NII Today

ISSN 1883-1974 (Print)
ISSN 1884-0787 (Online)
National Institute of Informatics News

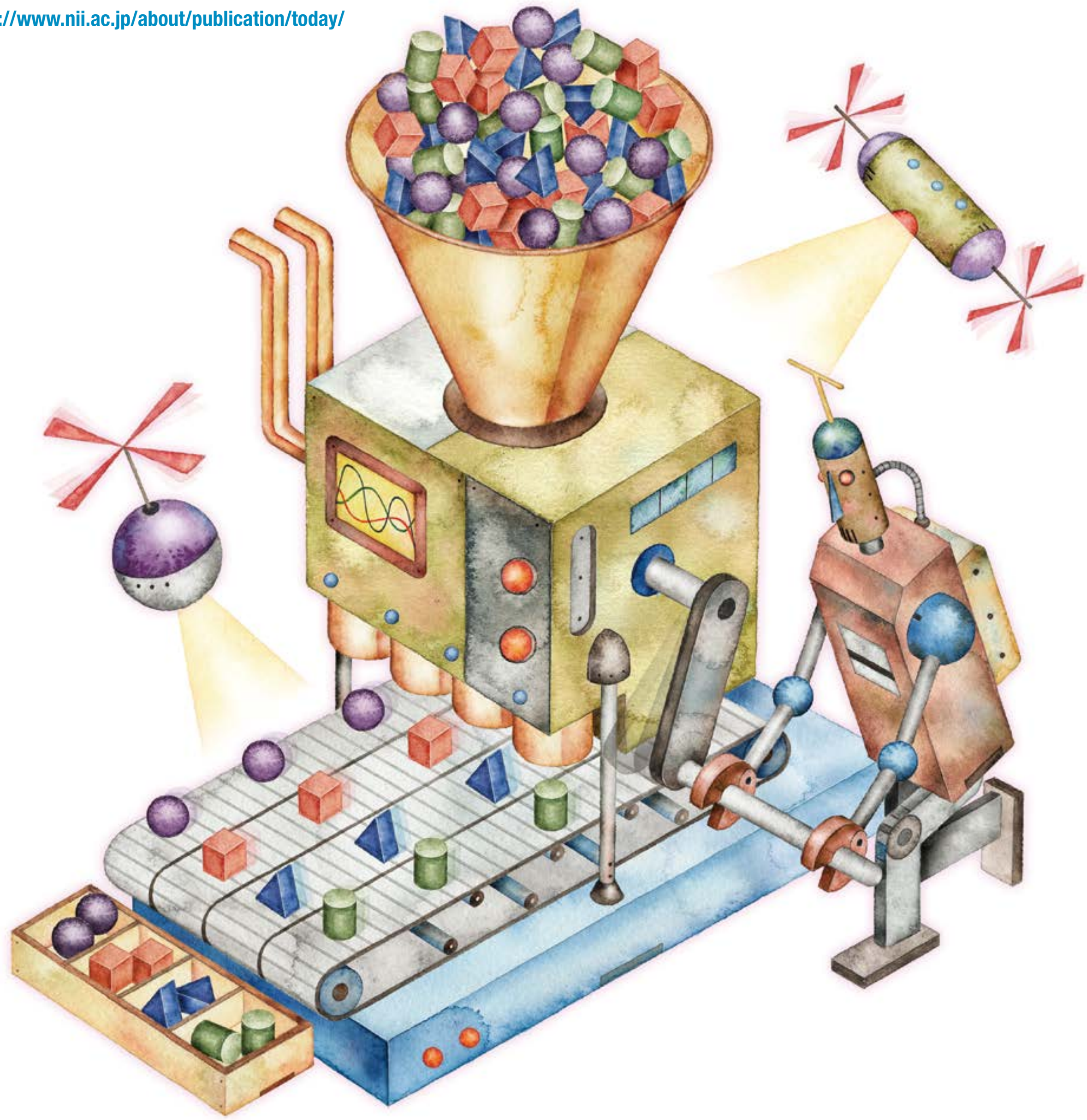
52
Jan. 2015

National Institute of Informatics News

Feature

Combining Research on Algorithms and Mathematics

A digital book version of "NII Today" is now available.
http://www.nii.ac.jp/about/publication/today/



This English language edition
NII Today corresponds to
No. 66 of the Japanese edition

NII Interview

Tackling the Challenges of a Big Data Society Using Algorithms and Mathematics

From Internet searches and purchase histories to smartphone location information, our everyday lives generate enormous amounts of data. Can this "big data" be used to create new value that will change society? We asked Takeaki Uno, a professor at the National Institute of Informatics (NII) who studies algorithms, and also Kenichi Kawarabayashi, a professor at NII who specializes in the discrete mathematics that underpins theoretical computer science, about how to tackle the challenges facing society in the age of big data.

Takita: It is said that the rapidly growing volume of data now exceeds computer processing power. What kinds of approaches do you think are necessary to analyze this data quickly and extract meaningful information?

Uno: First, we have to come up with ways to minimize the number of calculations. An algorithm is a plan for directing a computer to perform a particular calculation. We want to derive optimal solutions to real-life problems without performing huge calculations that use up the entire computer memory. It is also important to discard information-poor "noise" from the data. For example, if we are examining recorded data from a call center to try to determine whether a customer is angry, we may be able to tell this just from the voice in the first syllable of "Hello". This means that the rest of the data can be discarded.

Kawarabayashi: The question of where we should focus to enable us to distinguish between data that are unnecessary and data that should be kept can, I think, be approached from a mathematical perspective and formularized. I am studying "graph structure", which consists of sets of vertices and edges, like a rail or highway network, so when I look at data, I can see that structure. I can tell what is important and what can be omitted. This is due in large part to many years of experience. My reason for being interested in working with algorithm researchers is to make use of this ability in a way that will benefit society.

Takita: I understand that you are also conducting joint research with business.

Kawarabayashi: Japanese manufacturing is shifting from semiconductor development to information and communications technology (ICT); however, in ICT, there is as yet no basic theory like the theory behind semiconductors. The experience of success in semiconductor development, in which manufacturing technology was developed upon a complete basic theory, has no meaning in ICT. I think that scientists must take on the fundamental areas of theoretical research more.

Uno: Business people tend to understand the meaning of data intuitively, but as the volume of data increases, it becomes necessary to verbalize the meaning using numbers or mathematics. That is difficult to do. As researchers, we aim to build trusting relationships with companies and provide good mathematical models.

Takita: The United States is overwhelmingly strong in the field of ICT. Does Japan stand a chance?

Kawarabayashi: The 2020 Tokyo Olympics will be a game changer for Japan. A huge volume of data in the form of videos, photographs, and so on will be amassed during the Olympics. The question is how to make use of it.

Takeaki Uno
Professor, Principles of Informatics Research Division, National Institute of Informatics
Professor, Department of Informatics, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Ken-ichi Kawarabayashi
Professor, Principles of Informatics Research Division, National Institute of Informatics
Professor, Department of Informatics, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Uno: The keyword of the Olympic bid was omotenashi, which is the Japanese approach to hospitality and service, and this omotenashi involves information. A shop assistant determines what a customer wants from their appearance and conversation. We can perhaps create new services by verbalizing this kind of careful attention and on-site information using mathematics. For example, at a sporting venue, when the temperature rises, signs around the venue could show where to buy drinks, or, if information from cameras indicates that there are large numbers of children present, the signs could show the children's play areas.

Kawarabayashi: The development of ICT infrastructure for the Olympics is important, but the key to the future will be the extent to which human resources capable of utilizing the collected data are developed. A mathematical mindset is acquired partly through years of practice as a researcher and tackling many problems. From a business perspective, this may not yield benefits in the short term. However, in the long term, the basic strength of a company will be completely different depending on whether it has this kind of human resource. The companies in the United States that have succeeded in ICT are those where the working environment allows employees the time to enjoy research in a free atmosphere.

Uno: In the manufacturing and service industries, everyone is required to raise productivity in the same way. In contrast, the scientific world that is fundamental to ICT is similar to the arts or sports, where if one person in ten produces a good result, that's fine. That doesn't mean that the same result could be obtained by just employing the top person. It is only possible to reach the summit because the other nine people are there. In a business that depends on a star, the other employees can survive. This way of thinking doesn't exist in Japanese companies.

Takita: So, there is a need for development of human resources with a sense of mathematics.

Kawarabayashi: The question of whether mathematics is useful in the real world has long been debated, and I think that these days everyone agrees that it is useful. However, people do not consider mathematicians to be useful. That perception has to be changed. Of course, there have been various problems on the part of earlier generations of mathematicians. It is not good for mathematicians to ask business people questions such as "What is the definition of that, in terms of mathematics?" They must discover, assimilate, and formularize the problems facing society themselves. By doing this, I hope that mathematicians will show that they can solve social problems.

Uno: Until now, mathematics has involved trying to prove a particular phenomenon, with the work ending when a solution was found. That is of no use in the industrial world. What is important is putting the way of thinking and the way of looking at things while proving the phenomenon into practical use. Mathematics is often used with a view to increasing business efficiency, but going forward, it should be used to create new value. Mathematics is not simply calculations, it is deeply connected to human creativity. I think that communicating the appeal of mathematics will change the way it is viewed in society.

A Word from the Interviewer

Although a big data society in which computers present us with recommended clothes, books, and travel destinations is convenient, it also has a certain emptiness. Sometimes we feel uneasy about speeds of analysis that are far from our natural senses. I hope that incorporating subjective on-site views and the spirit of omotenashi into algorithms and mathematics will create a big data society that has people, rather than computers, at its core.

Kyoko Takita
Editorial Writer and Staff Writer, Yomiuri Shimbun Tokyo Headquarters

Joined the Yomiuri Shimbun after graduating from the Faculty of Foreign Studies, Sophia University. After working in the Hachioji branch office, graduated from the Graduate School of Journalism, University of California, Berkeley, in 2000. Worked in Yomiuri Shimbun's science department from 2002, responsible for reporting on science and technology policies, IT, space, environment, medicine, and disasters. Covered the Great East Japan Earthquake. Present post since 2014.

The Path to the Development of the World's Fastest Data Analysis Algorithm

Mathematics, Algorithms, and Passion Open New Doors

That's **Collaboration**

Data mining is a heuristic data analysis technique used to discover and classify frequent patterns and correlations in huge volumes of data in order to identify trends or even predict the future. Professor Hiroki Arimura and Professor Takeaki Uno work at the forefront of research on data mining. In 2004, the "world's fastest" data analysis program developed by them both won an international contest for data mining implementations. Since then, they have continued to contribute greatly to the analysis of big data by working hand in hand to develop numerous data analysis algorithms. We spoke with them about the progress of their research. They reveal the close relationship between mathematics, algorithms, and application.

Hiroki Arimura
Visiting Professor, National Institute of Informatics
Professor, Division of Computer Science, Graduate School of Information Science and Technology, Hokkaido University

Takeaki Uno
Professor, Principles of Informatics Research Division, National Institute of Informatics
Professor, Department of Informatics, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

An Encounter Between a Proponent of "Knowledge Creation Science" and a Mathematics Researcher

— Tell us how you came to collaborate together in your research.

Arimura: I was a kid who loved science, but I wasn't actually very good at mathematics. However, at some point, I became aware that there was a different type of mathematics to the basic mathematics that deals with changes in continuous quantities. That was "discrete mathematics", which investigates the relationships between things and puts quality before quantity, so to speak. A visit to Professor Setsuo Arikawa, an expert in computer science at Kyushu University who later became the President of Kyushu University, was what led me to become involved in my current research. From my conversation with Professor Arikawa, I learned that it is possible to prove mathematically what can and cannot be done using a computer, and also that capturing things discretely without doing calculations such as differentiation and integration is very useful to society. I became captivated by discrete mathematics—the mathematics of things. Although at that time I had already found other employment, I decided to instead join Professor Arikawa's laboratory and become a researcher. I feel like I have spent my time since then enjoying discrete mathematics and its application, without ever being aware that I was studying.

Later, around 1996, I encountered data mining. In the United States, research was being done that looked at both the entirety of and the individual details of huge volumes of accumulated data. Because this was attractive to me, I started working on research into artificial intelligence and machine learning. Since moving to Hokkaido University, I have cultivated relationships with scientific researchers in different fields, and I am promoting collaborative research by applying data mining techniques to research projects that straddle disciplines. These projects include the collection and analysis of a colossal amount of genetic information on organisms, the integration of morphological information/genetic information on deep-sea organisms obtained from deep-sea research vessels, and data analysis related to robots and communication. I have given the term "knowledge creation science" to research that creates new knowledge by collecting findings from different fields, which is what I am working on now.

Uno: As for me, from elementary school age, I loved puzzles and was keen on creating computer programs. I was also interested in the mathematics behind the programs, the mathematical optimization. Mathematical optimization is finding optimal solutions within a certain set rules and data framework. Even if the answer found is not the final solution, it might be sufficient to solve the problem.

Incidentally, I think that there are two types of mathematics researcher—"mathematical theory" researchers and "application" researchers. The difference between them is in how they think about "things". The former devise methods of finding ultimate answers with no basis on "things". The latter, based on the premise that there are data, devise a sequence of steps for processing the data within a finite time, and apply these steps to finding useful answers. There will not necessarily be just one answer, so the application researcher focuses on structures for deriving optimal solutions with limited time and scarce labor.

This is the basis of algorithms. Algorithms are a part of mathematics, but algorithm research differs from mathematical theory research in that it includes the concepts of time and process; in mathematics, an equation is always true, but in algorithms, equations and status can change as time and process go. A program can be considered to be a collection of algorithms, and it is this kind of research that I have always been involved in.

I was introduced to Professor Arimura by Professor Ken Satoh of NII. Professor Satoh hoped that amazing things could be achieved by bringing together Professor Arimura, with his awareness of issues in a wide range of areas, and myself, whose strength is in mathematical research aimed at problem resolution—in other words, someone who finds problems and someone who finds solutions. Since then, we have conducted our research by Professor Arimura raising questions and myself proposing solutions. Professor Satoh also made another great suggestion. He encouraged us to participate in an international data mining implementation contest that was held in the UK at that time.

The Fastest Algorithm—Discovered between Mathematics and Application

— And I understand that you won that international contest.

Arimura: The contest was to determine how fast a provided problem could be solved using a data mining program. In the world of data mining, the degree of sophistication of the algorithm can produce thousand-fold differences in speed. Professor Satoh said that if the high-speed algorithm "MAximal Clique Enumerator" developed by Professor Uno was applied to data mining, it definitely stood a chance of winning the contest. Professor Uno's immediate reply was, "We can do it!"

Uno: Because the application of that algorithm would enable frequent patterns to be derived from a database at high speed. In other words, it would enable meaningful information to be easily extracted from huge volumes of data.

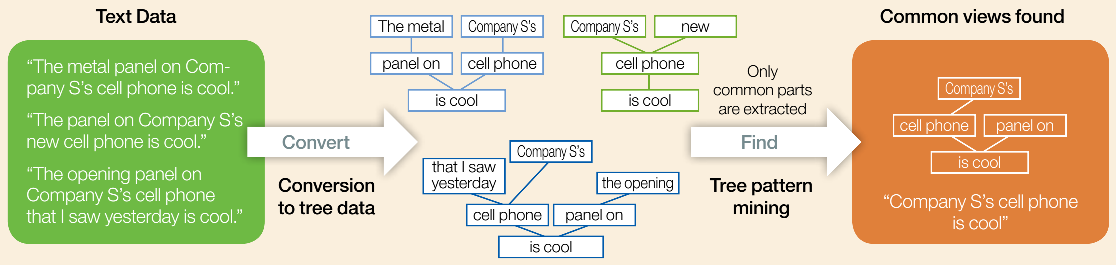
Arimura: When we first attempted it in 2003, Professor Uno researched the algorithm and I researched the model. At that time, we didn't win first place.

Uno: We had confidence in the algorithm, but lost because we straightforwardly processed all of the data. Most of the other participants were database engineers skilled in data handling, so they discarded the unnecessary parts of the huge volume of data at the data loading stage.

Arimura: In 2004, after reflecting on what had occurred, we worked on tuning the data. Professor Uno did the coding while I and a coworker ran experiments. We ended up doing experiments and coding in three shifts around the clock. Even when Professor Uno was on work trips overseas, we would send him the experimental results and then he would return the corrected results eight hours later. That hard work led to our fantastic win.

Our program LCM (Linear-time Closed itemset Miner) is publicly available. The algorithm used in the program is still said to be the world's fastest in the field of association data mining. This experience taught us that to open new doors, in addition to mathematics and algorithms, you also need passion. (Laughs)

Figure 1 Example of application of tree mining to text mining



The Latest Techniques for Extracting Useful Information from Big Data

— Could you tell us about your most recent research topics?

Arimura: The spotlight these days is on big data analysis. Huge volumes of diverse data are being generated and stored, such as SNS posts and links, mobile terminal location information, and data from all kinds of sensors installed in buildings. With these kinds of data, it is difficult to extract knowledge that humans can understand. Therefore, the aim of data mining technology is to extract information of a size (granularity) that is understandable to human beings using methods such as clustering, which finds frequent patterns and determines similarities, and decision trees, which classify data by dividing them into cases.

For example, we developed an algorithm that detects characteristic substructures as small tree patterns in web text data, graph structures of chemical compounds, dependency tree structures of natural language, and so on. Figure 1 shows an example of finding a template structure that appears repeatedly in the tree structures of an online database.

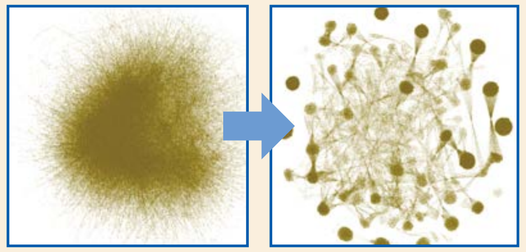
Uno: To obtain meaningful information from big data, the data have to be polished and classified into broad clusters for the data to be easy to understand. One method of doing this is to focus on the connections between the data. Figure 2 uses large-scale data on business-to-business transactions to show relationships between companies in graph form. The dots represent the companies and the lines represent transactions. The figure on the left is a plot of the actual data, and the figure on the right shows the result of polishing the data with an algorithm that we released recently. The business relationships are easier to understand in the right-hand figure. This algorithm can also be applied to searching for "communities composed of friends" on SNS. For example, analyzing 1 million items of data requires around 1 trillion data comparisons, but using this algorithm, it requires around 100 million data comparisons at most. The appeal of algorithm research is that it can be carried out using an ordinary computer, rather than a supercomputer.

Arimura: The current mainstream method for standard data analysis is to build a hypothesis and then test it, but this method has its limits when you are dealing with big data. We want to let the data speak for itself about the mathematics concealed within it in a way that humans can understand. I think that this dream can be realized by pushing forward with research into algorithms and improving the accuracy of machine learning.

(Interview/Report by Masahiro Doi)

Figure 2 Example of facilitating understanding of business-to-business transaction data by clustering

	Before polishing	After polishing
No. of vertices	3,282	3,282
No. of branches	35,168	73,132
No. of cliques	32,953	343



Column

What is algorithm?

Do you know the origin of the term "algorithm"? The term is said to come from the name of a mathematician active in Baghdad in the first half of the 9th century, Al-Khwarizmi, or in full, Abu Abdullah Muhammad ibn Musa Al-Khwarizmi*. Since the Latin translation of his textbook on calculation methods in India began with the phrase "algoritmi dicti" ("so says AlKhwarizmi"), it became the origin of the term "algorithm".

There have been algorithms in existence that were invented before the birth of computers, and even before the Common Era (BCE). A typical example is the Euclidean algorithm. This is a procedure for finding the greatest common divisor of two integers, and it is something that you learned about in your junior high school age. The Euclidean algorithm is still used in modern society: it forms the basis of public-key cryptosystem that supports secure communications, as well as being used in digital signatures. The inventor of the algorithm, Euclid, is best known for Euclidean geometry. He was a mathematician in the 3rd century BCE, and could surely never have dreamt that his algorithm would contribute to future society more than 2,000 years after his time.

The modern definition of algorithm as an expression that sets out a procedure for solving a problem in fields such as computing or mathematics did not come about until the 1930s. However, algorithms do not have to be presented as programs. For example, the way of doing multiple-digit addition, multiplication, and division by hand that is learned in elementary school can be described as algorithms because the answers can be reached by following certain procedures. Cooking recipes can also be described as algorithms if they are looked at as procedures for solving problems for making particular dishes. The process for delivering postal mail can also be described as an algorithm in the sense that the address and zip code on a postcard or letter is read, then it is sent between post offices, and finally it is delivered to the address according to a fixed procedure. Also, the delivery of Internet communication packets was initially designed based on the postal service procedure, a.k.a. the postal service algorithm, so an algorithm in the real-world has become a useful algorithm for telecommunications.

Then, what is algorithm research? Algorithm research includes studies aimed at finding solutions to problems that have not yet been solved, as well as studies to reduce the time and cost required to solve the problems. Algorithms for processing large volumes of data, such as scientific data and "big data", can take days rather than hours. Speeding up these algorithms, even if only slightly, is a very significant achievement.

As various algorithms come to be used in the real and cyber world around us, the term "algorithm" is also starting to be used more widely. Up until a decade or so ago, it was mainly used as a technical term in computer science, but now the term appears in our daily lives and society—from the "Algorithm March" (a Japanese dance fad) on TV, to algorithmic trading used in financial transactions. Over the next decade, what algorithms will be invented and in what meaning will the term "algorithm" be used?

(Written by Ichiro Satoh)

* The word "algebra" is also said to derive from the Latin transliteration of this mathematician's name.

The "Algorithm March", a dance that is familiar to Japanese TV viewers.

Crossing Boundaries to Develop Better Algorithms

Using Findings from Network Science and Theoretical Computer Science

Yuichi Yoshida
Assistant Professor, Principles of Informatics Research Division, National Institute of Informatics

Assistant Professor Yuichi Yoshida's research draws upon results in theoretical computer science to develop high-speed algorithms. His main research theme is constant-time algorithms, which are amazingly powerful in solving real-world problems. Using constant-time algorithms, computational problems that would take tens of thousands of years to solve by conventional methods can be completed in just a few milliseconds. The key to his achievements was taking notice of findings in fields "next door" to his own area of specialization.

Achieving Results by Noticing Findings in Fields "Next Door"

Assistant Professor Yuichi Yoshida's area of specialization is theoretical computer science. This is a discipline that, for example, investigates the amount of computation required if a certain problem is to be solved using a computer and examines whether the problem is solvable or unsolvable.

Assistant Professor Yoshida is currently branching out to research into "practical algorithms", i.e., algorithms that are high speed when run in practice using a computer. This is because when he looked at the world of algorithms in computer science from his theoretical computer science perspective, he noticed that it was "amazingly untouched". In other words, it looked like a treasure trove of unsolved problems.

For example, research on high-speed algorithms was being done in the field of databases, but there were many untouched areas of research in explaining theoretically why such algorithms are high speed.

Assistant Professor Yoshida has said, "If we can achieve theory-based high-speed algorithms, the usability of algorithms will increase. The development of a theory would also be a great stimulus for researchers in computer science dealing with real computers. I want to stimulate researchers in traditional fields by introducing theory into algorithms, just as Commodore Perry's 'black ships' stimulated Japan into modernizing."

The background to Assistant Professor Yoshida's comment is a flourishing of research that crosses over the boundaries between different disciplines.

"Coming into contact with various fields is essential for achieving breakthroughs. It is easier to find good problems and good tools in different fields for the very reason that they are different. In fact, researchers in the mathematical disciplines of representation theory and algebraic geometry sometimes solve problems in theoretical computer science. Thus, they are sometimes the 'black ships' visiting us! (Laughs)"

Physics is another example. Currently, many physicists are interested in complex networks and are studying their properties. One concept that has come from this area of research is betweenness centrality, which evaluates the importance of a graph's vertices by the number of shortest paths that pass through them (Figure 1) (Figure 2). If this concept were used in modeling infectious disease epidemics, for example, its scope of application would be wide, but the amount of computation would be very large, and it is therefore a potential subject for high-speed algorithm research. Assistant Professor Yoshida stresses, "It would be a waste not to use findings from other fields in algorithm research."

The Power of Constant-Time Algorithms

The main theme of Assistant Professor Yoshida's research is constant-time algorithms. The fundamental idea behind this kind of algorithm is to solve problems in a fixed time, regardless of problem size. This means that the more enormous the problem becomes, the more the constant-time algorithm displays its formidable power.

For example, if the SNS Facebook is modeled as a graph, it has approximately 1 billion vertices, which is equivalent to its number of users. To accurately determine the "diameter" of this graph (the greatest distance among all combinations of vertices on the graph) would require a computation time of around 20,000 years, assuming a computer processing speed of 100 million steps per second. This makes it impossible to solve the problem in practice.

However, if an error of $\varepsilon = 0.01$, i.e., an error of 1%, is permitted and a constant-time algorithm is used, the same problem can be solved almost instantaneously in 1.2 milliseconds. Reducing the computation time from 20,000 years to 1.2 milliseconds means that a problem that could not be realistically solved becomes solvable. This is the power of constant-time algorithms.

If these algorithms can be applied to the real world, processing that was impossible before will become possible. The concept will no doubt have tremendous potential in terms of usability as well.

However, Assistant Professor Yoshida's interest lies in theory rather than application to the real world. His most recent research investigated the nature of problems that can and cannot be solved in constant time, and he succeeded in "providing the necessary and sufficient conditions". He proved a theory concerning the essence of constant-time algorithms, using in the course of his proof the mathematical tool harmonic analysis, meaning that he introduced a finding from mathematics into theoretical computer science. Assistant Professor Yoshida says, "This is another example of research that makes use of results from a field 'next door'."

Connections Made through Programming Contests

Currently, Assistant Professor Yoshida acts as group leader of the Complex Network and Map Graph Group in the Exploratory Research for Advanced Technology (ERATO) Kawarabayashi Large Graph Project. The group members are algorithm experts who have participated in programming contests.

Pruned Landmark Labeling is a new algorithm created by this research group that finds the shortest distance connecting any two points on a graph at a speed one hundred times faster than the existing methods. The group incorporated observations from the real-world Web into this algorithm. A webgraph, which models Web links, has a structure made up of a "core", where many vertices cluster, and a "fringe", which is the peripheral region (Figure 3). The shortest path between two points on the graph will often pass through the core. The group made use of a theory based on this finding to succeed in increasing the speed of the algorithm.

A paper on this algorithm was written by Assistant Professor Yoshida and two doctoral students at the University of Tokyo, Takuya Akiba and Yoichi Iwata. The three had all taken part in the ACM International Collegiate Programming Contest (ACM-ICPC). These young people, who understand algorithms and are trained to create programs, are now standing at the forefront of algorithm research in theoretical computer science.

Algorithm research based on the latest theories has the potential to greatly expand the scope of what computers can do, thereby bring about dramatic changes in the real world. As IT spreads into every corner of our society, expectations of achievements in theoretical computer science continue to grow.

(Interview/Report by Akio Hoshi)

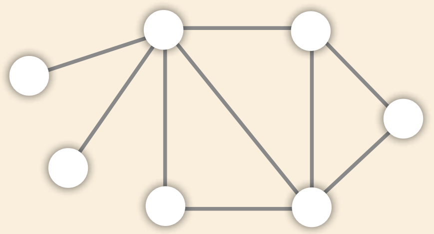


Figure 1 A graph consisting of 7 vertices and 9 edges. Graph theory examines the nature of graphs of this kind made up of vertices and connecting edges.

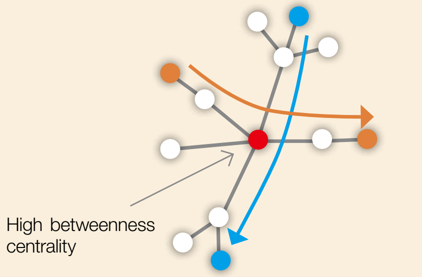


Figure 2 Overview of betweenness centrality. Both the shortest path between the orange vertices and the shortest path between the blue vertices pass through the red vertex, meaning that the red vertex has high betweenness centrality.

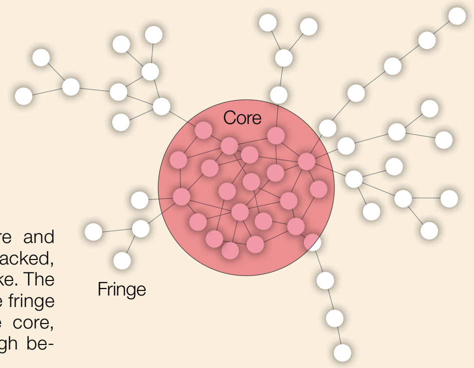


Figure 3 Overview of core and fringe. The core is densely packed, whereas the fringe is tree-like. The shortest paths between the fringe vertices pass through the core, giving the core vertices high betweenness centrality.

NII Special

2

Extracting New Algorithms from the Brain

Integration of Neuroscience and Informatics, and Approaching New Technologies for Information Processing

Dr. Ryota Kobayashi in NII is attempting to extract new algorithms from the brain by applying informatics. Results of the research are expected to have applications in medical technology, such as in treatments for incurable neurological disorders. Algorithms extracted from the brain will enable us to develop an efficient information-processing device in the near future.

Extracting information in the brain by brain activity measurement

In the United States, computer scientists have been closely collaborating with neuroscientists for a decade or more.

According to Dr. Ryota Kobayashi in NII, “One of the difficulties in neuroscience is measuring what is going on in the brain. Hypotheses have been proposed and tested through collaboration between computer scientists and neuroscientists.”

For example, when functional magnetic resonance imaging (fMRI) signals were measured in a subject viewing alphabet letters, it was possible to predict what letter he or she was looking at just from the fMRI signals (Miyawaki et al., Neuron, 2008). This means that we can read out information in the brain through measurements. This work is a result of the integration of fMRI measurement techniques in neuroscience and data mining in informatics.

Due to inspiration from the brain, an artificial intelligence algorithm called “deep learning” has been developed in recent years. Deep learning is outperforming traditional algorithms in many tasks, including image processing and speech recognition.

However, according to Dr. Kobayashi, “While computer scientists and neuroscientists have collaborated closely in the US and Europe, there are few collaborations in Japan.” Compared to the US and Europe, there are unfortunately fewer efforts to initiate government-led research projects in brain informatics in Japan.

Ryota Kobayashi
Assistant Professor, Principles of Informatics Research Division, National Institute of Informatics
Assistant Professor, Department of Informatics, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies

Utilizing Simulation Results for Medical Treatment

Dr. Kobayashi has studied a mathematical model of single neurons (elements of the brain) and the synapses that connect neurons. The research results will help to overcome incurable neurological disorders, and to develop new algorithms for information processing.

Dr. Kobayashi has developed the multi-timescale adaptive threshold (MAT) model, which won an international competition for predicting neural spikes that was organized by the Swiss Federal Institute of Technology in Lausanne (EPFL) from 2007 to 2009. The aim of the competition was to predict spike timings of a real neuron, and the participants competed on the accuracy of their predictions. Spike timings are communication signals used in the brain. Dr. Kobayashi’s neuron model can predict the spikes with a high degree of accuracy (90%), one which far exceeds that of the Nobel Prize-winning Hodgkin–Huxley model (60%) (Kobayashi et al., Front. Comput. Neurosci., 2009). Consequently, he received worldwide attention for building a realistic model neuron.

Dr. Kobayashi is also developing a method for estimating how neurons process information, as well as how information is transmitted between neurons by synapses.

He said about the aim of the research: “How strongly do synapses connect? What is the function of the synaptic connectivity? Investigating these questions will improve efforts to simulate brain activity realistically and the brain simulation will be applicable to various fields.”

One example of the application is medical treatment of Parkinson’s disease.

Symptoms of Parkinson’s disease can be reduced by implanting electrodes in the patient’s brain to stimulate the areas that are behaving abnormally. This is known as deep brain stimulation (DBS) and is a treatment that is attracting much attention. Research on the simulation of the brain activity during DBS has been started, and the results might guide medical researchers to more efficient treatments.

Informatics can also be applied to diagnosing Alzheimer’s disease and epilepsy—conditions that have traditionally been difficult to diagnose—by analyzing electroencephalogram (EEG) data. The rate of correct diagnosis is particularly high for Alzheimer’s disease, at 83% (Dauwels et al., Neuroimage, 2010), and early diagnosis/medication has the potential to extend the period during which a patient is able to live a normal life.

According to Dr. Kobayashi, “In addition to diseases of the brain, the simulation study was launched to understand the mechanisms of mental diseases, such as bipolar disorder and autism.”

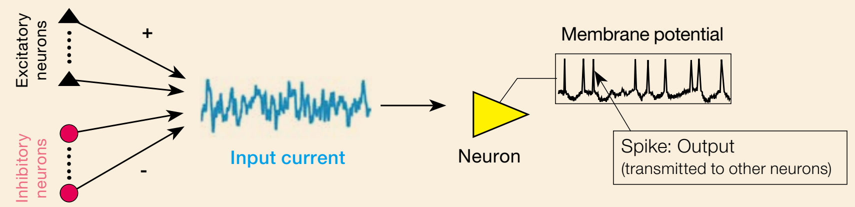


Figure 1 Information processing of a neuron. A neuron receives input currents from thousands of excitatory and inhibitory neurons. It processes information by generating spikes and transmitting information of spike timings to other neurons.

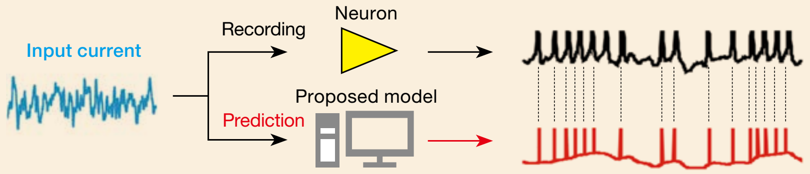


Figure 2 Predicting spike timings of a neuron. The upper panel shows membrane potential recorded from a cortical neuron (black), while the lower panel shows simulated results by using the proposed model (red). Spikes predicted within an error of 0.002 seconds are connected by dotted lines. This model can predict approximately 90% of spikes, and it won an international competition for predicting neural spikes organized by the Swiss Federal Institute of Technology in Lausanne (EPFL). This work is a collaborative project with Shigeru Shinomoto (Kyoto University) and Yasuhiro Tsubo (Ritsumeikan University). For model details, see Kobayashi, Tsubo, Shinomoto, Front. Comput. Neurosci., 2009; for contest details, see Gerstner & Naud, Science, 2009.

Developing New Computing Models

Research on brain activity will lead to the creation of new computing models.

Dr. Kobayashi said, “There are 200 billion neurons and 1 quadrillion synapses in the human brain, and even using the K computer, simulating one second of brain activity in 1% of the human brain requires 40 minutes for computation. This means that the human brain has much more computational resources than the supercomputer. The synapses between the neurons change during the learning process. If we can extract learning algorithms from the brain, we may be able to create new computing models.”

Facing the age of Big Data, enormous computing power is required to find knowledge from exploring data. However, if algorithms based on the brain are developed, the algorithms will be able not only to extract features that are impossible with conventional algorithms, but also to process information in a more sophisticated way. Also, if we can imitate computations in the brain, the development of an energy-efficient computer will no longer be just a dream: high-performance computers that are not based on enormous power consumption will be achieved.

More close collaboration between computer scientists and neuroscientists will facilitate the extraction of new algorithms from the brain, which will enrich our lives, such as through their application to medical treatments.

(Interview/Report by Katsuyuki Ohkawara)

NII essay

Richardson's Dream

Ken Hayami

Professor, Principles of Informatics Research Division, National Institute of Informatics

Two late-season typhoons swept across the Japanese archipelago. I felt their power, and at the same time felt that the forecast of a typhoon's path has become more accurate compared to in the past.

The emergence of numerical weather prediction.

A long time ago, even before the age of electronic computers, the English mathematician and meteorologist Richardson conceived a numerical method for predicting the weather based on equations governing fluids, and tried to realize it. He would discretize space into cells, and compute so that air would flow from a cell with high pressure to a neighbouring cell with lower pressure. However, it was impossible to compute quickly enough for prediction by manual computation relying on slide rules: it took six weeks to predict the weather just six hours ahead. What was worse, the prediction was far from the actual weather.

So what he dreamt of was to place 60,000 human "computers" seated in a large theatre-like building, each assigned to the computations for one cell, under the direction of a "conductor", such that the whole computation could be performed in parallel and speeded up sufficiently for prediction. This is what is called "Richardson's dream".

The role of mathematical models and algorithms.

This dream materialized with the advent of electronic computers, and nowadays is realized by supercomputers and parallel computers, so that fairly accurate weather forecasts and medium-range climate predictions have become possible. However, there is another reason why Richardson's dream was not realized within his own time. His equations (mathematical model) did not describe actual weather phenomena appropriately. With the development of computers, the mathematical model was also continuously improved.

Moreover, methods were developed to augment data when the observed data, on which the prediction is based, are incomplete. Also, numerical algorithms which fully utilize the power of parallel computers were developed. It is the combination of these developments in mathematical models, algorithms, and computers that has made today's weather and climate prediction possible.

As this example shows, predictions using computers become possible through joint efforts in mathematical models, algorithms, and computers.

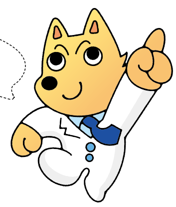
By the way, the interests of Richardson were not limited to weather. Being a Quaker and a pacifist, he analyzed wars by mathematical models and developed the first notion of the mathematical concept of fractals through his analysis of the lengths of borders between countries. It is interesting to note that he applied mathematics to various problems in real life, and developed new mathematics and algorithms through doing so.

Weaving Information into Knowledge
NII

National Institute of Informatics News [NII Today]

No. 52 Jan. 2015 [This English language edition NII Today corresponds to No. 66 of the Japanese edition.]
Published by National Institute of Informatics, Research Organization of Information and Systems
Address: National Center of Sciences 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430
Publisher: Masaru Kitsuregawa Editorial Supervisor: Ichiro Satoh
Cover illustration: Toshiya Shirotani Photography: Yusuke Sato
Copy Editor: Madoka Tainaka Production: Nobudget Inc.
Contact: Publicity Team, Planning Division, General Affairs Department
TEL: +81-3-4212-2164 FAX: +81-3-4212-2150 E-mail: kouhou@nii.ac.jp

A digital book version of NII Today is now available!



Bit (NII Character)

http://www.nii.ac.jp/about/publication/today/