Target scope of Conference

Deep learning plays a central role in recent developments in research in artificial intelligence (AI). Various ideas based on physics are found in the research of deep learning, and consequently, deep learning and physics are related intimately. This international conference is dedicated to (1) applications of deep learning to physics, (2) discovering similarities among deep learning and physics, and (3) leading to new paradigm in physics motivated by deep learning. Researchers in related fields are welcome to attending discussions at the conference.

Organizers

Koji Hashimoto (Osaka U), Masatoshi Imada (Toyota RIKEN / Waseda U), Kouji Kashiwa (Fukuoka Institute of Technology), Yuki Nagai (JAEA), Masayuki Ohzeki (Tohoku U), Enrico Rinaldi (Riken & Arithmer Inc.), Akinori Tanaka (RIKEN AIP), Akio Tomiya (Riken BNL)

Date and Place

31 Oct - 2 Nov 2019
Panasonic Auditorium, Yukawa Hall, Yukawa institute for theoretical physics, Kyoto university

Schedule

Schedule

<table>
<tr><th>First day: Oct 31(Thu)</th><th>Second day: Nov 1 (Fri)</th><th>Third day: Nov 2 (Sat)</th></tr>
<tr><td rowspan="5"></td><td>9:30-10:20 H.Y. Hu</td><td>9:30-10:20 F. Ruhle</td></tr>
<tr><td>10:20-11:10 G. Kanwar</td><td>10:20-11:10 J. Liu</td></tr>
<tr><td rowspan="2">Lunch break</td><td>Break</td></tr>
<tr><td>11:40-12:30 Y. Yamaji</td></tr>
<tr><td>"Physics, AI, Company" session</td><td rowspan="3">Lunch break</td></tr>
<tr><td>13:10 Opening Welcome</td><td>13:00-13:35 R.K. Seong</td></tr>
<tr><td>13:20-14:10 J. Halverson</td><td>13:35-14:10 Y. Miyake</td></tr>
<tr><td>14:10-15:00 S. Krippendorf</td><td>14:10-14:45 M. Ohzeki</td><td>14:00-14:50 Z.Y. Meng</td></tr>
<tr><td>15:00-15:50 G. Yang</td><td>14:45-15:00 Discussion</td><td>14:50-15:40 H. Yoshino</td></tr>
<tr><td>Break</td><td>Break</td><td>Break</td></tr>
<tr><td>16:20-17:10 L. Wang</td><td>15:30-18:00 Poster session</td><td>16:10-17:00 G. Shiu</td></tr>
<tr><td>17:10-18:00 S. Amari</td><td></td><td>17:00 Closing</td></tr>
</table>

Abstracts

First day: Oct 31 (Thu)

1. *13:20-14:10* James Halverson (Northeastern University)
   Generative Models and Statistical Predictions in String Theory
   String theory is a theory of quantum gravity and is the leading candidate for unifying observed particle physics and cosmology. If it is correct, its many solutions means that "fundamental physics" is an intrinsically complex system with many metastable states, motivating the use of techniques from deep learning and raising questions regarding the dynamics of complex systems. In this talk I will review aspects of the string landscape, how computationally complex problems arise in it, and how generative models could be used to overcome some of the difficulties. Specifically, a variety of generative adversarial networks will be used to approximate Kahler metric data on Kahler moduli space, which are relevant for axion physics in cosmology, including interactions of axion-like-particles with the photon. From another perspective, the utilized data techniques provide methods for learning random matrix approximations to matrix ensembles, in this case those that arise in string effective Lagrangians.
2. *14:10-15:00* Sven Krippendorf (Arnold Sommerfeld Center)
   Dualities in and from Machine Learning
   Dualities play a key role in our understanding of dynamical systems appearing in effective field theories from string theory and in general in fundamental physics. Dualities in field theory and string theory are a tool to obtain models describing data at a precision beyond a level standard effective field theory techniques allow. At this stage these techniques are not used explicitly in machine learning.
   Based on examples, we identify data questions which are easily answerable in one duality frame but not in the other. Simple neural networks confronted with data in this favoured duality frame can answer data questions, in the other frame they cannot identify the correct answer.
   I then discuss the question whether a deeper neural network can learn this favourable dual data representation? I present two strategies, how a dual representation can be learned without knowing about its existence.
3. *15:00-15:50* Greg Yang (Microsoft Research)
   Tensor Programs: The Feynman Diagrams of Deep Learning
   Feynman diagrams are used in physics to compute correlation functions in an interacting theory. In analogy to Feynman diagrams, we introduce *tensor programs* in this talk as a mechanism to compute the “correlation functions” of deep neural networks, which can be thought of as having an “interacting theory” due to their nonlinearities. We start with the motivating example of the Semicircle Law, which describes the eigenvalue distribution of a large matrix with random

Gaussian entries. We show a classic proof using Feynman diagrams, and then compare it against a new proof using Tensor Programs. We discuss similarities and differences in the “interactivity” handled by Feynman Diagrams and Tensor Programs. Next, we describe several consequences the Tensor Programs framework in deep learning: 1) wide neural networks of any architecture are Gaussian processes when randomly initialized; 2) when trained with small learning rate, wide neural neural networks of any architecture evolve via functional gradient descent with a kernel, called the “Neural Tangent Kernel”; 3) when trained with large learning rate, wide neural networks learn an embedding function of inputs into a space of random variables.

4. *16:20-17:10* Lei Wang (Chinese Academy of Sciences)
   Neural Canonical Transformation
   Canonical transformation plays a fundamental role in simplifying and solving classical Hamiltonian systems. We construct flexible and powerful canonical transformations as generative models using symplectic neural networks. The model transforms physical variables towards a latent representation with an independent harmonic oscillator Hamiltonian. Correspondingly, the phase space density of the physical system flows towards a factorized Gaussian distribution in the latent space. Since the canonical transformation preserves the Hamiltonian evolution, the model captures nonlinear collective modes in the learned latent representation. We present an efficient implementation of symplectic neural coordinate transformations and two ways to train the model. The variational free energy calculation is based on the analytical form of physical Hamiltonian. While the phase space density estimation only requires samples in the coordinate space for separable Hamiltonians. We demonstrate appealing features of neural canonical transformation using toy problems including two-dimensional ring potential and harmonic chain. Finally, we apply the approach to real-world problems such as identifying slow collective modes in alanine dipeptide and conceptual compression of the MNIST dataset.
   Ref: Shuo-Hui Li, Chen-Xiao Dong, Linfeng Zhang, Lei Wang, arXiv:1910.00024
5. *17:10-18:00* Shun-Ichi Amari (RIKEN)
   Deep Random Neural Field
   We begin with a short history of AI and neural networks starting with multilayer perceptrons and their learning. We then review recent trends of deep learning with extremely wide layers, showing that random connected networks play a fundamental role. We thirdly review a short history of statistical neurodynamics of random networks. We show how signals are transformed through random deep networks. We also show the Fisher information of random networks. We finally consider randomly connected neural fields having infinitely many neurons.

Second day: Nov 1 (Fri)

6. *9:30-10:20* Hong-Ye Hu (UC San Diego)
   Machine Learning Holography
   During this talk, I am going to review the recent progresses made in the field of machine learning holography. Holographic duality is a beautiful relation between certain quantum field theories on a d-dimensional boundary and classical gravity theories on a (d+1)-dimensional bulk. In general, finding this explicit duality map is challenging.
   In the first part of this talk, I am going to discuss designing the optimal holographic mapping by introducing the neural network renormalization group as a universal approach. The mapping is constructed using a flow-based hierarchical deep generative model. And I am also going to discuss the unification of renormalization group and generative model as the forward and backward flow of a holographic mapping. We applied this approach to the complex phi-4 theory in two dimensional Euclidian spacetime in its critical phase, and show that the emergent bulk geometry matches the three-dimensional hyperbolic geometry in saddle point approximation.
   In the second part of this talk, I am going to discuss how to use neural network to solve the holographic QCD problem. Suppose we have the lattice QCD data, can we find a gravity theory that can explain the experimental data? Moreover, can the gravity theory of the bulk predict new experimental phenomenon of the system? We used both feedforward neural network and newly developed neuralODE technique to solve this problem. By providing the lattice QCD chiral condensate data, the machine can learn a metric for the bulk field. We observed the coexistence of confinement and deconfinement of the bulk field. Moreover, the machine could predict a Q-Qbar potential that matches the experimental result.
7. *10:20-11:10* Gurtej Kanwar (MIT)
   Flow-based generative models for lattice theories
   Monte Carlo sampling is a powerful approach to computing partition functions, thermodynamic observables, and operator expectation values in quantum and statistical field theories. As critical points are approached in parameter space however, the cost of drawing independent samples typically diverges due to critical slowing down. Typically, solving or reducing this slow-down must be done on a case-by-case basis by exploiting features of the problem at hand, and in many cases no good approaches are known. I discuss our recent work on applying generative machine learning techniques to the sampling problem, including a case study of scalar phi^4 theory in which critical slowing down of sampling is removed by employing trained generative models.

"Physics, AI, Company" session

8. *13:00-13:35* Rak-Kyeong Seong (Samsung SDS)
   An Overview of AI at Samsung SDS
   At a time when advances in AI are shaping the world around us, Samsung is making key contributions in the endeavour of improving lives and businesses. The talk will give a broad overview of what recent contributions Samsung has made with the help of AI and what we can expect in the future.
9. *13:35-14:10* Youichiro Miyake (Square Enix)
   AI Technologies in Digital Games
   There are two flows in AI technologies in 60 years such as Symbolism and Connectionism. Digital Game AI uses both two technologies. In the lecture, many cases of AI in digital game titles including FINAL FANTASY XV (SQUARE ENIX, 2016) are introduced. One of purposes of digital game is to make an autonomous AI such as to recognize a game world, make a decision, and generate a body motion by itself. Understanding an environment around itself for AI is similar to physics for a human to understand the universe. And it includes also a philosophical problem such as "What is a world? ","What is a self ?" and "What is a life" ? Making an interactive AI in real-time with body is a science, engineering. and.philosophy. In the lecture, a whole figure of digital game AI is showed.
10. *14:10-14:45* Masayuki Ohzeki (Tohoku Univ.)
    Quantum annealing and its application to deep learning
    - new direction of quantum annealing -
    Quantum annealing is a generic solver of combinatorial optimization problem and is implemented by a hardware known as the D-Wave quantum annealer.
    On the other hand, the neural network, which is a big success in developing the artificial intelligence and data science, is also attained via solving optimization problem.
    In this talk, by taking the quantum annealer as an optimizer, we introduce several directions of its application.
    One of the main topic would be an application of the quantum annealer to the deep neural network although the standard one only deal with the binary variables.

Third day: Nov 2 (Sat)

11. *9:30-10:20* Fabian Ruhle (CERN)

    Machine learning for Calabi-Yau metrics

    I will discuss how neural networks can be used to study metrics on Calabi-Yau manifolds. These manifolds feature prominently in string theory, and their metrics determine physical properties of string theory models.

12. *10:20-11:10* Junwei Liu (Hong Kong University of Science and Technology)

    Self-learning Monte Carlo method and all optical neural network

    Self-learning Monte Carlo (SLMC) method is a general-purpose numerical method to simulate many-body systems. SLMC can efficiently cure the critical slowing down in both bosonic and fermionic systems. Moreover, for fermionic systems, SLMC can generally reduce the computational complexity and speed up simulations even away from the critical points. For example, SLMC is more than 1000 times faster than the conventional method for the double exchange model in 888 cubic lattice. In addition, SLMC also provides a general framework to naturally integrate the advanced machine learning techniques into Monte Carlo. In this talk, I will give an introduction about the background, basic idea and the design principle of SLMC. Later, I will explicitly show how to use SLMC and its great accelerations in classic systems, free fermions coupled with classical spins systems, and interacting fermion systems. At the end, I will talk our recent developed all optical neural networks, which can realize the intrinsic parallel calculations at the speed of light and are expected outperform the electronic neural networks with large system size.

    References:

    [1] Self-Learning Monte Carlo Method, PRB 95, 041101(R) (2017)

    [2] Accelerated Monte Carlo simulations with restricted Boltzmann machines, PRB 95, 035105 (2017)

    [3] Self-Learning Monte Carlo Method in Fermion Systems, PRB 95, 241104(R) (2017)

    [4] Self-learning Monte Carlo with Deep Neural Networks, PRB 97, 205140 (2018)

    [5] All optical neural network with nonlinear activation functions, Optica 6, 1132-1137 (2019)

13. *11:40:12:30* Youhei Yamaji (Department of Applied Physics, the University of Tokyo)

    Origin of High-Temperature Superconductivity Revealed by Boltzmann Machine Learning

    Recently machine-learning approaches are widely utilized as tools to analyze accumulated data across various research domains. Machine learning also has innovated ways of exposing physical observables, which is invisible in direct scientific measurements. Solving the underdetermined inverse problem is a crucial step towards the machine learning innovation of the scientific measurements, which enables us to extract fundamental physical quantities entangled with each other in existing experimental data. A combination of properly chosen prior knowledge and machine learning holds the key to find a solution to the inverse problem.

An enigmatic inverse problem in condensed matter physics is found in open issues of high temperature superconductors, which has long been a major challenge in physics. How electrons are mutually interacting is the key to identify the origin of the superconductivity. We are, however, able to observe the motion of electrons only after projection onto experimentally accessible degrees of freedom.

In this study, we utilized the Boltzmann machine, combined with physically sound prior knowledge, to solve the inverse problem and extract physical quantities hidden in experimental data. The method is applied to the angle-resolved photoemission spectroscopy spectra [1] of copper-oxide high-temperature superconductors. From a spectrum [2], we extracted metallic (normal) and superconducting (anomalous) components of the self-energy separately, in which mutual interactions among electrons are encoded. We found prominent peak structures emerging both in the normal and anomalous self-energies, respectively, which are canceled in the total self-energy and hence invisible in experiments, as the origin of high-temperature superconductivity [3]. The present achievement may open avenues for innovative machine-learning spectroscopy.

[1] A. Damascelli, Z. Hussain, and Z.-X. Shen, Rev. Mod. Phys. 75,
473-541 (2003).
[2] T. Kondo, et al., Nature 457, 296-300 (2009).
[3] Y. Yamaji, T. Yoshida, A. Fujimori and M. Imada, arXiv:1903.08060.

14. *14:00-14:50* Zi Yang Meng (Chinese Academy of Sciences)

What we talk about When we talk about learning in many-body physics

In this talk, I will review recent developments in a priori and a posteriori learning strategies in dealing with classical and quantum many-body systems. Thanks to these philosophical and numerical advancements, new paradigms in condensed matter and high energy physics such as non-Fermi-liquid, quantum criticality and emergent gauge-field coupled with matter field can be accessed with large-scale numerical simulations. These results in turn inspire further analytical and numerical progress towards the complete form of few important many-body physics problems.

Reference:

[1] TOPICAL REVIEW, J. Phys.: Condens. Matter 31, 463001 (2019)
https://dx.doi.org/10.1088/1361-648X/ab3295
[2] PNAS August 20, 2019 116 (34) 16760-16767
[3] Phys. Rev. X 9 , 021022 (2019)

15. *14:50-15:40* Hajime Yoshino (Osaka Univ.)

From complex to simple : hierarchical free-energy landscape renormalized in deep neral networks

We develop a statistical mechanical approach based on the replica method to study the phase space structure of deep neural networks. Specifically we analyze the configuration space of the synaptic weights in a simple feed-forward perceptron network within a Gaussian approximation.

By increasing the strength of constraints, i. e. increasing the number of imposed random input/output patterns, successive glass transitions take place layer-by-layer starting next to the input/output boundaries going deeper into the bulk. For deep enough network the central part of the network remains in the liquid phase. The successive glass transitions bring about a hierarchical free-energy landscape which evolves in space: it is most complex close to the boundaries but becomes renormalized into progressively simpler one in deeper layers. Similarly, in an idealized teacher-student setting, the student machine reproduces the teacher machine around the input/output boundaries but leaves behind the central part as very different from the teacher. These observations provide clues to understand why deep neural networks operate efficiently and also insights for related problems in glass physics and biology. Finally we present some results of a set of simple numerical simulations to examine the theoretical predictions.

16. *16:10-17:00* Gary Shiu (Univ. Wisconsin)
    Decoding the Shape of the String Landscape with Data Science
    What gave string theory its appeal is that it provides a consistent framework for unifying quantum mechanics with gravity. However, the richness of string theory also suggests an enormous number of solutions, often known as the string landscape. In some corner of string theory, the number of solutions was estimated to be 10500 and recent work suggested that there may even be more (~$10^{272,000}$). These solutions constitute a huge set of "theoretical data”, possibly far beyond what commonly referred to as "big data” in other data intense fields. When the dimension of the data space is huge, we cannot visualize the "shape” of data and decode the underlying physics. Searching for solutions with desired properties would also be extremely challenging. In this talk, I’ll discuss how topological data analysis and genetic algorithms can help uncover the structure of the string landscape.