

Person and AI
Yuko Murakami
(Rikkyo University)

We cannot live without technological supports in the current society. We implicitly assume the whole system work properly and we can rely on any services we usually use. We have no difficulties to come to the conference venue: we walk on the street, climb up and down the stairs, ride trains, exchange emails and messages with mobile phones and other ICT tools, and find where we are on a map with GPS. All these actions appeal to the services and infrastructures with the current society and technology. You might feel awkward if you do not read Japanese. A language is a social infrastructure. If you might not rely on it, you need help. Our physical ability is enhanced by those social systems. The line of thought reaches the observation that our agency is not simply based on our own physical body, while the current notion of agency is strongly connected to the notion of natural person, which is based on our physical body. In this paper, I argue to extend the notion of agency, and illustrate how it affects morality.

Traditional cognitive science assumes that the brain plays an essential role for performing mental activities. It is also a common belief that the mind supervenes to the body. Researchers of artificial intelligence (AI) focuses brain science, because they believe that the mind is reduced to, or at least corresponds to, the activity of the body, in particular the brain. Observational results of vital signals, such as cerebral blood circulation, are considered as physical counterparts of mental activities. For example, Kamitani and Tong (2005) decodes fMRI results to display what the experimental subject sees or even dreams. There seems direct correspondence between the mind and the brain.

The notions of mind and person involve tacit assumptions, while artificial intelligence seems to undermine intuitive justification of each set of such assumptions. It is illustrated by the series of externalization of the mental contents and extension out of the human body in the history of philosophy.

There have been trends toward externalization in the history of philosophy from the mid twentieth century (Lau and Deutch 2016). First meaning is externalized. Before Putnam (1973), the meaning of the language, which is the basis of language use regarded as activity in the mind, has been considered mental content. He argues, however, that meaning is not in the brain, but external conditions determine meaning (externalism of meaning) by the twin earth thought experiment. It is a thought experiment, so be careful not to think whether it is physically feasible or not.

The next externalization was memory, which has also been considered to be “mind-related.” Memory, as well as language use, has been regarded as human heart activity. Clark and Chalmers (1998) claimed the Extended Mind Thesis (EMT), inspired by the example that the note written in the notebook extend one’s memory. He formulates the Parity Principle (Clark 2008).

[Parity Principle] When we work on something, if we accept without hesitation as part of the cognitive process, as part of the world functions as if it gets in our mind (At that time) part of the world is part of our cognitive process.

I would like to claim here that the parity principle can be extended to agency as we extend our own body to involve our social systems. Driving a car, riding a public transport system in Tokyo or other big cities, and using information device such as smartphones and tablet computers in fact require training before entertaining them. Once you have access to these systems and device, it is almost trivial to use them to do what you want to do. Such a public transport system is not owned by you (unless you are the owner of the transportation company).

Each part of such a system are often interdependent. Adjusting one part affects other parts, but changes in other parts also bounce back to the original part. By putting dependencies, access relations, and reference relations together and describing them as a subsystem, there is no guarantee that atomic elements exist in the "part" of the agents.

[Parity Principle of agency] When we act, if we accept without hesitation as part of the cognitive process, as part of the world functions as if it gets in our mind (at that time) part of the world is part of the process of our agency.

Reference

- Kumagai, S. 2009, *Rihabili no yoru*. Igaku Shoin. (Nights of Rehabilitation, in Japanese)
- Lau, J and Deutsch, M 2016, "Externalism About Mental Content", *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition)*, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2016/entries/content-externalism/>>.
- Kamitani, Y., & Tong, F. 2005, Decoding the visual and subjective contents of the human brain. *Nature neuroscience*, 8(5), 679-685. doi:10.1038/nn1444
- Putnam, H. 1973 "Meaning and Reference." *Journal of Philosophy* Vol. 70, No.19, pp.699-711. doi: 10.2307/2025079
- Clark, A. and D. Chalmers 1998, "The Extended Mind." *Analysis*, Vol. 58, No. 1, pp. 7-19.
- Clark, A. 2008, *Supersizing the Mind*. Oxford University Press, Oxford.
- Belnap, N. 1991, "Backward and Forward in the Modal Logic of Agency." *Philosophy and Phenomenological Research*, Vol. 51, No. 4, pp. 777-807.
- Belnap, N. 1991, "Before Refraining: Concepts for Agency." *Erkenntnis*, Vol. 34, No. 2, pp. 137-169.
- Nakayama, Y. 2013 "The extended mind and the extended agent." *Social and Behavioral Sciences* 97, pp.503-510.
- Parfit, D. 1984, *Reasons and Persons*. Oxford University Press.
- Parfit, D. 1987, Divided Minds and the Nature of Persons. In Colin Blakemore & Susan A. Greenfield (eds.), *Mindwaves*. Blackwell.
- Prior, A. 1967, *Past, Present and Future*. Oxford University Press.
- Thomason, R.H. Indeterminist time and truth-value gaps. *Theoria*, 36:264-281.