Peter Gärdenfors

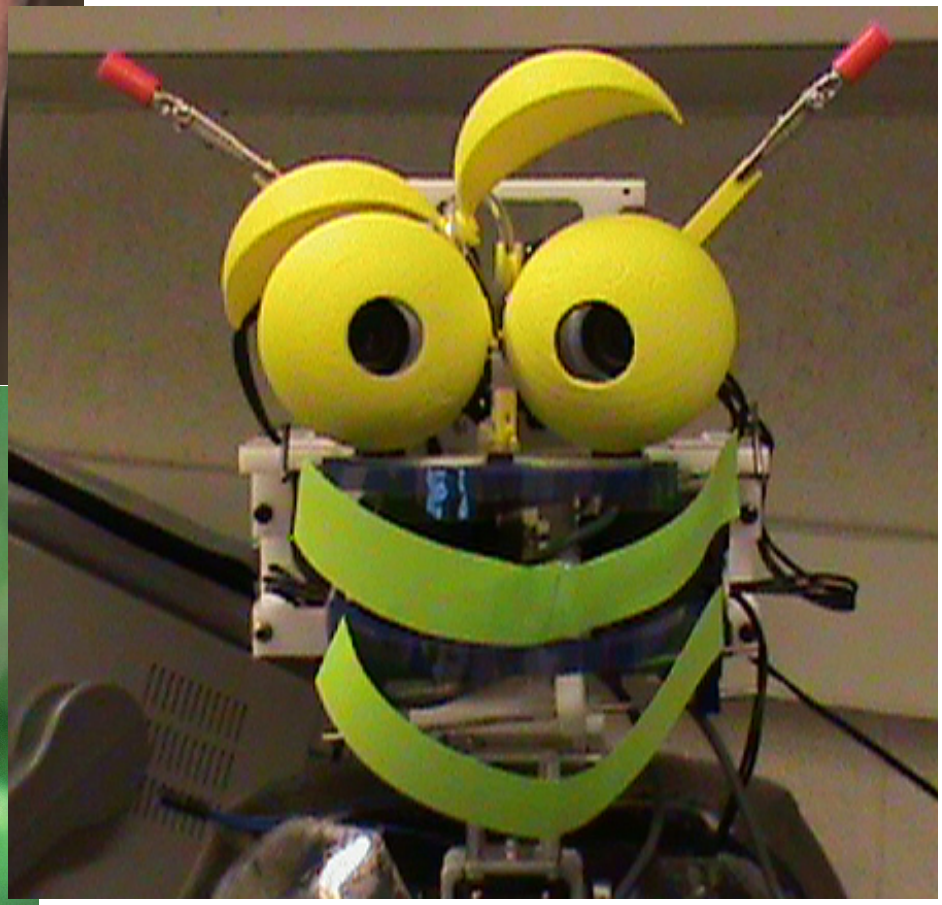
Lund University
Cognitive Science

Mind-reading robots

Mind-reading

- Nothing magical
- The ability to understand what others feel, want and think
- Philosophers: ”Theory of mind”
- Psychologists: ”Intersubjectivity”
- How can we make a robot that understands what humans feel, want and think?

Three kinds of mind-readers

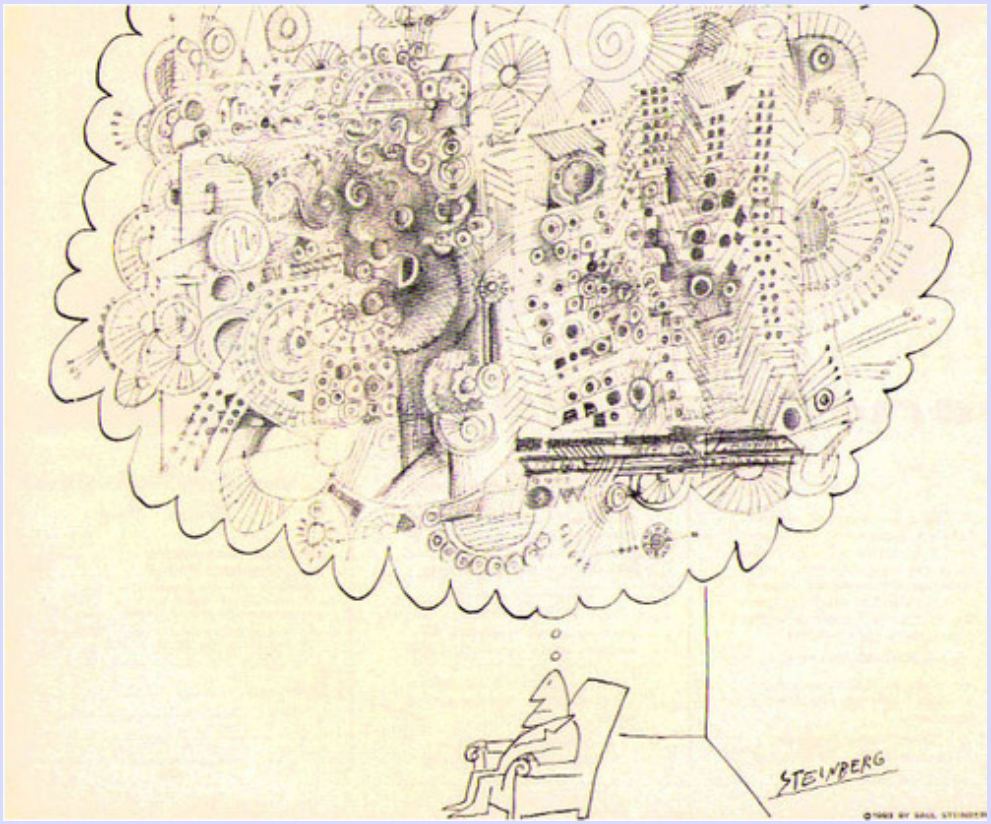


Humphrey: *A History of the Mind*, 1993

- *Sensations* are the *immediate* sensory impressions.
- *Perceptions* are *interpreted* sensory impressions.

From sensations to perceptions

- *Sensations* are the *immediate* sensory impressions.
- *Perceptions* are *interpreted* sensory impressions.
- *Imaginations* are experiences that are not directly governed by sensory impressions.
- Imaginations and perceptions are *representations*

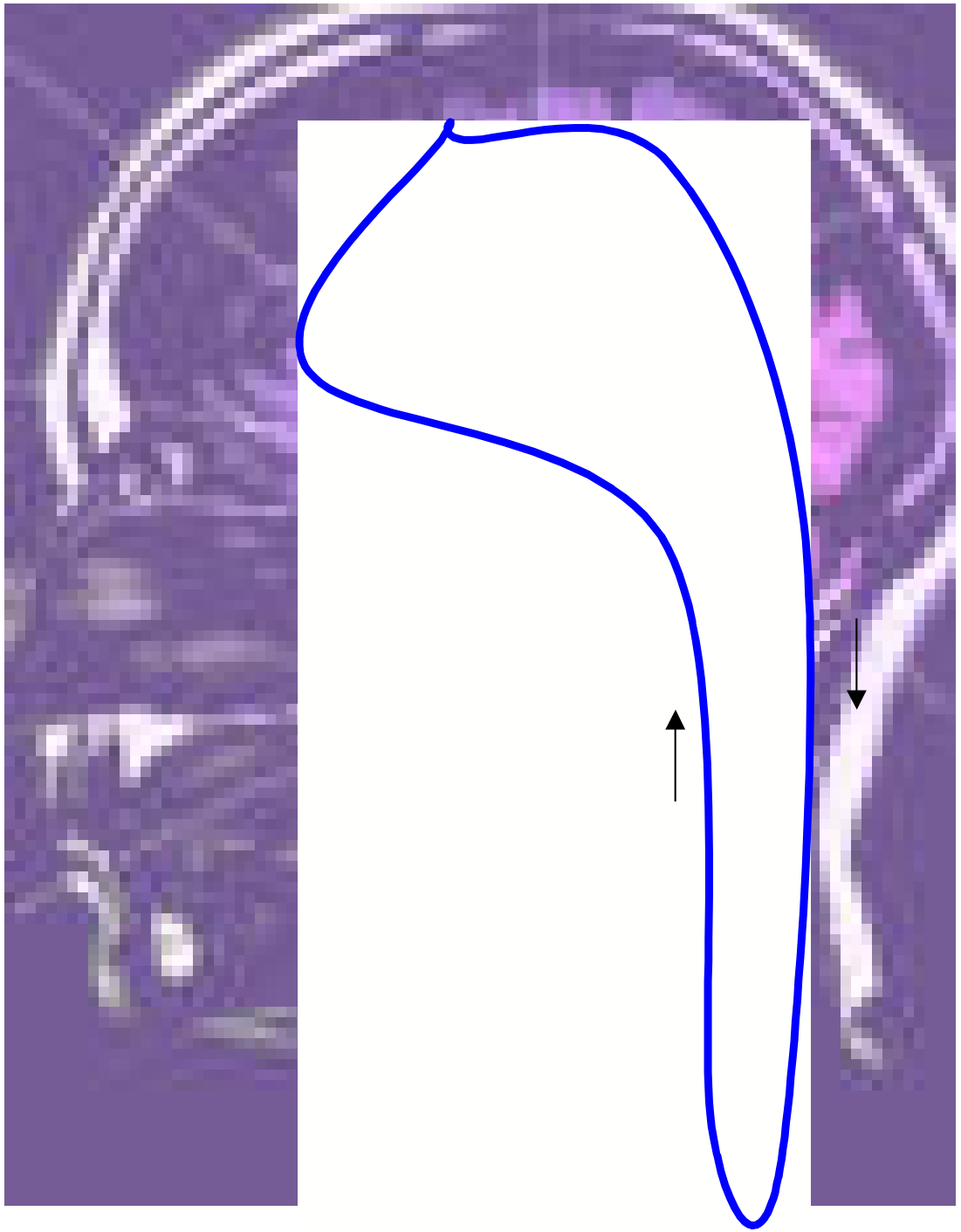


The mechanism of representations

Hypothesis:

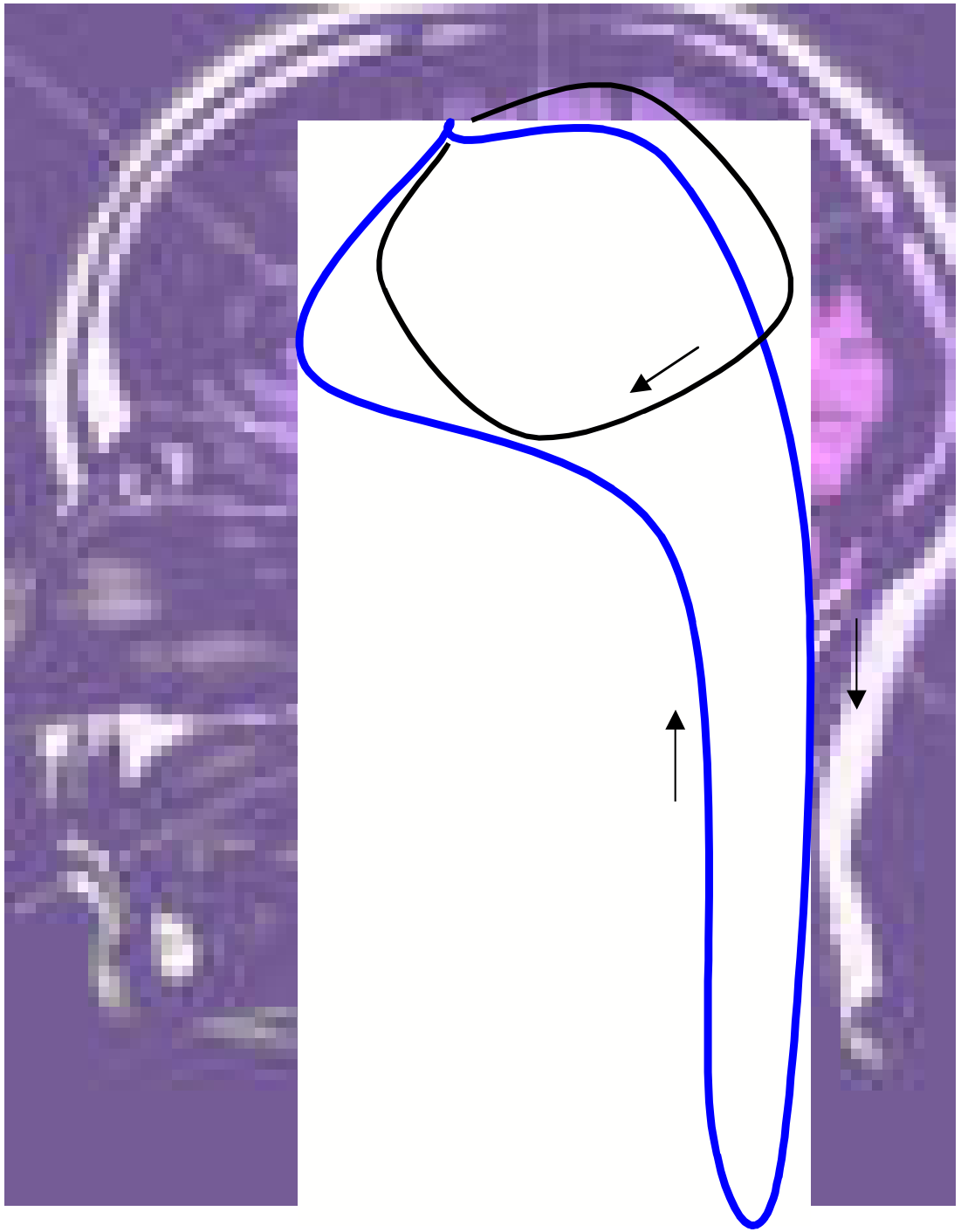
Perceptions and imaginations are created by *emulators*

The sensori-motor loop



Emulators running in parallel with the sensori-motor loop

Emulators help predicting the future



The emulators generate ”hidden variables” that explain *causal mechanisms*

- Physical causality (the emulators help us perceive the *forces* behind events)

The emulators generate ”hidden variables” that explain *causal mechanisms*

- Physical causality (the emulators help us perceive the *forces* behind events)
- Mental causality (the emulators help us perceive the *mental forces* – the emotions, beliefs, desires and intentions – that govern the behaviour of others)

Components of mind-reading

- Representing the *emotions* of others
- Representing the *attention* of others
- Representing the *intentions* of others
- Representing the *beliefs* of others
- Self-consciousness

- All (?) mammals
  Children 3 months (?)
- Primates (and others)
  Children 6-12 months
- Chimps (to some extent)
  Children 9-14 months
- Children 36-48 months
- Children from 36-48 months

Gärdenfors*: How Homo Became Sapiens*, 2003

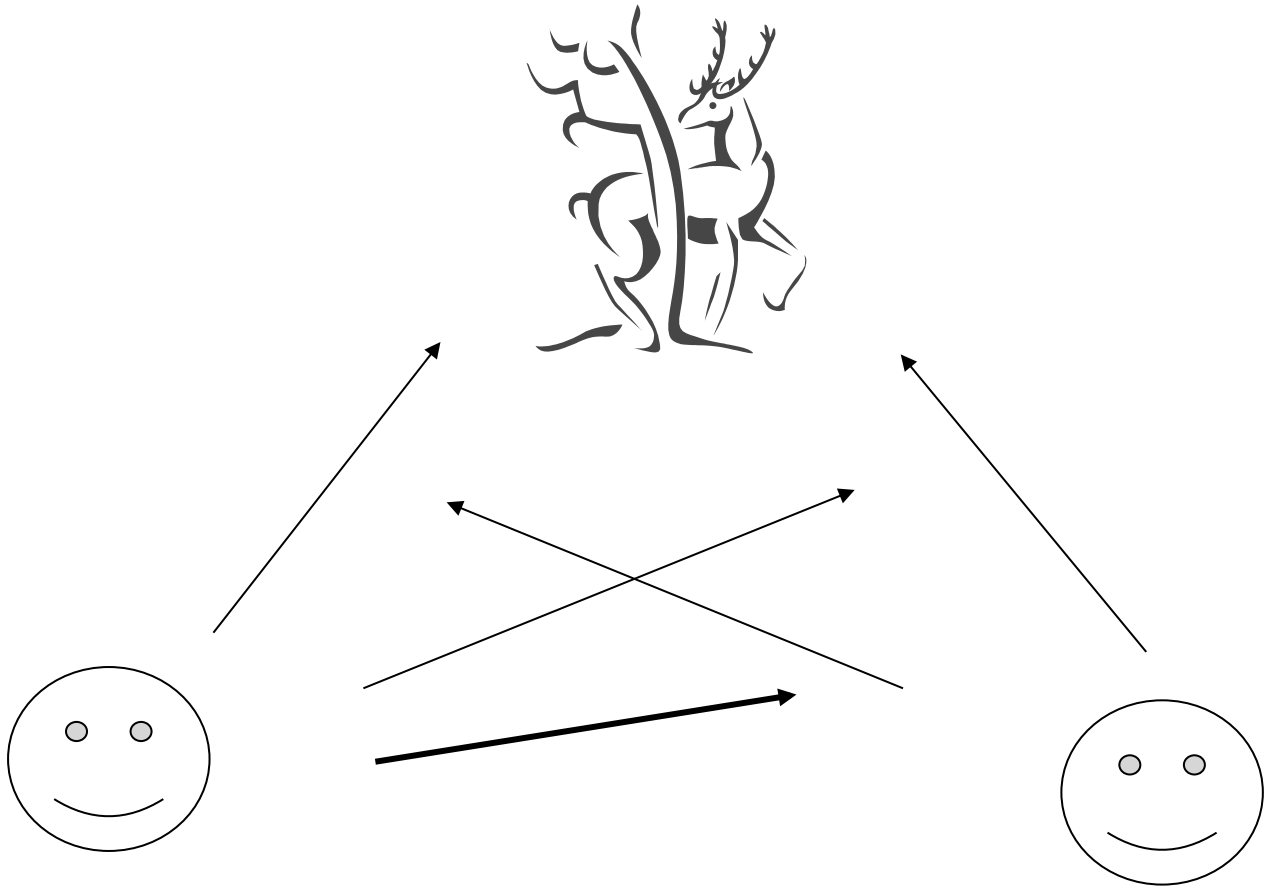
Representing the emotions of others: Empathy

- Empathy: perception of emotion in another activates the same emotion in the receiver
- Evidence for empathy in mammals and birds
- May depend on mirror neurons

Representing the attention of others

- Children at 6 months can follow the gaze of their mother if she turns her head
- At 12 months they can follow the gaze of their mother if she just moves her eyes
- At 18 months they can follow the gaze of their mother if she looks outside their field of vision (requires allocentric representation of space)
- Apes, dogs, goats, etc, can also follow gazes in an allocentric way
- Robots have problems exploiting gaze information

Joint attention is central for human collaboration and communication



Representing the intention of others

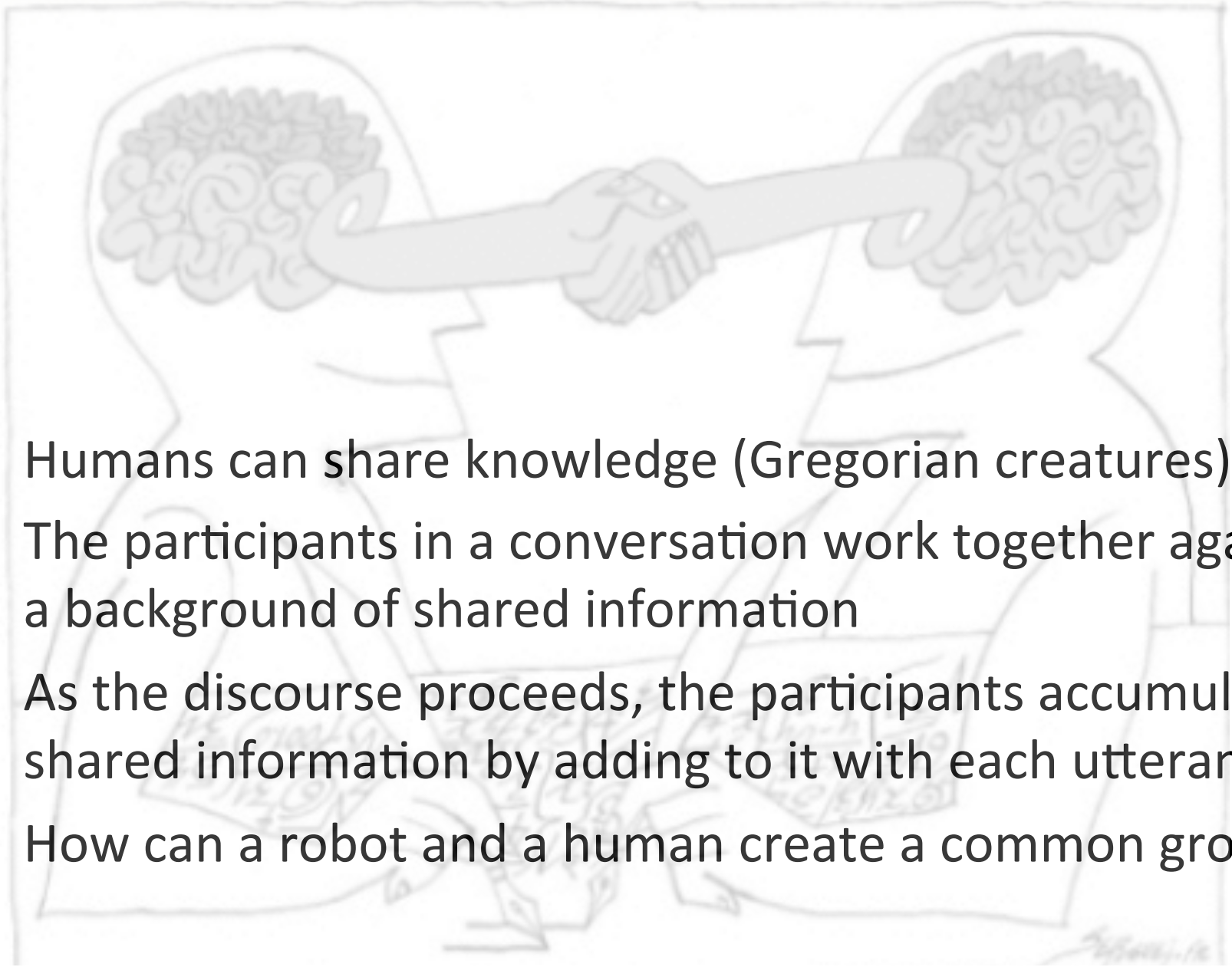
- Can non-human animals take the intentional stance (Dennett)?
- Experiments where an adult (1) deliberately avoids handing over or (2) fails to hand over a reward (a toy or food)
- Children from 9 months and chimps react differently to (1) and (2), i.e. to whether the failure was deliberate or not
- How can the robot understand the intention of a human user?

From joint attention to joint intention

Representing the beliefs of others: False belief tasks

0. Subjects are three- to five-year-old children.
1. The children are first shown a Smarties tube and then asked what they think is in it. All the children reply “Smarties” (or “sweets”).
2. When the tube is opened it is found to contain pencils.
3. Then the tube is closed.
4. The children are now asked what Bert, who has not yet seen what is in the tube, will say that it contains.
5. The three-year-olds generally answer “pencils” whereas most of the older children say “Smarties.”

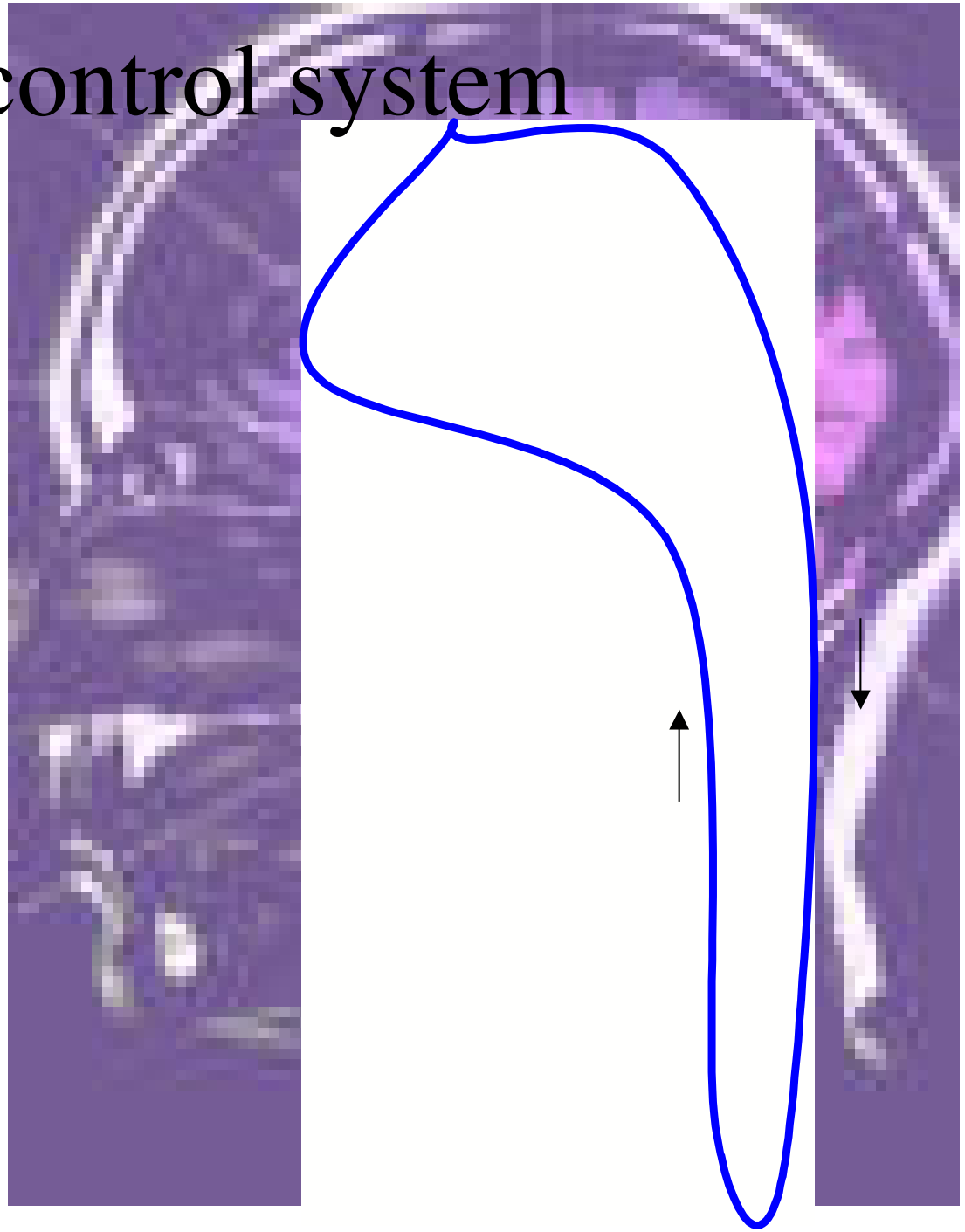
Common ground (joint knowledge)



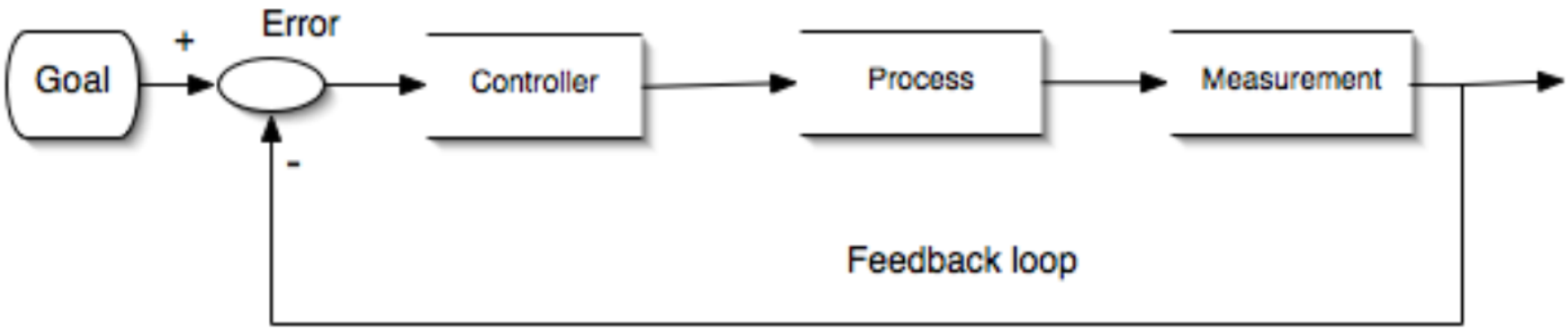
- Humans can share knowledge (Gregorian creatures)
- The participants in a conversation work together against a background of shared information
- As the discourse proceeds, the participants accumulate shared information by adding to it with each utterance
- How can a robot and a human create a common ground?

The brain as a control system

The sensori-motor loop

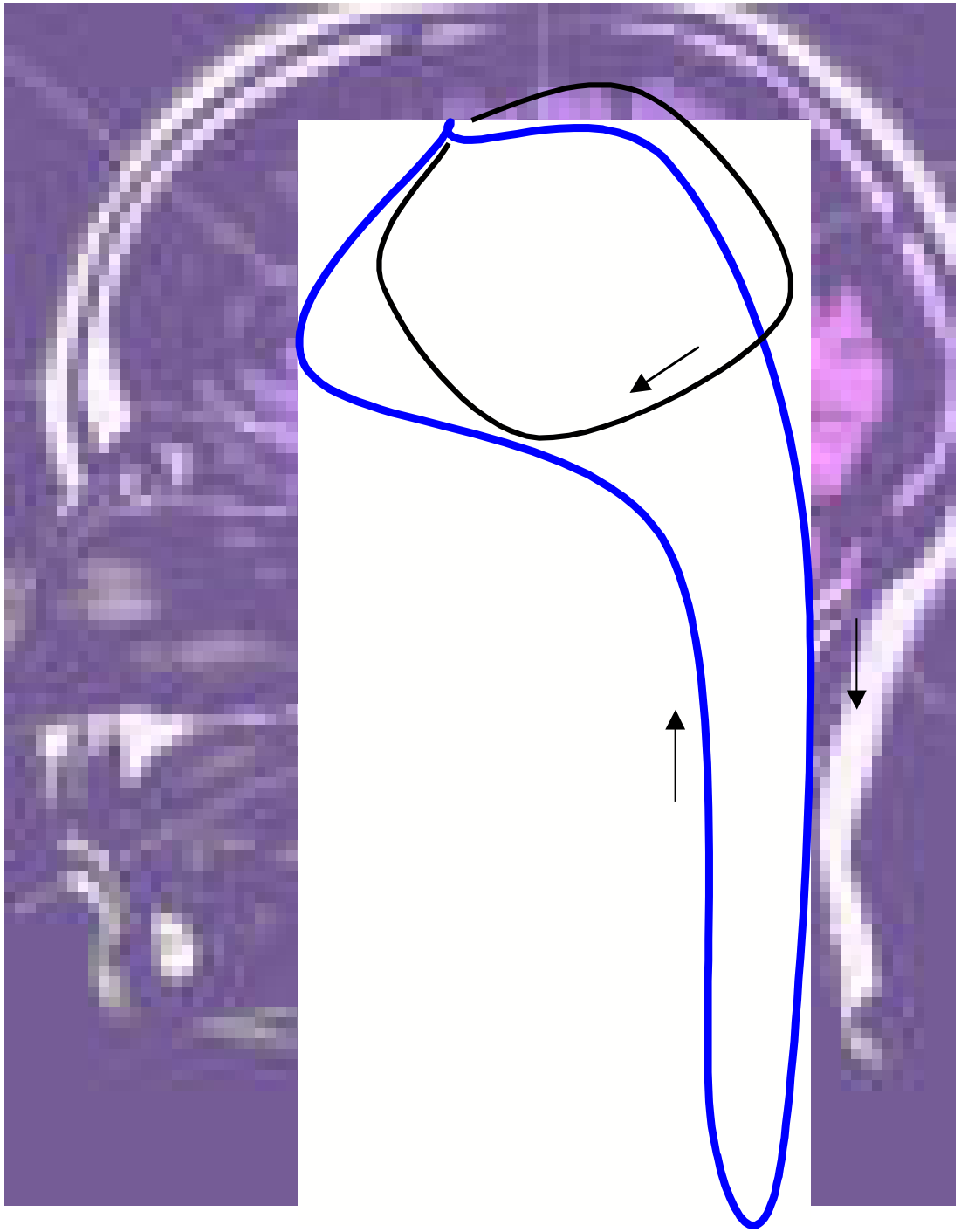


Feedback control

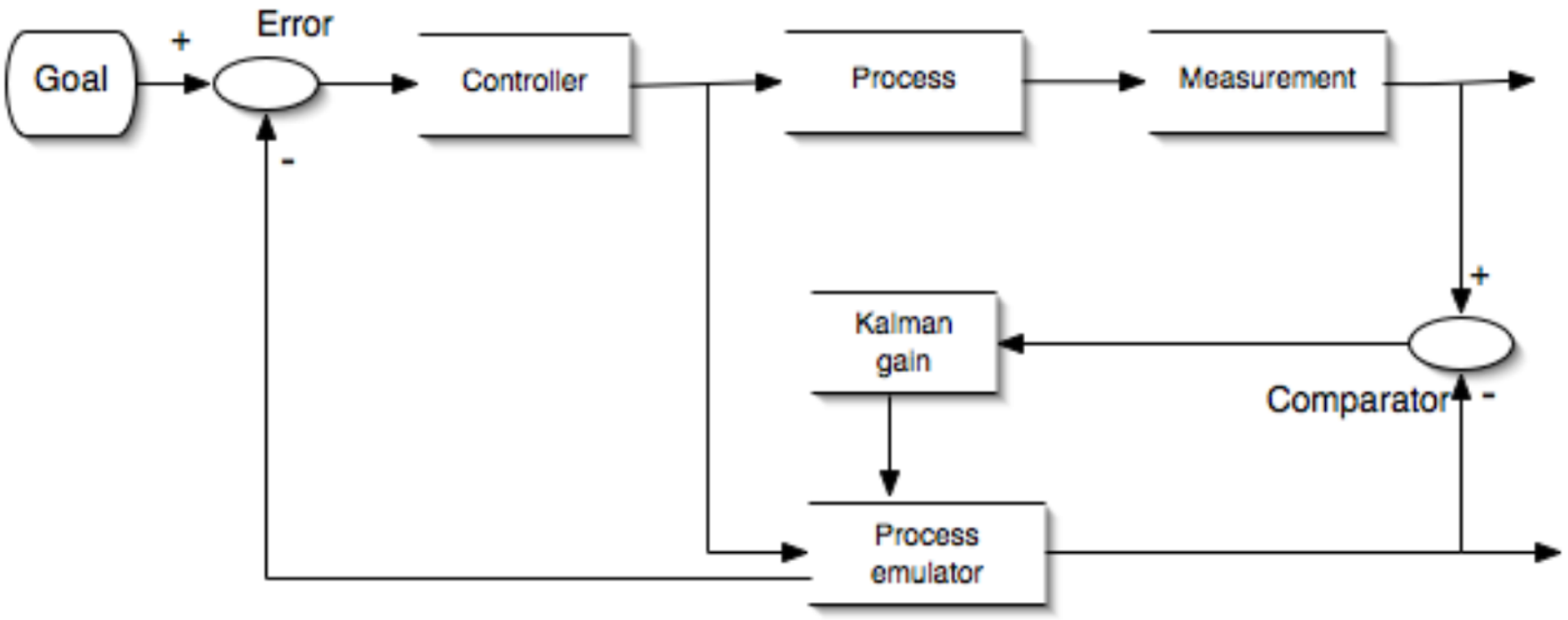


Emulators running in parallel with the sensori-motor loop

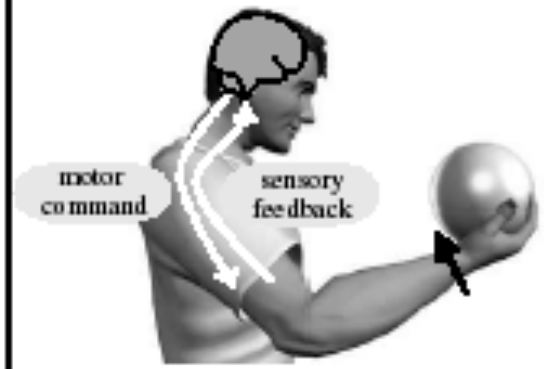
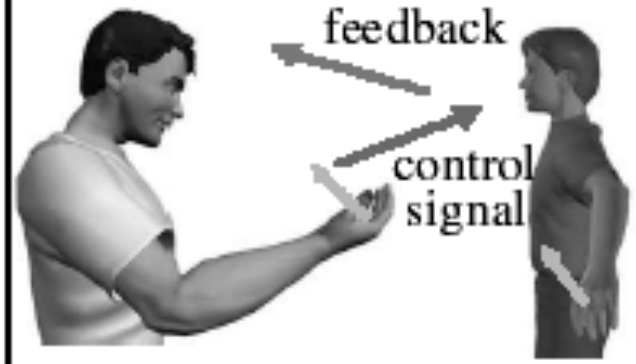
Emulators (forward models) help predicting the future



Emulator based on feed-forward model



An analogy between motor control and mind-reading

	motor control	social interaction
loop	(*a*) 	(*b*) 
control signal	motor command	communicative actions e.g. speech, gesture
consequences	change in my body's state	change in your mental state
state	configuration of my body	mental state of your mind

Wolpert, Doya and Kawato (2003)

Representing the emotions of another

Empathy doesn't need emulators (?)

It seems sufficient to *correlate* the facial and bodily emotional expressions of the other with your own emotional states.

(Gallese: Mirror neurons may provide a mechanism for this crossmodal task. Also fMRI studies by Wicker et al. 2003, Singer et al. 2004)

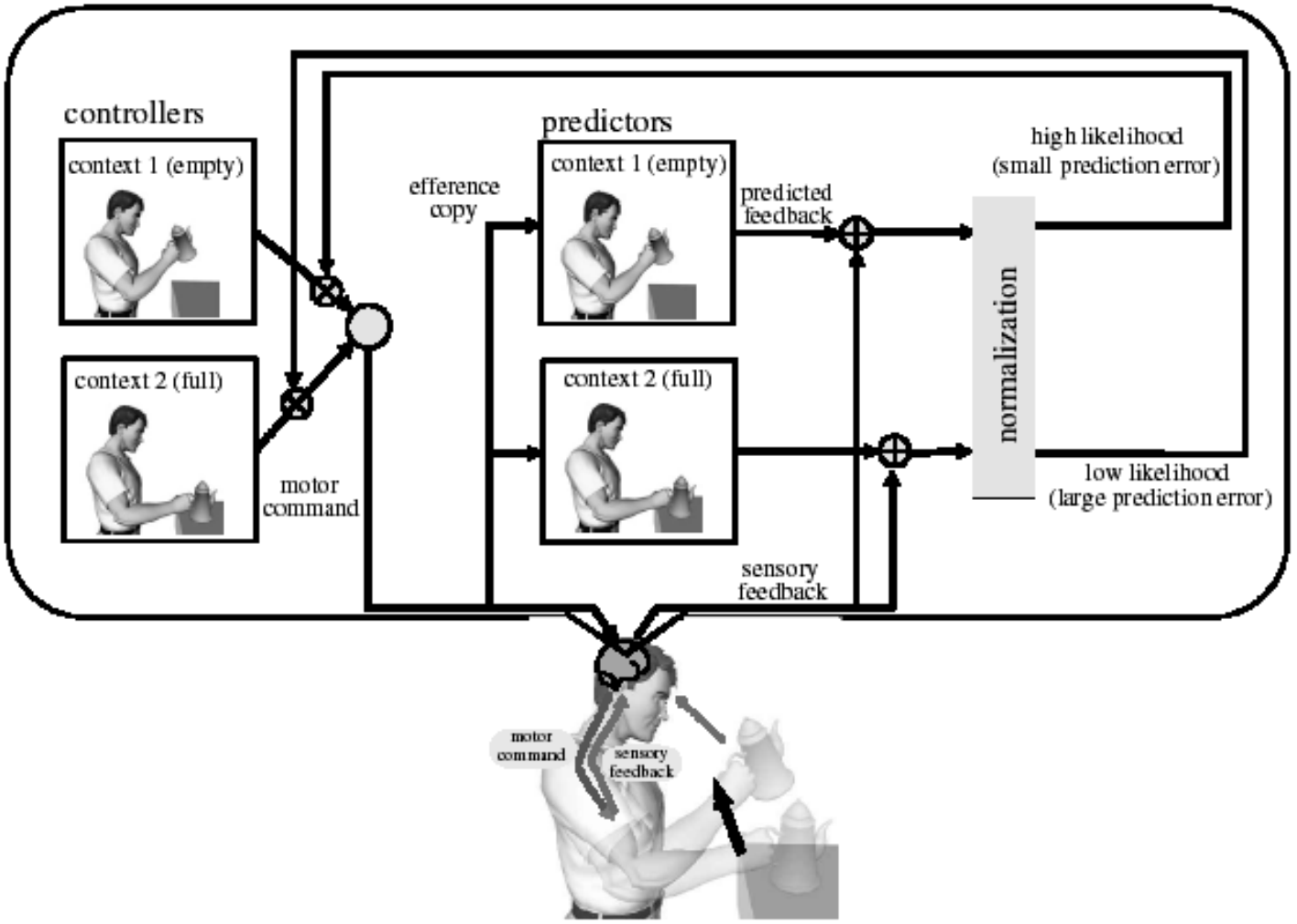
Representing the attention of another

Requires *coordinate transformations*. The direction and focus of the other's gaze must be transformed to your own gaze control. More difficult transformation if the attended object is outside your field of vision (requires allocentric representation).

How are intentions represented cognitively?

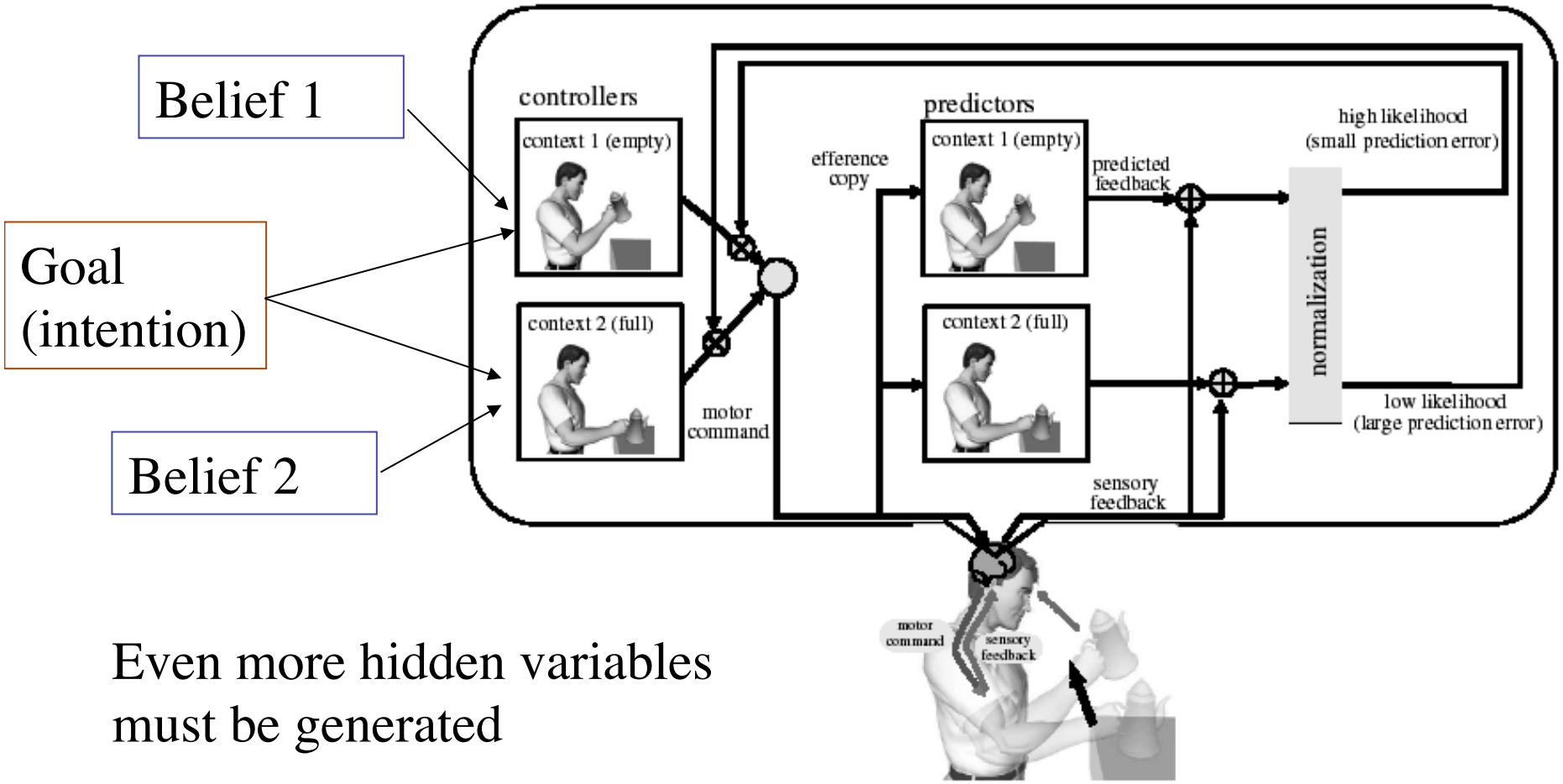
- From gaze following to mind following
- Brentano's directedness
- How can a robot read the intention of humans?
- What is the *context*?
- What is *valuable* for the human?

Bringing in context in the control loop



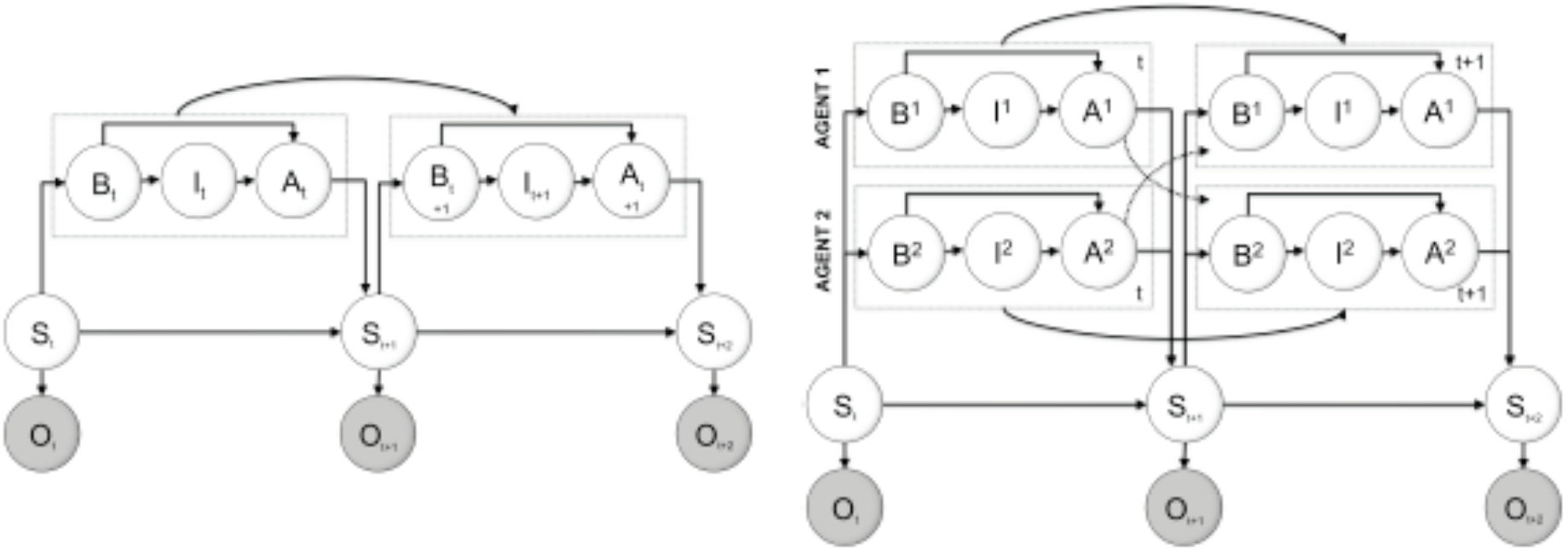
Wolpert, Doya and Kawato (2003)

Beliefs and intentions as part of the context



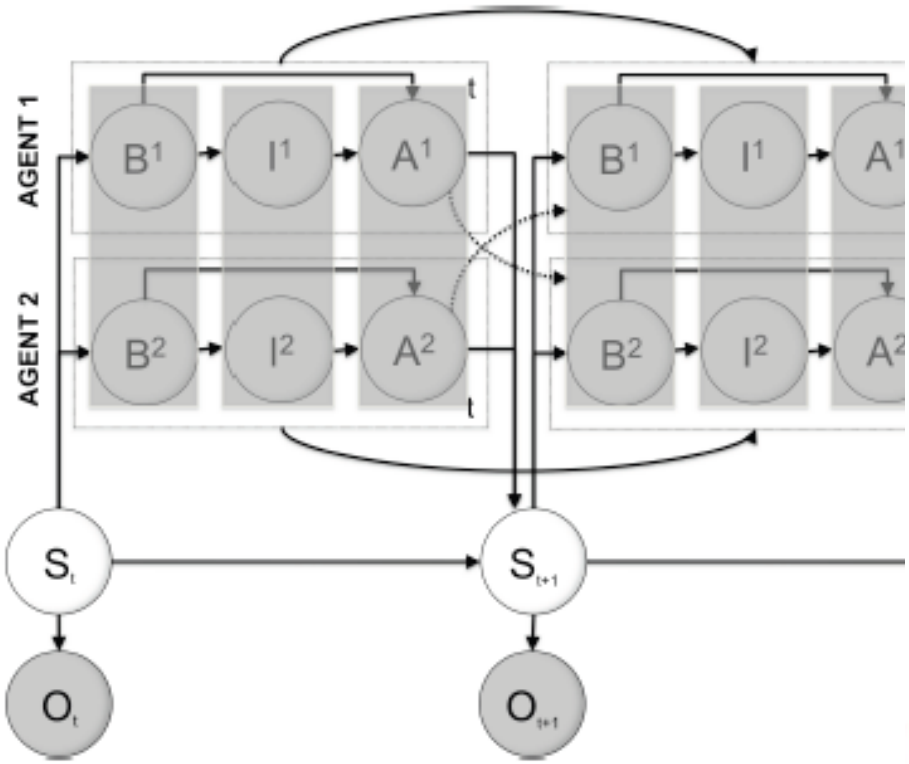
Even more hidden variables must be generated

Belief-intention-action models of one and two persons



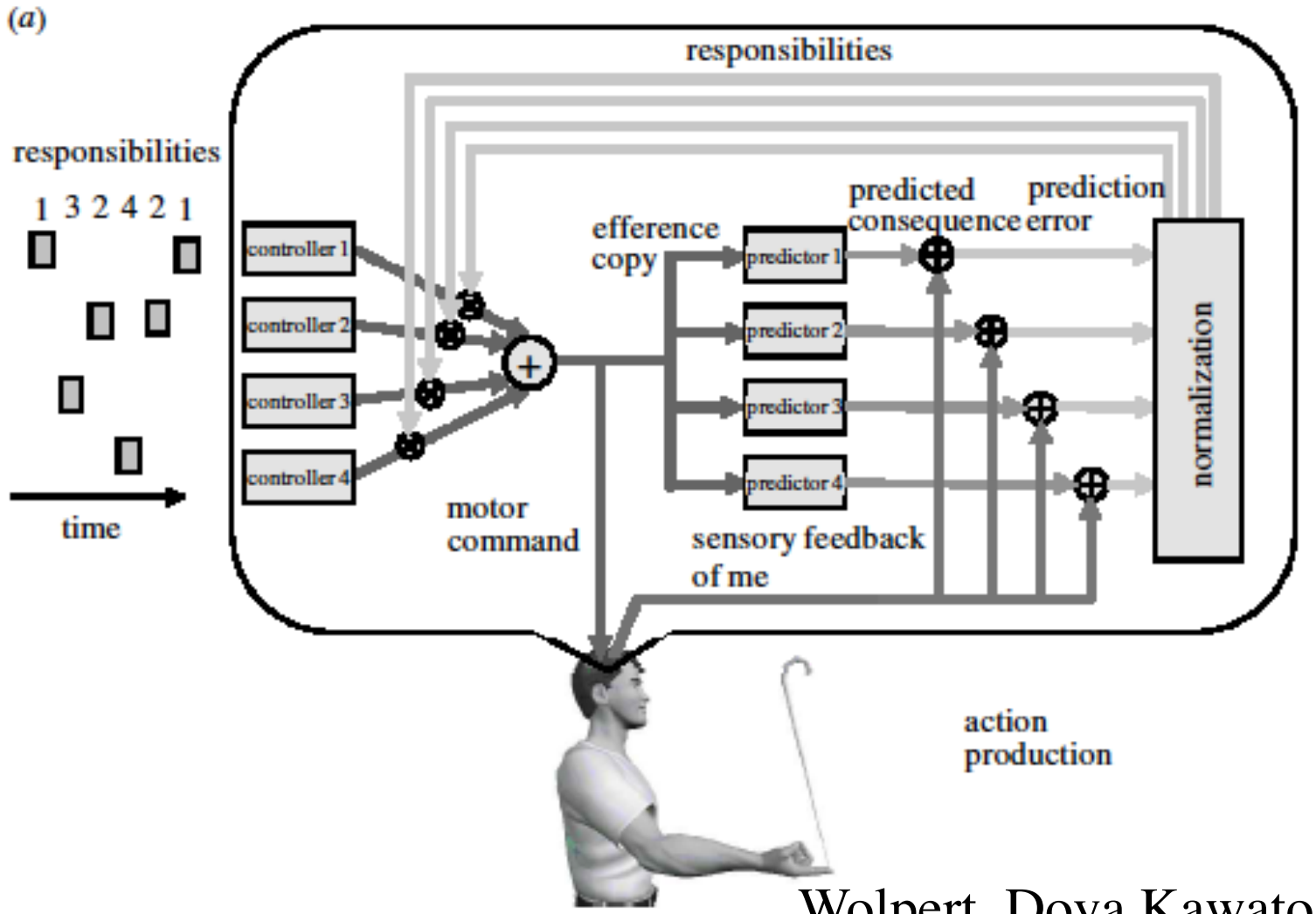
Pezzulo (2012)

Model of joint action



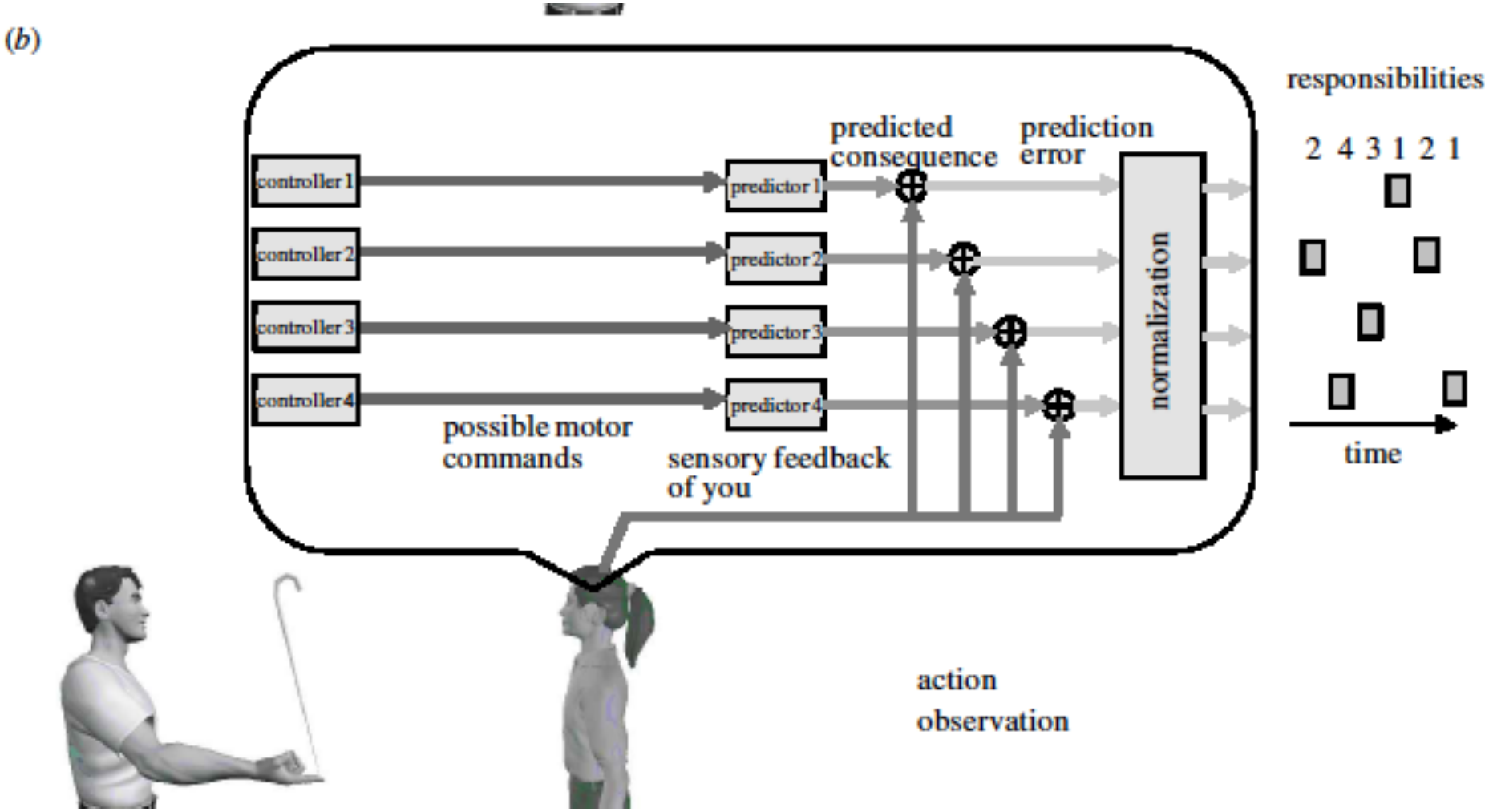
Pezzulo (2012)

Model of action control



Wolpert, Doya Kawato (2003)

Model of action percaption



Wolpert, Doya Kawato (2003)

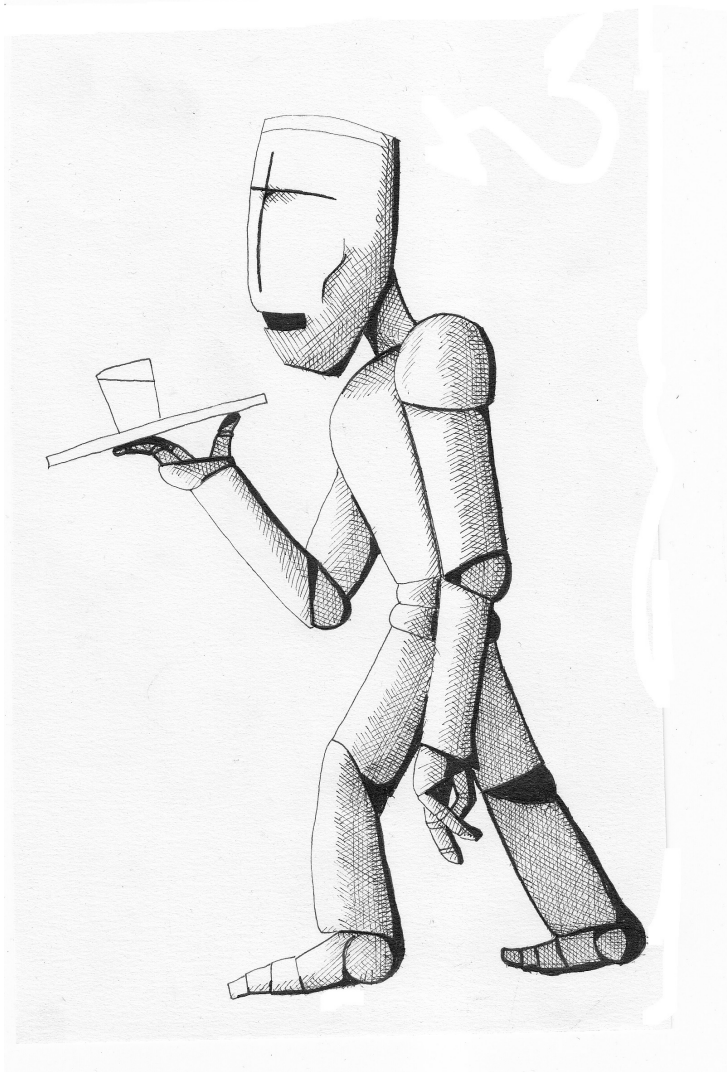
Tasks for a mind-reading robotics

(1) Develop human-robot joint attention techniques

(2) Model a flexible system for reading intentions

(3) Model joint intentions

(4) Model joint beliefs (common ground)

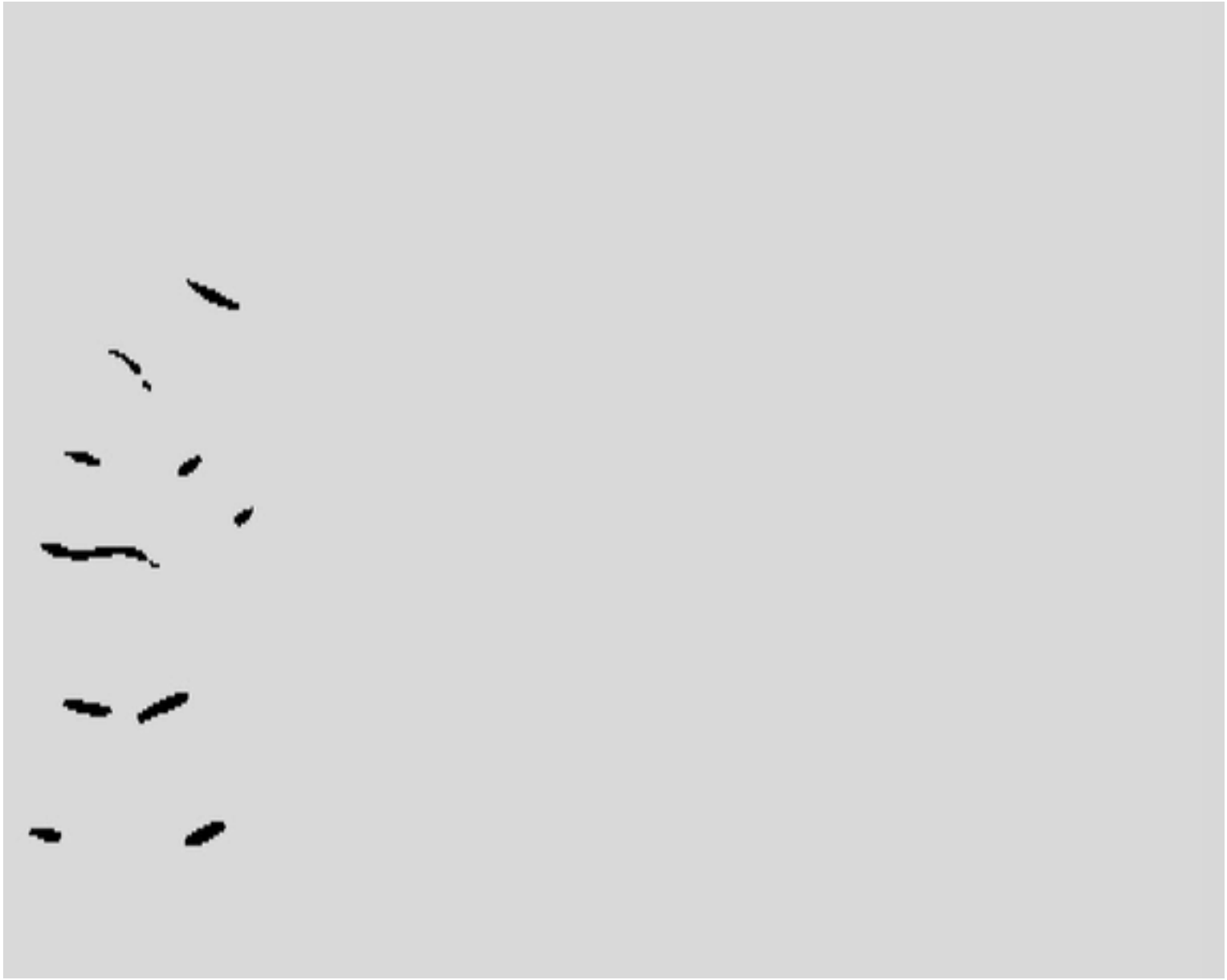


Peter Gärdenfors

Lund University
Cognitive Science

Mind-reading robots

Humans are excellent at identifying actions



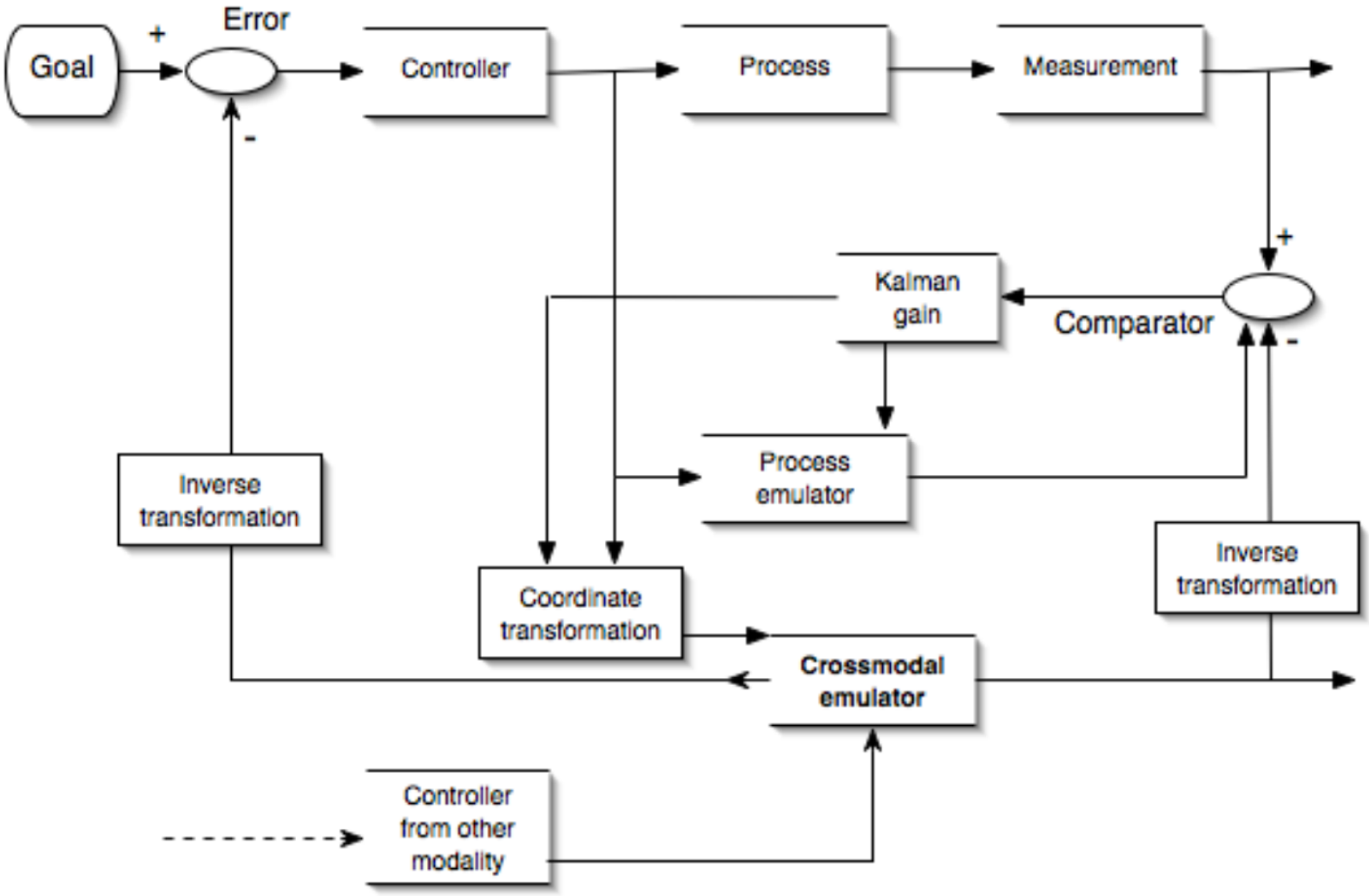
Pezzulo's analysis

	Tasks of perceptual processes	Computational mechanisms
Individual scenario	(i) estimating the state of the observed system (i.e., hidden environmental variables)	Kalman filtering, particle filtering, Luenberger observer
	Tasks of the observer	**Neuro-cognitive processes**
Interaction scenario, non communicative aspects	(i) mindreading (estimating cognitive variables of another agent)	motor resonance, action simulation, emulation, action and intention understanding, inverse planning
Interaction scenario, communicative aspects	(i) mindreading for recognizing communicative intentions	the same mechanisms as above
Joint action scenario	(iii) formation of shared representations (SRs)	behavioral entrainment, mutual emulation, joint attention; the explicit goal of forming SRs
Linguistic scenario	(i) mindreading for recognizing communicative intentions in speech acts, (iii) formation of shared communicative context	language understanding as mental simulation, interactive alignment, mechanisms for maintaining reference

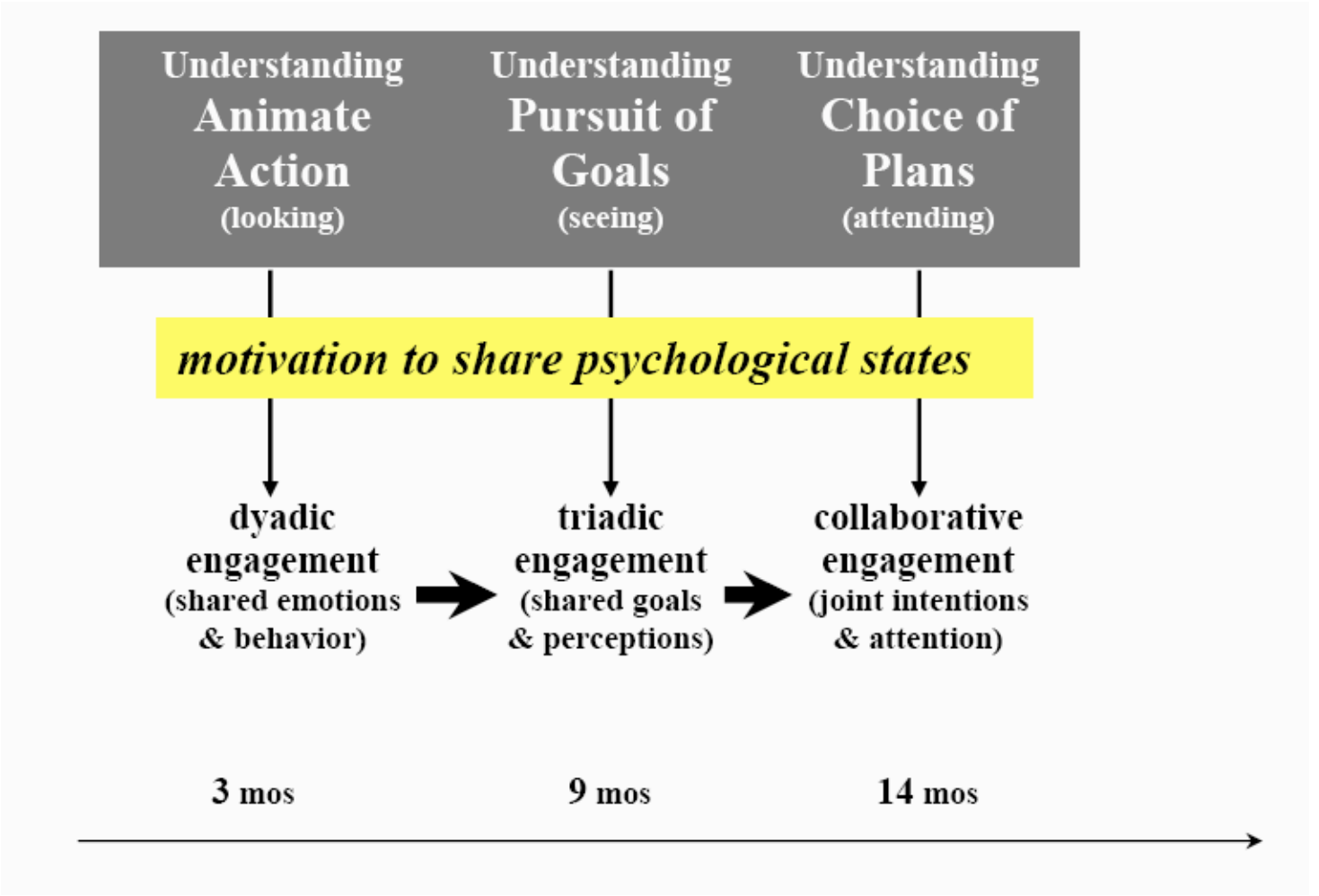
TABLE II
FORMAL SIMILARITY OF PROBLEMS ACROSS INDIVIDUAL, INTERACTIVE, JOINT ACTION AND LINGUISTIC SCENARIOS: ACTOR

	Tasks of action processes	Computational mechanisms
Individual scenario	(ii) achieving goals relative to the environment (changing environmental dynamics)	inverse modeling, (chains of) forward models, MAP, policy iteration
	Tasks of the actor	**Neuro-cognitive processes**
Interaction scenario, non communicative aspects	(ii) achieving goals relative to another's actions (e.g., helping, hindering, imitating)	action planning and execution; prediction and prospection mechanisms (for understanding action effects)
Interaction scenario, communicative aspects	(ii) achieving goals relative to another's internal variables (changing mental states of another agent);	planning and execution of communicative goals; recipient design
Joint action scenario	(iii) joint action control (takes joint goal into consideration, uses shared representations)	planning and execution of joint goals and of signaling actions; creation of affordances for others
Linguistic scenario	(ii) using language to achieve goals relative to another's internal variables, (iii) common ground formation	planning speech acts

An amodal emulator



Joint intentions



Tomasello et al. BBS 2005

The brain as a control system

- Control of self: Meta-cognition
- Control of others: Intersubjectivity (”theory of mind”)
- Have these systems evolved in parallell?

The brain adds lines