INTERSUBJECTIVITY, "OTHER INTELLIGENCES" AND THE PHILOSOPHICAL CONSTITUTION OF THE HUMAN-ROBOTICS-INTERACTION[1]

Bernhard Irrgang

University of Technology, Dresden

Abstract

Questions of validity are not questions of nature, but rather emerge from the conditions of constitution of the anthropological potentials and subjectivity of the embodied-human subject. Through the body as coping-mastery, the lifeworld of humans is both socially and technically signed. Essential for an analysis of intersubjectivity is a threefold hermeneutics of the other as natural, artificial and human. The human lived body is a special case, in which all three dimensions of this hermeneutics are incorporated. Hence, the body can serve as a model for intersubjectivity.

Heidegger's ontological evaluation of practical rationality (praktische Vernunft) is not unproblematic. It leads to a certain marginalization of theoretical and practical rationality (Vernunft). Therefore, supplements are needed. (Figal 2006, 25-57). Among these, first and foremost, would be a "thesis of handling" ("Umgangsthese"). The development of such a thesis requires the development of a phenomenological conception of intersubjectivity. Questions of validity (Geltungsfragen) are not questions of nature, but rather emerge - within the post-phenomenological perspective - from the questions regarding the conditions of constitution of the anthropological potential and subjectivity of the embodied-human subject. Corresponding to my thesis, through being-body (Leibsein) as "coping-mastery" (Umgehen-Können), the Lifeworld (Lebenswelt) of humans is both social and technically formed. Essential for an analysis of Intersubjectivity is a threefold hermeneutics of "the other" (der Andere) as nature, artificial and human. The human (lived) body (Leib) is a special case, in which all three dimensions of this hermeneutics are incorporated. Hence, the body can serve as a model for Intersubjectivity.

From Intersubjectivity to the Hermeneutics of "Other Intelligences"

In the phenomenology of Husserl, intersubjectivity is the term for all forms of togetherness of multiple transcendental or mundane egos (Iche). This togetherness, is rooted in a collectivisation, which comes from transcendental Ego (Ich). The archetype of this

collectivisation is the encounter of the other, which means the constitution of the first not-I or other. The constitutive course of the not-I (Nicht-Ich)-experience leads us from the collectivisation of the transcendental Monads (Monaden) to the all-Monads and from their mundane objectivation to the constitution of the world for everyone. For Husserl, this world is the actual objective world (Heldt 1976, 521). Shaun Gallagher talks about pragmatic intersubjectivity (Gallagher 1996) and takes a first step in my direction, because he interprets personal intersubjectivity, from the 2nd-person-perspective. By the time children start to integrate actions in pragmatic contexts, they achieve some sort of second intersubjectivity. Approximately at the age of one, children are able to go beyond direct and immediate person-person-relation of primordial intersubjectivity, leave it and capture other contexts.

In his treatise "The Philosopher and His Shadow" Merleau-Ponty outlines the philosophy of the (lived) body (Leib). His assumption is that there is an objective interpretational dilemma when it comes to the arbitrary, accidental or ambiguous. This dilemma is closely linked to the philosophy of the body (Leib). To comprehend the body in the Husserlian sense, we need a reduction that leaves the natural attitude behind. The natural attitude refers to the human body in a way opposed to nature. Furthermore, we need an attitude with regard to the body, which implies a certain relation between nature and the human mind (Merleau-Ponty 1960, 202-205). My body is integrated in the visible world. Aside from that, my body is able to feel, for instance, my right hand. In this process the (lived) body (Leib) is perceived. The body is characterized by its fleshy texture, and respective structure. But the human being can become an alter ego for others as well. In this respect he/she is not just flesh (Merleau-Ponty 1960, 215).

Important for the body (Leib) is the co-presence of my consciousness and my body (Körper). This body extends to the bodies beyond me, i. e. into intersubjectivity. For this, empathy is needed (Merleau-Ponty 1960, 220f.). Within the spectacle of the world, I have to find means to gather my thoughts and make them a piece of my life (Merleau-Ponty 1960, 224). Whether with or against his will, Husserl is revealing a wild world and mind. Things can only be understood from perspectivity, as the Renaissance illustrated fairly well. There emerges a projective world according to requirements of a panorama. This baroque world is not merely the minds concession to nature. Rather this world is rediscovered by the mind, namely pure mind without affection from culture. In fact, this pure mind has to build a new culture. The non-relative however, is neither given by nature itself, nor by the absolute mind's

systems. It is grounded in Being (Sein), which precedes the human (Merleau-Ponty 1960, 228).

The heterophenomenology of other humans is indeed based on the second-person-perspective (2 pp). While heterophenomenology of other intelligences, namely the intelligence of animals and machines, is a result of the third-person-perspective (3 pp). Therefore, it is not right to call it heterophenomenology. Instead, next to a phenomenology of other humans, we should develop a phenomenology of other non-human intelligences. Dennett's differentiation between a phenomenology of the "I" and a phenomenology of others misses the self-conception of phenomenology. Phenomenological inquiries on the presence of others turn against the illegitimate occlusions of Cartesian philosophy of consciousness, which created the problem of foreign psyche. Wittgenstein pointed out already, that the everyday speech provides no clue for sceptical considerations about foreign psyches. The human infant discovers the 2 pp or other humans as something special between the ages of 9 to 12 months. This is important for the development of morality and an adequate consciousness of self and ego too. During these months, the child realizes the other (usually the mother) as a human person, distinct from other things and creatures, equipped with own intentionality, norms and value.

The heterophenomenology suggested by Dennett implies the expansion of the concept of intersubjectivity to discussions about other intelligences. Given the different achievements of the species, it is absurd to compare chimpanzees with humans. They don't have a human body (Leib) and corresponding human competences. While we have perhaps underestimated the mental competences of animals, nevertheless they don't come close to human bodily competences. To restrict oneself to an atomization of mental states misses the real problem: namely the problem of human embodiment (Leiblichkeit) and the interpretation of that embodiment in the framework of a corresponding theory of subjectivity.

A Robot behaves within a frame. An animal behaves within a frame of behavioural patterns. But Humans act inside a horizon. For this reason, it seems to me, the problem of intersubjectivity and heterophenomenology arises against the background of the perspectivity of human subjectivity. The question regarding perspectivity of human intersubjectivity constitutes the problem of horizon for human-embodied action. Competences have horizons; they are not constituted by omnipotence. The frame-problem in the context of a functionalistic

theory of representation needs to be questioned for its grounds through a phenomenological reduction. If we do this, we will see that active, perspectival and embodied personality designs its own emotional-mental horizons. Situated robotics is in need of programmed frameworks. These frameworks are given through technology and the technician. Abstractly speaking, the behaviour of the animal is determined by nature. But human action already calls for greater active participation and has to be partially shaped by human itself.

Heterophenomenology, in a time of the technological superman, involves new dimensions. These are threefold: the heterophenomenology of other bodies (Leiber) (human intelligences), other living intelligences (organic intelligences, mostly in an animal manner), and other technical intelligences. The gradualism used here operates according to analogy. It involves other bodies (Leiber) in the sense of others, who are sensitive to pain. It is about other bodies (Leiber) in terms of the possession of emotions. And other bodies (Leiber) in terms of behaviour which can be modelled through simulations of action. The phenomenology of the given, in terms of the subject taking place, needs to be separated from a phenomenology of the created.

The heterophenomenological perspective in terms of a transclassical phenomenology allows three types of other intelligences:

(1) The other human being (with awareness of future and death and a model of self); human intelligence
(2) Other biological intelligences in terms of animals (with consciousness, and perhaps self-consciousness); natural intelligence
(3) Autonomous operating machines (without a model of self and consciousness), artificial intelligence.

We don't need a heterophenomenology but a new hermeneutics of the living, of nature, and of technology. Based on our embodied (leiblich) nature, we can allow animals and things to speak. So a hermeneutics of other intelligences, of animals as well as of technology or nature is absolutely necessary. The advance of human potentials for action will create new changes in old values and also create new values. Therefore, humans create new matters of fact. But these new matters of fact alone do not move values forward, but rather, this is provided by interpretation. The unification of the laboratory and discourse is the way in which

science proceeds. The constitution of the lifeworld (Lebenswelt) happens by means of constant expansions with interpretations, valuations, discussions, conflicts and integration into hierarchies and institutionalization. Collectives, as well as praxis and actions, are based on acceptance. Lifeworlds and cultures are worlds we always take for granted, until we reflect on them.

On Interpretation of Natural Intelligences (Organic Intelligence)

In many ways we have to be careful about the interpretation of animal behavior. We must first decide, which facts are really facts. In any case, we assume a deep and close connection between mental states and behavior. We observe, following our theoretical arguments, that many of our mental states are correlates of characteristic behavior. We observe other people and detect the same features of behavior or patterns of behavior. Therefore, we inductively presume that they are accompanied by various mental states. An inductive argument based on just one case, allows no such generalization (Gertler, Shapiro 2007, 407-410). Even if chimpanzees can pass the mirror-test, there are substantial doubts that they can identify mental states of others or even their own (Gertler, Shapiro 2007, 431). What seems clear is that research has underestimated the whole field of implicit (tacit) knowledge and competence in mammals, and even in humans. But humans generated implicit cultural and strategic competences based on sensorimotor dexterity and technical competences and handling of speech resulting from that dexterity. These implicit competences and knowledge have no complement in the realm of the animals. Humans are more open and adaptive, but are also more prone to illnesses and abnormal behavior. Already the surplus in tacit knowledge and competence, incorporates the potentials of the "homo faber" - explicit knowledge, reflection and self-reflection, rationality and freedom of action - all exclusively count as specific human. Even what we consider as animalistic in our human nature, is specifically human. This includes "homo technicus" as well.

We have certain traits of tacit knowledge in common with animals. This includes even potential self-knowledge. We have also implanted a part of our instrumental tacit knowledge into machines. The development of the human body is far less determined by genetic dispositions than that of other biological organisms. Therefore in contrast to other creatures, the development of the human brain and competences are more contingent on the environment and hence more individualistic. These are first clues and aspects of freedom and

action-competence (Handeln-Können) in humans. It seems to me that chimpanzees have varieties of mental states, but that they don't build up the subjectivity of an embodied human subject, with its extensive sensorimotor, linguistic and theoretical competences. An objection against the so-called one-case-induction is that one achieves knowledge not based on one case, but based on many states involved in the 1st person-perspective. This includes not only mental states, but also the effects of certain actions. Basically, we need to accept that the attributions of mental states, as well as the attributions of values, are just that: attributions and not descriptions. We can't describe the mental states of other humans, animals or machines. We can only describe our own perceptions, not the ones we make available to others. Indeed, there are good and plausible reasons for attributing human intelligence only to humans and not to animals, and for denying this kind of intelligence when it comes to machines.

Thus, animals have non-human minds or intelligences. They may perhaps possess some kind of consciousness, some kind of intentionality and in part self-consciousness. Based on genetic dispositions and instructions, they are able to learn certain actions. Machines exhibit a non-human mechanical mind without conscious, intentionality, or self-consciousness. They can behave according to programming. Compared to other humans, humans exhibit a different intentionality, self-awareness, fine motor skills and a verbal speech, which is capable of abstraction, graphical representation and formalization. It is important to locate the problem of other intelligences or other minds in the context of praxis. "Intelligence" embedded in machines; embedded in biological organisms and embedded in human bodies (Leiber) is always something different. Animals have merely a 1st-person-perspective, robots have no perspectivity. Obviously, the human-animal-comparison gives support for the assumption, that human mind is a construct of interpretations. The human-animal-comparison atomizes the characteristics and behavior and compares only particular individuals. Instead the "thesis of handling" considers machines, animals and humans as a whole.

Handling of Artificial Intelligence: How Autonomous Can Robots Become?

The basic idea of the agent was created in the 1950s. John McCarthy's software pioneered the underlying system, but the term came to be used decades later, at the time Apple developed the Knowledge Navigator in 1989. Agents are basically digital valets or butler. The metaphor of the "butler" has turned out to be useful in this context. It's a matter of

digital butlers, Info-butlers and agents in man-shape. This visual metaphor is malleable. It supersedes the simulation (Johnson 1997). The idea of agents, censors, zombies, which came to the fore in the AI, replaced philosophical ideas of the acting "I" (Ich) with a conglomeration of pseudo-"I´s" and names. What is accomplished by this replacement? The acting "I" is interpreted as a robot. The basis of the connectionist model in AI is a positivistic-mechanistic and reductionistic model of the human mind. The question then emerges: Is there a human thinking, which can escape embeddedness in the lifeworld? A thinking that is not libidinously emotionally or communicatively embedded? Such a thinking, has been constituted in the application of our technologies. The construction of houses, temples, weir systems and irrigation systems required a technical-constructive thinking, which was increasingly translatable to mathematics and typified the simulation of technical praxis.

Many want to assign moral status to machines and their "actions". As for humans it is often vice versa. To epiphenomenalism the "will" is an epiphenomenon of neuronal activity. Therefore the neural impulse generates two different effects: A mental and a physical effect, an act of volition and physiological body movement. We mistake the will as cause of the body movement (Zoglauer 1998).

The delegation of the chain of means and ends to technical systems and machines leads to an instrumental reduction of the schemes of action (Handlungsschemata). Robots are upgraded machines, but they remain machines. There is no fundamental difference between an automatic loom and a humanoid robot. Robots which can act in an "autonomous" way, or more precisely, which in many situations act like an animal, remain technical products. They are mere tools and not acting subjects. The acting of a robot is a case of action without acting subject. The reproduction of technical schemes of action (Handlungsschemata) took place since the Industrial Revolution. There, with the replication of the spinning activity and connection of looms, the first reproduction of complex technical action occurred. Therefore, the modeling and simulation of human competence including tool-use isn't something new. But indeed, what could be possible in robotics in the future, is extension of the basic schemata to non-technical action-sequences, such as perception, mobility etc., which can be used for non-technical ends. With it, a further mechanization (Technisierung) of everyday life will take place. But this is not a fundamental new fact (Irrgang 2008).

The different levels of intelligence in humans and machines can be found in the fringe-consciousness, to allow ambiguity, to differentiate between essential and non-essential aspects and in the graspable composition. To find a plausible register of terms (Begriffsverzeichnis) for just one domain is associated with enormous effort. The natural language shaped up as more complex than assumed. Without the existence of a global context, reciprocal understanding is impossible. Insofar, a contextual frame is to be sought. This frame is the shared culture. Human intelligence has the ability to unlock the sense of the words out of their context. Therefore it is not possible to analyze human behavior as a mere rule-guided processing of a certain amount of elements (Dreyfus 1985, 154-173).

Very helpful for an understanding of some issues in robotics and AI is Ricoeur's concept of action without an acting "I" (Ricoeur 1990, 73). Following Ricoeur, we have to ask, whether robots are able to act and how we can evaluate the actions of robots. In the framework of possible behavior patterns of a robot, it is undeniable that a robot in general can fulfill action-schemata (Handlungsschema). Action-schemata in general meaning is given, when intentionality and end of the action are shaped in advance. Furthermore, the idea of a machinelike action should be introduced. Afterwards, machine-actions and human actions seem determined. Actions which at first seem not to follow any explicit rules, may be subsumed under a rule afterwards. Rules formulated in retrospect do not apply for the future (Collins 1991). Robot-actions are a case of action without an acting subject.

But the question is still open, of whether a machine could possess an awareness of its own schematics of action through programming. We can't assume this, because the types of action performed by a certain computer should simulate human behavior as objective schemes of behavior (3 PP). The nature of the execution (Vollzugsaspekt) of a human agent or a human scheme of action can't be simulated in a machine, because it is a first person perspective (1 PP), in other words, the perspective of the execution. There may be comparability between the external aspects of human behavior and robot-behavior due to the similarity of the action-procedure (Irrgang 2005a). But without the aspect of execution (Vollzugsaspekt) given in 1 PP, one cannot speak of human action.

In the action-schemata meaning is given, when intentionality and end of action are shaped in advance. For robots, the structure of action in terms of action-procedures and the goal of action need to be programmed and created beforehand. Therefore an actual action-I

(Handlungs-Ich) is not necessary, if there is a created goal-structure. Thus the features of action seen from the 3 PP can be given, without an acting I (1 PP). The goal-structure of an action-schemata can be ethically assessed without an acting I (Ich). I can judge if a robot is performing ethically positive or negative action-schemata. However, a non-embodied (nicht-leiblich) action is action only in the abstract sense of the word, and whose consequences, the programmer who created the action-schemata and implemented it is responsible, and not the robot. Action itself is a construct of interpretation, a phenomenon of attribution. Viewed as pure physical procedures (3 PP), actions don't differ from events (Ereignisse). That's why positivists claim that moral action and human freedom don't exist. But our self-awareness tells us the opposite (Irrgang 2005a). While the engineer makes it possible for the machine to operate, by switching it on and off, or pausing and resuming operations (Irrgang 2005c), this is much different from living bodily execution (leiblicher Vollzug).

Routine is operating according to a plan, innovative action involves acting for an end not in the same way. Based on dispositions we acquire certain competences during action. Already existing skills are being updated, extended and advanced. The handling-thesis (Umgangsthese) calls for a phenomenological conception of technological artifacts (i.e. orientation towards surface-structures) and not a causal theory. The objective part of technical competence- action-schemata- can be implemented into machines. Insofar, spinning machine and weaving machine have been created through transfer of human patterns of manufacturing into mechanic production. Therefore the basis of technology is laws of nature and the objectifiable human action procedures and their possible implementation into technical artifacts. The coping-mastery (Umgehen-Können) of machines, instruments and infrastructure (with its measurable results) is a part of technical, and hence, technological praxis. Technology is objectification of both technical coping-knowledge (Umgangswisssen) and technical competence, the crosslinking of technical structures with modes of human-machine-interaction. Robots as well as synthetic life-forms are consequent results of this technological development. Technological understanding and natural scientific explanation should be applied together, in the sense of "...as well as..." (sowohl-als-auch). Technical praxis is always present in the creative process of construction, invention and research. For a general technology, the object of investigation is the relation between technical routine and technical innovation. The technical potential of an artifact is based on its inherent action-schemata. In the case of the robot, an individualization of general technical operation-schemata takes place (Irrgang 2008).

A formal action-structure of a robot or of its process sequence, detached from body (Leib), is imaginable and maybe programmable. The robot will behave in the frame of a predetermined action-pattern (Handlungsmuster). A robot can't give instructions to itself, but is bound to a given structure of instruction due to his construction. The "action" is pre-interpreted through the style of programming. Robots don't have "world". They do not orient themselves within the framework of human rationality, but within the framework of a given action-frames (Handlungsrahmen). Ultimately this is good, because AI is technical intelligence. Also the fact that there is a limit in the formalization of implicit (tacit) knowledge speaks against the assumption, that AI someday will achieve the status of natural or even human intelligence. If we accept the limitations of technicality we don't lose anything, except some of our prejudices.

REFERENCES

Collins, Harry M. 1991: "Artificial Experts." *Social Knowledge and Intelligent Machines*. Cambridge Mass./London (1990).

Dennett, Daniel 1993: *Consciousness Explained*. 1991, London.

Dennett, Daniel 2007: "Heterophenomenology, Reconsiderd" in: *Phenomenology of Cognition and Science*. 2007 6, 247-270.

Figal, Günter 2006: *Gegenständlichkeit. Das Hermeneutische und die Philosophie*. Tübingen.

Gallagher, S. 1996: "The Moral Significance of Primitive Self-Consciousness" in *Ethics* 107, 129-140.

Gallagher, Shaun 2004: "Hermeneutics and the cognitive sciences" in *Journal of Consciousness Studies* 11/2004.

Gallagher, Shaun, Dan Zahavi 2008: *The Phenomenological Mind: An introduction to Philosophy of Mind and Cognitive Science*. London, New York.

Gertler, Brie, Lawrence Shapiro 2007: (Hg.) *Arguing about the Mind*. New York, London.

Heidegger, Martin 1972: *Sein und Zeit*. Tübingen.

Heidegger, Martin 1994: *Zollikoner Seminare. Protokolle, Zwiegespräche, Briefe*. Hg. v. Medard Boss; (1987) Frankfurt.

Heidegger, Martin 2002: *Phänomenologische Interpretationen zu Aristoteles: Ausarbeitung für die Marburger und die Göttinger Philosophische Fakultät* (1922) Hg. von G. Neumann. Stuttgart.

Heldt, K. 1976: "Art: Intersubjektivität" in *HWP* IV, 521f.

Husserl, E. 1952: *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie* Bd. 2; ed. Mary Biemel, Husserliana Bd. IV, Den Haag.

Husserl, E. 1973: *Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlaß*. Teil 2 (1921-1928), Husserliana Bd. XIV, ed. I. Kern, Den Haag.

Husserl, E. 1976: *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*. ed. W. Biemel, Husserliana Bd. VI; Den Haag 1976.

Irrgang, B. 2005a: Posthumanes Menschsein? Künstliche Intelligenz, Cyberspace, Roboter, Cyborgs und Designer-Menschen - Anthropologie des künstlichen Menschen im 21. Jahrhundert; Stuttgart.

Irrgang, B. 2005b: *Einführung in die Bioethik*. München.

Irrgang, B. 2005c: "Ethical acts (actions) in Robotics" in Philip Brey, Frances. Grodzinsky, Lucas Introna (Hg.) *Ethics of New Information Technology. Proceedings of the Sixth International Conference of Computer Ethics* (CEPE 2005); Enschede 2005, 241-250.

Irrgang, B. 2005d: "Der Cyborg als der Übermensch Friedrich Nietzsches? Anmerkungen zur Posthumanismusdiskussion" in. R. Kaufmann, H. Ebelt (Hg.) *Scientia et Religio. Religionsphilosophische Orientierungen.* Fschr. für Hanna-Barbara Gerl-Falkovitz; Dresden 2005, 315-333.

Irrgang, B. 2007a: *Hermeneutische Ethik. Pragmatisch-ethische Orientierung für das Leben in technologisierten Gesellschaften*. Darmstadt.

Irrgang, B. 2007b: *Gehirn und leiblicher Geist. Phänomenologisch-hermeneutische Philosophie des Geistes*. Stuttgart.

Irrgang, B. 2008: *Philosophie der Technik*. Darmstadt.

Jaffard, R. 2005: "Das facettenreiche Gedächtnis" in *Spektrum der Wissenschaft Spezial 2/2005*. Gedächtnis 6-9.

Meltzoff, A. N.; R. Brooks 2001: "Like Me as a Building Block for Understanding Other Minds: Body, Acts, Attentions, and Intentions" in B. Malle, L. J. Moses; D. A. Baldwin: (Hg.) *Intentions and Intentionality: Foundation of Social Cognitions*. Cambridge Mass., 171-191.

Merleau-Ponty, Maurice 1960: *Signes*. Paris.

Merleau-Ponty, M. 1966: *Phänomenologie der Wahrnehmung*. übersetzt von Rudolf Böhm; Berlin.

Merleau-Ponty, M. 1976: *Die Struktur des Verhaltens*. übersetzt von Bernhard Waldenfels; Berlin, New York.

Merleau-Ponty, M. 2000: *Die Natur. Aufzeichnungen von Vorlesungen an Collège de France 1956-1960*. übersetzt von M. Köller ([1]1995).

Ricoeur, P. 1990: *Soi-même comme un autre*. Paris .

Stern, D. 1985: *The Interpersonal World of the Infant: Review from psychoanalysis and developmental psychology*. New York.

Zoglauer, Thomas 2002: *Konstruiertes Leben. Ethische Probleme der Humangentechnik*. Darmstadt.