Joint Attention Realized in a Robot with Intentional Agency

ECCS'10 LISBON
14/SEPTEMBER/2010

Joint Attention Realized in a Robot with Intentional Agency

S. Nagataki, M. Shibata, T. Konno,
T. Hashimoto, and H. Hattori

Background and Purpose

- Human beings have a distinctive ability to understand what is going on in others' minds.
- This ability, usually called "folk psychology," comprises a system of understanding others' intentions of varied levels.

- Humans (both we and our ancestors) presumably depend on the common basic mechanism for communication underlying the linguistic interaction.
- We deploy what might be called "primitive" parts of folk psychology for the purpose of communication.
- The aim of our presentation is to show the basic functional structure of primitive folk psychology as a pre-linguistic **mind-reading system**.
- As an example of mind-reading function, we have selected "**joint attention**".

- In order to attain this task, we adopt a "**constructive approach**" to human communication.

- This approach attempts
  (1) to construct a computational model of the phenomena in question,
  (2) to implement it into artificial systems, and then,
  (3) to grasp the mechanism on the part of the human communication system by examining the behaviors of those systems.

An experiment of human-robot Interaction

- Joint attention is an action of communicative eye gaze, which is defined operationally as "looking where someone else is looking."(Butterworth, 1991).

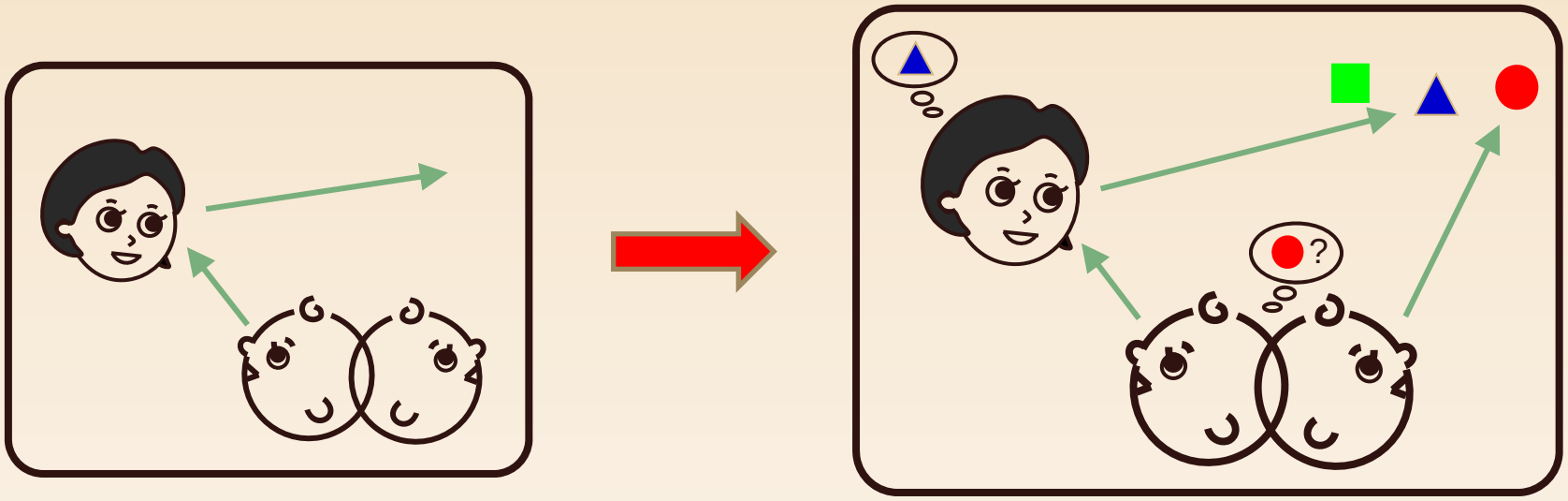
- This action begins as a reflexive action.
- Then infants come to understand and share others' attentions in terms of their intentions .

(Tomasello, 1995; Emery, 2000).

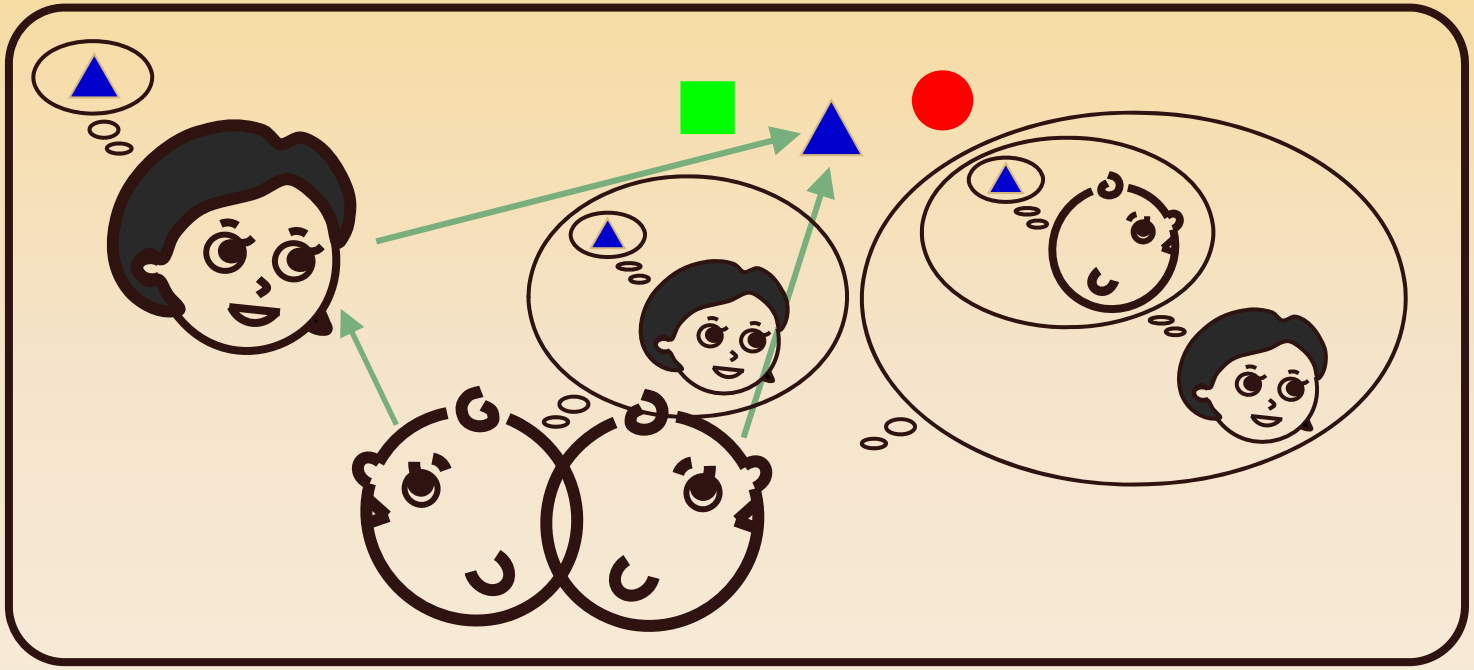
Reflexive Gaze Following → Joint Attention

How does joint attention arise from reflexive action ?

Infants need to become intentional agents in order to understand others as having intentions.
(cf. Tomasello, 2000)



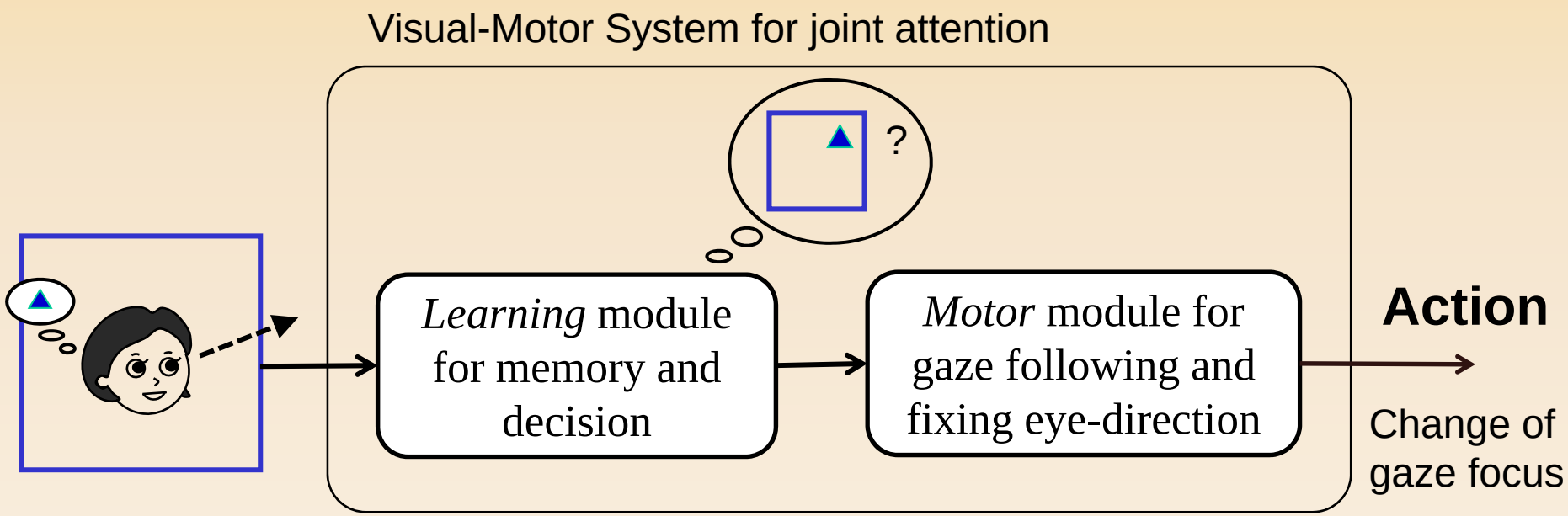
Nested Structure of Intentional Actions: in the case of joint attention



He knows that she understands that he is focusing on the object she is now looking at.

Computational Model of Joint Attention

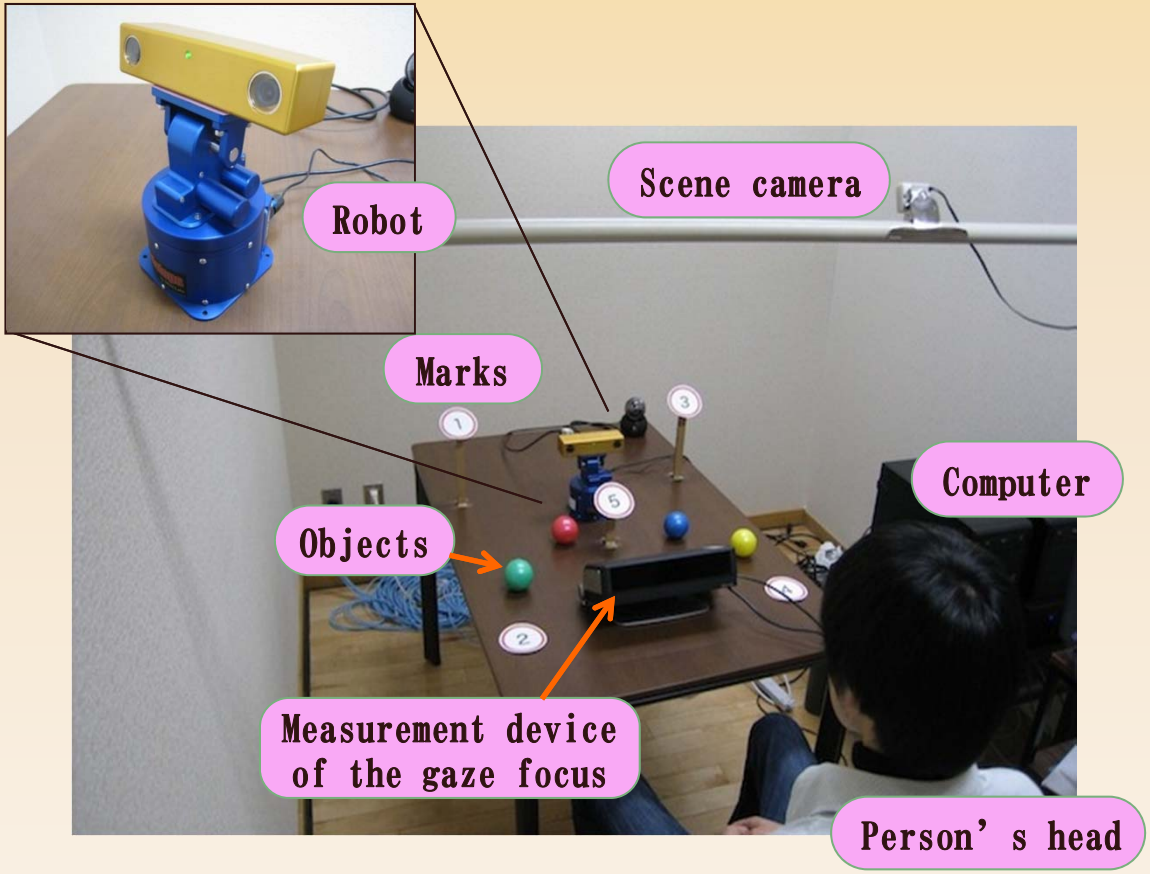
- Visual-Motor System consists of two modules.



Functions of the motor module(1~2) & Functions of the learning module(3~4)

1. Gaze following : The system turns the visual field to the caregiver's eye-direction.
2. Fixing its eye-direction: It turns its visual field to the target objects based on its decisions.
3. Associative memory: It remembers connections between directions of the caregiver's gaze and objects she gazes at.
4. Deciding its target object: It selects the target object from its memory of the connections.

Our Robot for Human-Robot Visual Interaction



- A person sits in front of the robot, and looks at the four balls.
- Positions of the balls are fixed.
- Resolution of eye-direction is 30 [deg].
- The robot gazes at the person and objects alternately.

Three Types of Robots

- **Type A:** Gaze following Model-I

  Robot turns his eyes (his camera) to follow person's eye-direction, then the robot gazes at the ball nearest to him.

- **Type B:** Gaze following Model-II

  Robot turns his eyes to follow person's eye-direction, then the robot gazes at the ball the person gazes at by using the measurement device of the person's gaze focus.

- **Type C:** Intentional Agent Model

  Robot determines a target ball **on the base of his memory** of connections between directions of person's gaze and locations of balls, and then robot turns his eyes to the target ball and gazes at it.

Important Characteristics of Type C: Immature but Intentional Joint Attention

1. The type C robot gazes at a remembered ball, therefore it does not follow the person's eye-directions by mere reflection.
2. Its association mechanism might cause a disagreement between the person's intended object and the robot's one.
3. The robot quickly turns his eyes to the object because he can determine the target directly from his memory.

Two Main Questions in Our Experiment

1. Do humans notice differences among three robots?
2. To what extent do humans think three robots' behaviors are close to human intentional actions?

- Participants were 9 (students and researchers).
- Participants interact serially with each robot for three minutes.

Experiment and Results

Questionnaire:

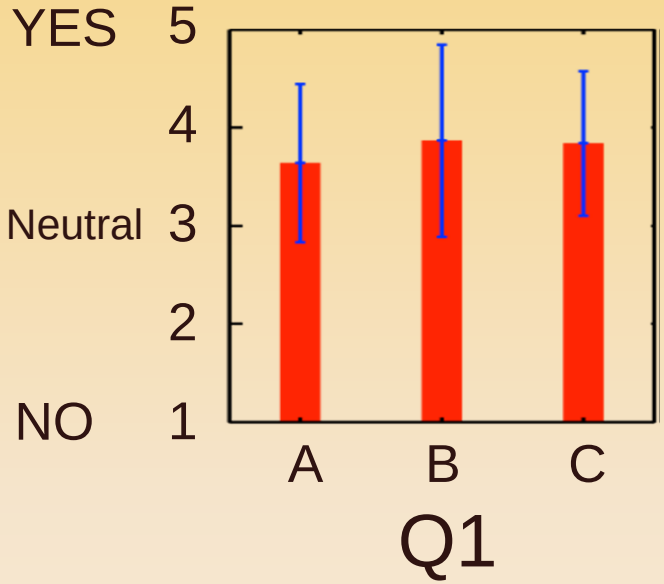
Q1. Could you make the robots identify your gazed object?

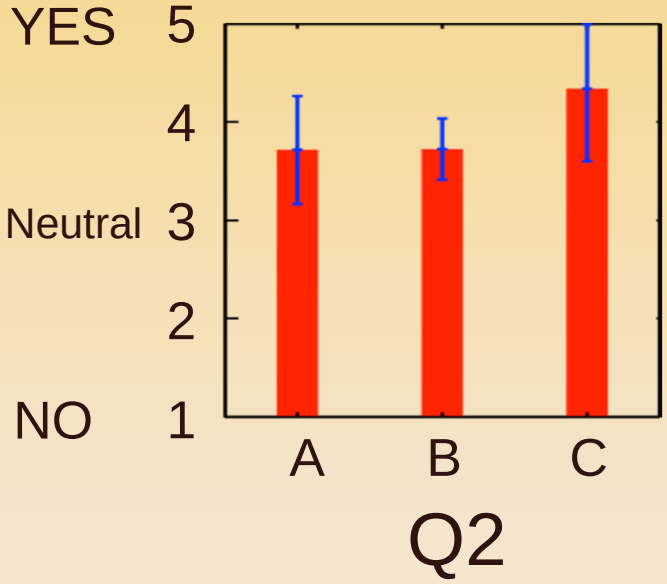
Q2. Could you identify the robots' gazed object?

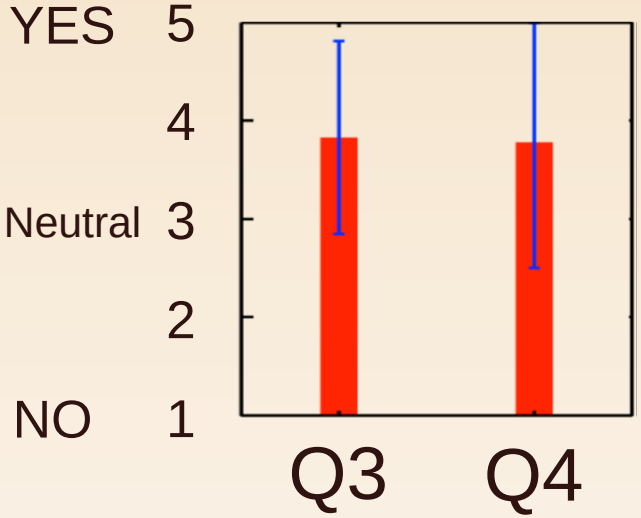
Q3. Did you notice some differences among three robots?

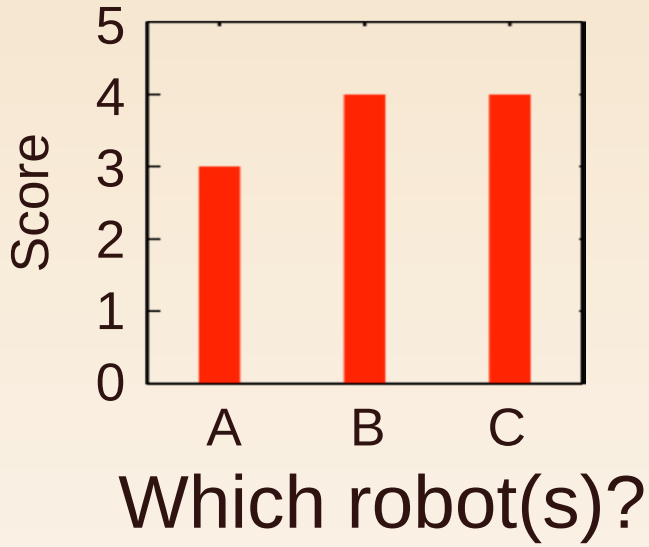
Q4. Did you feel as if the robots intentionally determined to gaze at some objects?
If so, which robot(s)?

Results


YES
Neutral
NO
5
4
3
2
1
A
B
C
Q1


YES
Neutral
NO
5
4
3
2
1
A
B
C
Q2


YES
Neutral
NO
5
4
3
2
1
Q3
Q4


Score
5
4
3
2
1
0
A
B
C
Which robot(s)?

The Role of Internal Mechanism of type C Robot not Fully Clarified

- In order to estimate to what degree Type C Robot attained genuine joint attention, we must see its behaviors in radically different situations from ours.
- But at this time, we will show a conceptual sketch for making robots realizing joint attention as an intentional action, rather than report the results of treating those situations.

Joint Attention with Understanding of Others' Intentions

- Our **type C** robot does not intend to understand humans' intentions to see something.
- Our next robot, **type D**, must be able to infer humans' intended objects (not without any failure), using the inference mechanism, in order to make their intended objects his desired objects.
- The inference mechanism allows him to reach those objects from his past experiences.

Joint Attention with Shared Intentions

- **Type D** robot doesn't care whether his inference is right or not.
- Because he doesn't have a desire to confirm that humans understand his intention to infer their intended objects.
- Our next, **type E**, robot must be able to correct his inference and change his desired object, by some assessment mechanism, when his inference was rejected by humans.
- Because this robot wants to confirm that humans understand that he intends to share their intentions, it could be said that he realizes a kind of nested structure of intentions and understandings.

Joint Attention as Having Multi-Realizable Multi Functions

- We don't think an evolution from **Type C** to **E** is the only line which realizes human joint attention, because the joint attention has some other functions than we have explored.
- Joint attention might require "discovering-salient-features and avoiding-danger" mechanism rather than our "remembering-familiar-objects and being-devoted-to-them" one.
- Even in our **type C–E**, other additional functional elements are to be found through the constructive approach in our future research.

References :

- G. Butterworth and N. Jarrett. What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. British Journal of Developmental Psychology, 9:55–72, 1991.
- N. Emery. The eyes have it: the neuroethology, function and evolution of social gaze. Neuroscience and Biobehavioral Reviews, 24(6):581–604, 2000.
- T. Konno and T. Hashimoto. Developmental construction of intentional agency in communicative eye gaze. Proceedings of the International Conference on Development and Learning, ICDL '06, CD– ROM, May 31 - Jun.3 2006.
- M. Tomasello. Joint attention as social cognition. In C. Moore and P. Dunham, editors, Joint Attention: Its Origins and Role in Development, pages 103–130. Lawrence Erlbaum, 1995.
- M. Tomasello. The cultural origins of human cognition. Harvard University Press, Cambridge, 2000.