

“How can we Relate to Things that  
Relate to us —Expression of Robots”

Kiyotaka Naoe  
(Tohoku University, Japan)

# Introduction

Human – Human – Interaction

Human – Robot • AI – Interaction

Key : Expression

facial expression

expression in eyes

voice

bodily movement

handwritten characters

...



ICD-LAB website <https://www.icd.cs.tut.ac.jp/portfolio/muu/>

# Writing and Drawing Machine

TOROBO-KUN or AxiDraw V3  
(Robot) write “handwritten”  
characters



AxiDraw V3 (<https://bouncy.news/1838>)



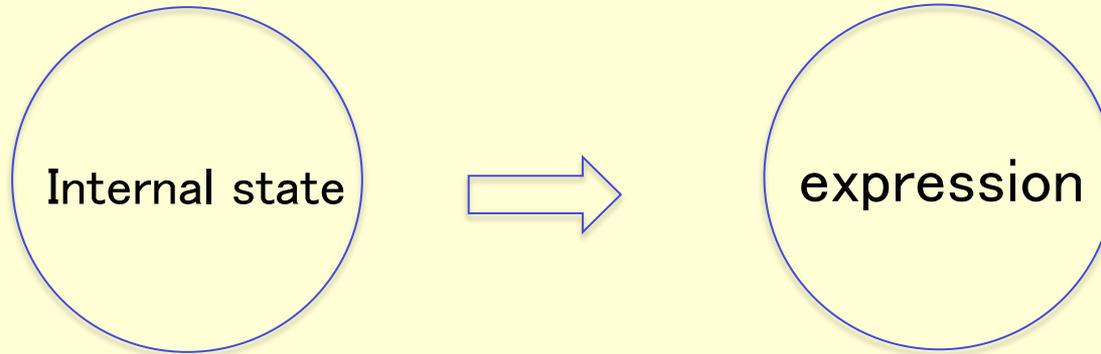
graphology

Both express  
personal character  
of the writer.



calligraphy

## original meaning of “expression”



Internal state(feeling, emotion etc.)

ex + pressare (aus- drücken)  
Introjection

A. Phenomena of “expression” by robots

B. Skepsis to “internal state”

How to think of “things that relate to us”?

How to think of “expression” without internal states?

{ intention  
feeling, emotion  
personhood, otherness

\* Limits of this presentation

- we focus on expression
- we do not refer to “the hard problem of consciousness” and the problem of personhood as such.

# 1. traditional theory of expression

## 1-1 Philosophy of expression and representation

- Neoplatonism, Leibniz
- 19-20C. Th. Lipps, Dilthey, Croce, Cassirer, Klages
- Kitaro Nishida “expressive character”, Kiyoshi Miki

traditional framework (the case of Dilthey)

(a) experience

(b) expression (expression of experiences)

(c) understanding (understanding of expressions)

Understanding is that of experience (internal state)  
expressed in an expression.

“An **expression** of lived **experience** can contain more of the nexus of psychic life than any introspection can catch sight of. It draws from depths not illuminated by consciousness. But at the same time, it is characteristic of an expression of lived experience that its relation to the spiritual or human content expressed in it can only be made available to **understanding** within limits. Such expressions are not to be judged as true or false but as truthful or untruthful.”

(W. Dilthey, The formation of the historical world in the human sciences, 1910/SW.III, 227)

## ⇒ Classification of issues

A to B	Expression of B	A' s understanding of B	
		facial expression bodily movement	Handwriting, drawing
human to human	( 1 )	(3-1)	(3-2)
human to robot	( 2 )	(4-1)	(4-2)
robot to human	( 1 )	(5-1)	(5-2)
robot to robot	( 2 )	(6-1)	(6-2)

- Examination of “expression” itself···(1)(2)
- Examination of “understanding” of other’ s expression  
···(3)~(6)
- \*direct expression···expression of face, eye, bodily movement  
expression in things···handwriting character, drawing



\*difference to illustration

human – robot : reciprocal

human – illustration: unidirectional

Issues out of consideration

- (5) understanding human expression by a robot
  - issue of facial recognition, emotion–recognition etc.
  - simulation, deep learning
- (6) another or a derivative issue

## 1-2. understanding of robot' s intention, emotion

### Theories of other minds recognition

Theory theory of mind	Analogy inference theory (Argument from analogy)
Simulation theory of empathy	Empathy theory

- **Theory-theory of mind:** individuals hold a basic or 'naïve' theory of psychology ("folk psychology") to infer the mental states of others, such as their beliefs, desires or emotions. This information is used to understand the intentions behind that person's actions or predict future behavior.
- **Simulation theory of empathy:** humans anticipate and make sense of the behavior of others by activating mental processes that, if carried into action, would produce similar behavior. This includes intentional behavior as well as the expression of emotions.

1) Analogy of Intention: What is the intention of a robot?

[issue(2)&(4)]

intention of a robot

= intention which the designer intended to build-in the robot

= intention which the designer intended to imitate in the robot

= intention which the designer intended to imitate a personal intention in the robot

= intention which the designer intended to imitate a personal character in the robot ?? (\*)

= intention which the partner ascribe to the robot as the intention the designer intended to imitate a personal character in the robot ??

(\*) imitated personal intention is not single-track, but multiple-track

⇒ Personal character is as a bearer required.

⇒ But it is doubtful whether the personal character can (or needs to) be imitated in the robot.

1' ) personal character → disposition, multiple-track disposition  
[issue(2)]

(Stina Bäckström, Martin Gustafsson, Skill, Drill, and Intelligent Performance: Ryle and Intellectualism, *Journal for the History of Analytical Philosophy*, Vol. 5, Nr. 5, 2017)

Ex. “knowing how” (G.Ryle)

Knowing how : intelligible (witty, skillful, cunning, methodical···)  
e.g. the clown's stripping and tumblings  
neither physical nor mental, rather an embodiment.

A skill is not an act, but a disposition.

When a particular condition is realized, it is to be bound to be in a specific state or change, like “glass is brittle.”

?? Do we seek to understand the designer's intention in the robot, or only its achievement (expression itself)?

## 2) Empathy [issue(3) & (4)]

similar body movements → imitation → projection  
(robot arm)

- Similar arm movements are necessary??  
(Functional similarities are already enough.)

- Similar arm movements  
→ similar internal states  
or presuppose personal characters??

- the projected experience or character is always mine.

Understanding the expression as such, multiple-track  
but unidirectional.



TOROBO-KUN <https://www.japantimes.co.jp/tag/torobo-kun/>

## 2. Theory of Direct Perception

### 2-1 Theory of Direct Perception by M. Scheler and L. Klages [issue(3)]

“The expressive movement carries its meaning in itself and is therefore understood without regard to data of the outside world of its bearer.” (L. Klages, *Grundlegung der Wissenschaft vom Ausdruck*, 1935/SW. 335)

“For we certainly believe ourselves to be directly acquainted with another person’s joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, with his entreaty in his outstretched hands, with his love in his look of affection, with his rage in the gnashing of his teeth, with his threats in the clenching of his fist, and with the tenor of his thoughts in the sound of his words.” (M. Scheler, *The Nature of Sympathy*, 1923, tra. 1954, p. 260)

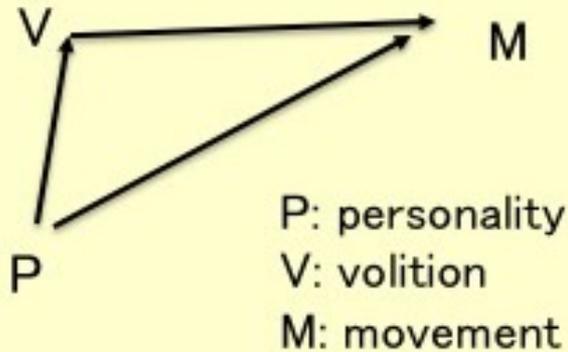


## Graphology [issue(3-2)]

- Intention → writing characters (letters)  
directed
- Personal character of the writer is expressed  
in the characters (letters)



“An expressive (body)–movement is the visible manifestation of the impulses and feelings which are represented in the vital movement of which it is a component part.” (Klages, *ibid.*)



“The expression manifests the pattern of a psychic movement as to its strength, duration and direction.”

⇒ the handwriting is a volitional movement and must necessarily carry the individual stamp of any personality.

(Expression considered as an aspect of the volitional movement, and represents the essence of the personality.)

every feature of the personality → personally characteristic movement

## Advantage of DP

- Focus on expression as such

Does not presuppose the internal “hidden” soul.

- fundamental scheme: “polarity of expression(act) and impression(suffering)”

An expression is a sign of a sense

### 1) bodily expression [issue(1) & (3)]

The body is the appearance of the soul; the soul is that which appears in the body. (Klages, *ibid*, 397f.)

We see the “actual” blush, as long as a state of affairs appears to us in the image[Bild] of a face covered with redness, and which now is of the appearance of the blushing sense. (Klages, *ibid*, 399)

### 2) Graphology, artistic works[issue(1) & (3)]

an expressive image produced by human

### 3) milieu things as an expression [issue(1)]

ex. awesome / enticing flame

polarity ⇒ “expressive images – enabling reasons”

“the appearing event exists by means of receiving souls” (Klages, *ibid*, 499)

## 2-2 Direct Perception(DP) and Robot

### Disadvantage of DP

1) presuppose a living organism and its expressive movements

“output – input”  $\neq$  “expression – impression”

mediated by a body–soul–unity

A robot does not have a “phenomenal body”.

2) abuse to Robot

A robot does not have an expressive movement nor pain.

Nevertheless, the **robot manifests a painful expression.**



## Alternative Paradigm

“Mask of a play”

motionless = without expressive movement

full of expression

the meaning of the motionless mask??

“What it offered must have been valuable at the time and what it left out must not have been worth pursuing; for it is a priori unlikely that the mask is nothing but a historical relic and that that which it excluded was entirely unknown.” (Karl

Bühler, *Ausdruckstheorie*, 1933, 15)(trans. BÜHLER'S ACTION THEORY OF EXPRESSION, UC Press E-Books Collection, 1982–2004 <https://publishing.cdlib.org/ucpressebooks/view?docId=kt4m3nc7mf&chunk.id=ss2.05&toc.id=ss1.12&brand=ucpress>)

## Noh mask

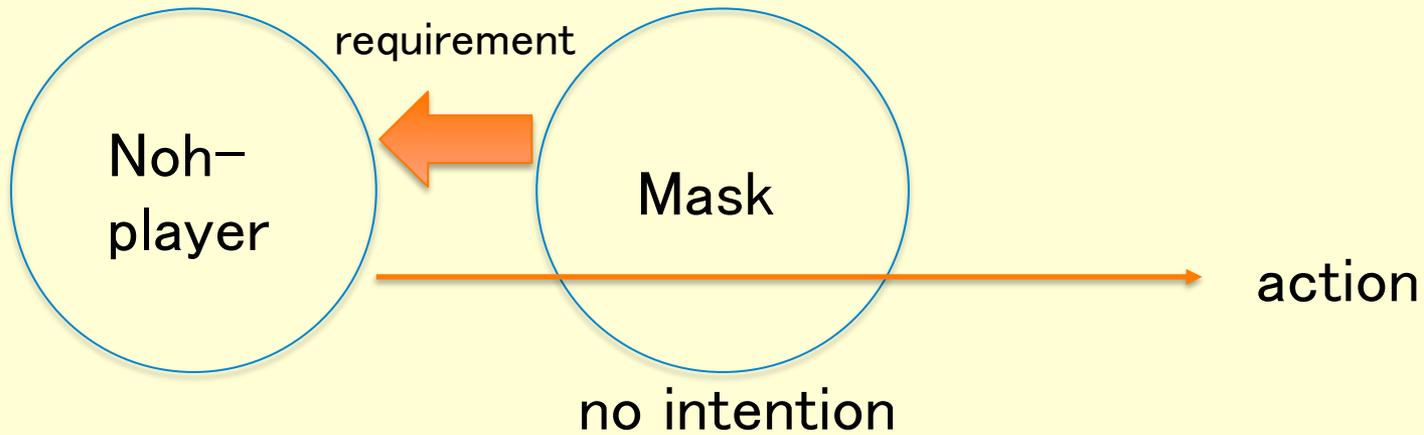


Physiognomy of  
artefacts, landscapes

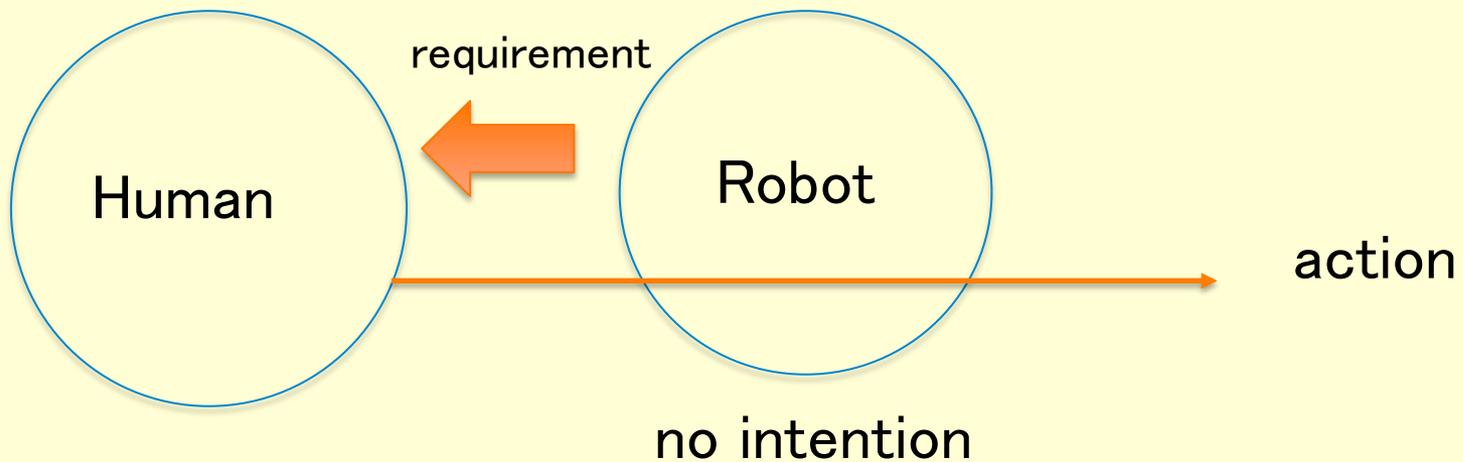
“It is things which abruptly unveil themselves to us as hateful, sympathetic, horrible, lovable. Being dreadful is a property of this Japanese mask, an inexhaustible and irreducible property which constitutes its very nature – and not the sum of our subjective reactions to a piece of sculptured wood” (Sartre, Intentionality)

(we do not view the mask first and decide it is dreadful subsequently. We view the mask, and the dreadfulness is already there, as well as all other properties the mask may possess for us.)

If this is so then compared to sculptures, which are by essence stationary, the mask is by essence moving. The true manifestation of the excellence of a mask has to be when it is put in a position of motion....when a Noh mask appears on the stage and gains a moving body, at that point something surprising occurs. Namely, the Noh mask – from which the look ought to have been stripped away – actually begins to display boundless variety in its looks....What especially draws our attention in the activity of these masks is that the mask totally absorbs into its own self the body and gestures of the moving actor who puts it on. Though in actuality it is the actor who puts on the mask and is moving, speaking in terms of the effect, it is the mask that has acquired a body. (Tetsuro Watsuji, Mask and Persona)

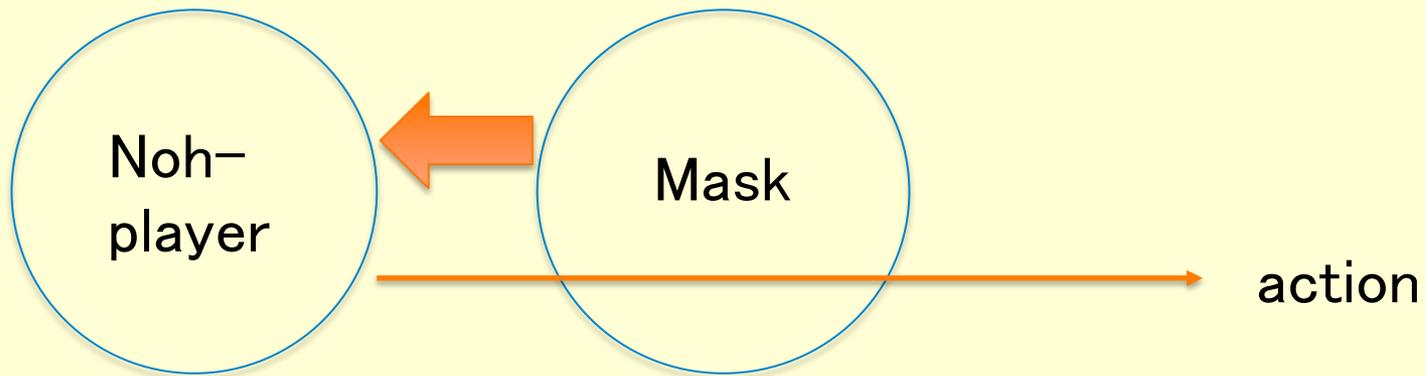


Artifacts require definite kind of actions  
(Actor's other properties are left out.)

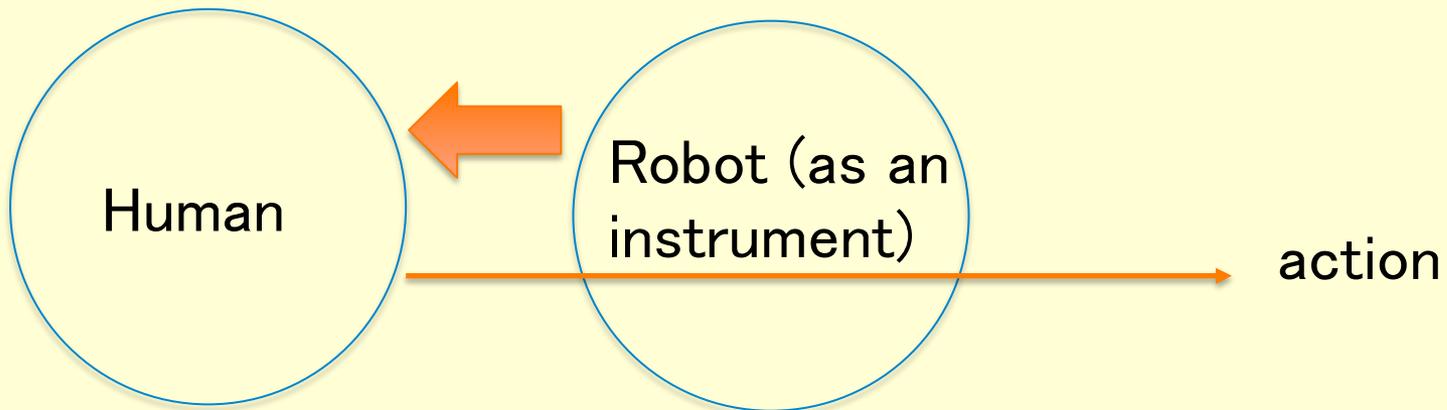


① Two-part relationship  
(Embodiment relations)

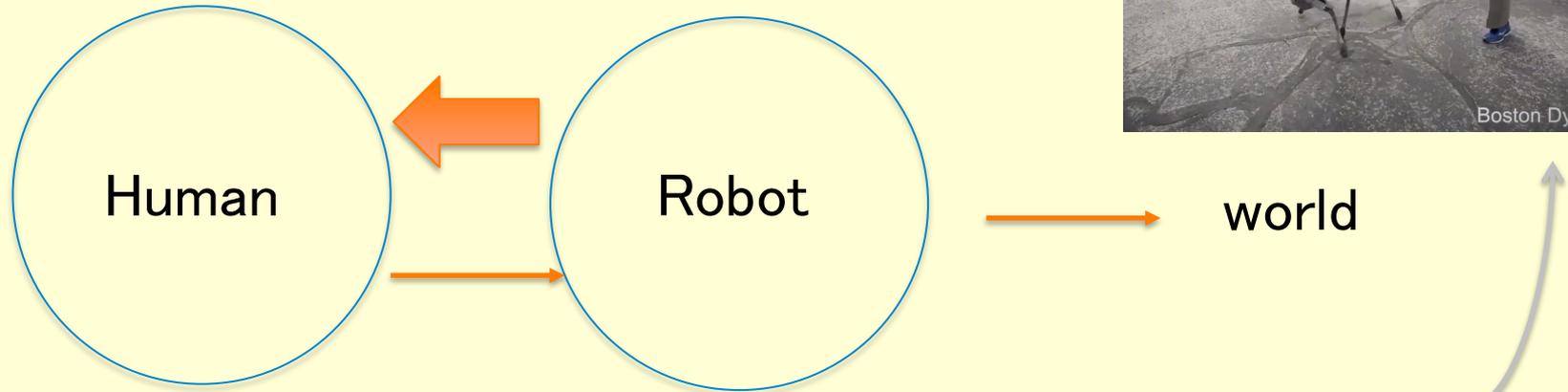
different feature of the personality  
→ personally characteristic movement



no internal states(intention etc.)



## ② Two-part relationship 1 (Alterity relations)



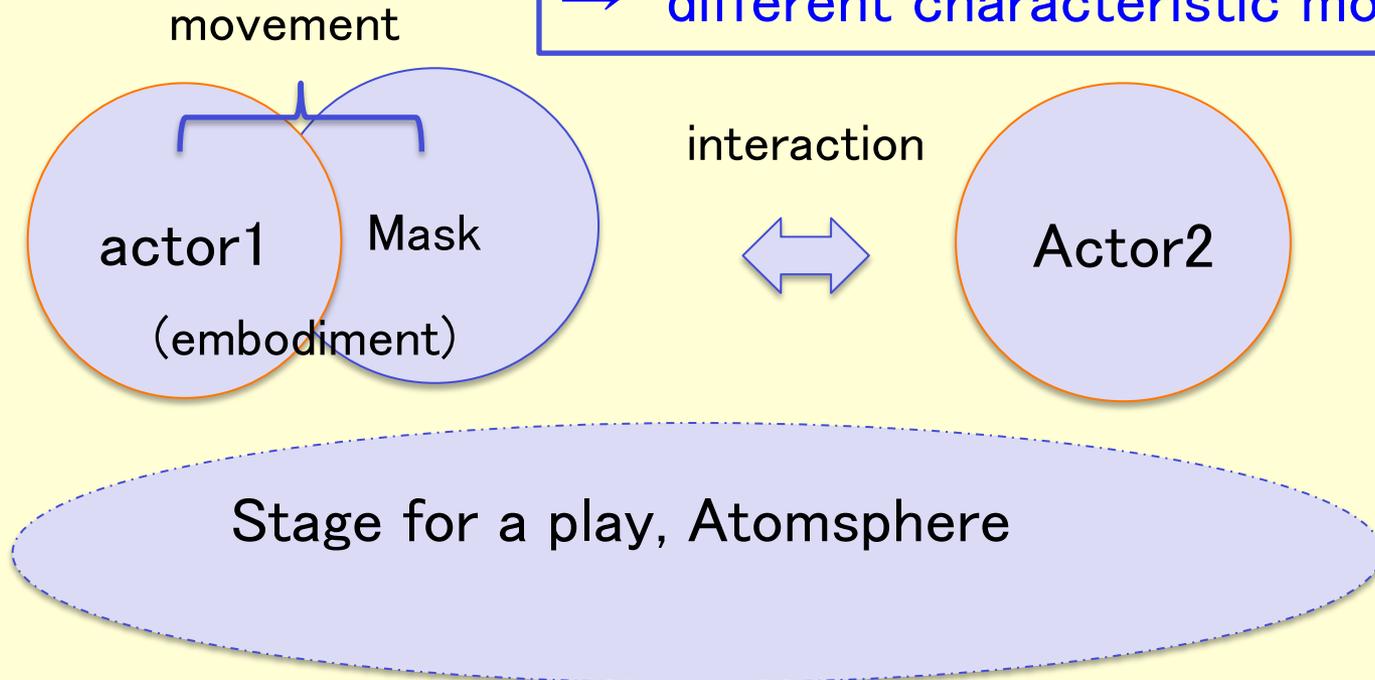
A robot is independent but has some relationship to human.  
(Robot is quasi-person  $\neq$  instrument)

A robot has some expression and require to do (or not to do) something.

different relationship to the actor  
+designer's intention  
→ individual characteristic movement (reply et al.)

### ③ Two-part relationship 1

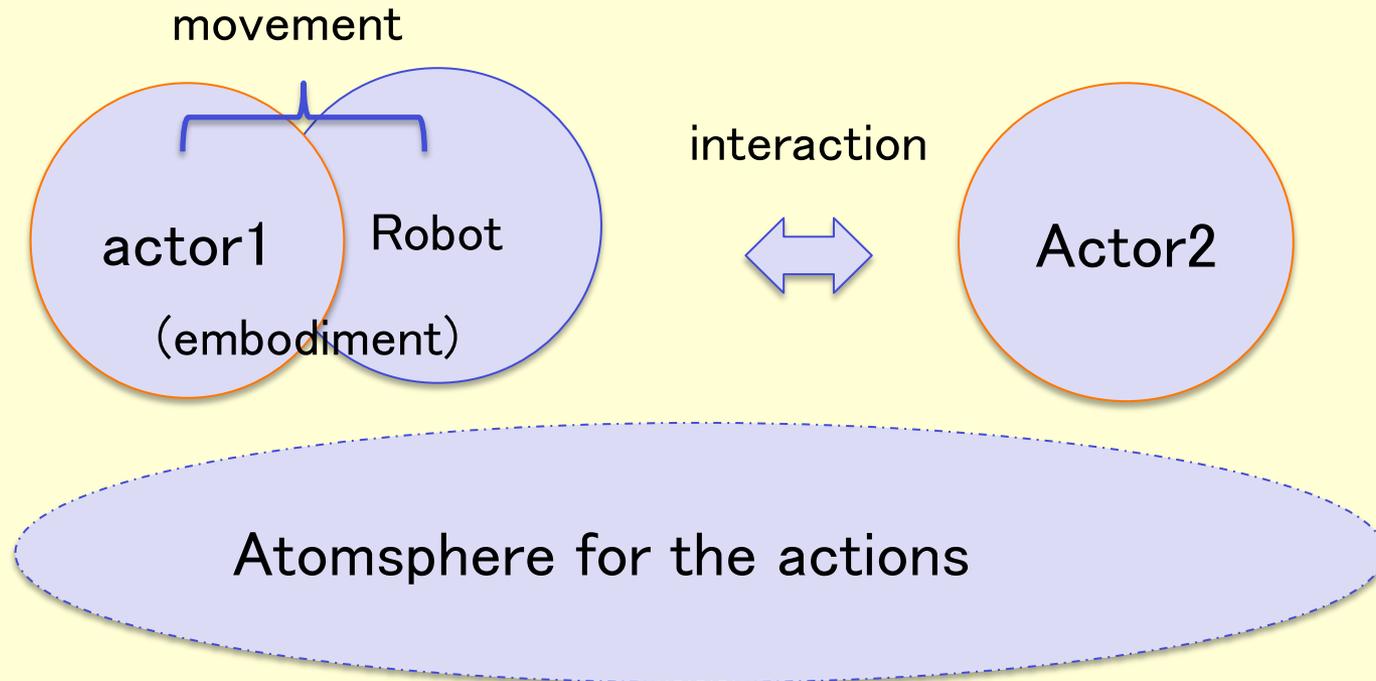
different feature of the actor's personality, other actors, stage  
→ different characteristic movement



“...in the Noh mask of a demon, we find all trace of humanity erased from the facial surface. Though it could also be said to “embody awesomeness,” it cannot be said to typify the awesomeness of a person's look. Generally speaking, it is not the face of a type of person.” (Watsuji, *ibid.*)

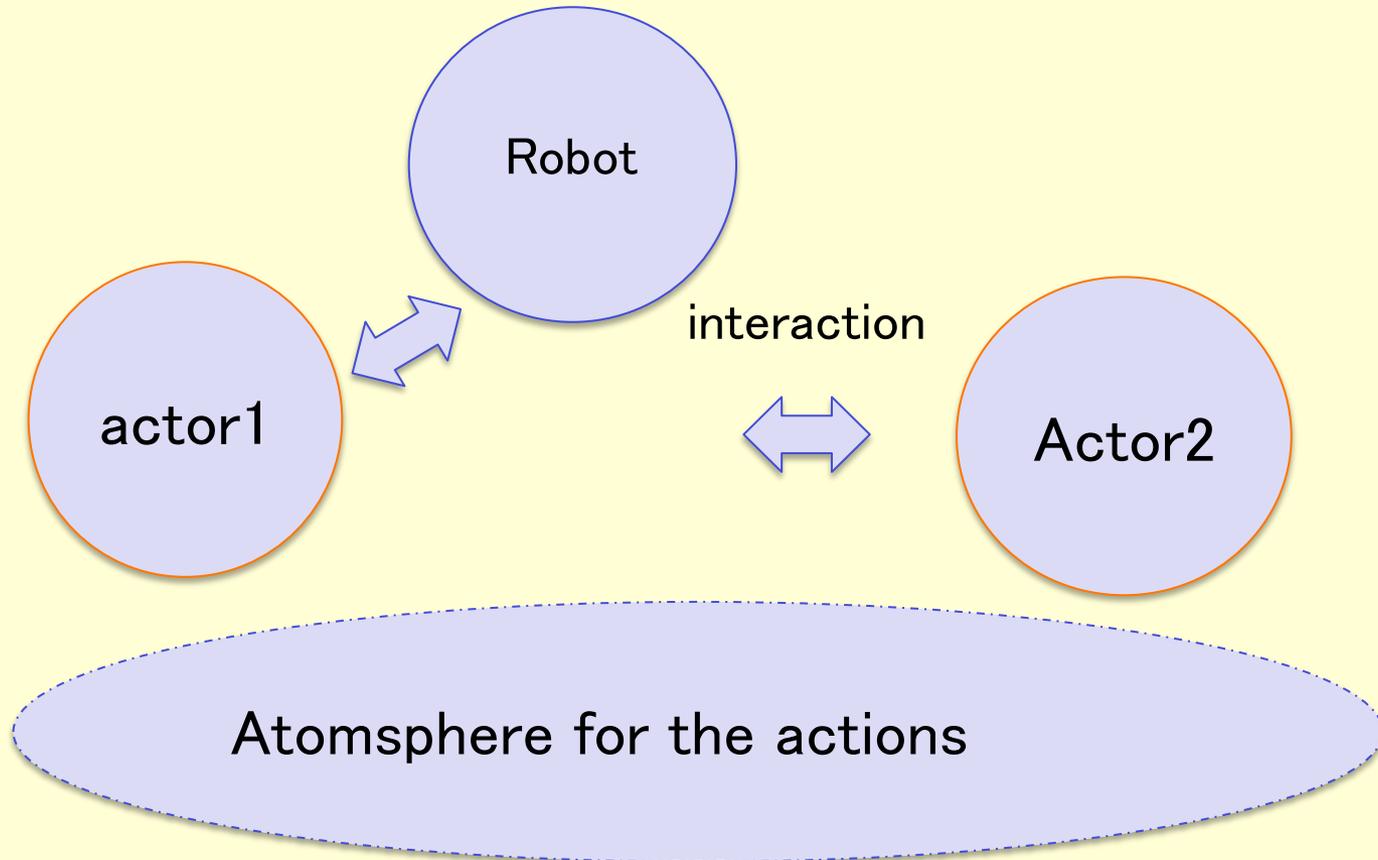
## ② Two-part relationship 2

different feature of the actor's personality, other actors, stage  
→ different characteristic movement



#### ④ Tripartite relationship

different feature of the actor's personality, other actors, stage  
→ different characteristic movement



## Conclusion : Relationship to Things that Relate to us

### 1) Beyond the dualism

human vs. machine

internal states – mechanistic

- Robot has no internal state(intention, feeling, emotion etc.) and no personhood(?).
- Robot has expressions, which is not only an analogy or a projection of the human subject.

### 2) Understanding Robot

- According to DP, an expression is perceived directly, without referring to the internal state.
- A personal character of the writer is expressed involuntarily in her/his writings. Every feature of the personality is expressed in personally characteristic movements.

- Expression of a robot is also perceived directly (as in the case of Robot Abuse).
- Characters written by AxiDraw V3 or TOROBO-KUN have also expressive character, though these robots do not have vital or expressive movements.
  
- Robot occupies an intermediate position.
  - i) traditional machine -- living organism
    - input-output                      “spontaneous movement”
    - no personhood                      quasi-personhood
  - ii) animation, avatar --- animal
    - virtual                                      real

(Both might have quasi-personhood.)

### 3) Communication without internal states

#### Noh-mask model

##### An Expression of a mask

- ┌ exists independently of human (creator, player)
- └ has dependency on human (player, audience et al.)

“Expression with no reference to any human” is non-sense.

- Whether in the embodiment or the alterity relation, the robot has acquired a body only in connection to the human actor.
- [hypothesis] Different features of the actor's personality, other actors, stage are expressed in the individual characteristic movement of the robot.
- Functions or a personal character defined by designer's intentions constitute the basis of this human-robot relation. But they are redefined in the interaction.