



8 Days at ICA 4 First session in Paris

Oraee A.

PUBLICATION DATE

12/7/2021

KEYWORDS

Artificial Intelligence	Data science	Deep learning	Responsability	Algorithms	Language	Robots
-------------------------	--------------	---------------	----------------	------------	----------	--------

ABSTRACT

This 4th edition of the ICA hosted by the Paris Institute for Advanced Studies (Paris IAS) from October 18 to 27, 2021, explores interdisciplinary issues at the intersection of cognitive science, neuroscience and artificial intelligence. Decisive advances over the last decades in the analysis of brain activity and its behavioral counterparts, as well as in information processing complementarities between neuroscience/cognitive science and artificial intelligence allow us to explore synergies and raise between these disciplines, which present considerable challenges and opportunities for the progress of society.

Introduction

The Intercontinental Academia (ICA) creates a global network of future researchers in which some of the very best young academics work together on paradigm-disciplinary research, mentored by eminent researchers from across the globe.

The ICA was established in 2016 through the University-Based Institutes for Artificial Intelligence (UBIAIS) coalition which has 44 member institutes around the world ([Cordelois](#),

During each edition of Intercontinental Academia, participants get together in person over the course of one year ¹.

Previous editions of ICA have focused on "Time", "Human Dignity" and "Laws of Dynamics".

The 4th edition of the ICA explores the complementarities between artificial intelligence, neuroscience and cognitive science and the tremendous challenges and opportunities for humanity. Fellows and mentors initially met online and in cyberspace, and now in Paris, from October 18 to 27. They shall meet again in cyberspace in the next few months and then finally, in Belo Horizonte in Brazil next June.

The first session, hosted by Paris Institute for Advanced Study (Paris IAS), includes 10-days of scientific sessions, discussion forums as well as scientific exchange with the Paris Saclay, Sorbonne Center for Artificial Intelligence and Ecole Normale Supérieure.



Each day at the [Paris IAS](#), [ICA4 Fellows](#) meet with their [Mentors](#) for a closed 30-minute seminar, during which two mentors launch the discussion with a presentation. Upon completion of the seminar, the Fellows then meet for 45 minutes to list the key takeaways and ideas that emerged from the discussion, followed by a collective brainstorming session in which the output of collective intelligence is collected, formatted and capitalised.

The other half of the day is left free for participants to reflect on the scientific topics in small groups. Such discussions are occasionally complemented by lectures from

Day 1: "The future needs wisdom!"

The very first lecture of ICA4 - Session 1 came from [Robert Zatorre](#) who took us into the fascinating world of music while explaining the relationship between perception and pleasure!

This was followed by another lecture that introduced a rather different perspective. It was presented by [Eliezer Rabinovici](#). The lecture mostly explored the complexities of AI enquiries and methods in the context of AI.

Before leaving for cocktails and welcome speeches in the chambers of Paris, a lecture was given by [Helga Nowotny](#), who emphasised an urgent need for a context-sensitive AI control system.

Perception, prediction, and pleasure: What can music teach us about neurocognition/intelligence?

Presented by [Robert Zatorre](#)

It was stated during the seminar that the brain represents the properties of the environment and guides behaviour through evaluation and reward. Aesthetic pleasure can be attributed to a phylogenetically older system that is centred on the striatum.

Moreover, results of the relation between connectivity of the auditory cortex with behaviour and several behavioural results were presented (e.g. related to amusia, anhedonia). Dynamic causal modelling and predictive coding frameworks have been proposed as possible explanations of the relationship between learning and reward in music. The seminar made the rewards evolve from a biological event to the expectation of the event.

Through the post-seminar collective discussions, the relevance of affective states (pleasure and fear) in learning was emphasised. Discussions concluded with an open-ended question, leaving ICA4 Fellows wondering about whether or not AI is a similar system for learning, and how should the reward and punishment be controlled.



Maybe AI does not need to understand or experience human emotions; it just needs to be like a human by capturing the features of a dataset that correctly describes the behavior.

High Energy Physics: Successes, Challenges and Magic

Presented by [*Eliezer Rabinovici*](#)

It was discussed that observing natural phenomena can motivate scientific enquiry and lead us to understand the unknown. Moreover, equations are a way to increase our understanding. However, a single, compact and reductionist explanation for all phenomena in physics may not necessarily exist. The scientific method requires that results are reproducible and the correspondence principle requires that new theories can explain all phenomena that the preceding theory was valid. To understand a phenomenon, one has to identify the relevant players and determine the correct explanation scale.

In AI We Trust: Power, illusion and control of predictive algorithms

Presented by [*Helga Nowotny*](#)

The session began with introducing the concept of singularity and defining its critical point: a change of state that can lead to the collapse of a system. In an attempt to address Ethical AI, examples such as Transhumanism (ideas of transcending the limitations of the human body through information sharing) were discussed. Furthermore, the illusion that AI knows humans better than humans know themselves was elaborated, ultimately concluding that AI is a tool, mentioning the existence of a possibility for human beings to both profit or suffer from the AI system depending on how it is applied.

"The future needs wisdom": an urgent need to institutionalise context-sensitive governance for AI, creating a standardised system to control AI, was discussed and collectively decided. This led to further debates regarding the concentration of technology advancement and its deteriorating impact on inequality. Thus, a global agreement is necessary to ensure that AI, although it is currently almost impossible to obtain! Therefore, we should educate the next generation of child of humanity that can grow to contribute to society. AI research is undertaken on a massive scale that it requires global efforts which go beyond a single country. The research should be paid by society, and their curiosity-led work should return to society as a whole.

Day 2: "In AI we trust"...or not!**



The ICA4 continued onto the second day, through which three seminars to mentors who had joined the first session in Paris from around the world!

The first lecture was by [Robert Aumann](#), a Nobel prize laureate, who focused on the convoluted concept of consciousness and its counterparts.

This was followed by a lecture from [Karen Yeung](#), who offered a rather critical view on the prevalence of AI, as well as some of its surrounding myths and misconceptions. She then went on to explain how responsibility should be re-defined to consider the impact(s) of AI in human societies.

Finally, [Raouf Boucekine](#) took the fellows for an exploration into the world of economics and finance, using the concept of equilibrium as an example to illustrate the differences between disciplines: mainstream economics VS. statistical physics!

- Why Consciousness?

Presented by [Robert Aumann](#)

Essentially, the seminar was focused on the purpose which consciousness serves. Consciousness was defined as the ability to do the following:

- Perceive
- Feel (emotions)
- Think/intend
- Carry out intentions (volition)

Of all the above, perceiving, thinking/intending, and carrying out intentions may be done by machines. However, feelings and emotions belong exclusively to human beings. In this context, it may be argued that the evolutionary function of consciousness is the regulation of emotions. This being said, we currently have no idea about how consciousness works. Although considerable progress has been made in AI, Artificial Emotion (AE) has remained rather untouched.

Myths and misunderstandings about responsibility for the unintended impact of AI

Presented by [Karen Yeung](#)

The talk mostly focused on responsibility for the unintended impact of Artificial Intelligence based on the presenter's Council of Europe study. It was argued that Machine Learning capacity to enable task automation and machine "autonomy" raise important questions about responsibility. Thus, responsibility-relevant attributes of ML were identified,

illustration is the data-driven profiling of individuals, and other ML applications hold adverse impacts on human rights, on both individual and collective levels.

While responsibility is important for human beings, who are considered as moral agents, to maintain peaceful social co-operation within the community, only a few studies have focused on tackling the fundamental role of responsibility for individuals, as well as the role of institutions.

The impacts produced by complex socio-technical systems using ML technologies have generated a range of concerns that fall under the heading of "algorithmic discrimination". While existing laws have an important role to play in ensuring the accountability of algorithmic systems, the implications of these technologies for their interference with human rights need to be studied further. This has been the primary focus of Karen Yeung's work.

In a nutshell, two dimensions of responsibility are required:

- Historic or retrospective responsibility: responsibility for conduct and events that have occurred in the past
- Prospective responsibility: roles and tasks that look to the future

Finally, five common myths and misunderstandings concerning responsibility and the unintended adverse impacts of AI were identified:

- Need for effective and legitimate mechanisms to protect human rights from AI applications
- Identifications of the appropriate responsibility model for allocating, distributing and managing the various threats and risks.
- Responsibility of states to ensure that these policy choices are made in a transparent and democratic manner, in order to effectively protect human rights.
- Need for more interdisciplinary research
- Application of the fundamental principle of reciprocity so as not to allow those who develop and run our advanced digital technologies and systems to increase and exercise their power without responsibility.

Data science and deep learning vs theory: two examples from economics and finance

Presented by [Raouf Boucekine](#)

The session included discussions on Data Science, Machine Learning (ML), and theories in the field of economics and finance that share common disciplinary roots. Examples from macroeconomics, in which characteristics of the underlying mechanisms of complex systems are of great interest, were then discussed in more detail. In this session, a misunderstanding between different disciplines was highlighted: the concept of

of great significance in mainstream macroeconomics, whereas this is not the case in physics (e.g., the "equilibrium" bias outside the econ area, discussed by Boucek (2020)). Finally, the use of various methods and approaches, such as (Dynamic Stochastic General Equilibrium), ABM (Agent-Based Modeling) and Network-Based methods, in the field of macroeconomics were discussed.

Day 3: "What you do FOR people, you do TO people, so do it WITH people!"

Day 3 of the first session of ICA4 continued in Paris IAS, where the Fellows sat for more scientific seminars, followed by discussions and brainstorming sessions.

The day kicked off by a framework proposed by [Saadi Lahlou](#), called "Installation", which enables scientists to analyse and regulate human behaviour. This was combined with a new technique to capture the subjective perception of action, ultimately bringing psychological and behavioural sciences one step closer to what was once considered a technically impossible task: introspection!

[William Hopkins](#) then joined the discussions with some stimulating videos from experiments on apes, while exploring self-recognition and social cognition in animals.

Finally, Toshio Fukuda revealed the Moonshot project: a society where humans and machines live together in 2050 !

Distributed Intelligence & Distributed Agency

Presented by [Saadi Lahlou](#)

We want intelligence to perform relevantly adapted actions that change the situation for the better. To design intelligence, we must first understand the nature of human activity. In this sense, the behaviour was defined as what people do, seen from the outside. In other words, behaviour remains an external description of objective phenomena. Intelligence activity is how people subjectively perceive their action and how they see it from their own perspective.

Installations consist of components that simultaneously support and control. In specific, local, and societal settings where humans are expected to behave. e.g., airport, metro, cash machine, etc. Installations consist of three layers: affordances, physical environment, embodied competencies and social regulations. Intelligence is distributed over these three layers.

The question now is: why do we have these installations? Because installations of our behaviours and consequently make us very efficient, although our short cognitive processing are very limited compared to animals. Installations are a redundancy produces resilience and learning.

Moreover, certain questions on designing trade-off issues were raised: which whom? AI agent? what kind of competence for the AI? What affordances? What degree of awareness? To whom does the agent report? How is it evaluated? He is a “privacy dilemma.” In other words, for better service, one must disclose information an “agency dilemma”? Can we make it explicit? Because the agency is distributed responsibility is shared. It means that we now have the “many hands” problem. Credentials for AI were suggested, which include values (what does it try to achieve), ownership (who takes responsibility for its actions), principles of action (rule domain of awareness and action), track record (list of transactions executed, its training).

To conclude, ICA4 Fellows were left with some questions as food for thought: activity, do we want to augment existing agents with more agency? If so, Humans? Material objects? Social system? New agents? Who learns what? What want to foster? What do we want AI for must be addressed for each activity, state activity and discussed?

Perspective on Artificial Intelligence research from studies on Agency, self-recognition and social cognition in animals

Presented by [*William Hopkins*](#)

The session began by discussing humans constructed concepts to reflect intellect across various domains of cognitive functions. In this sense, we use tests like the WAIS-Binet to quantify and scale performance to standards for specific age classes. They heavily on language. There are many approaches to developing fair tests of cognitive species with different sensory and motor capabilities. It began with Darwin and I George Romanes (1884) focused on animal intelligence and later on, Kohler (1913) learning. Within the same field, Robert Yerkes (1916) worked on “The mental life of apes: a study of ideational behavior”. Yerkes later developed the IQ test used in WW1 (army alpha test).

Upon drawing on some of the literature, several videos were played, in which animals performed various tasks including retrieving a peanut in the bottom of a tube followed by an ape imitating a human, and so on. Several animals passed the mark test. E.g., a yellow stick in their neck can identify it and try to remove using the mark. The parcellation of chimpanzee brain - compared to humans, the ones that passed the test showed differences in some cortices. Grey matter differences between MSR+ vs. MSR-

analysed the anterior cingulate since such neurones are rather long and connect the cingulate with the insula.

Moreover, results from studies that showed that human children outperform on social, but not physical, cognition tasks were presented and discussed. Much like in AI, Most early comparative studies of cognition and intelligence were strong on associative learning theory. However, associative or operant theories of learning are notoriously anti-cognitive. In the 1960s, there were attempts to reach artificial communication systems. The goal of the ape language studies was to determine if language is uniquely human. The answer depends on how we define language.

However, is it language? There is very little evidence for declarative production (e.g., the TV, give me an onion) in communication signals by primates and other animals. A question is: are social stimuli rewarding? For chimpanzees, yes. Experiment: touch a button to see other chimpanzees or another button to see random animals. The chimpanzee touches the button to see other chimps. Thus, the role of reward guides the learning in animals. Although animal cognition is often used to explain animal behaviour, it is explained by an associative learning mechanism.

AI and Robots for Future: The Moon Shot Project

Presented by [Toshio Fukuda](#)

Robots are avatars that pop up to help when humans need them. There is an increasing physical interaction between robots and humans. Toshio showed several multi-robot systems, e.g., monkey-type robots, multi-locomotion, intelligent cane, etc. One of these is Brachiator I-III. Brachiation is a form of long-armed ape locomotion. It uses a double pendulum, under-actuated mechanical system, variable constraint system, machine learning, AI, reinforce learning, soft computing (fuzzy, genetic algorithm). Regarding multiple types, in many cases, one creature has multiple types of locomotion in order to increase mobility. The motivation of their study is to develop a robot mechanism architecture that can achieve multiple locomotions. Hybrid computational intelligence and brain interface were also commented upon by the speaker while showing related videos. An example of such videos illustrated the Boston dynamics with three robots dancing and jumping which was quite impressive!

Moreover, AI+Robot+IoT (Internet of Things), the use of robots in mega-urbanisation, food, ageing, global warming, robot, and AI) were discussed. This was followed by further discussions on autonomous cars, which may be safer than human driving. In reality, the Eve project (a transparent body that simulates the human body), cyborg (fusion of robot and animal), and multiple robots (communication among robots)

Day 4: A visit to The University of Paris-Saclay

![]

(https://lh6.googleusercontent.com/NlgQpGHaBLRystLQsRXAbmAklSDIPADYo2zzpizb3wV2JZt231Ei1_LKcE10ENUndEKH6G8APDMxkjIHUFiQ_eN9eDjj-Pw6qEoONXtNSNyGxelsEsn6APg =602x451)

The Fellows embarked on their first scientific trip for ICA4 and were hosted at **Normale Supérieure de Paris-Saclay** throughout Day 4. The sessions at Saclay featured thought-provoking talks by [Xiao-Jing Wang](#) and [Jay McClelland](#), both of which explored the principles underlying cognitive behaviours, as well as the difference between machine intelligence. These were followed by a symposium on AI at University of Paris-Saclay. The symposium was followed by a half-day event with multiple workshops in which mentors discussed major advances and issues surrounding AI with other researchers such as [Stanislas Dehaene](#). Finally, the intellectually intense day concluded with a talk in which [Zaven Paré](#) raised important questions regarding how we interact with AI algorithms and intelligent robotics in the decades to come...

Efforts to understand the computational principles underlying cognition

Presented by [Xiao-Jing Wang](#)

Deep neural networks, despite their recent success, differ from human cognition in that they have no internal mental life - instead, they act as complex, nonlinear input-output systems. In humans, the prefrontal cortex (PFC) is known to be crucial for cognitive functions such as working memory, decision making, and executive function. An early avenue of research involved understanding how persistent neural activity may underlie working memory. This persistence is linked to recurrent connectivity, which is lacking in most deep neural networks. I will describe his previous research using spiking networks and tools from dynamical systems to understand the attractor dynamics behind this form of memory. In the second half of the talk, he showcased his more recent work which uses recurrent neural networks (RNNs) as a model organism to probe how the PFC may perform multiple cognitive tasks simultaneously. These RNNs can then be used to address questions such as what neural codes encode cognitive building blocks in a compositional manner, similar to the concepts of schema.

A different distinction between human intelligence and machine intelligence



Presented by [James McClelland](#)

While the latter (in particular machine learning algorithms) learns from statistics input data, humans learn to learn from explanations structured by culturally induced. Indeed, humans fail to perform in systematic ways, which we would expect if they were built into our cognitive functionality. But, McClelland points out that simple structure, as proposed by the pioneers of GOFAL, limits flexibility. This structure, as argued, is built by culture. For example, he described a classic study by Scribner (1973) which showed that non-Western cultures often lack a concept of absolute size and tend to classify objects based on concrete situations rather than abstract categories. These authors proposed that Western education creates a context in which complex relational concepts are learned, consistent with McClelland's later work correlating puzzle performance to mathematical education level. McClelland closed by reiterating that machines learn by examples but humans learn by explanations and that his explanation-based learning (rather than built-in structure) may underlie our propensity for one-shot learning.

Upon completion of the talks by ICA4 Mentors, Paris-Saclay hosted a half-day of multiple workshops in which ICA4 mentors and Paris-Saclay researchers discussed advances and issues surrounding AI. [Stanislas Dehaene](#) presented a series of fMRI behavioural evidence that humans use symbolic and recursive strategies on problems with complex sequences, as compared with monkeys which seem to use a simple strategy. In a session focusing on AI and ethics, [Paola Tubaro](#) revealed the lives of workers who provide the hand-labelled training data for products such as Siri. She discussed companies and corporations needing a cheap workforce in the same language reproducing historic colonial patterns.

Finally, the intellectually intense day came to an end with a talk in which Paré discussed his artistic works based on electronic marionettes and his collaborations with robotics specialists in Japan. Paré's conception of automaton-centred theatre challenges the audiences while challenging our tendency towards anthropomorphisation. This raises questions regarding how we will interact with AI algorithms and intelligent robots in the decades to come...

Bibliography

Cordelois, A. (2020). How hard can it be to make an ideal publication website? *Nature*, 222.



Footnotes

1: This is the footnote everyone always dreamed about. It is lengthy yet not that plain good.
