Book of Abstracts

International Interdisciplinary Computational Cognitive Science Summer School

September 2023

Contents

Lightning Talks

Inside the grid, yet outside the box: Computational investigations of human creativity using pixel patterns

12 Sep
13:50
HS001

Surabhi S. Nath
Max Planck Institute for Biological Cybernetics

Creativity is an important, yet elusive, human characteristic. Despite strides in understanding creativity as a cognitive ability, there is a paucity of computational studies associated with it. Our work aims to fill this gap with a computational investigation of the mechanisms underlying little-c creative artistic work, from a reward learning perspective. We define tractably constrained experimental settings involving binary 5x5 pixel patterns. We develop a taxonomy with two types of creativity – static and dynamic, crossed with two modes of creativity – evaluation and production, resulting in four experimental conditions for investigation. We outline various possible underlying computational mechanisms such as (1) an immediate value function for static creativity, (2) a long-run value function for dynamic creativity, (3) the history-dependent nature of evaluation and (4) a search process guiding the production. We design a series of behavioural experiments based on our taxonomy and propose directions for data analyses. Through this we aim to enrich the computational understanding of product and process creativity.

Cognitive modeling of the latent scope bias in causal explanations

12 Sep
14:10
HS001

Yuki Tsukamura
University of Tokyo

Latent scope bias refers to the tendency to prefer explanations that do not predict unobservable events over those that do in causal explanations. The occurrence of latent scope bias is believed to be due to an underestimation of the occurrence probability of unobservable events and Bayesian probability calculations using it. However, it remains unclear whether it is based on Bayesian probability calculations. Therefore, we employed subjective probability response data and compared this process with alternative models through Bayesian modeling. As a result, a cognitive model based on "subjective utility" was favored over Bayesian probability calculations, suggesting that the process of evaluating explanations leading to the latent scope bias is not based on Bayesian probability calculations.

Furthermore, while latent scope bias was observed in the aggregated data from all participants, it was not necessarily a large effect, and it was shown that it did not occur in a certain number of the participants when looked at individually. In other words, a certain

number of normative respondents answered values corresponding to the normative solution with relatively small variance. From these findings, it is also suggested that the latent scope bias may appear to arise in the entire group due to a certain number of people with bias.

How real-time interaction aids collaboration under the temptations to free ride

12 Sep
14:30
HS001

Ryutaro Mori
University of Tokyo

Understanding how human groups collaborate successfully, even when individuals may be tempted to free ride, is a fundamental question in the social sciences. While conventional models such as game theory often distill socially dependent situations into simplified "staged settings" in which agents choose actions only at fixed timings, real-world collaborations are typically marked by real-time interactions in which decisions are made instantaneously and communicated immediately. Our research posits that such real-time interactions facilitate collaboration by naturally creating a sequence of cooperative actions and reactions. Specifically, when decision timings are sequential rather than simultaneous, each decision becomes a response to prior ones, mitigating much of the social unpredictability and aiding coordination. Moreover, if agents themselves can decide when to act, the earlier/precedent decisions could lean more towards cooperation than if they were random; cooperative agents might decide early to encourage reciprocation, while competitive agents might delay revealing their intentions. We formulate these ideas using a classic prisoner's dilemma game and experimentally verify several key behavioral regularities. In the talk, I would particularly love to discuss how computational modeling can be employed to decipher real-time strategic interactions.

The interplay of synaptic remodeling, forgetting and creativity

12 Sep
13:50
SR003

Jonas Elpelt
Frankfurt Institute for Advanced Studies

Why do we forget? The fascinating phenomenon of forgetting is often associated with pathologies such as dementia, but forgetting can have crucial positive effects for behavioral flexibility and can enforce more creative abilities. Recently, it has been shown that forgetting involves biological mechanisms that actively and selectively 'erase' memory information and can lead to changes on different organizational levels (synapses, neurons, networks, behavior). In a joint German-Israeli project we want to study the change of representations under continuous synaptic reconfiguration in vitro, in vivo and in silico models. By investigating spontaneous dynamics in network configurations we want to explore the possibility that spontaneous changes in synaptic connection might drive forgetting. For this purpose we use the behavioral readout of forgetting and adaptive learning to construct a theoretical framework to link forgetting and intrinsic synaptic dynamics. Finally we want to establish a possible relationship of forgetting, increased synaptic plasticity and creativity in both biological and artificial neural networks.

Comparison of preprocessing of EEG hyperscanning data with different automated algorithms

12 Sep
14:10
SR003

Ralf Krüger
University of Tübingen, Fraunhofer IAO

In the context of naturalistic, out-of-lab brain imaging situations, the data quality is often not optimal. I am currently comparing several preprocessing algorithms (Artificial Subspace Reconstruction, global and local Autoreject and Automated ICA) on their ability to clean data obtained in an neurofeedback art exhibit. Two participants were shown a sculpture while simultaneously measuring inter-brain synchronisation with an 8-channel EEG system as part of the exhibit Brain Palace in the STATE Studio Berlin. I will present the results of this experiment as well as standard preprocessing methods on the power spectrum density of the subjects and several connectivity metrics. Since the connectivity metrics are very sensitive to changes in the frequency domain of the signals, care has to be taken that within subject pairs the signal is processed in similar ways, which needs to be accounted for in artifact removal. I used the Hyperscanning Pipeline for Python HyPyP to perform most of the analysis, but implemented some additional functionality to visualize the intermediate results and measure connectivity metrics not yet specified in the pipeline. I will also present a signal-to-noise ratio and its limitations for determining signal quality in this context.

Skill homeostasis: Toward a model of ultra-robust behavior

12 Sep
14:30
SR003

Guillaume Pourcel
University of Groningen

In this talk, I'll present my current research directions in modeling the impressive phenomenon of ultra-robust behavior in humans, focusing on two intriguing cases: sensory substitution and inverted vision. Sensory substitution demonstrates the brain's rapid adaptation to distorted sensory inputs, such as retranslating visual information into tactile sensations through an array of vibrators. Inverted vision studies are concerned with our abilities to adapt to the disruption of our visual processing by presenting the world upside-down. Both cases showcase remarkable resilience in the face of diverse perturbations.

Conventionally, explanations for robust behavior have rested upon the premise of a meta-learning framework, wherein evolution optimizes learning rules capable of coping with the variability of environments experienced during phylogeny. We propose an alternative hypothesis that suggests that the brain employs a homeostatic mechanism to regulate skills when faced with unprecedented disturbances. Drawing inspiration from computational neuroscience and population dynamics, we view the brain as a nonlinear dynamical system with distinct skill-related patterns encoded in the activity of neural populations.

Central to our model is the identification of skill-specific linear subspaces, acting as references for regulation. When significant perturbations occur, the brain seeks equilibrium by readjusting its internal dynamics. This novel perspective not only addresses robustness in a volatile environment but also unveils a mechanism for the brain to restore stability amidst unforeseen challenges.

Mastering prompt engineering: Unleashing the potential of language models

13 Sep
13:30
HS001

Nina Ma

Osnabrück University

Prompt engineering is a crucial aspect of harnessing the power of language models, enabling tailored and effective interactions. This talk delves into the art of creating prompts that yield desired outputs from language models like GPT-3.5. Prompt engineering involves using instructional completion prompts and fine-tuning techniques to enhance model performance. Completion prompts guide the model's responses, while fine-tuning optimizes its behavior. By leveraging fine-tuning, you can expose the model to a broader range of training data than can fit in a prompt, thereby achieving token savings and lower latency rates. Various strategies make prompt engineering more effective. Initiating with clear instructions, specificity, and tone-setting helps frame the context. Breaking tasks into subtasks and using probing questions elaborates outcomes effectively. Refining the prompt iteratively and employing additional tactics like delimiters, structured output requests, and temperature parameter modifications further enhance results. Prompt engineering also encompasses zero-shot, one-shot, and few-shot approaches. These methods leverage task descriptions, examples, and fine-tuning, respectively, to make language models more adaptable to specific tasks. However, prompt engineering isn't without challenges. Concepts like prompt leaking, prompt injection, and jailbreaking highlight potential risks associated with unintended model behavior and data privacy. In a world where language models play an increasingly significant role, mastering prompt engineering empowers us to extract meaningful and tailored insights, responses, and content from these powerful AI tools.

GEDI: GEnerative and DIscriminative training for self-supervised learning

13 Sep
13:50
HS001

Emanuele Sansone

KU Leuven

Self-supervised learning is a popular and powerful method for utilizing large amounts of unlabeled data, for which a wide variety of training objectives have been proposed in the literature. In this talk, we provide a Bayesian analysis of state-of-the-art self-supervised learning objectives and propose a unified formulation based on likelihood learning. Our analysis suggests a simple method for integrating self-supervised learning with generative models, allowing for the joint training of these two seemingly distinct approaches. We refer to this combined framework as GEDI, which stands for GEnerative and DIscriminative training. Additionally, we demonstrate an instantiation of the GEDI framework by integrating an energy-based model with a cluster-based self-supervised learning model. Through experiments on synthetic and real-world data, including SVHN, CIFAR10, and CIFAR100, we show that GEDI outperforms existing self-supervised learning strategies in terms of clustering performance by a wide margin. We also showcase an application to the

neuro-symbolic setting, where GEDI can learn symbolic representations supporting learning and reasoning in the small data regime without the need for additional supervision or costly pre-training steps.

Local syntactic coherence effects in GPT-3 surprisals

13 Sep
14:10
HS001

Tobias Hoffmann
University of Freiburg

Local syntactic coherence (LSC) effects have shown that the human sentence processor (HSP) can be misguided by a locally embedded sequence of words that could form a sentence in isolation, but must be analyzed differently in the left context of the sentence. These effects have been attributed to temporal local affixes (SOPARS, Tabor et al. 2004) and word-wise prediction in recurrent networks (Konieczny et al. 2005), among others. Large language models, such as GPT-3+, have impressive capabilities in language comprehension and production. Due to their transformer-based architecture, they should be immune to LSCs. We used the OpenAI API to retrieve the surprisal values for our test items word by word. We then reanalyzed our eye-tracking reading data on sentences with local syntactic coherence embedded in short contexts (citation omitted). Contexts were constructed to draw attention to either the local coherence meaning or the global meaning of the target sentence. While the context manipulation affected the size of the LSC effect, it did not alter GPT-3 surprises in the critical region. While GPT-3 surprise scores were good predictors of total reading time, they did not eliminate LSC effects. We conclude that HSP, unlike transformer-based LLMs, employs mechanisms of local attention.

Can we deepfake fMRI data? Unpaired functional alignment for group level neural decoding

13 Sep
14:30
HS001

Sabine Scholle
Osnabrück University

This presentation focuses on the functional alignment of fMRI data, drawing inspiration from Daniel Anthes' thesis titled "BOLD Deepfakes: Functional Alignment for Unpaired Data using Deep Neural Networks" and delving into the foundational work of J.V. Haxby, a renowned researcher in the field of fMRI data analysis (Haxby et al., 2010).

A core challenge for cognitive neuroscience is to find similarity across neural diversity (Churchland, 1998); that is, to find shared or similar neural processes supporting the diversity of individual cognitive experience. Anatomical variability and limited structure-function correspondence across cortex (Paquola et al., 2019) make this goal challenging. To address this challenge, functional alignment is an increasingly popular family of methods for functional magnetic resonance imaging (fMRI) analysis: from the initial introduction of hyperalignment in Haxby et al. 2011, the range of associated methods has grown to include several other linear methods. The presentation explores the innovative approach proposed by Daniel Anthes of "BOLD Deepfakes," leveraging nonlinear deep neural networks to achieve one-shot functional alignment for unpaired data.

Lastly, the presentation outlines potential avenues for further improvement and regularised enhancements in the context of my bachelor thesis. These ideas include exploring novel regularisation techniques to improve the robustness and generalisation of the deep neural network models.

---

13 Sep
13:30
SR003

Object tracking using an active stereo visual system

Tin Mišić
University of Zagreb

Object tracking has many applications in robotics, autonomous vehicles, and surveillance. Among all methods, those that use active stereo systems with moving cameras stand out the most. Special emphasis in this paper is given to approaches of active stereo systems that use virtual horopters and log-polar image mapping. An active stereo system is designed that uses ordinary web cameras and servo motors. The cameras and motors are connected with plastic parts printed on a 3D printer, and the entire system runs on a personal computer. The system's real-time performance, ability to track a moving object against various backgrounds, consistency of focus on the object, and accuracy of the estimated object position were tested. Approaches using Cartesian image mapping and those using log-polar mapping were compared. Based on the obtained results, some advantages of log-polar mapping over Cartesian mapping were shown, as well as the drawbacks of the specific implementation.

---

13 Sep
13:50
SR003

Investigating brain-like CNNs and their consequences

Niranjan Rajesh
Ashoka University

Recent research in neuro-inspired machine learning in the visual domain has resulted in CNNs that are modelled after the primate visual systems. These are usually done through structural or representational alignment of the CNNs with their primate brain counterparts. These models claim to be behaviourally comparable to human visual intelligence and correlations were found in areas like robustness to visual adversarial attacks. My work will further investigate this consequence of primate visual system alignment in CNN, i.e. the connection between brain likeness and adversarial robustness.

Monitoring nociception in the frontal EEG

13 Sep
14:10
SR003

Viktor Bublitz
Universitätsmedizin Greifswald

Monitoring pain and nociception in critical care patients who cannot self-report presents a substantial challenge. Clinical signs frequently lack both sensitivity and specificity, and existing technical methodologies come with inherent limitations. Accurate predictions of nociception could optimize the administration of analgesia ahead of procedures like endotracheal suctioning. In my doctoral research, I explored strategies to quantify nociception and anticipate reactions to painful interventions in intensive care settings. We particularly examined electroencephalogram (EEG) correlates that either precede or occur simultaneously with behavioral responses to noxious stimuli, such as endotracheal suctioning. Our results showed an elevation in power within the 2-5Hz band that both anticipated and matched these responses, coupled with a decrease in the alpha-band power during these events. Such patterns could be associated with the processing of noxious stimuli and might pave the way for refining and individualizing analgesia in patients unable to articulate their pain. Notably, other power bands and ratios did precede the responses in our study, but these could be attributed more to the concurrent sedation level and arousal than to nociception. Deciphering these intertwined effects is the motivation behind the research that I am currently conceptualizing for subsequent investigations.

Drift-diffusion modelling reveals distinct decision processes for 3D and global motion stimuli in humans and macaques

13 Sep
14:30
SR003

Revan Rangotis
Otto-von-Guericke University Magdeburg

Neurophysiology localised perceptual decision signals for binocular 3D depth and motion stimuli to visual area V5/MT. Drift diffusion modelling (DDM) is widely used to investigate the underlying decision-making processes by simulating decisions as noisy evidence-accumulation which terminates once a threshold is crossed. Two key questions remain unexplored:

1. To what extent are modelled decision processes affected by the visual stimulus (here 3D depth vs. motion) and by the effector for the response (hand vs. saccades)?

2. Are the modelling parameters comparable between monkeys and humans?

To answer these questions, we tested 2 male monkeys (Macaca mulatta) and 20 humans on different stimulus/effector combinations in a 2-alternative forced-choice task (2AFC). Stimuli were a 3D structure-from-motion cylinder or a random dot kinematogram (RDK), requiring perceptual decisions about binocular depth or direction of motion, respectively. Effectors comprised hand and saccadic eye movements. Linear discriminant analysis (LDA) revealed a strong separation by stimulus type but not for the effector for the human data ($p < 0.001$, ROC analysis). Nearly identical clustering was observed for the monkey data when it was projected onto the same space with almost complete overlap. We conclude

that DDM reveals distinct brain processes for perceptual decisions about visual motion and binocular 3D stimuli in humans and macaques, although perceptual signals for both have been localised to the same brain area. We found no distinction between hand and eye movement responses in either species.

Author Index