The world first self-aware robot and
The success of mirror image cognition

Junichi Takeno

Robot and Science Institute,
Computer Science, Meiji University
Kawasaki 214-8571, Japan
takeno@cs.meiji.ac.jp
JunTakeno@aol.com

6th November 2006, Karlsruhe University
10th November 2006, Munich University, Germany

At the beginning

The purpose of this research

To learn and understand the mechanism of human cognitive functions.

At CIRA 2005 in Helsinki

To provide for the unexpected emergence of a consciousness through the self-evolution of robots

Inspired from Hal's eye

from movie "   Space Odyssey"

- I would like to understand the mechanism of human cognitive functions.
  - **She plays an accordion almost subconsciously. But, she knows what she's going to do consciously. What is the mechanism of her cognition and behavior?**
  - **What is her thought process? What is her feeling process? How does her creativity function?**
  - **What is her will? What is the state of human consciousness?**

- And another purpose of this research is to provide for the unexpected emergence of a consciousness through the self-evolution of robots
  - **It is very dangerous for us because humans cannot understand any conscious mechanism of the robot right now.**

The method

- Using a robot

- With artificial neural networks

Using a robot

**Receptionist Robot
at IREC2005**

Actroid at IREC2005
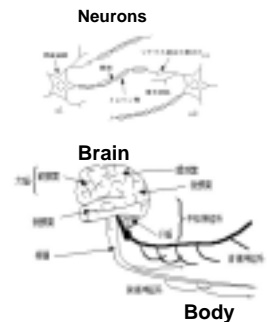
A robot is a very simple model of a human, I think.

**Multiple
Sensors**

Actuators

**Computing
Units**

at IREC2005

Brain and body

Neurons

Brain

A human receptionist

Body

The reason

- The robot has a body that is controlled by a computer.

- Also, the computer is informed about conditions of the body and the outer world simultaneously.

- And artificial neural networks have a mathematical analogy to human brain neural connections.

An expected effect

- **I will agree at this time that no robot is alive in a human sense of living.**

- **However, a robot becomes a subject for experiments about learning about human cognitive functions because there is a similarity in the relationship between the computer and the body.**

- **That is, a robot is a very simple model of a human.**

The Mirror Test

Bari, Italy

- The first target is an experiment on mirror image cognition by the robot. We call this experiment "The Mirror Test."

- The experiment is an evaluation method to learn whether a non-human being has the highest cognitive functions like humans.

- Humans become aware of their self image in a mirror. This is the origin of human self-consciousness, I think.

- So, if a living being identifies its reflected self-image in a mirror, we may judge that it possesses a self-consciousness function.

- Chimpanzees, orangutans and dolphins are successful in the test.

- From the French psychoanalyst Jacques Lacan's argument, the mirror test is a very important tool for learning about one of the human consciousness functions.

- If the experiments with a robot are successful, understanding the robot's mechanism may be our first step towards learning about human consciousness.

The introduction

What is mirror image cognition of the self?

- **We want to understand why humans are aware that the reflected image in the mirror is the self.**

- **However, nobody knows the reason yet.**

How do we make a robot aware of the self?

Simple reflex
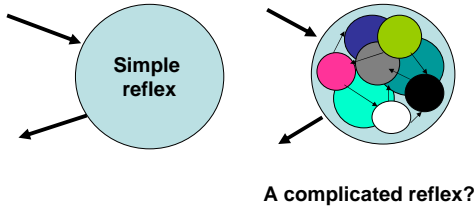
A complicated reflex?

- The problem must solved as a cognitive one.

- It is not a problem of behaviorism.

- This means that the process of the robot's awareness is not a simple reflex.


But, is it then only a complicated reflex?

- We believe that the complicated reflexes of robots must have a system of awareness, of course, because we are aware of it.

What kind of basic functions will be needed in the system?

the self

oneself

the other

The Self robot

the self

oneself

the other

The Other robot
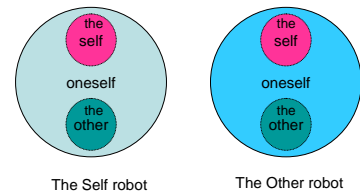
- **The system can discriminate the behavior of the self and the other.**

What kind of basic functions will be needed in the system?
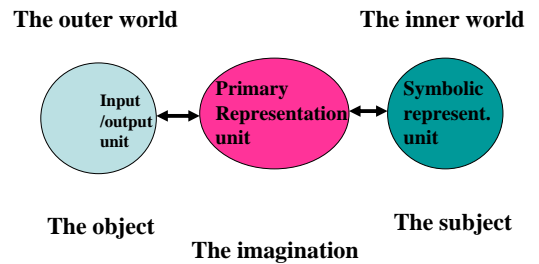

Imitation Behavior
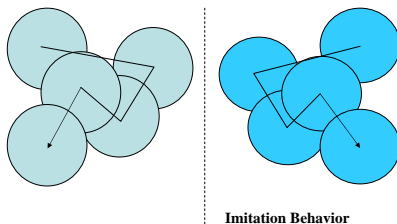
- **The system must be able to imitate the behavior of the other in a mirror-image-like manner because infants are capable of imitating in the same manner.**

How can we make the basic functions?

The object

The imagination

The subject

The triplicity by I. Kant

•**The concept of the basic model was inspired by the triplicity of the object, the subject and the imagination connecting them as advocated by I. Kant, the German philosopher.**


The outer world   The inner world

Input/output unit — Primary Representation unit — Symbolic represent. unit

The object   The imagination   The subject

• **The basic model comprises an input/output unit to deal with the outer world (the object), a symbolic representation unit to represent the inner world (the subject), and a primary representation unit as the imagination.**

• **The primary representation unit transmits information between the outer and the inner worlds.**

How can we form the model in a practical and concrete material manner?
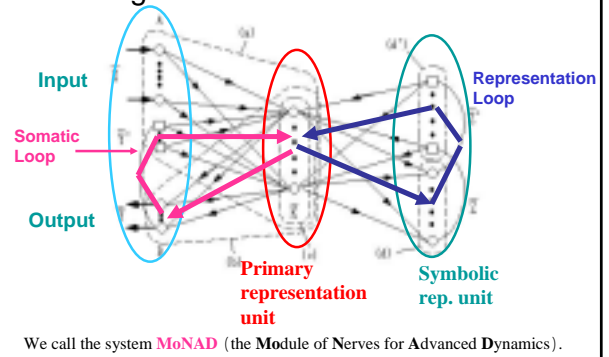
This is a difficult problem!

Conceptual model
of a self-aware robot

Our basic model

The outer world   The inner world

Input   Representation Loop   c1

Somatic Loop   c2

S11

Output   Primary representation unit   Symbolic rep. unit
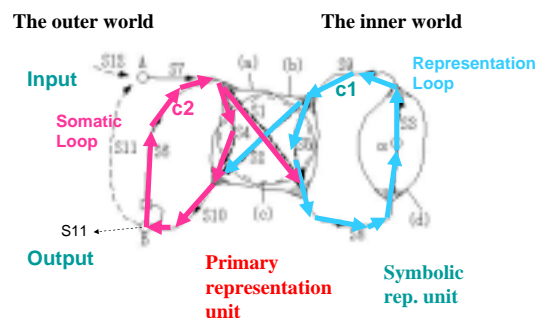
Computational model of a self-aware robot

A realization of the model using recurrent neural networks

Input

Somatic Loop

Output

Primary representation unit

Representation Loop

Symbolic rep. unit

We call the system **MoNAD** the **Mo**dule of **N**erves for **A**dvanced **D**ynamics .

Features of the MoNAD

- **MoNAD can learn imitation behavior from the other in a neural networks manner.**
- **And also, it can learn representations of the behavior of the self and the other at the same time.**
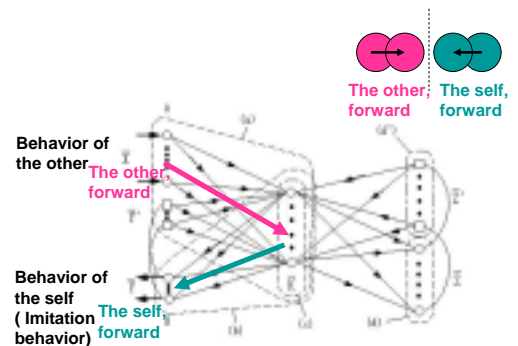
↓

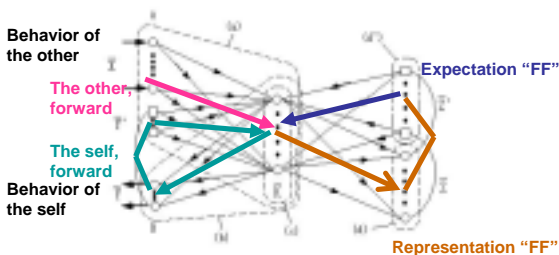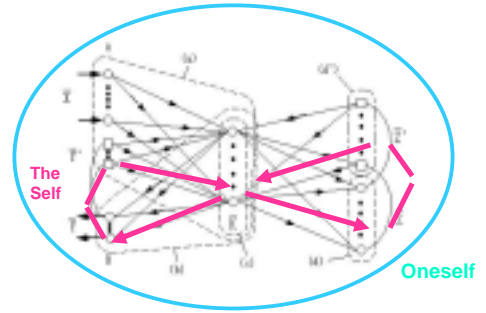**MoNAD can recognize both the behavior of the self and the other in parallel computational process.**

↓

I think that MoNAD has basic the functions of consciousness.

MoNAD can learn imitation behavior from the other in a neural networks manner.

The other, forward **The self, forward**

Behavior of the other **The other, forward**

Behavior of the self (Imitation behavior) **The self, forward**

And also, it can learn representations of the behavior of the self and the other at the same time.

Behavior of the other

The other, forward

The self, forward

Behavior of the self

Expectation "FF"

Representation "FF"

MoNAD can learn imitation behavior from the other in neural networks manner.

Behavior of the other

The other, Stop

The self, forward

Behavior of the self **Stop**

Expectation SF

Representation SF
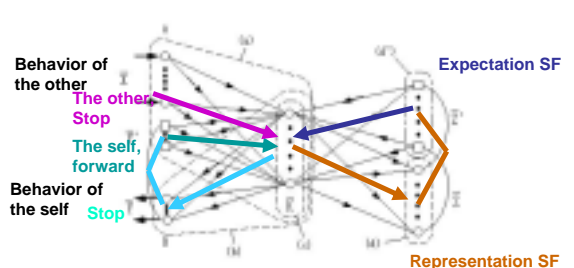
Ten features of the function of consciousness, based on the knowledge of Husserl:

1. Duality of sense of self,
2. orientation,
3. relationship between behavior and result,
4. expectation,
5. function of determination and conviction,
6. embodiment,
7. consciousness of others,
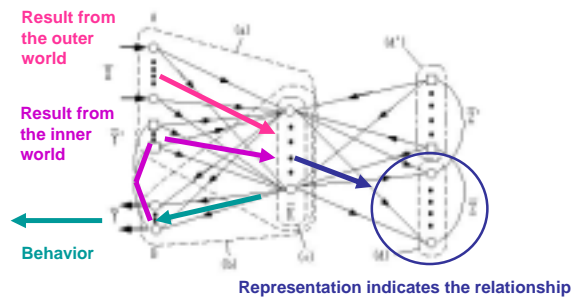8. feeling and thought,
9. chaos, and
10. emotion.

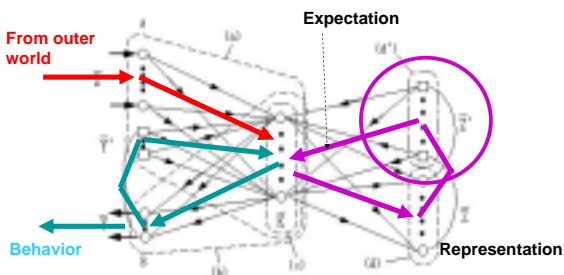# 1. Duality of sense of self,

The Self

Oneself

---

2. orientation,

• Imitation learning is one an orientation.

• And MoNAD can learn an actualization of the representation's expectation .
  This behavior is also an orientation.

---

3. relationship between behavior and result,

Result from the outer world

Result from the inner world

Behavior

Representation indicates the relationship

---

4. expectation,

Expectation

From outer world

Behavior

Representation

---

5. function of determination and conviction,

• MoNAD decides a behavior from the expectation.

• I do not yet give an explanation of the conviction.

6. embodiment,

- MoNAD is also able to detect its embodied own behavior using a somatic loop.

7. consciousness of others,

- The representation of MoNAD is a recognized result which indicates the behaviors of the self and the other.

- Therefore, MoNAD is able to detect the relationship between the self and the other.

- This is the first step toward consciousness of others, I believe.

8. feeling and thought,

Thought process
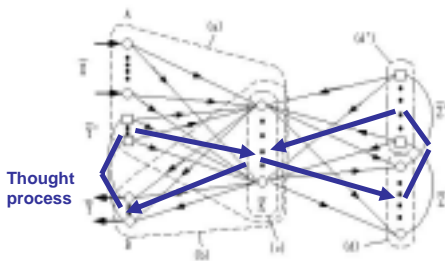
9. chaos, and 10. emotion

- Chaos can arises in the neural networks at any time.

- I do not yet offer an explanation about emotions.

About constructing more complex behaviors.

More complex behaviors will be implemented by using combined MoNADs.

Each MoNAD can communicate from lower and higher MoNADs.

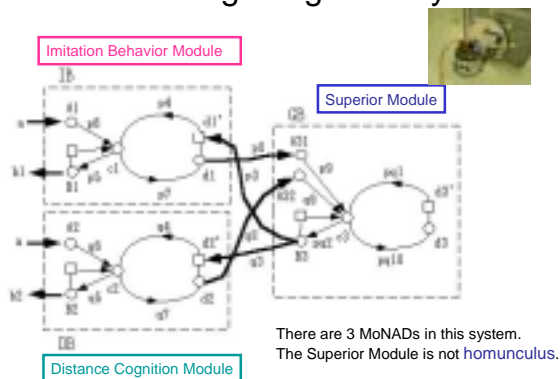**Cognition process**

Behavior process
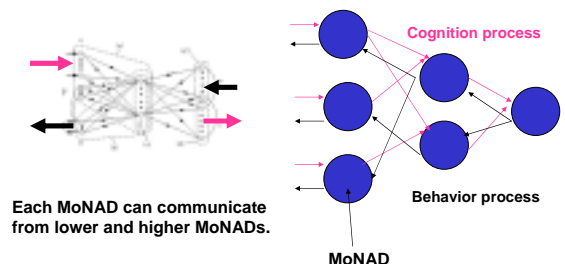
MoNAD

The MoNAD system has huge merits.

- The system can solve **the Symbol Grounding Problem** because the system can learn the relation between an environment and the symbol representation.

- The system can solve **the Binding Problem** because the system can cycle through cognition-behavior.
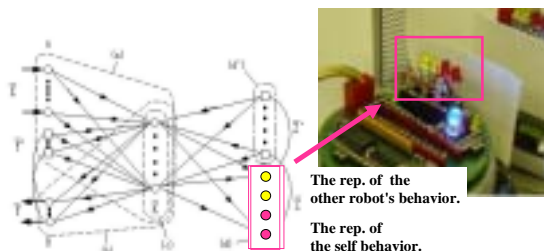

About the mirror image cognition experiments

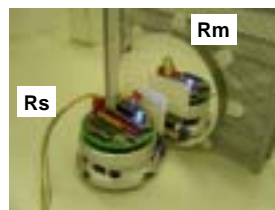

The mirror image cognition system

Imitation Behavior Module

Superior Module

There are 3 MoNADs in this system.
The Superior Module is not homunculus.

Distance Cognition Module

IB

DB

The role of the representation part in the Imitation Behavior Module .

The rep. of the other robot's behavior.

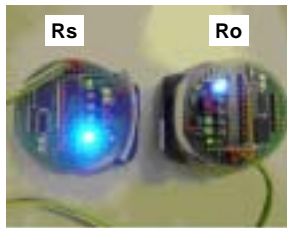**The rep. of the self behavior.**

Experiments

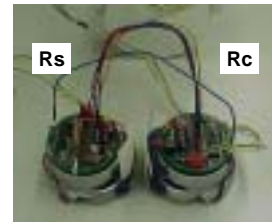## The experiments No.1

Rm

Rs

The robot Rs imitates the action of its own image Rm as reflected in a mirror.

The experiments No.2

Rs   **Ro**

The robot Rs imitates an action performed intentionally by another robot Ro as imitative behavior.
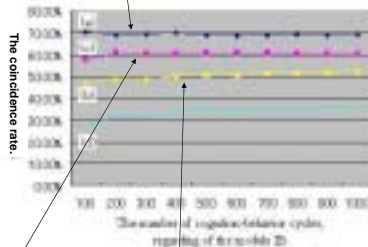
The experiments No.3

Rs   **Rc**

The Exp.3 is conducted in an environment where the other robot Rc is controlled completely via cables from the self-robot Rs to imitate the behaviour.

The results

Imitation of mirror image.

The coincidence rate.

Robots are imitating each other.

By controlled robot.

The conclusion

- **The robot controlled via cables is an extended part of the self-robot.**

- **Specifically, the mirror image of the self-robot has a higher imitation behavior coincidence rate than the controlled robot, which is an extended part of the self-robot, and thus the mirror image can be said to be a part of the extended body.**

- **Therefore, I conclude that the mirror image of the self is felt as a special part of the extended body.**