Computability in Developing Systems

Pavel B. Ivanov

Troitsk Institute for Innovation and Fusion Research (TRINITI)[†]
E-mail: *unism@narod.ru*

Written: 27 June 1996
Revised: 19 December 1996

Abstract

Human activity and cognition cannot be modelled without accounting for development. The principal directions of incorporating development in production systems, abstract automata and practical problem solvers are discussed in this paper. A few generalisations of a "physical" definition of Turing machine are described, including second-order and hierarchical models. Considering the modification of the machine's operation set and environment requires a new mathematics, which would incorporate self-modifying (reflective) axiomatic systems. Schemes of explicit and implicit definitions are discussed, and the necessity of a conceptual closure of hierarchical development in science is stressed. Universal features of development are described, distinguishing structural, systemic and hierarchical development. These features are indispensable to insure human-like behaviour in ensembles of Turing machines.

1. Introduction

Science of the XX century has been marked with intense interest to the problem of including the observer into the description of objective phenomena. It was argued that, in some cases, the act of observation would interfere with the process observed and the correct description of reality should account for such perturbations. The boundary between natural phenomena and their scientific description became vague, and this seemed to essentially limit human cognition. The famous Gödel theorem added to the idea of inherent logical insufficiency of the human mind, inspiring numerous discussions and interpretations. Some theoreticians still apply to incompleteness of formal systems in their metaphysical discourse (Penrose 1994, Stapp 1995). However, most scientists do not share these fears and continue their work fully convinced that science is able to find out everything people need in their life and activity. They understand that scientific methods cannot be reduced to formal systems, and even the development of mathematics follows the laws that have little in common with mechanical deduction.

Still, the ways science describes the world could be made the subject of a particular investigation, and nothing prohibits the formation of new sciences studying it. Actually, every science is aware of the limits of its applicability; moreover, the attempts to more exactly specify the applicability conditions often become a source of new discoveries. Such methodological study may be more obvious in sciences whose subject is originally related to the description of the observer. Thus, mathematicians intensely explore the foundations of mathematics, this activity being sometimes called metamathematics (Engeler 1983). The analogous branch in cybernetics is called "second-order" cybernetics (Foerster 1981). Also, many psychological works include an analysis of possible observer's involvement.

[†] The author is no longer affiliate to TRINITI

Mathematical theory of computability is a typical example of "second-order" approach. Generally, the subject of the theory is the operator's ability to achieve some goal that can be specified in a formal way, which assumes the description of both operation and goal specification. However, traditional models usually deal with formal systems with a predefined operation set, so that their behaviour basically remains the same. This does not exactly match our conceptions of human creativity, which imply the possibility of inventing something "new", not contained in the raw materials or instrumental skills. The most prominent feature of human activity is that it may change any rules, adopting them only temporarily, at will. This leads to the necessity of considering development, which may drastically change the very notion of computability. As the operation set expands, new goals become attainable, and a higher-order theory has to describe the rules governing such expansion.

The description of development may require new formal methods, since traditional logic demands the identity of the subject in any discourse, while development seems to repeatedly violate this identity, so that the basic axioms may become updated in the course of deduction. This may sound absurd — still, this holds for the development of mathematics too, which necessarily combines formal and informal elements.

To construct the models of developing problem solvers, one has to specify the properties of such "machines" related to development. These general definitions are then to be implemented in the formal descriptions. Thus, different kinds of development can be distinguished, including structural, systemic and hierarchical levels. It can be argued that hierarchical organisation is necessary for efficient operation in complex environment (Efimov 1982), and one comes to formulating the general principles of hierarchy. One of the most important features of hierarchical development is the interiorisation of any external activity through socialisation, as well as actualisation of internal processes, creativity.

In this article, I assess the current state of computability studies from the developmental viewpoint. I do not intend to derive any formal results, but rather indicate the search directions and explicate the ideas commonly implied. Section 2 summarises the traditional approach to computability on the example of the Turing machine, in a "physical" formulation more suited for further generalisation. Some directions of generalisation are discussed in Section 3, including multidimensional, non-local and analogue Turing machines. Second-order Turing machines are suggested as a more realistic model of human cognition. Developing Turing machines are considered in that section, with the stress on evolving environment, productive activity and abstraction. Formal aspects of development are discussed in Section 4, demonstrating the interlay of explicit and implicit definitions in both developing problem solvers and their mathematical descriptions. Implicit definitions are of primary importance for development, and I suggest several universal schemes applicable in developing machines. In conclusion, Section 5 specifies the general features of development to be implemented in the future models of human activity and cognition.

2. Traditional computability

Historically, computability problem assumed many forms. The well known formulations include Turing machines, RAM machines, recursive functions, λ-calculus, Post systems, Markov algorithms and others (Cutland 1980). The basic fact is that the class of "computable" elements is the same in all these theories. The well-known Church thesis postulates that any other theory of discrete computations would give the same computable class. This seems quite natural, since, in any case, one deals with finite sequences of operations from a numerable variety. The details of enumeration are of minor importance for computability.

Adopting Church thesis, one may consider any model of discrete calculation as a representative of the same general category. I choose the Turing machine representation, which stresses the idea of operation directly associated with certain aspects of human activity. To enhance the analogy, I will define a Turing machine in a "physical" manner, different from the mathematical formulations proper:

(1) The "universe" $U = W + M$.
(2) The "world" W:
     (a) a linearly ordered discrete set of possible locations $X = \{x\}$ ("configuration space");
     (b) a discrete set of possible states $S = \{s\}$ observable at any location;
(3) The "machine" M:
     (a) a discrete set of possible "internal" states C;
     (b) a finite set of possible "operations" R, represented by a relation $R: X \times S \times C \rightarrow X \times S \times C$;
(4) Current state of the universe U:
     (a) function $s: X \rightarrow S$, the overall state of the world W as a field of local states $s(x)$;
     (b) $x \in X$, the location of M;
     (c) $c \in C$, the state of M;
(5) Operation cycle:
     (a) the machine's input: $< x, s(x), c >$;
     (b) the choice of operation $r \in R$: $< x, s, c > \rightarrow < x', s', c' >$;
     (c) the machine's action: $x \rightarrow x' = x + dx$, $s(x) \rightarrow s'(x) = s(x) + ds$, $c \rightarrow c' = c + dc$;
     (d) the new input $< x', s'(x'), c' >$ initiates the next reaction, and so on.

The machine's operation is *local*: it may only change the state s of the world W in the point x where M is currently located. In the classical definition, Turing machine may move to the neighbouring locations only, so that $x' =$ either x, or $x+1$, or $x-1$ (that is, $dx = 0, +1, -1$). Evidently, the operation set R may be thought of as a set of corteges $r = < dx, ds, dc >$.

Each operation cycle may be associated with the time increment $dt = 1$, so that the *kinematics* of U can be described by three functions $x(t)$, $c(t)$ and $s(x,t)$ representing the external movement of M, the change of its internal state and the modification of the overall state of W. Virtually, one might consider the sequence of operations $r(t)$ (the "computation process"), and the *dynamics* of Turing machine could then be described by the equation set

$$dx/dt = r_x(x,s,c; t)$$
$$ds/dt = r_s(x,s,c; t)$$
$$dc/dt = r_c(x,s,c; t)$$

At any moment, dynamics implies choice between the currently admissible operations ("decision"). This choice may depend on the current state of U in a more or less definite way. Thus, one can consider a *stochastic* Turing machine, with every state $< x, s, c >$ assuming a number of admissible operations $< dx, ds, dc >$, which may be chosen with definite probabilities. In the opposite case, the behaviour of the machine may be entirely *causal*, so that the change of the state would be completely defined by the history of previous interactions.

In this model, *computability* means the existence of a process $r(t)$ transforming an initial state of the world $s_i(x)$ into a final state $s_f(x)$. In the process of computation, M would pass from the state $< x_i, c_i >$ to the state $< x_f, c_f >$, with c_f occasionally supposed to terminate (or suspend) the machine's operation (Mendelson 1964; Gorbatov 1986). The latter requirement agrees with the typical usage of a computer, while modelling human-like behaviour does not need such a restriction. Human activity never stops, and any particular state is transitory in it. Hence, computability is reduced to the

existence of a trajectory passing through two pre-defined points. Such trajectory may be not unique; for stochastic Turing machines, one may associate every allowed trajectory with a non-zero probability (or a probability amplitude), whereas the selection of a definite trajectory (method of computation) in causal machines depends on the initial conditions.

There are two major trends in computability studies. One of them is primarily concerned with the kinematics of operation, investigating the possibility of constructing an operation set enough to transform one state of the world W into another. However, practical applications require *efficient procedures*, rather than proofs of principal solvability (Chang&Lee 1973); this leads to a closer investigation of dynamics. Accordingly, one might distinguish two kinds of uncomputability: (a) kinematic uncomputability due to the essential incompleteness of the operation set, and (b) dynamic uncomputability due to the violation of time restrictions imposed. However, the traditional theory of computability does not consider development, since the sets X, S, C and R always remain the same in the course of computation.

3. Generalised Turing machines

The Turing-machine model of discrete computation can be generalised in many ways. Some of them may lead to new classes of computable functions.

The first evident generalisation is multidimensionality: locations x may be vectors, rather than single numbers. Accordingly, the change in the machine's location would mean that each coordinate is either increased by one, or diminished by one, or left the same. It is rather evident that such generalisation would not change the computable class, as long as spatial points remain numerable.

Also, one can remove the locality restriction, and consider machines which may:

    (a) observe the state of W in the points other than the current location of M;
    (b) change the state of W in several points, rather than a single point;
    (c) move to a spatial point not adjacent to the current location.

Again, this would not affect computability, as long as the operation set R remains finite. The allowance for infinite operation sets would lead to a wider class of computable functions, since such infinity is implicitly reflective, as will be discussed in the following sections. Infinite operation sets may provide a structural (static) description of development; however, such description is insufficient for higher-order research.

One more generalisation that may extend the computability class is analogue Turing machines in continuous space-time. Continuum requires a different concept of computability, replacing discrete sequences of events with smooth trajectories. Depending on the model of system dynamics, they will be either usual trajectories in the configuration space X accompanied with some "internal" movement in M, or evolution of virtual intermediate states in the stochastic (quantum) case, or any combination of these mechanisms. As the time increment dt tends to zero, kinematic computability will be associated with the trajectories of finite length, rather than with finite sequences of acts. As for the ordinary Turing machines, dynamic computability differs from kinematic computability, demanding that the trajectory connecting two spatial points should be passed in a reasonable time; moreover, in the continuous case, there can be *absolute* dynamic uncomputability when a finite-length trajectory requires infinite time to pass due to the singularity of machine's velocity distribution along the trajectory depending on the external force fields. Analogue Turing machines may provide more realistic models of human reasoning, eliminating many restrictions of discrete models (Mulhauser

1994). Still, the behaviour of analogue Turing machines remains, in a sense, predefined. The focus just shifts from discrete orbits in a set to the topology of a manifold.

One could consider Turing machines in an evolving environment, so that $s(x,t)$ would change in time due to some external process beside the machine's operation. The equations of motion would then be rewritten as

$$dx/dt = r_x(x,s,c;\ t)$$
$$ds/dt = r_s(x,s,c;\ t) + f_s(x,s;\ t)$$
$$dc/dt = r_c(x,s,c;\ t) + f_c(c;\ t)$$

The last of these equations accounts for the spontaneous evolution of the internal state of M. Such time-dependent formalism requires a revision of the very notion of computability, since the initial and final states of the world W are no more stationary. Thus, one might consider uniform computability, adiabatic computability, computability in average etc. Still, passive evolution is a rather restricted sort of development lacking essentially human creativity.

Within the Turing-machine model, development would affect either of:

(a) configuration space X;
(b) the set of possible states S;
(c) the set of internal states C;
(d) operation set R.

The possibility of a change in X and S is most important: it means that M can produce something new that had never existed in the world W before. This is one of the distinctive features of specifically human behaviour, and any other changes can be argued to originate from that "material" productivity. People do not merely rearrange the world — they replace it with another one, more suited for their life and activity. The change in C and R means that M is apt to internal reorganisation, which would make it another machine, from the traditional viewpoint. In human behaviour, such internal development may be relatively independent of external productivity.

Development drastically changes the notion of computability, since *any* state of U becomes accessible, as soon as a special operation is added to the original set. In common words, if something can be imagined, it can be done. Sooner or later, someone would find the way to achieve what has been thought of as unachievable. For a developing Turing machine, there are *no* uncomputable functions, and any proofs of unsolvability may only refer to a definite stage of development.

However, the development of Turing machines should not be considered as arbitrary. The very ability of breaking the rules may be governed by some "second-order" rules. To specify them, one might follow the traditional line of mathematical development, specifying minimal generalisation of the traditional Turing model that would lead to a kind of self-modification.

There is an evident possibility which implies one additional operation: if a state of U has proved to be inaccessible from some initial state using the current set of operations R, the corresponding transition $r = <dx, ds, dc>$ may be added to R. However, this generalisation may be not minimal when movement obeys the "smoothness" restriction: $dx = +1, 0, -1$ (with dc linking the "adjacent" states of M only). In this case, one will have to modify the "expansion rule": if an uncomputable state is adjacent to a computable state, the transition to the former from the latter may be added to the operation set. In general, consideration of such "minimality" will lead to discussing the reducibility of computational tasks (Cutland 1980; Efimov 1982; Gorbatov 1986).

This mechanism is much like space completion in mathematics, when every fundamental sequence is identified with a point of the space, thus being made convergent. The key feature of such development is that *M* should be able to somehow "know" about an inaccessible point before introducing a new operation making it accessible. In other words, the internal state of the machine should reflect the machine's operation, including the formation of computational tasks. Hence, Turing machine must be *hierarchical*, the higher level forming the "goals" for the lower-level computations. One might consider such machine as a combination of two different machines: one machine is to perform "calculations", while another is to set up computational tasks (Efimov 1982). So, the principal mechanism of development is integration of functionally different machines moderating each other's operation. From the viewpoint of every one of these united machines, a part of the machine's environment is *interiorised*, forming a new level of operation.

The definition of such "second-order" Turing machine includes:

(1)   The world $W = <X, S, s: X \to S>$,
(2)   Level-1 machine *M1* operating on *W*,
(3)   Level-2 machine *M2* operating on the "internal world" of *M1*: $W1 = <S, C, R>$.

The equations of motion for the two-level machine could be written as

$$dx/dt = r_{1x}(x,s,c; t)$$
$$ds/dt = r_{1s}(x,s,c; t)$$
$$dc/dt = r_{1c}(x,s,c; t)$$
$$dr_{1x}/dt = r_{2x}(s,c,r_1; t)$$
$$dr_{1s}/dt = r_{2s}(s,c,r_1; t)$$
$$dr_{1c}/dt = r_{2c}(s,c,r_1; t)$$

Of course, the structure of the equations may be more complex, the last three equations being functional rather than ordinary differential equations. The internal state of *M1* may include the kinematics of *M2*, the equation of motion being essentially nonlinear (recursive). Higher-order Turing machines can be constructed in the same way.

The machine's ability to change the external world and itself is related to the process of *abstraction*. The traditional definition of Turing machine implies explicit enumeration of operations from *R*, which is inappropriate for infinite sets *X*, *S* and *C*. Thus, to specify increment by one as an elementary operation on $S = \{0,1,2,...\}$, one has to make *R* infinite; however, one could introduce the *abstract* operation INC instead, transforming every $s \in S$ into $s+1$, which would mean only one operation instead of the infinity of similar acts. Evidently, abstraction is mostly due to machine's "self-observation" (reflectivity), a scheme derived from a number of special cases (Pospelov 1986). Since abstract operations can be generated from any finite number of samples, there may be cases when the result of operation is undefined, because there are no respective elements in *X*, *S* or *C*. In human behaviour, there are two basic reactions to such a situation: the formal result may be treated as *ideal*, and added to the collection of internal states *C*, or it may be *implemented* in external activity producing new elements of *S* or *X* (*intensive* or *extensive* development of the world). In particular, external activity may become mediated by a series of internal transformations.

4. Development in formal systems

Formal description of developing Turing machines may require new mathematical methods, since the allowance for self-modification leads to apparent contradictions violating the identity of the subject, the axioms being updated during the discourse (Pospelov 1986). Seeking for more rigor, one might

suggest that mathematics do not deal with development at all, and that mathematical models refer to a single level of development and work inside it. Any axiom modification will be hence considered as switching to another model that should be treated separately. However, this attitude does not help much since the very specification of the working level implies distinguishing it from the other levels, while considering any other levels would inevitably violate the "purity" of theory. For instance, the mathematical idea of an operation assumes its complete specification: one should exactly define how the operation is performed, and what would be its result for any given "arguments". This cannot be done in real life, where operations are *never* defined completely. As a rule, the introduction of an operation involves some informal descriptions, which complement the explicit and implicit definitions.

Explicit definition constructs a new operation as a combination of already existing operations, the rules of composition forming a higher level of operation. Several levels may be specified, but the highest level can never be introduced explicitly. Some mathematical theories merge these levels into a single "universe" with rather exotic properties (higher-order logic, Henkin 1950; nonstandard analysis, Davis 1977); hierarchical representation might make the problems more tractable.

A generalisation of the notion of Turing machine may include the rules of explicit definition in the operation set, which can thus be made extendible. However, explicit construction of operations does not give much freedom for development, since one might consider complete (closed) operation sets only, where any composition of operations gives an operation from the same set. However, the transitory processes preceding the formation of such complete sets may be of interest under special circumstances, when the rate of the machine's operation is much higher than the rate of completing the operation set.

In the implicit definition, a new operation is defined "by analogy", using the constructions like: "Act in the same manner as when transforming x into y." The actual ways of performance stay outside the theory, admitted they could be somehow discovered in practice; in other words, implicit definitions are to be further explicitly reduced to some other primary operations, which will have to be unfolded in their turn. Logically, an implicit definition of operation is a tautology: operation g transforming input x_{in} into output x_{out} is just the ordered pair $< x_{in}, x_{out} >$. Any set of such pairs may be called a "generic" operation, relation $g: X \rightarrow X$. However, this approach is not applicable to infinite sets, or analogue machines, and the traditional theory of computability generally sticks to finite alphabets and operation sets. Still, implicit definitions model an important feature of real behaviour: it is quite common in animals that reflex formation occurs via a momentary circuit closure, in a single try; even if thus acquired behaviour proves to be inadequate, animals do not completely dismiss the first impression, other reflexes just moderating its action.

Implicit definitions are most popular in mathematics, though they often lead to methodological problems. The definition of a set by its characteristic property is a typical example: "Let A be a set of all a that possess the property p." A cortege of n elements referred to as $< x_1, x_2, \ldots, x_n >$ is one more instance of implicit definition. Such constructions violate the traditional deductive scheme, but mathematics can never get rid of them. Inductive definitions are employed for more rigor, but the induction rule itself contains a logical loop, since the presumed n-th step of induction is an implicit reference too. So, to proceed with explicit constructions, one has to postulate some notions and rules appealing to "intuitive" abilities presumably shared by other people. Therefore, mathematical reasoning does not produce any "truths", and mathematical deduction is a method of elaborating a hypothesis rather than validating it.

Generally, with the allowance for implicit definitions, the process of completing the operation set via explicit combination of primary operations cannot be considered as transient, since the addition of a new implicit definition (or a new abstraction) will start the process of accumulating explicit constructions employing it. This process may be not finished to the moment when another primary operation appears, the state of the machine thus becoming always "transitory". Of course, one may study *adiabatic* abstraction, when the operation set becomes complete before every new implicit definition. Such smooth development would model some real cases, though it cannot cope with the revolutionary changes in human knowledge.

Though development through explicit construction is more important in the machines able of implicit definition, it still remains just quantitative, lacking "true" novelty. On the other side, the arbitrariness of implicit definition can only be tolerated to a definite extent, and one would necessarily try to investigate the mechanisms of implicit definition too. There is a danger of an infinite succession of the "levels of implication", when every newly discovered mechanism initiates the search for "deeper" mechanisms underlying it. Modern physics may be an example, with molecules consisting of atoms, atoms constructed of nuclei and electrons, nuclei containing hadrons and mesons, which become reduced to quarks and gluons, and some physicists suggest that quarks might be "made" of still "simpler" particles (or fields).

To avoid this "bad" infinity one might to arrange a "feed-back" from explicit to implicit definitions, so that accumulation of new combinations would result in a qualitative "jump" to a new kind of operation. This effect is quite common in human activity. Usually, single act is not enough to become aware of it as a representative of a new mode of action. It is its reproduction in different circumstances that would stress the universality. What looks like "instantaneous" solution, "insight", is rather people's communication folded in mental activity.

In general, implicit definitions admit many possible explications, since the transition from x to y can be performed in many different ways. This fact is quite obvious with mathematical functions defined as the sets of pairs $< x, f(x) >$: there are many functions that assume value y in the point x. In physics, such situations often arise in quantum theory, when a Hermitian operator is to be continued from the "on-shell" to the "off-shell" (unphysical) energies. As a rule, the selection of the "true" branch requires considering asymptotic conditions imposed by the macroscopic measurement procedure.

An implicit definition can be treated as a void position in a formal scheme. Thus, if there is an n-place formula assumed to be valid in all the cases under consideration, every one of n positions can be empty, which would define the entity completing the formula when substituted to the void position. For example, the dyad (a,b) assumes two forms of implicit definition, $(?,b)$ and $(a,?)$; the triad (a,b,c) leads to implicit definitions of the form $(?,b,c)$, $(a,?,c)$ and $(a,b,?)$, etc. In a sense, implicit definition is opposite to the statement of existence: instead of "There is a such that (a,b)", one says: "Let a be such that (a,b)". The well-known examples come from the school mathematics: $-x$ is such number that $x+(-x) = 0$, i is such number that $i^2 = -1$, etc. More academic definitions of these entities just explicate the implicit definitions introducing more implications elsewhere. However, the void-position form of implicit definition is transparent and convenient for implementation in artificial problem solvers.

Of course, the above schemes do not exhaust the possible forms of implicit definition. Two more classes of importance are folding and unfolding schemes. Thus, a new entity can be introduced as a "shortcut" for a n-place formula: for example, let y stand for (a,b). In programming, such folding may be illustrated by macro definitions, or inline functions. However, inline functions may become

out-of-line in some circumstances (for example, in the debugging environment), or even grow into separate program modules, as the whole project develops. The opposite case of unfolding scheme expands an element of *a* formula into a sub-expression. For example, in the triad (a,b,c) the link b can be represented as $(b(a),b',b(c))$, distinguishing the sides of b relating to its connections with a and c, as well as the way b' of linking $b(a)$ to $b(c)$ "inside" b. In particular, designation and substitution are special cases of folding/unfolding which are the basic operations of mathematical reasoning, though these operations are never explicitly defined in mathematics.

In modern mathematics, some classes of problems become substituted with alternative "linguistic" formulations, when one does not prove a theorem, but rather proves that the proof can be formulated using a specially designed formal language. Thus, the formulation of Gödel theorems by Mendelson 1964 and Cutland 1980 replaces arithmetic with one of its formal descriptions, and thus the result obtained is the incompleteness of this type of description, rather than of arithmetic itself. Another example is provided by nonstandard analysis, where all the theorems concern a formal description of the universe introduced rather than this universe proper (Davis 1977). This is a "second-order" tendency in mathematics, since the focus of investigation shifts from the observable world to the observer.

However, observer belongs to the same world, and one can notice that the transition from the "first-order" to "second-order" science may be treated as a change in the subject producing another (interdisciplinary) science, which is as "positive" as the original, "first-order" research. The "second-order" version of this new science can be built in its turn — and so on to infinity. An infinite sequence of levels can thus be unfolded starting from any science, either "positive" or not. Accordingly, there is a hierarchy of computability (solvability, verification etc.), since the formalism of any level may be found incomplete judging from the higher level only, which might be just an evidence of its own incompleteness. Inversely, the very existence of the higher level assumes that some problems cannot be resolved at the lower levels, which means that they are essentially incomplete.

From this viewpoint, it seems quite natural that, in seeking for the foundations of mathematics, any attempt to build a *rigorous* theory leads to the construction of many alternative theories, which are as acceptable, or as arbitrary. The commonly known history of non-Euclidean geometry finds its continuation in the diversity of the definitions of fractal dimension (Fractals 1985), alternative set theories (Zadeh 1973, Vopenka 1979), the categorial formulation of mathematics (Goldblatt 1979).

The infinite chain of incomplete theories can be folded in the abstraction of a new science describing the very method of generating the higher levels. Such science would be completely characterised by any two adjacent levels, thus being a *hierarchical synthesis* of them. On the other side, its distinction from the "first-order" approach would be completely defined by the "second-order" theory, while the "first-order" science would define the transition to the synthetic level from the "second-order" approach. No infinite regression occurs in this way.

An analogous triad of the levels of activity is well known in general psychology (Leontiev 1978); recently, Leontiev's theory has been revised within hierarchical approach and applied to aesthetic perception (Ivanov 1994, 1995). The conceptual cycle described would provide a kind of basic model for hierarchical mathematics, like the heating–cooling cycle of the ideal heat machine was a germ of modern thermodynamics. The logic of this article demands that the same triadic organisation should be inherent in Turing machines if they are expected to exhibit human-like behaviour.

5. Phenomenology of Development

A detailed investigation of developing Turing machines is yet to be performed. Still, some general directions of research and features to describe may be predicted beforehand. In this concluding section, I will give a brief summary of the hierarchical approach to the problem.

Changing environment

Individual development is only possible when an individual is placed in a developing environment. If the world W remains qualitatively the same, one may only speak about *transitory* development, which is bound to stop when all of the world becomes "comprehended" by the machine. The assumption of infinite W does not alter this conclusion, as far as this infinity is merely quantitative. Thus, one cannot enumerate all the integer numbers in a finite time — still, their "idea" is readily comprehended in the operation of enumeration. Actually, one does not need to further count integers after this basic operation has been formed.

Three ways of the involvement of machine's development in the development of the world can be considered:

(1) *Syncretic development.* Everything changes, and the machine changes as a part of the world.
(2) *Analytical development.* The machine changes its environment which demands for the new ways of adaptation.
(3) *Synthetic development.* Each action implies changes in both the environment and the organisation of the machine.

It may be readily seen that the first way corresponds to the level of physical existence, while the second and third levels could be associated with biological and conscious development respectively. Of course, the latter case attracts much more interest since it has to do with modelling human cognition.

Levels of coherence

As it has been noted for higher-order Turing machines, the basic mechanism of development is the synthesis of different objects into a kind of integrity. Such synthesis may proceed in different ways:

(1) *Structural development.* New elements or new relations between elements are added. Every element is principally equivalent to any other element, and all the relation are of generally the same kind.
(2) *Systemic development.* The components of the compound are functionally different. Thus, systemic synthesis of two independent structures may assume that one of them becomes "input" while the other becomes "output", the relation between them within the system differing from the uniformity of elements and relations within structure.
(3) *Hierarchical development.* This is the synthesis of structural and systemic development assuming that systemic development is reflected in the structures, and structural development leads to the distinctions in functionality. Here, the integration of several objects into the whole means that they become the different levels of hierarchy, so that the structures and behaviour at the lower levels depends on the "control parameters" prescribed by the higher levels.

Actually, development cycles through these stages, structures and systems becoming *hierarchical*. The hierarchical structure differs from the "plain" structure in that the elements become grouped into several distinct classes, so that the relations between the classes are distinguished from the relations

inside each class. Naturally, the classes may be further grouped into "higher-level" categories, forming a multilevel hierarchical structure. Hierarchical systems may be described in much the same way, their input and output structures becoming hierarchical.

Reflectivity

To insure that the synthesis of a few machines be more "powerful" than every one of the original machines, the machines to be combined should be different enough, possessing complementary sets of operations. That is, different machines should be able to compare themselves, so that any one of them could borrow lacking operations from its neighbours. In particular, a machine should be able to analyse its own operation set.

Such an ability implies advanced communication between the machines. Of course, the first premise is interaction. However, it is not enough. To serve for development, interaction must be reflexive, that is, the behaviour of one machine must influence its state through the feed-back from another machine. To make development intelligent, one more condition must be satisfied: the feed-back must involve the internal image of the machines in any one of them, and the internal image of their interaction. It should not be mere dependence on the environmental changes caused by the agent, in which case only "biological" development is possible.

Two implications are then to be considered. First, the comparison of different machines should be performed on a common basis, which is provided by cultural integrity in the case of human activity. In other words, human-like development implies the existence of a "collective" level uniting many individual machines in a kind of society. The interaction of any two individuals (and individual reflection in particular) is therefore never direct, but rather mediated by social organisation. It means that either interaction of the machine with the world should be non-local, or (which is more likely) the immediate environment of the machine is "subjectively" perceived as the topmost level of a hierarchy, any physical event being just an indication of a global process.

Furthermore, such socially mediated reflection becomes interiorised into a special operation of relating different individuals to each other. That is, the individual operation set R should contain the *operations of reflection*, along with the traditional "productive" operations and the operations of self-modification discussed above. Hence, the internal hierarchy of the machine tends to reflect the hierarchy of the world. In particular, every act of communication involves an internal image of the partner, actual behaviour being dependent on many folded communications with this internal model.

References

Chang, Ch., and Lee, R. (1973) *Symbolic Logic and Mechanical Theorem Prooving*. (New York, San Francisco, London: Academic)

Cutland, N. (1980) *Computability: An introduction to recursive function theory*. (Cambridge: Cambridge Univ.)

Davis, M. (1977) *Applied nonstandard analysis*. (New York: Wiley)

Efimov, E.I. (1982) *Intellectual Problem Solvers*. (Moscow: Nauka)

Engeler, E. (1983) *Metamathematik der Elementarmathematik*. (Berlin: Springer)

Foerster, H. von. (1981) *Observing Systems*. Selected Papers of Heinz von Foerster (Seaside, CA: Intersystems)

Fractals in Physics. Proc. of the VI Int. Symp. on Fractals in Physics (Trieste, Italy, 1985) (L.Pietronero and E.Tosatti, eds.) (Amsterdam, Oxford, New York, Toronto: North Holland, 1986)

Goldblatt, R. (1979) *Topoi: The categorial analysis of logic*. (Amsterdam, New York, Oxford: North Holland)

Gorbatov, V.A. (1986) *Foundations of Discrete Mathematics*. (Moscow: Vysshaya Shkola)

Henkin, L. (1950) *J. Symbolic Logic*, **15**, 81

Ivanov, P.B. (1994) *Leonardo*, **27**, no. 5, 417

Ivanov, P.B. (1995) *Leonardo Music Journal*, **5**, 49

Leontiev, A.N. (1978) *Activity, Consciousness and Personality*. (Englewood Cliffs, NJ: Prentice Hall)

Mendelson, E. (1964) *Introduction to mathematical logic*. (Princeton, NJ: D. van Nostrand)

Mulhauser, G.R. (1994) *Mind Out of Matter: Topics in the Physical Foundations of Consciousness and Cognition*. Online: International Philosophy Preprint Exchange, < http://cogsci.l.chiba-u.ac.jp/IPPE/Phil_of_Mind/Mulhauser.Mind_Out_of_Matter/ > (no longer available online)

Penrose, R. (1994) *Shadows of the Mind*. (New York: Oxford University)

Pospelov, D.A. (1986) *Situation-Driven Control: Theory and Practice*. (Moscow: Nauka)

Stapp, H.P. (1995) *Psyche*, **2** ; Online: < http://psyche.cs.monash.edu.au/volume2-1/psyche-95-2-05-qm_stapp-1-stapp.html >

Vopenka, P. (1979) *Mathematics in the alternative set theory*. (Leipzig: Teubner)

Zadeh, L.A. (1973) *The concept of a linguistic variable and its application to approximate reasoning*. (New York: Elsevier)

Comments

Below, find a few remarks made on this paper by the reviewers of *Cybernetics and Human Knowing*, followed by my replies.

Reviewer:

This paper is a discussion of computing and Turing machines from a somewhat "second order" perspective. I am not too excited about it. It seems more or less journalistic to me, without any really new ideas. It does discuss issues that should be discussed in this journal, and it may stimulate conversations.

Reply:

Since the journal is quite popular and concerning the philosophy of "second-order cybernetics" rather than any formal issues, it would be strange if my submission were of a different kind. The paper was not intended to be an ordinary mathematical publication, with much technicality and poor understanding of the sense of the results derived. Rather, it had to suggest an alternative to the blind manipulation with symbols, up to the necessity of employing a way of reasoning different from the traditional deductive scheme. The reviewer did not notice that in the text, since he shared the common prejudices about "science-like" discourse and could not conceive any other forms of reasoning that are as precise, and more adequate for the subject under consideration.

The end of the paragraph is a contradiction to the previous sentence: if the paper contained nothing new, how could it stimulate conversations?

Reviewer:

The author airs ideas which in general are of interest to the readership. However he is unable to simultaneously be precise and understandable, that is, he takes pains to provide (reasonable) mathematical definitions, but as soon as the conceptual terrain gets rough, he retreats to vague and airy prose. Moreover, I find serious problems in his discussion of and arguments regarding extendible Turing machines. But even if I'm mistaken here, he has a lot of explaining to do.

Reply:

Yes, the text is not readily understandable, since it is based on the notions far from the mathematical tradition, which could not be described in full in a paper subject to the usual size limitations (<5000 words). Moreover, this new approach can hardly ever be formulated completely, since self-development is one of its basic principles. There is a lot of explaining to do — but one of the ideas discussed in the paper is exactly that there can be *no* completely explicit formalism, and any mathematical paper is bound to contain something implicit in it, thus requiring a lot of explaining too. My paper is more honest than the bulk of other "scientific" works, since it does not try to hide the narrow places "under the carpet".

Indeed, I did not "take pains to provide (reasonable) mathematical definitions" — the definitions in the beginning of the paper just describe the usual notions, to fix the terms. There are no mathematical definitions in the paper, quite intentionally. The paper says that mathematical defitions are a very particular case of definition in general — so, I tried to avoid too much dependence on the traditional formalism, to stress the basic ideas. What seems "vague and airy prose" to the reviewer is actually never less precise than the usual mathematical methods, accounting for the implicitness inherent to all

the mathematical knowledge. I explicate what mathematicians try to hide behind the superficial formality.

Furthermore, what sense would there be in publications which do not leave room for further development? Why not first announce a viewpoint. and then get engaged in explanations? The demand to present a "ready-to-consume" work looks most strange, since it is the basic mechanism of science to suggest hypotheses which are to be verified later on.

Of course, I can agree that my style is not too crisp and clear — but I read many papers written in a language much more obscure and heavy, and this cannot be the true reason of the reviewer's negative attitude.

Reviewer:

Although I appreciate from my own experience how difficult this topic is, I don't feel that the author, in the final analysis, succeeds in presenting anything genuinely new. That is, anyone who has thought seriously about this topic arrives at much the same thoughts as the author, which accomplishment is not to be denigrated, but the problem is to get *further*.

Reply:

If the views I present are that obvious, why there are so few traces of them in the literature, and why there are so many quite opposite views, indicating a rather poor understanding of the problems related to development? Where are those who "thought seriously about this topic" and "arrived at much the same thoughts"? Are these thoughts forbidden to publicly discuss? I'm afraid, the reviewer's awareness of these issues was the result of reading the paper, rather than his own achievement — and his desire to get further was an intended effect. This is what I wrote the paper for.

Reviewer:

I have suggested that the middle of the paper be deleted, and the author concentrate on his opening theme. But whether this suggestion is of any use or not, the paper as it stands must be rejected.

Reply:

The suggestion to delete the parts of the text that contain the new ideas might be an indication of the reviewer's fear of them. I am advised to do nothing beyond the scope of traditional mathematics, and to never try to point to its insufficiency. But my interpretation of the traditional line seems too dangerous to the reviewer as well, and he says that the paper should not be published even if I removed any "bad" ideas from it. No further comment is needed.

http://unism.pjwb.net/arc
http://unism.pjwb.org/arc
http://unism.narod.ru/arc