_1 w_2 w_1 w_2,\ w_1 w_2 w_1 w_2 w_1 \ldots$$

There is no return. Action is cumulative.

In the contemporary science, real time shows up in probability theory and its applications, such as statistical physics. Probability is a characterization of certain actions, not states. In the theory of probability we deal with acts of choosing of a number of possibilities. If there are ten possible actions then the probability of actually doing one of them is 10_{-1}. If in the second step there is again the same number of options, the probability of the composition is 10_{-2}. Suppose that the first step was from state w_1 to w_2, and the second step takes the system back to w_1. The probability of each choice in one more step will be 10_{-3}, not 10_{-1}. In the next step it will be 10_{-4} etc. It is cumulative; there is no way back. This is why the time of statistical physics is irreversible.

In the world where the ultimate reality is action, time differs very much from the time of classical mechanics. The former is irreversible, the latter reversible: we can designate this difference as a difference in macro-structure. There is even a greater difference in micro-structure. Our real time is closer to Bergson's *duration* than to mechanical infinitely divisible time. Unlike the mechanical (model) time, real time is quantized: a quantum of time is the time involved in one action. An elementary action has a duration within which it makes no sense to speak about distinct moments of time. Real time does not consist of a continuum of infinitely small time coordinates. To make small times real, corresponding actions must exist. As it happens, this is not quite simple. We know from physics that there is no universal quantum of time; time quanta depend on actions. To perform an action with the duration t (by the order of magnitude), we must involve an amount of energy h/t, so the expression 'a point in time', which presumes a duration of $t = 0$, presumes an infinite energy. If it has any meaning, then, possibly, only for the state of the world just before the big bang.

9 Individual physical body

We saw that an object is more than just a complex of sensations: it includes a certain measure of stability with regards of our cognitive actions. When we speak of a physical body we have in mind more than just an object. We include the object's *historical record*. Hence 'a cat' is an abstraction applicable to any cat. The pet cat I may have at home is an on-going historic record of my having it.

The historic record of a physical body may not be actually known in full, not even partially. The important thing is that to speak of a physical body we must associate some, may be abstract, historical record with our direct perception of it. This record makes physical bodies, in principle, *identifiable*. Thus we consider the terms 'identifiable' or 'individual' body or object as synonymous with 'physical' body.

One of the most beautiful features of quantum mechanics is the way a philosophical analysis is translated into theory and experimentally confirmed. We can speak of individual physical bodies only if we can actually identify them by some means. We cannot do that to elementary particles when there are no impenetrable boundaries between them. Therefore they must be considered indistinguishable. This requires symmetrization of the wave function, which leads to various physical consequences brilliantly confirmed by experiment. Electrons are not physical bodies in the usual sense, because they are not individually identifiable bodies.

Acknowledgement It is my pleasant duty to thank my colleagues in the Principia Cybernetica Project Francis Heylighen and Cliff Joslyn for intensive discussions of various aspects of the project, including those addressed in this paper.

References

[1] Turchin, V.F. Cybernetics and Philosophy, in: *The Cybernetics of Complex Systems: Self-organization, Evolution and Social Change*, Felix Geyer, editor. Intersystem Publications, Salinas, California, 1991, pp.61-74 .

[2] Russell, Bertrand, *Human Knowledge: Its Scope and Limits*, Simon and Schuster, New York, 1948, p.269 etc.

[3] Aune, Bruce, Action and Ontology, *Philosophical Studies*, vol. 54, 1988, pp. 195-213.

[4] Turchin, V.F. *The Phenomenon of Science*, Columbia University Press, New York, 1977.