The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0368-492X.htm

K
51,13

186

Received 30 January 2022
Revised 29 May 2022
Accepted 16 July 2022

Contradiction of modern and social-humanitarian artificial intelligence

Alexander Nikolaevich Raikov
Laboratory of Modular Data Processing and Control Systems, Trapeznikov Institute of Control Sciences RAS, Moscow, Russian Federation, and
Massimiliano Pirani
DII–Department of Information Engineering, Polytechnic University of Marche, Ancona, Italy

Abstract

Purpose – The purpose of the paper is to propose an effective approach of artificial intelligence (AI) addressing social-humanitarian reality comprising non-formalizable representation. The new task is to describe processes of integration of AI and humans in the hybrid systems framework.

Design/methodology/approach – Social-humanitarian dynamics contradict traditional characteristics of AI. Suggested methodology embraces formalized and non-formalized parts as a whole. Holonic and special convergent approaches are combined to ensure purposefulness and sustainability of collective decision-making. Inverse problem solving on topology spaces, control thermodynamics and non-formalizable (considering quantum and relativistic) semantics include observers of eigenforms of reality.

Findings – Collective decision-making cannot be represented only by formal means. Thus, this paper suggests the equation of hybrid reality (HyR), which integrates formalizable and non-formalizable parts conveying and coalescing holonic approaches, thermodynamic theory, cognitive modeling and inverse problem solving. The special convergent approach makes the solution of this equation purposeful and sustainable.

Research limitations/implications – The suggested approach is far reaching with respect of current state-of-the-art technology; medium-term limitations are expected in the creation of cognitive semantics.

Practical implications – Social-humanitarian events embrace all phenomena connected with individual and collective human behavior and decision-making. The paper will impact deeply networked experts, groups of crowds, rescue teams, researchers, professional communities, society and environment.

Originality/value – New possibilities for advanced AI to enable purposeful and sustainable social-humanitarian subjects. The special convergent information structuring during collective decision-making creates necessary conditions toward the goals.

Keywords Artificial intelligence, Convergent methodology, Holonic approach, Hybrid reality, Eigenform, Third-order cybernetics, Collective behavior, Constructivism, Autonomous agent, Reflexive-active environment, Embodied intelligence, Cognitive semantics

Paper type Conceptual paper

1. Introduction

Social-humanitarian objects and events embrace all the phenomena connected with individual and collective human behavior and decision-making in entities like international organizations, governmental departments, companies and other organized groups.

This work was funded in part by the Russian Science Foundation "Socio-Humanitarian Foundations of Criteria for Evaluating Innovations Using Digital Technology and Artificial Intelligence" (No: 21-18-00184) and in part by the University Scientific Research (RSA) by Marche Polytechnic University.

Kybernetes
Vol. 51 No. 13, 2022
pp. 186-198
Emerald Publishing Limited
0368-492X
DOI 10.1108/K-01-2022-0057

Humanitarian principles are taken as a basis, for example humanity, neutrality, impartiality and operational independence. The purpose of humanitarian action has the priority to protect life and to ensure respect for humans (Mackintosh, 2000).

All these aspects of social-humanitarian reality cannot be fully formalized with computer models and traditional artificial intelligence (AI), which operate with logic and neuronal networks, still weak and narrow in scope. Current AI can recognize, predict and answer easy questions. But thinking, understanding, explaining, posing and solving problems and making correct and at the same time uncaused decisions depend on observer's behavior. Notably, the dominance of observers' effect has been today rendered evident even in physical processes that do not involve direct human participation (Rovelli, 1996). In addition, human has intuition, can meditate, fall into a trance and feel conscious. These are aspects of possible future development in AI (Raikov, 2019, 2021) to help human society solving complex social problems and difficulties.

To build such an AI, it is necessary to change the paradigm of traditional AI itself by explicitly considering an observer or an agent that reflexively and cognitively influences situations. The inclusion of purposive agents in cognitive processes elicits inverse problem solving through new semantics, possibly based on quantum, wave and relativistic effects (Raikov, 2019, 2021), which informational power arrives to 30–50 orders of magnitude higher than the traditional ones. This power can be conveyed into the social-humanitarian phenomenon that is characterized by collective consciousness and non-consciousness at the same time. Any attempt of formalization of this phenomenon distorts its emotional and transcendental integrity, namely as follows:

(1) Hidden thoughts, emotions and intentions of human;

(2) Lack of similarity between the language and consciousness;

(3) Deep and non-formalizable interaction between subjects;

(4) Ability of explosive mental activity and

(5) Capturing consciousness through the behavior of neighbors, etc.

The task of this paper is then to provide a position for a framework that aims to bridge the gap between formalizable and non-formalizable components of social-humanitarian AI.

2. Related works

Examples of application of traditional AI in social-humanitarian sphere are found in the acceleration of networked democracy processes (Raikov, 2018) or in raising the level of social responsibility (Raikov, 2020), using the special convergent approach, based on the methods of inverse problem solving on topology spaces, control thermodynamics, cognitive modeling, genetic algorithms, etc. Other AI methods typically incur in the difficulties as follows:

(1) Incorrect diagnosis of the state and behavior of individuals in the community;

(2) Impossibility of eliminating informational interference arising in the diagnosis of the behavior of groups of people;

(3) The need to use different means and methods to assess the behavior of various subjects, objects and events;

(4) Limited ability to use visual diagnostics to implement the functions of expensive sensors;

(5) Incompleteness of knowledge about the deep structure of the human body in all its complexity in the context of non-invasive visual diagnostics;

(6) Unreliability and latency of data (“dark data”) and

(7) Impossibility in complex integration of the operation of various neural networks based on an algorithmic approach, etc.

Although the effectiveness of the current AI resides in its abandoning of trying to imitate non-formalizable human thinking and emotions (Esposito, 2017), social-humanitarian AI still need efforts on the strong AI side. Machines should be endowed with the capability to extend their limits beyond any specific logic and formal system, embracing many of them swiftly and seamlessly in order to account for superveniences or emergences of meta-levels of consciousness over cognition. The vision here proposed looks for realms that consider frameworks like the Varela’s and Merleau-Ponty’s neurophenomenology (Bitbol, 2021; Beaton *et al.*, 2013), or others like panabstractism (Pierce, 2021), in order to overcome mind–body duality with a dynamical open process.

Within this vision, the third-order cybernetics is a most promising and now mature framework (Lepskiy, 2018; Espejo and Lepskiy, 2020; Umpleby *et al.*, 2019; Raikov, 2019), in which first contribution is a switch from classical to post-non-classical scientific rationality, where the reality is not static and immutable, but co-created in relation to a self-developing reflexive-active environment made of entities and their explicit or tacit relations.

A new transdisciplinary connection between all the societal and natural entities has to be built, with new approaches that penetrate deeper into the secrets of the sensual and emotional layers of human consciousness and the collective unconscious (Avdeeva *et al.*, 2020). The new vison of hybrid reality (HyR) and its framework in its renovated sense (Perko, 2020) is the realm where the contact point between humans and machines happen as a dynamic unstable system in which people and AI technology coexist and affect each other.

In this context, the concept of holon (Koestler, 1970) receives a new life and a central role for technological developments. The holon can be associated to any of the vertical recursive connection entities that intervene to integrate hierarchy levels of an organized system of systems. For example, the role that was assigned to Markov blankets in (Palacios *et al.*, 2020) can constitute an instance of holons that can take, in many cases, the function of membranes between different layers of organization in biological context. The holon is the interpreter and mediator between two realities, between macro and micro level, between two formal systems or even between two not representable but enacting and “live” objects in the state of affairs.

The holon is any entity (abstract or natural) that simultaneously straddles two adjacent levels of a hierarchy with associated reality objects, thanks to the so called *Janus effect* (Koestler, 1970): each element of the hierarchy has a face downward the sub-level that sees a self-contained whole of sub-assemblies and a face upward the apex acting the role of the part. In current industrial engineering, holons often embody the contact point between the physical and the cyber in cyber-physical systems (Derigent *et al.*, 2020; Bonci *et al.*, 2018).

A most notable and practically used concept enabled by the holon is that of the holarchy. The holarchy is a temporary purposeful grouping of holons into a hierarchical relationship (or in general a directed acyclic graph relationship). The holarchy concept is particularly useful when used as a functional structure for the expression of autopoiesis and emergent self-organization in systems.

The openness in the granularity and in the number of the levels of the holarchy requires an epistemic invariant. The relational-model multi-agent system (RMAS) architecture (Pirani *et al.*, 2022) identifies the full relational model as such a candidate invariant. Agents that implement holons in RMAS are good candidates to become the amphibians across the multiple realities that are co-created in light of third-order cybernetics (Lepskiy, 2018; Raikov, 2019) in which an adequate framework could be established for a proactive explanant and witness of the HyR phenomenon. Such a framework for the deep comprehension of the HyR is highly desirable as, at the moment, there is a gap in systems theory and engineering when the contact point between humans and technology is concerned as a whole.

Currently, useable models of HyR are available as proposed for example by Perko (2020). However, they remain on the observational perspective of the human side. The epistemological perspective that goes in reverse direction, from machines and algorithms that observe and manipulate humans, is lagging behind. By the adoption of suitable holon implementations, as exposed in section 4, new ways of modeling and subsequent control opportunities are created in order for the dual nature of the HyR to be explained with the help of autonomous activity by AI entities.

3. Advanced AI approaches

The semantics of social-humanitarian AI models cannot be comprehensively described in a formalizable way. However, it can be covered in a non-direct way by approaches based on electromagnetic waves, quantum and relativistic effects, including the human into the process, in order to exploit non-locality effects (Raikov, 2021).

Non-formalizable and non-local means are used in the convergent method of inverse problem solving in non-metric topological spaces (Raikov, 2018, 2020). In this case, the goal has a poor definition. The problem-solving process can have a non-stable character, but the application of some special information structuring creates the necessary conditions for making the decision process more purposeful and sustainable toward a possibly ill-defined goal (Ivanov, 1969). During the inverse problem-solving process, a group of people introduces information into the problem-solving process, providing sustainable purposefulness (convergence) through human thoughts and emotions. Some rules can ensure the convergence of the collective decision-making process. The rules for structuring information in the special convergent way are as follows:

(1) A tree of goals should be ranked by levels and importance;

(2) The spaces of AI model's semantics interpretations should be Hausdorff separable;

(3) The space of resources should be covered by a finite number of classes (the compact space) and

(4) Decision-makers should introduce in the process information, which lies beyond logic and formalisms.

Since neurons in the brain and cells in the body are made of atoms, quantum effects can also influence the thought and decision-making processes in HyR (Raikov, 2018, 2020, 2021). In this case, the quantum equations, such as the Schrödinger one and Heisenberg's uncertainty principle, can help in describing states and behavior of the thought process. As a result, the non-formalizable cognitive semantics of AI models based on quantum and wave-particle aspects of the field theory cannot be fully described using the classical logical and neuro-net approaches to AI.

Figure 1 illustrates the main aspects of this case. Space of trust among people with motivation is necessary for creating HyR. There are two fundamental pillars of collective decision-making on the left side of the figure. The first one, mathematical, forms the abovementioned convergent rules. The second one, thermodynamic, suggests rules for ensuring the sustainability of the collective decision-making process, namely a delicate balance between the formalizable (P) and chaotic (S) components ensure sustainability. The top part of the figure illustrates that the atomic and wave nature of the human brain's neurons and body's cells, influence collective thinking and decision-making processes.

These components form the cognitive (non-formalizable) semantics of Artificial Strong Intelligence (ASI). ASI is intended as opposed to Weak AI in which the agent can produce meaningful results to the observer even without any understanding of the meaning of its

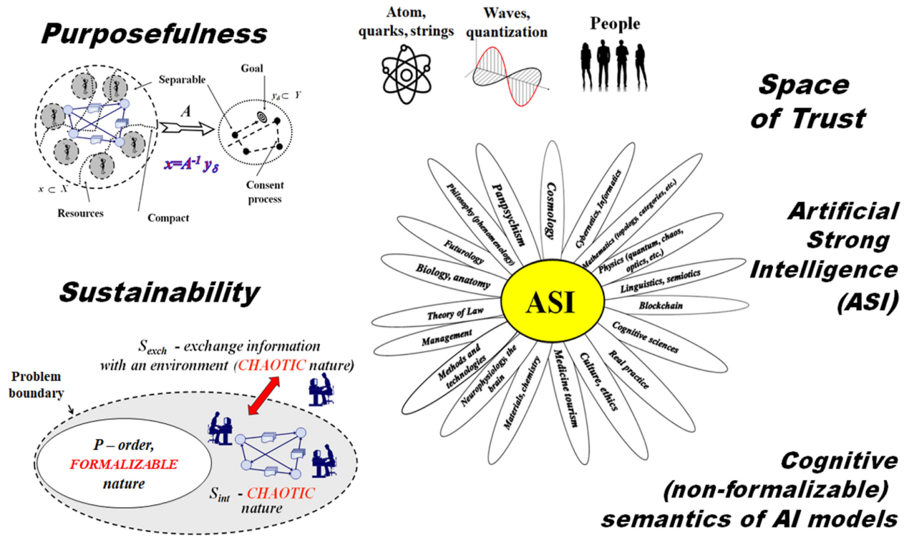

Figure 1.
Fundamental aspects of advanced artificial intelligence

operations. An ASI agent can make good decisions as with or without human help by immersing itself in HyR. A notable example of this kind of Weak AI processor was given in the famous Chinese Room argument of Searle. An ASI entity has to feature understanding as a conscious self. The detection, model and control of the onset of such consciousness in an artificial entity, whether it may ever appear or not (Fjelland, 2020), is an open problem. Such ASI capabilities can be described by using and crossing many different disciplines (Raikov, 2021), which here are shown as chamomile in Figure 1.

If a wave characterizes the cognitive semantics, it can be quantized, which is an immanent property of quantum particles that feature fluctuations following Heisenberg's uncertainty principle. Such a model and analogies can be useful for explaining human insight too. Before words can express the thought, the process of thinking looks chaotic. The purpose of thinking is to create an order in the chaos. Simultaneously, if this order is constructed based on existing knowledge, then the expected insight of a new idea will not occur because the thought will be frustrated by the retrospective logical stereotypes. When only deduction and induction rules of inference work, then the process becomes divergent and purposefulness is lost.

In addition, a logic formula cannot directly describe thoughts and emotions because they would immediately be abiding by a symbolic scheme with unavoidable inconsistencies. Verbalization is always doomed to neglect the potential of emotions and thoughts. However, some fundamental laws can help to represent a chaotic thought indirectly. For example, the chaotic peculiarity of thought can be represented by a corresponding dynamic system, but having some multidimensional space state from group theory. Two events are linked to one another, forming connected Lorentz subgroups: one rotation is translated into another by a continuous change. With this interpretation, the state of a thinking system is specified by a set of positions and vectors of events, for example in the system of generalized coordinates and velocities of the Minkowski space or vectors and their derivatives in an infinite-dimensional Hilbert space. States of these spaces are turning one into another over time. These changes can be defined in formalized and in non-formalized ways as well.

Taking these aspects into account, the describing processes of integration of AI and human beings in the united, or hybrid system, can be represented and materialized by the special *equation of HyR*.

4. The equation of hybrid reality

An agent tries to use models of realities in order to learn cumulatively and arrive to a cognitive goal. However, models based on mathematics hold only at the bifurcation of convergence and divergence of organization and disorganization in general systems (Bogdanov, 1984). Indeed, stability is a prerequisite for the existence and representation of reality. Agents have to go beyond any mathematics representation in order to materialize purposeful construction of organization as a natural "instinctive" embedded mechanism that creates a stable state of knowledge and being as an eigenform (Von Foerster, 2003).

While symbolic AI or mathematics systems handle snapshots or symbolic representations of things, an eigenform *is* the thing. Such an entity, immersed in its environment, makes continuous queries to the environment with senses and actuators. Accordingly, it receives replies as information and associated change of internal state: this is the act of knowing. The entity knows something when repeated queries provide the same answers. In this case, the agent finds (or feels) itself in a stable state of *being*. If some answers do not steadily arrive or are not as expected, the entity experiences an "irritation" (Füllsack, 2014). The irritation triggers a new knowing procedure in order to reach another stable state of being.

An example of eigenform is an autonomous agent trying to be in control of the state of affairs; she/he/it has to experience equilibrium between the *being* (i.e. acting effectively in the environment) and *knowing*. Agent's organization (or consciousness in some sense) can be modeled as a transformation (in eigenspace) of the eigenfunction of querying (knowing) about being: $K\boldsymbol{b} = \lambda \boldsymbol{b}$, where K is the operator associated to the act of knowing, corresponding to each observable quantity λ (eigenvalue), by querying the reality (sensing, learning, trying, etc.) and $\boldsymbol{b}$ is the eigenvector or formal vector of representation of *being* (having a contextualized local meaning).

In general, the former simple linear eigenvalue λ problem is to be coped as a fixed-point problem of mathematics. The solution of this problem does not exist in closed form in general, but the eigenform itself can realize (embody) a living form of it without explicit symbolic means (as tacit knowledge, inner experience, qualia etc.). An eigenform captures practically the major and dynamic visions in constructivism like neurophenomenology since Varela and Merleau Ponty (Bitbol, 2021; Beaton *et al.*, 2013) or in relational interpretation of quantum physics (Rovelli, 1996). This is a constructive approach for which higher-order cybernetics can be defined as *pragmatic constructivism* (Lissack, 2016).

The action that generates agent's reality object (the eigenform) can be expressed mathematically. If X is the *being* and f the act of *knowing*, we have $X = f(X)$. When f is applied indefinitely (though at established sampling events), the X is said the fixed point for the invariant operator f, $X = f(f(f(f(f(f(f(f(\ldots))))))))$. X is a fixed point of f as a value that is mapped indefinitely into itself by the function. Due to this property, an eigenform is a fixed point for a transformation. Moreover, f is at a level where the level and the meta-level are one and f is both object and subject of the discourse (Kauffman, 2005).

As for third-order cybernetics, the agent is immersed in a reflexive-active domain. This domain can be consistently defined with the use of the Church-Curry fixed-point theorem (Kauffman, 2016): "*Let D be a reflexive domain and let T be any element of the domain D. Then there is an element X in D such that X is a fixed point for T. That is, TX = X. X is then the eigenform for T in D.*"

In particular, in third-order cybernetics, the "object" is a stable relation between X, the observable; E, the observer or experimenter and S, the society or environment. This relation can be expressed by a triple $<X,E,S>$. Thus, in a third-order cybernetics context, the reality is materialized by the following fixed-point equation, through the T transformation:

$$\langle S, E, X \rangle = \mathrm{T}(\langle S, E, X \rangle) \tag{1}$$

The solution of Eq. (1) is an eigenform, which can exist in closed form or not. Any "computation" that solves the "object[s] as token[s] for eigenbehaviors" (Von Foerster, 2003).

The last step here put forth is clarifying the role of *holon* entities as the membrane between different realities consisting of eigenforms. As a particular case, for the HyR focus (Perko, 2020), it means treating the digital and the physical as two co-existing realities, continuously influencing and reflexing each other, as happens between denotative (formalizable, digital) and cognitive (non-formalizable) semantics of AI models (Raikov, 2021).

With the use of category theory, in the membrane across the two realities passes a *structuring functor* that transforms any symbolic system (a computational process, a state machine, a Büchi automata, etc.) into a dynamic system process. In this direction, a computation is embodied into "laws of physics." In the reverse, a *forgetful functor* determines an abstraction, which is the lifting from physical world into some logic or symbolic system. These functorial relations, when used in two directions, create the possibility of a *bisimulation*: the entities of *World #1* can see and use abstractions of *World #2*; the entities of *World #2* can see and use abstractions of *World #1*. The holonic nature of the membrane between the two worlds will handle and give body and meaning to this dual and continuous transformation as represented in Figure 2.

The structuring functorial transformation corresponds to an embodiment of an eigenform into a dynamical system, such as follows:

$$\dot{x}(t) = f(x(t), t) + g(x(t), u(t), t) + d(w(t), t), \tag{2}$$

where typically x is the state realization, f is the component related to state expression, g conveys the contribution of the input u and d accounts for exogenous environmental effects as w. If the system in (2) is made stable (through many of the feedback stabilization methods like Lyapunov-based techniques), then equation (2) equals to 0 asymptotically or semi-globally. In case of stability (2) can be rewritten into (1) and its solution is the solution of the fixed-point problem:

$$\langle S(w(t), t), E(u(t), t), X(x(t), t)\rangle = \mathrm{T}(\langle S(w(t), t), E(u(t), t), X(x(t), t)\rangle) \tag{3}$$

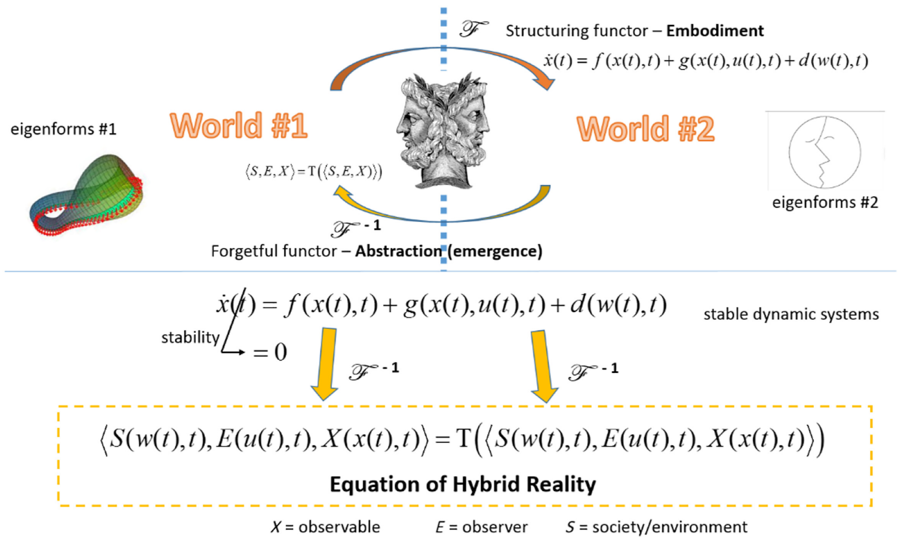

Figure 2.
The holonic membrane between two realities, whose meaningful and purposeful objects are materialized by the solution of the third-order fixed-point problem, deemed the *equation of HyR*

The fixed-point problem of eigenforms of (3) is deemed the *Equation of Hybrid reality*, just to express the huge impact that its solution can bring in science and its implications in technology that aims to realize it. Any implementation that achieves to solve the fixed problem (3) is an organized and stable entity.

4.1 Further implications from the 3rd order cybernetics framework

The methodology here proposed has been mostly contextualized in von Foerster's vision of stability of living systems under second-order cybernetics, which is an expression of a dynamic coupling between a system and a metasystem. The notion of *eigenform* embodies the relationship between a subject and object within the context of an observer, and it can be mapped to a stable living system that has an eigenvalue that is the outcome of the *eigenbehavior* that occurs for the metasystem–system couple. This suffices for the living system to feature autopoies and self-organization. The extension to 3rd order cybernetics has then been achieved by considering just another dimension to the reality, concerning the context of being of the living system subject to social conditioning.

This very simplification here performed is at the same time a limit and strength. It is strength in the sense of aiming at a first tentative toward practical materialization of the problem with a solution that can count on a well-defined expression as the one in (3).

Nevertheless, this expression still fails to address completely the self-organization required by aforementioned "irritation" event as envisioned by Füllsack (2014). This limit can be addressed relying to richer or merely alternative conceptions of third-order cybernetics that consider autopoiesis just a first step in the overall organization of a living being.

For example, in the Eric Schwarz's third-order cybernetics, the complementary autogenesis concept is brought about to widen the reach of self-organization to a meta-level that can be used to reduce uncertainties due to the variety of environmental interactions. Autogenesis addresses self-creation, which is what makes a system autonomous and able to create its own laws (Schwarz and Yolles, 2019). Valuably, the autogenesis is considered also a basis for an extension of paradigms of cybernetics up to the n-th order (Yolles, 2021). Another example is the third-order cybernetics definition of Ashby in which the ethics system constitutes the augmentation of the second-order into a cybernetics of ethical systems in which the third observer level is adjoint and different from the first and second observers (Ashby, 2020, 2022).

Due to the many shapes that the third-order cybernetics can assume and the richness of its definitions, further interesting and challenging work on the expression of the *equation of HyR* can be envisioned and anticipated since now.

4.2 Implications and future work on semiotics aspects of human-machine contradiction

The envisioned technology for the realization of the *equation of HyR* brings about the necessity to look for solid logic-mathematics groundings in the structuring of the semiotic systems that the aforementioned set of eigenbehaviors is prone to achieve when enacting in the formalizable part of reality.

Such a solid grounding can be found for example in the works on Deontic Impure Systems of Nescolarde-Selva and Usó-Doménech (2014), Nescolarde-Selva (2010), Nescolarde-Selva and Usó-Doménech (2012) and references therein. Holonic agents that aim to materialize the solution of the *equation of HyR* are going to construct the semiotic environment that is formed by the epistemologies and the associated belief systems surrounding a holonic subject. With so doing the holonic entities are expected to mutuate what is supposedly possible only by human subjects (Nescolarde-Selva and Usó-Doménech, 2014; Usó-Doménech and Nescolarde-Selva, 2016).

This human-artificial mapping becomes possible when second-order (and higher-order) cybernetics are put into suitable homomorphic relations for the construction of a hybrid semiotic environment. The recursive cybernetics property that links semiosis with eigenbehavior has been studied in some depth by authors like Gasparyan (2015, 2020) and Kauffman (2016).

The holon that is able to solve the *equation of HyR* is endowed with the mission to handle the transformation between an abstract system and an impure system as defined by Nescolarde-Selva and Usó-Doménech (2014). The *relative being* acting in an impure system is doomed to confront semantic incompleteness principle in which "any proposition based on a certain language L that includes the reality contradicts to itself" (Nescolarde-Selva, 2010), which derives in essence from the famous incompleteness theorem of Gödel. This might mean that if a realism-based framework is adopted, the systemic semiotic approach tends to see reality as an illusion.

If we adopt constructivism, this illusion becomes an active solid ground on which an agent can thrive and evolve. In addition, reality contradiction happens if we look at only one language and worldview at time. The solution to this contradiction is indeed to consider all the possible languages at the same time (although locally one by one) and switch among one and another with the same flexibility as humans seemingly do (Carruthers, 2017). In this reasoning agility that traverses multiple semiotic systems reside the secret of general AI and maybe also some seeds of consciousness for it.

Reasoning under classical ways cannot prove the beliefs it is based upon as beliefs arise through agent's experience, previous beliefs and reason to be assimilated. But most of all, reason needs experience to be formed, and beliefs need reason as well in a circular way in which reason and experience are based upon each other. This implies that the agent's context is dynamic, and formed upon beliefs, reason and experience (Nescolarde-Selva and Usó-Doménech, 2014), but implies also that the relative being is the subject of a second-order cybernetics recursion that should be handled with the ouroboros nature of the semiotic discourse exposed by Gasparyan (2020). In second-order and higher-order cybernetics and their vertical and horizontal recursive structures (Yolles, 2021), it is rather natural that significance at one level of context becomes a significant for a superior level in the semiotic system definition problem.

In a systemic conception of reality, the subjective agent is hypothetically found in a certain state characterized by several belief conditions like the Omphalical Belief and the Janus Belief conditions (Nescolarde-Selva and Usó-Doménech, 2014) that are integrated in the essential definition of the holon.

The semiotic vision usually adopted is anthropocentric as the observer subject is considered human. For sure, machines of today cannot construct a functionally autonomous semiotic system nor they have a belief system. They simply reflect, enable or amplify semiotics made by human subjects (Esposito, 2017). This is a first source of conflict. Currently machines are only able of intelligent communication rather than AI. Thus, it is necessary to look for a different stance in which machines are able to construct and communicate their own semiotic constructions and put them into symbiosis with the humans'.

The use of modal logic (logic of first and higher orders) cannot handle alone the unformalizable phenomena and their semantics that constitute relevant part of the socio-humanitarian phenomenology. The convergent methodology here reclaimed aims to provide the means to understand and control the unification between the human and the artificial into a new hybrid semiotic system (the semiotic system in HyR). The construction of such a system can be better achieved if some of the *super-human* capabilities of machines will be exploited at their best. The convergent method aims at reuniting into a whole (though keeping components well-distinguishable) formalizable or non-formalizable consciousness of the humans and of the machines. While human consciousness is immediately immanent to us, the shape and nature of machines' consciousness, whatever the shapes it will take, is a completely open field of urgent research.

5. Verification of the approach in real practice

The suggested approach is used for creating different strategies in the real economy. The social-humanitarian AI helps to accelerate getting an agreement between the participants of teams during strategic meetings in regional governments. Outside experts and civil representatives are invited.

The strategic planning process immerses itself into social-humanitarian context. It usually deserves a socio-economic system analysis by a team with using computer modeling methods, and a HyR perspective has to embrace both formalized and non-formalized ways of representation of information. Semiotic systems are at the basis of the formalizable and logic (denotative semantics) treatment of the sheaves of relations that are systemically established between objects and subjects of reality (Nescolarde-Selva and Usó-Doménech, 2014). In contrast, the non-formalizable aspects of the social-humanitarian system and models (cognitive semantics) needs activating enaction in the process of solving inverse problems by immersing the subject into the system in order to activate the mechanisms of convergence in decision-making. By means of the approach here suggested, the integration of AI (modeling) and human beings in a new symbiotic HyR framework is expected to provide an enhancement of accuracy and effectiveness in modeling and control of otherwise divergent dynamics.

Let us take a case of preparing a three-year strategy of development of youth policy in one of the regions of Russia, targeting also the strengthening of the atmosphere of leadership and team trust. A strategic meeting was held indeed to develop priorities for the regional youth policy. The specifics of the meeting were involving the use of group situational analysis technology and social-humanitarian AI that is designed in order to quickly coordinate the strategic interests of its participants.

Youth policy is formed in the context of young people's manifestation of independence and interest in politics. The intentions of young people are characterized by patriotism, pragmatism, dynamism, orientation toward results and success, optimism and willingness to take risks. The outcome of election campaigns and the direction of innovative, social, economic and social transformations in the region largely depend on the political, economic, social and technological preferences of young people.

Under these conditions, the vector of the main goals of the regional youth policy continues to turn toward the creation of conditions and support for various forms of socio-political, civil-patriotic, marketing-economic and innovative-technological self-realization of the young people.

Dozens of factors influence the effectiveness of youth policy, which differently and contradictorily characterize the political, economic, social and technological situation related to the achievement of the goals of the youth policy of the region. The objectives of the meeting included the search for optimal ways to achieve goals, including at the verbal, project and motivational levels.

The meeting was held for two days in a country hotel complex in which one of the small halls was equipped as a situational room. A sports hall, a swimming pool and a couple of rooms were also used to organize discussions of individual problems. The schedule of the meeting and the processing of the results were determined by the method of organizing strategic meetings using cognitive computer modeling tools. The use of the gym and swimming pool was used for intense relaxation, as well as the use of game methods of team building. This mode of operation made it possible to start the meeting at 7.00 a.m. and end at 11.00 p.m. The order of strategic meeting included the steps as follows:

(1) Characterize the current state and formulate a vision of the youth policy of the region;

(2) Identify and formulate factors that characterize the external and internal possibilities of youth policy;

(3) Identify and formulate the factors that characterize the external and internal weaknesses of the youth policy;

(4) Formulate a list of problems in the development of youth policy with an assessment of their importance;

(5) Formulate lists of strategic priorities for the development of youth policy with an assessment of their importance;

(6) Offer examples of the most promising projects according to strategic priorities (in the form of training) and

(7) Build examples of effective action chains (processes) of project team members.

The report was prepared after the strategic meeting containing the main results that were then used to:

(1) Preparation of proposals to clarify the organizational and functional structure of the Committee on Youth Policy of the regional government;

(2) Organization of marketing monitoring in the field of solving issues of increasing the effectiveness of youth policy and

(3) Development of the concept and strategic plan for the development of regional youth policy.

The strategic plan defined development goals agreed upon among stakeholders and ways to ensure the achievement of these goals. The priorities of the youth policy determine the preferences in applying the efforts (material, technical, financial and intellectual) of the authorities and local self-government in the implementation of the tasks of the youth policy of the region for a certain period.

6. Discussion and conclusion

At the beginning of the paper, it was noticed that modern AI can only recognize, predict and answer far too easy questions. It can operate only symbolic and formalized data. However, solving social and humanitarian problems with the help of AI requires obtaining a coordinated solution by groups of people and artificial actors in the HyR context. These hybrid systems, individuals or actors are usually asked questions that point toward problems that cannot be resolved neither by people nor by artificial actors alone. This is an important aspect for the prospective use of social and humanitarian AI in the field of solving long-standing fundamental problems in science and obtaining new scientific discoveries.

In order to solve social and humanitarian problems, traditional formalized, logical and symbolic tools, which are added to the neural networks, should be enriched with the convergent and the holonic multi-agent approaches, cognitive and quantum semantics, genetic algorithms and other natural computations methods. In this context, current AI technologies, which allow only symbolic realizations, can emerge as eigenforms into non-formalizable objects.

The *equation of HyR* here proposed, as a schema of strong AI, can have divergent or convergent types of solving, but it has the potential to ensure conditions for purposeful and sustainable collective decision-making, being it a unifying ground of the authors' convergent and holonic approaches.

The represented practical example showed the fruitfulness of the proposed approach to use the social-humanitarian AI that operates in a hybrid (human-machine) reality to build a successful strategy of the region's economic development branch in 2–3 weeks, as opposed to traditional strategy preparation, which can take 5–7 months.

References

Ashby, M. (2020), "Ethical regulators and super-ethical systems", *Systems*, Vol. 8 No. 4, p. 53, doi: 10.3390/systems8040053.

Ashby, M. (2022), "Problems with abstract observers and advantages of a model-centric paradigm", *Systems*, Vol. 10 No. 3, p. 53, doi: 10.3390/systems10030053.

Avdeeva, Z.K., Kovriga, S.V., Lepskiy, V.E., Raikov, A.N., Slavin, B.B. and Zatsarinny, A.A. (2020), "The distributed situational centers system as an instrument of state and corporate strategic goal-setting in the digital economy", *IFAC-PapersOnLine*, Vol. 53 No. 2, pp. 17499-17504, doi: 10.1016/j.ifacol.2020.12.2156.

Beaton, M., Pierce, B. and Stuart, S. (2013), "Neurophenomenology – a special issue", *Constructivist Foundations*, Vol. 8 No. 3, pp. 265-268, available at: http://constructivist.info/8/3/265

Bitbol, M. (2021), "The tangled dialectic of body and consciousness: a metaphysical counterpart of radical neurophenomenology", *Constructivist Foundations*, Vol. 16 No. 2, pp. 141-151, available at: https://constructivist.info/16/2/141.bitbol

Bogdanov, A. (1984), *Essays in Tektology: The General Science of Organization*, 2nd ed., Intersystems Publications, Seaside, CA, Transl. Gorelik, C.

Bonci, A., Pirani, M., Cucchiarelli, A., Carbonari, A., Naticchia, B. and Longhi, S. (2018), "A review of recursive holarchies for viable systems in CPSs", *2018 IEEE 16th International Conference on Industrial Informatics (INDIN)*, IEEE, pp. 37-42, doi: 10.1109/INDIN.2018.8472055.

Carruthers, P. (2017), "The illusion of conscious thought", *Journal of Consciousness Studies*, Vol. 24 Nos 9-10, pp. 228-252.

Derigent, W., Cardin, O. and Trentesaux, D. (2020), "Industry 4.0: contributions of holonic manufacturing control architectures and future challenges", *Journal of Intelligent Manufacturing*, Vol. 32, pp. 1797-1818, doi: 10.1007/10.1007/s10845-020-01532-x.

Espejo, R. and Lepskiy, V. (2020), "An agenda for ontological cybernetics and social responsibility", *Kybernetes*, Vol. 50 No. 3, pp. 694-710, doi: 10.1108/K-06-2020-0390.

Esposito, E. (2017), "Artificial communication? The production of contingency by algorithms", *Zeitschrift für Soziologie*, Vol. 46 No. 2, pp. 249-265, doi: 10.1515/zfsoz-2017-1014.

Fjelland, R. (2020), "Why general artificial intelligence will not be realized", *Humanities and Social Sciences Communications*, Vol. 7 No. 1, pp. 1-9, doi: 10.1057/s41599-020-0494-4.

Füllsack, M. (2014), "The circular conditions of second-order science sporadically illustrated with agent-based experiments at the roots of observation", *Constructivist Foundations*, Vol. 10 No. 1, pp. 46-54, available at: http://constructivist.info/10/1/046

Gasparyan, D. (2015), "What can the global observer know?", *Constructivist Foundations*, Vol. 10 No. 2, pp. 227-237, available at: http://constructivist.info/10/2/227

Gasparyan, D. (2020), "Semiosis as eigenform and observation as recursive interpretation", *Constructivist Foundations*, Vol. 15 No. 3, pp. 271-279, available at: https://constructivist.info/15/3/271

Ivanov, V.K. (1969), "Incorrect problems in topological spaces", *Siberian Mathematical Journal*, Vol. 10, pp. 785-791, Novosibirsk, Russia (In Russian).

Kauffman, L.H. (2005), "Eigenform", *Kybernetes*, Vol. 34 Nos 1/2, pp. 129-150, doi: 10.1108/03684920510575780.

Kauffman, L.H. (2016), "Cybernetics, reflexivity and second-order science", *Constructivist Foundations*, Vol. 11 No. 3, pp. 489-497, available at: http://constructivist.info/11/3/489

Koestler, A. (1970), "Beyond atomism and holism—the concept of the holon", *Perspectives in Biology and Medicine*, Vol. 13 No. 2, pp. 131-154, doi: 10.1353/pbm.1970.0023.

Lepskiy, V. (2018), "Evolution of cybernetics: philosophical and methodological analysis", *Kybernetes*, Vol. 47 No. 2, pp. 249-261, doi: 10.1108/K-03-2017-0120.

Lissack, M.R. (2016), "Shed the name to find second-order success: renaming second-order cybernetics to rescue its essence", *Constructivist Foundations*, Vol. 11 No. 3, pp. 470-472, available at: http://constructivist.info/11/3/470

Mackintosh, K. (2000), *The Principles of Humanitarian Action in International Humanitarian Law*, Overseas Development Institute, London, available at: http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/305.pdf

Nescolarde-Selva, J.A. (2010), "A systemic vision of belief systems and ideologies", Doctoral dissertation, Universitat d'Alacant-Universidad de Alicante.

Nescolarde-Selva, J. and Usó-Doménech, J.L. (2014), "Reality, systems and impure systems", *Foundations of Science*, Vol. 19 No. 3, pp. 289-306, doi: 10.1007/s10699-013-9337-8.

Nescolarde-Selva, J. and Usó-Doménech, J.L. (2012), "An introduction to alysidal algebra (III)", *Kybernetes*, Vol. 41 No. 10, pp. 1638-1649, doi: 10.1108/03684921211276800.

Palacios, E.R., Razi, A., Parr, T., Kirchhoff, M. and Friston, K. (2020), "On Markov blankets and hierarchical self-organisation", *Journal of Theoretical Biology*, Vol. 486, p. 110089, doi: 10.1016/j.jtbi.2019.110089.

Perko, I. (2020), "Hybrid reality development - can social responsibility concepts provide guidance?", *Kybernetes*, Vol. 50 No. 3, pp. 676-693, doi: 10.1108/K-01-2020-0061.

Pierce, B. (2021), "Can panabstractism offer an alternative approach to the hard problem?", *Constructivist Foundations*, Vol. 16 No. 2, pp. 161-163, available at: https://constructivist.info/16/2/161

Pirani, M., Bonci, A. and Longhi, S. (2022), "Towards a formal model of computation for RMAS", *Procedia Computer Science*, Vol. 200, pp. 865-877, doi: 10.1016/j.procs.2022.01.284.

Raikov, A. (2018), "Accelerating technology for self-organising networked democracy", *Futures*, Vol. 103 No. 17, pp. 17-26, doi: 10.1016/j.futures.2018.03.015.

Raikov, A. (2019), "Post-non-classical artificial intelligence and its pioneer practical applications", *IFAC-PapersOnLine*, Vol. 52 No. 25, pp. 343-348, doi: 10.1016/j.ifacol.2019.12.547.

Raikov, A. (2020), "Convergent social responsibility as the key to corporate strategic success", *Kybernetes*, Vol. 50 No. 3, pp. 785-793, doi: 10.1108/K-12-2019-0853.

Raikov, A. (2021), "Cognitive semantics of artificial intelligence: a new perspective", *Topics: Computational Intelligence*, Springer Singapore, Vol. 17. doi: 10.1007/978-981-33-6750-0.

Rovelli, C. (1996), "Relational quantum mechanics", *International Journal of Theoretical Physics*, Vol. 35 No. 8, pp. 1637-1678, doi: 10.1007/BF02302261.

Schwarz, E. and Yolles, M. (2019), "The third order cybernetics of Eric Schwarz", available at: https://ssrn.com/abstract=3427683, doi: 10.2139/ssrn.3427683.

Umpleby, S.A., Medvedeva, T.A. and Lepskiy, V. (2019), "Recent developments in cybernetics, from cognition to social systems", *Cybernetics and Systems*, Vol. 50 No. 4, pp. 367-382, doi: 10.1080/01969722.2019.1574326.

Usó-Doménech, J.L. and Nescolarde-Selva, J. (2016), "What are belief systems?", *Foundations of Science*, Vol. 21 No. 1, pp. 147-152, doi: 10.1007/s10699-021-09824-y.

Von Foerster, H. (2003), "Objects: tokens for (eigen-) behaviors", in *Understanding Understanding: Essays on Cybernetics and Cognition*, Springer-Verlag, New York, Berlin, Heidelberg, pp. 261-271.

Yolles, M. (2021), "Metacybernetics: towards a general theory of higher order cybernetics", *Systems*, Vol. 9 No. 2, p. 34, doi: 10.3390/systems9020034.

Corresponding author
Massimiliano Pirani can be contacted at: massimiliano.pirani@gmail.com