## Faculty Members and Labs in Department of Computer Science

<table>
<thead>
<tr>
<th>Imai Lab</th>
<th>Professor</th>
<th>Hiroshi Imai</th>
</tr>
</thead>
</table>

*No student recruitment in this Academic year*

Our laboratory aims to introduce new computational models such as quantum computers with new quantum algorithms, to develop effective algorithms for fundamental and new research field such as huge network, machine learning, and to analyze complexities to represent difficulty of computing in sophisticated problems. We undertake a variety of studies ranging from fundamental theories to applied fields. Presently the issues we study include:

1. **Design and analysis of algorithms** to solve large graph problems, hard optimization problems such as SAT and Max Cut, graph decomposition, and machine learning.

2. **Research on Quantum Computing and Quantum Cryptography**: in particular, we have been investigating the following topics, fundamental theory for realizing quantum computers, quantum algorithms, quantum complexity theory to analyze computational ability of quantum computers, quantum information theory for quantum cryptography, and theoretical analysis of quantum information such as quantum entanglement. IBM Quantum Computers will be used to reveal the state-of-the-art power of the modern real quantum computer.

3. **Analyzing graphs and matroids via their Tutte polynomial**: This research includes many typical problems such as Ising mode in physics, which has direct connection with graph max cut, Jones polynomial of a knot, chromatic number, reliability function, etc. BDDs are used to represent this invariant polynomial in an implicit and compact manner, and are extended to quantum tree tensor network and other physical problems.

4. **Machine learning algorithms for graphs and their optimization**, including graph mining, graph neural network. Optimization problems related to this direction will be explored.

   http://www-imai.is.s.u-tokyo.ac.jp/

<table>
<thead>
<tr>
<th>Suda Lab</th>
<th>Professor</th>
<th>Reiji Suda</th>
</tr>
</thead>
</table>

*No student recruitment in this Academic year*

Our laboratory researches on (1) parallel and high performance computing, (2) numerical algorithms and (3) data structures, aiming to higher speed, higher precision and higher reliability in large scale scientific computations.

One of our approaches for forthcoming supercomputers, which may contain millions of processors, is communication avoiding algorithms. They are designed to reduce the amount of inter-processor communication, which will be a significant cost in a large scale parallel computer. Effective and efficient utilization of 100+ CPU cores, their SIMD parallelism, and computational accelerators such as GPGPU are also important targets.

Automatic performance tuning or autotuning is a mechanism that chooses the best implementation after several trial executions on the real machine. We are investigating mathematical methods to model the performance from the settings of tunable parameters, and to optimize tunable parameters efficiently, based on Bayesian statistics.

We have worked on fast numerical algorithms with lower computational complexities. Mixed precision, higher precision for very large computations, and utilization of fast lower precision are interesting topics.

As for data structure, our main interest is in compact data structures, which is highly relevant to bigdata analyses, e.g., in bioinformatics. Bigdata applications require optimizing data structure space including the constant factor. When the data itself is prohibitively large, we need to compress it. The challenge is to do so in such a way that various queries can be supported efficiently. We are also working on unconventional data structures with applications to security and privacy protection, e.g., ORAM.

   http://sudalab.is.s.u-tokyo.ac.jp/~reiji/lab-e.html
Kobayashi Lab

Professor Naoki Kobayashi

Our research group is studying theoretical foundations for software and their applications such as program verification. On the one hand, an increasing number of important systems such as transportation systems, medical devices, and banking systems are now controlled by computer software, and a bug of such software can cause a serious disaster. On the other hand, such software is becoming more and more complex and larger, and it is difficult to maintain the quality of software by using traditional software engineering technologies like testing. In view of these situations, we aim to improve the reliability and efficiency of software by developing automated techniques for program verification and transformation based on rigorous mathematical methods. To achieve the goal, we also need to study and advance many research topics in theoretical computer science, such as type theory, formal languages and automata, and automated theorem proving. It is a pleasure of our research to find out that deep mathematical results, which initially seem to be only of theoretical interests, are actually quite useful for the practically-motivated research mentioned above. See http://www-kb.is.s.u-tokyo.ac.jp for details.

Recent research topics include:

1. Higher-order model checking: This is an extension of traditional model checking that has been successfully applied to system verification. We have recently constructed the first higher-order model checker in the world.
2. Automated program verification: By applying the higher-order model checking mentioned above, we are developing fully-automated program verification tools for programming languages like ML and Java.
3. Data compression: String and tree data can be compressed in the form of programs that generate them. The higher-order model checking above can be used to transform such compressed data without decompression.
4. Type theory and program semantics: We are studying intersection type systems and game semantics as foundations for program verification and transformation.
5. Protocol verification: Cryptographic communication protocols are used, for example in Internet shopping, for safely exchanging confidential data. We are developing a method for automated verification of such protocols.

Igarashi Lab

Professor Takeo Igarashi

(1) User Interface: We are working on user interfaces for information appliances ranging from personal computers, smartphone, robots, and self-driving automobiles. We not only develop techniques to improve efficiency but also propose new ways of interaction.
- User interfaces for machine learning and artificial intelligence. We focus on the preparation of training data and interactive learning.
- Interaction techniques for novel appliances such as smartphone, smart watches, AI speakers and smart glasses.
- Interaction techniques for giving directions and controlling real-world systems such as robots and self-driving automobiles.

(2) Graphics: We mainly work on interactive shape modeling. We work on content creation for digital media such as movies and games. We also work on interaction techniques for digital fabrication using 3D printers and laser cutters.
- Content creation such as 3D modeling and 2D animation using recent technologies such as sketching and machine learning. We also work on interaction techniques for medical imaging.
- Shape modeling for real world objects such as musical instruments, clothes, and toys leveraging real-time physical simulation.
- Novel techniques for 3D scanning and 3D printing to support personal fabrication.

We can provide opportunity for students to collaborate with research groups in other countries and productions. http://www-ui.is.s.u-tokyo.ac.jp/
Miyao Lab

Professor Yusuke Miyao

Natural Language Processing, Computational Linguistics, and Human Cognition

The goal of our research group is to make computers understand natural language. People communicate with each other using language, while its mechanism is still largely mysterious. We apply theories and technologies of computer science and linguistics to modeling the process of understanding and generating natural language. Specific research themes include: syntactic parsing to compute latent structures of sentences, semantic analysis to compute semantic structures, recognizing textual entailment to determine whether two texts express the same meaning or not. We also study research on grounding, which aims to connect language with non-verbal information such as vision and numerical data.

Our group is also engaged in developing real-world applications of natural language processing, with the focus on applying abovementioned fundamental technologies. Examples include dialog systems to communicate with computers using natural language, data-to-text to automatically explain various non-verbal information in the world in language.

Refer to our homepage for the details of the research: https://mynlp.is.s.u-tokyo.ac.jp/en/

Syntactic/semantic parsing using
Combinatory Categorial Grammar

Issei Sato Lab

Professor Issei Sato

We humans decide our behaviors on the basis of knowledge learned and abstracted from past experiences and current situations. Machine learning is a field of artificial intelligence for improving the performance and behaviors of a machine through the use of data accumulated from past human experiences and current human interaction.

Currently, we are working on the following four main research themes.
1. Generalization and memorization: Generalization is the property of predicting unknown data that does not exist in the training data. Memorization refers to storing the training data in the memory in learning algorithms: however, simply memorizing the training data does not lead to generalization, and over-abstracting the training data reduces the prediction performance. Analyzing the relationship between generalization and memorization is the most fundamental problem in machine learning research.
2. Perturbation and uncertainty. How perturbations to training data and model parameters, as well as uncertainty in predictions, affect learning algorithms.
3. Representation learning. What kind of abstract representation of training data leads to generalization.
4. Robustness. Property that is necessary for learning appropriately even when the training data or the data to be predicted are different from what is expected.

Required knowledge.
We use probability & statistics, linear algebra, functional analysis, and optimization as tools to construct the theory of machine learning. Our lab. assumes that the students are familiar with these mathematics. In addition, programming skills, primarily in Python, are also needed, because we perform exhaustive empirical analysis of the theory. See https://www.ml.is.s.u-tokyo.ac.jp/ for details.
Yoshimoto Lab
Associate Professor Yoshihide Yoshimoto

One of the major motivations to invent electronic computers was the application to science and technology. After the invention, the performance of the computers improved dramatically with the exponential development of semiconductor technology: Moore’s law. Computational science, which advances science with computation has benefited greatly from the development. Nowadays, however, because the limitation of semiconductor technology is coming up to the surface, the complexity of computer systems such as parallelization etc. is so increasing that both computational science and computer science have to cooperate once again for the further progress. This laboratory was set up under the above background to perform education and research which connect computational science and computer science. (https://www.cp.is.s.u-tokyo.ac.jp/)

Yoshimoto himself is specialized in solid-state physics which elucidates the properties of materials such as semiconductors, metals, dielectrics, and magnets. Especially, he is specialized in the first-principles electronic structure calculations which accurately simulates the quantum mechanics of electrons which dominates most of the properties of materials. He has developed and open to the public a program package xTAPP for this purpose. The right-hand side figures are examples of visualization of an electronic structure calculation. ( http://xtapp.cp.is.s.u-tokyo.ac.jp/ )

From the electronic structure calculation as a hometown, the aim of the laboratory set up between computational science and computer science is as follows:
1. Perform mutual exchange between wide range of fields in computational science by re-interpretation of the methods developed individually in each field from viewpoints of computer science.
2. Understand needs of computational science from viewpoints of computer science and propose more essential solutions.

Takamaeda Lab
Associate Professor Shinya Takamaeda-Yamazaki

CASYS: Laboratory for Computer Architecture and Systems

Our main research interest is the future computer architecture, including (1) custom computing using FPGA and specialized LSI, (2) algorithm/hardware co-design for machine learning, (3) high-level synthesis compiler for hardware design. We are also pursuing software researches on compilers, programming models, and frameworks for user/programmer friendly computers. We have various active collaborations with companies. We tackle real problems with our innovative architectural technologies.

Recent research topics: (new topics are welcome)
  - High-performance and low-power machine learning HW and architecture
  - DNN model and training algorithm for embedded hardware
  - Fast image processing algorithm and FPGA system
  - Communication-efficient federated learning algorithm
  - Secure and high-performance microprocessor and memory system
  - Hardware obfuscation compiler for IP protection
  - Multi-paradigm high-level synthesis HW compiler in Python
  - Customizable HW compiler for deep neural networks

Lab WEB: https://sites.google.com/view/casys/
GitHub: https://github.com/PyHDI/ https://github.com/NNgen/
<table>
<thead>
<tr>
<th>Ma Lab (Momentum Lab)</th>
<th>Associate Professor  Lei Ma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Since the past decade, data-driven AI is gradually eating software, which leverages the data-driven programming paradigm to develop software by learning from data. As real-world intelligent systems are often complex that contains both traditional software and AI components, designing fundamental methodology and engineering techniques for building trustworthy intelligent software could be a key challenge and theme for both research community and industries.</td>
<td></td>
</tr>
<tr>
<td>(1) Our lab focuses on both sides of software theory and software engineering, to design novel methodology, engineering techniques and toolchains towards building trustworthy intelligent software systems.</td>
<td></td>
</tr>
<tr>
<td>(2) We also actively adapt and adopt intelligent software techniques to diverse cyber-physical systems and cyber-cyber systems across industrial domains to solve real-world challenges.</td>
<td></td>
</tr>
<tr>
<td>Our lab particularly encourages impactful research, and collaboration with researchers around the world to tackle challenging industrial problems with trustworthy software-driven techniques for social good. More details and concrete topic examples of our research can be found at: <a href="https://www.malei.org/lab.html">https://www.malei.org/lab.html</a>, Email: <a href="mailto:malei@is.s.u-tokyo.ac.jp">malei@is.s.u-tokyo.ac.jp</a></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yanaka Lab</th>
<th>Lecturer  Hitomi Yanaka</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language allows people to communicate with great precision, and most of the information is composed of language. With the development of information technology, natural language processing (NLP) applications based on machine learning models with large corpora, such as information retrieval and machine translation, have been widely used in our daily life. However, it is not clear whether such applications understand the meaning of natural language in the same way that we do. We are working on NLP applications that realize human-like natural language understanding by integrating logic-based approaches with machine learning-based approaches based on theoretical linguistics.</td>
<td></td>
</tr>
<tr>
<td>(1) Probing language models through theoretical linguistics</td>
<td></td>
</tr>
<tr>
<td>Do current language models understand the various meaning of natural language? They might learn undesired biases and heuristics from the data. In this laboratory, we are analyzing language models from multiple perspectives of theoretical linguistics, cognitive science, and philosophy.</td>
<td></td>
</tr>
<tr>
<td>(2) Hybrid inference systems by integrating machine learning-based approaches and logic-based approaches</td>
<td></td>
</tr>
<tr>
<td>To calculate the semantic relatedness between texts, machine learning models have been used for representing the meaning of texts. However, these models often use shallow information, and it remains unclear whether they are capable of handling functional meanings of texts such as negation and quantifiers. By contrast, logic-based approaches have been successful in representing such functional meanings as logical formulas. To have advantages over both logic-based and machine learning-based approaches, we are studying hybrid semantic parsing and inference systems by integrating machine learning-based and logic-based approaches.</td>
<td></td>
</tr>
<tr>
<td>(3) Hybrid NLP applications by combining human knowledge with hybrid inference systems</td>
<td></td>
</tr>
<tr>
<td>Depending on applications, such as document checking, it is not necessary to automate all processes. Human-in-the-loop frameworks enable us to realize more effective NLP applications. We are developing hybrid NLP applications by combining human knowledge with hybrid inference systems.</td>
<td></td>
</tr>
<tr>
<td>See <a href="https://ylab.mystrikingly.com/">https://ylab.mystrikingly.com/</a> for details.</td>
<td></td>
</tr>
</tbody>
</table>
With the dramatic performance improvement of information and communication technology, intelligent information processing that was done only by humans is becoming possible also by computers. Under the theme of "how intelligent can computers be?", our laboratory is working on various research topics related to intelligent data analysis, called **machine learning**, in the field of artificial intelligence.

1. **Construction of Learning Theory**
   Generalization is the ability to cope with unknown situations, and is indispensable for computers to behave intelligently. We are theoretically investigating the mechanism of acquiring the generalization capability based mainly on probability and statistics.

2. **Development of Learning Algorithms**
   Machine learning involves various subjects such as supervised learning (learning from input-output paired data), unsupervised learning (learning from input-only data), and reinforcement learning (learning through interaction with an environment). We are developing practical and theoretically motivated machine learning algorithms.

3. **Application of Machine Learning Technologies to Real World**
   Growth and spread of the Internet and sensor technologies allow us to collect a huge amount of data in engineering and fundamental sciences such as documents, audio, images, movies, e-commerce, electric power, medicine, and biology. We are collaborating with industry partners and applying state-of-the-art machine learning technologies to solving real-world challenging problems.

[http://www.ms.k.u-tokyo.ac.jp/](http://www.ms.k.u-tokyo.ac.jp/)
Yokoya Lab

We study image processing for computational imaging and image analysis based on computer vision and machine learning. In particular, we work on intelligent information processing to automatically extract map information, such as land cover labels and elevation models, from remote sensing images acquired by spaceborne and airborne sensors.

1) **Image processing for computational imaging**

Computational imaging, which integrates sensing and computation, allows us to acquire information that cannot be obtained by hardware alone and to overcome hardware limitations, such as resolution and noise. Based on machine learning, optimization, and signal processing, we build mathematical models and develop algorithms to recover unknown original signals from incomplete observation data.

2) **Remote sensing image analysis**

Remote sensing enables us to observe places that are inaccessible to humans; however, it is difficult to collect enough training data due to the limitations of field surveys and visual interpretation. We work on mapping and 3D reconstruction by using synthetic data from simulations and inaccurate labels with low collection costs as training data. We also work on data fusion based on deep learning to handle multimodal data obtained from different spaceborne sensors in an integrated manner.

3) **Image analysis for solving global issues**

We promote projects to solve global issues, including environmental problems, climate change, large-scale natural disasters, and food problems. Our goal is to contribute globally to the realization of the SDGs by solving real-world problems, such as assessing building damage during disasters, estimating biomass and carbon stocks in forests, and mapping crop types, in collaboration with related institutions and researchers in Japan and overseas.

Nakai and Park Lab

Our lab is one of the laboratories specializing in bioinformatics on the Shirokanedai campus ([https://fais.hgc.jp/](https://fais.hgc.jp/)). The basic motivation of our research activities is a naive wish to understand how biological information is encoded as DNA sequences, which are the data type of a simple one-dimensional string. Particularly, we try to understand the regulatory modes of genes that are encoded in the genome with various approaches from computer science. In other words, we aim to understand how the common genome DNA sequence can specify a variety of cell types that constitute our body.

For example, we challenge to characterize the higher-order genome structure that is responsible for the regulation of gene expression, to analyze the temporal and spatial patterns of gene expression at the single-cell resolution, to clarify the mechanism of RNA splicing that occurs post-transcriptionally, and so on. All these studies are designed and performed on the supercomputer system, typically using a combination of natural language models and deep learning techniques. We also pursue collaborations with experimental researchers to verify our in silico prediction results and to contribute to finding ways to treat incurable diseases and boost regenerative medicine. Thus, we not only analyze data and/or develop algorithms but also construct and maintain necessary databases that are open to the public.

Our lab members have a variety of research and international backgrounds, but they all share the common mindset of advanced computational genomics. Although they pursue their own interests and goals, where more emphasis is placed on biological importance, they are interested in the research topics of other members and help each other. Thus, we try to respect each student's research interests and motivation as much as possible when she/he chooses her/his own research topic.
Imoto and Katayama Lab

We are a data science laboratory in the field of health and medicine. Using a wide variety of genome-related big data such as whole genome, transcriptome, epigenome, and metagenome of intestinal microflora, and large-scale data related to health and medicine with time axis such as receipt information, specific health checkups, and wearable devices, we develop statistical/machine learning data analysis techniques. We perform large-scale analysis using supercomputers and cloud computing. We collaborate with various domestic and international companies and researchers in diverse fields to promote social implementation of big data analysis in health and medical fields. Our major research themes are listed below (but not limited):

**Research Theme 1: Development of Genome Big Data Analysis Technology.** We are analyzing the immune system and cancer genomes. We are analyzing genomic regions related to human immunity, analyzing integrated immune systems such as gut microbiota metagenomics, and developing new statistical analysis models to analyze cancer genomes. We also serve as a representative organization of the analysis group of the national project (Whole Genome Analysis Project) in cancer genomic medicine.

**Research Theme 2: Development of Modeling Technology for Biological Systems.** We are developing modeling and simulation technology for biological systems formed by tens of thousands of molecules including genes and non-coding RNA. We are developing mathematical analysis techniques for predicting the effects of drugs and the future progress of diseases.

**Research Theme 3: Development of Artificial Intelligence for Realizing Genomic Medicine.** In the analysis of cancer genomes, thousands to tens of thousands of genomic mutations are found from the whole genome sequence data. To identify cancer-causing (pathogenic) genomic mutations from among these, information published as scientific papers is used. However, the number of papers in the life science field will exceed 30 million by 2021, and it is already beyond the capability of human intelligence to cover all of them. To solve this problem using artificial intelligence, we are conducting joint research with companies and conducting research involving medical practice at the Institute of Medical Science Hospital.

**Research Theme 4: Infectious Disease Research.** New Coronaviruses. Virus genome mutations. Infection risk assessment and control at mass gathering events. Sewage epidemiology studies in the Olympic athletes’ village. Social implementation of infection control by collaborating with the NPB and J-League.

Our laboratory's web page ([https://www.hm-intelligence.com/](https://www.hm-intelligence.com/)) introduces recent activities and research topics (including many themes other than those mentioned above). Please take a look if you are interested.

Shibuya Lab

Medical Science is now dramatically changing. It is mainly due to the advent of next generation sequencers that can sequence DNA with incredibly high speed and low cost. State-of-the-art computer science techniques are required to deal with the enormous amount of medical big data. We aim to develop fundamental big data technologies that could drive medical science to the next stage.

Research topics in our lab:

1. **Big data indexing / searching**
   Development of faster, more accurate, and scalable indexing/searching technologies are highly desired in the era of individual genomic medicine, where we might need to deal with tremendous country-wide individual genomic data.

2. **Privacy preserving computation**
   Genome is the ultimate individual information to be protected. We need to develop privacy preserving computation technologies, such as differential privacy and secret computation, to meet ethical requirements from genome data.

3. **Machine learning / AI algorithms for medical science data**
   Development of accurate prediction algorithms for medical big data is always required.

4. **Application to clinical sequencing**
   We need to apply our results to actual applied medicine, such as clinical sequencing.

Our theme is not restricted to the above. We think we have infinite possibilities to change the world of medical science, molecular biology, drug discovery, etc., by developing novel big data technologies for them.

Home page: [http://shibuyalab.hgc.jp](http://shibuyalab.hgc.jp)
Our laboratory’s research is focused on natural language technologies to assist human intelligent activities. Our major challenges include the following subjects in text and media studies that are based on machine learning including deep learning, statistical modeling and analysis, or annotation and corpus analysis.

(1) Text mining: document structure analysis and information analysis; information identification and entity disambiguation; automatic construction of linguistic resources

(2) Machine reading comprehension: Exploiting the semantic structure of natural language; extracting knowledge and information from natural language

(3) Modeling human language activities: Measuring and analyzing language activities in text media

Our recent research topics are as follows: machine reading comprehension, design and analysis of natural language understanding tasks, real document processing, retrieval and understanding of mathematical expressions, information identification and linkage, and recommendation and English writing assistance for academic researchers. We also welcome new research topic proposals related to natural language processing and information retrieval.

Our research facilities include variety of information resources and a large-scale computation platform at NII. Students are encouraged to act as independent researchers/engineers by being allowed to join seminars and discussion groups with interdisciplinary and international researchers and to participate in related joint research activities at NII.

http://www-al.nii.ac.jp/en/

Imari Sato Lab

Analyzing Light Transport for Scene Understanding

The appearance of an object changes significantly depending on its shapes, surface reflectance properties, and the lighting condition of the scene in which it is placed. Given an unknown picture, it might not so difficult for us to obtain different kinds of information about the objects in the image. However, it turns out to be an difficult task for a computer to figure out. In the past, techniques for automatically modeling and analyzing the photometric and geometric information of a scene have been studied in the fields of both computer vision and computer graphics research.

Furthermore, some researcher showed that the spectral reflectance and emission of objects provides innate information about material properties that have proven useful in applications such as classification, synthetic relighting, and medical imaging to name a few. In our lab, we analyze different types of optical phenomena including spectral reflectance, absorption, subsurface scattering, and fluorescence emission for modeling and understanding real scenes.

Message

Seeing is believing: the goal of computer vision is to understand how humans process and use visual information for understanding the surrounding world. There are so much excitement in analyzing real scenes using cutting edge technologies. Please join us and share a feeling of accomplishment with our lab members!

http://research.nii.ac.jp/pbv/index.html
Our laboratory focuses on "algorithms", "discrete mathematics", and their applications. In particular, we plan to study the following topics, focusing mainly on "graphs".

1. Develop theoretically fast and accurate algorithms for graph problems or prove NP-hardness for them

2. Develop scalable and accurate algorithms that can be implemented, using theoretical tools such as discrete mathematics, for graph datasets in the real world.

3. Develop algorithms that can run theoretically (or practically) quite fast (and accurate), when limiting graph family (for example, planar graphs or social networks).

4. Apply knowledge and implementation techniques of graph algorithms to machine learning, especially online learning and deep learning

5. Research on graphs that appear in the fields of natural language processing, machine learning, databases, data mining, programming languages, etc.

6. Combinatorial optimization and/or discrete mathematics

In our laboratory, we plan to interact with top researchers for theory in overseas for the above "theoretical" research, and jointly work with researchers in Japan and overseas, including industries, for other research topics.

Algorithm innovations such as current information retrieval technology (Google's PageRank) and security technology (Apple's (Local) Different Privacy) have led to the creation of national-scale businesses. What is important here is that both PageRank and Differential Privacy are basic and theoretical researches of algorithms and discrete mathematics, not application-oriented work from the beginning. In our laboratory as well, we plan to focus on basic and theoretical research, rather than conducting research with applications in mind from the beginning. E-mail: k_keniti@nii.ac.jp