Workshop Philosophy in Technology
31st of May, 2022

Workshop is organized by Warsaw University of Technology and Pontifical University of John Paul II

Dr. Roman Krzanowski (Pontifical University of John Paul II),
Dr. Pawel Stacewicz (Warsaw University of Technology),
Prof. Pawel Polak (Pontifical University of John Paul II)
Phil.in.tech.2022[at]gmail.com

The Ontophobic Turn (or how much philosophical is current philosophy of technology

Agostino Cera

My paper sketches a *historicization of the newest philosophy of technology, starting from the so-called Empirical Turn* (according to Hans Achterhuis' definition). My basic assumption is that the philosophy of technology is living an *epistemic crisis*. This means that many prominent scholars in this field don't believe anymore in both "philosophy" (as a form of knowledge) and "technology" (as something in itself). To avoid being deterministic, the current mainstream in the philosophy of technology has become apologetic, i.e. no more able to be critical.

The hermeneutical hypothesis at the basis of my historicization is that after 40 years *the Empirical Turn proved to be an Ontophobic Turn*. By this expression, I mean an *over-reaction* against the transcendental approach to the question concerning technology, in particular *against Heidegger's* (i.e. continental) *legacy* in the philosophy of technology. This over-reaction consists of the *transition from an over-distance to an over-proximity*: from an almost disinterest in the empirical/ontic dimension of technology (and therefore an over-distance), typical of the first generation, to an almost absolute interest in this empirical/ontic dimension (and therefore an over-proximity) with a consequent a priori disinterest in any ontological/transcendental implication of technology. This disinterest finally takes the form of a taboo, i.e. a real Onto-phobia.

With regard of this situation, my objection is the following. If "Technology" with capital T becomes nothing, i.e. if the philosophy of technology becomes a problem-solving activity in the presence of concrete problems emerging from the single technologies, then it must be admitted that this activity can be performed much better by specialists (scientists, engineers, politicians…) than by philosophers. As a consequence, the ontophobic turn in philosophy of technology *culminates in the* disappearance of the reason for a philosophical approach to the question of technology. Given this assumption, *the paradoxical accomplishment of the empirical turn should be the final self-suppression of the philosophy of technology*.

Abduction Based on Conditionals as a Foundation for Artificial Intelligence

Rolf Pfister

Abduction is considered as an uncertain but very powerful inference because it is the only one that can introduce new concepts. However, abduction is also highly debated. One reason for this is that the term 'abduction' is used for many, quite different theories, e.g. some use it for Peirce's retroduction, while others use it for Lipton's Inference to the Best Explanation. Moreover, some theories interpret abduction as a logical syllogism, while others view it primarily as a computational method or as a process of epistemic change.

To overcome the limitations of current theories of abduction, a new theory is proposed. It aims to cover both the context of discovery, i.e. how a concrete hypothesis is generated, and the context of justification, i.e. how the quality of a hypothesis is rated. Furthermore, it aims to formalise the process of abduction, which would allow its application in computer science and artificial intelligence.

Unlike most other theories on abduction the proposed theory considers abductive inferences not as intrinsically explanatory but as intrinsically conditional: For a given fact, an abductive inference infers a fact that implies it. The implication is represented by a conditional, where the implying fact is the antecedent and the given fact is the consequent. Conditionals are understood in the definition of relevance conditionals, i.e. a conditional is considered true iff there is a relevant connection between the antecedent and the consequent.

There are three types of abduction: Selective abduction selects an already known conditional whose consequent is the given fact and infers that its antecedent is true. Conditional-creative abduction creates a new conditional in which the given fact is the consequent and a defined fact is the antecedent that implies the given fact. Propositional-conditional-creative abduction assumes that the given fact is implied by a hitherto undefined fact and thus creates a new conditional with a new proposition as antecedent.

Different procedures can be used to select or create the propositions and the conditionals. The procedures provide guidance on how to perform a specific abductive inference and are called patterns. A pattern consists of a set of rules for both generating and justifying an abductive conclusion and covers the whole inference process. As generative and justificatory rules are based on different assumptions, e.g. on the principle of causality or on statistical methods, patterns are highly theory-dependent. Consequently, there is an infinite number of patterns that rely on different theories and use different methods. As a result, the various patterns differ in their applicability, efficiency and persuasiveness.

Since the content of abductive inferences can be formally represented in the form of propositions, it follows that an abductive inference is formalisable iff its pattern is formalisable. A pattern is formalisable iff every rule of the pattern, whether it concerns the generation or the justification, is formalisable and the pattern covers the complete inference process.

Based on this, a logic of abduction will be developed that enables probabilistic and non-monotonic reasoning in artificial intelligence. In contrast to older approaches, it allows more powerful abductive reasoning, especially in regard to creative abduction. Furthermore, I investigate why the formalisation of abductive reasoning has mostly failed so far.

The talk is based on an article that is forthcoming in Synthese (Rolf Pfister: Towards a Theory of Abduction Based on Conditionals, https://philsci-archive.pitt.edu/20201/) as well as on more recent work.

On the Controversial Analogy Between Biological and Artificial Cognitive Systems

Arianna Pavone

The analogy between brain and computer has always been a driving force in the field of Artificial Intelligence (AI), as a source of inspiration for the design of any algorithm underlying intelligent systems on the one hand, and as a justification for why some algorithmic technologies were able to achieve similar (or superior) performance to human ones in solving certain tasks on the other. Both of these accounts of the brain-computer analogy, however, seem to be called into question by recent advances of modern AI based on deep learning techniques. On the one hand, it has emerged that the adoption of algorithmic strategies other than those presumed to be adopted in a biological brain has led to wider success, with results that significantly exceed human performance; on the other hand, the analogy between modern artificial neural models and biological ones appears increasingly inconsistent [4], in the physical structure, in the type of abstraction operated on the input data and in the implemented algorithmic solutions, so much so as to argue that the two cognitive models may be incommensurable [2]. In this work we addresses the issue of giving some explanations of the reason for the failure of the analogy between biological cognitive models and the algorithmic logics adopted by modern deep learning systems, embracing the emerging thesis of incommensurability of the two cognitive models proposed by the scholar Beatrice Fazi [2], which is in contrast to the position of many scholars. On the other hand, in this work we also discuss the consequences that the weakening of this analogy has brought in the scenario of AI technologies. Specifically, with the abandonment of the analogy with the brain at all costs in the design of the algorithms underlying a cognitive architecture, we have returned to an opportunistic attitude, whereby the effectiveness of a cognitive model is measured again only based on its performance: if it fulfills the task for which it was designed, then it is a good cognitive architecture, otherwise not. Regarding the second aspect, the need for a plausible explanation of the efficiency of the method seems to be a purely theoretical instance, but there are also more practical implications. For several applications there is a requirement that, since [5] has become known as “Explainable AI” (XAI).

Despite their opacity [1], the analogy with the functioning of a biological brain allowed the algorithms ancestors of modern deep learning systems to be more acceptable, and not fully understanding them could be justified by our limited knowledge of how a biological brain works [4]. As a result of this argument, if we accept the fact that artificial cognitive models are incommensurable compared to the human cognitive model and, consequently, that they are able to develop independent solutions significantly divergent from those adopted by a biological brain, we could be able to exploit them as a source of inspiration to guide human investigation. After all, alternative cognitive models available in nature (implemented in different animal species) have often inspired our technology and science in general. It follows therefore that incommeansurable cognitive models, precisely because they are such, can contribute substantially and productively to science, technology, engineering and mathematics to ”provide value, optimize results and stimulate inspiration” [3].

References

[1] Burrel, J.: How the machine ’thinks’: Understanding opacity in machine
learning algorithms. Big Data & Society, 1–12 (2016)
[2] Fazi, M.B.: Beyond Human: Deep Learning, Explainability and Representation.
Theory, Culture & Society (2020)
[3] Holm, E. A.: In defense of the black box. Science 05 Apr 2019: Vol. 364,
Issue 6435, pp. 26-27 (2019)
[4] Perconti, P., Plebe, A.: Brain Inspiration Is Not Panacea. In Brain- Inspired Cognitive Architectures for Artificial Intelligence - 2 vol. Advances in Intelligent Systems and Computing, vol. 1310, pp.359-364 (2021).
[5] van Lent, M., Fisher, W., Mancuso, M.: An Explainable Artificial Intelligence System for Small-unit Tactical Behavior, in Proc. of the AAAI Conference on Artificial Intelligence, pp. 900–907 (2004)

Turing and the Anatomy of Thinking Machines

Hajo Greif, Adam Kubiak, Paweł Stacewicz

In this paper, we inquire into the role of Turing’s biological thought in the development of his concept of intelligent machinery. We trace the possible relations between his anticipation of connectionism in Turing (1948) on the one hand and his theory of organic pattern formation or morphogenesis (1952) on the other. These works were concerned with distinct fields of inquiry and followed distinct paradigms of biological theory, respectively postulating analogues of Darwinian selection in learning and mathematical laws of form in pattern formation. Still, these strands of Turing’s work are related, first, in terms of being amenable in principle to his (1936) computational method of modelling. Second, they are connected by Turing’s scattered speculations about the possible bearing of learning processes on the anatomy of the brain. We argue that these two theories form an unequal couple that, from different angles and in partial fashion, point towards cognition as an embodied phenomenon while, for reasons inherent to Turing’s computational approach to modelling, not being capable of directly addressing it as such.

Mr. Moneybags is a Robot: Karl Marx and the possibility of artificial capitalists

Jamie Kelly

The news, science fiction, and academic writing are all awash in discussions of robot workers. Most fear them as usurpers, some embrace them as liberators, but there is no shortage of theorizing about the possibility that artificial intelligence might soon replace humans in factories, offices, and drivers' seats around the world. It is otherwise with capitalists: no one seems worried that artificial intelligence will take over the role of investors, owners, or titans of industry. I think that this is a mistake, as there is much to be gained from considering how artificial intelligence can change the way we think about the role of capital in our economy. The rapid emergence of algorithmic trading, the applicability of deep learning to investing, and the possibility of long-term financial planning by artificial intelligence ought to make us sanguine about the possibility that computers might be able to make better investment decisions than humans. And yet, while rumors abound that artificial intelligence is about to replace doctors and lawyers, professors and programmers, no one seems worried about robots replacing capitalists. In this paper, I provide a Marxist account of what it might mean for robots to replace capitalists in our economy, explain why no one seems particularly worried about this possibility, and explore some of the problems and possibilities opened up by the possibility of artificial investors. Throughout, my goal is to explore the some of the neglected social and legal assumptions undergirding contemporary discussions of automation.

Husserl and Philosophy In Technology

Ken Archer

Contemporary ethics discourse is shaped fundamentally by the modern divorce of the humanities, including ethics, from the sciences, both natural and information sciences, and technology. This divorce casts science as instrumental reasoning, isolated from ethical considerations, and technology as the application of such reasoning. The role of ethics, as a result, is conceived as orienting science and technology from the outside.

Advocates for AI ethics, as an example, tend to teach a set of exogenous ethical principles – fairness, accountability, trust, privacy - that statisticians and engineers are expected to apply. The implication of such advocacy is that, whereas AI systems may not be morally neutral, those who build them are applying instrumental, calculative reasoning to build towards a design and must be taught the ethical ramifications of their finished designs. This formalist framing of AI unwittingly adopts a self-understanding of statistics work that AI ethics advocacy should in fact resist. By buying into formalist assumptions around probability, AI ethics perpetuates the self-understanding of technical work within AI as an applied science which allows little room for agency, such that builders of AI systems have agency primarily in the decision whether to build something.

The divorce of ethics and science, a divorce mutually agreed upon by advocates and skeptics of modern technology, thus perpetuates the ethical crises that it seeks to solve, leading to the intractability we experience today. However, this divorce is not actually true to the history or practice of the technologies and sciences, including AI and statistics, posing these ethical challenges.

As Husserl demonstrates, both the humanities and the sciences are grounded in a life-world of intersubjective experience from which the ethical dimension of science and technology emerges. For Husserl, ethics, science and technology are not competing modes of reason, but aspects of the unity of reason when considered phenomenologically. Fundamentally, science is the self-responsible pursuit of practical concerns according to evidence, and thus emerges within a pre-scientific life context.

While the divorce between ethics and science agreed upon by promoters and critics of modern technology perpetuates the ethical crises that it seeks to solve, leading to the intractability we experience today, Husserl calls for a critique and renewal, a desedimentation, of narrow conceptions of scientized technology to uncover their original ethical meaning in primordial experience. Husserl's philosophy thus combines a pessimistic assessment of the spiritual and ethical damage of Western technology, not unlike the assessment of Heidegger, with a more optimistic call for critique and renewal that is itself a continuously necessary renewal of the sedimented lifeworld.

What can a philosopher learn from, or find in computer games? Some remarks on the state of research and possible perspectives

Łukasz Mścisławski

The phenomenon known as computer games can be considered a fully-fledged and highly specific area of philosophical research. Philosophical inquiries cover various areas of research, from the most fundamental ones, connected with methodological solutions, ontology, through cognitive issues, to those connected with the ethical dimension or cultural role of this phenomenon. The present work aims not only at an overview presentation of the state of research, referring to works of Aarseth, Saeng and others, but also at proposing the development of some threads of ontology research, especially with the help of two tools. If we consider computer games as technical projects, with a particular, multi-objective specificity, it turns out that the first tool - Bromm's concept of attributed worlds - can significantly enrich the insight into ontological problems. The second of the proposed tools - Perzanowski's concept of combinatorial ontology - on the other hand, allows us to grasp the richness of the ontological problematic of computer games. Further, more general directions of research, taking into account that of ontological and semantic character of the whole area of numerical simulations, are also briefly sketched.

Engineering knowledge as proper subject of epistemology

Mihály Héder

This presentation aims to add a new angle of investigation to the topic in analytic epistemology known as the *value of knowledge* problem. The question is whether there is any epistemic value beyond truth, taking also into account that truth can be accidental. The investigations always assume natural scientific knowledge as their subject while my work focuses on the knowledge of engineering designs which will be represented as a particular kind of belief sets.

The core of the argument is that these particular belief sets are always produced by a reliable process. In this understanding of the situation, being reliably-formed is not a normative requirement against knowledge (cf. Unger 1968), one that needs further justification, rather, it is the case that all such design belief sets in existence are reliably formed and thus the machine-product (Zagzebski 2003) type objections are idle. In other words, there is an important property in which reliably-formed and non-reliably-formed design belief sets differ, and that is their existence. Therefore we could say that his account of engineering knowledge falls into the category of causal theories of knowledge (Goldman 1967).

Based on the insight that design documents are to be interpreted as sets of beliefs, we will be able to define a specific property that all of these sets of beliefs share even before any justification attempt is made. We call this property validity (different from logical validity), a separate dimension of design-as-belief-sets that is the precondition of both truth and justification. The key move of our argument is that unlike guessing the number of coins in our pocket, it is not possible to accidentally arrive at the complex set of beliefs a non-trivial valid design represents.

References

Goldman, A. I. (1967). A causal theory of knowing. The journal of Philosophy, 64(12), 357-372.
Unger, P. (1968). An analysis of factual knowledge. The Journal of Philosophy, 65(6), 157-170.
Zagzebski, L. (2003). Intellectual motivation and the good of truth, in DePaul & Zagzebski 2003: 135–154.

Philosophy in Technology – Selected Polish Examples

Maksymilian Smolnik

The paper is devoted to the certain issues of the presence of philosophy in engineering and technical sciences. The main aims of the work are as follows:

- to constitute a methodological scheme of relations between philosophy, technical sciences and engineering activities as a framework for the later discussion
- to concisely characterise the examples of the use of philosophical concepts by selected Polish scientists and engineers (e.g. Józef Konieczny, Janusz Dietrych)
- and finally, to suggest valuable fields of future research for the Philosophy in Technology.

When aiming at constituting the Philosophy in Technology approach, a methodological discussion is necessary to clarify the relations between the main concepts used in the research. Therefore, the first main part of the paper includes a presentation of a proposed structure of such relations. The key concepts present here are: philosophy, technical sciences, engineering (i.a. designing technical objects), particular methodologies (e.g. design methodology). They may be defined and understood in different ways but still all of them may be considered as certain areas of human actions. The latter fact allows one to attempt to develop a classification of the objects of those actions, and adopting some classification is necessary to begin an ordered discussion on the examples of the links between philosophy and technology.

The second main part of the paper includes the characteristics of applications of philosophical concepts by the selected Polish scientists and (at the same time) engineers. Józef Konieczny (1936 – 1984) is known as a military specialist, scientist and philosopher, famous for his outstanding concepts in the fields of the theories of: conflict, use and maintenance of technical objects, action and support. He suggested several new areas of scientific and philosophical research intended mostly for military and engineering purposes. Janusz Dietrych (1907 – 2001) is recognised as a mechanical engineer, designer and design methodologist, famous for developing the science of construct strongly related with ethical concepts and moral evaluation of engineering activities.

All the discussed issues deliver new areas of further, advanced research of (at least) philosophical, methodological and historical character, which could be conducted being considered as Philosophy in Technology.

Completeness in Information Systems Ontologies

Timothy Tambassi

According to Cumpa, an uncritical belief which silently encompasses the categorial debate from Plato to contemporary substantialists is that

[1] the completeness of systems of ontological categories should be (somehow) justified.

By "completeness", refers to three different criteria of adequacy, which Thomasson presents as follows:

[2] comprehensiveness, that is, providing, at a high level of abstraction, categories for everything there is (or might be);

[3] exhaustivity, which indicates that whatever there is (or might be) should find its place in one and only one category;

[4] hierarchical organization, according to which no category can be in more than one level in the hierarchy of ontological system.

The notion of completeness is not, however, monopoly of ontological philosophy. Computer science, for example, has internalized and adapted such a notion to the peculiarity of its own area of investigation – an area that, ontologically speaking, can be divided into (at least) two different sub-branches.

[5] The first one deals with problems concerning specification, programming, implementation, verification, and testing of computational systems, which are traditionally seen as composed of two ontologically distinct entities: software and hardware;

[6] The second sub-branch concerns information systems ontologies, conceived as formal representations of conceptualizations aimed at representing and systematizing the contents of a specific domain of interest, and making such contents processable (also) by artificial agents.

The purpose of this talk is to analyze the notion of completeness, by distinguishing different varieties of completeness and by questioning its consistency with the open-world assumption, which formally assumes the incompleteness of conceptualizations on information systems ontologies.

Anarchy in Cyber Age

Lizaveta Mazuryna

This presentation is about the anarchy movement and philosophy in the digital age. It presents crypto-anarchy as the central idea behind the Cypherpunk movement, the group insisting on the radical changes in the political and economic realm realised, primarily, through the tools of cryptography and blockchain technology. The core idea: - "Crypto anarchy is liberating individuals from coercion by their physical neighbors—who cannot know who they are on the Net—and from governments" - this group is aiming to increase personal freedom through administration of technology. NB: this group defines anarchy as the "absence of government", in place of a popular association with "lawlessness, disorder, chaos". Their main idea - "freedom from external coercion". This idea will be realised through Virtual Communities - a new form of "cyberspacial realities", whose "walls" are directly held by crypto. Those are all sorts of social organisations which are characterised by the sense of "belonging" and which, through crytpography, will "the transcend national borders". Blockchain will play an instrumental role in their formation: the idea of global village gets very accute - there, people share a mental, not physical, connection. Digital Money - it acquires a new form of realisation through "untraceable and anonymous" digital cash, whose ultimate form of realization is "Swiss banks in cyberspace". Anonymous systems are considered as a powerful antidote to surveillance - U.S.S.R., Iraq, China are listed as examples of police states. Colonization of Cyberspace: this becomes possible with advancement of World WIde Web to Web 3.0 which, in synchrony with cryptographic tools, could create a safe space for virtual communities to flourish. Though operated by different laws (in jurisdiction too), they will still have a direct connection with the physical space - at least with regard to market relations

What is Philosophy in Technology

Roman Krzanowski, Pawel Polak

The key to the workshop is “in” not “of”. In the workshop we search for the studies of salient philosophical dimensions, or underpinnings, of technology and demonstrating how philosophical insights may shed new light on what technology does or misses. Some examples of PinT are: the concept of knowledge underpinning knowledge engineering, the nature of free will, personhood, and autonomy probed by the autonomous machines, the essence of ethics as interpreted in AI systems, abduction and induction and knowledge assumptions behind natural resources explorations, the nature of human mind underpinning AGI studies, the deep (natural) concept of computing underpinning Turing computing paradigm, philosophy of methodological aspects of technology- these just being few examples of potential topics to be explored by PinT.

We claim that the best technology is created when its philosophical foundations are (critically) realized and the best engineers are usually (knowingly or not) philosophers. But of course, the reverse claim may not hold. We also claim that critical discussion on philosophical underpinnings of technology is necessary to rationally guide the development of technology and to realize the limitations, prospects, dangers and benefice of the technology on which we all depend.

What is philosophy in technology and what this workshop is about? Engineers rarely are academic philosophers, but they all have to use (mostly without any reflection) some philosophical assumptions in their professional activity. This philosophy internal to technology we call Philosophy IN technology (PinT). It is typically beyond the scope of philosophy OF technology (PoT). Philosophy IN technology asks questions about philosophy internal to technology, esp. about deep philosophical underpinnings of technology, about philosophical assumptions, usually implicit, not articulated, hidden from practitioners of technology, yet the assumptions that determine and create the horizon of technological activities. These assumptions often, if not always, determine the direction of the development of technology and by this are responsible, partially or not, by its success or failure. For example, philosophy in technology asks what are assumptions underlying the GOFAI model of AI and how they influenced the development of AI? Philosophy in technology does not ask what is technology qua technology, what is artifact, what is the role of technology in society, what are ethical aspects of technology, and similar questions - the typical problems philosophy OF technology deals with. Philosophy of technology elaborates on technology from the vantage point of epistemology, ontology, methodology, ethics, and axiology. Philosophy IN technology probes deep, usually not explicated philosophical assumptions underlying technology.