

Machinery, Intelligence and Our Intentionality Grounds for Establishing Paradoxical Discourses

C.T.A. Schmidt

Le Mans University, Laval France
Colin.Schmidt@univ-lemans.fr

Abstract: Attaching the robotic body to the artificial brain (the computer) is a poor way of going about constructing autonomous mentality. It represents nothing more than an extension of the brain and succumbs to using experience as a confirmation of the scientist's belief that he may speak in artificial terms of mentality. This naturally leads to producing a paradoxical discourse on the subject of robotics

and thereby leads to confusion. The author indicates readings of paramount importance for disentangling the language involved in this special form of evolutionary computation.

Keywords: Robotics, AI, Philosophy, Mind/body problem, Extra-domain vision

The original target of the artificial sciences is 1/maintained but 2/heavily questioned in this article. The author feels there is no use sweeping strong Artificial Intelligence under the carpet as it remains very present in the minds of the specialists despite an apparent swing to realism. Strong AI, supported by newer fields such as evolutionary computing and cognitive, humanoid or epigenetic robotics, and yet others is questioned on the grounds of the fundamental aspects of communication. Our understanding of the function of language and what it means *to mean* something while interacting with others is a major attribute of human life. The ultimate aim herein is to enquire into the current state of knowledge in the remnants of strong AI (its implicitness, suitability...). Its implicit nature in the minds of the forerunners in the community is particularly bothersome inasmuch as it hinders other scientists from sensing it and thus apprehending its actual role.

1 The Dual Message

Scientists of the artificial have become fantastic 'modern archers'. In comparison to early work in building computers, newer sub-fields of the artificial sciences — cognitive ergonomics, simulations theory, artificial life creatures, brain-inspired and humanoid robotics— have resulted in the development of various corporal extensions of the artificial. This is no doubt to thwart, or rather adjust to, the onslaught of neo-Dreyfusian literature produced since 1972. It would seem that the ability to *learn* via the tools of modernity, *develop* oneself/itself and to some extent *adapt* has been pinpointed as key issues in this area; hence the importance of Internet connectivity and meaning itself for the intellectual elements and behaviourist approaches for the physical ones. Even fusion here is being attempted. Over the years, these scientists have been 'shooting' even higher and higher, have recognised cognitivism as a necessary component in reactivating their (behaviour-heavy) theories and have covered a lot of ground in the name of progress.

For the sake of the artificial sciences, I would like to explore territories beyond cognitivism; but in order for the robotics community to follow me in my adventures, they will have to admit to holding parallel adherence to the following two statements, adherence that I shall briefly demonstrate. As a matter of fact, there seems to be a Jekyll and Hyde-like discourse going on amongst specialists implicated in robotics. Though they do not exactly hold the same positions, they (Balkenius, Kozima, Prince, Kezuma and specifically R. Brooks, T. Ziemke, J. Zlatev and yet others) say:

- 1/they do not want to wake us from our narcissistic dreaming;
- 2/they want to recreate Man in his own image.

They conduct hands-on and/or theoretical research on embodying artificial minds to come in line with the phenomenological necessities for the correct functioning of a mind as Dreyfus suggested, but then say it is *just* for modelling cognition. *Just for modelling cognition?* This is double talk, though they probably do not realise that they are doing this. Ziemke for example, while (seemingly) arguing against strong AI-related technology in his 2002 *Sign Systems Studies* article, says that a mind's character is determined by the body that supports it and goes on to say: "if cognition is dependent on body and sensorimotor capacities, then the only way to achieve or study human-level or human-like intelligence in artefacts is to equip them with human-like bodies and sensorimotor capacities, i.e. to build *humanoid robots*¹". What does the author not rule out here? Notice that he *does* in fact allow room for achieving or recreating the full intellectual capacities of *homo sapiens* in his discourse. Essentially, the fundamental long-term goal Ziemke must have is to reconstruct *human* reality in robots. In the Introduction of the Proceedings of the 1st *International Conference on Epigenetic Robotics*, J. Zlatev and C. Balkenius take up a similar stance and discussing the possible existence of "true intelligence" in artificial systems; they put brackets around the term "possible" as if it had nothing to do there.

On a more practical note, things are today as they are, and so the fact that the *artificial "cloning"* of man has not happened yet, this scientific community has reverted to the lesser task of just building 'spare parts', spin-offs products or debating policies and theories. In light of the standpoint they find themselves at, perhaps all they can do for now is generally just study human cognition.

Realism is growing in the artificial sciences but I must ask this question: why would *roboticists* want to study cognition? Just for fun? Is not the main goal in life of a roboticist to build robots? The idea is to put the knowledge acquired from their studying cognition into practice, even if at a later date. For the time being they are modelling cognition and looking into the question of how to get the machine to learn progressively; presently, "only the software not the hardware develops" according the Ziemke. At a later time when all the spare parts have been perfected, they will take up the huge and "more noble" task of connecting them all together, say with some serious help from biologists, plastic surgeons and other body specialists, adding a sprinkle of Shellyan "theory"², and introduce us to their brand spanking new... Frankenstein (or rather the doctor's monster).

The reader must please excuse this caricature, but is that not roughly the aim of the project? How do the general public interpret their work? Of course the instigators of like projects will be better equipped to avoid the creature's possible upheavals —at the design stage. And the release of such 'enhanced units' will benefit from the newly acquired knowledge on learning and intentionality in the robotics community and the formal *building-blocks approach* brought forth by M. Minsky. But why hide behind current-day technical issues? In the end, the ultimate goal roboticists have in mind, however far off it may be, will cause us to lose our narcissistic hold on the world. Why don't they just come out and say it? The reader must understand that I am not addressing the here-and-now technical difficulties stated (ever so precisely) in the community's literature. In fact I can accept the presupposition according to which they *are* able to achieve true fully human intelligence, but what do they wish society to become with these new inventions?

¹ The emphasis put on humanoids is not even mine.

² Mary Shelly (1818) is the author of the novel about the coming to life and emotional swings of the famous monster build by doctor Frankenstein.

I must reiterate the fact that the issues herein concern their *ultimate goal* to recreate humans. Of course the current climate in applied research (especially in Europe) is so that it does not sensibly allow these scientists to argue or publish on the aims of their science beyond 4-6 years from now. T. Ziemke, Editor of the well known *Connection Science* journal, does find the time to point out that MIT's *Cog* and Kozima's *Infanoid* projects are supposed to develop ontogenesis and that Zlatev's approach, which purports that "robotogenesis could possibly recapitulate [human] ontogenesis, leading to the emergence of intentionality, consciousness and meaning in robots", is taking things "to the extreme". As far as I can see, the idea is *still* being expressed in its most supreme form, despite the 'realistic varnish' (*pace* Ziemke, Zlatev).

1.1 Are we still on target? About otherhood

I think we can safely say we were right to wonder about the goals of robotics and their related products and prototypes, as well as their possible effects. Ethics aside, we must compare with the past to evaluate our position. When I look at the 'curve' the artificial sciences and technologies are following in general, I cannot avoid thinking that we are off the *neo-human* target man has held in his imagination for years (i.e. more advanced than the golden protocol droid C-3PO from Star Wars, RoboCop, etc.) and is now trying to implement, whatever the methods and materials. If one looks at things from a non-behavioural perspective, we are missing the bull's-eye: Artificial Selfhood. Any inkling of personhood in an entity relies on the possibility to communicate to support its very existence. This is why Turing was keen on the dialogical capabilities of the machine, perhaps more so than the actual performance in dialogue. Does concentrating on building limbs, axis torsos and eyes, etc. bring us closer to creating otherhood? Fundamentally, when a human has his leg amputated or undergoes a kidney transplant, his selfhood remains the same.

Humanoid form does not support personhood.

If we consider mental capabilities other than linguistic competence, we notice that work in Robotics has started to integrate learning, emotions etc. into the systems-building process in order to create *autonomy*, *self-will*, etc. But are these machines really independent of their creator, Man? Need they be autonomous³? The dialogical dimension of cognition should play a role in specifying human intentionality. Some seem to suggest that the goal of recreating intellect is currently limited to the level of a child's learning (J. Zlatev, C. Breazeal and —to some extent and a while ago— S. Turkle) and others more recently opt for even baby thought⁴. Strategically, if the goal is really to rebuild man, the community must put the emphasis on research on the *higher* cognitive existence of humans so that it may catch up to the physical side of robotics. Humans exist within and thanks to society.

The original goal was to create a machine in the image (and reason) of Man. Something that perhaps has little to do with the question considered makes me think that we are going even further off-topic. The strong '*animalistic turn*' in the artificial world —Artificial Life— suggests an intentional avoidance of the target (i.e. Sony *Aibo*⁵ ERS-7). What's the matter, are we not up to it? Is personhood too intangible? Whatever the problem is, when H.L. Dreyfus cried out for corporeal extension, I doubt he had electronic pets in mind! But then again, perhaps today's aim in robotics is to reproduce the whole evolution of life, starting with beings —animals— that possess notions of otherhood *only* (i.e. no selfhood), as P. Gärdenfors puts forth (p. 157-158).

³ Cf. Schmidt C.T.A. & Kraemer F. (2004). We address both the appropriateness and inappropriateness of using the term "autonomy" in speaking about artificial systems.

⁴ For details cf. citations in my forthcoming AAAI Fall Symposium address, *Technical Report FSS-04*.

⁵ AIBO is an acronym for "Artificial Intelligence RoBot".

1.1.1. What is the Target?

One objection to the type of investigation proposed here could be raised: time has changed the course of *scientific action*. We no longer need to reconstruct man, these "lesser devices" suffice our immediate needs and soothe our desires of controlling *'technical procreation'*. But what about *scientific thought* or the *personal will* of the scientist of the artificial? Can the scientist following in the footsteps of those making world-renown predictions as did Turing and Simon with respect to the future capacities of machines (very human), settle for toddler-like and pet-like implementations? It is clear to me that —if one is willing to consider industrial research— some of the hefty players in the corporate world have not lost sight of the original God-playing goal: take for example Sarcos, Honda P4, Asimo and the domestication initiative at Fujitsu who wish to put *HOAP-1*, (a miniature Humanoid) *into our homes*, or when S. Amagai, President of the Entertainment Robot Company at Sony Corporation, says that his robot *QRIO* "*will be your friend*"⁶.

Missing the target perhaps reflects our inability to define ourselves. In order to reconstruct an autonomous human being with full physical and mental capabilities, without forgetting the joint nature of intentionality, we will need a proper definition of Man himself. Do we have this in a usable format? Otherwise we will not be able to attain the ultimate goal we continue to express, however awkwardly, subconsciously or in a hidden fashion.

What do we wish to achieve in employing the artificial sciences? What are the technological productions supposed to represent? The motivation may be more personal. Perhaps, quite simply, what we desire in working towards the goal of "giving birth" to a human (J. Weng *et al*, in *Science*) is recognition from others or a new meaning in life —that of *Creator*— or... *could it be just for the challenge*? Amagai says the first and foremost reason for creating *QRIO* was "technological prowess"⁶. I would not consider this drive to be automatically utilitarian.

Realising the contradictions that exist is very likely to help a body of knowledge to progress. P. Gärdenfors comforts me in my anti-block building theoretical stance by stating that, if you *insist* in building with blocks, you must start with a representation of the Other, which in turn allows the addition of one's self-awareness, one's consciousness, one's language, etc., right on down to very basics (this process has been referred to as "reverse engineering" by Zlatev). Minsky, McCarthy and others did this but the other way round for fifty-some years and this is what the scientific community comes up with? "Oh, you got it the wrong way around, just reverse it". Who is wrong here? Even though the dream of rebuilding man is still around, we *are* downshifting. But not only to *Artificial Wildlife*. K. Dautenhahn conducts research seeking change or at least to find some practical uses of our current robot technology on a social level. She believes that robotic agents appearing on a screen may help in getting people to express personal facts that would be difficult to relate otherwise. Copying human beings is not paramount for her and the challenge factor is subdued in her discourse. I think machines in general should neither be smart nor social in the truly human sense as this would hinder the creation of more efficient exchanges between Humans and artificial objects. Many traditional beliefs we have about machines *and people* would be suspended, and this could render everyday communication impractical.

Why would we want to recreate man? The *Festival de Cannes du Robot* took place in June this year; obviously the dream is still lurking. Also relatively recent are events like the 8^{ème} RoboCup football competition⁷ in Lisbon, which tends to suggest that robotics has fully developed its playful aspects while leaving to its wayside the nourishing of a utilitarian goal. But their advent does represent progress in that the community itself is becoming other-aware. Mind theory must have a dialogical component. Football is not played alone. Avoiding running into other players on the field (some use input from cameras on the ceiling) is not yet up to the level of addressing the issues of language use but it does at least pertain to the

⁶ QRIO is the official name for the Sony Dream Robot which means "Quest for Curiosity", *cf.* <http://www.sony.net/SonyInfo/QRIO/interview/index2.html>. Even "Aibo" means "companion" in Japanese.

⁷ *Cf.* http://robo-erectus.org/robocup_2003.php

sociological dimension of cognition. It does represent a start to addressing the really difficult issues of recreating the human being.

1.1.1.1. Processing the Relation: a Priority

I would like to point out that, if strong robotics and hard-core AI is the chosen path, the main issue that has to be dealt with is the reciprocal exchange of knowledge. Does the superficial structure of alternating utterances evoke for today's top scientists the dialogical mind structure we must seek for this? Supposedly not. Why then the mix-up between the non-manifest phenomena —thought, intentionality... — and the manifest —language use, playing with blocks, building human-like structures...? Sometimes in newly established disciplines, discernment takes time.

Effectively, at the dawning of communication theories, dialogism was absent from the model of inspiration —human beings. With their analyses of language usage, Grice and Searle with their *Speech Acts* came close to obtaining the necessary discernment between the formations of back-and-forth language usage (manifest behaviour) and the dialogical mind (non-manifest society-based activity); they did succeed in drawing attention to the pragmatic difficulties of accounting for inter-human conversational processes. However, their efforts supplied *ex post facto* explanations of human interlocutors in practice, thus further perpetuating the monological model that was taken up, not only by people working in AI, Natural Language Processing (NLP), Automatic Translation (AT) —thus for machines and robots—, but also by those in linguistics proper trying to explain how humans themselves use language. In my mind, a later vision of how our language systems works dismantles Grice and Searle's monumental philosophies of language. F. Jacques feels it is futile to consciously amass hypotheses —i.e. "I know that you know that I intend to..."—as Grice did with his *Implicatures* in "*Logic and Conversation*". According to Jacques: "Neither the speaker nor the listener need to entertain intentions as complicated as this in order to communicate [...] we are not aware of these intentions when speaking [...] reciprocity would not live on in separation"⁸. As a matter of fact, one can only carry out an analysis of intentions-in-a-structure on the basis of linguistic expressions from the speaker and *one's own* 'calculations' of them: hence the listener has no role in this. Definitively, there exists a non-manifest dialogical realm (call it mysterious if you like) between speakers yet to pierce through to with our Science. And opening up the brain to look for intentions entails refuting the "emergence" of truly higher order cognition many roboticists are hoping to see come of their contraptions.

So we run into a big problem here, one for which I do not have the answer: how would one implement a language processing system for a robot on the basis of speakers co-producing elusive intentions, beliefs, etc.? Jacques abolishes the conceptual structure that was designed for rebuilding the human linguistic apparatus. But on an explanatory level, is he wrong? No, this should not come as a surprise as intentionality in general has always been invisible. That's because intentionality is co-intentionality. Although I do not believe in a block-building structural theory for overcoming this problem, in citing Jacques here I realise that I would have to side with Gärdenfors and Zlatev concerning the *order* of block treatment if block-building were the only way to go. According to this order, the physical side of the construction can wait. Can it not? It is expensive and there is no reason to put 'socially-lobotomised' creations on this earth if we are seeking GOFAI strength robots. Priority must given to analysis of the relation (the *Self-Other cohesion* which of course operates through other-awareness) rather than that of the poles (the *Self* and the *Other*) for a theory of communication to have the necessary explanatory power to handle co-intentionality.

⁸ My translation, from p. 185-187.

1.1.1.1.1. Methodology, Intentionality and Fiction

The release of the film *I, Robot* (2003) took place relatively recently; in this adventure picture, Will Smith faces a whole armada of robots human society does not trust. The fascination society has with robots shows that we do not fully reject the idea of rebuilding lives, including Man's. As such, I do not think we could say that the notion of robot does not stimulate the mind. Many novels carry the same type of matter.

The human imagination helps fix scientific goals like investigating analogies that may exist between organic- and inorganic-supported functions. As understanding intentionality in its conjoint form is important to the success of robotics, examples of its presence and problems must be explored to ameliorate its explanation and "implementability". In this area, robotics must be aware that it is working with the *development of systems of beliefs and values* and this is why the realist Zlatev does not rule out the Agent Smith from *The Matrix* type scenarios in the future (p. 287)]. In Section 1, I drew a parallel between Frankenstein's —which would seem to be based on fictitious events— and building robots. Galvanism, however, is not fiction—it has been performed on a human cadaver—, whereas Shelley's extrapolation on the actual experiment performed 15 years before her novel came out⁹ is. So the general idea of constructing a human being from lifeless materials, whether inorganic or not, has been around for a long time (i.e. the Shelleyan example comes relatively late in the course of history). Quite intriguingly, Dr. Victor Frankenstein had the same sort of hopes for his "contraption" as do many a doctor of Robotics or AI today. Despite a few malfunctions, his monster did demonstrate a strong *will* to learn and *imitate* his human counterparts at a behavioural level; he was able to *develop* language skills too. But perhaps the most interesting aspect of the monster's "life" was that, although he was an artefact (man-made), he was able to *experience emotions* in comparing himself to man (the Other), and was thus *conscious* of the horrible life he had, life that was plagued by his own disgracefulness. The fact that he was also able to *sense the feelings of others* and that he *realised* he was repugnant from the Other's point of view shows he had *you-awareness*, and of course, a very impressive start to possessing a theory of mind. If the doctors in robotics and AI could reach—or start out with? (as Gärdenfors points out)— the inspection of otherhood as in this sad story, this would be considered an enormous victory. This is perhaps drawing a layman's parallel of no value, but nevertheless...

The robotics/AI community needs to start elucidating some of the mystique of its profession for those who have trouble adhering to the inventions of the artificial. Public accessibility and acceptability do go hand in hand. Otherwise the public might just take their methodology to be the contemporary version of Dr. Frankenstein's. Cold soup reheated.

The beginnings of AI and robotics did really start during the "boom period" (50s to 70s), so I admit that the requirements on science and technology may have changed since veritable GOFAI days. But there are paradoxical positions being adopted by scientists, peculiarly, both *for* and *against* deranging our narcissistic self-worship. Could the shift in epochs (i.e. the increased speed) be the cause of this confusion? In any event, if they do want to wake us from our narcissistic dreams (i.e. with prototypes that deform, or rather, redefine our human value system), they will have to account for the veritable shadows lingering in the human communication process exposed above.

There is most certainly something else I do not understand about "Phenomenologic Robotics" of the authors adhering to epigenetic meaning: *they say the material elements of a human body (flesh, blood...) are not sufficient enough to explain mind so we can attempt to get mind into another entity made with other materials (iron, plastic, glass...), but in building such an entity we must not forget to add human-like eyes, ears and limbs etc., otherwise it will not be able to think properly.* So perplexing is the logic here, I wonder whether we should not totally separate the discussions on experiential meaning from those on bodily existence and not, for the time being, 'force' them together with unified theories (*pace* Zlatev).

⁹ In 1803 Giovanni Aldini, Professor of Physics in Bologna and nephew of medical doctor and physician Luigi Galvani, conducted an experiment on an executed man in which he obtained a convulsive reaction; he did so by wiring up the body of the deceased to a battery, cf. the terminological notes J. Malrieu offers us, p. 312 in the 1997 French edition of Frankenstein.

Welcome progress that is coming out of the debate is that the scientific community seems to be acquiring an extra-representationalist vision on the matter.

References

- Brooks R. (2002). *Robot: The Future of Flesh and Machines*, London: The Penguin Press.
- Dautenhahn K. (2000). *Human Cognition and Social Agent Technology, Advances in Consciousness Research Vol. 19*, Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Dreyfus, H. (1972). *What Computers Still Can't Do: A Critique of Artificial Reason* (1994, Fourth printing), Cambridge, MA: The MIT Press.
- Gärdenfors P. (1996) "Language and the Evolution of Cognition", in RIALLE V. & FISETTE D., *Penser l'esprit : des sciences de la cognition à une philosophie cognitive*, Grenoble : Presses universitaires de Grenoble.
- Grice, P. (1975). "Logic and Conversation", in Cole, P. & Morgan J. (Eds.), *Syntax and Semantics : Speech Acts, Vol. 3*, (pp. 41-58), New York/London: Academic Press.
- Fujitsu's Home page: <http://pr.fujitsu.com/>
- Honda's Home page: <http://asimo.honda.com/>
- Jacques, F. (1982). *Différence et subjectivité : Anthropologie d'un point de vue relationnel*, Paris: Aubier Montaigne.
- Sarcos' home page: <http://www.sarcos.com/>
- Scassellati B. (2001). *Foundations for a Theory of Mind for a Humanoid Robot*, Ph.D. Dissertation, May 6: Massachusetts Institute of Technology.
- Schmidt C.T.A. & Kraemer F. (2006), "Robots, Dennett and the Autonomous. A Terminological Investigation", *Minds and Machines. Journal for Artificial Intelligence, Philosophy and Cognitive Science*, Vol. 20 N° 1, Kluwer/Springer.
- Schmidt C.T.A. (2005) "Of Robots and Believing", *Minds and Machines. Journal for Artificial Intelligence, Philosophy and Cognitive Science*, Kluwer.
- Schmidt C.T. (2004), "Humanoids, from Interfaces to Intelligence. Really? A Philosophical Statement on Retrograding or Scientists Caught Back-peddling", *The American Association of Artificial Intelligence Fall Symposium Series on 'The Intersection of Cognitive Science and Robotics: From Interfaces to Intelligence'*, October 22-24 Washington D.C., *Technical Report N° FSS-04*, Menlo Park CA: The AAAI Press.
- Schmidt C.T.A. (2006), "A Relational Stance in the Philosophy of Artificial Intelligence", in L. Magnani (ed.), *Computing and Philosophy*, Associated International Academic Publishers, Pavia Italy.
- Schmidt C.T. (2004), "Let Me Introduce You to My Non-Agent", in Schaub H., Detje F. & Brüggemann U. (Eds.). *The Logic of Artificial Life*, Berlin: Aka-Verlag/IOS Press, p. 122-127.
- Shannon C.E. (1948). *A mathematical theory of communication*. In *Bell System Techn. Journ.* July & Oct.
- Shelley M. (1818). *Frankenstein*, Standard Novels; [(1988/1997) *Frankenstein ou le Prométhée moderne*, Rocher/Gallimard].
- Sonystyle home page: <http://www.sonystyle.com/is-bin/INTERSHOP.enfinity/>
- Turing A. (1950), "Computing Machinery and Intelligence", *Mind*, vol. LIX n° 236.
- Turkle S. (1984), *The Second Self, Computers and the Human Spirit*, New York: Simon & Schuster.
- Vernant D. (1996), « Intelligence de la machine et sa capacité dialogique », Rialle V. & Fiset D., *Penser l'esprit : des sciences de la cognition à une philosophie cognitive*, Grenoble : Presses universitaires de Grenoble.
- Watanabe T. (2002), "InterActor: Speech-Driven Embodied Interactive Actor", *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication RO-MAN 2002*, Sept. 25-27 2002, Berlin Germany: IEEE, p. 430-435.
- Weng J., McClelland J., Pentland A., Sporns O., Stockman I., Sur M. & Thelen E. (2001), "Autonomous Mental Development by Robots and Animals", *Science Magazine*, Vol. 291 N° 5504, The American Association for the Advancement of Science, pp. 599-600.
- Wittgenstein L. (1953). *Philosophical Investigations*, Oxford: Blackwell.
- Ziemke T. (2002), "On the epigenesis of meaning in robots and organisms: Could a humanoid robot develop a human(oid) Umwelt?", *Sign Systems Studies*, n° 30.
- Zlatev J. (2001), "The Epigenesis of Meaning in Human Beings, and Possibly in Robots", *Minds and Machines* n° 11, Kluwer, p. 155-195.
- Zlatev J. & Balkenius C. (2001), "Why Epigenetic Robotics?", Introduction in the *Proceedings of the 1st International Conference on Epigenetic Robotics*.
- Zlatev J. (2002), "Meaning = Life (+Culture). An outline of a unified biocultural theory of meaning", *Evolution of Communication*, n° 4 vol. 2, John Benjamins, p. 255-299.