

Disposition and Mind
KyungHee University
2012. 05. 30.-31.

a distributed-but-unified mind

Insok Ko

Inha University, Korea

insok@inha.ac.kr

aim of the talk

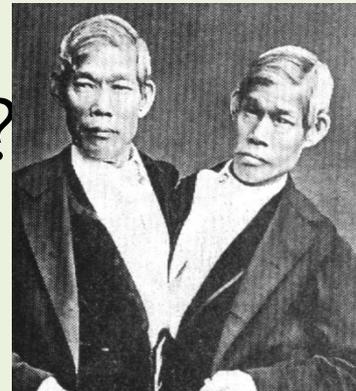
1. to analyze a familiar, but mostly implicit and unquestioned thesis
2. to discuss the place (or status) of intelligent robotic beings in our world

inviting questions

How many minds are in this room?

How many persons are in this room?

How many persons.. in
case of conjoined twins?



How to count the number of *res cogitans*?

Or .. **How are we doing it? And why?**

some background motivation

- Personal experience in a “Committee for Robot-Ethics Charter”
- From roboethics to ontology of robotic beings: “What kind of ontological status shall be applied to robots or robotic systems?” as a social & philosophical question
- We will be living in a society prevailing with highly sophisticated and clever robotic systems very soon. As what should we take such clever artifacts?

[Situations]

- A1: "Oops, a hand-shaped thing dangled from my right side hit on your head!"
- A2: "Sorry, my two-year-old daughter destroyed your car. I will pay for it."
- A3: "Sorry, a TF team of my company let you suffer such terrible, undeserved loss."
- A4: "Everyone who had handled the tool several times got severe damage in the hand. Definitely, the design of the tool is the problem"
- A5: "Sorry, the automatic railway control system killed your innocent husband while saving other five lives in the trolley accident."

[Question] Who (or what) is responsible in what way?

grammar of 'responsibility' analysis (including liability or accountability)

We seek two things:

- 1) the boundary of responsible agents
- 2) mode of responsibility for each of those agents: What kind of & How much responsibility is there?

Basic assumptions thereby:

- 1) individuation of (responsible) agents
- 2) only human agents have genuine agency in full-blown sense

grammar of 'responsibility' analysis (continued)

We ask “**What** caused the accident?”, but not “What is responsible for the accident?”. Instead, we ask “**Who** is responsible for the accident?” It is often the case, that the **What** and the **Who** in a given context designate different things. We discriminate ‘who’ from ‘what’. What is needed for a who-status? (How about the ‘Médecins Sans Frontières’, who got Nobel Peace Prize in 1999?)

[Situations (with robotic beings)]

- B1: A robot rescued a boy from drowning.
- B2: A robot tried to rescue a boy from drowning but failed, because it was inappropriately programmed.
- B3: The robotic system detected the situation and took the rescue action automatically, but it was too late. If it had moved more swiftly, it would have rescued the boy for sure.
- B4: The robotic system detected the situation and took the rescue action automatically, but due to some unpredictable fluctuation of environment the rescue robot bumped heavily into the boy's face to make him unconscious, which was the primary cause of the fatal accident.

[Question] Who, or what, was the agent in action?

What is a robot?

- [Merriam-Webster] 1: a machine that **looks like a human being** and performs various complex acts (as walking or talking) of a human being 2: a device that automatically performs complicated often repetitive tasks
- [<http://en.wikipedia.org/wiki/Robot>] **While there is no single correct definition of "robot", a typical robot will have several or possibly all of the following properties:**
artificially created; can sense its environment, and manipulate or interact with things in it; **makes choices** based on the environment, often using automatic control or a preprogrammed sequence; (re)programm-able; moves with one or more axes of rotation or translation; makes dexterous coordinated movements; **moves without direct human intervention; appears to have intent or agency**

ontological placement of (intelligent) robotic beings

- intelligent robotic beings as a form of extended mind?
- or as externalized mind, as a special type of extended mind
- sometimes in a form of distributed (but unified) mind

- “Consciousness is not something that happens inside us. It is something we do or make. Better: it is something we achieve.” (Noë, 2009)
- “Where does the mind stop and the rest of the world begin?” (Clark&Chalmers, 1998)

debate about extended mind (1)

- The parity principle
"If, as we confront some task, a part of the world functions as a process which, *were it done in the head*, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world *is* (so we claim [for that time]) part of the cognitive process." (Ibid.)

debate about extended mind (2)

- The coupling-constitution fallacy
“The fallacious pattern is to draw attention to cases, real or imagined, in which some object or process is coupled in some fashion to some cognitive agent. From this, one slides to the conclusion that the object or process constitutes part of the agent's cognitive apparatus or cognitive processing.”
(Adams&Aizawa, 2010)

evaluating Otto's notebook

- What kind of (active!) cognitive function (or cognitive role) does Otto's notebook take?
- Let's suppose that Otto's notebook is (upgraded to) a notebook PC, a smart one. It would be a part of Otto's cognitive apparatus, but still only of a complementary mode and not of a constitutive mode with respect to Otto's cognition. It is functional only when it is called to action by Otto and **its content is interpreted by him.**

Chinese room revisited

- [contra “systems reply” to Searle’s C-Room Argument] Even the whole room [equipped with the alleged system] does not understand meaning of any sentence, though it has some phenomenal “linguistic capacity”. (This is nothing but the original message of the argument.)
- A clever computer (refinement of C-room or Otto’s notebook) with some learning mechanism is different from and more than Leonardo’s tattoos in *Memento*. It automatically processes and reconstructs the available input data to produce new output data.

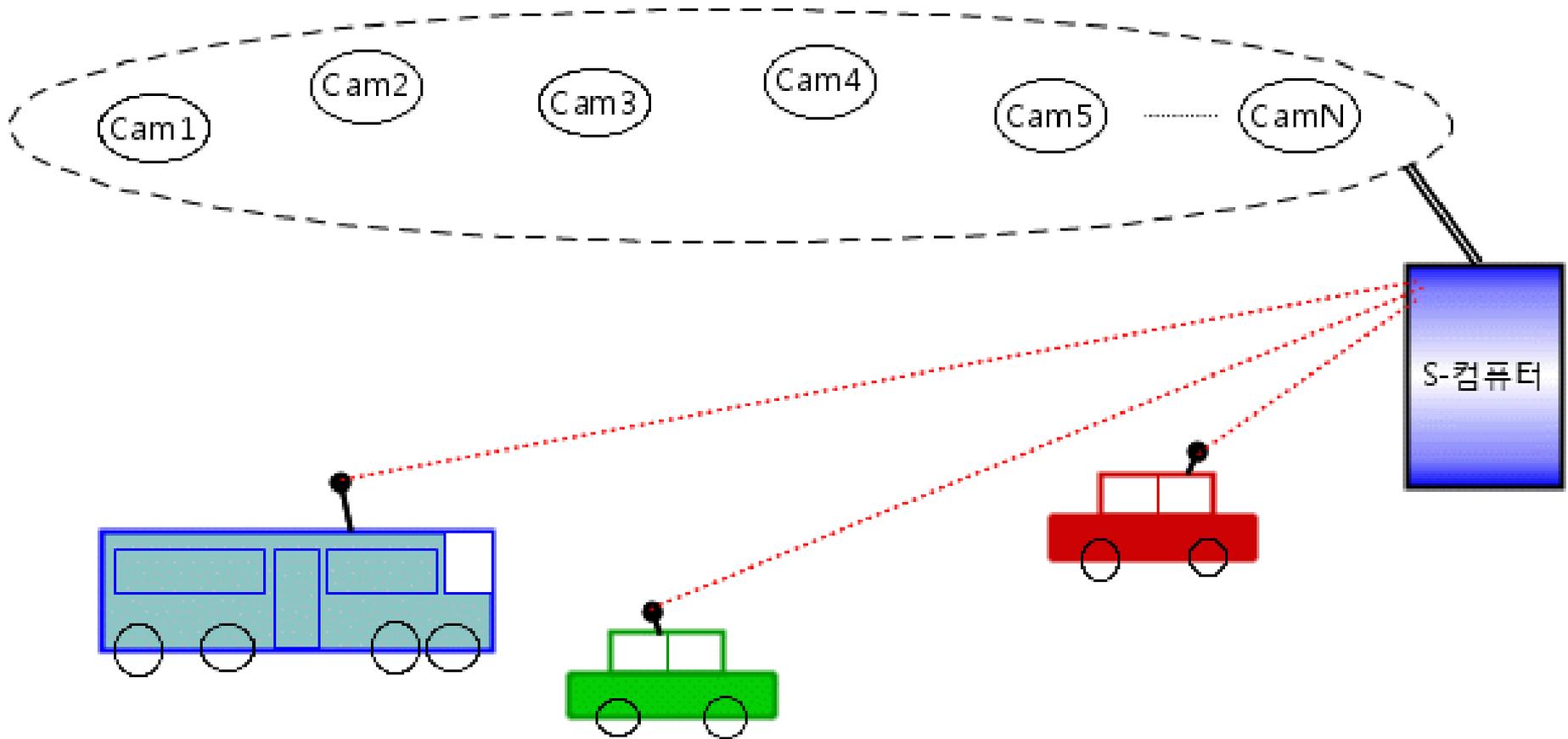
evaluating Otto's notebook again

- Its function depends essentially on the man-made program, in which the 'interpretive base' is integrated. But while it runs, it runs by its own nature without human intervention.
- This suggests the possible status of Otto's notebook, whether it is integrated in his skull or just remote-controlled by him: It is not a genuinely but virtually (at the phenomenal level) autonomous cognitive agent which functions to some complementary effect to Otto's own cognition.
- Its status lies between Otto's old notebook and a genuinely autonomous cognitive agent.
- The parity is not that strong but some coupled thing constitutes genuine parts.



A woman, paralyzed from the neck down, uses a robot arm to serve herself a drink, a first for her in 15 years since a stroke. (NY Times, 2012.05.16)

a case of application of robotics (general urban transportation control system)



unity of the distributed system

- The number of participating computers, local surveillance systems, smart vehicles, etc. can be arbitrary. They should only be attuned so as to effectively cooperate in a common project or task.
- The degree of such attunement, or the degree of unity, of the whole system is at our disposal. It depends also on technology and constraints in the given situation.
- Such an attuned system of distributed objects would work as if it has a mind. The function of this apparent (or virtual) mind does not depend on human intervention. I.e., it will function as a quasi-autonomous agent.

intelligent robotic system as externalized social mind

- Recall the question at the beginning of the talk: "As what should we take such clever artifacts?"
- My suggestion: Such robotic system is to be interpreted as a form of externalized social mind.
 - extended mind vs. distributed mind
 - extended mind vs. externalized mind
 - Why should it be a "social" mind?

conclusions

1. A robotic system equipped with some level of artificial intelligence will function as an autonomous agent on the phenomenal level while it runs. It will perceive, calculate, decide, make and perform decisions.
2. Whether it has genuine autonomy is debatable, but it does not have to be a genuine autonomous agent in order to fulfill its role in our world.
3. Such system would consist of spatiotemporally distributed parts. It will manifest a distributed-but-unified mind.

some additional thoughts

4. We shall entrust certain agential roles to such an artificial system. We may grant it a sort of **entrusted authority** within the boundary of its functionality.
5. Such robotic system is to be interpreted as a form of externalized [social] mind.
6. In order that such intelligent robotic systems, especially the ones embedded in public domain, are sound ones, they **should** be the externalization of social mind and not of some specific individual minds.

- I maintain that our society in near future shall entrust certain restricted authority to some of the artifacts we produce.
- For instance some intelligent robotic systems would take care of the whole urban transport system including the safety management.
- We would then be obliged to conform to the rule and order realized in the intelligent mechanical system.

remaining question

- Is there certain element of responsibility corresponding to the entrusted authority?
I am skeptical.
- But who/what carries then the responsibility for the dispensable damages caused by application of the entrusted authority of intelligent artificial system?
- This question is of a practical sort, and we should deal it also from a pragmatic viewpoint: What is the best way to distribute and attribute responsibility in the given cases?

references

- F. Adams & K. Aizawa (2010), "Defending the Bounds of Cognition", in: R. Menary (ed.), *The Extended Mind*, MIT Press.
- A. Clark & D. J. Chalmers (1998), "The Extended Mind", *Analysis* 58.
- A. Noë (2009), *Out of Our Heads*, Hill and Wang.
- W. Wallach & C. Allen (2009), *Moral Machines: Teaching Robots Right from Wrong*. Oxford UP.
- <http://en.wikipedia.org/wiki/Robot>
- <http://plato.stanford.edu/entries/chinese-room>
- http://www.nytimes.com/2012/05/17/science/bodies-inert-they-moved-a-robot-with-their-minds.html?_r=1&ref=global-home