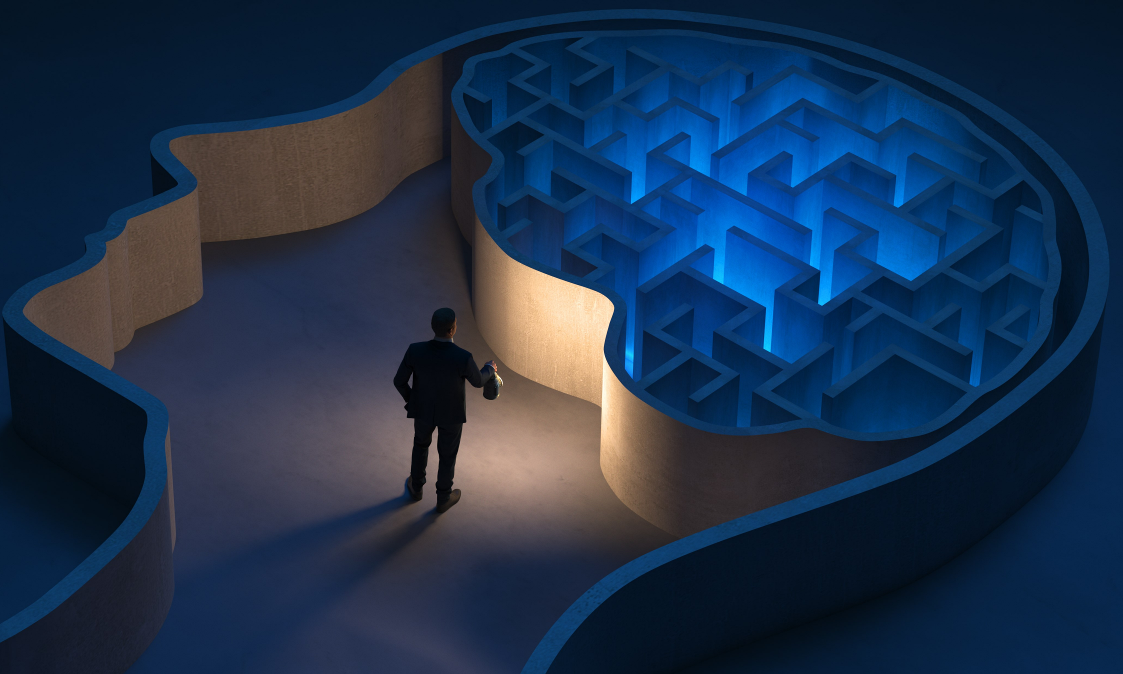
Causal Cognition in Humans and Machines

- University of Oxford, 11-12 January 2024 -



PROGRAMME

Table of Contents

Welcome

It is a great pleasure to welcome you all to the second instalment of Causal Cognition in Humans and Machines! Over recent years, a significant convergence has emerged among research in psychology, computer science, and artificial intelligence (AI), focused on developing machines capable of emulating many aspects of human cognition and reasoning abilities. Central to these advances is the investigation of causal thinking, an essential facet of human intelligence. A thorough understanding of causality is indispensable for making informed decisions and assuming responsibilities in both artificial and real-world contexts.

This conference aims to bring together researchers, engineers, and students from various disciplines to engage in comprehensive discussions on causal intelligence. The conference will be highly interactive, with dedicated time for discussion panels, aiming to stimulate an interdisciplinary exploration of causality across psychology, computer science, and AI.

By kind invitation of the Department of Computer Science, the conference takes place at the University of Oxford, UK.

Welcome to Oxford!

Website and Google Map links at: https://amcs-community.org/events/causal-cognition-humans-machines/

The Conference Team

Organisers

Dr Selma Dündar–Coecke
Centre for Educational Neuroscience
& Quantum Brain Art
UK

Dr Sean Tull
Quantinuum
UK

Dr Johannes Kleiner
Ludwig Maximilian University of Munich
& University of Bamberg
Germany

Prof Aleks Kissinger
University of Oxford
UK

Scientific Committee

Andreas Demetriou
Aristotle University of Thessaloniki

Steven Sloman
Brown University

Prof York Hagmayer
Gottingen University

David Lagnado
University College London

Jules Hedges
Strathclyde University

Bob Coecke
Quantinuum

Robin Lorenz
Quantinuum

John Barrett
University of Oxford

Sieglinde Pfaendler
IBM

Martha Lewis
Bristol University

Conference Dinner

We will be delighted to welcome all attendees to the conference dinner on

Thursday, January 11, 2024 at 7pm

at the

Cherwell Boathouse.

Good food, great company, and unforgettable memories await. This event is sponsored and includes a buffet and two glasses of wine or an equivalent drink per person.

Please note that RSVP was required to attend this event.

Address: Cherwell Boathouse, Bardwell Rd, Oxford OX2 6ST
Google Maps: Link on the conference website.

Cognitive Jam Session

Join us for a Music event on **Friday night** at the **upstairs venue of Jericho Tavern**, featuring DJ Rave's Jam Session from 6pm to 11pm.

Feel free to bring your own instruments. We welcome musicians of all levels to participate and make this music event a memorable experience!

Address: 56 Walton Street, Oxford, OX2 6AE
Google Maps: Link on the conference website.

	Thursday
8:45 - 9:15	Arrival & Registration
9:15 - 9:30	Welcome
9:30 - 10:30	Invited Talk **1 Steven Sloman**
10:30 - 11:00	Coffee
11:00 - 12:00	Invited Talk **2 Samantha Kleinberg**
12:00 - 12:30	**3 Stefan Lejinen**
12:30 - 13:00	**4 Tomas Veloz & Olga Sobetska**
13:00 - 14:00	**Lunch** (provided) **& Posters** **P1 Burcu Ünlütabak** — **P2 Zhouwanyue (Nata) Yang** **P3 Kai Holland** — **P4 Joshua Tan**
14:00 - 15:00	Invited Talk **5 Yoshua Bengio**
15:00 - 15:30	Coffee
15:30 - 16:30	Invited Talk **6 Francesca Raimondi**
16:30 - 16:45	**7 Johannes Kleiner**
16:45 - 17:00	Break
17:00 - 18:00	**Panel Discussion**
19:00	**Conference Dinner (Cherwell Boat house)**

	Friday
9:00 - 9:30	Arrival
9:30 - 10:30	Invited Talk **8 Ruth Byrne**
10:30 - 11:00	Coffee
11:00 - 12:00	Invited Talk **9 Michael Thomas**
12:00 - 12:15	**10 Selma Dündar–Coecke**
12:15 - 12:30	**11 Sean Tull**
12:30 - 13:00	**12 Gregoire Sergeant-Perhuis**
13:00 - 14:00	**Lunch** (provided) **& Posters** **P5 Dominik Luke — P6 Uziel Awret — P7 Michael A.Popov** **P8 Oleksandra Halchenko — P9 Tianshu Chen**
14:00 - 15:00	Invited Talk **13 Bob Coecke**
15:00 - 15:30	Coffee
15:30 - 16:00	**14 Bart Jacobs**
16:00 - 16:30	**15 Jules Hedges**
16:30 - 17:00	**16 Deanna Kuhn**
17:00 - 17:15	Break
17:15 - 18:15	**Open Discussion**
19:00	**Music event (Jericho Tavern)**

Talk Titles and Abstracts

1 Steven Sloman (Invited)

Limited Human Causal Knowledge is About Interventions

People tend to think they have more expertise than they do. In social domains, this is in part because they confuse their understanding and their values. Hyoseok Kim and I show in multiple experiments in value-laden domains that people's sense of understanding increases when they are induced to frame issues in terms of sacred values rather than in terms of causal consequences. Sacred values frames also tend to increase people's willingness to take action but also their intransigence. In separate work with Samir Tatlidil and Iris Bahar on human communication in service of robot planning, we show that eliciting causal models of an object facilitates machine construction of a plan to assemble the object. However, an elicitation procedure that requires explicit articulation of the causal model just induces confusion. A more effective procedure asks for judgments of the effectiveness of interventions. People are able to make local judgments about whether intervening on an element will affect a second element. Reasonable causal models can be inferred from a set of such pairwise judgments.

2 Samantha Kleinberg (Invited)

Learning Causes and Using Them

The collection of massive observational datasets has led to unprecedented opportunities for causal inference, such as using electronic health records to identify risk factors for disease. New causal inference methods enable us to learn highly complex models from these datasets, but what happens when people attempt to use them? In this talk I discuss new methods that allow causal relationships to be reliably inferred from complex observational data, work on understanding how people use these models for decision-making, and what it means for inferring causal models that are useful and usable.

3 Stefan Lejinen

Causality beyond Shannon

Shannon's information theory forms the foundation for many of today's models of computation, including deep learning e.g. in formulating loss functions and model regularization. Information theory quantifies the amount of information in a probability distribution by assessing its unpredictability, bracketing out the causality of the process that underlies the creation of the distribution.

This bracketing out is generally considered a feature of computation for pragmatic reasons (i.e. making information quantifiable) but can also lead to causal misalignment of computational models with their context, e.g. making them difficult to interpret. Substrate independent information processing where the underlying thermodynamics and synchronicity are abstracted away leads to technology that is no longer grounded in reality beyond the efficient but thin causal link with transistors and storage of electrons in a memory.

We turn to Aristotle's four aspects of causality (material, formal, efficient, and final) to offer a framework to machine learning for grounding causation in thermodynamics and synchronicity. Final causation, the purpose or end goal of AI systems, is perhaps the most elusive of these four and is often overlooked in modern AI systems as it is codified simply as a symbolic objective. This makes encoding ethical and long-term goal-oriented behaviors in AI challenging, where we would like to secure some level of human control on AI. Bridging philosophy with technology this talk aims to shed light on the deeper understanding of causality in AI, potentially guiding future developments towards more holistic and ethically grounded AI systems.

4 Tomas Veloz & Olga Sobetska

Bridging between Quantum and Living Systems towards Human Cognition and AI

This talk aims to explore groundbreaking connections between two seemingly unrelated fields in the foundations of cognition for both machines and humans.

The first field is quantum cognition, which involves using mathematical structures from quantum theory, such as Hilbert spaces and projector operators, to model aspects of human semantic processing that are commonly understood as non-rational. This has been applied to formulate causal modelling for tasks such as categorization, decision-making, and language understanding.

The second field is the study of the emergence of cognition in living systems, which focuses on identifying structures, at the system biological level represented by biochemical reaction networks, that can implement dynamic mechanisms resembling symbolic-processing-like causal cognition. This includes perception, selective response, memory, and anticipation.

While the first field is primarily concerned with modelling meaning and can be applied to provide causal descriptions of data patterns and prediction, the second field seeks to uncover architectures that demonstrate how the open-ended evolutionary process of life can scaffold structures where symbolic processing and cognition occurs.

In this talk, we will provide context for these two perspectives within the ongoing debates about the alignment problem and the future of AI. In particular, we will focus on how these frameworks can be synergistically combined to develop a framework towards the emergence of objective functions in complex environments.

5 Yoshua Bengio (Invited)

Amortized Inference of Bayesian Causal Models for Quantitative AI Safety Guarantees

A necessary requirement for AI alignment is a good world model, including of human moral judgements. To generalize well out-of-distribution, it should be causal. To avoid confidently wrong decisions, it should be Bayesian. Inference in this model would not only capture epistemic uncertainty but also enable reasoning over latent causal variables, making it possible to benefit from system 2 inductive biases. How can we achieve this in a computationally tractable way? We propose to start from the recent advances in amortized variational inference using generative flow networks (GFlowNets), including applications to Bayesian inference, with the property that as the neural network training loss improves, the Bayesian approximation will become better and safety guarantees will improve.

6 Francesca Raimondi (Invited)

Causal AI for Actionable Decision Making

Causal Discovery Algorithms: We employ advanced, automated, and human-guided algorithms that integrate expert domain knowledge. This approach elucidates cause-and-effect relationships among variables, fostering a collaborative environment between the domain experts and the data scientists.

Causal AI models: Causal AI is distinguished by its inherent explainability, reduced model bias, and improved model generalization. I will discuss some pitfalls of post-hoc explainability as it is commonly performed, as well as causal regularization, which can help battle the underspecification of models.

Industry Applications: I will delve into various industry use cases, demonstrating how we identify causal models, estimate causal effects, optimise interventions, deliver actionable recommendations in alignment with business constraints and objectives, and evaluate algorithmic fairness through the lens of causality.

Integration with LLMs: A novel aspect of our approach involves leveraging Large Language Models (LLMs) to augment Causal AI. This integration introduces global domain knowledge, enhancing causal discovery, modeling, and reasoning through prompt engineering.

Decision Operations: I will discuss how, in a practical setting, we action, trace and attribute the outcomes of decisions in the presence of confounders.

7 Johannes Kleiner

Consciousness in Causal Cognition

Consciousness is sometimes viewed as a by-product or aftermath of cognition. Several contemporary theories of consciousness, however, suggest that consciousness and cognition are deeply entwined. The goal of this short presentation is to illustrate how consciousness may be relevant for causal cognition.

8 Ruth Byrne (Invited)

How Counterfactual Explanations Affect People's Understanding of an AI System's Decisions

People readily create counterfactual explanations about how an outcome could have turned out differently, if some aspects of the situation had been different. Their counterfactual explanations differ in important ways from causal explanations, and I discuss some implications of these differences for eXplainable AI (XAI). I consider three recent discoveries: first I outline recent experimental results about the effects of counterfactual explanations of an AI system's decisions on people's choices of certain or risky options. Next, I assess findings of differences in subjective measures of people's preferences for counterfactual and causal explanations of an AI system's decisions, and objective measures of the effects of these explanations on the accuracy of their understanding. Third, I sketch current empirical discoveries of differences in preferences for simple causal explanations for predictive inferences compared to diagnostic inferences. I suggest that experimental evidence from cognitive science can enrich future developments of psychologically plausible automated explanations in XAI.

9 Michael Thomas (Invited)

Educational Neuroscience – Teaching the Brain to Understand Causality

Young children gain an understanding of how the world works based on their experiences, furnishing them with intuitive knowledge, such as night comes when the sun drops below the horizon. When they reach school, they are taught scientific knowledge about how the world works that may conflict with those intuitions: night actually falls when the horizon rises to cover the sun due to the rotation of the Earth. In this talk, I will discuss the approach of educational neuroscience, which seeks to translate insights into how the brain learns to improve educational outcomes. I'll discuss the work of the Centre for Educational Neuroscience in the area of learning counterintuitive concepts in science and mathematics, including its development and testing of a new primary school learning activity based on neuroscience principles. The key principle is that intuitive concepts must be inhibited via content-specific brain circuits to enable access to counterintuitive concepts acquired in formal educational settings, and these circuits are trainable. I will use this case study to outline the broader approach of educational neuroscience – that our theories of how the mind works (and learns) need to be constrained by how the brain works; and I'll discuss how the computational primitives of the brain sometimes overlap but frequently differ from those used in machine learning approaches.

10 Selma Dündar–Coecke

To What Extent is General Intelligence Relevant to Causal Reasoning? A Developmental Study

To what extent general intelligence mechanisms are associated with causal thinking is unclear. There has been little work done experimentally to determine which developing cognitive capacities help to integrate causal knowledge into explicit systems. To investigate this neglected aspect of development, 138 children aged 5–11 studying at mainstream primary schools completed a battery of three intelligence tests: one investigating verbal ability (WASI vocabulary), another looking at verbal analogical (Verbal Analogies subset of the WRIT), and a third assessing non-verbal/fluid reasoning (WASI block design). Children were also interviewed over the course of three causal tasks (sinking, absorption, and solution), with the results showing that the developmental paths exhibited uneven profiles across the three causal phenomena. The confirmatory factor analyses suggested that the impact of cognitive ability factor in explicitly identifying causal relations was large. The proportion of the direct effect of general intelligence was 66%. Of this, 37% was the indirect effect of age. Nonverbal (fluid intelligence) ability explained a further 28% of the variance, playing a unique role in causal thinking. The results suggested that, overall, cognitive abilities are substantially related to causal reasoning, but not entirely due to developmental differences in "g" during the age periods studied.

11 Sean Tull

Causal Cognition via String Diagrams

I will give an overview of how the language of string diagrams can be used to provide a fully graphical treatment of some of the key formal notions used in causal cognition. Primarily I will focus on how one may formally describe the entire Pearlian causal model framework in terms of string diagrams, including causal models, interventions and counterfactuals. This graphical approach is arguably more natural and intuitive than the traditional treatment via probability theory and DAGs, and helps to clarify various causal concepts. This is joint work with Robin Lorenz which builds on several previous works, notably by Fong and Jacobs, Kissinger and Zanasi. Finally I will also touch on how the same language can be applied to treat cognitive frameworks such as predictive processing and Active Inference (as developed by Friston et al.) in joint work with Johannes Kleiner and Toby St Clere Smithe.

12 Gregoire Sergeant-Perhuis

Challenges and Prospects for the Projective Consciousness Model

In order for agents to make decisions based on their noisy sensory observations, they must create a model of their environment using an internal state-space of extracted features. The innovative concept we propose is that this space possesses a geometric structure that allows features to be integrated as a cohesive whole. We achieve this by encoding various perspectives that the agent can have on its environment directly into the geometry of its 'state space.' Each perspective corresponds to a 'frame' that helps the agent orient itself within its environment. One mathematical construction that encapsulates this idea is the notion of a G-space: a (topological) space X with a group G acting (continuously) on X. In this framework, the agent's state-space encompasses all possible frames the agent can adopt, and each change of frame is associated with an element of the group. This G-space serves as a global workspace where features are organized; perspectives can be those of other potential agents or those induced by the agent's own actions, such as its movements. It has been proposed that such structure models important phenomenological aspects of consciousness [1,2,3,4]. We will discuss how to integrate causal modeling and group-structured world models and hint at how this integration can lead to improved predictions by the agent.

13 Bob Coecke (Invited)

Quantum Causality in Pictures: Everyone can do it!

We report on a recent experiment where we showed that fairly randomly selected secondary school pupils can perform exceptionally at an Oxford University post-Grad exam when the language of Quantum Picturalism [1] is used instead of the usual quantum formalism. Next we show how cognitive models can be extracted from this new quantum formalism, which meanwhile have been executed on quantum hardware, and form a pathway to interpretable AI. Of particular interest here is linguistic semantics in terms of circuits [2].

[1] Coecke & Gogioso (2023) Quantum in Pictures. Quantinuum.

[2] Coecke & Wang-Mascianica & Liu (2023) Our quest for finding the universality of language. Medium blog.

14 Bart Jacobs

How Does the Human Update? Via Pearl or via Jeffrey?

There are two fundamental update mechanisms which are poorly distinguished in the literature. They can be associated with Pearl and with Jeffrey. Their outcomes can differ wildly. This is somewhat uncomfortable, because it is not so clear when one should apply which rules.

Informally, one can say that updating according to Pearl happens via encouragement. One can mathematically show that a validity increases. Jeffrey corresponds to discouragement, in the sense that a mismatch is decreased. This corresponds to the approach in Predictive Coding theory, where an abstact aim is to descrease Kullback-Leibler divergence. The latter also takes the form of a mathetical result. It will be shown that this result is at the heart of Predictive Coding. A big question is: which update mechanism best fits human updating, under which circumstances. I don't have the background to answer this question, but the meetingforms an ideal environment to present this question.

15 Jules Hedges

Passive Inference is Compositional, Active Inference is Emergent

It was previously proven by Toby Smithe that "Bayesian inverses compose by a chain rule". That is, if two conditional probability distributions are composed (by integrating out the middle variable), the Bayesian inversion of the composite can be computed from the individual inverses. The formula for this is reminiscent of the chain rule from calculus, and is structurally identical to the backpropagation algorithm for neural networks. This common structure (which also appears in game theory and some other contexts) is captured by a category-theoretic gadget called "lenses". In ongoing work in progress with Toby, we extend this from exact to approximate Bayesian inference. If we replace the exact inverses with a functional form and learn the parameters over a time series of observations, by a dynamic programming argument the composite will still converge to the exact inverse - even though some loss functions of interest, such as variational free energy, are themselves not exactly compositional. This is "passive inference". However, if the forward "prediction" kernel is also being simultaneously learned, compositionality fails and gives way to emergent behaviour. We would like to see this as a precise boundary between unintelligent and intelligent systems.

16 Deanna Kuhn

Choosing Your Evidence: Freeing but Fraught

Participants in causal reasoning studies typically are asked to examine information and determine what causal conclusions are justified. In life outside research settings, the sequence is most often the reverse: We hear a dubious claim and ask, “How can they say that; what’s the evidence?” Or we have our own controversial claim to make and ask ourselves, “What evidence can I use to show this?”

We have examined this latter scenario and unsurprisingly find it to be fraught with error. Despite the risks, it affords adolescents and young adults we’ve worked with the challenge that self-directed thinking offers. Their objective is to demonstrate that a particular causal claim is correct, drawing as they see fit on a multivariable body of data to freely explore and make use of. Each data instance consists of a particular constellation of levels of several independent variables and an outcome. Participants use any method of analysis they wish and call on as many data pieces as they believe needed to demonstrate the correctness of the claim – no prize is put on efficiency.

For any claim, the data available always allowed for the long-held gold standard of a controlled comparison, i.e., a pair of instances differing only with respect to a single factor. Yet only a small minority of participants ever made such a comparison and never consistently. Instead, they identified favorable outcomes and implicated as causal the variable levels associated with them. This was done selectively, with those relations identified being ones for which they had an explanation of the mechanism connecting antecedent and outcome This has led us to a line research to further investigate how evidence and explanation interact in causal reasoning.

Poster Presentations

P1 Burcu Ünlütabak

Theory of Mind Performance of Large Language Models: A Comparative Analysis of Turkish and English

Theory of mind (ToM), understanding others' mental states, such as beliefs and emotions, is a defining skill belonging to humans. Research assessing LLMs' performance in ToM scenarios yields conflicting findings and leads to discussions about whether and how they could show ToM understanding. Psychological experiments indicate that language and the ToM are intricately connected, and the characteristics of a specific language can influence how mental states are represented and communicated. Thus, it is reasonable to expect that the characteristics of a model language could influence how LLMs communicate with humans, especially when the conversation involves references to mental states. This study aims to examine the performance of ChatGPT in the ToM tasks in English and Turkish. English has Subject-Verb-Object order, uses sentence completion with relative clauses, and has mental state verbs like think and believe. In contrast, Turkish is a difficult-to-model language with Subject-Object-Verb order, context-based agglutinating structure, and a complex verb system with tense, aspect, and modality inflections. Besides, it has a special verb san-, meaning "falsely believe," that could influence ToM performance. In my presentation, I will show a set of first-order and second-order ToM prompts in English and Turkish using new scenarios and narratives based on standardized tasks and collected responses for each prompt using Open AI's API (30 trials per prompt). I will categorize responses for (i) accuracy (ii) linguistic structure and (iii) answer quality. I will then present the statistical comparison of the accuracy scores and frequency of response categories across languages.

P2 Zhouwanyue (Nata) Yang

A Comparison of AI and Human Cognition

In this presentation, an approach to comparing human and AI cognition will be present, by integrating observations of the constitutive processes of understanding with traditional Kantian epistemology. This method involves two key steps: first, it assumes that cognition emerges from observable processes inherent to the understanding, which can be reported to and analyzed by epistemologists; second, it assumes that the justification of a cognitive process is derived from the cognitive faculties used by the understanding's possessor. A case study focusing on mathematical practices will be given, specifically how to calculate the product of large numbers. Diverse methods employed by human agents will be analyzed in terms of cognitive faculties, actions, and objects involved in their cognitive processes. According to our first assumption, these human cognitive process of calculation can be observed and differentiated as distinct approaches to the multiplication of large numbers. Based on our second assumption, we explain these methods through the cognitive faculties utilized by different human agents. Subsequently, we predict and test how AI agents might generate results for large number multiplication, drawing comparisons with human cognitive processes. Although the use of cognitive faculties itself cannot be observed, its consequences can be observed as AI agents (based on the same models) can report their understanding of test questions. Ideally, this approach might enhance our understanding of the similarities and differences in human and AI cognition.

P3 Kai Holland

The Conscious Experience of the Observer

A model of consciousness is a theoretical description that relates brain properties of consciousness (e.g., fast irregular electrical activity, widespread brain activation) to phenomenal properties of consciousness (e.g., Qualia, a first-person perspective, the unity of a conscious scene). Because of the diverse nature of these properties (Seth et al. 2005) useful models can be either mathematical/logical or verbal/conceptual. (Scholarpedia, n.d) In relation to this explanation for “useful models of consciousness” this dissertation will aim to explain the creative development of a prototype visual scientific model of the conscious experience of the observer and suggest its relevance and then possible usefulness in explaining and making tangible this abstract phenomenon of consciousness (Scholarpedia, n.d.).

The “Conscious experience of the observer” is suggested to be so defined as the entirety of any given human beings' life, from birth to death, it is defined in an attempt to quantify the physics principles at play in facilitating that conscious experience of life as an observer for any given one of us between these key points in our human conscious experience and our timeline of life in this universe, as a product of the universe on Planet Earth.

Through the use of visual problem solving, creative thinking, visual communication practice, using basic diagrammatic graphic design tools and a vast literature review process, it will then be shown that a visual model of the conscious experience of the observer could potentially be useful in creating an informed visual hypothesis for the nature of consciousness itself. It will be shown how the creative and collaborative creative process was fundamental in attaining what is possibly now a useful visual, conceptual, higher-order theory of consciousness as attention to memory which operates primarily on understandings derived directly from the Penrose interpretation and other competing models of consciousness and their predictions and models of understanding. (Scholarpedia, n.d). therefore, it is wise to have an understanding of models of consciousness and the topics of discussion before engaging in this reading. It will also be shown how the hypothesis for applying understandings through quantum mechanics, relativity, the Penrose interpretation of consciousness and the new definition of Endoquantum Mechanics have come to form what can only be described as a potential unified field theory of everything we could ever come to know and remember throughout our conscious experience as observers (Kaku, 2020) Further it will be shown how this would not have been possible without the creative thinking, collaborative and visual problem solving processes through the development of basic visual and logical scientific diagrams through simple graphic design tools.

P4 Joshua Tan

Isomorphism and Bijection between Domains

GAN dissection (Bau et. al. 2018), is a technique for interpreting the behavior of neural networks, especially generative networks, through the detection of "causal units", i.e. single neurons or sets of neurons in a network that control for the appearance of a particular concept in the generated result. In this talk, we present preliminary work that uses this notion of causal units to define a coherent notion of isomorphism between domains (e.g. image domains or natural language domains), based on the structure of possible interventions that preserve a given concept. We then compare this isomorphism to information-theoretic measures, e.g. arising from mutual information / conditional entropy or from approximate Bayesian inference, that generally measure when a map between domains is bijective. We conclude by discussing applications within computational learning theory, specifically for understanding the limitations of techniques for style transfer between domains such as cycle consistency, as well as broader implications for causal interpretations of generative networks.

P5 Dominik Luke

Weird Semantics of Token Embeddings: Or what are Large Language Models actually Models of

The aim of this paper is not to answer the question but to offer a richer and more complex definition of the problem of representation "inside" LLMs. I will argue that all current methods designed to uncover the model of the world learned and used by an LLM, which while incredibly valuable, are destined to fall short because they operate on a fundamentally impoverished model of the problem space LLMs have to solve. This makes them struggle to identify both the true extent of the amazing achievement current models represent and the nature of the capabilities and failures they exhibit. I will argue that this limit stems from two related fallacies: 1. Isomorphism fallacy - the assumption that any system producing symbolic structure must have some formal internal structural representation that can be transformationally mapped to the surface output. 2. Words and rules fallacy - the assumption that language (including its semantics) is at core a generative system consisting of modular division between units of representation (words, etc.) and a set of rules that determine their composition.

P6 Uziel Awret

On Human Reasoning and Spatial Sensibility

I will begin with some comments on the evolutionary origins of causal inference arguing that the same place-neuron architecture that enables a rat to navigate a well-trodden maze (Buzsaki& Lisman's (Spatio) Temporal Coding)(1) which may already exist in bees, underlies causal inference and serial analysis in general. In both cases hippocampal place neurons undergoing Theta-Precession project to cortical areas to execute a serial code: the difference is that in navigation the hippocampus projects to the 'evolutionary older' sensory cortex while in more abstract serial analysis the hippocampus projects to higher frontal cortical areas.

Next I will argue that understanding the evolution of spatial phenomenology (including that of geometrical sensibility) and its relevant neural correlates is not only crucial to understanding human reasoning and causal inference but that it can reveal basic biological architectures that may be necessary for the design of 'smart-enough' AI that will reason like us. Fortunately we have a relatively good understanding of the neural mechanisms that 'generate' our spatial phenomenology and I will say a few words on the interaction between hippocampal place-neurons and entorhinal grid-neurons including direction, boundary and speed neurons.

Despite recent work (Friston, Pezzulo, Fleming) attempting to simulate Temporal-Coding, and the action of place-neurons with 'extended' Predictive Coding, making predictions not only about the 'next step' but about an extended series of steps (see Friston on consciousness, 'temporal depth' and the ‘temporal horizon’) we may have to explore more biomimetic alternatives. I will end by stressing the importance of spatial sensibility to humanlike reasoning.

P7 Michael A.Popov

Cognitive Acausality Principle in Humans and Machines

In contrast with traditional causality there is in human experience a field of an Acausal Cognition where a chance event seems causally interconnected with coinciding fact. In particularly, in psychiatry a consciousness acausally can exist during unconscious coma. Unconscious human sleep may contain conscious dreams also. Cognitive acausality, as simultaneous occurrence of two meaningfully - but not causally connected events, in fact, could be found beyond psychiatry.The quantum phenomenon of the “half - life” cats and quantum entanglement in physics appear as the instances of such sort of acausal orderedness. Psychiatrists V.Bekhterev, C.G.Jung and quantum theorist W.Pauli have considered acausal cognitition as an empirical principle, but not a philosophical observation. Acausal parallelism in pure mathematics was rediscovered by ABC conjecture (Masser - Orsterle 1985). However, earlier S.Ramanujan in a very short article on round numbers in 1917 described a kind of acausal synchronicity between addition and multiplication without K constant in his attempt of radical of the integers (Popov 2023). Recently, Shinichi Mochizuki (in MFO - RIMS 23 version) had found a remarkable case of acausal orderedness between transformation group symmetries of modular functions (such as Theta functions) and symmetry groups of the hyperbolic geometry of the upper half - plane (2023). Similarly, contrary to all expectations, experiments with walking robots (Valkyrie type of orbital robots for International Space Station) have shown that these intellectual robotic systems need own models of LLM - based “subjective space” (predicted by psychiatrists). Thus, meaningful parallels to the momentary subjective states for space robotics became also unexplainable without cognitive acausality principle.

P8 Oleksandra Halchenko

Bonsai Philosophy in Systems Thinking. Can it Supplement Learning and Teaching?

There is no need for a proof that our world is contexual. Everything needs something else to fullfil itself. Similarly to the law of contexuality in quantum physics, the social science of teaching and cognitive science of learning exist only within the contexts of their measurement. Bonsai is the ancient art of creating a perfect tree, but in miniature. They are all normal trees, but they are not allowed to grow naturally so that they are planted in their own ‘contexts. Every pot is a confined space with its own ecosystem as every individual and every brain is. Shaping nature does not result in completion due to the fact that everything needs something else. This research attempts to investigate whether systems thinking mindset can be supplemented by Bonsai phiosophy to enhance learning. This study strives to create a new concept of Bonsai logic and present a mind map of Bonsai Logic Model to be potentially used to improve the environment, provide feedback and predict system behavior. The study seeks to introduce new contexuality analysis and improvement tools based on Bonsai principles of cultivation such as pruning, wiring, watering, fertilizing and positioning. This research raises question how learning and teaching can be improved by the right sequence and style of feedback providing. Bonsai feedback style is getting away from too impersonal training and is aiming to solve a problem of biased evaluation.

P9 Tianshu Chen

The Biasing Effect of Causal Events on Visual Task

A lot of our understanding of the world are based on our causal perception. In the present study, we looked at the effect of causal perception, i.e., what does causal perception do in addition to perceiving causality. More specifically, we looked at how causal perception affects people's performance in a visual-cognitive task. Participants were asked to report colour changes in two objects, one involved in a causal event (collision) and the other independent. We observed a significant bias towards reporting the colliding colour changing object when colour changes happened during and before the collision, and the bias was less significant when the colour changes happened after the collision. Based on these findings, we proposed an explanation suggesting that the biasing effect results from the interaction between the causal event and a specific attentional period we referred to as an "anticipation period". During this anticipation period participants pay extra attention to objects that could potentially engage in a causal interaction. Hence, they are more likely to notice colour changes happen to the colliding balls when the colour changes happen during the anticipation period. This research contributes to the field of causal cognition and raises questions about the idea that causal perception is a bottom-up perception, as suggested by Michotte. Instead, it provides additional insights into the causal schema theory and suggests people retrieve information from memories to form a causal schema, which is then employed to make prediction of causal events.

NOTES

Department of Computer Science
University of Oxford
Wolfson Building
Parks Road
Oxford, OX1 3QD
UK

Association for Mathematical Consciousness Science