ISSN 1883-1974 National Institute of Informatics News
No.45 Apr. 2013
This English language edition NII Today corresponds to No.59 of the Japanese edition

NII Today
National Institute of Informatics News

FEATURED TOPIC | **The New World Being Created by Ultra-High Speed Database Engines**

Ultrafast Database Engine will Change

It is said that we are now in the age of "big data". "Development of Ultrafast database engine" was selected by the Cabinet Office as one of the 30 research projects to be supported by the Funding Program for World-Leading Innovative R&D on Science and Technology, whose goal is to take a leading global position. The research project aims to develop the technologies that will serve as the core of big data processing, and investigate their application into society. Professor Kitsuregawa, the principal investigator, discussed with us the position of the age of big data within the history of computer development, how it will change future society, and what role the new technologies will play.

Tsuji: What is behind the current interest in "big data"?

Kitsuregawa: There has been an increased focus on big data since President Obama announced at the end of March, 2012, that America would be investing 200 million dollars in big data projects. Big data has an impact on society equivalent in scale to the impact the Internet has had. In other words, it has the potential to fundamentally change not only science and industry, but society itself. He declared this is why he would set up a national project. This created a big data boom around the world. I believe his message provided an insight into the views of the White House.

Tsuji: So you think it was a historic message?

Kitsuregawa: Yes. When Thomas Watson, founder of IBM, began selling computers, he said "I think there is a world market for maybe five computers". 100 years later, analytical technologies have progressed to the point where they are today, and I think we need to reflect on the weight of that history. A good example is "Watson", the computer developed by IBM which became reigning champion on "Jeopardy!" I think it's an outstanding achievement that showed how data can become knowledge.

Tsuji: Your project, part of the Cabinet Office's Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST)(*1) was underway before the announcement by the American government, right?

Kitsuregawa: They selected my project in 2009, before big data became a global phenomenon, and for that I am very grateful. What the project calls the "age of ultra-large database" is what would be called the "age of big data" in modern parlance. It was the only IT related theme out of all the 30 FIRST themes, and I am acutely aware of the great deal of responsibility that entails. Members of industry and researchers from universities across the nation are working in the project, which is being supported by the University of Tokyo and NII.

Tsuji: So what is the age of big data?

Kitsuregawa: In America, it's also called the "age of observation"; the world could be made "observable" that is, advanced sensors will be able to observe industries from health care to transportation and electrical systems, with a great deal of detail. It will provide very close views of a wide range of activities. This expansive set of data is called "big data". This big data will be uploaded to the cloud and analyzed there, and then fed back into the real world. I think many business segments will shift to

Masaru Kitsuregawa
Director General, National Institute of Informatics
Professor, Institute of Industrial Science, The University of Tokyo

Atsuko Tsuji
Staff writer at Op-Ed Section, the Asahi Shimbun
Bachelor's Degree, University of Tokyo, She has been working as a science journalist at the Asahi Shimbun. She has occupied various positions such as a science correspondent at the Asahi Shimbun American General Bureau and an editorial writer of science before taking her current position in February, 2013. Her areas of coverage as a science journalist are science, technology, and medicine. She was granted a Knight Science Journalism Fellowship at MIT in 1989, and a Reuters Fellowship at Oxford University in 2004.

Society

this type of structure over roughly the next ten years.

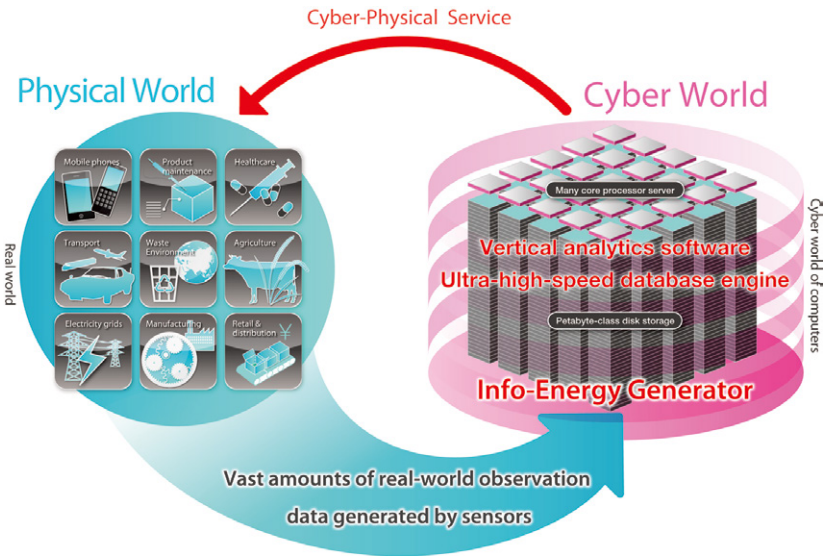
Tsuji: You have worked on the Information Explosion and Information Grand Voyage projects in the past. What are the goals of this project?

Kitsuregawa: Until now uplinking, sending data into the cloud, has taken far, far more time than downlinking, receiving data. Data insertion has been performed by so-called "sneakernet" — transporting it on physical media. According to Ministry of Internal Affairs and Communications statistics, the ratio of downloading to uploading is roughly 10 to 1. If the data inserted in the future becomes big data, this relationship will be reversed. So how could I contribute to this sweeping change? Originally I made a plan to develop ultrafast engine for processing database, the core of analytics, and to conduct demonstration experiments to see how their results could be made use of in society. However, after the change in political administration, our budget was cut by over half, so we decided to focus on the former part, and spend roughly 10% budget on the latter part.

Tsuji: What are ultrafast database engine?

Kitsuregawa: Increasing database engine speed by 1,000-fold would create a qualitative transformation that would change the world. That's why we set as our objective to increase performance 1,000-fold. We are taking a new approach, called out-of-order execution, to squeeze even greater performance out of the same hardware than conventional software. With the traditional sequential approach, programs always runs in the same order. No matter how many times you execute a program, it always performs the same deterministic processing steps in the same sequence. Our out-of-order execution performs those steps non-sequentially. That is, the execution order might be different from execution to execution. Of course, the final result could be the same. This non-determinism is source of the great performance.

Tsuji: That sounds almost like magic. How far have you gotten?



Fusing cyber information (the web) and sensor information (the real world) to create new value and test its effectiveness in the worlds of medicine, economics, and agriculture

Kitsuregawa: It's something like an egg of Columbus. We've already achieved 300-fold performance, and Hitachi, who we're working with, commercialized a product providing 100-fold performance in May of last year. Our goal is to achieve 1,000-fold performance on multi-core processors by March, 2014. I don't want that 1,000-fold figure to take on a life of its own, I'm just saying that our goal is to achieve speeds 1,000 times higher than Hitachi's current HiRDB. Also, it is impossible to achieve a 1,000-fold speed increase for queries(*2) only looking at a few records, and the conventional hash clustering approach is the best approach for queries reading entire databases, and cannot be made 1,000 times faster for such queries. Our approach will prove its effectiveness for the access necessary for the age of big data, not looking at small subsets, but not looking at entire datasets either.

Tsuji: How will it be used within society?

Kitsuregawa: I think it holds the most promise for the health care field. Demonstration experiments in developing countries are showing that providing data to individuals results in health improvements. Experiments performed in Japan as part of the Information Grand Voyage project indicated that clear communication of information is effective in improving health — that it serves as so-called "information medicine". However, people are complex, making them difficult to use as testing targetsand there is also a need to make changes to existing systems, so we cannot expect the kinds of rapid changes that we have seen with IT technology. We must get back to the basics of big data, collecting data, and work forward from there.

Tsuji: I'm really looking forward to seeing how your research results will change society.

Kitsuregawa: Perfect prediction is impossible regarding how our developed software could be finally contributing to society because our research is still on-going and the software is very partially released. The young university researchers and industry members engaged in the project have said that this research is fascinating, and they're all very excited to be working on it. I hope that the project, and the researchers, will be able to contribute to the age of big data.

A Word from the Interviewer

Big data is something I've often heard discussed, but wasn't very knowledgeable about. Throughout my interview with Professor Kitsuregawa, the excitement felt by the project's researchers was palpable. The interview impressed on me the need to understand its position within the overarching structure, as well as the need to see it from a historical perspective. Big data is a field of fierce international competition. I look forward to seeing the future of this great endeavor.

***1: Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST)**
A national research and development program whose objectives are the middle- and long-term improvement of Japan's underlying and competitive strengths in fields such as manufacturing and safety assurance through the promotion of world-leading research, as well as the steady and ongoing usage of the results of this research and development to benefit the Japanese people and Japanese society. 30 projects were selected from a pool of 565 applicants. The program began in March, 2010, and will end in March, 2014.

***2: Query**
Character string expressing the processing request (query) to be issued to a database. Queries are issued to systems as commands to perform operations on data such as searching, modification, and deletion.

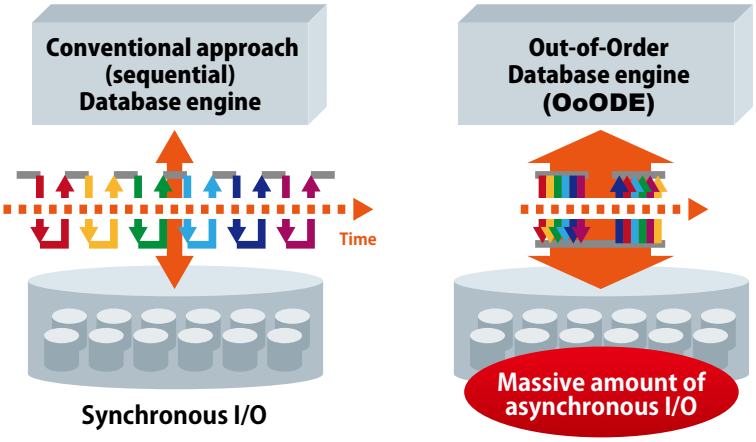
The New Principle of "Out-of-Order Makes it Possible to Increase Data Pro

Database engine which can quickly process massive amounts of data are essential for the big data utilization. In recent years, database engine technology has generally been considered as having reached maturity, but a team led by Professor Kitsuregawa is developing new ultrafast database engine based on the new principle of "out-of-order execution". We talked with Professor Kitsuregawa about the state of the Funding Program for World-Leading Innovative R&D on Science and Technology, which is striving to achieve search speeds 1000 times faster than conventional search technologies by using out-of-order execution.

Why is Out-of-Order Execution Fast?

The execution principle of out-of-order execution forms the core of the ultrafast database engine that has been developed in the Cabinet Office's Funding Program for World-Leading Innovative R&D on Science and Technology project for "Development of the Fastest Database Engine for the Era of Very Large Database and Experiment and Evaluation of Strategic Social Services Enabled by the Database Engine ". This project is being carried out by a team of researchers, led by Professor Kitsuregawa as a principal investigator. What is this execution principle that makes it possible to achieve data processing speeds 1000 times faster than conventional processing approaches? What makes out-of-order execution so fast? Professor Kitsuregawa explains.

"At present, database processing is generally based on the principle of sequential processing. When a command is received, it is executed. When it is completed, the system then executes the next command. This process is repeated in order. This means that the commands must be issued in order. With out-of-order execution, however, all of the commands to be executed are issued at once. You can issue data requests before the system completes accesses of other data. This is how the data may come back out of order, but that's not a problem. As long as everything matches up in the end, then it's fine — that's the concept. That's what makes it fast."



(Left) Conventional database engine, and (Right) Out-of-Order database engine

With conventional sequential execution, data is normally processed in a pre-defined order. When one part is performing a process, other parts must wait. Most of the time taken in data processing actually consists of this wait time, with hardware left unused and inactive. Dramatic processing time reductions could be achieved if many parts functioned simultaneously.

Rewriting Source Code from the Ground Up

When one hears about processing multiple parts at the same time, "distributed processing" and "parallel processing" pop to mind. However, according to Preofessor Kitsuregawa, out-of-order execution is entirely different.

"Both distributed computing and parallel computing divides a given processing into multiple processing instances, each of which executes sequential processing. The processing itself is performed in the pre-defined order, and, in that sense, it's the same as the conventional approach. However, with out-of-order execution, processing patterns change each time a command is executed. Yet, in the end, the results all line up."

This approach might not seem likely to prove successful. "This approach was first conceived around 2005, and, at the time, even we were not fully convinced of its feasibility. When we actually applied the approach to database engines, though, the results were very promising. We decided to work together with Hitachi, Ltd. to develop a new database engine based on this principle. We weren't going to just modify existing database engines. We were going to take on a much greater challenge, requiring hard work and tough decisions — rewriting the source code from the ground up," explains Professor Kitsuregawa.

Execution"
cessing Speeds 1000-Fold

Increasing the Speed by Three Orders of Magnitude Would Change the World

Their work is becoming visible. At the end of May, 2012, Hitachi commercially launched a new database engine that gave 100 times higher performance for a wide range of queries than their conventional database products. It used conventional hardware; the only difference was in the software.

"Our goal, though, was a 1000-fold speed increase. Increasing the speed by three orders of magnitude would change the world. When we try to lead the era of big data, it is essential to make it possible to handle data sets 1,000 times larger than current data sets, with the same level of convenience and speed. This requires improvement of hardware performance as well, so we are now using multi-core servers and entering the last phase of development as we work to achieve our goal."

The most common form of database is what is called a relational database. The principle that underpins relational databases has not changed in over 30 years. They have been considered almost "complete", but the out-of-order execution approach tries to change the core parts of the relational database engines. In the modern age of big data, the amount of data grows explosively, from giga to tera, peta, exa, and zetta. The ultrafast database engine could open a gate for this new era.

(Written by Interviewer Akiko Seki)

University Industry Collaboration

Implementing a new paradigm-changing principle, developed in Japan and promoting the use of big data, into high-speed-data-access infrastructures

Utilizing the outcome of joint research and development with the University of Tokyo in a commercial product

The high-speed data access infrastructure released by Hitachi, Ltd. in late May, 2012, generated a great deal of interest throughout the IT industry. Hitachi Advanced Data Binder Platform boasts high-speed search capabilities approximately 100 times faster(*1) than their conventional products. This commercial product used the results of joint research on ultrafast database engine(*2) between the University of Tokyo and Hitachi as part of the Funding Program for World-Leading Innovative R&D on Science and Technology.

Against the backdrop of growing interest in the use of big data in business and social activities, Hitachi Advanced Data Binder Platform has been launched as a powerful commercial tool for supporting data search and analytics. The key to an ultrafast database engine is, of course, the principle of out-of-order execution(*3) developed by Professor Kitsuregawa's team. Hitachi's development team has dedicated their efforts to contributing to the use of big data by producing and releasing products based on this revolutionary principle from Japan. This was not easy to accomplish.

The mission given to the team was to implement out-of-order execution in Hitachi's HiRDB database engine. This meant that the team had to throw out their conventional code body, and rewrite the code based on the new principle. For an IT vendor, a decision such as this requires a great deal of bravery. Nonetheless, the developers took on the challenge, believing that such a bold move was necessary to make something truly innovative. If successful, the project might even change the very paradigm of databases. It was certainly a worthwhile and rewarding challenge for the young researchers and engineers.

The developers have succeeded in using the results of their research and development to create and release a database engine that could search data approximately 100 times faster than their conventional products. As the final goal of their research project, the University of Tokyo and Hitachi now aim to develop a database engine offering search speeds 1,000 times faster than conventional database engines by the end of fiscal year 2013.

Only time can tell what business and social revolutions will emerge from this fertile new land cultivated by the ultrafast database engine.

***1:** Performance was measured for analytics queries that needed to access a specific amount of data that met specific conditions. Experiments were carried out based on a standard analytics database benchmark. Note that there was some variance in execution speeds, depending on the type of data analysis request.

***2:** Carried out as part of the Kitsuregawa FIRST Project.

***3:** A new principle invented by Professor Kitsuregawa and Project Associate Professor Goda (The University of Tokyo).

Cyber Physical Systems, Creating New

One of the sub-themes being advanced as part of the Kitsuregawa FIRST Project, in addition to database engine development, is Subtheme 2: "Large Scale Cyber Physical Application Verification Testing". We talked with Naonori Ueda, leader of Subtheme 2, about examples of social applications of database engines in fields such as medicine and agriculture, as well as machine learning, a technology vital to big data handling.

Striving to Create New "1+1=3" Value

"The goal of the large scale Cyber Physical application verification testing is to use the database engine developed by Professor Kitsuregawa in ways that can benefit society. Cyber Physical Systems (CPS) attempt to combine various real-world data obtained from sources such as sensor networks with the computing capabilities of the cyber world in order to build systems which contribute to society. The CPS concept arose out of our new ability to acquire and store massive amounts of data regarding people and objects thanks to advances in sensor network and cloud technologies. We want to deeply analyze this big data to create new, '1+1=3'-like value, which would in turn create new data, in a positive feedback cycle," explained Naonori Ueda, Director of Machine Learning • Data Science Center, NTT Communication Science Laboratories and Visiting professor, NII.

Naonori Ueda
Visiting Professor, National Institute of Informatics
Senior Distinguished Scientist,
Director of Machine Learning • Data Science Center
NTT Communication Science Laboratories

Data analysis technologies have existed before the age of big data, but until now they have consisted mainly of analysis technologies for hypothesis verification of single kinds of data. With big data analysis, which merges and analyzes a wide range of types of data, it is unusual to start with a hypothesis. Instead, the data is used to discover hypotheses.

This "1+1=3 value" refers to a butterfly effect, the creation of unexpected and novel value. Hopes are high that carrying out in-depth analysis of the massive amounts of diverse data people create will make it possible to identify unexpected value that conventional technology was unable to discover. Mr. Ueda points out that in-depth scenario creation is essential to achieving this.

"You can't strike gold just by randomly gathering and analyzing data. You need to carefully consider what value you want, how you will extract it, and create corresponding scenarios."

Effectiveness in Forecast Overall Patient Recovery

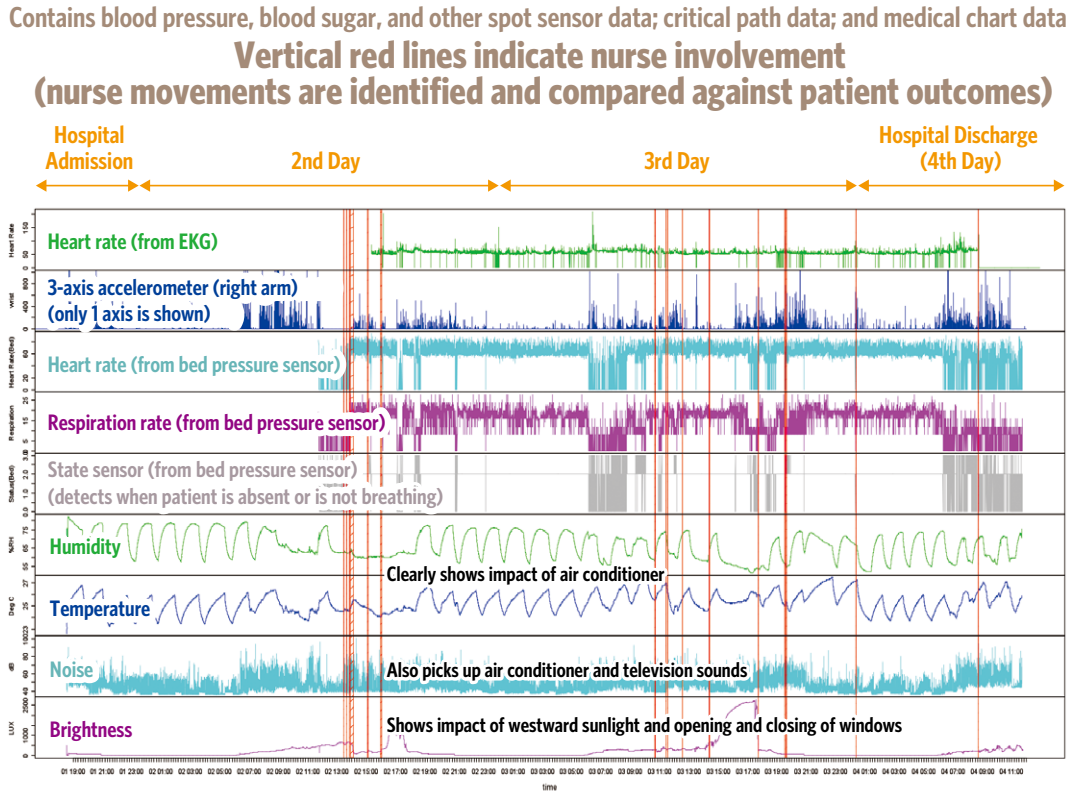
Mr. Ueda and his fellow project members are developing important fundamental technologies for handling big data, such as compression and anonymization. They are also designing and implementing CPS scenarios for the medical and agricultural fields. For example, in the medical field, at the Saiseikai Kumamoto Hospital, they are beginning to use data acquired from sensors attached by nurses and patients undergoing cardiac treatment.

"Normally, at medical sites, treatment is directed by a treatment schedule called the 'clinical path' (*1). However, sometimes patients' conditions take a sudden turn for the worse, and treatment becomes more long-term. Conversely, some patients recover and are discharged from the hospital earlier than expected. Until now, predicting these outcomes has come down to the gut feeling of those working in medical institutions, but we would like to use CPS to predict these outcomes."

There are various reasons for differences in patient recovery rates. Many times, they are the result of multiple, complex intertwined factors. In the past, doctors and nurses had no other choice than to make predictions based on their own experience when performing regular checkups and when offering nursing care. In this demonstration experiment, a variety of sensors, taking data on patient biometrics and hospital room environments, monitor patients 24 hours a day, predicting the course of their overall recovery. "The nurses are also outfitted with 3-axis accelerometers, making it possible to automatically identify nurses' activities, such as taking temperatures, giving medical interviews, taking blood, and the like. We are combining an enormous and diverse range of data, not only from patients but from nurses as well, and hope to use it to predict patient recoveries, feeding these results back into treatment and home care."

In the agricultural CPS scenario, the project is performing a demonstration experiment of "genomic selection technology" (*2), which makes predictions and selections based on crop DNA data to create new breeds.

Value from Sensor Network Information



(Ex.) Measurements for percutaneous coronary intervention (PCI) patient, from admission to discharge

"We are in the midst of a global food shortage, and it is urgent that we be able to cultivate high crop yields. However, conventional breed improvement, which takes time and money, will not produce results fast enough. We believe genomic selection technology can contribute to solve these problems. Over the remaining year of the project, we would like to carry out full-fledged analysis and produce a body of results."

Probabilistic Model-Based Machine Learning

These application verification tests are supported by database engines, and by findings from machine learning (*3), Mr. Ueda's speciality.

For example, consider a newspaper article saying, 'Baseball player Ichiro drank delicious wine at a restaurant.' This wouldn't be a baseball article, but a gourmet article. Most search engines, though, use keyword searches, making it difficult to find multifaceted articles like this. With machine learning, which uses a statistical approach called probabilistic modeling (*4), to more accurately deduce the contents of articles based on factors such as word frequencies and distributions."

This probabilistic model approach helps categorize the movements of nurses, described above. Mr. Ueda came up with the approach of using a probabilistic model to comprehensively evaluate the answers provided by multiple learning methods, and having a computer learn the results in advance.

"For example, we would teach it that if A says 1, B says 2, and C says 3, the answer is actually 5. This is called meta-learning (*5), or, in other words, learning how to learn. It holds that things are not decided by majority vote. One of the key aspects of meta-learning is that even if the accuracy of each individual learning method is low, the accuracy can be greatly improved by combining the results."

Meta-learning holds promise not only for human action classification, but for creating unexpected new value from big data. Mr. Ueda conveyed his dedication to further improving human action sensing technologies in order to contribute to solve a host of societal issues.

(Written by Interviewer Madoka Tainaka)

***1: Clinical pathway**
Tables containing standardized treatment and testing schedules. Treatment plans.

***2: Genomic selection technology**
Method of predicting and selecting individual strains and genetic capacities using the many DNA base sequences distributed throughout entire genomes.

***3: Machine learning**
One of the fields of artificial intelligence, machine learning is a technology which replicates the learning functions of people on computers.

***4: Probabilistic model**
Model for generating data based on probability distributions. This mathematical statistical model is effective for quantifying the effect of uncertain future events, such as in weather forecasting or stock predictions.

***5: Meta-learning**
Learning the learning process itself. The prefix "meta-" means "beyond" or "high order".

Large Scale Health Care Study in Bangladesh, and the Potential for Big Data Use

One of the sub-themes of the Kitsuregawa FIRST Project currently underway is the Portable Clinic Project, being carried out in Bangladesh by Kyushu University Hospital associate professors Naoki Nakashima and Ashir Ahmed. Let's look at the potential for BOP (Base of the Pyramid) business and social contribution system construction using big data by examining the example of this demonstration experiment.

Naoki Nakashima
Associate Professor,
Medical Information Center,
Kyushu University Hospital

Providing Medical Services to the Poor and Collecting Big Data

Sadakane: Please tell me the objectives and background of this large-scale health care study in Bangladesh.

Nakashima: Bangladesh lacks a robust medical infrastructure, so this project aims to use health checkups which utilizes telemedicine, implementing measures based on individual checkup results (risks) in order to improve the standard of health. It is a social business(*1) creation experiment, and at the same time has as its goal the collection of massive amounts of real-world data in order to create a cyber-physical system. Portable clinic studies have been carried out in Japan in the past, but due to the complexities of the legal system, as well as social customs, they have only been able to gather data for around 100 people each year. We needed to acquire a large amount of data, so we decided to work with the Grameen Group(*2) in Bangladesh, and perform health checkups for 10,000 patients. There are few doctors in Bangladesh, with most villages having none, but there are pharmacies across the nation. 98% of the country has cell phone coverage, so the necessary infrastructure elements were all in place.

A doctor performs remote diagnosis via Skype

Sadakane: What led to the collaboration with the Grameen Group?

Nakashima: Associate Professor Ashir Ahmed, at our university, is also a project director for the group, so we've had an exchange agreement between our university and Grameen Communications(*3) since 2007, and we have been engaged in joint research into social business, its creation, and its popularization. This project is one facet of that.

Sadakane: Please discuss specifically how you're carrying out the demonstration experiment.

Nakashima: First, we're having local people come in for health checkups, and local nurses, called "Grameen Medical Ladies"(*4) measure their heights, weights, etc., and, using a sensor set they carry in an attaché case, they also measure their temperatures, blood pressure, blood sugar levels, blood oxygen levels, and the like. Patients with moderate or severe health issues undergo remote diagnosis by a doctor, using Skype™. Between July and December of last year, we completed checkups for 5,500 people.

Producing Health Checkup Results and Looking at Potential for Business Roll-Out in Other Countries

Sadakane: What is novel about this project?

Nakashima: One thing that makes it novel is that it uses global standards. Other novel points of the project from a social system perspective are that it is being carried out under the various constraints placed by legislation, transmission bandwidth, and the like, and that medical advice is being given not only to those presenting symptoms, but those without presentation as well.

Sadakane: What results has the project produced?

Nakashima: We divided the patients into four risk groups, and found that 14% were healthy, 66% had mild health problems, 17% had moderate health problems, and that 3% had severe health problems. We performed remote diagnosis and prescribed medicine for those with moderate and severe health problems, as well as offering lifestyle advice, and urged them to receive a second checkup in two months. The second checkup was only performed for those who had been diagnosed with moderate or severe health problems during the first checkup. Approximately half of them actually underwent the checkup, and many of them showed an improvement at the end of those two months.

Sadakane: So it showed potential for a new medical service in developing countries. Last, could you please tell us about the issues the project is currently facing, and your hopes for its future?

Nakashima: We are behind schedule in reaching our goal of 10,000 checkups, and this is because the health checkups don't align well with peoples' needs. People who want diagnoses for headaches, injuries, and the like have interpreted the project as only identifying chronic conditions, and not offering diagnoses for headaches, injuries, etc., and therefore deciding not to have the health checkup. In order to resolve this mismatch problem, we plan to develop a CDSS(*5) system for nurses, which efficiently connects local nurses with remotely situated doctors, and to carry out a survey to lay the groundwork for meeting patients' needs. To do so, we need to collect even more data. In the future, we'd also like to establish some sort of basic clinics. There are a lot of people with obesity-induced high blood pressure in Bangladesh. This is a trend that spans developing countries. We think it would be eminently feasible to customize the system to the conditions in individual countries, and expand it into a business. After establishing a business model in developing countries, it would also be possible to implement the system in developed countries as well. In areas struck by the Great East Japan Earthquake there have been problems with nutrition and access to medical treatment. I think this system could be used to maintain peoples' health in emergency situations such as this as well.

(Written by Interviewer Yuko Sakurai)

Kunihiko Sadakane

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

***1: Social business**
Activities which apply business methods in order to resolve social problems. Social enterprises.

***2: Grameen Group**
Corporate group, including non-profit entities, created by Grameen Bank to provide financing to support the autonomy of those suffering from poverty. The Group is led by Nobel Peace Prize winner Dr. Muhammad Yunus.

***3: Grameen Communications**
Non-profit information technology company engaged in IT-related operations. The company works to improve information access and offer IT education in agricultural villages.

***4: Grameen Medical Ladies**
The women who receive self-reliance support from the Grameen Group, such as being employed by Grameen Group companies, are called "Grameen Ladies". In this demonstration experiment, women with nursing certifications perform health checkups as "Grameen Medical Ladies".

***5: CDSS (Critical Decision Support System)**
System used to collect various information and use it to supplement doctors' diagnoses.

NEXT PAGE

Shonan Meeting Topics

NII 湘南会議 SHONAN MEETING

List of Past Shonan Meetings

Feb. 18-20, 2013
Multi-activity in Interaction: A Multimodal Perspective on the Complexity of Human Action

Dec. 17-20, 2012
Social Issues in Computational Transportation Science

Nov. 12-15, 2012
The NII Shonan Configurable Computing Workshop

Nov. 9-12, 2012
Quantitative methods in security and Safety critical applications

Nov. 3-6, 2012
The Future of Multimedia Analysis and Mining

Oct. 15-18, 2012
Grid and Cloud Security: A Confluence

Oct. 9-12, 2012
Whole-Session Evaluation of Interactive Information Retrieval Systems

July 30-August 2, 2012
A Meeting of IT-enabled Services (ITeS)

May 28-31, 2012
Parallel Methods for Constraint Solving and Combinatorial Optimization

May 19-22, 2012
Bridging the theory of staged programming Languages and the practice of High-performance computing

May 15-17, 2012
Engineering Autonomic Systems (EASy)

Apr. 23-26, 2012
Hybrid Systems: Theory and Practice, Seriously

Jan. 11-15, 2012
Large-Scale Distributed Computation

Nov. 14-17, 2011
International Symposium on Symbolic Systems Biology (ISSSB'11)

Nov. 5-9, 2011
Hybrid Quantum Devices

Sept. 23-26, 2011
Automated Techniques for Higher-Order Program Verification

Sept. 18-21, 2011
Knowledge-leveraged Computational Thinking through Natural Language Processing and Statistical Logic

Sept. 14-17, 2011
Dependently Typed Programming

Sept. 7-14, 2011
The 14th Agda Implementors Meeting

Feb. 13-18, 2011
Graph Algorithm and Combinatorial Optimization

For details, please see the NII Shonan Meeting website.
http://www.nii.ac.jp/shonan/

Expectations Rise for NII Shonan Meeting

The NII Shonan Meeting, which began in February, 2011, was held for the 20th time in February of this year. Last November, the commemorative symposium "Creating Global Future Value in Asia" was also held. Professor Zhenjiang Hu, now the chair of the academic review committee, dedicated himself to creating the Shonan Meeting. We talked with him about its objectives, results, and his hopes for the future.

The 20th Seminar of Asia's First Dagstuhl-Style Workshop

The NII Shonan Meeting (hereinafter referred to as the "Shonan Meeting") is Asia's first residential international seminar. It was established to provide a place for top researchers in the field of informatics from various countries to come together and exchange ideas and discuss the issues they are currently working on. It was modeled on the Dagstuhl seminars held in southwest Germany. The Dagstuhl seminars are an international workshop, supported by the German government, which has promoted the field of informatics in Germany and around the world for over 20 years. NII's Professor Zhenjiang Hu has attended the seminar 6 times, and, keenly aware of its importance and significance, determined to create a similar center for informatics research in Japan.

"One of the things that sets Dagstuhl apart is that it is not a place for announcing research results, but for discussing promising future issues. Top level researchers gather in this pleasant small town, staying there and exchanging ideas with each other. It is extremely productive. However, there was no comparable top level international informatics workshop in Asia. NII believed that by creating a center for this kind of dynamic meeting of the best minds in informatics in Japan as well, it could contribute to the globalization of Asian informatics.

This vision came to fruition in February, 2011, with the first Shonan Meeting. In the two years since then (effectively one and a half years due to a six month suspension following the Great East Japan Earthquake), 20 seminars have been held. Guided by Special Appointment Professor Yoh'ichi Tohkura, the Shonan Meeting's steering committee chairman, members of the administrative committee and international members in charge of overall operations have succeeded in smoothly carrying out the seminar, led by the strong drive and dedication of all involved.

New topics emerge from the seminars, attended by over 400 people from over 30 countries

Each seminar of the Shonan Meeting is attended by 20 to 35 people, and over 400 people have attended in total. Over 60% are non-Japanese, representing 34 nationalities.

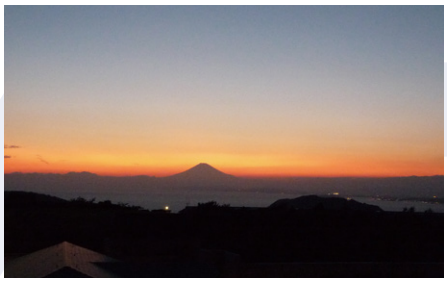
"Fixed themes aren't decided for the seminar in advance. Instead, they are set flexibly at the seminar itself, based on the interests of the participants. The themes themselves are decided upon discussion by all the seminar's participants. The Shonan Meeting is certain to have a strong influence on informatics research in Japan and throughout Asia."

As a general rule, the residential periods are from Monday to Thursday. This is to enable participants to meet with researchers who are not participating in the seminars. Past participants have been extremely satisfied. The main reason for their satisfaction is the community forming between researchers that the seminar makes possible, but that is not the only attraction of the Shonan Meeting.

"First, the fact that we have scrupulously maintained the Dagstuhl approach has been very well received. Another appealing thing about the Shonan Meeting is the allure of Japanese culture, such as the attention to detail in the

Zhenjiang Hu
Professor, Information Systems Architecture Research Division, National Institute of Informatics

as a Center of Asian Informatics Research



Left: View of Mt. Fuji from the Shonan Village Center, the site of the Shonan Meeting.
Right: A meeting in progress. During the Shonan Meeting, the meeting room is open 24 hours a day.

services offered, the delicious food, the view of Mt. Fuji from the conference site, and the excursion to Kamakura."

The fact that some people have already attended the seminar multiple times is a testament to the satisfaction of its participants. Each seminar creates new projects, and there have been moves to establish another organizer? to lead further, more focused discussions. This shows that the Shonan Meeting has succeeded in what it set out to do.

In addition to its main objective of creating a new center for global research, there is another major goal of the Shonan Meeting.

"Compared to westerners, Asians don't tend to be as good at taking risks and exhibiting leadership. Holding the seminar in Japan by its very nature creates more opportunities for Japanese and other Asian researchers to act as organizers. The Shonan Meeting provides chances for researchers to be leaders at conferences in which cutting-edge issues are discussed, creating a stronger presence among the global research community."

The Shonan Meeting is also considering new, uniquely Japanese approaches, such as coordination between industry and academia, and accepting open applications

Last November, the highly successful commemorative symposium "For Making Future Value From Asia" was held at the Gakushi Kaikan in Chiyoda-ku, Tokyo, attended by almost 100 people. The keynote lectures were given by an impressive lineup of speakers: world-famous database theory researcher J.D. Ullman (professor at Stanford University), top virtual logic programming researcher R.A. Kowalski (professor at Imperial College London), and big data research pioneer M. Kitsuregawa (Professor, Institute of Industrial Science, The University of Tokyo). Special sessions in the latter half of the symposium included lectures by R. Wilhelm (scientific director of Schloss Dagstuhl and professor at Saarland University) and professor Hu, and a session in which professors Wilhelm, Ullman, and Kowalski, as well as S. J. J when (the chairman of the supervisory board of the Dagstuhl Seminar and professor of Technische Universitt Berlin), discussed their hopes for the future of the Shonan Meeting.

"I was the emcee, but the discussion was dynamically led by the interests of the participants. That's the Shonan style."

During the session, in addition to praise for the way the Shonan Meeting was carried out, some suggested not merely following in the footsteps of Dagstuhl, but adding new and unique elements.

"We would like to add new, uniquely Japanese and Asian elements to the Shonan Meeting. For example, there is little corporate involvement at Dagstuhl, but Japan has many leading information companies, so we would like to take advantage of this aspect of our region through collaboration between industry and academia. Also, right now the event is by invitation only, but we would like to open it partially to open application. This will increase the amount of thematic variety, as well as further improving the quality of the conference."

The Shonan Meeting is becoming more and more well-known, and the range of corporate applicants and themes has expanded. However, Professor Hu states, "it will take a bit longer before we get everything completely on track, including funding and personnel." The societal need for this kind of international conference is enormous, and it serves a very important role, both for Japan and the world. The new challenges being taken on by the Shonan Meeting, such as its leveraging of the regional characteristics of Asia, hold tremendous potential.

(Written by Interviewer Yuko Sakurai)

The successful commemorative symposium held in November, 2012, was attended by almost 100 people.

"The Ordinary and Extraordinary World of Powers"

Akiko Aizawa

Professor,
Digital Content and Media Sciences Research Division,
National Institute of Informatics

There is a famous film called "Powers of Ten". It's a short film—less than 10 minutes long—that was made in 1968, but is a classic science film that remains fresh no matter how often one sees it. It begins by showing a single square meter of park, where a person is enjoying a picnic. The camera begins zooming out, faster and faster, showing 10 times the area, then 100 times. Once the camera has pulled out so far that it encompasses the whole of space, it begins zooming back in, showing 10x the magnification, then 100x. The camera zooms in to the skin of the napping picnicker, then into the atoms of which the skin is composed, before finally arriving at the quark level. It is a journey from the world of 10^{26} to 10^{-18}.

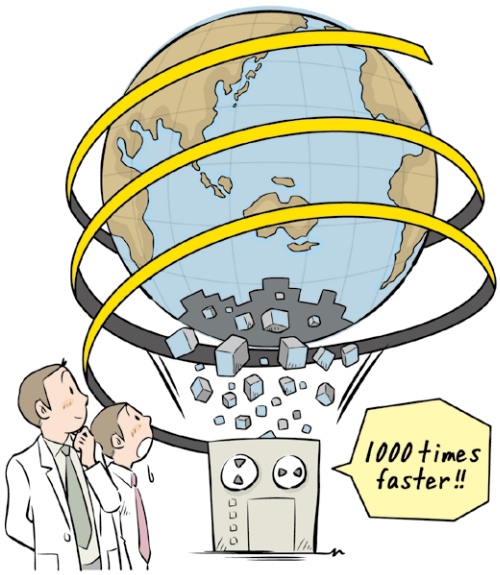
Logarithms were discovered almost four centuries ago, and expressions of "10^n" have become common in daily life.There are central pressures of 950 hecto-(10^2)pascals, disks with 1 tera-(10^{12})byte of capacity, CO_2 concentrations of 390ppm (10^{-6}), and nano-(10^{-9}) technology. These would not exist without the scientific developments that have made it possible to measure with these degrees of accuracy, and they are deeply tied to the natural and social recursive structures called powers.

However, reflecting on these powers, the sense of scale expressed by 10^n is one that exceeds our imaginations. If the Earth were a ball right in front of you, 1 centimeter in diameter, the sun would be a 109 centimeter ball, 115 meters from you. It is difficult for us to visually take in these differences. Logarithms can be considered convenient tools for enabling us to conceptualize the otherwise hard to perceive world of 10^n.

Using logarithms, multiplying 10^1 by 10^2 consists of adding 1+2, tripling the number of zeroes. However, an actual 1,000-fold increase would be the difference between making 1,000 yen an hour and 1,000,000 yen an hour! How could you spend your money if you made that much? There are few people who could really conceive of how a change in income like that would change their lives.

What about if computers ran 1,000 times faster? Computers running 1,000 times as fast as modern computers would result in a far greater degree of precision in our ability to understand our world and control our environment. From the distance between stations to a single block, from a day to a minute, from every customer of a store to me, personally. If flexible responses were provided in everyday searches, based on the exact conditions at the time, it would be fair to call it computer intelligence resulting from computing power.

The question, then, is: how? The 1,000,000 yen hourly wage, and the 1,000x speed computer, are both challenges to our powers of imagination.



Weaving Information into Knowledge

NII

NII Today No.45, Apr. 2013

This English language edition NII Today corresponds to No. 59 of the Japanese edition

Published by National Institute of Informatics, Research Organization of Information and Systems
http://www.nii.ac.jp/
Address: National Center of Sciences
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430
Chief editor: Yoh'ichi Tohkura Cover illustration: Makoto Komori
Photography: Yuki Akiyama and Akira Taniguchi Copy Editor: Madoka Tainaka Production: inscape 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

Cover Illustration

The cover shows the cycle of sensors (depicted as countless people) gathering various data from actual society, and a database engine (depicted as a massive brain) performing high speed processing, creating new value and feeding back into actual society.