Cognitive Robotics to Understand Human Beings

KAYOKO ISHII
Life Science Unit

1 Introduction

The question of what human beings (self and others), the mind, and the world are has always been of great interest to humankind. Most academic disciplines originated to answer these questions. As neuroscience has advanced, the notion that brain function is closely related to the mind has become more widely accepted, increasing the expectation that unknown aspects of the mind could be explored by neuroscience. However, a long-standing question regarding the mind is that one's mind seems to be associated with oneself as a physical existence, yet its content seems not to be expressed physically. Simply accumulating knowledge acquired through external analysis and observation of the brain as a physical entity is not sufficient to elucidate the essential functions of the brain and the nature of the mind.

The brain has many parallel units (modules) that represent different parts of the body or participate in different functions. When neuroscientists study the properties of a module, they apply a controlled stimulus to the subject, so that it perturbs only the targeted module (or limited numbers of modules, including the target). When human beings engage in usual activities, however, many different modules work in an autonomous and distributed manner. Particular ideas or actions are generated either by the exchange of information between specific modules or the selective involvements of certain modules. Unless the algorithms for these information exchanges and selections can be elucidated, observation of the physical state of the brain at a given time cannot lead to an understanding of the information processing taking place at that time.

In Japan, research on the computational theory of the brain[1] and research combining theory and physiological experiments[2] has been carried out. One of the major themes in Japanese brain science in the 1990s was “creating the brain” beside analytical experimental sciences (“understanding the brain”) and research oriented to medical applications (“protecting the brain”). This theme was significant in that it not only expressed the concept of understanding brain functions through “cycles of creation of models of brain, computational theory and neural networks, their verification through experimental science, and improvement of theories and models,” but also expressed the unconventional orientation of creating new systems inspired by the brain. Furthermore, computational neuroscience was defined as “to investigate information processing of the brain to the extent that artificial machines, either computer programs or robots, can solve the same computational problems as solved by the brain, essentially in the same principle” [3]. Based on this conceptual framework, innovative researchers, although still few in number, are engaging in studies to elucidate human brain functions through “cycles of creation of brain algorithms, their verification through robots, noninvasive measurements of brain activities, psychology, and experimental sciences, and further improvement of the algorithms.”

From the perspective of ordinary Japanese sensibilities as well, the mind cannot be considered in isolation from the body, the environment, and the existence of other people. In other words, attempting to create the brain alone will not elucidate the essential functions of the brain itself or the mechanisms of the mind. Embodiment and context dependence are key concepts in cognitive science and neuroscience, and robots given bodies to interact with the environment are serving as effective simulation tools [4].

2 The Field of Cognitive Robotics

Since their beginning, robots have been constructed to imitate, replace, and supplement human beings or a part of human functions. Since 1960, the mainstream of robot development has been oriented to industrial applications - manufacturing robots. In recent years however, we have witnessed a rapid increase in the development of robots designed to serve ordinary people rather than experts [5-7]. Traditionally, robotics referred to a combination of science, engineering, psychology, sociology, and other disciplines necessary "for the development, construction, and dissemination of practical robots," with particular emphasis on the engineering aspects.

During the process of seeking the necessary conditions for robots to act as flexibly, smoothly, and autonomously as human beings in the real world, robotics researchers began to turn their attention to human cognitive mechanisms, learning, recognition of others, and social behaviors. In Japan, since around 1994, robotics researchers have organized research groups, such as the Keihanna Research Group for Sociointelligenesis, with the primary aim of elucidating human cognition, development, and behaviors by using robots. These researchers have adopted a "constructivist" approach, which aims to explain human cognitive mechanisms by creating and testing robots that can develop humanlike cognitive abilities to cope with the real world ("cognitive developmental robotics")[8]. These researchers have the advantage of having not only advanced knowledge and experiences in physics and engineering, which share a common basis in mathematics, but also a broad knowledge and understanding of biology, the humanities, and the social sciences, and of having a solid verification platform, such as robots.

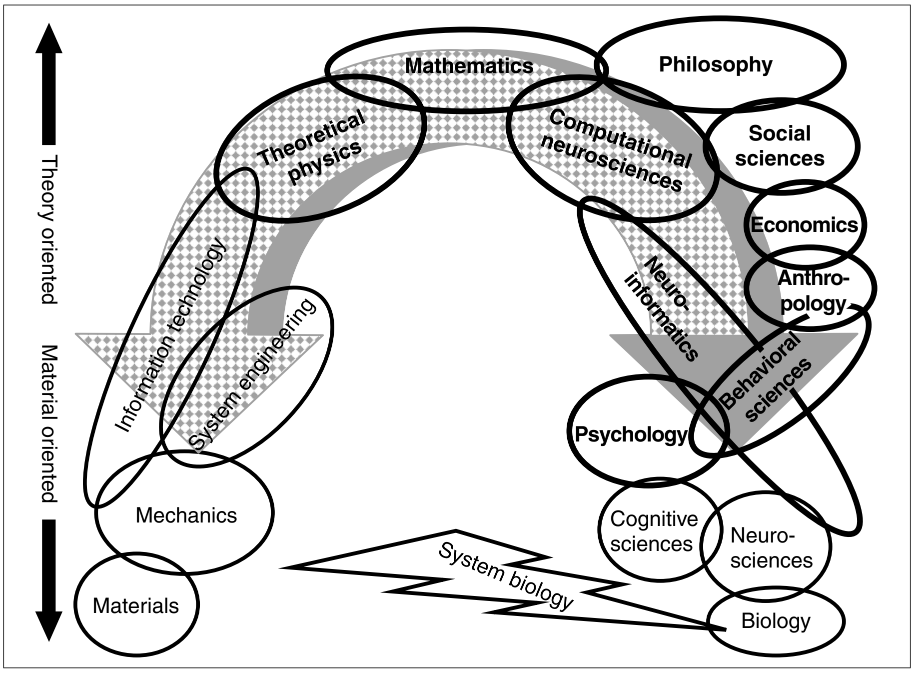
2-1 Cognitive robotics

"Cognitive robotics" in this report refers to an comprehensive science in which robotics, as described above, neurosciences (ranging from the experimental to the theoretical or mathematical variety and neuroinformatics), cognitive science, psychology (psychophysics and behavioral measurement) and behavioral sciences seamlessly collaborate in unity while keeping variations in perspective, closely connecting with fields such as philosophy, social sciences, anthropology, and economics; exchanging their knowledge and methodologies, and executing mutual verification. Robotics herein represents an expectation of interdisciplinary integration, rather than the simple collection of independent research fields, and will be realised by using robots as a common verification platform to highlight weaknesses and errors in research processes in individual disciplines and contradictions among different disciplines,

For the development of commercial robots, a demand-oriented perspective, based on future prospects and on a broad understanding of humans and society, is required. For humanoid robots (humanoids hereafter), the initial phase of development has been completed for structural modules and actuators, which correspond to their bodies, and the computers to control them. Currently, humanoid hardware developed in Japan is widely used, both in Japan and abroad as platforms for the development of software to serve as the cognitive mechanisms. Certainly, future success in the development of practical humanoids will depend on the improvement of their "cognitive functions."

Some research laboratories dedicated to the development of practical robots have taken up parallel research in cognitive robotics, which offers this fundamental knowledge. Even researchers specializing in the development of structural modules and actuators must take compatibility with next-generation cognitive functions into consideration. Energy consumption is another important future issue expected to accompany advances in robotic cognitive function. Research to solve this issue will become necessary.

Figure 1 : Cognitive robotics



Prepared by the STFC

Energy consumption problems associated with improvement in cognitive functions

Neural systems are enormous energy consumers in animals. Humans have remarkably large brains for their bodies (approximately 2.5% of body weight), and the brain consumes 20% of the body's total energy. The cerebral neocortex of primates increased exponentially in volume as their social behaviour became more complex (the social brain hypothesis, Reference 14). In order for robots to work in complex human society, improvement of their cognitive functions is indispensable. This raises the problem of energy consumption for information processing. Furthermore, because of size restrictions and mobility requirements for humanoid robots, it is critical to invent new materials and structures enabling flexible and efficient information processing within limited spaces. This means that humanoid robots can be most desirable test beds for the creation of new paradigms for computational theories and materials/structures.

Although noninvasive methods to measure brain activity have progressed, certain constraints remain. Currently, interpretation of results requires either (1) statistical analysis of data from multiple measurements or (2) training subjects for considerable hours to ensure reproducible responses prior to a single measurement and interpreting the result. In order to decipher spontaneous information processing in a subject's brain by each single measurement, algorithms of the information processing are a prerequisite. Currently, algorithms are being proposed by mathematical studies. Before they can become worthwhile for practical use however, the algorithms must be repeatedly verified and improved through simulations using robots. Such algorithms, if available, would serve as the foundation for the development of new computers and human-machine interfaces.

2-2 Global progress of robotics

In October 2005, the European Commission published a report on the growth of global robot markets[15].

A) Since the European Commission's report focuses on robot markets, it deals with major areas of robotics without a category for cognitive robotics. In Japan, a market for humanoid robots, such as personal and home robots, and service robots has already started. Although cognitive robotics can be regarded as a research area in basic science, not necessarily oriented to applications, it is also recognized as a fundamental necessity for development targeted toward commercial robots.

B) In Europe, because of the cultural backgrounds of monotheism and the belief that "God created human beings in his own image," researchers feel guilty and reluctant to create humanoid robots (artificial humans)[16]. Therefore, there is little expectation for the commercialization of humanoid robots. However, for the purpose of basic research in cognitive science, neuroscience, and certain medical fields, since 2004, the European Commission has promoted projects similar to the cognitive developmental robotics projects of the Keihanna Research Group for Sociointelligenesis. European researchers also conduct questionnaires and other surveys, before and after robot demonstrations, to evaluate how contact with actual robots can decrease public aversion to

Table 1 : Examples of research areas in cognitive robotics

Knowledge to be obtained	Research area	Subject of analysis
Properties of elementary functions of unconscious and autonomous congnition / behaviours	Experiment combining experimental animals and robots	Based on the neural activities in an experimental animal's brain, a behaviour predicted from the neural activities is reproduced simultaneously in a robot.
	Physiological and psychological experiments on conscious experimental animals	Relationships between temporal profile, properties, intensity of neural activities and the manifestation of cognition and behaviour (cause-effect relationship).
	Primatology, anthropology	Development, learning and social behavior of monkeys and humans.
	Cognitive archaeology, anthropology, history	Changes in human cognitions associated with evolution and environmental and social diversification.
	Cognitive developmental robotics	A robot that has a cognitive framework and can achieve route-finding through physical interaction with the environment.
	Psychophysics, behavioral measurement	Time sequence, correlation and regularity of unconscious perception/behaviour. Control of behavior and cognition as a result of perturbation of perceptions and behavioral patterns.
Mechanisms of development in intelligence for interaction with others and society; Mechanism of developmental disorders of social abilities, e.g. autism	Genetics, evolution, anatomy, physiology	Motivation; selective attention; recognition of the novelty and regularity of stimuli; imitation.
	Fetology, baby science	Spontaneous motion; response to a caregiver's cyclic repeated actions.
	Theory-of-mind	Pointing, joint attention, false belief task (estimation of others' expectations and predictions).
	Mirror neurons-, analysis of perception-behavior relationship	Common neural information processing in perceiving others' actions and expression of emotions and in evoking/performing/expressing the same actions and emotions in the self.
	Noninvasive brain activity measurement	Location, strength and temporal changes of brain activities during cognition and behaviour associated with others or self.
	Computational neuroscience	Close forward/reverse relationship between brain algorithms for perception and behavior and perceptive and behavioral models.
	Cognitive developmental robotics	Development of cognitive patterns through physical interactions with others and the environment.
	Philosophy	Relationship of neuroscientific functions and physiological meanings of emotion and sensation with actual feelings and senses or "the experience of reality."
Mechanism of the formation of norms of social conduct; Mechanism of the expression of impulsive acts and depression	Neuropharmacology, psychoneurology	Perturbation of parameter molecules that control brain activities and the mechanism of deficits in social behaviour.
	Cognitive developmental robotics	Simulation of perturbation of parameter molecule and changes in individual/cooperative behaviour.
	Economics	Effect of prediction/evaluation of advantages and disadvantages on human behaviour; the role of values, motivation and emotion in decision-making.
	Social sciences, social psychology	Tools for human interaction: objects, gestures, languages, technologies and regimes that have accumulated through history and are shared in society; Caregivers treat their children as more mature and older than their actual states in order to involve them in a communication game.

Prepared by the STFC based on References [9-13] and other material them.

C) European industries are currently developing home robots but carefully refrain from giving them a hint of resemblance to human beings[15].

D) In the U.S, NASA announced in December 2005 that it would promote the development of humanoid robots that could aid construction of a lunar base as a step toward manned Mars exploration. NASA explains why construction robots need to be humanoid as follows. Humanoids can use the same tools and equipments as human crews. Furthermore, since programming all tasks is impossible, robots must learn their work, and humanlike shape will make it easier for human crews to teach and remotely control them. Based on the concept that “it may not be the human capability to learn, but to teach, that has contributed most to our progress,” NASA cites the following as the basic concepts of its robotics research[17]. (1) Robots need to be fostered/taught rather than to learn; (2) robots need to be able to “teach” other robots rather than to simply transfer data to them; (3) robots’ ability to “teach” is proof that “learning took place”; and (4) research must be practical.

The U.S. has been systematically applying cognitive science to human education for many years. It is difficult to objectively evaluate how much and when learning of a child has been completed. In the U.S, sound accumulation of experiments and observation in human education enable criteria to be set as (1) - (3) above. In addition, the ability to “teach” is a concrete indication of the ability to recognize self action. To achieve practical purposes as mentioned in (4), the contribution of broad basic research is, in fact, indispensable. The U.S. can establish such a policy because it has such a large pool of researchers in social science, anthropology, psychology, and philosophy, who can contribute for applied research while carrying out basic research.

Table 2 : Comparison of developing robot markets

Area of activity	Degree or level of activity			
	Japan	Korea	Europe	U.S.
Manufacturing robotics	++++	±	+++	±
Humanoids A)	++++	+++	± B)	±
Personal/home robotics A)	++++	+++	± C)	±
Service robotics A)	++	+++	++	++
Biological & medical applications	±	±	+++	+++
Security and space robotics	±	–	++	++++ D)

++++: Excellent; +++: Very Good; ++: Good;
± : Fairly Good; -: No Remarks
A~D : see text.

Prepared by the STFC from Reference[15]

3 Comparison between Humans and Robots as Systems

3-1 Beyond differences in materials and structure

Many people believe that robots can never have the “same kind” of mind as human beings, because although the mind is not physical in principle, it is an attribute of human beings that consists of biological materials and structures. It is not known, however, how and to what extent mental functions are dependent on biological material and structure, and in what way. On the other hand, those who try to develop anthropomorphic robots are aware of the limits of existing mathematical computations and materials/structures and are exploring new materials and structures inspired by biological systems. The basic principle in robotics as a science for understanding human beings is, while understanding fundamental differences, to study human information processing and behavior with simulation on robots and to seek better conditions and principles of simulation.

3-2 Changes in human-robot comparison

Robots are basically described as machines with computers for information processing and with input and output devices, which are auto-regulated by the computer. Human beings can also be considered as systems, with the brain that processes information and auto-regulates sensory inputs and motor outputs. Comparisons between humans and robots as systems have been changing, as described below.

(1) The age of artificial intelligence

In the early age of robot studies, intelligence alone was emphasized among human characteristics and it was compared with symbol processing by artificial intelligence, which corresponds to the brain in robots. The brain was regarded as responsible for the entire process of perception, recognition, planning, and decision-making (top-down approach), while the body was merely a device for inputs and outputs. Robots based on this concept were incapable of adapting to unpredictable changes in their environments.

As tasks became more related to the real world, weaknesses of computers and robots emerged, one after another[18]. The fact that what human beings do naturally is in fact very intricate functions thus came to be recognized for the first time as a topic of scientific research. For researchers in modern philosophy and cognitive science, a new set of subjects were presented by robotics research [19].

(2) The age of neural network

Robots were designed according to the concept of neural networks. Circuits of information processing were formed and reinforced according to experiences and their frequency. These robots were no longer dependent on symbol processing that assumed mental representations. In contrast to (1), these systems were formed in a bottom-up manner, triggered by inputs of stimuli. Although these systems were appropriate for modeling insects and other creatures without central nervous systems and were robust against environmental changes, they were unable to elaborate higher-order functions, such as those seen for vertebrates.

(3) The age of combining top-down and bottom-up approaches

In recent years, embodiment, interaction with the environment, and development have become key concepts in cognitive science and philosophy. From their viewpoints, both human beings and robots have bodies that move and have diverse interactions with environments. Unlike computers, they must be able to find solutions within a limited time to complex problems occurring incidentally. The solutions must be valid in the real, physical world feasible under constraints imposed by the physical and functional properties of their own bodies.

As human beings mature over a period of years, they formulate “self” algorithms to integrate the outside world, input processing, and output production by repeated information processing in the neural system and physical interaction with the environment. Likewise, robots for understanding human beings must have the ability to autonomously change their own

Examples of human functions that seem ordinary but are actually remarkable

[Problem setting ability] Machines can process symbols quickly but cannot set problems by themselves.

[Domain-specific knowledge] The more closely related a problem is to the real world, the more human beings utilize their wide repertoires of domain-specific knowledge to solve them. Most of this is implicit knowledge that is held unconsciously or recruited according to physical or environmental cues.

[Heuristic knowledge] To solve problems in the actual world, human beings quickly select a finite number of information items required at a given moment. Machines cannot do this (the frame problem). Although heuristic solutions may be difficult, even for humans in novel complex situations, humans can avoid being brought to a standstill by acting as if the frame problem did not exist.

[Symbol grounding problem] Machines cannot associate symbols used for language processing or computation with actual objects and phenomena in the world.

[Binding problem] Humans can process multiple characteristics of an object in a parallel and distributed manner and finally bind them all together as the characteristics of the object (e.g. processing elementary information of an apple: “redness, brightness, size, roundness, hardness, smell, taste, etc.,” and rebinding them as “an apple”).

information processing methods (algorithms) and to develop intelligence through their physical interaction with the environment. In order to behave adaptively, they must also selectively perceive the world according to their genetic traits (initial conditions), experience, and memory, and according to predictions, motivation, and purpose. In other words, both top-down and bottom-up approaches of research are indispensable.

4 Ways of Understanding the Mind

Many people vaguely hold the agnostic view that the nature of the mind is hard to understand and will therefore never be elucidated scientifically. One opinion is that "The very moment a mental function is programmed, people stop considering it as an essential element of 'true thinking.' An indispensable core of knowledge always resides in the next thing to be programmed"[20].

Even natural scientists sometimes hold the implicit bias that "the mind, biological systems, and humanity are something special." This may hinder the elucidation of the mind. When analyzing the brain, neuroscientists with little knowledge of psychology, behavioral science, and philosophy may derive mental processes from "naive psychology," a set of commonplace theories without scientific bases. Efforts to recreate human cognitive functions in non-biological robots can be a means of escaping such biases.

Questions such as "in the end, can robots have minds?" or "do we want robots to have minds?" are not necessarily common interests of robotics researchers. The basic principle of cognitive robotics for understanding humanity is that even if researchers personally predict that the human mind cannot be completely recreated in robots, they should attempt to understand humans through the process of creating homologues of the human mind.

4-1 Substantialistic attempts to create a mind

From substantialistic viewpoints, researchers "postulate" that components of the mind are basically intrinsic in individual humans, and "assume" that robots can (be created to) have similar intrinsic components. Inquiry into what kind of principles should be used to create close homologues of the human mind's components and attempts to configure robots based on such principles will further improve understanding of the nature of the human mind. In psychology and neuroscience, where it is said that consciousness is only the tip of the mind's iceberg, verification that most mental processes occur unconsciously has begun. Since unconscious cognitive and behavioral processes are known to be relatively "mechanical" and closely related to physical states and the environment[11], they are appropriate to be built into robots.

As neuroscience revealed most macro structures and functional localization in the brain, it was disclosed that each function unit is working in an autonomous, distributed, and recurrent manner. Engineering attempts to reproduce conscious/unconscious systems to integrate autonomous, distributed, and recurrent processing are now carried out[12].

4-2 Relation Theory to validate the "substantial" mind

Now that the human mind is far from being elucidated completely and the robot "mind"

Turing Test

The Turing test was proposed in 1950 by mathematician Alan Turing as an answer to the question "Can machines think?" Instead of directly answering the question, he invented a verification method that uses an imitation game to distinguish humans from computers. A human inspector, a test subject (a machine), and a control subject (a human) go into separate rooms and communicate with each other through teletyped text. The inspector asks various questions to determine which subject is the human. If the machine can make the inspector judge it as a human, it is acknowledged that the machine has demonstrated the ability to think.

consists of materials and principles differing to those of the human mind, methods based on the Relation Theory are used to evaluate the "mind" of robots constructed in the manner described in 4-1. The mind of robots will be ameliorated by repeated cycles of construction from a substantialist perspective, evaluation with Relational methods and modification with the substantialist approaches.

5 Robots as Social Members

5-1 Humans interpret the world intentionally

When a computer creates a combination of characters and words that meets the requirements of a haiku (i.e., meets its substantial requirements) by chance during random symbol manipulation, one who reads it without knowledge of the process might assume the presence of an author and recognize the author's intentions, implications, and metaphors in the "haiku" [20]. This may occur because humans have a propensity to try to find meanings of any subject encountered [21]. For the time being, when human beings find intentions, feelings, and other mental properties in robots, this can be attributed to human empathy and projection of emotions.

(1) Projection of emotions

Human beings can empathize with or project emotions, even onto non-human creatures, natural structures, artificial tools, and vehicles. Young children often see faces, expressions, and emotions on objects. This is a normal phenomenon for infants, who are still developing the ability to promptly identify the faces and voices of fellow humans from among diverse stimuli from the outside world and to infer other people's intentions ("theory of mind"). This tendency disappears as children grow. In certain societies, where empathy with dolls and toys is an implicit taboo for adults, resistance or rejection may be encountered when these objects are used for psychotherapy [22].

(2) Animism

In anthropology and archaeology, it is known that in hunter-gatherer societies and traditional societies preserving close relationships with nature, people, including adults, tend to recognize spirits in many subjects (animism). Traces of animism can be seen even in some modernized countries, such as Japan, where indigenous beliefs have survived or have not been suppressed by ideas imported from abroad (i.e., Buddhism, Confucianism, and Taoism). For example, some Japanese perform rituals for worn-out tools and captured animals and fish (e.g. bonito burial mounds ["Katsuozuka"] and the Ainu bear ritual ["Kumaokuri"]) to cordially send their spirits into another world. Such traditions may be the basis for the Japanese tendency to not resist finding emotion in robots, which are mere machines [23, 24].

On the other hand, there are areas where new or imported religions repressed indigenous animism as taboo or heresy. In such areas, especially those where a monotheistic religion believing that humans were created in the image of an omnipotent Creator and all other creatures were created for the use of humans prevailed, people tend to think that "creation of humanoid machines is disobedience to God," "humanoid machines are harmful and dangerous to human beings," and "robots can never have a soul because they are not created by God" [16].

5-2 Appropriate distance between humans and robots

(1) The Uncanny Valley

According to psychological studies, people in general unconsciously feel affection towards an artificial object if it reminds them of a human or living creature. Robots with nuts-and-bolts appearances are treated relatively roughly, while those with relatively humanlike appearances evoke the kind of attitudes and responses akin to those seen among humans [25].

However, as early as 1970, a Japanese robotics researcher suggested the possibility that excessive similarity to a human might elicit repulsion [26]. His hypothesis was as follows: (1) As robots become more humanlike in appearance, human beings feel friendlier towards them; (2) however, when the resemblance to human beings exceeds a certain level, people become uncomfortable, falling into the so-called Uncanny Valley; and (3) as the resemblance further increases, friendly feelings increase again.

It was suggested that every human has his or her own comfortable physical distance from others (any other human's conspecifics) that varies according to social relationships and degree of intimacy. People feel anxiety or repulsion when someone approaches too closely[27]. This suggests that there may be critical distances or conditions of cohabitation between robots and humans that divide the reactions of the latter, as to whether they feel comfortable or not.

The development of robots is advanced in Japan, Europe, and the U.S. for the moment People's attitudes toward humanoids vary among these areas. The Europeans and Americans are skeptical or passive towards the development itself or the releasing of robots into the public. In Japan, on the other hand, so-called pet robots and communicating robots are already commercially available and have been accepted favourably to date. The Japanese in general tend to avoid precautions against the possible risks of humanoids and discussing countermeasures against them. However, as humanoids become more widespread and more humanlike, even the Japanese public's favourable attitudes toward robots could sour. Another possibility is that Japan may find an original way to develop robots, based on the traditional Japanese emphasis on cooperation and refrain from encountering problems of their uncontrollability, even in a complex real environment, and that such robots could be applied in usual social lives. If this could be realised, it would develop as a unique area of research.

(2) Humans can adapt even to inappropriate science and technology

When emerging sciences or technologies are discussed in terms of human adaptability, invasiveness, and usefulness, their promoters often claim that children can adapt themselves to them easily, although adults may have troubles, and that problems will disappear once those who have adapted themselves in childhood become the majority of the population.

In principle, the brain develops in accordance with its genetic frameworks. However, if a newborn organism is exposed to certain artificial stimuli from birth and before a certain developmental stage (critical period), it comes to recognize the given condition as natural. If the stimulus is removed early enough before the critical period, the organism may recognize the world as intact siblings do. The phenomenon is known as plasticity. If a stimulus persists until after the critical period, its influences are fixed for the rest of life. Because of plasticity, organisms may adapt, even to an entity that is meaningless or harmful to their survival. The fact that an immature organism can adapt to a given stimulus or environment does not necessarily warrant the harmlessness or usefulness of that stimulus or environment.

A classic experiment well known in neuroscience

When a kitten is raised from birth through a critical period in a visual environment in which it is exposed to unique visual stimuli, vertical or horizontal stripes for example, it becomes unable to recognize any visual stimuli other than those presented in the experimental environments, vertical or horizontal patterns, respectively.

Long-term prediction and careful analysis are essential for resolving the difficult questions of whether people adapting to the new stimulus of robots from infancy would benefit from them throughout their lives, and whether allowing many people over several generations to grow up with such a stimulus as a present environmental factor would work to the benefit of human society and humanity as a living species. There is a need to embark on broad follow-up surveys (cohort studies) while many members of society are still from generations that do not accept robots as a pre-existing environmental factor.

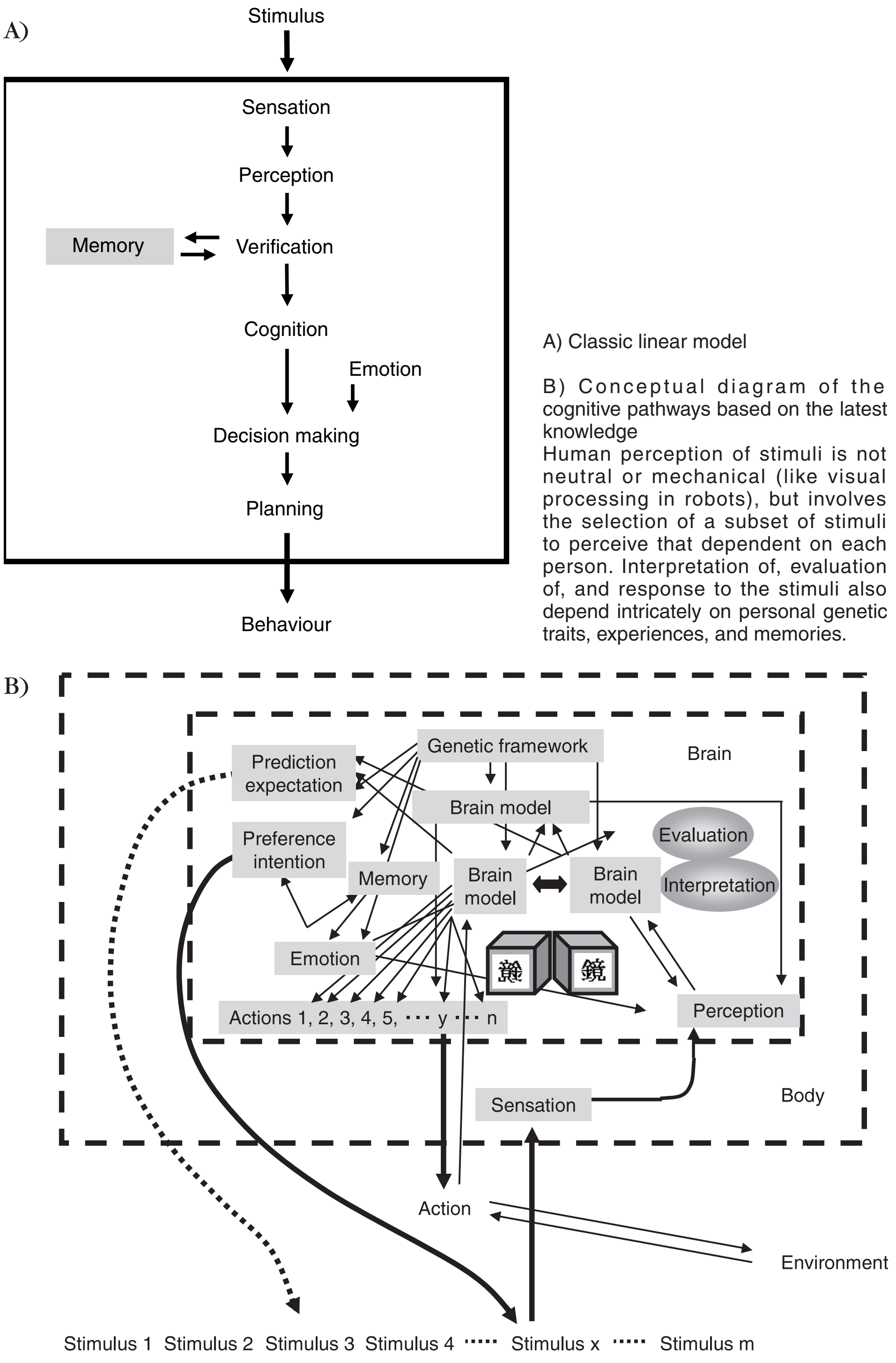
6 Examples of Projected Future Cognitive Robotics Research

6-1 *Mechanisms of unconscious/conscious autonomous behavior*

Most information processing in the brain and behaviour takes place in an unconscious manner. When people walk along a familiar path or engage in a skilled task, or when sleepwalkers return to their beds after wandering around, they are behaving and perceiving unconsciously and autonomously. If robots that can act autonomously are to be manufactured, those in the first stage will perceive and behave unconsciously.

Recent improvements in neuropsychological and psychophysical methodology and behavioral measurement methods, combined with the development of noninvasive brain activity measurements, have promoted rapid advances in the elucidation of unconscious cognitive and behavioral processes [9-11]. As a result, the linear model proposed for processing from stimulation to action (Figure 2, A) based on

Figure 2 : Models of cognitive pathways



A) Classic linear model

B) Conceptual diagram of the cognitive pathways based on the latest knowledge
Human perception of stimuli is not neutral or mechanical (like visual processing in robots), but involves the selection of a subset of stimuli to perceive that dependent on each person. Interpretation of, evaluation of, and response to the stimuli also depend intricately on personal genetic traits, experiences, and memories.

Prepared by the STFC

conscious behavior has been proven invalid. Research, although still in rapid development, has suggested that multiple diverse pathways of information expression and processing take place simultaneously and concurrently (Figure 2, B). During development, physical interactions with others and the environment lead to the construction of strongly interconnected perception and behavior models in the brain, formulating algorithms for information processing.

Humans acquire the ability to unconsciously and autonomously act and perceive, partly because their motor and neural systems are endowed with structures that enable functions profitable for survival. These structures are being analyzed via molecular-biological analyses in neurobiology and neuropathology. Another reason is that they formulate perception and behavior models and algorithms in the brain through interaction with the environment. Such models and algorithms have been proposed from studies in computational neuroscience and are being refined by testing on robots. According to a hypothesis "certain orders of unconsciousness are self-organized in a bottom up manner, despite being a system that works automatically and purposelessly, because recurrent cycles of information processing are embedded" [12].

It is supposed that unconscious cognitive processes can be analyzed scientifically, rather more easily than conscious ones are, because the former are more directly linked to physical state and environmental conditions, and they function passively and mechanically according to type of stimuli and the states of subjects and environments [11]. An effective means of analysis is to reproduce perception and motion models of the human brain and recurrent information processing cycles in robots, and to evaluate the generation of autonomous actions. Robots for practical use are also expected to act autonomously, without depending exclusively on human-made programs.

A major challenge of Japanese society is to enable elderly and disabled people to live as independently as possible in an ordinary environment. It is thus meaningful to seek out prerequisites for the autonomous actions of humans and to develop supporting technologies.

According to a theory, it is only when people face external changes, which prevent them from continuing with their ongoing unconscious behaviour, that the information processing in the brain changes and hence the state known as consciousness is generated.

It has emerged that even when people take an action intentionally or focus on a single stimulus from several available, the result of their decision is already "determined" in term of neural activities in the brain areas that modulate behaviour a few hundred milliseconds before they become aware of the decision or in terms of the manifestation of elementary actions. Experimental modification of subjects' neural activities or behavioral patterns can change

Examples of definitions of consciousness and unconsciousness:

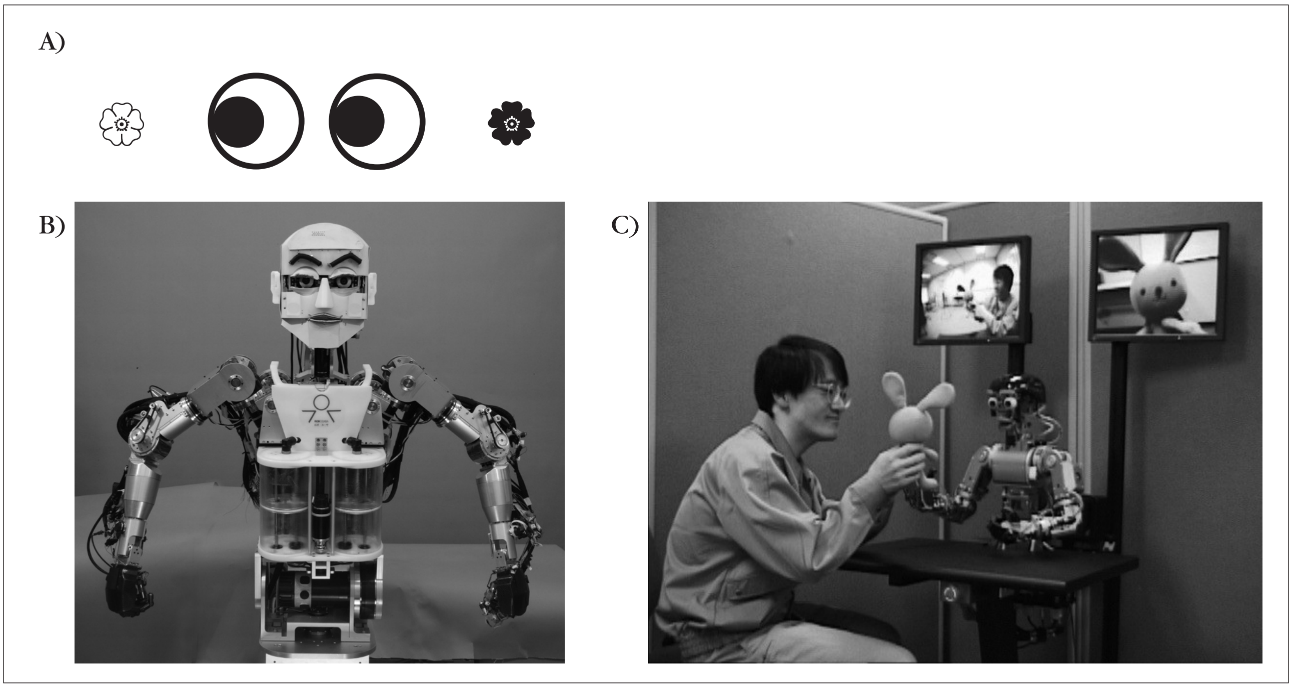
(1) Consciousness probably refers to a loose set of many interrelated, heterogeneous things rather than to a specific state or function. Consciousness can only be shaped against a "background" of unconsciousness. Unconsciousness precedes consciousness; either ontogenetically (developmentally) or phylogenetically (evolutionally) [11].

(2) Consciousness is a process of becoming aware of any inhibition against thinking or introspection, and a process of introspection, in which such awareness elicits past inhibitions against behavior (physical and mental) [30].

(3) Consciousness is defined as making approximations by performing a highly simplified fictitious series of computations to solve unconsciously generated improper configurations, associated with massive and parallel sensori-motor integration [3].

(4) Consciousness is not the cause of cognition, but only a result. Consciousness is a specific state of working memory and is meant to model unconscious manipulations as simply as possible and to store the results as episode memory [12].

Figure 3 : Examples of research on basic functions for understanding and expressing emotions



A) Humans respond sensitively to eyes and gazes, even if the eyes are presented in quite simplified form. If someone is gazing at an object, people assume that the person is interested in it.

B) Emotion Expression Humanoid Robot WE-4RII chooses and displays one of seven predefined facial patterns of emotional expression in response to external stimuli or gaze tracing. Evaluation by humans judges have shown 100-percent recognition of the state of "anger" on WE-4RII [31].

C) The infant-like robot "Infanoid" can display "eye contact," by detecting a human's frontal facial patterns from video images taken by its cameras and directing its eyes toward the detected face. Infanoid also can display "joint attention," by detecting the location and orientation of a human face or the direction of a pointing finger (based on wide-angle images for peripheral vision), searching in that direction to find an object and directing its own eyes and hands in the same direction (based on narrow-angle images for fovia vision) [13].

Prepared by the STFC from References [13,31]

the contents of decisions without the subject being aware [9,28]. Even in conscious behaviour, the processes preceding self-awareness may be unconscious and represented only physically. Algorithms of processes to find another stable behavioral pattern, even when perturbations are introduced during unconscious, autonomous action, can be both formulated and verified in robots.

Philosophers conduct thought experiments on subjects, such as whether changes occurring within a robot during this process can be regarded as the generation of the equivalence of human "consciousness" [29].

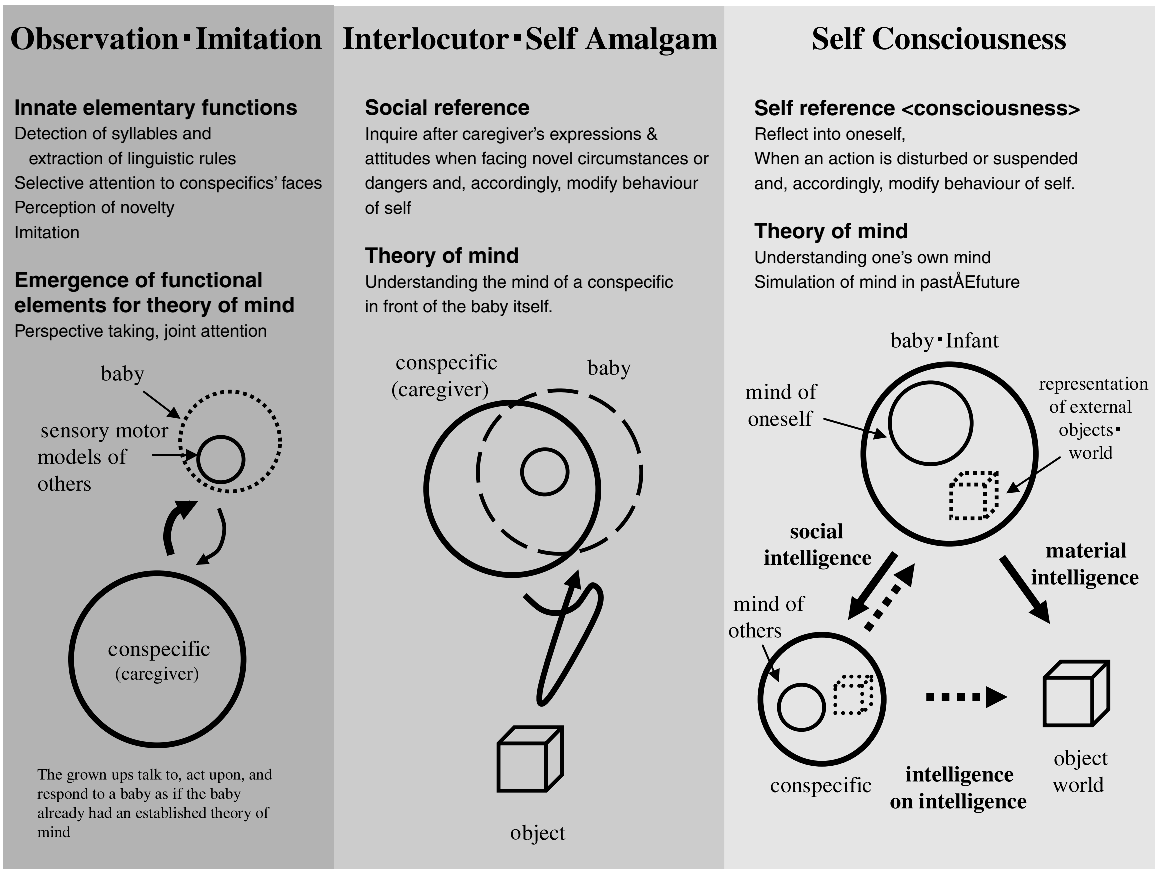
Modern law is based on the existence of free will and the concept of personal responsibility for acts based thereon. Even if science accepted a paradigm that human decision-making, whether performed consciously or unconsciously, is a mechanical process, one's own acts would still appear to be decided by one's own will from the perspective of everyday intuition and "naive psychology,". This shows that legal research is needed on such future problems as how human personal responsibility should be defined under such situations, and whether one can investigate a robot for "personal responsibility" of its acts when robots become able to act autonomously in ordinary society.

6-2 Understanding and expressing emotions

There are three potential subjects in robotics research that attempts to address "emotion": (1) elementary functions constituting emotional self-experience, (2) expression of emotions, and (3) the recognition of others' facial expressions and emotions. Research can be carried out with robots in all of these areas, and it is already underway for (2) and (3). Such research sometimes goes no further than the reproduction of experiments, which are already done with human subjects, simply replacing the latter with robots. In the near future, however, robotics research on emotion will examine how robots as manufactured products are accepted by users.

The most significant of these three concerns for cognitive robotics as a means for human understanding is (1), which deals with attempts

Figure 4 : Mind development



Observation・Imitation

Innate elementary functions
Detection of syllables and
extraction of linguistic rules
Selective attention to conspecifics' faces
Perception of novelty
Imitation

Emergence of functional elements for theory of mind
Perspective taking, joint attention

The grown ups talk to, act upon, and respond to a baby as if the baby already had an established theory of mind

Interlocutor・Self Amalgam

Social reference
Inquire after caregiver's expressions & attitudes when facing novel circumstances or dangers and, accordingly, modify behaviour of self

Theory of mind
Understanding the mind of a conspecific in front of the baby itself.

Self Consciousness

Self reference <consciousness>
Reflect into oneself,
When an action is disturbed or suspended and, accordingly, modify behaviour of self.

Theory of mind
Understanding one's own mind
Simulation of mind in pastÅEfuture

A) While babies have innate devices for mental development, they need caregivers who actively act upon and interpret them.
B) Babies first learn the significance of social interaction tools (gestures, language, objects, etc.) through interaction with their conspecifics. Since babies cannot clearly distinguish their own minds from others', they may cry in response to another's pain as if it were their own.
C) Babies soon begin to apply these tools to themselves and use them as tools to think. They also come to utilize interpretation of the contents of others' minds as knowledge. Even for adults, the self is not completely distinct from the other as represented in the chart, and this tendency is more evident when the other is affectionally close to oneself. When one observes someone close to oneself subject to a painful stimulus, it evokes the same brain activities as if one's own body were exposed to the same stimulus. An established model to distinguish the self from others can regress due to artificial factors, such as confinement and brainwashing.

Prepared by the STFC based on References [13][32]

to construct and verify functions of emotional self-experience in robots. Advances in this research will also deepen research on (2) and (3).

Chomsky's suggestion that "humans have the innate ability to voluntarily acquire language" has had an impact on various fields of research. "Baby science" and fetology have shown that humans start actively seeking stimuli and constructing their worlds as early as immediately after birth or even at the late embryonic stage. Newborns display innate functions (genetically obtained anatomical/physiological properties), such as imitation, selective attention to human faces, identification of novel stimuli from the outside world, and syllable identification and rule extraction from speech (Figure 4, A).

In psychology, on the other hand, Vygotsky proposed an "outside-to-inside" model to explain his idea that "during development, humans initially learn the significance of social interaction tools (gestures, language, objects, etc.) through interaction with others, and eventually begin to apply these tools for themselves and use them as thinking tools"[13] (Figure 4, A through C). The theory has been revaluated recently, and the importance of intervention by caregiver is emphasized in the "zone of proximal development", which children cannot reach easily by themselves. Interaction from others must initially take place in order that babies come to understand others' emotions and to express their own emotions. Self-experience of emotions, such as understanding, recognizing, and expressing the same, is established by applying the cognitive procedures one has developed to understand others' emotional expressions to oneself. Human babies initially gaze selectively into the eyes of their conspecifics, mainly caregivers, and then initiate eye contact, identify the targets of others' gazes, and point to a subject of interest to attract others' attention to it (joint attention) (Figure 4, B). This behaviour does not occur spontaneously in autistic people, who lack the ability to understand others' emotions due to developmental neurological factors.

Humans display emotion-related physical responses to stimuli before they become aware of their own subjective emotions. It is known in experimental psychology that manipulation of such emotional physical events can artificially evoke or modify subjective emotions. The mechanisms of subjective emotion largely depend on physical changes. For example, consider the process that leads to the expression of pleasant feelings through smiling.

(1) At first, the caregiver smiles at the baby. Because of an innate device to imitate others, the baby mimics it and then experiences pleasure induced by the action itself. As a result, the baby learns to smile repeatedly just to be rewarded with the pleasant sensation.

(2) The caregiver responds to the baby's smile as if it were the her/his "expression of emotion" (e.g. as if the baby were interested in something funny). Smiles are thus exchanged between the caregiver and the baby. This interaction expands when the caregiver and the baby involve outside objects as players (e.g. by pointing to, grabbing, or shaking them). Even before the baby is able to understand language, the caregiver casts it various words related to pleasure and smiling, etc. and responds to her/his actions and vocalizations as if the baby understood the linguistic expressions and their meaning. The caregiver thus integrates the baby into a language game. The baby's perception of others' facial expressions (perception model) forms in close connection with the expression of its own emotions (behavior model).

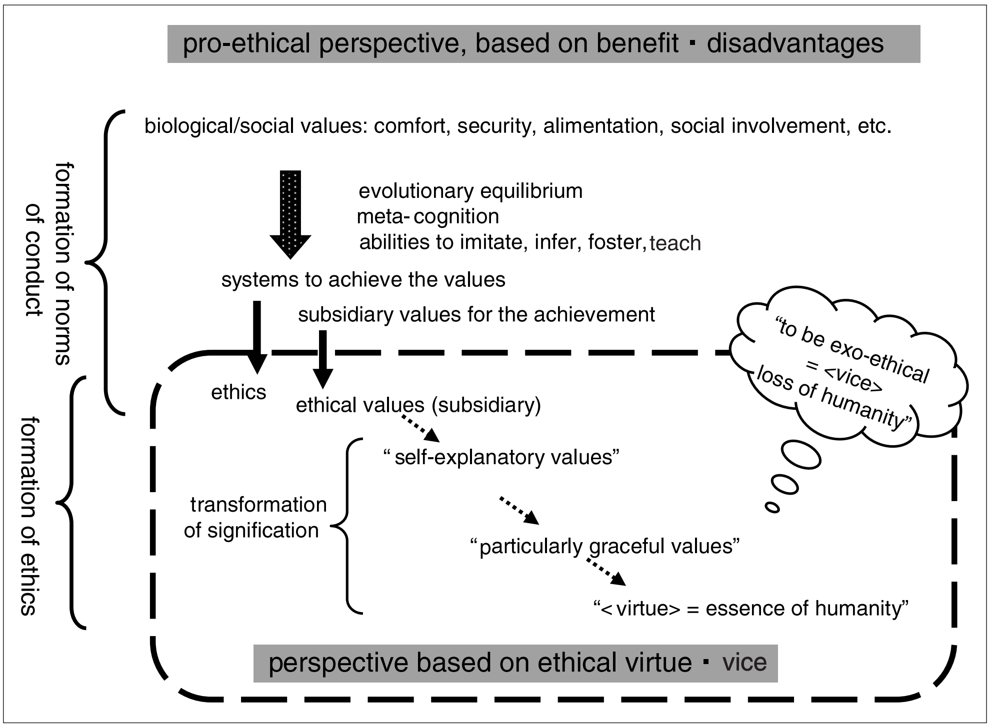
(3) When the baby sees someone smiling, it uses a reverse model of its own behavioral model and assumes that the person is experiencing a pleasant feeling (understanding others' emotions). Linguistic expressions such as "smile," "funny," "happy" and "pleasant" are linked with these assumptions through the language game.

(4) When the baby smiles (with a behavioral model), it uses understanding of others (perception model) and associated linguistic expressions to form understanding of its own emotions, i.e., "I smile because I'm happy."

It is expected that such a series of phenomena and brain algorithms can be constructed in robots to evaluate the mechanisms by which human interaction develops the elementary functions of emotion.

Children are believed to develop a "theory of mind" (a theory about the state of others' minds, which is unverifiable) to estimate and understand the mental states of others. Because it can be said that one "has a theory of mind" when becoming capable of estimating another's beliefs, intentions, and knowledge, this can be tested with the ability to discern a false belief in another (a false belief task). Research on monkeys showed that a theory of mind could not be detected with false belief tasks, even though they demonstrated joint attention and other abilities necessary to establish a theory of mind[32]. Human babies seem to develop such abilities spontaneously. However, to allow children to fully develop a theory of mind, it is important that parents and caregivers talk to and respond to them as if they already had a complete theory of mind, even before children have actually developed one. Monkeys and current robots cannot develop their own theory of mind in response to humans, even if humans project emotions onto them and interact with them as if they had minds. Simulations of human development with robots will shed light on mechanisms (1) to perceive communication behaviour from others and to form in oneself a model for the same behavioiur and another model of one's own mental state, induced by communication, and then (2) to act upon younger

Figure 5 : Mechanisms for the generation of norms of conduct



Prepared by the STFC based on References [33, 34]

conspecifics to help them develop a theory of mind.

6-3 Research on the mechanism for establishing social codes of conduct

Since human beings are highly social animals, codes of conduct are significantly influenced by the interests of the community as well as those of individuals. Social psychologists believe that algorithms concerning the generation of social comportments and social codes can be mathematically formulated, while taking into consideration evolution and the balance of rewards resulting from selfish behavior and altruistic behavior (evolutionary stable equilibrium[33] in Figure 5). For instance, individual drivers on a freeway choose their lanes and speeds based on their own best interests. On a macroscopic level, however, the movement of a group of vehicles on a freeway can best be described by fluid dynamics. While it is almost impossible to convey an algorithm for the relationship between the behavior and interests of an individual driver to the mass of drivers, developing an algorithm for the optimization of the macroscopic movement of a group of vehicles is relatively easy, and it could be "genetically" passed on to all drivers. This also applies to the generation of behavior patterns in many non-human organisms as well.

Because such genetic factors impose only loose constraints on individual human beings' behavior, diversity is generated. In humans, neural circuits have developed that enable humans to (1) imitate the actions of those who are closely related to themselves, (2) observe others and imitate actions that have brought them profits, (3) integrate the actions of others into their own inventory of actions (learning), (4) reverse the neural model for behavioural learning and use it to interpret others' behaviour and to infer their mental state, (5) change their own behaviour according to such understanding, and (6) generalize cognitive and behavioral procedures acquired in actual situations to apply them to different situations (meta-recognition). At the same time, humans have acquired the computational ability to predict and evaluate, from distant temporal and spatial perspectives as well as actual ones, rewards (safety and comfort) and punishments (risk, hunger, isolation, instability of the community as a whole, etc.) resulting from their own actions.

Robots in the form of experimental rodents were designed to simulate the neural networks responsible for reward prediction and evaluation (Cyber Rodents). Parameters of their expectation of rewards, patience for reward acquisition, and sensitivity to punishments are manipulated and the various behavioral patterns are analysed[35]. The rodents are designed to give and receive battery chips (foods) and to imitate others' behaviour, so that their learning, propagation, and genetic transfer of social behaviour can be simulated. Through the fusion of research of evolution in biology and anthropology, social psychology and cognitive robotics, it is expected to clarify the genetic bases and algorithms of cooperative behaviour.

The algorithm to optimize the interests of the community is rather something to minimize the possibility of the majority of people suffering from evil deeds than to prohibit specific individuals from committing an evil deed. If one cannot infer another's mind that "wants not to suffer" and cannot control one's behaviour based on estimation and understanding, her/his evil deed will never be self-controlled. Groups of robots that simulate cooperative behaviour can facilitate exploration of the relationships between the theory of mind and ethical behaviour.

It may be possible that human specific psychoneurosis and behavioral disorders have resulted from an enormous increase in the amount and complexity of information processing in the brain and patterns of social behaviour. For example, it is supposed that "schizophrenia is a by-product of the evolution of linguistic ability," and that "the dissemination of phonemic characters, printed documents, and new technologies have changed human sensory and cognitive patterns" [36]. In the course of increasing the complexity of cognitive functions of robots to make them more humanlike, researchers may accidentally and unexpectedly encounter robots that display cognitive/behavioral patterns inconvenient to others. Given that advances in neuroscience have been driven in part by the elucidation of neurological disorders and developmental diseases, scientists can build robots that simulate diseases that may cause deviations from normal cognitive/behavioral patterns under specific conditions, and analyze them to elucidate how the brain shifts between normal and abnormal states. They can also develop systems that are robust against deviation from normal states.

Most scientific findings on the unconscious sides of mental processing, decision-making, and the manifestation of codes of conduct may not be compatible with "naive psychology," relying upon the conventional understanding of the world. Although people may benefit from scientific knowledge, it is not always easy for them to use such knowledge to interpret their daily lives. Indeed, people may not need such knowledge in their everyday lives.

It is important that philosophers and social scientists engage in research in cognitive robotics by analysing findings from their own viewpoints in order to formulate theories and models compatible with the real world.

7 Ethical Debates Raised from Cognitive Robotics

In a stable society, perhaps "the fact that morals are nothing but means (to realize a safe and comfortable society) is better kept as tacit knowledge. Therefore, people may prefer to be passively convinced of morals"[34] ("formation of ethics" in Figure 5). This is a reason why explanations of human behaviour with pro-ethical values, such as profit/loss, or comfort/discomfort, are in themselves considered a vice. Ordinary people may not necessarily feel agreeable about sciences clarifying the mechanism by which norms of conduct are formed.

In human history, there have been many occasions when existing moral systems had to be reviewed and reconstructed facing the collapse of social orders or in the course of social reform. Traumats of these events have been passed down in the form of myths, legends, and histories. In general, Japanese natural scientists know little about the relationship between the formation mechanisms of ethics and norms of conduct, although they are usually eloquent in explaining the latter. It is imperative to involve social scientists and cultured persons in debates on desirable procedures to communicate scientific knowledge to the public and on the social impact of scientific knowledge. Political efforts will be needed in such areas to organize forums collecting participants from diverse fields.

Those who believe in the existence of virtue and vice without questioning their origin may assume that advances in genetics, neuroscience, and behavioural sciences will make it possible, through appropriate use of medication and gene manipulation, to prevent individuals "born with the propensity to commit crimes" from committing actual deeds. However, if neither virtue nor vice is inherent in each individual, this prediction or expectation is misplaced.

For academic disciplines intended to improve our understanding of humanity, broad and serious discussions should be held not only on how far we can explore scientifically, but also on how far we can conduct research and whether there are areas, topics, and depths that must not be reached. If such discussions result in a decision whereby some restrictions should be imposed on research activities, actual procedures to implement them would then have to be discussed, and conclusions should be made publicly available to ensure effective implementation.

8 Policies to Promote Cognitive Robotics Research

8-1 Fundamental concepts in cognitive robotics

For new sciences and technologies to emerge and advance, conceptual bases of development is required. Crucial concepts arise in the very early stages of research. They influence the selection of noteworthy issues, effective methods, and feasible goals, and impact on the feedback circuits of verification processes. Researchers should recognize a comprehensive framework of concepts from their earliest inception and intentionally construct theory likely to lead in the direction of their ultimate goals.

Humanoid robot research is based on a variety of visions and concepts (Table 3). For example, in Japan the concept of "mechatronics" was proposed in the 1970s as a complete fusion of mechanics and electronics. The idea of compromising quite different elements has been described as very Japanese[16]. One Japanese robotics researcher states, "although the creation of the concept of mechatronics may have had no direct influences on society, it helped Japan decide in which direction to go, and with this reliable guide into the future, people felt at ease to follow the direction" [37].

Mechatronics was a product of the fusion of scientific concepts among different areas of engineering. By contrast, cognitive robotics to understand human beings, as an emerging discipline, requires the incorporation of concepts from fields other than the natural sciences and engineering, interaction among vague notions that have yet to be established as academic

Table 3 : Fundamental concepts of humanoids

Concepts	Effects • Results
A biological system can be studied, in principle, relative to the analogy of machines (Macy conferences, 1946 - 53)	Feedback
Based on biological principles, machines can be designed to have biological functions.	Bionics (1960 -)
In order to release humans from labour that is unsuitable for humans (monotonous or tough), machines should be created to undertake it (Wiener)	Automation
Mechatronics The unification of mechanics and electronics (1972 -)	Amelioration of structure / actuator / control systems of robots
Autopoiesis, neural network, connectionism	Robustness of robots against incidental changes
Embodiment, embedding, situated	Emergence
To investigate information processing of the brain to the extent that artificial machines, either computer programs or robots, can solve the same computational problems, as the brain, essentially in the same principle	Cognitive developmental robotics (1994 -)
Comprehensive studies, including theories & the humanities	Cognitive robotics

Prepared by the STFC based on References [3,37] theories (requiring the creation of new fields), and interaction with real-world knowledge, culture, and naive psychology, which are not academic disciplines.

8-2 Emphasis on theory-oriented research

In Japan, research in both neuroscience and robotics often focuses on material oriented subjects, as described in Figure 1. For example, neuroscience focuses on analysis of the brain's structure and functions, while humanoids research emphasizes the development of structures and actuators, and the analysis of their motor controls. The same tendency is seen for research budgets. While large budgets are appropriate for research projects in experimental brain sciences and research intended to lead to physical products rather easily, it is hardly the case for those in theoretical and mathematical research. Psychological, theoretical and computational research is essential for a comprehensive understanding of brain functions and the mechanisms of the mind based on physically oriented research, as well as for building robots with cognitive patterns so humanlike as to be able to walk autonomously in actual societies. However, such less material-orientated researches are often despised. Although Japan has had outstanding researchers in computational neuroscience for many years, they remain few in number and have little opportunity to interact with most experimental neuroscientists and robotics researchers. Participation in brain science by mathematicians and theoretical physicists is also insufficient. Some Japanese humanoids manufacturers take advantage of European strengths in mathematics and physics by assigning European labs to perform research on cognition and theories of cognition-behavior correlation, while their domestic laboratories develop exclusively structures and actuators. While broad cooperation with overseas researchers may be important, Japan should first consolidate its own human resources for research, and facilitate domestic exchanges of knowledge and ideas for further refinement. As research in brain systems, including their interaction with the body and the environment, advances, research may reach a point where it cannot progress farther without breakthroughs in research on functional units, such as visual or auditory, and modules. A potential solution is ensuring an environment in which research projects adopting diversified approaches, ranging from analytical study of individual functions to comprehensive research on systems, progress in parallel while interacting with each other. This is essential for the culture of science to thrive in this country.

Japan should therefore first build a continuous chain of knowledge sharing and cooperation on various levels, through material-oriented research in biology, theoretical and mathematical research, to material-oriented research in engineering. For example, researchers in neuroscience, psychology, and engineering may as well provide opportunities to clarify the theoretical problems that must be solved for the construction of robots, to consult extensively with mathematicians and theoretical physiologists on how these problems can be solved, and to cooperate with these scientists in solving the problems.

8-3 Selective promotion of philosophy and social science

The relationships between mind and body, the self, conspecifics, or the environment are subjects of research, not only in the natural sciences, but also in the humanities and social sciences. Scientific findings on the human mind usually require verification by philosophy and social science before they can be seen as truly applicable knowledge in society. Natural scientists would have to be reinforced with philosophers and the social scientists in order to clarify visions and conceptual framework of their research.

However, in today's Japan, there are not enough research activities in philosophy and the social sciences that could immediately make such contributions. An effective solution is to clearly distinguish research that rigidly adheres to the history of philosophy (or social science), to the interpretation of preceding studies, to imported knowledge or to issues focusing only on lectures, from that research endowed with flexible thinking and the potential for verification that allows research to address actual problems in the world. At the same time, research organizations should selectively hire researchers with flexibility and criticism. This will help foster philosophers and social scientists who can derive original ideas, using cognitive robotics as a practical test platform.

Philosophical research should not necessarily be confined to the department of literature of universities. The efficiency of philosophical research could be improved by moving philosophy laboratories from environments that can no longer encourage the proposal of new questions and solutions to centers of active research in natural sciences and engineering, such as cognitive robotics and neuroscience. This would allow philosophers to perform research that involves interaction with researchers in other fields. Philosophers could verify their theories through psychological experiments and simulations with robots, while researchers in robotics, cognitive science, and neuroscience could seek advice and criticism from a philosophical perspective on their research concepts and interpretation of experimental results. Providing philosophy laboratories that adopt such research systems with as much research funding as natural science research labs receive, namely to a level sufficient for conducting simulation experiments by themselves, would raise the self-esteem of active philosophers. Another effective measure would be to allow philosophy majors to be fostered in such research laboratories, after being taught in basic undergraduate programs for two years, so that they could develop new areas of research outside the traditional world of philosophy research.

8-4 Establishing a virtual research center

When there are several organizations on the frontiers of research, they can operate more efficiently when combined into a research center where they could develop comprehensive visions and conceptual framework and verification systems. It would also serve as a driving force for other research projects and organizations. In a situation where this is not immediately feasible, an effective alternative is the establishment of a virtual research center that can provide ample funds, better research environments, and release from routine burdens, as well as assists in contacting and cooperating with other domestic or overseas organizations. Guidelines, objectives, and evaluation of the outcomes of research conducted there must be made widely available to the public.

In this virtual research center, graduate students and postdoctoral researchers (e.g. in philosophy, psychology, theoretical biology, mathematics, etc.) would be allowed to constantly participate in research activities of engineering and information technology labs. For example, a cognitive robotics research project at a mechanical engineering lab may need the in-situ participation of researchers in philosophy, psychology, and theoretical biology, but it is often not easy for the lab to hire specialists from other fields or to grant their academic degrees. By hiring researchers and assigning them to such a lab (or labs), the virtual research center could benefit by flexibly utilizing competent human resources and broadening researchers' perspectives.

8-5 Developing science policy perspectives based on advanced research

The paradigm "to investigate information processing of the brain to the extent that artificial machines, either computer programs or robots, can solve the same computational problems as solved by the brain, essentially in the same principle," has inspired research attempts that are quite radical, albeit scarce. It spurred movements to verify theories and models of the human brain by integrating findings in diverse fields and constructing robots, even before brain structure and functional modules were completely elucidated. This approach of "constructing robots at first"[8] and attempting verification second assumes that robots can be made with minds, based on a broad understanding of what the mind is. This yields higher intellectual productivity than the conventional approaches of "taking an action only after it has been enforced with theories in a specific field."

Many of these researchers have promptly introduced findings and perspectives in psychophysics into their research. As a result, they have come to address consciousness and unconsciousness and the boundary between subject and object, i.e., the self, in addition to traditional approaches on individual function, such as sight, hearing, motion, memory, or emotion.

The significance of such radical approaches originated by Japanese researchers has caught the eyes of European and American researchers and science policymakers more quickly than their Japanese counterparts. For example, the European Commission launched research projects such as Neurobot and Cognon in 2004 under the financial support of its 6th Framework Program, forming consortia to promote them. Unlike Japan, Europe is skeptical of the idea of producing humanoids for the use of the general public, but in 2004, it organized RobotCub, a consortium to create infant-like robots as an open platform for research in cognitive science and neuroscience. Now that humanoid robots made in Japan are employed at overseas research institutes, European and American researchers are going to use them actively as platforms for developing software to serve as the cognitive mechanisms of robots.

Many administrators working in the European Commission have doctorates in specialized areas, such as science, engineering, medicine, and psychology. They have carried out research and written scientific articles by themselves. These experts are engaged in diverse areas of work, ranging from general research funding to science and technology foresight. When innovative articles and reviews are published, these officials can detect their potential and significance at an early stage and plan and promote research projects in those areas. Although Japanese organizations participate in the European Commission's consortia, so far there has been no Japanese equivalent. There is even the possibility that the buds sprouting from the seed of Japan's advanced research will be harvested in Europe. As certain Japanese researchers stand at the forefront of the world research community, the Japanese government should have a comprehensive plan to promote their research by clarifying what kind of system of knowledge should be constructed and how scientific findings should be applied for future social systems.

9 Conclusion

Knowledge on the human mind and behaviour has to be shared by all human beings. During past debates over the Human Genome Project and the patenting, industrial application, and market value of the DNA sequence, a major concern was, "to whom does the human genome belong?" The question "to whom does knowledge of the human mind belong, or can it belong to anyone at all?" is likely to spur even more serious debates. Japan should take leading roles in building a system that allows all people worldwide to share accurate knowledge. Particularly in the 21st century, applied technologies that make use of findings on human cognition or that control cognitive processes - so-called science and technology to exploit the brain - will no doubt bring advances in many ,fields such as economics, manufacturing, labor, entertainment, medicine, education, politics, and diplomacy. Preceding the spread of this sort of practical applications, the promotion of "studies to explore the brain" has to vitally deepen our fundamental understanding of the mind as actually grasped by human beings.

In cognitive robotics, scientists from different disciplines can unite, scientifically elucidate questions on the mind by using robots as a common concrete platform, and interpret those issues in social contexts. Because robots easily attract public attention, comparison of robots and human beings is expected to facilitate the discussion and examination of questions on the mind.

References

[1] Shunichi Amari, "Mathematics of Neural Networks: The Cerebral Information Processing System," Sangyo Tosho, 1978. (Japanese)

[2] Masao Ito, ed., "Higher Cerebral Functions and Central Programming," Sangyo Tosho, 1976; Nakaakira Tsukahara, ed., "Cerebral Information Processing," Asakura Publishing, 1984. (Japanese)

[3] Mitsuo Kawato, "The Computational Theory of the Brain," Sangyo Tosho, 1996.

(Japanese)

[4] J. Tani, “An Interpretation of the ‘Self’ from the Dynamical Systems Perspective: A Constructivist Approach,” Journal of Consciousness Studies, No. 5, pp. 5-6, 1998.

[5] Hiroshi Komatsu, “Research and Development Trends in Robot Technology — Toward Promoting the Commercialization of Personal Robots,” Science & Technology Trends — Quarterly Review, No. 10, pp. 37-49, January 2004.

[6] Kumi Okuwada, “Trend in the Introduction and Development of Robotic Surgical Systems,” Science & Technology Trends — Quarterly Review, No. 10, pp. 75-91, January 2004.

[7] Masakazu Ejiri, “Challenges and Prospects for the Promotion of Robotics,” National Institute of Science and Technology Policy, lecture anthology, Vol. 116, 2003. (Japanese)

[8] Hideaki Sena, Minoru Asada, Kenji Doya, Jun Tani, Kenichiro Mogi, Kazuo Hiraki, Hideyuki Nakashima, Hiroshi Ishiguro, Yasuo Kuniyoshi, and Tomohiro Shibata, “Mystery of Knowledge: A Challenge in Cognitive Developmental Robotics,” Keihanna Research Group for Sociointelligenesis, comp., Kodansha, 2004. (Japanese)

[9] B. Libet, “Mind Time: The Temporal Factor in Consciousness (Perspectives in Cognitive Neuroscience),” Harvard University Press, 2005

[10] Shinsuke Shimojo, “The Future of Subliminal Mind and Potential Human Perspectives,” Chuo Koronsha, 1996. (Japanese)

[11] Shinsuke Shimojo, “What is Consciousness?,” Kodansha, 1999. (Japanese)

[12] Takashi Maeno, “How to Make a Conscious Robot: Fundamental Idea based on Passive Consciousness Model,” Journal of the Robotics Society of Japan, January 2005 (Vol. 23, No. 1). (Japanese)

[13] Hideki Kozima and Akira Takada, “Developmental Approach to Social Interaction,” Journal of the Japanese Society for Artificial Intelligence, Vol. 16, No. 6, 2001 (Japanese); Hideki Kozima, “What Can Robots Do for the Education of Handicapped Children,” included in Shinichi Watabe, ed., “Education of Handicapped Children in the 21st Century’s Technological Society,” pp. 105-113, 2004. (Japanese)

[14] R. I. M. Dunbar, “Origin of Language,” Kagaku, Vol. 67, pp. 289-296, 1997. (Japanese)

[15] European Commission Profile IST 2005: http://europa.eu.int/information_society/research/robotics/docs/robotics_profile_052.pdf

[16] P. Dario, et al., in the Italy-Japan 2005 Workshop “The Man and the Robot: Italian and Japanese Approaches”: http://www.robocasa.net/workshop2005/index.php?lang=jp&page=program

[17] A. Stoica, et al., “Humanoids for Lunar and Planetary Surface Operations,” Proceedings for 2005 5th IEEE-RAS International Conference on Humanoid Robots, pp. 345-350, 2005.

[18] Hubert L. Dreyfus, “What Computers Can’t Do: The Limits of Artificial Intelligence” (Revised Edition), Harper & Row, 1979

[19] Yukihiro Nobuhara, ed., “Philosophy of Mind: Robot Version,” Keisoshobo, 2004. (Japanese)

[20] D. R. Hofstadter, “Gödel, Escher, Bach: An Eternal Golden Braid,” Basic Books, Inc., New York, 1979

[21] Masao Kurosaki, “Did Philosophers Dream of Androids? —Philosophy of Artificial Intelligence,” Tetsugakushobo, 1987. (Japanese)

[22] S. Turkle, “Relational Artifacts/Children/Elders: The Complexities of CyberCompanions,” in Proceedings, CogSci-2005 Workshop: Toward Social Mechanisms of Android Science: http://www.androidscience.com/proceedings2005/TurkleCogSci2005AS.pdf

[23] Takeshi Umehara and Hisakazu Fujimura, ed., “The Dawn of Ainu Studies,” Shogakukan, 1994. (Japanese)

[24] Atsuo Takanishi, et al., in Italy-Japan 2005 Workshop “The Man and the Robot: Italian and Japanese Approaches.”

[25] B. Lee, “Interpersonal Perception and Android Design”:

http://www.androidscience.com/proceedings2005/LeeCogSci2005AS.pdf

[26] Masahiro Mori, "The Uncanny Valley," Energy Vol. 7, No. 4, pp. 33-35, 1970. (Japanese)

[27] E.T. Hall, "The Hidden Dimension: Man's Use of Space in Public and Private," The Bodley Head Ltd, London, 1966.

[28] Shinsuke Shimojo, "The Brain's Decision-Making Mechanism and the Body," in the Lecture Anthology for the 23rd Annual Conference of the Robotics Society of Japan, p. 13, 2005. (Japanese)

[29] Masayoshi Shibata, "The Robotic Mind: Seven Philosophical Stories," Kodansha, 2001. (Japanese)

[30] Tadashi Kitamura, "Can a Robot Have a Mind? — Introduction to Cyber Consciousness Theory," Kyoritsu Shuppan, 2000. (Japanese)

[31] Atsuo Takanishi, et al.: http://www.takanishi.mech.waseda.ac.jp

[32] T. Matsuzawa, ed., "Primate Origins of Human Cognition and Behavior," Springer, Tokyo, 2001.

[33] Tatsuya Kameda and Koji Murata, "Social Psychology to Solve Complexity: Humans as Adaptive Agents," Yuhikaku Publishing, 2000. (Japanese)

[34] Hitoshi Nagai, "Philosophy for Children," Kodansha, 1996 (Japanese); "Manga Philosophize," Kodansha, 2004. (Japanese)

[35] C. Capi and K. Doya, "Evolution of Neural Architecture Fitting Environmental Dynamics," Adaptive Behavior, Vol.13 (1), pp. 53-66, 2005; K. Doya, "Metalearning and Neuromodulation," Neural Networks, Vol.15, No. 4/5, pp. 495-506, 2002.

[36] M. McLuhan, "The Gutenberg Galaxy — The Making of Typographic Man," University of Toronto Press, 1962

[37] Susumu Tachi, "Introduction to Robots — Philosophy for creation and wisdom for use," Chikuma Shinsho, 2002. (Japanese)

Kayoko ISHII, PhD
Life Science Unit, Science and Technology Foresight Center

(Original Japanese version: published in March 2006)