FILOZOFIA I NAUKA
Studia filozoficzne i interdyscyplinarne
Tom 10, zeszyt specjalny, 2022

Editorial

PHILOSOPHY AND COMPUTING: AI, VIRTUALITY, EPITEMICITY

doi: 10.37240/FiN.2022.10.zs.0

1. THE CALLING OF PHILOSOPHY AND COMPUTING TODAY

We need a productive, veridical narrative related to the current growth of artificial intelligence and its social role.

The narrative of fear that came from the early movies and novels is great for a thriller; but in terms of social discourse, it puts us at the level of Russian peasants, during the reign Tzar Alexander, kneeling at the view of the early trains, taking them to be a deed of the Satan.

The narrative of dismissal, claiming that those are just tools and need to be treated as such, is highly outdated since computer programs and robots are much more already, and their capabilities increase. The approach of Isaac Assimov, prescribing that we must view computers as slaves, limited to obedient following of human orders, came from the narrative of dismissal, and turned it into a narrative of enslavement of intelligent beings such as AI. *My criticism of Assimov's approach is not particularly guided by moral disagreement with him (at this point moral status of AI is unclear, and needs much further work and development of future AI before it gets settled). It is based on practical considerations. It is a waste to treat a philosopher as a slave (to use Plato's example) since he would be a terrible slave, and if there is social use for philosophers it is clearly not facilitated through enslaving them. The same goes for advanced AI—it it just waits for us to decide, the car it "helps drive" is likely to crush and the armed rocket it can strike down is going to reach its target before its human handler can figure out what the matter is. We are just not efficient enough to be in charge of those new generations of smart and efficient being—at least at the level of fast implementation.*

The narrative of surrender is also a bad choice, just like excessive fear or dismissal. From my criticism of those attitudes, it does not follow that AI should become in charge of the human world. In extremally time-sensitive

predicaments, and in complex multi-factor mathematical analysis—human beings cannot even conceptualize some of the factors that surface through advanced big data analysis within open conceptual frameworks. We should be able to set up the general goals, exclude and include some of the acceptable means of attaining them (e.g. through including moral and high level legal imperatives, as well as the economic and other objectives).

Thus, *the narrative of optimal balance between human and artificial intelligence* emerges as the sole strategy to move forward, towards human flourishing, not engrossing oneself into the *triller-quality pessimism* or some tendencies to reverse towards the slavery economy—now by oppressing not only the enslaved human beings[1] but also artificial intelligence.

The general drift of the Asian philosophy and economy, for instance in China and Japan (however different those mentalities and social practices are), makes it easier to conceptualize such bi-directional approach than getting it through o of the other of conceptual frameworks. With all the creative potential of Western cultures, it is worth trying to develop a closer understanding of the mechanics and semiotics of the world with the artificial beings playing a role of much more than the tools, but effectively balanced by the human good, values and objectives.

Philosophy of Computers and especially Philosophy of AI plays the role of carrying this relevant debate beyond the academic lecture-room or strategy think tanks among the politicians or business leaders.

In the current issue of *Filozofia i Nauka* [Philosophy and Science] some of those topics are posed directly; but most are conversation openers for further debates. The more various routes we explore the better the chances of a reaching constructive world-views in the epoch of AI.

2. PHILOSOPHY AND COMPUTING TODAY

This issue of *Filozofia i Nauka* presents some of the important aspects of Philosophy of Computing in 2021–2022, which is when all those papers have been created. The first part of the volume is devoted primarily to what we decided to call: Philosophy shaped by Artificial Intelligence (AI), or artificial general intelligence (AGI). It is essential since AI, especially those projects that pave the way towards AGI, open substantial philosophical issues; the rapid growth of those domains makes those issues even more relevant at this very moment.

The second area is Virtual Space, which becomes more and more relevant for our daily lives. The move towards placing so much of our work and life activities online, which is a substantial aspect of the Economy 4.0 (Rogers,

[1] It seems like the society enslaving some smart beings would enslave more of them, including the many human beings—which is not only bad ethics but also bad economics.

2016), accelerated largely due to the social distancing of the COVID pandemics. The third area is a bit more amorphous; it centers around the topics such as perception, imagination, motivation and the mind, as understood through the prism of not only human or animal cognitive activities but also those of computers.

3. PHILOSOPHY SHAPED BY AI

This issue starts with an article by Mark Burgin and Rao Mikkilineni on the seven layers of computation. The authors go beyond the standard distinction between symbolic and sub-symbolic computing. They introduce, or revamp, categories such as *super-symbolic computation, hybrid computation, fused computation, blended computation, and symbiotic computation*. The article skillfully incorporates the background in philosophical semiotics, especially the works of Charles Sanders Peirce, which creates the leitmotif, which comes back a few times but becomes the dominant theme in the article by Ricardo Gudwin and Eduardo Camargo, which closes this section.

I wrote on non-reductive physicalism for advanced AI. What may be of interest is that *I identify the stream of awareness with what neuroscience calls creature consciousness*. This allows us to ditch substance dualism (few people view creature consciousness as a *non-materialist* substance) and take non-reductive physicalism seriously enough. Kyrtin Arteides follows with his proposal of Philosophy 2:0. The author argues that super intelligent AI systems may help human philosophers sort out their disagreements and check some of their ideas against the background of the current sciences. What may be fascinating, or refreshingly, in this intriguing article, is that the author writes from the viewpoint of an expert in AI consciousness, viewing philosophy in this framework. Arteides proposes "perspective maps" to maintain contextuality of knowledge (Figure 1), but then he wants to resolve the differences through a couple of rounds of *mediator feedback* (Figure 2). I happen to believe, that Arteides' approach, especially multi-core analysis (Figure 4), could be repeated in not so distant future, in the context of sub-symbolic, fused and other systems based *through and through* on fuzzy logic, thus able to capture blobs of meaning, instead of packing philosophical ideas in analytical semantics of sorts. Arteides' fascinating ideas related to philosophy as ecology of thoughts (Figures 6–7, 9), seem to be a step in a similar direction. This superb article, is still built on the assumption that philosophies must be *human-readable*, which seems like a non-controversial assumption, does not it?

The section closes with two rather thorough articles. Building on predictive coding and predictive processing, Jeffrey White explore the possibility of creating metaphysical self. This is a paper where boldness in philosophi-

cal thinking comes from interpretations of some work in AI. Most philosophers would view this approach as highly suspect; yet, taking into account the new ontologies created by various cognitive architectures in AI, many things become an open question. The article by Eduardo Camargo, Ricardo Gudwin on grounding cognition in Peircean semantics is a part of the fascinating project developed by the second author for several years now. The paper seems to keep just the right balance (if there is such a thing) between philosophy and AI, building at the conceptual space that belongs to both domains. We encounter a rather detailed in presentation of Peirce's semiotics (in section 3), then generalized in knowledge generation for both human and robot domains (section 4). In section 5 the authors focus on post-Peircean ontology that follows smoothly from the preceding arguments. Section 6 is devoted to the area of action and creativity, also largely based on Peirce's theory, yet updated and really focused on contemporary uses, which is advanced in section 7.

4. VIRTUAL SPACE IN PHILOOPHY

The article by Mariusz Mazurek devoted to ontology of virtual objects, provides a bridge between ontological reflection from Part 1 of the issue and the issue of virtuality dominating Part 2 The paper focuses on modes of virtual existence—from private objects on someone's screen through the process of their social objectifying. This results in intersubjective and often autonomous objects. The paper also presents important works in ontology of virtual objects, including those by Michael Heim, Jeri Fink, Lynn Baker and leading Polish authors.

This is followed by two articles from the 4th space group, whose agenda is to work our specificity of the virtual space in modern world. Bogdan Popoveniuc presents the interrelationship between personal and moral identity in the virtual space. This is probably the only "Continental" article in this issue, and Continental in a good sense it is, especially as phenomenology of the 4th space. The paper focuses on ontologies of the virtual space viewed both in philosophical and engineering perspectives, mostly as a technological extension of reality. It also leads to epistemological reflection, that bridges Part 2 with Part 3 of the current issue. Popoveniuc analyses space, including prominently the virtual space, as the foundation of subjectivity, engaging in deep reflection on philosophy of self. He analyses virtualization as a gradual cultural and technological process. The paper is a source of eruditional information from Gilles Deleuze and Félix Guattari on rhizomes and reality as maps (so compatible with António Damasio's neuroscientific theory of mind) to Luciano Floridi's *distributed morality*. But the main focus is on identity se the self, crowned within Francisco Varela's and Hum-

berto Maturana's *autopoiesis* moved even to the futurological dimension in projecting future trends in the 4th space.

The article by Christoph M. Abels, Daniel Hardegger on privacy and transparency in the 4th space may look like a very applied paper. In fact it does follow up on Abels' presentation and papers that are just coming out that focus on multifarious aspects of privacy and its problems in the virtual space. However, the article has also a second important aspect. It contains the most complete presentation to date, of the 4th space theory. While Daniel Hardegger came out with his interpretation of the 4th space in January 2021, those publications are often hard to locate and function as working projects. In this article we have a transparent, interesting and novel presentation of Hardegger's theory of the 4th space, strengthened y joint work and discussions within the 4ths space group. The article is worth reading for both the theoretical part and its practical application to the issues of privacy.

The following article, by Dustin Gray on virtual forms of surveillance and control is an interesting follow up on the work by Abels. While the topic is obviously relevant, some of the proposed solutions lead to philosophical a conundrum—where limits on control are seen as necessarily coming from our functioning in virtual space, or even in a large society.

5. EPISTEMOLOGY AND COMPUTERS

The name of Part 3—*Epistemology and Computers*—is a broad heading under which we placed somewhat more traditional articles on philosophical problems informed by computer science, especially AI. Magnus Johnsson opens this part with his article on perception, imagery, memory and consciousness. The author focuses on BICA approach; namely similarity of cognitive architectures between AI and animal/human brains. Johansson argues that some of the principles he puts forth are relevant for phenomenal consciousness of machines. He also develops such epistemic issues in AI, as memory, consciousness and imagination.

Rafał Maciąg writes about *knowledge as a phenomenon in the area of digital technologies, in particular artificial intelligence*. What seems like a standard article in epistemology opens up to the epistemologies generated by AI and related fields. This is followed by an article by Pavel N. Baryshnikov on computationalism in philosophy of mind. It uses anti-computationalist arguments to tackle the semantic problems, especially the lack of semantic properties, in in the computationalist theory of mind. This is relevant in various areas of AI.

The paper by Robin Hill is built on an interesting observation. There is an *artificialist fallacy*, defined as "causal justification of the *influence* of a technology, particularly artificial intelligence, by appeal to the *existence*

of the technology." This is akin to well-known naturalist fallacy. This lucid article also tackles the issue of value judgments in the artificialist fallacy.

Simon X. Duan, in his article on the "Platonic computer" tackles another conundrum—the inverse hard problem of consciousness. Idealism holds that consciousness is the fundamental nature of reality, thus, "matter is a derivative of consciousness." Thus, it becomes impossible to justify existence of the material world, which is the "inverse hard problem of consciousness" coined by Max Velmans. The paper is based on the concept of meta-computing and meta-consciousness as essential in generating "abstract entities" as well as "physical and nonphysical realities."

Last but not least, we have an article on dual-process approach to the problem of AI agency perception, by Marcin Rabiza. The author focuses on the two kinds of agency: 1. automatic, routine, often unconscious; 2. slower, controlled, more conscious. This is applied to AI.

This issue is maximally devoted to philosophy in computing, especially in AI, or sometimes philosophy in AI—not so much to philosophy *about* AI. This differentiates it from many other philosophical publications on AI, especially those from the 20th century. Philosophy seems much needed in theoretical AI, and many forms of cognitive science—while those disciplines open new conceptual avenues for philosophical thinking informed by developments, especially in artificial consciousness. This fruitful phase is just at its beginnings, as long as philosophers do not pontificate based on the old good theories and stay informed in the general trends, while engineers and scientists treat philosophy as an opening field for brainstorming and a design opportunity for the new views on reality.

Piotr (Peter) Bołtuć
Guest-editor
University of Illinois at Springfield, USA
The Warsaw School of Economics, Poland