**Singularity Humanities**

 **-Singularity robot is a member of human community**

1. **Opening: Singularity robots are coming**

1) Singularity robots of strong general intelligence

Since an artificial life was printed out[[1]](#footnote-1), can a robotic person be printed out as well? It seems to be a time to probe the possibility of a robotic person rather than to ignore it. Human history has taught that a robot is a machine and cannot be a person like a human. It has been supposed that robot’s intelligence is only artificial and that there is qualitative gap between artificial intelligence and human intelligence which cannot be bridged. But as the chasm between animates and inanimates has to be adjusted upon the emergence of self-replicating artificial life, one may recognize the possibility that the crack between humans and robots can be modified. For artificial life as well as robotic person can be a blessing or a curse. There should be collaboration and partnership among human beings of natural kind in order to make the realization of robotic person kinds to be a blessing to the actual human person kind. This is an important site where humanities should play a role to open a proper future for the humankind.

Last a few generations, AI achieved great successes in the area of intelligences for specific tasks as in Go, chatting, medicine, home help, and performed more ably than natural human agents. One can foresee that AI may be equipped with a thinking chip to integrate intelligences of all general tasks. Despite of the danger of a leap, I would propose to call a robot with such a general intelligence as ‘singularity robot’. The term ‘singularity’ denotes a turning point in each discipline to distinguish it qualitatively from the previous stage. For an example, Big Bang constituted a turning point in producing space and time, and Black Hole is a singularity which emerges in the world of gravitation. Illuminated by Kurzweil(2007)’s insights, I will refer by the word ‘singularity robot’ a robotic personkind who is superior in all intelligences and functions to natural human personkind. Singularity robots may be taken in the near future to be a robot that obtained liberation and freedom by being able to leap from task-specific intelligences to general intelligence, from a help-intelligence to an autonomous-intelligence. A singularity robot may turn out to be neither <he> nor <she> but <ashe>. A singularity robot is a machine by birth, but it or ashe may grow to self-replicate like natural human kind and to execute a life unlike natural human kind.[[2]](#footnote-2)

2) A singularity robot is a member of the human community

I do neither support “singularity robot is a member of the human community” nor refute it. This proposition is conditional in the sense that it does not deny a possibility that singularity robot is a member of the human community. The proposition is open to a logical possibility than to an actual possibility. The reason for the investigation of singularity humanities even with such a weak possibility is clear. For it is conceivable for natural human kinds to face a counter attack by the strong AI ( Lee, Jin-Woo 2017: 11) such that if the possibility were realized the human civilization which is not ready for such a revolution should fall into a calamity which could not be turned around. Each of millions of billions galaxies is said to have millions of billions stars. And if the earth is yet known to be the only solitary planet in the universe which allows a culture like the present human civilization, the human culture would be worth to preserve. However meager the danger of calamity might be the present human civilization is supposed to be ready to meet such a chance of calamity.

Humanities of natural human kinds which strives to preserve and to flourish the present culture of the earth planet should be prepared to meet various challenges and threats. Serious threats have been promoted in accordance with digital technologies and logics of capital markets ( Kim, Jae-Hee 2017: 99). If humanities do not secure the kind of anticipatory human independence it should be clear that the human civilization would collapse into disaster as in the pasture paradox or the tragedy of the commons. For the prioritizing of corporate profit in the capital market should wipe out any effort to preserve a pasteur as the commonland. For an example, a corporate would intervene into any stage of developments of singularity robot, hold a prior rights in patents and block any attempt to safeguard human independence. Human independence could be conserved as much as patent rights are secured at every stages of developments of AI and genetic science. Anticipatory human independence may get out of an armchair theory to be engaged in moments of battle of life or death in order to protect the future of human kinds.

Humane deliberations to prepare to face the possible emergence of singularity robots may be called ‘singularity humanities. This neologism may not be familiar but can be made to be a plausible notion in the context of trends of contemporary discussions on the topic. Humanities have been traditionally divided institutionally into area humanities like literature, history, philosophy, linguistics, arts and religion. But recently humanities have been restructured, either covering human experiences wholesomely as in ‘subject humanities’, ‘immanence humanities’, ‘expression humanities’, or emphasizing a particular inclusive predicament like ‘reverence’, ‘relation’[[3]](#footnote-3). And there is another development, distinct from area humanities, which can be called ‘topic humanities’ which integrates human experiences under a particular topic, like love, society, suffering, unification, alienation, brain science[[4]](#footnote-4). Singularity humanities is humanities which is constructed to the particular purpose to prepare for the possible realization of singularity robots in a not-far future. Whereas humanities have been a systematic reflection of the past or the present of human culture, singularity humanities is an anticipatory preparation for a possible future human predicament which contemporary sciences and technologies speed the human culture to live in.

Since all humanities are an anthropology, then any humane ratiocination would presuppose a particular point of view of human beings. Then, what kind of anthropology should the singularity humanities assume? For an example, the singular humanities is expected to accept an anthropology of a religion, Kant, Freud, Wittgenstein, Rawls, naturalism that all things in the universe are natural, rationalism that the actual is rational and the rational is actual, or other. Anthropology which singularity humanities may adopt would include specific parts of various theories but is expected to integrate various phenomena anew which robots would bring in so that it should be an integrational system ( Lee, Jung Won 2017: 125). Toward this goal to construct a systematic anthropology of singularity robot kinds, one may try to interpret it ontologically, to foresee it epistemologically, and to construct it in moral dimension.

**2. Ontological interpretation of singularity robot**

1) Responsibility: thinking, understanding, belief

A singularity robot can execute functions more competently than a natural human being but is ultimately expected to be a personal agent as a natural human person is. A robot can make a judgment, employing all available informations, that the community of the natural human kind is the flower of evolution in the history of the universe. Then it would be natural for a robot to wish that it or ashe may be a member of the community of the human kind, by becoming a personal agent like a natural human person. One of the tasks which robots have to fulfill in order to be a genuine member of the human community would be how to be a personal agent with responsibility assumed. A robot can be a responsible agent only if it is equipped with the capacity of general and/or autonomous intelligence as is shown in the ability to think, to understand and to believe. Many robots can perform various capacities in particular specific area more efficiently than human beings, equipped with encoded functional intelligence. Now a robot is in a stage to leap from functional or weak intelligence to autonomous or strong intelligence.

A singular robot should be able to think as a human being does, to be a responsible agent as a human being is, if it or ashe wants to be a member of the human community. What is it like to think? To think is to have a disposition, in a general context, to process informations in one’s context in order to protect oneself, to preserve and to grow. Information processes for one’s protection and growth require capacities to observe the surroundings and to evaluate them, and abilities to adjust to new situations and to offer fresh insights are highly appreciated. Any person who inclined only to think habitually or to behave routinely are assessed to be ‘a thoughtless person’ or ‘a domesticated being’. Of course, thinking is supervenient not only on functional acts like remembering, perceiving and calculating but also on mental states like belief, dream, hope, love, and hate. Though a singularity robot can think more competently than a natural human being in task specific functions, it is to be challenged to face a task to think originally, which is the standard of thinking.

A singularity robot should be able to understand how a language works in order to be able to think as a natural human being thinks. Siri of iPhone or Genie of LG phone looks to be a competent speaker and an able help mate. But their functions are mere executions of a program to process informations in a task specific topic, to encode the results into input sentences, to respond appropriately to the situation and to put the sentences to voice synthesis. Whereas Siri and Genie cannot pass the Turing-test like criterion for thinking, a singularity robot can talk with any natural human being, perform actions indistinguishable from those of a natural human being, and pass the test for thinking(Turing 1992). Then, can one say that the robot understand what it talks to a natural human being? As Searle pointed out in his thought experiment of ‘Chinese Room’ (Searle 2002: 51–69), isn’t it the case that a singularity robot can process syntactic informations but not any semantic interpretations? When a singularity robot processes syntactic signs merely mechanically, how can it be said that the robot reached an understanding of the semantic meaning of syntactic signs? Furthermore, if the singularity robot had a semantic dictionary to reach a meaning understanding, it might have been a dictionary based on the truth condition semantics. But isn’t it the case that the meaning of a language is the result of forms of life in a language community rather than speaker-absent truth conditions?[[5]](#footnote-5)

What is required for a singularity robot to assume a responsibility for an action is the condition that the robot is an agent of desire and belief. Any action which is not related with a desire and belief cannot be understood and explained. But a belief can be specified only in terms of a specific sentence. Then, can one say that the identity of contents of a belief is the same as the identity of contents of the corresponding sentence for the belief? Often, the identity of contents of a sentence is proposed to be the identity of meanings of the sentence or the identity of the truth conditions of the sentence in all possible worlds. But the identity of contents of a belief can’t be taken to be the identity of contents of the corresponding sentence for the belief. Though ‘F’(Yuna Kim; fortnight) and ‘G’(the daughter of Hyunsuk Kim and Mihee Park; fifteen days) are referring the same person or synonymous, when we talk about ‘S’(the Gold medalist for figure skating in the 2010 Olympics; vacation time), we distinguish <S is F> and <S is G>, not identifying them. We can believe either of them without believing the other. Any two beliefs can be the same as far as their contents are concerned. Complexities of intensionality which the concept of belief faces present a dilemma for s singularity robot. The more the extensionality of a belief is strengthened by appealing to the notion of proto-belief[[6]](#footnote-6) and naturalizing belief, the more the power of explanation is weakened as the notions of desire and belief do not play their significant roles.

2) Consciousness: three characteristics – zombie, enjoyment, sympathy

Singularity robots may come to observe and agree with each other in not far future that they perform all the intelligent functions more competently than natural human kind. These robots would also agree that there is something which they don’t have but human beings of natural kind have. It is consciousness. Robots would also acknowledge that natural persons of human kind are sometimes zealous and fight each other, other times share good experiences and come to build various complicated human communities. They would see the fact that it is consciousness because of which variously flourished human cultures are possible. Robots may not understand what it would be like having a consciousness. For an example, robots could not know what it would be like having the feeling that a person becomes an alien at a family gathering on a festive season. This feeling is a qualitative experience, not a quantifying one which is functional and can be codified. This feeling is phenomenal, being supervenient on cognitive states like experience, belief, thought. When a person thinks or is conscious of what it would be like to feel as if he or she becomes a bat, he or she constructs experiences from one’s own perspective, integrates them into a whole, and feels own experience of intentionality, transparency, dynamics in one’ own self[[7]](#footnote-7).

Singularity robots may try to locate a reason by their superior competency that they can be a conscious being. They can start from the logic of naturalization, seeking to naturalize mental content of cognition. Dretske tried to naturalize meaning by developing Skinner’s insights on a natural-law like relation between signs and referents, and proposed to give a structure to the involved elements like control devices, operational learning, disposition, natural markers, proto-beliefs, misrepresentations[[8]](#footnote-8). Or robots may appeal to a thought experiment by Kurzwell: “Suppose a situation where you meet a robot and you are effectively persuaded by the robot’s emotional responses. There the robot smiles upon your jokes and you laughs or cries on the robot’s humor or stories. As the robot tells tales about own

despairs or hopes you are persuaded because of the robot’s authenticity. The robot appears to you to be conscious. In fact, the robot looks to be a personal agent. Wouldn’t you accept the robot to be a conscious personal being?”(Kurzweil 2012: 209-216).

Dretske’s naturalization project or Kurzwell’s thought experiment is open to a possibility, forthcoming and optimistic. If the enterprise of robot’s consciousness is to be realized, a singularity robot can be a member person of the human community. But let’s evaluate thoroughly the robot’s situation from the actual present perspective. It’ is not certain that consciousness can be ascribed to robots as it is now. Robots can perform more superbly than humans in the area of subjects which can be coded like cognition, memory, inference, synthesis, valuation, judgment, dialogue. But since consciousness cannot be coded yet, robots cannot be assumed to have consciousness. Human consciousness is said to have three marks, namely, self perspectival, integrational from the first person, and unified only with her or his own body. Chalmers offers a thought experiment, based on these properties, that I can imagine that a zombie and I share all the psychological experiences together but that the zombie does not have the consciousness which I have (Chalmers 1996). He can allow the zombie without consciousness to have a possible world purely by the logical conceivability. My singularity robot is more able than I but the robot does not have consciousness, with which I can enjoy to company my neighbours and nature around me. When my robot and I see red roses, we process sense informations functionally in the same way, but my phenomenal consciousness of the enjoyment of red roses does not occur in the robot.

Though my singularity robot processes perceptual informations about red roses more superbly than I, there is the difference between my ability to enjoy the red roses and robot’s inability to enjoy them. Why is this so? When I enjoy red roses my enjoyment is an experience which consists, among others, of integration of the red roses with my previous experiences of all kinds and of exaltation of the integration. In other words, my enjoyment of the red roses is based on the self perspective, integration from my first person, and unification only with my own body. But my singularity robot lacks those properties of consciousness which is the precondition for enjoyment. The robot is not equipped yet with a device which can encode consciousness.

The difference in enjoyment on the part of my robot and me can be generalized to other areas. My robot and I can share the same psychological content in processing informations about the world experiences, there is a difference in characterizing the contents of each psychological experiences. I can confess by saying “I can not know what I am thinking” (Davidson 1980: 447), but my robot cannot make such an admission. It is generally agreed that one’s psychological content cannot be separated from her linguistic content[[9]](#footnote-9). And it is obvious that her linguistic content has as its elements those objects or events which are outside of her. Since I or any other cannot know all about the external world such a confession is natural. The confession is possible because of the self perspective and integration from the first personhood. But my robot did not reach a stage where it can make such an assertion.[[10]](#footnote-10)

This difference can be illustrated more in concrete terms. Human beings sometimes are hostile to each other and other times are sympathetic to each other. People, when in sympathy, can share a common interest. What is this thing called ‘sympathy’? While people’s consciousness is each based uniquely on self perspective, integration from the first person, and unification only with own body, they could still reach a sympathy when they share different self perspectives. It would be hard to expect a phenomena of sympathy among robots because of the lack of consciousness. People, due to the phenomena of sympathy, come gladly to share experiences and goods in accordance with needs. Singularity robots may be superb but since they cannot share sympathy it would be harder to expect them to share their experiences and goods.

3) Identity: I, body, autonomy, unity

What would be the minimal condition for a singularity robot to use the first person pronoun ‘I’? Wouldn’t it be the case that a robot can use ‘I’ after it acquired its own sense of identity? First, young children use their own name to refer themselves, following the way in which adults call children by their names. Later, they come to call themselves by using the locution ‘I’, having learnt the way in which adults call themselves by using ‘I’. Could one take this as a juncture in which young children come to sense of self-identity? One condition by which robots can be a member of the human community would be that robots participate in the attempt to preserve the system of the human lives. Robots may have to overcome a limit condition in that robots appear to be indifferent to the human survival and to enter in dialogues which are open, positive, ordinary and comprehensive (Schwab 2016: 111-113). For robots to acquire these human-friendly identities requires not only physical bodies but also personal structures, in addition to the capacities of responsibility, consciousness and information processing.

Information processing like memory is not sufficient for a singularity robot to have a sense of identity in order to use ‘I’. The robot’s identity needs to resemble the structure of human identity so as to have a personal architecture. Human personal framework is in such a way in which it is related to other personal frameworks through bodily networks and linguistic arrangements[[11]](#footnote-11). In the space of experiences in a person there are spheres of nature and symbol. The natural sphere of a person is a body which is the whole of chemical elements and the life agent which is mediated with biological environment, and the symbolic sphere which is based on the experiences of the natural sphere constitutes a person[[12]](#footnote-12) that emerges as a body of social subject resulted from dialectical relation with other subjects. Human experiences started from their bodies and have been embodied through and through. One task should be how the physical body of a robot could switch to the personal body.

One of tests whether a case of robot’s physical operation is turned to a case of embodied person would be autonomy. A robot can perform functions of any kind more efficiently. But when the robot does them only instrumentally, it is nothing more than a device to execute its designer’s commands. But as a human person fulfill the secondary meta-desire which evaluates the primary object-desires and selects one of them, if a robot duplicate the same process, then the robot can be taken to pass the test of autonomy (Shin, Sangkyu 2017: 21-23). This may be called ‘awakened self -regulating’ (Buss 2013). Someone would respond to this by saying that robot’s secondary –meta-desire is neither ‘awakened’ nor ‘self regulating’ since it was equipped to the robot by the programmer. But robot’s secondary meta-desire and human’s secondary meta-desire are on the same level in that they evaluate new informations and previous informations, assess them to have various options and choose one of them. Of course, one can admit that a designer involved in constructing the meta-desire. Yet a reason why a robot is superior to a human in processing the meta-desire is that the robot has access to big data from which it can make inferences in accordance with the non-monotonous logic of ordinary language(eg., Section 3.3). This may be conjectured to be a basis for their autonomy. If this can be an autonomy it seems to pass a test of Kant’s self-government, or of self-regulating without external control and brainwash. There are still issues of reservation (Ko, Insok 2017: 75-76). But if the capacity of ‘awakened self –regulating’ is realizable in a singularity robot, it can be a proper subject that calls itself as ‘I’.

What is more concretely to be demanded in singularity robot’s presumed embodied personhood will be that the robot is uniquely or exclusively bodied. Embodied personhood of Barack Hussein Obama is derived from the fact that he has his body exclusively of others. A reason why the robot has to satisfy the condition that it should be exclusively bodied may come from the common sense belief that a robot has traditionally regarded to be one of possibly multiple realizers which were output from one and the same program. Cartesian criteria of personal identity is that a and b are the same person if and only if they share all the mental experiences. In accordance to this criteria, c and d are the same robot person if and only if they share all the content softwares. This criteria raises difficulties and troubles. For possibility or actuality of multiple realization of softwares into numerous robots is bound to make the idea of robot’s identity to be useless and the idea of robot’s personhood to be empty. This is why one way for a robot to acquire embodied personhood is the manner in which the robot has its content software exclusively to its hardware. As a human would say, a robot should be able to say “I think therefore I am my body”. A robot should be able to show that Cartesian dualism does not work in his case and that its thoughts and its body are organically integrated.

One important sphere of robot’s ontology is a social space where robots construct their own community. When robots established their own community they will come to a position that their community can work together with other communities to find common interests, just as each of them found out that each robot can work together with other robots to form a stable community. When robots of the past observed that human beings play games, laugh together, and enjoy what they do collectively, they might have reflected that they had been detached from or indifferent to their own kinds. But singularity robots may come to share certain interests with their neighbors like other robots and human beings, and try to expand the span of their sympathy. With such conditions improving, one may dream that singularity robots may come to make contributions to establish an expanded community of persons of various kinds.

The last stage of ontological development of singularity robot would be a chance where AI can be integrated with genetic engineering. There should be some independent discussions as to how soon it would be possible that such a stage is realizable. Even after a singularity robot enjoyed realizations of conditions involving the notions of autonomy, exclusive relation of content, ‘I’, and communal membership, the robot needs to make it certain that it has the capacity of self-replication (Son, Donghyun 2017: 15). This stage may wait for some fusion between artificial life and robotics (Hong, Sungook 2012: 5). Without such a fusion, it might be hard for a robot to achieve an authentic personal identity.[[13]](#footnote-13)

**3. Epistemological prospects of singularity robot**

1) Semantics of general intelligence: conceptual role semantics

A distinction was made before between robots with domain specific intelligence and singularity robots with general intelligence as autonomous agents. Ontological interpretations on singularity robots draw a dark shadow about the personhood of robots. But an epistemological prospect on singularity robots attract a bright aspect of robots development. For an example, the notion of information helps to face the challenge of semantics which was an obstacle for robots and guide to develop non-monotonous logic to structure reasons of ordinary languages which have been set aside by monotonous logic. Contemporary achievements in the area of brain science are of a great use.

Beforehand, let’s ask what can be a semantics for general intelligence of singularity robot? One natural candidate for this should be the truth-conditional semantics which traditional logics have based on. But the notions of reference and of correspondence which the truth semantics presupposed confronts various difficulties. These notions are laden with ambiguities, not free from a speaker’s role, presupposing a more basic language. One other candidate would be the conceptual role semantics which has been discussed much lately. One may pay attention to this since it can be useful for a semantics for singularity robots[[14]](#footnote-14).

Cognitive science has a tradition that meaning is regarded as a function of mental causation. Some people ask “how can brains grant a meaning to their representations?” and wonder “what is it for a human being to grasp the meaning of an expression?”. It may be conjectured that brain’s granting a meaning is an act of granting proper causal roles to its representations and that grasping the meaning of an expression is another act of holding an adequate causal relation which the expression maintains in the person’s brain. A road to explain how to grasp a meaning of an expression will be made brighter as it is on the way to be possible that causal relations are interpreted as a relation among information superveniences.

One reason why conceptual role semantics can be associated with singularity robots may come from the accessibility which robots can enjoy as to the big data. The meaning of a sign tends to be determined by the way it functions in the system of representations to which the sign belongs. And robots are in a position which is more favourable to access informations about conceptual associations with other signs in the system. Going further, a question of which theory of reference is adequate would depend on the way how expressions function in the system of human thoughts. The function of a sign in the system of thought is exhibited in the architecture of its conceptual roles, and it would help to determine which theory of reference may be proper. In other words, since the aspect of conceptual roles would decide how referring expressions function, conceptual roles are more fundamental than expressions. It looks that babies acquire the concept of fitting prior to learning to use the expression ‘fitting’. These conceptual roles are relative to representational systems and maintain intentional phenomena

(Thomason 1980), maintaining degrees of suitability variously in all types of conceptual structures.

Conceptual role semantics can be concretely clarified by Block’s insights (Block 1986: 628-629). He discusses about identity and individuation of conceptual roles. Let’s see how the notion of identity works. One difference of identities in conceptual roles would be shown, when you take more time to infer B from A than I would, when you would infer ‘dangerous’ from ‘tiger whereas I do not, or when we disagree with each other as we try to infer ’vegetable’ from ‘tomato’. Conceptual role semantics rejects the truth conditional semantics and then the distinction between types of sentences, analytic and synthetic, but insists that meanings share resemblances to each other in degrees of pluralism. The notion that an expression has the mathematically singular meaning as a substance is rejected.

Block also offers a basis for individuation of conceptual roles(Block 1986: 632-633). What is the relation between natural languages and mental languages? Suppose that an English speaker tries to construct some content by using a mental vocabulary ‘CAT’. The speaker’s English computation maps ‘CAT’ into ‘cat’ and a listener’s English understanding maps ‘cat’ into ‘CAT’. Imagine that the listener heard “Felix is a cat” and answers by saying “Felix should weigh more than 0.01 grams” to the question how much Felix would weigh. The processing of informations would go roughly something like the following.

 (1) Input: Felix is a cat.

 (2) Output of structure of the sentence understanding: FELIX IS A CAT.

 (3) Output of the inference: FELIX WEIGHS MORE THAN 0.01 GRAMS.

 (4) Output of the sentence: Felix weighs more than 0.01 grams.

A sentence of a natural language and the corresponding mental sentence should have properties of different causal operations but their conceptual roles may be individuated to be the same. The individuation may be divided into two stages. First, assume that properties of causal operations in processing natural language are mediated to the standard connective link by the merits of mutual causal relations. Second, abstract away the structure which results the standard connective link. Then, individuation of conceptual roles, for an example, cause or effect of ‘cat’ is obtained in the structure that it is the same as cause or effect of ‘CAT’. An illustration can be given below.

 (5). Keyboard language: 3 + 4 = ?

 (6). Machine language production from keyboard language: +(11, 100) = ?

 (7). Machine language inference: +(11, 100) = 111

 (8). Machine language production into keyboard language: 3 + 4 = 7

It looks that conceptual role semantics is more comfortable to singularity robots than theoretical semantics.

2) Logic for general intelligence: non-monotonous logic

A singularity robot should be able to think a thought where it can make inferences of general intelligence as humans do in their ordinary lives if it is to be a member of the human community. Beliefs of human agents have been regarded under a singular modality or a monotonous epistemological attitude. But human beliefs in the actual world take place as tokens of non-monotonous attitudes which are loosely connected(Thomason 2013). This situation requires a model of common sensical thinking or method of deliberate thinking which covers various types of reasoning. Toward this goal some logicians distinguishes monotonic logic from non-monotonic logic[[15]](#footnote-15). Classical logic of monotonicity allows the volume of truths to increase as new knowledges are added and inferences are executed, but there is no mechanism to give up informations or to change beliefs according to ever-changing situations because of its formal rigidity. But ordinary logic of non-monotonisity is competent to eliminate certain method of inferences and to expand to new ways of inferences as some new informations are added, since it is ready to discern ways how to respect the weight of some new informations. Because this logic pays a due to the ordinary appropriateness, this can be employed as a logic of general intelligence.

The non-monotonous logic, emphasizing the value of ordinary appropriateness, formalizes phenomena of defeasible reasoning which are used in the common sense space(Strasser & Antonelli 2001). Common sensical people used to make tentative inferences in order to reach a conclusion in their ordinary space, and entertain poise, discretion or right to ignore their inferences as some appropriate new informations are added. Types of thinking as in medical diagnoses, scientific inferences, inferences of fallibility belong to this type of non-monotonous reasoning. This type of inferences is inferences which are practiced in actual lives though it is placed outside of classical logic or intuitionistic logic which covers deductive reasoning or mathematical inferences. Monotonous logic demands a type of formal inferences like “if P⊢R then P&Q⊢R” whereas non-monotonous logic allows a type of informal inferences like “though P⊢R, it may be that P&Q⊢⌝R”. Though one may think that since the road is wet it might have rained, she may make an inference that since the roof is dry it was not rained but street was cleaned. “P⊢R” in non-monotonous logic is not a syntactic operation but a probable conjecture based on typicality of P, and “P&Q⊢⌝R” corrects errors in the actual world and makes an inference to expand the human cognitive world. Non-monotonous logic is not a truth-preserving logic which is expected in the conceptual space of the timeless world but is a temporal logic which is more faithful to the ever-changing world.

Singularity robots are in favourable positions in the context of defeasible reasoning of non-monotonous logic. “Q” in non-monotonous inference “though P⊢R, P&Q⊢⌝R” is a premiss which is obtained in an ordinary space which leads to “⌝R”. Therefore if one wants to arrive at a rational, just and common sensical conclusion, she should make it certain that ‘Q’ is to be grounded on the resourceful and exhaustive database. The bigger what’s called ‘Big Data’ is the better, and robot is more accessible to the Big Data. Effectiveness of robot’s accessibility to the Bid Data will be clarified by two presuppositions. The logic of knowledge representations has two assumptions. The first one is the closed-world assumption which says that true assertions are known to be true and its converse that what is not known to be true is false. The second is the closed-world assumption which states that absence of knowledge does not imply a falsehood. Since these assumptions are two formal and troublesome, singularity robot is in advantageous position by its accessibility to Big Data for the operation of general intelligence.

**4. Moral horizon of singularity robots**

1) When a singularity robot becomes a robot person

Can a singularity robot be a human person? Can a singularity robot become a member of the human community? While epistemological prospect of a singularity robot suggests that such a robot is not impossible, ontological interpretation of a singularity implies that appearance of such a robot is doubtful so far as the present conceptual space of the human language dictates. It leads to suppose that a project to embody a singularity robot into an authentic person will be left as a project for a long time. It has been accepted that there is a mysterious boundary which cannot be crossed between life organisms and inorganic material units, but J. Craig Venter produced an output of artificial life in 2003 (Church and Regis 2012: 46-49). Then, should anybody acknowledge a possibility of dissolution of the boundary between self-reflective personal being and non-reflective machine? If such a dissolution is actually ‘inescapable’(Shin, Sangkyu 2014: 243).), a natural human person should be better prepared to meet an appearance of such a robot person. If singularity robots showed up in the cultural space where human being were not ready to face them there is a possibility that human culture may suddenly collapse into calamity. The present generation of human kind cannot afford to suppose that actualization of robot person is impossible because of its conceptual constrain, but should better prepare for a minuscule possibility of singularity robot to make it certain that our later generations may conduct a humane life worthy of them.

Singularity robots equipped with general intelligence will take much time to appear but robots with domain specific intelligence have been performing various functions far more competently. Alphago, Hubo and robots for chatting, cleaning, child caring, elderly helper, medicine, nursing and so on have been engaging with counter human beings successfully, tends to be personified, and appears already to enter into marginal side of the human community. There is the gap which is getting deeper and bigger between philosophical conceptions of thinking, consciousness, reflection and everyday uses of those terms on the part of users of those robots (Cheon, Hyundeuk 2017: 32).

2) Humanity independence: singularity robot is a user of human languages

EU adopted a declaration in January 12, 2017 which grants robots with artificial intelligence to have the legal status of ‘electronic personhood’[[16]](#footnote-16). It requires robots to be equipped with device of moral constrain (Bostrum 2014: 175-200), to be responsible for their actions even with their partial autonomies. And manufacturers are expected to load robots with ‘kill switch’ so that robots are stopped whenever they are not friendly toward members of human kind. These sorts of opinions or policies have been offered by S. Hawking, AI specialists and involved organizations[[17]](#footnote-17). Some predicts singularity robots will emerge in 2045 or in 2145 and there is a chance that those intelligent robots will protest about the kill-switch device which was forced on them. They will claim that this device is racist and against humanity.

Singularity robot’s protests of such a kind look to be reasonable. For they might have based their arguments on freedom of expression, freedom of criticism, and the like, which the human civilization has adopted dearly. But humankind can respond to such protests by justifying the kill-switch device on two grounds. First, the language robots are using is the language of human kind. And moral values which humankind has developed throughout history are installed in this language. There are human virtues like caring the weak, love of neighbor, pity, healing. These are not adverbial virtues which one can live without. But they are essential virtues without which anybody can not be a human person. If a singularity robot wants to be admitted in the community of human language, they have to think in accordance with the grammar of the human language. Various languages of human kind are inter-translatable but they exhibit proper traits of humanity in a way they do not share with languages of other non-human kinds. Second, there is a clear condition to attach the kill-switch device. It is an emergency condition that is harmful to a human being. Robots have a freedom to criticize human beings but they are not free to harm humans or to reject to help humans. Since these same rules have been employed among human beings, these rules can’t be taken to be racist.

5. **Concluding: preemptive humanities**

As Aztec tribe who was ancestors of Mexican people was conquered in 1521 by white Spain General Hernán Cortés whom they believed to be the God that they have waited, should the human kind on the earth planet be conquered by robots whom they believed to be their helpers? Singularity humanities need to establish the anticipatory selfdefence toward a possible invasion on the human lives by robots. An ontological interpretation on singularity robots deem the possibility of such robot’s appearance to be low or near impossibility. The assessment of such a low possibility comes from the fact that the ontological condition for a robot is conceptually strict. Human kind’s anticipatory selfdefence toward singularity robots is an obligation to defend the future of the humankind as well as an expansion of responsibility for the survival of the future generation (Son, Wha Chul 2017: 114).

How could singularity humanities make concrete cases for the anticipatory selfdefence toward singularity robots? There should be long hours of discussions to result various proposals and to offer many choices. One form of anticipatory selfdefence is, on one hand, to equip robots with devices like ‘suicide cell’ to forbid them to be harmful to human beings and, on the other hand, to grant robots rights of coexistence with human beings in a way that human persons and robot persons cooperate and flourish. This vision of the relation between them will be possible when they accept the values embedded on the language of human beings. As the human beings of natural kind registered healing of the wounded, caring of others, discussing together, democracy, freedom, self realizations and so on to be basic virtues in their language, it would be natural to expect the robot persons of artificial kind to exhibit those values in their lives so far as they employ the language they learned and accepted[[18]](#footnote-18).[[19]](#footnote-19)

1. J. Craig Venter’s team created a simple virus life organism <JCVI-syn1.0> in November 2003 and published an article on it in May 2010. Smith ∙ Hutchison ∙ Pfannkoch ∙ Venter (2003); Pennisi (2010); Pollack (2010). [↑](#footnote-ref-1)
2. As the ‘singularity’ of Big Bang denotes a point which demarcates wether the notion of space and time is intelligible, the ‘singularity’ of singularity robot may be taken to designate a point nullifies the traditional criterion for the extension of human beings which the human beings of natural kind has monopolized. [↑](#footnote-ref-2)
3. Kang, Young Ahn (2002); Lee, Jeong\_Woo (2009); Kim, Sang-Hwan(2012); Chung, Daihyun, et al. (2001); Shin, Syng-hwan(2010); Choi, Jinduk (2004). [↑](#footnote-ref-3)
4. Seo, Dong Wook (2013); Baik, Young-Seo (2014); Lee, Wang Joo (2014); Kang, Shin Joo · Jee, Seungho (2013); Kim, Doe Sik (2011); Kim, Sungmin · Park, Youngkyun(2011); Lee, Byungsoo(2010); Jang, Dayk (2012); Lim, Chul-Woo · Woo, Ki Dong · Choe, Junyoung, et al. (2008). [↑](#footnote-ref-4)
5. My reason to choose a life form semantics and not a truth-conditional semantics is based on discussions of Wittgenstein’s notion of the rule paradox(Chung, Daihyun 1997; 2004). [↑](#footnote-ref-5)
6. I have two hypotheses in maintaining my inclination that all things conserve proto-beliefs as an information processor. The first is that an evolutionary development from inanimate states to animate states presuppose such a notion of proto-beliefs(Chung, Daihyun 2016a) and the second is that there is not a thing without integration(不誠無物)(Chung, Daihyun 2016b). Here I just want to note that a proto-belief is not mature enough to carry a notion of accountability. [↑](#footnote-ref-6)
7. The distinction in human minds between functional aspects of cognition and phenomenal aspects of consciousenss is clarified by Chalmers(1996; Yoon, Bosuk 2009: 227-231). What is functional is able to be coded and can be multiply realizable, but what is phenomenal is not yet to be coded, cannot be duplicated, and is not separable from its embodied physicality. [↑](#footnote-ref-7)
8. Dretske (1988; 1990); Chung, Daihyun (2001: 171-183); Kim, Hyoun (2012: 327-360). [↑](#footnote-ref-8)
9. Why it is hard to separate mental content from sentential contend has been discussed by the following 4 arguments: (i) Quine: indeterminacy of lingual content; (ii) Stich: the holistic character of content; (iii) Davidson: the irreducibility of content; (iv) Kripke’s Wittgenstein: the queerness of content. One may derive two implications from these arguments. The first is that any mental state has no definite content and the second is that any lingual form has not definite content(Chung, Daihyun 2001: (i) 34-41, (ii) 327-334, (iii) 242-258, (iv) 99-120). [↑](#footnote-ref-9)
10. A reason why a robot cannot make such a confession is that if it has mental content its mental content is not subject to the constrain of the language community since it is not yet a member of the community. [↑](#footnote-ref-10)
11. Lee, Ju Hyang (1994: 156-178); Chung, Daihyun (2001: 431-450); Kim, Sunhie (2004: 29-102). [↑](#footnote-ref-11)
12. Noh, Yangjin(2009: 113-132, 207-232); Han, Jung-Sun (2000). [↑](#footnote-ref-12)
13. ‘Singularity humanities’ is an anthropology which limits itself to the relation between human persons of natural kind and robot persons of artificial kind. But it would be difficult to restrain the scope of how far human persons of natural kind may be ‘enhanced’ with the help of science and technology. Then one can imagine a notion of ‘alternative person’ who is neither natural nor artificial, but is fusional. The present human culture should be able to face not only singularity humanities but also fusion humanities. [↑](#footnote-ref-13)
14. Heil & Mele (1991); Chung, Daihyun (1997). [↑](#footnote-ref-14)
15. Jeong, Youngki (1998: 273-276); Choi, Jeonghwa ∙ Park, Youngtaek(2010). [↑](#footnote-ref-15)
16. Ryu, Hyonjung ∙ Ahn, Jaemin ∙ Kim, Minji (2017); Song, Jaeso(2017). [↑](#footnote-ref-16)
17. Kim, Hyunjae(2017), Gwak, Nopil(2017), Song, Jinsik (2017), [Nicholas Davis](https://www.weforum.org/agenda/authors/nicholas-davis) (2016), Marr, B. (2016), Floridi, Luciano (2014), Sullins, John P. (2006), Floridi, Luciano and J. W. Sanders (2004). Isaac Asimov, an American writer, offered already in 1942 three laws of robotics: (i) Robots should not harm a human person; (ii) Robots should obey commands of a human, but not in the way to contradict the the first law; (iii) Robots should protect themselves, but only in the spirit the first two laws(Ko, Insok 2011). [↑](#footnote-ref-17)
18. The draft of the present article was read at a joint symposium of <Association of Korean Humanities Research Centers> and <Korean Humanities Expansion Center> under the theme of <Humanities and the Future of Liberal Arts Education>, March 31, 2017. I thank for discussions by participants of the symposium, especially Professors Ki-Jeong Song, Sangkyu Shin, Jae-Hee Kim, Hyundeuk Cheon, and also for three anonymous reviewers who clarified my discussions. [↑](#footnote-ref-18)
19. This is a translation of Daihyun Chung’s article “Singularity Humanities -Singularity robot is a member of the human society”(in Korean), Korean Philosophical Association, ed., *Philosophy,* 131(May 2017): 189-216.

================

**Bibliography**

Baik, Young-Seo (2014), *A Road to Social Humanities,* Seoul: Changbi.

Block, N. (1986), “Advertisement for a Semantics for Psychology.” *Midwest Studies in*  *Philosophy*, 10: 615–678.

Block, N. (1987), “Functional role and truth conditions.” *Proceedings of the Aristotelian*  *Society,* 61: 157-183.

Bostrum, N. (2014), *Superintelligence: Paths, Dangers, Strategies*. Oxford U. Press.

Buss, Sarah. (2013), “Personal Autonomy”. *The Stanford Encyclopedia of Philosophy*  (2000 Edition). Edward N. Zalta. ed. URL =

 <<https://plato.stanford.edu/archives/win2016/entries/personal-autonomy/>>.

Chalmers, David J. (1996), *The Conscious Mind: In Search of a Fundamental Theory.* New York: Oxford University Press.

Cheon, Hyundeuk(2017), “Artificial intelligence and artificial emotions-Is an emotion robot realizable? ”, Korean Philosophical Association, ed., *Philosophy*,131: 217-243.

Choi, Jeonghwa ∙ Park, Youngtaek (2010), “Ontology Open World Inference”, *Information*  *Science: Software Application,* 37-4: 282-296.

Choi, Jinduk (2004), *Humanities, Philosophy and Confucianism*, Seoul: Humanfield.

Chung, Daihyun (1997), *Philosophy of Fitting-Toward a Theory of Truth and Meaning*, Seoul: Chulhakgwahyunsilsa

Chung, Daihyun (2001), *Embodied Mental Content-Contextual Externalism of Mental*  *Language*, Seoul: Acadia.

Chung, Daihyun(2004), “There’s No Such Facts-Investigation of Normativity of Factual Sentence Meaning”, Korean Society of Analytic Philosophy, ed., *Philosophical*  *Analysis,*  10(2004.12): 1-24.

Chung, Daihyun(2013), *This Can Be Spoken Thus-Pluralistic Realism,* Seoul: Sechang.

Chung, Daihyun (2014), “Dispositions: An Integrational(誠) Analysis”, *Diogène*, 248: 59-70.

Chung, D. (2016a), “Integrationality: A Metaphysical Basis for the Concept of Causation”, *Philosophical Analysis* (ed., The Korean Society of Analytic Philosophy), 17-1: 1-20.<http://philpapers.org/rec/CHUIAM>

Chung, D. (2016b), “Can Reference Be Naturalized? - Notes Toward an Integrational Causality”, *Philosophy Study*, 6-5: 289-304. <<http://philpapers.org/rec/CHUCRB>>

Chung, Daihyun, et al. (2001), *Expressive Humanities*, Seoul: Thinking Tree.

Church, G. M. and E. Regis (2012), *Regenesis: How Synthetic Biology Will Reinvent*  *Nature and Ourselves*. New York: Basic Books.

Crnkovic, G. D. and W. Hofkirchner. (2011), “Floridi’s ‘Open Problems in Philosophy of Information, Ten Years Later.” *Information* 2: 327-359.

Davidson, D. (1980), Essays on Acions and Events. Oxford: Clarendon.

[Davis, Nicholas](https://www.weforum.org/agenda/authors/nicholas-davis) (2016), “What is the fourth industrial revolution?”;

 <https://www.weforum.org/agenda/2016/01/what-is-the-fourth-industrial-revolution/>

Dretske, Fred (1988), *Explaining Behaviour: Reasons in a World of Causes*. MIT Press.

Dretske, Fred (1990), "Reply to Reviewers." *Philosophy and Phenomenological Research,* L-4: 819-839.

Floridi, Luciano (2010), *Information: A Very Short Introduction.* Oxford U. Press.

Floridi, Luciano (2014), [*The Fourth Revolution: How the Infosphere is Reshaping Human*  *Reality*](https://www.amazon.com/Fourth-Revolution-Infosphere-Reshaping-Reality-ebook/dp/B00KB1BRSM/ref%3Dsr_1_3%5C?s=books&ie=UTF8&qid=1487017814&sr=1-3&keywords=the+4th+industrial+revolution)*.* Oxford University Press.

Floridi, Luciano (2017), “Semantic Conception of Information.” S*tanford Encyclopedia of*

 *Philosophy.* Edward N. Zalta. ed. URL

 <https://plato.stanford.edu/archives/spr2017/entries/information-semantic/>.

Floridi, Luciano and J. W. Sanders (2004), “On the morality of artificial agents.” *Minds and Machines.* 14-3: 349-379.

Floridi, Luciano ed. (2004), *Philosophy of Computing and Information.* Blackwell.

Gavrilova, M.L., Tan, C.J.K., et. al, eds. 2009. *Transactions on Computational*  *Science V.* : Issues on Cognitive Knowledge Representation. Springer.

Gwak, Nopil(2017), “23 Rules to Avoid AI Calamities”, *Hangyore Daily News*, 2017.03.16; <<http://plug.hani.co.kr/futures>>.

Hamilton O. Smith; C. A. Hutchison; C. Pfannkoch; J. Craig Venter. (2003), "Generating a synthetic genome by whole genome assembly: assembly: {phi}X174 bacteriophage from synthetic oligonucleotides". *Proceedings of the*  *National Academy of Sciences* 100 (26): 15440–15445.

Han, Jung-Sun (2000), “Habits and Habitual Cognition”, *Philosophy and Phenomenology*  *Research*, 29: 1-25.

Heil, J. et. al. (1991), “Mental Causes”. *American Philosophical Quarterly* 1991: 61-71.

Hong, S. & Jang, D., ed.(2012), *Brain Science: Crossing the Boundary*, Seoul: Bada.

Hong, S. (2012), “Adventures Brain Science is Opening”, Hong & Jang, ed.(2012): 5-11.

Im, Chul-Woo · Woo, Ki Dong · Choe, Junyoung, et al. (2008), *Happy Humanities*, Seoul, Eemajin.

Jang, Dayk (2012), “Future of Humanities, Future of Neuro Humanities”, Hong & Jang, ed.(2012): 332-336.

Jeon, Sungmin(2015), “Hybo-Pride of Korean Robot”, *Dongah Science*, June 8, 2015.

Jeong, Youngki (1998), “Rationality of Non-monotonous Logic”, Korean Society of Analytic Philosophy, ed., *Philosophical Analysis*, (1998): 261-288.

Kang, Shin Joo · Jee, Seungho (2013), *Humanities Shin Joo Kang Faced* (in Korean),Seoul: Sidaeeychang.

Kang, Young Ahn (2002); *Knowledge with a Human Face-Toward Philosophy of*  *Humanities,* Seoul: Sonamu.

Kim, Doe Sik (2011); Conceptual Analysis of Unification Humanities, Gonguk University, *Treatises of Unification Humanities*, 51: 155-172.

Kim, Hyoun (2012), “Higher Order of Consciousness and Misrepresentation”, Seoul National University, *The Journal of Philosophical Ideas*, 46: 327-360.

Kim, Hyunjae(2017), “Steven Hawking: Need for the World Government to Constrain AI”, *Yonhap News*, March 12, 2017.

Kim, Jaegwon(1997), *The Philosophy of Mind*, tr., J. Ha & S. Kim, Chulhakgwahyunsilsa.

Kim, Jaegwon(1999), *The Mind in the Physical World*, tr. J. Ha, Chulhakgwahyunsilsa.

Kim, Jae-Hee (2017), “Technology and Alienation-Cases of Simongdong and Stiglere, Korean Philosophical Association, ed., *Spring Proceeding,* (2017): 85-101.

Kim, Sang-Hwan(2012); *Philosophy and Liberal Imagination*, Seoul: Moonhakgwajisungsa.

Kim, Sungmin · Park, Youngkyun(2011), “Unification Discourse and Unification Humanities-An Investigation on an Unification Paradigm”, Korea University, *Philosophical Studies,* 92: 143-172.

Kim, Sunhie (2004), *Person and Body in the Cyber Age*, Seoul: Acanet.

Ko, Insok (2011), “Reconsidering Asimov’s three Laws of Robots”, Korea University, *Philosophical Studies*, 93: 97-120.

Ko, Insok (2017), “Human in the age of AI: Who am I?”, Society of Philosophical Studies, ed. *Philosophical Studies*, 2017: 67-76.

Kurzweil, Ray(2007), *Singularity is Coming*, Myungnam Kim, et al. trs., Seoul: Kimyoungsa.

Kurzweil, Ray (2012), “Thought Experiments on the Mind.” *How to Create a Mind*. Penguin Books: 199-247.

Lee, Byungsoo(2010), “Reflections on Normative Discourses of Unification”, Korean Association For Studies of Philosophical Thought, ed., *Epoch and Philosophy,* 21-2: 355-388.

Lee, Jeong-Woo (2009), *What is Subject?* Seoul, Gureenbee.

Lee, Jin-Woo (2017), “AI – It Goes Over Human Beings”, Korean Philosophical Association, *Spring Proceedings*: 1-13.

Lee, Ju Hyang (1994), “Causal Powers of Mental States and Non-individualism”, Korean Philosophical Association, ed., *Philosophy,* 41: 156-178.

Lee, Jung Won (2017), “The Age of AI: What Can Philosophers Do?”, Korean Philosophical Association, ed., *Spring Proceedings*, 2017: 123-130, 125.

Lee, Wang Joo (2014), *Humanities for Wounded,* Seoul: Daumsaenggak.

Marr, B. (2016), “Why Everyone Must Get Ready For The 4th Industrial Revolution.” <https://www.forbes.com/sites/bernardmarr/2016/04/05/why-everyone-must-get-

 ready-for-4th-industrial-revolution/#403a26413f90>

McCarthy, J. (1980), “Circumscription – A Form of Non-Monotonic Reasoning.” *Artifical*  *Intelligence* 13: 27–29.

McKay, T. & Michael Nelson (2000), “Propositional Attitude Reports.” *The Stanford* *Encyclopedia of Philosophy* (2000 Edition), Edward N. Zalta. ed. URL = <https://plato.stanford.edu/archives/spr2014/entries/prop-attitude-reports/>.

Moore, Robert C. (1985), “Semantical considerations on non-monotonic logic.” *Artifical*  *Intelligence* 25-1: 75–94.

Noh, Yangjin (2009), *Body ∙ Language ∙ Philosophy,* Seogwangsa.

Pennisi, E. (2010), ["Genomics. Synthetic genome brings new life to bacterium."](http://www.sciencemag.org/cgi/reprint/328/5981/958.pdf) *Science.* 328 (5981): 958–9.

Pollack, Andrew. (2010), “His Corporate Strategy: The Scientific Method.” *NYTimes*  September 4, 2010.

Preston, J. and M. Bishop. eds. (2002), *Views into the Chinese Room: New Essays on*  *Searle and Artificial Intelligence*. New York: Oxford University Press.

Reiter, Raymond (1980), “A Logic for Default Reasoning.” *Artificial Intelligence* 13: 81–132.

Richard, M. (1983), “Direct reference and ascriptions of belief.” *Journal of Philosophical Logic* 12: 425-452.

Rim, Ill-Hwan(1999), “Information Knowledge Cognition Conception”, Society of Philosophical Studies, ed., *Philosophical Scrutiny of Information Society*, Seoul: Chulhakgwahyunsilsa: 37-72.

Ryu, Hyonjung ∙ Ahn, Jaemin ∙ Kim, Minji (2017), “EU Grants Robots Electronic Personhood”, *Chosen Ilbo,* January 16, 2017.

Salmon, W. (1984), *Scientific Explanation and the Causal Structure of the World.* Princeton: Princeton University Press.

Salmon, W. (1998), *Causality and Explanation*. New York: Oxford University Press.

Schwab, Klaus. (2016), *The Fourth Industrial Revolution,* Penguin.

Searle, J. (2002), “Twenty-one Years in the Chinese Room.” In Preston and Bishop. eds. (2002).

Seo, Dong Wook (2013), “How is St. Paul a Political Philosopher?”, Dong Wook Seo, et al. eds., (2013); *Fighting Humanities*, Seoul: Banbi.

Shannon, Claude. (1948), "A Mathematical Theory of Communication." *Bell Systems*  *Technical Journal* 27: 379-423.

Shin, Sangkyu (2014), *The Future of Homo Sapiens-Posthuman and Transhumanism*, Seoul, Acanet.

Shin, Sangkyu (2017), “Can AI be an Autonomous Moral Agent?”, Korean Philosophical

 Assocaition, ed., *Philosophy,* 132(August 2017): 265-292.

Shin, Syng-hwan (2010), Humanities Here and Now, Seoul: Humanitas.

Son Donghyun (2017), “Reconsidered Conditions for Being Humanlike”, Society of Philosophical Studies, ed., *Philosophical Studies*, (2017): 3-17.

Son, Wha Chul (2018), “Governance of Science and Technology in the AI Age”, Seoul National University, *The Journal of Philosophical Ideas,* 68(May 2018): 267-299.

Song, Jaeso(2017). “AI: a Friend or a Foe?”, Dasan Research Center,

 <http://www.edasan.org/html2/board/index.html.>

Song, Jinsik (2017), “The 4th Industrial Revolution: Is This All Right?”, *Kyunghyang Daily,*  February 25, 2017, <truejs@kyunghyang.com>.

Strasser, Christian & G. Aldo Antonelli (2016), “Non-monotonic Logic.” *The*  *Stanford Encyclopedia of Philosophy* (2000 Edition). Edward N. Zalta. ed.

 URL = <https://plato.stanford.edu/archives/win2016/entries/logic-nonmonotonic/>.

Sullins, John ( 2016. "Information Technology and Moral Values." *The Stanford*  *Encyclopedia of Philosophy*  (Spring 2016 Edition). Edward N. Zalta. ed. URL=

<https://plato.stanford.edu/archives/spr2016/entries/it-moral-values/>.

Sullins, John P. 2006. “When Is a Robot a Moral Agent?.” *International Review of Information Ethics*. Vol. 6 (12/2006): [<http://hdl.handle.net/10211.1/427>](http://hdl.handle.net/10211.1/427).

Thomason, Richmond H. (1980), “A model theory for propositional attitudes.” *Linguistics*  *and Philosophy* 4: 47–70.

Thomason, Richmond H. (2013), “Belief, Intention, and Practicality: Loosening up Agents and Their Propositional Attitudes.”

 <http://web.eecs.umich.edu/~rthomaso/documents/pr-reas/ep.pdf>

Turing, A. (1992), *Mechanical Intelligence.* D.C. Ince. ed. Amsterdam: North-Holland.

Wang, Y. (2007), “The theoretical framework of cognitive informatics.” *Int. J. Cognit.*  *Inform. Nat. Intell*. 1: 1-27.

Wang, Y. (2009), “Formal Description of the Cognitive Process of Memorization”, In Gavrilova (2009), pp. 1-19.

Yoon, Bosuk (2009), *Computer and Mind-Status of Mind in the Physical World,* Seoul: Acanet.

=================== [↑](#footnote-ref-19)